System Analysis and Information Technology

SAIT 2017

May 22 – 25, 2017
Kyiv, Ukraine

IASA

Institute for Applied System Analysis

National Academy of Sciences of Ukraine
Ministry of Education and Science of Ukraine
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Nataliya D. Pankratova (Ed.)

System Analysis and Information Technologies

19-th International Conference SAIT 2017
Kyiv, Ukraine, May 22 – 25, 2017

Proceedings

IASA

Institute for Applied System Analysis
at the Igor Sikorsky Kyiv Polytechnic Institute

UDC [519.7/.8:(004+007)](100)(06)
ББК 22.18я43+72я43
С40

Volume editor:
Nataliya D. Pankratova, Dr.Sc., Prof.

Editorial board:
Petro I. Bidyuk, Dr.Sc., Prof.
Nataliya D. Pankratova, Dr.Sc., Prof.
Anatoliy I. Petrenko, Dr.Sc., Prof.
Yuriy P. Zaichenko, Dr.Sc., Prof.
Elena L. Oparina

Revising:
Gennadii D. Kiselyov, Ph.D.
Mykola A. Murga, Ph.D.
Elena L. Oparina
Illia O. Savchenko, Ph.D.
Oleksandr M. Terentiiev, Ph.D.

Design and typesetting:
Illia O. Savchenko
Mykhailo P. Makukha

System analysis and information technology: 19-th International conference SAIT 2017, Kyiv, Ukraine, May 22 – 25, 2017. Proceedings. – ESC "IASA" NTUU "Igor Sikorsky Kyiv Polytechnic Institute", 2017. – 340 p.

С40 **Системный анализ и информационные технологии**: материалы 19-й Международной научно-технической конференции SAIT 2017, Киев, 22 – 25 мая 2017 г. / УНК "ИПСА" НТУУ "КПИ им. Игоря Сикорского". – К.: УНК "ИПСА" НТУУ "КПИ им. Игоря Сикорского", 2017. – 340 с. – Текст: укр., рус., англ.

С40 **Системний аналіз та інформаційні технології**: матеріали 19-ї Міжнародної науково-технічної конференції SAIT 2017, Київ, 22 – 25 травня 2017 р. / ННК "ІПСА" НТУУ "КПІ ім. Ігоря Сікорського". – К.: ННК "ІПСА" НТУУ "КПІ ім. Ігоря Сікорського", 2017. – 340 с. – Текст: укр., рос., англ.

This book of abstracts includes issues connected with the research and development of complex systems of various nature in conditions of uncertainty and multifactor risks, Grid and high performance computing in science and education, intelligent systems for decision-making, progressive information technologies for needs of science, industry, economy, and environment. The problems of sustainable development and global threats estimation, forecast and foresight in tasks of planning and strategic decision making are investigated.

В сборнике рассматриваются вопросы, связанные с разработкой и исследованием сложных систем разной природы в условиях неопределенности и многофакторных рисков, Grid и систем высокопроизводительных вычислений в науке и образовании, интеллектуальных систем поддержки принятия решений, прогрессивных информационных технологий для потребностей науки, промышленности, экономики, окружающей среды. Исследуются вопросы устойчивого развития и оценивания глобальных угроз, прогноза и предвидения в задачах планирования и принятия стратегических решений на уровне регионов, больших городов, предприятий.

У збірнику розглядаються питання, що пов'язані з розробкою та дослідженням складних систем різної природи в умовах невизначеності та багатофакторних ризиків, нових інформаційних технологій, Grid і систем високопродуктивних обчислень в науці і освіті, інтелектуальних систем підтримки прийняття рішень, прогресивних інформаційних технологій для потреб науки, промисловості, економіки та навколишнього середовища. Досліджуються питання сталого розвитку та оцінювання глобальних загроз, прогнозу та передбачення в задачах планування та прийняття стратегічних рішень на рівні регіонів, великих міст, підприємств.

© Institute for Applied System Analysis at the Igor Sikorsky Kyiv Polytechnic Institute, 2017

ISBN 978-966-2748-93-2 (print)
ISBN 978-966-2748-94-9 (ebook)

http://sait.kpi.ua

Table of contents · Содержание · Зміст

Plenary talks · Пленарные доклады · Пленарні доповіді **11**
Lyubchyk L.M., Galuza A.A., Grinberg G.L. Ranking Model Real-Time Adaptation via Preference Learning Based on Dynamic Clustering 12
Згуровський М.З. Загальна концепція періодичного розвитку системних світових конфліктів 13
Кулик А.С. Диагностирование как средство уменьшения апостериорной неопределённости в функционировании объектов автоматического управления 14

Section 1. System analysis of complex systems of various nature
Секция 1. Системный анализ сложных систем разной природы
Секція 1. Системний аналіз складних систем різної природи **16**
Baranovska L.V. Method of resolving functions for the pursuit game with a pure time-lag . 18
Chugay A.M. Mathematical modelling of homothetic polytopes packing 19
Cojocaru Sv., Petic M., Albu V. Developing classifiers for information related to disaster topics 20
Gorban N.V., Khomenko O.V., Paliichuk L.S. Asymptotic behavior of solutions for climate energy balance model on manifold without boundary 22
Kaniovska I.Y., Zaitsev A.A. Implementation of Markov random processes for studying risks in credit portfolios 23
Nadtochii O.V., Bidyuk P.I. Modeling and forecasting demographic processes in Ukraine . . 24
Pankratova N.D., Kondratova L.P. The model for guaranteed functioning of complex engineering systems taking into account the margin of permissible risk 26
Pavlenko V.D., Salata D.V. Identification Eye-Motor System with Using Volterra Model . . 28
Soloviev V.N., Romanenko Y.V. Economic analog of Heisenberg uncertainly principle and financial crisis 32
Titchiev I.I. Scenario Analysis Related to the Field of Social Disasters by means of High Level Petri Nets 34
Безсмертна А.В. Особливості побудови рекомендаційної системи для користувачів телебачення 36
Бігун С.О., Іваненко Д.О. Реалізація алгоритму перевірки ефективності методу оцінювання в статистичних моделях, керованих шумом Леві 37
Білущак Ю.І., Гончарук В.Є, Чернуха О.Ю., Чучвара А.Є. Математичне моделювання дифузійних процесів у випадкових шаруватих структурах за явного врахування стрибків концентрації та її похідної на міжфазних границях 38
Бойко І.О., Бідюк П.І. Застосування скорингового аналізу для оцінювання платоспроможності клієнтів 40
Болдак А.О., Лазаренко І.С. Алгебраїчний підхід до верифікації моделей складних систем на прикладі побудови індексу нерозповсюдження зброї 41
Болдак Л.О., Горох О.С. Моделювання сценаріїв за допомогою баєсівських мереж довіри 42
Болжеларський Я.В., Дмитрук В.А., Гера Б.В. Математична модель динамічної поведінки рухомої фізико-механічної системи при дії зосереджених та розподілених сил опору рухові 43

Бондаренко Н.С. Огляд Skip-gram моделі нейронної мережі для технології Word2Vec . 45
Боровская Т.Н., Северилов П.В., Северилов В.А. Оптимальное агрегирование систем с произвольными ресурсными связями для задач переработки отходов 46
Буценко Ю.П., Лабжинський В.А. Модифікована процедура мажоритарного голосування для діагностування техногенно небезпечних систем 48
Васильев В.И., Виштальд Д.М., Любашенко Н.Д. Структурные функции в задачах оценки доступности сетевой услуги в классе бинарных стохастических моделей 49
Войтех Д.В. Визначення оптимальної кількості кластерів при кластеризації 51
Волкова В.Н., Ефремов А.А., Логинова А.В. Модели и технологии для проектирования автоматизированных информационных систем 53
Гайвась Б.І., Дмитрук В.А., Торський А.Р., Дмитрук А.А. Пакет програм для дослідження сушіння зерна пшениці в газозваженому стані за імпульсної дії сушильного агента 55
Горбань Н.В., Палійчук Л.С., Хоменко О.В. Асимптотична поведінка розв'язків неавтономних дисипативних еволюційних задач в нескінченновимірних просторах 57
Горелова Г.В. Анализ вариантов когнитивной модели рынка труда 58
Городецкий В.Г., Кулиш Р.Д. Алгоритм идентификации систем с периодическим внешним воздействием 61
Дідковський О.І. Модель побудови датасету для вирішення задачі оптимізації енергоспоживання 62
Донченко В.С. Синтез двох фундаментальних підходів розв'язання задачі лінійної дискримінації 64
Донченко В.С., Зінько Т.П. SVD та псевдообернення в обробці сигналів: векторні та матричні представники 65
Донченко В.С., Назарага І.М., Тарасова О.В. Матрична множинна регресія для прогнозування макроекономічних та медійних показників: приклади використання . . . 66
Дорофеев Ю.И., Любчик Л.М., Никульченко А.А. Дескрипторный подход к синтезу управления запасами в цепях поставок с неопределенными временными задержками 67
Ємельянова Д.І. Інформаційно-програмне забезпечення комплексної оцінки рівня екологічної безпеки системних об'єктів 68
Єремеєв І.С., Єщенко О.І. Про один підхід до аналізу систем теплопостачання 70
Заводник В.В., Ноголь О.С. Оценивание макроэкономической динамики страны на базе статистики межотраслевого баланса и когнитивного моделирования 72
Захарчук Н.Г., Ладогубець Т.С., Фіногенов О.Д. Визначення індексу випадкової узгодженості в програмному забезпеченні, що реалізує MAI 74
Кангун Г.І., Бідюк П.І. Моделювання військових конфліктів методами інтелектуального аналізу даних 76
Каніщева О.В., Главчева Ю.М., Висоцька В.А. Визначення стилю автора для виявлення плагіату в академічному середовищі 78
Кобера В.В., Каніовська І.Ю. Застосування реальних опціонів в управлінні ризиками інвестиційних проектів 80
Козуля М.М., Моїсеєв В.Ф. Програмна підтримка ентропійно-інформаційної оцінки відповідності стану складних систем 81
Козуля Т.В., Білова М.О. Графологічні моделі з аналізу взаємодії «системний об'єкт – навколишнє середовище» при вирішенні завдань безпеки 83
Коршевнюк Л.О. Ранжирування ризиків в задачах ризик-менеджменту 85
Куценко А.С., Товажнянский В.И. Упрощенная математическая модель тепловых процессов зданий 87
Лопатин А.К., Курач Я.С. Устойчивые аттракторы в динамической модели ценообразования Вальраса с линейной функцией спроса и функцией предложения в виде логистического тренда Ферхюльста 88

Лопатин А.К., Черненко О.Б. Динамическая модель ценообразования с функцией предложения и с функцией спроса в виде растущего и убывающего логистических трендов Ферхюльста . . . 90
Лубинець Ю.М., Федюков В.Д., Каніовська І.Ю. Байєсівське усереднення для оцінки параметрів линійної регресії . . . 92
Минько О.В., Джигирей І.М. Аналіз програмних пакетів підтримки оцінювання життєвого циклу продукційних систем . . . 93
Морозов К.В., Романкевич В.А., Сапсай Т.Г. О модификации графо-логической модели для систем и их компонентов с множественными состояниями . . . 94
Недашківська Н.І. Порівняльний аналіз методів розрахунку пріоритетів альтернатив рішень на основі нечітких експертних оцінок парних порівнянь . . . 95
Новицький Є.А., Шубенкова І.А. Оцінка ризиків проектів за умови відсутності повної інформації . . . 96
Опарина Е.Л. Поиск рационального компромисса требований при создании сложных технических объектов . . . 97
Панкратов В.А. Стратегия инновационного развития социально-экономических систем 98
Панкратова Н.Д., Демьянчук А.З., Мерзляков А.Д. Применение полиномов Якоби при восстановлении функциональных зависимостей . . . 99
Пархоменко А.В., Путренко В.В. Аналіз регіональних збройних конфліктів з використанням методів просторової кластеризації . . . 101
Пашинська Н.М., Путренко В.В. Просторово-часове моделювання забруднення повітря внаслідок аварій на хімічних об'єктах з використанням інструментарію ГІС . . . 102
Писарчук О.О., Лагодний О.В. Статистичний аналіз часових рядів експериментальних даних активності тематичного контенту в інтернет-середовищі . . . 103
Пичугина О.С. Оптимизация на евклидовом комбинаторном множестве $E_{nk}^{\pm}(G)$ . . . 105
Піддубна Ю.І., Пишнограєв І.О. Оцінювання ефективності та інвестиційної привабливості стартапів . . . 107
Притоманова О.М., Журавель С.В., Киселев М.Я. Управление проектами в IT-сфере методами PMI PMBOK на основе нейронечетких технологий . . . 109
Рак О.О. Побудова мережі маршрутів громадського транспорту як задача багатокритеріальної оптимізації . . . 110
Романенко В.Д., Мілявський Ю.Л., Мокін В.Б. Автоматизація управління водним балансом водогосподарської ділянки з нейтралізацією обмежених природних збурень на основі когнітивних моделей . . . 111
Рудакова А.В., Терновая Т.И., Олокоба О.Й. Методы предварительной обработки сигналов в автоматизированной системе диагностики для функционального контроля электромеханического оборудования . . . 112
Савченко А.В. Інформаційно-експериментальна підтримка процесів моделювання екоефективного згоряння палива в дизелі . . . 113
Савченко І.О. Оцінювання чутливості двохетапної процедури модифікованого методу морфологічного аналізу . . . 115
Северилов П.В., Боровская Т.Н., Колесник И.С., Гришин Д.И. Оптимизация производственных систем с учетом ресурсных обратных связей . . . 116
Сербул Р.С., Данилов В.Я. Системний підхід в управлінні IT-проектами . . . 118
Смирнов С.А., Коваль Л.В. Встановлення вагових коефіцієнтів критеріїв за допомогою порядкових експертних оцінок . . . 120
Статкевич В.М. Регулярные выражения для языков сети Петри "производитель/потребитель" с ограниченным буфером размера 1 и 2 . . . 122
Стефанишин Д.В. Про врахування ризику при формуванні оптимальної стратегії введення нових агрегатів на гідроакумулюючих електростанціях в Україні . . . 123

Стефанишин Д.В., Стефанишина-Гаврилюк Ю.Д. Про оптимізацію структури інвестиційного портфеля в природокористуванні на підставі попарного порівняння альтернатив з врахуванням ризику невикористаних можливостей . . . 124
Ткаченко В.Ю., Ткаченко Н.В., Жданова О.Г., Сперкач М.О. Задача мінімізації сумарного запізнення розкладу виконання завдань зі спільним директивним терміном паралельними пристроями . . . 125
Харчук В.В., Войтко С.В. Управління фінансовими потоками в кластерних структурах 127
Ходневич Я.В., Стефанишин Д.В., Корбутяк В.М. Про геоінформаційний аналіз та імітаційне моделювання взаємодії потоку і русла при обтіканні донних гряд з метою прогнозування місцевих розмивів берегів рік . . . 128
Шемеляк Р.О., Бідюк П.І. Системний підхід до моделювання фінансово-економічних процесів . . . 129

Section 2. Intelligent systems for decision-making
Секция 2. Интеллектуальные системы принятия решений
Секція 2. Інтелектуальні системи прийняття рішень **132**
Iievlieva S.M., Garev Y.V., Pikhota S.R. Development of intellectual audio-player based on spectral sound analysis . . . 134
Kuksa M.U., Norkin V.I. Numerical calculation of a distribution for the asset price minimum under the binomial asset price model . . . 136
Lazarenko D.V., Lyunda V.V. Facial Landamark Detection . . . 137
Naderan M., Zaychenko Y.P. Using neural network anfis and cascade for predicting futures contract prices . . . 139
Solodrai I.I., Marchenko O.I. Technique for combining neuroevolution and temporal-difference learning . . . 140
Zaychenko Y.P., Zaychenko E.Yu., Ovi Nafas Aghaei Agh Ghamish Survivability Analysis of New Generation Computer Networks (NGN) . . . 141
Zorin Yu. On Swarm Algorithms with controlled Levy Flight . . . 142
Архіпов Д.В. Прогнозування ходу виконання проекту як елемент управління проектами 144
Бідюк П.І., Крамаренко Ю.Ю. Прогнозування розвитку демографічних процесів в Україні 146
Буткалюк С.В. Система оптичного розпізнавання нотного запису . . . 147
Вішталь Д.М., Васильєв В.І., Любашенко Н.Д. Подання і застосування категорії часу (Date/Time) в технології публікації і споживання веб-даних . . . 149
Галатенко Д.В. Використання нейронних мереж для розпізнавання нот піаніно . . . 151
Гожий В.В., Чертов О.Р. Автоматичне визначення психологічного портрету особистості у соціальних мережах . . . 152
Горбань Г.В. Генерація асоціативних правил у багатомірних даних у межах одного виміру 154
Давидюк О.Ю. Вибір ефективних методів кластеризації та класифікації мережевої активності для набору даних KDD CUP 99 . . . 156
Єремеєв І.С., Дичко А.О. Використання методів нечітких множин у процесах біологічного очищення стічних вод . . . 158
Забелин С.И. Геометрические методы анализа Big Data . . . 159
Зайченко Ю.П., Гамидов Г.И. Каскадные нейро-нечеткие сети в задачах прогнозирования на рынках ценных бумаг . . . 160
Зайченко Ю.П., Гула І.С. Прогнозування ризику банкрутства банків в умовах невизначеності . . . 161
Заковоротний О.Ю., Харченко А.О. Нейромережеве прогнозування швидкості в системах підтримки прийняття рішень . . . 162
Зражевська Н.Г. Моделювання і прогнозування мір ризику VaR і CVaR для часового ряду фондових індексів . . . 163

Коляденко А.В. Система підтримки прийняття рішень комплексного обслуговання колцентру . . . 164
Кондратенко Н.Р., Снігур О.О. Моделювання багатовимірних залежностей на основі нечітких множин типу 2 на недовизначених гідрогеологічних даних . . . 166
Кравець І.О. Дослідження чітких та нечітких нейронних мереж для ідентифікації та керування слабоформалізованими об'єктами . . . 168
Куца К.В. Анализ кредитоспособности малого и среднего бизнеса в Украине с использованием нечеткой логики . . . 169
Мурга М.О. Побудова обгортки процесу в просторі зі змінним базисом приналежностей ознакам нестаціонарних процесів . . . 170
Мысник Б.В., Снитюк В.Е. Мультиагентная оптимизация функционирования предприятий отрасли. Восстановление пропусков в данных . . . 171
Перова И.Г., Бодянский Е.В. Об одной метрике в задачах интеллектуального анализа медицинских данных . . . 173
Поворознюк А.И., Филатова А.Е. Полутоновая морфологическая фильтрация на основе локальных статистик . . . 174
Поворознюк Н.І., Гритчук Д.Т. Вейвлет-аналіз сигналів для моніторингу якості електроенергії . . . 176
Савельев Ю.Д. Анализ методов распознавания лиц с помощью нейронных сетей . . . 177
Сащук І.М., Чепков І.Б. Методика визначення інформативних ознак розпізнавання стану об'єктів моніторингу . . . 180
Скляр А.В. Сравнение интеллектуальных систем принятия решений в играх с неполной информацией . . . 182
Соловйов С.О., Якименко І.В., Кир'яченко С.П. Моделювання біохімічного резерву організму при лікуванні хворих на ІХС . . . 185
Суржан А.О., Бідюк П.І. Оцінювання інвестиційних ризиків в стартап-проектах . . . 187
Таранян А.Р. Ускорение локализации высококонтрастных движущихся объектов в видеопотоке . . . 188
Тимошенко Ю.О., Ніколаєв С.С., Ходак А.Р. Порівняння моделей на основі нейронних мереж для задачі машинного розуміння тексту . . . 190
Чапалюк Б.В., Зайченко Ю.П. Аналіз алгоритмів детектування обличчя . . . 191
Шаповалова С.І., Мажара О.О., Москаленко Ю.В. Екстракція правил виведення нейронних мереж з представленням в форматі CLIPS . . . 192
Шулькевич Т.В., Селін О.М., Селін Ю.М. Математичний апарат інтелектуального аналізу даних для прогнозування нелінійних нестаціонарних процесів різної природи 194
Щеголь Ю.М., Бідюк П.І. Система підтримки прийняття рішень для прогнозування нестаціонарних процесів в екології . . . 195

Section 3. High Performance Computing and Microsystems Engineering
Секция 3. Высокопроизводительные вычисления и разработка микросистем
Секція 3. Високопродуктивні обчислення та розробка мікросистем **196**
Artuhov V.G., Brytov O.A. Possibility of sigma-delta modulator noise transfer function zeros optimization . . . 198
Kachko N.A. Genetic algorithms over distributed systems with MapReduce model . . . 199
Kucher V.O. High performance RDMA-based Dataflow Management System . . . 201
Pechurin N.K., Kondratova L.P., Pechurin S.N. Reliable balanced UAV computing cluster and V.Leontyev model . . . 202
Yaremenko V.S. An approach for data clustering CURE algorithm implementation using the MapReduce technology . . . 204
Гіоргізова-Гай В.Ш., Кульбачний К.А. Організація обліку електронних медичних карт на базі хмарних сервісів Amazon AWS . . . 205

Єршов С.В., Пономаренко Р.М. Паралельні обчислення багаторівневих нечітких систем Такагі–Сугено . . . 207
Капшук О.О. Реалізація біометричних технологій аутентифікації на мобільних пристроях 209
Кирюша Б.А., Колінько А.М. Порівняльна характеристика бюджетних embedded рішень для обробки відеопотоку . . . 211
Марченко О.О., Марченко О.І. Модель динамічного розпаралелення пошуку в дереві методом Монте-Карло для grid-систем . . . 213
Орлова М.М., Телелейко І.С. Сумісність підходів хмарних обчислень та грід-технологій 215
Петренко А.І. Нова парадигма "Ontology-based Software Engineering": перехід від об'єктно-орієнтованого до процесно-орієнтованого програмування . . . 216
Романов В.В., Ткаченко В.А. Особенности проектирования преобразователей "Напряжение – Частота" . . . 218
Сергеєв Є.І., Романов В.В. Вибір складової фреймворку Xamarin при розробці кросплатформних додатків . . . 220
Шевченко Є.Т., Ліщук К.І. Балансування навантаження між кластером серверів з урахуванням типу задач . . . 222

Section 4. Progressive information technologies
Секция 4. Прогрессивные информационные технологии
Секція 4. Прогресивні інформаційні технології **224**
Basyuk T.M. Approach towards generating a software system of website popularizing . . . 226
Chereda H., Nikolaiev S. Comparison of resampling techniques and its effect on real time filtration of non-uniformly sampled bio-signal time series . . . 228
Didenko D.G. Fuzzy recovery database with implementation in Java . . . 230
Gorodko N.A., Boyarinova Yu.E. Using hyper complex numbers for modeling the quality management system of information services . . . 231
Kalinovsky Ya.A., Boyarinova Yu.E., Hitsko Ya.V., Sukalo A.S. Package hypercomplex symbolic computations in Maple environment . . . 232
Kryvolap D.V., Bidyuk P.I. Modeling and forecasting Ukrainian economy in transition . . . 233
Naumenko T.O. The analysis of fields of using ontologies in the construction of information systems . . . 234
Samuel Adegoke On implementing CARMA overlay in VOIP/SIP . . . 236
Trofimov A.S., Zenkovich M.V. Mobile secure runtime . . . 240
Антонюк С.А. Віртуальна реальність: новий погляд на web . . . 242
Арчвадзе Н.Н., Пховелишвили М.Г., Шецирули Л.Д. Об одном алгоритме параллельной сортировки . . . 244
Бакун С.А., Литвинюк А.А. Прогнозування врожайності сільськогосподарських культур з використанням технологій Machine Learning . . . 245
Бакурова А.В., Ропало А.М. Анализ существующих информационных систем в управлении аптечными сетями . . . 247
Бидниченко А.А., Безносик А.Ю. Использование объектной модели в ядрах современных Unix-подобных операционных систем . . . 251
Бідюк П.І., Терентьєв О.М., Просянкіна-Жарова Т.І., Савастьянов В.В. Застосування інструментів SAS BASE для дослідження ефективності методів обробки пропусків у вибірках даних з метою підвищення якості прогнозування показників продовольчої безпеки країни . . . 253
Біла А.І., Бідюк П.І. Побудова моделей гетероскедастичних процесів . . . 255
Бойко Р.С. Комп'ютерне бачення: метод розпізнавання джерел вогню та ознаки появи вогню в відеокадрі . . . 257
Вельбицкий И.В. Программирование без языков программирования (новая графическая полиглот-концепция) . . . 259

Виклюк Я.І., Гусак О.М. Інформаційна технологія раннього виявлення лісових пожеж . 261
Ворона М.В., Лазаренко И.С. Механизмы повышения успешности управления проектами по созданию WEB-сайтов . . . 262
Гнатенко В.Ю., Ситников В.С., Ступень П.В. Метод получения полиномов Фаулхабера для интерполяции сигналов . . . 264
Губар Б.Д., Олійник Ю.О. Архітектура рекомендаційної системи з вибору продуктів харчування . . . 266
Губарєв В.Ф., Комісаренко В.І. Визначення параметрів відносного руху космічних апаратів на основі аналізу зображень за допомогою інформаційних ознак . . . 267
Джамалов А.Т., Рагимов Р.М. Аэрокосмический мониторинг обеспечения экологической безопасности транспортировки энергоносителей по магистральным трубопроводам . . . 268
Залевська О.В., Котолупов О.В. Дослідження властивостей тривимірних клітинних автоматів . . . 273
Замекула О.І., Заболотня Т.М. Розроблення програмної системи для створення та аналізу результатів онлайн-опитувань . . . 274
Згуровський М.З., Павлов О.А. Методологічні основи побудови ПДС-алгоритмів для важкорозв'язуваних задач комбінаторної оптимізації . . . 275
Землянський Д.І., Ліщук К.І. Спосіб побудови структури бази даних для динамічно змінюваних об'єктів . . . 277
Зубок В.В. Виявлення мережевих атак на комп'ютерні системи за допомогою нейронних мереж . . . 278
Калинин В.В., Бояринова Ю.Е. Клеточные автоматы . . . 280
Касумова Т.А., Пашаева М.М., Джамалова А.Г., Аралис С.В. Роль информационной культуры в создании и развитии информационного космического общества . . . 281
Киричек Г.Г., Овчар А.О. Захист передачі мережних даних . . . 283
Кіріленко О.Г., Труш Г.О. Систематичний огляд екологічного програмного забезпечення 285
Копп А.М., Орловский Д.Л. Континуум предприятия как составная часть знание-ориентированного подхода к управлению предприятием . . . 287
Корнієнко О.О. Алгоритм апроксимації вейвлет-пакетного дерева декомпозиції . . . 289
Куцаченко Н.Г. Оценка эффективности реализаций алгоритмов трекинга движения средствами OpenCV . . . 290
Лавренюк А.М., Лавренюк С.І., Тульчинський П.Г. Оптимізація архітектури програмного забезпечення для аналізу великих даних базі мови програмування python . . . 291
Лисенко А.В. Порівняльний аналіз методів трекінгу об'єктів на відеоряді . . . 293
Лілікович С.О., Лютенко І.В. Суб'єктно-орієнтований підхід до забезпечення диверсності N-версійного програмного проекту . . . 295
Ляхевич О.В. Порівняльний аналіз методів відновлення зображення . . . 297
Максим К.Є. Дослідження сучасних методів стеганоаналізу . . . 298
Малишев А.І. Синхронізація анімації обличчя віртуального персонажа із голосом в реальному часі . . . 299
Мегрелишвили Р.П., Шенгелия С.В. Исследование и синтез матричных структур . . . 301
Москаленко В.В., Миронюк И.А. Построение иерархии ключевых показателей деятельности предприятия в рамках информационной системы управления эффективностью 303
Муха И.П., Вольский А.В. Анализ и оптимизация структуры сайта . . . 305
Мухін В.Є., Корнага Я.І., Яковлева А.П., Герасименко О.Ю., Козерацький М.С. Комплексний аналіз часу обробки запитів в розподілених базах даних . . . 306
Негода К.В., Чертов О.Р. Система ідентифікації психологічного стану дитини за її малюнками . . . 307
Нестюк О.М., Ліщук К.І. Технічна підтримка продукту з використанням автоматизованого пошуку відповідей за змістом та чат-ботів . . . 308

Орєхов О.А., Орєхова Н.А. Дослідження властивостей бібліотеки алгоритмів глибокого навчання Keras на прикладі задачі компрометації нейронної мережі 309
Осиюк Д.В., Шубенкова И.А. Сквозная аналитика для проектов сферы электронной коммерции на основе Google Analytics 310
Павлов О.А., Місюра О.Б., Мельников О.В., Лисецький Т.М. Система календарного та оперативного планування і прийняття рішень для підприємств дискретного типу 312
Погорелов В.В., Марченко О.І. Використання графа залежностей значень і станів для трансляції програм 314
Попеляев Д.П. Подходы к распознаванию огня на изображении. Сверточные нейронные сети 315
Прогонов Д.О. Вплив невідповідності областей приховання повідомлень та проведення стегоаналізу на ефективність статистичних стегодетекторів 317
Продан А.О. Дослідження законів Ципфа з метою визначення унікальності автора за формулою 319
Ревва Р.В. Подготовка интерактивных отчётов с использованием визуальной аналитики для последующего упрощения процесса принятия решений в бизнесе 320
Сергєєв Д.С. Виділення концептів у природно-мовному тексті як спосіб наповнення бази знань 321
Сергєєв А.В. Особливості проектування гібридних систем доставки контенту 322
Сибіряков П.А. Система автоматизації генерування тестів 323
Скукіс О.Є. Застосування інформаційних технологій для розв'язання задач транспортних перевезень 324
Сухоручкина О.Н., Прогонный Н.В. Информационные технологии интеллектуального управления мобильным роботом и их использование в учебных целях 325
Ткаченко К.П., Шубенкова І.А. Створення баз даних для медичних закладів у паліативній медицині 327
Федь В.В. Порівняльний аналіз криптографічних методів 328
Хает Ф.И., Алфимцев А.Н., Девятков В.В. Гендерные поведенческие модели работы с пользовательскими интерфейсами различных типов устройств 329
Хіміч А.В. Моделювання процесу отримання знань із тексту 330
Хурцилава К.В., Литвинов В.А., Майстренко С.Я. Про критерії вибору моделі представлення просторово-часових геоданих в ГІС лісового господарства 332
Шубенкова И.А., Клименко Ф.Т. Проект управления созданием программного кода с использованием систем контроля версий 334
Щука Р.В., Іванов С.С., Терентьєв О.М., Орешко В.А., Яворський Я.Б., Кузенко А.Ю. Побудова предиктивної моделі прогнозування злочинності на основі методу подібних траєкторій в системі SAS 335

Authors · Авторы · Автори **337**

Plenary talks

System analysis of complex systems of various nature

Intelligent systems for decision-making

High Performance Computing and Microsystems Engineering

Progressive information technologies

Plenary talks

Lyubchyk L.M., Galuza A.A., Grinberg G.L.
National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

Ranking Model Real-Time Adaptation via Preference Learning Based on Dynamic Clustering

Learning to rank has become an important research topic in many information retrieval problems, such document retrieval, collaborative filtering, search engine modeling, e-commerce and other internet-related applications [1]. A typical example of learning to rank problem is a learning model adaptation for an unknown ranking mechanism of search engine based on the information about the search results of the ranked set of documents that match a particular sequence of queries [2]. So the training data consists of queries and ranked sequence of documents characterized by their features. One of the most effective approaches to solving these problems is preference learning, based on the recovery of a scalar continuous function in the objects feature space, known as a function of preferences or utility function, which is used to rank objects considered. It is assumed that a certain object is preferable to another if its feature vector corresponds to a larger value of preference function. Preference function can be considered as a model of unknown ranking function, describes, for example, search engine operation mechanism or user preferences. The problem of preference function recovery is extensively studied in relation to the development of multi-criteria decision-making methods [3]. In practice frequently used a simple model of preference functions as a linear combination of features with weights that determine their relative importance, but linear model not always adequately reflects the complex nature of preferences, arising in real ranking systems. Various approaches to the preference function nonlinear model construction are considered, at that the most difficult is the problem of the choice of its structure ensuring the possibility of preference high accuracy of approximation with restricted complexity of the model. A perspective method of preferences function identification used kernel-based machine learning technique with regularization based on expert information about the features weights [4]. At that the identified model may be represented in a nonparametric form as linear combination of kernel function, though the weighting coefficients may be computed without making direct reference to high-dimensional vectors of model coordinate functions. The difficulty of solving the problem of real time ranking learning upon receipt of the current information in data stream scheme is that the relative rank of objects with similar feature vectors in different data series may differ significantly. In this regard, a new approach to preference ranking learning is proposed, based on pre-clustering of the objects followed by preference function identification using the information of average ranks of objects belonging to each of the clusters [5]. Thus the problem is reduced to preference learning to rank on clusters in the feature space of ranked objects, while aggregated training dataset is formed from the centers of clusters and estimates of the average rank of the objects from the cluster. The proposed preference learning on clusters method allows to fully realizing the advantages of the kernel-based approach. While the dimension of the model is determined by a pre-selected number of clusters and its complexity do not grow with increasing number of observations. Thus real-time preference function identification algorithm based on training data stream includes successive estimates of cluster parameter as well as average cluster ranks updating and recurrent kernel-based nonparametric estimation of preference model [5].

References. **1.** T.Y. Liu. Learning to Rank for Information Retrieval. Foundations and Trends in Information Retrieval, Vol. 3, No.3, pp. 225–331, 2009. **2.** L. Hang, A. Short. Introduction to Learning to Rank. IEICE Trans. Inf. and Syst., Vol. E94-D, No.10, 2011. **3.** J. Figueira, S. Greco, M. Ehrgott. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, Switzerland, 2005. **4.** L. Lyubchyk, G. Grinberg. Preference Function Reconstruction for Multiple Criteria Decision Making Based on Machine Learning Approach. In "Recent Developments and New Directions in Soft Computing", L.A. Zadeh et al. (eds), Springer, Switzerland, 2014, pp. 53–63. **5.** L. Lyubchyk, G. Grinberg. Real Time Recursive Preference Learning to Rank from Data Stream. The 1th IEEE Int. Conf. on Data Stream Mining and Processing, 23–27 August 2016, Lviv, Ukraine.

Згуровський М.З.
КПІ ім. Ігоря Сікорського, Київ, Україна

Загальна концепція періодичного розвитку системних світових конфліктів

Вивчення глобального еволюційного процесу розвитку цивілізації як складної, цілісної системи, що самоорганізується, передбачає врахування цілого ряду взаємопов'язаних процесів і чинників різної природи, серед яких одне з центральних місць займають світові конфлікти. Пошук і побудова загальних моделей, що адекватно описують закономірності перебігу світових конфліктів, була і залишається однією з найактуальніших, але невирішених проблем, що стоять перед наукою. Незважаючи на численні спроби, до цього часу так і не запропоновано адекватного науково-обґрунтованого метричного інструментарію вирішення завдань глобального прогнозування та передбачення розвитку світових конфліктів, що вкрай важливо в епоху входження світової цивілізації в XXI століття, як особливу критичну фазу свого розвитку.

У контексті пошуку шляхів вирішення даної проблеми, ґрунтуючись на системному узагальненні результатів дослідження великого емпіричного матеріалу [1–3], в роботах [4–6], виявлено і обґрунтовано гіпотетичну закономірність перебігу системних світових конфліктів, що базується на так званій динамічної моделі С-хвиль, в основі якої лежить метрика "золотого перетину".

В даному дослідженні виконано узагальнення та формалізацію підходів, розглянутих в роботах [4–6] і запропонована загальна концепція С-хвиль системних світових конфліктів. При цьому особливо виділено клас великих С-хвиль, що охоплюють надвеликі тимчасові інтервали, закономірність плину яких інваріантна до еволюції самої природи світових конфліктів.

Обґрунтовано, що якщо людство змінить парадигму "необмеженого споживання" в планетарному масштабі, наприклад, на парадигму "гармонійного співіснування і розумного споживання" і докорінно перетворить свою техносферу на "природоподібну" (дружню до середовища проживання людини, засновану на конвергенції нано-, біо-, інформаційних, когнітивних і соціогуманітарних технологій (НБІКС технологій)), то в майбутньому воно забезпечить сталий розвиток (sustainability) середовища свого проживання і продовжить творчу місію на планеті Земля.

Література. **1.** Яковец Ю.В. Прогнозирование циклов и кризисов. М.: МФК, 2000. – 426 с. **2.** Назаретян А.П. Цивилизационные кризисы в контексте Универсальной истории. М.: "Мир". 2004. – 367 с. **3.** List of wars: http://en.wikipedia.org/wiki/List_of_wars. **4.** Згуровский М.З. Закономерность течения системных мировых конфликтов и глобальные угрозы XXI столетия // Кибернетика и системный анализ. – 2007. – №5. – С. 87–99. **5.** Взаимосвязь больших кондратьевских циклов и системных мировых конфликтов // Кибернетика и системный анализ. – 2009. – №5. – С. 84–92. **6.** Згуровский М.З. Метрические аспекты периодических процессов в экономике и обществе // Кибернетика и системный анализ. – 2010. – №2. – С. 3–9.

Кулик А.С.
Национальный аэрокосмический университет им. Н.Е.Жуковского "Харьковский авиационный институт", Харьков, Украина

Диагностирование как средство уменьшения апостериорной неопределённости в функционировании объектов автоматического управления

Детерминистский и вероятностный принципы моделирования не позволяют обеспечить требуемую адекватность математических моделей в задачах автоматического управления современными и перспективными объектами. Качество моделей можно существенно повысить с использованием нового принципа нечёткости, позволяющего описывать неопределённости условий функционирования.

Неуклонное повышение конструктивной сложности, уровня многообразия выполняемых функций, масштабности и энергоёмкости автоматически управляемых динамических объектов обуславливает увеличение множества неопределённых факторов, оказывающих дестабилизирующее влияние на функционирование объектов автоматического управления, включающих объект управления, приводы и датчики.

Дестабилизирующие воздействия – это различные возмущения и помехи, а также неисправности, сбои и отказы. Неопределённость, возникающую в процессе функционирования, можно уменьшить посредством последовательного решения таких задач диагностирования как: 1) обнаружения дестабилизации, 2) локализации мест дестабилизации, 3) определения класса дестабилизации, 4) установления вида дестабилизации.

Решение задач автоматического диагностирования базируется на математических моделях. Наиболее адекватными представляются диагностические функциональные модели с нечёткими, интервально-размытыми значениями параметров [1]. В линейном приближении простейшая диагностическая модель для объекта диагностирования описывается в дискретном пространстве состояний следующей системой векторно-матричных уравнений:

$$\triangle x[(k+1)T_0] = G\triangle x(kT_0) + [A_i\widehat{x}(kT_0) + B_i u(kT_0)]\triangle\lambda_i; \triangle x(k_0T_0) = \widetilde{x}_0;$$
$$\triangle y(kT_0) = C\triangle x(kT_0) + [C_i\widehat{x}(kT_0) + D_i u(kT_0)]\triangle\lambda_i,$$

здесь $\triangle x(kT_0)$ – n-мерное отклонение вектора состояния объекта диагностирования от эталонного поведения, $\triangle x(kT_0) \in X^n$; $G - (n \times n)$-мерная матрица динамики; $C - (m \times n)$-мерная матрица выхода объекта; $\widehat{x}(kT_0)$ – n-мерный вектор состояния эталонной модели, $\widehat{x}(kT_0) \in X^n$; $u(kT_0)$ – r-мерный вектор управляющих воздействий, $u(kT_0) \in U^r$; A_i, B_i, C_i и D_i – матрицы соответствующих размерностей, отражающие чувствительность матриц номинальной модели объекта диагностирования к i-му прямому диагностическому признаку λ_i; $\triangle\lambda_i$ – отклонение прямого диагностического признака от номинального значения $\triangle\lambda_i \in \langle\triangle\lambda_i\rangle$, $\langle\triangle\lambda_i\rangle$ – интервальное число с нечётко-определёнными границами; $\triangle y(kT_0)$ – m-мерный вектор косвенных диагностических признаков, $\triangle y(kT_0) \in Y^m$.

Для каждой задачи диагностирования формируется диагностическая функциональная модель, отражающая специфику диагностирования конкретного дестабилизирующего воздействия. Диагностические функциональные модели отражают связь косвенных признаков, доступных измерению, с прямыми, недоступными измерению. Для получения оценочных значений прямых признаков дестабилизирующих воздействий необходимо использовать идентификационные методы к уравнениям диагностических функциональных моделей.

Результаты решения задач диагностирования позволяют сформировать базу знаний и структуру автоматического диагностического вывода для получения полного диагноза причин апостериорной дестабилизации объектов автоматического управления в реальном масштабе времени.

Литература. 1. Кулик А.С. Элементы теории рационального управления объектами. – Харьков: Нац. аэрокосм. ун-т им. Н.Е. Жуковского «ХАИ», 2016. – 255 с.

This page was left blank intentionally

Plenary talks

System analysis of complex systems of various nature

Intelligent systems for decision-making

High Performance Computing and Microsystems Engineering

Progressive information technologies

1 System analysis of complex systems of various nature

Section 1

System analysis of complex systems of various nature

1. System analysis methods for complex systems of various nature in conditions of uncertainty and risks.
2. Mathematical methods, models and technologies for complex systems' research.
3. Technology foresight system methodology in problems of planning and strategic decisions' making.
4. Theory and methods of optimal decision-making.
5. Problem oriented methods of complex systems' analysis and designing in conditions of uncertainty and risks.
6. Nonlinear problems of system analysis.
7. System methodology of sustainable development.
8. Text analytics and natural language processing.

Секция 1

Системный анализ сложных систем разной природы

1. Методы системного анализа сложных систем разной природы в условиях неопределенности и рисков.
2. Математические методы, модели и технологии исследования сложных систем.
3. Системная методология технологического предвидения в задачах планирования и принятия стратегических решений.
4. Теория и методы принятия оптимальных решений.
5. Проблемно-ориентированные методы анализа и проектирования сложных систем в условиях неопределенности и рисков.
6. Нелинейные задачи системного анализа.
7. Системная методология устойчивого развития.
8. Текстовая аналитика и обработка естественного языка.

Секція 1

Системний аналіз складних систем різної природи

1. Методи системного аналізу складних систем різної природи в умовах невизначеності та ризиків.
2. Математичні методи, моделі та технології дослідження складних систем.
3. Системна методологія технологічного передбачення в задачах планування та прийняття стратегічних рішень.
4. Теорія та методи прийняття оптимальних рішень.
5. Проблемно-орієнтовані методи аналізу та проектування складних систем за умов невизначеності та ризиків.
6. Нелінійні задачі системного аналізу.
7. Системна методологія сталого розвитку.
8. Текстова аналітика та обробка природної мови.

Baranovska L.V.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Method of resolving functions for the pursuit game with a pure time-lag

Consider a controlled object whose dynamics is described by the linear pure time-lag differential system in an Euclidean space R^n:

$$\dot{z}\,(t) = Bz\,(t-\tau) + \phi\,(u\,,\ v)\,, \qquad z \in \mathrm{R}^n\,,\ u \in U\,,\ v \in V\,, \tag{1}$$

where B is a square constant matrix of order n; U and V are nonempty compact sets; the function $\phi\,(u\,,\ v)\,,\ \phi : U \times V \to \mathrm{R}^n$, is jointly continuous function of its variables, u and v are the control parameters of the pursuer and the evader chosen from the control domains U and V, respectively; z is the state vector, involving geometric coordinates, velocities, accelerations of the pursuer and the evader; $\tau = const > 0$.

The initial condition

$$z\,(t) = z^0\,(t)\,, \qquad -\tau \leqslant t \leqslant 0, \tag{2}$$

is absolutely continuous function.

The terminal set M^* is a cylinder:

$$M^* = M_0 + M, \tag{3}$$

where M_0 is a linear subspace of the space R^n, and M is a compact set in the orthogonal complement L to M_0 in R^n.

The players choose their controls in the form of certain functions [1]. In this way the players affect the process (1) pursuing their own goals. The goal of the pursuer is to bring a trajectory of the process to the terminal set M^* (3) in the shortest time. The goal of the evader is to avoid the meeting of the trajectory of the process with this set on a whole semi-infinite interval of time or if it's impossible to maximally postpone the moment of meeting.

We consider the local problem of approach within fixed interval of time [1]. If the controls are chosen in the form of Lebesgue measurable functions $u\,(s)$ and $v\,(s)$ [2], which take its values in sets U and V, respectively, then a solution of the system (1) with the initial condition (2) can be presented by the formula:

$$z\,(t) = exp_\tau\{B,t\}z^0\,(-\tau) + \int_{-\tau}^{0} exp_\tau\{B,t-\tau-s\}z'^0\,(s)\,ds + \int_0^t exp_\tau\{B,t-\tau-s\}\phi\,(u,v)\,ds,$$

where $exp_\tau\{B,t\}$ is the time-lag exponential [3], $t \geqslant 0$.

The game is evolving on the closed time interval $[0;T]$, where T, $T > 0$, is the end point of game, and it is such that $z\,(T) \in M^*$ or $\pi z\,(T) \in M$. Here π is the orthogonal projector from R^n onto the subspace L.

Consider the multivalued mappings

$$W\,(t,v) = \pi exp_\tau\{B,t\}\phi\,(U,v)\,, \qquad W\,(t) = \bigcap_{v\in V} W\,(t,v)\,.$$

Pontryagin's condition. The mapping $W\,(t) \neq \emptyset$ for all $t \geqslant 0$.

For the game, described above, satisfying Pontryagin's condition, the scheme of the method of resolving functions is developed. The guaranteed time of the game termination is found, and corresponding control law is constructed.

References. **1.** Chikrii, A.A. (2013). Conflict-Controlled Processes. Springer Science & Business Media. **2.** Baranovskaya, G.G., & Baranovskaya, L.V. (1997). Group Pursuit in Quasilinear Differential-Difference Games. Journal of Automation and Information Science, 29(1), 55–62. **3.** Хусаинов Д.Я., Диблик Й., Ружичкова М. Линейные динамические системы с последействием. Представление решений, устойчивость, управление, стабилизация. – Киев: ГП Информ.-аналит. агенство, 2015. – 252 с.

Chugay A.M.
Institute for Mechanical Engineering Problems, Kharkiv, Ukraine

Mathematical modelling of homothetic polytopes packing

In this work we construct on the ground of phi-functions a mathematical model of the problem stated and investigate its characteristics. Based on the characteristics we offer a way of construction of starting points, a rapid algorithm of searching for local extrema and an original approach for non-exhaustive search of local extrema to obtain a good approximation to a global extremum.

The 3D packing problem belongs to a general class of combinatorial optimization problems which are concerned with packing a set of geometrical objects into container to minimize the waste or maximize profit. The 3D irregular packing problem occurs in many applications, for example: material science, powder technology, nano technology, container loading, layout design, building layout, 3D laser cutting. These problems are NP-hard, and, as a result, solution methodologies generally employ heuristics. Some researchers develop approaches based on mathematical modeling and general optimization procedures.

In order to construct a mathematical model of the problem stated it needs to describe interaction polytopes between themselves depending on their location in analytical form. As a tool for such description may be adapted phi-functions [1]. Phi-functions to describe non-overlapping and containment constraints for appropriate pairs of rotating polytopes are derived. Based on phi-functions a mathematical model of the problem is formulated as a non-linear optimization problem. The feasible region W of constructed mathematical model has a complicated structure: it is, in general, a disconnected set, each connected component of W is multiconnected, the frontier of W is usually made of nonlinear surfaces containing valleys, ravines. A matrix of the inequality system which specifies W is strongly sparse and has a block structure. The feasible region W is specified by a system of nonlinear inequalities with piecewise continuously differentiable functions, which involve operations of maximum and minimum of smooth functions. This means that the feasible region can be represented as a finite union of sub-regions. Each sub-region is described by a system of inequalities with smooth functions. In order to diminish solution time of problem it is desirable to give starting points belonging to the feasible region. At present construction of starting points are executed by greedy and heuristic algorithms. The algorithms do not permit to generate any starting points. It results in significantly restriction of diversity of local minimum points. Furthermore the algorithms demand essential time expenditures. The circumstances reduce to necessity to develop new approaches to derive starting points. We offer the approach which allows to construct starting points belonging to the feasible region without any limits. To realize the approach we cover polytopes by spheres of minimal radii and pack the spheres into cuboid of a minimal volume. Then we take center coordinates of the spheres and sizes of polyhedron obtained as a starting point and compute a local minimum point of the problem. In order to compute a new local minimum point it needs to build a new starting point. If the starting point is randomly generated then objective value at an appropriate local minimum point may be either improved or made worse or invariable. We suggest a way of a transition from a local minimum point to a new local minimum point at which the objective value is either improved or invariable. A hypothesis that homothetic coefficients of polytopes are variable allows to develop a new way to generate local extremum points. The way simplifies the solution process and improves the performance of results obtained.

References. 1. Chernov N., Stoyan Y., Romanova T. Mathematical model and efficient algorithms for object packing problem. Comput. Geom.: Theory and Appl. 43(5): 535–553 (2010).

***Cojocaru Sv.*[1] *Petic M.*[1,2] *Albu V.*[1]**
[1]*Institute of Mathematics and Computer Science of the Academy of Sciences of Moldova, Chişinău, Republic of Moldova;* [2]*Alecu Russo Balti State University, Bălţi, Republic of Moldova*

Developing classifiers for information related to disaster topics

The research is focused on solving the problem of identifying signals related that an accident occurred or may occur. For this purpose methods of selecting and processing texts from global network Internet are developed. Information in Romanian is of special interest for us. In order to obtain the mentioned tools we should follow several steps, divided into preparatory stage and processing stage. We also provide a classification of human behavior into typical or non-typical in relation to certain situations.

Introduction. One of the factors destabilizing the situation in society are social disasters caused both by natural factors and man-made (e.g., technogenic catastrophes) through conscious actions of individuals (terrorism). These events have a wide resonance in society through information networks and the media. However, due to lack of training and awareness, and due to time constraints, it is not always possible to find an appropriate solution to such disasters, leading to negative consequences after such events. To provide decision makers with a suitable support it is necessary to develop specialized analytical tools based on fast information processing.

In our research, conducted within NATO Science for Peace Project NUKR.SFPP 984877 – "Modelling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism", we will focus our attention on processing large volume of unstructured data available at global information networks in order to identify and analyze Romanian language texts, which refer to the subject mentioned above.

In this study we propose to identify some the useful data sources and accumulate a number of texts in order to create a basic lexicon of relevant words (markers) for disaster domains of different kinds: natural, technological or social. Thereafter the formed lexicon of markers will serve as a basis for development of tools to extract texts from the identified sources and to classify them. From the other hand we will also process the videos illustrating the corresponding texts in order to provide sentiment analysis.

Online news articles collection and processing. According to [1] we will go through the following steps: manual Romanian texts collection and their (manual) classification, lexicon of markers based on these texts and other available sources elaboration, automated enriching of the vocabulary by flexing and derivation and the database with classified texts creating. After this preparatory stage we will proceed with automated news texts selection, their filtering according to lexicon of markers and analysis. There are a number of Data Analytics suites which permit to solve many of these tasks. But the commercial ones being quite expensive [2,3] are not accessible to the entire spectrum of potential users. Therefore our approach is based on adapting and developing the existing Open Source tools.

The main part of our attention was paid to news articles from Internet. We analyzed two news sites and three TV sites, containing Romanian texts. The collected texts were processed and manually classified. 616 news articles were divided into 10 categories of disasters: railway, air and cars accidents, fire, earthquake, hurricanes, radioactive substances, attacks, flood and diseases. They contain 142840 words. All texts were lemmatized and annotated with the part of speech tagger. The obtained collection of Romanian texts about disasters permits to extract and annotate a lexicon of markers concerning the topic on disasters. We annotated manually those words and/or word expressions that express the meaning of social disasters. As we established 10 categories of articles of social disasters, each of them has its own set of lexical markers.

The work is carried out as part of two projects: "Modeling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism" supported by NATO.

We used Natural kit for NodeJS. In order to operate with it the collected texts are presented in NodeJS format, the data base was populated according to 10 categories of texts mentioned above. Using a Crawler based application, written using Node.js and Request library, texts are extracted and filtered according to the list of markers. After that they are cleaned (by removing different HTML tags etc.) and analyzed by comparing with the classified texts from the data base. The similarity score which can offer a statistic idea of how closed is the analyzed text to a selected topic is calculated.

Human behavior classification. Regarding to sentiment analysis we deal with human emotion and gesture classification. We propose a neural network model and demonstrate, that this architecture represents an effective tool for real-time processing of person's behavior [4]. The proposed approach combines most recent biometric techniques with the neural network approach for real-time emotion and behavioral analysis. Automated analysis of these behaviors with the machine learning techniques allowed us to train a complex convolutional neural network (CNN) to classify behavior of a user by classification both body movements and facial features. Such a feedback can provide important measures for user response to an interaction with any chosen system with a limited number of gestures involved. We use infrared cameras to automatically detect features and the movements of the limbs in order to classify user behavior into typical or untypical for the kind of task he is performing.

The architecture of a CNN can be described as following. A small pixel region goes to input neurons and then connects to a first convolution hidden layer. There we can involve a set of learnable filters, which are activated during the presentation some particular type of feature in pixel region in the input. On this phase, CNN does shift invariance, which is carried by feature map. Subsampling layer goes next. There we have two processes: local averaging and sampling. As a result, we get declining resolution of feature map. To correspond this task CNN needs supervised learning. Before starting the experiment, we gave a set of labeled videos with different emotional experience. The system analyses images and finds similar features. Then the system creates a map, where it arranges videos in accordance with similar features. Thereby, images with similar emotions form certain class. To test the system, we add other videos and correct the system when it refers them improperly.

In the series of experiments, emotions of human subjects were recorded, recognized, and analyzed to give statistical feedback of the overall emotions of a number of targets within a certain time frame. In this study, we developed a neural network model for recognition of body movements on two types (typical and non-typical) and facial expression accuracy (of 18% and 38% error rate, respectively). These results are achieved independently and combined afterwards with a simple classification algorithm. In this case the rate of accuracy increased up to 98% [5]. To improve system's results, the proposed model requires a large amount of training data, which we cannot be easily obtained. Therefore, the natural continuation of current research would be conducting further field tests to obtain more training data and improve the performance.

Conclusions. Our experience has shown that the proposed tool provides acceptable results.

References. **1.** R. Cody. An Introduction to SAS University Edition. SAS Institute, 2015. **2.** Sisense. https://www.sisense.com/. **3.** Sv. Cojocaru, M. Petic, Gr. Horoş. Tools for Texts Monitoring and Analysis Aimed at the Field of Social Disasters, Catastrophes, and Terrorism. Computer Science Journal of Moldova, vol.24, no.2(71), 2016, pp. 157–171. **4.** V. Albu, Sv. Cojocaru. Measuring human emotions with modular neural networks and computer vision based applications, Computer Science Journal of Moldova, v.23, n.1 (67), 2015, pp. 40–61. **5.** V. Albu. Measuring customer behavior with deep convolutional neural networks. BRAIN. Broad Research in Artificial Intelligence and Neuroscience. Volume 7, Issue 1, 2016, pp. 73–78.

Gorban N.V., Khomenko O.V., Paliichuk L.S.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Asymptotic behavior of solutions for climate energy balance model on manifold without boundary

We consider the climate energy balance model which was proposed in Budyko [1] and Sellers [2] and examined also in Díaz et al. [3–5]. Let $(\mathcal{M}, \mathbf{g})$ be a C^∞ compact connected oriented two-dimensional Riemannian manifold without boundary (as, e.g. $\mathcal{M} = S^2$ the unit sphere of $\mathbb{R}^3$). Consider the problem:

$$\frac{\partial u}{\partial t} - \Delta u + R_e(x,u) \in QS(x)\beta(u), \quad (x,t) \in \mathbb{R}_+ \times \mathcal{M}, \tag{1}$$

where $\Delta u = \mathrm{div}_{\mathcal{M}}(\nabla_{\mathcal{M}} u)$; $\nabla_{\mathcal{M}}$ is understood in the sense of the Riemannian metric $\mathbf{g}$. The unknown $u(x,t)$ represents the average temperature of the Earth's surface. In Budyko [1] the energy balance is expressed as

$$\text{heat variation} \;=\; R_a - R_e + D.$$

Here $R_a = QS(x)\beta(u)$. It represents the solar energy absorbed by the Earth, $Q > 0$ is a solar constant, $S(x)$ is an insolation function, given the distribution of solar radiation falling on upper atmosphere, β represents the ratio between absorbed and incident solar energy at the point x of the Earth's surface (so-called co-albedo function). The term R_e represents the energy emitted by the Earth into space, as usual, it is assumed to be an increasing function on u. We consider $R_e = Bu$, where B are some positive constant. The term D is heat diffusion, we assume (for simplicity) that it is constant.

Let $S : \mathcal{M} \to \mathbb{R}$ be a function such that $S \in L^\infty(\mathcal{M})$ and there exist $S_0, S_1 > 0$ such that $0 < S_0 \leqslant S(x) \leqslant S_1$. Suppose also that β is a bounded maximal monotone graph of $\mathbb{R}^2$, that is there exist m, $M \in \mathbb{R}$ such that for all $s \in \mathbb{R}$ and $z \in \beta(s)$ $m \leqslant z \leqslant M$.

We examine the regularity and long-term dynamics as $t \to +\infty$ of all weak solution for problem (1) under the assumptions listed above. For this aim we use the methods of multivalued analysis (developed by Zgurovsky M.Z., Melnik V.S., Kasyanov P.O., Kapustyan O.V. etc. [7]). The provided results where partially announced in Gorban et al. [6].

As a results of the investigations the existence and properties of weak solutions, and existence of Lyapunov function are obtained. We state that there exist trajectory and global attractors for all weak solutions of problem (1) and provide their structure and regularity properties. The thesis is prepared on the results of [8].

References. **1.** Budyko, M.I.: The effects of solar radiation variations, on the climate of the Earth. Tellus. Vol. 21. P. 611–619 (1969). **2.** Sellers, W.D.: A global climatic model based on the energy balance of the Earth-atmosphere system. J. Appl. Meteorol. Vol. 8. P. 392-–400 (1969). **3.** Díaz, H., Díaz, J.: On a stochastic parabolic PDE arising in climatology. Rev. R. Acad. Cien. Serie A Mat. Vol. 96. P. 123–128 (2002). **4.** Díaz, J., Hernández, J., Tello, L.: On the multiplicity of equilibrium solutions to a nonlinear diffusion equation on a manifold arising in climatology, J. Math. Anal. Appl. Vol. 216. P. 593–613 (1997). **5.** Díaz, J., Hernández, J., Tello, L.: Some results about multiplicity and bifurcation of stationary solutions of a reaction diffusion climatological model. Rev. R. Acad. Cien. Serie A. Mat. Vol. 96, Iss. 3. P. 357–366 (2002). **6.** Gorban, N.V., Gluzman, M.O., Kasyanov, P.O., Tkachuk, A.M.: Long-Time Behavior of State Functions for Budyko Models. Continuous and distributed systems III. Series studies in systems. Decis. Control. doi:10.10.1007/978-3-319-40673-2_18. **7.** Zgurovsky, M.Z., Kasyanov, P.O., Kapustyan, O.V., Valero, J., Zadoianchuk, N.V.: Evolution inclusions and variation Inequalities for Earth data processing III. Springer, Berlin, 2012. DOI:10.1007/978-3-642-28512-7. **8.** Gorban, N.V., Khomenko O.V., Paliichuk, L.S., Tkachuk, A.M.: Long-time behavior of state functions for climate energy balance model. DCDS. Series B. (2017) to appear.

Kaniovska I.Y., Zaitsev A.A.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Implementation of Markov random processes for studying risks in credit portfolios

The fundamental problem in the analysis of credit portfolios is to determine their cost distribution and to calculate the default risks both for a portfolio as a whole and every separate debtor in particular. When the portfolio is small it is possible to regard it in terms of individual debtors. Therefore it will be advantageous to use a separate analysis of each debtor and merge the results at the end. At the same time, large portfolios require a more complex system for risk assessment because there exist elements that cannot be seen in smaller ones [1].

There exist a fair amount of paradigms (mixed binomial model, dependent lifetime model, etc.) and methods (CreditMetrics, KMV etc.) used for investigations of portfolio's properties. CreditMetrics incorporates dependencies of debtors, which might be e.g. some common systematic factors. These dependencies are quoted as a pairwise correlation, which is translated into a correlation matrix indicating pairwise correlations throughout the whole portfolio. CreditMetrics is, therefore, a model rewarding well-diversified portfolio with a lower riskiness [2]. At the same time, it was disputed that such approaches suffer from a significant drawback: the necessity to assume the nature of the distribution which is often impossible to verify basing on empirical data.

Kaniovski and Pflug have introduced the model which considers every debtor in the portfolio to migrate between credit ratings following a Markovian rating process. After that, the portfolio was investigated as an n-dimensional random Markovian process each component of which depended on the other ones. The advantage this approach has is that there is no need to make any initial assumptions for the portfolio's distribution thus removing the necessity to verify it. The model makes use of determining deteriorating and non-deteriorating rating random tendency variables for the every debtor in a given portfolio [3].

In this research, we determine what happens when we use the tendency variable vector obtained by this method and incorporate them into one of the widely accepted analysis methods: CreditMetrics. The main results obtained give us a difference in accuracy between the standard model and modification [3, 4].

In the classic CreditMetrics model, it is possible to calculate the bound values for each debtor in the portfolio, basing on the company's performance and indices. Such factors as company's sector, credit rating, LGD, par, coupon and maturity are considered. After bound values for each debtor are calculated it is possible to obtain a value for the whole portfolio simply by adding them. In the second approach, we consider every debtor to follow a Markovian process with 8 states (AAA – CCC and the absorbing default state) with a transition matrix provided by "Standard and Poor's" and calculated the portfolio's expected value with respect to dependencies between processes. We have used several homogeneous and heterogeneous test portfolios of sizes 10, 20, 40 debtors. Over the course of Monte-Carlo simulations, we have discovered that classic CreditMetrics method performs best in cases of middle to big heterogeneous and homogeneous and small heterogeneous portfolios. On the other hand, the alternate method proved to be more reliable in cases of small to middle homogeneous portfolios [3, 5].

References. **1.** J.P. Morgan, CreditMetrics™– Technical Document, New York, 1997, pp. 35–47. **2.** Frey, R., Backhaus, J., Portfolio Credit Risk Models with Interacting Deafault Intensities: a Markovian Approach, University of Leipzig, 2004. **3.** Y.M. Kaniovski, G.Ch. Pflug Risk assessment for credit portfolios: a coupled Markov Chain model, Science Direct, 2007, pp. 2303–2323. **4.** Tomasz R. Bielecki, Stéphane Crépey, Alexander Herbertsson, Markov Chain Models of Portfolio Credit Risk, Illinois Institute of Technology, 2009, pp. 13–21. **5.** Giuseppe Di Graziano, L.C.G Rogers, A dynamic approach to the modelling of correlation credit derivatives using Markov chains, University of Cambridge, 2006.

Nadtochii O.V., Bidyuk P.I.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Modeling and forecasting demographic processes in Ukraine

Demographic models and the forecasts based upon them are important elements of general monitoring and management for the processes in society. They provide a possibility for directed influence on development of socio-economic components of society and to direct them into a side necessary for a country. The results of forecasting are necessary for reflection and understanding the future characteristics of demographic processes under consideration as well as for long term planning in the spheres of economy, sociology, education, medical service etc. Actually demographic forecasting is scientifically based foresight of population number and its components, tendencies of their future development in specific economic conditions [1]. The forecasts are helping to determine necessary rate of economy development, and strategic planning of social developments. They create a basis for analyzing requirements of goods and services, necessary resources planning, estimation of possible changes in international policy etc.

Most of modern actual demographic processes exhibit nonlinear nonstationary behavior that require constructing new high quality forecasting mathematical models. To solve the problem help decision support systems (DSS) that play very important role in processing data and expert estimates, adequate model constructing, estimating of quality forecasts for a given horizon and generating decision alternatives for decision makers (DM). DSS provide wide possibilities for computing high quality forecasts with the most different techniques and help to combine the estimates generated by different methods and models. Thus, modern DSS are considered as highly developed computer systems that are capable to perform multiple complicated computations directed towards providing maximum help for a DM at all stages of data and expert estimates processing. Features of adaptation are usually implemented in such systems aiming to constructing model with adaptive structures and parameters. Following this approach we can maintain model adequacy in conditions of analyzing nonlinear non-stationary processes characteristics. The DSS proposed in this study has an architecture including the following elements: the language subsystem (interface), the main processing unit that performs all necessary computations, data and knowledge base (DKB), and subsystem visualizing intermediate and final results of computing [2]. The DSS functionality is based on the following computational operations: data collecting and storing, preliminary data processing; statistical analysis of data aiming to preparing the data for model structure and parameters estimation, computing of short- and medium-term forecasts using constructed forecasting functions, and quality analysis subsystem that is based on three sets of statistical quality criteria (quality of data, models, and forecasts). In each case we hire separate sets of statistical parameters (statistics) directed towards solving specific monitoring problems for the computational processes [2, 4].

Among possible uncertainties that could take place in modeling, forecasting, risk estimation, and alternatives generation the following ones should be mentioned: data uncertainties; model structure and parameters uncertainties; forecasts uncertainties, and uncertainties related to computational process, for example, systematic computing and methodology errors. The data uncertainties are usually caused by the following reasons: incomplete measurements, short samples, missing data and outliers; influence of external random disturbances on functioning of a system or process under study; and measurement errors. Today there exist appropriate statistical instrumentations for identification and coping with the uncertainties. There is a large number of various techniques for imputation of missing data such as expectation maximization (EM) algorithm, moving average and regression based forecasting, optimization techniques, imputation from appropriate distribution etc. Negative influence of external random disturbances and measurement errors could be partially eliminated by application of optimal Kalman filter, elliptical filtering, exponential smoothing or other appropriate method. The Kalman filtering techniques are among the most often applicable and convenient due to availability of various useful possibilities of the filter functions including

estimation of non-measurable components of a state vector, estimation of statistical parameters for stochastic disturbances and measurement errors. On the stage of model structure and parameters estimation we meet another set of uncertainties caused by a model order estimation errors, lack of information about processes nonlinearity, delay time and type of data distribution. The model structure is usually refined thanks to implementation of appropriate adaptation scheme, application of appropriate statistical tests that help to reveal statistical data structure in the best possible way. The uncertainties of data and model structure certainly influence the quality of the model parameters estimates say they may result in higher variance and/or bias of the estimates. To reduce influence of the uncertainties we introduce adaptive estimation schemes for model structure, parameters, and consequently forecasts and risk estimates. A key point for correct estimation of mathematical model parameters is in correct determining of a distribution type for the data being processed. If it is far from normal (what happens rather often) then we apply appropriate Markov chain Monte Carlo (MCMC) algorithm adjusted to specific distribution of statistical data [3].

Such systemic approach is directed towards quality improvement for all possible computations programmed in the system. We also take into consideration the necessity of hiring different kinds of criteria dependently on linear or non-linear models are analyzed. For example, non-linear classification models require for application such criteria as Gini index, common accuracy, area under curve (AUC) etc.

The modeling and forecasting techniques implemented in the system are as follows: regression analysis (linear and nonlinear), the group method for data handling (GMDH and fuzzy GMDH), neural nets and neuro-fuzzy techniques, and Bayesian networks (static and dynamic). Bayesian networks and neuro-fuzzy approach help with eliminating the influence of probabilistic and amplitude type uncertainties. Besides Bayesian networks create a powerful instrumentation for combining quantitative and qualitative variables in one model. Besides the model could have a very high dimensionality, say, several hundred variables. The theory of Bayesian networks is quickly developing these days. Another possibility for solving the forecasting problem provide for such methods as optimal Kalman filtering techniques, hierarchical models, nonparametric and Bayesian regression.

Among the processes under study are the following: Gross Domestic Product (GDP) for Ukraine, general dynamics of population, dynamics of birth and death, number of registered marriages and divorces. Preliminary study of the demographic processes selected shows that the most appropriate forecasting models in this case are as follows: autoregressive, autoregressive with moving average, autoregressive with polynomial trend, and the group method for data handling method based models. To select the best models the following statistics were hired: determination coefficient, sum of squared residuals, Durbin–Watson statistic, F-statistic etc. The tests for stationarity and nonlinearity are applied for identifying time series structure. The best forecasts are selected with mean absolute percentage error (MAPE), mean squared errors (MSE), and Theil coefficient [2–4]. We have also applied combined criteria that include both model adequacy and forecasts quality statistics. In a case of studying nonlinear processes some other extra criteria were applied. The results achieved for short- and medium-term forecasting of the demographic processes are quite acceptable or high quality, say with MAPE changing in the range between 0.3% and 3.8%.

The future developments will be directed towards application of combined forecasting techniques with optimization of weighting coefficients. Additional efforts are required for identifying and taking into consideration existing specific uncertainties regarding data, models and forecasts. We are also aiming to automating all basic computational functions of the DSS.

References. **1.** Novoselova S.V., Denysenko M.B. Basic demography. – Minsk: Altiora, 2012 – 133 p. **2.** Bidyuk P.I. Design of Decision Support / P.I. Bidyuk, O.P.Gozhyj, L.O.Korshevnyuk. – Mykolaiv: Black Sea State University named after Petro Mohyla, 2012. – 380 p. **3.** Tsay R.S. Analysis of Financial Time Series / R.S.Tsay. – Hoboken (New Jersey): John Wiley & Sons Inc., 2010. – 715 p. **4.** Dovgyj S.O. Decision support systems based on statistical and probabilistic procedures / S.O. Dovgyj, P.I. Bidyuk, O.M. Trofymchuk. – Kyiv: Logos, 2014. – 419 p.

Pankratova N.D., Kondratova L.P.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

The model for guaranteed functioning of complex engineering systems taking into account the margin of permissible risk

In operation of the complex engineering systems (CES) a practical need arises for the systemic coherence of the diagnosis rates with the pace of work processes in different modes of CES operation. Such coordination could be one of the most important conditions to ensure the guaranteed security of high-risk objects, which include nuclear power plants, large industrial associations, complicated technological and technical systems of various purpose. Analysis of accidents and disasters allows to reveal the most important causes and disadvantages of the prevailing safety management principles of modern technique. One such reason consists in the peculiarities of diagnostic systems, focused on the detection of failures and malfunctions. The existing diagnostics tools are oriented to control of CES technical status and detect failures taking into account the uncertainties, where the failures are associated with the moment in time of its appearance, place and class. An examples of such diagnostics tools are [1–3].

The rational accompaniment of CES operation process will provide a certain margin of permissible risk with the required permissible values for the risk degree and the risk level. The implementation of the strategy of the guaranteed CES functioning assumes to use the principles of a system guaranteed diagnosing of the risk situations, timeliness and reliability of the estimation and forecasting of the risk situations [4]: the system efficiency indicators coherence of technical diagnostics systems with those of the safety control system on ensuring with the necessary reliability to prevent CES failures within the margin of permissible risk of various modes; timely detection, identification and assessment of abnormal mode risk over the forecast period of CES operation to guarantee the timely elimination of the causes of the risk before the failure appearance.

There is proposed the model for CES guaranteed functioning in real operating conditions under the influence of destabilizing factors and different kind uncertainties taking into account the margin of permissible risk. This model is designed to further research ideas proposed in [5], and is being implemented on the basis of timely detection and elimination of the causes of a possible transition of CES to inoperable state accounting the integral index of decision maker informedness (DMI) about the situations, as a combination of partial indicators DMI completeness, veracity, timeliness and unavoidable threshold limit time under the impact of destabilizing factors.

The margin of permissible risk is estimated for the situations of some S_j set of the nested sequence of $S_1 \subset S_2 \subset ... \subset S_j \subset ... \subset S_{n_s}$, which characterize CES state at a some time moment. The situations of S_j set are characterized by the time intervals, which are represented in matrix of $[T_R] = \left\{\left[\left(\ t_f^i\ \right), \left(\ t_r^i\ \right)\right]\right\} | \left(\ t_f^i\ \right) < \left(\ t_r^i\ \right), \left(\ t_r^i\ \right) < \left(\ t_{cr}\ \right), i = \overline{1,m}$, where $\left(\ t_f^i\ \right), \left(\ t_r^i\ \right)$ are row vectors of the time moments for the formation and the implementation of decision relatively to S_j situation set, $\left(\ t_{cr}\ \right)$ is the row vector of the boundary values for time taking decisions relatively to the S_j situation set. Taking into account the different kind uncertainties there is estimated in real time mode based on the functional dependencies of $\widetilde{y}_i = \phi_i\left(\{x_j, j = \overline{1,n}\}\right), i = \overline{1,m}$, restored with reasonable accuracy at a given time interval in the sample of real CES functioning indicators values of $y_i | i = \overline{1,m}$ and CES parameters of $x_j | j = \overline{1,n}$ accounting the forecast on some period, the rational level of the integral DMI index as the row vector of $\left(\ I_0^-\ \right) = \langle I_{0p}^- = \min_{i\in[1,m]}\left\{\overline{I}_0^{ip}(t_k) | t_k \in \left[\left(\ t_f^i\ \right), \left(\ t_r^i\ \right)\right], \left(\ t_r^i\ \right) < \left(\ t_{cr}\ \right)\right\}, p = \overline{1, S_j}\rangle$, providing the recognition of S_j set situations. In the expression of $\left(\ I_0^-\ \right)\ \overline{I}_0^{ip}(t_k)$ is the row vector of the integral DMI index average values for $\widetilde{y}_i$ about p-th situation of S_j set in the time moments of $t_k \in \left[\left(\ t_f^i\ \right), \left(\ t_r^i\ \right)\right], \left(\ t_r^i\ \right) < \left(\ t_{cr}\ \right), i = \overline{1,m}$.

The margin of permissible risk, represented by the vector of $\left(\ R_i^{perm}\ \right) = \langle R_{ip}^{perm}, p = \overline{1, |S_j|}\rangle$

$\forall i = \overline{1,m}$ is realized relatively on S_j set situations in the time intervals of $\left[\left(\ t^i_{a_1}\ \right), \left(\ t^i_{s_2}\ \right)\right], \left(\ t^i_{s_1}\ \right) \leqslant \left(\ t^i_{a_1}\ \right) \leqslant \left(\ t^i_{s_2}\ \right)\ \forall i = \overline{1,m}$ with the boundaries of $\left(\ t^i_{a_1}\ \right)$, $\left(\ t^i_{s_2}\ \right)$, accounting to the values of $\widetilde{y}_i\left(t^{ip}_{a_1}\right)$ with the risk degree of $\eta_{ip}\left(\widetilde{y}_i\left(t^{ip}_{a_1}\right)\right) < \eta_{ip}\left(\widetilde{y}_i\left(t^{ip}_{s_2}\right)\right)$ and the risk level of $W_{ip}\left(\widetilde{y}_i\left(t^{ip}_{a_1}\right)\right) < W_{ip}\left(\widetilde{y}_i\left(t^{ip}_{s_2}\right)\right)$ and $\widetilde{y}_i\left(t^{ip}_{s_2}\right)$ with $I^{ip}_0\left(t^{ip}_{s_2}\right) \geqslant I^{-}_{0p}$ for p-th situation of S_j set.

The coordination of the margin of permissible risk is implemented relatively S_j set of situations in according with conditions of $\left(\ I^i_0(t_k)\ \right) \geqslant \left(\ I^-_0\ \right)$, $t_k \in \left[\left(\ t^i_{s_1}\ \right), \left(\ t^i_{s_2}\ \right)\right]$, $i = \overline{1,m}$ for $\widetilde{y}_i$ in time intervals of $\left[\left(\ t^i_{ca_1}\ \right), \left(\ t^i_{cs_2}\ \right)\right] \subset \cup_{i\in[1,m]}\left[\left(\ t^i_{a_1}\ \right), \left(\ t^i_{s_2}\ \right)\right]$, where it is guaranteed the prevention of abnormal mode situations, and is defined by the vector of $\left(\ R^{perm}_c\ \right) = \langle R^{perm}_{ip}, p = \overline{1,|S_j|}\rangle$.

The model for guaranteed CES functioning in real operating conditions under the influence of destabilizing factors and different kind uncertainties describes the coordination of the margin of permissible risk for the functional dependencies of $\widetilde{y}_i$ as

$$R^{perm}_c = \langle R^{perm}_{ip} = \frac{\left|\left[\min_{i\in[1,m]} t^i_{a_1}, \min_{i\in[1,m]} t^i_{s_2}\right]\right|}{\left|\left[\min_{i\in[1,m]} t^i_{s_1}, \min_{i\in[1,m]} t^i_{s_2}\right]\right|}, p = \overline{1,|S_j|}\rangle; \tag{1}$$

$$t^{ip}_{a_1} = \min_{t^{ip}_k \in \left[t^{ip}_{s_1}, t^{ip}_{s_2}\right]} \left\{t^{ip}_k | t^{ip}_k < t^{ip}_{s_2}\right\}, p = \overline{1,S_j}, i = \overline{1,m}; \tag{2}$$

$$\eta_{ip}\left(\widetilde{y}_i\left(t^{ip}_{a_1}\right)\right) < \eta_{ip}\left(\widetilde{y}_i\left(t^{ip}_{s_2}\right)\right), W_{ip}\left(\widetilde{y}_i\left(t^{ip}_{a_1}\right)\right) < W_{ip}\left(\widetilde{y}_i\left(t^{ip}_{s_2}\right)\right), p = \overline{1,S_j}, i = \overline{1,m}, \tag{3}$$

$$I^{ip}_0\left(t^{ip}_{s_2}\right) \geqslant I^{-}_{0p}, p = \overline{1,S_j}, i = \overline{1,m}. \tag{4}$$

In (1), (2) $t^{ip}_{a_1}$ characterizes for $\widetilde{y}_i$ the low boundary of the time interval to realize the margin of permissible risk of abnormal mode relatively p-th situation of S_j set. The parameter of $t^{ip}_{a_1}$ is defined in the time intervals of $\left[\left(\ t^i_{s_1}\ \right), \left(\ t^i_{s_2}\ \right)\right] \subseteq \left[\left(\ t^i_f\ \right), \left(\ t^i_r\ \right)\right]$ to search the additional information on decision-making taking into account conditions of (3) on the risk degree of $\eta_{ip}\left(\widetilde{y}_i\left(t^{ip}_{s_2}\right)\right)$, risk level of $W_{ip}\left(\widetilde{y}_i\left(t^{ip}_{s_2}\right)\right)$ and conditions of (4) on integral DMI index of $I^{ip}_0\left(t^{ip}_{s_2}\right)$ in the final times to identify the situations of S_j set.

The study of CES operation process using the model described above was carried out on the example of a closed deep water system whose state at each time moment of a given operating period is characterized by the indicator values of the water level y_1, y_2 in the inlet and outlet reservoirs, and the water pressure y_3 and its temperature y_4 at the inlet of technological object. Taking into account obtained based on the model coordinated values of the margin of permissible risk of abnormal mode situations for the functional dependencies of $\widetilde{y}_1$, $\widetilde{y}_2$, $\widetilde{y}_3$, $\widetilde{y}_4$ it is possible to guarantee the prevention of transition to abnormal situations.

Thus, using of the proposed model allows ensure the survivability of CES functioning in the real operating conditions under the influence of destabilizing risk factors and different kind uncertainties.

References. **1.** Lanetsky B.N., Lukyanchuk V.V., Terebukha, I.N. Complex model to control the technical state and restore the efficiency of complex technical system with a multi-level structure // Systems of armament and military equipment. – 2011. –4(28). – P. 73–75. **2.** Lerner U., Bayesian R.P., Koller D., Biswas G. Fault Detection and Diagnosis in Dynamic Systems [Text] // Proceedings of the XVII-th National Conference on Artificial Intelligence (AAAI-00), Austin, Texas, August, 2000. – P. 531–537. **3.** Fagarasan I., Iliescu S.St. Parity equations for fault detection and isolation [Text] // International Conference on Automation, Quality and Testing. – Robotics. – 2008. – P. 99–103. – DOI:10.1109/ AQTR.2008.4588715. **4.** Pankratova N.D. Ensuring the security and survivability of the complex engineering systems operating // System analysis in the design and management. – SPb, 2016. - P. 266–273. **5.** Pankratova N.D., Kondratova L.P. System evaluation of engineering objects' operating taking into account the margin of permissible risk // Eastern-European Journal of Enterprise Technologies. – 2016. – T. 3, №4(81). – P. 13–19. – DOI:10.15587/1729-4061.2016.71126.

Pavlenko V.D., Salata D.V.
Odessa National Polytechnic University, Institute of Computer Systems, Odessa, Ukraine

Identification Eye-Motor System with Using Volterra Model

A new method of constructing nonparametric dynamic model of the human oculomotor system on the basis of experimental data "input-output". This takes into account nonlinear and inertial properties of the eye of the rectus muscles. A technology for tracking eye movement based on the videos. It is possible to determine the dynamic characteristics of the oculomotor system functions as a transition of the first and second order – integral transforms Volterra kernels.

Introduction. The innovative technology of Eye tracking which is rapidly developing nowadays – is the process of determination of the point where look being sent to or the determination of eye movements relatively to the head [1]. This high-tech innovation has been further developed and effectively used in the construction of a mathematical model of process of tracking eye movement to detect anomalies in data tracking to quantify the motor symptoms of Parkinson's disease [2, 3]. Using nonlinear dynamic model of Wiener and Volterra–Laguerre and their identification is based on a random effects test [4], which requires the application of methods of correlation analysis and generate a large amount of experimental data (long-term experimental studies).

In order to build of Volterra model oculo-motor system (EMS) a person is encouraged to use the test deterministic effects, for example, step signals (the most appropriate for the study of the dynamics of EMS) [5], which simplifies the computational algorithm to identify and significantly reduce the time of processing of experimental data. There is method, and computer algorithms identifying deterministic nonlinear dynamical systems in the form of Volterra models using multi-step signals [6].

The purpose of work is development method for constructing nonparametric dynamic model of oculo-motor system, taking into account its inertial and nonlinear properties, based on experimental studies of "input-output" and also computational tools and software for the information technology processing experimental data.

The purpose and research problems. The purpose of work is development method for constructing nonparametric dynamic model of oculo-motor system, taking into account its inertial and nonlinear properties, based on experimental studies of "input-output" and also computational tools and software for the information technology processing experimental data. To achieve this goal were set this following tasks:

- Development methods for constructing nonlinear dynamic model of EMS as a Volterra kernels which characterizing both nonlinear and inertial properties of the nature objects;
- Development information technology of obtaining experimental data for identification EMS based on pupil's movement tracking using video registration;
- Development computational methods of identification multidimensional dynamic (transient) characteristics EMS using test inputs as a Heaviside functions of different amplitudes;
- Verification constructed model EMS.

Computing Method of Multidimensional Transient Functions for Identification EMS. Taking into account specificity investigated object to identification used test multistage signals. If test signal $x(t)$ represents an identity function (Heaviside function) – $\theta(t)$, the result of identification the transition function of the first order and the diagonal section n-th order. To determine the sections subdiagonal transition functions n-th order ($n \geqslant 2$) EMS tested using the n step test signal with given amplitude and different intervals between signals. With appropriate processing responses get subdiagonal section n-dimensional transition functions $h_n(t-\tau_1, ...t-\tau_n)$ which

represent n-dimensional integral of Volterra kernel n-order $w_n(\tau_1...\tau_n)$ [5]:

$$h_n(t-\tau_1,...,t-\tau_n)=\int_0^{\infty}...\int_o^{\infty}w_n(t-\tau_1-\lambda_1,...t-\tau_n-\lambda_n)d\lambda_1...d\lambda_n. \quad (1)$$

Method for determination sections of n-dimensional transition functions based on the statement, proof of which is similar to that given in [6].

Is developing a method of constructing approximate Volterra model identification method based on the Volterra series time-domain response is based on the approximation $y(t)$ at an arbitrary deterministic signal $x(t)$ in the form of integral power of the polynomial N-th order (N – order approximation model)

$$\tilde{y}_N(t)=\sum_{n=1}^{N}\hat{y}_n(t)=\sum_{n=1}^{N}\int_0^t...\int_0^t w_n(\tau_1,...,\tau_n)\prod_{i=1}^{n}x(t-\tau_i)d\tau_i. \quad (2)$$

Let the input test signals EMS are fed alternately $a_1x(t), a_2x(t), , ..., a_Lx(t); a_1, a_2, ..., a_L$ – distinct real numbers satisfying the condition $\mid a_j \mid \leqslant 1$ for $\forall j$=1,2,...,L; then

$$\tilde{y}_N[a_jx(t)]=\sum_{n=1}^{N}\hat{y}_n[a_jx(t)]=\sum_{n=1}^{N}a_j^n\int_0^t...\int_0^t w_n(\tau_1,...,\tau_n)\prod_{i=1}^{n}x(t-\tau_i)d\tau_i=\sum_{n=1}^{N}a_j^n\hat{y}_n(t). \quad (3)$$

The partial components in the approximation model $\hat{y}(t)$ are found using the least square method (LSM). This makes it possible to obtain such evaluation in which the sum of squared deviations of responses identified the nonlinear dynamical system $y[a_jx(t)]$ on the model $\hat{y}_N[a_jx(t)]$ response is minimal, i.e. EMS provides a minimum criterion

$$J_N=\sum_{j=1}^{L}(y[a_jx(t)]-\tilde{y}_N[a_jx(t)])^2=\sum_{j=1}^{L}(y_j(t)-\sum_{n=1}^{N}a_j^n\hat{y}_n(t))^2\rightarrow min, \quad (4)$$

where $y_j(t)=y[a_jx(t)]$. Minimization of the criterion (4) is reduced to solving the system of normal equations of Gauss, which in vector-matrix form can be written as

$$\mathbf{A}^{'}\mathbf{A}\hat{\mathbf{y}}=\mathbf{A}^{'}\mathbf{y}, \quad (5)$$

where

$$\mathbf{A}=\begin{bmatrix}a_1 & a_1^2 & ... & a_1^N\\ a_2 & a_2^2 & ... & a_2^N\\ ... & ... & ... & ...\\ a_L & a_L^2 & ... & a_L^N\end{bmatrix},\quad \mathbf{y}=\begin{bmatrix}y_1(t)\\ y_2(t)\\ ...\\ y_L(t)\end{bmatrix},\quad \hat{\mathbf{y}}=\begin{bmatrix}\hat{y}_1(t)\\ \hat{y}_2(t)\\ ...\\ \hat{y}_N(t)\end{bmatrix}.$$

From (5) we obtain

$$\hat{\mathbf{y}}=(\mathbf{A}^{'}\mathbf{A})^{-1}\mathbf{A}^{'}\mathbf{y}. \quad (6)$$

Results of Identification EMS Transient Function. For successful implementation of the technology of experimental determination of the dynamic characteristics of EMS and human diagnostics in technical and medical applications, it's necessary to have data sets – the coordinates of the position of the pupil on plane, namely, the values of the horizontal and vertical eye rotation angles with respect to the initial position.

Tracking the angles of pupil rotation with the help of video recording involves the use of video cameras to produce images of the pupil over time at regular intervals, which would be clearly fixed the position of the pupil when it moves. We need a high-speed camera that would record not only visible to the human eye movement EMS (saccades), and more informative studies for micro-movements – tremors (small oscillations of the eye), drift (slow flowing micro-movements) and micro saccades (rapid eye movement lasting 10–20 ms).

For EMS studies the selected camera shooting speed is 120 (frames/sec). For an array of pupil coordinate values when moving eyes based on the footage, it is necessary to divide the video into

individual frames. One study time is 408,17 ms (50 frames). Processing video recording data associated with the release of each frame of video on the center of the pupil of the eye.

Result image analysis – the pixel coordinates of the pupil center on the frame of video x_i, y_i, i = 1,2, ..., m (m – the number of video frames, m=50). According to the coordinates of the pupil center and the results of the calibration is calculated direction of gaze, tied to the subject observer of the image. The accuracy assessment of coordinates in pixels ±2 pel.

Using tools developed by data processing means based nonparametric dynamic model of the human form of EMS in transition and transition of two-dimensional functions (1).

Approbation tracking technology of the pupil's behavior based on video registration is performed on the task of analysis of work the EMS along the horizontal axis. Result eye image analysis shown on (fig. 1).

Where in the input (test) signal – distance from the base of the perpendicular, dropped from the centre of pupil eye to the plane, in which is formed the perturbation – the light source to the point source (light spots) in the horizontal plane. Measure the response (the output) is a function of the current deviation of pupil in the frame image of the EMS from the starting point, depending on the time.

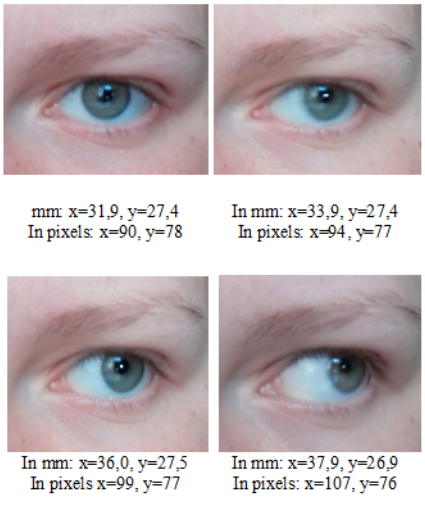


Figure 1. Result eye image analysis

To determine the diagonal section of the transient response second order object is tested at first step signal with an amplitude of the a (horizontal distance to light spot from the starting point, represents the original position the pupil).

Measured response of the eye $y_1(t), y_2(t), y_3(t)$ to the input test signals $a_1\theta(t), a_2\theta(t)$ and $a_3\theta(t)$ (L=3) for values of the test signal amplitudes a_1 = 0,33, a_2 = 0,66 and a_3 = 1 shown in (fig. 2).

The model response is calculated on the basis of estimates of the transient functions $\hat{h}_1(t)$ and $\hat{h}_2(t,t)$

$$\tilde{y}(t,a) = a\hat{h}_1(t) + a^2\hat{h}_2(t,t). \tag{7}$$

Comparison of the response of the constructed model $\tilde{y}(t,a)$ with the response of the identified EMS (with experimental data) $y(t,a)$ shown in (fig. 3).

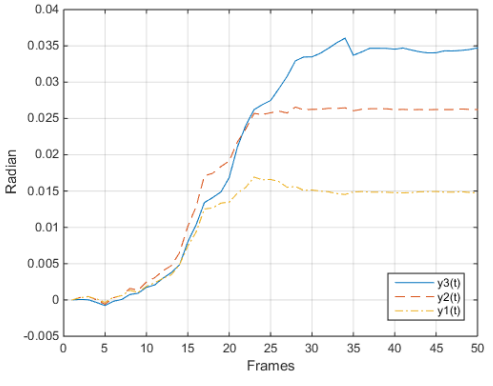


Figure 2. Responses EMS $y_1(t), y_2(t)$ and $y_3(t)$.

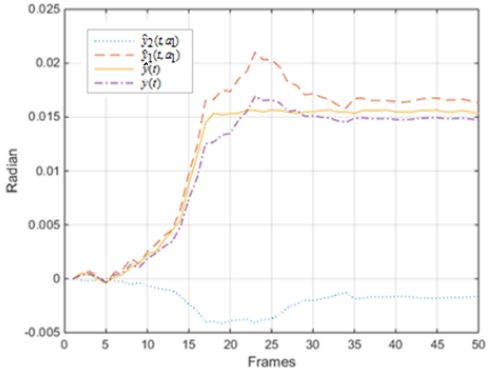


Figure 3. Responses EMS $y_1(t)$ and model $\tilde{y}(t,a_1), \hat{y}_1(t)$ and $\hat{y}_2(t)$.

Obtained graphs of EMS transient functions first $\hat{h}_1(t)$ and second $\hat{h}_2(t,t)$ shown in fig. 4 and fig. 5, respectively.

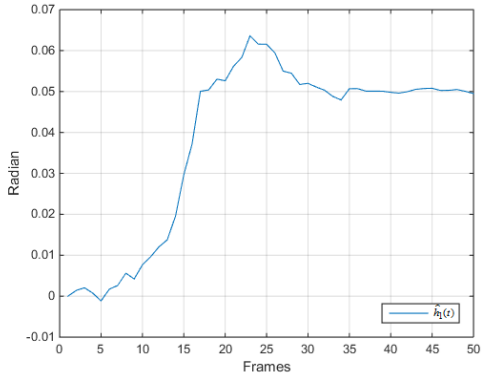


Figure 4. Transient functions $\hat{h}_1(t)$

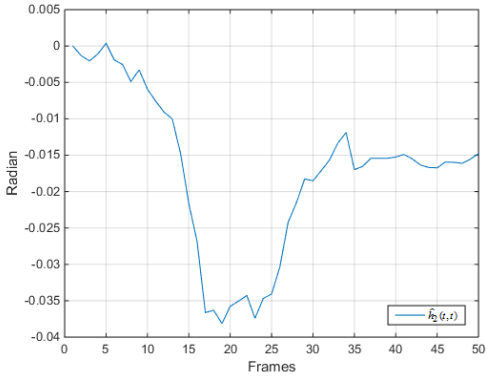


Figure 5. Transient functions $\hat{h}_2(t,t)$

Provided graphs are practically the same (standard deviation σ= 0.0009) which confirms effectiveness computational algorithm of identification and adequacy of the constructed model based on experimental data "input-output".

Acknowledgment. Proposed the method for constructing nonparametric dynamic models of human EMS taking into account its inertia and nonlinear properties based on the experimental data "input-output". Has been further developed technology tracking the pupil's behavior with help of video recording that has allowed to determine eye's dynamic characteristics. Proposed technology of tracking pupil's behavior does not need special equipment and laboratory experimental conditions and accessible for widespread use. Important feature of the technology is undemanding to the hardware, which opens up the possibility of its use in the modern mobile devices. Verification of the developed model showed adequacy her under the investigation object – virtually identical (within acceptable error) responses of the object and model at the same impact test.

References. **1.** J. Kepler, and U. Linz, "Biomechanical Modelling of the Human Eye." Netz Werk für Forschung, Lehre und Praxis, Linz, 231 (2004). **2.** D. Jansson and A. Medvedev, "Volterra Modeling of the Smooth Pursuit System with Application to Motor Symptoms Characterization in Parkinson's Disease." European Control Conference (ECC), 1856–1861. DOI: 10.1109/ecc.2014.6862207 (2014). **3.** D. Jansson, "Stochastic Anomaly Detection in Eye-Tracking Data for Quantification of Motor Symptoms in Parkinson's Disease". Advances in Experimental Medicine and Biology, 823, 63–82. DOI:10.1007/978-3-319-10984-8-4 (2015). **4.** P.Z. Marmarelis and V.Z. Marmarelis, "Analysis of Physiological Systems," The White Noise Approach, Plenum Press, New York, 487 (1978). **5.** F.J. Doyle, R.K. Pearson and B.A Ogunnaike, "Identification and Control using Volterra Models." Published Springer Technology & Industrial Arts, 314 (2001). **6.** V.D. Pavlenko, O.O. Fomin, A.N. Fedorova, and M.M. Dombrovskyi, "Identification of Human Eye-Motor System Base on Volterra Model." Herald of the National Technical University «KhPI». Subject issue: Information Science and Modelling, Kharkov, NTU «KhPI», 21 (1193), 74–85 (2016).

***Soloviev V.N.*[1], *Romanenko Y.V.*[2]**
[1]*Kryvyi Rih State Pedagogical University, Kryviy Rih, Ukraine;* [2]*Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine*

Economic analog of Heisenberg uncertainly principle and financial crisis

The Heisenberg uncertainty principle is one of the cornerstones of quantum mechanics. The modern version of the uncertainty principle, deals not with the precision of a measurement and the disturbance it introduces, but with the intrinsic uncertainty any quantum state must possess, regardless of what measurement is performed [1]. Recently, the study of uncertainty relations in general has been a topic of growing interest, specifically in the setting of quantum information and quantum cryptography, where it is fundamental to the security of certain protocols [2,3].

The aim of this study is to analyze the concepts and fundamental physical constants in terms of achievements of modern theoretical physics, they search for adequate and useful analogues in the socio-economic phenomena and processes, and their possible use in early warning of adverse crisis in financial markets. The instability of global financial systems depending on ordinary and natural disturbances in modern markets and highly undesirable financial crises are the evidence of methodological crisis in modelling, predicting and interpretation of current socio-economic conditions.

In paper [4] we have suggested a new paradigm of complex systems modelling based on the ideas of quantum as well as relativistic mechanics. It has been revealed that the use of quantum-mechanical analogies (such as the uncertainty principle, notion of the operator, and quantum measurement interpretation) can be applied to describing socio-economic processes. Methodological and philosophical analysis of fundamental physical notions and constants, such as time, space and spatial coordinates, mass, Planck's constant, light velocity from the point of view of modern theoretical physics provides an opportunity to search of adequate and useful analogues in socio-economic phenomena and processes.

Let us proceed to the description of structural elements of our work and building of the model. Suppose there is a set of M time series, each of N samples, that correspond to the single distance T, with an equal minimal time step Δt_{min}:

$$X_i(t_n), t_n = \Delta t_{min} \cdot n; n = 0, 1, 2, ..., N-1; i = 1, 2, ..., M. \quad (1)$$

To bring all series to the unified and non-dimentional representation, accurate to the additive constant, we normalize them, having taken a natural logarithm of each term of the series:

$$x_i(t_n) = \ln X_i(t_n), t_n = \Delta t_{min} n; n = 0, 1, 2, ..., N-1; i = 1, 2, ..., M. \quad (2)$$

Let us consider that every new series $x_i(t_n)$ is a one-dimensional trajectory of a certain fictitious or abstract particle numbered i, while its coordinate is registered after every time span Δt_{min}, and evaluate mean square deviations of its coordinate and speed in some time window ΔT:

$$\Delta T = \Delta N \cdot \Delta t_{min} = \Delta N, 1 << \Delta N << N. \quad (3)$$

The "immediate" speed i of particle at the moment t_n is defined by the ratio:

$$\nu_i(t_n) = \frac{x_i(t_{n+1}) - x_i(t_n)}{\Delta t_{min}} = \frac{1}{\Delta t_{min}} \ln \frac{X_i(t_{n+1})}{X_i(t_n)}. \quad (4)$$

with variance D_{ν_i} and mean square deviation $\Delta \nu_i$.

To evaluate dispersion D_{x_i} coordinates of the particle are used in an approximated ratio:

$$2D_{x_i} \approx D_{\Delta x_i}, \quad (5)$$

where

$$D_{\Delta x_i} =< (x_i(t_{n+1}) - x_i(t_n))^2 >_{n,\Delta N} -(< x_i(t_{n+1}) - x_i(t_n) >_{n,\Delta N})^2 = \\ =< \ln^2 \frac{X_i(t_{n+1})}{X_i(t_n)} >_{n,\Delta N} -(< \ln \frac{X_i(t_{n+1})}{X_i(t_n)} >_{n,\Delta N})^2, \quad (6)$$

which is derived from the supposition that x coordinates neighboring subject to the time of deviation from the average value $\bar{x}$ are weakly correlated. The chaotic nature of real time series allows to $x_i(t_n)$ as the trajectory of a certain abstract quantum particle (observed at Δt_{min} time spans). We

can write an uncertainty ratio for this trajectory [5]:

$$\Delta x_i \cdot \Delta \nu_i \sim \frac{h}{m_i}, or: \frac{1}{\Delta t_{min}}(< \ln^2 \frac{X_i(t_{n+1})}{X_i(t_n)} >_{n,\Delta N} - (< \ln \frac{X_i(t_{n+1})}{X_i(t_n)} >_{n,\Delta N})^2) \sim \frac{h}{m_i}, \quad (7)$$

where m_i – economic "mass" of an i series, h – value which comes as an economic Planck's constant.

Since the analogy with physical particle trajectory is merely formal, value, unlike the physical Planck's constant, can, generally speaking, depend on the historical period of time, for which the series are taken, and the length of the averaging interval (e.g. economical processes are different in the time of crisis and recession), on the series number etc. Whether this analogy is correct or not depends on particular series' properties.

To test the suggested ratios and definitions we have chosen stock markets time series with in one day for the period from January 3, 2004 to February 3, 2017 (https://finance.yahoo.com). Calculations were made in the model moving window. Elected window, for example, a 250 dots (approximate marketing year), it calculated value of the effective mass, then the window moved along a row with a step length in one day and again there was a calculation of the mass. This procedure continues until the exhaustion of the time series. The next step is to compare the dynamics of the original series with a corresponding effective mass dynamics, analysis of system changes on the market and the characteristics of effective mass behavior. Fig. 1 shows the "masses" US stock indexes (S&P 500 – sp), China (Hang Seng Index – hsi), Germany (DAX – dax) and Canada (S&P TSX Composite – gsptse). Obviously, there is a dynamic characteristic values M depending on the internal dynamics of the market. In times of crisis known (marked by arrows in the figures) "mass" is significantly reduced in the pre-crisis period.

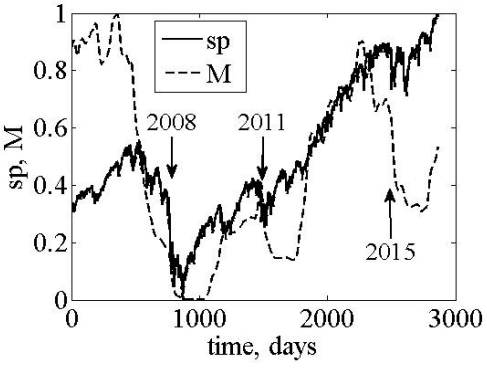


Figure 1

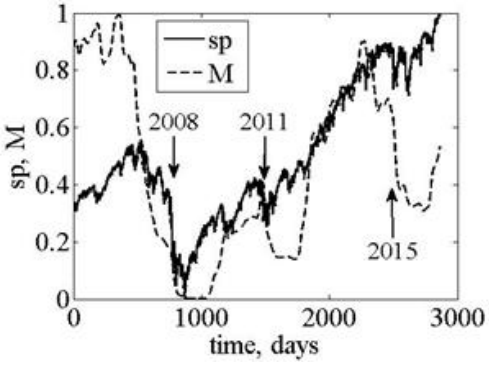


Figure 2

Comparison of original time series with the corresponding weight (Fig. 2) can watch proactive nature of mass change. This fact can be used as an indicator-precursor crisis. It should be noted that the results are robust with respect to the dynamic nature of the series (FX, spot markets, derivatives markets and so forth) and dimensioning procedures (window width, averaging procedures and so forth).

References. **1.** L.A. Rozema, A. Darabi, D.H. Mahler, A. Hayat, Y. Soudagar, A.M. Steinberg, "Violation of Heisenberg's Measurement-Disturbance Relationship by Weak Measurements", Phys. Rev. Lett., vol. 109, 100404, 2012. **2.** R. Prevedel R., D.R. Hamel, R. Colbeck, K. Fisher, and K.J. Resch, "Experimental investigation of the uncertainty principle in the presence of quantum memory", Nature Phys., vol.7, no.29, pp. 757–761, 2011. **3.** M. Berta, M. Christandl, R. Colbeck, J. Renes, R. Renner, "The Uncertainty Principle in the Presence of Quantum Memory", Nature Phys., vol. 6, no. 9, pp. 659–662, 2010. **4.** V. Saptsin, V. Soloviev (2009, Jul 7) Relativistic quantum econophysics – new paradigms in complex systems modelling, https://arxiv.org/0907.1142. **5.** V. Soloviev, V. Saptsin (2011, Nov 10) Heisenberg uncertainty principle and economic analogues of basic physical quantities, https://arxiv.org/1111.5289.

Titchiev I.I. [1,2]
[1] *Institute of Mathematics and Computer Science of the Academy of Sciences of Moldova, Chişinău, Republic of Moldova;* [2] *Tiraspol State Univesity, Chisinau, Republic of Moldova*

Scenario Analysis Related to the Field of Social Disasters by means of High Level Petri Nets

Lately humanity increasingly are facing with different kind of social disasters, which eventually can lead to new accidents and catastrophes. To analysis and mitigate their consequences, the formalism of High Level Petri nets has been proposed. It will examine the problem of people evacuation in useful time from the multi-storey buildings.

Introduction. Natural disasters, technogenic disasters, or executed by people disaster can occur daily. People expose their life to the dangerous impact of the environment and the economic factors, during these. Thus the problem of successfully evacuation of people from buildings is the actual one.

In [4] we study the scenario analysis for people evacuation from one storey buildings, but in many cases people working in multi-storey buildings. Evacuation from these is a very complicate and complex problem. In order to solve these one we propose to use the High Level Petri Nets (HLPNs). High Level Petri nets [1–3] provide a framework for the design, specification, validation, and verification of complex discrete event systems. They support data and functionality definitions, such as using complex structured data as tokens and algebraic expressions as transition formulas.

Creating large, complex nets can be a hard task. But similar to modular programming, the construction of complex nets can be broken into smaller pieces by utilizing the facilities within different tools for creating substitution transitions.

In hierarchical nets there is a method by which a transition can represent an entire piece of net structure, so that the net containing the transition executes as if the logic that the transition represents were physically present at the location of the transition. Such a transition is called a *substitution transition.*

Conceptually, nets with substitution transitions are nets with multiple layers of detail – you can have a somewhat simplified net that gives a broad overview of the system you are modelling, and by substituting transitions of this top-level net with more detailed pages, you can bring more and more detail into the model.

Relationship between ordinal nets and HLPNs:

- Given hierarchical net, it is easy to construct an equivalent ordinal (classical) net and vice versa.
- A hierarchical net and its equivalent non-hierarchical net have exactly the same set of markings, steps and occurrence sequences.
- A hierarchical net and its equivalent non-hierarchical net are behaviorally equivalent.
- Computer analysis with many software tools takes almost the same time for both a hierarchical net and its equivalent non-hierarchical net.

High-Level Petri Nets. High Level Petri-nets assets:

1. High Level Petri-nets have a intuitive graphical representation and a well-defined semantics which unambiguously defines the behaviour of each HLPNs.
2. They are very general and can be used to describe a large variety of different systems.
3. HLPNs have very few, but powerful, primitives, an explicit description of both states and actions.
4. Are stable towards minor changes of the modelled system.
5. A formal analysis methods allow to proved properties of HLPNs.
6. The two most important analysis methods are known as occurrence graphs and place invariants. Computer tools supporting their drawing, simulation and formal analysis, exist.

17.80013.5007.01/Ua „Development of modelling tools for mitigation strategies of social disaster caused by catastrophe and terrorism" (DIMSAD)

Definition 1. A *High-level Petri Nets* is a structure $HLPN = (P; T; D; Type; Pre; Post; M_0)$ where

- P is a finite set of elements called Places.
- T is a finite set of elements called Transitions disjoint from $P(P \bigcap T = \emptyset;)$.
- D is a non-empty finite set of non-empty domains where each element of D is called a *type*.
- $Type : P \bigcup T \to D$ is a function used to assign types to places and to determine transition modes.
- $Pre; Post : TRANS \to \mu PLACE$ are the pre and post mappings with
  $TRANS = \{(t; m) | t \in T; m \in Type(t)\}$
  $PLACE = \{(p; g) | p \in P; g \in Type(p)\}$
- $M_0 \in \mu PLACE$ is a multiset called the initial marking of the net.

A *Marking* of the HLPN is a multiset, $M \in \mu PLACE$.

A transition is enabled with respect to a *net marking* or in a particular *transition mode*. A transition mode is an assignment of values to the transition's variables, that satisfies the transition condition (i.e., the transition condition is true). The transition's variables are all those variables that occur in the expressions associated with the transition. These are the transition condition, and the annotations of arcs involving the transition.

A finite multiset of transition modes, $T \in \mu TRANS$, is enabled at a marking M iff $Pre(T_\mu) \leqslant M$.

A step may occur resulting in a new marking M' given by $M' = M - Pre(T_\mu) + Post(T_\mu)$.

Evacuation system. In any building their evacuation plan are meeting, in these planes we will denote rooms, doors like in [4], Multi-storey buildings modeled by Petri nets can be organized into hierarchies of layers like in Figure 1. This gives rise to hierarchical nets that, although in all perfectly equivalent to ordinary nets, can facilitate the management of such large-scale models.

A transition mode for the evacuation system from multi-storey building is the assignment of values to the transition's variables, that satisfies the transition condition [4] (i.e., speed of flows, walking speed, people's density in the rooms, transfer function, flow rate).

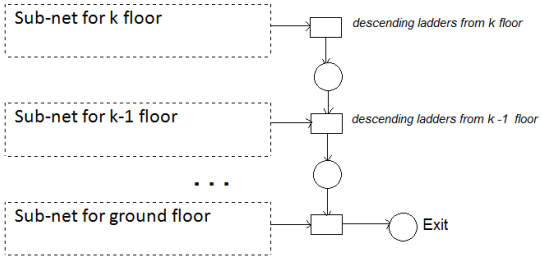


Figure 1. Evacuation system for multi-storey building

Conclusions. For the problem of mitigation of some consequences of disasters, a method of High Level Petri Nets was proposed. They allow to model and simulate of the system that represent emergency evacuation of people in case of social disaster from multi-storey buildings. By this method it possible to model such large systems in a manageable and modular way.

References. **1.** He X., Murata T. *High-Level Petri Nets – Extensions, Analysis, and Applications*, Electrical Engineering Handbook (ed. Wai-Kai Chen), Elsevier Academic Press, 2005, pp. 459–476. **2.** Jensen K., *An Introduction to High-level Petri Nets*. Proceedings of the 1985 International Symposium on Circuits and Systems: Kyoto 85, pp. 723–726, Kyoto, Japan, 1985. **3.** Jensen K., Rozenberg G. *High-level Petri Nets: Theory and Applications*. Springer-Verlag Eds., p. 724, London, UK, 1991. **4.** Cojocaru S., Petic M., Titchiev I., *Adapting Tools for Text Monitoring and for Scenario Analysis Related to the Field of Social Disasters*, In the proceedings of The 18th International Conference on Computer Science and Electrical Engineering (ICCSEE 2016), October 6–7, 2016, Prague, Czech Republic, pp. 886–892.

Безсмертна А.В.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Особливості побудови рекомендаційної системи для користувачів телебачення

Підходи до побудови рекомендаційних систем. РС зазвичай формують список рекомендацій одним з двох способів – за допомогою колаборативної та контентної фільтрації або на основі індивідуального підходу. Підхід колаборативної фільтрації [1, 2] будує модель на основі поведінки користувача в минулому (об'єкти, раніше куплені або обрані та/або оцінки, надані цим об'єктам), а також аналогічних дій інших користувачів. Ця модель використовується для прогнозування об'єктів (або оцінок для цих об'єктів), які можуть викликати інтерес у користувача. Контентний підхід використовує серію дискретних характеристик об'єкта, щоб рекомендувати додаткові об'єкти з подібними характеристиками.

Недоліком *контентної фільтрації* є необхідність глибокого розуміння характеристик об'єкта. А для збору описової інформації про об'єкти необхідні значні затрати ресурсів. Також рекомендаційні системи, побудовані за принципом контентної фільтрації мають малу степінь новизни. *Колаборативна фільтрація* дає досить велику точність передбачення уподобань користувача та базується на історії активності як користувача, якому власне і буде надаватись рекомендація, так і всіх інших користувачів з клієнтської бази.

В основі кожної рекомендаційної системи лежить певний рекомендаційний механізм, реалізований за допомогою методів інтелектуального аналізу даних. В залежності від наявних даних, предметної області, цільової аудиторії, способу застосування рекомендаційної системи, обирається підхід і відповідні йому методи аналізу.

Критерії оцінки якості рекомендаційних систем:

- Точність
- Швидкість адаптації
- Степінь новизни
- Вирішення проблеми "холодного старту"

Побудова моделі. Головною *проблемою* побудови рекомендаційної системи є динамічність телепрограми. Спостерігається деяка щотижнева періодичність, однак щодня виходять нові фільми, серіали та телепередачі. Тому використання класичного принципу колаборативної фільтрації в даному випадку не є ефективним. Для того, щоб позбутися недоліків колаборативної фільтрації та поєднати переваги колаборативної та контентної фільтрації, доцільно побудувати модифіковану модель.

Для того, щоб сформувати для кожного користувача набір рекомендацій на майбутній тиждень, необхідно провести колаборативну фільтрацію по користувачам на основі сусідства, часову категоризацію, жанрову кластеризацію.

Колаборативна фільтрація буде формувати рекомендації того телеконтенту, який вже був у показі і по якому вже є історія його перегляду користувачами. Для формування степені уподобання користувачем того чи іншого контенту використовуються дані про тривалість та частоту перегляду.

Для формування рекомендацій користувачам нового контенту, використовується *часова категоризація* і *жанрова кластеризація*. Користувачів поділяють на категорії в залежності від того, в який період доби він віддає перевагу дивитися телебачення, а також які жанри та який тип телеконтенту йому більше подобається. За допомогою цих методів можна реалізувати рекомендаційну систему, яка буде відповідати поставленим критеріям оцінки якості.

Література. **1.** Воронцов К.В. Колаборативная фильтрация и матричное разложение, 2016 [Електронний ресурс]. – Режим доступу: Voron-ML-CF.pdf. **2.** Хомутов Н.Ю. Системы прогнозирования предпочтений пользователей на основе их действий, 2015 [Електронний ресурс]. – Режим доступу: KhomutovNY.pdf.

Бігун С.О., Іваненко Д.О.
Київський національний університет ім. Тараса Шевченка, Київ, Україна

Реалізація алгоритму перевірки ефективності методу оцінювання в статистичних моделях, керованих шумом Леві

Для статистичних моделей зі сталим кроком спостережень процесу, заданого СДР з шумом Леві, реалізовано алгоритм перевірки ефективності методу оцінювання, в якому мінімізовано втрати інформації при умовному усередненні.

Розглядається процес X, що є розв'язком рівняння

$$dX_t = a_\theta(X_t)dt + dZ_t, \tag{1}$$

в якому Z – процес Леві без дифузійної компоненти, тобто

$$Z_t = ct + \int_0^t \int_{|u|>1} u\nu(ds, du) + \int_0^t \int_{|u|\leqslant 1} u\tilde{\nu}(ds, du).$$

Тут ν – пуассонова точкова міра з інтенсивністю $ds\mu(du)$, а $\tilde{\nu}(ds, du) = \nu(ds, du) - ds\mu(du)$ – відповідна компенсована міра.

В роботі [1] був запропонований алгоритм перевірки ефективності методу оцінювання невідомого параметра θ у випадку, коли спостереження відбуваються зі сталим кроком h. Алгоритм полягає в наступному:

1. вибирається метод оцінювання (наприклад, найменших квадратів) і будується оцінка $\hat{\theta}_n$ невідомого параметра θ_0;
2. генерується N траєкторій процеса X з $\theta = \hat{\theta}_n$ і по кожній з них будується вибірка розміру n;
3. обчислюються оцінки $\hat{\theta}_n^k$, $k = 1, \ldots, N$ і знаходиться вибіркова дисперсія $s_N^2(\hat{\theta}_n) = \frac{1}{N}\sum_{k=1}^N \left(\sqrt{n}(\hat{\theta}_n^k - \hat{\theta}_n)\right)^2$;
4. обчислюється n_0 з міркувань близькості X до стаціонарного розв'язку рівняння;
5. генерується N траєкторій X з $\theta = \hat{\theta}_n$, по кожній з них обчислюється деякий функціонал $\Xi_{h,k}^1(n_0)$, $k = 1, \ldots, N$, складниками якого є стохастичні інтеграли по пуассоновій мірі;
6. знаходиться вибіркове середнє $J_N(\hat{\theta}_n, n_0) = \frac{1}{N}\sum_{k=1}^N \Xi_{h,k}^1(n_0)^2$;
7. за значенням $\sqrt{J_N(\hat{\theta}_n, n_0)s_N^2(\hat{\theta}_n)}$ робиться висновок про ефективність (при фіксованих параметрах N і n_0) методу і про втрату випадковості при умовному усередненні.

Величина $J_N(\hat{\theta}_n, n_0)$ є оцінкою інформації Фішера. Згідно з [1] більш точною оцінкою є середнє умовних сподівань по так званому мосту від функціоналів $\Xi_{h,k}^1(n_0)^2$. При чому втрата інформації при заміні умовного сподівання на безумовне сягає в окремих випадках 75 відсотків. Однак, така заміна, була виправдана, оскільки для реалізації умовного усереднення виникала потреба у потужному машинному ресурсі і часі.

В даній роботі реалізується уточнення запропонованого в [1] алгоритму за рахунок розпаралелювання з виконанням на графічному адаптері. Найбільша кількість операцій в алгоритмі припадає на генерацію траєкторій процесу, що виконується в циклі з передумовою.

Через відсутність в стандартних бібліотеках OpenCL генератору псевдовипадкових послідовностей для цієї мети було використано генератор на основі XORShift. Крім того, розроблена реалізація надає можливість вибирати з доступного устаткування пристрій, на якому будуть виконуватись обчислення.

Література. **1.** Д.О. Іваненко, С.В. Боднарчук, Оцінка ефективності методу оцінювання в статистичних моделях, керованих шумом Леві, Теорія ймовірностей та математична статистика. 92 (2015) 9–22.

***Білущак Ю.І.*[1], *Гончарук В.Є.*[2], *Чернуха О.Ю.*[1], *Чучвара А.Є.*[1]**
[1] *Центр математичного моделювання Інституту прикладних проблем механіки і математики ім. Я.С.Підстригача НАН України, Львів, Україна;* [2] *Національний університет "Львівська політехніка", Львів, Україна*

Математичне моделювання дифузійних процесів у випадкових шаруватих структурах за явного врахування стрибків концентрації та її похідної на міжфазних границях

Природні та штучні матеріали, які широко використовують в інженерній практиці, мають складну, зокрема багатофазну, структуру. Їхнє застосування вимагає оцінки розподілів дифузійних полів та їх поведінки у залежності від умов внутрішнього міжфазного контакту, зовнішніх дій і можливих просторових реалізацій структури. При цьому для таких середовищ, як правило, є невідомими дані про конкретне просторове розташування окремих фаз, проте достатньо інформації про їхній дольовий вміст та основні фізико-хімічні властивості.

Нехай домішкові частинки мігрують у шарі товщиною z_0, який складається з підшарів двох фаз, розташування яких є невідомим. Вважаємо, що дифузійні властивості фаз (області Ω_0 і Ω_1) можуть суттєво відрізнятися, фази в тілі розташовані за рівномірним розподілом і об'ємна частка v_1 області Ω_1 є набагато меншою за об'ємну частку v_0 області Ω_0 ($v_1 << v_0$).

Концентрація домішкових частинок $c_j(z,t)$ в області Ω_j визначається з рівняння дифузії

$$\rho_j \frac{\partial c_j(z,t)}{\partial t} = d_j \frac{\partial^2 c_j(z,t)}{\partial z^2},\ z \in \Omega_j = \bigcup_{i=1}^{n_j} \Omega_{ij},\ t \in [0,\tau]\ (\tau < \infty),\ j = 0,1, \tag{1}$$

де ρ_j – густина області Ω_j; d_j – кінетичний коефіцієнт дифузії частинок в цій області, n_j – кількість підшарів фази j, Ω_{ij} – i-та однозв'язна область фази j, $i = \overline{1,n_j}, j = 0,1$.

Приймаємо нульові початкові умови, а також, що на границі тіла $z = 0$ підтримується постійне значення концентрації c_*, а на нижній границі $z = z_0$ – вона дорівнює нулю, тобто

$$c_0(z,t)\big|_{t=0} = c_1(z,t)\big|_{t=0} = 0, \quad c_0(z,t)\big|_{z=0} = c_* \equiv const, \quad c_0(z,t)\big|_{z=z_0} = 0, \tag{2}$$

умови неідеального контакту на концентрацію у вигляді

$$\big[c(z,t)\big]_{z=z_l} = k_0 c_0(z,t)\big|_{z=z_l+0} - k_1 c_1(z,t)\big|_{z=z_l-0} = \left(1 - \frac{k_0}{k_1}\right) c_0(z,t)\big|_{z=z_l-0}, \tag{3}$$

$$\left[\frac{\partial c(z,t)}{\partial z}\right]_{z=z_l} = \rho_0 d_0 \frac{\partial c_0}{\partial z}\Big|_{z=z_l+0} - \rho_1 d_1 \frac{\partial c_1}{\partial z}\Big|_{z=z_l-0} = \left(1 - \frac{\rho_0 d_0}{\rho_1 d_1}\right) \frac{\partial c_0(z,t)}{\partial z}\bigg|_{z=z_l-0}, \tag{4}$$

$$\big[c(z,t)\big]_{z=z_l+h_{l1}} = k_0 c_0(z,t)\big|_{z=z_l+h_{l1}+0} - k_1 c_1(z,t)\big|_{z=z_l+h_{l1}-0} = \left(\frac{k_0}{k_1} - 1\right) c_0(z,t)\big|_{z=z_l+h_{l1}+0}, \tag{5}$$

$$\left[\frac{\partial c}{\partial z}\right]_{z=z_l+h_{l1}} = \rho_1 d_1 \frac{\partial c_1}{\partial z}\Big|_{z=z_l+h_{l1}+0} - \rho_0 d_0 \frac{\partial c_0}{\partial z}\Big|_{z=z_l+h_{l1}-0} = \left(\frac{\rho_0 d_0}{\rho_1 d_1} - 1\right) \frac{\partial c_0}{\partial z}\bigg|_{z=z_l+h_{l1}-0}, \tag{6}$$

де $\big[c(z,t)\big]_{z=z_l}$ – стрибок функції $c(z,t)$ в точці z_l. За такої постановки задачі випадковими величинами є границі контакту $z = z_l$ та $z = z_l + h_{l1}$, тобто межі Ω_0 та Ω_1, які є внутрішніми для тіла. Це, в свою чергу, призводить до стохастичності поля концентрації домішки.

Розв'язок сформульованої контактно-крайової задачі (1)–(6) шукаємо у вигляді інтегрального ряду Неймана, для чого вводимо у розгляд випадкову функцію просторової координати $c(z,t)$, яка описує концентрацію в усьому тілі; з контактної задачі (1), (3)–(6) та рівняння балансу маси знаходимо диференціальне рівняння, якому підпорядковується функція $c(z,t)$. Запропоновано нове представлення оператора рівняння дифузії для всього тіла, яке явно враховує стрибки функції концентрації та її похідної на границях контакту фаз. Побудовано інтегро-диференціальне рівняння, еквівалентне вихідній контактно-карайовій задачі, яке розв'язано шляхом ітерування. Розв'язок отриманий у вигляді ряду Неймана. Усереднення поля концентрації проведено за ансамблем конфігурацій фаз з рівномірною функцією розподілу

включень. У результаті отримано розрахункову формулу вигляду

$$\frac{\langle c(z,t)\rangle}{c_*} \approx 1 - \frac{z}{z_0} - \frac{2}{z_0}\sum_{n=1}^{\infty} e^{-d_0 y_n^2 t/\rho_0}\sin(y_n z) + \frac{4v_1}{z_0^2}\sum_{n=1}^{\infty}[\sin(y_n z)\times$$

$$\times e^{-d_0 y_n^2 t/\rho_0}\left[\frac{d_0 R_n t}{8\rho_0 y_n}\left(1-\frac{\rho_1}{\rho_0}\right) + \sum_{\substack{m=1\\ m\neq n}}^{\infty}\sum_{n=1}^{\infty} A_{mn} y_m \left(\frac{d_1}{d_0}-\frac{\rho_1}{\rho_0}\right) A_e\right] + I_2(z,t);$$

$$I_2(z,t) = \frac{4v_1}{h z_0^2 \rho_0} K_d \left[\sum_{m=1}^{\infty}\sin(y_n z)(A_m + A_r) - \frac{2\rho_0}{d_0}\sum_{\substack{m=1\\ m\neq n}}^{\infty}\sum_{n=1}^{\infty} A_e \sin(y_m z) B_{nm}\right].$$

Ця формула відрізняється від аналогічної формули задачі [1] з іншим представленням оператора рівняння дифузії виникненням доданка I_2, який пропорційний коефіцієнту $K_d = 2 - \frac{\rho_0 d_0}{\rho_1 d_1} - \frac{k_0}{k_1}$. Коефіцієнт K_d може бути додатнім або від'ємним, відповідно збільшуючи або зменшуючи усереднену концентрацію. Причому коефіцієнти дифузії можуть різнитися на порядки, а коефіцієнти концентраційної залежності хімічного потенціалу – в рази. Тоді у випадках значно менших коефіцієнтів у включеннях ніж в матриці співвідношення $\frac{\rho_0 d_0}{\rho_1 d_1}$ набуває великих значень, а співвідношення $\frac{k_0}{k_1}$ стає нехтовно малим незалежно від значень k_j. Коли коефіцієнт k_1 у включеннях значно більший ніж в матриці – кількісне співвідношення $\frac{\rho_0 d_0}{\rho_1 d_1}$ може стати нехтовно малим.

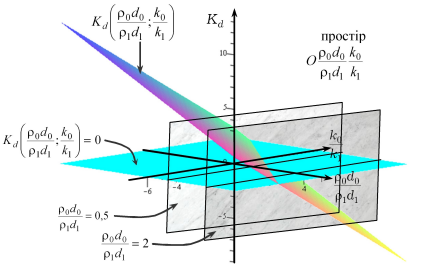


Рис. 1. 3D-поверхня K_d

На рис 1. наведено 3D-поверхню K_d. По осі абсцис відкладено відношення k_j, по осі ординат – відношення коефіцієнтів дифузії, а аплікатою є коефіцієнт K_d як функція від відношення коефіцієнтів дифузії та коефіцієнтів концентраційної залежності хімічних потенціалів в матриці та включеннях. Зміна знаку K_d відбувається при перерізі поверхні площиною $K_d(\frac{k_0}{k_1}, \frac{\rho_0 d_0}{\rho_1 d_1}) = 0$ (рис. 1). Щоби отримати залежність від відношення $\frac{k_0}{k_1}$ зроблено перерізи поверхні K_d паралельними площинами для $\frac{\rho_0 d_0}{\rho_1 d_1} = const$. На рис 2. наведено проекцію таких перерізів на площину $O(\frac{k_0}{k_1} K_d)$, де по осі абсцис відкладене відношення $\frac{k_0}{k_1}$, а прямі пораховані для різних значень $\frac{\rho_0 d_0}{\rho_1 d_1}$ =0.01; 0.1; 0.5; 1; 2; 5; 10 (криві 1–7).

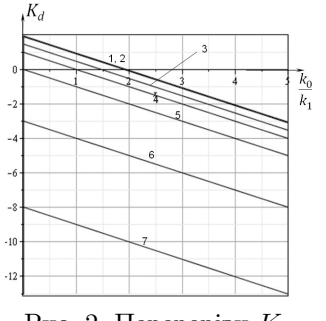


Рис. 2. Перерерізи K_d

Зазначимо, що чим більші значення приймають відношення коефіцієнтів дифузії або відношення k_j тим менших значень набуває коефіцієнт K_d. Причому K_d не міняє знак, якщо $\frac{\rho_0 d_0}{\rho_1 d_1} > 2$. В інших випадках зміна знаку цього коефіцієнта відбувається при значеннях $\frac{\rho_0 d_0}{\rho_1 d_1} < 2$ разом з відповідним зростанням значень $\frac{k_0}{k_1}$. Тобто чим менше відношення коефіцієнтів дифузії тим при більших значеннях відношень коефіцієнтів відбувається зміна знаку K_d. Зазначимо, що хоча відношення коефіцієнтів $\frac{k_0}{k_1}$ і $\frac{\rho_0 d_0}{\rho_1 d_1}$ є взаємно незалежними, але на зміну знаку K_d вони впливають у комплексі. Проведено числові експерименти для дослідження функції усередненої концентрації, які виявили три екстремуми в поведінці нового доданку функції $\langle c(z,t)\rangle$, причому величина останнього екстремуму є завжди є найбільшою. Встановлено області значень параметрів задачі, за яких цей доданок є нехтовно малим.

Література. 1. Chaplya Y., Chernukha O., Bilushchak Y. (2012) Contact initial boundary-value problem of the diffusion of admixture particles in a two-phase stochastically inhomogeneous stratified strip. J. Math. Sci., 183(1), P. 83–99.

Бойко І.О., Бідюк П.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Застосування скорингового аналізу для оцінювання платоспроможності клієнтів

Одним із перспективних напрямів оцінки платоспроможності позичальника є використання системи кредитного скорингу, яка базується на бальній оцінці факторів кредитного ризику. Скоринг є математичною або статистичною моделлю, за допомогою якої на основі кредитної історії "минулих" клієнтів банк намагається визначити, яка ймовірність того, що конкретний потенційний позичальник поверне кредит у визначений строк.

Для аналізу якості моделей і встановлення найкращої моделі для розв'язання даної задачі використаємо декілька критеріїв для оцінювання адекватності моделей [1]: загальна точність моделі; помилки I-го та II-го роду.

Реалізація моделі відбувалася у два етапи. Перший – оцінювання ймовірності повернення кредиту, другий – визначення класу позичальника. Всі моделі побудовано на реальних даних фінансової звітності банку.

Для оцінювання платоспроможності клієнтів було обрано такі моделі:

Нечітка логіка. Системи нечіткого логічного висновку перетворюють значення вхідних змінних у вихідні з використанням нечітких правил. До недоліків такої системи належить те, що вона не адаптується самостійно після отримання нових даних, а серед переваг можна зазначити те, що вона дає змогу отримати кількісну оцінку із лінгвістичних вхідних змінних, а також ефективне відображення залежності цих змінних у вигляді нечітких правил.

Дерева рішень. Дерева прийняття рішень – послідовні ієрархічні структури, що складаються з вузлів, які містять правила, тобто логічні конструкції вигляду "якщо ... то ...". Кінцевими вузлами дерева є "листя", відповідне знайденим рішенням і об'єднуючі деяку кількість об'єктів вибірки, що класифікуються. До переваг моделі можна віднести можливість одночасно працювати із змінними різних типів, відсутність попередніх припущень про закони розподілу даних. Було побудовано 4 типи дерев рішень: CHAID, Exhaustive CHAID, CART, QUEST. Показники точності та помилки було усереднено для подальшого порівняння з іншими моделями.

Рейтинговий підхід. Для побудови такої моделі обираються вхідні параметри (наприклад, дохід позичальника за певний період). Для кожного параметру задають експертні оцінки в балах. На виході маємо суму балів для позичальника по кожному з вхідних параметрів. Якщо сума балів позичальника більша, ніж поріг, визначений експертами, то клієнт визначається як платоспроможний, інакше – неплатоспроможний. У такої моделі є два недоліки – суб'єктивність оцінок експертів, а також відсутність її адаптивності. До переваг моделі належить її простота.

Табл. 1. Порівняння точності моделей та помилок 1-го та 2-го роду.

Модель/ Оцінка	Точність моделі	Помилка 1-го роду	Помилка 2-го роду
Нечітка логіка	0,755	0,175	0,07
Дерева рішень	0,7395	0,1603	0,1005

За результатами порівняльного аналізу характеристик моделей оцінювання платоспроможності клієнтів найкращою виявилась модель нечіткої логіки, з показниками: "Точність моделі" = 0,755; "Помилка 1-го роду" = 0,175; "Помилка 2-го роду" = 0,07.

Література. 1. Бідюк П.І. Порівняльний аналіз характеристик моделей оцінювання ризиків кредитування / Бідюк П.І., Кузнєцова Н.В // Наукові вісті НТУУ "КПІ". – 2010. – № 1. – С. 42–53.

Болдак А.О., Лазаренко І.С.
Світовий центр даних з геоінформатики та сталого розвитку КПІ ім. Ігоря Сікорського, Київ, Україна

Алгебраїчний підхід до верифікації моделей складних систем на прикладі побудови індексу нерозповсюдження зброї

Побудову ієрархічної моделі нерозповсюдження зброї характеризує ряд особливостей, притаманних складним системам:

- відсутність формалізованого математичного опису та загальної концепції побудови моделі;
- складність спостереження та управління системою, оскільки множина впливаючих факторів неоднозначно визначається в силу конфіденційності та специфіки об'єкта дослідження;
- нестаціонарність процесів в системі та еволюція в часі;
- неможливість проведення активних експериментів, та адаптації системи до поточного стану.

Тому доцільно на описовому етапі побудови моделі мінімізувати невизначеність та неповноту даних, підвищити адекватність експериментального моделювання за рахунок розробки та використання нових методів верифікації моделей, методів оцінювання невизначеності моделей а також методів та засобів узгодження даних різної природи.

Предметну область представляє собою частину реального світу, яка розглядається в границях певного контексту, пов'язаного з науковими дослідженнями та (або) певною діяльністю.

Прикладна теорія, в свою чергу, є сукупністю моделей та співвідношень між ними, які використовуються для вирішення задач, сформульованих в рамках певної предметної області.

Під моделлю будемо розуміти: `equation`:

$$M = \langle S, C(S) \rangle, \tag{1}$$

де $S = \{s_i\}, s = \overline{1,n}$ – сукупність властивостей, яка обирається при побудові моделі, а $C(S) = \{c_j\} : c_j \subseteq s_1^{N_1} \times ... \times s_n^{N_n}$ – контекст моделі, який задано в термінах властивостей S.

В даній постановці властивості s відповідає індикаторна функція $\chi_S : O \to Z$, яка для кожного об'єкту предметної області $o \in O$ визначає деякий ступінь прояву $z \in Z$.

Для даної моделі, в контексті теорії множин, визначено операції доповнення, узагальнення та агрегування.

Сформулюємо властивості прикладної теорії, яка задана в термінах моделей (1) і операцій доповнення, узагальнення та агрегування.

Твердження 1. Множина моделей прикладної теорії з операціями доповнення, узагальнення та агрегування утворюють асоціативне кільце з одиницею $M(S_U)$ та нейтральним елементом $M(\emptyset)$ відносно агрегування.

Твердження 2. Властивість є моделлю.

Твердження 3. Для будь-якої підмножини об'єктів предметної області може бути побудована модель.

Твердження 4. Універсальна множина властивостей може бути визначена, як

$$S_U = \bigcup_{i=1}^{n_1} S_i : S_i \subseteq S_{INPUT}, n_1 = |S_{IMPUT}|.$$

Застосування запропонованого підходу до побудови моделі неросповсюдження зброї дає змогу позбутися явища мультиколінеарності, повторного використання властивостей, які є складовими різних чинників, знівелювати неповноту даних на етапі побудови концептуальної моделі.

Болдак Л.О., Горох О.С.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Моделювання сценаріїв за допомогою баєсівських мереж довіри

Один з підходів до керування такими складними системами, якими є соціо-економічні системи, це *сценарне планування*, яке передбачає побудову оптимізованої стратегії керування на основі аналізу послідовності подій, які відображають керуючі впливи та реакції об'єкту керування. У цьому випадку під *стратегією керування* розуміється послідовність кроків (дій), які з певною ймовірністю можуть привести до бажаного або запланованого кінцевого стану об'єкту керування, а *сценарій* розглядається як послідовність згаданих вище подій, що відповідає критеріям правдоподібності та внутрішньої узгодженості [1].

Виходячи з системного підходу до моделювання об'єкту керування, будемо описувати керуючі впливи X та реакції об'єкту керування Y в термінах подій, а в якості моделі поведінки об'єкту керування використовувати модель його стану S. Для об'єктів керування, що розглядаються, в разі відсутності формальної моделі стану пропонується використати експертно-статистичну модель причинно-наслідкових зв'язків між змінними $s_i \in S$, $i = \overline{1,n}$, що в сукупності характеризують стан об'єкту керування, та для кожної з яких відома область дискретних значень. Події-керуючі впливи E_x і події-реакції E_y можуть бути внесені в таку модель за умови, якщо для кожної з них задати область значень як $\{0,1\}$ та визначені причинно-наслідкові зв'язки $E_x \to S$ та $S \to E_y$.

Якщо для E_x експертним шляхом визначені апріорні імовірності, а для S та E_y умовні імовірності для кожного з значень відповідних областей визначення, то можна використати модель *баєсівської мережі довіри* (БМД) для моделювання сценарію як часової послідовності логічно пов'язаних та взаємоузгоджених подій.

$$G \subseteq (V_a \times V_b \times P(b|a)), V_a = V_b = E_x \cup E_y \cup S,$$

де V_a, V_b – змінні, G – граф причинно-наслідкових зв'язків зважений умовними імовірностями. Очевидно, що таку модель можна розділити на три шари:

1. Шар регулювання об'єкту – це множина керуючих подій, впорядкування якої за часом описує *сценарій впливу*.
2. Шар стану об'єкту.
3. Шар реакції об'єкту – це множина відповідних подій, впорядкування якої за часом описує *сценарій поведінки*.

Використання відомих ефективних методів прямого та зворотнього імовірнісного виводу дає можливість планування послідовності керуючих впливів (побудови стратегії) та прогнозування поведінки об'єкту управління, виходячи з заданої обмеженої області значень стану або множини подій-реакцій.

Можливості використання описаної моделі як ефективного інструменту для систем підтримки прийняття рішень (СППР) на основі сценарного планування обмежуються методом гілки [1]. У цьому випадку СППР повинна підтримувати наступні бізнес-процеси.

Визначення та уточнення сценаріїв, який виконується експертною групою. Суть його полягає у побудові сценаріїв впливу, у побудові моделі стану об'єкта управління та БМД.

Сценарне планування, яке здійснюється особою, що приймає рішення, з метою прогнозування сценарію поведінки об'єкта керування та аналізу причин виникнення того чи іншого стану або події-реакції, що дозволяє їй визначити власну стратегію дій для досягнення чи уникнення певних вихідних подій.

Література. **1.** Форсайт та побудова стратегії соціально-економічного розвитку України на середньостроковому (до 2020 року) і довгостроковому (до 2030 року) часових горизонтах / наук. керівник проекту акад. М.З. Згуровський. – Київ: НТУУ "КПІ ім. Ігоря Сікорського", Вид-во "Політехніка", 2016. – 184 с. **2.** Тулупьев А.Л. Байесовские сети: логико-вероятностный подход / А.Л. Тулупьев, С.И. Николенко, А.В. Сироткин – СПб.: Наука, 2006. – 607 с.

***Болжеларський Я.В.*[1], *Дмитрук В.А.*[2,3], *Гера Б.В.*[1]**

[1] *Львівська філія Дніпропетровського національного університету залізничного транспорту ім. акад. В.Лазаряна, Львів, Україна;* [2] *Центр математичного моделювання Інституту прикладних проблем механіки і математики ім. Я.С.Підстригача НАН України, Львів, Україна;* [3] *Національний університет "Львівська політехніка", Львів, Україна*

Математична модель динамічної поведінки рухомої фізико-механічної системи при дії зосереджених та розподілених сил опору рухові

Дослідження стану фізико-механічної системи, що рухається під дією сили тяги, та зазнає дії різноманітних факторів, котрі чинять опір рухові, має важливе значення для об'єктів транспорту. При зростанні швидкості транспортних засобів сили опору зростають і суттєво впливають на деформації і напруження у їх елементах. Особлива увага при цьому приділяється нештатним ситуаціям: екстреному гальмуванню, потрапляння сторонніх об'єктів на колію, поломки рухомого складу тощо [1]. Опір рухові, що супроводжується втратами енергії руху, створюється як всередині системи через тертя між її елементами, так і внаслідок взаємодії системи з оточуючим середовищем. При математичному моделюванні силової дії на рухому систему отримуємо розподілені сили, що представляють усю сукупність сил протидії рухові.

На залізничному транспорті рухомі об'єкти мають видовжену форму. Розглядатимемо їх на моделі складеного в'язко-пружного поздовжньо неоднорідного стрижня, на одному з кінців котрого прикладена сила тяги спрямована вздовж осі, а у протилежному напрямку розподілені сили опору, залежні від характеристик елементів стрижня і швидкості поздовжнього руху системи в цілому.

Поздовжньо неоднорідний, складений з різнорідних елементів в'язко-пружний стрижень, що рухається під дією сили F_0, прикладеної до одного з його країв. Якщо встановилась постійна швидкість руху, то пружні динамічні ефекти в стрижні не проявляються і лише поява додаткового інтенсивного навантаження P_0 в напрямку протилежному рухові призводить до поздовжніх динамічних деформацій і напружень. Ці напруження накладаються на квазістатичні напруження, що виникають від сил опору ϖ, котрі можна записати виходячи з їх закономірностей при взаємодії системи із зовнішнім середовищем.

Рівняння для визначення переміщень точок неоднорідного стрижня [2], запишемо у вигляді

$$\rho(x)\frac{\partial^2 u(x,\,t)}{\partial t^2} - \frac{\partial}{\partial x}\left(k(x)\left(\frac{\partial u(x,\,t)}{\partial x} + \mu\frac{\partial^2 u(x,t)}{\partial x \partial t}\right)\right) = -\varpi(x,t) - \delta(x - x_*)P_0(t),$$

$$x \in (0;\ L), \quad t > 0, \tag{1}$$

де позначено x координату вздовж осі, t – час, $u(x,t)$ – переміщення точок стрижня, $\rho(x) = \rho_V(x)S(x)$ приведену до осі густину стрижня, де $\rho_V(x)$ функція густини, $S(x)$ – площа поперечного перерізу стрижня, $k(x) = E(x)S(x)$ приведений модуль пружності, $E(x)$ – модуль пружності на розтяг, μ – коефіцієнт непружного опору матеріалу при поздовжніх коливаннях, $\varpi(x,t)$ розподілені по довжині стрижня сили опору рухові, $P_0(t)$ – зосереджена в точці $x = x_*$ додаткова сила опору.

Рівняння (1) розглядаємо з граничними умовами, що відповідають відсутності зовнішньої дії на краю $x = 0$ і заданню сили $F_0(t)$ на границі $x = L$, а саме

$$\left(\frac{\partial u(x,t)}{\partial x} + \mu\frac{\partial^2 u(x,t)}{\partial x \partial t}\right)\Big|_{x=0} = 0, \text{ та} \quad k(x)\left(\frac{\partial u(x,t)}{\partial x} + \mu\frac{\partial^2 u(x,t)}{\partial x \partial t}\right)\Big|_{x=L} = F_0(t), \quad t > 0. \tag{2}$$

Початкові умови на переміщення і відносну швидкість наступні

$$u(x,\,0) = f_0(x), \quad \frac{\partial u(x,\,0)}{\partial t} = f_1(x), \quad x \in (0;\ L). \tag{3}$$

Функція $f_0(x)$ задає початковий розтяг стрижня, а $f_1(x)$ розподіл швидкості у початковий момент часу. Поздовжня сила вздовж стрижня виражається через переміщення наступним чином

$$\Phi(x,t) = k(x)\left(\frac{\partial u(x,t)}{\partial x} + \mu\frac{\partial^2 u(x,t)}{\partial x \partial t}\right).$$

Початково-крайову задачу (1–3) розв'язуємо використовуючи метод розділення змінних і представляючи розв'язок у вигляді ряду по власних функціях.

Для проведення числових розрахунків отримаємо розподілені сили опору, початкові умови і визначаємося з коефіцієнтами рівняння. Для прикладу рухомого поїзда записуємо дані стосовно стрижневої системи, що складається з n елементів і рухається зі сталою швидкістю. Функцію густини $\rho(x)$ на інтервалах, що займає кожен елемент, отримуємо у вигляді $\rho(x) = \rho_i \equiv \frac{m_i}{l_i}$, де m_i – маса, l_i – довжина i елемента при $x \in (x_{n-i}, x_{n-i+1})$, $i = \overline{0;n}$ $x_0 = 0$, $x_{n+1} = L$. При записі функції розподіленого опору рухові $\varpi(x)$ нехтуємо її зміною в часі, оскільки одразу після прикладання додаткової сили $P_0(t)$ загальна швидкість поїзда v_0 не встигає істотно змінитися. Залежність $\varpi(x)$ від основного питомого опору рухові та інших видів опору зовнішнього середовища для різних швидкостей v_0 записуються у вигляді многочленів зі знайденими експериментально значеннями коефіцієнтів [3]. Розподілений по довжині стрижня силовий опір рухові визначається формулою $\varpi(x) = \frac{w_{0i}}{l_i} m_i g$, $x \in (x_{n-i}, x_{n-i+1})$ $(i = \overline{0;n})$, де w_{0i} – загальний питомий опір i елемента. При такому означенні функція $\varpi(x)$ – кусково-неперервна на інтервалі $(0,\ L)$ і не залежить від часу t при русі системи з постійною швидкістю. Числові значення для вхідних даних можна отримувати з довідників на перевезення.

Проведено тестові обчислення для стрижня, що складається з трьох однорідних частин $(0,\ x_1)$, $(x_1,\ x_2)$, $(x_2,\ L)$, приймаючи умови ідеального пружно-деформованого контакту між ними. Сила тяги F_0 і розподілені сили опору $\varpi(x)$ урівноважуються, і розтягнутий ними стрижень у початковий момент часу рухається з постійною швидкістю.

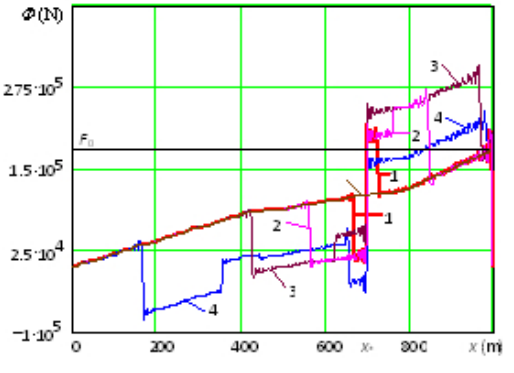


Рис. 1. Функція внутрішньої поздовжньої сили $\Phi(x,t)$ у пружному стрижні.

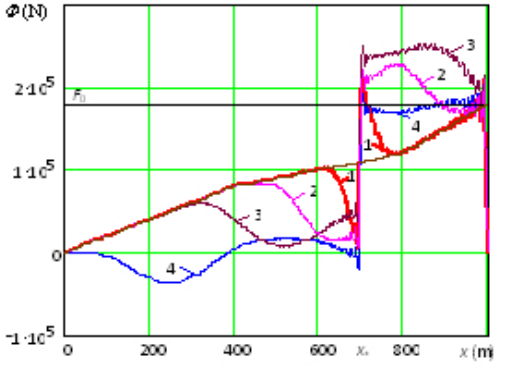


Рис. 2. Функція сили $\Phi(x,t)$ у в'язко-пружному стрижні μ=0,5 s.

Проведені розрахунки показали, що у перші секунди після раптового прикладання сили, напрямленої проти руху, динамічні ефекти призводять до місцевого зростання внутрішніх напружень. Урахування в'язкості не дає їх суттєвого зменшення. Воно приводить до згладжування фронту хвилі та її спотворення внаслідок відбивання від внутрішніх поверхонь при розгляді кусково-неоднорідного стрижня. Інтенсивна сила, локалізована всередині стрижня, призводить до зростання розтягу його частини, де прикладена сила тяги, та суттєвого зростання в ній внутрішніх сил. При цьому у набігаючій частині стрижня виникають стискаючі деформації і вд'ємні сили, які можна оцінити при наявності вхідних даних у кожному конкретному випадку.

Література. **1.** Блохин Е.П., Манашкин Л.А. Динамика поезда (нестационарные продольные колебания). – М.: Транспорт, 1982. – 222 с. **2.** Campos L.M.B.C., Santos A.J.P. On the propagation and damping of longitudinal oscillations in tapered visco-elastic bars // Journal of Sound and Vibration, 1988, Vol. 126, 1, p. 109–125. **3.** Гребенюк П.Т., Долганов А.Н., Скворцова А.И. Тяговые расчеты. – М.: Транспорт, 1987. – 272 с.

Бондаренко Н.С.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Огляд Skip-gram моделі нейронної мережі для технології Word2Vec

Word2Vec – це програмне рішення, а також клас моделей на базі векторного представлення слів, що використовує нейронні мережі для пошуку семантичних відношень у текстах. Результат роботи алгоритму може бути використаний в подальших дослідженнях у якості бази для інструментів машинного навчання та систем аналізу природних мов. Ця програмна розробка є відносно молодою (презентований командою Google у 2013 році) інструментом аналізу даних і має імплементації у вигляді бібліотек на C, Python та Java.

Word2Vec використовує Skip-gram архітектуру як один із двох способів генерації координатних представлень векторів-слів. Модель намагається спрогнозувати слова, що могли б оточувати надане слово – знаходимо слова W_{i-2}, W_{i-1}, W_{i+1}, W_{i+2} (контекст), використовуючи слово W_i [1]. Розглянемо структуру моделі.

Тренувальні дані. Тренувальними зразками для нейронної мережі слугують пари слів, отримані за допомогою попереднього аналізу текстів. Пари утворюються таким чином: кожне слово співставляється зі іншим таким, що потрапляє у його *вікно перегляду*. Розмір вікна перегляду N є одним з параметрів алгоритму [2].

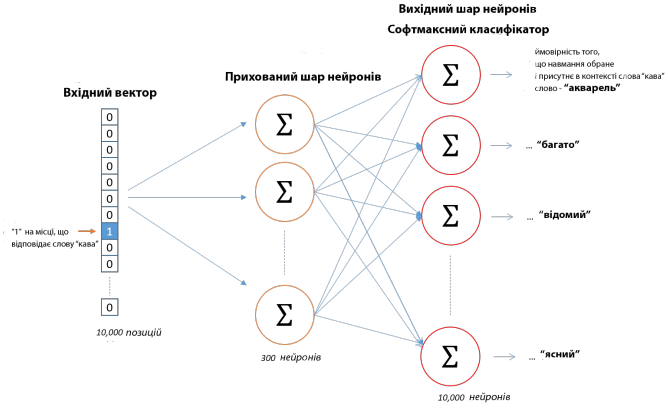


Рис. 1. Структура нейронної мережі

Вхідний вектор. Очевидно, що слово потребує кодування перед тим, як потрапити на вхід нейронної мережі. Для представлення такого слова був запропонований унітарний код – вектор розміру V, де кожному слову з тренувального документу відповідає розряд з одиницею на позиції, що відповідає вхідному слову.

Прихований шар. Розмірність матриці вагових коефіцієнтів внутрішнього шару залежить від кількості параметрів, яка визначає кількість нейронів внутрішнього шару та розміру алфавіту: результатом множення $V * N$ розмірної матриці вагових коефіцієнтів на вектор вхідних даних розмірності V є N-розмірний вектор внутрішнього шару, який насправді є строчкою матриці, що відповідає одиниці!

Вихідний шар. N-розмірний вектор внутрішнього шару далі множиться на $N * V$-розмірну матрицю, що містить інформацію про кожне слово алфавіту. Останнім кроком для отримання вектору результату є софтмаксна нормалізація. Результатом роботи нейронної мережі є вектор розміру V, де кожному слову відповідає ймовірність його потрапляння в контекст у розмірі від 0 до 1.

Основним недоліком даної моделі є велика кількість громіздких обчислень. Так як розмір алфавіту впливає на розмірність матриць вагових коефіцієнтів, одним із варіантів покращення є обробка певних словосполучень у якості одного "слова". Застосування цього підходу разом з декількома іншими дозволяють суттєво зменшити витрати на обчислення і роблять цю модель більш привабливою для використання.

Література. **1.** NOFREEHUNCH.GOOGLE WORD2VEC TUTORIAL https://nofreehunch.wordpress.com/2015/08/30/google-word2vec-tutorial-part-1/ **2.** Tensor Flow. Vector Representations of Words https://www.tensorflow.org/tutorials/word2vec.

Боровская Т.Н., Северилов П.В., Северилов В.А.
Винницкий национальний технический университет, Винница, Украина

Оптимальное агрегирование систем с произвольными ресурсными связями для задач переработки отходов

Вступление. Сегодня специализация производств достигла уровня отдельных государств. Существуют экономически эффективные государства, специализированные на бананах, банкинге, бензозаправках. Такая специализация региональных структур негативно влияет на объём и структуру рабочих мест, а затем и на социальную структуру в целом. В данной работе исследуются новые модели оптимизации экологической компоненты СТЭС – экономически эффективной переработке отходов агропроизводства. Первоисточник и назначение моделей – типичные задачи практики: – эффективная переработка бытовых отходов города, – региональная система сбора и переработки отходов агросектора в рыночные продукты, – автономная система производства биодизеля, электроэнергии и др. для территорий с дорогой логистикой. Проекты – базируются на инновационных технологиях и их носителях, интеграции исследований, проектирования, обоснования, строительства.

Новые математические модели работы: – метод разработки бинарных операторов оптимального агрегирования [1, 2] для новых видов ресурсных связей между подсистемами; – оптимальное агрегирование структур «K видов ресурсов, L технологий, M видов продукции».

Цель разработки – получение параметризованных моделей эффективности для оптимизации проектов при неполных данных по ценам и эффективностям технологий.

Методология оптимального агрегирования. Основа методологии оптимального агрегирования – интеграция в целостную информационную систему: «рабочая модель ⇔ «объект» ⇔ «компьютерная система», где: «рабочая модель» – программно реализованная математическая модель, «объект» – производство продуктов и услуг, включающее подсистемы «лаборатории», «стенды», «демонстраторы» и «компьютерная система» – АСУ, АСУТП, АСПР. Методы оптимального агрегирования базируются на алгебре, носители которой – функции «затраты, выпуск», операторы – аналогичны бинарным операторам «сумма», «произведение», «обратная связь». Результат агрегирования – оптимальная эквивалентная функция производства (ОЭФП) (рис. 1).

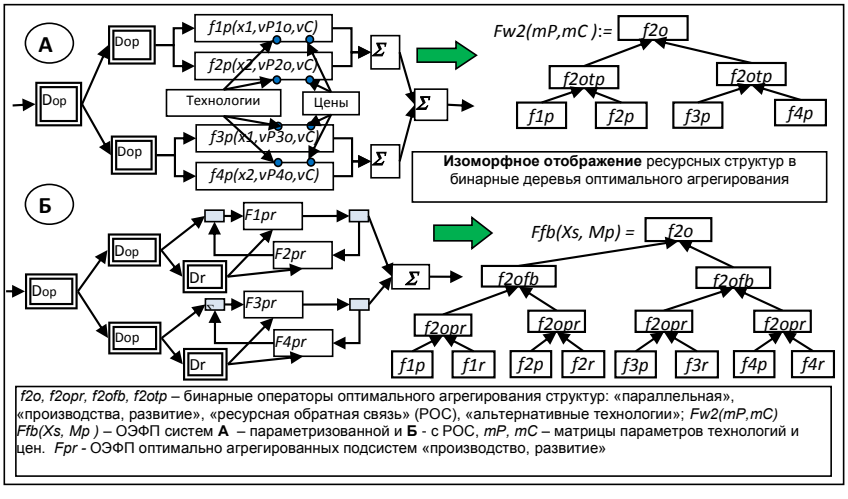


Рис. 1. Изоморфное отображение ресурсных структур в бинарное дерево оптимального агрегирования

Порядок разработки моделей анализа и оптимизации. Порядок разработки моделей рассмотрим на примере проектирования, разработки, обоснования и строительства автономной

системы обеспечения биодизелем и энергией. Методология оптимального агрегирования меняет акценты в научно-прикладных исследованиях, – снимаются проблемы: подгонки модели объекта под линейные, выпуклые, квадратичные функции и соответствующие методы, снимаются проблемы размерности и сходимости. Ассоциативность операторов оптимального агрегирования позволяет заменить многомерную задачу нелинейного программирования системой одномерных задач, решаемых векторизованным прямым перебором. Главные проблемы: – построение ресурсных моделей производственных систем методами прикладного системного анализа и программная реализация операторов и операндов. На рис. 2 приведены два шага из процедуры получения параметризованной ОЭФП для новой задачи.

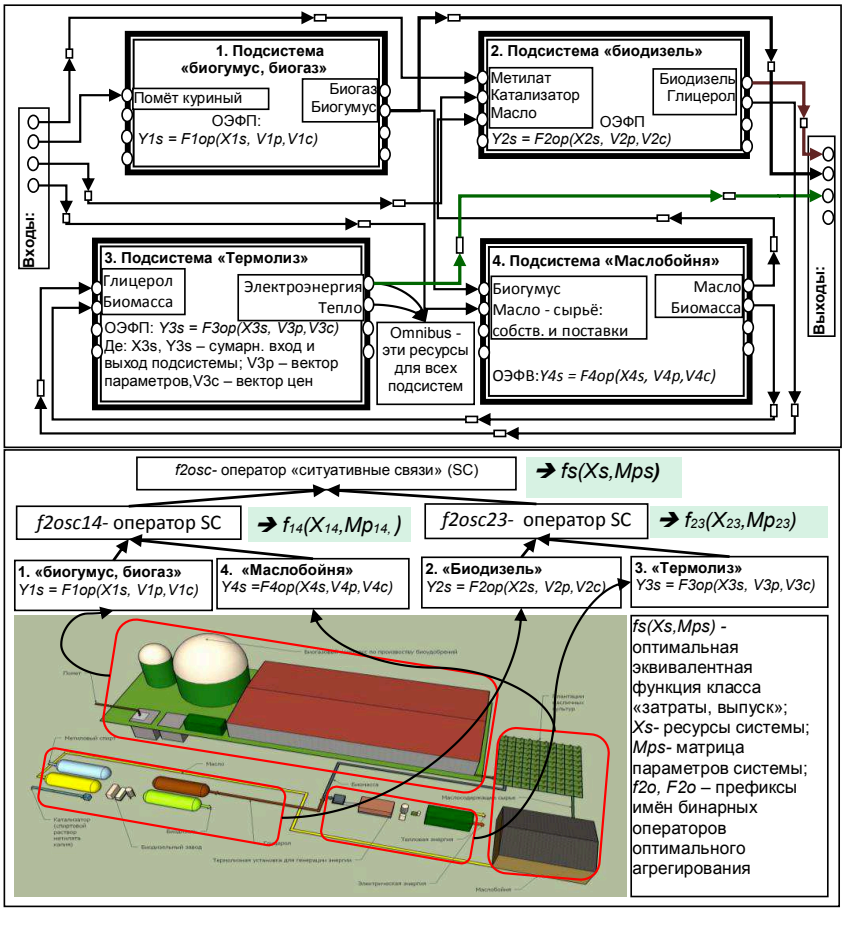


Рис. 2. Анализ ресурсных связей и оптимальное агрегирование подсистем

Выводы. В работе выполнена разработка параметризованных операторов оптимального агрегирования для автономной системы, проведено моделирование и анализ.

Литература. **1.** Боровська Т.М. Метод оптимального агрегування в оптимізаційних задачах: монографія / Т.М. Боровська, І.С. Колесник, В.А. Северілов. – Вінниця: УНІВЕРСУМ–Вінниця, 2009. – 229 с. – ISBN 978–966–641–285–3. **2.** Боровська Т.М. Моделювання і оптимізація процесів розвитку виробничих систем з урахуванням використання зовнішніх ресурсів та ефектів освоєння: монографія / [Т.М. Боровська, С.П. Бадьора, В.А. Северілов, П.В. Северілов]; за заг. ред. Т.М. Боровської. – Вінниця: ВНТУ, 2009. – 255 с. – ISBN 978–966–641–312–6.

Буценко Ю.П., Лабжинський В.А.
КПІ ім. Ігоря Сікорського, Київ, Україна

Модифікована процедура мажоритарного голосування для діагностування техногенно небезпечних систем

Техногенно небезпечні системи – це системи, експлуатація яких зазвичай пов'язана з ризиками для здоров'я і навіть життя людей та значними матеріальними збитками [1]. Це може бути наслідком використання помилкових даних про стан системи, які спричинені неадекватними показами датчиків, завадами у каналах зв'язку або програмними збоями. У техногенно небезпечних системах для підвищення вірогідності такої інформації використовують структурну (N-модульне резервування датчиків або каналів зв'язку) або алгоритмічну надлишковість (N-версійне програмування) [2]. Таким чином, техногенно небезпечна система після опитування N датчиків або отримання результатів виконання N програм діагностується шляхом голосування (вибору результату, "підтриманого" показами більшості датчиків або результатами виконання більшості програм). Результати голосування (участь або неучасть у прийнятті узгодженого рішення) можуть використовуватись для діагностики датчиків.

Розглянемо ситуацію, за якої кожен датчик системи (програма контролю) знаходиться під впливом чотирьох потоків випадкових подій:

- потік прихованих збоїв викликає одномоментні невірні результати, що не виявляються засобами діагностики, якщо при цьому датчик залишається працездатним і може використовуватись під час подальших вимірів (програма завершується нормально);
- потік діагностованих збоїв викликає одномоментні невірні результати, що діагностуються системою, при цьому покази не враховуються при поточному голосуванні, але можуть враховуватись при подальших;
- потік прихованих відмов викликає вихід датчика з ладу (аварійне завершення програми), що не виявляються діагностичними засобами системи, всі подальші результати невірні;
- потік діагностованих відмов викликає вихід датчика з ладу (аварійне завершення програми), який фіксується засобами діагностики. Як поточні, так і подальші результати не враховуються.

Якщо для оцінки якості системи після кожного опитування датчиків (виконання програм контролю) застосовується процедура голосування, то це дозволяє використовувати неучасть датчика (результатів виконання програми) в узгодженому голосуванні для формування його рейтингу у наступних голосуваннях. На відміну від використаного, наприклад, в роботі [3] правила формування рейтингу, пропонується враховувати у рейтингу умовну ймовірність того, що елемент вийшов з ладу або виконання програми завершилося аварійно (за умови, що результат k разів підряд не брав участі в результуючому голосуванні). Досліджуються варіанти асимптотичної поведінки таких імовірностей. Алгоритм керування системою включає в себе виключення з її конфігурації елементів, для яких значення цих імовірностей перевищують гранично допустиме. Запропонована методика дозволяє використовувати й інші критерії некоректних показань датчиків. Досліджується контрольованість функціонування системи за умови можливого виключення датчиків (контролюючих програм) з її складу. У припущенні незалежності функціонування датчиків (програм) допускається їх неоднорідність за показниками надійності.

Література. **1.** Панкратова Н.Д. Розпізнавання позаштатної ситуації в динаміці функціонування техногенно небезпечного об'єкта / Н.Д. Панкратова, А.М. Радюк // Наукові вісті НТУУ "КПІ". – 2008. – № 3. – С. 43–52. **2.** Иыуду К.А. Надежность, контроль и диагностика вычислительных машин и систем / К.А. Иыуду. – М.: Высшая школа, 1989. – 216 с. **3.** Latif-Shabgahi G. History-based weighted average voter: A novel software voting algorithm for fault-tolerant computer systems / G. Latif-Shabgahi, J.M. Bass, S. Bennett // Proc. of the 9th IEEE Euromicro Workshop on Parallel and Distributed Processing. – Mantova, Italy, 2001. – P. 402–409.

Васильев В.И., Висшталь Д.М., Любашенко Н.Д.
КПИ им. Игоря Сикорского, Киев, Украина

Структурные функции в задачах оценки доступности сетевой услуги в классе бинарных стохастических моделей

В качестве статической модели сети рассматривается размеченный конечный связный неорграф $G(Y;X)$, где $|Y|$ – число узлов графа, $|X|$ – число ребер графа. Узлы графа играют вспомогательную роль.

В качестве динамической модели сети рассматривается стохастический граф, элементы которого взвешены весовыми параметрами. Сеть предназначена для обеспечения сетевой услуги в любой момент времени для всех пар узлов. Произвольная пара узлов может получить сетевую услугу, если в данный момент времени найдется по крайней мере один простой путь, их соединяющий.

Сетевая услуга доступна для всех узлов, если в данный момент времени существует по крайней мере одно остовное дерево.

Элементы сети и структурные функции, соответствующие условиям доступности сетевой услуги, принимают бинарные значения. Пусть $\varphi(x_1(t),...,x_n(t))$ – структурная функция, соответствующая некоторому виду сетевой услуги.

Пространство состояний графа W_G можно рассматривать и как векторное пространство по $mod(2)$ [1] и как булеву векторную решетку B_2^n.

Использование пространства W_G удобно для алгоритма нахождения разрезающих множеств графа.

Изоморфизм булевой алгебры классов и соответствующей булевой алгебры свойств дает возможность алгебраизации необходимых вычислений для оценки характеристик доступности сетевой услуги. Функционирование сети рассматривается как процесс случайного блуждания по булевой векторной решетке.

Принятые допущения модели:

1. Зависимостью на уровне отдельных элементов можно пренебречь.
2. Эволюция элементов моделируется во времени соответствующими альтернирующими процессами восстановления с конечными средними значениями интервалов доступности (недоступности) для выполнения сетевой услуги.
3. Структурная функция, соответствующая сетевой услуге, монотонная функция алгебры логики.
4. Эволюция структурных функций во времени моделируется соответствующими альтернирующими процессами восстановления с конечными средними значениями интервалов доступности (недоступности) сетевой услуги.
5. Существует $\lim\limits_{t\to\infty} P\{x_i(t)=1\} = \dfrac{M\theta_i}{M\theta_i + M\xi_i} = P_i$, где $M\theta_i$, $M\xi_i$ – средние стационарные интервалы доступности (недоступности) i-го элемента для выполнения сетевой услуги.
6. Существует $\lim\limits_{t\to\infty} P\{\varphi(x_1(t),...,x_n(t))=1\} = \dfrac{M\theta_\varphi}{M\theta_\varphi + M\xi_\varphi}$, где φ – детерминированная монотонная функция случайных аргументов, $M\theta_\varphi$, $M\xi_\varphi$- средние стационарные интервалы доступности (недоступности) сетевой услуги, $M\theta_\varphi + M\xi_\varphi$ средняя стационарная длина цикла альтернирующего процесса восстановления.

В этих условиях требуется оценить точно или приближенно следующие сетевые характеристики:

- $h_\varphi(P_1,...,P_n)$ – стационарная вероятность доступности сетевой услуги;
- $M\theta_\varphi$, $M\xi_\varphi$ – средние стационарные интервалы доступности (недоступности) сетевой услуги.

Основная вычислительная проблема оценки сетевых характеристик обусловлена большой размерностью пространства состояний. Это обстоятельство влечет трудновычислимость сете-

вых характеристик. Можно поднять порог вычислимости, используя свойство монотонности структурной функции, свойство дуальных автоморфизмов булевой алгебры и алгоритмы представления структурной функции в ортогональной бесповторной форме. Свойство монотонности структурной функции позволяет осуществить структурно-фазовое укрупнение состояний сети.

Детерминированные свойства структурных функций определяются следующими тождествами:

$$\bigvee_{\{\alpha_i\}\in I}\left(\bigwedge_{k_i\in\{\alpha_i\}} x_{k_i}\right) = \varphi(x_1,...,x_n) = \bigwedge_{\{\beta_j\}\in J}\left(\bigvee_{r_j\in\{\beta_j\}} x_{r_j}\right), \quad (1)$$

$$\bigvee_{\{\beta_j\}\in J}\left(\bigwedge_{r_j\in\{\beta_j\}} x_{r_j}\right) = \varphi^*(x_1,...,x_n) = \bigwedge_{\{\alpha_i\}\in I}\left(\bigvee_{k_i\in\{\alpha_i\}} x_{k_i}\right), \quad (2)$$

$$\bigvee_{\{\beta_j\}\in J}\left(\bigwedge_{r_j\in\{\beta_j\}} \bar{x_{r_j}}\right) = \bar{\varphi}(x_1,...,x_n) = \bigwedge_{\{\alpha_i\}\in I}\left(\bigvee_{k_i\in\{\alpha_i\}} \bar{x_{k_i}}\right), \quad (3)$$

$$\bigvee_{\{\alpha_i\}\in I}\left(\bigwedge_{k_i\in\{\alpha_i\}} \bar{x_{k_i}}\right) = \bar{\varphi}^*(x_1,...,x_n) = \bigwedge_{\{\beta_j\}\in J}\left(\bigvee_{r_j\in\{\beta_j\}} \bar{x_{r_j}}\right), \quad (4)$$

где:

- $\bigwedge_{k_i\in\{\alpha_i\}} x_{k_i} = \alpha_i(x)$ – i-я простая импликанта функции $\varphi(x)$,
- $\{\alpha_i\}$ – набор индексов существенных переменных i-й импликанты,
- $\bigwedge_{r_j\in\{\beta_j\}} x_{r_j} = \beta_j(x)$ – j-я простая импликанта $\varphi^*(x)$,
- $\{\beta_j\}$ – набор индексов существенных переменных j-й импликанты,
- $\bigvee_{k_i\in\{\alpha_i\}} x_{k_i} = \alpha_i^*(x)$ – двойственная по отношению к $\alpha_i(x)$ функция,
- $\bigvee_{r_j\in\{\beta_j\}} x_{r_j} = \beta_j^*(x)$ – двойственная по отношению к $\beta_j(x)$ функция.

Интерпретация остальных компонент тождеств определяется аналогично.

Символ I используется и для обозначения полного списка простых импликант функции $\varphi(x)$ и для обозначения индексов семейств существенных переменных простых импликант. Символ J используется аналогичным образом.

Сделаем несколько замечаний:

1. Тождества (1), (2), (3), (4) дают возможность находить минимальное по числу литералов представление соответствующих структурных функций. Однако при этом утрачивается независимость функционирования, рассматриваемая на уровне простых импликант функций.
2. Тождества (1), (2), (3), (4) верны не только для бинарных значений переменных, но и для непрерывных неотрицательных значений переменных. С комбинаторной точки зрения теоретико-множественные семейства индексов простых импликант $I = \{\{\alpha_i\}\}_{i=1}^{|I|}$ и $J = \{\{\beta_j\}\}_{j=1}^{|J|}$ образуют два двойственных семейства минимальных трансверсалей.
3. Знание $M\theta_i$, $M\xi_i$, а также семейств множеств I и (или) J дает возможность выявить "узкие места" в обеспечении сетевых характеристик, определять степень влияния на эти характеристики как отдельных элементов так и определенных конфигураций элементов, что важно для развивающихся сетей.

Для оценки характеристик сети используются эффективные алгоритмы поиска семейств I и J, а также алгоритмы представления структурной функции в ортогональной бесповторной форме. В работе рассматриваются иллюстративные примеры оценки сетевых характеристик в зависимости от весовых параметров элементов и разных структурных функций.

Литература. 1. М. Свами, К. Тхуласираман. Графы, сети и алгоритмы. – М.: Мир – 1984.

Войтех Д.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Визначення оптимальної кількості кластерів при кластеризації

Постановка задачі. При вирішені задачі ітеративної кластеризації даних зазвичай постає питання вибору оптимальної кількості кластерів. При відносно малій розмірності екземплярів даних, відповідь на питання можна дати при їх візуальній оцінці. Проте при більш складних представленнях екземплярів цей підхід часто дає неоднозначні результати та може внести похибку в подальший аналіз даних. Оскільки кількості кластерів є одним з вхідних параметрів методів ітеративної кластеризації, важливим є визначення її оптимального значення.

Існуючі методи оцінки кількості кластерів.

1. *Правило великого пальця (“By rule of thumb”)* [1]. Найпростіший метод. Кількість кластерів K оцінюється за формулою:
$$\sqrt{\frac{n}{2}}, \tag{1}$$
де n – об’єм вибірки. Дана оцінка є грубою, проте її можна використовувати для будь-яких типів вибірки.

2. *“Ліктьовий” метод (“Elbow” method)* [2]. Найбільш універсальний метод. У методі для різної кількості кластерів K для даної вибірки розраховується оцінка сумарного квадратичного відхилення (Sum Squared Error):
$$SSE = \sum_{i=1}^{K} \sum_{x \in c_i} dist(x, c_i)^2, \tag{2}$$
де x – екземпляр вибірки, c – центроїд i-го кластера. Отримана залежність (SSE від K) зображується у вигляді графіка.
Очевидно, при збільшенні кількості кластерів значення оцінки SSE зменшується. Оптимальним вважають таку кількість кластерів, при якій додавання ще одного кластера не призводить до значного зменшення оцінки SSE. Зазвичай дана точка на графіку є точкою перегину кривої. Саме тому метод отримав назву *“ліктьовий”*.
Однак, іноді при виборі такої *“ліктьової”* точки також можуть виникати неоднозначності (наприклад, перегинів на графіку може бути декілька).

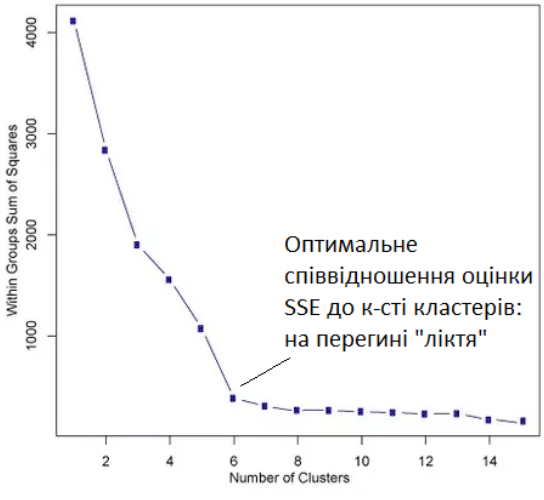


Рис. 1. Залежність оцінки SSE від кількості кластерів

3. *"Силуетний" метод (Choosing k Using the Silhouette)* [3,4]. У даному методі для кожного екземпляра вибірки розраховується т.з. *"силуетна"* відстань:

$$s(x_i) = \frac{b(x_i) - a(x_i)}{max(a(x_i), b(x_i))} \tag{3}$$

де $a(x)$ – середня відстань від екземпляра x до усіх інших екземплярів того ж кластера; $b(x)$ – мінімальна з середніх відстаней від x до усіх екземплярів кожного з інших кластерів.

Очевидно, що $s(x)$ приймає значення від -1 до 1. Якщо $s(x)$ наближається до -1, то x віднесено до невірного кластеру. Якщо $s(x)$ знаходиться в околицях 0, тоді x можна рівноцінно віднести до декількох кластерів. Якщо $s(x)$ наближається до 1, то x віднесено до вірного кластеру.

Після обчислення даної оцінки для всіх екземплярів знаходять їх середнє – *середню силуетну відстань*. Далі розрахунок виконують для різної кількості K кластерів. Обчислення показують, що зі зростанням кількості кластерів оцінка *середньої силуетної відстані* зменшується. Кількість кластерів, що відповідає найбільшій середній силуетній відстані вважається оптимальною.

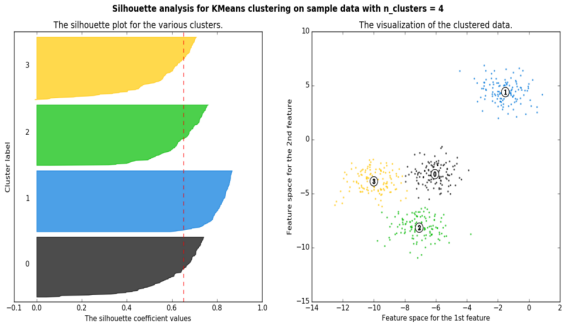


Рис. 2. Приклад графічного представлення *"силуетної"* відстані для 4-ох кластерів [5]

4. *Метод перехресної валідації (Cross-validation)* [6]. Відмінність даного методу полягає у тому, що у ньому для кожної кількості кластерів(з обраного діапазону) модель кластеризації будується багато разів, грунтуючись на різних фрагментах вибірки даних.

У процесі роботи методу вибірка даних випадковим чином розділяється на v частин. Кожну частину по черзі відкладають як тестову, а за допомогою інших v–1 (навчальних) частин будується модель кластеризації, далі для тестової частини, внесеної до побудованої моделі, розраховується значення цільової функції (наприклад, сумарне середньоквадратичне відхилення від центроїдів).

Отримані v значень усереднюються для різної кількості кластерів. К-сть кластерів, що відповідає мінімальному усередненому значенню ЦФ і обирається як оптимальна.

Результати. Метод кластеризації K-Means є одним з найбільш популярних методів кластеризації. Можливість оцінити оптимальну кількість кластерів для даного методу допомагає знайти збалансоване рішення задачі кластеризації. Розглянуті методи відрізняються за точністю оцінки та складністю обчислень, тому їх використання залежить від конкретної задачі.

Література. **1.** Hartigan JA Clustering Algorithms, New York: J. Wiley & Sons, 1975. **2.** Andrew Ng, Clustering with the K-Means Algorithm, Machine Learning, 2012. **3.** Pollard K.S., and Van Der Laan, M.J., A Method to Identify Significant Clusters in Gene Expression Data, U.C. Berkeley Division of Biostatistics, Working Paper Series, p. 107, 2002. **4.** Kaufman L., and Rousseuw P., Finding Groups in Data: An Introduction to Cluster Analysis, New York: J. Wiley & Son, 1990. **5.** http://scikit-learn.org/stable/auto_examples/cluster/plot_kmeans_silhouette_analysis.html. **6.** Wang, J., Consistent selection of the number of clusters via cross-validation. Biometrika, 2010.

Волкова В.Н., Ефремов А.А., Логинова А.В.
Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

Модели и технологии для проектирования автоматизированных информационных систем

На основе обобщения предшествующего опыта [1] сформировалось три основных подхода к проектированию информационных систем: а) системно-целевой подход, основанный на *структуризации* или *декомпозиции* целей и функций системы; б) подход, который называют *морфологическим* (в широком смысле), *лингвистическим, тезаурусным, терминальным*, методом *"языка"* системы, с помощью которого определяется «пространство состояний» системы и реализуется поиск взаимосвязей (мер близости) между элементами; в) *процессный* подход.

С самого начала разработки АСУ и АИС как первой очереди АСУ разных уровней осуществлялись на основе системно-целевого подхода, т.е. на основе структуризации целей и функций системы управления. Для управления разработками АСУ были подготовлены и изданы соответствующие общеотраслевые руководящие методические материалы, (ОРММ) в которых отражались теория и основная терминология, обязательная для использования при разработке и внедрению АИС и АСУ. На основе системно-целевого подхода формировалась и Государственная система научно-технической информации (ГСНТИ).

Для реализации этого подхода [1] предлагалось определять функциональную часть (ФЧ), обосновываемую путем анализа целей и функций системы управления с применением методик системного анализа, и обеспечивающую часть (ОЧ), что не утрачивает актуальности и в настоящее время при разработке любых автоматизированных информационных систем, особенно для крупных предприятий и организаций.

Процессный подход (который можно считать развитием функционально-технологического подхода) основан на структуризации информационных процессов во времени, на представлении процессов в форме графов, реализуемый на основе методология SADT с помощью функционально-ориентированных и объектно-ориентированных CASE- и RAD-технологий и их компьютерной реализации с использованием стандартов IDEF и DFD [2], реализуемых автоматизированными средствами BPWin, ARIS, язык UML и т.п. Подход применяется для разработки экономических информационных систем. Создана теория проектирования [3, 4], основанная на применении процессного подхода, анализе бизнес-процессов, представляемых специальными графами.

Таким образом, при разработке методик и моделей для проектирования и организации функционирования ИС применяются различные технологии, классификация которых предложена в [5, 6] – от технологий взаимодействия разработчиков и пользователей ИС с ЭВМ до технологий создания баз данных и обработки информации (рис. 1). Кроме того, наряду с этими технологиями необходимы новые технологии и автоматизированные средства их реализации (на рисунке обведены двойной линией), помогающие на этапе концептуального проектирования – при разработке структуры ФЧ ИС, определении очередности разработки информационных подсистем или выборе готовых программных продуктов и принятии решений по развитию ИС.

В числе таких технологий – автоматизированные диалоговые процедуры анализа целей и функций (АДПАЦФ) систем управления, автоматизированные диалоговые процедуры для реализации методов организации сложных экспертиз (АДПОСЭ), таких как метод решающих матриц Г.С. Поспелова, методы, основанные на информационном подходе А.А. Денисова, и т.п. Такие средства в настоящее время в большинстве своем еще не доведены на уровня совершенных программных продуктов. Но существуют экспериментальные разработки [6–8].

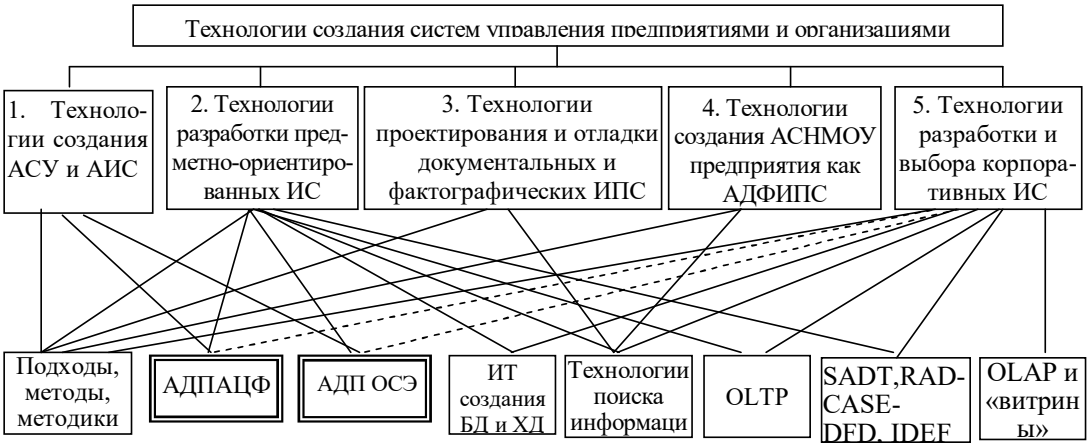


Рис. 1. Технологии создания информационных систем

Литература. **1.** Волкова В.Н. Теория информационных процессов и систем: учебник и практикум. – М.: Юрайт, 2014. – 502 с. **2.** Черемных С.В., Семенов И.О., Ручкин С.В. Структурный анализ систем: IDEF-технологии. – М.: Финансы и статистика, 2003. **3.** Тельнов, Ю.Ф. Инжиниринг бизнес-процессов / Ю.Ф. Тельнов. – М.: Финансы и статистика, 2004. **4.** Смирнова Г.Н. Проектирование экономических информационных систем: учебник / Г.Н. Смирнова, А.А. Сорокин, Ю.Ф. Тельнов / под ред. Ю.Ф. Тельнова. – М.: Финансы и статистика, 2005. **5.** Классификация информационных технологий / В.Н. Волкова, А.Ю. Васильев, А.А.Ефремов, Н.Б. Паклин, В.Н. Юрьев // Прикладная информатика. Том 10. №5(591) – 2015. М.: Market, 2054. – С. 127–147. **6.** Моделирование систем и процессов. Практикум: Под ред. В.Н. Волковой. – М.: Издательство Юрайт, 2016. – 295 с. **7.** Автоматизированные диалоговые процедуры анализа целей и функций систем управления / Под ред. В.Н. Волковой. – СПб.: Изд-во Политехн. ун-та, 2010. – 72 с. **8.** Волкова В.Н., Денисов А.А. Методы организации сложных экспертиз / СПб.: Изд-во Политехн. ун-та, 2010 г. – 128 с.

***Гайвась Б.І.*[1], *Дмитрук В.А.*[1,2] *Торський А.Р.*[1] *Дмитрук А.А.*[3]**
[1] *Центр математичного моделювання Інституту прикладних проблем механіки і математики ім. Я.С.Підстригача НАН України, Львів, Україна;* [2] *Національний університет "Львівська політехніка", Львів, Україна;* [3] *Львівський національний університет імені Івана Франка, Львів, Україна*

Пакет програм для дослідження сушіння зерна пшениці в газозваженому стані за імпульсної дії сушильного агента

Пшениця – одна з найважливіших продовольчих культур. Україна є одним з ключових виробників зерна пшениці на світовому ринку, щорічно нарощуючи обсяги його експорту. Зберігання зерна є важливою ланкою в зерновому бізнесі. Падіння закупівельних цін під час збору врожаю, в основному, пояснюється збільшенням пропозиції зерна з боку виробників, які не мають можливості зберігати зерно. Проте, процес зберігання потребує постійного вивчення, вдосконалення та пошуку нових можливостей. З метою підвищення стійкості зерна при зберіганні і консервуванні його сушать.

Сушка належить до найбільш енергоємних процесів. З метою економії вартісних енергоресурсів сучасні апарати сушіння повинні задовольняти ряд високотехнологічних вимог за мінімуму видатків. Для таких апаратів необхідно створювати відповідне програмне забезпечення комплексного аналізу вологих матеріалів та алгоритми дій відповідно до встановлених характеристик об'єктів сушіння. Використання певних методів сушіння зумовлено видом енергії зв'язку вологи з матеріалом. Крім того, вимагає застосування різних методів математичного моделювання кожний конкретний вид апарату сушіння, що пов'язано з різною взаємодією потоку частинок і агента сушіння (імпульсного, псевдозрідженого, сушка в гідродинамічних активних режимах і т.д.).

Для рівномірного нагріву всієї маси зерна необхідно забезпечити розділення зерен. Експериментально встановлено, що при правильно вибраному співвідношенні між кількістю зерна, проміжного теплоносія і величиною розходу повітря зерно не тоне, а бере участь в загальному пульсаційному русі суміші. Отриманий газозважений шар дозволяє забезпечити розділення зерен одне від другого і збільшити поверхні контактів з основним і проміжним теплоносіями В залежності від швидкості газу та характеристик матеріалу зважений шар може знаходитись в стані спокійного псевдорозрідження [1].

Львівським Центром математичного моделювання ІППММ ім. Я.С. Підстригача проводиться системне дослідження особливостей сушіння зерна пшениці в газозваженому стані за імпульсної дії сушильного агента, з якого, зокрема, виконано постановку нової моделі теплопровідності та механодифузії в зерні та міжзерновому шарі [2], а також запропоновано метод розв'язування зв'язаної задачі про зміну концентрації вологи та температури в зернині в часі з врахуванням неоднорідності структури зернини та швидкості продуву міжзернового середовища [3] та визначено оптимальні значення функцій керування (пароповітряної суміші та температури) сушильного агента за врахування основних характеристик зерна [4]. В роботі [5] розглянуто напружено-деформований стан зерна в газозваженому стані та встановлено закономірності виникнення напружень з урахуванням двошарової структури окремої зернівки. На основі отриманих розв'язків зв'язаної задачі про зміну концентрації вологи та температури в зернині створено «Пакет програм для дослідження сушіння зерна пшениці в газозваженому стані за імпульсної дії сушильного агента (скорочено – пакет «SuchGas»)» призначений для проведення кількісних та якісних досліджень таких процесів, а саме – аналізу змін концентрації вологи та температури в зернині в часі, в залежності від глибини розміщення її в шарі, неоднорідності її структури, швидкості продуву міжзернового середовища, керуючих функцій (пароповітряної суміші та температури) сушильного середовища та фізико-хімічних характеристик зерна.

Пакет «SuchGas» забезпечує розрахунок (рис. 1):

- концентрацій вологи та температури в зернині залежно від значень:
  - часу,

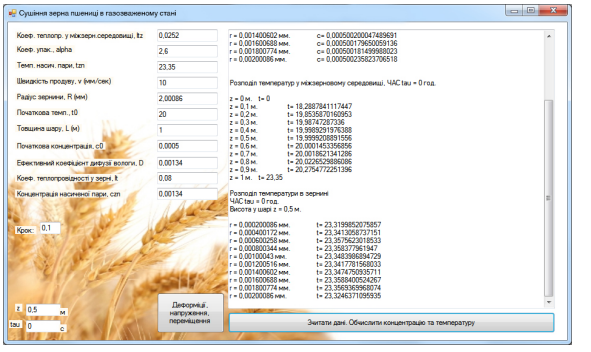


Рис. 1. Приклад роботи пакету програм

– збурюючих осцилюючих функцій концентрації та температури агента сушіння,
– швидкості продуву міжзернового середовища шару зерна,
– товщини шару;
– місця вимірів за товщиною зернини, а саме, окремо в ядрі і в оболонці;

За відомими значеннями концентрації та температури в зерні рахується

- сумарний потік вологи з зерна в процесі випаровування, який залежить від термоградієнтного коефіцієнта.

Далі за зміною концентрації та потоком вологи в зерні визначаються

- переміщення,
- деформації,
- напружений стан зернини:
  – в довільній точці за радіусом,
  – в залежності від її місця розміщення,
  – часу.

Гранично допустимі температури при сушці зерна залежать від роду зерна, його цільового призначення, вихідної вологості і конструкції сушарень. Крім температури має значення організація потоків дисперсної і газової фаз, місце, момент часу, тривалість контакту. Комп'ютерне моделювання сушіння зерна пшениці в газозваженому стані за імпульсної дії сушильного агента показало, що найсприятливіші умови взаємодії фаз досягаються в прямоточних сушарнях ідеального витіснення. Це обумовлено однаковою швидкістю частинок, тобто лінійністю профілю поршневого потоку. Основне співвідношення для моделі ідеального витіснення витікає з умови неперервного потоку. Але ця умова накладає обмеження на зміну швидкості на вході і виході зони ідеального витіснення.

Література. **1.** Муштаев В.И. Основные теоретические положения конвективной сушки и уточненный метод расчета сушильных аппаратов. – М.: МИХМ, 1971. – 81 с. **2.** Гайвась Б., Дмитрук В., Торський А., Дмитрук А. Модель імпульсного осушення шару зерна з врахуванням неоднорідної структури зернини // Матеріали 18-ї міжнар. наук-тех. конф. SAIT. – К.: НТК «ІПСА» НУТУ «КПІ», 2016. – С. 63–65. **3.** Гайвась Б., Дмитрук В. Сушіння зерна в імпульсному режимі агента сушіння з врахуванням шаруватості структури зернини // Фізико-математичне моделювання та інформаційні технології. 2015. – Вип. 22. С. 7–23. **4.** Гайвась Б., Дмитрук В., Дмитрук А. Сушіння зерна в активних гідродинамічних режимах з урахуванням шаруватості його структури // Фізико-математичне моделювання та інформаційні технології. 2015. – Вип. 23. С. 29–41. **5.** Гайвась Б., Дмитрук В., Дмитрук А. Визначення напружено-деформованого стану зернини в залежності від її місцерозміщення в газозваженому стані в імпульсному режимі осушення // Фізико-математичне моделювання та інформаційні технології. 2016. – Вип. 24. С. 14–26.

Горбань Н.В., Палійчук Л.С., Хоменко О.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Асимптотична поведінка розв'язків неавтономних дисипативних еволюційних задач в нескінченновимірних просторах

Вивчення глобальної поведінки розв'язків неавтономних дисипативних задач пов'язують, зокрема, з застосуванням теорії рівномірних глобальних та траєкторних атракторів в нескінченновимірних просторах. В [1–5] розроблені методи та підходи для дослідження задач виду

$$\partial_t u(t) = A_{\sigma(t)}(u(t)), \tag{1}$$

де $\sigma(s)$, $s \geqslant 0$, – функціональний параметр, так званий часовий символ. В конкретних задачах математичної фізики функція $\sigma(s)$ складається з усіх залежних від часу членів досліджуваної задачі (зовнішні сили, параметри середовища, функції взаємодії тощо). В вищезгаданих роботах припускалось, що параметр σ рівняння (1) належить хаусдорфовому топологічному простору Ξ_+ функцій, визначених на R_+, що приймають значення в деякому повному метричному просторі. Зазвичай, в застосуваннях в ролі топології простору Ξ_+ розглядають топологію локальної збіжності $\forall [t_1, t_2] \subset R_+$. Далі, розроблений в даних роботах підхід передбачає розгляд сім'ї рівнянь типу (1) з різними символами $\sigma(s)$, що належать $\Sigma \subseteq \Xi_+$. Σ називається простором символів сім'ї рівнянь (1). В застосуваннях, як правило, в ролі простору символів Σ природно розглядати ядро трансляційно компактної функції $\sigma_0(s)$ у відповідному хаусдорфовому топологічному просторі Ξ_+. Однак, строга реалізація даного підходу є досить проблематичною та вимагає досить жорстких додаткових умов на параметри вихідної задачі, зокрема умови трансляційної компактності функції взаємодії в деякому компактному хаусдорфовому топологічному просторі. Авторами запропоновано альтернативну схему дослідження поведінки неавтономних систем, в результаті чого доведено існування рівномірного глобального та траєкторного атракторів всіх слабких розв'язків задачі (1) та встановлено його топологічні властивості за природних умов, без припущень щодо трансляційної компактності та гладкості на коефіцієнти задачі. Встановлено, що умови:

(A1) існують такі трансляційно рівномірно інтегровна в просторі $L_1^{\mathrm{loc}}(R_+)$ функція $c_1 : R_+ \to R_+$ та стала $\alpha_1 > 0$, що для кожного $\tau \geqslant 0$, $y \in K_\tau^+$ і $t_2 \geqslant t_1 \geqslant \tau$ виконується така нерівність:

$$\|y(t_2)\|_H^2 - \|y(t_1)\|_H^2 + \alpha_1 \int_{t_1}^{t_2} \|y(t)\|_V^p dt \leqslant \int_{t_1}^{t_2} c_1(t) dt;$$

(A2) існують такі трансляційно рівномірно інтегровна в просторі $L_1^{\mathrm{loc}}(R_+)$ функція $c_2 : R_+ \to R_+$ та стала $\alpha_2 > 0$, що для кожного $\tau \geqslant 0$, $y \in K_\tau^+$ і $t_2 \geqslant t_1 \geqslant \tau$ виконується така нерівність:

$$\int_{t_1}^{t_2} \|y'(t)\|_{V^*}^q dt \leqslant \alpha_2 \int_{t_1}^{t_2} \|y(t)\|_V^p dt + \int_{t_1}^{t_2} c_2(t) dt.$$

є достатніми для існування рівномірних глобального та траєкторного атракторів для неавтономної дисипативної задачі. Тезу підготовлено за результатами роботи [6].

Література. **1.** Бабин, А.В., Вишик, М.И.: Аттракторы эволюционных уравнений. Наука, Москва (1989). **2.** Мельник, В.С. Семейства м-полупотоков и их аттракторы. Доклады РАН. Т. 353, №2. С. 158–162 (1997). **3.** Zgurovsky, M.Z. [et al.]: Evolution inclusions and variation Inequalities for Earth data processing III. Springer, Berlin (2012). **4.** Sell, G.R.: Global attractors for the three-dimensional Navier-Stokes equations. J. Dyn. Diff. Eq. Vol. 8, Iss. 12. P. 1–33 (1996). **5.** Chepyzhov, V., Vishik, M.: Trajectory attractor for reaction-diffusion system with diffusion coefficient vanishing in time. Discrete and Continuous Dynamical Systems. Vol. 27, Iss. 4. P. 1498–1509 (2010). **6.** Zgurovsky, M.Z. [et al.]: Uniform global attractors for non-autonomous dissipative dynamical systems. DCDS. Series B. (2017) to appear.

Горелова Г.В.
Институт управления в экономических, экологических и социальных системах,
Инженерно-технологическая Академия Южного федерального университета, Таганрог, Россия

Анализ вариантов когнитивной модели рынка труда

Введение. Проблемы рынков труда (РТ), в том числе, избытка или недостатка рабочей силы на рынке существуют постоянно. Можно ли, каким образом и кому регулировать этот рынок или адаптироваться к его условиям – тоже является проблемой. Для решения этих проблем и принятия соответствующих управленческих решений разными участниками этого рынка желательно понять его структуру и предвидеть возможные сценарии его развития. В работе предлагается использовать для этих целей инструмент когнитивного моделирования сложных систем (РТ – сложная система), разрабатываемый в Инженерно-технологической академии Южного федерального университета (ИТА ЮФУ).

Разработка и анализ когнитивных моделей рынка труда. Предлагаемое когнитивное моделирование сложных систем [1, 2] происходит в несколько этапов, основными являются: 1) разработка когнитивной модели (когнитивной карты или более сложной модели, например, в виде параметрического функционального графа), 2) всесторонний анализ модели (сложности, связности, устойчивости, сценарного анализа и др.), 3) разработка и анализ стратегий развития, принятия управленческих решений.

Проиллюстрируем эти этапы на ряде примеров исследования абстрактного рынка труда.

Первый этап. Разработка когнитивной модели. Существует много методов разработки когнитивных моделей (работы с экспертами, анализ текстов и статистических данных о системе, анализ теории и др.) Для разработки "стартовой" (начальной) когнитивной карты воспользуемся теоретическим описанием, которое существует в экономической литературе. Рассмотрим классическую ситуацию спроса-предложения на рынке труда в условиях совершенной конкуренции, отраженную рис. 1.

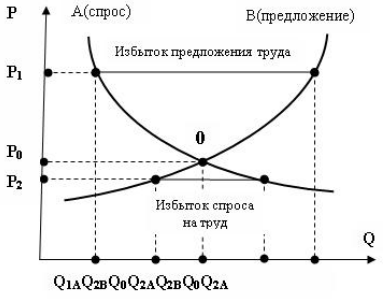


Рис. 1. Кривые спроса-предложения на рынке труда в условиях совершенной конкуренции

Рис. 1 является классическим примером из учебников. На нем по оси ординат откладывается уровень ставки заработной платы (P), а по оси абсцисс – количество предложения труда (Q). Как на товарном, так и на рынке труда присутствуют спрос (A) и предложение труда (B). Точка пересечения кривых спроса и предложения труда (точка 0 – спрос на труд равен предложению труда) соответствует уровню ставки заработной платы P0 и количеству предложения труда Q_0. То есть все работодатели, которые готовы платить заработную плату P_0 работников и их спрос на труд, таким образом, удовлетворен полностью. Также в данной точке все наемные работники трудоустроены и согласны работать при ставке заработной платы P_0. Наступает равновесие на рынке труда (equilibrium in labor market), достигается полная занятость. Но такое положение практически не встречается, всегда существует отклонение от точки равновесия. Если $Q_{1B} > Q_{1A}$, то возникает безработица, количество желающих найти работу и продать свои услуги труда при заработной плате P_1 будет чрезмерно высокое. Если же $Q_{2A} > Q_{2B}$,то возникает избыток спроса на труд.

Переведем эту вербальную и графическую информацию в форму когнитивныхкарт, рис.

2. Изображение получено с помощью программной системы когнитивного моделирования (ПСКМ) [1], в которой одновременно с построением когнитивной карты G (в диалоговом режиме с исследователем) происходит моделирование матрицы отношений AG этого знакового ориентированного графа.

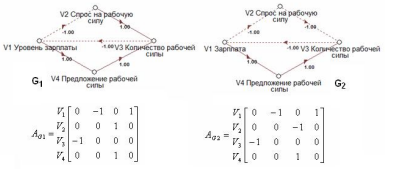


Рис. 2. Когнитивные карты: G_1–равновесный рынок, G_2 – неравновесный рынок

Второй этап. Анализ когнитивной модели. ПСКМ позволяет, производя соответствующие операции над матрицами A_G, определять различные свойства когнитивной модели. На рисунках 3 и 4 представлены результаты импульсного моделирования, анализа устойчивости к возмущениям, структурной устойчивости (анализ циклов) для G_1 и G_2.

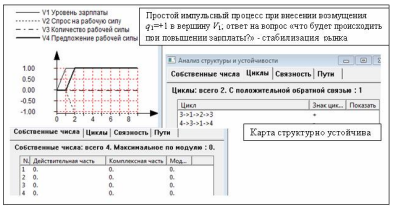


Рис. 3. Результаты анализа когнитивной карты G_1 “Равновесный рынок”

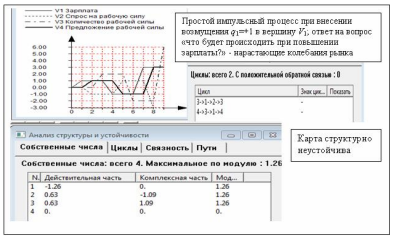


Рис. 4. Результаты анализа когнитивной карты G_2 “Неравновесный рынок”

Поясним график импульсного процесса. По оси абсцисс отложены такты моделирования, по оси ординат – величины импульсов, изменяющихся по мере обхода вершин когнитивной карты. Количество тактов может быть увеличено/уменьшено по желанию эксперта. Возможно внесение возмущений в несколько вершин и на желаемых тактах моделирования для более глубоко анализа возможных сценариев.

Приведенный пример можно рассматривать как определенную иллюстрацию возможностей ПСКМ. Но не только. Когнитивные карты спроса-предложения (не обязательно на РТ) с уже известными свойствами (рис. 3 и 4) могут стать элементами более содержательных когнитивных моделей (рис. 5).

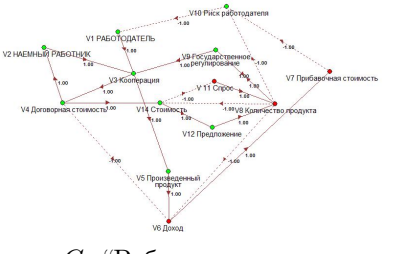


Рис. 5. Когнитивная карта G_3 “Работодатель – наемный работник – рынок”

Модели G_1 и G_2 не отражают всю содержательную глубину необходимого анализа рыночных отношений, поэтому предположения о том, как и почему происходит увеличение/снижение уровня зарплаты, моделируемое единичным импульсом $q_1 = +1$ или $q_1 = -1$, а также к каким последствиям это может привести для участников рынка, например, государства, работодателя, наемного работника, остаются не проанализированными. Поэтому, возвращаясь

к первому этапу исследования, можно уточнять цель исследования и, соответственно, изменять когнитивную карту.

Поясним смысл модели G_3. Основой для ее разработки послужил, главным образом, текст статьи "Источник социального оптимизма" [3]. Согласно ему, работодатель и наемный работник могут находиться в отношениях "*эксплуататор – эксплуатируемый*", поскольку наниматель [3, С. 509–510] "...присваивает себе прибавочную стоимость производимого продукта"; ... "совместно произведенный нанимателем и наемным работником *продукт* является их совместной собственностью, причем доля каждого пропорциональна материальному и трудовому вкладу его. В том, что вклад сторон нередко качественно различен ... – еще не содержится никаких элементов эксплуатации человека человеком. Это всего лишь разделение труда и кооперация. ... Если бы после реализации произведенного продукта доход делился пропорционально вкладу сторон в его производство, то не было бы никакой эксплуатации. Но дело в том, что ... наемный работник ...не хочет и не может ждать реализации продукта и требует немедленной выплаты своей доли, исходя из некоторой договорной более или менее гарантированной его стоимости. Однако, расплатившись с рабочим, хозяин становится единоличным собственником продукта и, пользуясь рыночной конъюнктурой, может продать его дороже той договорной стоимости, исходя из которой производился расчет с наемным работником, присвоив себе эту прибавку, хотя по справедливости должен был бы пропорциональную ее часть доплатить работнику. Вот это и есть эксплуатация". ... Но... "эксплуатация" выступает как форма оплаты коммерческого риска предпринимателя".... "Суть классовых противоречий состоит... в спорах вокруг договорной стоимости произведенного продукта" ... "если договорная стоимость окажется равной рыночной стоимости продукта, то наступает момент "социальной справедливости"...

Исследование одного из вариантов модели G_3 описано в [2]. Изображенная на рис. 5 модель дополнена раскрытием вершины "конъюнктура рынка" вышеприведенной моделью G_1 "спрос-предложение". На рис. 6 представлены некоторые результаты вычислительного эксперимента.

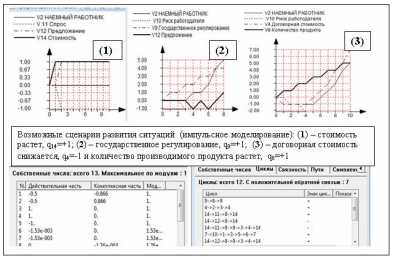


Рис. 6. Некоторые результаты анализа когнитивной карты G_3

Заключение. Результаты, приведенные на рисунках, в данном случае содержательно не интерпретируются. Это задача научного отчета и основа для начала принятия решений исследователем как о когнитивной модели, так и о самой изучаемой системе. Конечной целью их может быть разработка рекомендаций по управлению устойчивым развитием рынка (в том числе, рынка труда) в регионе, на определенной территории, в определенных условиях внешней среды. В этом случае, даже если реальный рынок не удовлетворяет сформулированным требованиям, в результате исследования может быть предложена его структура (когнитивная карта), для реализации которой на практике необходимо будет осуществлять определенные организационные решения.

Литература. **1.** Горелова Г.В., Захарова Е.Н, Радченко С.А. Исследование слабоструктурированных проблем социально-экономических систем: когнитивный подход. – Ростов н/Д: Изд-во РГУ, 2006. – 332 с. **2.** Инновационное развитие социально-экономических систем на основе методологий предвидения и когнитивного моделирования / Под ред. Г.В. Гореловой, Н.Д. Панкратовой. – Киев: Наукова думка, 2015. – 464 с. **3.** А. Денисов. Глазами народного депутата СССР. – СПб: Изд. СПбГТУ, 2006. – 660 с. – С. 509–510.

Городецкий В.Г., Кулиш Р.Д.
КПИ им. Игоря Сикорского, Киев, Украина

Алгоритм идентификации систем с периодическим внешним воздействием

Идентификация систем различного типа – одна из основных проблем теории моделирования. Такая задача предполагает определение параметров модели при наличии информации о входе и выходе системы [1]. В последнее время всё больше внимания уделяется исследованию нелинейных систем с внешним возбуждением, в том числе хаотических.

При этом наиболее интересной является задача реконструкция реальной модели, отображающей физику процесса. Пример такой задачи – параметрическая идентификация при заданном общем виде модели. Кроме этого, наряду с постоянными коэффициентами модели могут быть неизвестны и параметры входного воздействия. Последнее обстоятельство существенно усложняет задачу, даже если это воздействие имеет гармонический характер.

В данной работе для идентификации неизвестной модели предлагается простой алгоритм, требующий меньших затрат времени по сравнению с известными [2]. Особенность данного алгоритма заключается в том, что он не требует воспроизведения временного ряда, сгенерированного реконструированной системой, для его сравнения с исходным и последующей оценкой погрешности реконструкции. Для иллюстрации алгоритма была выполнена идентификация системы Дуффинга [3]

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = x_2, \frac{\mathrm{d}x_2}{\mathrm{d}t} = c_0 cos\omega t + c_1 x_1 + c_2 x_2 + c_3 x_1^3. \quad (1)$$

Реконструкция была произведена по временному ряду наблюдаемой переменной x_1 системы (1), представленному на рисунке 1.

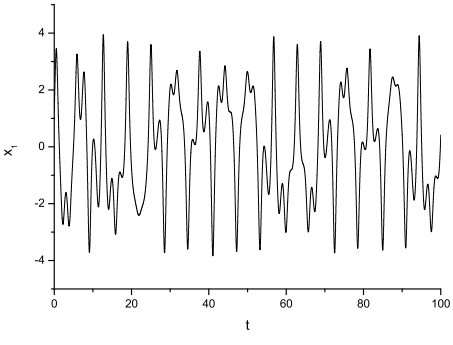


Рис. 1. Временной ряд наблюдаемой переменной x_1 системы (1)

Численные эксперименты подтвердили надежность алгоритма и его сравнительно высокое быстродействие.

Литература. **1.** Льюнг Л. Идентификация систем. Теория для пользователя. – М: Наука – 1991. – 432 с. **2.** Bezruchko B., Smirnov D., Dikanev T., Sysoev I. Construction of dynamical model equations for nonautonomous systems from time series (peculiarities and special techniques // Chaos and its reconstructions, ed. G. Gouesbet, Nova Science Publishers, 2003, pp. 215–243. **3.** Ueda Y. Randomly Transitional Phenomena in the System Governed by Duffing's Equation / Y. Ueda // J. Stat. Phys. – 1979 – vol. 20. – P. 181–196.

Дідковський О.І.
Київський національний університет ім. Тараса Шевченка, факультет інформаційних технологій, Київ, Україна

Модель побудови датасету для вирішення задачі оптимізації енергоспоживання

У доповіді показаний перехід від прогнозування обсягу споживання електроенергії великою кількістю споживачів до прогнозування часу роботи приладів кліматконтролю для кожного окремого споживача. Крім того, репрезентований метод побудови датасету, який забезпечує отримання результатів з високою точністю і низькою похибкою, та розглянуті моделі споживання електроенергії в періоди пікових навантажень.

Відомо, що рівень споживання електроенергії в значній мірі залежить від умов навколишнього середовища. Саме тому можливі значні коливання обсягу спожитої енергії. У випадку екстремальних умов в навколишньому середовищі рівень споживання електроенергії зростає експоненційно, оскільки вмикаються всі електричні прилади, які забезпечують належний рівень комфорту у приміщеннях (житлових, виробничих, тощо). Так, у періоди високих температур інтенсивно використовуються кондиціонери, а у періоди низьких температур – електричні обігрівачі. Нелінійний ріст вартості спожитої електроенергії також пов'язаний зі специфікою функціонування ринку електроенергії: додаткові потужності мають більшу ціну. Наприклад, під час каліфорнійської енергетичної кризи 2000–2001 років зниження споживання електроенергії на 5% приводило до зниження витрат на 50% [1]. Тому існує нагальна потреба в побудові достатньо точних (з невеликою дисперсією) прогнозів споживання електроенергії на основі адекватних моделей енергоспоживання. Крім того, такі моделі доцільно застосовувати і у випадках звичайних температур. Це також може забезпечити зниження витрат на енергозабезпечення.

Але вище згадані моделі не дозволяють будувати якісний прогноз обсягу витраченої електроенергії конкретним споживачем. Також, не завжди можливо отримати кількісні дані енергетичних витрат. Саме тому в цій роботі запропоновано пов'язати споживання електроенергії з температурним режимом у приміщенні (будівлі). Такий метод є доцільним, оскільки температурних даних більше і вони масштабовані. Крім того, коефцієнт кореляції між часом роботи приладів кліматконтролю та обсягом спожитої електроенергії досить значний (>0.8), що буде показано у доповіді. Зазначимо, що мінімальний час прогнозу енергоспоживання у працюючих моделях прогнозування загального рівня споживання становить 1 годину, а заміри температури фіксуються кожні 5 хвилин.

Впливати на обсяг споживання електроенергії, який пов'язаний з охолодженням (нагріванням) приміщень, простіше ніж на загальний обсяг споживання. Це пов'язане з тим, що використання електроенергії для потреб непов'язаних зі зміною температури сильно відрізняється для різних користувачів. Аналіз роботи [2] свідчить, що методи прогнозу електроспоживання за наявності змінних, що характеризують споживання у той самий час дня, але напередодні, працюють і містять багато інформації про міжзмінні зв'язки. Відсутність таких змінних у датасеті не є критичною, якщо прогнозується загальний рівень споживання електроенергії. У роботі [3] показане застосування моделі ARIMA та порівння з методами SVR та OLS. Отримані результати мають незначну похибку, разом з тим в ARIMA були використані дані лише про енергоспоживання. Але, якщо поставлена задача прогнозування для кожного споживача, то знаходження змінних, які зберігають інформацію про міжзмінний зв'язок або інформацію про попердні дні, є критичним.

У доповіді розглянута побудова моделі залежності $T = F(x_1, ..., x_n)$, де T – час роботи приладів кліматконтролю, а $x_1, ..., x_n$ – звичайні змінні, такі як температура у приміщенні, температура встановлена на приладах кліматконтролю, температура на вулиці, тощо. Зауважимо, що аналітичний вигляд залежності T від $x_1, ..., x_n$ нам невідомий.

Нехай T насправді залежить від $\ln(-x_1 \times x_2)$. Тоді, якщо змінної $x' = -x_1 \times x_2$ немає в датасеті, та застосовується метод, який базується на деревах прийняття рішень, потрібно буде як мінімум 2 листи для того, щоб виділити ті ж самі розбиття, які можна було б виділити одним за наявності змінної x' в датасеті.

Розглянемо розбиття листа X_m дерева прийняття рішень по змінній x_j з пороговим рівнем t ($[x_j \leqslant t]$). Тоді x_j, t знаходяться з критерію розбиття:

$$Q(X_m, j, t) \to min_{j,t},$$

де $Q(X_m, j, t)$ – функція похибки. Вона задається формулою:

$$Q(X_m, j, t) = \frac{|X_l|}{|X_m|} H(X_l) + \frac{|X_l|}{|X_m|} H(X_l)\,,$$

де $H(X_r)$ – критерій інформативності підмножини X_r. $|X_l|$, $|X_r|$ – кількість елементів у лівому та правому листі, після розбиття.

У випадку задачі регресії критерій інформативності має вигляд:

$$H(X) = \frac{1}{|X|} \sum_{i \in x} (y_i - \bar{y}(x))^2.$$

У задачах класифікації існує багато різних критеріїв інформативності. Найбільш вживаними є критерій Джині та ентропійний критерій. Нехай при класифікації об'єктів на K груп апостеріорна ймовірність належності об'єкта до групи має вигляд:

$$p_k = \frac{1}{|X|} \sum_{i \in X} [y_i = k].$$

Тоді критерій Джині має вигляд:

$$H(X) = \sum_{k=1}^{K} p_k (1 - p_k),$$

а ентропійний критерій:

$$H(X) = -\sum_{k=1}^{K} p_k \ln p_k.$$

В доповіді показано застосування вищенаведених критеріїв.

У випадку великої кількості змінних можуть виникати ситуації, при яких розділення по x_1, або x_2 не буде. Саме тому у доповіді запропонована модель визначення змінних датасету, яка дозволяє отримати максимальну інформацію з даних.

Дослідження представлене у доповіді дозволяє показати на конкретному прикладі яким чином потрібно змінювати та підбирати величини при побудові датасету. Крім того, така постановка задачі є вкрай важливою для обробки індустріальних даних: аналітики самі створюють датасет та кожного дня може надходити нова інформація, на відміну від ML-змагань, в яких початковий датасет не змінюється протягом часу.

Література. **1.** International Energy Agency (2003). The Power to Choose – Enhancing Demand Response in Liberalised Electricity Markets. *Presentation*, 37–66. **2.** Pierre Dagnely, Tom Ruette, Tom Tourwé, Elena Tsiporkova, Clara Verhelst (15 Dec 2015). Predicting Hourly Energy Consumption. Can Regression Modeling Improve on an Autoregressive Baseline? *Data Analytics for Renewable Energy Integration. Volume 9518 of the book series Lecture Notes in Computer Science (LNCS)*, 105–122. **3.** Hamid R. Khosravani, Maria Del Mar Castilla, Manuel Berenguel, Antonio E. Ruano and Pedro M. Ferreira (20 Jan 2016). A Comparison of Energy Consumption Prediction Models Based on Neural Networks of a Bioclimatic Building. *Energies* 9(1), 1–24.

Донченко В.С.
Київський національний університет ім. Тараса Шевченка, Київ, Україна

Синтез двох фундаментальних підходів розв'язання задачі лінійної дискримінації

Задача лінійної дискримінації (ЛД) за об'єднаною навчальною за двома класами вибіркою $x_i \in R^m, i = \overline{1,n}$ в статистичній постановці веде початок від Р.Фішера. В алгебраїчній постановці ЛД-задача (в дихотомічній постановці) зводиться за певної стандартизації (див. [1,2]) до СЛА Рівнянь (СЛАР) чи Нерівностей (СЛАН) щодо роздільного функціоналу w вигляду $X^T w \geqslant I, I^T = \left(\underbrace{1,...,1}_{n}\right)$. Віднесення до одного з двох класів відноситься за виконанням одної з двох виключних умов відповідно $w^T x \geqslant \Delta, \quad w^T x \leqslant -\Delta, \ \Delta > 0$ – параметр задачі. Відстань за нормаллю між роздільними гіперплощинами класів є пропорційною $\|w\|^{-2}$. В роботі [1] СЛАН стандартним чином зводиться до СЛАР $X^T w = I + \xi, \quad \xi \geqslant 0$ і за цих обмежень ЛД-задача зводиться до оптимізаційної задачі максимізації відстані між роздільними гіперплощинами, тобто – до мінімізації $\|w\|^2$, точніше – за скоригованим функціоналом $\|w\|^2 + c^T\xi : \ c \geqslant 0, c \in R^n$. В роботі [2] теж здійснюється перехід до дослідження розв'язності СЛАР $X^T w = y, \quad y \geqslant I$, але – вже засобами псевдо обернення (ПдО). В кінцевому рахунку розв'язок ЛД-задачі за цього підходу зводиться до пошуку мінімуму додатно визначеної квадратичної форми $y^T(E_n - X^+X)y$ в області $y \geqslant I$.

Ці два підходи можуть бути синтезовані у алгоритмі розв'язання ЛД-задачі, в якому оптимізація функціоналу $\|w\|^2 + c^T\xi$ першого підходу здійснюється у звуженій області. Це звуження полягає у введенні додаткового лінійного обмеження $(E_n - X^+X)(I + \xi) = 0$ в першому підході, еквівалентного умові $y^T(E_n - X^+X)y = 0$ розв'язності ЛД-задачі другого підходу.

Пропонований підхід допускає перенесення на випадок, коли навчальна вибірка складається з матриць на основі концепції кортежних операторів [3].

Література. **1.** Vapnik Vladimir. The Nature of Statistical Learning Theory / John Wiley & Sons. – 732 p. **2.** Кириченко Н.Ф., Лепеха Н.П. Применение псевдообратных и проекционных матриц к исследованию задач управления, наблюдения и идентификации. // Кибернетика и системный анализ. – 2002. – № 4. – С. 107–124. **3.** Донченко В.С., Зінько Т.П., Скотаренко Ф.М. "Вектори Ознак" в задачі групуванні інформації: вектори і матриці. – Problems of Computer Intellectualization. – Institute of Cybernetics NASU, ITHEA. – Kyiv, Ukraine–Sofia, Bulgaria. – 2012. – С. 111–124.

Донченко В.С., Зінько Т.П.
Київський національний університет ім. Тараса Шевченка, Київ, Україна

SVD та псевдообернення в обробці сигналів: векторні та матричні представники

SVD-розвинення лінійних операторів та теорія псевдо обернення (ПдО) за Муром – Пенроузом є потужним аналітичним засобом аналізу в R^n, оскільки надає вичерпні конструктивні можливості дослідження систем лінійних алгебраїчних рівнянь (СЛАР) та розв'язання задачі МНК наближення в таких системах. В евклідових просторах числових векторів SVD- та ПдО-теорії реалізуються як теорії, що описують властивості матриць, хоча SVD- та ПдО-теорії мають ширші, ніж теорія факторизації матриць. Обмеженість ширшого застосування згаданих теорій визначається відсутністю «матричної алгебри» у поданні слушних операторів. Концепція кортежних операторів (КорО) з матричними кортежами [1, 2] уможливила перенеснення можливостей SVD- та ПдО- аналізу матричні евклідові простори $R^{m\times n}$, що дозволяє розглядати матриці як представників замість векторів: векторів ознак. Це, зокрема, є важливим для розпізнавання аудіосигналів, де природним матричним представником аудіо об'єкта є його спектрограма. Те саме стосується і зображень. SVD та ПдО дозволяє реалізовувати процедури обробки сигналів, що ілюструється запропонованою в роботі фільтрація – «очистки» сигналу на основі SVD-подання слушних матриць. Вона полягає у відкиданні молодших членів у SVD-поданні слушного набру об'єктів: матриць чи матричних кортежів, що відповідають навчальній вибірці. Процедура описана на прикладі векторів, хоча техніка КорО дозволяє перенести її у тому самому вигляді на матричні представники. Результат процедури SVD-фільтрації на прикладі мовних сигналів «дерево» та «звук» представлено на Рис. 1.

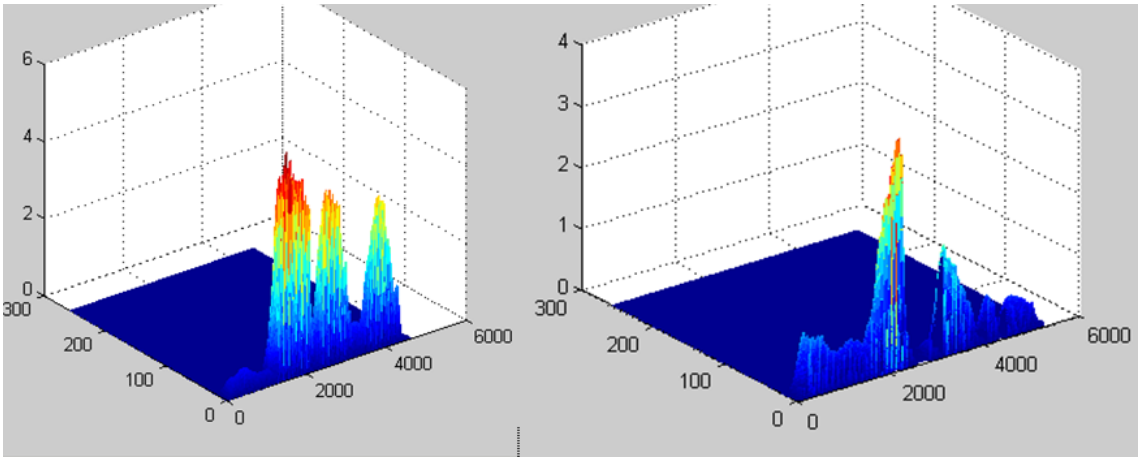


Рис. 1. Результат SVD-фільтрації слів «дерево» та «звук»

Література. **1.** Донченко В.С., Зинько Т.П. Концепция кортежности и ее реализация для матричных кортежей // Международный научно-технический журнал "Проблемы управления и информатики", 2016, №4. – С. 87–99. **2.** Donchenko V., Zinko T., Skotarenko F. Feature vectors in grouping information problem in applied mathematics: vectors and matrixes // Problems of computer Intellectualization. ITHEA, Kiyv, Ukraine–Sofia, Bulgaria, 2012. – P. 111–124.

Донченко В.С., Назарага І.М., Тарасова О.В.
Київський національний університет ім. Тараса Шевченка, Київ, Україна

Матрична множинна регресія для прогнозування макроекономічних та медійних показників: приклади використання

У доповіді розглядаються приклади застосування матричної множинної регресії [1, 2] для прогнозування макроекономічних та медійних показників. З метою розв'язання задачі оцінювання методом найменших квадратів для множинної матричної регресії використано розроблений у [2] математичний апарат сингулярного (SVD) розкладу і техніку псевдообернення (ПдО) за Муром-Пенроузом у рамках розвитку концепції кортежних операторів.

Зокрема, для розв'язання задачі прогнозування телемедійних та макроекономічних показників реалізовано запропонований у [2] алгоритм оцінювання за методом найменших квадратів для вектора невідомих параметрів класу матричних функцій.

Для випадку медіа-показників розв'язується задача прогнозування річної частки шести телеканалів [3] для годинних інтервалів у межах усередненого тижня 2013 року з дискретизацією за днями тижня (понеділок, вівторок, ..., неділя). Використано дані 2006–2013 років з погодинною дискретизацією за днями тижня для усередненого тижня; дані 2013 року використовуються для аналізу модельної точності прогнозування. Задача розв'язується у середовищі Maple із використанням алгоритму обчислення оцінки вектора невідомих параметрів [2], що застосовується для річних часток каналів, дискретизованих до матриць розмірності 24x7 (години x дні тижня). Точність прогнозування розраховано за критерієм APE (absolute percentage error), і у комерційних часових інтервалах значення цього критерію є прийнятним для комерційного планування бізнес-процесів. Також досліджено вплив розміру прогностичної бази даних за кількістю історичних періодів, що використовуються для прогнозування показників медіа-бізнесу, на точність прогнозування за критерієм APE. Як результат, оптимальним є використання історичного періоду у 7–8 років для прогнозування медійних показників. Історичний період такого розміру дозволяє прогнозувати річні медіа-показники у рамках 10% інтервалу точності за критерієм APE.

Для варіанту макроекономічних показників емпіричними даними для розрахунків стали дані Державної служби статистики України за 2007–2015 рр. із квартальною дискретизацією [4]. Запропонований у [2] алгоритм реалізовано у середовищі Wolfram Mathematica і застосовано для прогнозування внутрішнього валового продукту, оплати праці найманих працівників, кінцевих споживчих витрат, експорту товарів та послуг, імпорту товарів та послуг. Розглянуто декілька варіантів групування спостережуваних даних. Згідно із обчисленими значеннями критерію APE (що для різних показників знаходиться в межах від 0% до 20%), точність прогнозування значень річних економічних показників є доброю. Перевищення 10% значення похибки можна пояснити використанням статистичних даних кризових 2008–2009 років, а також даних без урахування тимчасово окупованої території АР Крим, м. Севастополя та частини зони проведення антитерористичної операції.

Таким чином, проведені розрахунки показують, що матрична множинна регресія може бути використана для розв'язання задач прогнозування з прийнятною для планування точністю в розглянутих предметних областях.

Література. **1.** Donchenko V. "Feature Vectors" in Grouping Information Problem in Applied Mathematics: Vectors and Matrixes / V. Donchenko, T. Zinko, F. Skotarenko // Problems of Computer Intellectualization: international conference, Institute of Cybernetics NASU, ITHEA. – Kyiv, Ukraine, – Sofia, Bulgaria. – 2012. – P. 111–124. **2.** Донченко В.С. Матрична множинна регресія / В.С. Донченко, О.В. Тарасова // Вісник КНУ імені Тараса Шевченка. Серія: фіз.-мат. науки. – 2015. – № 2. – С. 133 – 138. **3.** Назаров М. Зарубежные рынки телевизионной рекламы: сравнительное исследование / М. Назаров – Москва: ООО «НИПКЦ Восход–А», 2011. – 364 с. **4.** Офіційний сайт Державної служби статистики України [Електронний ресурс] – Режим доступу: www.ukrstat.gov.ua. – Назва з екрану.

Дорофеев Ю.И., Любчик Л.М., Никульченко А.А.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Дескрипторный подход к синтезу управления запасами в цепях поставок с неопределенными временными задержками

Цепь поставок (ЦП) представляет собой интегрированный производственный процесс, включающий взаимосвязанные производственные узлы, склады и каналы распределения ресурсов, которые добывают сырье, перерабатывают его и поставляют готовую продукцию в розничную сеть либо конечным потребителям. Для управления цепями поставок разработано множество подходов [1]. Большинство публикаций по управлению ЦП посвящено разработке эвристических подходов либо методов математического программирования. Однако, в настоящее время стало понятно, что оптимизация потоков материальных ресурсов в ЦП невозможна без применения системного подхода и методологии теории автоматического управления.

Управление запасами заключается в определении размеров заказов на их пополнение и моментов времени формирования заказов. Выбор стратегии управления запасами определяется характером внешнего спроса. В условиях неопределенности спроса используется концепция «неизвестных, но ограниченных» воздействий, при этом соответствующая модель спроса характеризуется интервальной неопределенностью. Другим источником неопределенности является наличие временных задержек, которые возникают вследствие затрат времени на транспортировку ресурсов между узлами ЦП, затрат времени на переработку сырья и полуфабрикатов, наличия человеческого фактора. Предполагается, что номинальные значения задержек известны. Однако, в процессе функционирования ЦП они могут меняться. В результате возникает необходимость обеспечения робастности системы управления запасами относительно вариаций указанных параметров. Таким образом, определение оптимальных уровней страховых запасов ресурсов в узлах цепи поставок и разработка методов синтеза робастного управления запасами в условиях неопределенности спроса и интервалов временных задержек является актуальной научной и практической проблемой.

В последнее десятилетие значительное внимание уделялось вопросам анализа устойчивости и синтеза управления для систем с запаздыванием. В итоге предложено множество методик и опубликовано достаточно много результатов (см. [2] и многочисленные ссылки в ней). Полученные результаты можно разделить на две категории: те, которые не зависят от временных задержек, и, соответственно, которые зависят. В работах первой категории величина задержки не принимается во внимание. Как следствие, полученные результаты являются консервативными, особенно для систем с малыми задержками. В работах второй категории при синтезе условий устойчивости используется вся доступная информация о задержках, вследствие чего полученные результаты являются менее консервативными. В последнее время для уменьшения степени консерватизма результатов анализа и синтеза систем управления с временными задержками предложен подход на основе дескрипторного описания системы [3].

Дескрипторными называют модели объектов управления, в которых вводятся дополнительные переменные состояний, которые имеют описательный смысл. Благодаря наличию дополнительных алгебраических связей между переменными состояний дескрипторные модели обладают свойствами, которые не характерны для традиционного способа описания систем. В результате дескрипторный подход позволяет получить достаточные условия существования регулятора на основе обратной связи, который обеспечивает робастную устойчивость замкнутой системы и гарантированное качество управления запасами для цепей поставок с дискретным временем и интервально неопределенными временными задержками.

Литература. **1.** Лотоцкий В.А. Модели и методы управления запасами / В.А. Лотоцкий, А.С. Мандель. – М.: Наука, 1991. – 188 с. **2.** Liu P.-L. Further results on the stability analysis of systems with time-varying delay / P.-L. Liu // Int. J. Analysis. – 2013. – ID 721407. – P. 1–11. **3.** Zhang W. Robust stability test for uncertain discrete-time systems: a descriptor system approach / W. Zhang, H. Su, Y. Liang et al. // Lat. Am. Appl. Res. – 2011. – Vol. 41. – No. 4. – P. 359–364.

Ємельянова Д.І.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Інформаційно-програмне забезпечення комплексної оцінки рівня екологічної безпеки системних об'єктів

Актуальність дослідження. Актуальність дослідження обумовлена необхідністю комплексного оцінювання рівня екологічної безпеки об'єктів природно-техногенного походження на основі запровадження інформаційно-методичної підтримки рішень щодо сталого розвитку системних об'єднань.

Мета дослідження. Метою дослідження є визначення складових інформаційно-методичної підтримки комплексної оцінки рівня екологічної безпеки природно-техногенних об'єктів на основі показників екологічної відповідності в системі «техногенний об'єкт – навколишнє природне середовище». Реалізація мети пов'язана з вирішенням таких завдань:

1. Визначення на комплексній основі функціональних можливостей інформаційно-програмного забезпечення з екологічних досліджень складних природно-техногенних утворень «об'єкт – навколишнє середовище» на основі MIPS-аналізу, загально-детальної і детальної оцінки змін в системах за результатами ризик-аналізу «стан$_1$ – процес – стан$_2$»;
2. Апробація методично-програмної розрахункової системи комплексної оцінки рівня екологічної безпеки об'єктів еколого-економічної діяльності.

Методика дослідження і аналіз результатів. Загальна оцінка відповідності стану систем вимогам екологічної якості визначається за MIPS-аналізом з урахуванням матеріальних потоків, пов'язаних з навантаженням на природне середовище з метою виявлення деструктивних порушень, що кількісно характеризуються величинами MI-індексів, які відповідають за матеріальну інтенсивність чинника навантаження. MI-індекс зазначає рівень екологічності фактора дії на об'єкт. При визначенні рівня екологічності природно-техногенних об'єктів за MIPS-аналізом враховуються питома вага чинників навантаження і кількість матеріальних потоків у компонентах навколишнього природного середовища [1].

Відповідність факторів і процесів вимогам стабільного функціонування об'єкта визначається на основі комплексного ризик-аналізу «стан$_1$ – процес – стан$_2$». Кількісна ризик-оцінка встановлюється за індексним методом на основі нормативно-допустимого значення факторів навантаження з урахуванням їх класу небезпеки. Ризик змін у внутрішньому просторі системи визначається за логарифмічною функцією з урахуванням індексів, що становлять величини негативної дії з позиції імовірності трансформації дестабілізуючих факторів, які призводять до стабілізації або дестабілізації кінцевого стану «система – навколишнє середовище (НС)».

На практиці запропоноване методичне забезпечення на основі MIPS- і ризик-аналізу надано для систем «об'єкт – НС», послідовне використання методів дозволяє підтверджувати результати загальної оцінки рівня екологічної безпеки за MIPS-аналізом загально-детальною і детальною ризик-оцінкою. Такий системний підхід забезпечує комплексну оцінку природно-техногенних об'єктів, яка становить інформаційну основу для прийняття рішення щодо регулювання стану системи при підтримці процесів її стабілізації [2].

Висновки. Таким чином, на основі комплексного підходу визначення інформаційно-методичної підтримки прийняття рішення сформовано упорядковану систему аналізу «стан$_1$ – процес – стан$_2$» для управління якістю і безпекою системних об'єктів з урахуванням їх внутрішніх механізмів стабілізації і особливостей зовнішнього зв'язку систем з навколишнім середовищем відповідно до отриманих результатів визначення рівня екологічності природно-техногенних комплексів і виявлення ризик-факторів дестабілізації за розробленим алгоритмічно-програмним забезпеченням комплексної оцінки екологічної безпеки складних об'єктів (рис. 1).

Результати програмної реалізації показали перспективність запровадження комплексного підходу для розв'язання практичних задач оцінки рівня безпечності функціонування природно-

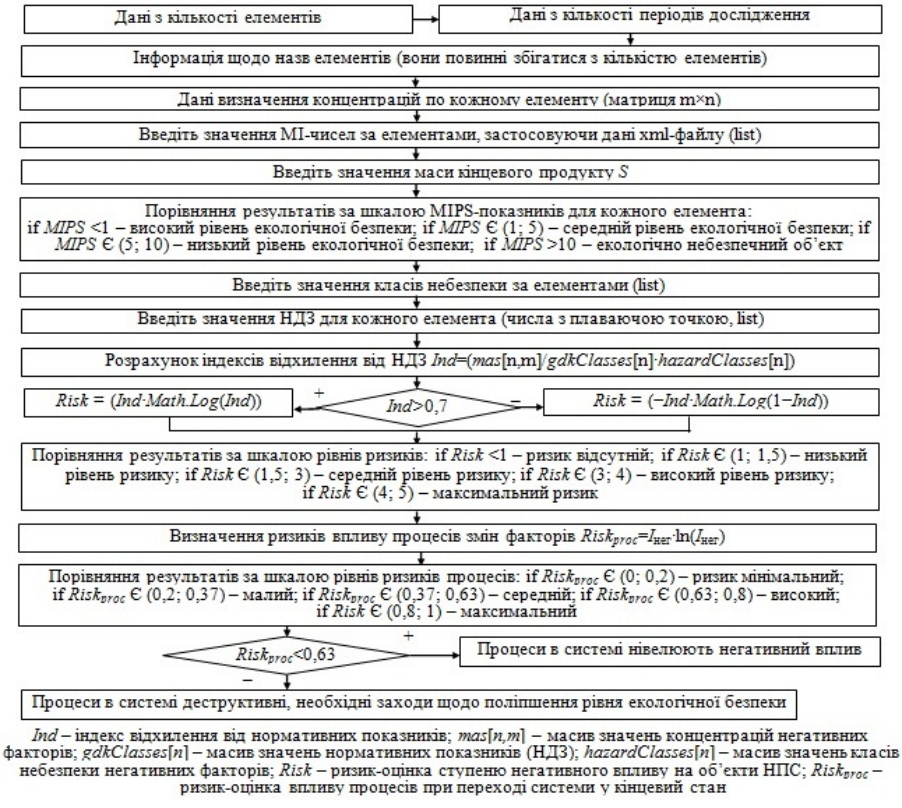
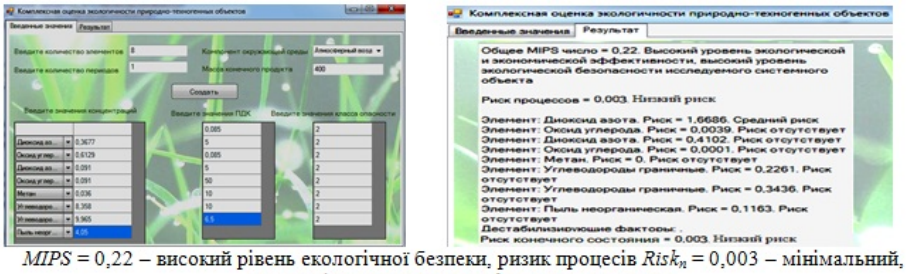


Рис. 1. Схема алгоритму програмного розрахунку комплексної оцінки екологічної безпеки системних об'єктів

техногенних об'єктів (рис. 2).



MIPS = 0,22 – високий рівень екологічної безпеки, ризик процесів $Risk_n$ = 0,003 – мінімальний, ризик кінцевого стану $Risk_2$ = 0,03 – низький

Рис. 2. Фрагмент програмної реалізації визначення рівня екологічної безпеки природно-техногенного об'єкта

Література. **1.** Козуля Т.В. Теоретико-методичні основи комплексного аналізу та оцінювання екологічності природно-техногенних об'єктів / Т.В. Козуля, Д.І. Ємельянова // Системні дослідження та інформаційні технології – К.: ННК «ІПСА» НТУУ «КПІ», 2016. – № 1. – С. 73–84. **2.** Козуля Т.В. Екологічний ризик на різних рівнях дослідження природно-техногенних систем, інформаційне забезпечення його оцінки/ Т.В. Козуля, Д.І. Ємельянова // Проблеми інформаційних технологій. 2015. – № 17. – С. 138–144.

***Єремєєв І.С.*[1], *Єщенко О.І.*[2]**
[1] *Таврійський національний університет імені В. І. Вернадського, Київ, Україна;* [2] *КПІ ім. Ігоря Сікорського, Київ, Україна*

Про один підхід до аналізу систем теплопостачання

Останнім часом у світі багато уваги приділяють концепції сталого розвитку. Ця концепція передбачає поєднання природних і техногенних циклів у єдину систему з метою керування цими циклами за рахунок «буферних» зон, які існують об'єктивно (природні цикли), або створення технологічних циклів [1].

Системи теплопостачання (СТП) можна розглядати як складні системи, що здатні до адаптації в умовах змін зовнішніх та внутрішніх чинників, причому структура СТП до певної міри може розглядатися як фрактальна структура, тобто структура, що складається з самоподібних елементів, які відрізняються лише масштабами (наприклад СТП будинку нагадує СТП мікрорайону і централізовану СТП цілого міста: усі вони містять як складові генератори тепла, розподільчі мережі, локальні теплообмінники, лічильники тепла тощо).

На будь-якому рівні деталізації (індивідуальний обігрівач, домова, мікрорайонна чи централізована система) можна спостерігати однакові компоненти циклів-«вісімок» [2]: розвиток (**r**), консервація (**K**), криза (ω) і реорганізація (α), у тому числі й реструктуризація (**s**). Початок кожного з циклів (сезонного, чи багаторічного або життєвого) починається з ***розвитку***: встановлення відповідних технічних засобів, накопичення ресурсів (енергії), зв'язків, розгалуження структури (мережі) тощо, після чого настає період ***консервації*** усього накопиченого, сталого функціонування і, одночасно, перманентного локального доопрацювання (на підставі отриманого досвіду та з урахуванням зміни вимог, зв'язків, ресурсів) до моменту початку ***кризи*** (соціально-економічні і суто технічні проблеми, природні чи антропогенні катаклізми тощо), коли технічні засоби вимагають масової заміни або суттєвого налаштування (у тому числі переходу до нових джерел палива), накопичені ресурси витрачаються для компенсації кризових впливів, а встановлені зв'язки порушуються, нарешті починається період ***реорганізації*** (адаптації), коли система або пристосовується до умов, що склалися, і повертається до старої «орбіти», або починає реструктуризацію, в наслідок якої може перейти на іншу «вісімку» (**s**).

Процес «функціонування» моделі СТП при цьому можна представити наступним чином (Рис. 1). Якщо приростання ресурсу (ефективності) ПР відбувається одночасно з приростанням зв'язності (складності) ПЗ, що відповідає «розвитку» системи (РЗВ) – модель функціонує адекватно. Якщо у якійсь точці траєкторії (наприклад **A**) виявиться від'ємне приростання ресурсу, це стає сигналом кризи (КР), що наближається, і необхідності проведення локальної штучної кризи (ЛШКР) з метою адаптації (АД) до нових умов чи вимог з урахуванням ціни ризику (ЦРЗ), після чого система може вийти на більш високий рівень як з точки зору ресурсу, так і зв'язності. Це сприятиме більш ефективному і довгому нормальному функціонуванню системи в режимі консервації (КС).

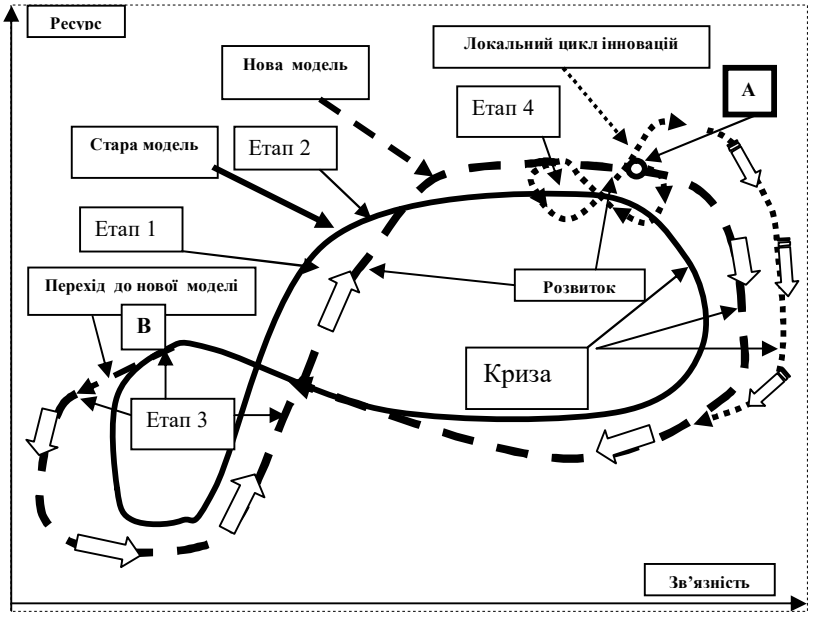


Рис. 1. Функціонування СТП у разі, коли стара модель вже не відповідає сучасним вимогам і треба її змінити

Траєкторія руху згідно старої моделі (етапи 1 і 2)
Траєкторія руху згідно нової моделі (етапи 3 і 4)
Траєкторія локального циклу (локальна «вісімка»)

A – точка на траєкторії, в якій ПР має від'ємний знак внаслідок локальних кризових явищ

B – точка переходу на нову модель (етап 3 – реорганізація)

Сформульовано наступні евристики для забезпечення адаптаційних процесів в системі:

ЯКЩО ([ПР >0] ТА [ПЗ >0]), ТО (РЗВ);
ЯКЩО ([ПР ~ 0] ТА [ПЗ >0]), ТО (КС);
ЯКЩО ([ПР <0] ТА [ПЗ <0]), ТО (КР);
ЯКЩО ([ПР >0] ТА [ПЗ <0]), ТО (РОР);
ЯКЩО ([ПР <0] ТА [ПЗ ~ 0]), ТО (АД);
ЯКЩО ([ПР <0] ТА [ПЗ >0] ТА [ЦРЗ $\leqslant$ ЦРЗ$_{\text{прип}}$]), ТО (ЛШКР).

Тут ЦРЗ$_{\text{прип}}$ – максимально прийнятна ЦРЗ за даних умов.

Зазначений підхід допоможе своєчасно «відчути» проблеми, що накопичуються під час експлуатації СТП, а також обумовлені впровадженням нових ідей і технологій, і наблизить еру широкого впровадження методів та засобів створення «розумних» будинків не тільки у сфері приватного сектору, але й у більш розгалуженій сфері багатоквартирного житла.

Література. **1.** Статюха Г.А. Устойчивое развитие – концепция, подходы и модели. Системний аналіз та інформаційні технології. Матеріали Міжнародної конференції SAIT 2011. – К.: УНК «ИПСА» НТУУ «КПИ», 2011, с. 38. **2.** Holling C.S. Understanding the Complexity of Economic, Ecological and Social Systems. Ecosystems, 2001, #4, pp. 390–405.

Заводник В.В., Ноголь О.С.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Оценивание макроэкономической динамики страны на базе статистики межотраслевого баланса и когнитивного моделирования

В докладе авторами представлено решение задачи оценивания динамики изменения темпа роста ВВП страны (рассмотрена макроэкономическая статистика национальных счетов Украины 2003–2015 гг., с пробелом, приходящимся на 2012 год, и России 1998–2006 гг.) по средствам исследования межотраслевого баланса, таблица «расходы – выпуск», проведенного на базе методологии когнитивного моделирования.

Макроэкономика государства, как сложная многоуровневая, иерархическая система, функционирующая в условиях информационной неопределенности (неполноты, недостоверности и несвоевременности) и довольно влиятельного человеческого фактора является идеальным объектом для применения методологии и принципов когнитивного моделирования. Что, в свою очередь, позволяет построить и оценить сценарии стратегического планирования и проектирования развития динамики экономических процессов в краткосрочной и долгосрочной перспективе.

В работе [1] подробно изложены алгоритм построения когнитивной карты взаимодействий, элементов исследуемого объекта, ее исследования и анализа в рамках когнитивного моделирования. На первом этапе реализации предлагается использование методологии предвидения на базе различных методов качественного анализа. Как правило, в качестве исходного материала предлагается использовать результаты экспертных оценок.

Авторы доклада считают, что эксперт хоть и ориентируется на свою интуицию и жизненный опыт не является источником объективной информации. Во-вторых, для того, чтобы предвидеть, нужно придумать «новый» мир, а экономическая наука может только анализировать и моделировать динамику текущих экономических процессов, количественно измеряемую на множестве нормативно определенных параметров. Поэтому предложено уйти в количественную конкретику реальных статистических данных, за достоверность которых отвечают не эксперты, а государственные инстанции.

Данные, в свою очередь, были взяты из двух источников [2,3] и представляют собой развернутые таблицы межотраслевого взаимодействия субъектов экономики двух стран, Украины и России. В них выбирается внутреннее межотраслевое потребление, то есть материальные расходы отраслей для производства нового продукта, товаров или услуг. По этим данным строятся когнитивные карты взаимодействия. Номенклатура отраслей за указанный временной период 3 раза менялась и охватывала от 15 до 24 отраслей. Это представляет собой базовый набор. Существует еще расширенный набор, состоящий из 80 подотраслей. Были рассмотрены два варианта: 15 отраслей; агрегированный – 5 отраслей (сельское, рыбное и лесное хозяйство, промышленность, строительство, транспорт и связь, услуги).

Промышленность, например, включает в себя: добывающую промышленность; обрабатывающую промышленность; поставку электроэнергии, газа и кондиционированного воздуха; поставку воды и отведения отходов. В свою очередь, обрабатывающая промышленность включает в себя: добычу угля, торфа, урановых руд; добычу нефти и газа; добычу металлических руд; разработку карьеров камня, песка и глины; добычу соли.

Построенные когнитивные матрицы, содержащие стоимостную информацию, довольно большие цифры, взаимодействия указанных отраслей, то есть использование для производства продукции друг друга, нормировались согласно формуле $b_{ij} = \frac{0.9\star(a_{ij}-\min_{i,j} a_{ij})}{\max_{i,j} a_{ij}-\min_{i,j} a_{ij}}, i, j = \left[\frac{1,5}{1,15}\right]$, а затем исследовались на устойчивость [1].

В результате проведенного анализа данных, полученных на основе расчетов, выявлены некоторые закономерности изменения динамики, структуры годовых когнитивных карт, построены графики изменения по годам норм матриц A5 и A15 для Украины (рис. 1) и России (рис. 2).

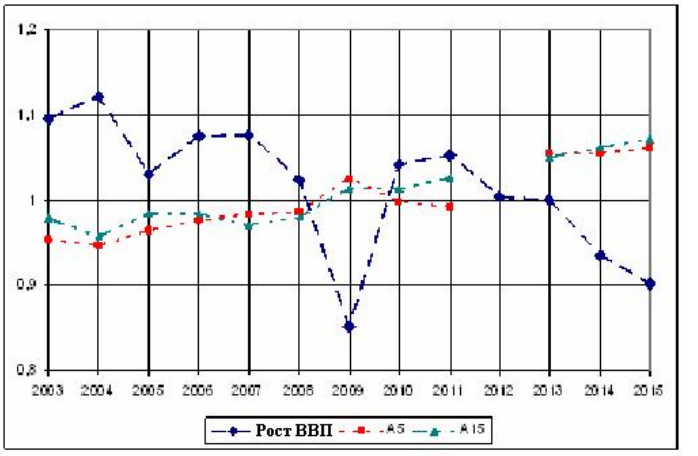


Рис. 1. Украина, динамика: рост ВВП; нормы матриц A5 и A15

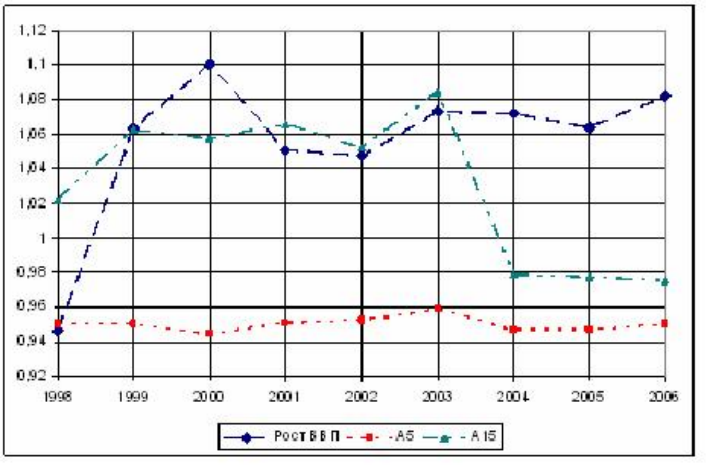


Рис. 2. Россия, динамика: рост ВВП; нормы матриц A5 и A15

Выявлены значительные расхождения результатов исследования матриц. Если по российской статистике матрицы A5 по всем годам удовлетворяют критериям 1, 2 [1], является устойчивой по возмущению и по значению (норма матрицы A5 менше 1), что в свою очередь, отражают состояние устойчивого экономического роста по указанным годам, то для матриц A15 эти условия удовлетворяются только для 2003–2006 гг. Для Украины, результаты исследования матриц A5 и A15 практически совпадают и полностью отражают динамику изменения темпа роста ВВП Украины.

Литература. **1.** Горелова Г.В., Панкратова Н.Д. Инновационное развитие социально- экономических систем на основе методологий предвидения и когнитивного моделирования. – К.: Наукова думка, 2015. – 464 с. **2.** Государственная служба статистики Украины, электронный ресурс, www.ukrstat.gov.ua **3.** Федеральная служба государственной статистики РФ, электронный ресурс, http://www.gks.ru.

***Захарчук Н.Г.*[1], *Ладогубець Т.С.*[2], *Фіногенов О.Д.*[1]**
[1]КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна; [2]КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Визначення індексу випадкової узгодженості в програмному забезпеченні, що реалізує MAI

Теоретичні відомості. Індекс випадкової узгодженості (ІВУ) в MAI використовується для визначення узгодженості думок експертів. Найбільш поширеними є 2 шкали ІВУ, які були наведені Т. Сааті [1,2]. Ряд авторів при відтворенні експериментів з визначенням ІВУ, отримали відмінні значення для матриць відповідної розмірності, огляд деяких представлений [3]. Відмінність в результатах викликана різницею методів обчислення максимального власного числа та розмірністю вибірок (К) випадкових матриць (від К = 50 до К = 50000 та більше). Тому в наукових статтях та при використанні програмного забезпечення, що реалізує MAI, можливе виникнення суперечливості результатів внаслідок використання різних констант для обчислення оцінки узгодженості (ОУ).

У більшості СППР ІВУ не відображається у явному вигляді, тому необхідний непрямий метод його визначення. У випадку якщо надані числові значення максимального власного числа матриці парних порівнянь або індекс узгодженості (ІУ) і оцінка узгодженості (ОУ), то ІВУ можна обчислити наступним чином (1):

$$\text{ІВУ} = \frac{\text{ІУ}}{\text{ОУ}} = \frac{\lambda_{max} - n}{n-1} \cdot \frac{1}{\text{ОУ}} \tag{1}$$

Опис тестового прикладу. Найпростіше визначити шкалу при оцінці узгодженості матриці N=3, для якої в обох шкалах, визначених Т. Сааті, спостерігається найбільша відмінність: ІВУ(N=3) = 0.58 [1], ІВУ(N=3) = 0.52 [2]. Тестовий приклад має наступний вигляд (рис. 1, табл. 1).

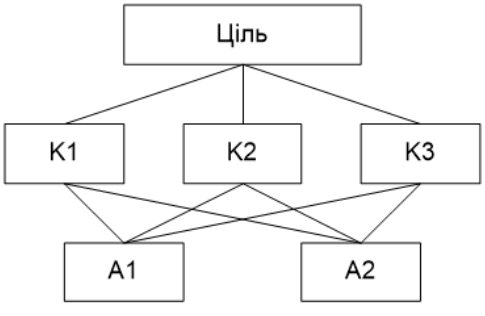


Рис. 1. Схема ієрархії

Табл. 1. Матриця парних порівнянь критеріїв

К	К1	К2	К3
К1	1	1/9	1/6
К2	9	1	4
К3	6	1/4	1

Нехай за критеріями К1-К3 альтернативи рівноправні. У даному випадку неузгодженість ієрархії визначатиметься неузгодженістю матриці критеріїв, так як матриця при N=2 завжди узгоджена. Для розглянутої матриці максимальне власне число $\lambda_{max} = 3.107847$. Оцінка узгодженості при використанні шкал з [1, 2] буде дорівнювати:

Табл. 2. Оцінка узгодженості для тестового прикладу

	ІВУ = 0.58 (шкала [1])	ІВУ = 0.52 (шкала [2])
ОУ	0.093	0.104

Зазвичай рівень порогу узгодженості прийнято вважати за 0.1. Таким чином у першому випадку (з використанням шкали [1]) матриця буде узгодженою (0.093 < 0.1), в іншому – ні. У випадку коли СППР не надає дані про оцінку узгодженості, проте контролює рівень порогу

(0.1) даний приклад також дозволить виявити використовувану шкалу.

Аналіз СППР. Для прикладу були розглянуті наступні програмні продукти (табл. 3):

- NooTron [4] – українське програмне забезпечення, розроблене в Національній металургійній академії України (м. Дніпро).
- Імператор [5], MPriority [6], СППР «Вибір» [7] – засоби, розроблені російськими вченими та студентами.
- BPMSG AHP priority calculator [8], Transparent choice [9] – веб-застосунки.
- Priority Estimation Tool (AHP) [10], Expert Choice [11] – іноземне ПЗ.

Табл. 3. Результати визначення ІВУ

№	Назва програмного продукту	λ_{max}	ІУ	ОУ	ІВУ
1	NooTron	*	*	*	0.58
2	Імператор 3.1	*	0.0539	0.11	0.49
3	Priority Estimation Tool (AHP)	*	*	0.093	*
4	BPMSG AHP priority calculator	*	*	0.112	*
5	Transparent choice	*	*	*	*
6	MPriority 1.0	3.1078	0.0539	0.0929	0.5802
7	СППР “Вибір”	3.111	0.056	0.096	0.583
8	Expert Choice	*	*	0.1	*

* – системи не надають інформацію про відповідні дані.

Іноді інформація про шкалу ІВУ, що використовується в програмних продуктах, надається в документації до системи (наприклад, [4]). СППР «Вибір» та Trasparent Choice не проводять контролю узгодженості, тому визначити шкалу для ІВУ неможливо. Тільки в BPMSG AHP priority calculator матриця критеріїв (табл. 1) позначалась як неузгоджена, а це свідчить про те, що для ІВУ була використана шкала [2]. Програма Імператор надає тільки значення ІУ та ОУ (0.0539 та 0.11) відповідно. Врахувавши їх, обчислене значення ІВУ рівне 0.49, що не відповідає жодній зі шкал, запропонованих Т. Сааті. Всі ж інші програмні продукти використовують шкалу [1], для якої ІВУ = 0.58.

Висновки. Наведені результати обчислення ІВУ в програмних продуктах показують, що результат вибору залежить не тільки від оцінок експертів, а і від використовуваного програмного забезпечення з огляду на відмінності у шкалах ІВУ та у значенні порогу узгодженості.

Література. **1.** Саати Т. Принятие решений. метод анализа иерархий / Томас Саати. – Москва: Радио и связь, 1993. – 278 с. **2.** Saaty T.L. Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process / Thomas Saaty., 2000. – 478 с. **3.** Панкратова Н.Д. Моделі і методи аналізу ієрархій. Теорія. Застосування: навч. посібник / Панкратова Н.Д., Недашковська Н.І. – К.: ІВЦ “Видавництво «Політехніка», 2010. – 372 с. **4.** NooTron [Електронний ресурс]. – Режим доступу: http://nootron.net.ua. **5.** Імператор [Електронний ресурс]. – Режим доступу: http://kazus.ru/programs/viewdownloaddetails/kz_0/lid_6134.html. **6.** Діалогова система MPriority 1.0. [Електронний ресурс]. – Режим доступу: http://www.tomakechoice.com/mpriority.html. **7.** СППР “Выбор” [Електронний ресурс]. – Режим доступу: http://www.softportal.com/software-7763-sppr-vibor.html. **8.** BPMSG AHP priority calculator [Електронний ресурс]. – Режим доступу: http://bpmsg.com/academic/ahp_calc.php. **9.** Transparent choice [Електронний ресурс]. – Режим доступу: https://www.transparentchoice.com. **10.** Priority Estimation Tool (AHP) [Електронний ресурс]. – Режим доступу: https://sourceforge.net/projects/priority. **11.** Expert Choice [Електронний ресурс]. – Режим доступу: http://expertchoice.com.

Кангун Г.І., Бідюк П.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Моделювання військових конфліктів методами інтелектуального аналізу даних

В наш час військові конфлікти стали частиною повсякденності для багатьох країн світу, включаючи Україну. Якісне та всебічне моделювання може стати важливим кроком на шляху до їх вирішення через відображення найважливіших аспектів та виявлення глибинних, неочевидних залежностей і прогнозування можливих сценаріїв розвитку.

Моделювання являє собою дослідження будь-яких явищ, процесів або систем об'єктів шляхом побудови і вивчення їх моделей; використання моделей для визначення або уточнення характеристик і раціоналізації способів побудови нових об'єктів. Моделювання – одна з основних категорій теорії пізнання: на ідеї моделювання по суті базується будь-який метод наукового дослідження – як теоретичний (при якому застосовуються різного роду знакові і абстрактні моделі), так і експериментальний, який використовує предметні моделі. Коли мова заходить про відображення політично-воєнних процесів, загальнонауковий підхід до моделювання вимагає деякого уточнення. Моделювання в політиці є описом політичних процесів за допомогою обмеженої кількості значущих чинників. Скорочення числа факторів проводиться за рахунок абстрагування від тих з них, чиїм впливом на процес можна з будь-яких підстав нехтувати.

Основною метою роботи є огляд існуючих прикладів застосування методів інтелектуального аналізу даних до моделювання військових конфліктів та отриманих у роботах, що розглядаються, результатів.

Складність моделювання військового конфлікту полягає в першу чергу в тому, що він є складною системою з багатьма невідомими або невимірюваними параметрами. Крім того, необхідно враховувати неповноту та нечіткість інформації. Аналіз та моделювання подібної системи має на меті не тільки виявлення існуючих взаємозв'язків та залежностей, а і передбачення та попередження майбутніх конфліктів на основі результатів аналізу.

Одним з методів аналізу та моделювання військових конфліктів є метод експертних оцінок, приклади практичного застосування якого розглядаються у роботах [1,2]. Авторами роботи [1] метод експертних оцінок було ефективно застосовано для цілей воєнно-політичного прогнозування на прикладі Росії. Недолік, пов'язаний із суб'єктивністю оцінок, було ліквідовано значною кількістю експертів. В свою чергу у роботі [2] для дослідження можливих стратегій Росії для вирішення проблеми Південних Курил було сформовано ієрархічну цільову структуру досліджуваної проблеми та застосовано метод аналізу ієрархій Т. Сааті. Представлені результати використання методу аналізу ієрархій показали його достатню ефективність. Формування ієрархічних цільових структур забезпечило чітку структуризацію досліджуваної проблеми, що дозволить суттєво спростити роботу аналітиків та експертів. Кількісні оцінки елементів "дерева цілей" дають можливість приймати більш обгрунтовані рішення при виборі тієї чи іншої стратегії поведінки в конфлікті.

Можливість застосування нейронних мереж та, зокрема, Баєсівської нейронної мережі досліджується в роботах [3, 4]. Незважаючи на велику кількість зібраних даних та проведення складного аналізу, статистичні показники варіюються у різних джерелах і емпіричні результати часто незадовільні. Автори [3] пропонують нейромережеву модель, що робить ефективний прогноз, базуючись на простій ідеї того, що стабільність, сила та повторюваність причин конфлікту напряму впливає на вірогідність його розвитку. Мережа у [4] моделює відносини між набором вхідних параметрів та одним виходом, що може приймати значення «мир» та «конфлікт». Для початкового вибору вхідних параметрів було застосовано підходи когнітивного моделювання, застосування якого до проблеми військових конфліктів будо докладно розглянуто в роботах [8,9]. В результаті було обрано наступні параметри: Демократія, Союзники, Непередбачені обставини, Відстань, Можливості, Основна Сила. Було виявлено найважливіші з обраних вхідних параметрів та їх ваги щодо кінцевого результату. Відповідно,

найважливішим виявився параметр Залежності, після нього було розташовано Можливості, Демократія та Союзники.

Робота [8] окреслює підхід до формалізації конфлікту, побудови його математичної моделі та дослідження конфлікту як багатомірного, неоднозначного та слабко передбачуваного явища. Отримана модель дозволяє бачити тенденції розвитку та можливості процесів, а спеціальні процедури дозволяють уникнути необхідності попереднього накопичення статистичного матеріалу. Методику формування множини факторів воєнного конфлікту та системи взаємозв'язків між ними наведено у [9]. Викладено метод формалізації негативних та різнознакових взаємозв'язків між елементами когнітивних карт з використанням логіки антонімів. З метою виразу негативних та різнознакових зв'язків між елементами когнітивних карт введені відповідні оператори, сформовані їх основні властивості.

Згідно з авторами статей [5–7], у вирішенні проблеми моделювання військових конфліктів себе позитивно показує застосування підходів теорії ігор. Гра – це теоретичне уявлення ситуації, розгляд якої дозволяє практикуватися з метою отримання досвіду і придбання майстерності в справі прийняття рішень, а також дає основу для проведення експериментів по виробленню нових стратегічних і тактичних концепцій і їх перевірці. У [5] представлено динамічну гру з багатьма акторами. Підхід дає базис для формування коаліцій. Симуляції в роботі [6] дають правдоподібні шаблони динаміки конфліктів, які в'яжуться як з історичними конфліктами, так і зі стандартними постулатами воєнної теорії. Робота [7], в свою чергу, пропонує узагальнену модель військового конфлікту, в якій рівнозначно враховувалися рішення щодо просторового руху і часового розподілу ресурсів, а також різні послідовності взаємодій.

В даній роботі було розглянуто різні підходи до моделювання військових конфліктів. Отримані результати складаються як з теоретичних здобутків, так і конкретних – у вигляді обраних раціональних стратегій для конкретного актору у конкретній ситуації (наприклад, стратегія Росії щодо Курильського конфлікту [2]). Отже, незважаючи на те, що військовий конфлікт є складною багаторівневою системою, а дані по ньому завжди є неповними, нечіткими або частково невимірюваними, його моделювання, а також прогнозування на базі побудованої моделі за допомогою методів інтелектуального аналізу є цілком реальною задачею. Створення такої моделі зробить можливим глибше розуміння та раціональне вирішення сучасних конфліктів, полегшить та прискорить процес прийняття рішень у військовій сфері. Базуючись на моделях та аналізі минулих та сучасних конфліктів, стане можливе прогнозування та попередження конфліктів у майбутньому.

Література. **1.** Богатырёв Э.Я., Макиев Ю.Д., Малышев В.П. Анализ методов военно-политического прогнозирования. / Богатырёв Э.Я. // Стратегия гражданской защиты: проблемы и исследования. – 2013. – Том 3. – №2(5). **2.** Абаев Л.Ч. Моделирование кризисов и конфликтов в международных отношениях: опыт практического применения. / Абаев Л.Ч. // Проблемы национальной стратегии. – 2012. – № 1. – С. 57–72. **3.** Beck N., King G., Langche Z. Improving Quantitive Studies of International Conflict: A Conjecture. / Beck N. // American Political science review. – 2000. – Vol. 94. – №1. – Pp. 21–35. **4.** Marwala, T. and M. Lagazio, Modelling and Controlling Interstate Conflict, Budapest / Marwala T. // Hungary: IEEE International Joint Conference on Neural networks. – 2004. **5.** Sheffran J. Calculated Security? Mathematical Modelling of Conflict and Cooperation. / Sheffran J. // Gut-zum-Druk. – 2003. – Pp. 390–412. **6.** Klaus B. Beckmann and Lennart Reimer Dynamics of Military Conflict: an Economics Perspective / Klaus B. // Review of Economics. – 2014. – Bd. 65. – H. 2. – Pp. 193–213. **7.** Sheeba P.S., Ghose D. A resource based game theoretical model for military conflicts / Sheeba P.S. // DRDO-IISc Programme on Advanced Research in Mathematical Engineering. – 2005. **8.** Рома О.М., Толубко В.Б., Лєнков С.В. Спосіб формалізації конфлікту та його системне моделювання / Рома О.М. // Ukrainian Information Security Research Journal. – 2009. – Том 11. – № 3(44). **9.** Адаменко А.А. Методика формування множини факторів воєнного конфлікту та системи взаємозв'язків між ними / А.А. Адаменко // Наука і техніка Повітр. сил Збройн. сил України. – 2011. – № 1. – С. 141–145.

***Каніщева О.В.*[1], *Главчева Ю.М.*[1], *Висоцька В.А.*[2]**
[1] *Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;*
[2] *Національний університет "Львівська політехніка", Львів, Україна*

Визначення стилю автора для виявлення плагіату в академічному середовищі

Актуальність роботи. Стратегічне завдання наукових та освітніх закладів – формування нової наукової еліти, яка спроможна внести вклад у розвиток держави шляхом упровадження інноваційних передових розробок у промислове виробництво, сільське господарство, медицину, IT сектор тощо. Ці заходи сприятимуть сталому розвитку економіки держави.

Більшість урядових та громадських організацій спрямовують свої зусилля на розгляд питань, вирішення яких сприятиме створенню умов для підвищення ефективності науки та освіти. Явище плагіату є тим фактором, що негативно впливає на якість науки та освіти [1, 2]. Академічна доброчесність (подібностей) у наукових роботах має сьогодні широке поняття. Це сукупність етичних принципів та визначених законом правил, якими мають керуватися учасники освітнього процесу під час навчання, викладання та провадження наукової (творчої) діяльності з метою забезпечення довіри до результатів навчання та/або наукових (творчих) досягнень. Забезпечення академічної доброчесності вимагає системного підходу до вирішення та реалізації комплексу організаційних, навчальних, технічно-технологічних заходів. Серед останніх важливу роль відіграють: створення депозитарію академічних текстів установи та Національного репозитарію академічних текстів (НРАК); використання інформаційних систем (ІС) виявлення плагіату в наукових роботах, яка буде виявляти подібності на основі даних НРАК та Інтернет.

У роботі [3] плагіат поділяють на чотири види, де кожен має своє цільове призначення. В залежності від типу діяльності та сфери застосування: професійний (присвоєння інтелектуальних, творчих, професійних здобутків інших у професійних цілях), освітньо-науковий (присвоєння чужого інтелектуального майна виключно у процесі здобуття наукового ступеня, освітнього кваліфікації або визнання у цих напрямках), соціальний (аналогічно професійному, але у побутових відносинах і не відноситься до фахової діяльності); нормативний (привласнення законодавчих, юридичних, методичних, наукових, практичних напрацювань).

В залежності від форми авторами цієї роботи виділено наступні види плагіату:

1. Точне/часткове копіювання авторської роботи у мовній, лексичній, технологічній інтерпретації (ідентифікують більшість антиплагіатних ІС).
2. Повторення основної ідеї (визначити доволі складно у зв'язку із відсутністю методів видобування та аналізу змісту документу антиплагіатними ІС).
3. Плагіат, який передбачає наявність певних посилок: у посиланнях; у визначені цитат; посилання на неіснуючі джерела; у наведенні точних фактів з не власних досліджень без конкретизації джерела; у поданих джерелах (частково реалізовано у деяких антиплагіатних ІС).
4. Плагіат як результат замовлення деяким псевдоавтором у комерційній організації або у приватної особи та видачі її за власну роботу (визначення стилю автора, на основі його вже існуючих робіт, що є актуальним завданням і досліджується в межах даної роботи).

Метою даної роботи є визначення стилю автора наукових статей для визначення плагіату на більш глибокому рівні. Цю задачу розглянемо як задачу класифікації текстів (рис. 1), тобто кожний документ відноситься тільки до конкретної категорії/класу (автора). У роботі у якості даних використано одноосібні наукові статті 100 науковців України, з кожного по 10 статей. Таким чином отримано колекцію з 1000 текстів. Для кожної категорії визначено еталонні показники, які характеризують його стиль. Для кожного з досліджених текстів визначено коефіцієнт належності до конкретної категорії. Для визначення стилю автора використані методи машинного навчання (Байєсівський класифікатор та метод опорних векторів). Також досліджено за допомогою нашого корпусу кількісні оцінки мови, такі як

коефіцієнт різноманітності тексту, ступінь синтаксичної складності, зв'язність мовлення, індекс винятковості і концентрації тексту [4], які використовують для визначення стилю автора.

Рис. 1. Загальна схема роботи для визначення стилю автора

Етап предобробки тексту досить трудомісткий та складається з первинного опрацювання лінгвістичних даних (побудова рядів розподілу, обчислення статистик, статистичних оцінок та інші параметри лінгвометрії), лексикографічного опрацювання текстових даних (створення частотних і алфавітно-частотних словників конкретного автора, словників-конкодансів, слововказивників, зворотних словників, словників ключових слів стилю автора тощо) та методів стилеметрії для визначення авторської атрибуції тексту. Без статистично опрацьованого доробку автора (еталону) це практично зробити неможливо. Аналіз та інтерпретація на лінгвістичному рівнів стилістичних особливостей і закономірностей письменницького стилю конкретного автора побудований на методі контент-аналізу та складається з наступних етапів: коректний відбір текстів; лематизація текстових одиниць; усунення неоднорідності текстових одиниць; побудова словників та організація їх на основі статистичних розподілів у потрібних у необхідних частотних словникових шкалах; пошук параметрів, що адекватно відображають структуру частотного словника; перевірка параметрів на ефективність; математичне моделювання та аналіз лексико статистичних розподілів; побудова статистичних класифікацій; інтерпретація результатів та визначення еталонних коефіцієнтів різноманітності тексту конкретного автора. Важливим фактором є побудова коректних словників ключових слів стилю автора. Для досягнення мети дослідження було розроблено систему, що дозволяє обрати мову або мови, з яких складений текст. Доступ до процесу знаходження множини ключових слів з врахуванням основ тематичних слів можна отримати на ресурсі Victana. При побудові будь-яких словників з україномовних текстів необхідно враховувати основи цих слів без флекцій.

Висновки та основні результати. У роботі розглянуто актуальність задачі визначення плагіату, проаналізовано види плагіату, описано підхід для визначення стилю автора. Авторами досліджені методи машинного навчання та кількісні оцінки мови для вирішення задачі визначення стилю автора на прикладі розроблено корпусу українських академічних робіт. Окремою проблемою є не одноосібні публікації, а авторського колективу. Визначення стилю в такому випадку є трудомістким, так як один стиль автора накладається на інший стиль. В такому випадку визначити, що така робота є «комерційною» (замовленою) досить складно.

Література. **1.** Лупаренко Л.А. Інструментарій виявлення плагіату в наукових роботах: аналіз програмних рішень // Інформаційні технології і засоби навчання. – 2014. – Т. 2. – №. 40. – С. 151–169. **2.** Болілий В.О. Перевірка унікальності тексту при оцінюванні студентських робіт творчого або дослідницького характеру / В.О. Болілий, В.В. Копотій // Наукові записки НДУ ім. Гоголя, Психолого-педогогічні науки // Збірник наукових праць. – 2011. – № 7. – С. 134–145. **3.** Петренко В.С. Поняття та види плагіату. / В.С. Петренко // Часопис цивілістики. – 2013. – Вип. 14. – С. 128–131. **4.** Математична лінгвістика. Кн. 1. Квантитативна лінгвістика: навчальний посібник для вузів / За ред., передмова Володимир Володимирович Пасічник; Юрій Миколайович Щербина, Вікторія Анатоліївна Висоцька, Тетяна Валеріївна Шестакевич. – Львів: Новий Світ–2000, 2012. – 358 с.

Кобера В.В., Каніовська І.Ю.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Застосування реальних опціонів в управлінні ризиками інвестиційних проектів

Розглядаються фінансові ризики інвестиційного проекту і управління ними за допомогою методу та інструментарію реальних опціонів.

Предметом дослідження є фінансові відносини між державою і приватними інвесторами щодо реалізації інвестиційних проектів.

Метою дослідження є розробка моделі управління фінансовими ризиками інвестиційного проекту, що передбачає використання реальних опціонів в умовах невизначеності.

Застосування реальних опціонів для оцінки ризиків інвестиційних проектів дозволяє об'єктивно оцінити вплив різних економічних чинників на фінансові результати в умовах невизначеності і прийняти обґрунтовані управлінські рішення на ранніх стадіях розробки проектів.

В ході дослідження було розкрито зміст процесу управління фінансовими ризиками як сукупності заходів впливу на грошові потоки з метою мінімізації збитків і підвищення прибутку, включаючи виявлення та оцінку наслідків ризику для економічних суб'єктів, визначення ймовірності настання ризику, прийняття рішень про входження в проект або відмову від нього, проведення попереджувальних, захисних та компенсаційних заходів, вибір альтернативних варіантів.

Вплив фінансового ризику на результати інвестиційної діяльності в умовах невизначеності відображено через побудовану економіко-математичну модель управління фінансовими ризиками з використанням опціонів. Дана модель передбачає використання реальних опціонів як фінансового інструменту для обґрунтування доцільності укладення контрактів з інвесторами – учасниками проекту. В рамках моделі розроблена схема, що відображає послідовність етапів застосування реальних опціонів в якості фінансового інструменту в цілях мінімізації ризиків інвестиційних проектів. Дано рекомендації по впровадженню запропонованої моделі з урахуванням наявності стимулів підвищення ефективності реалізації інфраструктурних проектів, що дозволяє збільшити грошові потоки і майбутню доходність інвестицій.

Теоретична значимість дослідження полягає в розвитку аспектів теорії державних і корпоративних фінансів; обґрунтуванні доцільності використання реальних опціонів в сфері управління фінансовими ризиками, збільшенні наукових знань теоретичного характеру, розкривають сутність, зміст і роль інвестицій в інфраструктурні проекти.

Практична значимість отриманих результатів полягає в доказі ефективності застосування моделі управління фінансовими ризиками з використанням методу реальних опціонів при оцінці інфраструктурних проектів.

Література. **1.** Воронцовский, А.В. Управление рисками с помощью условных и безусловных срочных контрактов на продажу продукции / А.В. Воронцовский // Финансы и бизнес. – 2006. – № 3. – 130–141. **2.** Грачева, М.В., Ляпина, С.Ю. Управление рисками в инновационной деятельности / М.В. Грачева, С.Ю. Ляпина. – М.: Юнити-Дана, 2010. – 352 с. **3.** Маевский, Ф.В. Риск-менеджмент. Превентивное управление: учебно- практическое пособие / Ф.В. Маевский, В.Н. Вяткин, В.А. Гамза. – Днепропетровск: Наука и образование, 2013. – 264 с. **4.** Семенкова, Е.В. Операции с ценными бумагами / Е.В. Семенкова. – М.: Дело, 2009. – 488 с. **5.** Халл, Дж.К. Опционы, фьючерсы и другие производные финансовые инструменты / Дж.К. Халл // ИД "Вильямс". – Москва–Санкт-Петербург–Киев. – 2008. – 890 с. **6.** Чернова, Г.В., Кудрявцев, А.А. Управление рисками / Г.В. Чернова, А.А. Кудрявцев // ТК Велби, Изд-во Проспект. – М., 2008. – 190 с.

Козуля М.М., Моїсеєв В.Ф.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Програмна підтримка ентропійно-інформаційної оцінки відповідності стану складних систем

Актуальність і мета роботи. Сучасна теорія системного аналізу передбачає розгляд і надання відповідної аналітичної оцінки стану складних утворень у вигляді соціально-економічних, соціально-екологічних, еколого-економічних і природно-техногенних систем, які об'єднуються у соціально-еколого-економічні об'єкти дослідження з інформаційною компонентою, яка враховує якісні характеристики як окремих елементів, так і їх відношень. Отже, актуальним для вирішення проблемних завдань у соціальній, екологічній, економічній діяльності та в сферах їх взаємодії є розробка комплексного методичного забезпечення щодо надання системних характеристик стану та динаміки складних об'єктів дослідження на основі впровадження методів математичної логіки та теорії інформації [1, 2].

Для прийняття рішень з розв'язання задачі оцінки відповідності стану природно-техногенних систем вимогам безпеки сформувати комплекс системних методів оцінки якості системних об'єктів на основі об'єктивної обробки даних і подання результатів, надати інформаційно-програмну підтримку методичного забезпечення для вирішення практичних завдань [1].

Розв'язання завдання дослідження. Зазвичай при дослідженні природно-техногенних комплексів використовуються моніторингова інформація і база знань (феноменологічний підхід) для системного аналізу і формування в його межах моделей і рішень з управління якістю «об'єкт – навколишнє соціально-природне середовище». Кількість інформації для встановлення точного опису ξ при наявності відомого і достатнього обсягу значень $\eta = y_j$ визначається за функцією

$$S(\xi|\eta = y_j) = -\sum_i p_{i|j} \log_2 p_{i|j},\ MS(\xi|\eta) = -\sum_j P(\eta = y_j) \sum_i p_{i|j} \log_2 p_{i|j} \tag{1}$$

де $S(\xi|\eta)$ – умовна ентропія η при $\xi = x$; $MS(\xi|\eta)$ – математичне очікування умовної ентропії при змінній ξ [3].

Відповідно до постановки задачі якості складних систем кількість інформації відносно ξ, що міститься в отриманих результатах попереднього етапу розв'язку задачі чи визначається відносно характеристик системи η, дорівнює різниці між такими залежностями [1, 4]:

$$I(\xi|\eta) = S(\xi) - MS(\xi|\eta);\ I(\eta|\xi) = -\sum_i p_{ij} \log_2 \frac{p_{ij}}{P\{\xi = x_i\} P\{\eta = y_j\}} \tag{2}$$

де $I(\xi|\eta)$ – кількість інформації в ξ відносно η; $I(\eta|\xi)$ – кількість інформації в η відносно ξ; $S(\xi)$ – ε-ентропія; $MS(\xi|\eta)$ – математичне очікування умовної ентропії при змінній ξ; p_{ij} – сумісний розподіл ймовірностей.

Оскільки ξ і η визначають однакові співвідношення: кількість інформації в η відносно ξ і кількість інформації в ξ відносно η, то $I(\eta|\xi) = S(\xi)$. Ця послідовність аналізу стану елементів аналітичної системи N, що складає множину X змінних x, дозволяє отримати оцінку відповідності через ентропію, яка визначається за інформацією для певної ситуації, тобто $S = I = \log_2 N$.

На практиці даний підхід об'єктивного системного дослідження складних об'єктів апробовано при розв'язанні задач екологічної безпеки при дослідженні модуля очищення шламових вод діючої збагачувальної фабрики ЦОФ «Кураховська» і установки з перезбагачення шламів зі шламонакопичувача. Необхідно визначити оптимальні умови кінетики осаду завислих речовин з шламових вод за станом кінцевих зразків очищення при використанні мішалки і за умов природного осадження. При визначенні оптимальних умов осаду для отриманих зразків установлюється міра невідповідності цільовим вимогам за ентропійною функцією n виміряних

характеристик розрахунку L:

$$S = \frac{L}{\sum_{i=1}^{n} L_i} \tag{3}$$

Отримані результати для кожної ентропійної характеристики аналізують відповідно до пошуку відрізків, де система досягає свого максимуму та мінімуму:

$$\left[\min; \min + \Delta S \cdot 0,38\right]; \ \left[\max - \Delta S \cdot 0,38; \max\right] \tag{4}$$

де $\Delta S = S_{\max} - S_{\min}$ – довжина відрізку аналізованого розподілу значень. Якщо розглянута характеристика потрапляє до зазначеного інтервалу $\left[a, b\right]$ – інтервал для max/min (4), то отримують оптимальне значення ідентифіковано за правилом:

$$K = \begin{cases} 1, x \in \left[a, b\right]; \\ 0, x \notin \left[a, b\right], \end{cases} \tag{5}$$

Так як у системі необхідно провести градуювання одночасно за декількома параметрами, доцільно ввести кон'юнкцію для виконання одночасно всіх обмежень:

$$t^{\min} \wedge H_{\text{о}}^{\max} \wedge H_{\text{з}}^{\min} = 1, \tag{6}$$

де t – час кінетики осаду, хв.; $H_{\text{о}}$ – висота освітленого розчину; $H_{\text{з}}$ – висота згустку.

Для забезпечення безперервності виробничих процесів, оперативного реагування на небезпечні ситуації на промислових об'єктах при взаємодії «система – навколишнє середовище (НС)» запропонована програмна підтримка контролю безпеки у вигляді мобільного додатку. Програмний продукт орієнтовано на користувачів гаджетів з операційною системою Android, яка поширена на багатьох розповсюджених електронних пристроях, що надає можливість оперативно визначати оцінки стану складних природно-техногенних утворень. При формуванні програмного продукту врахована різниця між версіями операційної системи з можливістю встановити додаток на різні електроні пристрої, починаючи з версії Android 4.1 і до останньої версії 6.0. У розробці враховані ситуації залучення до роботи телефонних пристроїв і планшетних комп'ютерів з операційною системою Android, що дозволяє отримати кількісні розрахунки відповідності виробничих процесів цільовій функціональності – оцінка рівня безпеки (рис. 1).

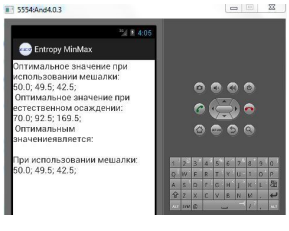


Рис. 1. Значення точок дослідження

Висновки. Таким чином, на основі запропонованого методичного забезпечення з ентропійно-інформаційної оцінки відповідності стану систем установленим вимогам, програмної розробки підтримки вимірювально-розрахункових операцій створена контрольна система безпеки для дослідження виробничих процесів і управління їх якістю на рівні «система – НС».

Література. **1.** Козуля М.М. Знання-орієнтована комплексна методика оцінки стану складних систем / М.М. Козуля – Вісник НТУ «ХПІ» Серія: Інноваційні дослідження у наукових роботах студентів – Харків, НТУ «ХПІ», – 2016. – №19(1191) – С. 39–44. **2.** Козуля М.М. Знание-ориентированные системы и энтропийно-информационные модели в определении качества сложных систем / М.М. Козуля // Материалы III Всероссийской конференции и школы для молодых ученых «Системы обеспечения техносферной безопасности». – Таганрог: ЮФУ, 2016. – С. 146–147. **3.** Моклячук Т.О. Методи оцінки екологічного ризику від забруднення стійкими пестецидами / Т.О. Моклячук // Збалансоване природокористування. – 2014. – № 2. – С. 135–142. **4.** Петров В.В. Энтропийные методы проектирования сложных информационных систем / В.В. Петров, В.М. Агеев, Н.В. Павлова // Итоги науки и техники. Техническая кибернетика. – М., 1985. – Т.18. – С. 78–123.

Козуля Т.В., Білова М.О.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Графологічні моделі з аналізу взаємодії «системний об'єкт – навколишнє середовище» при вирішенні завдань безпеки

Актуальність дослідження. Визначається необхідністю комплексної оцінки стану та рівня рівноважної функціональності систем навколишнього природного середовища (НПС) при вирішенні завдань екологічної безпеки та управління на основі послідовного застосування відповідним чином удосконалених методів аналізу складних об'єктів. Проблемні задачі з визначення загальної відповідності вимогам екологічної безпеки різнорідних системних об'єктів дослідження спрямовані на встановлення факторів з дестабілізації стану об'єктів навколишнього середовища (НС) і гармонізації соціально-економічного розвитку. Одним із природних механізмів досягнення гомеостатичної взаємодії «складна система – НС» є посилення самоорганізаційних факторів, що забезпечують перебіг процесів, які призводять до рівноваги «природна (екологічна) система – соціально-економічна система», що потребує нових підходів для дослідження взаємодії і зв'язків між системою і середовищем, в якому вона функціонує.

У роботі пропонується система інформаційно-методичного забезпечення оцінки ступеня екологічної безпеки природно-техногенних об'єктів у межах аналітичної системи «стан – процес – стан» з виявленням факторів підтримки їх сталого розвитку. Метою роботи є запровадження топологічного моделювання як складової комплексного оцінювання відповідності стану і функціональності системного об'єкта вимогам екологічної якості на основі системного аналізу взаємодії «об'єкт – НС» [1].

Методика дослідження і аналіз результатів. Прийняття рішень з управління якістю об'єкта згідно з вимогами екологічної безпеки має здійснюватися відповідно до його структури, характеристик взаємодії його складових частин між собою і зовнішнім середовищем. При структуруванні системи дослідження запропоновано використати графологічні засоби моделювання – топологічні графи. Загалом ці моделі застосовують для оцінки стану комп'ютерних, електричних мереж тощо [2]. У роботі для інформаційного забезпечення оцінки впливу НС на природно-техногенні об'єкти використані ГІС технології з надання структури «об'єкт – НС» у вигляді топологічного графа.

Процес інформаційної взаємодії «система – НС» є неперервним, характеризується певними особливостями функціонування системи та співвідношенням з елементами НС, що встановлюється за результатами системного аналізу графічних моделей.

Просторове положення досліджених об'єктів характеризується тематичними складовими даних ГІС і включає в себе дві взаємопов'язані частини – позиційну (тополого-геометричну) і непозиційну (атрибутивну). Топологічна модель відображає зв'язки систем з НС у вигляді впорядкованого набору ланок, взаємозв'язок між якими визначається дугами. Топологічність означає зберігання в моделі об'єкта цілісності «система – НПС» відповідно до встановлених безперервних просторових зв'язків (рис. 1).

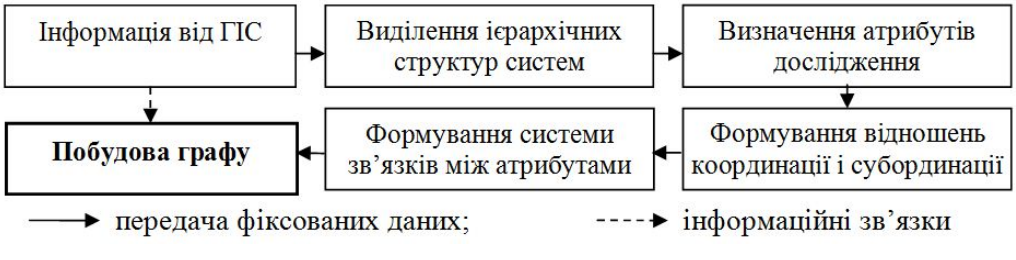


Рис. 1. Побудова топологічного графу як моделі «система – навколишнє середовище»

Застосування топологічної моделі дає можливість використати джерела ГІС для встановлення зв'язку між різнорідними системами об'єкта дослідження: природними, тобто екосистемами, соціальними і економічними, для структуризації природно-техногенного об'єкта і проведення аналізу «стан – процес – стан». Такий підхід дозволить надати інформаційне забезпечення для виявлення процесів, які є ваговими та відповідальними за досягнення стаціонарності в системі «об'єкт дослідження – НС» [3]. Відповідно до класичних положень екологічної безпеки, якість НС залежить від рівня впливу виробництва на природні системи довкілля. Специфіка функціонування промислового підприємства визначається умовами стану «система – навколишнє середовище», що встановлюється на основі даних за графоаналітичною моделлю у вигляді топологічного графу відповідно до даних ГІС (рис. 2).

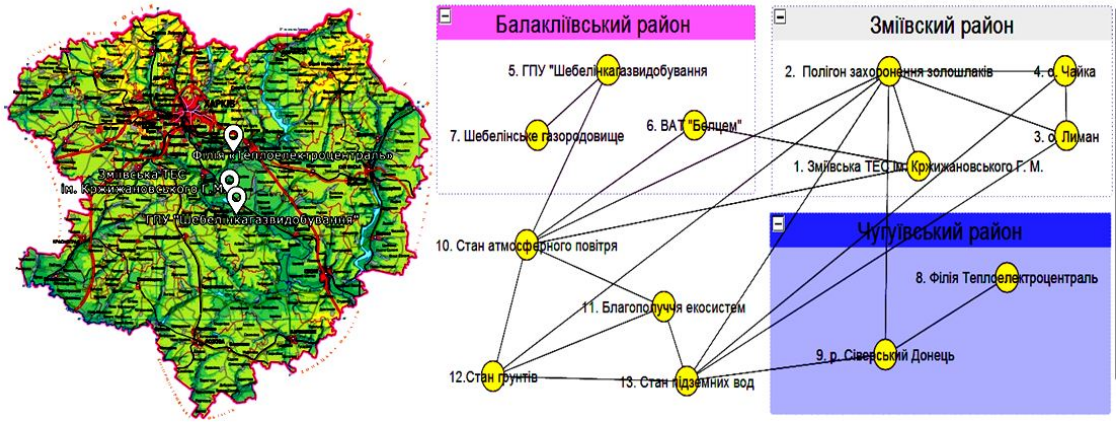


Рис. 2. Оцінка даних «об'єкт – навколишнє середовище» за ГІС у вигляді топологічного графу

Таким чином, формування топологічного графу дозволяє визначити особливості функціонування об'єкта дослідження, обумовлені станом середовища, у якому він перебуває та з яким пов'язаний зв'язками різної природи.

Висновки. У роботі визначено особливості структуризації природно-техногенного об'єкта дослідження еколого-соціально-економічного змісту з позицій вивчення взаємодії «система – НС» для вирішення завдань комплексної оцінки екологічної якості системного угруповання на основі запровадження в інформаційне-методичне забезпечення графологічних методів моделювання, а саме моделей виду топологічних графів.

Література. **1.** Козуля Т.В. Інформаційно-методична основа визначення системи оцінки екологічності складних об'єктів / Т.В. Козуля, М.О. Білова // Системный анализ и информационные технологии: материалы 18-й Международной научно-технической конференции SAIT 2016. – К.: УНК «ИПСА» НТУУ «КПИ», 2016. – С. 96–97. **2.** Основы топологии [Электронный ресурс]. – Режим доступа: http://desktop.arcgis.com/ru/arcmap/10.3/manage-data/topologies/topologybasics.htm. **3.** Белова М.А. Комплексный подход к оценке безопасности состояния системных объектов различного уровня организации / М.А. Белова // Материалы III Всероссийской конференции и школы для молодых ученых «Системы обеспечения техносферной безопасности». – Таганрог: ЮФУ, 2016. – С. 34–36.

Коршевнюк Л.О.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Ранжирування ризиків в задачах ризик-менеджменту

Вступ. При розв'язанні практичних задач ризик-менеджменту в системах різної природи, як правило, стикаються з принциповими обмеженнями на ресурси, що можуть бути використані для запобігання та зменшення негативних проявів ризиків. Це спричиняє неможливість своєчасної обробки та формування адекватної відповіді на всі виявлені загрози та нові ризики. Саме тому актуальною постає проблема ранжирування ризиків в задачах ризик-менеджменту за ознаками їх критичності та значущості щодо обсягів потенційних втрат та можливості загибелі системи внаслідок настання ризикових ситуацій.

Ранжирування ризиків та визначення пріоритетів ризикових ситуацій надають можливість встановити послідовність які ризики необхідно опрацювати і врахувати в першу чергу, які відкласти, а які взагалі доцільно проігнорувати.

Розв'язання задачі. Позначимо несприятливу подію, за якої відбувається або можливий ризик втрати певного ресурсу Z системи як ризикову ситуацію S. Ресурсами, що знаходяться під ризиком можуть бути, наприклад, відстань, час роботи пристрою, запас певних матеріалів, кількість працівників, фінансові активи, температурний режим, обсяг інформації та будь-який інший важливий і вимірюваний показник для функціонування системи. Якщо певна несприятлива подія може викликати втрати одночасно декількох ресурсів, то відповідно будемо розглядати декілька окремих ризикових ситуацій для кожного такого ресурсу.

Ризикова ситуація характеризується множиною факторів причин, що можуть викликати цю ситуацію, часом її настання, ризиком, який може викликати дана ситуація, та значеннями часу настання такого ризику. Фактори ризику характеризуються відповідними вимірюваними показниками. Відхилення вимірюваних значень за даними показниками від заданих або досягнення ними певної комбінації означає фактичне настання конкретної ситуації ризику.

Опис ризикової ситуації S для досліджуваної задачі ранжирування за пріоритетом врахування ризиків пропонується формалізувати на основі суттєвих інформативних ознак відповідного їй ризику у такий спосіб [1, 2]:

$$S =< B, Z_{CR}, D, P, L, LD >, \tag{1}$$

де S – достовірність даних; Z_{CR} – ступінь критичності ресурсу Z системи, який знаходиться під ризиком втрат; D – стрімкість настання ризику; P – ступінь ризику – ймовірність настання ризику; L – рівень ризику – обсяг потенційних втрат ресурсу Z; LD – витрати на запобігання проявам ризику та на відновлення системи після заподіяних ризиком втрат.

Всі суттєві ознаки ризикових ситуацій, за якими виконується їх ранжирування, приводяться до вигляду, при якому більше значення ознаки відповідає більшій значущості та більшому пріоритету врахування ризику. Тому показник часу, через який може настати ризик, тобто може відбутись втрата ресурсу, у формальному описі перетворений на зворотній показник D названий стрімкістю настання ризику. Менший період часу до настання ризику буде відповідати більшій стрімкості настання ризику, яка може бути визначена у такий спосіб:

$$D = \frac{T}{t}, \tag{2}$$

де T – горизонт прогнозування, тобто найбільший період часу, на який в системі можуть прогнозуватись ризики; t – період часу, через який може відбутись втрата певного ресурсу Z від моменту настання ризикової ситуації.

Задача ранжирування ризикових ситуацій S полягає у тому, щоб за відомими значеннями ознак з (1) розрахувати відповідний кількісний показник пріоритету PRT ситуації та відповідного їй ризику. Для цього визначення пріоритету кожної ризикової ситуації відбувається в два етапи: 1) надання оцінок виявленому ризику за кожною його ознакою і 2) подальше об'єднання усіх оцінок ризику у єдине значення пріоритету PRT.

Необхідно зазначити, що на етапі об'єднання оцінок різноманітних ознак дослідники стикаються з проблемою порівняння та агрегування значень довільних несумісних одиниць виміру і різних за своєю природою та сенсом. Так,

- достовірність даних B – безрозмірна якісна величина або значення у порядковій шкалі для бального оцінювання;
- критичність ресурсу Z_{CR}, який знаходиться під загрозою втрат, – також безрозмірна якісна величина або значення у порядковій шкалі для бального оцінювання;
- стрімкість настання ризику D – відносна безрозмірна величина;
- ступінь ризику P – ймовірність настання ризику є безрозмірною величиною в інтервалі $[0; 1]$;
- рівень ризику L – обсяг потенційних втрат вимірюється у власних одиницях виміру відповідних ресурсу Z, наприклад: пропускна полоса каналу зв'язку у Гц; об'єм машинної пам'яті у гігабайтах; час роботи у хвилинах; практична дальність дії апарату у метрах; тиск у мм ртутного стовпа або у Паскалях; обсяг фінансів у грошових одиницях, тощо;
- витрати на зменшення ризику LD – вимірюється у одиницях виміру відповідних ресурсу Z або у єдиному мірилі вартості ресурсів системи, наприклад, у грошових одиницях.

Для ранжирування ризикових ситуацій і відповідного визначення їх пріоритету на основі значень ознак, що мають різні одиниці вимірів та природу, пропонується підхід кількісної уніфікації оцінок за ознаками. За таким підходом, ще не етапі оцінювання всі важливі ознаки ризикової ситуації приводяться до єдиних відносних одиниць виміру, та оцінки надаються у відсотках від мінімально можливого до максимально можливого значення. Або вже надані оцінки ознак у своїх відповідних одиницях вимірів нормуються до одиниць виміру у відсотках і шкали від 0 до 100 відсотків. Нормування відбувається за формулою [2]:

$$X^{norm} = \frac{X - X^{min}}{X^{max} - X^{min}} 100\%, \tag{3}$$

де X – значення оцінки певної ознаки за її базовими одиницями виміру; X^{min} і X^{max} – відповідно найменше та найбільше можливі значення ознаки ризикової ситуації для даного режиму функціонування системи. Зважена сума значень оцінок за ваговими коефіцієнтами ознак ризикової ситуації дасть значення її пріоритету [2]:

$$PRT = a_B B + a_Z Z_{CR} + a_D D + a_P P + a_L L + a_{LD} LD, \tag{4}$$

де a_X – ваговий коефіцієнт відповідної ознаки X. Вагові коефіцієнти ознак є нормованими і їх сума дорівнює одиниці:

$$a_B + a_Z + a_D + a_P + a_L + a_{LD} = 1. \tag{5}$$

Такий спосіб додавання непорівнюваних величин різної природи у єдиний показник пріоритету виглядає не достатньо коректним у сенсі об'єднання оцінок багатокритеріального вибору, хоча формально значення будуть приведені до однакових відносних шкал. Зазначимо, що для задачі визначення пріоритетів ризикових ситуацій важливим є ранжирування значень пріоритетів та упорядкування між собою, а не їх абсолютні значення. Більше значення пріоритету ризикової ситуації буде відповідати більшій загрозі від даного ризику, які обробляються послідовно від першого з найбільшим пріоритетом до останнього.

Висновки. Запропонована і формалізована множина суттєвих ознак ризикових ситуацій для встановлення значень пріоритетів ризиків. Запропоновано та обґрунтовано підхід ранжирування ризиків на основі зваженої суми значень оцінок ознак ризикових ситуацій приведених до єдиних відносних одиниць виміру.

Література. **1.** Коршевнюк Л.О. Формалізація постановки задачі керування ризиками в системах різної природи / Л.О. Коршевнюк, П.І. Бідюк // Наукові вісті НТУУ «КПІ». – 2013. – № 6. – С. 49–54. **2.** Коршевнюк Л.О. Визначення пріоритетів ризикових ситуацій при керуванні динамічними системами / Л.О.Коршевнюк // Механіка гіроскопічних систем: Системи та процеси керування. – 2015. – № 30 (2015). – С. 5–16.

Куценко А.С., Товажнянский В.И.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Упрощенная математическая модель тепловых процессов зданий

Математическая модель является необходимым компонентом при современной постановке задачи управления. В то же время, реальные технологические процессы зачастую не поддаются адекватному математическому моделированию, ориентированному на специфику задач управления. К такому классу управляемых процессов относится процесс теплоснабжения зданий. Предлагаемые в [1] математические модели тепловых процессов зданий основаны на представлении ограждений в виде системы дифференциальных уравнений в частных производных, что затрудняет их использование при решении задачи синтеза закона управления тепловой мощностью. Применение комбинированного принципа управления предъявляет менее жесткие требования к адекватности математической модели, используемой в канале управления по возмущению, поскольку ее неточность компенсируется каналом управления по отклонению. Возможность представления распределенного теплового процесса как конечномерного обоснована в [2]. Там же показано, что во многих случаях можно ограничиться одномерным представлением динамики теплопередачи через ограждение. Опираясь на эти результаты, предлагается упрощенная математическая модель теплового процесса отапливаемого здания, имеющая структуру системы линейных дифференциальных уравнений с постоянными коэффициентами:

$$\frac{dT}{d\vartheta} = -[1 + 2(\rho + \omega)]T + \rho T_p + q + (1 + \rho + \omega)T_o,$$

$$\frac{dT_p}{d\vartheta} = \xi T - \xi(1 + \omega)T_p + \xi\omega T_o + \xi q,$$

$$T_v = \frac{1}{1 + \rho + \omega}(q + T + \rho T_p + \omega T_o),$$

где ξ, ρ, ω – безразмерные критерии подобия зданий, определяемые на основе теплофизических свойств материалов, конструкций ограждений, перегородок и остекления здания, ϑ-безразмерное время, q – тепловая мощность (в температурных единицах измерения); T, T_p, T_v, T_o – усредненные температуры ограждения, перегородок, внутреннего воздуха и окружающей среды соответственно. Приведенная математическая модель имеет второй порядок, содержит одно управляющее – q и одно возмущающее воздействие T_o, а также скалярный выход T_v. Особенностью предлагаемой математической модели является обоснованный отказ от учета аккумулмрующих свойств внутреннего воздуха в связи с его незначительной теплоемкостью по сравнению с теплоемкостями ограждения и перегородок. Предложенная математическая модель позволяет качественно, а во многих случаях и с достаточной степенью точности количественно прогнозировать температурное состояние отапливаемого помещения в зависисмости от изменения погодных условий при выбранном законе регулирования тепловой мощности, а также синтезировать системы управления теплоснабжением на основе индивидуальных тепловых пунктов зданий. В связи с невысоким порядком и линейностью математической модели она может быть основой для реализации метода управления с прогнозирующей моделью, который, в свою очередь, может быть использован в канале управления но возмущающему воздействию.

Литература. **1.** Табунщиков А.Ю. Математическое моделирование и оптимизация тепловой эффективности зданий / Ю.А. Табунщиков, М.М. Бродач. – Москва: АВОК – ПРЕСС, 2002. – 194 с. **2.** Куценко А.С. Системный подход к математическому моделированию тепловых процессов зданий / А.С. Куценко, С.В. Коваленко, В.И. Товажнянский // Восточно-Европейский журнал передовых технологий. – 2014. – №4/4 (70). – С. 9–12.

Лопатин А.К., Курач Я.С.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Устойчивые аттракторы в динамической модели ценообразования Вальраса с линейной функцией спроса и функцией предложения в виде логистического тренда Ферхюльста

Основополагающую роль в классической современной экономике играет концепция равновесия. Полная формулировка общей концепции равновесия может быть отнесена на счет Вальраса и других авторов [1, 2]. Дискретная модель Вальраса для одного товара может быть записана следующим образом:

$$x(t+1) = x(t) - w(S(x,\alpha) - D(x,\beta)). \tag{1}$$

где w – положительная константа, x – цена, α, β – векторы параметров, соответствующие экзогенным факторам, а $D(x,\beta)$ и $S(x,\alpha)$ представляют соответственно спрос и предложение.

В классической модели ценообразования «чистой» рыночной экономики (1) Вальраса делается два априорных допущения:

а) положение равновесия единственное;

б) выполняются условия асимптотической устойчивости положения равновесия.

В предлагаемой работе рассматривается модель Вальраса (1) с линейной функцией спроса и нелинейной функцией предложения в виде растущего логистического тренда Ферхюльста:

$$D(x(t)) = a - b(x(t) - m) + u, S(x(t)) = u + \frac{A_0}{1 + A_1 \exp(-\alpha(x(t) - m))}.$$

где a, b – параметры прямой, параметр α отражает скорость роста функции, параметр A_0 является верхней горизонтальной асимптотой, параметры u, m отвечают за смещение графика функции.

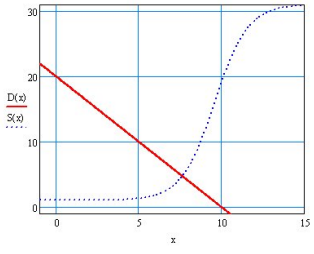


Рис. 1. Функции $D(x)$ и $S(x)$. Пересечение кривых дает координаты точки равновесия (статическое равновесие)

Показывается, что наряду с единственным устойчивым положением равновесия существуют другие устойчивые аттракторы в виде периодических функций с конечным периодом и устойчивым хаотическим аттрактором. Это явление впервые описано Chiarella [3] и Hommes, C.H. [4] с помощью паутинообразной модели с адаптивным ожиданием и нелинейным предложением в виде S-кривой в форме арктангенса. Уравнение (1) после подстановки значений функций спроса и предложения принимает вид:

$$x(t+1) = (1 - wb)x(t)) + wa - wm - w\frac{A_0}{1 + A_1 \exp -\alpha(x(t) - m)}. \tag{2}$$

Уравнение (2) исследуется методом имитационного моделирования с использованием программы $E\&F$ Chaos.

На первом этапе подбираются основные параметры (см. таблицу ниже)

A_0	A_1	α	a	b	m	u
30	5	1	3	2	8	1

Далее строится бифуркационная диаграмма (цена x по оси ординат) по параметру α (ось абсцисс). Варьируемым является параметр a.

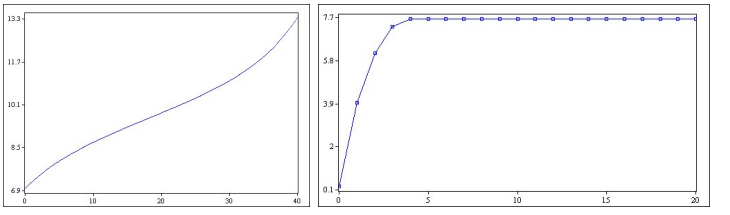


Рис. 2. Слева: при $\alpha = 1$ единственное устойчивое положение имеет место для всех значениях $0 < a < 40$. Справа: траектория цены изображена для $a = 3$

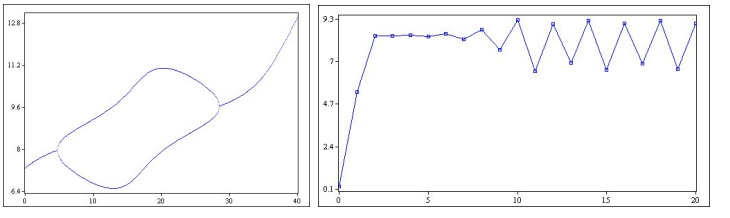


Рис. 3. Слева: при $\alpha = 2$ имеется два интервала устойчивого положения равновесия и один интервал устойчивого 2-периодического решения. Справа: показано 2-периодическое решение при $a = 10$

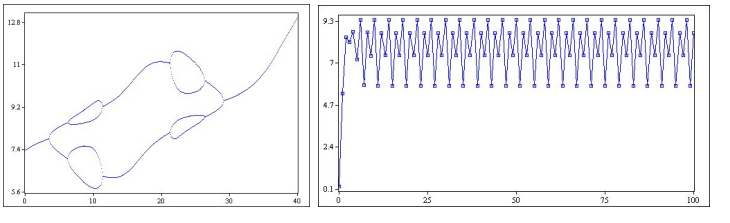


Рис. 4. Слева: при $\alpha = 2.4$ имеется два интервала устойчивого положения равновесия, три интервала устойчивого 2-периодического решения, два интервала, устойчивого 4-периодического решения. Справа: показано 4-периодическое решение при $a = 10$

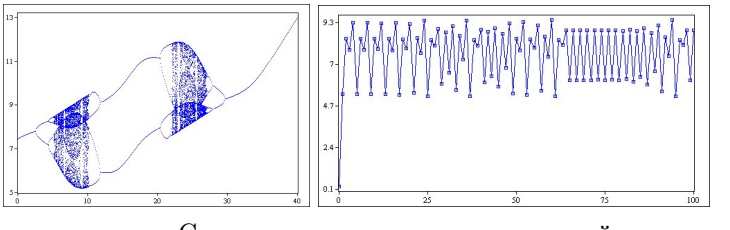


Рис. 5. При $\alpha = 3$ имеется хаос. Справа: показано хаотическое устойчивое решение при $a = 10$

Вывод: В модели Вальраса (1) устойчивые положения равновесия могут перейти в новые устойчивые k – периодические режимы через удвоение периодов колебаний ($k = 2, 4, 8, \dots$) и хаос при малом изменении параметра a.

Литература. **1.** Занг В.Б., Синергетическая экономика. Время и перемены в нелинейной экономической теории, Мир, Москва, 1999. **2.** Данилов Н.Н., Курс математической экономики, Высшая школа, Москва, 2006. **3.** Chiarella C., The cobweb model: Its instability and the onset of chaos. Economic Modeling, 1988, 5, 377–384. **4.** Hommes, C.H., Dynamics of the cobweb model with adaptive expectations and nonlinear supply and demand. Journal of Economic Behavior & Organization, 1994, 24, 315–335.

***Лопатин А.К.*[1] *Черненко О.Б.*[2]**
[1]*КПИ им. Игоря Сикорского, ИПСА, Киев, Украина;* [2]*Национальная академия управления, Киев, Украина*

Динамическая модель ценообразования с функцией предложения и с функцией спроса в виде растущего и убывающего логистических трендов Ферхюльста

Основополагающую роль в классической экономике играет концепция равновесия. Полная формулировка общей концепции равновесия может быть отнесена на счет Вальраса и других авторов [1, 2]. В классической модели ценообразования "чистой" рыночной экономики делается два априорных допущения:

а) положение равновесия единственное, то есть является глобальным;
б) выполняются условия асимптотической устойчивости положения равновесия.

Это значит, что при любом эндогенном возмущении существует точка равновесия и при любом отклонении от равновесного положения со временем цена снова примет равновесное значение.

В то же время дело обстоит таким образом, что модель "чистой" рыночной экономики основывается на ряде нереалистических предположений, характерных для так называемой "совершенной конкуренции". В настоящее время большинство рынков в существующих экономических системах не отвечают жестким условиям и ограничениям "чистой" рыночной экономики.

В отличие от классической модели, в настоящей работе изучаются устойчивые аттракторы в модели Вальраса, для которой положение равновесия является частным случаем.

В качестве функции спроса и предложения примем убывающий и возрастающие логистические тренды Ферхюльста соответственно.

$$D(x(t)) = u_2 + \frac{B_0}{1 + B_1 \exp(\alpha_2(x(t) - m_2))}, S(x(t)) = u_1 + \frac{A_0}{1 + A_1 \exp(-\alpha_1(x(t) - m_1))}. \quad (1)$$

Здесь положительные параметры α_2, α_1 отражают скорость убывания и возрастания функций. Параметры B_0, A_0 являются верхними горизонтальными асимптотами. Параметры u_2, m_2, u_1, m_1 отвечают за смещение графиков функций относительно начала координат. Изобразим функции $D(x)$ и $S(x)$ на Рис 1.

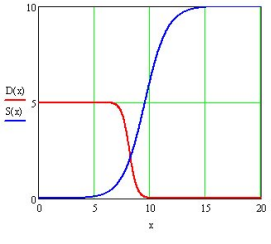


Рис. 1. Функции $D(x)$ и $S(x)$. Пересечение кривых дает координаты точки равновесия (статическое равновесие)

Уравнение Вальраса для функций спроса (1) и предложения (2) принимает вид:

$$x(t+1) = x(t) + w(u_2 - u_1) + w\frac{B_0}{1 + B_1 \exp + \alpha_{21}(x(t) - m_{21})} - w\frac{A_0}{1 + A_1 \exp - \alpha_1(x(t) - m_1)}. \quad (2)$$

Уравнение (2) исследуется методом имитационного моделирования с использованием программы E&F Chaos. На первом этапе подбираются основные параметры, которые представлены в таблице.

$x(0)$	w	m_1	m_2	u_1	u_2	A_0	A_1	B_0	B_1	α_1	α_2
0.25	0.2	8	8	1	0	10	5	4	0.5	0.5	3

Далее строится бифуркационная диаграмма (цена x по оси ординат) по параметру α_2 (ось абсцисс). B_0 является варьируемым параметром.

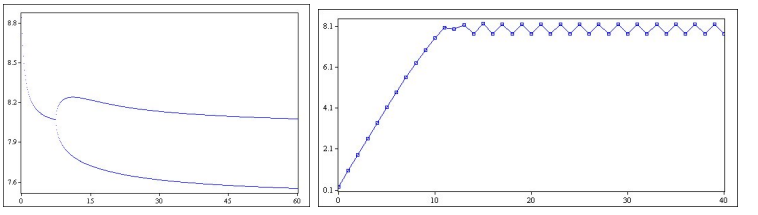


Рис. 2. Слева: $B_0 = 5$. При α_2 устанавливаются 2-периодические колебания. Справа: траектория цены изображена для $\alpha_2 = 15$

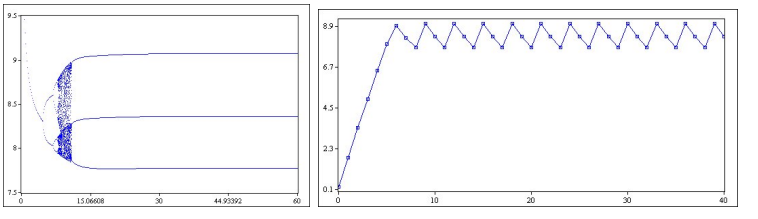


Рис. 3. Слева: $B_0 = 9$ при $\alpha_2 > 11$ устанавливаются 3-периодические колебания. Справа: траектория цены изображена для $\alpha_2 = 15$

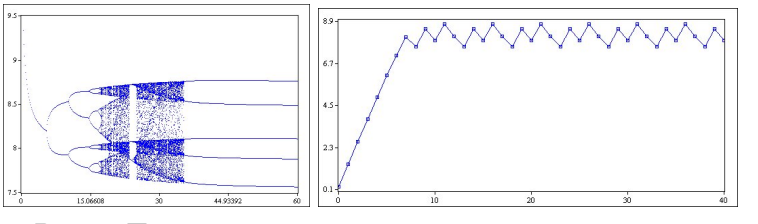


Рис. 4. Слева: $B_0 = 7$. При $\alpha_2 > 35$ устанавливаются 5-периодические колебания. Справа: траектория цены изображена для $\alpha_2 = 40$

Следующий рисунок на примере 3-периодичных колебаний демонстрирует, что построенные стационарные колебательные режимы устойчивы к внешним возмущениям.

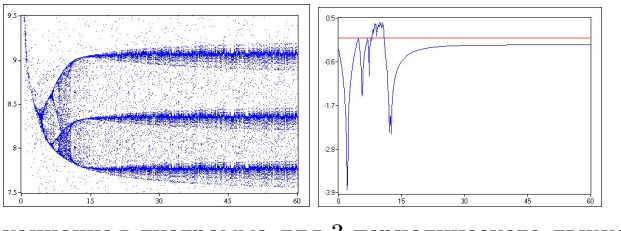


Рис. 5. Слева: бифуркационная диаграмма для 3-периодического движения, подверженная малым случайным возмущениям. Справа: Диаграмма старшего показателя 3-периодического невозмущенного движения

Вывод: в приведенной модели режим с неподвижной точкой сменяется устойчивым k-периодическим решением ($k = 2.3, 5$) путем прохождения через режим удвоения периода колебаний. Нам не известны аналогичные результаты в научной литературе для cob-web модели.

Литература. **1.** Занг В.Б. Синергетическая экономика. Время и перемены в нелинейной экономической теории. Мир, Москва, 1999. **2.** Данилов Н.Н. Курс математической экономики. Высшая школа, Москва, 2006.

Лубинець Ю.М., Федюков В.Д., Каніовська І.Ю.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Байєсівське усереднення для оцінки параметрів линійної регресії

Баєсівська статистика має давнє походження, однак активне її використання почалось порівняно недавно. Головними перевагами такого методу є можливість поєднувати попередні та поточні дані, коли нові дані стають доступними, то попередній апостеріорний розподіл стає апріорним. Також забезпечуються висновки, які залежать від кількості даних та є точними без асимптотичної апроксимації. Баєсівський метод підпорядковується принципу правдоподібності: якщо дві різні вибірки згенерували однакову функцію правдоподібності для деякого тета, то всі висновки про тета повинні бути однаковими для обох конструкцій. Проте наявні і мінуси байєсівських моделей. Наприклад, немає точної відповіді, де взяти апріорну ймовірність. Це потребує обережної математичної формалізації суб'єктивних переконань. Байєсівське усереднення звертається до моделі невизначеності в розв'язанні канонічної проблеми регресії.

Нехай дана лінійна модель, де y – залежна змінна, α_γ – константа, β_γ – коефіцієнти регресії та ε – випадкова похибка з нормальним розподілом, тобто $\varepsilon \sim N(0; \sigma^2)$:

$$y = \alpha_\gamma + \ X_\gamma \beta_\gamma + \varepsilon. \tag{1}$$

Проблема виникає, коли в матриці X є багато потенційних пояснюючих змінних: які з $X_j \in X$ потрібно включити в модель? Зробити висновок, чи потрібно включати всі змінні в лінійну модель, неефективно та майже неможливо за обмеженого числа спостережень. Байєсівське усереднення розв'язує цю проблему шляхом оцінки моделі для всіх можливих комбінацій із X та побудови середньозваженого за всіма з них. Наприклад, якщо X містить K потенційних змінних, то необхідно оцінити 2^K можливих комбінацій і 2^K моделей. Модель оцінюється за її середнім показником апостеріорної модельної ймовірності, яка випливає із теореми Байєса:

$$p(M_j|y, X) = \frac{p(y \vee M_j, X)p(M_j)}{p(y \vee X)} = \frac{p(y \vee M_j, X)p(M_j)}{\sum_{k=1}^{2^K} p(y \vee M_k, X)p(M_k),}. \tag{2}$$

$p(y|X)$ – узагальнена ймовірність, яка є константою для всіх моделей, тобто є мультиплікативним поняттям. Апостеріорна модельна ймовірність $p(y \vee M_j, X)$ пропорційна до відособленої правдоподібності моделі $p(M_j \vee y, X)$ (ймовірність даних в моделі M_j) та апріорної модельної ймовірності $p(M_j)$ – тобто, наскільки ймовірно дослідник розуміє модель M_j до вивчення необхідних даних. Ренормування приводить до апостеріорної модельної ймовірності, а тому розподіл зваженої апостеріорної моделі для будь-якої статистики Θ (наприклад, коефіцієнта β):

$$p(\Theta|y, X) = \sum_{\gamma=1}^{2^K} p(\Theta \vee M_\gamma, y, X)p(M_\gamma \vee X, y). \tag{3}$$

Для демонстрації ми застосували даний метод для визначення найбільш оптимальної моделі регресії, яка описує рівень злочинності по різних районах країни в залежності від характеристик цих районів. В силу того, що незалежних змінних дуже багато, робити регресійну модель на них всіх було б нераціонально, тому доречніше визначити, яка модель була б найкращою. Порівнюється вибір моделі за допомогою байєсівського усереднення та марківських ланцюгів методом Монте-Карло. З'ясовано, що метод байєсівського усереднення показує такі ж результати, як і ітераційний метод, проте працює швидше. Також за допомогою цього метода можна відновлювати щільність розподілу залежних змінних, тобто прогнозувати їх, на основі інших спостережень.

Література. **1.** Albert, J. (2009). Bayesian computation with R. London, UK: Springer. **2.** Zeugner, S. (2011). Bayesian Model Averaging with BMS for BMS Version 0.3.0. **3.** Yuan, Y., & MacKinnon, D.P. (2009). Bayesian mediation analysis. Psychological Methods, 14, 301–322. doi:10.1037/a0016972.

***Минько О.В.*[1] *Джигирей І.М.*[1,2]**
[1]КПІ ім. Ігоря Сікорського, Київ, Україна; [2]Світовий центр даних з геоінформатики та сталого розвитку КПІ ім. Ігоря Сікорського, Київ, Україна

Аналіз програмних пакетів підтримки оцінювання життєвого циклу продукційних систем

У сучасних умовах все актуальнішою постає потреба в підтримці прийняття рішень щодо заходів екологізування виробництв, обґрунтованих як з точки зору екологічної, так і промислової ефективності. Така задача потребує формалізованого методу оцінки впливу промислових та господарських об'єктів на навколишнє середовище.

Серед таких методів найбільше розповсюдження отримав метод оцінки життєвого циклу (ОЖЦ), в основу якого покладено врахування можливих впливів промислово-продукційної системи на довкілля протягом життєвого циклу, від видобутку ресурсів до утилізації. Методологія збирання даних, їх аналізу та інтерпретації регламентується серією стандартів ISO 1404X. Відповідно до ДСТУ ISO 14040:2013 та ДСТУ ISO 14044:2013 ОЖЦ складається з чотирьох фаз: визначення цілі та сфери застосування, інвентаризаційного аналізу, оцінки впливу та інтерпретації. З прийняттям цих стандартів та затвердженням формалізованої процедури визначення екологічних впливів з'явилась можливість автоматизувати процес оцінювання, що і стало поштовхом до розвитку відповідних програмних продуктів, що дають змогу зменшити трудомісткість та вартість проведення таких досліджень. Розглянемо найбільш поширені.

SimaPro. Цей програмний пакет здатен визначати обсяги викидів двоокису карбону та багато інших показників життєвого циклу різноманітних продукційний систем. Можна створювати власні продукційні системи з базових компонентів, що значно поширює сферу застосувань додатку. В комплекті поставляються бази даних US LCI, ILCD, Ecoinvent та LCAfood [1]. До недоліків слід віднести високі вимоги до кваліфікації користувача та високу вартість.

GaBi. Програмний пакет, що охоплює всі необхідні функції для проведення екологічного аудиту, включаючи модуль створення звітів [2]. Важливими перевагами додатку є функція запису набору даних, що дає змогу змінювати вхідні дані протягом процесу оцінки чи розробки продукту, та можливість наочно створювати діаграми потоків у виробничому процесі. До недоліків слід віднести фрагментованість пакету, необхідність вручну оновлювати бази даних та відсутність прозорості в розрахунках екологічних показників.

Umberto. Програмний засіб, що підтримує сумісні з ISO бази даних ОЖЦ та екологічних слідів [3]. Основною відмінністю додатку є гнучкість: користувач може створювати як прості лінійні моделі процесів, так і складні моделі з багатьма параметрами та нелінійні моделі. До переваг додатка можна віднести те, що це був один з перших програмних пакетів для ОЖЦ, а отже, наявна велика база знань та документація. Також програма має сучасний зручний користувацький інтерфейс, що спрощує використання. Найбільшим недоліком є ціна – 15 тис. дол. США за ліцензію без обмежень у часі.

Використання інструментів екологічного керування, зокрема ОЖЦ, з метою ухвалювання рішень для екологоорієнтованого трансформування вітчизняних підприємств може стати одним з рушіїв змін, які уможливлять вирішення проблем у сферах ресурсо- і енергозбереження, керування відходами, зменшення шкідливості виробництв, збереження природного довкілля. На ринку наявний широкий вибір програмних продуктів для автоматизованого проведення ОЖЦ та аналізу результатів. Проте більшість таких продуктів створено з огляду на умови в країнах-розробниках, а отже, щонайменше, близькість до дійсності отримуваних результатів залежать від локалізації використовуваних баз даних. Враховуючи вищезазначене, актуальним є розроблення вітчизняного програмного продукту, що відповідатиме еколого-економічному стану України, національним екологічній політиці та стратегії розвитку.

Література. **1.** SimaPro, simapro.com. **2.** Life Cycle Assessment LCA Software – GaBi Software, www.gabi-software.com/international/index. **3.** Software for Life Cycle Assessment, Material and Energy Flow Analysis (MFA) and Carbon Footprint Calculations, www.umberto.de/en.

Морозов К.В., Романкевич В.А., Сапсай Т.Г.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

О модификации графо-логической модели для систем и их компонентов с множественными состояниями

Рассчет надежности отказоустойчивых многопроцессорных систем может производиться при помощи графо-логических (или GL-моделей). Такая модель представляет собой неориентированный граф, каждому ребру которого соотвествует некоторая реберная функция, зависящая от вектора состояния системы, каждый из элементов которого принимает значение 1 или 0, в зависимости от того, исправен или нет соответствующий компонент системы. Если функция принимает значение, равное нулю, то соответствующее ей ребро исключается из графа. Потеря графом связности соответствует выходу системы из строя. Расчет показателей надежности системы производится путем проведения статистических экспериментов с моделью.

Такие модели позволяют выполнять расчеты для систем, которые, как и их компоненты могут находиться только в одном из двух состояний: полностью работоспособном или неработоспособном. В реальности для некоторых систем можно выделить больше двух состояний: добавляются состояния частичной работоспособности. Также более двух состояний может быть и у компонентов системы (в частности, это может быть актуально для верхних уровней иерархических систем). Причем количество состояний системы и количество состояний у каждого из ее компонентов в общем случае могут отличаться. Для таких систем может потребоваться, например, находить вероятность их пребывания в том или ином состоянии.

В докладе предлагается обобщение графо-логической модели, позволяющее выполнять такой расчет для систем, имеющих более чем два состояния. В таком случае, каждый из элементов вектора состояния системы будет представлен целым числом из множества $\{0, 1, ..., s_i - 1\}$, где (s_i – количество состояний соответствующего компонента системы). Реберные функции будут представлены функциями многозначной логики, которые принимают целые значения из множества $\{0, 1, ..., s - 1\}$, где s – количество состояний системы. Отметим, что значение 0 будет соответствовать полностью неработоспособному состоянию системы, а $s - 1$ – полностью работоспособному. Каждое ребро графа, таким образом, будет иметь некоторый целый вес. При этом, ситуации, когда система находится в состоянии с номером z, будет соответствовать наличие в графе остова, состоящего из ребер с весами, не меньшими, чем z. Легко заметить также, что в случае двух состояний предложенная модель вырождается в “классическую” GL-модель.

Для определения состояния системы достаточно в графе модели выбрать такой остов, ребро с наименьшим весом которого будет иметь наибольший возможный вес. Этот вес и будет соответствовать состоянию системы. Можно также показать, что искомый остов, в свою очередь, будет соответствовать остову максимального веса (стоимости), который можно найти одним из известных методов, в частности, при помощи алгоритмов Прима и Крускала. Кроме того, эти алгоритмы могут быть легко модифицированы так, чтобы сразу определять состояние системы.

В частности, модифицированный алгоритм Прима будет иметь следующий вид (изначально множество вершин U – пусто, а искомое состояние $r = s - 1$):

1. В множество вершин U добавляем произвольную вершину графа.
2. Среди всех ребер графа, которые инцидентны одной вершине из множества U и одной – не входящей в U, выбираем такое, вес которого максимален.
3. Если вес этого ребра w_i меньше текущего значения r, то присваиваем $r = w_i$.
4. Вершину, инцидентную этому ребру и не входящую в U добавляем в данное множество.
5. Если множество U еще не содержит все вершины графа, то переходим к п. 2.
6. Конец.

После выполнения алгоритма значение r будет соответствовать номеру состояния системы.

Недашківська Н.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Порівняльний аналіз методів розрахунку пріоритетів альтернатив рішень на основі нечітких експертних оцінок парних порівнянь

В [1] запропоновано технологію моделювання роботи експерта високої компетентності при оцінюванні альтернатив рішень методами парних порівнянь в шкалі Сааті. Використовуючи цю технологію, отримано оцінки рівня неузгодженості експертних парних порівнянь, який вноситься шкалою Сааті, а також отримано оцінки помилок пріоритетів (ваг), обчислених методами типу «трикутник» і «лінія» парних порівнянь. Показано [1], що коли експерт виконує оцінювання за методом парних порівнянь в шкалі Сааті, відносно малий рівень неузгодженості (CR=0.01) його оцінок є не лише прийнятним, але й бажаним, оскільки вимога повної узгодженості (CR=0) може внести додаткову помилку при побудові матриці парних порівнянь і, як наслідок, в результуючі ваги.

Проведено порівняльний аналіз сучасних моделей GPM, LUAM, TLGP, 2SLGP, та сімейства адитивних моделей, які обчислюють інтервальні пріоритети чи ваги альтернатив рішень на основі інтервальних матриць парних порівнянь (ІМПП). Моделі GPM, LUAM, TLGP та адитивні застосовуються як для узгоджених, так і для неузгоджених ІМПП. Моделі GPM і LUAM розраховують нормовані інтервальні ваги. Інтервальні ваги за TLGP потребують подальшого використання спеціальних методів нормування з метою їх подальшого використання для розв'язання багатокритеріальної задачі. Моделі GPM і TLGP, на відміну від LUAM, дозволяють в деякій мірі оцінити неузгодженість ІМПС (початкових експертних оцінок). Оцінками неузгодженості в цих моделях можуть служити оптимальні значення цільових функцій [2]. Проте, визначається тільки наявність неузгодженості в ІМПС. Моделі GPM і TLGP не дозволяють оцінити допустимість цієї неузгодженості для її використання в процесі прийняття рішень. Перевагою моделі GPM є можливість визначити найбільш неузгоджений елемент ІМПП [2]. Також модель GPM досить легко розширити на випадок нечітких матриць парних порівнянь, елементи яких задані трикутними або трапецієвидними нечіткими множинами.

Виявлено особливості роботи моделей GPM і LUAM на ІМПП з різною мірою неузгодженості [2,3]. Результати аналізу показують, що для слабо неузгодженої ІМПП з циклом не існує ранжирування альтернатив рішень, яке задовольняло б усім елементам цієї матриці. Ваги, обчислені методами GPM і LUAM на основі слабо неузгодженої ІМПП, в загальному випадку призводять до різних ранжирувань альтернатив рішень. Встановлено, що властивість слабкого збереження порядку в основному дозволяє виявити порядково нетранзитивні елементи слабо неузгоджених ІМПП. Після коригування цих елементів виключаються протиріччя в результуючих вагах і підвищується достовірність отриманих результатів.

В [4,5] запропоновано новий метод визначення довірчих інтервалів для ваг альтернатив рішень, що враховує невизначеність, присутню в матриці парних порівнянь і обгрунтовану шкалою Сааті, яка використовується експертом при виконанні оцінювання, а також такими особистими якостями експерта, як песимізм і оптимізм. Запропоновано технологію моделювання процесу оцінювання методом парних порівнянь експертами реалістом, песимістом та оптимістом. Використовуючи результати моделювання, проведено порівняльний аналіз та показано переваги запропонованого в [4,5] методу над іншими відомими методами на основі ІМПП та НМПП.

Література. **1.** Недашковская Н.И. Сравнительный анализ методов парного экспертного оценивания альтернатив решений // Системні дослідження та інформаційні технології. – 2014. – №4. – С. 35–44. **2.** Недашковская Н.И. Модели парных сравнений на основании интервальных оценок экспертов // Питання прикладної математики і математичного моделювання. Збірник наукових праць. – 2015. – Випуск 15. – С. 121–137. **3.** N. Pankratova, N. Nedashkovskaya. Estimation of decision alternatives on the basis of interval pairwise comparison matrices // ICA – 2016.- Vol.7, No.2. – P. 39–54. **4.** Недашковская Н.И. Построение доверительных интервалов для весов альтернатив решений на основе экспертных парных сравнений // Системні дослідження та інформаційні технології. – 2015. – №3. – С. 121–130. **5.** Недашковская Н.И. Метод оценивания неопределенности экспертных оценок парных сравнений при вычислении весов альтернатив решений // Проблемы управления и информатики. – 2015. – №5. – С. 130–142.

Новицький Є.А., Шубенкова І.А.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Оцінка ризиків проектів за умови відсутності повної інформації

Оцінка рівня ризиків є одним з найважливіших етапів економічного аналізу, оскільки для управління ризиком його необхідно перш за все визначити, проаналізувати і оцінити.

Аналіз ризику – це застосування системи спеціальних знань з дослідження економічних явищ і процесів за умов невизначеності та конфліктності з метою отримання якісної та кількісної інформації, необхідної для прийняття управлінських рішень.

Метою проведення аналізу ризиків є отримання необхідної інформації для прийняття управлінських рішень щодо можливості передбачення та відповідного захисту компанії від наслідків ризикових подій. Зазвичай аналіз ризику проводять у такій послідовності:

1. визначення внутрішніх та зовнішніх чинників, що збільшують чи зменшують ступінь певного виду ризику;
2. аналіз виявлених чинників;
3. оцінювання певного виду ризику;
4. встановлення допустимого ступеня ризику;
5. аналіз окремих операцій щодо обраного ступеня ризику;
6. розробка заходів щодо зниження ступеня ризику.

Для аналізу ризику використовують кількісні та якісні методи. **Якісний аналіз ризику** передбачає виявлення джерел та причин ризику процесів і робіт, ідентифікацію зон та видів ризику, виявлення практичної користі та можливих негативних наслідків, які можуть виникнути у процесі реалізації проектів (робіт, процесів), що містять ризик. Більшість вчених, які займаються проблемою оцінки ризику, констатують, що якісний аналіз є найскладнішим етапом загального аналізу ступеня ризику. **Якісна оцінка ризиків** – це процес ідентифікації та визначення ризиків, що вимагають швидкого реагування.

Найпростіша схема якісного аналізу ризику передбачає:

- ідентифікацію окремих видів ризику;
- вибір техніки аналізу ризику;
- встановлення факторів ризику та їх значимості;
- створення моделі механізму дії ризиків;
- встановлення взаємозв'язку окремих ризиків та сукупного ефекту від їх дії;
- формування загального портфеля господарських ризиків.

У результаті якісного аналізу необхідно отримати чітке уявлення про всі можливі ризики, виявлені потенційні зони ризику, а також негативні наслідки або додаткові вигоди, що можуть виникнути при реалізації певного рішення.

Метою кількісного аналізу є отримання числового вираження окремих ризиків із визначенням характеристики ймовірності і можливих втрат. Кількісне значення рівня ризику часто визначають як деяку функцію від добутку показників наслідків ризикової ситуації та ймовірності її виникнення. Для цього формується набір сценаріїв і можуть бути побудовані для окремих видів ризику функції розподілу ймовірності настання втрат залежно від їх розмірів.

Кількісні методи передбачають оцінку ризику в абсолютному і відносному вираженні. В абсолютному вираженні ризик вимірюють частотою чи розмірами можливих збитків у грошовому еквіваленті. У відносному вираженні ризик вимірюють різними безрозмірними показниками, що є відношенням двох чи більше показників.

Література. **1.** Шоломицкий А.Г. Теория риска. Выбор при неопределенности и моделирование риска: [учеб. пособие для вузов] / А.Г. Шоломицкий; Гос. ун-т – Высшая школа экономики. – М.: Изд. дом ГУ ВШЭ, 2005. – 400 с. **2.** Лук'янова В.В. Економічний ризик: [навч. посіб.] / В.В. Лук'янова, Т.В. Головач. – К.: Академ-видав, 2007. – 464 с. **3.** Пісаревский І.М. Управління ризиками: [навч. посібник] / І.М. Пісаревський, О.Д. Стешенко. – Харків: ХНАМГ, 2008. – 124 с.

Опарина Е.Л.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Поиск рационального компромисса требований при создании сложных технических объектов

Системные согласования различных требований и показателей прочности и надежности при создании сложных технических объектов (СТО) выполняются на основе определения компромиссов противоречивых целей при формировании множества Парето. Согласования различных требований прочности и надежности СТО базируется на регулировании критериев, представляющих собой взаимосвязанную систему оценивания вероятностей отказов СТО в процессе эксплуатации. При выработке таких критериев целесообразно обеспечить на основе условий мониторинговых информаций соответствующие свойства и цели для выполнения решений обратных задач теории прочности и надежности. Учитывая взаимосвязанные системы оценивания вероятностных отказов СТО, целесообразно для известных внешних воздействий определить границы изменения входных параметров: допусков на геометрические параметры; на свойства физико-механических характеристик материалов и конструкций; параметры материалоемкости и грузоподъемности и ряд аналогичных свойств. Параметры и свойства обеспечивают заданную вероятность отказа или заданный коэффициент запаса [1], что соответствует минимуму или максимуму целевой функции.

При этом в методах формирования оптимальных по Парето решений многокритериальных задач создания СТО априорно полагается, что множество Парето существует [2]. Вместе с тем, в реальных системных задачах формирования множества компромиссов противоречивых целей априорно нельзя исключить ситуацию, когда исходные условия эксплуатации и заданные требования к надежности СТО являются принципиально несовместимыми в рамках рассматриваемого проекта. При использовании принципа Парето в процедурах решения многокритериальных задач создания СТО требуется, прежде всего, установить необходимость и доказать возможность формирования непустого множества компромиссов целей при заданных требованиях, ограничениях и условиях. Следует обратить внимание, что эти задачи принципиально отличаются от типовых задач системной оптимизации тем, что целевые функции, области определения и множества значений априорно не задаются. Они должны быть определены на основе неполной, зачастую разнородной исходной информации о создаваемом СТО и условиях его эксплуатации.

Сущность метода нахождения рационального компромисса и возможность его применения для решения многокритериальных задач создания СТО, системного согласования противоречивых целей в условиях неопределенности позволяет проанализировать и рассмотреть примеры различных задач для согласования требований к внешним и внутренним показателям исследуемых объектов при априорно известных ограничениях на основные показатели внешнего воздействия [3, 4]. Исследуются основные системные свойства и факторы при согласовании различных требований и показателей технического состояния объектов в условиях противоречивых исходных данных при рациональной реализации заданных ограничений. Оцениваются возможности и результаты рационального компромисса при разработке и применении СТО в условиях неопределенности. Анализируются для конкретного СТО требования к свойствам и показателям возможностей реализации компромиссов противоречивых целей на основе формирования множества Парето.

Литература. **1.** ДСТУ3900-99 Государственный стандарт Украины. Безопасность в чрезвычайных ситуациях. Основные положения. К.: Госстандарт Украины, 1999. – 5 с. **2.** Подиновский В.В. Ногин В.Д. Парето-оптимальные решения многокритериальных задач. – М.: Наука, 1982. – 256 с. **3.** Опарина Е.Л. Рациональный компромисс в задачах системного согласования противоречивых целей // Системні дослідження та інформаційні технології. – 2004. – №4. – С. 90–101. **4.** Опарина Е.Л. Математические основы решения многокритериальных задач синтеза иерархических структур сложных систем // Матеріали X Міжнародної науково-технічної конференції «Системний аналіз та інформ. технології», ІПСА НТУУ «КПІ», 20–24 травня, м. Київ. – 2008, с. 112.

Панкратов В.А.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Стратегия инновационного развития социально-экономических систем

Моделирование поиска новых путей развития крупных социально-экономических систем на уровне компании, мегаполиса, крупного предприятия, региона является достаточно актуальной проблемой. Технологическая или организационная структура таких систем является иерархической, многоуровневой, взаимосвязанной системой такого же типа систем или различных типов функциональных элементов, которые могли бы быть сфокусированы в некотором пространстве или разнесены на значительные расстояния.

При создании формализации математической модели сложной системы необходимо учитывать сложность, неформализуемость, неопределенность, разнообразие концептуальных взаимосвязей организационных, технических, технологических, экономических и других задач на различных этапах жизненного цикла сложных систем и самой высокой цены возможной ошибки. Отсюда следует практическая необходимость поиска новых концепций, принципов, моделей и подходов к созданию стратегии инновационного развития крупных социально-экономических систем. Наиболее целесообразным является решение проблемы на основе синтеза различных методологий с современными информационными технологиями.

Для создания стратегии развития крупных социально-экономических систем предлагается синтез методологии предвидения [1, 2] с методологией когнитивного моделирования, которая открывает уникальную возможность в рамках единого программно-аналитического комплекса решения задач стратегического планирования и быстрого реагирования.

В соответствии с методологией предвидения на последнем этапе процесса предвидения лицам, принимающим решения, предлагаются 3–4 сценария, которые, вообще говоря, являются сложными слабоструктурированными системами. В настоящее время в рамках методологии предвидения выполнена формализация ряда методов качественного анализа (SWOT-анализ, метод анализа иерархий и его модификации, методы Delphi, перекрестного анализа, морфологический анализ и другие), которые стали основой набора инструментов создания альтернатив сценариев. Полученные на этапе предвидения характеристики являются исходными данными для начальной итерации когнитивного моделирования.

Для обоснованного осуществления одного или другого сценария целесообразно привлекать когнитивное моделирование, которое позволяет на основе знаний и опыта построить причинно-следственные связи, проанализировать поведение сложной системы и предложить научно обоснованную стратегию для реализации приоритетного сценария [3].

Показана способность разработанного математического и программного обеспечения реализации стратегии инновационного развития социально-экономических систем, базирующейся на объединении методологий предвидения и когнитивного моделирования, к решению проблем по выявленным приоритетным отраслям и исследованы ряд задач по социально-экономическим системам в пределах приоритетных отраслей.

Литература. **1.** Zgurovsky M., Pankratov V.A. Strategy of innovative development of region on the basis of foresight methodologies and cognitive modeling synthesis // System research&Information Technologies, №2, 2014. P. 7–20. **2.** Pankratov V. The creation of strategy for innovation development of socio-economic systems // International Journal. “Information technologies&knowledge”. ITHEA. SOFIA, V.3, №1. – 2014. – P. 84–99. **3.** Pankratov V. To the estimation of cognitive maps sustainability // Intern. Conf “X Szkola Geomechaniki 2011”. Materialy Naukowe, Gliwice-Ustron, V.2, 18–21 pazdziernika 2011. – 2011. – P. 119–128.

Панкратова Н.Д., Демьянчук А.З., Мерзляков А.Д.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Применение полиномов Якоби при восстановлении функциональных зависимостей

При реализации различных прикладных задач наиболее часто в качестве базовой аппроксимирующей функции при восстановлении функциональных закономерностей используют полиномы Чебышева, Лаггера, Лежандра, Эрмита и др., применение которых не всегда позволяет получить одинаково требуемую точность приближения. Актуальным является нахождение такого общего вида полиномов, которые за счет вариации своих параметров смогли бы не допустить снижения точности приближения при аппроксимации различного вида функций.

Последнее обуславливает целесообразность рассмотреть более общую структуру базовых аппроксимирующих функций – многочлены Якоби, а также функции Гаусса. В данной работе проводится исследование выбора базовой аппроксимирующей функции $\phi_{j_s p}$, применяемой на третьем уровне в модели восстановления функциональных зависимостей в аддитивной [1] и мультипликативной [2] формах, в виде полинома Якоби. Далее проводится сравнительный анализ с наиболее часто используемыми полиномами: Чебышева, Лежандра, Лагерра и Эрмита, с целью подбора для многочленов Якоби оптимального значения параметра α при аппроксимации различных функций (тригонометрических, экспоненциальных).

Полиномы Якоби – это одни из классических ортогональных полиномов, которые имеют много важных приложений в таких областях как: математическая физика, численные методы, и многие другие. Он определен и неотрицателен на интервале $[-1;1]$.

Полиномы Якоби $P_n(z)$ получаются из гипергеометрических функций и в общем случае имеют вид:

$$P_n^{(\alpha,\beta)}(z) = \frac{\Gamma(\alpha+n+1)}{n!\Gamma(\alpha+\beta+n+1)} \sum_{m=0}^{n} \binom{n}{m} \frac{\Gamma(\alpha+\beta+n+m+1)}{\Gamma(\alpha+m+1)} \left(\frac{z-1}{2}\right)^m. \quad (1)$$

Первые три полинома Якоби имеют вид:

$$P_0^{(\alpha,\beta)}(x) = 1, \quad (2)$$

$$P_1^{(\alpha,\beta)}(x) = \frac{1}{2}(\alpha-\beta+x(\alpha+\beta+2)), \quad (3)$$

$$P_2^{(\alpha,\beta)}(x) = \frac{x^2}{8}(\alpha+\beta+3)(\alpha+\beta+4) + \frac{x}{4}(\alpha-\beta)(\alpha+\beta+3) + \frac{1}{8}((\alpha-\beta)^2-(\alpha+\beta)-4). \quad (4)$$

На границах интервала полиномы Якоби принимают конечные значения:

$$P_n^{(\alpha,\beta)}(-1) = (-1)^n \frac{1}{n!}(\beta+n)(\beta+n-1)\dots(\beta+1), \quad (5)$$

$$P_n^{(\alpha,\beta)}(+1) = \frac{1}{n!}(\alpha+n)(\alpha+n-1)\dots(\alpha+1). \quad (6)$$

Рекуррентная формула полиномов Якоби:

$$P_{n+1}(x) = (A_n x + B_n)P_n(x) - C_n P_{n-1}(x), \quad (7)$$

где

$$A_n = \frac{(2n+1+\alpha+\beta)(2n+2+\alpha+\beta)}{2(n+1)(n+1+\alpha+\beta)}, \tag{8}$$

$$B_n = \frac{(\alpha^2-\beta^2)(2n+1+\alpha+\beta)}{2(n+1)(2n+\alpha+\beta)(n+1+\alpha+\beta)}, \tag{9}$$

$$C_n = \frac{(n+\alpha)(n+\beta)(2n+2+\alpha+\beta)}{(n+1)(n+1+\alpha+\beta)(2n+\alpha+\beta)}. \tag{10}$$

Некоторые другие полиномы могут быть выведены с полиномов Якоби. Если в полиномах Якоби приравнять параметры α и β, то будут получены полиномы Гегенбауэра $C_n^{(\alpha)}(x)$:

$$C_n^{(\alpha)}(x) = \frac{\Gamma(2\alpha+n)\Gamma(\alpha+\frac{1}{2})}{\Gamma(2\alpha)\Gamma(\alpha+n+\frac{1}{2})} P_n^{(\alpha-\frac{1}{2},\alpha-\frac{1}{2})} \tag{11}$$

Частным случаем полиномов Гегенбауэра с параметром $\alpha = \frac{1}{2}$ являются полиномы Лежандра $P_n(x) = C_n^{(\frac{1}{2})}(x)$. Рекуррентная формула:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x) \tag{12}$$

Частным случаем нормированного полинома Гегенбауэра для параметра $\alpha \to 0$ также являются полиномы Чебышёва $T_n(x) = \lim_{\alpha\to 0} n\Gamma(\alpha)C_n^{(\alpha)}$. Рекуррентная формула:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \tag{13}$$

Также эффективность исследуемых полиномов будет сравнена с полиномами из других классов ортогональных полиномов, а именно:

- с полиномами Лаггера $L_n^{(\alpha)}(x)$, где параметр α вещественное число больше –1. Рекуррентная формула:

$$(n+1)L_{n+1}^{(\alpha)}(x) = (2n+1+\alpha-x)L_n^{(\alpha)}(x) - (n+\alpha)L_{n-1}^{(\alpha)}(x); \tag{14}$$

- с полиномами Эрмита $H_n(x)$, рекуррентная формула которых имеет вид:

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x). \tag{15}$$

В данном исследовании предлагается вычислительный алгоритм и программная реализация восстановления функциональных зависимостей в аддитивной [1] и мультипликативной [2] формах, где в качестве базовых полиномов используются полиномы Якоби.

Проводится вычислительный эксперимент, чтобы показать приоритетность использования полиномов разной сложности в зависимости от вида задачи.

В результате исследований даны рекомендации относительно целесообразности использования полиномов, полученных в процессе декомпозиции полиномов Якоби.

Литература. **1.** Панкратова Н.Д. Формирование целевых функций в системной задаче концептуальной неопределенности // Доп. НАН України. – 2000. – №.9. – С. 67–73. **2.** Панкратова Н.Д., Опарина Е.Л. Восстановление многофакторных закономерностей в условиях концептуальной неопределенности // Системні дослідження та інформаційні технології. №3. -2004. – С. 103–114. **3.** Восстановление зависимостей по эмпирическим данным. Вапник В.Н. Главная редакция физико-математической литературы издательства «Наука», М., 1979, – 448 с. **4.** Свешников А.Г., Боголюбов А.Н., Кравцов В.В. Лекции по математической физике: Учеб. пособие. – М.: Изд-во МГУ, 1993. – 352 с.

Пархоменко А.В., Путренко В.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Аналіз регіональних збройних конфліктів з використанням методів просторової кластеризації

Останніми роками в світі збільшується кількість регіональних збройних конфліктів. Тому постає питання аналізу причин їх виникнення та динаміки розповсюдження. З цією метою аналізуються дані про світові збройні конфлікти, що мали місце з 1989-01-01 по 2015-12-31, та розробляються інструменти візуалізації результатів аналізу. В якості вхідних даних було взято масив UCDP GED – набір геоприв'язаних даних про конфлікти від кафедри досліджень питань миру і конфліктів університету Упсала (Швеція) [1, 2]. Основною одиницею аналізу для цього набору є "подія", індивідуальний інцидент (явище) смертельного насильства, що відбулось в певний момент часу в певному місці. Набір даних містить 128 264 подій. Вміст набору даних можна розділити на 7 категорій: ідентифікатори подій; учасники; джерела інформації про подію; географічні дані; час; точність; летальні цифри. Основною метою є проведення просторово-часового аналізу методом "Гарячих точок" та відображення його результатів у вигляді інтерактивних 3D-візуалізацій з географічною прив'язкою.

Метод "Гарячих точок" розраховує статистичний показник для кожної події в наборі даних. Підсумкові p-значення (ймовірності) та z-оцінки (стандартні відхилення) говорять про те, в якій області простору кластеризуються події з високими або низькими значеннями [3]. Метод працює шляхом аналізу кожної події в контексті сусідніх за географією подій. Щоб бути статистично суттєвої гарячою точкою, подія повинна мати високе значення і бути оточена іншими подіями з також високими значеннями. Локальна сума для події та її сусідів порівнюється пропорційно із сумою всіх подій; коли локальна сума дуже відрізняється від очікуваної локальної суми, і коли ця відмінність є занадто великою, щоб бути результатом випадкового процесу, отримується статистично значуща z-оцінка. Статистика гарячих точок використовує формулу:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \overline{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}}, \quad (1)$$

де x_j – атрибутивне значення для подій j, $w_{i,j}$ – просторова вага між подіями i та j, n – загальна кількість подій, $\overline{X}$ – середнє арифметичне значень подій, S – дисперсія.

Статистична величина G_i^*, для кожної події в наборі даних, є z-оцінкою. Для статистично значущих позитивних z-оцінок, чим більша z-оцінка, тим інтенсивніша кластеризация високих значень (гаряча точка). Для статистично значущих негативних z-оцінок, чим менше z-оцінка, тим інтенсивніша кластеризация низьких значень (холодна точка).

На виході отримуємо новий набір даних про події з z-оцінкою, p-значенням і рівнем достовірності Gi_Bin для кожної події у вхідному масиві UCDP GED. Для візуалізації результатів аналізу та надання доступу до них через мережу Інтернет, на клієнтській частині використовуються засоби відкритого проекту WebGL Globe та бібліотеки візуалізації three.js. Це дозволяє завантажувати дані у форматі JSON та забезпечувати інтерактивну взаємодію користувача з ними. На стороні серверу використано СУБД MySQL для збереження вихідного набору даних, веб-сервер Apache для обробки асинхронних запитів клієнтської частини та власне програмне забезпечення на мові PHP, яке виконує процедуру аналізу за методом "Гарячих точок".

Література. **1.** Sundberg, Ralph, and Erik Melander, 2013, "Introducing the UCDP Georeferenced Event Dataset", Journal of Peace Research, vol.50, no.4, 523–532. **2.** Croicu, Mihai and Ralph Sundberg, 2016, "UCDP GED Codebook version 5.0", Department of Peace and Conflict Research, Uppsala University. **3.** Getis, A. and J.K. Ord. 1992. "The Analysis of Spatial Association by Use of Distance Statistics" in Geographical Analysis 24(3).

***Пашинська Н.М.*[1], *Путренко В.В.*[2]**
[1]*Київський національний університет ім. Тараса Шевченка, Київ, Україна;* [2]*КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна*

Просторово-часове моделювання забруднення повітря внаслідок аварій на хімічних об'єктах з використанням інструментарію ГІС

Значна концентрація промислових підприємств та високий рівень техногенного навантаження на територію України підвищують ризики забруднення повітря внаслідок аварій на хімічних підприємствах. Зростання ризиків таких небезпек зумовлюють створення автоматизованих систем моніторингу та прогнозування забруднення з використанням моделювання. Важливе значення має задача оперативного моделювання наслідків викидів забруднюючих речовин в результаті аварій на хімічно небезпечних підприємствах, що певною мірою розглядається у роботах [1, 2]. Однак складність процесу просторового розповсюдження забруднення потребує єдиних підходів до його моделювання в режимі реального часу.

Метою просторово-часового моделювання є надання просторової прогнозної інформації відносно можливих наслідків аварії на хімічно небезпечних об'єктах для проведення заходів щодо упередження виникнення аварії та підготовки сил та засобів щодо ліквідації її наслідків.

В Україні у практичній сфері використовується методика МНС. Для цієї методики розраховуються глибина поширення хмари, площа зони можливого хімічного забруднення (ПЗХЗ), площа та ширина прогнозованої зони хімічного забруднення, час випаровування та час підходу хмари небезпечних хімічних речовин до населеного пункту, прогноз площі забруднення населеного пункту, прогноз втрат населення за кількістю населення в ПЗХЗ, виконується присвоєння ступеня хімічної небезпеки для адміністративно-територіальних одиниць, що потрапили в ПЗХЗ [3].

У загальному вигляді задача прогнозу забруднення повітря математично може бути визначена як розв'язання за певних початкових і граничних умов диференційного рівняння:

$$\frac{\partial q}{\partial t} + \sum_{i=1}^{3} u_i \frac{\partial q}{\partial x_i} = \sum_{i=1}^{3} u_i \frac{\partial}{\partial x_i} K_i \frac{\partial q}{\partial x_i} - \alpha q, \qquad (1)$$

де t – час; x_i – координати; u_i, K_i– складові середньої швидкості переміщення забруднення і коефіцієнта обміну, що відносяться до напряму осі x_i (i = 1, 2, 3); α – коефіцієнт, що визначає зміну концентрації за рахунок перетворення забруднення.

Геоінформаційне моделювання поєднує в собі просторове моделювання з обробкою та аналізом просторових та атрибутивних даних. ГІС забезпечують надання необхідної інформації про топографічні особливості місцевості, векторних шарів областей, районів, населених пунктів з інформацією про кількість населення, рельєфу та лісових масивів території, для якої виконується прогнозування за рахунок інтеграції даних дистанційного зондування, цифрових моделей місцевості, векторних шарів об'єктів та статистичних даних на основі сучасних розподілених веб-орієнтованих платформ.

Таким чином, поєднання сучасних обчислювальних потужностей з математичними просторово-часовими моделями та геоінформаційними технологіями на базі веб-рішень дозволяє досягнути нової якості реагування на загрози хімічних аварій та попередження населення.

Література. **1.** Майстренко С.Я. Прототипна версія системи «Повітря» для розрахунку зон уражень внаслідок викидів небезпечних речовин в атмосферу на основі використання скринінгових моделей та web-технологій / С.Я. Майстренко, І.В. Ковалець, В.П. Беспалов, Т.О. Загреба, К.В. Хурцилава // Математичні машини і системи. – 2016. – № 4. – С. 33–41. **2.** Снитюк В.Є. Технологія постпрогнозування хімічного зараження навколишнього середовища з використанням нечітких даних / В.Є. Снитюк, О.М. Землянський // Вісник Хмельницького національного університету. – 2013. – №6. – С. 272–278. **3.** Методика прогнозування наслідків виливу (викиду) небезпечних хімічних речовин при аваріях на промислових об'єктах і транспорті. Зареєстровано в Міністерстві юстиції України 10 квітня 2001 р. за № 326/5517.

Писарчук О.О., Лагодний О.В.
Житомирський військовий інститут ім. С.П. Корольова, Житомир, Україна

Статистичний аналіз часових рядів експериментальних даних активності тематичного контенту в інтернет-середовищі

В умовах збройної агресії Російською Федерацію під час ведення гібридної війни проти України виникає необхідність прогнозування розвитку інформаційних загроз в інтернет-середовищі, які своїм деструктивним впливом на соціум призводять до невдоволення серед населення, акцій протестів проти влади, революції, терористичних актів, збройного конфлікту, тощо. Інформаційні загрози представляють собою трансформовані інформаційні повідомлення, які мають конкретну тематику і розміщуються на контентах інтернет-середовища в дискретні проміжки часу [1]. Динаміка збільшення появи повідомлень за обраною тематикою свідчить про проведення противником інформаційно-психологічних акцій проти населення нашої країни.

Метою роботи є вибір адекватної моделі прогнозування розвитку інформаційних загроз в інтернет-середовищі, яка обирається після статистичного аналізу вихідних даних. Статистичний аналіз являє собою первину обробку експериментальних даних і необхідний для:

1. Виявлення закону розподілу випадкової складової, яка обумовлена похибками вимірювання в інформаційних повідомленнях.
2. Розрахунку характеристик закону розподілу випадкової величини за параметрами: математичного сподівання m_y, дисперсії D_y, середнього квадратичного відхилення σ_y, достовірності апроксимації R^2.
3. Виявлення у вибірці вимірів систематичної складової (тренду), на яку накладається закон розподілу випадкової складової.

Для проведення статистичного аналізу в результаті моніторингу відкритих джерел інформації інтернет-середовища за певний період з дискретними моментами часу було отримано вибірку вимірів наростаючого об'єму значень параметра досліджуваного процесу. Отримана вибірка вимірів являє собою кількість появи інформаційного повідомлення (одиниці аналізу – «Економічне блокування території Криму») за контентом K_1 (інформаційні сайти інтерент-середовища) в тематиці «Політичні новини»

$$y = \{y_i\}_{i=1...m}, \tag{1}$$

де: y_i – виміри; m – об'єм вибірки для контенту K_1 складає 210 значень за період з 14.09.15 р. по 15.10.15 р. (тобто 30 діб (7 раз на добу)) у дискретні проміжки часу з інтервалом 2 год. $t_1 = 8{:}00$, $t_2 = 20{:}00$ (рис. 1).

Виявлення закону розподілу вибірки вимірів є першим кроком статистичного аналізу даних. Від того, до якого закону розподілу відноситься вибірка вимірів залежить тактика подальшого аналізу. Виявлення закону розподілу проводиться за допомогою побудови гістограми щільності розподілу вибірки і за виглядом її форми робиться припущення про закони розподілу, яким можуть підпадати її значення. Якщо гістограма має єдиний чітко виражений максимум і є приблизно симетричною, то припускають, що випадкова величина розподілена за нормальним законом. Також при висуненні гіпотези про закон розподілу випадкової величини враховуються апріорні дані про закони розподілу, що притаманні обраній області дослідження. Для отриманої вибірки вимірів наростаючого об'єму побудовано гістограму (рис. 2).

В ході проведення вимірювань ми отримуємо певну закономірність в якій одні значення зустрічаються частіше інші рідше. Все це говорить про те, що ми маємо справу з випадковою величиною, яка розподілена за деяким законом з певними характеристиками. Характеристики цього закону нам не відомі, але ми їх можемо оцінити на основі результатів вимірювання. *Розрахунок характеристик закону розподілу* випадкової величини проводився для знаходження параметрів невідомого законів розподілу математичного сподівання m_y, дисперсії D_y,

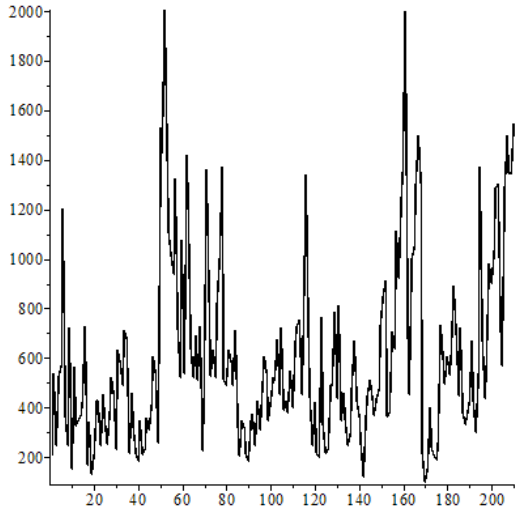


Рис. 1. Вибірка вимірів

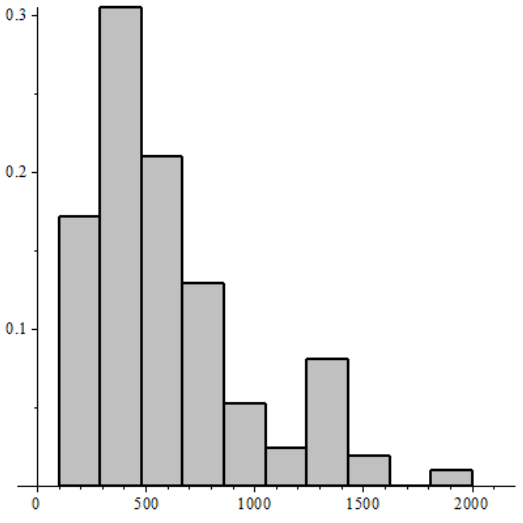


Рис. 2. Гістограма вибірки вимірів

середнього квадратичного відхилення σ_y, достовірності апроксимації R^2 за виразами (2)–(5):

$$m_y = \frac{1}{n}\sum_{i=1}^{n} y_i. \tag{2}$$

$$D_y = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - m_y)^2. \tag{3}$$

$$\sigma_y = \sqrt{D_y}. \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{5}$$

де $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, y_i – виміряне значення, $\widehat{y_i}$ – оцінка виміряного значення.

Виявлення систематичної складової (тренду) у вибірці вимірів реалізується на знаходженні невідомих параметрів нелінійної моделі з використанням методу найменших квадратів, суть якого полягає в згладжуванні вибірки вимірів. Критерієм оптимальності МНК є вимога мінімуму дисперсії згладжених параметрів рис. 3.

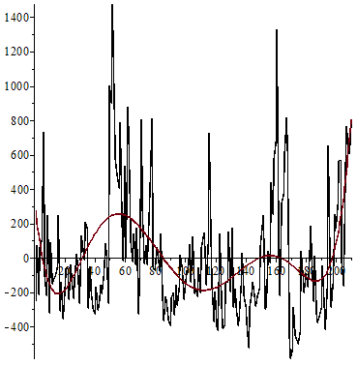


Рис. 3. Виявлення тренду в вибірці вимірів

Таким чином застосування статистичного аналізу для експериментальних даних активності тематичного контенту в інтернет-середовищі дає змогу виділити систематичну складову (тренд), який можна використовувати для прогнозування розвитку інформаційних загроз, що трансформовані в інформаційних повідомленнях.

Література. 1. Горбулін В.П. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія / В.П. Горбулін, О.Г. Додонов, Д.В. Ланде. – К.: Інтертехнологія, 2009. – 164 с.

Пичугина О.С.
Харьковский национальный университет радиоэлектроники, Харьков, Украина

Оптимизация на евклидовом комбинаторном множестве $E_{nk}^{\pm}(G)$

Рассмотрим следующую задачу дискретной оптимизации:

$$f(x) = x^T Ax + b^T x \to \min, \tag{1}$$

$$\sum_{i=1}^{n} x_i^{2j} = c_j,\ j \in J_n, \tag{2}$$

где $A = [a_{ij}]_{n\times n}$ – симметричная матрица, $b \in R^n$, $J_n = \{1, ..., n\}$,

$$c^j \in R^1,\ j \in J_n \text{ – положительная, монотонная последовательность.} \tag{3}$$

Проанализируем дополнительные ограничения. Как видно, система (2) содержит симметричные полиномы, а именно это степенные суммы n первых четных степеней. Тот факт, что число дополнительных ограничений совпадает с размерностью пространства, позволяет предположить, что в результате пересечения поверхностей, заданных каждым из уравнений системы (2), может образовываться дискретное множество. Поверхность $S_j = \left\{ x : \sum\limits_{i=1}^{n} x_i^{2j} = c_j \right\}$ представляет собой суперсферу радиуса $\sqrt[2j]{c_j}$ с коэффициентом деформации j ($j \in J_n$). Ясно, что если выбрать все n суперсфер одинакового радиуса ($\forall i \neq j\ \sqrt[2i]{c_i} = \sqrt[2j]{c_j} = c > 0$), то, за счет разницы в коэффициентах деформации, сфера S_1 будет вписана в суперсферу S_2, S_2 в S_3 и т.д. При этом в их пересечении будет образовываться дискретное множество вершин гипероетаэдра $CE_n(c) = \{(\pm c, 0, ..., 0), ..., (0, ..., 0, \pm c)\}$. Выбирая теперь радиус произвольной суперсферы значительно больше или существенно меньше c, мы приходим к ситуации, когда данная суперсфера не будет иметь общих точек с остальными, соответственно система ограничений будет (2) несовместна в пространстве R^n. Чтобы избежать описанную выше ситуацию мы добавили условие (3). Имеет место следующая теорема:

Теорема 1. Область (2) допустимых решений задачи (1)–(2) представляет собой дискретное множество следующего вида:

$$E_{nk}^{\pm} = \left\{ z \in R^n :\ \mathrm{z}=(z_i)_{i\in J_n},\ z_i = x_i y_i, i \in J_n,\ \ \mathrm{x}=(x_i)_{i\in J_n} \in E_{nk}(G),\ \mathrm{y}=(y_i)_{i\in J_n} \in B_n^{'} \right\}, \tag{4}$$

где $E_{nk}(G)$ – общее множество перестановок [1] из мультимножества $G = \{g_i\}_{i\in J_n} \geqslant 0$ неотрицательных решений системы $\sum\limits_{i=1}^{n} g_i^{2j} = c_j,\ j \in J_n$, среди которых k различны; $B_n = \{-1, 1\}^n$ – бинарное множество.

Итак, задача (1)–(3) представляет собой дискретную задачу на дискретном множестве (4), сформированном отражениями $E_{nk}(G)$ относительно всех координатных плоскостей. Заметим, что множество $E_{nk}(G)$ в большинстве случаев содержит неполиномиальное число элементов. С другой стороны, при выборе $c_j = n$, $j \in J_n$ оно превращается в бинарное множество, мощность которого 2^n. Таким образом, $E_{nk}^{\pm}(G)$ – это расширение двух известных комбинаторных множеств $E_{nk}(G)$ и $B_n^{'}$, а его мощность может достигать $|E_{nk}(G)| \cdot \left|B_n^{'}\right| = 2^n n!$. Итак, исходная задача (1)–(3) позволяет эквивалентную формулировку в виде квадратичной задач (1),

$$x \in E_{nk}^{\pm}(G). \tag{5}$$

Учитывая непрерывное функциональное представление (2) множества $E_{nk}^{\pm}(G)$, она может решаться как классическая задача на условный экстремум. Поскольку дополнительные ограничения в ней отсутствуют, различные эвристические схемы типа метода случайного поиска, генетические алгоритмы и т.п. применимы в этом случае. Мы же предлагаем подходы, позволяющие находить как точные решения, так и приближенные с оценкой точности. Эти

методы сочетают исследованные свойства $E_{nk}^{\pm}(G)$ и его выпуклой оболочки – многогранника $P_{nk}^{\pm}(G) = conv\ E_{nk}^{\pm}(G)$ – со свойствами функций на них.

1. Некоторые их полученных комбинаторных свойств $E_{nk}^{\pm}(G)$, $P_{nk}^{\pm}(G)$:
   a) Мощность $E_{nk}^{\pm}(G)$;
   b) Вершинная расположенность $E_{nk}^{\pm}(G)$, т.е. его совпадение с $vert\ P_{nk}^{\pm}(G)$;
   c) Критерий смежности вершин;
   d) Разложимость $E_{nk}^{\pm}(G)$ по семействам параллельных гиперплоскостей и декомпозиция на комбинаторные множества такого же вида меньшей размерности;
   e) Гиперграни $P_{nk}^{\pm}(G)$ представляют собой подобные многогранники либо общие перестановочные многогранники размерности на 1 ниже;
   f) Получена несводимая система $P_{nk}^{\pm}(G)$.
2. Экстремальные свойства $E_{nk}^{\pm}(G)$:
   a) Явное решение линейной задачи;
   b) Аналитическое решение задачи поиска проекции произвольной $x \in R^n$ на $E_{nk}^{\pm}(G)$.

Заметим, что $E_{nk}^{\pm}(G)$ вписано в сферу S_1, т.е. имеет место его полиэдрально-сферическое представление:

$$E_{nk}^{\pm}(G) = S_1 \cap P_{nk}^{\pm}(G), \tag{6}$$

Это позволяет адаптировать к задаче (1), (5) полиэдрально-поверхностный метод (ППМ) решения нелинейных комбинаторных задач [2]. Данная группа методов включает ППМ-схемы ветвей и границ [2], ППМ-жадные алгоритмы и т.п.

Основные требования к ним:

- Знание полиэдрально-поверхностного представления, т.е. линейной системы многогранника и описанной поверхности, задаваемой строго выпуклой функцией. В нашем случае это представление (6), которое нам известно согласно (2) и Свойству 1.f.
- Возможность сведения задачи $f(x)$ к оптимизации выпуклой дифференцируемой функции $F(x)$ (в силу Свойства 1.b). В данном случае такая функция может быть найдена в виде $F(x) = f(x) + \left(\|A\| - \sum\limits_{i:a_{ii}>0} a_{ii}\right)\left(x^2 - \sum\limits_{i=1}^{n} g_i^2\right)$.
- Эффективный способ решения полиэдральной релаксационной задачи (1), $x \in P_{nk}^{\pm}(G)$. Мы предлагаем использовать модификацию [2] метода условного градиента, в которой вспомогательные линейные задачи решаются на основе Свойства 2.a.
- Наличие эффективного пути решения сферической релаксационной задачи (1), $x \in S_1$. Такой способ для оптимизации квадратичной функции на сфере существует [3].
- Декомпозиция на подобные задачи меньшей размерности (см. Свойства 1.d, 1.e).
- Возможность постепенного уточнения верхней оценки в течение всего итерационного процесса. Это осуществляется путем проектирования решений релаксационных задач на $E_{nk}^{\pm}(G)$ (см. Свойство 2.b).

Выводы. Мы исследовали класс задач полиномиальной оптимизации, дополнительные ограничения которой, как оказывается, задают дискретное множество $E_{nk}^{\pm}(G)$. Этот класс NP-трудных задач, поскольку он включает безусловные квадратичные задачи. Мы исследуем комбинаторные и экстремальные свойства $E_{nk}^{\pm}(G)$ и соответствующего многогранника и предлагаем модификацию полиэдрально поверхностного метода к решению поставленной задачи.

Литература. **1.** Y.G. Stoyan, O.A. Yemets, Theory and methods of Euclidean combinatorial optimization, Kiev: ISSE, 1993. **2.** O. Pichugina, S. Yakovlev, "Convex extensions and continuous functional representations in optimization, with their applications," J. Coupled Syst. Multiscale Dyn, 4, pp. 129–152, 2016. **3.** J. Dahl, "Convex optimization in signal processing and communications," Ph.D. dissertation. Dept of Communication Technology, Denmark: Aalborg University, 2003.

***Піддубна Ю.І.*[1], *Пишнограєв І.О.*[2]**
[1]*КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна;* [2]*КПІ ім. Ігоря Сікорського, ФММ, Київ, Україна*

Оцінювання ефективності та інвестиційної привабливості стартапів

Стартапи, завдяки своїй мобільності в плані втілення нових ідей, не втрачають своєї популярності вже більше 10 років. Жорстка конкуренція, швидке насичення цільової аудиторії, складність «зачепити» та зацікавити потенційного користувача у сукупності з стрімким розвитком технологій створюють складні умови для розвитку інновації. Стартап асоціюється з високотехнологічним бізнесом, цілі створення якого є актуальними та корисними, здатними вирішити певну проблему й покращити життя, а також у швидкі терміни окупити інвестиції та отримати прибуток [1, 2].

Основними факторами, що визначають привабливість стартапу для інвестора, є [2]:

1. Доцільність проекту, його потенційних попит та цільова аудиторія;
2. Ризики проекту та їх мінімізація, а також можливість контролю за ходом проекту;
3. Передбачувана норма прибутку проекту;
4. Глибина аналізу прогнозів по фінансовим і маркетинговим параметрам проекту;
5. Стабільність зростання області бізнесу, на яку націлено проект;
6. Можливість виходу з проекту (мінімізація збитків в разі провалу).

Безумовно, глибина аналізу та точність відповідей на вищезгадані питання сильно залежить від стадії, на якій знаходиться проект.

Питання оцінювання ефективності проекту хвилює як автора проекту, так і потенційного інвестора. Вартість самої ідеї (якщо вона не оформлена як інтелектуальна власність у вигляді патенту, авторського свідоцтва, ліцензії і т.д.) приймається рівною нулю. Оцінці підлягає тільки творчий потенціал автора/команди і підприємницька ініціатива.

Існує багато методів оцінювання ефективності стартапу [3, 4]:

1. Витратний метод. Дозволяє оцінити реальні фінансові витрати на створення аналогічного стартапу виходячи з поточної ринкової вартості спеціалістів проекту, витрат на ліцензування, наявних активів, витрат на рекламу та ін. Перевагою методу є можливість оцінки ефективності витрат грошових засобів командою стартапу, а суттєвим недоліком – він не враховує вартість інтелектуальної власності, особистої ініціативи команди проекту та інших нематеріальних цінностей. Витратний метод є базовим для деяких інших методів оцінювання стартапів;
2. Метод «зверху вниз» – використовується на ранніх стадіях стартапу, коли інформація про проект ще дуже обмежена. Такий аналіз є узагальненим, оцінювання проводиться по одному показнику та не вимагає значних зусиль та часу. Недоліком є невисока точність;
3. Метод Беркуса. Аналіз потенціалу стартапу проводиться за допомогою деяких емпіричних коефіцієнтів до витратного методу. Наприклад: надбавка за привабливу ідею становить 20%–40%, за професіональний менеджмент проекту – від 20% до 60%. Суттєвим недоліком методу, як будь-якої емпіричної оцінки, є особистісна необ'єктивність і спроба підлаштуватися під поточний ринковий момент;
4. Модель Ave Maria (вільне скорочення від слів Acquisition, Value, Engagement, Monetization, Retention, Intellectual Property). Даний метод не дає оціночних характеристик проекту, але зручний для порівняння різних проектів при виборі пріоритетного напрямку інвестування;
5. Метод скорингу (бенчмаркингу). Аналіз проводиться шляхом порівняння стартапу з вже існуючими подібними проектами з використанням наступних коефіцієнтів, що впливають на вартість оцінки:
   - наявність сильної команди: 0–30%;
   - розмір ринку: 0–20%;
   - новизна продукту та технології: 0–15%;
   - конкурентне оточення: 0–10%;

- маркетинг, канали продажів, партнерство: 0–10%.

Суттєвим недоліком методу є суб'єктивне ранжування факторів, а також висока ймовірність неправильної середньої оцінки подібних проектів.

При оцінці витрат зазвичай оперують трьома основними параметрами – вартістю, перевагами і гнучкістю. Для кожного з них окремо визначається свій рівень ризику.

Оцінка переваг повинна проводитися з точки зору вартості проекту і стратегічних вкладень, що виходять за рамки інформаційних технологій. Гнучкість визначається з використанням методологій розрахунків ф'ючерсів і опціонів, наприклад моделей Блека–Шоулза, або оцінки справедливої ціни опціонів (Real Options Valuation). Для інвестицій в інформаційні технології аналіз ризиків повинен передбачати доступність і стійкість параметрів виробників, продуктів, архітектури, корпоративної культури, обсягу і тимчасових рамок реалізації проекту.

На базі витратного методу оцінки побудована методологія сукупної вартості володіння (TCO) – ефективний підхід до визначення найкращого співвідношення ціна/якість на основі розгляду таких ключових бізнес-процесів, як відновлення після збоїв, управління модернізацією та технічна підтримка. TCO підходить для підрахунку поточних вартісних параметрів, з її допомогою можна досить повно проаналізувати ефективність виконання якихось окремих функцій або набору функцій. Однак дана методологія абсолютно не враховує ризики і не дозволяє співвіднести виробничу технологію зі стратегічними цілями подальшого розвитку підприємства та рішенням завдання підвищення конкурентоспроможності. Оцінити інноваційний проект з іншого ракурсу дозволяє методологія TEI (сукупний економічний ефект), яка призначена для підтримки прийняття рішень, зниження ризиків і забезпечення «гнучкості», тобто очікуваних або потенційних переваг, які залишаються за рамками аналізу витрат.

Так як методологія TCO розглядає тільки витрати, але повністю ігнорує тривалість і ризики проекту, має сенс використовувати модель оцінки, яка враховує ці чинники. Одним з варіантів є поділ проекту на задачі за допомогою автоматизованих засобів. Таким чином, використовуючи поєднання методів TEI та TCO в сукупності з стандартом PMBOK [2] можна отримати досить достовірний результат оцінки стартапу.

Висновки. Таким чином, є велика кількість підходів до оцінювання ефективності стартапів. Розглянуті методи не є універсальними та мають свої недоліки й особливості. На жаль, існуючі методи оцінювання ефективності стартапів носять у більшості своїй описовий характер. Наявної в них інформації недостатньо для визначення коштів, які необхідно закласти у проектний бюджет для компенсування цін на проектні роботи, витрат на можливі ризики та маркетингові комунікації. Внаслідок чого набувають великого практичного значення дослідження, які спрямовані на розробку універсальних та високоточних методів оцінювання ефективності та інвестиційної привабливості стартапів в умовах невизначеності. Так як стартап є завжди унікальним, статистика по оцінюванню, впровадженню та супроводу проекту відсутня, що також значно ускладнює процес аналізу.

Пропонується об'єднати дві методології TEI та TCO, спільне використання яких дозволить оцінити інноваційний проект як з точки зору витрат та ефективності їх використання командою стартапу, так і з урахуванням потенційних ризиків, прихованих стратегічних переваг та недоліків.

Література. **1.** Згуровский М.З., Панкратова Н.Д. Системная стратегия технологического предвидения в инновационной деятельности // Системні дослідження та інформаційні технології. – 2003. – №3 – С. 7–24. **2.** PMBOK. Руководство к Своду знаний по управлению проектами, 5-е изд./Project Management Institute. PMI, 2013. 589 с. **3.** Global Risks 2013. Eighth Edition // World Economic Forum, 2013. **4.** Севастьянов А.К. Постоянное улучшение качества на пути к доверию и безопасности // Бизнес и безопасность. – Киев, издательство «ШАНС» ООО, №4, 2002. – С. 12–17.

***Притоманова О.М.*[1], *Журавель С.В.*[1], *Киселев М.Я.*[2]**
[1] *Днипровский национальный университет им. Олеся Гончара, Днепр, Украина;* [2] *Днепропетровский национальный университет железнодорожного транспорта имени академика В.Лазаряна, Днепр, Украина*

Управление проектами в IT-сфере методами PMI PMBOK на основе нейронечетких технологий

Под проектом в современном менеджменте понимают временное предприятие, направленное на создание уникального продукта, услуги или результата. Управление проектом – это процессы планирования, передачи, отслеживания и контроля всех аспектов проекта, а также мотивация всех участников проекта для достижения целей проекта в ожидаемые сроки, с определенными затратами, качеством, объемом, выгодами и рисками.

Одним из необходимых условий успеха проекта является нахождение баланса между конкурирующими базовыми (объем работ, ресурсы и расписание) и дополнительными ограничениями (качества, рисков, вовлеченности заинтересованных сторон, коммуникации, экспертизы команды и др.). Но, при создании проектов в сфере информационных технологий, классическое управление проектами становится малоприменимо в связи с уникальностью большинства программных продуктов, а также быстрой изменчивостью данной сферы. В связи с этим возникла модель Agile и методологии/фреймфорки Scrum, Kanban, PRINCE2, а так же гибкие методы в Американском национальном стандарте управления проектами PMI PMBOK (Project Management Institute – Project Management Body of Knowledge).

Применение методов математического моделирования для управления проектами позволяет определить последовательность улучшений процессов проекта для достижения его успеха. В работе предложены входные параметры математической модели управления проектом на основе адаптированных процессов PMBOK:

1. Анализ требований;
2. Отслеживаемость требований;
3. Обоснованность оценки задач;
4. Соблюдение графика поставок;
5. Качество программного кода;
6. Комфортность работы с командой;
7. Комфортность работы с руководителем команд;
8. Информативность отчетов о выполнении работ.

Очевидно, что количественная оценка входных параметров модели является сложной задачей, но качественно оценить данные параметры можно адекватнее. Например, задать входные параметры как лингвистические переменные с единой шкалой из пяти качественных термов: <отлично> – заказчик полностью доволен выполнением задач или процессов, нет необходимости улучшений; <хорошо> – заказчик доволен выполнением задач или процессов, возможны улучшения; <нормально> – выполнение задач или процессов рассматривается как достаточное с необходимостью улучшений; <плохо> – выполнение задач или процессов рассматривается как несоответствующее ожиданиям; <очень плохо> – выполнение задач или процессов полностью не соответствует ожиданиям. Выходной параметр проекта (успех проекта) может принимать значения <да> или <нет>.

Для моделирования сложных процессов, когда нет явной аналитической зависимости, описывающей поведение исследуемого объекта, или, если экспертные знания об объекте можно сформулировать только в лингвистической форме, целесообразно применять системы с нечеткой логикой. Математическая модель управления проектом, построенная на основе нейронечетких технологий, позволяет моделировать влияние каждой составляющей проекта на его результат, а также получать различные возможные оценки состояния проекта. Модель на основе нейронечеткого подхода может служить основой для создания систем поддержки принятия решений при управлении проектом.

Рак О.О.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Побудова мережі маршрутів громадського транспорту як задача багатокритеріальної оптимізації

У англомовній літературі задача знаходження оптимальних маршрутів громадського транспорту називається здебільшого UTRP (urban transit routing problem).

UTRP – це перша задача в рамках задачі створення мережі громадського транспорту (urban transit network design problem – UTNDP). Наступні задачі: вибор типів транспортних засобів для обслуговування маршрутів, створення розкладу руху та розкладів для працівників підприємств-перевізників [3].

Якщо аналізувати задачу знаходження оптимальних маршрутів громадського транспорту з боку обчислювальної складності, то це задача з категорії NP-повних, що на великих обсягах даних вирішується здебільшого евристичними та метаевристичними методами – шляхом постійної генерації мереж та їх валідації [1].

UTRP складно сформулювати математично, адже на відміну від задачі комівояжера (travelling salesman problem) маршрути можуть бути довгими чи короткими, прямими чи циклічними, одна зупинка може використовуватись різними маршрутами.

Крім генерації маршрутів також складної задачею є їх валідації, адже, наприклад, для знаходження оптимального шляху від початку до кінця подорожі та обрахування значення критерію середнього часу подорожі пасажира потрібні доволі значні обчислювальні ресурси.

Вхідними даними для задачі є дані про існуючу інфраструктуру (топологію) [2, 3]: граф вулично-дорожньої мережі (ВДМ), час проїзду між вузлами графа ВДМ, зупинки громадського транспорту, попит на перевезення з однієї точки до іншої – матрицю кореспонденцій.

Шляхом аналізу джерел були визначені найбільш часті критерії оптимізації: відсоток задоволеного попиту без пересадок (з 1 чи 2 пересадками), незадоволеного попиту, а також середній час на поїздку на одного користувача.

Найбільш розповсюдженим прикладом для тестування алгоритмів є Mandl's Swiss Network. З ним можна ознайомитись за посиланням – http://users.cs.cf.ac.uk/C.L.Mumford.

В цілому, це різновид задачі багатокритеріальної оптимізації [4], для вирішення якої у сучасних працях використовують здебільшого такі методи як генетичні алгоритми (genetic algorithms), пошук з заборонами (tabu search), імітація випалювання (simulated annealing). Найбільш популярними стали алгоритми колективного інтелекту (swarm intelligence): алгоритм колонії мурах та алгоритм колонії бджіл.

Це пов'язано з припущенням, що люди можуть приймати рішення про вибір оптимального маршруту подібно до інших людей і з урахуванням їх маршруту, навіть якщо це не найбільш оптимальний маршрут. Власне, при вирішенні UTRP це полегшує задачу, адже якщо деяка частина людей обирає однаковий маршрут, то один маршрут зможе перевезти більшу кількість людей, аніж якщо кожна людина обирала індивідуально оптимальний маршрут. Також це краще і для людини, бо на наступних етапах UTNDP будуть обраховуватись оптимальні інтервали руху, а, очевидно, чим більше пасажирів перевозить маршрут, тим менший інтервал руху буде для нього оптимальним.

Література. **1.** Mumford, "Research on the Urban Transit Routing Problem (Bus Routing)", 06-Apr-2016. [Online]. Available: https://users.cs.cf.ac.uk/C.L.Mumford/Research%20Topics/UTRP/Outline.html. [Accessed: 10-Mar-2017]. **2.** P. Chakroborty and T. Dwivedi, "Optimal Route Network Design for Transit Systems Using Genetic Algorithms" in Engineering Optimization, vol. 34(1), 2002, pp. 83–100. **3.** V. Guihairea and Jin-Kao Hao, "Optimal Route Network Design for Transit Systems Using Genetic Algorithms" in Transportation Research Part A: Policy and Practice, vol. 42, pp. 1251–1273, Dec. 2008. **4.** O.J. Ibarra-Rojas, F. Delgado, R. Giesen and J.C. Munoz, "Planning, operation, and control of bus transport systems: A literature review" in Transportation Research Part B: Methodological, vol. 77, pp. 38–75, Jul. 2015.

***Романенко В.Д.*[1], *Мілявський Ю.Л.*[1], *Мокін В.Б.*[2]**
[1]*КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна;* [2]*Вінницький національний технічний університет, Вінниця, Україна*

Автоматизація управління водним балансом водогосподарської ділянки з нейтралізацією обмежених природних збурень на основі когнітивних моделей

Когнітивне моделювання застосовується у тих випадках, коли об'єктом дослідження є складна система великої розмірності, структуру якої можна представити зваженим орієнтованим графом. Цей граф прийнято називати когнітивною картою (КК) складної системи. При цьому вершини КК відображають координати (фактори), а ребра – зв'язки між цими координатами.

У даній доповіді розроблено КК причинно-наслідкових зв'язків для оцінки впливу об'ємів води різних водних об'єктів у межах водогосподарської ділянки (ВГД). До вимірюваних координат x_i належать:

- об'єм води, отриманий в резерв за рахунок дотаційних перекидань ззовні від даної ВГД;
- об'єм води, отриманий зі стоку річок і атмосферних опадів для наповнення водойм;
- приріст стоку річок на даній ВГД;
- об'єм води, переданий за межі ВГД через зовнішні та внутрішньобасейнові перекидання;
- об'єм (резерв) води у поверхневих водоймах;
- об'єм води, що випадає у вигляді атмосферних опадів на даній ВГД та що повертається у водотоки цієї ВГД у результаті танення криги;
- об'єм води, що випаровується в атмосферу з даної ВГД;
- об'єм води, отриманий з підземних джерел для наповнення водосховищ.

До невимірюваних координат вершин КК y_i належать:

- об'єм води в ґрунті, отриманий шляхом фільтрації з водосховищ та ставків;
- об'єм (рзерв) води у підземних водних об'єктах на даній ВГД;
- об'єм води бічної припливності, що формується у межах ВГД;
- об'єм зворотних та стічних вод підприємств;
- об'єм води при водозаборі підприємств;
- об'єм води, що використовується підприємствами на безповоротне споживання;
- об'єм підземних вод, гідравлічно пов'язаний зі стоком річок.

Для вимірюваних та невимірюваних координат вершин КК розроблено моделі взаємопов'язаних імпульсних процесів відповідно:

$$\Delta\overline{X}(k+1) = A\Delta\overline{X}(k) + D\Delta\overline{Y}(k), \tag{1}$$

$$\Delta\overline{Y}(k+1) = C\Delta\overline{Y}(k) + R\Delta\overline{X}(k), \tag{2}$$

де $\Delta\overline{X}(k) = \overline{X}(k) - \overline{X}(k-1)$, $\Delta\overline{Y}(k) = \overline{Y}(k) - \overline{Y}(k-1)$ – перші різниці, A, B, C, R – вагові матриці суміжності. У моделі імпульсного процесу (1) невимірювані прирости $\Delta\overline{Y}(k)$ відіграють роль обмежених збурень. Для автоматизації управління у модель (1) вводиться вектор управління у такий спосіб:

$$\Delta\overline{X}(k+1) = A\Delta\overline{X}(k) + B\Delta\overline{U}(k) + D\Delta\overline{Y}(k). \tag{3}$$

Задача полягає у синтезі вектора управління регулятора стану $\Delta\overline{U}(k) = -K_p\Delta\overline{X}(k)$, що заглушає збурення $\Delta\overline{Y}(k)$ у моделі (1). Синтез виконується на основі мінімізації розміру інваріантних еліпсоїдів динамічної системи, що зводиться до виконання еквівалентних умов у вигляді лінійних матричних нерівностей.

Рудакова А.В., Терновая Т.И., Олокоба О.Й.
Херсонский национальный технический университет, Херсон, Украина

Методы предварительной обработки сигналов в автоматизированной системе диагностики для функционального контроля электромеханического оборудования

Для определения режима функционирования (допустимого, критического, аварийного) электромеханических систем в процессе эксплуатации широко используются методы виброакустодиагностики.

Системы виброакустодиагностики основываются на анализе вибраций и шумов, исходящих от движущихся механизмов [1]. С помощью анализа вибрационных и акустических сигналов можно не только выявлять повреждения деталей механизмов и конструкций, ослабление крепления, отложения на стенках, коррозию и т.п., но и идентифицировать режимы работы двигателей в зависимости от нагружения.

Задача акустической диагностики заключается в определении технического состояния контролируемого объекта по зарегистрированному шуму. Производственный шум имеет широкий частотный спектр сигнала и каждому состоянию работоспособности соответствует свой характерный спектр.

Классический подход к выделению информативных признаков, на основе которых проводится оценка ситуации (распознавание конкретного состояния), базируется на анализе частотного спектра сигналов, получаемого с помощью преобразований Фурье [2]. Недостатком традиционного подхода является вычислительная сложность при осуществлении дискретного преобразования Фурье и необходимость значительных объемов памяти для хранения спектров эталонных сигналов.

Для сравнения эффективности различных методов предварительной обработки диагностических сигналов использован ряд реализаций акустических сигналов, полученных экспериментально при всевозможных условиях нагружения электропривода. При исследованиях расчеты выполнялись в следующем порядке: 1) построение гистограммы распределения, определение числовых характеристик случайного сигнала и проверка гипотезы о законе распределения [3]; 2) получение частотного спектра и проведение спектрального анализа сигнала [2]; 3) проведение кратно-масштабного анализа [4]; 4) временной анализ случайных сигналов в приращениях разной степени [5].

Целью исследований является проверка целесообразности использования новейших методов цифровой обработки сигналов для обнаружения аномалий, свидетельствующих о каких-либо неполадках в электромеханических системах, и выявление их влияния на изменение значений информативных признаков.

Интеграция систем виброакустодиагностики в компьютеризированные системы управления сложными многоприводными установками открывает возможность в реальном времени оценивать и выявлять критические режимы работы электромеханического оборудования, и своевременно формировать управляющие воздействия, позволяющие стабилизировать функционирование производственных комплексов.

Литература. **1.** Азовцев Ю.А., Баркова Н.А., Гаузе А.А. Вибрационная диагностика роторных машин и оборудования целлюлозно-бумажных комбинатов. / Ю.А. Азовцев, Н.А. Баркова, А.А. Гаузе // СПб.: СПбГУРП, 2014. – 127 с. **2.** Бендат Дж. Применение корреляционного и спектрального анализа / Дж. Бендат, А. Пирсол. – М.: Мир, 1983. – 312 с. **3.** Вентцель Е.С. Теория случайных процессов и ее инженерные приложения / Е.С. Вентцель, Л.А. Овчаров. – М.: Высшая школа, 2000. – 384 с. **4.** Новейшие методы обработки изображений / Под ред. А.А. Потапова – М.: ФИЗМАТЛИТ, 2008. – 496 с. **5.** Прикладной анализ случайных процессов / Под ред. С.А. Прохорова // Самара: СНЦ РАН, 2007. – 582 с.

Савченко А.В.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Інформаційно-експериментальна підтримка процесів моделювання еко-ефективного згоряння палива в дизелі

Актуальність. Під час розробки робочих процесів перспективних дизелів намагаються зменшити максимальну температуру з метою зниження обсягів утворення оксидів азоту, що відповідає сучасним вимогам еко-ефективності енергетичних установок. Перебіг таких процесів відповідає певному алгоритму окиснення палива в циліндрі дизеля. У даній роботі досліджені питання підтримки швидкості реакції на максимально можливому рівні, що відповідає скороченню часового інтервалу згоряння і дозволяє знизити утворення оксидів азоту NO_x [1, 2].

Мета та задачі. Для розв'язання задачі моделювання еко-ефективного згоряння палива в дизелі необхідно визначити алгоритм ідентифікації початку процесу окиснення палива за відповідною поведінкою тиску у циліндрі на основі теоретичної обробки результатів експериментального дослідження режимів роботи дизеля на різних видах палива.

Розв'язання завдання дослідження. В межах даного дослідження звернено увагу на ідентифікацію точки початку згоряння, що фіксується за відповідною інформацією особливостей поведінки тиску в циліндрі.

За експериментальними даними встановлюється функція наростання тиску в циліндрі, що за багатьма результатами досліджень визначається таким чином:

- тиск адіабатичного стискання за теоретичним розрахунком

$$P_{teor}(x) = \frac{P(x-1)V(x-1)^K}{V(x)^K}; \tag{1}$$

- невідповідність зафіксованих виміряних величин тиску теоретичним розрахунковим значенням

$$\Delta P_{teor}(x) = P_{\text{ц}}(x) - P_{teor}(x), \tag{2}$$

де $P_{teor}(x)$ – розрахунковий тиск у точці x за умови адіабатичного стискання; $V(x)$ – надпоршневий об'єм в точці x; K – показник адіабати, $K = 1,4$ для повітря; $P_{\text{ц}}(x)$ – виміряний тиск у циліндрі в точці x.

Для фіксації початку процесу згоряння достатньо індикаторної діаграми дизеля. За характерним підйомом тиску визначають початок процесу окиснення, який необхідно певним чином контролювати для досягнення зростання його швидкості з відповідним скороченням часового інтервалу, що забезпечить інгібування утворення оксидів азоту та збільшення тиску до 30% від показників дизеля на стандартному дизельному паливі. При цьому утримується необхідна температура окиснення палива, однак вона нижча за стандартну, що дозволяє знизити вміст NO_x у відпрацьованих газах дизеля (рис. 1) [1].

Надійність запропонованих засобів визначення моменту початку впорскування палива забезпечується визначенням на основі похідної функції тиску палива за кутом повороту колінчастого валу. В межах даної задачі прийнято, що максимум похідної тиску палива відповідає початку подачі палива в камеру згоряння і дотримання еко-ефективного перебігу процесу окиснення (рис. 2).

Висновки. Таким чином, запропонований інформаційно-експериментальний підхід до розв'язання задач дослідження особливостей еко-ефективного спалювання палива в дизелях дозволив підтвердити ствердження, що зниження максимальної температури в циліндрі та зменшення тривалості процесу згоряння мають вчинити позитивний вплив на зниження рівня викидів оксидів азоту з відпрацьованими газами дизеля. Практична цінність проведених робіт визначається розробкою загального алгоритму обробки даних роботи дизеля.

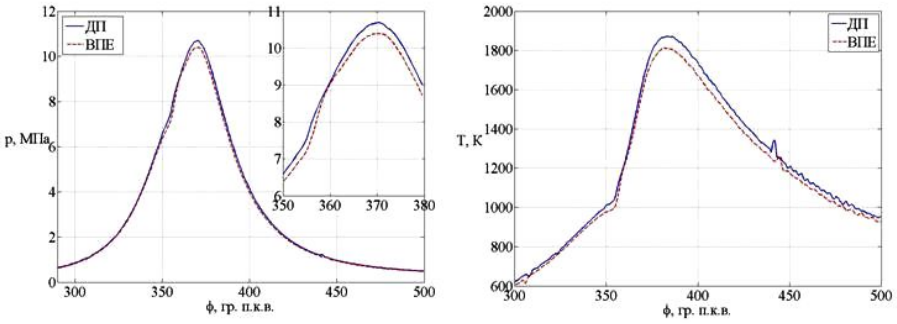


Рис. 1. Результати обробки експериментальних даних залежності тиску та температури від кута повороту колінчастого валу

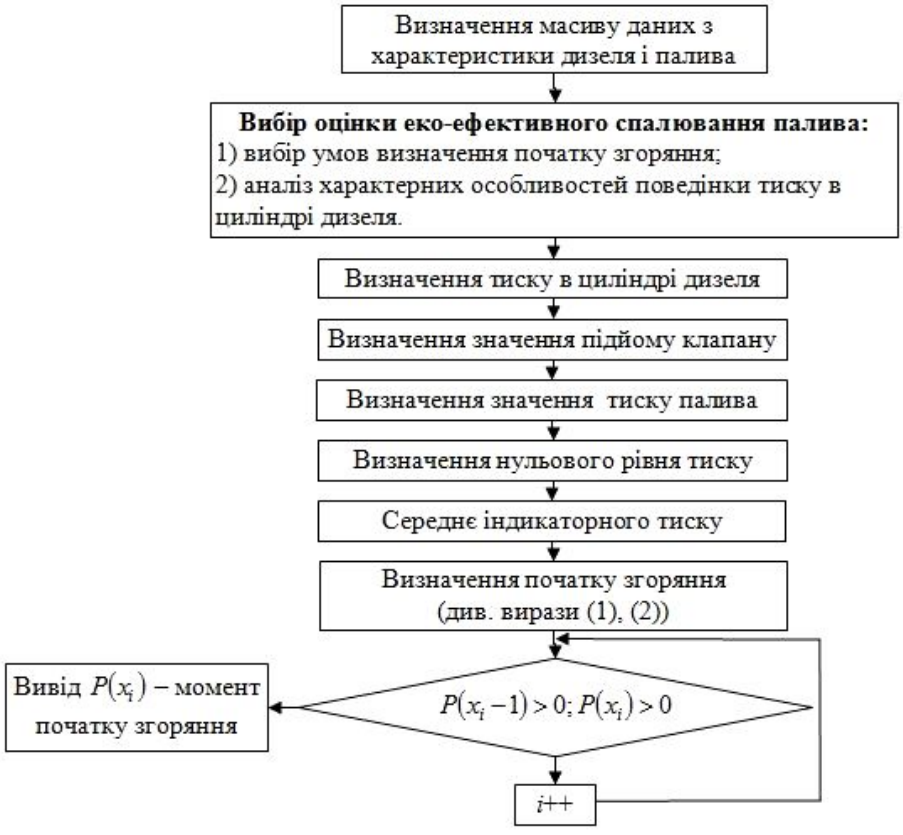


Рис. 2. Загальний алгоритм обробки даних експериментального дослідження роботи дизеля для визначення початку еко-ефективного згоряння палива

Література. **1.** Марченко А.П. Особливості процесу згоряння в дизелі при роботі на водопаливній емульсії / А.П. Марченко, І.В. Парсаданов, А.О. Прохоренко, А.В. Савченко, О.О. Осетров, Д.В. Мешков // Двигатели внутреннего сгорания. – 2016. – № 1. – С. 3–10. **2.** Марченко А.П. Экспериментальное определение энергии активации водотопливной эмульсии / А.П. Марченко, А.А. Прохоренко, А.В. Савченко // Двигатели внутреннего сгорания. – 2016. – № 2. – С. 9–14.

Савченко І.О.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Оцінювання чутливості двохетапної процедури модифікованого методу морфологічного аналізу

У задачах передбачення більшість об'єктів, процесів, явищ, що досліджуються, характеризуються неточністю, невизначеністю, неповнотою інформації, тому традиційні математичні моделі і підходи неможливо або недоцільно використовувати. Це обумовило появу ряду методів якісного аналізу, одним з яких є модифікований метод морфологічного аналізу (МММА) [1].

Експертне оцінювання, що є основою методів якісного аналізу, часто є єдиним джерелом інформації про об'єкти передбачення, однак експертні оцінки суб'єктивні і відображають реальний світ лише в деякому наближенні. Завжди існує можливість того, що неточність вхідних даних призведе до хибного ранжування альтернатив. Для вирішення цієї проблеми іноді застосовують методику аналізу чутливості. Зокрема, існують дослідження в цьому напрямку для методу аналізу ієрархій [2]. Такий підхід дозволяє визначити, які вхідні оцінки є більш критичними, тобто призводять до зміни рангів альтернатив при менших відхиленнях. На ці оцінки слід звернути особливу увагу і, можливо, провести уточнююче оцінювання, оскільки ця інформація може виявитись визначальною для прийняття рішення.

В [3] показано, що розв'язок системи рівнянь МММА може бути знайдений у вигляді

$$\vec{x}_i = diag(C_{i|k}\vec{p}_k)\vec{p}_i,\ i,k \in [1;N],\ k \neq i,$$

де $\vec{x}_i$ – вектор шуканих імовірностей, $\vec{p}_i$ – вектор вхідних (початкових) імовірностей, $C_{i|k}$ – матриця, побудована спеціальним способом із матриці взаємозв'язків альтернатив параметрів, N – кількість параметрів морфологічної таблиці. Цей вираз дає можливість оцінити три типи чутливості відносно вхідних даних в МММА:

1. Чутливість ймовірності альтернативи параметра відносно зміни значень незалежної ймовірності альтернатив цього ж параметра;
2. Чутливість ймовірності альтернативи параметра відносно зміни значень незалежної ймовірності інших альтернатив іншого параметра;
3. Чутливість ймовірності альтернативи параметра відносно зміни значень матриці взаємозв'язків.

Важливо також оцінити чутливість результатів, отриманих в ході двохетапної процедури МММА [1], оскільки на другому етапі оцінюються саме альтернативи рішення. Можна показати, що результат оцінювання результативності альтернатив параметрів другого етапу для морфологічної таблиці з двох параметрів може бути знайдений у вигляді

$$R_j^{(i)} = \vec{p}_1^T C_R \vec{p}_2,\ i \in [1;N'],\ j \in [1;n_i], \tag{1}$$

де $R_j^{(i)}$ – результативність j-ї альтернативи i-го параметра, N' – кількість параметрів морфологічної таблиці другого етапу, n_i – кількість альтернатив i-го параметра, C_R – матриця, сформована на основі матриці взаємозв'язків і матриці зв'язків. Використовуючи міркування, описані в [3], можна вираз (1) поширити на випадок довільної кількості параметрів. За допомогою (1) можна сформувати вирази для оцінки чутливості значення $R_j^{(i)}$ відносно незалежних оцінок ймовірності або значень матриці взаємозв'язків.

Література. **1.** Савченко І. О. Методологічне і математичне забезпечення розв'язання задач передбачення на основі модифікованого методу морфологічного аналізу // Системні дослідження та інформаційні технології. – 2011. – №3. – С. 18–28. **2.** Nataliya D. Pankratova, Nadezhda I. Nedashkovskaya. Sensitivity analysis of a decision-making problem using the analytic hierarchy process // International Journal "Information Theories and Applications", Vol. 23, Number 3, 2016. – P. 232–251. **3.** Savchenko, I.O. Estimating the Solution Sensitivity in Application of the Modified Morphological Analysis Method // Cybernetics and Systems Analysis. – September 2016, Volume 52, Issue 5, pp. 782–790.

Северилов П.В., Боровская Т.Н., Колесник И.С., Гришин Д.И.
Винницкий национальний технический университет, Винница, Украина

Оптимизация производственных систем с учетом ресурсных обратных связей

Вступление. Потенциальная эффективность производственной системы определяется эффективностью технологий и продуктов производства, но реальные результаты существенно зависят от связей с окружающей средой – экологией, поставщиками ресурсов, потребителями и конкурентами. Преимущество данной разработки относительно аналогов обусловлено интеграцией математических моделей с программами пакетов моделирования. Такие рабочие модели позволяют устранить проблемы аппроксимации, размерности, сходимости. В [1, 2] была поставлена и решена задача оптимального агрегирования для одноконтурной системы с ресурсной обратной связью (РОС). В данной работе поставлена новая задача разработки рабочей модели и метода оптимального агрегирования производственных систем [1] с учётом множественных ресурсных обратных связей. Примеры систем с РОС: – синтез этилена, где непрореагировавший компонент возвращается в реактор; – биореактор имеет два контура РОС – возвращение воды в систему реактора биогумуса в агропроизводство; – возвращение капиталовложений и затрат производства. Более сложные обратные связи – информационно-ресурсные – возникают при взаимодействии с потребителями, персоналом и конкурентами.

Оптимальное агрегирование структур с РОС. Использование методологии оптимального агрегирования меняет содержание научного исследования: собственно, оптимизация – рутинная операция, главная часть решения задачи – анализ объекта и конструирование ресурсной модели объекта. По содержанию это *прикладной системный анализ.* Новизна работы – расширение базовой алгебры оптимального агрегирования за счёт введения параметризованного ассоциативного бинарного оператора оптимального агрегирования структур РОС. В отличие от аналогов, в методологии оптимального агрегирования критический этап – это построение модели производственной системы (ПС) как технологического преобразователя «ресурс – продукт». На рис. 1 представлены: а) – базовая ресурсная структура одноконтурной ПС, б) – структура двухконтурной ПС, в) – отображение ресурсной структуры б) в бинарное дерево оптимального агрегирования системы – формулу решения задачи – оптимальную эквивалентную функцию производственной системы с двумя контурами РОС: функционирования и развития для базовых задач.

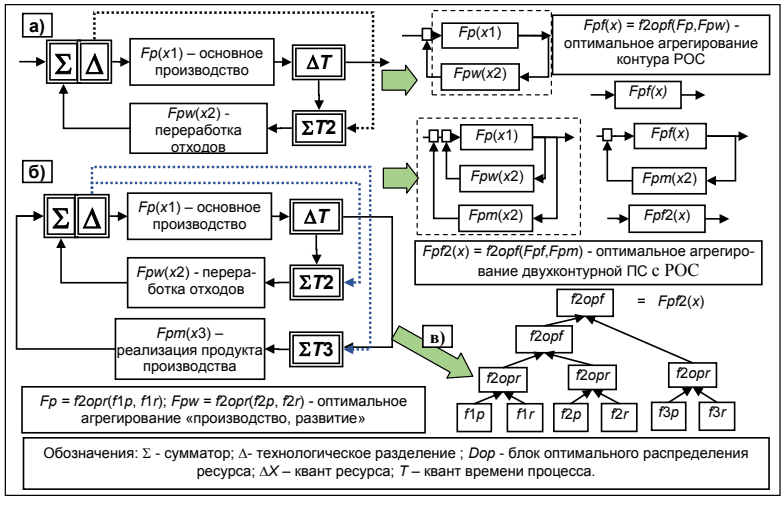


Рис. 1. Изоморфное отображение ресурсных структур в бинарное дерево оптимального агрегирования

На рис. 2 представлены альтернативы управления процессами производства и развития.

Рис. 2. Альтернативные структуры управления производством и развитием системы с РОС

Специфика функций развития в том, что существуют множества альтернативных вариантов развития – от ремонта наличного оборудования, покупки и монтажа нового до реализации собственных инноваций. На рис. 3 представлена схема формирования функций развития. Выбор альтернативы развития зависит от доступных ресурсов и времени на реализацию.

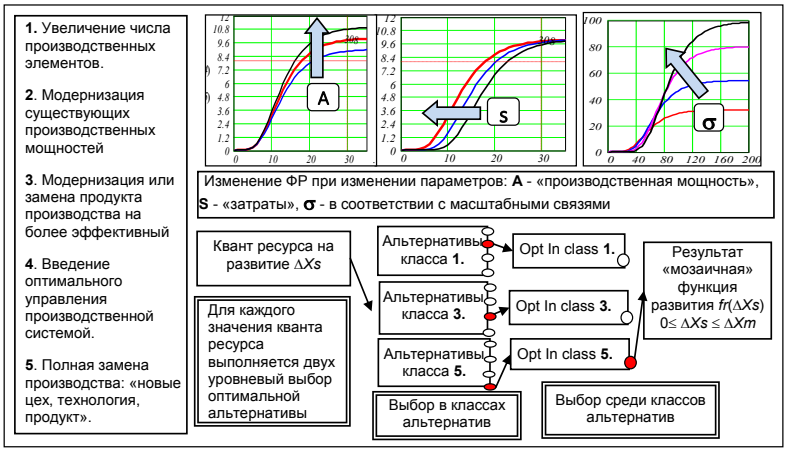


Рис. 3. Построение функций развития систем с РОС на множестве альтернатив развития

Выводы. Выполнена разработка операторов оптимального агрегирования многоконтурных структур. В работе представлены исследования рабочих математических моделей с многоконтурными РОС, позволяющих управлять трудноформализуемыми информационными процессами в средах потребителей, персонала и конкурентов.

Литература. **1.** Боровська Т.М. Моделі оптимального інноваційного розвитку виробничих систем / Т.М. Боровська, І.С. Колесник, В.А. Северілов, П.В. Северілов // Східно-Європейський журнал передових технологій: Математичне та інформаційне забезпечення комп'ютерно-інтегрованих систем управління. – 2014. – Т. 5, № 2(71).– С. 42–50. DOI: 10.15587/1729-4061.2014.28030. **2.** Borovska T.N. Modeling and optimization of agrarian systems with waste recycling in bioreactors / P.V. Severilov., T.N. Borovska, Yu.N. Dmytryk, E.P. Khomyn // Nauka i studia (Poland). – 2014. – № 16 (126). – P. 42–50. – ISSN 1561-6894.

Сербул Р.С., Данилов В.Я.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Системний підхід в управлінні IT-проектами

Вступ. Менеджмент, як технологічний процес, є основним і невід'ємним фактором розвитку проектів. У переважній більшості випадків для невеликих компаній найняти досвідченого менеджера є не просто – послуги гідного фахівця коштують недешево, та й довіряти на ранньому етапі сторонній особі учасникам проекту буде складно. Тому менеджментом займаються, як правило, самі учасники проекту.

Основний процес менеджменту. Менеджер працює з процесами. Процеси є складовою частиною проектів. Процес може бути разовим або безперервним, але він в будь-якому випадку є ітеративним. Це означає, що у кожного процесу є циклічні властивості – він легко може бути повторений, і навіть для початку нового процесу можливо застосовувати напрацьований досвід – академічні методики, особистий досвід, досвід колег і так далі [1, 2].

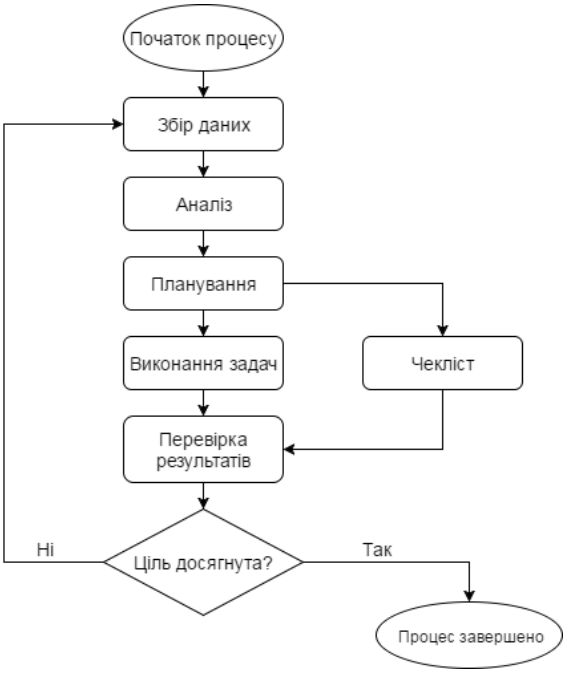


Рис. 1. Алгоритм виконання процесу

Розглянемо основні кроки виконання процесу:

- до початку процесу необхідно формалізувати вихідні дані і виділити цілі;
- етап аналізу є опціональним. Він проводиться, в залежності від масштабів і ціни процесу. Якщо процес дорогий – всі вихідні дані піддаються деталізації, інформація доповнюється схемами і резюме;
- на етапі планування вибираються методи розв'язання задачі, визначається, як саме буде здійснюватися процес;
- для забезпечення коректності приймання ще на етапі планування складається чекліст – список критеріїв, який однозначно дає зрозуміти, що проект завершений;
- якщо процес не є безперервним – по досягненню цілей він може бути завершений;
- при повторному виконанні процесу до вихідних даних додаються результати попередньої ітерації.

Один з найпопулярніших і в той же час суперечливих методів традиційного менеджменту – делегування. Є купа академічної інформації про те, як, кому і коли доручати завдання [3].

В умовах невеликої IT-компанії, як правило, делегування в загальному розумінні недоступно. Занадто мало грошей, занадто мало людей. Безпосередніх учасників у новонародженого проекту, як правило, мало. Наймати експертів з боку – дорого, та й до того ж загрожує витоком інформації і додатковими тимчасовими витратами.

Тому, для забезпечення ефективності менеджменту (вдалого управління проектом) доцільно приділяти більшу увагу іншим доступним методикам:

- системний аналіз;
- інтерактивний контроль;
- управління ризиками.

«Розділяй і володарюй», заповідали нам стародавні правителі. І до цього дня ця методика управління проектами виявляється однією з найефективніших [6].

Методику поділу завдань на підзадачі можна застосовувати як для розробки робочого плану, так і для аналізу інших аспектів і ситуацій.

Для кожного аналітичного елемента, для кожного фактора і підзадачі формалізуються три питання: мета, об'єкти і методи. Оцінюється масштаб, важливість і складність завдання. Якщо завдання є складним, масштабним або критичним – воно ділиться на підзадачі або виділяється в окремий проект [4, 5].

Сильні корпорації, такі як Google, використовують метод малих груп для вирішення практично всіх ключових завдань. Виділяється колектив зацікавлених фахівців, який працює над проектом. Як показує практика, завдання вирішуються, і проекти виявляються дуже успішними.

Висновки. Завдяки методу системного аналізу малі робочі групи вирішують складні завдання. Причому роблять це швидко і дешево. Не потрібно створювати тонни не потрібної документації. Але використання навіть деяких основних методик дасть проекту життя. Основні методики:

- Поділ завдань на підзадачі;
- Виділення субпроектів;
- Запис (логгіювання) всього.

Цілями системного підходу є:

- Отримання прозорого і очевидного уявлення всіх деталей проекту;
- Виявлення потенційно вузьких місць;
- Виявлення прихованих чинників, перш за все витрат;
- Узгодження тактичного і стратегічного бачення між усіма учасниками проекту.

Система управління проектом – це взяті в своїй єдності методи, процедури, програми. Менеджери повинні представляти специфіку управління людськими ресурсами організації (і, в тому числі, команди) в сучасних умовах: пріоритет психологічним факторам, мотивації, стимулювання, розвитку, виховання самоорганізації, самоконтролю, професійної самореалізації.

Література. **1.** Вучкович-Стадник, А.А. Оценка персонала: четкий алгоритм действий и качественные практические решения [текст] / А.А. Вучкович-Стадник. – М.: Эксмо, 2008. – 215 с. **2.** Гринберг, Дж. Организационное поведение: от теории к практике [текст] / Дж. Гринберг, Р. Бэйрон. – М.: ООО «Вершина», 2004. – 912 с. **3.** Дульзон А.А. Управление проектами: учеб. пособие / А.А. Дульзон; Нац. исслед. Том. политехн. ун-т. 3-е изд., перераб. и доп. Томск: Изд-во Том. политехн. ун-та, 2010. 334 с. **4.** Biafore B. Microsoft Project. 2013 / B. Biafore. The Missing Manual PDF O'Reilly Media. 2013. 812 p. **5.** Project management. A Guide to the Project Management Body of Knowledge: PMBOK guide. 3rd ed. 2004. 506 p. **6.** Campbell C.A. The One-Page Project Manager for Execution: Drive Strategy and Solve / C.A. Campbell, M. Campbell // Problems with a Single Sheet of Paper PDF Wiley. 2nd ed. 2012. 210 p.

Смирнов С.А., Коваль Л.В.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Встановлення вагових коефіцієнтів критеріїв за допомогою порядкових експертних оцінок

Задачі багатокритеріальної оптимізації на практиці найчастіше розв'язуються методом лінійної згортки. Він базується на використанні вагових коефіцієнтів критеріїв, проте процедури їхнього отримання є суб'єктивними й не достатньо обґрунтованими [1]. Кількісне оцінювання їхніх значень здійснюється в процесі спеціальних експертних процедур. Незважаючи на те, що такі процедури містять у собі певний елемент довільності, часто вони дають змогу знайти задовільні практичні рішення багатокритеріальних задач.

Однак, є дві обставини, що обмежують сферу коректного застосування цього методу. По-перше, не можна з певністю стверджувати, що глобальна згортка саме лінійно залежить від початкових критеріїв. Тим не менше, низка результатів теорії корисності дає змогу побудувати локальні функції корисності із початкової множини критеріїв та обґрунтувати лінійну залежність глобальної функції корисності від них [2]. Отже, метод лінійної згортки є досить універсальним та теоретично обґрунтованим.

По-друге, є обмеження, що стосується процедур експертного оцінювання вагових коефіцієнтів. Мета цих процедур – якомога точніше чисельно виразити неформальні, якісні уявлення експертів. Такі процедури не можуть бути повністю алгоритмізовані й завжди матимуть машинно-людський характер. Методика організації та проведення таких процедур спеціально розробляється та вважається цілком надійною. Було проведено безліч психологічних експериментів, за яких порівнювалися різні методи встановлення ваги критеріїв. Проте вони дають різні результати [3]. Саме використання в таких методах кількісних точкових оцінок, можливо, і є причиною такої неузгодженості. До того ж, такі оцінки не завжди цілком відповідають суб'єктивним можливостям експертів, а також об'єктивним особливостям предметної області.

Інтервальний підхід дозволяє підвищити надійність процедури оцінювання вагових коефіцієнтів. Крім того, такі оцінки є більш зручними для експерта та задовільними для поставленої задачі. Проте вони все ще є кількісними, тому ризик помилки експерта залишається значним [4].

Тому зараз пропонується використання порядкових експертних оцінок, через їхню більшу надійність. Водночас виникає проблема перетворення якісних оцінок у кількісні, яка вирішується за допомогою методів колективного прийняття рішень. Експертам пропонується упорядкувати критерії за важливістю. Результатом буде профіль переваг, до якого можна застосувати методи визначення переможця за кількісною оцінкою. До таких методів належать метод Борда та деякі методи змістовні за Кондорсе. Отримані кількісні оцінки після нормування дають вагові коефіцієнти.

У метода Кондорсе відсутня суб'єктивність, що властива методу Борда, проте досить часто під час аналізу реальних профілів переваг виникає ситуація, коли переможця немає. Це відбувається через виникнення циклу, тобто порушується транзитивність. Це можна спостерігати на прикладі парадоксу Кондорсе.

Нехай маємо 60 виборців, і вони голосують за кандидатів A, B, C так:

- 23-є: A≻B≻C
- 10-о: C≻A≻B
- 8-о: C≻B≻A
- 2-є: B≻A≻C
- 17-о: B≻C≻A

Тоді маємо такі по-парні порівняння: A≻B 33, B≻A 27; C≻A 35, A≻C 25; B≻C 42, C≻B 18. Отже, у результаті маємо такі пари: A≻B, B≻C, C≻A. Отримали суперечність.

У деяких ситуаціях змістовні за Кондорсе методи дають змогу уникнути парадоксу. Зокрема, часто це дозволяє зробити метод Шульце. За його допомогою можна визначити переможця,

коли виборці під час голосування впорядковують кандидатів згідно зі своїми перевагами. Шляхом від кандидата C_1 до кандидата C_n називаємо ситуацію, за якої є орієнтований граф від C_1 до C_2, від C_2 до C_3, ..., C_n. У цьому випадку також є непряма перемога C_1 над C_n. Силою непрямої перемоги є найслабша пряма перемога одного кандидата над іншим на цьому шляху. Силою найсильнішого шляху від кандидата A до кандидата B буде максимальне значення із сил усіх можливих шляхів від A до B, тобто всіх непрямих перемог. Якщо шляху між якимись кандидатами немає, то сила найсильнішого шляху становить 0. Переможцем вважається кандидат, що перемагає будь-якого іншого кандидата. Можливим переможцем є кандидат, який перемагає або дорівнює кожному іншому.

Перевагами методу Шульце є точніше та повніше врахування реальних побажань виборців, чого зазвичай вистачає для подолання парадокса Кондорсе. Недоліком є складність аналізу, унаслідок повноти матриці прямих перемог та великої кількості непрямих перемог. Розглянемо приклад.

Табл. 1. Матриця сил прямих перемог

	*,A	*,B	*,C	*,D
A,*	-	**9**	3	**8**
B,*	2	-	**6**	4
C,*	**8**	5	-	**6**
D,*	3	**7**	5	-

Табл. 2. Матриця сил непрямих перемог

	*,A	*,B	*,C	*,D
A,*	-	**9**	6	*6*
B,*	6	-	6	6
C,*	**8**	**8**	-	**8**
D,*	*6*	**7**	6	-

Табл. 3. Матриця найсильніших шляхів

	*,A	*,B	*,C	*,D
A,*	-	A≻B	A≻B≻C	A≻B≻C≻D
B,*	B≻C≻A	-	B≻C	B≻C≻A≻D
C,*	C≻A	C≻A≻B	-	C≻A≻D
D,*	D≻B≻C≻A	D≻B	D≻B≻C	-

Тому, за методом Шульце переможцем є кандидат C, бо C≻A≻D≻B. Вважаємо кандидатів за критерії, а виборців – за експертів. Індивідуальна сила для кожного критерія обраховується як сума сил непрямих перемог за рядком: A=21, B=18, C=24, D=19. Загальна сума сил непрямих перемог становить 82. Вагові коефіцієнти критеріїв знаходимо нормуванням індивідуальних сил: c_A=0,26; c_B=0,22; c_C=0,29; c_D=0,23. Так, критерій C має найбільшу вагу серед усіх інших.

Отже, головною ідеєю роботи було знаходження якісного визначення вагових коефіцієнтів. У доповіді проведено застосування та порівняння методів Борда, Копленда, Сімпсона, Шульце та ін. З іншого боку, багатокритеріальний вибір є результатом агрегування індивідуальних переваг критеріїв. Однак, розглянуті методи не враховують кількісні чинники вибору. Крім того, наявні засоби, що дозволяють звести значення критеріїв до єдиної шкали та оцінити їхню "вагу" (наприклад, метод аналізу ієрархії Сааті), містять у собі людський чинник (вербальна шкала рівнів важливості тощо). Урахування рейтингу певних кандидатів дозволить точніше визначати переможця, відповідно до побажань виборців.

Література. **1.** Подиновский В.В. Количественная важность критериев // Автоматика и телемеханика – 2000, – № 5. – С. 110–123. **2.** Фишберн П.С. Теория полезности для принятия решений. – М.: наука, 1978. – 358 с. **3.** Ларичев О.И. Теория и методы принятия решений, а также Хроника событий в Волшебных Странах: Учебник. – М.: Логос, 2000. – 296 с. **4.** Смирнов С.А., Гонтаренко И.С. Гарантированный синтез скалярного критерия для решения задачи многокритериальной оптимизации // Систем. дослідж. та інформ. технології. – 2006, – № 2. – С. 99–106.

Статкевич В.М.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Регулярные выражения для языков сети Петри "производитель/потребитель" с ограниченным буфером размера 1 и 2

Сети Петри, предложенные К. Петри, являются удобным средством для моделирования различных процессов, систем и сетей [1,2]. Они обладают такими свойствами, как недетерминированность, асинхронность, параллелизм и конфликтность. Э. Дейкстрой предложена задача о производителе/потребителе [1] с неограниченным буфером и с ограниченным буфером p'_5 размера n, соответствующие сети изображены на рис. 1. Рассмотрим языки L-типа для данных сетей: переходам t_1, t_2, t_3 и t_4 соответствуют символы a_1, a_2, a_3 и a_4 алфавита, множество заключительных маркировок состоит из единственной маркировки, совпадающей с начальной маркировкой μ_0 (детальнее см. [1]). Для варианта с неограниченным буфером язык является контекстно-свободным, его допускает, например, следующий автомат с магазинной памятью (детальнее см. [3]):

$$\langle\{q_0,q_1,q_2,q_3\},\{a_1,a_2,a_3,a_4\},\{A\},\Delta,\{q_0\},\{q_0\}\rangle,$$
$$\Delta=\{((q_0,a_1,\varepsilon),(q_1,\varepsilon)),((q_0,a_3,A),(q_2,\varepsilon)),((q_1,a_2,\varepsilon),(q_0,A)),((q_1,a_3,A),(q_3,\varepsilon)),$$
$$((q_2,a_4,\varepsilon),(q_0,\varepsilon)),((q_2,a_1,\varepsilon),(q_3,\varepsilon)),((q_3,a_4,\varepsilon),(q_1,\varepsilon)),((q_3,a_2,\varepsilon),(q_2,A))\}.$$

Цель данного сообщения – указание регулярных выражений для языков сетей L-типа с ограниченным буфером размера $n=1$ и $n=2$. Обозначим эти языки L_1 и L_2 соответственно. Для случая $n=1$ на рис. 2 изображён граф достижимости соответствующей сети, являющийся конечным. По графу достижимости построен конечный автомат M: вершинам графа сопоставлены состояния автомата, а рёбрам графа – переходы автомата [4]. Следовательно, язык L_1 регулярен. По автомату M последовательным исключением состояний построен обобщённый конечный автомат (детальнее см. [3]), искомое регулярное выражение имеет вид

$$[a_1a_2[(a_1a_3+a_3a_1)(a_2a_1a_4a_3)^*(a_2a_4+a_4a_2)]^*a_3a_4]^*.$$

Для случая $n=2$ подобная схема приводит к регулярному выражению

$$[a_1a_2[a_1a_2a_3a_4+(a_1\|a_3)(a_2\|a_4)+(a_1a_2(a_1\|a_3)+(a_1\|a_3)a_2a_1)\cdot$$
$$\cdot[(a_2\|a_4)(a_1\|a_3)+a_2a_1a_4a_3+a_4a_3a_2a_1]^*(a_4a_3(a_2\|a_4)+(a_2\|a_4)a_3a_4)]^*a_3a_4]^*,$$

где $\|$ – оператор параллельной композиции [1]: $a_1\|a_3=a_1a_3+a_3a_1$, $a_2\|a_4=a_2a_4+a_4a_2$.

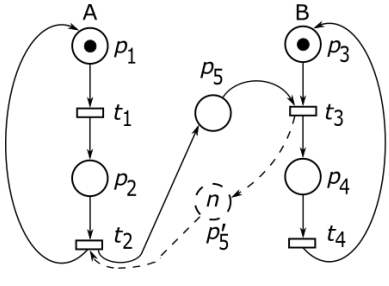


Рис. 1. Сеть Петри производитель(A)/потребитель(B)

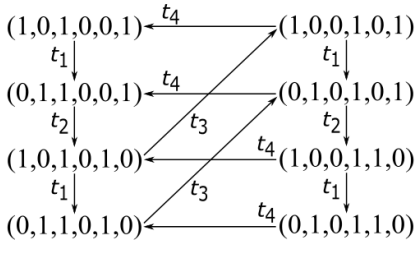


Рис. 2. Граф достижимости ($n=1$)

Литература. **1.** Питерсон Дж. Теория сетей Петри и моделирование систем. – М.: Мир, 1984. – 264 с. **2.** Котов В.Е. Сети Петри. – М.: Наука. Гл. ред. физ.-мат. лит-ры, 1984. – 160 с. **3.** Пентус А.Е., Пентус М.Р. Теория формальных языков. – М.: Изд-во ЦПИ при механико-математическом ф-те МГУ, 2004. – 80 с. **4.** Valk R., Vidal-Naquet G. Petri nets and regular languages. – Journal of Computer and System Sciences, 1981, **23**, p. 299–325.

Стефанишин Д.В.
Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна

Про врахування ризику при формуванні оптимальної стратегії введення нових агрегатів на гідроакумулюючих електростанціях в Україні

На сучасному етапі розвитку електроенергетики гідроакумулюючі електростанції (ГАЕС) розглядаються як одні з найефективніших і найперспективніших джерел регулювання навантажень в об'єднаних енергетичних системах (ОЕС) в контексті забезпечення стійкості, живучості, надійності і безпеки ОЕС, особливо в країнах з суттєвою долею теплової і, особливо, атомної енергетики [1, 2], до яких відноситься і наша країна. Досвід експлуатації численних ГАЕС в різних країнах світу показує, що на ГАЕС найбільш повно, в порівнянні з іншими технологіями генерації і акумулювання енергії, можна забезпечити якнайширші можливості для маневрування встановленими потужностями при вирішенні складних завдань безперебійного регулювання навантаження, особливо в умовах різко-змінних графіків роботи ОЕС.

В Україні натепер працює три ГАЕС: Київська – перша в Україні і на території країн бувшого СРСР, яку було введено в експлуатацію ще в 1972 р.; експлуатуються три агрегати на Дністровській і два агрегати на Ташлицькій ГАЕС. Основними їх функціями є регулювання частоти і графіка навантажень в ОЕС країни, формування аварійного резерву електроенергії. В планах, зокрема в Програмі розвитку гідроенергетики на період до 2026 р. [3], розглядається можливість добудови і введення в експлуатацію ще чотирьох агрегатів на Дністровській і чотирьох агрегатів на Ташликській ГАЕС, а також будівництво Канівської ГАЕС.

Введення нових гідроенергетичних потужностей пов'язується з ризиком, детальний аналіз, адекватна оцінка і належне врахування якого при прийнятті рішень являє собою надзвичайно складну задачу [4]. Відповідна задача, обтяжена ризиком, дещо спрощується при виборі оптимальних стратегій реалізації попередньо прийнятих рішень на підставі порівняння альтернатив з метою його мінімізації. Подібні задачі вперше вирішувалися нами при порівнянні варіантів розвитку гідроенергетики в Росії на перспективу до 2030 р. [5] та при виборі оптимального варіанту подальшого розвитку Дніпровського каскаду гідроелектростанцій [6]. При цьому розглядалися різні альтернативи щодо черговості введення нових потужностей гідрогенерації, серед яких в якості оптимальних відбиралися обтяжені меншим сукупним ризиком, включно і ризик невикористаних можливостей, варіанти.

В роботі, що пропонується, розглядається задача формування оптимальної стратегії вводу нових потужностей на ГАЕС в Україні в рамках виконання прийнятої урядом Програми [3]. Задача вирішується на підставі попарного порівняння можливих альтернатив з урахуванням ризику невикористаних можливостей [7], де в якості розрахункових характеристик, за якими формувалися складові сукупних ризиків, приймалися енергетичні показники та вартість.

Література. **1.** Pumped storage plants – Status and perspectives / P. Vennerman, K.H. Gruber, J.U. Haaheim and al. // VGB Power Tech, 2011. No. 4. – P. 32–38. **2.** Розвиток теплоенергетики та гідроенергетики / Є.Т. Базеєв, Б.Д. Білека та ін.; наук. ред. В.М. Клименко та ін., 2013. – 399 с. **3.** Програма розвитку гідроенергетики на період до 2026 року. Схвалено КМУ від 13.07.2016 р. № 552-р. (http://zakon2.rada.gov.ua/laws/show/552-2016-%D1%80#n7). **4.** Векслер А.Б. Надежность, социальная и экологическая безопасность гидротехнических объектов: оценка риска и принятие решений / А.Б. Векслер, Д.А. Ивашинцов, Д.В. Стефанишин. – СПб.: ВНИИГ им. Б.Е. Веденеева, 2002. – 591 с. **5.** Стефанишин Д.В. Вибрані задачі оцінки ризику та прийняття рішень за умов стохастичної невизначеності / Д.В. Стефанишин. – К.: Азимут-Україна, 2009. – 104 с. **6.** Стефанишин Д.В. Про перспективи гідроенергетики в Україні та вибір варіанту розвитку Дніпровського каскаду з врахуванням ризику / Д.В. Стефанишин // Гідроенергетика України. – 2010. – №3. – С. 5–11. **7.** Стефанишина-Гаврилюк Ю.Д. Прийняття рішень у природокористуванні з урахуванням ризику невикористаних можливостей на підставі попарного порівняння альтернатив / Ю.Д. Стефанишина-Гаврилюк, Д.В. Стефанишин // Системні дослідження та інформаційні технології. – 2016. – №3. – С. 51–62.

Стефанишин Д.В., Стефанишина-Гаврилюк Ю.Д.
Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна

Про оптимізацію структури інвестиційного портфеля в природокористуванні на підставі попарного порівняння альтернатив з врахуванням ризику невикористаних можливостей

Важливою вимогою щодо природокористування в сучасних умовах є його раціоналізація на основі задоволення різних цілей та критеріїв: економічних, соціальних, екологічних. При цьому питання соціальної стабільності та збереження якості довкілля при використанні природних ресурсів стали такими ж важливими для суспільства, як і отримувані економічні вигоди.

Відповідна стратегія, направлена на раціоналізацію природокористування, отримала в світі назву концепції «стійкого» (або «сталого») розвитку [1, 2]. Її реалізація передбачає зростання економіки при одночасному поліпшенні екологічної ситуації й вирішенні актуальних соціальних проблем. В цілому, концепція сталого розвитку співзвучна з концепцією «практично досяжного мінімального ризику» в природокористуванні [3,4], адже обидві виходять з того, що природокористування, так чи інакше, обтяжене проблемними ситуаціями і, відповідно, ризиком, який має враховуватися і мінімізуватися – за рахунок економічних можливостей, нових технологій, прийняття більш складних рішень, зокрема тих, що орієнтуються на комплексне використання природних ресурсів, перерозподіл можливостей і, відповідно, на диверсифікацію ризику на нових умовах – соціальної справедливості і збереження довкілля [1–5].

Одним із способів диверсифікації інвестиційного ризику є портфельне інвестування [5], метою якого є зниження сукупного ризику шляхом розподілу наявних активів між кількома напрямками діяльності. При цьому слід враховувати, що особливістю природокористування в сучасних умовах є необхідність включення в структуру інвестиційних портфелів окрім економічної складової, соціальної та екологічної складових [6], і що множина альтернативних портфелів, серед яких має здійснюватися вибір, може бути нескінченною [5].

Запропоновано метод формування скінченної множини ефективних портфелів та вибору оптимальної структури інвестиційного портфеля за критерієм мінімального ризику на підставі попарного порівняння альтернатив з врахуванням ризику невикористаних можливостей. Задача розв'язується на прикладі порівняння за ризиком структур інвестиційних портфелів, які названо базовими, що можуть поєднувати в собі три напрямки інвестування в природокористуванні з врахуванням відношення носія рішення до ризику: несхильності до ризику, байдужості і схильності. Показано, що у випадку трьох напрямків інвестування різними способами з врахуванням ризику невикористаних можливостей формується множина з 16 альтернативних структур ефективних портфелів (альтернатив) в контексті врахування психологічної позиції носія рішення щодо ризику загалом і ризику невикористаних можливостей зокрема, серед яких існує оптимальна на підставі попарного порівняння за критерієм мінімального ризику [3].

Література. **1.** Згуровський М.З. Сталий розвиток у глобальному і регіональному вимірах: аналіз за даними 2005 року / М.З. Згуровський. – К.: Політехніка. – 2006. – 84 с. **2.** Бідюк П.І. Дослідження сталого розвитку макроекономічних процесів в Україні / П.І. Бідюк, В.Я. Данилов, Йошіо Мацуокі // Системні дослідж. та інформ. техн. – 2015. – №3. – С. 84–97. **3.** Стефанишина-Гаврилюк Ю.Д. Прийняття рішень у природокористуванні з урахуванням ризику невикористаних можливостей на підставі попарного порівняння альтернатив / Ю.Д. Стефанишина-Гаврилюк, Д.В. Стефанишин // Системні дослідж. та інформ. техн. – 2016. – №3. – С. 51–62. **4.** Дрозд І.П. Концепція прийнятного ризику та проблеми забезпечення техногенної безпеки в Україні / І.П. Дрозд, А.С. Охота // Еколог. безпека та природокор. Зб. наук. праць. – 2011. – К.: КНУБА, ІТГІП. – Вип. 7. – С. 82–108. **5.** Вітлінський В.В. Економічний ризик: ігрові моделі / В.В. Вітлінський, П.І. Верченко та ін. – К.: КНЕУ, 2002. – 446 с. **6.** Стефанишин Д.В. Метод формування інвестиційного портфеля в природокористуванні з врахуванням ризику невикористаних можливостей / Д.В. Стефанишин, Ю.Д. Стефанишина // Еколог. безпека та природокор. Зб. наук. праць. – Вип. 9. – К.: КНУБА, ІТГІП НАНУ, 2012. – С. 201–212.

Ткаченко В.Ю., Ткаченко Н.В., Жданова О.Г., Сперкач М.О.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Задача мінімізації сумарного запізнення розкладу виконання завдань зі спільним директивним терміном паралельними пристроями

Постановка задачі. Задано множину завдань J, число ідентичних пристроїв m, для кожного $j \in J$ відомо тривалість виконання t_j. Всі завдання мають спільний директивний термін D. Необхідно побудувати розклад σ виконання завдань $j \in J$ на пристроях такий, щоб досягнути мінімуму функціоналу $F(\sigma) = \sum_{j \in J} \max[0; C_j(\sigma) - D]$, де $C_j(\sigma)$ – момент завершення виконання завдання j. Всі завдання множини J надходять на виконання одночасно і обслуговуються без переривань [1]. Задача є NP-складною. Пронумеруємо завдання по неспаданню значень t_j і розіб'ємо J на підмножини $J_1, J_2, \ldots, J_m$, що не перетинаються, де J_i – множина завдань, що обслуговуються i-м пристроєм, $i = \overline{1,m}$. Розіб'ємо множину завдань J_i на підмножини $P_i(\sigma)$, $S_i(\sigma)$, $Q_i(\sigma)$ за наступними ознаками: $P_i(\sigma)$ – множина завдань, що не запізнюються в розкладі пристрою i; $S_i(\sigma)$ – множина завдань, що запізнюються в розкладі пристрою i, для яких виконуються умови: $S_i^H < D$, $C_j > D$, $\forall j \in S_i(\sigma)$ (S_i^H – момент початку виконання завдання j); $Q_i(\sigma)$ – множина завдань, що запізнюються в розкладі пристрою i, для яких виконуються умови: $S_i^H > D$. $P = \bigcup\limits_{i=\overline{1,m}} P_i$; $S = \bigcup\limits_{i=\overline{1,m}} S_i$; $Q = \bigcup\limits_{i=\overline{1,m}} Q_i$. $R_i = D - \sum\limits_{j \in P_i(\sigma)} t_j$ – резерв часу пристрою i в розкладі σ; $\Delta_i = \sum\limits_{j \in P_i(\sigma) \cup S_i(\sigma)} t_j - D$ – запізнення виконання завдання $j \in S_i(\sigma)$ відносно директивного терміну. Позначимо $\mu = |P(\sigma) \cup S(\sigma)|$. Із теорем, сформульованих і доведених в [1, 2], виведена достатня умова оптимальності розкладу: оптимальним є рівномірний розклад $\sigma \in \Psi_P \subseteq \Psi_{PS}$. До класу Ψ_{PS} відносять розклади, в яких до $P(\sigma) \cup S(\sigma)$ входять найкоротші завдання; до класу $\Psi_P \subseteq \Psi_{PS}$ – розклади, в яких до $P(\sigma)$ входять найкоротші завдання. Опис класів розкладів Ψ_{PS}, Ψ_P наведений в [2].

Нехай $\overline{\mu}(\underline{\mu})$ – найбільше (найменше) значення μ, при якому $\sigma \in \Psi_{PS}$:

$$\underline{\mu} = \min\left\{\mu \Big| \sum_{j=1}^{\mu-1} t_j < m \cdot D \leqslant \sum_{j=1}^{\mu} t_j\right\}; \qquad \overline{\mu} = \max\left\{\mu \Big| \sum_{j=1}^{\mu-m} t_j < m \cdot D \leqslant \sum_{j=1}^{\mu} t_j\right\}.$$

Згідно з [1], розклад $\sigma \in \Psi_P$ оптимальний, якщо в ньому завдання з множин $S(\sigma)$, $Q(\sigma)$ рівномірно розподілені між m пристроями, тобто за умови кратності m кількості завдань, що належать множинам $S(\sigma)$ та $Q(\sigma)$. Позначимо $|S(\sigma)| = s$, $|P(\sigma)| = p$. Очевидно, що: $|Q(\sigma)| = n - \mu$; $\mu = p + s$; $0 \leqslant s \leqslant m$. Розклад є рівномірним, якщо $(n - \mu + s) \mod m = 0$. Для кожної конкретної задачі $n, m, \overline{\mu}, \underline{\mu}$ – const. Тому подільність на m числа $n - \mu + s$ повністю залежить від пари μ, s. Лише для таких μ, що $(n - \mu) \mod m = 0$, існують два допустимих значення s: $s = 0$ та $s = m$, які не можуть бути справедливими одночасно для одного розкладу. Для всіх інших значень μ відповідне значення $s = m - (n - \mu) \mod m$.

Під *ідеальним контуром* (ІК) будемо розуміти такий частковий розклад σ^* виконання завдань з множини $P(\sigma)$, за якого можна гарантувати $\sigma \in \Psi_P$ та його рівномірність. Для кожної з пар чисел s та μ, які гарантують рівномірність розкладу, справедливо наступне: кількість пристроїв, для яких $S_i(\sigma) = \varnothing$, дорівнює $m - s$, тобто момент закінчення виконання завдань з множини $P(\sigma)$ на цих $m - s$ пристроях буде збігатися із директивним терміном; для решти s пристроїв $|S_i(\sigma)| = 1$. В рамках розгляду часткового розкладу σ^* можна констатувати наступне: для $m - s$ пристроїв, у яких в повному розкладі $S_i(\sigma) = \varnothing$, в частковому розкладі $\Delta_i = R_i = 0$; для решти s пристроїв у частковому розкладі сумарний резерв $U = m \cdot D - \sum\limits_{j=1}^{p} t_j$. Для того, щоб повний розклад σ належав класу Ψ_P, необхідно гарантувати виконання умови 2 "найменша тривалість виконання роботи $j \in S(\sigma)$ більша за найбільший резерв пристроїв" [2] для розкладів з класу $\sigma \in \Psi_P \subseteq \Psi_{PS}$. Очевидно, якщо виконується $\lceil \frac{U}{s} \rceil \geqslant t_{p+1}$ (де $\lceil \frac{U}{s} \rceil$ – найменше ціле число, при якому $\frac{U}{s} \leqslant \lceil \frac{U}{s} \rceil$), то умова 2 порушується, і повний розклад $\sigma \notin \Psi_P$.

Тобто, для пари чисел s та μ, що розглядаються, не можна побудувати рівномірний розклад $\sigma \in \Psi_P$. Якщо умова 2 не порушується, для задачі можна побудувати ідеальний контур. В такому частковому розкладі σ^* i-й пристрій $(i = \overline{1,s})$ матиме резерв $1 \leqslant R_i \leqslant U - s + 1$, при чому $\sum\limits_{i=1}^{s} R_i = U$. Тепер наведемо підхід до побудови ІК.

Крок 1. Обчислити $\overline{\mu}$ та $\underline{\mu}$.

Крок 2. Знайти такі пари чисел s та μ, для яких виконується $(n - \mu + s) \mod m = 0$. Обчислити $p = \mu - s$.

Крок 3. Для кожної пари s та μ обчислити сумарний резерв $U = m \cdot D - \sum\limits_{j=1}^{p} t_j$.

Крок 4. Побудувати ІК: частковий розклад із кількістю пристроїв $m - s$, що мають нульові резерви, і s пристроїв, сумарний резерв яких $U = m \cdot D - \sum\limits_{j=1}^{p} t_j$.

Крок 5. Сформувати множину $S(\sigma)$ таким чином, щоб виконувалась умова $S^H_{j_k} \leqslant S^H_{j_l}$, якщо $t_{j_k} \leqslant t_{j_l}$, $\forall j_k, j_l \in S(\sigma)$.

Крок 6. Сформувати множину $Q(\sigma)$: Q_i містить ті і тільки ті елементи, котрі відрізняються від $|P \cup S| + i$ на величину, кратну m, $i = \overline{1,m}$.

Сформована множина ІК може бути використана в будь-якому алгоритмі розв'язання задачі. Якщо в процесі роботи алгоритму буде знайдено такий частковий розклад σ^*, який співпадає з ІК, то за побудовою повний розклад σ належатиме класу Ψ_P, буде рівномірним, як наслідок – оптимальним. Отже, існування такого часткового розкладу σ^* є достатньою умовою оптимальності повного розкладу. В рамках описаного підходу кожному обраному $\mu \in [\underline{\mu}; \overline{\mu}]$ для ідеального контуру можна однозначно поставити у відповідність числа $\{p, s, U\}$. Побудуємо дерево, листами якого будуть можливі ідеальні контури (рис. 1), які характеризуються набором значень $\{p, s, U\}$.

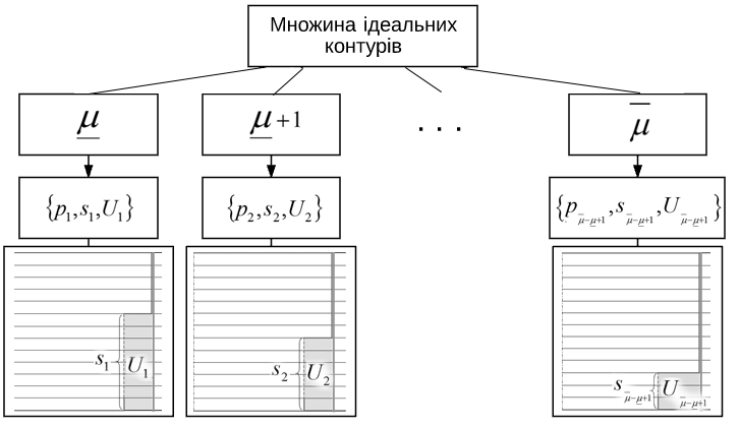


Рис. 1. Дерево ідеальних контурів

Висновок. Задача звелася до побудови часткового розкладу σ^* виконання завдань з множини $P(\sigma)$, який буде найближчим до деякого ІК. В якості алгоритмів для її вирішення можна брати, наприклад, метод табу-пошуку або алгоритм симуляції відпалу. Зауважимо, що наведений підхід можна розширити для задачі із пристроями різної продуктивності завдяки введенню при розрахунках коефіцієнтів продуктивності цих пристроїв.

Література. **1.** Танаев В.С. Введение в теорию расписаний / В.С. Танаев, В.В. Шкурба. – М.: Наука, 1975. – с. 256. **2.** Основы системного анализа и проектирования АСУ: Учебн. пособие. / А.А. Павлов, С.Н. Гриша, В.Н. Томашевский и др. Под общ. ред. А.А. Павлова, К.: Выща шк., – 1991. – 367 с.

Харчук В.В., Войтко С.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Управління фінансовими потоками в кластерних структурах

Дослідженню фінансових потоків, а також ролі і місця фінансового потенціалу кластера у формуванні фінансового потенціалу регіону в даний час приділяється мало уваги. У зв'язку з цим управління функціонуванням кластеру на основі перерозподілу фінансових потоків розглядатися як один із шляхів досягнення всіх основних цілей створення ефективної економіки.

Політика управління фінансовими потоками являє собою сукупність методів, інструментів і прийомів цілеспрямованого впливу з боку фінансових служб підприємства на рух грошових коштів для досягнення їх оптимальної структури з метою підвищення ринкової та економічної доданої вартості підприємства та підвищення його конкурентоспроможності. Основною метою управління фінансовими потоками є забезпечення фінансової рівноваги підприємства в процесі його діяльності і розвитку шляхом балансування обсягів надходження і витрачення фінансових коштів підприємства та формування необхідної звітності, що забезпечить проведення усестороннього аналізу фінансових потоків [1]. Метою аналізу є оцінка здатності підприємства генерувати фінансові потоки в обсягах та за строками здійснення необхідних планових платежів, забезпечення оптимізації фінансового потоку. Місце аналітичних досліджень в управлінні фінансовими потоками можна визначити за етапами цього процесу (рис. 1).

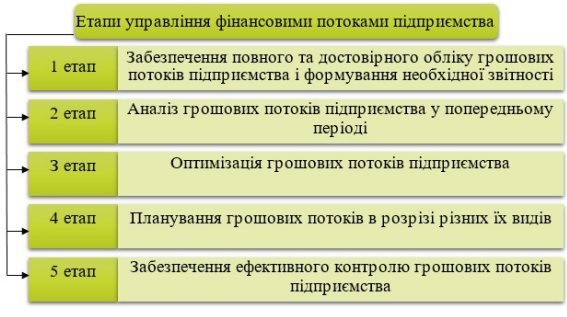


Рис. 1. Основні етапи управління фінансовими потоками підприємства

Поетапне здійснення процесу управління фінансовими потоками підприємства забезпечить постійний моніторинг рівномірності та синхронності формування фінансових потоків у розробці окремих інтервалів часу [2].

Аналіз фінансових потоків підприємства доцільно проводити за такими етапами: 1) вивчення динаміки обсягу формування додатного чистого грошового потоку, 2) дослідження збалансованості фінансових потоків, 3) вивчення синхронності фінансових потоків та оцінювання рівня ефективності фінансових потоків. Дана методика забезпечує комплексне вивчення характеристик фінансових потоків максимізації чистого грошового потоку, зростання якого забезпечує підвищення рівня самофінансування [3].

Підводячи підсумки, ефективно організовані фінансові потоки на підприємстві є передумовою досягнення високих кінцевих результатів його господарської діяльності в цілому.

Література. **1.** Пінчук С.С. Управління та оптимізація грошових потоків на підприємстві як передумови його ефективної діяльності / С.С. Пінчук, О.М. Врадій – Економіка і управління. – 2012. – Вип. 21–22(2). – С. 185–189. **2.** Сологуб С.М. Грошові потоки та їх аналіз на підприємствах залізничного транспорту / С.М. Сологуб, О.А. Стужний – Економіка і управління. – 2012. – Вип. 21–22(1). – С. 167–173. **3.** Лигоненко Л.О. Управління грошовими потоками [Текст] / Л.О. Лигоненко, Г.В. Ситник. Навч. посіб. – К.: Нац. торг. екон. ун-т, 2005. – 255 с.

***Ходневич Я.В.*[1], *Стефанишин Д.В.*[1], *Корбутяк В.М.*[2]**
[1] *Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна;*
[2] *Національний університет водного господарства та природокористування, Рівне, Україна*

Про геоінформаційний аналіз та імітаційне моделювання взаємодії потоку і русла при обтіканні донних гряд з метою прогнозування місцевих розмивів берегів рік

В структурі та організації сучасного управління складними об'єктами та процесами особливу роль відводять математичному моделюванню, за результатами якого проводиться контроль, аналіз явищ і процесів, здійснюються прогнози та приймаються відповідні рішення.

Задачі моделювання, які при цьому постають, важко формалізуються, характеризуються неповнотою, неточністю і суперечливістю даних і знань, великою розмірністю простору рішень, і для їх розв'язання немає чітко заданих алгоритмів. Такі задачі, з різними і не завжди чіткими цілями і критеріями, прийнято називати інтелектуальними. Класичні підходи до їх розв'язання часто призводять до створення занадто складних математичних моделей, які важко забезпечити необхідними вхідними даними, точність яких відповідала б можливостям цих моделей [1]. Тому успіх в розв'язанні таких задач лежить в умінні використовувати різну інформацію, різні підходи та моделі, і, залежно від ситуації, застосовувати різні критерії, щоб приймати обґрунтовані рішення. Одним з таких підходів може бути імітаційне моделювання, яке дозволяє розширити предметну область досліджень.

Прогнозування місцевих розмивів русел рік є однією з таких інтелектуальних задач. Її рішення неможливе без використання знань з різних предметних областей, таких, наприклад, як гідрологія (стік води та наносів), гідравліка (гідродинаміка, рух води у відкритих руслах), механіка ґрунтів, геоморфологія, прикладна математика (рівняння математичної фізики, чисельне і комп'ютерне моделювання), геоінформатика (дистанційне зондування Землі) тощо.

В роботі, що пропонується, задача прогнозування місцевих розмивів дна і берегів рік розв'язується шляхом імітаційного моделювання сил опору дна русла ріки в області течії за природною донною перешкодою (грядою наносів тощо) в межах розв'язання задачі чисельного моделювання турбулентного тривимірного водного потоку, яке ґрунтується на системі рівнянь Рейнольдса [2,3]. Припускається, що сила опору дна в цьому випадку в цілому характеризує умови, за яких інтенсифікується місцевий розмив русла ріки за донною перешкодою, що відхиляється від берега вниз за течією, внаслідок розвитку місцевої турбулентної течії.

Дослідження проводились для умов річок Українських Карпат. Охоплено ділянки з різними типами руслового процесу, кожна з яких описана відповідними гідравлічними критеріями, – незавершене та обмежене меандрування, руслова багаторукавність, вільне меандрування [4]. При проведенні імітаційного моделювання та верифікації його результатів використовувались дані натурних обстежень більш ніж 50-ти ділянок річок. Встановлено, що відповідні їм дані геопросторового аналізу забезпечують достатню для практики точність результатів моделювання, які дозволяють контролювати і прогнозувати перебіг локальних деформацій у природному руслі, виявляти потенційно небезпечні ділянки, де можлива інтенсифікація місцевої руслової і берегової ерозії, в залежності від основних гідроморфологічних характеристик.

Література. **1.** Моисеев Н.Н. Математические задачи системного анализа / Н.Н. Моисеев. – М.: Наука, 1981. – 487 с. **2.** Ходневич Я.В. Методика чисельного моделювання турбулентних течій у відривних зонах при оцінці розмиваючої здатності потоку гірських рік / Я.В. Ходневич, О.Є. Щодро, В.М. Корбутяк // Зб. наук. пр. «Гідромеліорація та гідротехнічне будівництво». – Рівне: НУВГП, 2008. – Вип. 33. – С. 81–86. **3.** Khodnevich Y.V. Mathematical modeling the conditions of intensification of the riverbed local erosion behind of obstacle that deviates from the shore downstream / Y.V. Khodnevich, D.V. Stefanyshyn // Zesz. Nauk. Inżynieria Lądowa i Wodna w Kształtowaniu Środowiska. Nr 10. – Kalisz, 2014. – S. 7–18. **4.** Корбутяк В.М. Гідроморфологічні аспекти пропуску паводкових витрат води гірськими руслами / В.М. Корбутяк // Вісник НУВГП. – Рівне. – 2007. – Вип.3 (39), Ч. 2. – С. 92–101.

Шемеляк Р.О., Бідюк П.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Системний підхід до моделювання фінансово-економічних процесів

Готові до використання моделі фінансово-економічних процесів (ФЕП) зустрічаються дуже рідко. Навіть існуючі апробовані моделі потребують корегування їх структури та/або параметрів з метою їх адаптування до конкретних даних і зовнішніх випадкових збурень. Тому у більшості випадків необхідно будувати нову модель на основі наявних статистичних даних. А побудова нової адекватної моделі не може виконуватись без застосування системного підходу. Структурна схема такого системного підходу до моделювання ФЕП зображена на рис. 1.

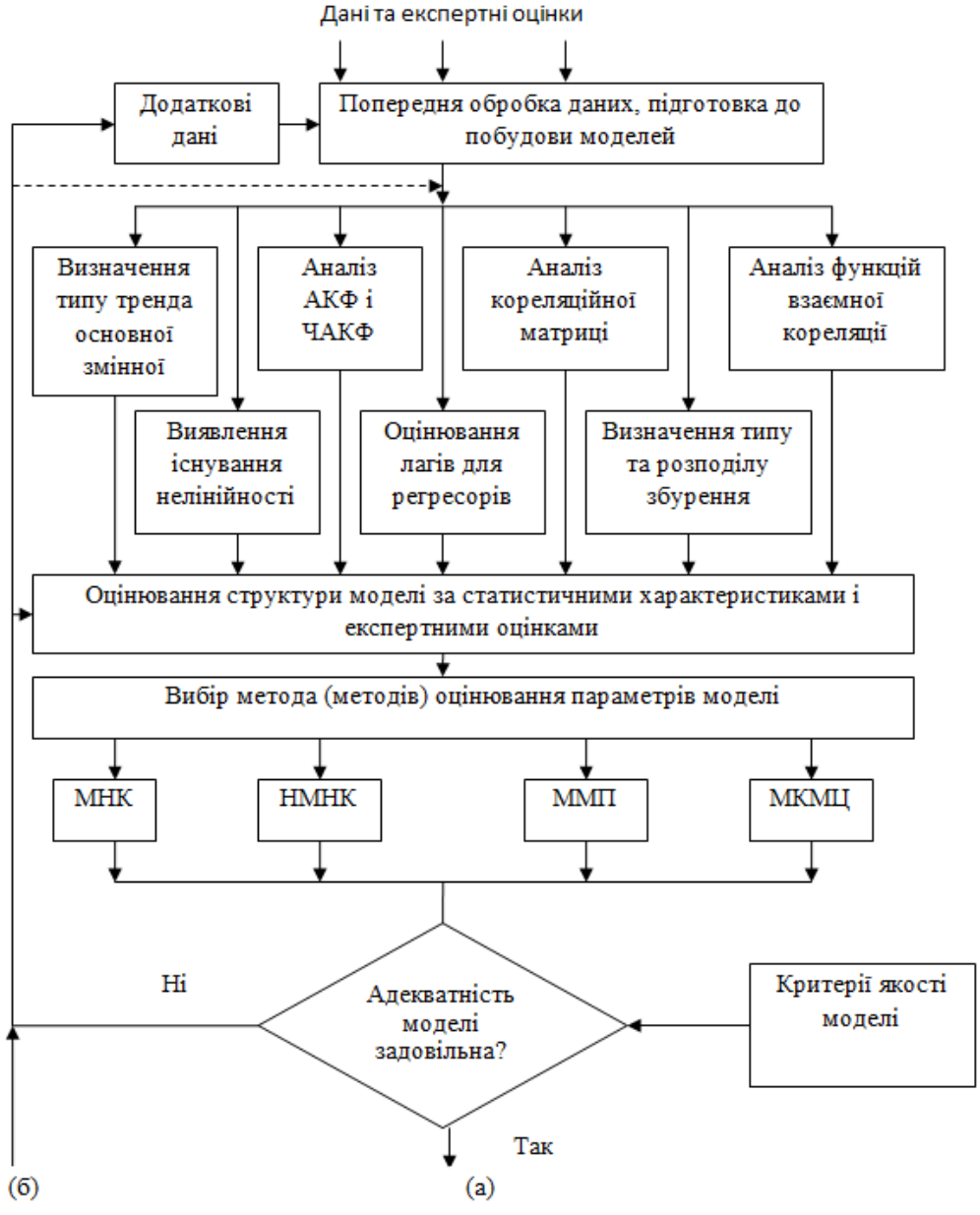


Рис. 1. Структурна схема системного підходу до моделювання

Попередня обробка даних необхідна для приведення їх до форми, яка забезпечить можливості коректного застосування методів оцінювання параметрів моделі та отримання їх статистично значущих оцінок. На основі коректно підготовлених даних оцінюють структури та параметри математичних моделей-кандидатів процесів, вибраних для прогнозування та керування. Вибір (оцінювання) структури моделі – ключовий момент її побудови. Нагадаємо, що структура моделі включає п'ять елементів: (1) розмірність (кількість рівнянь, які утворюють модель); (2) порядок – максимальний порядок диференціальних або різницевих рівнянь моделі; (3) нелінійність та її тип (нелінійності стосовно змінних або параметрів); (4) час затримки (лаг) реакції на виході відносно входу та його оцінка; (5) зовнішнє збурення процесу та його тип (детерміноване або випадкове, тип розподілу). Як правило, для одного процесу оцінюють кілька моделей-кандидатів, а потім вибирають з них кращу за допомогою множини статистичних параметрів якості моделі. Для вибору моделі рекомендують використовувати не менше трьох статистичних критеріїв якості: коефіцієнт детермінації, суму квадратів похибок моделі і статистику Дарбіна-Уотсона [1, 2]. Для оцінювання якості прогнозів також рекомендують використовувати не менше трьох статистичних критеріїв якості, а саме: середньоквадратичну похибку, середню абсолютну похибку у процентах і коефіцієнт Тейла (Theil).

З метою забезпечення адекватності моделі в умовах моделювання нестаціонарних процесів запропонований підхід передбачає застосування структурно-параметричної адаптації моделей до нових даних. Структурна адаптація передбачає повторне обчислення статистичних характеристик, необхідних для уточнення структури моделі, а параметрична забезпечує повторне оцінювання параметрів моделі альтернативними методами. Оскільки моделі можуть бути лінійними, псевдолінійними (моделі, нелінійні за змінними) і нелінійними стосовно параметрів, то відповідна система підтримки прийняття рішень для моделювання і прогнозування містить такі методи оцінювання параметрів: звичайний метод найменших квадратів (МНК), метод максимальної правдоподібності (ММП), нелінійний МНК і метод Монте-Карло для марковських ланцюгів (МКМЛ). Останній метод представляє собою ітераційну процедуру Монте-Карло, яка адаптується до оцінювання параметрів моделей. Перевага МКМЛ полягає у тому, що він є фактично універсальним стосовно можливих структур моделей і типів зовнішніх випадкових збурень, але його основний недолік полягає у тому, що час, необхідний для отримання високоякісних оцінок параметрів, може бути неприйнятним для використання цього методу у реальному часі.

Системний підхід до моделювання і прогнозування процесів довільної природи передбачає можливість виявлення та врахування невизначеностей, які зустрічаються у процесі побудови математичних моделей, оцінювання прогнозів та керування. Так, невизначеності даних проявляються у вигляді пропусків вимірів, наявності екстремальних значень, шумів вимірів та коротких вибірок. Невизначеності структури зумовлені неможливістю точно оцінити порядок моделі, час запізнення по входу і точно встановити належність випадкового збурення до конкретного типу розподілу ймовірностей. Параметричні невизначеності зумовлені некоректним вибором методу оцінювання параметрів моделі, методичними похибками, високим рівнем шумів вимірів, а також короткими вибірками даних. Вплив невизначеностей на результати моделювання і прогнозування можна суттєво зменшити завдяки застосуванню відповідних методів боротьби з ними, а саме: методів заповнення пропусків, обробки екстремальних значень, цифрової та оптимальної фільтрації, розмноження значень вибірок. Вплив структурних і параметричних невизначеностей суттєво зменшується завдяки використанню структурно-параметричної адаптації математичних моделей до даних та перемінних режимів функціонування об'єктів, що описуються математичними моделями.

Література. **1.** Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів / навчальний посібник. – Київ: НТУУ «КПІ», 2010. – С. 34–53, 108–123. **2.** Tsay R.S. Analysis of Financial Time Series / R.S. Tsay. – Hoboken (New Jersey): John Wiley and Sons Inc., 2010. – 715 p.

This page was left blank intentionally

Plenary talks

System analysis of complex systems of various nature

Intelligent systems for decision-making

High Performance Computing and Microsystems Engineering

Progressive information technologies

2 Intelligent systems for decision-making

Section 2

Intelligent systems for decision-making

1. Intelligent decision-making systems (IDMS) in finance-economical area (micro- and macro economical systems, banks, stock exchanges, insurance companies, etc.).
2. Decision-making systems in social processes' management.
3. Decision-making systems in technological processes' management in industry.
4. Intellectual analysis of data and knowledge; problems of data and knowledge mining, knowledge bases for IDMS.
5. Mathematical modelling and forecasting of complex objects and processes.
6. Decision-making under the data uncertainty conditions (systems of fuzzy logical deduction).
7. Modern methods and algorithms for IDMS (genetic and evolutional algorithms, neural networks, etc.).

Секция 2

Интеллектуальные системы принятия решений

1. Интеллектуальные системы принятия решений (ИСПР) в финансово-экономической сфере (микро- и макроэкономические системы, банки, биржи, страховые компании и т.п.).
2. Системы принятия решений в управлении социальными процессами.
3. Системы принятия решений в управлении технологическими процессами в промышленности.
4. Интеллектуальный анализ данных и знаний; проблемы добывания данных и знаний (Data&Knowledge Mining), базы знаний для ИСПР.
5. Математическое моделирование и прогнозирование сложных объектов и процессов.
6. Принятие решений в условиях неопределенности данных (системы нечеткого логического вывода).
7. Современные методы и алгоритмы ИСПР (генетические и эволюционные алгоритмы, нейронные сети и т.п.).

Секція 2

Інтелектуальні системи прийняття рішень

1. Інтелектуальні системи прийняття рішень (ІСПР) в фінансово-економічній сфері (мікро- та макроекономічні системи, банки, біржі, страхові компанії і т.д.).
2. Системи прийняття рішень в управлінні соціальними процесами.
3. Системи прийняття рішень в управлінні технологічними процесами в промисловості.
4. Інтелектуальний аналіз даних і знань; проблеми добування даних і знань (Data&Knowledge Mining), бази знань для ІСПР.
5. Математичне моделювання та прогнозування складних об'єктів і процесів.
6. Прийняття рішень в умовах невизначеності даних (системи нечіткого логічного виведення).
7. Сучасні методи й алгоритми ІСПР (генетичні та еволюційні алгоритми, нейронні мережі і т.д.).

Iievlieva S.M., Garev Y.V., Pikhota S.R.
Kharkiv National University of Radioelectronics, Kharkiv, Ukraine

Development of intellectual audio-player based on spectral sound analysis

In this work in a first time was proposed a way of development of intellectual audio-player used for ranging based on a sound spectral analysis.

Nowadays the issues devoted to the optimization of searching the information in the internet are dramatically actual. Big companies spend huge budgets on research and development of algorithms which are able to suggest their users a relevant advertisement. One of the most trending steam for today is smart things. Smart house, smartphone, smart car etc. Thus the issue of searching a soundtrack which is suit to user's preferences flawless is also a part of modern trendo.

Formulation of the problem. The analysis of internet resources has shown that the new way of user's preferences predicting is able to improve a quality and unlock the potential of such audio-services as Spotify, SoundCloud, Apple and Play Music and the long list of similar services. At this moment the most popular way of solving this problem is a multidimensional statistical analysis of user preferences and personal data to build a preferences profile. Everything mentioned above determine the relevance of selected development subject.

Usually to solve an issue connected with a process of building a preferences profile, the Neural Networks (NN) are being used. Once should know that a NN is able to provide its own ways of solving an assigned task after a certain period of time needed to train the neurons. There is no need to work with training set after the process of the NN training is finished, thus the experimental data set can be used as well. Furthermore the accuracy will be the higher the bigger data set of highly correlated values would be provided on the NN's input.

Numerous modern systems based only upon a statistics analysis processing the data about committing selections, their repetition frequency, refusals or adding to personal lists of preferences (playlists) etc.

Due to this, we can conclude that user's experience is one of the most important part of any service. The general purpose of given research and development is providing people with the new user's experience of using the sound-services.

More detailed research of sound-services' UX question shows that suggesting the new playlists is a very popular feature. On a reasonable question "Why?" there is a simple answer. In a modern stream of data it is almost impossible for usual user to find completely new song that will fit to all his requirements

This way we are coming to an issue of recognizing user's preferences.

To achieve a high level of UX it is essential to follow the way people use a service and find points where it can be improved.

In a given case the consideration should start with the influence of the season, weather, weekday, daytime and the rest of other similar indicators on the decision making process. Obviously these factors could completely change person's preferences! Thus it is noticeable that taking into account the additional statistics which is also very influential at UX the quality of user's preferences predicting is raised.

Consider a person who wants to listen to a rock-n-roll music and turns on a random song. User possibly can estimate a random song as a "too hard for listening" or maybe as a "rhythmically too fast". That means we should suggest him something something less hard and more slow.

Talking mathematically we should find another combination of frequencies in our songs bank which is brings hardness and rhythm to a lower level

This consideration can lead us to a new valuable feature – ranging songs by their spectral characteristics.

Any song consists of the nine main frequency ranges: Low-low frequencies, low-middle frequencies, low-high frequencies, middle-low frequencies and so on the same way till the high-high frequencies. Each frequency has its own interval (e.g. low-low frequencies lay between 8–10 Hz).

To define the difference between one song and another we can write the frequency's values of each range at the each moment of the song to our sound bank. After that any song can be described as a set of different frequencies with their repetition frequency. Then using the apparatus of data analysis, particularly data clustering songs can be ranging and dividing into a few parts. Thus if user would like to change a distortion level he will be switched over the next range and the random song from it would be suggested.

Why is that needed and how it influences the UX?

Here is an example. Let's assume that the user listens to the wide variety of music, starting with the simple acoustic one and up to the progressive death metal. It's obvious that there is a huge difference in the way the person perceives those two genres. Thus it is important to understand how users form his preferences at least verbally.

Obviously, he wouldn't be classifying or ranging song in any sense, instead, he will formulate his query as a phrase, such as: "I would like to listen to some light calm jazz". From the mathematical point of view this problem belongs to the decision making on the fuzzy sets, as the terms "light" and "calm" could differ from user to user. Research being conducted is to introduce a ranking system that will allow to personalise the ranking of "hardness", tempo and other song parameters using Neural Networks. Furthermore, having this data it will be possible to define user preferences more accurately. Users could formulise their queries differently, but the task assigned to the system is to guess the possible choice as accurately as possible. So the query "Turn on the music" doesn't provide any particular specifications, whilst "Play Sad But True by Metallica from the Black Album" is a direct request for playing a specific song. Thereby, user queries should be processed by the different modules. The first one will process the queries containing the elements of fuzzy sets. Such queries are: emotional preferences ("something funny"), "hardness" or tempo preferences ("something harder", "something slower"). The other module will be processing specific requests. For example, search for the particular song, particular artist, request for the specific genre. To teach the system, the separate module will process the feedback received from user after him got the suggested song from the system. During the feedback processing the evaluation inquiries will be used: "Next", "Perfect", "Misfire" and so on. To make user's attitude to the particular song more detailed, there will also be an optional 5-point scale to evaluate the result received. These days the key point is to develop a module for processing the fuzzy preferences and "hardness"-ranking.

Conclusion. The new approach to building the intellectual audio player is introduced in this application. It will allow the musical services to process user queries more qualitatively and give more precise answers.

References. **1.** habrahabr.net/thread/1565. **2.** en.wikipedia.org/wiki/Fourier_series.

Kuksa M.U. [1] ***Norkin V.I.*** [1,2]
[1] *Igor Sikorsky Kyiv Polytechnic Institute, Faculty of Applied Mathematics, Kyiv, Ukraine;* [2] *V.M. Glushkov Institute of Cybernetics of NAS of Ukraine, Kyiv, Ukraine*

Numerical calculation of a distribution for the asset price minimum under the binomial asset price model

Binomial model of asset price volatility is a tree of asset price evolution scenarios, where each branch is subdivided into two subbranches with some fixed returns a, b and corresponding probabilities, $\pi, 1-\pi$ [1, 2].

Denote ξ a random return of the asset. Then the average return equals $E\xi = a\pi + b(1-\pi)$. If $E\xi < 0$, then buying such asset does not make sense, and if it is already in the portfolio, it is advisable to sell it. If $E\xi > 0$ and $\xi \geqslant 0$ almost sure, such asset is attractive for investors. But if $E\xi > 0$ and $\xi < 0$ with positive probability, then decision process regarding investment in the asset is more complicated. In such a case it is useful to know how likely the asset price may fall below some specified threshold over time.

Let us calculate the distribution function of the minimal asset price $\min_{0<\tau\leqslant t} x^\tau$ by time t, i.e. probability of the event that the asset price falls down below some positive threshold x during time t. If $\xi^1, \xi^2, ...\xi^t$ is a sequence of asset returns over time, then the asset price at time t is easily calculated: $x^t = x^0(1+\xi^1)(1+\xi^2)...(1+\xi^t)$. So we need finding the probability $\psi(x,t) = Pr\{\min_{0<\tau\leqslant t} x^0(1+\xi^1)(1+\xi^2)...(1+\xi^\tau) < x\} = \psi(x,t) = Pr\{\min_{0<\tau\leqslant t}(x^0/x)(1+\xi^1)(1+\xi^2)...(1+\xi^\tau) < 1\}$.

Denote an auxiliary stochastic process: $X^\tau = u(1+\xi^1)(1+\xi^2)...(1+\xi^\tau)$, where $\tau = 1, 2, ...t; X^0 = u$.

If we find the probability $\varphi(u,t) = Pr\{\min_{0<\tau\leqslant t} X^\tau \geqslant 1 \forall t | X^0 = u\}$, then we can obtain $\psi(x,t) = 1 - \varphi(x^0/x, t)$.

By definition, $\varphi(u<1,t) = 0$ and $\psi(x \geqslant x^0, t) = 1$. Probabilities $\varphi(u,\tau), \tau = 1,2,3,...$, satisfy the following recurrent equations:

$\varphi(u,\tau) = E\varphi(u(1+\xi),\tau-1) = \pi\varphi(u(1+a),\tau-1) + (1-\pi)\varphi(u(1+b),\tau-1)$, $0 \leqslant u \leqslant (x^0/x)(1+\max\{a,b\})^{t+1}$, $\tau = 1,2,...,t$;

$$\varphi(u,0) = \begin{cases} 0, & u < 1, \\ 1, & u \geqslant 1, \end{cases} \qquad 0 \leqslant u \leqslant (x^0/x)(1+\max\{a,b\})^{t+1}.$$

Functions $\varphi(u,t), \tau = 1,2,...,$ can be calculated step by step using the above recurrence relation. As result we obtain values $\varphi(u,t)$ for $u \in [0, x^0/x]$.

Example. Suppose that the initial asset price $x^0 = 100$. Also let $a = 0,2$; $b = -0,1$ and $\pi = 0,3$.

Table 1. Probability of asset price falling below level x during time t

	$x = 40$	$x = 60$	$x = 80$	$x = 100$
$t = 5$	0	0,1681	0,5282	0,5282
$t = 10$	0,0282	0,1493	0,3828	0,6496
$t = 15$	0,1268	0,2969	0,5155	0,7216
$t = 20$	0,2375	0,1464	0,6080	0,7723

A similar approach can be applied to calculating the probability that the asset price exceeds level $x \geqslant x^0$ by time t, $\chi(x,t) = Pr\{\max_{0<\tau\leqslant t} x^0(1+\xi^1)(1+\xi^2)...(1+\xi^\tau) > x\}$. Remark also that the above numerical method is applicable not only to the binomial model but to any one step discrete distribution asset price model.

References. **1.** Cox J.C., Ross R.A., Rubinstein M. Options pricing: a simplified approach. J. of Financial Economics. 1976. V. 7. P. 229–263. **2.** Anthony M., Biggs N. Mathematics for Economics and Finance. Methods and Modelling. – Cambridge: University of Cambridge, 1996.

Lazarenko D.V., Lyunda V.V.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Facial Landamark Detection

Facial landmark detection is a fundamental component in many face analysis tasks, such as facial attribute inference [1], face verification [2,3], and face recognition [4]. Facial landmark detection is traditionally approached as a single and independent problem. Popular approaches include template fitting approaches [5,6] and regression-based methods [7].

In this paper we will describe Active Appearance Model (AAM) a widely used regression approach. It was proposed by T.F. Cootes, G.J. Edwards and C.J. Taylor in 1998. Originally active external models were used to estimate the parameters of face images, but later they were adapted by other fields, in particular by medicine for the analysis of X-ray images and images obtained by magnetic resonance imaging.

How appearance model are generated. The models are generated by combining a model of shape variation with a model of the appearance variations in a shape-normalised frame. It is necessary to have a training set of labelled images, where key landmark points are marked on each example object. Image marking is a manual or semi-automatic process. Each label has its own number and determines the characteristics of a point we want to find using the model while adapting to a new image. The procedure of education starts with normalization, in order to compensate the difference in a scale, incline and offset. After this, all sets are aligned in a common coordinate frame and are represented by a vector x. Then a principal component analysis (PCA) is applied to the data.

$$x = \overline{x} + P_s b_s, \tag{1}$$

where $\overline{x}$ is the mean shape, P_s is a set of orthogonal modes of variation and b_s is a set of shape parameters. Changing the parameter vector b_s we can receive various kinds of shape deformations to adjust its under real image. Below we can see an example of this shape.

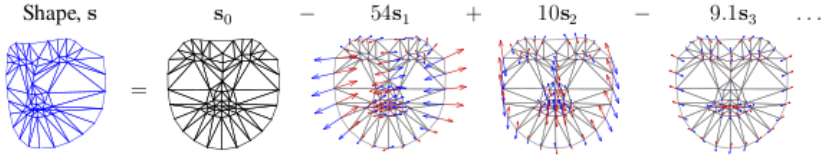


Figure 1. Directions of principal components

By applying PCA to the normalised data we obtain a linear model:

$$g = \overline{g} + P_g b_g, \tag{2}$$

where $\overline{g}$ is the mean normalised grey-level vector, P_g is a set of orthogonal modes of variation and b_g is a set of grey-level parameters.

Since there may be correlations between the shape and grey-level variations, PCA is applied to the data follows. We generate the concatenated vector.

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \overline{x}) \\ P_s^T (g - \overline{g}) \end{pmatrix}, \tag{3}$$

where W_s is a diagonal matrix of weights for each shape parameter, allowing for difference in units between the shape and grey models. We apply PCA on these vectors, giving a futher model

$$b = Qc, \tag{4}$$

where Q are the eigenvectors and c is a vector of appearance parameters controlling both the shape and grey-levels of the model.

$$x = \overline{x} + P_s W_s Q_s c, \tag{5}$$

$$g = \overline{g} + P_g Q_g c, \tag{6}$$

where

$$Q = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix}. \tag{7}$$

Implementation. Most models are trained to detect mouth, nose, eyes, eyebrows. In this case we lose a lot of information about the skin on the forehead. In many situations we need to detect as much facial skin as possible. The tricky moment is detect the hairline. To detect landmarks of the forehead we use empiric formulas to get the point of the upper limits of the skull and then use skin detection to cut the area above the hairline and correct existing points. The model was trained on a set of 800 images of face, each labelled with 70 points around the main features. Figure 2 shows the results.

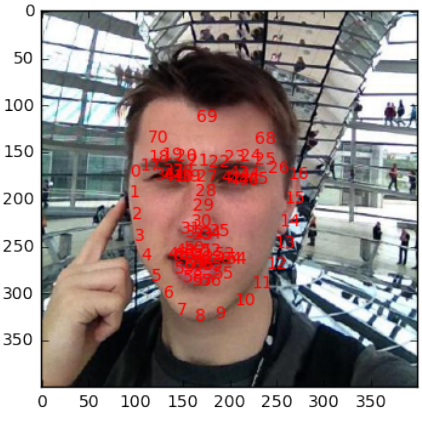


Figure 2. Results of the model

Conclusion. The AAM algorithm is is a very effective for landmarking deformable objects and has many applications. For example, results of the work can be widely used in medical sphere: to the human eye, no matter how long and hard you stare at your wrist or someone else's face, you would struggle to detect any change in color. For a computer, however, the tiniest per-pixel fluctuations are easy to detect. As our heart pumps blood around body, arteries swell with bright red blood, which changes the color of skin slightly. Thus we can measure accurately our heart rate.

Future work will explore deep multi-task learning for dense landmark detection and other vision domains.

References. **1.** Luo, P., Wang, X., Tang, X.: A deep sum-product architecture for robust facial attributes analysis. In: CVPR. pp. 2864–2871 (2013). **2.** Lu, C., Tang, X.: Surpassing human-level face verification performance on LFW with GaussianFace. Tech. rep., arXiv:1404.3840 (2014). **3.** Xiong, X., De La Torre, F.: Supervised descent method and its applications to face alignment. In: CVPR. pp. 532–539 (2013). **4.** Zhu, Z., Luo, P., Wang, X., Tang, X.: Deep learning multi-view representation for face recognition. Tech. rep., arXiv: 1406.6947 (2014). **5.** Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. PAMI 23(6), 681–685 (2001). **6.** Zhu, X., Ramanan, D.: Face detection, pose estimation, and landmark localization in the wild. In: CVPR. pp. 2879–2886 (2012). **7.** Cao, X., Wei, Y., Wen, F., Sun, J.: Face alignment by explicit shape regression. In: CVPR. pp. 2887–2894 (2012). **8.** T.F. Cootes, G.J. Edwards, C.J. Taylor Wolfson: Active Appearance Models. pp. 5–6 (1998).

Naderan M., Zaychenko Y.P.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Using neural network anfis and cascade for predicting futures contract prices

The article describes the fuzzy neural network ANFIS and Cascade for forecasting prices on futures Coffee Historical Prices / ICE futures. For the training of neural networks chosen daily futures contract price Coffee / ICE futures for 5 years.

Introduction. Currently, fuzzy neural network (FNN) are increasingly used in predicting problems in the financial field. Their distinctive features are: the possibility to predict nonlinear non-stationary processes, flexibility, a wide arsenal of learning algorithms of network settings, the possibility of expert information accounting during formulization of the fuzzy rule base of conclusion, the use as quantitative as qualitative information [1, 2]. Also, by the absence of necessity for training parameters, a membership function is an advantage of Cascade neo-fuzzy. Only linear weight need to be trained in Cascade NN [3].

Fuzzy systems. Common sense, human thinking and judgment are the lures of fuzzy logic. What after all does fuzzy logic bring to the party is a question that has been hotly debated upon during the past several years, and probably will continue to remain controversial in the near future. The Neuro Fuzzy architecture is shown below. The circular nodes represent nodes that are fixed whereas the square nodes are nodes that have parameters to be learnt.

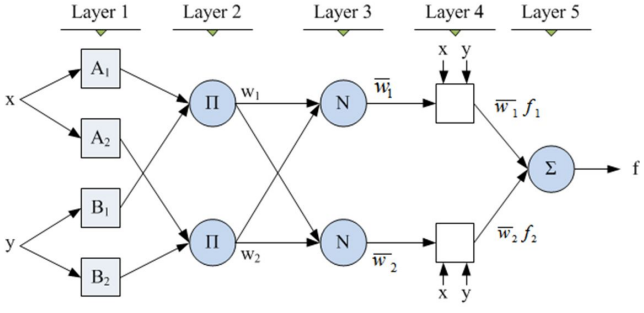


Figure 1. Neuro Fuzzy architecture

The primary reason for this is that the ultimate universal proof of why a fuzzy logic solution is 'better' than a conventional one, for whatever reason, does not exist. First of all, this is because fuzzy logic does not replace conventional control techniques. Instead, it renders possible new solutions that were not feasible to implement earlier. Hence a direct comparison between a fuzzy logic solution and a conventional one rarely exists. Even when a conventional solution for comparison exists, the advantages of the fuzzy logic solution depend very much on the type of the application. Thus, it is impossible to quantify the 'benefits in general' of using fuzzy logic [4].

The work is dedicated to predict the prices for futures contracts for coffee with neural networks ANFIS and Cascade. It is executed in this study a comparative analysis of application of ANFIS and Cascade for the assigned task, and also it is introduced the networks architecture. In the most studies in the field of forecasting of prices for shares that currently exist, are given the prices for the entrance to the neural network. In this work, informative signs, that are applied for the entrance of the neural network, are invariant to the price volatility.

Conclusion. Thus, the considered system allows effectively solve the problem of prices prediction for the futures contracts, as it does not depend on the price volatility.

References. **1.** Zaichenko Yurii Petrovich. Research of the cascading neo-fuzzy of neural networks in predicting problems in the financial field. **2.** Naderan Edris. International Journal "Information Content and Processing", Volume 1, Number 3, 2014. **3.** Naderan Svetlana. International Journal "Information Content and Processing", Volume 1, Number 2, 2014. **4.** Hemanth Kumar P., Prashanth K.B., Nirmala T.V., S. Basavaraj patil. Neuro Fuzzy based Techniques for Predicting Stock Trends. IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012.

Solodrai I.I., Marchenko O.I.
Igor Sikorsky Kyiv Polytechnic Institute, Faculty of Applied Mathematics, Kyiv, Ukraine

Technique for combining neuroevolution and temporal-difference learning

In this paper we propose a technique of combining evolutionary approach with temproal-difference learning, one of classical methods for solving on-line reinforcement learning task.

Traditionally the evolutionary approach for solving reinforcement learning task considers the search in space of agent behaviours by some specific evolutionary algorithm [1]. The key disadvantage of such approach is in fact that the space of behaviours is infinite. Moreover, a simulation of evolution does not take into account environment interaction experience of individual agents. This experience potentially can be used to improve agent's policy during its lifetime and as a result to improve average policy of population.

On the other hand, temporal-difference learning (and many other reinforcement learning methods) does not perform search in space of policies. Instead it continuously approximates a current policy value-function to real values [2].

$$\hat{v}_\pi(s) \approx v_\pi(s) \equiv \sum_{k=0}^{\infty} \gamma^k R_{k+1}. \tag{1}$$

Then it improves the policy using this approximate function. The problem here lies in representation of approximate $\hat{v}_\pi(s)$. The representation depends only on the intuition of the designer and can be far from optimal. We propose a technique that aims to avoid mentioned disadvantages by combining these two approaches.

Combining technique. One can differently interpret outputs of neural network that represents policy π. It may be the probability of taking particular action, or its value. The value of $\hat{v}_\pi(s)$ by definition means *value of state*. But one can use it to determine the best action in the current state:

$$a' = \pi(s) = \text{argmax}_a \left(\sum_{s'} \left[P_{sa}(s') * v_\pi(s') \right] \right). \tag{2}$$

So, the value-greedy policy just takes action that will lead to the state s with the maximum value of $\hat{v}_\pi(s)$. However, to be able to do that, it's necessary to have a way for estimating a next state, which is not always possible.

If we apply the temporal-difference update rule to a neural network using backpropagation algorithm, it will output $\hat{v}_\pi(s)$. To avoid estimation of a next state we can adjust learning rule so that neural network would output $\hat{v}_\pi(s) \to \left(A \to \mathbb{R}\right)$ instead of $\hat{v}_\pi(s) \to \mathbb{R}$. Then the value of ith output of the network will mean the expectation of $\hat{v}_\pi(s')$ if action a_i was taken.

$$N_i(s) = \sum_{s'} \left[P_{sa_i}(s') * v_\pi(s') \right]. \tag{3}$$

By applying periodical evolution of such neural networks, we get a system that performs search of agents that are *able to learn better* through interaction with environment.

Conclusions. The idea of combining evolutionary approach with other learning methods for solving reinforcement learning problems has already shown promising results [3]. We expect proposed technique to be effective for tasks with huge state space and especially for tasks with dynamic environment. Improvements may be obtained by applying evolutionary process not only to representation of value-function, but also to various parameters of learning.

References. **1.** K.O. Stanley, R. Miikkulainen. Evolving neural networks through augmenting topologies. Evolutionary Computation, 10(2):99–127, 2002. **2.** R. Sutton. Learning to predict by the methods of temporal differences. Machine Learning, 3:9–44, 1988. **3.** Sh. Whiteson. Evolutionary Function Approximation for Reinforcement Learning. Journal of Machine Learning Research, 7:877–917, 2006.

***Zaychenko Y.P.*[1], *Zaychenko E.Yu.*[1], *Ovi Nafas Aghaei Agh Ghamish*[2]**
[1]*Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine;* [2]*Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine*

Survivability Analysis of New Generation Computer Networks (NGN)

Burst growth in amount of Internet users, and ever-increasing demands for quality of service in computer networks has led to appearance of new class of networks – new generation networks (NGN). NGN have some specific features: a) high speed of data transmission; b) different classes of service (CoS); c) high demands on Quality of Service (QoS): packets transfer delay (PTD), packets delay variance (PDV) and packets loss ratio (PLR). Each type of flows (CoS) sets corresponding demands to the values of QoS. NG computer networks are based on modern communication technology – MPLS and GMPLS (Generalised MPLS) [1].

The appearance of NG computer networks stimulated the acute necessity of development of corresponding models and methods for analysis and design of NGN. One of the important problems connected with NGN networks design is the development of models and algorithms for survivability analysis of New Generation Computer Networks.

Earlier the problem of survivability analysis and optimization for ATM networks was investigated in [2] and for networks with MPLS technology in [3]. In order to perform such analysis it's necessary to introduce adequate survivability indices which would take into account the peculiarities of this technology and algorithm for its estimation to develop. The goal of the present paper is to develop and investigate the model and method of NGN survivability analysis.

Let's consider the survivability analysis problem statement.

Let it be NGN network which is defined by oriented graph $G = \{X, E\}$ where $X = \{x_j\}$ is a set of network nodes, $E = \{(r, s)\}$ is a set of channels, μ_{rs} are channels capacities.

Assume that in network K classes of service (CoS) are transmitted due to so-called demand matrices $H(k) = ||h_{ij}(k)||_{i=\overrightarrow{1,N},j=\overrightarrow{1,N}}$, (Mbits per sec). For each flow class CoS the corresponding QoS are introduced as a given value mean delay time $T_{m,k}$ estimated by the following expression [3]:

$$T_{m,k} = \frac{1}{H_{\Sigma}^{(k)}} \sum_{(r,s)\in E} \frac{f_{rs}^{(k)} \sum_{l=1}^{k} f_{rs}^{(l)}}{\left(\mu_{rs} - \sum_{l=1}^{k-1} f_{rs}^{(l)}\right) \cdot \left(\mu_{rs} - \sum_{l=1}^{k} f_{rs}^{(l)}\right)} \tag{1}$$

where $H_{\Sigma}^{(k)} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{(k)}$ is a total intensity of input flow K-th CoS, μ_{rs} – is channel (r, s) capacity, $f_{rs}^{(k)}$ is a K-th class of flow value in the channel (r, s).

It's necessary to determine the survivability factors for a given network.

In paper [3] the following complex factor was suggested for survivability analysis of MPLS networks:

$$P\left\{H_{\Sigma}^{\Phi}(1) \leqslant r\%H_{\Sigma}^{0}(1)\right\}, P\left\{H_{\Sigma}^{\Phi}(2) \leqslant r\%H_{\Sigma}^{0}(2)\right\}, ..., P\left\{H_{\Sigma}^{\Phi}(k) \leqslant r\%H_{\Sigma}^{0}(k)\right\} \tag{2}$$

where $H_{\Sigma}^{0}(k)$ is k-th class flow value in the faultless state; $H_{\Sigma}^{\Phi}(k)$ is real flow value of class k in case of failures, $r = (50 \div 100)$, $k = \overrightarrow{1,k}$, $P\left\{H_{\Sigma}^{\Phi}(k) \leqslant r\%H_{\Sigma}^{0}(k)\right\}$ is the probability that flow value of k-th class transmitted in a network would be not less than fraction r of nominal flow value in the faultless state $H_{\Sigma}^{0}(k)$.

The algorithm for estimation and analysis of complex survivability factor (2) for NGN was developed and investigated. The experimental investigations of the developed algorithm have been carried out which are presented in report and perspectives of its application for NGN networks analysis are discussed.

References. **1.** Олвейн Вивьен. Структура и реализация современной технологии MPLS. Изд. Дом «Вильямс», 2004. – 480 с. **2.** Е.Ю. Зайченко. Сети АТМ: моделирование, анализ и оптимизация. – Киев, 2003. – 216 с. **3.** Е.Ю. Зайченко, Ю.П. Зайченко. Сети с технологией MPLS: моделирование, анализ и оптимизация. – К.: НТУУ «КПИ», 2008. – 240 с.

Zorin Yu.
Igor Sikorsky Kyiv Polytechnic Institute, Faculty of Applied Mathematics, Kyiv, Ukraine

On Swarm Algorithms with controlled Levy Flight

The paper presents an approach to algorithm performance improvement due to applying Levy Flight with controlled lambda value. The controlled movement step length of Levy Flight ensures fair balance between exploration and exploitation of the search space, and enables an algorithm to avoid trapping into local optimums on the other hand. Computational experiments have shown that proposed approach leads to algorithms performance improvement in regard of result accuracy and robustness.

In the last decade a few metaheuristic Swarm Algorithms for global optimization have been developed [1–3]. They have been derived from the behaviour of biological systems in nature and based on swarm intelligence. Behavior of the swarm agents is determined by specific rules inherent by given biological species and randomization component which controls the diversity of solutions. As a latter the sequence generated by the Levy Flight [4] is used.

The Levy Flight

$$Levy \sim u = t^{-\lambda}, (1 < \lambda \leqslant 3)$$

is a random walk in which the step lengths have a stable probability distribution. It has an infinite variance with an infinite mean. In general, a random walk is a Markov chain whose next location depends only on the current position and the transition probability. When defined as a walk in a space of dimension greater than one, the steps made are in isotropic random directions. In all abovementioned algorithms λ is considered an algorithm's parameter and, therefore, has a constant value. Due to fixed step length, the algorithms might miss better local search capabilities and sometimes they trap into several local optimums. Furthermore fixed step length narrows down an algorithm capabilities to explore remote parts of the search space.

Let's assume that all the agents are ranked in descending order accordingly to the value of their fitness function. As it known, the step length of the Levy Flight substantially depends on the value of λ. The less the value of λ, the bigger is the step length. In order to increase the quality of swarm in searching the problem space efficiently and finding the optimal solution we suppose that "worse" agents should perform the Levy Flight with bigger swing (less λ value) to explore remote parts of the search space. On the other hand, "better" agents are intended to search more thoroughly the perspective areas found in the previous iterations. This means that their step length should be smaller, or the value of λ must be larger. It may be achieved if λ belongs to the interval $[\lambda_{\min}; \lambda_{\max}]$ and the dependence of λ value on the agent's rank is as follows [5]

$$\lambda = \lambda_{\max} - i(\lambda_{\max} - \lambda_{\min})/(m-1),$$

where m is the swarm size, i – rank of i-th agent.

There is another way to control the step length of the Levy Flight [6]. Suppose best_ever is the agent that found the global best solution so far among all the agents. We will also assume that another agent is the more "worse", the farther its current position from the best_ever is. So, it seems quite logical to increase its Lévy Flight step length (by reducing λ value) for more extensive exploration of the search space. Thereby, we will calculate the value of λ for i -th agent as follows

$$\lambda = Xe^{((-dist_{i,be}/D)Y)},$$

where $dist_{i,be}$ is a distance from i-th agent to best_ever, D is the maximum distance in all dimensions, X and Y are the algorithm parameters.

Both ways the Levy Flight with controlled step length ensures that: a) a swarm of random walkers will be concentrated around recently found extremum (due to greater value of λ), as it should in evolutionary algorithms, and b) at the same time few swarm members will always explore

more distant regions of search space (due to less value of λ).

The proposed approach has been applied to the Firefly Algorithm [1], the Bat Algorithm [2] and the Cuckoo Search Algorithm [3]. In all cases, it has proved to be highly efficient.

Computational experiments. To conduct algorithms statistical analysis, widely used benchmark functions have been chosen [7]. Table 1 presents the results of tests that have been conducted on original and modified algorithms. Each algorithm has been run 50 times upon all the test problems with the number of decision variables equal to 30. In Table 1 the mean results obtained by each algorithm are shown. The result standard deviations are shown in brackets. The designations used in Table 1 are as follows: FA – Firefly Algorithm, CS – Cuckoo Search Algorithm, BA – Bat Algorithm; FAm, CSm and BAm – their modified versions respectively.

Table 1. Comparative analysis of original and improved algorithms.

Test function	FA	FAm	CS	CSm	BA	BAm
Ackley	3.21e-4 (3.01e-1)	1.1e-07 (1.3e-06)	4.2e-5 (1.1e-03)	2.9e-11 (4.4e-09)	5.6e-11 (7.2e-08)	0.4e-17 (6.1e-15)
Sphere	1.95e-11 (2.9e–9)	1.4e-17) (3.1e-15)	(2.7e-07) (6.4e-04)	(4.1e-13) (7.4e-11)	3.2e-12) (1.4e-09)	(8.3e-21) (5.1e-18)
Griewank	1.3e-05 (3.2e-02)	4.4e-09 (1.4e-07)	6.1e-07 (8.1e-04)	2.5e-11 (7.1e-09)	6.3e-09 (5.4e-06)	2.2e-14 (1.8e-12)
Rastrigin	4.85e-3 (3.08e-1)	(3.3e-07) (1.7e-05)	(5.2e-05) (6.1e-02)	(7.5e-09) (4.3e-07)	5.1e-09) (2.4e-05)	(3.6e-12) (3.8e-10)
Rosenbrock	9.25e-4 (3.43e-1)	5.6e-06 (2.4e-04)	1.3e-07 (4.1e-04)	5.6e-11 (9.4e-10)	8.7e-11 (6.1e-08)	2.4e-15 (7.1e-12)
Schwefel	1.2e-3 (4.4e-1)	1.2e-07 (1.3e-05)	7.1e-07 (2.2e-04)	3.2e-11 (8.7e-10)	1.1e-09 (7.2e-06)	6.1e-13 (4.2e-11)

Conclusions. The statistical results obtained from the benchmark functions demonstrate the superiority of the modified algorithms to their original versions in regard of results accuracy and robustness. In particular the values of the result standard deviation demonstrate more robust behavior of the modified algorithms. The most substantional performance improvement has been achieved when optimizing functions with uneven landscape and multiple minima. The introduction of the Levy Flight with controlled step length enables modified algorithms to find globally and locally improved solutions.

References. **1.** Yang, X.-S. Firefly algorithm, stochastic test functions and design optimization, Int. J. Bio-inspired Computation, vol. 2, no. 2, pp. 78–84, Feb. 2010. **2.** Yang, X.-S. A New Metaheuristic Bat-Inspired Algorithm, Nature Inspired Cooperative Strategies for Optimization, Springer Press, vol. 284, pp. 65–74, 2010. **3.** Yang, X.-S. Nature-Inspired Metaheuristic Algorithms, 2nd Edition. Luniver Press, pp. 105–116, 2010. **4.** A.M. Reynolds, C.J. Rhodes. The Lévy flight paradigm: random search patterns and mechanisms, Ecology. v 90. pp. 877–887, 2009. **5.** Zorin Yu. A Statistical Analysis of the Bat Algorithm, Сборник трудов XVI Международной научной конференции "Интеллектуальный анализ информации (ИАИ-2016)" им. Т.А. Таран. Сборник трудов. – К.: Просвіта, С. 75–79, 2016. **6.** Zorin Yu. A metaheuristic algorithm for multimodal functions optimization, Сборник трудов, Международная научная конференция имени Т.А. Таран Интеллектуальный анализ информации ИАИ2015, с. 88–92, 2015. **7.** Virtual Library of Simulation Experiments: Test Functions and Datasets. Available at http://www.sfu.ca/~ssurjano/index.html.

Архіпов Д.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Прогнозування ходу виконання проекту як елемент управління проектами

Прогнозування є невід'ємною частиною процесу прийняття рішень в умовах невизначеності. Необхідність в прогнозуванні виникає тоді, коли є невпевненість в майбутньому і деякі аспекти майбутнього не контрольовані [1]. Якщо всі факти, що відносяться до події, вже відомі, а майбутнє події вже детерміністично передбачене, або їм можна керувати з оглядом на відоме на момент прогнозування, то для досягнення бажаного результату можуть бути прийняті будь-які відповідні рішення. В іншому випадку, рішення потрібно приймати ґрунтуючись на прогнозах, що враховують невизначеність майбутнього.

Межі детерміністських підходів і необхідність імовірнісних моделей у прийнятті технічних та управлінських рішень неодноразово перевірялися протягом останніх чотирьох десятиліть [2–4]. Однак, практики часто уникають використання імовірнісних методів, здебільшого через відсутність відповідних інструментів, брак достатньої кількості достовірних даних та додаткові труднощі в оперуванні невизначеністю в кількісному вираженні в ході процесів прийняття рішення.

Для прийняття технічних та управлінських рішень існує безліч методів прогнозування, які можна охарактеризувати декількома способами. Макрідакіс та співавт. [5] розбивають методи прогнозування на чотири групи: чисто суб'єктивні підходи, причинно-наслідкові або пояснювальні методи, екстраполяційні методи (або методи часових рядів), а також будь-які комбінації з цих трьох. З іншого боку, Аль-Табтабай та Дикман [6] класифікують моделі прогнозування як економетричні моделі, моделі часових рядів і суб'єктивні моделі (методи експертних оцінок). З такою кількістю доступних альтернатив вибір відповідного методу для конкретної проблеми є складним рішенням. В літературі можна знайти декілька всебічних оглядів методів прогнозування та алгоритмів вибору правильного методу [1]. Наприклад, для створення довідника з вибору найкращого методу або їх комбінацій Георгоф і Мердік [7] оцінили 20 методів прогнозування за 16 критеріями.

Деякі спеціалісти з прогнозування вважають, що використання суб'єктивної інформації є ненайкращою альтернативою, яку можна виправдати тільки відсутністю об'єктивних даних [8], в той час як інші пропонують використання суб'єктивної інформації як способу підвищення ефективності прогнозування [7]. Аль-Табтабай та Дикман [6] провели дослідження застосування теорії соціальних суджень для прогнозування вартості комплексу будівельних робіт і стверджують, що історичні дані та компетентні судження на основі досвіду і знання повинні бути включені в методику прогнозування.

Таким чином стає очевидним, що існує потреба в надійних і послідовних методах прогнозування в управлінні проектами. Найбільш популярні на цей час детерміністські методи, такі як метод критичного шляху або метод освоєного обсягу, є надмірно спрощеними або непослідовними у застосуванні та зроблених припущеннях. Це робить їх ненадійними і непрактичними. Розвинуті методи прогнозування, які об'єднають всю необхідну інформацію та невизначеність у прогнозах ефективності проекту в математично обгрунтований спосіб, дозволять менеджерам проектів приймати більш раціональні та ефективні рішення.

За математичне підгрунтя для таких методів можна прийняти два відомі алгоритми: фільтр Калмана та Байесівський логічний вивід.

Фільтр Калмана дозволяє побудувати ймовірнісну модель, яка приймає до уваги різні види попередньо відомої інформації – наприклад, минулий досвід ведення подібних проектів. Прогнозування може враховувати історичні показники, поточні фактичні дані про хід виконання

проекту та невизначеність, що випливає з випадкового характеру помилок вимірювання та шуму процесу. В рамках проектного прогнозування, шум процесу можна інтерпретувати як відхилення у ході виконання проекту, що виникає через невизначеності при плануванні та під час виконання.

Інший метод адаптивного прогнозування дає можливість побудувати регресійну модель, параметри якої розраховуються з використанням Байесівського виводу.

З досвіду відомо, що типові графіки залежності прогресу проектів від часу мають форму латинської S незалежно від одиниці виміру – будь це сукупні витрати, робочі години або відсоток виконаних робіт [9]. Отже, в якості початкового прогнозу можна обрати модель S-кривої з урахуванням характеристик проекту, історичних даних та плану проекту.

Якщо менеджер проекту має первісний план виконання проекту, і якщо підібрана одна з відомих моделей S-кривої з відповідними параметрами Θ, що максимально відтворює цей план, тоді ці параметри можна уточнювати з використанням фактичних даних про хід проекту D. Теорему Байєса для цього випадку можна записати у вигляді:

$$P(\Theta|D) = \frac{P(D|\Theta)P(\Theta)}{P(D)}, \tag{1}$$

де $P(\Theta)$ – апріорний розподіл параметрів до одержання нових результатів; $P(D|\Theta)$ є умовною ймовірністю того, що результати D будуть отримані при використанні параметрів Θ; $P(D)$ – розподіл фактичних спостережень ходу виконання проекту D; а $P(\Theta|D)$ є апостеріорним розподілом параметрів Θ.

Для отримання висновку щодо ефективності запропонованих методів порівняно з детерміністськими, необхідно дослідити і порівняти результати прогнозів з використанням реальних даних.

Література. **1.** Armstrong J. Principles of Forecasting: A Handbook for Researchers and Practitioners [Електронний ресурс] / J.S. Armstrong // Springer. – 2001. – Режим доступу до ресурсу: https://www.gwern.net/docs/predictions/2001-principlesforecasting.pdf. **2.** Ang A. Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering / A. Ang, W. Tang. – Wiley: Technology & Engineering, 2007. – 406 с. **3.** Hertz D. Risk analysis in capital investment [Електронний ресурс] / David B. Hertz // Harvard Business Review. – 1979. – Режим доступу до ресурсу: https://hbr.org/1979/09/risk-analysis-in-capital-investment. **4.** Spooner J. Probabilistic models for construction estimating and bidding [Електронний ресурс] / J.E. Spooner // Doctoral Dissertations. – 1974. – Режим доступу до ресурсу: http://scholarsmine.mst.edu/cgi/viewcontent.cgi?article=2861&context=doctoral_dissertations. **5.** The accuracy of extrapolation (time series) methods: Results of a forecasting competition [Електронний ресурс] / [S. Makridakis, A. Andersen, R. Carbone та ін.] // Journal of Forecasting. – 1982. – Режим доступу до ресурсу: https://www.researchgate.net/publication/229864129_The_Accuracy_of_Extrapolation_Time_Series_Methods_Results_of_a_Forecasting_Competition. **6.** Al-Tabtabai H. Judgemental forecasting in construction projects / H. Al-Tabtabai, J. Diekmann. // Construction Management & Economics. – 1992. – №10. – С. 19–30. **7.** Georgoff D. Manager's Guide to Forecasting [Електронний ресурс] / D. Georgoff, R. Murdick // Harvard Business Review. – 1986. – Режим доступу до ресурсу: https://hbr.org/1986/01/managers-guide-to-forecasting. **8.** Chau K. Monte Carlo simulation of construction costs using subjective data [Електронний ресурс] / Kwong Wing Chau // Construction Management and Economics. – 1997. – Режим доступу до ресурсу: https://www.researchgate.net/publication/24077222_Monte_Carlo_simulation_of_construction_costs_using_subjective_data. **9.** A Guide to the Project Management Body of Knowledge, Third Edition [Електронний ресурс] // Project Management Institute. – 2004. – Режим доступу до ресурсу: http://kholid.lecturer.pens.ac.id/misc/PMBOK3rdEnglish.pdf.

Бідюк П.І., Крамаренко Ю.Ю.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Прогнозування розвитку демографічних процесів в Україні

Для будь-якої країни її населення є основою суспільно-економічного розвитку. Від розвитку демографічних процесів значною мірою залежить майбутнє життя суспільства. Дослідження перспектив регіонального розвитку не можна виконати без ретельного аналізу його демографічної складової. Населення виступає одним із найважливіших факторів економічного розвитку суспільства. Тому моделювання демографічних процесів є важливою ланкою системного аналізу складних соціально-економічних об'єктів.

У літературі «демографічний прогноз» визначається як науково-обґрунтоване бачення майбутньої демографічної ситуації (чисельності, віково-статевої та сімейно-шлюбної структури населення), виявлення можливих несприятливих або небезпечних тенденцій, для уникнення яких необхідні певні заходи тривалого характеру [1]. Демографічні прогнози використовуються для визначення кількості та структури трудових ресурсів, потреб суспільства в різних товарах та послугах, розвитку та розташуванні об'єктів соціальної інфраструктури.

Метою виконаної роботи є дослідження методології демографічного моделювання; побудова моделей вибраних демографічних процесів на основі наявних статистичних даних; прогнозування за допомогою моделей майбутніх значень досліджуваних часових рядів та вибір кращої моделі для застосування.

У результаті досліджено такі методи побудови математичних моделей:

- методи регресійного аналізу;
- метод групового урахування аргументів (МГУА);
- метод експоненційного згладжування та інші.

Для вибору кращої моделі з моделей-кандидатів використаємо такі статистичні параметри, що характеризують адекватність моделі, а саме [2]: коефіцієнт множинної детермінації; сума квадратів похибок; статистика Дарбіна–Уотсона. А також параметри для оцінювання якості прогнозу: середня абсолютна похибка в процентах (САПП); коефіцієнт Тейла.

В якості прикладу побудови моделей обрано такі демографічні процеси: чисельність населення, народжуваність, смертність, динаміка шлюбів та розлучень. Для побудови моделей використано дані Державної служби статистики України.

В результаті виконаних обчислювальних експериментів встановлено, що кращі результати короткострокового прогнозування надає МГУА (САПП = $(0.8 \div 5)$ %), а довгострокового – модель авторегресії з трендом (САПП = $(1.8 \div 7.5)$ %).

Отже, обчислення оцінки прогнозів майбутньої демографічної ситуації, тобто як зміниться чисельність населення, народжуваність, смертність – все це важливі складові для розвитку і планування різних сфер економіки та діяльності суспільства (соціальної інфраструктури, охорони здоров'я, освіти, пенсійного забезпечення). А отримані моделі можуть бути використані при прийнятті важливих управлінських рішень.

Література. **1.** Прогнозування розвитку територій. Демографічне прогнозування: навч. посіб. / Н.І. Балдич, О.В. Берданова, Н.М. Гринчук [та ін.]; за заг. наук. ред. М.М. Білинської. – К.: НАДУ, 2013. – 120 с. **2.** Бідюк П.І. Аналіз часових рядів / Бідюк П.І., Романенко В.Д., Тимощук О.Л. – К: Політехніка, 2010. – 317 с.

Буткалюк С.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Система оптичного розпізнавання нотного запису

В час постійного здешевлення та збільшення потужностей комп'ютерів більшість медіа зберігається, редагується та поширюється у цифровому вигляді. Як і оцифровування друкованих книг за допомогою OCR-систем, є необхідність у зберіганні партитури в цифрових форматах.

Опис задачі. Викликані труднощі. Оптичне розпізнавання нот (англ. OMR – optical music recognition) або Music OCR (optical character recognition – оптичне розпізнавання символів) – одне із застосувань OCR для розпізнавання та збереження нотного запису у придатній для подальшого програвання або редагування формі. OMR має достатньо спільних підходів з OCR. Однак, враховуючи специфіку правил нотного запису – має значення положення знаків і по горизонталі і по вертикалі, структура розпізнавання OMR є більш складною.

Окрім того, ускладнює процес розпізнавання те, що один і той же за значенням символ може бути зображений різними способами – можуть бути присутніми різний розмір штилів, дуг та з'єднань, віддзеркалення символів, зв'язаність символів тощо.

Найбільшою проблемою можуть бути так звані super-imposed об'єкти – різні знаки, зображення яких суміщене між собою [1].

Опис загального підходу до вирішення. Деталі побудови системи оптичного розпізнавання нотного запису можуть відрізнятися, однак, можна виділити загальну схему обробки:

Завантаження та попередня обробка зображення – проводиться бінаризація, вирівнювання деформованого зображення, пошук та видалення текстових елементів тощо. В побудованій системі використовуються проекційний метод вирівнювання зображення (вузькі вертикальні полоси суміщуються за піками проекцій) та адаптивна бінаризація. Для видалення тексту було використано Tesseract OCR.

Пошук та видалення ліній нотного стану – проводиться знаходження позицій ліній нотного стану на зображенні (для подальшого визначення точного положення/висоти нот) та видалення нотних лінійок (для подальшого виділення об'єктів та їх класифікації). В нашому випадку використовувався модифікований XOR-проекційний метод [2].

Знаходження значущих об'єктів на та біля нотного стану, використовуючи алгоритми сегментації об'єктів та різноманітні евристики для з'єднання пошкоджених об'єктів.

Класифікація знайдених об'єктів – зазвичай застосовуються методи класифікації, засновані на порівнянні шаблонів або специфічних характеристик (дескриптори Фур'є, компактність об'єкту, моменти Церніке, центральні моменти тощо) і використанні класифікаторів типу k-NN, методів опорних векторів, нейронних мереж. В роботі використаний класифікатор k-NN з використанням моментів Церніке, Фур'є, об'єму, діагональних проекцій та відношень.

Створення нотної моделі – представлення символьного зображення враховуючи правила побудови музичного тексту. Для задачі можуть використовуватися модернізовані граматики визначених тверджень [3] (DCG – Definite Clause Grammar), двовимірні граматики та різноманітні евристики.

Застосування описаних підходів до розробки системи OMR. При побудові системи попередньої обробки зображення, виділення та класифікації символів були використані різні підходи до описаних вище етапів.

Дослідження якості роботи програмного продукту проводилося на 110 тестових семплах з роздільною здатністю близькою до 300 DPI (dots per inch – точок на дюйм). Результати дослідження якості аналізу наведемо у таблиці 1.

Як видно з таблиці 1, точність аналізу зображень в програмному продукті більше 95%. Потрібно враховувати те, що тренування k-NN класифікатора проводилося на п'яти зображеннях, що є замалим значенням для адекватного класифікування символів.

Табл. 1. Результати роботи програмного продукту

Типи об'єктів	Всього	Правильно розпізнані	Похибка 1 роду	Похибка 2 роду	Точність (precision)	Повнота (recall)	F_1 – міра
Нотні голівки	2372	2342	10	30	0.9957	0.9874	0.9915
Паузи	65	65	0	0	1.0000	1.0000	1.0000
Альтерації	242	236	8	6	0.9672	0.9752	0.9712
Ключі	110	110	5	0	0.9565	1.0000	0.9778
Роздільники	721	710	13	11	0.9820	0.9847	0.9834

Більшість помилок при класифікації спричинені відсутністю відповідних об'єктів в базі даних класифікатора (знак альтерації бекар, нота довжини 1/4 з хвостиком вниз). Проблеми з такими символами вирішуються тренуванням k-NN класифікатора.

Інший тип помилок виникає під час класифікації з причини використання лише інваріантних за розміром характеристик об'єктів, що змішує схожі за виглядом символи різного розміру. Для виправлення таких помилок необхідно перед класифікацією приводити зображення до відомого масштабу та враховувати характеристики неінваріантні до розмірів символів.

З аналізованих зображень видно, що не завжди спрацьовує система аналізу та видалення текстового шару на основі tesseract-ocr, що спричиняє велику кількість шумової інформації під час класифікації символів у k-NN класифікаторі (відсутня можливість класифікувати символи тексту, оскільки вони сильно фрагментовані від процесу попередньої підготовки зображення).

Висновки. Проведено дослідження існуючих підходів до вирішення задачі розробки системи автоматизованого оцифровування музичної нотації. Проведений аналіз алгоритмів попередньої обробки зображення та алгоритмів класифікації музичних об'єктів на нотному листі з виділенням їх недоліків. Запропонований стійкий алгоритм видалення ліній нотного стану з зображення та алгоритм класифікації музичних об'єктів обробленому зображенні.

За аналізом отриманих результатів роботи програмного продукту можна зробити висновок, що обрані алгоритми дозволили досягти виконання поставлених цілей з точністю більше 95%. Результати роботи можуть бути покращені при використанні більш складних евристичних алгоритмів вибору та збільшенні вибірки k-NN класифікатора.

Література. **1.** Bainbridge, David. Dealing With Superimposed Objects In Optical Music Recognition / David Bainbridge, Tim C Bell // Sixth International Conference On Image Processing And Its Applications. – 1997. [Електронний ресурс]. – Режим доступу: http://www.cs.waikato.ac.nz/~davidb/publications/ipa97/. **2.** George, Susan Ella. Visual perception of music notation: on-line and off-line recognition / Susan Ella George. – IRM Press, 2005. – P. 984. **3.** Bainbridge, David. An extensible Optical Music Recognition system / David Bainbridge, Tim Bell // Nineteenth Australasian Computer Science Conference. – 1997. [Електронний ресурс]. – Режим доступу: http://www.cs.waikato.ac.nz/~davidb/publications/acsc96/.

Віштальь Д.М., Васильєв В.І., Любашенко Н.Д.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Подання і застосування категорії часу (Date/Time) в технології публікації і споживання веб-даних

В роботі досліджено сучасний стан проблеми по створенню єдиного веб-простору даних. Обґрунтовується доцільність розробки системи підтримки процесу публікації і споживання даних у Всесвітній павутині (WWW). Теоретичною і практичною основою створення системи є стандарти технологій Semantic Web і Linked Data. З метою підвищення ефективності пошукових запитів до WWW проаналізовано обробку часових даних.

Для розміщення і пошуку даних у веб-просторі існують і розвиваються технології Semantic Web (Семантична павутина) і Linked Data (Зв'язані дані), які забезпечують створення і розвиток глобального веб-простору даних, в якому структуровані дані є читабельними як для людей, так і для комп'ютерів. В основі цих технологій лежить ідея забезпечення найширшого доступу до даних [1–3], коли будь-який користувач (людина, комп'ютер) може виявляти дані, отримувати доступ до них, обробляти їх і використовувати знову і знову [1]. При реалізації цієї ідеї практична цінність даних (зокрема, отримання нових знань там, де вони навіть не були передбачуваними) безмежно зростає.

В даній роботі було проведено аналіз існуючих практичних і теоретичних результатів в області вказаних технологій. Виявилось, що на даний момент опубліковані досягнення по створенню єдиного глобального веб-простору даних мають рівень практичної реалізації в світовому масштабі [3].

Проблема полягає в тому, що для подальшого розвитку мережі даних потрібно залучати нових користувачів (окремих людей, організації, бізнес-структури, тощо). Але для них це означатиме прикладання додаткових зусиль і часу для вивчення нових способів роботи з даними. Процес навчання ускладнюється через велику кількість он-лайн документації, яка міститься на різноманітних сайтах.

Як один з варіантів вирішення цієї проблеми, пропонується розробка системи підтримки процесу розміщення даних у Всесвітній павутині і подальшого їх використання. Структуру системи в укрупненому вигляді подано на рис. 1.

Підсистема підтримки публікації веб-даних	Вибір формату подання RDF моделі
	Опис ресурсу (набору даних)
	Опис метаданих
	Шаблони публікації даних
Підсистема підтримки споживання веб-даних	Пошук даних у веб-просторі
	Розробка застосувань Linked Data
	Публікація і споживання статичних даних
	Споживання даних в реальному часі
Технічна підтримка	Клієнт-серверне забезпечення

Рис. 1. Система підтримки роботи з даними у веб-середовищі

В даній роботі основну увагу приділено категорії часу в контексті публікації і використання даних у Всесвітній павутині (WWW).

При роботі людей і комп'ютерів з набором даних по технологіях Semantic Web і Linked Data поняття часу зустрічається в наступних випадках:

1. Час як частина самого набору даних.
2. Час як елемент метаданих.
3. Час як один з параметрів пошуку даних.

Випадок 1. Час як частина самого набору даних. Аналіз існуючих словників, онтологій показав наявність достатньої кількості формальних описів часу (date, time) як класу, як типу даних, як властивості для побудови RDF (Resource Description Framework) моделі даних. Проблема полягає в ефективному їх застосуванні при публікуванні користувачем своїх даних. Це

обумовлено, зокрема, наступними причинами: публікатор може не усвідомлювати важливість роботи з часовими елементами для організації пошуку даних в майбутньому; для освоєння існуючих можливостей потрібні вагомі зусилля з боку не підготовленого користувача.

Випадок 2. Час як елемент метаданих.

За результатами досліджень міжнародної відкритої організації DCMI (The Dublin Core Metadata Initiative) [5] існує 4 рівні сумісності та взаємодії програм, які використовують метадані про ресурси на WWW. Зокрема, рівень 2 (сумісність за формальною семантикою) зараз розвивається найбільшими темпами. Як результат, змінився ідейний напрям діяльності DCMI з "Розробка простих метаданих для виявлення інформаційного ресурсу" на "Структуровані дані для оптимізації пошуку". Рівень 2 базується на формальній моделі RDF.

До формування метаданих є ряд вимог [3], з яких наступні безпосередньо пов'язані з часом:

- Які метадані треба прийняти до уваги, коли описувати веб-дані?
- Як відслідковувати різні версії набору даних і управляти ними?
- Коли і як архівувати веб-дані?

Щоб всі вищенаведені вимоги виконувались, пропонується реалізувати наступні механізми для здійснення цільового доступу людини і програми до елементів date/time:

1. При публікації набору даних надати програмну підтримку у формі рекомендацій щодо включення відповідних описів.

2. Автоматизувати формування деяких елементів в RDF моделі в процесі публікації користувачем своїх даних (bootstrap).

Обидва механізми передбачено в системі підтримки процесу публікації і споживання веб-даних.

Випадок 3. Час як один з параметрів пошуку даних.

Для запитів, які містять інтервали часу, пошук може виявитися задовгим. Тому, наприклад, якщо дані, що публікуються в мережі, описують одяг мешканців визначеної території в 19-му столітті, то цей період часу бажано додати до метаданих у відповідному форматі. Це підвищить ефективність обробки запитів на пошук веб-даних про одяг мешканців в різні періоди історії. Для запитів, які стосуються споживання даних в режимі реального часу, додавання показників таймера забезпечить можливість наукової обробки отримуваних даних.

Отже, аналіз випадків 1–3 показав, що для обслуговування створення і обробки даних в частині врахування часу виникає потреба у відповідній підсистемі.

Структура і функції такої підсистеми передбачають, зокрема:

- виокремлення часових термів у кореневий клас TIMECATEGORY,
- автоматичне формування підтримки часу для відповідних елементів моделі RDF з підключенням таймеру,
- підтримка часових властивостей і характеристик як опції,
- виведення окремих даних типу date/time в метадані з метою оптимізації пошуку за відповідними запитами.

На даний момент розроблено скрипти (bootstrap datetime picker) [4], які дозволять технічно реалізувати цю підтримку.

Отже, актуальною є задача розробки системи підтримки публікування і споживання даних у Всесвітній павутині. Підсистема підтримки часу стає її невід'ємною складовою.

Якщо при публікації даних будуть витримуватись вимоги щодо опису часових характеристик, тоді і при споживанні цих даних користувач зможе отримати максимальний ефект у вигляді отримання потрібної інформації, нових знань, закономірностей тощо.

Література. **1.** Tom Heath and Christian Bizer (2011). Linked Data: Evolving the Web into a Global Data Space (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1–136. Morgan @ Claypool. **2.** Bob DuCharme Learning SPARQL. Querying and Updating with SPARQL 1.1 O'Reilly Media July 2013. – 386 p. **3.** Data on the Web Best Practices. W3C Recommendation 31 January 2017. **4.** https://tarruda.github.io/bootstrap-datetimepicker/. **5.** http://dublincore.org/metadata-basics/.

Галатенко Д.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Використання нейронних мереж для розпізнавання нот піаніно

Нейронні мережі відіграють все більш важливу роль в житті людей. Вони використовуються в розпізнаванні візуальних образів, тексту, мови, та багато іншого. Така тенденція росту дає підстави говорити про інтелектуалізацію обчислювальної техніки в цілому.

В даній роботі розглядається проблема розпізнавання нот піаніно із застосуванням нейронних мереж. Розпізнавання проходить на основі спектральних особливостей нот. Такий підхід може бути використаних для навчання користувача грі на піаніно його улюблених мелодій, навіть якщо людина не має професійної музичної освіти, подібно до навчальних відео Piano Easy [1]. Зазвичай розбір однієї мелодії проходить в ручному режимі і займає значний проміжок часу, в той час як нейронна мережа дозволить автоматизувати цей процес до розпізнавання нот на льоту.

Загальна техніка вирішення цієї проблеми полягає у розпізнаванні звуків мелодії для співставлення їх із клавішами які мають бути натиснені в даний момент часу. Музичний звук має характерну структуру в спектральному діапазоні. В його склад входить основний тон як правило із найвищою амплітудою і подальші гармоніки. Найбільш важливими є декілька перших гармонік, саме вони визначають ноту звуку. Проте звуковий спектр являє собою великий набір даних який необхідно мінімізувати для розпізнавання на льоту. Ця проблема вирішується за допомогою використання мел-частотних кепстральних коефіцієнтів (MFCC) визнаних найкращими для розпізнавання музичних нот в роботі [2]. MFCC представляють собою нелінійний спектр спектра, чудово апроксимуюуть слухову систему людини. Застосовуються для вирішення задач NLP. Обчислення MFCC відбувається наступним чином:

1. віконне перетворення Фур'є;
2. нелінійне розбиття спектру на N частин за допомогою застосування мел-шкали;
3. обчислення енергії сигналу для кожного інтервалу із застосування трикутних фільтрів;
4. дискретне косинусне перетворення.

Для мінімізації отриманого набору даних використовуємо метод головних компонент (PCA). Цей метод дасть можливість зменшити кореляцію ознак та виділити найбільш значимі з них. Сам алгоритм визначається такими кроками [3]:

1. обчислення матриці кореляції;
2. знаходження власних значень і відповідних власних векторів матриці;
3. сортування власних векторів по спаданню відповідних їх власних значень;
4. проекція вхідних даних (MFCC) на власні вектори;
5. відсікання останніх M проекцій (найменш значимих компонент).

Для класифікації використовується штучна нейронна мережа (ШНМ) прямого поширення, тобто напрямок розповсюдження інформації не зміюється. Для навчання нейронної мережі використаний спосіб зворотнього розповсюдження помилки. Нейронна мережа містить один внутрішній шар і має наступні параметри конфігурації: 24 нейрони у внутрішньому шарі; допустима похибка класифікації – 1%; швидкість навчання 0,05 с. Для оцінки процесу навчання використана крос-валідація, де в якості валідаційного набору використано 30% навчальної вибірки. Вибірка являє собою марковані ноти піаніно з кодом відповідної клавіші.

В результаті отримуємо систему з навченою ШНМ, яка здатна правильно розпізнавати ноти піаніно в режимі реального часу.

Література. **1.** Adele – Someone Like You – Piano Tutorial Easy – how to play (synthesia). https://www.youtube.com/watch?v=rOKP6gwF7RI. **2.** 2. Eronen A. Workshop on Signal Processing for Audio and Acoustics, pp. 19–22 http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=969527. **3.** 3. Simon O. Haykin Neural Networks and Learning Machines Third Edition https://cours.etsmtl.ca/sys843/REFS/Books/ebook_Haykin09.pdf.

Гожий В.В., Чертов О.Р.
КПІ ім. Ігоря Сікорського, ФІПМ, Київ, Україна

Автоматичне визначення психологічного портрету особистості у соціальних мережах

Дані тези присвячені автоматичному визначенню психологічних рис особистості шляхом інтелектуального анлізу її профілю у соціальних мережах.

Big Five. Є різні методи визначення психологічного портрету (наприклад, визначення темпераменту через опитування [2], тест Роршаха [1]), але найзручнішим для автоматичного визначення є так звана Five Factor Model, яка базується на 5 персональних характеристиках, які людина може мати або не мати. Ці п'ять характеристик (Big Five [3]) це: екстраверсія (коли людина орієнтована на навколишнє оточення – друзі, нові знайомства тощо), відкритість досвіду (цікавість проти обережності), сумлінність (організованість проти легковажності), приємність (дружня і компанійська людина чи аналітична і стоїть окремо) та нейротицизм (людина є чутливою/нервовою чи твердою/впевненою).

Видобування психологічних рис. Згідно із [4–7], ми маємо наступну інформацію про риси із Big Five у соціальних мережах:

- Екстраверти мають більше друзів, витрачають більше часу на соціальні мережі, роблять пости в більшості про їхнє щоденне життя, завантажують багато фотографій та часто змінюють картинку профілю.
- Відкриті для досвіду користувачі роблять пости про їхні інтелектуальні інтереси.
- Сумлінні користувачі у загальному випадку залишаються осторонь соціальних мереж, але у момент навчання вони завантажують більше відео, створюють більше фотоальбомів та організовують їх більш старанно.
- Невротичні користувачі завантажують у більшості фотографії (порівнюючи із екстравертами – вони завантажують більше фотографій у свої альбоми).
- Приємні у житті користувачі отримують більше "лайків" та коментарів під їхніми постами та більше всіх відзначаються на фотографіях інших користувачів.

Детальніше про вагомість тих чи інших властивостей профілю див. Рис. 1, про більш детальни вплив того чи іншого чиннику див. Табл. 1.

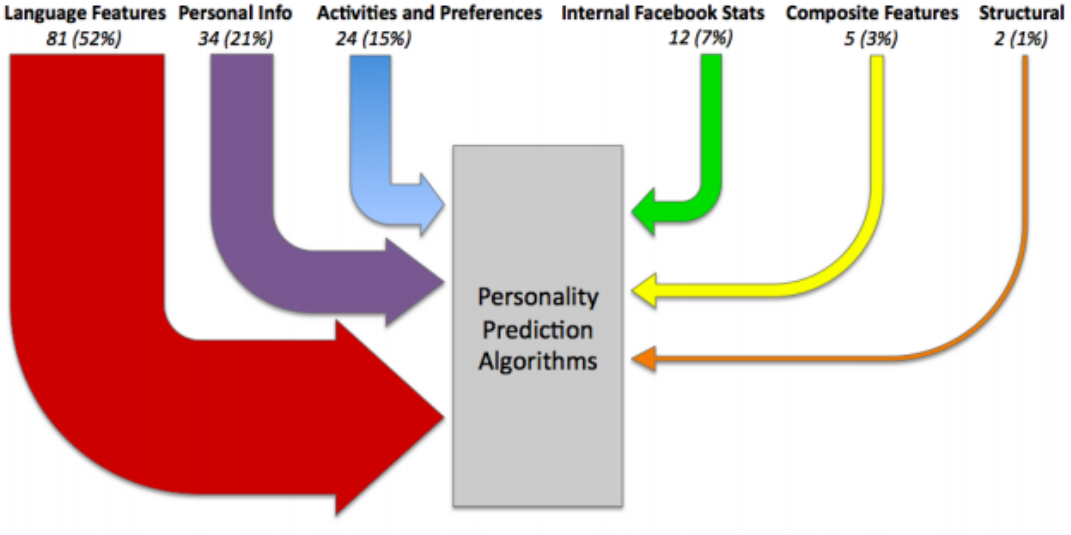


Рис. 1. Розподіл видів інформації за вагомістю

	Open.	Consc.	Extra.	Agree.	Neuro.
Linguistic Features					
Swear Words	0.006	-0.171	0.032	-0.084	-0.120
Social Processes (e.g. Mate, talk, they, child)	0.010	0.264	0.091	-0.022	-0.142
Human Words (e.g. baby, man)	0.078	0.203	0.070	-0.050	-0.062
Affective Processes (e.g. Happy, cried, abandon)	0.105	-0.009	0.136	0.203	0.038
Positive Emotions (e.g. Love, nice, sweet)	0.052	0.045	0.117	0.167	-0.013
Anxiety Words (e.g. Worried, fearful, nervous)	0.044	-0.150	0.008	0.101	0.192
Perceptual Processes (e.g. Observing, heard, feeling)	-0.040	-0.195	-0.163	-0.027	0.096
Seeing Words (e.g. View, saw, seen)	0.060	-0.227	-0.112	0.013	0.067
Biological Processes (e.g. Eat, blood, pain)	-0.014	0.042	0.038	0.154	0.067
Ingestion Words (e.g. Dish, eat, pizza)	-0.098	-0.050	0.029	0.031	0.207
Work Words (e.g. Job, majors, xerox)	0.134	0.096	0.154	0.048	-0.044
Money Words(e.g. Audit, cash, owe)	-0.161	0.024	0.012	-0.006	0.029
Structural Features					
Number of Friends	-0.094	-0.078	0.186	0.013	-0.069
Egocentric Network Density	-0.152	0.050	-0.224	0.059	0.032
Activities and Preferences					
Activities (char length)	0.115	0.095	0.188	0.066	-0.145
Favorite Books (char length)	0.158	-0.093	0.019	0.082	0.028
Relationship Status (none listed, single, not single)	0.093	0.071	0.194	0.040	-0.036
Last Name length in characters	0.012	-0.111	0.000	-0.044	0.184

Табл. 1. Взаємозв'язок між чинниками із профілю та психологічними рисами [8]

Таким чином, здійснивши аналіз профілю у соціальній мережі, проаналізувавши зміст постів (наприклад, сумлінні люди не використовують нелітературних слів у своїх постах, вони більше говорять про себе, про дітей тощо; відкриті до досвіду користувачі уникають розмов про гроші, натомість люблять говорити про книги), проаналізувавши інші статистичні показники (кількість фотографій та відео, кількість друзів, кількість альбомів), можна автоматично визначити психологічні риси користувача, при цьому він не буде змушений витрачати час на опитування.

Література. **1.** "Тест Роршаха" Wikipedia. Available: https://en.wikipedia.org/wiki/Rorschach_test. **2.** "Темперамент" Wikipedia. Available: https://en.wikipedia.org/wiki/Temperament. **3.** "Big Five personality traits" Wikipedia, 8 Mar 2014. Available: https://en.wikipedia.org/wiki/Big_Five_personality_traits. **4.** S.K.W. Ph.D., "What is Your Facebook Personality?," Psychology Today, 20 Dec 2011. [Online]. Available: https://www.psychologytoday.com/blog/fulfillment-any-age/201112/what-is-your-facebook-personality. **5.** A. Morin, "What Your Facebook Use Reveals About Your Personality And Your Self-Esteem," Forbes, 31 October 2014. [Online]. Available: http://www.forbes.com/sites/amymorin/2014/10/31/what-your-facebook-use-reveals-about-your-personality-and-your-self-esteem. **6.** G.S. Ph.D., "What Can You Learn About People from Facebook?," Psychology Today, 02 Jul 2015. [Online]. Available: https://www.psychologytoday.com/blog/close-encounters/201507/what-can-you-learn-about-people-facebook. **7.** A. Blaszczak-Boxe, "What Your Facebook Photos Say About Your Personality," Live science, 5 Aug 2014. [Online]. Available: http://www.livescience.com/47191-facebook-photos-reveal-personality-traits.html. **8.** Golbeck J., Robles C., Turner K., "Predicting Personality with Social Media", HCIL-2010-30, November 2010.

Горбань Г.В.
Чорноморський національний університет імені Петра Могили, Миколаїв, Україна

Генерація асоціативних правил у багатомірних даних у межах одного виміру

Однією з важливих задач Data Mining у даний час є пошук асоціативних правил, що дозволяє виявити між статистичними даними залежності у вигляді $X \to Y$, де X є умовою (або антецедентом), а Y – наслідком (або консеквентом) [1, 2].

Звичайно методи пошуку асоціативних правил застосовуються для даних у вигляді плоских таблиць. Однак асоціативні правила можуть бути знайдені і у багатомірних даних, які звичайно представляються у вигляді OLAP-кубів, складовими яких є виміри та міри.

У багатомірних даних можливо знайти асоціативних правил, що можна класифікувати на наступні типи [3]:

1) міжвимірові асоціативні правила (inter-dimensional association rules) – правила між атрибутами (значеннями) різних вимірів:

2) асоціативні правила в межах одного виміру (intra-dimensional association rules):

3) гібридні асоціативні правила (hybrid association rules) – правила між атрибутами різних вимірів, до яких входити декілька атрибутів одного виміру.

Дана доповідь присвячена методу пошуку другого типу зазначених вище правил, оскільки даний тип ще не досліджений достатньо добре на відміну від міжвимірових асоціативних правил, які були детально розглянуті у дисертаційному дослідженні автора та інших його публікаціях.

Асоціативні правила у межах одного виміру можливо знайти у кожному вимірі, який є складовою частиною кубу. У кожного виміру є множина значень (атрибутів), яку можна представити наступним чином [4]: $A_I = \{A_I^1, A_I^2, ..., A_I^k, ..., A_I^m\}$, де A_I^k – k-й атрибут I-го виміру, m – кількість атрибутів в I-му вимірі.

Розглянемо пошук асоціативних правил у межах одного виміру у загальному випадку. Подібні правила представляються у вигляді:

$$A_I^x \wedge ... \wedge A_I^y \to A_I^z \wedge ... \wedge A_I^v,$$

де I – індекс певного виміру OLAP-кубу, x, y, z, v – індекси певних атрибутів I-го виміру. При цьому одні атрибути входять до антецеденту, а інші – до консеквенту.

Представимо множини відповідних індексів атрибутів як Ant та $Cons$:

$$Ant = \{x, ..., y\};$$
$$Cons = \{z, ..., v\}.$$

Мінімальна кількість атрибутів виміру у правилі становитиме 2, коли в антецеденті і консеквенті буде знаходитись рівно по одному атрибуту: $A_I^x \to A_I^y$.

У свою чергу, максимальна кількість атрибутів може становити їх загальну кількість у вимірі m. Для того, щоб було отримане повноцінне асоціативне правило, потрібно, щоб в антецеденті та консеквенті був хоча б один атрибут, що в даному випадку представлятиме їх кількість:

$$A_I^x \to A_I^y \wedge ... \wedge A_I^z;$$
$$A_I^x \wedge ... \wedge A_I^y \to A_I^z;$$

Більшість знайдених асоціативних правил будуть мати наступний вигляд:

$$A_I^x \wedge ... \wedge A_I^y \to A_I^z \wedge ... \wedge A_I^v,$$

коли і в антецеденті, і в консеквенті буде більше ніж один атрибут виміру.

Таким чином, можливо згенерувати асоціативні правила між кількістю атрибутів одного виміру від 2 до m: $k = 2..m$. Для того, щоб почати генерацію асоціативних правил, спочатку

необхідно сформувати множину Ant, а потім із атрибутів, що залишились ще невикористаними, сформувати множину $Cons$.

У множину Ant може входити від 1 до $k-1$ атрибутів ($j = 1..k-1$). У випадку правил із загальною кількістю 2 атрибутів, до Ant може входити тільки один атрибут. Формування множин Ant та $Cons$ буде здійснюватись наступним способом (на прикладі виміру з 5 атрибутами):

$$k = 2; j = 1;\ A_I^{'} = \{\{A_I^1\}, \{A_I^2\}, \{A_I^3\}, \{A_I^4\}, \{A_I^5\}\};\ P(A_I^{'}) = C_5^1 = 5.$$

Почергово витягуючи один елемент множини $A_I^{'}$, буде отримана множина Ant для окремого правила. Подібним чином для конкретного правила отримується і множина $Cons$, тільки беруться комбінації з різниці загальної множини атрибутів та множини Ant.

Для правил, у антецеденті яких буде знаходитись атрибут A_I^1

$$B_I = A_I \backslash Ant = A_I \backslash \{A_I^1\} = \{\{A_I^2\}, \{A_I^3\}, \{A_I^4\}, \{A_I^5\}\};$$
$$B_I^{'} = \{\{A_I^2\}, \{A_I^3\}, \{A_I^4\}, \{A_I^5\}\};\ P(B_I^{'}) = C_4^1 = 4.$$

Таким чином, будуть сформовані наступні правила:

$$A_I^1 \to A_I^2; A_I^1 \to A_I^3; A_I^1 \to A_I^4; A_I^1 \to A_I^5.$$

Так само можна згенерувати інші асоціативні правила в межах виміру I між двома його атрибутами, в антецеденті яких будуть інші атрибути. За описаною методикою можливо знайти також правила між більшою кількістю атрибутів певного виміру.

Для обчислення підтримки та достовірності асоціативних правил у реляційній моделі даних використовуються кількість транзакцій, що присутні у предметному наборі [1–3]. Для багатомірних даних такою кількісною характеристикою є значення міри кубу, що відповідає конкретним значенням його вимірів. Для прикладу візьмемо OLAP-куб з трьома вимірами, в межах першого з яких потрібно знайти асоціативні правила. Значення інших двох вимірів при цьому мають фіксовані значення, наприклад a та b.

Підтримкою асоціативного правила в межах одного виміру буде відношення суми всіх мір, що мають у вимірі, за яким проводиться аналіз, атрибути, що належать як до умові, так і до наслідку, до підсумкового значення міри (ALL) за відповідним виміром:

$$Supp = \frac{\sum_{k=0}^{p} M_{Ant_k,a,b} + \sum_{l=0}^{r} M_{Cons_l,a,b}}{M_{ALL,a,b}}.$$

У свою чергу, достовірність асоціативного правила дорівнюватиме відношенню суми мір, що містять атрибути аналізованого виміру, що присутні як в умові, так і в наслідку, до суми мір, що містять атрибути виміру, які присутні тільки в умові

$$Conf = \frac{\sum_{k=0}^{p} M_{Ant_k,a,b} + \sum_{l=0}^{r} M_{Cons_l,a,b}}{M_{Ant_k,a,b}}.$$

У подальші плани автора входить розробка методів генерації асоціативних правил у межах одного виміру з подальшою їх реалізацією, а також дослідження гібридних асоціативних правил між багатомірними даними.

Література. **1.** Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. – СПБ: БХВ-Петербург, 2004. – 336 с. **2.** Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям: Учебное пособие. 2-е изд., перераб. и доп. – СПб.: Питер, 2010. – 704 с. **3.** Zhu, H. Online analytical mining of association rules. Master's thesis, Simon Faster University, Burnaby, British Columbia, Canada, 1998, 117 p. **4.** Fisun, M. Generation of the Association Rules among Multidimensional Data in DBMS Caché Environment / M. Fisun, H. Horban // Advances in Intelligent Systems and Computing. Selected Papers from the International Conference on Computer Science and Information Technologies, CSIT 2016, September 6–10 Lviv, Ukraine. – 2016. – pp. 63–79.

Давидюк О.Ю.
КПІ ім. Ігоря Сікорського, ФТІ, Київ, Україна

Вибір ефективних методів кластеризації та класифікації мережевої активності для набору даних KDD CUP 99

Через постійне збільшення кількості інцидентів, пов'язаних з порушенням інформаційної безпеки в комп'ютерних мережах, виявлення мережевих атак та аномалій є важливою проблемою, яка вивчається в різних сферах науки. В проведеному дослідженні автором запропонована комплексна модель машинного навчання, яка комбінує переваги навчання з учителем та без учителя, а також дозволяє з високою точністю класифікувати кластери мережевої активності.

Об'єкт дослідження. Набір даних мережевої активності *KDD CUP 99* [1].

Предмет дослідження. Моделі та методи виявлення мережевих атак та аномалій.

Мета роботи. Розробити ефективну модель кластеризації вхідних даних з використанням машинного навчання, а також модель класифікації отриманих кластерів на основі машинного навчання з учителем.

Дослідження. В якості вхідних даних було обрано запис мережевої активності *KDD CUP 99* який являє собою набір векторів, кожен з яких репрезентує один запит в мережі. Кожен вектор складається з 41 числової координати, які були отримані після обробки мережевого запиту, а також одного класифікуючого запису, який визначає до якого класу поведінки відноситься даний запит. Всього в наборі існує 23 різних типи поведінки, включаючи 22 типи атак та один загальний тип нормальної поведінки.

Наявні моделі та методи виявлення аномалій та кластеризації мережевої активності не розглядають питання класифікації отриманих кластерів на достатньому рівні. Наприклад в роботах [2, 3] пропонується визначати тип поведінки, яку описує кластер по відносним розмірам кластерів. Такий підхід реалізується лише в окремих випадках: при значній частці (>98%) нормального трафіку та малій кількості аномальної активності. В інших випадках класифікація кластерів за таким принципом не можлива з огляду на те, що відносний розмір всіх кластерів буде однаковим в межах похибки.

Для підвищення якості машинного навчання було проведено початкову обробку даних, яка включала в себе нормалізацію векторів. Кожен вектор нормалізувався згідно з наступними правилами:

$$avgvector[j] = \frac{1}{N}\sum_{i=1}^{N} vector_i[j],\ stdvector[j] = (\frac{1}{N-1}\sum_{i=1}^{N}(vector_i[j] - avgvector[j])^2)^1/2,$$
$$newvector[j] = \frac{vector[j] - avgvector[j]}{stdvector[j]}.$$

Початковим етапом обробки даних є кластеризація. Основна ціль цього етапу полягає у виокремленні та групуванні різних типів мережевих запитів по їх поведінці.

Для проведення кластеризації векторів автором було обрано алгоритм *DBSCAN (Density-based spatial clustering of applications with noise)* [4]. Даний вибір обумовлено тим, що цей алгоритм проводить найбільш чітку кластеризацію завдяки його властивості знаходити та відкидати так званий шум (вектори, які знаходяться задалеко від всіх існуючих кластерів і водночас не можуть сформувати окремий кластер). Ця важлива властивість алгоритму є визначальною у виборі методу кластеризації, оскільки інші розглянуті алгоритми (*k-Means, k-Means++, fuzzy k-Means* [5–7]) давали не достатньо чітко окреслені кластери, що в свою

чергу не дозволило провести їх класифікацію з високою ймовірністю розпізнавання. Ще однією важливою особливістю алгоритму є те, що кількість кластерів не потрібно визначати при ініціалізації алгоритму, що дасть змогу в майбутньому обробляти різнорідні дані. В результаті обробки мережевої активності були виділені чіткі кластери, кожен з яких містив в собі абсолютну більшість (99–100%) векторів однієї і тієї ж категорії.

При кластеризації для визначення відстані між векторами було використано *Евклідову метрику*:

$$distance(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}.$$

Для забезпечення максимальної загальності алгоритму кожна координата вектору при обрахуванні мала однакову вагу, хоча для покращення точності алгоритму в специфічних умовах є сенс використовувати змінений підхід визначення відстані, при якому різні координати мають різну вагу в результуючому підрахунку.

Після кластеризації та ручної класифікації (по типу атак) отриманих кластерів було використано алгоритм машинного навчання з учителем *kNN (k-nearest neighbors algorithm)* [8]. Використання цього алгоритму дозволило побудувати класифікатор, який з високою ймовірністю (*>80%*) передбачає, до якого типу атак відноситься побудований кластер.

Висновок. В ході дослідження була проведена обробка та інтелектуальний аналіз вхідних даних – мережевої активності *KDD CUP 99*. Після кластеризації векторів з використанням алгоритму *DBSCAN* було отримано набір кластерів, кожен з яких практично повністю складається з векторів які належать до одного і того ж типу мережевої активності. Такий чіткий розподіл дозволив побудувати класифікуючу модель з використанням алгоритму *kNN*. Дана модель дозволила з високою ймовірністю (>80%) визначати, до якого з існуючих типів мережевих атак належать знайдені (використовуючи не розмічені дані), але не класифіковані кластери.

Отримані моделі пропонується використовувати в подальших дослідженнях в якості основи для обробки некласифікованої мережевої активності. А також для побудови багаторівневих моделей та методів виявлення мережевих атак та аномалій.

Література. **1.** KDD CUP 99, kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. **2.** Intrusion detection with unlabeled data using clustering http://cs.fit.edu/~pkc/id/related/portnoy-thesis00.pdf. **3.** Evaluation of K-Means Clustering for Effective Intrusion Detection and Prevention in Massive Network Traffic Data pdfs.semanticscholar.org/7d91/9cb6ddb47d4ce71a4163cc5a94bf31dab41e.pdf. **4.** Ester, Martin; Kriegel, Hans-Peter; Sander, Jörg; Xu, Xiaowei (1996). Simoudis, Evangelos; Han, Jiawei; Fayyad, Usama M., eds. A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press. pp. 226–231. **5.** MacQueen, J.B. (1967). Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. 1. University of California Press. pp. 281–297. **6.** Arthur, D.; Vassilvitskii, S. (2007). "k-means++: the advantages of careful seeding" (PDF). Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics Philadelphia, PA, USA. pp. 1027–1035. **7.** Dunn, J.C. (1973-01-01). "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters". Journal of Cybernetics. 3 (3): 32–57. **8.** Altman, N.S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression". The American Statistician. 46 (3): 175–185.

***Єремєєв І.С.*[1], *Дичко А.О.*[2]**
[1] *Таврійський національний університет імені В. І. Вернадського, Київ, Україна;* [2] *КПІ ім. Ігоря Сікорського, Київ, Україна*

Використання методів нечітких множин у процесах біологічного очищення стічних вод

Як відомо, не можна розглядати інформацію, не розглядаючи які-небудь ситуації невизначеності. Система біологічного очищення стічних вод (СВ) є класичним зразком системи, що працює в умовах невизначеності, оскільки її стани визначаються, як правило, не у реальному часі, як і зовнішні збурюючі чинники; існує щільний, але неоднозначний зв'язок між зовнішніми чинниками (температурою довкілля, атмосферним тиском, інтенсивністю опадів, часом, коли протікають процеси очищення тощо, причому ще й із значним часовим лагом між подіями та зміною станів системи). Управління такою системою має передбачати: спостереження станів та ступенів можливості цих станів; визначення станів, які є неможливими згідно з додатковою інформацією (наприклад, показники СВ не можуть бути кращими, або такими ж, як показники природних вод); передбачення станів, які не спостерігаються, але принципово можливі і яким надається ненульова ступінь можливості $fM(c)$, менша, ніж мінімальна ступінь можливості $fM(a)$, яка обчислюється для станів, що спостерігаються.

Можливість передбачення станів, які не спостерігаються, припускає наявність інформації ззовні системи, або інформаційний зв'язок між подією, явищем і станом системи. Такий зв'язок дійсно існує, але він не є однозначним і дозволяє лише на якісному рівні оцінювати можливі стани системи, причому із значною (до декількох десятків відсотків) похибкою.

Якщо скласти таблицю (отриману на підставі багаторічних досліджень очисних споруд), в яку входять страти (межі, всередині яких можуть спостерігатися характеристики СВ, з відповідними показниками імовірності спостереження); множина станів, які мають можливість спостереження міри нуль (межі, за які ніколи не виходять змінні, які спостерігаються); обчислені можливі стани (які реально не спостерігалися) і відповідні ступені можливості їхньої реалізації та імовірні наслідки цієї реалізації, то можна створити умови для управління процесами на підставі евристик.

Пропонується визначення приростання евклідової відстані між парами реальних станів, що спостерігаються і знаходяться поруч один до одного на цьому часовому інтервалі, а також імовірність та можливість такого переходу і рушійні сили, що спричиняють.

Для визначення та прогнозування впливу принципово можливих станів, які не можна зафіксувати онлайн, але які можуть вплинути глобально на систему та її показники, на першому етапі необхідно робити визначення на підставі даних, що є у розпорядженні (даних спостереження поточного процесу на певному інтервалі спостережень). Далі, на різних рівнях уточнення визначаються найкращі гіпотези щодо оцінок можливості реалізації інших станів узагальненої системи, формуються думки про те, наскільки ці гіпотези впливають на реальні властивості змінних, що розглядаються (ці думки формуються на підставі відповідних експериментальних характеристик та конкретних функцій). Кінець-кінцем, узагальнене обмеження, яке задано, доповнюється (або замінюється) обмеженнями, що реконструйовані за допомогою найкращих гіпотез, причому з кожною пов'язується певна ступінь довіри.

Рішення про необхідні регулюючі дії приймається на підставі вибору (методом Монте-Карло) тих, чи інших страт забруднювачів (меж існування) з урахуванням імовірності їхньої реалізації. Отриманий таким чином спектр віртуальних забруднень дасть змогу підібрати адекватні регулюючі дії для мінімізації цих забруднень. Імовірність Pe ефективної дії евристик:

$$Pe = (1/K) * \sum_{k=1}^{K} Pjk, \tag{1}$$

де Pjk – імовірність реалізації j-ї страти ($j = [I, J]$) k-го параметру (забруднювача), K – кількість параметрів (забруднювачів).

Забелин С.И.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Геометрические методы анализа Big Data

Вступление. Современная наука использует геометрические методы для нахождения структурных особенностей в массивах данных для дальнейшего более детального анализа. Математический формализм позволяет эффективно анализировать облачные массивы данных, т.е. конечный массив точек. Облако точек (англ. point cloud) – набор вершин в трёхмерной системе координат. Эти вершины, как правило, определяются координатами X, Y и Z и, как правило, предназначены для представления внешней поверхности объекта.

Big Data. Big Data находится везде, так как большие объёмы разнообразных четких и неопределенных данных могут быть собраны или сгенерированы с большой скоростью с помощью современных приложений. Взрывной рост облачных хранилищ, обработки данных, легкодоступность социальных сетей – всё это привело к необходимости фундаментального понимания процессов, которые порождают эти данные [1].

Линейные геометрические методы. Методы уменьшения размерности позволяют уменьшить число размерностей в идеале до 3D или 2D для более удобной визуализации данных. Одним из наиболее используемых является метод главных компонент (англ. principal component analysis, PCA). PCA – один из основных способов уменьшить размерность данных, потеряв наименьшее количество информации [2]. Вычисление главных компонент может быть сведено к вычислению сингулярного разложения матрицы данных или к вычислению собственных векторов и собственных значений ковариационной матрицы исходных данных.

Нелинейные геометрические методы. Последние исследования в области декомпозиции данных предоставили новые методы, которые позволяют обеспечить эффективное построение моделей. Одним из наиболее интересных является метод Isomap. Isomap – это комбинация уже существующих классических методов, используя их, Isomap строит взвешенную матрицу попарных расстояний между всеми точками модели [3]. А далее, используя многоразмерное шкалирование для расчета позиций точек в уменьшенном по размерности пространстве [4, 5].

Isomap Modified. В результате исследований был разработан модифицированный метод уменьшения размерностей многомерных наборов данных названный Isomap Modified. В его основе лежит использование парной системы уточнения, которая обеспечивает равномерное распределение точек на стадии уточнения метода Isomap. Как и в оригинальной системе уточнения, движение точки следует перпендикулярно градиенту стресс функции, при этом постоянно уменьшая шаг, но модифицированная схема не исключает возможности увеличения шага, для избегания локальных минимумов. Так же модифицированная схема производит изометрическое отображение, в отличие от оригинальной схемы [4].

Выводы. С экспоненциальным ростом объёма данных, Big Data, становится очевидной проблема построения новых моделей для обработки данных. Был разработан модифицированный Isomap для уменьшения размерностей многомерных наборов данных.

Литература. **1.** Wikipedia contributors. (2016, April 12). Big data [Online]. Available: en.wikipedia.org/wiki/Big_data. **2.** Wikipedia contributors. (2016, May 1). Principal component analysis [Online]. Available: en.wikipedia.org/wiki/Principal_component_analysis. **3.** Balasubramanian and E.L. Schwartz. The Isomap algorithm and topological stability. Science, 295, 2002. **4.** Cox, T. and Cox, M. Multidimensional Scaling. Chapman & Hall, London, 211, 1994. **5.** Roweis, S. and Saul, L. Nonlinear dimensionality reduction by locally linear embedding. Science, 290, 2000.

***Зайченко Ю.П.*[1], *Гамидов Г.И.*[2]**
[1]*КПИ им. Игоря Сикорского, ИПСА, Киев, Украина;* [2]*ОАО "АзерИшыг", Баку, Азербайджан*

Каскадные нейро-нечеткие сети в задачах прогнозирования на рынках ценных бумаг

Задачи прогнозирования курсов акций и биржевых индексов на рынках ценных бумаг в последние годы вызывают значительный интерес среди инвесторов и руководства иновационных фондов. Для прогнозирования на финансовых рынках традиционно использовались методы регрессионного анализа, метод экспоненциального сглаживания [1] и метод МГУА [2]. Однако в последние годы все шире для этих целей начинают использоваться системы с нечеткой логикой и нечеткие нейронные сети. Их достоинствами являются возможность работы с нечеткой и качественной информацией, а также использование экспертной информации в виде базы нечетких правил вывода.

Вместе с тем для организации применения ННС в задачах прогнозирования необходимо настроить базу правил и провести обучение параметров функций принадлежности нечетких правил вывода. Это сопряжено с большими вычислительными затратами и требует значительного времени обучения.

В последние годы появился новый класс ННС – каскадные нео-фаззи нейронные сети (CNFNN), их достоинством является отсутсвие необходимости обучения базы правил и параметров функций принадлежности, нечетких множеств, а обучаются по выборке только веса связей, что позволяет существенно сократить вычислительные затраты и время обучения [3], а также решать задачи большой размерности. Это дает возможность применения каскадных нео-фаззи сетей в задачах анализа высокоразмерных данных (Big Data Mining).

Целью работы является исследование CNFNN в задачах прогнозирования курсов акций и индексов на финансовых рынках. В качестве объекта исследовани й был выбран рынок ценных бумаг Германии. В процессе исследований варьировалось число входных переменных, число каскадов, число правил, а также соотношение между обучающей и проверочной выборками. В качестве критериев точности прогнозирования были выбраны СКО (MSE) и САПП (MAPE). В результате исследований было определено оптимальное число каскадов сети, чило входных переменных и рациональное число правил для данной задачи прогнозирования. Проведено сравнение эффективности прогнозирования с использованием каскадных нео-фаззи сетей и нейронной сети ANFIS. В докладе приводятся результаты экспериментальных исследований и обсуждаются направления дальнейших исследований.

Литература. **1.** Льюис К.Д. Методы прогнозирования экономических показателей: Пер. с англ. – М.: Финансы и статистика, 1986. – 133 с. **2.** Зайченко Ю.П. Нечеткий метод индуктивного моделирования в задачах прогнозирования макроэкономических показателей. // Системні дослідження та інформаційні технології. – 2003. – №3. – С. – 25–45. **3.** Ye. Bodyanskiy, Ye. Viktorov, I. Pliss The cascade NFNN and its learning algorithm // Вісник Ужгородського національного університету. Серія «Математика і інформатика». – 2008. – Вип. 17. – С. 48–58. **4.** Bodyanskiy Ye, Zaychenko Yu., Pavlikovskaya E., Samarina M., Viktorov Ye. Neo-fuzzy neural network structure optimization using GMDH for solving forecasting and classification problems // Proc. Int. Workshop on Inductive Modeling 2009. Krynica, Poland, 2009. – 77–89.

Зайченко Ю.П., Гула І.С.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Прогнозування ризику банкрутства банків в умовах невизначеності

Задача прогнозування ризику банкрутства банків є однією з найважливіших в банківській сфері. Для стабільного функціонування банку потрібен апарат для своєчасного виявлення банкрутства. В даний момент існує велика кількість методик для визначення рейтингу банків. Але практично всі вони працюють за повних та достовірних даних. У зв'язку з цим важливою являється розробка методів прогнозування, які працюють в умовах невизначеності. Особливо ці методи є актуальними для банків України, оскільки банки часто надають недостовірну інформацію про свої фінансові показники.

Банківські установи в процесі своєї діяльності мають справу із різними категоріями ризику, що керуються по-різному в залежності від цілей і стратегій. В залежності від різних критеріїв ризики можна первісно класифікувати по різному: згідно сфери виникнення на внутрішні і зовнішні; згідно етапу виникнення на етапі рішення та на етапі реалізації; згідно можливості диверсифікації на систематичний та специфічний. Причому кожен із видів може мати свою подальшу класифікацію. Слід зазначити, що зміна одного із численних видів ризику викликає зміну майже всіх інших видів ризику.

Для використання ННМ в задачах прогнозування потрібно побудувати базу правил і провести навчання параметрів функції приналежності. ННМ включають можливість використання знань експертів у вигляді нечітких правил виводу.

Ключовим питанням було, які саме показники фінансової звітності українських банків необхідно застосовувати в якості вхідних даних для моделей, щоб досягти максимальної ефективності.

Ціллю роботи є дослідження методів та методик аналізу фінансового стану і прогнозування ризику банкрутства банків на прикладі України за допомогою нечітких методів. Для дослідження були використані дані квартальної фінансової звітності банків, які взяті з сайту Національного банку України [4]. В процесі досліджень варіювалась кількість вхідних змінних, кількість правил, а також відношення між навчальною та тестовою вибірками. В результаті досліджень було визначено оптимальну кількість вхідних змінних та раціональну кількість правил для даної задачі прогнозування. Був проведений порівняльний аналіз результатів прогнозу з використанням ННМ ANFIS та TSK. Доповідь містить міркування стосовно напрямків подальших досліджень.

Литература. **1.** Зайченко Ю.П., Оценка кредитных банковских рисков с использованием нечеткой логики // Системні дослідження та інформаційні технології. К: ВПК «Політехніка» ВПІ НТУУ «КПІ» – 2010. – №2. – С. 37–54. **2.** Згуровский М.З., Зайченко Ю.П. Основы вычислительного интеллекта. – К.: Изд « Наукова Думка»,. 2013.-406 с. **3.** Bodyanskiy Ye, Zaychenko Yu., Pavlikovskaya E., Samarina M., Viktorov Ye. Neo-fuzzy neural network structure optimization using GMDH for solving forecasting and classification problems // Proc. Int. Workshop on Inductive Modeling 2009. Krynica, Poland, 2009. – 77–89. **4.** https://bank.gov.ua/control/uk.

Заковоротний О.Ю., Харченко А.О.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Нейромережеве прогнозування швидкості в системах підтримки прийняття рішень

Одним з поширених методів розв'язання задач прогнозування є використання штучних нейронних мереж. При цьому поведінка певного процесу найчастіше перетворюється у часовий ряд та потім, за допомогою нейронної мережі, прогнозується поведінка отриманого ряду. Задача прогнозування в нашому випадку необхідна для отримання оптимальних значень швидкості на наступних ділянках шляху та перевірки стійкості руху в системі підтримки прийняття рішень машиністом. Однак, в даній задачі ми зіштовхуємося з неперіодичними процесами, що відносно швидко змінюються в часі, а саме – відбувається втрата стійкості при русі на нерівних ділянках шляху [1].

Для підвищення якості прогнозу необхідно проводити попередню обробку інформації. Кожна вибірка часового ряду представляє собою дискретну функцію, що містить інформацію про параметри профілю шляху. При розв'язанні задач прогнозування, роль нейронної мережі полягає в отриманні майбутньої реакції системи на основі її попередньої поведінки. Використовуючи інформацію про значення змінної x в попередні моменти часу $x_{(t-1)}$, $x_{(t-2)}$, ..., $x_{(t-n)}$, мережа отримує рішення про найбільш ймовірне значення послідовності $x(t)$ на даний момент часу t. Для адаптації вагових коефіцієнтів мережі використовується фактична похибка прогнозування та значення даної похибки в попередні моменти часу.

При розробці архітектури нейронної мережі та алгоритму стосовно її навчання, використаємо МГУА-нейронні мережі, основу яких складають ідеї методу групового урахування аргументів О.Г. Івахненко [2]. На рис. 1 наведено МГУА-нейронну мережу для нашої задачі.

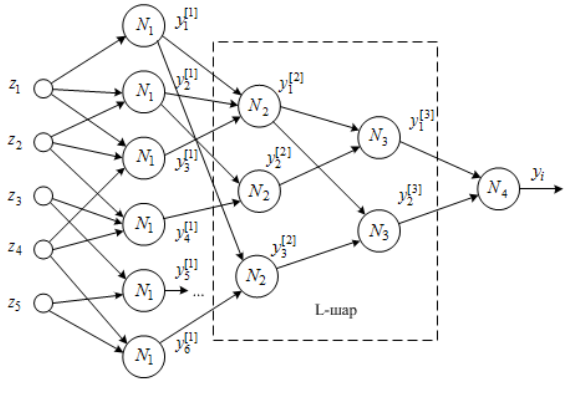


Рис. 1. Нейронна мережа на основі методу групового урахування аргументів

На відміну від нейронних мереж з фіксованою архітектурою, типу класичного багатошарового персептрону, МГУА-мережа має змінну структуру, яка може змінюватися в процесі навчання.

Системи управління, що використовують штучні нейронні мережі, є однією з можливих альтернатив класичним методам управління. Можливість використання нейронних мереж для розв'язання задач керування в цілому базується на можливості апроксимації будь-якої функції дійсних чисел з заданою степінню точністю. Дане положення доведено за теоремою Вейєрштраса. Таким чином для розв'язання задач керування можуть бути використані нейронні мережі навіть з одним прихованим шаром.

Література. **1.** Дмитриенко В.Д. Моделирование и оптимизация процессов управления движением дизель-поездов / В.Д. Дмитриенко, А.Ю. Заковоротный. – Х.: Изд. центр «НТМТ», 2013. – 248 с. **2.** Ивахненко А.Г. Системы эвристической самоорганизации в технической кибернетике / А.Г. Ивахненко – Киев: Техника, 1971. – 372 с.

Зражевська Н.Г.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Моделювання і прогнозування мір ризику VaR і CVaR для часового ряду фондових індексів

Метою даної роботи є отримання прогнозних оцінок динамічних мір ризику часового ряду логарифмічної доходності на денній основі індексу РТС за 2014 рік (часовий ряд *RTSI*) за запропонованою авторами методикою [1]. Для отримання прогнозу динамічних мір ризику *VaR* і *CVaR* був реалізований наступний алгоритм:

1. Для часового ряду *RTSI* будуємо часовий ряд оцінок дисперсій (ЧРД).

2. Доводимо сильну залежність ЧРД з використанням п'яти стандартних методів [1, 2]. В якості оцінки параметра Херста вибираємо агреговане значення $\widehat{H}_{mean} = 0,8692$.

3. Аналіз ЧРД на залежність членів ряду від своїх попередніх значень з використанням тесту Льюнг-Боксу [2] підтверджує наявність залежності. Таким чином, для моделювання ЧРД обираємо модель *FIGARCH* [1].

4. На основі гетероскедастичної моделі часового ряду виду:

$$X_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t Z_t, iid,$$

оцінюємо динамічні міри ризику [2]:

$VaR^t_\alpha = \mu_t + F^{-1}(\alpha)\sigma_t = \mu_t + VaR_\alpha(Z)\sigma_t$, $CVaR^t_\alpha = \mu_t + CVaR_\alpha(Z)\sigma_t$.

Отже, прогнозування динамічних мір ризику на P періодів вимагає побудови моделі прогнозування σ_t на P періодів, а також обчислення $VaR_\alpha(Z)$ і $CVaR_\alpha(Z)$ для випадкової величини Z, що має такий же розподіл, що і будь-яка випадкова величина із $\{Z_t\}$:

$$VaR^{t+p}_\alpha = \mu_{t+p} + VaR_\alpha(Z)\sigma_{t+p}, CVaR^{t+p}_\alpha = \mu_{t+p} + CVaR_\alpha(Z)\sigma_{t+p} \quad (1)$$

Для розглянутого часового ряду тренд відсутній (близький до нуля).

5. Виконуємо актуалізацію моделі шляхом зведення її до моделі $AR(\infty)$. Знаходимо розв'язок редукованої системи рівнянь Юла-Уолкера, де для знаходження коефіцієнтів використовується запропонована авторами модель АКФ [1]. Визначаємо лаг редукованої авторегресійної моделі, використовуючи інформаційні критерії Акаіке, Хеннена-Куїна і Байеса: $M = 55$. Перевіряємо, що залишки моделі належать класу *iid* (критерій відношення дисперсій). За отриманою моделлю здійснюємо прогнозування σ_t:

$$\sigma^2_{l+p} = \sum_{i=o}^{M} \widehat{a}_i \sigma^2_{l-i-1}, l = N+1, ..., p = 1, ..., P.$$

6. Знаходимо реалізації $\widehat{Z}_t = X_t/\widehat{\sigma}_t$, проводимо їх аналіз на *iid* і нормальність розподілу. Для знахождення оцінок $\widehat{VaR_\alpha(Z)}$ і $\widehat{CVaR_\alpha(Z)}$ використовуємо явні формули, отримані з використанням функції розподілу *GPD*, параметри функції оцінюються по ММП.

7. Будуємо прогноз по (1), $P = 5$, вікно з накопиченням. Застосовуючи бектестінг, проводимо кількісний (похибки ME, MAE. MSE) і якісний аналізи (тести Купеца, Крістофферсена і V-тест) отриманих оцінок.

У роботі на прикладі побудови прогнозу динамічних мір ризику для часового ряду *RTSI* продемонстрований, запропонований авторами, алгоритм побудови прогнозу мір ризику *VaR* і *CVaR* для часових рядів на основі гетероскедастичної моделі часового ряду. Розглянута процедура прогнозування добре себе показала для даного числового прикладу, що підтверджено тестами, проведеними на різних етапах алгоритму.

Література. **1.** Zrazhevska, N.,& Zrazhevsky, A. (2017). Construction of prediction of dynamic risk measures var and cvar for financial time series with different volatility. European Journal of Basic and Applied Sciences, 4, 13–22. **2.** Tsay, R.S. (2010). Analysis of Financial Time Series. Hoboken: John Wiley & Sons, Inc.

Коляденко А.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Система підтримки прийняття рішень комплексного обслуговання кол-центру

Система підтримки прийняття рішень комплексного обслуговання кол-центру призначена для спрощення діяльності кол-центру та покращення його діяльності. Створення такої системи є актуальним, оскільки в Україні активно розвиваються як власні кол-центри компаній, так і аутсорсингові контакт-центри, які базуються в Україні, але приймають зарубіжні дзвінки.

Задача створення системи підтримки прийняття рішень комплексного обслуговання кол-центру полягає в тому, що потрібно створити таку систему, щоб обслуговування дзвінків клієнтів відбувалося вчасно та в повному обсязі. Дана задача зводиться до того, щоб визначити таку кількість операторів кол-центру, яка здатна ефективно обслуговувати всі дзвінки, що надійшли.

Дана задача складається з 3 етапів:

1. Прогнозування кількості дзвінків, що надійдуть в кол-центр.
2. Визначення мінімальної кількості операторів, що необхідно для обслуговування даної кількості дзвінків.
3. Розроблення графіка виходу операторів на роботу.

Перший етап. На першому етапі завдання – спрогнозувати кількість дзвінків. Прогноз потрібно мати для загальної кількості дзвінків на наступний місяць, на кожен день цього місяця, та на кожну годину кожного дня місяця.

Розглянемо фактори, що можуть впливати на кількість дзвінків.

1. Помісячна сезонність – кількість вхідних дзвінків змінюється із врахуванням сезонності. Сезонність залежить від діяльності компанії, яка обслуговує колл-центр. Конкретну залежність слід визначати, виходячи із наявних історичних даних. Для коректного визначення сезонності потрібні дані за 3 повних історичних періоди.
2. День тижня – кількість вхідних дзвінків значно залежить від дня тижня. Конкретна залежність може сильно варіюватись із специфікою діяльності компанії, яка обслуговує колл-центр. Зазвичай, на понеділок приходиться пік дзвінків, у неділю йде значний спад. Конкретну залежність слід визначати, виходячи із наявних історичних даних.
3. Година дня – кількість вхідних дзвінків значно залежить від дня тижня. Яскраво виражена залежність – вночі дзвінків завжди менше, ніж удень. Конкретну залежність слід визначати, виходячи із наявних історичних даних.
4. Святковий день – кількість вхідних дзвінків значно зменшується в свята. А на наступний день після них йде незначний приріст. Також святковий день значно впливає на ті години, на які приходиться під викликів. Наприклад, на новий рік значно більша кількість дзвінків в нічні години, аніж зазвичай.
5. Акція – якщо компаніє планує певну акцію, або технічне обслуговування, що зачепить велику кількість клієнтів компанії, то цілком прогнозованим є значне підвищення кількості вхідних викликів. Історичні дані допомагають при прогнозуванні того, наскільки зміниться потік дзвінків.
6. Форс-мажор. Даний фактор впливу неможливо передбачити завчасно, але потрібно закласти ці дані в місячний прогноз.

На першому етапі порівнюються існуючі методи прогнозування. Розглядається згладжування даних і виділення лінії тренду, прогнозування на один крок за допомогою експоненціального згладжування та прогнозування з використанням моделі ARIMA [1,2]. Досліджується точність прогнозу за різними підходами – щоденний, щотижневий, щомісячний. На основі найточнішого щоденного прогнозу складається погодинний прогноз.

Другий етап. Другий етап полягає у знаходженні необхідного числа операторів, що здатні обслужити отриману на попередньому етапі кількість дзвінків.

Розглянемо фактори, що впливають на роботу операторів.

1. Тривалість дзвінка. Тривалість дзвінка – один із найбільш фундаментальних факторів, що впливає на роботу усього кол-центру. Тривалість дзвінка сама, в свою чергу, залежить від багатьох факторів.
2. Показник *Service Level*. Service Level – це відношення вірогідності того, що клієнт дочекається з'єднання з оператором, на час очікування з'єднання клієнта із оператором.
3. Постобробка дзвінка. Оператори мають не лише відповісти на дзвінок, а і внести відповідні дані в інформаційну систему. Час на постобробку зазвичай коливається в межах 5–30 секунд на один дзвінок.
4. Час з'єднання – це час, який витрачається на з'єднання клієнта із оператором.
5. Перерви в роботі оператора.
6. Форс-мажорні обставнини.

Припустимо, що потік дзвінків можна описати, як *потік Пуассона*. Пропонується перейти від одного великого нестаціонарного потоку до багатьох менших, кожен з яких буде відповідати за в межах однієї години, і кожен з яких буде стаціонарним. За таких умов можливе застосування формули *Ерланга-С* [3]

$$E(m,u) = \frac{\frac{u^m}{m!}}{\frac{u^m}{m!} + (1 - \frac{u}{m}) \sum_{k=0}^{m} \frac{u^k}{k!}}, \tag{1}$$

де m – кількість операторів; u – інтенсивність потоку Пуассона.

Задача знаходження кількості операторів зводиться до пошуку такого їх числа, при якому показник service level, що знахидиться за формулою Ерланга, був би максимально наближеним до заданого аналітиком.

Третій етап. Третій етап – знаходження розкладу, за яким оператори ходитимуть на роботу в кол-центр. Обчислення попереднього розкладу відбувається за допомогою методу одного з методів лінійного програмування, а саме – лінійного симплекс-методу. Розглянемо дану задачу лінійного програмування [4]

$$F = a_{0,1}x_1 + a_{0,2}x_2 + ... + a_{0,n}x_n + b_0 \to \min.$$
$$a_{1,1}x_1 + a_{1,2}x_2 + ... + a_{1,n}x_n \geqslant b_1$$
$$...$$
$$a_{m,1}x_1 + a_{m,2}x_2 + ... + a_{m,n}x_n \geqslant b_m$$
$$x_1 ... x_n \in \mathbb{Z}$$

Отримаємо оптимальний розклад, що враховує як умову мінімізації кількості операторів, і також є зручним для співробітників колл-центру. Цей розклад аналітик зможе доповнити за необхідності.

Висновки. Розглянута задача створення системи підтримки прийняття рішень комплексного обслуговання кол-центру, показано її актуальність. Для даної задачі показано математичний апарат.

Література. **1.** Бідюк П.І. Аналіз часових рядів / П.І. Бідюк, В.Д. Романенко, О.Л. Тимощук. – К.: Політехніка, 2010. – 317 с. **2.** Бідюк П.І. Часові ряди: моделювання і прогнозування / П.І. Бідюк, О.І. Савенков, І.В. Баклан. – К.: ЕКМО, 2003. – 144 с. **3.** Вентцель Е.С., Овчаров Л.А. Теория вероятностей. М., 1969, 368 стр. **4.** Зайченко Ю.П. Дослідження операцій. – Київ: ЗАТ "ВІПОЛ", 2000. – 688 с.

Кондратенко Н.Р., Снігур О.О.
Вінницький національний технічний університет, Вінниця, Україна

Моделювання багатовимірних залежностей на основі нечітких множин типу 2 на недовизначених гідрогеологічних даних

Недовизначеність вхідних даних – одна з проблем, характерних для моделювання складних природних систем. До таких належать екологічні, соціальні, економічні, технічні системи різної природи, для математичного опису функціонування яких складно побудувати один аналітичний закон [1].

Виділяють низку т. зв. НЕ-факторів [2], що мають визначальний вплив на адекватність моделі, побудованої за вибіркою експериментальних даних, як-то неточність вимірювань, відсутність можливостей для безпосереднього спостереження об'єкта, неповнота та неоднозначність знань про предметну галузь і розв'язувану задачу, вплив неврахованих (прихованих) параметрів, брак експертних знань про предметну галузь або неможливість їх формалізувати, а також невизначеність, пов'язана з розмірністю простору вхідних ознак (надлишковість та зашумленість) [2, 3]. Всі ці фактори тією чи іншою мірою властиві природним системам та процесам. Особливості моделювання систем цього класу показано на прикладі оцінювання стану артезіанської свердловини в довільний момент часу в межах від початку гідрогеологічної розвідки до її повного завершення.

Задача оцінювання запасів підземних вод характеризується важкодоступністю експериментальних даних, оскільки отримання вхідних даних, необхідних для роботи будь-якої моделі, вимагає виконання значного обсягу робіт. Невипадково тривалість геологорозвідувальних робіт, що передують введенню свердловини в експлуатацію, складає від 6 місяців до кількох років [4]. Тому важливо протягом усього періоду розвідки володіти якнайповнішими знаннями про стан об'єкта, щоб у разі виникнення труднощів мати змогу переглянути рішення щодо місцерозташування свердловини, використовуваних технологій, провести додаткові спостереження тощо. Задача вимагає побудови моделей, що адекватно відображають об'єкт на будь-якому етапі його дослідження, навіть за умови неповноти або недовизначеності вхідних даних. Це зумовлює необхідність у розробці моделей та методів їх розв'язання за допомогою технологій інтелектуального аналізу даних. Моделі цього класу, зокрема математичний апарат кластерного аналізу та нечітких множин, надають потужний інструментарій для опису та моделювання складних систем, що не піддаються повній формалізації. Підземні води в цілому та артезіанська свердловина як частинний випадок належать до таких систем.

Метою дослідження є розширення можливостей існуючих нечітких моделей та методів прийняття рішень в умовах недовизначеності вхідних даних. Побудовано систему нечітких моделей, що об'єднує в собі як переваги системи підтримки прийняття рішень на основі інтервального нечіткого логічного висновку, так і технологій інтелектуального аналізу даних [5].

Вихідний вектор експериментальних даних, що може містити пропуски, подається на вхід інтервальної нечіткої логічної системи типу 2 без змін, а також пройшовши через процедуру виділення інформативних ознак. Зазначена процедура може виконуватись одним або кількома експертами в предметній галузі. У разі, коли експертів декілька, кожен із них генерує свій набір ознак, та як наслідок окрему модель. В такому поданні описаний раніше метод автоматизованого виділення ознак з точки зору моделі виступає в ролі ще одного експерта. Інтервальний вихід нечіткої логічної системи за повним вхідним вектором та виходи результуючих моделей після відкидання неінформативних ознак об'єднуються за агрегуючим критерієм Y_A:

$$Y_A = Y_{full} \bigcap \left(\bigcup_{i=1}^{n} Y_i \right). \quad (1)$$

Запропонована система моделей допускає існування альтернативних моделей на основі інших технологій прийняття рішень або формальних процедур, що вже існують для розв'язання поставленої задачі. В рамках даної роботи показано процедуру прийняття рішень на основі методу кластерного аналізу для поточного моніторингу стану свердловини в ході її експлуатації (альтернативна модель 1) та прямого оцінювання стану свердловини на основі законодавства України (альтернативна модель 2). У випадку, коли інтервальний вихід альтернативної моделі має таку саму розмірність і якісну природу, що й інтервали на виході нечіткої логічної системи, він також враховується при побудові агрегованого інтервалу за критерієм Y_A. В іншому разі виходи альтернативних моделей подаються користувачу системи окремими інтервалами, незалежно від основного виходу.

Досліджено роботу множини моделей, утворених в результаті виділення інформативних ознак експертами, та в результаті автоматизованого виділення ознак. Показано, що в ряді випадків уже на початку геологорозвідувальних робіт можна зрозуміти, що в подальшій роботі на вибраній ділянці очікуються ускладнення. На основі цієї інформації можуть бути проведені додаткові дослідження та прийняте рішення про припинення робіт, що дозволить зберегти суттєві матеріальні та людські ресурси (табл. 1).

Табл. 1. Експеримент – виявлення труднощів на початковому етапі

№ етапу ГРР	Експерт 1 Визн. ознак/ всього: Вихід моделі	Експерт 2 Визн. ознак/ всього: Вихід моделі	Автом. метод Визн. ознак/ всього: Вихід моделі	Повний вектор Визн. ознак/ всього: Вихід моделі	Вихід альтерн. моделі 1	Вихід альтерн. моделі 2	Агрегований вихід
1	9/39: [1, 39; 1, 51]	8/41: [0, 43; 0, 45]	9/32: [1, 37; 1, 42]	18/84: [1, 39; 4, 51]	x	x	[1, 39; 1, 51]
2	16/39: [0, 24; 0, 27]	19/41: [0, 01; 0, 01]	16/32: [0, 09; 0, 03]	35/84: [0, 24; 6, 27]	x	x	[0, 24; 0, 27]
3	39/39: [0, 09; 3, 97]	41/41: [0, 43; 0, 69]	32/32: [0, 23; 0, 33]	84/84: [3, 49; 8, 7]	H:[0, 59; 0, 86]	недостатня	[3, 49; 3, 97]

Результати дослідження потенційно можуть бути застосовані в довготривалих програмах вивчення природних процесів задля скорочення терміну прийняття рішення та заощадження необхідних для цього ресурсів [1].

Література. **1.** Kondratenko N. Interval Fuzzy Modeling of Complex Systems under Conditions of Input Data Uncertainty / N. Kondratenko, O. Snihur // Eastern-European Journal of Enterprise Technologies. – vol. 4/4 (82). – 2016. – Pp. 20–28. **2.** Бардачев Ю.Н. Методологическая предпочтительность интервальных экспертных оценок при принятии решений в условиях неопределенности / Ю.Н. Бардачев, В.В. Крючковский, Т.В. Маломуж // Вісник Харківського національного університету. – № 890. – 2010. – С. 18–28. **3.** Панкратова Н.Д. Системный анализ и оценивание динамики экологических процессов / Н.Д. Панкратова, В.В. Заводник // Системні дослідження та інформаційні технології. – 2004. – № 2. – С. 47–59. **4.** Боревский Б.В. Оценка запасов подземных вод. / Б.В. Боревский, Н.И. Дробноход, Л.С. Язвин – 2-е изд., перераб. и доп. – К.: Выща шк. Головное изд-во, 1989. – 407 с.: ил. **5.** Кондратенко Н.Р. Використання інтервальних нечітких множин типу 2 в умовах недовизначеності вхідних даних / Н.Р. Кондратенко, О.О. Снігур // Системний аналіз та інформаційні технології: матеріали 18-ї Міжнародної науково-технічної конференції SAIT 2016, Київ, 30 травня – 2 червня 2016 р. / ННК "ІПСА" НТУУ "КПІ". – К.: ННК "ІПСА" НТУУ "КПІ", 2016. – С. 224–225.

Кравець І.О.
Чорноморський національний університет імені Петра Могили, Миколаїв, Україна

Дослідження чітких та нечітких нейронних мереж для ідентифікації та керування слабоформалізованими об'єктами

Багато об'єктів керування є специфічними, і для них неможливе використання стандартних методів керування, які використовуються для автоматичного керування лінійних технічних об'єктів. Для таких об'єктів характерні таки особливості: математичний опис моделі відсутній, або частково відсутній, не стаціонарність моделі, необхідность швидкої адаптації до змін зовнішнього середовища. Вхідні та вихідні характеристики системи описуються нестаціонарними часовими рядами.

Для ідентифікації та керування такими об'єктами краще всього використовувати системи штучного інтелекту, насамперед нейронні мережі завдяки їх можливості навчання та пристосування к змінам зовнішнього середовища.

Застосування нейронних мереж для розв'язання задач управління дозволяють виділити два етапи проектування: етап ідентифікації керованого процесу та етап синтезу регулятора. Так як модель об'єкта невідома, то у системі керування використовується адаптивне керування. Для контролера на основі еталонної моделі використовується непряме адаптивне управління та дві нейронних мережі, одна – для ідентифікації та передбачення поведінки об'єкту, друга – у складі регулятора. Для задач ідентифікації та адаптивного керування можливо використовувати рекурентні нейронні мережі типу NARX з затримкою у часі між виходом мережі і входом та з часовою затримкою вхідних даних. Але у режимі адаптивного керування необхідним є процес перенавчання нейронних мереж у процесі роботи, тому зручним є використання нечітких нейронних мереж. Для нейрорегуляторів на базі нечітких нейронних мереж використано нечітку нейронну мережу з логічним висновком Сугено, як найбільш точну.

У роботі проведено порівняний аналіз використання чітких та нечітких нейронних мереж для адаптивного керування економічним об'єктом. В результаті дослідження нейромережевих регуляторів слабоформалізованими об'єктами було встановлено.

1. Для об'єктів, які можуть бути представлені рівнянням стану (диференціальними або різницевими з невідомими коефіцієнтами) можливе використання нейромережевих регуляторів як на базі чітких рекурентних нейронних мереж, так і нечітких нейронних мереж з висновком Сугено.
2. Для об'єктів, які можуть бути представлені тільки базою правил, кращим є використання нечітких нейронних мереж з висновком Сугено та попередня кластеризація вхідних даних, яка дозволяє значно зменшити базу правил.

Література. **1.** Hagan M.T., Demuth H.B. Neural networks for control. //Proceedings of the 1999 American Control Conference. SanDiego:CA1999. P. 311–340. **2.** Зайченко Ю.П. Нечеткие модели и методи в интеллектуальних системах. – К., "Слово", 2008. – 333 с. **3.** Simon Haykin. Neural Networks. A comprehensive foundation. Second Edition. McMaster University Hamilton, Ontario Canada. // Prentice Hall Upper Saddle River New Jersey 2006. **4.** Takagi T., Sugeno M. Fuzzy identification of system and its application for modeling and control. //IEEE Trans. Systems, Man and Cybernetics. 1995 V.15 N1 P. 116–132.

Куца К.В.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Анализ кредитоспособности малого и среднего бизнеса в Украине с использованием нечеткой логики

Актуальность задачи. В настоящий момент наиболее перспективными субъектами экономики Украины являются малые и средние предпринимательства. Данный сегмент рынка занимает важное место в развитии экономики страны в целом. По оценкам Национального Банка Украины, в 2017 году в первую очередь будет расти кредитование малого и среднего бизнеса (МСБ). Выдавая кредит субъектам МСБ, банк рискует выдать кредит потенциальному банкроту. Кредитору приходится принимать решение о выдаче кредита в условиях неопределенности и недостоверной информации. Банкам необходимо адаптировать кредитную стратегию с учетом специфики клиентских сегментов.

Таблица 1. Определение МСБ (для Украины) [3].

Категория компании	Количество сотрудников	Оборот
Средний бизнес	$\leqslant$250 чел.	$\leqslant$ € 50 млн.
Малый бизнес	$\leqslant$50 чел.	$\leqslant$ € 10 млн.
Микро бизнес	$\leqslant$10 чел.	$\leqslant$ € 2 млн.

Постановка задачи. Введем переменные, описывающие кредитоспособность предприятия [2]: Кредитная история(x1)=(хорошая, плохая, нет). Финансово-экономическое состояние(x2)= (стабильность, отрицательные тенденции, положительные тенденции). Вероятность несвоевременного погашения(x3)=(низкая, средняя, высокая, очень высокая). Риск убытков(x4)= (высокий, средний, низкий). Документация(x5)=(плохая, нормальная, хорошая). Банкрот(x6)=(да, нет). Обеспечение(x7)=(первоклассное, ликвидное, неликвидное, отсутствует, возможность проблем). Платежеспособность(x8)=(высокая, средняя, низкая). Необходимо определить уровень кредитоспособности заемщика с помощью методов нечеткого логического вывода.

Решение задачи. В качестве алгоритма нечеткого вывода будем использовать алгоритм Мамдани. Механизм нечеткого вывода в своей основе имеет базу знаний, формируемую экспертами в виде совокупности нечетких правил вида:

$$\Pi i\text{: если } x1 \text{ есть } Ai,\ x2 \text{ есть } Bi,\ \ldots,x8 \text{ есть } Hi,\text{ то } x9 \text{ есть } Zi;\ x1\in X1,\ldots,x9\in X9,$$

где x1, ..., x8 – имена входных переменных, x9 – имя переменной вывода. Ai, Bi,...,Zi – функции принадлежности, заданные на множествах X1, ..., X9 соответственно. На первом этапе (фаззификация) находятся степени истинности для предпосылок каждого правила. На втором этапе находятся уровни «отсечения» для предпосылок каждого из правил. Затем находятся «усеченные» функции принадлежности. Третий этап – композиция. Производится объединение найденных усеченных функций. Четвертый этап – дефаззификация. Выполняется процесс перехода от функции принадлежности выходной переменной к ее числовому значению (например, центроидным методом) [1].

Результаты работы. На основании вышеперечисленных критериев с использованием нечетких правил выводится итоговая оценка кредитоспособности заемщика. Т.о., мы можем сделать вывод о выдаче кредита заемщику в условиях неопределенности и недостоверной информации.

Литература. **1.** Зайченко Ю.П. Нечеткие модели и методы в интеллектуальных системах. – К.: Изд. дом «Слово». 2008. – 92 с. **2.** Зайченко Ю.П. Оценка кредитных банковских рисков с использованием нечеткой логики // Системные исследования и информационные технологии. 2010, №2. – С. 44–47. **3.** Закон Украины «О развитии и государственной поддержке малого и среднего предпринимательства в Украине» 22.03.2012 №4618-VI.

Мурга М.О.
КПІ ім. Ігоря Сікорського, Київ, Україна

Побудова обгортки процесу в просторі зі змінним базисом приналежностей ознакам нестаціонарних процесів

Доповідь присвячена представленню результатів розробки методу побудови границі області руху кривої належностей процесу у просторі зі змінним базисом приналежностей ознакам.

Простір зі змінним базисом приналежностей ознакам нестаціонарних процесів. В роботі [1] представлено теоретичні результати та експериментальні дослідження нового методу прогнозування, здебільшого, фінансово-економічних процесів, який ґрунтується на припущенні про існування декількох їх моделей одночасно, прогностична здатність яких з часом змінюється (зазвичай, в гіршу сторону), але є сама прогнозованою. Як вказується в роботі, прийняття рішення про включення нової моделі/ознаки до базису можливо на основі динаміки перетинань границі руху кривої приналежностей процесу, яка в роботі названа обгорткою процесу. Таким чином, актуальним є питання розробки методу побудови обгортки процесу, якому і присвячено дану доповідь.

Метод побудови обгортки процесу у просторі приналежностей зі змінним базисом. Проблема при побудові обгортки (G – у позначеннях роботи [1]) полягає у наявності великої кількості шумових чинників реальних процесів, які призводять до непрогнозованих великих амплітуд викидів значень їх показників, які не доцільно включати до області простору, яку обмежує обгортка процесу. Таким чином, обчислення обгортки, ґрунтуючись на максимальних значеннях амплітуд, є недоцільним. Робити ж це пропонується наступним чином.

Нехай $g(t, w_1, ..., w_j, f_1, ..., f_j) = \{g_k(t, w_1, ..., w_j, f_1, ..., f_j), k = \overline{1, K}\}$.

Нехай $g_k(t, w_1, ..., w_j, f_1, ..., f_j) \leqslant 0$ у випадку, коли значення кривої належності не перетинає границю g_k або лежить на ній та $g_k(t, w_1, ..., w_j, f_1, ..., f_j) > 0$ в іншому випадку. Нехай $I(t) = 0$, якщо $\forall k = \overline{1, K} g_k(t, w_1, ..., w_j, f_1, ..., f_j) \leqslant 0$ та $I(t) = 1$ в іншому випадку.

Нехай інтервал $[0; T]$ поділено на M рівних за величиною інтервалів $[0; \tau], [\tau; 2 \cdot \tau]; ..., [(M - 1)\tau; T]$.

Нехай $g^m_{mean} = \frac{1}{\tau} \int_{t \in [(m-1)\cdot\tau; m\cdot\tau]} \frac{1}{K} \sum_{k=1}^{K} g_k(t, w_1, ..., w_j, f_1, ..., f_j)$.

Тоді задачу побудови обгортки процесу в просторі приналежностей можна сформулювати наступним чином:

$$\begin{gathered} min \sum_{m=1}^{M} \int_{t \in [(m-1)\cdot\tau; m\cdot\tau]} \sum_{k=1}^{K} (g_k(t, w_1, ..., w_j, f_1, ..., f_j) - g^m_{mean})^2, min \sum_{m=1}^{M-1} \varepsilon_m, \\ \sum_{t \in [(m-1)\cdot\tau; m\cdot\tau]} I(t) - \varepsilon_m \leqslant \sum_{t \in [m\cdot\tau; (m+1)\cdot\tau]} I(t), \\ \sum_{t \in [m\cdot\tau; (m+1)\cdot\tau]} I(t) \leqslant \sum_{t \in [(m-1)\cdot\tau; m\cdot\tau]} I(t) + \varepsilon_m, \forall m = 1, \bar{M} - 1, \\ \varepsilon_m \geqslant 0 \end{gathered} \tag{1}$$

Оскільки на практиці t належить дискретній множині, до дана задача є задачою багатокритеріального нелінійного дискретного програмування і може бути вирішена, наприклад, методом штрафних функцій з використанням квазіградієнтів, або генетичним алгоритмом.

Література. 1. Мурга М.О. Простір зі змінним базисом приналежностей ознакам для прогнозування показників фінансово-економічних процесів / М.О. Мурга // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. пр. – К.: Век+, – 2016. – № 64. – С. 138–144.

***Мысник Б.В.*[1], *Снитюк В.Е.*[2]**
[1] *Черкасский государственный технологический университет, Черкассы, Украина;* [2]*Киевский национальный университет им. Тараса Шевченко, Киев, Украина*

Мультиагентная оптимизация функционирования предприятий отрасли. Восстановление пропусков в данных

Экономическое развитие страны неразрывно связано с государственной поддержкой малого и среднего бизнеса. Его часть в ВВП составляет 50–60%, что сравнимо с данными европейских стран. Увеличение такой части ВВП по сравнению с 2014 г. (до 15 %) приводит к возрастающей конкуренции. Особенно такая конкуренция ощутима среди предприятий отраслей, производящих однотипную продукцию. Для городов – это производство мебели, окон, дверей, инфраструктурная деятельность и т.п. Перед руководителями таких предприятий возникают задачи развития, расширения, ликвидации или перепрофилизации. Принятие соответствующих решений происходит в критических условиях или условиях неопределенности, что зачастую уменьшает эффективность их результатов.

Ранее [1] предложено для объективизации принимаемых решений использовать мультиагентный подход. Его главная идея заключается в рассмотрении каждого предприятия как некоторого интеллектуального агента, имеющего структуру, элементный базис и функционирующего по определенной программе. Взаимодействие предприятия с внешним миром формализуется в виде входящих и исходящих факторов и характеристик. Их численные значения заносятся в базу данных, которая идеализированно считается релевантной всем предприятиям отрасли. Дискретность данных определяется временем осуществления любой транзакции на любом предприятии.

Известно, что эффективность функционирования предприятия является функцией показателей эффективности выполняемых отдельных функций. И показатели эффективности, и критерий эффективности являются некоторыми константами или моделями. В предположении, что их идентификация выполнена, можно решить задачу прогнозирования показателей эффективности и критерия эффективности функционирования предприятия как функциональных зависимостей от показателей деятельности других предприятий и времени. Такие данные динамично накапливаются в базе данных. Полученные модели могут использоваться лицом, принимающим решения, для анализа типа "Если $x \in A$, то $y \in ?$" и предназначены для поддержки принятия решений.

Формирование базы данных сопряжено с проблемой инсайдерского характера информации о деятельности предприятий. Частично такую информацию можно получить из открытых источников: ассортимент продукции, акции и скидки, элементы и комплектующие, затраты на рекламу и т.п. В силу разных причин зачастую время транзакций является известным, но неизвестны значения соответствующих параметров или известны значения не всех параметров. В таких случаях рационально осуществлять восстановление или прогнозирование отсутствующих данных, исходя из той информации, которая есть на данный момент времени.

Для решения задачи осуществим дальнейшее развитие и частично модифицируем метод, предложенный в [2]. Проведенный анализ методов восстановления пропусков в таблицах данных свидетельствует об их низкой точности, что связано с предположениями о линейной зависимости и корреляции факторов. В реальных задачах эти предположения чаще всего неверны. Рассмотрим подход, в котором такие допущения не играют никакой роли.

Итак, предположим, что данные сосредоточены в таблице $\Omega = \{X_1, X_2, ..., X_n, Y\}$, где X_i – факторы, $i = \overline{1,n}$, Y – зависимая характеристика, количество строк в таблице Ω соответствует количеству наблюдений или экспериментов и равно k. Пусть l строк в таблице не содержат пропусков, соответственно, в $k-l$ строках пропуски присутствуют. Тогда таблицу Ω разделим

на Ω_k и Ω_o, где Ω_k – контрольная (комплектная) таблица, не содержащая пропуски, Ω_o – таблица, содержащая данные с пропусками и используемая далее для обучения. Во всей таблице в столбце Y пропусков нет.

Главная идея исследования заключается в комбинировании нейросетевых технологий идентификации неизвестных зависимостей и их эволюционной оптимизации. Будем считать, что таблицы Ω_k и Ω_o содержат хотя бы по одной строке каждая (при определенных условиях). Представим пропуски из Ω_o в виде вектора $d = (d_1, d_2, ..., d_p)$, который в дальнейшем и будет потенциальным решением задачи, $d_i \in [a_i, b_i]$, где a_i и b_i определяются лицом, принимающим решения, $i = \overline{1,p}$. Возможно еще несколько подходов к их определению. В первом случае для каждого столбца X_i матрицы Ω находят x_{imin} и x_{imax} – минимальный и максимальный элемент столбца данных, $i = \overline{1,n}$. Далее достаточно положить $a_i = x_{imin} - \varepsilon$, $b_i = x_{imax} + \varepsilon$, где ε – некоторое заданное положительное число. Во втором случае $a_i = \overline{x_i} - 3\sigma_i$, $b_i = \overline{x_i} + 3\sigma_i$, где σ_i – выборочное среднеквадратическое отклонение.

Определив координаты вершин прямоугольного гиперпараллелепипеда – области поиска решения рассматриваемой задачи, генерируем q потенциальных векторов-решений $d^1, d^2, ..., d^q$ На следующем шаге производим обучение q нейросетей, где в качестве исходных данных используются данные таблицы Ω_0 со сгенерированными значениями вместо пропусков. Обучение нейросети заканчивается, если значение $E_0^j = \frac{1}{k-l}\sum_{i=1}^{k-i}(y_i - \widehat{y}_i)^2$ будет меньшим некоторого заданного значения, y_i – табличные значения, $\widehat{y}_i$ – значения, вычисленные нейросетью, $j = \overline{1,q}$.

Далее на вход обученной нейросети подаются данные из таблицы Ω_k и вычисляются значения $E_k^j = \frac{1}{k}\sum_{i=1}^{k}(y_i - \widehat{y}_i)^2$, которые и будут значениями фитнесс-функции в эволюционном (генетическом) алгоритме, $j = \overline{1,q}$. Используемая фитнесс-функция не является единственной. Она преимущественно отражает точность прогнозирования на контрольной выборке. В то же время, в обучающей и контрольной выборке могут содержаться данные, которые не только отличаются наличием или отсутствием пропусков, но и имеющие разную внутреннюю структуру и особенности. В качестве примера можно указать разные величины дисперсий обучающей и контрольной выборки. Осуществить учет особенностей выборок частично можно, используя в качестве фитнесс-функции аддитивную свертку $E = E_0^j + E_k^j = w_1 * \frac{1}{k-l}\sum_{i=1}^{k-l}(y_i - \widehat{y}_i)^2 + w_1 * \frac{1}{k}\sum_{i=1}^{k}(y_i - \widehat{y}_i)^2, w_1 + w_2 = 1$. Весовые коэффициенты определяются экспертом и указывают на важность величины ошибки на обучающей и контрольной выборках. На этапе предварительного анализа их значения могут быть определены пропорционально величине дисперсии соответствующих выборок. Однако заметим, что необходимы дополнительные исследования и не исключено, что необходимо учитывать особенности решаемой задачи и наличие выбросов в данных.

В докладе будут рассмотрены результаты экспериментов, полученные при разном процентном соотношении пропусков в таблице данных, при взвешенном использовании критериев E_0^i и E_k^j. Исследована проблема и условия разрешимости задачи восстановления пропусков в данных, если они содержатся в каждой строке.

Литература. **1.** Voloshyn O. Intellectual Information Support of Branch Enterprise Executives Decision Making Processes / O. Voloshyn, B. Mysnyk, V. Snytyuk // International Journal "Information Theories & Applications", 2014. – Vol. 21, № 4. – P. 303–113 (ISSN 1310-0513). **2.** Снитюк В.Е. Эволюционный метод восстановления пропусков в данных / В.Е. Снитюк // Сборник трудов VI-й Межд. конф. "Интеллектуальный анализ информации". – Киев. – 2006. – С. 262–271.

Перова И.Г., Бодянский Е.В.
Харьковский национальный университет радиоэлектроники, Харьков, Украина

Об одной метрике в задачах интеллектуального анализа медицинских данных

В задачах обработки медицинских данных часто возникает необходимость кодирования результатов наблюдений, которые изначально измерены в разных шкалах. В докладе рассмотрены основные особенности кодирования медицинских данных различного происхождения с целью дальнейшей их обработки.

В медицинских выборках данных информация часто представлена в виде так называемой таблицы «объект-свойство», в которой строкам соответствуют пациенты, а столбцам – признаки, которые их характеризуют. В эту таблицу информация может поступать из абсолютно разных источников и быть уже предобработанной по ранее заданным, а иногда неизвестным шкалам. В любом случае перед какой-либо обработкой этой таблицы (прежде всего, с помощью методов вычислительного интеллекта [1]) необходимо провести преобразование всех входящих в нее данных для чего предлагается использовать унифицированный подход, описанный ниже. Предположим, что таблица «объект-свойство» представляет из себя матрицу $X = \{x_{ij}\}$ размерностью $N \times n$, где N – количество пациентов, а n – количество характеризующих их признаков. Пусть в этой таблице признаки измеряются в разных шкалах: категориальной (номинальной или бинарной), ранговой (порядковой) и числовой (интервальной и отношений). Таким образом каждый из пациентов описывается вектором $x_i = (x_{i1}, \ldots, x_{ij}, \ldots, x_{in})^T \in R^n$. Для кодирования информации кроме центрирования и нормирования необходимо, чтобы все данные принадлежали одному интервалу, то есть $x_{jmin} \leqslant x_{ij} \leqslant x_{jmax}$ В качестве минимального и максимального диапазонов удобно использовать значения $x_{min} = 0$ и $x_{max} = 1$, таким образом, чтобы все данные попадали в гиперкуб $[0; 1]^n$.

Количественная метрика. Для количественных (числовых) данных обычно используется метрика Минковского $d_{ij}^{num} = (\frac{1}{N}\sum_{j=1}^{N}|x_{ij} - x_{lj}|^p)^{\frac{1}{p}}$, автоматически совпадающая с эвклидовой при $p = 2$ и манхэттенской при $p = 1$, для всех $i, l = 1, 2, \ldots, n$.

Ранговая метрика. Для ранговых переменных предполагается, что каждый признак x_i имеет R_i рангов, $r_i = 1, 2, \ldots, R_i$, что означает, что объектом анализа является не собственно x_{ij}, а соответствующий ранг $x_{ij}^{r_i}$ для всех $j = 1, 2, \ldots, N$. Далее в рассмотрение вводится распределение частот $f_i^{r_i} = \frac{N_i^{r_i}}{N}$, где $N_i^{r_i}$ – число появления r_i-го ранга в i-м столбце. Далее вводятся в рассмотрение их частотные числовые характеристики вида $x_{ij}^1 = 0,5 f_i^1$, $x_{ij}^{r_i} = x_{ij}^{r_{i-1}} + 0,5(f_i^{r_{i-1}} - f_i^{r_i})$, которые кодируются на гиперкуб $x_{ijr_i} = \frac{x_{ij}^{r_i} - x_{ij}^1}{x_{ij}^{R_i} - x_{ij}^1}$. Таким образом, расстояние между x_{ij} и x_{lj} записывается в форме $0 \leqslant d_{ij} = \frac{1}{N}\sum_{j=1}^{N}|x_{ijr_i} - x_{ljr_l}| \leqslant 1$.

Категориальная метрика. В качестве категориальной метрики используется выражение

$$0 \leqslant d_{ij}^{cat} = \frac{1}{N}\sum_{j=1}^{N}\delta(x_{ij}, x_{lj}) \leqslant 1;\ \delta(x_{ij}, x_{lj}) = \begin{cases} 1, & \text{если } x_{ij} \neq x_{lj}, \\ 0, & \text{если } x_{ij} = x_{lj}. \end{cases}$$

Далее несложно ввести в рассмотрение общее расстояние между объектами анализа x_i и x_l в виде $0 \leqslant d_{ij} = (\frac{1}{n_{num}} d_{ij}^{num} + \frac{1}{n_{rank}} d_{ij}^{rank} + \frac{1}{n_{cat}} d_{ij}^{cat}) \leqslant 1$, где n_{num}, n_{rank}, n_{cat} – количество числовых, ранговых и категориальных переменных соответственно в исходной таблице «объект-свойство».

Предобработанная таким образом информация далее может быть проанализирована стандартными методами Medical Data Mining [2], включая классификацию, кластеризацию, идентификацию и прогнозирование.

Литература. **1.** L. Ruthowski Computational Intelligence. Methods and Techniques, Berlin-Heidelberg: Springer-Verlag, 2008. – 514 p. **2.** C. Fernandez-Llatas, J.-M. Garcia-Gomez Data Mining in Clinical Medicine Humana Press: Springer, 2015. – 270 p.

Поворознюк А.И., Филатова А.Е.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Полутоновая морфологическая фильтрация на основе локальных статистик

Постановка проблемы и анализ литературы. Большая часть медицинской информации, получаемая в результате инструментальных методов исследования, регистрируется в виде рентгеновских изображений, таких как маммограммы, флюорограммы, томограммы, рентгенограммы грудной клетки, конечностей и т.д. От качества визуализации исследуемых органов и опорно-двигательного аппарата напрямую зависит точность и своевременность диагностики. Поэтому задача повышения качества визуализации рентгенологических изображений за счет цифровой обработки изображений, в частности с помощью морфологического анализа, является актуальной научно-практической задачей.

В наиболее общем смысле Пытьев Ю.П. определил морфологический анализ как группу методов решения задач узнавания, классификации объектов, оценки параметров объектов, выделения различия в сценах по их изображениям (сигналам), основанных на понятии формы сигнала [1]. Еще более общее определение морфологического анализа дал в [2] Визильтер Ю.В. Морфологический анализ – схема анализа данных, которая в качестве обязательного этапа предполагает обоснованное (в некотором смысле оптимальное) построение модельного описания гипотетического (скрытого) прообраза наблюдаемых данных. В работе [3] предложен компаративный фильтр на основе взаимного контрастирования следующего вида:

$$\varphi_a^w(f,g)[x,y] = g_0^{w[x,y]}[x,y] + a(f, g^{w[x,y]})(g[x,y] - g_0^{w[x,y]}[x,y]), \quad (1)$$

где $g_0^{w[x,y]}[x,y]$ – локальное среднее значение изображения $g[x,y]$ в окне $w[x,y]$; $a(f, g^{w[x,y]}) \in [0;1]$ – локальный коэффициент взаимного сходства фрагмента изображения $g[x,y]$ с элементами изображения f, такой, что $a(g, g^{w[x,y]}) = 1$ и $a(o, g^{w[x,y]}) = 0$, если $o[x,y] = const$.

В качестве коэффициента $a(f, g^{w[x,y]})$ могут использоваться локальный нормированный и поисковый коэффициенты корреляции, локальный и поисковый морфологические коэффициенты корреляции Пытьева, «тепловые» ядра взаимного сходства фрагментов изображения на основе сравнения векторов локальных признаков и другие подобные меры сходства изображений. Независимо от такого, каким способом вычислен коэффициент $a(f, g^{w[x,y]})$, для применения фильтра (1) необходимо иметь эталонное изображение или модель формы, так как в [3] в качестве коэффициента используются различные локальные и поисковые коэффициенты геометрической корреляции форм.

Поэтому для решения задачи повышения качества визуализации биологических объектов на рентгенологических изображениях такие фильтры непригодны, так как из-за высокой вариабельности форм объектов или их структурных элементов, которые могут анализироваться на рентгенологических снимках, невозможно задать эталонное изображение или модель формы.

Цель работы. Разработка метода морфологической фильтрации полутоновых изображений для решения задачи повышения качества визуализации биомедицинских рентгеновских изображений.

Разработка полутонового морфологического фильтра на основе локальных статистик. Рассмотрим адаптивный фильтр Винера (АФВ), отклик $\phi[x,y]$ которого определяется следующими выражениями:

$$\phi[x,y] = \mu[x,y] + K(g^{w[x,y]})(g[x,y] - \mu[x,y]); \quad (2)$$

$$K(g^{w[x,y]}) = \begin{cases} \dfrac{\sigma^2[x,y] - \nu^2}{\sigma^2[x,y]}, & \sigma^2[x,y] > \nu^2; \\ 0, & \text{в противном случае;} \end{cases}$$

где $\mu[x,y] = \frac{1}{|\eta|} \sum\limits_{(x',y')\in\eta} g[x',y']$ – локальное среднее изображения $g[x,y]$ в окне $w[x,y]$ размером $|\eta|$ пикселей; η – множество точек изображения $g[x,y]$, попавших в окно фильтра $w[x,y]$; $\sigma^2[x,y] = \frac{1}{|\eta|-1} \sum\limits_{(x',y')\in\eta} (g[x',y'] - \mu[x,y])^2$ – локальная дисперсия изображения $g[x,y]$ в окне $w[x,y]$; ν^2 – дисперсия шума. Если дисперсия шума неизвестна, то она вычисляется как среднее значение локальных дисперсий изображения $g[x,y]$ как $\nu^2 = \frac{1}{|M|} \sum\limits_{(x,y)\in M} \sigma^2[x,y]$, где M – множество точек изображения $g[x,y]$.

Очевидно, что в выражениях (1) и (2) $g_0^{w[x,y]}[x,y] = \mu[x,y]$. Легко заметить, что $K(g^{w[x,y]}) = 0$, если $g[x,y] = o[x,y]$ или изображение $g[x,y]$ не содержит полезного сигнала, а лишь шумовую составляющую, т.е. $\sigma^2[x,y] \leqslant \nu^2$. При этом $K(g^{w[x,y]}) = 1$, если изображение $g[x,y]$ не содержит шум, т.е. $\nu^2 = 0$. Т.о., коэффициент $K(g^{w[x,y]}) \in [0;1]$, и его свойства аналогичны свойствам коэффициента $a(f, g^{w[x,y]})$. Следовательно, АФВ (2), по сути, является морфологическим фильтром. При этом модель формы задается не явно, а косвенно за счет вычисления локальных статистик сигнала, попавшего в окно фильтра. Т.о., форма полезного сигнала в окне определяется законом распределения яркости пикселей, попавших в окно морфологического фильтра. В (2) в качестве таких статистик используются локальное среднее $\mu[x,y]$, локальная дисперсия $\sigma^2[x,y]$ фильтруемого изображения, а также дисперсия шума ν^2.

Однако такие статистики корректно использовать, если закон распределения яркостей пикселей, попавших в окно, подчиняется закону Гаусса. Для реальных сигналов закон распределения яркостей пикселей неизвестен, а при относительно малых размерах окна фильтра предположение о нормальном законе распределения яркостей пикселей фильтруемого сигнала может быть некорректным. Поэтому предлагается использовать такие локальные статистики, вычисление которых не требует знание закона распределения случайно величины. В работе предлагается использовать соответствующие локальные квантили, которые определяются по локальной выборке, т.е. по значениям пикселей, попавших в окно фильтра. Тогда отклик предлагаемого морфологического фильтра $\phi_m[x,y]$ определяется следующими выражениями:

$$\phi_m[x,y] = q_{0.5}[x,y] + K_m(g^{w[x,y]})(g[x,y] - q_{0.5}[x,y]); \tag{3}$$

$$K_m(g^{w[x,y]}) = \begin{cases} \dfrac{\Delta q[x,y] - \Delta q_{\text{ср}}}{\Delta q[x,y]}, & \Delta q[x,y] > \Delta q_{\text{ср}}; \\ 0, & \text{в противном случае;} \end{cases}$$

где $q_{0.5}[x,y]$ – локальная медиана изображения $g[x,y]$ в окне $w[x,y]$ размером $|\eta|$ пикселей; $\Delta q[x,y] = q_b[x,y] - q_a[x,y]$ – локальное межквантильное расстояние в окне $w[x,y]$; q_a, q_b – локальные a-квантиль и b-квантиль изображения $g[x,y]$ в окне $w[x,y]$, причем $a < b$; $\Delta q_{\text{ср}} = \frac{1}{|M|} \sum\limits_{(x,y)\in M} \Delta q[x,y]$ – среднее локальное межквантильное расстояние.

Фильтр, отклик которого определяется по (3), получил название полутонового морфологического фильтра (ПМФ) на основе локальных статистик (ЛС).

Выводы. В работе разработан ПМФ на основе ЛС, который позволяет без необходимости знания закона распределения полезного сигнала выполнять фильтрацию полутонового изображения с учетом формы объекта. Разработанный ПМФ на основе ЛС может быть использован для решения задачи повышения качества биомедицинских рентгенологических изображений.

Литература. **1.** Пытьев Ю.П. Методы морфологического анализа изображений / Ю.П. Пытьев, А.И. Чуличков. – М.: Физматлит, 2010. – 336 с. **2.** Визильтер Ю.В. Теория и методы морфологического анализа изображений: автореф. дис. д-ра физ.-мат. наук: 05.13.17 / ГНЦ ФГУП ГосНИИАС. – М., 2008. – 34 с. **3.** Лебедев М.А. Выделение отличий на мозаичных изображениях на основе референтных фильтров / М.А. Лебедев, А.Ю. Рубис, В.С Горбацевич, Ю.В. Визильтер, О.В. Выглов // Компьютерная оптика. – 2016. – Т. 40, № 1. – С. 80–86.

Поворознюк Н.І., Гритчук Д.Т.
КПІ ім. Ігоря Сікорського, ФЕА, Київ, Україна

Вейвлет-аналіз сигналів для моніторингу якості електроенергії

Вступ. Сучасний стан електроенергетики характеризують дві тенденції, які впливають на якість електроенергії. З одного боку спостерігається широке запровадження перетворювачів електричної енергії з ключовими режимами роботи їх елементів, а також нелінійних навантажень, які спотворюють форму кривих струмів і напруг. З іншого боку зростає частка споживачів, які чутливі до спотворень у мережі живлення. Наприклад, комп'ютерні і мікропроцесорні системи, технологічні лінії, вимірювальні прилади і системи. Низька якість електроенергії в електричних мережах, що живлять такі чутливі споживачі, призводить до значних економічних втрат [1].

Рішення задачі. Щоб знизити втрати від низької якості електроенергії, здійснюється моніторинг мереж електроживлення. Одним з найважливіших завдань систем моніторингу є визначення і класифікація показників якості електроенергії, які нормуються міжнародними [2–5] і вітчизняними [6] стандартами. На першому етапі здійснюється збір первинної інформації про стан мережі електроживлення за допомогою системи датчиків. Сигнали з датчиків нормуються і перетворюються у цифровий вид. Щоб значно зменшити обсяг оброблюваної інформації здійснюється сегментація сигналів, тобто виділення лише спотворених ділянок сигналів. Виділені у результаті сегментації спотворені ділянки сигналів підлягають подальшій обробці, у ході якої здійснюється вилучення найважливіших параметрів, які є діагностичними ознаками виду спотворень. Такі сигнали є нестаціонарними і несинусоїдними, тому їх доцільно аналізувати як у часовій, так і у частотній області. Останнім часом розроблено багато методів представлення і аналізу сигналів у частотно-часовій області [7]. Одним з таких методів є вейвлет-аналіз сигналів. На відміну від Фур'є-аналізу, у якому сигнал розкладається по системі синусоїдних функцій нескінченої тривалості і тому неможливо відстежувати зміну параметрів у часі, у вейвлет-аналізі використовується функція скінченої тривалості, яка називається материнською. Застосування цієї функції для аналізу розтягнутих і стиснутих у часі копій досліджуваного сигналу дає змогу виявити спектральний склад сигналу, а переміщення материнської функції по осі часу – виявити зміну параметрів сигналу у часі. Для дослідження якості електроенергії у системах електропостачання у середовищі Matlab/Simulink були синтезовані сигнали з типовими спотвореннями напруг електромережі. У тому ж середовищі була створена програма вейвлет-аналізу цих спотворень.

Висновок. Результати комп'ютерного моделювання показали високу роздільну здатність вейвлет-аналізу у частотно-часовій області, що дає змогу отримати широкий спектр діагностичних ознак спотворень і тим самим істотно підвищити точність їх ідентифікації.

Література. **1.** Chattopadhyay, S.; Mitra, M.; Sengupta, S. Electric Power Quality, 1st ed.; Springer: New York, NY, USA, 2011. **2.** IEEE Std 1159–2009, IEEE Recommended Practice for Monitoring Electric Power Quality; IEEE Power & Energy Society: New York, NY, USA, 2009. **3.** CEI/IEC 61000–4-30 International Standard. Testing and Measurement Techniques–Power Quality Measurement Methods, 1st ed.; International Electrotechnical Commission: Geneva, Switzerland, 2003. **4.** EN 50160: 2007: Voltage characteristics of electricitysupplied by public distribu-tion systems. **5.** IEEE Std 1459-2010: IEEE Standard Definitions for the Measurement of Electric Power Quantities Under Sinusoidal, Nonsinusoidal, Balanced, or Unbalanced Conditions, Piscataway, USA, March 2010. **6.** ГОСТ 13109-97 "Нормы качества электрической энергии в системах электроснабжения общего назначения". **7.** A. Augustine et al. Review of Signal Processing Techniques for Detection of Power Quality Events / American Journal of Engineering and Applied Sciences 2016, 9 (2): pp. 364–370.

Савельев Ю.Д.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Анализ методов распознавания лиц с помощью нейронных сетей

Распознавание лиц на основании данных, полученных средствами захвата изображений, различных сенсоров и баз данных, является серьёзной проблемой из-за большого количества вариантов особенностей лица, освещения и фона изображения. Но механизм данный механизм находит применение в различных приложениях, а задача является одной из основополагающих в комъютерном зрении. Ниже представлен обзор различных актуальных подходов к распознаванию лиц, методов и алгоритмов с использованием искусственных нейронных сетей для выделения ключевых особенностей, применение которых является одним из наиболее эффективных подходов обработке изображений и распознавании образов в последние года. Рассмотрено то, как используются нейронные сети для задач распознавания лица, кратко описаны особенности различных методов и их общий обзор.

Структура системы распознавания лиц. Любой из рассмотренных алгоритмов базируется на архитектуре, в основе которой лежит следующие четыре этапа:

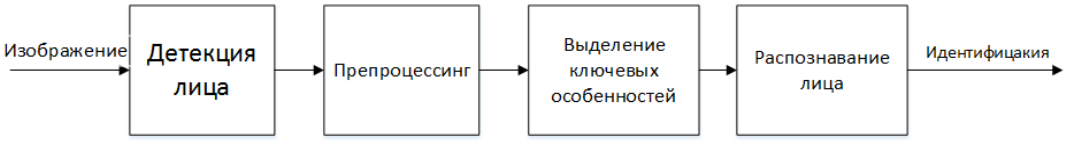


Рис. 1. Общая структура системы распознавания лиц

Первым шагом является детекция лица из входящего изображения. Суть этапа состоит в подтверждении, имеет ли входное изображение лицо, и локализация этой области. После этого эта область подаётся на препроцессинг.

Во время препроцессинга изображение лишается нежелательного шума, размытия, различных условий освещённости, теней с помощью различных техник. Необходимость его применения определяется подходом, используемым для выделения ключевых особенностей.

На третьем этапе происходит выделение ключевых особенностей с помощью выбранного для этого подхода. Результатом этого чаще всего является вектор фиксированного разрешения или набор определённых точек с соответствующим расположением.

На четвёртом этапе происходит анализ отображений каждого лица. Здесь происходит распознавание личностей на основе имеющейся базы данных.

Преимущество использования нейронных сетей. Задача выделения лица человека в различных условиях и последующей идентификации была одной из первых и она стимулировала развитие машинного зрения, теории распознавания объектов. Тем не менее, множество исследований, проводящихся в ведущих научных центрах всего мира в течение нескольких десятилетий, так и не привело к созданию реально работающих систем компьютерного зрения, способных обнаруживать и распознавать человека в любых условиях. Наиболее сложной задачей компьютерного зрения является проблема устранения неоднозначности, возникающей при проектировании трехмерных объектов реального мира на плоские изображения. Цвет и яркость отдельных пикселей на изображении также зависит от большого количества трудно прогнозируемых факторов. В число этих факторов входят:

1. число и расположение источников света;
2. цвет и интенсивность излучения;
3. тени или отражение от окружающих объектов.

Также проблемой является различные углы обзора лица. Многие популярные алгоритмы, например Виолы–Джонса, не обладают необходимой гибкостью и при угле наклона больше 30 градусов их точность существенно падает. И это не позволяет в стандартной реализации детектировать повернутое лицо человека под произвольным углом, что в значительной мере затрудняет или делает невозможным использование алгоритма в современных производственных системах с учетом их растущих потребностей. Нейронные сети позволяют избежать выше приведённых проблем с выделением ключевых точек при достаточной глубине, большой выборке данных и правильном походе к обучению. Чем больше выборка данных, которая имеется у вас для обучения, чем более разнообразная она, тем более гибкую и универсальную модель вы сможете обучить. Тем не менее, проблема получения выборки является основной при использовании нейронных сетей, так как объёмы необходимых данных для обучения действительно велики.

PCA с использованием искусственных нейронных сетей. PCA (метод главных компонент) [1]. Это метод уменьшения размерности, основанный на преобразовании Карунена–Лоева. В задаче распознавания лиц его применяют главным образом для представления изображения лица вектором малой размерности (главных компонент), который сравнивается затем с эталонными векторами, заложенными в базу данных.

Главной целью метода главных компонент является значительное уменьшение размерности пространства признаков таким образом, чтобы оно как можно лучше описывало «типичные» образы, принадлежащие множеству лиц. Используя этот метод можно выявить различные изменчивости в обучающей выборке изображений лиц и описать эту изменчивость в базисе нескольких ортогональных векторов, которые называются собственными (eigenface). В данном подходе для каждого лица происходит обучение нейронной сети на основе eigenface, которая в дальнейшем будет определять схожесть лиц. Метод имеет высокую точность, но недостатком является то, что для каждого лица нужна своя нейронная сеть.

Deep Dense Face Detector. Данный алгоритм базируется на глубокой свёрточной нейронной сети и служит для детекции лиц [2]. Основными идеями является высокая способность сетей в классификации и выделении особенностей для построения единого классификатора для детекции лиц и минимизация вычислительной сложности за счёт упрощения архитектуры детектора. Автор метода использовал архитектуру AlexNet и обучил её для задачи детекции именно лиц. Лица одного и того ж класса (человека) в разных условиях служили обучающей выборкой. Лица были модифицированы с помощью различных методов аугментации данных для увеличения размеров самой выборки

Детектор анализирует оценки DDFD детектора и ищет кореляцию между этими оценками и распределением положительных примеров в обучающей выборке. Метод показал себя хорошо во многих случаях и условиях, был устойчив к различным типам освещения, хотя процент точности распознавания значительно упал при сильном повороте лица. Хотя автор указал возможные пути решения данной проблемы с помощью различных подходов к работе с входными данными.

Недостатком является необходимость полного переобучения сети для добавления новых лиц.

Radial Basis Function Neural Networks. Данный подход отлично себя зарекомендовал и имеет одни из самых высоких показателей в точности [3]. Как часть препроцессинга, используется гибридный подход на базе Модели Активной Формы (Active Shape Models) и PCA (Principal Component Analysis). С помощью ASM происходит выделение контура и форм лица для создания персонального профиля, а после используется PCA алгоритм, чтобы уменьшить

размерность отображений лиц.

В распознавательной части используются Radial Basis Function Neural Networks (RBF NN) для идентификации уникального шаблона ассоциированного с каждым человеком. Данная архитектура состоит из трёх модулей, занимающихся фазой состояния, фазой вывода и фазой заключения, которая работает с помощью правил нечёткой логики взамен стандартных if-then условий.

Выводы. Несмотря на то, что проблема распознавания лиц является одной из старейших в компьютерном зрении и распознавании образов, она ещё не решена до сих пор. Наиболее сложной задачей компьютерного зрения является проблема устранения неоднозначности, возникающей при проектировании трехмерных объектов реального мира на плоские изображения. Цвет и яркость отдельных пикселей на изображении также зависит от большого количества трудно прогнозируемых факторов. Поэтому методы, которые сильно зависят от этих условий, не являются эффективными.

Применение для решения проблемы распознавания лиц нейронных сетей вывело задачу на новый уровень. Абсолютно разные варианты их использования для решения разных подзадач проблемы, от генерации структурной модели лица, подхода к задаче распознавания как к задаче классификации до использования классификаторов с нечёткой логикой на базе нейронных сетей показывает их универсальность. Варианты компоновки классических подходов в компьютерном зрении с нейронными сетями также даёт большое количество возможных вариантов решения этой проблемы. И именно подобные методы кажутся наиболее перспективными в решении поставленной проблемы.

Таблица 1. Сравнение средней точности методов распознавания

Метод	Точность в процентах
PCA with ANN Face recognition system	95.45
DDFD	91.79
RBNNs	97.56
CNN	85.1
BPN+RBF	98.88
RCNN	90.3
RINN	90.6
Gabor Wavelet Faces with ANN	93

Литература. **1.** N. Jindal, V. Kumar, “Enhanced Face Recognition Algorithm using PCA with Artificial Neural Networks”, International Journal of Advanced Research in Computer Science and Software Engineering, vol 3, no. 6, (2013), pp. 864–872. **2.** S. Sudhakar Farfade, M. Saberian, L.-J. Li, “Multi-view Face Detection Using Deep Convolutional Neural Networks”, International Conference on Multimedia Retrieval, Shanghai, China, (2015). **3.** S.-H. Yooa, S.-K. Oha, Witold Pedrycz,” Optimized face recognition algorithm using radial basis function neural networks and its practical applications”, International journal on Neural Networks, volume 69, (2015), pp. 111–125.

***Сащук І.М.*[1], *Чепков І.Б.*[2]**
[1] *Науково-дослідний інститут Міністерства оборони, Київ, Україна;* [2] *Центральний науково-дослідний інститут озброєння і військової техніки Збройних Сил України, Київ, Україна*

Методика визначення інформативних ознак розпізнавання стану об'єктів моніторингу

Завдання моніторингу різноманітних об'єктів, подій і явищ (далі – об'єкт спостереження ОС) з метою визначення змін у їх стані, поведінці, характері діяльності тощо постають у багатьох сферах людської діяльності. До таких завдань, наприклад, належать діагностування організму людини в медицині, моніторинг потенційно небезпечних природних явищ та об'єктів техногенного характеру, контроль за використанням радіочастотного ресурсу. В [1] показано взаємозв'язок діагностування з завданнями розпізнавання образів. Для розпізнавання стану та визначення змін у характері діяльності ОС в [2] запропоновано використовувати інформаційно-сигнатурні технології (ІСТ) – прийоми, способи та методи автоматизованого збору, обробки, накопичення, зберігання та передачі інформації про ОС на основі виділення та аналізу його сигнатури $\boldsymbol{S}_{OC}$. Під частинними сигнатурами s_i, $i = \overline{1,N}$ довільного ОС розумітимемо характеристики, описи, параметри, які і утворюють множину $\boldsymbol{S}_{OC} = \{s_1, s_2, ..., s_i, ..., s_N\}$ – його повну сигнатуру. Визначення елементів S_{OC} здійснюється за результатами моніторингу параметричними або іконічними (відео-, фотозображення) методами діагностування шляхом формування множини характерних ознак $\boldsymbol{X}^{(M)}$, M – кількість елементів в $\boldsymbol{X}^{(M)}$. Залежно від складності ОС, його функціональності, впливу на оточуюче середовище та конкретних завдань моніторингу ті чи інші характерні ознаки можуть бути більше або менше інформативними (корисними). З огляду застосування ІСТ, природним є прагнення відбору найбільш інформативних (для виконання конкретного завдання) ознак, які складатимуть відповідну підмножину (підпростір) ознак $\boldsymbol{C}^{(P)} = \{s_1, s_2, ..., s_p, ..., s_P\}$, $P \leqslant M$, $\boldsymbol{C}^{(P)} \in \boldsymbol{X}^{(M)}$ та забезпечуватимуть розпізнавання стану ОС із заданими показниками якості. Постановку завдання визначення інформативних характерних ознак можна описати таким чином:

- зменшення до мінімуму кількості необхідних для опису класів ознак без суттєвого погіршення показників якості розпізнавання;
- можливість використання відносно простих і нересурсоємних алгоритмів розпізнавання, що особливо актуально для іконічних методів, первинними інформаційними матеріалами (вхідними даними) для яких є файли великого формату [3];
- забезпечення розпізнавання з найкращою якістю.

Серед численних методів розпізнавання (методи на основі байєсівського підходу, методи статистичних рішень, методи розділення у просторі ознак, логічні методи розпізнавання та інші) найбільшого поширення набули методи статистичних рішень. Для них якість розпізнавання характеризується ймовірністю правильного розпізнавання P_0 або ймовірністю помилки розпізнавання $P_e = 1 - P_0$.

Формування множини корисних характерних ознак розпізнавання здійснюватимемо за їх інформативністю. В загальному випадку міру інформативності $I_k(L_u)$ k-ої характерної ознаки x_k з множини $\boldsymbol{X}^{(M)}$ щодо розпізнавання u-го стану деякого ОС L_u можна оцінити як різницю початкової ентропії рішення $H_0(L_u)$ про стан ОС та ентропії рішення про стан ОС за цією характерною ознакою

$$I_k(L_u) = H_0(L_u) - H(L_u/x_k),$$

де

$$H_0(L_u) = -p(L_u)\log_2 p(L_u);$$

$$H(L_u/x_k) = \sum_{j=1}^{J_k} p(x_{kj}) H(L_u/x_{kj}) = -\sum_{j=1}^{J_k}\sum_{u=1}^{U} p(L_u, x_{kj}) \times$$

$$\times \log_2 p(L_u, x_{kj}) + \sum_{j=1}^{J_k} \sum_{u=1}^{U} p(L_u, x_{kj}) \sum_{u=1}^{U} \log_2 p(L_u, x_{kj});$$

$$p(L_u, x_{kj}) = p(L_u) p(x_{kj}/L_u);$$

$p(L_u)$ – апріорна ймовірність розпізнавання u-го стану ОС, $p(x_{kj})$ – апріорна ймовірність появи характерної ознаки x_k зі значенням j, $p(x_{kj}/L_u)$ – умовна ймовірність правильного визначення j-го значення k-ої характерної ознаки в стані ОС L_u.

Для параметричних методів розпізнавання практично важливим є визначення ймовірнісних характеристик розпізнавання станів ОС по довільному параметру λ. Безумовний розподіл параметра λ описується виразом

$$P(\lambda) = \sum_{u=1}^{U} q_u P_u(\lambda),$$

де q_u – апріорна ймовірність наявності параметра λ для u-того стану ОС, $P_u(\lambda)$ – розподіл ймовірності появи параметра λ в u-тому стані ОС.

Враховуючи багатомодальність $P(\lambda)$ у випадку розпізнавання за одним параметром, для рівноймовірних станів ОС можна записати

$$P(\lambda) = \frac{1}{2\sigma\sqrt{2\pi}} \left(\exp\left(-\frac{(\lambda - m - (2l-1)\Delta m)^2}{2\sigma^2} \right) + \exp\left(-\frac{(\lambda - m + (2l-1)\Delta m)^2}{2\sigma^2} \right) \right),$$

де $l = 1, 2, 3, ...; m, \sigma$ – математичне очікування та середнє квадратичне значення похибки визначення параметра λ; Δm – усереднена по станах ОС відстань між суміжними законами розподілу. Інформативність відповідної характерної ознаки за результатами визначення λ буде тим вищою, чим більше Δm відрізняється від σ за умови $\Delta m > \sigma$.

Для іконічних методів моніторингу у якості P_0 доцільно прийняти ймовірність правильного дешифрування об'єктів на зображеннях, яка визначається за формулою [4]

$$P_0 = \exp\left(\frac{\ln \alpha}{\lg \frac{1+K_0}{1-K_0}} \left(\frac{d}{d_0} \right)^2 \right),$$

де K_0 – радіометричний контраст об'єкта на зображенні, α – заданий рівень достовірності, d_0 – характерна детальність ОС на зображенні, що потрібна для їх виявлення, d – роздільна здатність знімальної апаратури. Середня ймовірність похибки розпізнавання для U станів ОС за умови апріорної невизначеності визначається як

$$P_e = \sum_{x \in X} 0.5 \left[1 - \frac{\max_{1<u\leqslant U}\{P(L_u)p(x/L_u)\}}{\sum_{u=1}^{U} p(L_u)p(x/L_u)} \right].$$

Таким чином, критерієм відбору корисних характерних ознак розпізнавання стану ОС є максимізація P_0 (мінімізація P_e).

Література. **1.** Файнзильберг Л.С. Математические методы оценки полезности диагностических признаков. Монография. / Л.С. Файнзільберг. – К.: Освіта України. 2010. – 152 с. **2.** Сащук І.М., Романов О.М. Застосування системного підходу для синтезу структури апаратно-програмних комплексів технічного аналізу сигналів на основі інформаційно-сигнатурних технологій / І.М. Сащук, О.М. Романов // Системний аналіз та інформаційні технології: матеріали 18-ї МНТК SAIT 2016, Київ, 30 травня – 2 червня 2016 р. / ННК "ІПСА "НТУУ "КПІ". – К.: ННК "ІПСА" НТУУ "КПІ", 2016. – С. 153–154. **3.** Анисимов Б.В., Курганов В.Д., Злобин В.К. Распознавание и цифровая обработка изображений: Учеб. Пособие для студентов вузов. / Б.В. Анисимов, В.Д. Курганов, В.К. Злобин. – М.: Высшая школа, 1983. – 295 с. **4.** Станкевич С.А. Уточнення відомої емпіричної формули оцінки імовірності правильного дешифрування об'єктів на аерокосмічному зображенні / С.А. Станкевич // Праці наукового центру ВПС України. – Вип. 7. – К.: НЦ ВПС, 2004. – С. 242–246.

Скляр А.В.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Сравнение интеллектуальных систем принятия решений в играх с неполной информацией

Вступление. На сегодняшний день одной из наиболее важных сфер, в которых применяется искусственный интеллект, является оптимальная стратегия в играх с нулевой суммой и неполной информацией, в частности, в покере. Учитывая то, что исследования таких игр позволяют моделировать режимы работы различных систем [1], а также сложность изучения в силу неполноты информации, изучение данной проблемы является очень существенной задачей. Для исследования игр с неполной информацией применяются различные подходы, в частности различные алгоритмы обучения с подкреплением, глубинного обучения, а также обучение при помощи баесовских сетей. Наиболее перспективными и показывающими наилучший результат на сегодняшний день являются следующие алгоритмы: байесовские сети принятия решений [2], глубинное обучение с подкреплением [4], алгоритм минимизации противодействия [3]. В связи с важностью изучаемого вопроса, актуальной задачей, которая выполняется в этой работе, является сравнение вышеперечисленных алгоритмов для нахождения наиболее оптимального.

Байесовская сеть принятия решений. Байесовская сеть принятия решений – один из наиболее хорошо изученных и алгоритмически простых алгоритмов. Байесовская сеть – это направленный ациклический граф, каждой вершине которого соответствует случайная переменная, а дуги графа кодируют отношения условной независимости между этими переменными. Сеть для игр с неполной информацией, показана на примере покера на рис. 1.

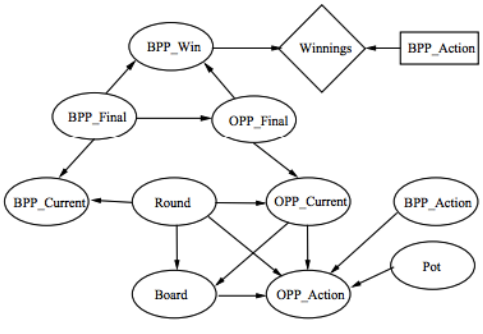


Рис. 1. Байесовская сеть принятия решений.

Здесь

- *BPP_Next_Action* – узел принятого решения.
- *OPP_Final* – тип конечной руки оппонента.
- *BPP_Final* – тип конечной руки игрока.
- *BPP_Win* – булевая переменная, показывающая, выиграет ли игрок в этой раздаче.
- *OPP_Current* – тип текущей руки оппонента.
- *BPP_Current* – тип текущей руки игрока.
- *BPP_Action* – узел, хранящий предыдущеее принятое решение.
- *OPP_Action* – узел, хранящий предыдущее принятое решение оппонента.

Байесовская сеть – довольно сильный механизм принятия решений в играх с неполной информацией, однако он обладает недостатками по сравнению с другими алгоритмами, так как в силу простоты модели, действия агента являются легко предсказуемыми, в частности, в случае сильной руки, агент будет играть сильно, в случае слабой руки – наоборот, слабо.

Глубинное обучение с подкреплением. Данный алгоритм не обладает недостатками, которые присущи байесовским сетям, однако он является более сложным как в отношении реализации, так и в отношении сложности работы алгоритма. В качестве примера мы будем исследовать “Алгоритм фиктивных нейронных игр” [5]. Особенностью алгоритма является

использование алгоритма "Полнотекстовой формы экстенсивных игр", который гарантирует линейную сложность и время благодаря обновлению стратегии игрока следующим образом:

$$\sigma \propto \lambda_1 x_{\pi_1}(s)\pi_1(s,a) + \lambda_2 x_{\pi_2}(s)\pi_2(s,a), \quad \forall s, a. \tag{1}$$

Для тренировки агента аткже используются нейронные сети. Для этого данные разбиваются на два набора:

- Память M_{RL} используется для обучения с подкреплением и содержит кортежи переходов $(s_t, a_t, r_{t+1}, s_{t+1})$, где
  - s_t – состояние, в котором находился агент в момент времени t
  - a_t – действие, которое совершает агент
  - r_{t+1} – мгновенное подкрепление, которое получает агент после своего действия
  - s_{t+1} – состояние, в которое переходит агент
- Память M_{SL} используется для обучения с учителем и содержит кортежи вида (s_t, a_t).

Алгоритм обучает нейронную сеть $Q(s, a|\theta^Q)$ на данных, находящихся в M_{RL}, после чего выбирается лучшая стратегия противодействия, $\beta = \epsilon$-жадная(Q), которая выбирает случайное действие с вероятностью ϵ, а в противном случае – выбирает действие, которое максимизирует прогнозируемое подкрепление. Агент также тренирует отдельную нейронную сеть $\Pi(s, a|\theta^\pi)$ для имитации собственной лучшей стратегии при помощи обучения с учителем на данных из набора M_{SL}. Эта сеть связывает состояния с вероятностями действия и определяет стратегию в среднем, $\pi = \Pi$. Во время игры, агент выбирает свои действии из смеси двух стратегий, β и π. На каждой итерации:

- θ^Π обновляется при помощи метода стохастического градиентного спуска со следующей функцией потерь:
$$L(\theta^\Pi) = E_{(s,a)\sim M_{SL}}[-log\Pi(s, a|\theta^\Pi)]. \tag{2}$$
- θ^Q обновляется при помощи метода стохастического градиентного спуска со следующей функцией потерь:
$$L(\theta^Q) = E_{(s,a,r,s')\sim M_{RL}}[(r + max_{a'}Q(s', a'|\theta^{Q'}) - Q(s, a|\theta^Q))^2]. \tag{3}$$

Основной проблемой данного алгоритма является практическая невозможность одновременного обучения нескольких агентов для ускорения тренировки. В связи с этим, а также в связи со сложностью сети, которую необходимо обучать, алгоритм является достаточно затратным в отношении времени и, поэтому, зачастую неэффективным.

Алгоритм минимизации сожаления противодействия. Наиболее интересным на сегодняшний день является алгоритм минимизации противодействия, так как для обучения он не использует нейронные сети, что позволяет агенту обучаться быстро, однако показывает результаты не хуже, чем более сложные алгоритмы. Основным понятием в данном алгоритме является минимизация сожаления, которое является обобщением функции потери для теории игр. Для определения данного понятия рассмотрим многоразово повторяющуюся экстенсивную игру. Пусть σ_i^t – стратегия, которую использует игрок i в t-м раунде. Среднее обобщенное сожаление игрока i на момент времени T будет равняться:

$$R_i^T = \frac{1}{T} \max_{\sigma_i^* \in \sum_i} \sum_{t=1}^{T} (U_i(\sigma_i^*, \sigma_{-i}^t) - u_i(\sigma^t)). \tag{4}$$

Также определим $\bar{\sigma}_i^t$ как средняя стратегия для игрока i с момента времени 1 по T. В частности, для каждого набора информации $I \in \Im_i$, для каждого $a \in A(I)$ определим:

$$\bar{\sigma}_i^t(I)(a) = \frac{\sum_{t=1}^{T} \pi_i^{\sigma^t}(I)\sigma^t(I)(a)}{\sum_{t=1}^{T} \pi_i^{\sigma^t}(I)}. \tag{5}$$

Теорема 1. В игре с нулевой суммой в момент времени T, если среднее обобщенное сожаление обоих игроков меньше ϵ, то $\bar{\sigma}^T$ – равновесие уровня 2ϵ.

Благодаря приведенной выше теореме и свойствам алгоритма минимизации сожаления, в играх с самим собой, данный алгоритм может использоваться для подсчета приближенного

равновесия по Нэшу. Далее определим сожаление противодействия. Основная идея подхода заключается в том, чтобы провести декомпозицию обобщенного сожаления в набор аддитивных компонент, которые могли бы быть минимизированы по отдельности. Сначала рассматривается информационный набор $I \in \Im_i$ и все варианты выбора, сделанные игроком i в этом наборе. Определяется полезность противодействия $u_i(\sigma, I)$ как ожидаемоая полезность в случае, если был получен информационный набор I и все игроки, кроме игрока i, который играет для достижения набора I, играют используя стратегию σ, или, более формально:

$$u_i(\sigma, I) = \frac{\sum_{h \in I, h' \in Z} \pi^{\sigma}_{-i}(h)\pi^{\sigma}(h, h')u_i(h')}{\pi^{\sigma}_{-i}(I)}, \tag{6}$$

где $\pi^{\sigma}(h, h')$ – вероятность перехода от истории h к истории h'. Наконец, для всех $a \in A(I)$ определим $\sigma|_{I \to a}$ как стратегию, идентичную с σ, за исключением того, что игрок всегда выбирает действие a, находясь в информационном множестве I. Непосредственное сожаление противодействие рассчитывается по формуле:

$$R^T_{i,imm}(I) = \frac{1}{T} \max_{a \in A(I)} \sum_{t=1}^{T} \pi^{\sigma^t}_{-i}(I)(u_i(\sigma^t|_{I \to a}, I) - u_i(\sigma^t, I)). \tag{7}$$

Так как нас больше всего интересуют случаи, когда сожаление положительно, то обозначим положительное сожаление следующим образом: $R^{T,+}_{i,imm}(I) = max(R^T_{i,imm}(I), 0)$. Это приводит к следующему выводу:

Теорема 2. $R^T_i \leqslant \sum_{I \in \Im_i} R^{T,+}_{i,imm}(I)$

Благодаря теореме мы видим, что сожаление можно уменьшить изменяя только $\sigma_i(I)$. Таким образом, мы определяем стратегию для времени T+1 следующим образом:

$$\sigma^{T+1}_i(I)(a) = \begin{cases} \frac{R^{T,+}_i(I,a)}{\sum_{a \in A(I)} R^{T,+}_i(I,a)} & \text{если} \sum_{a \in A(I)} R^{T,+}_i(I,a) > 0 \\ \frac{1}{|A(I)|} & \text{в противном случае} \end{cases} \tag{8}$$

Уравнение (8) мгновенно приводит к следующему выводу:

Теорема 3. Если игрок i выбирает действия согласно последнему уравнению, то $R^T_{i,imm}(I) \leqslant \Delta_{u,i}\frac{\sqrt{|A_i|}}{\sqrt{T}}$ и, соответственно, $R^T_i \leqslant \Delta_{u,i}|\Im_i|\frac{\sqrt{|A_i|}}{\sqrt{T}}$, где $|A_i| = max_{h:P(h)=i}|A(h)|$.

Именно на последнем утверждении базируется утверждение о том, что стратегия, приведенная в уравнении (8) может использоваться в игре с самим собой для достижения равновесия Нэша. В работе показывается практическое применение приведенного алгоритма для игры в покер.

Выводы. В данном исследовании анализируются существующие алгоритмы для принятия решений в играх с неполной информацией, их преимущества и недостатки, проводятся вычисления, чтобы показать, насколько выгодным может быть использование алгоритмов разной степени сложности для игр с неполной информацией на примере покера.

В результате исследования получены результаты и рекомендации относительно оптимальности использования вышеперечисленных алгоритмов для задач интеллектуального анализа в играх с неполной информацией.

Литература. **1.** Parsons, S., Wooldridge, M. Game theory and decision theory in multi-agent systems. Autonomous Agents and Multi-Agent Systems (JAAMAS), 2002. **2.** Nicholson, A., Korb, K., Boulton, D.. Using bayesian decision networks to play texas holdem poker. International Computer Games Association (ICGA) Journal. Monash University, Victoria, Australia, 2006. **3.** Lisy, V., Davis, T., Bowling, M. Counterfactual regret minimization in sequential security games. AAAI Conference on Artificial Intelligence (AAAI), 2016. **4.** Heinrich, J., Silver, D. Deep Reinforcement Learning from Self-Play in Imperfect-Information Games. University College London, UK, 2016. **5.** Heinrich, J., Lanctot, M., Silver, D., Fictitious self-play in extensive form games. Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 2015.

***Соловйов С.О.*[1,2] *Якименко І.В.*[1] *Кир'яченко С.П.*[2]**
[1]*КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна;* [2]*Центральна науково-дослідна лабораторія НМАПО імені П.Л. Шупика, Київ, Україна*

Моделювання біохімічного резерву організму при лікуванні хворих на ІХС

Вступ. Ішемічна хвороба серця – це комплексне захворювання, при якому порушується кровопостачання органу, тобто порушення відбуваються в коронарних артеріях. Ішемічна хвороба серця в Україні займає провідні позиції в структурі причин смертності від захворювань. За даними Всесвітньої організації охорони здоров'я серед країн Європейського Союзу, країн СНД, та інших європейських країн – Україна посідає одне з перших місць серед смертності населення від ІХС. Так, наприклад, за показником Раптової серцевої смертності від ІХС у віці від 0 до 64 років на 100 000 населення – Україна посідає перше місце серед цих країн (143,78 / 100 000 населення). З цього можна зробити висновок, що ефективне лікування при ішемії необхідно при перших же ознаках захворювання [1]. Оцінка застосування певної терапії на основі математичної моделі допоможе вчасно скорегувати хід лікування, тим самим забезпечить найбільш ефективне лікування.

Мета роботи. Створити модель біохімічної кінетики клінічних показників при різних схемах терапії, враховуючи функціональні резерви організму людини при ішемічній хворобі серця.

Опис моделі. У дослідження були включені пацієнти, які знаходились на стаціонарному лікуванні у І терапевтичному відділенні КЗ «Київська міська клінічна лікарня №7» за період 2016 року.

Для хворих на ІХС застосовувались різні схеми терапії (наприклад застосування L-аргінін аспартату, триметазидину, ранолазину і т.д.). При визначенні ефективності застосування певної схеми лікування доцільно слідкувати за змінами стану організму людини, а головне – біохімічного резерву [2], і в даному випадку будуть визначатись зміни показників ліпідного обміну, білкового обміну та клітинного резерву на основі моделей біохімічної кінетики [3].

Запропоновано чотири рівняння кінетики зміни взаємопов'язаних показників ліпідного обміну (1)–(4): ліпопротеїди високої щільності, ліпопротеїди низької щільності, холестерин та тригліцериди; два рівняння кінетики зміни показників білкового обміну (5)–(6): загальний білок та альбумін; три рівняння зміни показників клітинного резерву (7)–(8): нормальні клітини, апоптичні клітини, мітохондріальний мембранний потенціал. Кожне рівняння представляє собою залежність значень кожного конкретного показника після лікування від початкових даних (до лікування). Причому невідомі константи k – константи впливу певної терапії.

Рівняння кінетики змін показників ліпідного обміну:

$$(HDL)_{\text{к.}} = \left(\frac{k'_1}{(LDL)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k'_2}{(Trigl)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k'_3}{(LDL)_{\text{п.}} \cdot (Trigl)_{\text{п.}}} \right) \cdot (HDL)_{\text{п.}} \quad (1)$$

$$(LDL)_{\text{к.}} = \left(\frac{k''_1}{(LDL)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k''_2}{(Trigl)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k''_3}{(LDL)_{\text{п.}} \cdot (Trigl)_{\text{п.}}} \right) \cdot (LDL)_{\text{п.}}, \quad (2)$$

$$(Chol)_{\text{к.}} = \left(\frac{k'''_1}{(LDL)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k'''_2}{(Trigl)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k'''_3}{(LDL)_{\text{п.}} \cdot (Trigl)_{\text{п.}}} \right) \cdot (Chol)_{\text{п.}}, \quad (3)$$

$$(Trigl)_{\text{к.}} = \left(\frac{k''''_1}{(LDL)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k''''_2}{(Trigl)_{\text{п.}} \cdot (Chol)_{\text{п.}}} + \frac{k''''_3}{(LDL)_{\text{п.}} \cdot (Trigl)_{\text{п.}}} \right) \cdot (Trigl)_{\text{п.}}. \quad (4)$$

- HDL – ліпопротеїди високої щільності
- LDL – ліпопротеїди низької щільності
- $Chol$ – холестерин
- $Trigl$ – тригліцерини

- $()_{\text{п.}}$ – початкове значення (до лікування)
- $()_{\text{к.}}$ – кінцеве значення (після лікування)

Рівняння кінетики змін основних показників білкового обміну:

$$(Alb)_{\text{к.}} = \frac{k'_1}{(Prot)_{\text{п.}}} \cdot (Alb)_{\text{п.}} \tag{5}$$

$$(Prot)_{\text{к.}} = \frac{k''_1}{(Alb)_{\text{п.}}} \cdot (Prot)_{\text{п.}} \tag{6}$$

- $Prot$ – загальний білок
- Alb – альбумін
- $()_{\text{п.}}$ – початкове значення (до лікування)
- $()_{\text{к.}}$ – кінцеве значення (після лікування)

Рівняння кінетики змін клітинного резерву:

$$(N)_{\text{к.}} = \left(\frac{k'_1}{(Ap)_{\text{п.}} \cdot (MMP)_{\text{п.}}} + \frac{k'_2}{(App)_{\text{п.}}} + \frac{k'_3}{(MMP)_{\text{п.}}} \right) \cdot (N)_{\text{п.}}. \tag{7}$$

$$(Ap)_{\text{к.}} = \left(\frac{k''_1}{(N)_{\text{п.}} \cdot (MMP)_{\text{п.}}} + \frac{k''_2}{(N)_{\text{п.}}} + \frac{k''_3}{(MMP)_{\text{п.}}} \right) \cdot (Ap)_{\text{п.}}. \tag{8}$$

$$(MMP)_{\text{к.}} = \left(\frac{k'''_1}{(N)_{\text{п.}} \cdot (Ap)_{\text{п.}}} + \frac{k'''_2}{(N)_{\text{п.}}} + \frac{k'''_3}{(Ap)_{\text{п.}}} \right) \cdot (MMP)_{\text{п.}}. \tag{9}$$

- N – нормальні клітини
- Ap – Апоптичні клітини
- MMP – Мітохондріальний мембранний потенціал
- $()_{\text{п.}}$ – початкове значення (до лікування)
- $()_{\text{к.}}$ – кінцеве значення (після лікування)

Маючи набір даних про пацієнтів, застосовану терапію та показники кожного до та після лікування, можна скласти системи рівнянь, беручи за основу по черзі кожне з (1)–(9) рівнянь. Тобто для групи людей, що мали однакову схему лікування, з відомої вибірки вхідних даних (параметри ліпідного обміну до та після лікування; параметри білкового обміну до та після лікування; параметри зміни клітинного резерву) для кожного запропонованого рівняння кінетики можна скласти системи рівнянь для визначення невідомих. Кількість невідомих змінних (констант) від одного до трьох, в залежності від рівняння кінетики. Вибірка пацієнтів набагато більша ніж кількість констант, тому відповідно рівнянь в кожній системі більше ніж невідомих значень. Отримуємо перевизначену систему лінійних рівнянь, з якої невідомі значення коефіцієнтів доцільно визначати за допомогою методу найменших квадратів [4].

Визначивши коефіцієнти моделі впливу терапії на показники біохімічного резерву, отримуємо рівняння біохімічної кінетики для певної схеми лікування.

Висновок. Завдяки отриманим рівнянням кінетики можливо спрогнозувати, як конкретна терапія вплине на показники пацієнта, маючи дані про його початковий стан. Визначаючи рівняння кінетики з коефіцієнтами впливу різних схем лікування, можна порівнювати їх вплив на пацієнта та обирати найефективніше лікування. Таким чином модель може використовуватись лікарем на практиці при прийнятті лікувальних рішень.

Література. **1.** Інформаційний портал про здоров'я. Лікування ІХС, www.a-betka.in.ua/2016/04/preparaty-dlja-likuvannja-ihs.html. **2.** Функціональні резерви організму людини, http://pti.kiev.ua/korysna-info/rizni/319-funkcionalni-rezervi-organizmu-ljudini..html. **3.** С.Д. Варфоломеев, К.Г. Гурич. Биокинетика. Практический курс. – Москва, 1999. **4.** Лоусон Ч., Хенсон Р. Численное решение задач методом наименьших квадратов. – М.: Наука, 1986.

Суржан А.О., Бідюк П.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Оцінювання інвестиційних ризиків в стартап-проектах

Сьогодні, в епоху стрімкого розвитку інноваційних технологій відбувається жорсткий перерозподіл ринку і дедалі меншими стають шанси новостворених бізнес-структур отримувати достатній дохід. Саме тому підприємці мусять приділяти значну увагу впровадженню нових технологій та виведення на ринок інноваційних продуктів. Нині замало створити бізнес-структуру, необхідно забезпечити її унікальність. При цьому дедалі частіше йде мова не лише про нові підходи до розвитку «підприємництва», а саме про створення «стартапу», який, в свою чергу, завжди найбільш асоціюється із двома словами: унікальність та ризик. Означене й обумовлює актуальність даного дослідження.

У науковому вжитку термін «стартап» з'явився порівняно недавно й одне з них має таке тлумачення: процес виходу на ринок новоствореного підприємства з інноваційним проектом, зазвичай, у короткотривалий термін і з мінімальними капіталовкладеннями [1]. Проте інвестиції в стартапи завжди є ризиком, адже неможливо однозначно визначити успіх проекту, а отже і повернення інвестицій інвесторам. Виділяють такі ознаки стартапу: інноваційність; обмеженість початкових інвестицій; швидкий розвиток; низькі шанси на успіх [3]. Також виділяють такі етапи розвитку стартапу: ідея – команда – ринок – концепція – технічне завдання – прототип – тестування – альфа-версія – закрита бета-версія – публічна бета-версія – запуск – пошук інвесторів [2].

Фінансування стартап проектів – ризикова діяльність. Існують різні варіанти пошуку інвестицій для стартапів: бізнес-ангели, венчурні фонди, інвестиційні фонди, краудфандинг. Перші три установи пов'язує одне: обрати та профінансувати найуспішніший на думку експертів проект при мінімальних ризиках фінансових втрат.

З ризиком тісно пов'язується поняття економічної безпеки проекту, причому як безпека суб'єкта, що представляє проект, так і безпека інвестора. Безпека суб'єкта полягає у тому, що ризиковий і неуспішний проект призведе до збитків інвесторів. Безпека інвестора прямо залежить від адекватної оцінки проекту і суб'єкта, що представляє проект. Підвищення безпеки інвестиційних проектів забезпечує стійкість регіональної економіки [4]. Актуальною є задача побудови математичних моделей та інформаційних технологій на їх основі для оцінювання ризику стартап-проектів відносно рівня безпеки їх фінансування. Популярні підходи до розв'язання цієї задачі ґрунтуються на використанні нечіткої математики, нейронечітких моделей, байєсівських мереж для різних інвестиційних суб'єктів (інвестиційні фонди, венчурні фонди, бізнес-ангели та ін.).

Також все більшого поширення набувають підходи на основі комбінування математичних моделей, побудованих на основі ідеологічно різних підходів, наприклад, комбінування регресії з методами штучного інтелекту (байєсівські мережі, нейронечіткі моделі і т. ін.). Це забезпечує можливість компенсації похибок, які характерні для моделей конкретного типу. Загалом, для успішного розв'язання задач менеджменту інвестиційних ризиків необхідно будувати спеціалізовані системи підтримки прийняття рішень на основі принципів системного аналізу.

Література. **1.** Мрихіна О.Б. Перспективи стартап-компаній у контексті конкурентоспроможного розвитку українського ринку високих технологій / О.Б. Мрихіна, А.Р. Стояновський, Т.І. Міркунова // Актуальні проблеми економіки. – 2015. – № 9. – С. 215–225. **2.** От идеи к реализации: 12 этапов проектирования стартапа [Електронний ресурс]. – Режим доступу: https://kontur.ru/articles/1493. **3.** Этапы развития стартапов [Електронний ресурс]. – Режим доступу: http://startup-house.ru/stati-pro-startapy/link/etapy-razvitiya-startapov **4.** Kelemen, M. Problems of Protected Interests in the Security Sectors [Text] / M. Kelemen // Warszaw: Wydawnictwo Wyzszej szkoly menedzerskiej w Warszawie im. Prof. Leszka J. Krzyzanowskiego, 2015. – 114 p.

Таранян А.Р.
Московский государственный технический университет им. Н.Э.Баумана, Москва, Россия

Ускорение локализации высококонтрастных движущихся объектов в видеопотоке

Роль интеллектуальных систем в нашей жизни возрастает с каждым днем. Важнейшим компонентом многих интеллектуальных систем является подсистема локализации, распознавания и отслеживания различных объектов в видеопотоке.

Данная работа посвящена разработке метода ускорения локализации определенного класса объектов – *высококонтрастных объектов*. К классу высококонтрастных можно отнести следующие объекты:

- Автомобильные номерные знаки;
- Дорожные знаки;
- QR-коды.

Разработанный метод. Наиболее очевидным способом ускорения локализации объектов в видеопотоке является анализ изменений между соседними кадрами видеопотока и поиск новых объектов лишь в измененных областях очередного кадра. Поскольку распространенные методы вычитания фона (среди которых можно выделить метод [1], основанный на смеси Гауссовских распределений) плохо подходят для обработки видеопотоков, снятых нестатической камерой, а их ресурсоемкость порой сравнима с ресурсоемкостью поиска объектов на полном кадре, с учетом характеристик высококонтрастных объектов был разработан новый метод вычисления измененных областей между соседними кадрами видеопотока.

Вычисление измененных областей кадра, содержащих искомый высококонтрастный объект, можно осуществить путем попарного сравнения всех пикселей с одинаковыми координатами. Однако, если учесть минимальный размер искомого объекта в кадре и характер контрастных переходов на типовом изображении объекта в вертикальном и горизонтальном направлениях, то можно сильно ограничить множество сравниваемых пикселей. Так, для локализации номерного знака, при минимальном размере локализуемого знака в кадре 14x65 пикселей, предлагается сравнивать лишь те пиксели соседних кадров, которые находятся на сетке с горизонтальным шагом в 32 пикселей, и вертикальным шагом в 14 пикселей.

Для устранения небольших дрожаний видеокамеры и шумов предлагается сравнивать не значения пикселей непосредственно, а значения v^1_{ij} и v^2_{ij} для вертикальных линий и h^1_{ij} и h^2_{ij} для горизонтальных линий:

$$v^t_{ij} = \sum_{l=i-1}^{i+1} (im^t[l,j] + \alpha * (im^t[l,j-1] + im^t[l,j+1])), t = 1,2,$$

$$h^t_{ij} = \sum_{k=j-1}^{j+1} (im^t[i,k] + \alpha * (im^t[i,k-1] + im^i[i,k+1])), t = 1,2,$$

где im^1 – это полутоновое представление первого кадра, im^2 – это полутоновое представление второго кадра, а α – коэффициент, определяющий значимость значений пикселей, находящихся в непосредственном соседстве с пикселями построенной сетки.

Пиксель (i,j) считается изменившимся, если

$$|v^1_{ij} - v^1_{ij}| > \varepsilon \text{ либо } |h^1_{ij} - h^1_{ij}| > \varepsilon$$

Значения параметров α и ε для задачи локализации номерных знаков эмпирическим путем были выбраны равными 0.2 и 92, соответственно.

После вычисления измененных пикселей сетки определяются клетки сетки, которые могли измениться между соседними кадрами. Производится это путем пометки клеток, которые

находятся рядом с изменившимся пикселем сетки (2 либо 4 клетки) (рис. 1).

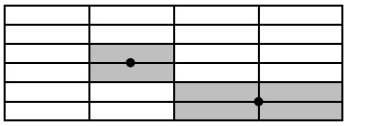


Рис. 1. Определение измененных клеток

Далее, на основе информации о потенциально изменившихся клетках сетки строится интегральная сетка изменений между соседними кадрами. Идея интегральной матрицы, предложенная в работе [2], заключается в построении представления изображения, которое позволяет быстро (за константное количество операций) вычислять сумму пикселей любой прямоугольной области изображения. Значение каждой клетки интегральной сетки изменений представляет собой количество потенциально измененных клеток, которые находятся левее и выше текущей клетки. На рис. 2 представлен пример кадра видеопотока и интегральной сетки изменений, построенной для представленного и предшествующего ему кадров.

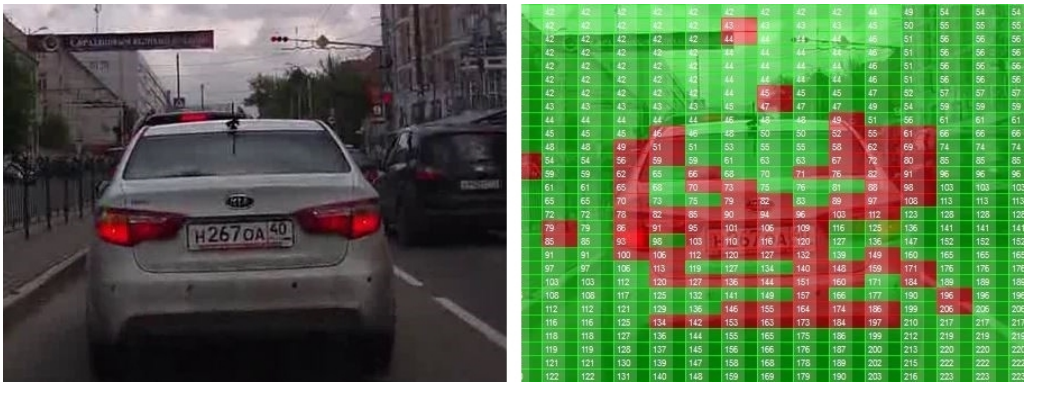


Рис. 2. Пример кадра видеопотока и интегральной сетки изменений

Многие алгоритмы локализации объектов на изображении, к числу которых относится и алгоритм Виолы–Джонса [3], используют метод “скользящего окна” для поиска объектов фиксированного размера. Для поиска объектов различных размеров производится множество сканирований изображения с применением “скользящих окон” соответствующих размеров. Предложенная интегральная сетка внедряется следующим образом: при проходе скользящего окна над очередной областью кадра эта область обрабатывается только в том случае, если она была изменена по сравнению с предыдущим кадром. Факт изменения области вычисляется просто – определяются значения клеток интегральной сетки изменений, соответствующих левому верхнему левому и нижнему правому углам данной области, и если данные значения не равны, то рассматриваемая область изменилась по сравнению с предыдущим кадром видеопотока.

Разработанный метод был апробирован на задаче локализации автомобильных номерных знаков. Результаты тестирования показали высокую эффективность данного метода:

- Увеличение производительности на 104% на видеопотоках с неподвижных камер;
- Увеличение производительности на 68% на видеопотоках с авторегистраторов;
- Увеличение производительности на 14% на видеопотоках, отснятых с руки;
- Обнаружение 99.1% номерных знаков, обнаруженных немодифицированным алгоритмом.

Литература. **1.** Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction // Proceedings of International Conference on Pattern Recognition, 2004. – P. 28–31. **2.** Crow F. Summed-area tables for texture mapping // ACM SIGGRAPH Computer Graphics, 1984. Vol. 18. N. 3. – P. 207–212. **3.** Viola P., Jones M.J. Robust real-time face detection // International Journal of Computer Vision, 2004. Vol. 57. – P. 137–154.

Тимошенко Ю.О., Ніколаєв С.С., Ходак А.Р.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Порівняння моделей на основі нейронних мереж для задачі машинного розуміння тексту

Знаходження відповідей на питання (Question Answering, QA) – це одне з ключових завдань обробки природньої мови, яке вимагає як розуміння природньої мови, так і зовнішніх загальних знань. Відповіді на питання на основі розуміння прочитаного (Reading comprehension-based question answering, RCQA) – це споріднена задача знаходження відповіді на питання за допомогою уривку тексту з відповідного документу. Одним із чинників, які до недавнього часу гальмували поступ у сфері RCQA, була нестача навчальних даних. Попередні набори даних мали високу якість через ручну людську підготовку, однак вони були малими за розмірами, тож не довзоляли використовувати жадібні до даних моделі на кшталт глибоких нейронних мереж [1]. Нещодавно Райпуркар та ін. (2016) [2] випустили Стенфордський набір даних для відповідей на питання (Stanford Question Answering dataset, SQuAD), який за розміром у сотні разів перевищує усі попередні набори даних та включає в себе велике різноманіття особливостей, необхідних для всеохопного машинного розуміння (МР) природньої мови.

У загальному випадку, об'єкт вибірки для задачі RCQA включає в себе питання, уривок тексту, який містить відповідь, та проміжок всередині уривку, який є відповіддю на питання. Набір даних SQuAD представлений множиною трійок (Q, P, A), де $Q = (q_1, ..., q_i, ..., q_M)$ – це питання довжиною M слів, P – уривок довжиною N слів, $A = (a_b, a_c)$ – відрізок для відповіді, $1 \leqslant a_b \leqslant a_c \leqslant N$. Задача RCQA може бути представлена як оцінка умовної імовірності $Pr(A|Q, P)$ та передбачення відповідей для тестових елементів $A^* = \arg\max_{A \in A(P)} Pr(A|Q, P)$,
де $A(P)$ – множина відповідей-кандидатів в P.

Для досягнення високої точності МР на наборі даних SQuAD було розроблено низку моделей. Найкращих результатів на цьому наборі досягають моделі на основі нейронних мереж. У даній роботі буде розглянуто та порівняно результати роботи на наборі даних SQuAD трьох таких моделей:

1. Мережа двонапрямленого потоку уваги [3] – модель, яка представляє багатокроковий ієрархічний процес, який використовує механізм двонапрямленого потоку уваги для отримання контекстної репрезентації запиту без раннього сумування.
2. Глибока LSTM мережа [4] – особливий вид рекурентних нейронних мерж, здатних запам'ятовувати інформацію за довгий період часу.
3. Читач сумованої уваги [5] – модель, яка використовує увагу для прямого вибору відповіді з контексту напротивагу обчисленню відповіді з використанням змішаного представлення слів у документі як у схожих моделях.

Детальне порівняння результатів роботи цих моделей дозволить виокремити структурні елементи нейронних мереж, які особливо ефективні в задачах QA.

Література. **1.** Caiming Xiong, Victor Zhong and Richard Socher. Dynamic Coattention Networks For Question Answering, 2016; arXiv:1611.01604. **2.** Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev and Percy Liang. SQuAD: 100,000+ Questions for Machine Comprehension of Text, 2016; arXiv:1606.05250. **3.** Minjoon Seo, Aniruddha Kembhavi, Ali Farhadi and Hannaneh Hajishirzi. Bidirectional Attention Flow for Machine Comprehension, 2016; arXiv:1611.01603. **4.** Karl Moritz Hermann, Tomáš Kočiský, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman and Phil Blunsom. Teaching Machines to Read and Comprehend, 2015; arXiv:1506.03340. **5.** Rudolf Kadlec, Martin Schmid, Ondrej Bajgar and Jan Kleindienst. Text Understanding with the Attention Sum Reader Network, 2016; arXiv:1603.01547.

Чапалюк Б.В., Зайченко Ю.П.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Аналіз алгоритмів детектування обличчя

На сьогоднішній день забезпечення безпеки в місцях масового скупчення, таких як метрополітен, залізничні вокзали та інші, відбувається за допомогою систем автоматизованого моніторингу. Основною задачею таких систем є аналіз потокової інформації з камер відеоспостереження та аналіз отриманого потоку в режимі онлайн. Зазвичай, в рамках аналізу відбувається детекція рухомих об'єктів, розпізнавання обличь людей та пошук їх в доступній базі даних злочинців.

В рамках дослідження було проведено аналіз існуючих алгоритмів детекції обличчя в режимі онлайн. Метою дослідження є виявлення найбільш точного та достатньо швидкого для режиму онлайн алгоритму детекції.

Більшість сучасних алгоритмів та підходів для класифікації об'єктів на зображеннях базуються на згорткових нейронних мережах, так як вони показують найбільш точні результати в різноманітних задачах аналітики зображень. Хоча такі алгоритми показують дуже гарні результаті в розпізнаванні, вони вимагають наявності достатньо великої вибірки даних для навчання, а також серйозних обчислювальних ресурсів. Тому, для перевірки роботи алгоритмів, були вибрані добре відомі набори даних вже з розміченими зображеннями, які були зроблені спеціально для розробки систем детекції обличчя. Для даного дослідження було обрано фреймворк FDDB [1], що містить велику кількість розмічених зображень з різним рівнем складності, у тому числі зображення з великим скупченням людей. Також, додатково був вибраний більш сучасний набір даних WIDER [2], де присутня велика кількість складних для детектування облич зображень, з урахуванням складних позицій обличчя, різного фону, рівня освітленості та присутності макіяжу на обличчі.

В рамках дослідження для вирішення задачі детекції обличчя використовувалися глибинні нейронні мережі Faster-RCNN [3] та Faceness [4], а також алгоритм Віола–Джонса [5]. Було проведено дослідження точності детектування на вибраних вибірках даних цими алгоритмами. В результаті, найвищу точність показала глибинна мережа Faster-RCNN, а найменша точність була показана алгоритмом Віола–Джонса.

В результаті проведеного дослідження було виявлено, що найкращі результати в детектуванні обличчя на сьогоднішній день дають глибинні нейронні мережі. Хоча такий підхід вимагає більше обчислювальної потужності від апаратних засобів, глибинні нейронні мережі працюють достатньо швидко для того, щоб їх можна було застосовувати в програмному забезпечені для онлайн-відеоаналітики.

Література. **1.** V. Jain та E. Learned-Miller, FDDB: A benchmark for face detection in unconstrained settings, Technical Report UM-CS-2010-009, University of Massachusetts, Amherst, 2010. **2.** S. Yang, P. Luo, C.C. Loy та X. Tang, WIDER FACE: A face detection benchmark, CVPR, 2016. **3.** H. Jiang та E.G. Learned-Miller. (2016, May) Face Detection with the Faster R-CNN [Online]. Available: https://arxiv.org/abs/1606.03473. **4.** S. Yang, P. Luo, C.C. Loy та X. Tang, From facial parts responses to face detection: a deep learning aproach, ICVV, pp. 3676–3684, 2015. **5.** P. Viola та M. Jones, Rapid object detection using a boosted cascade of simple features Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR, 2001.

Шаповалова С.І., Мажара О.О., Москаленко Ю.В.
КПІ ім. Ігоря Сікорського, Київ, Україна

Екстракція правил виведення нейронних мереж з представленням в форматі CLIPS

Нейронні мережі здобули широке поширення для вирішення прикладних проблем штучного інтелекту, зокрема сфера їх застосування поширюється на системи ідентифікації, контролю, управління технологічними процесами, прийняття рішень, розпізнання шаблонів, інтелектуального аналізу даних, медичної та технічної діагностики. Однак концептуальним недоліком цього підходу є відсутність інструментів пояснення логічного висновку, особливо у випадку систем з підвищеними вимогами щодо безпеки [1]. Тому вдосконалення засобів екстракції правил (rule extraction) з навченої нейронної мережі є актуальною задачею та має практичне значення.

Згідно з визначенням [2] екстракція правил передбачає формування пояснення гіпотези на основі навченої нейронної мережі та множини прикладів, на яких проводилося навчання. При цьому сформоване пояснення повинно не тільки бути зрозумілим, а й з великою ймовірністю передбачати поведінку нейронної мережі.

Правила, здобуті в процесі екстракції, формують базу знань та можуть застосовуватися як основа для створення незалежних експертних систем або як механізм пояснень в системах підтримки прийняття рішень на основі нейронних мереж. Дослідники відзначають, що ці правила потребують додаткової формалізації для подальшого використання в логічному виведенні [3]. Крім того, з широким поширенням систем глибинного навчання (deep learning), які передбачають обробку надвеликих масивів даних, виникає проблема обсягів правил, отриманих на основі даних систем. Остання проблема стосується як формалізації та верифікації, так і ефективності обробки сформованих баз знань. Це зумовлює необхідність розробки засобів перетворення виокремлених на основі нейронних мереж правил на існуючі декларативні мови.

Метою даної роботи є створення програмного забезпечення перетворення правил, отриманих на основі нейронних мереж, на мову оболонки продукційних систем CLIPS.

Існує три базові підходи до екстракції правил з нейронних мереж: декомпозиційний (decomposition), педагогічний (pedagogical) та еклектичний (eclectic) [4]. Алгоритми декомпозиції працюють шляхом поділу мережі на рівні нейронів. Результат, отриманий від кожного нейрона, агрегується для представлення мережі в цілому. Педагогічний підхід розглядає нейронну мережу як "чорний ящик". Алгоритм намагається сформувати правила, які безпосередньо відображують кореляцію між вхідними характеристиками та вихідним значеннями. Еклектичні алгоритми передбачають поєднання вищезгаданих підходів.

Безпосередні алгоритми екстракції правил залежать від архітектури нейронної мережі та класу вирішуваної задачі. Для багатошарових перцептронів, рекурентних нейронних мереж та карт самоорганізації існують запропоновані рішення, що знайшли використання на практиці [5, 6]. Однак для більш сучасних мереж глибинного навчання проблема ефективної екстракції правил залишається не розв'язаною в повній мірі [7].

Алгоритми екстракції знань з нейронних мереж можуть представляти правила як математичні вирази, вирази символьної логіки, нечіткої логіки або дерева рішень, використовуючи різні формати. Проте найчастіше зустрічаються логічні правила двох типів: if-then та M-of-N. В загальній формі if-then правило представляється наступним чином [7]:

$$IF \quad X \in S(i) \quad THEN \quad Y = y(i), \tag{1}$$

де S – суматор вхідних сигналів, Y – вихідний сигнал. Наприклад, для випадку одного нейрону функцією активації f, функція суми для нейрону $S_i = \sum_i^n (X_i * W_i^j)$, де X_i – вхідні характеристики, W_i^j – відповідні ваги. Вихідним значенням нейрону є функція від зваженої суми вхідних параметрів: $Y = f(S_j)$.

M-of-N передбачає істинність умовної частини правила у випадку виконання M з N множини

булевих виразів, які відображають вхідні сигнали мережі [7]:

$$IF \quad M \quad \{N\} \quad THEN \quad Z. \tag{2}$$

де Z – заключення, яке трактується за сигналом виходу нейрону. При цьому існує можливість перетворення M-of-N правил в if-then.

З метою забезпечення узагальнених рішень дослідники [8, 9] запропонували алгоритми екстракції з нейронних мереж правил у вигляді виразів логіки предикатів першого порядку (здебільшого з деякими обмеженнями).

В даній роботі для підвищення ефективності використання сформованих правил пропонується формалізувати їх за синтаксисом спеціалізованих програмних засобів обробки інформації в експертних системах. Серед існуючих моделей представлення знань в найбільш повній мірі заданому формату if-then правил відповідає продукційна модель. Продукційні системи вважаються адаптацією представлення класичної логіки до систем штучного інтелекту. Саме тому, в даному дослідження пропонується представляти правила, отримані з нейронних мереж у форматі синтаксису мови розробки продукційних систем CLIPS [10]. Вибір CLIPS зумовлено можливістю побудови логічного виразу в антецеденті правила у відповідності до виразів логіки першого порядку (з застосуванням кванторів існування (умовний оператор системи exists) та узагальнення (умовний оператор системи forall), а також з використанням базових математичних та користувацьких функцій. Умові елементи правила в середовищі представляються у вигляді шаблонів фактів. Факти можуть формуватися за заздалегідь заданим форматом (впорядковані), так і у процесі генерації правила (невпорядковані). Для спрощення процедури генерації антецеденти правил дозволяється формувати на основі структур невпорядкованих фактів, однак попередня обробка даних з виокремленням значимих структур спрощує сприйняття правил [8]. Для задач з чітким визначенням вихідного формату (наприклад, класифікації) вихідні значення пропонується представляти у вигляді впорядкованих фактів. Подібна концепція знаходить своє відображення у підході до генерації предикативних правил в роботі [8], де цільовому предикату (Target predicate) відповідає шаблон впорядкованого факту мови CLIPS.

Представлення правил, сформованих на базі нейронних мереж, у вигляді продукцій CLIPS дозволяє безпосереднє застосування апробованих механізмів логічного виведення, а також проведення додаткових експериментів з вибору оптимальної стратегії розв'язання конфліктів. Крім того, більшість сучасних обгорток продукційних систем мають синтаксис мови подібний до CLIPS, що в подальшому дозволить міграцію створених баз знань на інші експертні системи.

Література. **1.** M. Darrah, B. Taylor, and S. Skias, "Rule extraction from dynamic cell structure neural networks used in a safety critical application," Proc. Seventeenth Int. Florida Artif. Intell. Res. Soc. Conf. FLAIRS 2004, vol. 2, pp. 629–634, 2004. **2.** M.W. Craven, "Extracting Comprehensible Models from Trained Neural Networks", Ph.D. dissertation, Department of Computer Sciences, University of Wisconsin-Madison, 1996. **3.** F. Qiu and John R. Jensen, "Opening the black box of neural networks for remote sensing image classification," Int. J. Remote Sens., vol. 25, no. 9, pp. 1749–1768, 2004. **4.** M.G. Augasta and T. Kathirvalavakumar, "Rule extraction from neural networks – A comparative study," 2012 Int. Conf. Pattern Recognition, Informatics Med. Eng. PRIME 2012, March 21, 2012 – March 23, 2012, no. February, pp. 404–408, 2012. **5.** H. Jacobsson, "Rule Extraction from Recurrent Neural Networks: ATaxonomy and Review," Neural Comput., vol. 17, no. 6, pp. 1223–1263, 2005. **6.** M. Supriyaningsih and F. Hakim, "Applying Fuzzy Self-Organizing Map and SPADE Algorithm in Indonesian Natural Disasters," pp. 109–118, 2015. **7.** T. Hailesilassie, "Rule Extraction Algorithm for Deep Neural Networks: A Review," IJCSIS) Int. J. Comput. Sci. Inf. Secur., vol. 14, no. 7, pp. 376–381, 2016. **8.** R. Nayak, "Generating rules with predicates, terms and variables from the pruned neural networks," Neural Networks, vol. 22, no. 4, pp. 405–414, 2009. **9.** M. Guillame-Bert, K. Broda, and A. d'Avila Garcez, "First-order logic learning in Artificial Neural Networks," 2010 Int. Jt. Conf. Neural Networks, pp. 1–8, 2010. **10.** CLIPS Online Documentation http://clipsrules.sourceforge.net/OnlineDocs.html.

***Шулькевич Т.В.*[1], *Селін О.М.*[2], *Селін Ю.М.*[2]**
[1]*КПІ ім. Ігоря Сікорського, Київ, Україна;* [2]*КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна*

Математичний апарат інтелектуального аналізу даних для прогнозування нелінійних нестаціонарних процесів різної природи

Сутність і мету технології інтелектуального аналізу даних (ІАД) можна описати таким чином: технологія, яка призначена для пошуку у великих інформаційних масивах даних неочевидних, об'єктивних, корисних на практиці закономірностей. ІАД здійснюється за допомогою використання технологій розпізнавання шаблонів, а також статистичних і математичних методів. При розвідці даних багаторазово виконуються операції перетворення "сирих" даними (відбір ознак, стратифікація, кластеризація, візуалізація і регресія), що призначені для знаходження: структур, які інтуїтивно зрозумілі для людей і краще розкривають суть процесів, що лежать в основі їх протікання; моделей, які можуть передбачити результат або значення певних ситуацій, використовуючи історичні або суб'єктивні дані. Задачі що вирішує ІАД поділяються за призначенням на описові і предиктивні. Описові, або дескриптивні, задачі пов'язані з поліпшенням розуміння аналізованих даних. Ключовий момент у таких моделях – простота і прозорість результатів для сприйняття людиною. До такого типу задач належать кластеризація і пошук асоціативних правил. Рішення предиктивних, або прогнозуючих, задач поділяється на два етапи. На першому етапі на підставі набору даних з відомими результатами будується модель. На другому етапі вона використовується для прогнозу результатів на підставі нових наборів даних. Вимагається, щоб побудовані моделі працювали максимально точно. До цього типу задач відносять задачі класифікації і регресії. Сюди можна віднести і задачу пошуку асоціативних правил, якщо результати її рішення можуть бути використані для прогнозу появи деяких подій. В доповіді коротко розглянуто деякі методи ІАД, що можна застосовувати для аналізу та прогнозування нелінійних нестаціонарних процесів в задачах прогнозування. До них можна віднести наступні методи:

- Приховані марковські моделі [1, 2];
- Метод подібних траєкторій [3];
- Лінгвістичне моделювання [4–6].

В доповіді наведено математичний апарат, що об'єднує в собі три види подачі інформації за для її аналізу та прогнозування. Перший вид – звичний числовий, другий – графічний, що ще й досі отримуються з аналогових приладів реєстрації та третій – символьний (метод лінгвістичного моделювання), що ще не є розповсюдженим. Для методу лінгвістичного моделювання сформульована змістовна та математична постановки задачі знаходження лінгвістичних образів часових рядів. Було введено поняття лінгвістичного моделювання та поетапно описано спосіб застосування даного підходу для розв'язання поставленої задач. Також, для більш повного розгляду задачі, були введені додаткові підзадачі, описані їх математичні постановки та наведені приклади їх виконання. Приведено правила переходу усіх трьох видів отримання інформації одне в одне. Наведені методи є універсальними як з боку виду отриманої інформації так і з боку наявності в цій інформації нелінійностей та нестаціонарностей.

Література. **1.** Juang, B.H., Rabiner, L.R. (1991) Hidden Markov models for speech recognition, Technometrics. **2.** Rabiner L.R., «A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition», Proceedings of the IEEE, vol, 77. no.2, February 1989, pp. 257–284. **3.** G. Kollios, D. Gunopulos, and V.J. Tsotras. Nearest neighbor queries in a mobile environment. In Spatio-Temporal Database Management 1999, pp. 119–134. **4.** Fu K.S. (1968) Sequential Methods in Pattern Recognition and Machine Learning. – Academic Press. **5.** Баклан І.В. Математичні моделі прогнозування часових рядів різної природи / Баклан І.В., Селін Ю.М., Шулькевич Т.В.// Вестн. Херсонского национального техн. ун-та. – Херсон: ХНТУ, 2014. – Вып. 3 (50). – С. 213–218. **6.** Баклан І.В. Лінгвістичне моделювання: основи, методи, деякі прикладні аспекти. Систем. технології. – 2011. – № 3. – С. 10–19.

Щеголь Ю.М., Бідюк П.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Система підтримки прийняття рішень для прогнозування нестаціонарних процесів в екології

Сьогодні існує нагальна проблема дослідження процесів в екології навколишнього середовища. Умовний розподіл навколишнього світу на системи забезпечує можливість аналізу процесів, що відбуваються в ньому. Зазвичай під системою розуміють сукупність одного або декількох елементів середовища, що взаємодіють між собою та з довкіллям, від якого система відокремлюється фізичними або умовними границями.

У даному дослідженні розглянуто методи, які дають можливість моделювати і прогнозувати нелінійні нестаціонарні процеси (ННП) в екології, а також дають змогу оцінити якість отриманих прогнозів і порівняти з існуючими результатами.

Створена СППР забезпечує реалізацію системного підходу до обробки даних, побудови математичних моделей та оцінювання прогнозів на їх основі. Основними відмінностями запропонованої системи від існуючих є такі: (1) забезпечення можливості високоякісної попередньої обробки даних, спрямованої на коректне нормування, фільтрацію шумів вимірів, врахування зовнішніх випадкових збурень, заповнення можливих пропусків вимірів, оцінювання невимірюваних компонент вектора стану досліджуваного об'єкта, обробку екстремальних значень і т. ін.; (2) використання трьох множин критеріїв: якості даних, адекватності математичних моделей, що будуються на основі статистичних даних, і критеріїв якості оцінок прогнозів розвитку досліджуваних процесів; (3) реалізація адаптивного режиму функціонування системи, який забезпечує підтримку високого рівня адекватності моделей в умовах зміни режиму функціонування процесу і наявності не стаціонарності (інтегрованість та/або гетероскедастичність).

Досліджено такі методи: побудовано множину регресійних моделей (за методикою регресійного аналізу даних), нечіткий метод групового врахування аргументів (НМГВА), нейромережевий підхід та деякі інші. На основі виконаного дослідження здійснено порівняльний аналіз результатів методів. Отримані результати попередніх досліджень свідчать про те, що кращі результати моделювання і прогнозування можна отримати, частіше всього, за допомогою нечіткого МГУА, нейромережевого підходу та нелінійних регресійних моделей (поліноми і логіт).

За критерії якості побудованих моделей обрані такі: коефіцієнт множинної детермінації, статистика Дарбіна–Уотсона, сума квадратів похибок і F-статистика. Оцінювання якості прогнозів виконано за такими критеріями: середньо-квадратична похибка, середня абсолютна похибка, середня абсолютна похибка у процентах і коефіцієнт Тейла. Отримані результати демонструють високу якість побудованих моделей та можливість їх практичного застосування для прийняття відповідних управлінських рішень. Практично показано, що вирішення проблем моніторингу довкілля потребує застосування сучасної методології системного аналізу, яка надає можливість отримати коректні та об'єктивні оцінки стану навколишнього середовища і згенерувати необхідні управлінські рішення.

Широке застосування математичних моделей взаємодії суспільства і природи на основі системного підходу обумовлене, з одного боку, швидким зростанням використання і можливостей обчислювальної техніки та різноманітністю обчислювальних програмних систем, порівняно низькою вартістю розробки моделей і високою точністю розрахунків, а з другого – суттєво обмеженими можливостями лабораторного або натурного моделювання складних процесів і систем навколишнього середовища.

Література. **1.** Бідюк П.І. Аналіз часових рядів (навчальний посібник) / Бідюк П.І., Романенко В.Д., Тимощук О.Л. – К.: Політехніка, 2010. – 317 с. **2.** Бараннік В.О. Моделювання і прогнозування стану довкілля (навчальний посібник) – Харків: Харківська національна академія міськаго господарства, 2007. – 85 с.

Plenary talks

System analysis of complex systems of various nature

Intelligent systems for decision-making

High Performance Computing and Microsystems Engineering

Progressive information technologies

3 High Performance Computing and Microsystems Engineering

Section 3

High Performance Computing and Microsystems Engineering

1. High performance computing: Supercomputing, distributed, cloud, Grid, and quantum computing, high performance systems and networks.
2. Microsystems engineering and design.

Секция 3

Высокопроизводительные вычисления и разработка микросистем

1. Высокопроизводительные вычисления: Суперкомпьютерные, распределенные, облачные, Грид- и квантовые вычисления, высокопроизводительные системы и сети.
2. Разработка и проектирование микросистем.

Секція 3

Високопродуктивні обчислення та розробка мікросистем

1. Високопродуктивні обчислення: Суперкомп'ютерні, хмарні, Грід- та квантові обчислення, високопродуктивні системи та мережі.
2. Розробка та проектування мікросистем.

Artuhov V.G., Brytov O.A.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Possibility of sigma-delta modulator noise transfer function zeros optimization

The possibility of noise reduction due to zeros locations optimization is discussed. A sigma-delta modulator structure with complex feedback is analized. It is shown that for this structure zeros optimization is possible with some limitations.

The most important parameter of sigma-delta analog to digital converters is effective number of bits (ENOB). Usually it is defined as

$$ENOB = (SNR - 1.76)/6.02(bits), \quad (1)$$

where SNR is signal to quantization noise ratio, P_s is signal power, P_n – quantization noise power.

Maximization of ENOB in sigma-delta ADCs is made possible by sigma-delta modulator. The modulator structure is designed in such a way that useful signal experiences no change while quantization noise spectrum is shaped to shift the most noise power to high frequency band and minimize the noise power in the signal band. That is from the point of view of the quiantization noise the modulator is a high pass filter characterized by noise transfer function [1,2].

A variety of sigma-delta modulator structures are proposed [1,2]. The unification of their analysis can be achieved by matrix representation of the structures [3]. The matrix model provides derivation of signal and noise trunsfer functions as ratio of polynomials in z domain.

Sigma-delta modulator design involves a procedure of zeros optimization that allows to reduce quantization noise power by appropriate placing of noise transfer function zeros. The noise trunsfer function analysis shows that not all the structures allow to change zeros positions because the transfer function numerator polinomial coefficients do not depend on the structure coefficients as in [3]. To obtain movable zeros some additional feedbacks should be added to the structure. This gives the modulator structure with complex feedbacks shown in [1,2].

Development of the noise trunsfer function for the complex feedback modulator shows that the numerator coefficient at z^{-1} is constant and equals to negative of the modulator order. Therefore the transfer function zeros choise is limited by the zeros whose real parts sum to the modulator order. This makes impossible placing zeros exactly on the unit circle. As a result there are no pure nulls in the magnitude response. Instead there are notches whose depth depend on the distance from the zero to unit circle.

Thus it's importatn to compare different allowed zeros placements and choose the better. As a measure of zeros placement effectivness we use noise reduction factor caculated as the integral of the squared magnitude response over useful signal frequency band.

As the simplest zero placement method, used in [1], it is possible just to replace all zeros real parts by 1. In this case minimization of noise reduction factor is performed by sliding the zeros along the vertivcal line passing through the point $z = 1$. The same effect can be achieved by moving zeros along unit cuicle and replacing real parts by 1.

Another possibility can be used with modulators order equal or greater than 5. In this case every odd complex conjgate pare of zeros is placed on the unit circle while every even pare is scaled to obtain the sum of the foure real partes qual to 4. The singe real zero stays at the point $z = 1$.

Calcultions performed for the complex feedback sigma-delta modulators of 3rd and 5th order show that for zero frequencies less than $0.01F_s$ both methods of zero placement mentioned above give very close results that in turn are very close to those obtained when zeros are placed directly on the unit circle.

References. **1.** R. Schreier, G.C. Temes, Understanding Delta-Sigma Data Converters. Wiley-IEEE Press, 2005. **2.** O. Bajdechi, J.H. Huijsing, Systematic Design of Sigma-Delta Analog-to-Digital Signal Converters. Kluwer Academic Publishers, 2003. **3.** V. Artuhov, O. Brytov, Sigma-Delta Modulators Transfer Functions Derivation, IEEE 36th International Conference on Electronics and Nanotechnology (ELNANO), April 19–21, 2016, Kyiv, Ukraine, pp. 370–373.

Kachko N.A.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Genetic algorithms over distributed systems with MapReduce model

There is a need to develop an algorithm that predicts a value of some variable by having previous set of values over a span of time. More than that, it should be able to take into consideration values of the next timestamps. A crude oil or stock market prices can be used as an example. So there are two key goals to achieve in developed method: due to massive count of previous data it should be able to handle large chunks of values, and because of nature of trading market it has to predict the price before it actually changes, otherwise there is no sense in such method.

To meet first requirement a distributed computing exists. It is meant to take advantage of several nodes in order to increase overall performance for the price of small communication overhead. MapReduce, in particular, is an implementation of a distributed algorithm that consists of two parts. Map performs actual operation such as sorting or filtering, Reduce performs a summary operation, such as combining all items into one output [1].

To meet second requirement genetic algorithms can be used. They are proven to show mere precision in adequate terms. To satisfy this need there are quite a few methods that could be used as a solution.

It is known that fitness evaluation is the most CPU intensive tasks and can be very time consuming and therefore become a bottleneck in the overall performance [2], so using a number of resources would deal with potential degradation. But there is an issue: how a two-step computation algorithm can combine with a generation producer with an ability to constantly evolve and take into consideration new pieces of data?

To create a solution for this problem a new, modified genetic method should be developed. In order to ensure that each node of cluster takes approximately equal amount of data and performs same amount of computational work, the next algorithm of distribution, based on one described in [3], is proposed:

1. Each node of distributed network begins by creating a random initial population, then produces next generations via permutation, mutation, crossover and elite passing. In other words it proceeds with developing a prediction path as usual genetic algorithm, creating generations of candidates to use and sorting them by fitness function. At the last step a sorted list with best individuals should be produced. An example for 3 nodes is shown on fig. 1. Use of several calculation nodes is proven to produce greater precision and performance of process.

Node 1	
Candidate	Fitness function result
Candidate 10	0.95
Candidate 11	0.93
Candidate 12	0.91
Candidate 13	0.89
Candidate 14	0.87
...	...

Node 2	
Candidate	Fitness function result
Candidate 20	0.96
Candidate 21	0.93
Candidate 22	0.9
Candidate 23	0.87
Candidate 24	0.84
...	...

Node 3	
Candidate	Fitness function result
Candidate 30	0.94
Candidate 31	0.93
Candidate 32	0.92
Candidate 33	0.91
Candidate 34	0.9
...	...

Figure 1. Example of nodes computation result

2. Every node chooses several best candidates (shown bold red) and send them to the root node, where ones with more satisfying score (shown bold blue) are used as output, as shown on fig. 2.
3. As the new actual values of the variable come in, the prediction path of the output candidate should be corrected in a such way that the first predicted item becomes exact value. An example with three training and three predicted values is shown in the fig. 3–5.
4. As the new data comes in or more precision is needed a new instance of computation (or MapReduce routine) starting with first step of this algorithm may be launched.

Node 1	
Candidate	Fitness function result
Candidate 10	0.95
Candidate 11	0.93
Candidate 12	0.91
Candidate 13	0.89
Candidate 14	0.87
…	…

Node 2	
Candidate	Fitness function result
Candidate 20	0.96
Candidate 21	0.93
Candidate 22	0.9
Candidate 23	0.87
Candidate 24	0.84
…	…

Node 3	
Candidate	Fitness function result
Candidate 30	0.94
Candidate 31	0.93
Candidate 32	0.92
Candidate 33	0.91
Candidate 34	0.9
…	…

Root node	
Candidate	Fitness function result
Candidate 20	0.96
Candidate 10	0.95
Candidate 30	0.94
Candidate 11	0.93
Candidate 21	0.93
Candidate 31	0.93
Candidate 32	0.92
Candidate 12	0.91
Candidate 22	0.9

Figure 2. Summary of client nodes

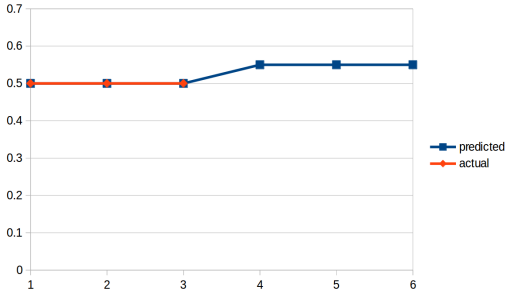


Figure 3. Predicted path

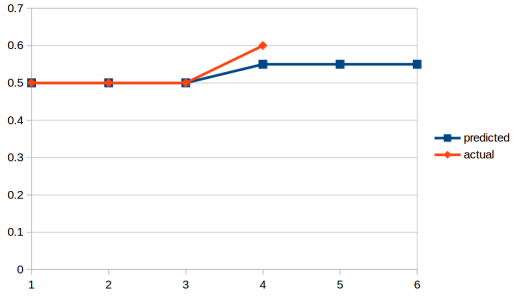


Figure 4. Next actual value differs from predicted

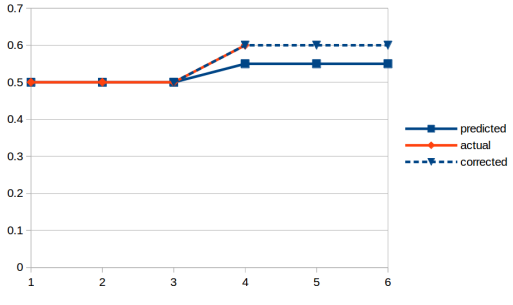


Figure 5. Correction of candidate's predicion

The outcome of proposed algorithm is improved performance and precision of prediction algorithms built over genetic methods that can be used in virtually unlimited number of appliances.

References. **1.** Wikipedia, the free encyclopedia (2017, Feb 14). *MapReduce* [Online]. Available: https://en.wikipedia.org/wiki/MapReduce. **2.** R.Nedunchelian et al., "Dynamic task scheduling using parallel genetic algorithms for heterogeneous distributed computing," M.S. thesis, Dept. Comp. Sci. Eng., Sri Venkateswara College, Pennalur, Sriperumbudhur, 2005. **3.** N. Kachko and V. Yaremenko, "Primality Test Problem", unpublished.

Kucher V.O.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

High performance RDMA-based Dataflow Management System

One of the modern research areas is the study of nuclear matter properties under extreme conditions. Experiments on heavy ion collisions at high densities and temperatures are planned at the GSI facility in Darmstadt, Germany, where a new international accelerator facility called FAIR (Facility for Antiproton and Ion Research) is being built. One of the major planned scientific pillars of FAIR is the CBM (Compressed Baryonic Matter) experiment [1].

The CBM experiment is searching for rare events so it has to be run repeatedly. It is planned to run 10 million times per second for 2 months with an estimated data generation rate of 1 TB/sec, which challenges todays possibilities in data processing, as there is no technology to transfer that much data over a single connection or to store it in a single place so it has to be processed on the fly. Due to free-running detectors the classical event building is replaced by a timeslice building. For a full event reconstruction each detector sensor data has to be exchanged and combined with data from other sensors, related to the same event or a group of events, called intervals within one Compute Node. Different processing intervals are distributed to different Compute Nodes [2]. This exchange is to be done over a high-speed low-latency network within a HPC cluster.

The designed dataflow management system is using the RDMA [3] protocol for Network abstraction and aims to apply its capabilities to the data exchange. It consists of a single helping service in addition to the applications on each of the connecting parties. The only functionality of the service is to keep track of the network and provide information for newly connecting applications.

The design of the RDMA-based Dataflow Management System [4] was targeted at solving two major requirements: high network bandwidth and reliability of operation. These were implemented by means of the following:

- By means of network stacks bypass and exchanging data directly between applications providing low latency the CPU utilization is reduced and high bandwidth is ensured.
- Memory is pinned using the memory registration procedure of the Operating System allowing remote access to selected memory regions.
- The designed system allowed extending and shrinking of the operating network without restarting.
- Extending is done by exchanging connectivity information through a central service.
- Shrinking is done by monitoring the connection status of working connections.
- Data distribution and combination is offloaded to the RDMA protocol, which notifications are used for maintenance.

The systems performance and reliability was verified by the results of a network testrun with 8 clients at maximal (for Infiniband) bandwidth for an hour. Connection management was performed by the server, which kept track of the network and provided connectivity information for newly connecting parties. The ability to reach the server at any time at a permanent location makes the network dynamic, being able to expand and shrink during the operation without restarting.

This system is currently unique in its behavior and is fault-tolerant while maximizing bandwidth.

References. **1.** FAIR. (2017) The compressed baryonic matter experiment. [Online]. Available: http://www.fair-center.eu/for-users/experiments/cbm.html. **2.** J. de Cuveland and V. Lindenstruth, "The CBM First-level Event Selector (FLES)," *ZIB*, Sep. 2013. **3.** Mellanox Technologies, *RDMA Aware Networks Programming user manual*, 1.7 ed., 2015. [Online]. Available: https://www.mellanox.com/related-docs/prod_software/RDMA_Aware_Programming_user_manual.pdf. **4.** V. Kucher, "Crow10: RDMA-based Dataflow Management System with Grid-style interconnection handling," Dec. 2016.

***Pechurin N.K.*[1] *Kondratova L.P.*[2] *Pechurin S.N.*[1]**
[1]*National aviation university, Kyiv, Ukraine;* [2]*Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine*

Reliable balanced UAV computing cluster and V.Leontyev model

Computing clusters are the set of connected computer systems, based on the use of widespread and relatively cheap information and telecommunication technologies, the corresponding software and hardware. Computing cluster for the user is presented as a single resource, which is carried by a related set of computer systems (complete computers), that own all resources necessary for their full operation, including processors, memory, input/output subsystem, etc.

This cluster resources generalization are necessary to the (parallel) solving as functional tasks, as processing streams of photo and video frames (eg, parallel execution of procedures of photos' zoom in JPEG, which received from UAV).

The balancing (alignment) of a load to a computing cluster is performed in order to optimize the use of resources, reduce query maintenance time, horizontal cluster scaling, as well as providing fault tolerance (redundancy) [1]. In this case the load is distributed between several computing elements, such as computers, telecommunication infrastructure (the main numerical characteristics are the delay time of the frames, the capacity of channels, the speed of data transmission over communication channels), central processors or disks and data warehouse management is improved, productivity, reliability and availability in data storage are increased [2].

A feature of the computing cluster for the solutions of UAV's functional problems is rapid changes, sometimes even in the decision process of the separate general problem, the topology of the computing and telecommunication system. Supplying of computing cluster reliability may be realized, for example, by RAID technology. As RAID technology implementation requires additional load to a computing cluster, fast (for UAV computing cluster especially) planning (balancing) instrument is necessary to design.

The balancing system is used to forward users' requests to the least loaded or most suitable complete computer with mirror copies of the information resource, which includes RAID systems.

Planning (parallel) applied problem solving using the available single computational cluster resource associated with the distribution of computing resources on the full computer of cluster in providing solutions of common problem.

The quick changeability of field communication system topology, even in the process to solve a separate general problem, leads to the necessity to set up operatively the parameters of the model to plan the performance process of full connected UAV (see fig.1).

Here there is proposed based on the V. Leontiev model, to realize a copying productivity distribute on complete computers method were RAID system are used [3].

The UAV cluster for the supplying its reliability by RAID requires n computing sources (complete computers).

It is given the intensity of producing process of computational resource, which we denote $X = [x_i]\,|_{1xn}$, where $x_i | i = \overline{1,n}$ is the intensity of producing process.

We are using a matrix of (technological) coefficients $A_{nxn} = [a_{ij}]$, where a_{ij} is the value of computational resource of i-th complete computers, which spends to copying (mirroring) data by RAID, per unit of computational resource of j-th complete computers.

Assuming linear function, describing i-j linking, that is

$$f_{ij}\,(a_{ij}) = a_{ij} * x_j,$$

obtain the model of distributed computing resource similar to V.Leontyev's model of interbranch

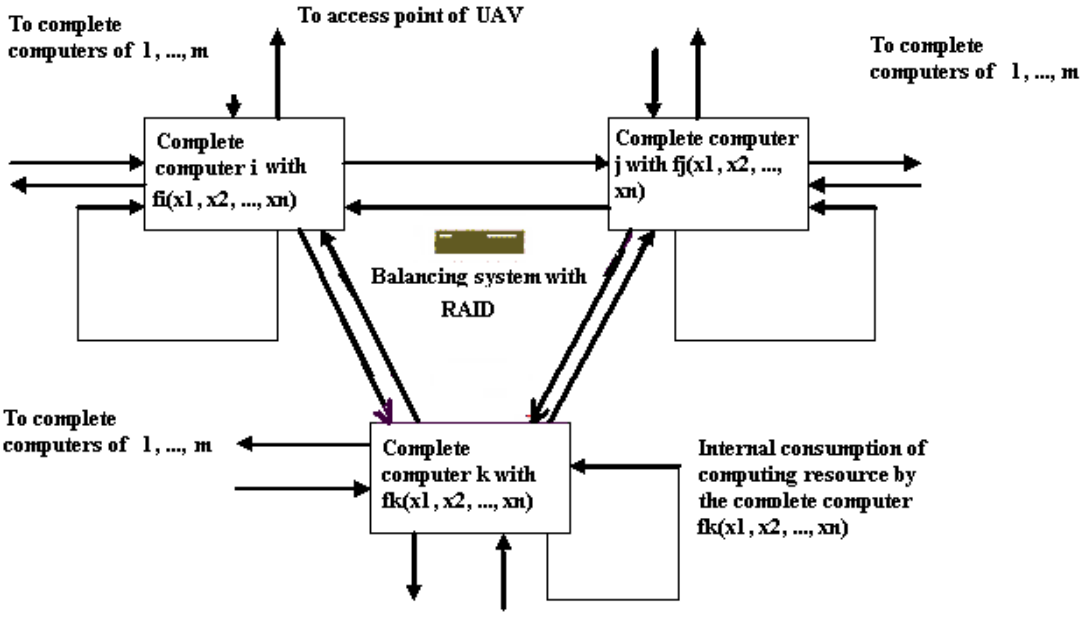


Figure 1. Fully connected computing cluster with RAID system of balancing.

balance as constrains of

$$(E - A) * X^T - R^T = 0.$$

Here E is the auxiliary identity matrix, R represents differs from UAV computing cluster absolute balance matrix.

For the identification of the parameters of A, X there is proposed to carry out by the method of principal components at $R = 0$. When $R = const \neq 0$, to determine A, X there was utilized a nonlinear equation system solving method.

The goal function formalizing lead to the special nonlinear programming model; to determine A, X were utilizing corresponding methods [4].

To supply reliable balance UAV computer cluster may be used RAID system; to supply RAID technology program implementation up UAV computer cluster may be used similar V. Leontyev model.

References. **1.** E. Dubrova. Fault-Tolerant Design. Springer, 2013. **2.** D. Patterson, G. Gibson, R. Katz, The types of the redundant arrays of inexpensive disks. – Berkeley: University of California, 1987. **3.** I.A. Zhukov, N.K. Pechurin, L.P. Kondratova, S.N. Pechurin. The distribution of resources in computing cluster for UAV // Problems of informatization and control: Proceedings, 2016, 3 (55), P. 1–5. **4.** V.V. Leontyev. Interbranch Economics, M.: 1997, 315 p.

Yaremenko V.S.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

An approach for data clustering CURE algorithm implementation using the MapReduce technology

CURE (Clustering Using REpresentatives) is large-scale-clustering algorithm in the point-assignment class. Instead of representing clusters by their centroid, it uses a collection of representative points, as the name implies [1]. For more efficiency CURE algorithm needs some modifications to be run in the distributed system and in this case the MapReduce [1] technology could be used.

The general model is presented on figure 1. It consists of N nodes that run Map program and 1 node that runs a Reduce program. Chunks from datasets are input to the map program. Map program finishes its work after the full dataset on this node has being processed.

The output of map program is a set of key-value pairs, where the key is equals to 1 and the value is a set of representative points from cluster(i,j). In this case i – it is a number of the node and j – a cluster number from the ith node.

As illustrated on the figure 1, the reduce program gets the output of the map program. The main task of reduce program is to merge clusters produced from all datasets. After merging is done, the reduce program produces the resulting dataset which contains a description of each cluster: centroid and representative points. Algorithms for map and reduce tasks are described in table 1.

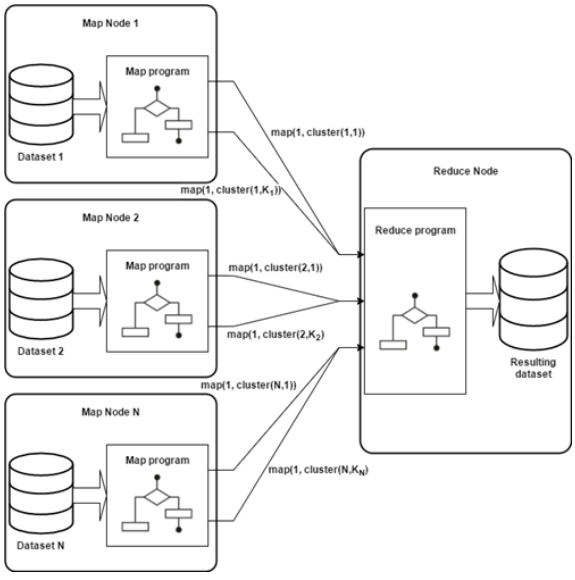


Figure 1. Data clustering model using the MapReduce technology

The advantage of the described approach is in distributed calculations on each node separately. Also, it is not necessary to load a full dataset into the RAM because map task processes a chunk of data in a correct way.

Table 1. Algorithms description

Map	Reduce
For each chunk in dataset 1. Load chunk to RAM (Random Access Memory) 2. Hierarchy clustering of the data in RAM 3. Centroid calculation for each cluster 4. Choose representative points for each cluster 5. Move each representative point to the centroid direction For each obtained cluster 1. Create key-value pair, key equals to 1 and value is a list of representative points 2. Send the created pair	For each input cluster 1. If it can be merged with any other one – step 2, otherwise – step 5. 2. Merge two clusters 3. Recalculate centroid 4. Choose representative points and continue loop 5. Save as a new cluster For each obtained cluster 1. Add to the list with representative points

References. 1. J. Leskovec, A. Rajaraman, J.D. Ullman (2014) Mining Of Massive Datasets, Second Edition. Cambridge University Press. ISBN-13: 978-1107077232. Print.

Гіоргізова-Гай В.Ш., Кульбачний К.А.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Організація обліку електронних медичних карт на базі хмарних сервісів Amazon AWS

Створення єдиної системи обліку електронних медичних карт (ЕМК) надає змогу оперативно і мобільно приймати медичні рішення, сприяє покращанню їх якості, є базою для проведення наукових і статистичних досліджень, тощо [1].

Система обліку ЕМК є високонавантаженою розподіленою системою. Перевагами рішень для побудови її серверної архітектури на основі хмарних технологій у порівнянні з традиційною фізичною інфраструктурою є нижча собівартість надання послуг завдяки відсутності необхідності мати декілька окремих мереж або центрів обробки даних, меншій кількості ліцензій на програмне забезпечення, меншій кількість обслуговуючого персоналу. Постачальник хмарних технологій може забезпечувати кваліфіковане обслуговування, високу безпеку, надійність і доступність даних [2].

За кількістю та вартістю сервісів, які необхідні для розгортання серверної інфраструктури системи обліку ЕМК у хмарному середовищі, і що надаються відомими постачальниками хмарних послуг, Amazon AWS виглядає найбільш привабливо завдяки більш простому регулюванню роботи усіх сервісів, більшій кількості підтримуваних баз даних, можливості реалізувати синхронну реплікацію у кластері Galera і т.д. [3].

На рис. 1 представлена архітектура серверної чистини системи обліку ЕМК, що забезпечує відмовостійкість системи, безперервну та швидку роботу додатків.

1. Сервіс VPC – надає мережеві послуги (розбиття на підмережі, налаштування безпеки трафіку).
2. Архівне сховище даних побудовано на основі сервісу Amazon RDS і використовує механізм синхронної реплікації Galera on MariaDB. Сховище відповідає за довготривале збереження інформації ЕМК, є базою для статистичного обслуговування, також через нього змінюється загальна інформація про пацієнта.
3. Оперативне сховище даних побудовано на основі СУБД NoSQL MongoDB через швидку генерацію документів та можливість їх зміни (орієнтованість на Web серверні додатки). Сховище отримує інформацію з архівного сховища на час лікування пацієнта, де до Mongo Document вносяться корективи, а після закінчення лікування вже скореговані дані повертаються назад. Для розгортання сховища використовуються спеціально підготовлені інстанси (екземпляри) Amazon AWS.
4. Сховище даних великого об'єму побудовано на основі Amazon S3. Тут зберігається мультимедійний контент, пов'язаний з ЕМК (фото, 3D томографічні зображення, аудіозаписи, тощо). У БД MariaDB та Mongo будуть зберігатися тільки гіперпосилання на файли, що лежать у бакетах (контейнерах) S3.
5. Інстанси EC2 з бекендом і фронтендом/або з фронтендом під ОС MS Server 2016 Enterprise – це місце розташування ПЗ фронтенду додатку, що створюється на платформі .NET (обрано через можливість інтеграції з MongoDB, Galera, S3 та можливість каскадного оновлення за допомогою Chocolatey). Інстанси, що тут знаходяться, є одночасно бекендом і фронтендом (сервер IIS, що своїм пулом налаштований на теку додатку). Бекенд має прямий доступ до баз даних та S3, і за бажанням користувача в залежності від типу операції може створювати або змінювати інформацію у будь-якому сховищі.
6. Балансувальник навантаження Load Balancer побудовано на основі Amazon ELB. Він використовується для рівномірного розподілу запитів клієнтів до інстансів EC2 з бекендом /фронтендом та їх динамічного масштабування в залежності від навантаження системи. Сховища даних на Galera on MariaDB та MongoDB є саморегульованими системами, які самостійно розподіляють навантаження між узлами і мають функцію знаходження та

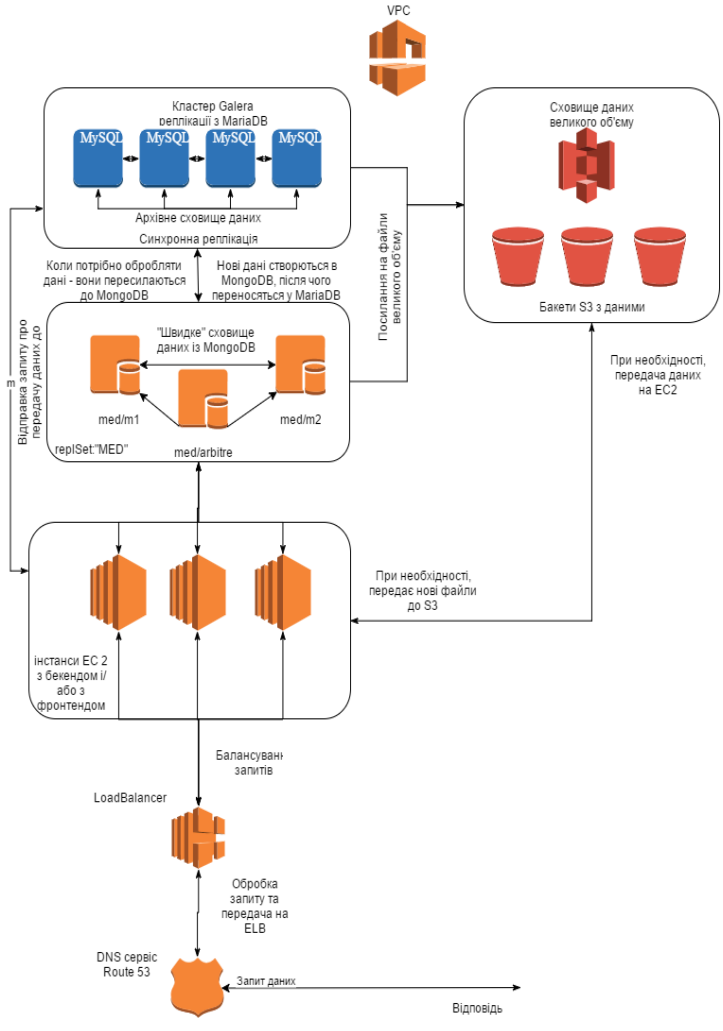


Рис. 1. Загальна архітектура серверної чистини системи обліку ЕМК

перезапуску пошкоджених вузлів.

7. Route53 – служба реєстрації імен. Вона слугує для того, щоб надавати сервісам, на яких працює тека, логічні DNS імена, а також для створення розподіленої системи імен, що надасть кожному регіону держави свою точку доступу на порталі медичних карток. Через route53 йде звернення до додатку через DNS ім'я балансувальника. Трафік, що поступає через route53 додатково фільтрується за допомогою політик безпеки VPC.

Запропонована архітектура для обліку електронних медичних карт на базі хмарних сервісів Amazon AWS може бути використана регіональною громадою, крупною спортивною або іншою установою, члени якої мають потребу у швидкому і якісному наданні медичних послуг.

Література. **1.** Супруненко О.О. Проблеми впровадження інформаційних комп'ютерних технологій в медичну галузь. // Тезисы и доклады Международной конференции «Интегративная медицина». – Київ: Алтимед, 2009. – С. 164–166. **2.** Rodney Lusk, MD. Cloud Computing May Be a Simple Solution for Your EHR Needs, http://www.enttoday.org/article/cloud-computing-may-be-a-simple-solution-for-your-ehr-needs/2/. **3.** Офіційний сайт Amazon AWS, http://aws.amazon.com/.

Єршов С.В., Пономаренко Р.М.
Інститут кібернетики ім. В.М.Глушкова НАН України, Київ, Україна

Паралельні обчислення багаторівневих нечітких систем Такагі–Сугено

Розглянуто концепцію та реалізацію ярусо-паралельної та динамічної моделей для здійснення нечіткого логічного виведення в експертно-діагностичних програмних системах, бази знань яких ґрунтуються на нечітких правилах. Побудовані оцінки ефективності ярусно-паралельної та динамічної схем обчислень графів залежностей різної складності між блоками нечітких правил Такагі–Сугено. Проведено порівняльну характеристику ефективності ярусно-паралельної та динамічної моделей.

На даний час для побудови експертних та діагностичних інтелектуальних систем застосовеються моделі та методи, що грунтуються на системах нечіткого логічного виведення. Такі системи нечіткого виведення складаються з набору правил якщо-то. Нечітка модель Такагі–Сугено представлена нечіткими правилами [1]:

$$R_j : if\ x_1\ is\ A_{1j}\ and\ x_2\ is\ A_{2j}\ and\ ...\ and\ x_n\ is\ A_{nj},$$
$$then\ y = g_j(x_1,\ x_2,\ ...,\ x_n),\ j = 1, 2, ..., N.$$

Разом зі зростаючим розповсюдженням таких систем, зростає їх складність. Вона пов'язана головним чином з великою кількістю входів нечіткої системи і, як наслідок, різким збільшенням кількості правил для обробки вхідних даних, що мусить писати експерт. Для вирішення проблеми розмірності було запропоновано використання багаторівневих ієрархічних систем нечіткого логічного виведення, ідея яких полягає в тому, що на вхід однієї системи подається вихідне значення попередньо обробленої нечіткої системи, що логічно зв'язана з нею (рис. 1). Такий підхід дозволяє контролювати складність кожної елементарної нечіткої системи, проте виникає наступна проблема. Це значно зростаюча кількість обчислень, які необхідно виконати разом із зростанням часу виконання таких програмних систем. В даній роботі запропоновано два підходи до зменшення часу виконання багаторівневих нечітких програмних систем на основі паралельної обробки незв'язаних між собою систем. Багаторівнева архітектура використовується при побудові нечітких мультиагентних систем, що функціонують в розподіленому середовищі [2].

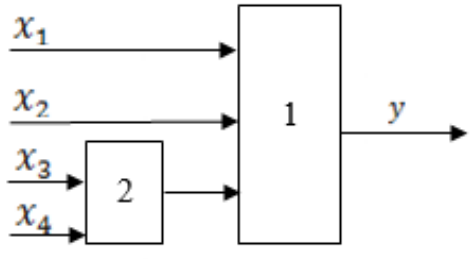


Рис. 1. Приклад багаторівневої системи нечіткого логічного виведення

Зазначимо, що при реалізації даних схем розпаралелювання була використана технологія MPI на основі обміну повідомленнями між процесами. Розглянемо обидві схеми детальніше [3].

Ярусно-паралельна схема. Принцип роботи даної схеми розподіляється на два етапи.

1. Створюється план розпаралелювання. Будується дерево ярусів системи, вершинами цього графа виступають системи нечіткого логічного виведення. Після чого також визначається кількість ітерацій, за яку буде розрахована вся багаторівнева система. Зазначена кількість залежить від числа ярусів та кількості робочих процесорів. У даному випадку кількість робочих процесорів дорівнює N, де N – кількість задіяних процесорів в цілому. Кожній вершині присвоюється номер процесу, на якому вона буде виконана.
2. Виконуються обчислення багаторівневої системи згідно вищезазначеного плану розпаралелювання. Кожний процесор, після обчислення чергової нечіткої системи, за допомогою повідомлень MPI передає результати обчислень всім процесорам, що виконують обчислення систем, які ієрархічно вищі та залежать від даної нечіткої системи.

Динамічна схема. Особливістю динамічної схеми є відсутність побудови плану розпаралелювання. Одразу вибирається кількість доступних вершин та розпочинається їх обчислення. По мірі виконання, вільним процесам надається команда на виконання наступних вільних вершин. Таким чином, зводиться до мінімуму час простою вільних процесорів. Але в даній схемі один процес власне обчислень не виконується. Він займається розподілом завдань та обміном повідомленнями між робочими процесами, тому кількість робочих процесів тут $N-1$.

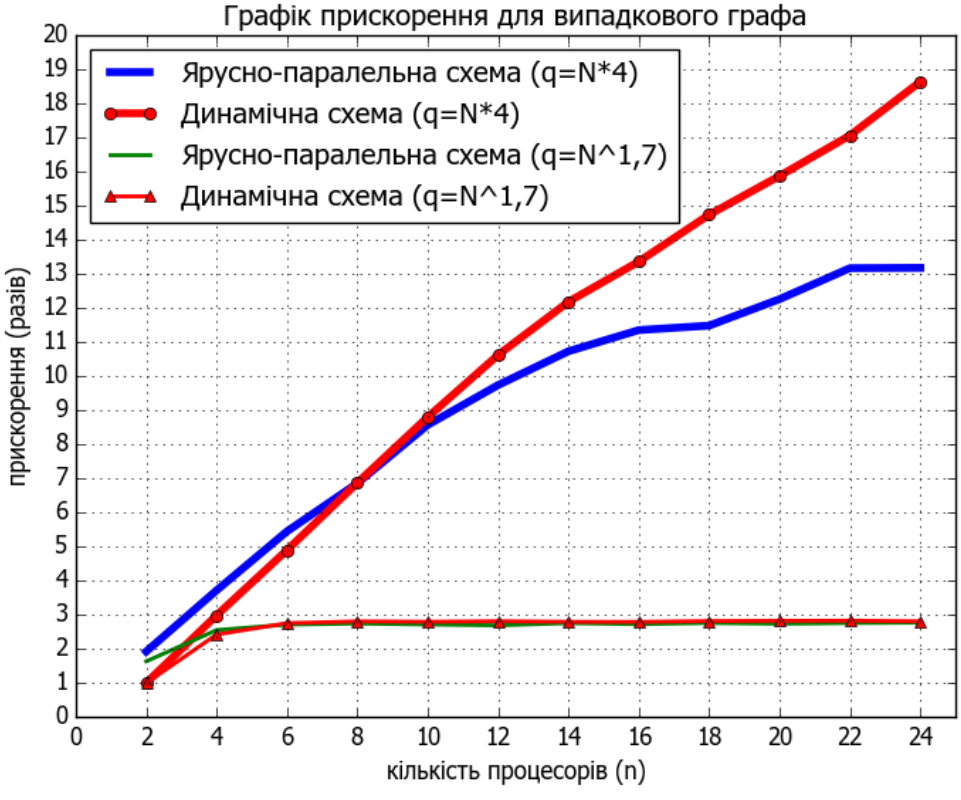


Рис. 2. Графік прискорення для двох паралельних схем, де N – кількість нечітких систем (вершин графа), q – кількість ребер в залежності від N (кількість зв'язків між системами)

Висновки. Зі збільшенням кількості процесорів зростає прискорення для обох схем розпаралелювання (рис. 2), але при суттєвому їх збільшенні динамічна схема має кращі показники прискорення, ніж ярусно-паралельна схема за рахунок відсутності ярусів та мінімального простою процесорів. Всі обчислення проводилися на суперкомп'ютері СКІТ Інституту кібернетики НАН України [4]. Проте у випадку, коли кількість процесорів досить невелика, ефективнішою є саме ярусно-паралельна, адже у неї для обчислень нечітких систем задіяні всі наявні процесори.

Література. **1.** Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А.В. Леоненков. – СПб.: БХВ, 2005. – 736 с. **2.** Ершов С.В. Принципы построения нечетких мультиагентных систем в распределенной среде / С.В. Ершов // Компьютерная математика. – 2009. – № 2. – С. 54–61. **3.** Ершов С.В., Пономаренко Р.М. Методи організації паралельних обчислень для багаторівневих нечітких систем Такагі–Сугено // Проблеми програмування. – 2016. – № 1. – С. 141–149. **4.** Суперкомпьютеры ИК НАН Украины, icybcluster.org.ua.

Капшук О.О.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Реалізація біометричних технологій аутентифікації на мобільних пристроях

В роботі розглядаються питання реалізації біометричних технологій аутентифікації на мобільних пристроях (смартфони, планшети) для отримання доступу до веб-сервісів. Основна увага приділяється забезпеченню високої надійності, безпечності та зручності процессів аутентифікації за рахунок використання багатофакторних та мультимодальних біометричних технологій, які спрощують процес отримання доступу до веб-сервісів і усувають проблеми використання традиційних паролів.

Використання мобільних пристроїв (МП) для отримання доступу до багатьох веб-сервісів (ВС) стає в даний час все більш популярним в усьому світі. Наявність великої кількості мобільних пристроїв, здатних забезпечити доступ до ВС постачальників різних послуг, призводить до швидкого зростання ринку мобільних сервісів. За даними дослідження Biometrics Research Group, Inc. об'єм сегменту мобільної біометрії в усьому світі оцінюється у $9 млрд. дол. до 2018 р. і $45 млрд. до 2020 року [1]. Такий великий об'єм ринку ВС стає дуже привабливим для шахраїв. За прогнозами компанії Cybersecurity Ventures до 2021 року збитки від кібершахрайства збільшаться удвічі в порівнянні з 2015 роком і досягнуть $6 трильйонів доларів [2]. Протидія загрозам інформаційній безпеці (ІБ), що виникають при використанні мобільних сервісів, вимагає впровадження надійних систем мобільної аутентифікації користувачів (СМАК), що поєднують в собі зручність використання і посилений захист.

Біометрична аутентифікація може бути використана для забезпечення кращого захисту паролем облікових записів онлайн сервісів або особистого обладнання (наприклад, телефонів, планшетів або ПК). Біометричні дані все частіше використовуються для аутентифікації фінансових операцій, особливо платежів. Слід зазначити, що протягом тривалого часу в системах обмеженого доступу стаціонарного типу успішно застосовуються біометричні технології, в яких для ідентифікації і аутентифікації людей використовуються різні біометричні характеристики. Вибір біометричних факторів для реалізації СМАК на МП можна зробити з урахуванням можливостей МП (вбудовані відео камери, мікрофони, сенсори відбитків пальців), а також вимог до безпечності та зручності використання.

В роботі розглядаються питання надійності мобільної біометричної аутентифікації користувачів, а також безпечності використання мобільних пристроїв і сервісів.

Аналіз існуючих джерел дозволяє зробити висновок, що найбільш популярними на даний час є реалізації СМАК з використанням багатофакторних мультимодальних біометричних технологій. Існує ряд переваг, властивих мультимодальній біометрії, причому найбільш вагомими з них є підвищені рівні безпеки і надійності (в разі багатофакторних рішень).

Однією з проблем використання СМАК є можливість обману (спуфінгу). Як відомо, захист від спуфінгу в таких системах забезпечується додатковою перевіркою на живучість (належність біометричних параметрів живій людині). Використання такої перевірки суттєво підвищує надійність аутентифікації. Все більш популярним методом боротьби зі спуфінгом є поєднання біометричних факторів. Багатофакторний підхід до аутентифікації експоненціально підвищує рівень безпеки СМАК. Захист СМАК від атак типу «АТА» і «Man-In-The-Middle»

На сьогоднішній день практично всі смартфони та планшети стали носіями важливих персональних або корпоративних даних. Також за допомогою МП власників можна легко отримати доступ до їх облікових записів, таких як Gmail, FaceBook, інших соціальних мереж і навіть до корпоративних сервісів. Тому наразі гостро стоїть проблема забезпечення конфіден-

ційності таких даних і розробки засобів для захисту МП від несанкціонованого доступу (атак типу«АТА», After-Theft-Attacks) у разі їх викрадення або втрати.

Атаки типу MitM («Man-In-The-Middle») дуже поширені і, як свідчать чисельні публікації, успішні. В 2016 році за даними робочої групи по боротьбі з фішингом (The Anti-Phishing Working Group) було встановлено рекорд за кількістю фішингових атак – 1 220 523 атак [3]. Це на 65% більше, ніж у 2015 р. Найбільша кількість атак припадає на ритейл (майже 42%), на другому місці виявився фінансовий сектор, на третьому і четвертому місцях, відповідно, – інтернет-провайдери та платіжні сервіси. Велика частина (понад 6 тис.) фішингових доменів стабільно розташовується в доменній зоні .com. За останні дванадцять років кількість інцидентів, пов'язаних з фішингом, зросла більше ніж в п'ятдесят разів.

В роботі проведено аналіз засобів багаторівневого захисту від зазначених вище атак типу АТА та MitM за допомогою прив'язки зареєстрованого мобільного пристрою до певного облікового запису користувача шляхом надання персоналізованого криптографічного ключа, прихованого в декількох рівнях шифрування.

Слід зауважити, що з кожним роком технічні характеристики МП змінюються, збільшується потужність процесора, з'являються вбудовані 3D камери та сенсори глибини, які дозволяють при аутентифікації використовувати 3D технології розпізнавання обличчя користувачів і, як наслідок, підвищити надійність аутентифікації і зменшити імовірність спуфінгу. Існує багато компаній, які пропонують засоби для підтримки та розробки СМАК, які можна розглядати як біометричні сервіси для реалізації різних протоколів біометричної аутентифікації користувачів ВС.

При проектуванні СМАК в залежності від необхідного рівня безпеки можуть використовуватись наступні підходи:

1. Повністю децентралізована архітектура безпеки.
2. Конфіденційність проектування. Анонімна аутентифікація без використання персональних даних.
3. Безпечність за рахунок використання багатомодальної біометрії з врахуванням ознак належності біометричних даних живій людині та персональний захист біометричних даних.
4. Проектування на базі хмарних технологій, яке легко імплементується і доступне будь-де, у будь-який час, на будь-якому пристрої.
5. Вбудоване в пристрій користувача розпізнавання біометричних параметрів.
6. Використання для СМАК загального призначення таких біометричних характеристик користувача, як голос та обличчя.
7. Біометричні характеристики користувача не зберігаються у відкритому вигляді, зберігаються тільки зашифровані шаблони на пристрої користувача.

Література. **1.** Biometrics Research Group, Inc releases report on mobile biometrics market [Електронний ресурс] // Режим доступу: – www.biometricupdate.com/201510/biometrics-research-group-inc-releases-report-on-mobile-biometrics-market. – Дата доступу 10.10.2016. – Назва з екрану **2.** Убытки от киберпреступности увеличатся вдвое к 2021 году [Електронний ресурс] // Режим доступу – http://www.beznal.pro/news/9002-ubytki-ot-kiberprestupnosti.html. – Дата доступу 10.10.2016. – Назва з екрану. **3.** Для бизнеса по обману интернет-пользователей 2016 г. стал лучшим в истории. [Електронний ресурс] // Режим доступу: – http://safe.cnews.ru/news/top/2017-03-02dllyabiznesapoobmanuinternetpolzovatelej. Дата доступу 10.10.2016. – Назва з екрану.

Кирюша Б.А., Колінько А.М.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Порівняльна характеристика бюджетних embedded рішень для обробки відеопотоку

Для порівняння було виділено 4 бюджетні варіанти:

1. модуль ESP32, камера OV7670;
2. модуль Arduino Uno та ArduCAM-M-2MP Camera Shield;
3. модуль Mini Openwrt Dev Board, кабель OTG Host Power Splitter Micro USB Male to USB A Female And Micro USB Female Cable та камера TeckNet C016 USB HD 720P Webcam;
4. модуль Spartan 6 XC6SLX9 FPGA Development Board та камера ov7670;
5. модуль Raspberry Pi 3 Model B Motherboard та камера Raspberry Pi Camera Board.

Модуль ESP32, камера OV7670. Модуль ESP32 має тактову частоту (англ. clock rate) 240 MHz (в деяких моделей – 160 MHz) [1], що дає досить широкі можливості для обробки відеопотоку з камери. Також ESP32 підтримує безпровідні стандарти Wi-Fi та Bluetooth, що надає додаткові можливості для IoT розробок. Сама камера представляє собою сенсор з процесором цифрової обробки сигналів (англ. digital signal processor, DSP). Таким чином, конфігуруючи відповідні регістри, можна задавати формат зображення та швидкість відеопотоку (кадрова частота – англ. Frames per Second (FPS)) ще до відправки даних на ESP32. Для конфігурації даних регістрів використовується протокол Serial Camera Control Bus (SCCB) [2].

Модуль Arduino Uno та ArduCAM-M-2MP Camera Shield. ArduCAM для передачі даних використовує SPI, а для конфігурації регістрів камери – I^2C. Камера може працювати в 7 режимах [3]: Single Capture Mode, Multiple Capture Mode, JPEG Compression, Normal Read and Burst Read Operation, Rewind Read Operation, Low Power Mode, Image Sensor Control.

Модуль Mini Openwrt Dev Board, кабель OTG Host Power Splitter Micro USB Male to USB A Female And Micro USB Female Cable та камера TeckNet C016 USB HD 720P Webcam. Mini Openwrt Dev Board має 512 Mbit оперативної пам'яті та досить потужний процесор з тактовою тастотою 600MHz, що дозволяє опрацьовувати більші (в порівнянні з попередніми варіантами) обсяги відеоданих. Проте для підключення камери в цьому випадку треба додатковий кабель. Mini Openwrt Dev Board має встановлену OpenWRT, проте можна оновити заводську прошивку [4]. Після підключення камери, встановлення відповідного SDK та перепрошиття Mini Openwrt Dev Board, для відео трансляції можна використовувати пакет MJPEG-Streamer.

Модуль Spartan 6 XC6SLX9 FPGA Development Board та камера OV7670. Spartan 6 XC6SLX9 має 9152 логічних комірок (англ. logic cell). Кожна комірка має таблицю істинності (англ. Look-Up Tables), D-тригер, мультиплексор та зв'язки з сусідніми комірками. Код для Spartan 6 XC6SLX9 пишеться на мові опису апаратури (Verilog/VHDL). Це є одночасно перевагою і недоліком даного рішення: написання коду, як правило, займає більше часу, проте за рахунок низького рівня абстракції досягається висока швидкодія. В Spartan 6 XC6SLX9 досить мало пам'яті (90Kbit розподіленої пам'яті та 576Kbit блочної), тому потрібно організувати швидку передачу даних з Spartan 6 XC6SLX9 на USB-UART (чи на інший кінцевий блок), аби не допустити переповнення буферів і, як наслідок, втрати даних. Для конфігурації регістрів DSP камери (див. варіант 1) існує модуль ініціалізації [5], проте він написаний для Nexys4-DDR board.

Модуль Raspberry Pi 3 Model B Motherboard та камера Raspberry Pi Camera Board. Такий варіант є широко поширеним і добре документованим [6]. Raspberry Pi не має передвстановленої операційної системи, проте її можна встановити з SD-карти. Виробник рекомендує 6 операційних систем [7]: Raspbian, Pidora, Arch Linux, Risc OS, ARM-версія OpenELEC, RaspBMC. Після встановлення операційної системи можна писати код, наприклад, на Python, де вже реалізовані спеціальні бібліотеки для роботи з камерою. Проте можна обійтися і без

операційної системи, використовуючи GNU ARM Embedded Toolchain [8]. В цьому випадку можна отримати більшу швидкодію, але такий спосіб вимагає написання коду на C та аналізу документації Raspberry Pi та Raspberry Pi Camera Board.

Порівняльна таблиця. При виборі того чи іншого варіанту потрібно шукати компроміс між тактовою частотою модуля і кількістю кадрів за секунду, які цей модуль встигне опрацювати. Якщо модуль не встигає опрацювати відеопотік, варто зменшити fps і/або оптимізувати код, використовуючи, де можливо, апаратну реалізацію.

Також варто враховувати співвідношення пам'ять/якість результуючого зображення. Якщо оперативної пам'яті не вистачає для збереження зображення даного формату, можна перейти до Grayscale зображень і/або зменшити роздільність. Також можна передавати на кінцевий блок (USB-UART, напр.) не всі дані кадру, а частинами. Проте в цьому випадку деяка кількість тактів модулю буде використана для частішої передачі даних.

Табл. 1. Порівняльна таблиця

Варіанту	clock rate (MHz)	RAM capacity (Mbit)	Формати	max fps [1]	Розмір матриці	Wi-Fi
ESP32, OV7670	240	4.0625	VGA QVGA CIF QCIF	30	640 x 480	Підтримується
Arduino Uno, ArduCAM-M-2MP	16	0.015625 (2Kbyte)	RGB565/555 GRB422 YUV YCbCr JPEG	60	1600×1200	Не підтримується
Openwrt Dev Board, TeckNet C016	600	512	MJPEG JPEG	30	1280x720	Підтримується
Spartan 6 XC6SLX9 OV7670	100	≈ 0.65 (90kb + +576kb)	VGA QVGA CIF QCIF	30	640x480	Не підтримується
Raspberry Pi, Raspberry Pi Camera	1200	8192 (1 GB)	JPEG H.264	90	2592×1944	Підтримується

[1] max fps – максимальна кадрова частота, яку підтримує камера.

Література. **1.** V. Shymanskyy, "ESP32: знакомимся, пишем и запускаем первую прошивку", Habrahabr.ru, 2016. [Online]. Available: https://habrahabr.ru/post/309746/. [Accessed: 22- Feb- 2017]. **2.** "Serial Camera Control Bus (SCCB) Functional Specification", Ovt.com, 2007. [Online]. Available: http://www.ovt.com/download_document.php?type=document&DID=63. [Accessed: 22- Feb- 2017]. **3.** "ArduCAM-M-2MP Camera User Guide", http://www.arducam.com, 2015. [Online]. Available: http://www.arducam.com/downloads/shields/ArduCAM_Mini_2MP_Camera_Shield_DS.pdf. [Accessed: 22- Feb- 2017]. **4.** "Refresh the firmware – WRTnode Wiki Site", Wiki.wrtnode.com, 2014. [Online]. Available: http://wiki.wrtnode.com/index.php?title=Refresh_the_firmware. [Accessed: 22- Feb- 2017]. **5.** "GitHub – westonb/OV7670-Verilog: Verilog modules required to get the OV7670 camera working", Github.com, 2016. [Online]. Available: https://github.com/westonb/OV7670-Verilog. [Accessed: 22- Feb- 2017]. **6.** "Worksheet – Getting started with picamera | Raspberry Pi Learning Resources", Raspberrypi.org, 2017. [Online]. Available: https://www.raspberrypi.org/learning/getting-started-with-picamera/worksheet/. [Accessed: 22- Feb- 2017]. **7.** А. Гагарин, "Raspberry Pi: первое знакомство. Урок 1. Устройство и установка ОС", Занимательная робототехника, 2014. [Online]. Available: http://edurobots.ru/2014/05/raspberry-pi-pervoe-znakomstvo-urok-1-ustrojstvo-i-ustanovka-os/. [Accessed: 23- Feb- 2017]. **8.** "Программируем Raspberry Pi на голом железе", Habrahabr.ru, 2014. [Online]. Available: https://habrahabr.ru/post/208726/. [Accessed: 23- Feb- 2017].

Марченко О.О., Марченко О.І.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Модель динамічного розпаралелення пошуку в дереві методом Монте-Карло для grid-систем

Розробка нових методів та алгоритмів, що можуть прискорити обчислення задач, час рішення яких зараз є занадто великим, була і залишається актуальною проблемою сьогодення. Особливо це важливо в галузі задач штучного інтелекту, яка відзначається надзвичайно великим обсягом даних і високими вимогами до швидкості їх обробки. Однією з найбільш поширених форм подання інформації в задачах штучного інтелекту є форма у вигляді дерев послідовних рішень, а одним з методів штучного інтелекту, що був запропонований відносно недавно і призначений для виконання швидкого пошуку правильних рішень у дереві інформації, є метод пошуку в дереві з використанням методу Монте-Карло (Monte Carlo Tree Search – MCTS) [1]. Цей метод вже встиг суттєво вплинути на розв'язання задач штучного інтелекту, а особливо складних ігрових задач, таких як ігри в шахи та Go. Метою даної доповіді є аналіз існуючих рішень з розпаралелення пошуку по дереву методом Монте-Карло та узагальнення різних підходів до розпарелелення у вигляді моделі.

Структурно-алгоритмічні особливості методу MCTS для розпаралелення. На сьогоднішній день для розпаралелення пошуку в дереві методом MCTS були запропоновані три загальних підходи: листкова паралелізація, коренева паралелізація і деревна паралелізація [1]. Для реалізації цих підходів вже було запропоновано досить багато способів [1–7], але в жодному з них не використовуються динамічна зміна політик кожного кроку MCTS під час роботи пошуку, а також динамічний розподіл апаратних ресурсів, як з метою кращого завантаження цих ресурсів (що дає додаткове прискорення процесу пошуку), так і з метою динамічного контролю за формою дерева MCTS (як воно зростає швидше, вглибину чи вширину) [8].

Модель динамічної паралелізації пошуку в дереві методом Монте-Карло в grid-системах. Для того, щоб побудувати узагальнену модель динамічної паралелізації пошуку в дереві методом Монте-Карло в grid-системах, спочатку визначимо її дві складові підмоделі, необхідні для прийняття рішення про подальшу паралелізацію у конкретний момент часу, а потім розширимо її іншими параметрами.

Такими двома підмоделями є:

STM (Search Tree Model) – модель дерева пошуку MCTS;

GAM (Grid Archtecture Model) – модель архітектури grid-системи.

Модель STM дерева пошуку MCTS визначимо як

$$STM(t) = (STW(t), STD(t), NW(t), NP(t), t),$$

де

STW (Sub-Tree Width) – ширина піддерева MCTS від вершини дерева, де приймається поточне рішення в момент часу t;

STD (Sub-Tree Depth) – глибина піддерева MCTS від вершини дерева, де приймається поточне рішення в момент часу t;

NW (Number of Wins) – кількість виграшів, отриманих від вершини дерева, де приймається поточне рішення в момент часу t;

NP (Number of Playouts) – кількість вже виконаних разів моделювання гри, починаючи від вершини дерева, де приймається поточне рішення в момент часу t.

Модель GAM архітектури grid-системи визначимо як

$$GAM(t) = (NN, Node[1...NN], t),$$

де

NN (Nodes Number) – загальна кількість grid-вузлів;

$Node[1...NN]$ – масив дескрипторів grid-вузлів, в якому кожен i-й grid-вузол визначається як

$Node[i] = (NCC, NGC, Free(t))$,

де

NCC (Number of CPU Cores) – кількість ядер центрального процесора CPU;

NGC (Number of GPU Cores) – кількість ядер графічного процесора GPU;

$Free$ (Flag "free/busy") – прапорець "вільний/зайнятий" в момент часу t.

На основі цих підмоделей і загальних принципів роботи пошуку методом MCTS, визначимо модель динамічної паралелізації DPM (Dynamic Parallelization Model) пошуку в дереві методом Монте-Карло в grid-системах як

$$DPM(t) = (STM(t), GAM(t), TP(t), EP(t), SP(t), BPP(t), PP(t), t),$$

де

STM (Search Tree Model) – модель дерева пошуку MCTS в момент часу t;

GAM (Grid Archtecture Model) – модель архітектури grid-системи в момент часу t;

TP (Tree Policy) – політика дерева (політика вибору вершини дерева для розширення) в момент часу t;

EP (Expansion Policy) – політика кроку розширення в момент часу t;

SP (Simulation Policy) – політика кроку моделювання в момент часу t;

BPP (BackPropogation Policy) – політика зворотного обчислення параметрів передбачення (фактично є певною формулою переобчислення) в момент часу t;

PP (Parallelization Policy) – політика паралелізації в момент часу t.

Політика паралелізації PP, в свою чергу, також є моделлю, яка описує види паралелізації MCTS, та їх залежність від інших параметрів описаної вище моделі.

$$PP(t) = (RPD(t), LPD(t), TPD(t), DWC(t), t),$$

де

RPD (Root Parallelization Degree) – степінь кореневої паралелізації в момент часу t;

LPD (Leaf Parallelization Degree) – степінь листкової паралелізації в момент часу t;

TPD (Tree Parallelization Degree) – степінь деревної паралелізації в момент часу t;

DWC (Depth-Width Criterion) – критерій «глибина-ширина» форми піддерева MCTS від поточної вершини [8] в момент часу t.

Висновки. Паралелізація пошуку в дереві з використанням методу Монте-Карло MCTS може бути виконана згідно декількох різних підходів. На основі аналізу цих підходів запропонована модель динамічної паралелізації пошуку в дереві методом Монте-Карло в grid-системах. Напрямком подальшого дослідження може бути розробка нового методу розпаралелення MCTS на основі запропонованої моделі.

Література. **1.** Cameron Browne and others. A Survey of Monte Carlo Tree Search Methods // IEEE Trans. on Computational Intelligence and AI in Games. – vol. 4. – no. 1. – March 2012. **2.** 2. G.M.J.-B. Chaslot, M.H.M. Winands, and H.J. van den Herik. Parallel Monte-Carlo Tree Search // Proc. Comput. And Games, LNCS 5131, Beijing, China. – 2008, pp. 60–71. **3.** T. Cazenave and N. Jouandeau. On the Parallelization of UCT // Proc. Comput. Games Workshop, Amsterdam, Netherlands. – 2007, pp. 93–101. **4.** H. Kato and I. Takeuchi. Parallel Monte-Carlo Tree Search with Simulation Servers // Proc. Int. Conf. Tech. Applicat. Artif. Intell., Hsinchu City, Taiwan. – 2010, pp. 491–498. **5.** A. Fern and P. Lewis. Ensemble Monte-Carlo Planning: An Empirical Study // Proc. 21st Int. Conf. Automat. Plan. Sched., Freiburg, Germany. – 2011, pp. 58–65. **6.** Y. Soejima, A. Kishimoto, and O. Watanabe. Evaluating Root Parallelization in Go // IEEE Trans. Comp. Intell. AI Games. – vol. 2. – no. 4, – 2010, pp. 278–287. **7.** S. Gelly, J.-B. Hoock, A. Rimmel, O. Teytaud, and Y. Kalemkarian. The Parallelization of Monte-Carlo Planning // Proc. 5th Int. Conf. Inform. Control, Automat. and Robot., Funchal, Portugal. – 2008, pp. 244–249. **8.** Марченко О.О., Марченко О.І. Критерій «глибина-ширина» для контролю форми дерева пошуку при використанні методу Монте-Карло // Комп'ютерно-інтегровані технології: освіта, наука, виробництво, 2016. – № 24–25. – С. 42–57.

Орлова М.М., Телелейко І.С.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Сумісність підходів хмарних обчислень та грід-технологій

Існує безліч обчислювальних систем різного масштабу для обробки інформації, які застосовуються в різноманітних сферах діяльності. Найбільший інтерес представляють системи, які обробляють надвеликі обсяги даних.

Для грід-обчислення характерними є: відсутність центру управління оброблюваних ресурсів, використання відкритих стандартів, використання спеціалізованого рівня обслуговування; передбачає колективний розподілений режим доступу до ресурсів і пов'язаних з ними послуг у межах глобально-розподілених віртуальних організацій. За допомогою грід-обчислень є можливість надавати обчислювальні ресурси як утиліти, які можуть бути як активні, так і пасивні. Структурою грід є віртуальна організація, задачею якої є створення середовища, де фрагменти одного додатку можуть оброблятися на різних комп'ютерах та мають взаємозв'язок між собою. Віртуальна організація утворюється над простором реальних комп'ютерів, мереж, адміністративних зон, центрів обчислювання даних [1].

Хмарна інфраструктура суттєво спрощує та зменшує вартість розробки грід-додатків. Якщо існують дані, які потребують обробки, необхідно відправити запит на сервер для обробки цих даних. Після завершення обробки цей сервер можна зупинити або можна відправити запит для обробки нової порції даних. Хмарні обчислення можна розглядати у двох напрямах: як власник обчислювальних ресурсів та як користувач. У першому випадку, вони орієнтовані на надання інформаційних ресурсів зовнішнім користувачам. У другому – це отримання їх у вигляді послуги зовнішнього постачальника, обробка та передача якої залежить від кількості спожитих ресурсів відповідно до встановленого тарифу [2].

Обидва типи обчислень включають в себе багатопоточність та багатозадачність. Безліч користувачів можуть виконувати різні завдання, маючи доступ. Спільне використання ресурсів серед великого потоку завдань користувачів призводить до зниження витрат на стіорення інфраструктури та максимальну пропускну спроможность каналу. Хмара та грід-обчисленння забезпечують домовленість про рівень обслуговування Service Level Agreement для гарантованої доступності безперебійної роботи, близько 99%. Якщо обслуговування послуг нижче рівня гарантованого обслуговування безперебійної роботи, користувач отримує кредит на сервіс для прийому даних. Таким чином, хмарні та грід-обчислення доповнюють один одного. Грід-обчислення забезпечує об'єднання гетерогенних ресурсів у єдине обчислювальне середовище, забезпечуючи те, на чому базуються хмарні обчислення: більш високий рівень абстракції, надаючи обчислювальні ресурси кінцевим користувачам у вигляді сервісів. Грід та хмари функціонують на різних рівнях організації масштабованих обчислювальних процесів та процесів обробки даних. Тому доцільно інтегрувати ці два підходи для ресурсів мінімізації витрат на обчислення та підвищення ефективності використання обладнання.

Обмін контенту між хмарою та призначеним для користувача терміналом призводить до додаткових витрат, які часто не компенсуються підвищеною швидкодією хмарних серверів. Крім того значно ускладнюють глобальне впровадження такої інфраструктури складність конфігурування, концептуальну складність, необхідність у модифікації існуючого чи створення власного, спеціалізованого програмного забезпечення. Але хмарні обчислення вже на рівні інфраструктури як послуги надають масштабовані віртуальні апаратні ресурси у вигляді віртуальних машин та додаткових двох рівнів: платформи як послуги та програмного забезпечення як послуги. Об'єднання переваг хмарних та грід-обчислень призводить до гібридної інфраструктури накладання структур, прикладом якої є застосування грід на хмарній системі на рівні інфраструктури, або реалізація хмарної платформи у складі інфраструктури грід.

Література. **1.** Риз Дж. Облачные вычисления (Cloud Application Architectures). – СПб.: БХВ-Петербург, 2011. – 288 с. **2.** Antonopoulos, N., Gillam, L. Cloud Computing: Principles, Systems and Applications. Springer, 2010. – 414 p.

Петренко А.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Нова парадигма "Ontology-based Software Engineering": перехід від об'єктно-орієнтованого до процесно-орієнтованого програмування

Розглядаються сучасні дослідження з галузі сервіс-орієнтованих обчислень.

Вступ. Сьогодні у світі набирає темпи новий підхід до проектування архітектури програмних комплексів як динамічної системи взаємодіючих незалежних веб-сервісів, доступних через мережу, що дозволяє розробникам прикладного ПЗ зосередитися, головним чином, на пошуку і комбінуванні сервісів, які відповідають бізнес/технічним характеристикам, і звести до мінімуму розробку нового коду [1–3]. Архітектура ПЗ будується на основі бази знань, яка містить онтології предметних галузей. Моделі сервісів представлені у формі, що дозволяють автоматично перевести їх в виконуваний код через перетворення послідовності моделей шляхом зміни мов їх опису: CIM (обчислювальна незалежна модель) –> PIM (незалежна від платформа модель) –> PSM (модель для конкретної платформи) –> BPEL, UML, Java. Нові сервіси також можуть динамічно створюються з використанням існуючих [4]. В результаті отримуються рішення, які обслуговують кілька доменів і призначені для повторного довгострокового спільного використання багатьма проектантами і організаціями.

Застосування розробленої технології та розподілених репозиторіїв сервісів дозволяє значно знизити собівартість та підвищити якість прикладного програмного забезпечення, що створюється, а також забезпечує шляхи інтегрування додатків в межах сукупності великої кількості інформаційних систем підприємства чи кількох підприємств в межах країни або континенту.

Європейська наукова хмара. Саме на такому сервіс-орієнтованому підході базується найбільший європейський проект вартістю 2 млрд. євро, підготовкою якого займаються зараз всі відомі організації і інфраструктури (EUDAT, CENT, OpenAIRE, GÉANT, EGI, ESFRI, e-IRG, ESA та інші), зі створення ресурсів європейської відкритої науково-дослідницької хмари (**European Open Science Cloud for Research**), що буде обслуговувати 1,7 млн. вчених і 80 млн. професіоналів з різних галузей науки і технологій. За минулий короткий термін – близько п'яти років – новий напрям науки про сервіси, їх менеджмент і інженерію (**Service Science, Management, Engineering, SSME**) набув широкого поширення в провідних університетах світу і Європи. Сьогодні відомі 102-і SSME навчальні програми по всьому світу (88 магістерських та 14 бакалаврських) з підготовки підготовці фахівців нового типу – сервісних спеціалістів, про яких в Україні майже нічого невідомо [5]. В НТУУ "КПІ" з 2016 офіційно відкрита спеціалізація "Системне проектування сервісів" в межах спеціальності 122 "Комп'ютерні науки і інформаційні технології".

Порівняння підходів до створення ПЗ. Зараз настав час переходу на нову парадигму програмування, пов'язану не з об'єктами, а з бізнес-процесами і їх складовою частиною – бізнес-функціями. В табл. 1 наведені основні переваги, що з'являються у цьому випадку.

Таблиця 1. Порівняння підходів до розробки ПЗ

№	Об'єктно-орієнтований підхід	Процесно-орієнтований підхід
1	**ВІДПОВІДАЛЬНІ ВИКОНАВЦІ**	
	Розробники ПЗ, що більшу частину часу працюють з кодом	Архітектори ПЗ, які більшу частину часу працюють з онкології і моделями спільно з експертами в предметній галузі
2	**ЕТАП АНАЛІЗУ ТЗ**	
	Вимоги формуються на природній мові для системних аналітиків, які перетворюють їх в технічні характеристики для того, щоб дати можливість розробникам / дизайне-рам зрозуміти їх.	Аналіз домену виконується постачальни-ками, що дозволяє розробникам додатків зосередитися тільки на пошуку і комбінуванні сервісів, які відповідають бізнес / технічним характеристики.
3	**ЕТАП ПРОЕКТУВАННЯ**	
	Уточнення структури класів, проектування архітектури системи (структури компонентів і модулів), вибір будівельних блоків. Наступне оновлення їх з урахува-нням технічних змін є складним завданням	Вибір або побудова онтології прикладного домену і проектованого додатку. Моделі представлені у формі, пов'язаної зі специфікаціями, які дозволяють розробни-кам системи сервісів автоматично перевести їх в виконуваний код (CIM-> PIM -> PSM-> UML).
4	ЕТАП РЕАЛІЗАЦІЇ	
	Кодування на конкретній мові програмування (Java, C ++, C #). Розробка виконується однією командою, яка створює функції, класи, модулі.	Складні інструменти і мови, такі як MS.NET, BizTalk, WebSphere, Oracle SOA Suite і BPEL використовуються для розробки сервісів і робочих процесів. Нові сервіси можуть динамічно створюються з використанням існуючих.
5	ЕТАП ТЕСТУВАННЯ І ПІДТРИМКИ	
	Тестування зазвичай проводиться в тій же організації на основі вихідного коду і функціональних специфікацій. Тестові сценарії визначаються розробниками / тестувальниками.	Тестування поділиться між постачальником сервісів, брокером і клієнтом, практично без взаємодії між ними. Сервіси можуть бути перевірені до їх реєстрації та використання. Тестові сценарії можуть бути автоматично згенеровані обома брокером і клієнтом.
6	ЗАГАЛЬНА ХАРАКТЕРИСТИКА РІШЕННЯ	
	Проектно-залежні рішення, які в основному не використовуються повторно і мають відносно коротко-строкове ізольоване застосування в обраному домені.	Крос-проектні рішення, які обслуговують кілька доменів і призначені для повторного довгострокового спільного використання багатьма проектантами і установами.

Література. **1.** Петренко А.А. Цели и объекты науки о сервисах/ Петренко А.А. // Системні дослідження і інформаційні технології. – №2, 2015. – С. 75–82. **2.** Петренко О.О. Порівняння типів архітектури систем сервісів // Системні дослідження і інформаційні технології – №4, 2015. – С. 48–62. **3.** Петренко І.А. Автоматизовані методи пошуку і відкриття необхідних сервісів/ Петренко І.А., Петренко О.О. // Вісник Університету "Україна", Серія "Інформатика, обчислювальна техніка та кібернетика", №1(17), 2015, С. 55–64. **4.** Петренко А.А. Семантическое модельно-управляемое моделирование архитектуры системы сервисов на основе доменных онтологий / Петренко А.А, Петренко А.И. // Электронное моделирование, ISSN 0204–3572, 2016. – Том.38. №5. – С. 1–17. **5.** Петренко О.О. Підготовка кадрів для індустрії сервісів / Петренко А.А. // Information Technologies in Education. – №22, 2015. – P. 158–164.

***Романов В.В.*[1], *Ткаченко В.А.*[2]**
[1]*КПИ им. Игоря Сикорского, ИПСА, Киев, Украина;* [2]*ООО НПО "Кристалл", Киев, Украина*

Особенности проектирования преобразователей "Напряжение – Частота"

Данная работа посвящена проектированию одного из видов преобразователей, а именно преобразователям типа "Напряжение – Частота" (ПНЧ). Рассмотрены схемотехнические, конструктивно-технологические решения, рассчитаны основные электрические параметры отдельных блоков микросхемы преобразователя с единой целью – получение минимальных значений нелинейности преобразования входной информации в виде напряжения U_i или тока I_i в частоту F_i.

Анализ целевой функции – преобразование исходной информации в виде напряжения U_i или тока I_i в частоту $U_v = F_i = f[U_i(l_i)] = S_{pr} * U_i(I_i)$ при $\frac{dF_i}{dT^oC} = 0$ и $\frac{dF_i}{U_{cc}} = 0$, U_{cc} – напряжение питания преобразователя, $S_{pr} = \frac{F_i}{I_i}$ крутизна преобразования напряжения (тока) в частоту позволяет провести пошаговый синтез преобразователя и представить один из возможных вариантов структурной схемы ИМС ПНЧ (рис. 1). При этом частота преобразования после ряда промежуточных преобразований:

$$F_i = \frac{2 * U_i}{(3 * R_T) * 4 * C_T * U_R} = \frac{U_i}{6 * R_T * C_T * U_R} = \omega_0 * \frac{1.0472 * U_i}{U_r}, \quad (1)$$

где $\omega_0 = \frac{1}{2*\pi*R_T*C_T}$; $U_r = \varphi_E = \frac{1}{2} * (\varphi_c + \varphi_v)$, где U_r – напряжение полуволны меандра; φ_C – дно зоны проводимости; φ_V – вершина валентной зоны; φ_E – электростатический потенциал проводника [1]. Из этого выражения следует, что при $U_i = 0$ частота преобразования $F_i = 0$. Числитель U_i ,определяющий частоту преобразования, никогда не может быть равным нулю, т.к. не учтены параметры операционного усилителя, шумы, напряжение смещения, входные токи и их разность, коэффициент усиления и т.д. Величины этих параметров вносят ошибку в общую величину управляющего напряжения или тока, что сказывается на нелинейность частоты преобразования при малых значениях напряжения преобразования U_i. Схемотехнически, конструкциями элементной базы, его технологической настройкой стремятся уменьшить влияние параметров входного операционного усилителя на точность частоты преобразования F_i.

В нормальных условиях для кремния $\varphi_E = 833$ мв. Величина U_r должна изменяться по определенному закону, чтобы обеспечить минимальную нелинейность преобразования напряжения U_i или тока I_i в частоту F_i. Это требование возникает из-за изменения ширины запрещенной зоны полупроводника от температуры окружающей среды и в соответствии с математической моделью прямо смещенного p-n перехода база-эмиттер U_{be}=φ*$ln(\frac{I_K}{I_O}+1)$ транзисторов схемы ИМС приводит к изменению электрического режима элементной базы ИМС и коэффициента h_{21e}. Тепловой ток I_O p-n переходов является функцией многих переменных. Для уровня технологии ИМС 1101ПН02 он лежит в пределах $[2 \div 4\sqrt{2}\pi]$ * 10^{-15} А. Обычно при расчетах используется величина $I_O = 2 * 10^{-15}$ А.

Назначение приведенных на рис. 1 блоков следующее:

1. **Входной усилитель** представляет собой операционный усилитель (ОУ), который преобразует входное напряжение U_i в управляющий ток I_i с помощью внешнего прецизионного термостабильного резистора R_T.
2. **Прецизионный мультивибратор М** управляется входным током I_i и выдает пропорционально I_i импульсное напряжение (меандр) с частотой F_i.
3. **Драйвер D** усиливает выходное напряжение мультивибратора М и подает его на выходной ключ К (транзистор Т0) с открытыми коллектором и эмиттером.
4. **Ключ К** – выдает в нагрузку необходимый ток I_0.
5. **Источник опорного напряжения** – U_{REF} построенный по схеме BANDGAP, должен выдавать выходное напряжение U_{REF}, изменяющееся по определенному закону (адаптивные системы) в соответствии с изменениями внешних условий, в частности, температуры

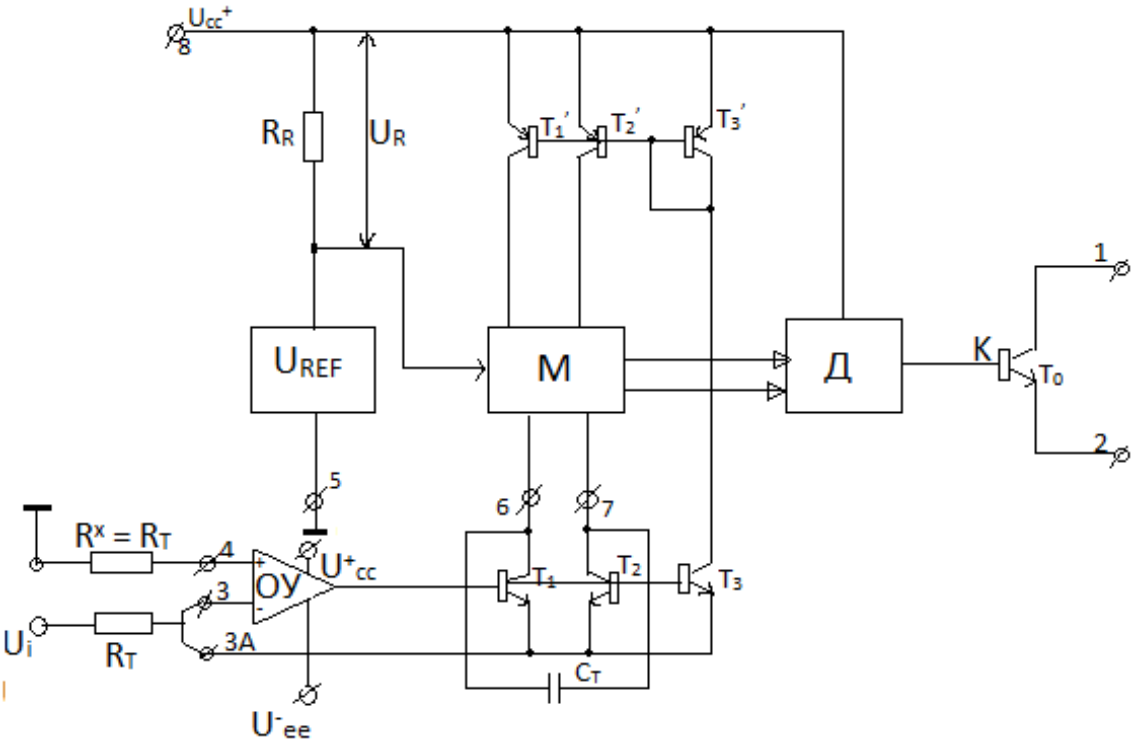


Рис. 1. Структурная схема ИМС ПНЧ

в диапазоне изменений управляющих входных токов (напряжений).

Ниже приводятся сравнительные данные, полученные в результате измерений опытных образцов ИМС 1101ПН02 и одной из лучших зарубежных микросхем подобного класса – АД654 (Voltage to Frequency Converters – VFC) [2].

В табл. 1, 2 приведены экспериментальные данные зависимостей частоты преобразования F_i от управляющего тока I_i и поведения F_i при изменении температуры окружающей среды T_A.

Таблица 1

$I_i(\mu A)/F_i$(КГц)	0.01	0.1	0.2	1.0	10.0	100.0	200.0	500.0	1000
AD654	-	0.0408	-	0.462	4.678	46.78	93.57	233.9	467.8
1101ПН02	0.0043	0.0502	0.095	0.502	5.067	50.74	101.5	253.4	505.7

Таблица 2

№	$T_A{}^0$C/F_i	–40	–30	–20	–10	25	50	70	$E_R/\Delta T_A$
1	AD654, $F_i = F_o = 250$ кГц	239.1	239.2	239.9	240.1	240.6	236.8	234.0	0.02%
2		240.3	239.9	239.7	239.8	238.8	236.3	235.9	0.016%
1	1101ПН02, $F_i = F_o = 250$кГц	244.6	244.0	244.3	244.7	244.5	239.7	239.0	0.02%
2		255.1	253.7	254.7	255.5	256.0	249.5	249.8	0.022%

Примечание. $\frac{E_r}{\Delta T_A}$ определяется по формуле: $\frac{Er}{T_A} = \frac{(|F_{i.max} - F_{i.min}| * 100)}{F_0 * |T_{A.max} - T_{A.min}|}$; где F_0 – калибровочная частота; Er – нелинейность преобразования

Литература. **1.** Научно-технический отчёт по ОКР ООО НПО “Кристалл” “Преобразователь - 5”, Киев-Москва, 2013. **2.** Date Sheet IC AD654, со “Analog Device”, USA.

Сергеєв Є.І., Романов В.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Вибір складової фреймворку Xamarin при розробці крос-платформних додатків

Для написання програм під кожну з операційних систем (ОС) потрібно витрачати багато ресурсів, а також часу. Оскільки кожна ОС має різні API інтерфейси, специфічну логіку роботи, нативну мову програмування. Для того, щоб створити програмний продукт (ПП), який буде доступний для різних ОС компанії-розробнику потрібно тримати штат програмістів, спеціалізованих для різних ОС, наприклад, Java команду для написання Android програм, C# – команду для написання програм під Windows Phone і т.д. Тому є фреймворки, які дозволяють позбутися від такої залежності і писати код під різні ОС, використовуючи одну мову програмування. Звичайно, у такого підходу є свої недоліки. По-перше, програми, які написані на нативній мові програмування для даної ОС, працюють швидше. По-друге, потрібно знайти спеціалістів, які знають усі платформи, які підтримує даний фреймворк, звісно, спеціаліст, який спеціалізується на одній конкретній платформі, буде знати саме цю платформу набагато краще. Проте, при такому підході можна позначити й деякі переваги. Першою такою перевагою, звісно, є те, що розробникам не потрібно міняти бізнес-логіку ПП, оскільки код пишеться з використанням однієї мови програмування, звісно, не весь код буде спільним, проте за статистикою (яка надана спеціалістами з Xamarin) такий код займає в середньому 20% від усього програмного коду. Другою перевагою такого підходу є те, що час на створення ПП значно скорочується, оскільки потрібно писати менше коду, скорочується й кількість конфліктів між командами.

На сьогоднішній день існує багато фреймворків, які орієнтуються на різні ОС і дозволяють розробляти кросс-платформні додатки, одним з них є Xamarin – унікальний фреймворк, який дозволяє будувати такі додатки з використанням єдиної мови програмування – C#. З використанням данного фреймворка компілюється так званий "нативний" код для операційних систем Android, Windows, iOS. Для підтримки цих трьох операційних систем, фреймворк розділяється на відповідно **три складові – Xamarin.Android, Xamarin.iOS та Xamarin.Forms**.

Звичайно Xamarin – не єдиний фреймворк, який орієнтується на декілька ОС, є багато аналогічних фреймворків, наприклад Ionic, OnsenUI, Sencha Touch, Robo VM, NativeScript зі своїми перевагами та недоліками. Ми детально зупинемось на фреймворкі Xamarin, як перспективному фреймворку, що не уступає своїм конкурентам, тим паче, що Xamarin тепер є повністю безкоштовним фреймворком (анонсовано на конференції Build 2016) і підтримується Microsoft.

Основними продуктами, що було створено за цей період часу є: Xamarin Platform, Xamarin.Android, Xamarin.Forms, Xamarin Test Cloud, Xamarin for Visual Studio, Xamarin.Mac і т.д.

Xamarin.Android додатки виконуються в середовищі виконання (CB) Mono. Mono працює пліч-о-пліч з віртуальною машиною Android Runtime (ART). Xamarin.Android надає розробнику такі можливості: створення спеціальних провайдерів контенту та резолверів контенту, що надають можливість отримувати та надавати дані, а також шукати дані з інших додатків у системі; створення сервісів, що мають свій життєвий цикл та дозволяють виконання деяких складних задач в іншому потоці; використання фрагментів замість декількох activity для створення швидшого додатку, оскільки не потрібно довантажувати додаткові activity у відповідь на дії користувача; Xamarin.Android підтримує усі API рівня ОС Android, отже розробникам завжди доступні усі можливості даної платформи; як і в звичайному додатку під Android, Xamarin.Android надає змогу використання ресурсів – зображень, елементів інтерфейсу, аудіо та відео, файлову систему та інше [1].

Отже, Xamarin.Android наділена багатим функціоналом для розробки додатків під ОС Android, що не гірші від Java аналогів, дозволяє використовувати усе, що надає платформа, а

додатки по швидкодії не програють аналогам.

Технологія Xamarin.iOS значно відрізняється від Xamarin.Android. ОС iOS не дозволяє використовувати середовища виконання. Для iOS використовуються спеціальні прив'язки, для нативних компонентів та можливостей даної ОС. Однак, неможливість використання середовища виконання – не проблема, оскільки навідміну від використання JIT (Just-In-Time) компіляції, додаток компілюється одразу в байт код. Xamarin.Android та Xamarin.iOS є складовими Xamarin Native, що використовується для написання чистого нативного додатку з використанням C#.

Xamarin.Forms – це зовсім несхожа технологія на дві попередні. Різниця в тому, що Xamarin.Forms дозволяє писати нативні додатки з більшою кількістю спільного коду, надає розробнику абстракції графічного інтерфейсу для кожної з платформ з використанням мови XAML (Extensible application markup language). Дана технологія використовує нативні компоненти під час роботи додатку, тобто код графічного інтерфейсу буде перетворено на нативний код окремої конкретної платформи. Код Xamarin.Forms має можливість взаємодіяти з операційною системою, використовувати її програмний інтерфейс. Даний підхід є набагато простішим, на рис. 1 зроблено порівняння цих підходів [2].

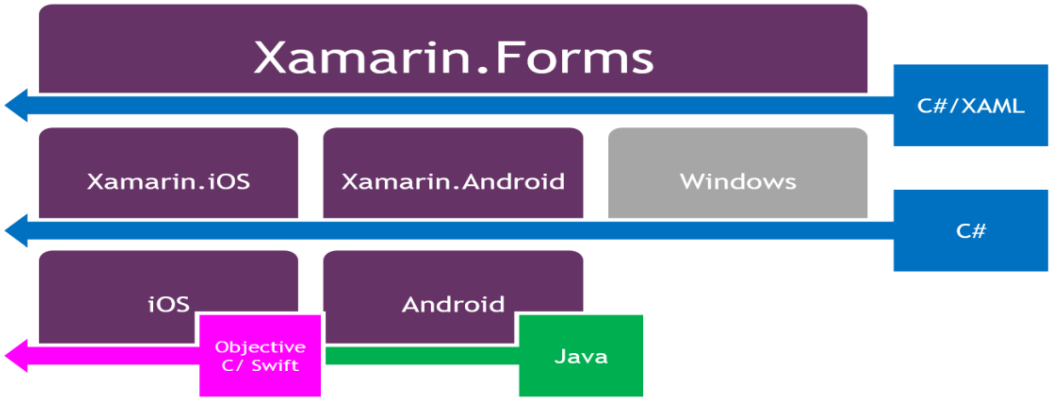


Рис. 1. Схема покриття Xamarin.Forms та Xamarin Native

Як видно з рисунка, Xamarin.Forms повністю "покриває" усі три платформи та дозволяє розробнику зосередитися на бізнес логіці додатку та його структурі і економити час розробки. Такий додаток буде також кросплатформним, однак він матиме багато обмежень з точки зору функціональної частини додатку, оскільки неможливо об'єднати усі можливості ОС і використовувати як одну спільну, адже вони реалізовані по-різному та й деякі з них не наявні у інших ОС.

Можна зробити висновки, що точних рекомендацій для розробника, який з даних підходів використовувати – Xamarin Native (Xamarin.Android та Xamarin.iOS) чи Xamarin.Forms немає, усе залежить від бажання та можливостей розробника, масштабів проекту. Є лише декілька критеріїв селекції одного з цих підходів – час розробки та можливості додатку. Якщо програмний продукт повинен мати широкий функціонал та використовувати усі можливості програмного інтерфейсу (API) платформи, то гарним вибором, звичайно буде Native підхід, якщо ж основним критерієм є час розробки, то єдиним логічним вибором є Xamarin.Forms.

Література. **1.** Jonathan Peppers. Xamarin Cross-platform Application Development Second Edition / Peppers J. – Birmingham: Packt, 2015. – 459 с. **2.** Nilanchala Panigrahy. Xamarin Mobile Application Development for Android Second Edition / Panigrahy N. – Birmingham: Packt, 2015. – 296 с.

Шевченко Є.Т., Ліщук К.І.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Балансування навантаження між кластером серверів з урахуванням типу задач

Наведено опис оптимізованого алгоритму балансування навантаження між серверами з урахуванням типу задачі для визначення члену кластеру, який являється найменш завантаженим в момент виникнення такої задачі.

З кожним роком кількість нових інтернет-ресурсів зростає. Лише за 2016 рік було зафіксовано приріст користувачів мережі близько семи відсотків. Але зростає не тільки кількість нових ресурсів, а й навантаження на ресурси, які вже існують. Для підвищення надійності та доступності інтернет-ресурсів використовуються методи горизонтального масштабування, а саме збільшення кількості серверів для обробки даних [1].

Для розподілення навантаження між кластерами серверів застосовується проміжне програмне забезпечення – балансувальники (Load Balancers). В більшості випадків такі балансувальники працюють методом простого перебору всіх доступних серверів, тобто при надходженні нової задачі балансувальник нічого не враховують, окрім черговості серверів.

Опис проблеми. Уявімо ситуацію, коли у розпорядженні балансувальника є 5 серверів, які виконують задачі, що потребують значних ресурсів. Задачі можуть займати від кількох секунд до кількох годин роботи серверу в режимі максимального завантаження. Метод простого перебору серверів буде не оптимальним в разі, коли перша задача була максимально ресурсоємна, а всі наступні були швидкі по часу виконання. В такому випадку шоста задача буде довгий час очікувати вирішення першої, хоча задачі з другої по п'яту до цього моменту були вже вирішені.

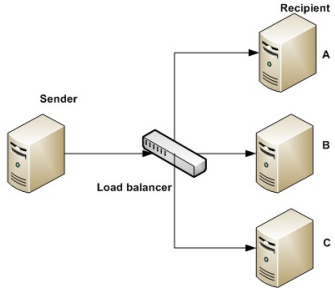


Рис. 1. Загальна схема роботи простого балансувальника

Відомі рішення. Одним з найпопулярніших рішень цієї проблеми є введення максимального часу відповіді серверу про прийняття задачі на вирішення (timeout). Тобто це час, після якого балансувальник спробує передати задачу наступному серверу [2]. Цей підхід тільки частково вирішує проблему, так як:

1. Вводиться певний timeout, під час якого задача буде простоювати.
2. Додаткове навантаження на оперативну пам'ять балансувальника.

Пропоноване рішення. Для більш ефективного розподілу навантаження між серверами пропонується розробити програмне забезпечення балансувальника, яке не буде зберігати в пам'яті нічого, що стосується самої задачі, а лише буде підтримувати певний протокол правил розподілення задач по кластеру. Засобом передачі метаданих для визначення найбільш ефективного сервера вирішення задачі пропонується використовувати заголовки (headers) http-запиту. Перевага способу полягає в тому, що він є швидким і не потребує додаткових дій з мережею для досягнення кінцевого результату роботи балансувальника. Зменшення кількості запитів призведе до покращення важливого параметру такого типу систем – швидкості відповіді. Зі сторони балансувальника (після відповіді) буде застосовано набір математичних моделей. Результати моделей можна буде використати при наступному розподілі задач подібного типу. Такий підхід можливий тільки для балансування на прикладному рівні, а це означає, що програмне забезпечення буде окремо налаштовуватися для кожного набору задач.

Література. **1.** Бунин О. (2012). Разработка высоконагруженных систем. По материалам конференции HighLoad++. Издательство Олега Бунина. http://coollib.net/b/280449. **2.** Оллспоу Дж. (2011). Искусство планирования мощностей. Питер. bukvy.net/books/web/117369-iskusstvo-planirovaniya-moschnostey.html.

This page was left blank intentionally

Plenary talks

System analysis of complex systems of various nature

Intelligent systems for decision-making

High Performance Computing and Microsystems Engineering

Progressive information technologies

4 Progressive information technologies

Section 4

Progressive information technologies

1. Support (mathematical, algorithmic, linguistic, informational-organizational, technical, software) of the control systems, information processing, and development technologies.
2. E-commerce.
3. Information security and guarding.
4. Data and knowledge bases as the environment of information support for control and design.

Секция 4

Прогрессивные информационные технологии

1. Обеспечение систем управления (математическое, алгоритмическое, лингвистическое, информационно-организационное, техническое, программное), обработка информации и технологии их создания.
2. Электронная коммерция.
3. Информационная безопасность и защита информации.
4. Базы данных и знаний как среда информационной поддержки управления и проектирования.

Секція 4

Прогресивні інформаційні технології

1. Забезпечення систем управління (математичне, алгоритмічне, лінгвістичне, інформаційно-організаційне, технічне, програмне), обробка інформації та технології їх створення.
2. Електронна комерція.
3. Інформаційна безпека та захист інформації.
4. Бази даних і знань як середовище інформаційної підтримки управління та проектування.

Basyuk T.M.
National University "Lviv polytechnic", Lviv, Ukraine

Approach towards generating a software system of website popularizing

Together with the advent and development of search engines in the mid-1990s search engine optimization has appeared. At the beginning of creation, search engines provided the highest priority of content and internal factors which could be easily manipulate by owners. This led to the fact that set of pages appeared in search results that did not contain the relevant information request, which drastically reduced the quality of the search engines. Developing and patenting by the owners of search engine algorithms set was the way out of the situation [1], which radically changed the approach to the SERPs and web sites evaluating. This became the beginning of search engine optimization (SEO), which is the process of adjusting html code, text content (content), site structure, control of external factors to meet the requirements of search engines. The main goal was raising of site position in search results for defined requests by users.

Based on conducted analysis it turned out [2,3] that almost all search engines take into account such parameters of website in the calculation of its relevance as: the frequency of keywords (available algorithms allow conducting semantic text analysis, the purpose of search of spam page) and indexing of website citation (determines the number of external resources relevant topics that link to this site). However, due to closure of algorithms operation of search engines, known methods of evaluating priority is issuing local efforts of individual researchers and determine the impact of individual factors of ranking that can not pretend for the universality of the results or the ranking mechanisms. The other side is the process of monitoring the web site that currently is carried out by the majority of SEO in manual mode, which does not assess the set of parameters in real time. Using of automated tools for promoting allows saving considerable time for analysis of the resource, and can get rid of the set of errors. However, existing software (CS Yazzle, Social Media Marketing, CY-PR, SEOGadget) reflects not only general information without recommendations promoting websites in search results [4]. Given that the creation of an approach towards designing a software system of recommendations provision to the popularization of sites in the global network is the urgent task. Generalization of influence factors of Internet resources ranking and methodological basis formation for system design of web sites popularization. The study will provide basis for creating recommendation system that will determine the impact of major optimization factors and provide recommendations for entry of website in top search engine ranking positions. In order to achieve the goal the following tasks should be solved: the analysis of the process of optimizing web sites and indexes of popularity resource; identify the parameters that most influence the position of the web site and take them into account in the projected system; carry out a systematic analysis of the subject area that will create a basis for further system design according to the methodology of object-oriented programming. The approach provides the means of formation of recommendations for improving the visibility of resource mechanisms of search engines and thus create a basis for its entry in the top SERPs. The results of the work done in its entirety solve current scientific and practical task of creating recommendation system to promote web sites.

The conducted analysis showed that the optimization of any website lies in work with both internal factors (reduction of text and page layout in accordance with the selected requests, the improvement of the quality, optimization of structure, navigation and internal links site) and external (exchange by links, registration in catalogs and other measures to improve incentives and increase the number of links in the resource). However, they neglect facilitate the emergence of a plurality of conflicts that will prevent effective popularization of resource.

Whereas, it is proposed to apply the following as internal factors of assessment: placing of keywords in the tag title (it is displayed in search results as page title) the presence of keywords in the URL of the website (positively affects the search results, provided that the request is present not only in the address, but in text filling) achievement of optimum keyword density (3–7%); the use of tags h1, h2; prescription of keywords in the meta tags that are indexed by search engines; determination of the uniqueness of content; information update rate (affects both the speed of indexing and ranking level); term of resource existence and domain level . Based on the conducted

studies, for the design of future systems, system analysis of subject area with the construction of the set of UML diagrams, two main of which is shown in Fig. 1 and Fig. 2 was carried out.

Fig. 1 shows the case diagram, displaying a plurality of entities (actors) that interact with the system through the use of options. Actor on it, is a specialist in optimizing websites and the following use options: Check Right domain – verifies for the correct (according to the requirements of search engines) writing domain; Check duplicate contents – provides control for content uniqueness, and is expanded by test of URL structure and meta tags search; Check for right URL structure – verifies for the correct choice of resource structure.

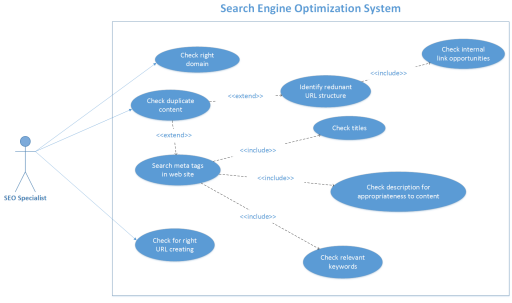


Figure 1. Use case diagram

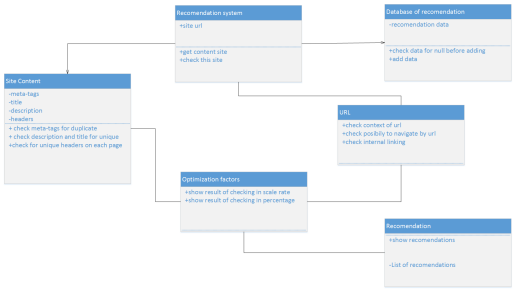


Figure 2. Class diagram

Fig. 2 shows the diagram of classes of designed system that reflects the static structure of the system in terms of object-oriented programming. This diagram contains the following classes: Recommendation system – contains, as an attribute of a link to the site and the procedures for obtaining content for its review; Site content – contains attributes such as meta tags, headings and their description, that is the main internal factors of ranging that are tested according to frequency of meta tags, headlines, etc. uniqueness; URL – This class contains methods for checking the content of URL, for correcting spelling, navigation capabilities, etc; Optimization factors – class of optimization factors includes two methods: Display of the site conformity according to the requirements of the SEA in percentage and due to 5-point scale; Recommendations – class contains attributes for the list of recommendations concerning the optimization; Recommendation database – class includes methods for checking the data on empty meanings and data addition. Applying UML design to this problem provides a set of tools to further support of the development of software that lies primarily in optimizing for use at the development of complex software systems and tools provision both for the maximum acceleration and quality improvement of designed recommendation system.

Conclusion. The analysis of the subject area, due to which the main factors were managed to be determined influencing the as resource ranking and its popularity as well was carried out as a result of conducted study. The analysis of the impact of various parameters on the position of the web site, which served as a methodological basis for the construction of conceptual models of the process of making recommendations in the popularization of web sites was carried out.

Further research will be focused on the process of developing decision making support systems concerning the promotion of a web site that will provide required device for the improvement of resource ranking in the global network.

References. **1.** Grant Kennedy. 2016. Seo: Marketing Strategies to Dominate the First Page (SEO, Social Media Marketing). CreateSpace Independent Publishing Platform, 122 p. **2.** J. Jantsch and P. Singleton. 2016. SEO for Growth: The Ultimate Guide for Marketers, Web Designers & Entrepreneurs. Wiley, 238 p. **3.** Taras Basyuk. 2015. The main reasons of attendance falling of internet resource. International Scientific and Technical Conference «Computer science and information technologies CSIT-2015», 91–93. IEEE (DOI:10.1109/STC-CSIT.2015.7325440). **4.** A. Clarke. 2016. SEO 2017 Learn Search Engine Optimization With Smart Internet Marketing Strateg: Learn SEO with smart internet marketing strategies. CreateSpace Independent Publishing Platform, 194 p.

Chereda H., Nikolaiev S.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Comparison of resampling techniques and its effect on real time filtration of non-uniformly sampled bio-signal time series

Introduction. Sequences of events occur in multiple applications across different domains including, but not limited to, bio-signals in healthcare, complex telecommunication networks, user interface studies, video processing etc. These sequences can be presented as non uniformly sampled time series. To obtain any useful insights from the signals, these time series have to be firstly filtered out from noise. But frequency filtering operation is defined only on uniformly sampled data. To provide the analysis of such signals, the signal have to be resampled with uniform sampling rate in order to apply well-known digital signal processing techniques. There are variety of resampling techniques presented in [1] and [2]. The most well known ones are linear interpolation, cubic spline interpolation and optimal sinc interpolation. Thus, the main question to be answered is how different resampling techniques influence on the filtered signal quality.

Research description. The goal of the research is to compare influence of the different uniform resampling techniques on the quality of heart beat events extraction from real time video through digital signal processing. As an example in this paper we will use oxy-hemoglobin level signals from the remote plethysmographic face imaging module using video-camera and ambient lightning. These bio-signals have extreme noise level and their sampling rate is equal to the video-camera frame rate which is on average approximately 15–20 Hz. In real time video recording, the moment of video frames capturing is not uniformly distributed in time. The frames appearance in time usually has Gaussian distribution and it significantly varies depending on camera vendor and video recording conditions. After the spatial preprocessing filters are applied to the series of frames then only one dimensional time series of hemoglobin levels is left. The main goal is to obtain person's heart-rate and RR-intervals using these oxy-hemoglobin level sequence. As it was stated above, before the temporal filtration procedure the uniform resampling should be applied. The chosen uniform sampling rate for the resampled signal is 30 Hz that is higher than the mean video-camera frame-rate. To perform resampling we should choose the best interpolation approach that does not change the spectral properties of the original noisy signal.

Resampling stage. The resampling technique was applied in the real time to non-uniformly sampled signals. This technique is based on different interpolation approaches: linear interpolation, cubic spline interpolation and sinc optimal interpolation [2]. As a raw signal the non-uniformly sampled oxy-hemoglobin level time sequence is used in this research (shown on figure below).

In case of the linear interpolation approach we may lose some precision in data peaks locations. On the other hand, the linear interpolation is computationally fast and can be easily implemented for the real time processing. Also it provides smoothing effect for the time series.

The cubic spline interpolation is significantly more time consuming and the resulting resampled signal becomes more noisy compared to the linear interpolation.

As for the sinc optimal interpolation, it is the most time consuming approach that is not convenient for the real time processing. It was discovered that sinc interpolation should be used only for downsampling purposes and it is not suitable for our application because the downsampling approach decreases the precision of heartbeat event time extraction. Moreover attempts to use sinc filtration lead to noise level increase in case of resampling with the sampling rate equal to the raw signal average sampling rate.

On the figure below linearly resampled filtered final signal, sinc interpolated, cubic spline interpolated, linear interpolated and raw non-uniform original signals' plots are shown (from top to bottom). The OX axis represent time in seconds, the OY axis represent signal amplitude (in relative units).

Filtration stage. Real time filtration can be performed by IIR filters or FIR filters. There are several requirements applied to the filtration stage. Such filtration requires filter's linear phase

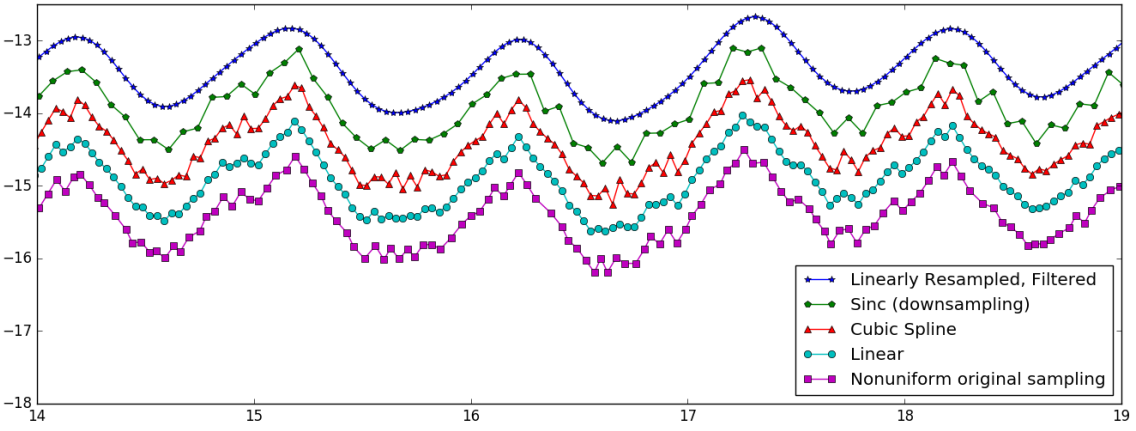


Figure 1. Comparison of raw non uniform signal and its real time resampled versions with different types of interpolations

response and relatively thin bandpass in frequency domain – from 0.6 to 2 Hz, compared with the provided uniform sampling rate sr, which is equal to 30 Hz. The analog Bessel filter and FIR filters have linear phase response. The digital version of Bessel filter created through the bilinear transform has approximately linear phase response for the frequencies below $sr/4$. But the smooth amplitude response of the bandpass Bessel filter does not provide required passband's borders sharpness. The Butterworth digital filter has as flat a frequency response as possible in the passband, but its non-linear phase response generates distortions in filtered signal. Therefore, it was decided to use FIR filters that meet demands.

Quality measurement stage. The criterion used for the filtration and resampling quality evaluation is a measure of the similarity between series of contactless extracted heartbeat events using video camera and heartbeat events directly measured from the person's Holter recordings. The criterion is presented in [3].

Conclusion. All resampling techniques were compared with given criteria and it was shown that the final quality achieved is almost the same.

The criterion shows average quality value of 95% computed on 6 experiments including 3 people where cubic spline interpolation performance is 94.97% and linear interpolation performance is 95.23%. Sinc interpolation showed unstable results depending on resampling rate. Besides it is computationally expensive and highly inefficient in real-time applications. It was discovered that sinc interpolation is more suitable for downsampling which decreases heart event peaks time appearance precision.

So we have chosen to use the linear interpolation because of its computational speed and slightly higher quality metric value.

The obtained results are in accordance with the theory which states that high frequency uniform resampling does not influence on low frequencies of noisy signal or its impact is insignificant.

References. **1.** Frida Gunnarsson, Fredrik Gustafsson, Frequency Analysis Using Non-Uniform Sampling With Application To Active Queue Management, 2004, URL: citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.1582. **2.** Francis P. Dawson, Linus Klaffke, Variable-Sample-Rate Delayless Frequency-Adaptive Digital Filter for Synchronized Signal Acquisition and Sampling, in IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, vol.46, NO. 5, OCT. 1999, pp. 889–897. **3.** Sergii Nikolaiev, Metric and algorithm for similarity between two temporal event sequences calculation, in System research and information technologies, NO. 1, 2017, to be published.

Didenko D.G.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Fuzzy recovery database with implementation in Java

Introduction. OLTP systems are replenished daily with new data by using a large number of operators and end-users. This handmade containing 5% error in the input forms, which then during the next import-export fall in OLTP systems [1]. For error correction can be applied various methods: clustering, search for a guide, genetic algorithms and neural networks [1, 2]. Nevertheless, these approaches are very labor-intensive.

Formulation of the problem. The problem of qualitative data analysis is important, especially for clusters and cloud databases are very large. In this case, you must recreate the dictionary with T columns "ID" and "NAME", which does not contain duplicate "NAME" and "like duplicate". This problem was seen and solved earlier in [3] using the formalism duplicate similarity, which is determined by the distance between tuples so-called "editorial distance" [4] but different implementation of the algorithm.

Implementation. After the incoming import file as XML-file, in sequential version follow these steps:

1. Create a dictionary table T ("ID", "NAME") with primary key "ID", and an alternate key until "AK_NAME" is created, because it will prevent further make deliberate duplication.
2. Import incoming XML-file using JPA [5].
3. Ukrainianization in dictionary – replace all occurrences of English letters in Ukrainian.
4. Restore data register if all data in the column only capitalized, then replace them at appropriate using neural network.
5. Remove from the column repeats "NAME" with automatic restoration of values in the dependent foreign key tables. This is most likely done with the cursor inside a nested cursor.
6. Finding the distance between lines allows you to find similar value in the column "NAME", and remove them as well as in step five.
7. Create a unique (alternate) key "AK_NAME" onto the field "NAME".

In parallel execution using Java language created streams that run in parallel, and then combine the results at the end [5].

Conclusion. The algorithm used when working with large amounts of data for quality improvement database and deleting unnecessary duplicates. This algorithm has sufficient performance for large data sets. You can also add language support (English, German and Russian), each of which is necessary to build your neural networks. Further development of MapReduce architecture is the implementation of a parallel approach to distributed systems using Hadoop framework with Java-implementation. This is possible because the algorithm is executed over 1–7 replicated copy of the data from the database.

References. **1.** Барсегян А.А., Куприянов М.С., Холод И.И., Тесс М.Д., Елизаров С.И. Анализ данных и процессов: учеб. пособие. – 3-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2009. – 512 с.: ил. **2.** Пасічник В.В. та ін. Глобальні інформаційні системи та технології: моделі ефективного аналізу, опрацювання та захисту даних. Монографія / В.В. Пасічник, П.І. Жежнич, Р.Б. Кравець, А.М. Пелещишин, Д.О. Тарасов. Львів: Видавництво Львівської політехніки, 2006. 348 с. **3.** Didenko D.G. Fuzzy recovery dictionary by the id and the name into database. //«Системний аналіз і інформаційні технології (CAIT-2016)», м. Київ. – 2016. – С. 303. **4.** Райкова О.А. Нестрогое сопоставление записей реляционных баз данных с использованием редакционного расстояния между кортежами и ключевого набора атрибутов [Текст] / О.А. Райкова // Технические науки: проблемы и перспективы: материалы междунар. науч. конф. (г. Санкт-Петербург, март 2011 г.). – СПб.: Реноме, 2011. – С. 137–140. **5.** Schildt Herbert. Java 8. Ninth Edition. – New York. McGraw Hill Professional, 2014. – 1376 p.

***Gorodko N.A.*[1], *Boyarinova Yu.E.*[1,2]**
[1]*Institute for Information Recording of the NAS of Ukraine, Kyiv, Ukraine;* [2]*Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine*

Using hyper complex numbers for modeling the quality management system of information services

In connection with the steady development of information technologies in various fields of activity all the more urgent becomes the question about the quality of the information needed to solve a particular problem or achieve a certain goal.

Despite the fact that there are plenty of quality management systems, unfortunately, in many cases, developers are faced with the choice of the system, which meets all requirements and is able to satisfy established or projected needs of users.

The main objective of the simulation information services management systems (ISMS) is defined as the search for the best or the best (suboptimal) results of process performance management. To do this, you must define a set of components of a process, procedure and the rules for their implementation, monitoring and control mechanisms within the process parameters characterizing the performance of procedures and process in general, as well as the results of process performance.

Thus, each sequence of the variables in the expression that characterizes the process of ISMS can be described by hypercomplex numbers.

The transition from graphic obtained by one of the computer-aided design systems, hypercomplex to provide information about ISMS processes, carried out in accordance with the following rules.

- Basis of hypercomplex number system is formed of the points scheme and the required number of ways. Moreover, the path are chosen so that among them was not linearly dependent. It follows that the maximum path length can not exceed $n-1$, where n – dimension hypercomplex number system.
- The point of the graph-scheme – is a way of zero length.
- The law of interaction of the basic elements is reduced to multiplication of ways, ie, the product is the total path ways, if the end of the first path coincides with the beginning of the second track. Otherwise, the product is zero.
- Each element hypercomplex number system is a linear combination of basic elements that must include all points of the process flow-chart.

It should be noted that the effective use of HNS in modeling ISMS process should be applied algebra such that can be arithmetized, that is, they must contain a single element. Analysis of the results of the research showed that this property is possessed by the associative algebra over the field of real numbers [1,2]. When their construction method is usually used ways algebra on graphs. With it you can build a finite algebra over the field of real numbers by the original construction of the circuit with the subsequent generation of algebra multiplication table, which corresponds to the transition scheme above.

As a reporting system for modeling ISMS, it proposed to use hypercomplex number systems. This approach will allow for the formal description of the parameters, constraints, and the connections between them, taking into account all available information, as well as the well-known laws, etc., that describe the processes of information services.

References. **1.** Kalinovskiy Ya.A. Isomorphic higher dimensional hypercomplex number systems and their applications / Ya.A. Kalinovskiy, Yu.E. Boyarinova – K.: NAS of Ukraine, IIR, 2012. – 128 p. **2.** Sinkov M.V. The finite hypercomplex number systems. Basic theory. Applications / M.V. Sinkov, Ya.A. Kalinovskiy, Yu.E.Boyarinova– K.: NAS of Ukraine, IIR, 2010. – 389 p.

***Kalinovsky Ya.A.*[1] *Boyarinova Yu.E.*[1,2] *Hitsko Ya.V.*[2] *Sukalo A.S.*[1]**
[1]*Institute for Information Recording of the NAS of Ukraine, Kyiv, Ukraine;* [2]*Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine*

Package hypercomplex symbolic computations in Maple environment

Hypercomplex number systems (HNS) have been widely used in various fields of science and technology, namely mathematical modeling of various processes and phenomena [1, 2]. However, operating with hypercomplex numbers raises some difficulties with their multidimensionality. For example, to multiply two quaternions with numerical coefficients, which are four-dimensional hypercomplex numbers, you need to 16 real multiplications and 12 additions. But hypercomplex numbers of coefficients may be not only the real numbers and algebraic expressions with polynomials, various functions and hypercomplex numbers, including a character variables and coefficients. Therefore, manipulation of such objects requires the use of tools and computer algebra systems.

Many of the currently used computer algebra systems (or symbolic computations) are the most common Mathematica, Matcad, Matlab, Derive, Maple and others [3]. All systems have means for symbolic processing of complex numbers, some systems can be operated with quaternion variables, expressions and functions and in Mathematica have a specialized package for working with Clifford algebras. However, this package is designed primarily for applications in theoretical physics and mathematics.

The authors have created a package hypercomplex symbolic computations in Maple environment [4]. This package consists of procedures written in Maple, which are incorporated into the library. The library can be installed on any computer with Windows operating system and the Maple computing is not lower than version 5. The library is called from a Maple session by joining the program. Package hypercomplex computing open to updating and editing by any user.

Structurally package hyper computing consists of the following subsystems:

- perform algebraic operations in the HNS;
- manipulation of HNS and Cayley tables;
- definition of algebraic expressions hypercomplex performance;
- storage of frequently used expressions;
- implementation of modular operations hypercomplex expressions;
- visualization and service.

Hypercomplex information in the package appears in two forms. For a visual representation of the hyper expressions used in the conventional form of a linear combination of the basis elements. For the internal representation used by the presentation of a list of factors. This can greatly simplify the structure of the hypercomplex data processing algorithms.

Below is an example of calling and connecting hyper computing package that is stored on the disk in D:, Giper folder, the library name – HYP1.mw:

```
restart
read(D://Giper//HYP1.mw)
with(HYP1)
```

In the future it is expected to expand the package as the composition of the subsystems, as well as individual procedures.

References. **1.** Sinkov M.V. The finite hypercomplex number systems. Basic theory. Applications / M.V.Sinkov, Ya.A. Kalinovskiy, Yu.E.Boyarinova – K.: NAS of Ukraine, IIR, 2010. – 389 p. **2.** Kalinovskiy Ya.A. Hypercomplex number systems and fast algorithms for digital processing of information / Ya.A.Kalinovskiy, D.V. Lande, Yu.E. Boyarinova, Ya.V. Hitsko – K.: NAS of Ukraine, IIR, 2014. – 130 p. **3.** Edwards Ch.H. Differential equations and boundary value problems. Modeling and computation using Mathematica, Maple and MATLAB/ Ch.H. Edwards, David E.Penny – M.: Willyams, 2008. – 1104 p. **4.** Dyakonov V.P. Maple 10/11/12/13/14 in mathematical calculations / V.P. Dyakonov – M.: DMK-Press, 2011. – 800p.

Kryvolap D.V., Bidyuk P.I.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

Modeling and forecasting Ukrainian economy in transition

Most of actual processes of economy and finances in transition exhibit nonlinear non-stationary behavior. Both linear and non-linear processes can be stationary or non-stationary. Non-stationary processes can exhibit trend (linear or non-linear) and/or time dependent variance (heteroscedastic behavior). Nonlinearity of the processes can also be of the two basic types: nonlinearity with respect to variables and nonlinearity with respect to parameters (according to known general classification) [1]. Nonlinearity with respect to variables can be formally described by polynomials, bilinear equations etc. In this case we can apply ordinary LS technique for estimation of the model parameters. On the other hand nonlinearity with respect to parameters can be described by logistic regression, appropriate nonlinear state space equations, some special functions etc. Heteroscedastic behavior means that some processes volatility is changing it time what requires special models for description dynamics of their variance. The most popular today are the following models of the heteroscedastic processes: autoregression with conditional heteroscedasticity (ARCH), generalized ARCH (GARCH), exponential GARCH (E-GARCH), and the model of stochastic volatility. The main purpose of using these models is in prediction of variance (or volatility) of the processes under consideration. The prediction estimates are useful for forecasting future prices on various products, decision making regarding trading operations, analysis of economic stability etc.

Non-stationary processes with trends are also called integrated of various order; for example, designation I(2) means that the process contains trend of second order. Quite adequate models for describing dynamics of integrated economic processes turned out to be autoregression (AR) plus members that describe linear or nonlinear trend. Pure trend models are useful for long term forecasting. A special class of macroeconomic processes create co-integrated (CI) processes. In this case we consider at least two processes with trends of similar order. For identification of existing co-integration between the processes under consideration a special statistical test is applied. One of the processes can be used for managing (control) of another one. As an example of such processes could be mentioned gross domestic product (GDP) and total tax deductions what had been proved by appropriate modeling with the use of actual statistical data from Ukrainian economy. These two processes turned out to be co-integrated with order CI(1,1), and the tax deductions could be formally used for increasing (decreasing) GDP. Generally speaking Ukrainian GDP is non-stationary processes that exhibits nonlinear behavior at some periods of time. For example, its trajectory is nonlinear at the very beginning of the transition period starting from 1990 and up to 1995. This means that development of the GDP is rather sophisticated in nature and actually requires model constructing for different modes of its trajectory.

Another point that is considered in the process of modeling the economy in transition is model adaption procedure that is required for adjusting models structure and parameters to changing modes of the processes dynamics. Such procedure is implemented in the frames of decision support system (DSS) constructed. The system is based upon preliminary data processing procedures, model structure estimation techniques using correlation functions and expert estimates, model parameters estimation procedures that include OLS, RLS, NLS, ML, and Monte Carlo for Markov Chains (MCMC) techniques. This set of methods makes it possible correct parameter estimation for the linear and nonlinear models mentioned above. The computational experiments carried out with the use of actual statistical data show that it is possible to get high quality short term forecasts with mean absolute percentage error in the range of about 0,5 – 7.5%.

References. **1.** Bidyuk P.I., Romanenko V.D., Tymoshchuk O.L. Time series analysis. – Kyiv: NTUU "KPI" (in Ukrainian), 2013. – 600 p.

Naumenko T.O.
Igor Sikorsky Kyiv Polytechnic Institute, Institute for Applied System Analysis, Kyiv, Ukraine

The analysis of fields of using ontologies in the construction of information systems

The first step in the conceptual modeling activities of Information Systems is the transformation of the perceived real-world into a model of the world it intends to represent [1]. Ontology – based on description logic knowledge representation system, widely used in solving problems of cataloging, classification and presentation friendly information for a person. We are interested in ontologies in the context of Ontology-Driven Information Systems, where they can be used at development time and at run time [1].

There is a big area of using the ontology for knowledge representation in information systems, for example:

- **tasks of the analysis of texts** – BIOAP (Basic Integrated Ontological Analytic Processor) 1.0
  The ontology defines the terms used for the description and represent knowledge of a particular subject area. It is necessary for people, for applications of database systems and different other information systems which share specific information in any subject domain. Ontologies include the available for computer processing definitions of basic concepts in the domain and the relationships between them [2].
- **evaluation of Health Information System** – FI-STAR (Future Internet Social and Technological Alignment Research)
  The ontological representation of a health information system gives a computable structure from which several indications, including evaluation aspects, can be extracted. Functions can be defined on this ontology that quantify, combine, compare, or select some of the nodes or branches. The ontology itself can be extended by assigning values to its nodes and edges, giving the possibility of further inferences [3].
- **construction of semantic information systems**
  Use of semantic technologies gives indisputable advantage at all main analysis stages, design, implementation, testing and attending of information systems, including the description of semantics of data domain with use of ontologies as variable part of an information system. Approach to creation of semantic information systems on the basis of the Repository of IT infrastructure of researches of power engineering allows quickly enough and it is rather simple to develop them [4].

Today, more and more ontologies are used to store representations of the world in multi-agent systems (MAS). It allows to apply output algorithms in the description logic based on knowledge of the agent, and also provides greater standardization and ordering knowledge that simplifies the communication between agents. The task consists in a study of methods of application of ontologies in case of design and implementation of MAS, problems and advantages of representation of knowledge of agents in the form of ontologies, and also the review of perspectives of further development of integration of knowledge representation systems into the control unit of agents of MAS.

Ontologies are widely applied at different stages of development and operation the multi-agent systems. Ontologies use for design of different MAS, and also in case of implementation of MAS for the decision of tasks of control of sales and rent, trade at the exchange, maintenance of safety of information resources, information search and other tasks.

- **an architecture for sensor web applications** – ODMAS SWAP (Ontology Driven Multi-Agent System Sensor Web Agent Platform)
  While the ontology based systems have shown some success in modeling geospatial data, none provide a comprehensive framework for representing all aspects of geospatial data (space, time, theme and uncertainty) and the system entities (agents, services and processes) that will serve

and process this data. The agent based approaches provides better abstractions for modeling and managing system entities. However, current multi-agents system infrastructures are inflexible and static and do not address ongoing application deployment and reconfiguration. Furthermore, most agent based architectures assume that ontologies already exist, and provide little support for creating and managing ontologies [5].

- **intelligent system for distributed management of small spacecraft group** – Smart Satellite
  Ontology application principles for the solution of new class of problems of intelligent moving objects behavior management that can act both fully autonomously and as a group in the context of solving tasks of Earth remote sensing by a group of small-scale spacecrafts are considered in this paper. Ontology structure and ontology/scenes editor main functions are described. An ontological description of observed objects and target recognition rules and method of objects search satisfying the ontological description are suggested. Examples of the scenes, developed on basis of the ontology that corresponds with Earth remote sensing problems are given [6].
- **multi-agent system for production control in real time** [7]

In implementations of multi-agent systems, ontologies are used to organize domain knowledge and available agents. There are three qualitatively different approaches to ontology integration into the work of multi-agent systems: either each agent stores its ontology containing available to him the knowledge and concepts or ontology is the same for all agents and stored centrally (centralized MAS), or ontology part one, and part – allocated (hybrid MAS).

As a rule, work with MAS used taxonomic ontology actually representing knowledge in a hierarchy. The perspective direction of application of ontologies in MAS are groups of **robots**: objects the component Smart House system, robots assistants of people and production in all spheres, computer-aided testing and also unmanned aerial vehicles (UAVs).

The analysis of the majority of the tasks solved with use of single UAVs shows that they can be solved more effectively group. Thus, for unmanned aerial vehicles there can be urgent problems of group interaction in the nearest future. One their possible methods of the solution of this problem and planning of strategy of behavior of each certain agent use of ontologies not only as knowledge bases about the outside world, but also as bases of the management system applying an automatic output on the basis of ontological knowledge as natural and certainly incorrect method of a choice of behavior to the agent is.

Also application of ontologies will provide standardization of representation of knowledge that, presumably, will simplify information exchange between heterogeneous agents, and also will allow in case of the situation of impossibility of communication which is often found in real projects between agents to recover or predict beforehand behavior of other agent with some accuracy on the basis of the known parts of his ontology.

References. **1.** Soares A. Building Ontologies for Information Systems: What we have, what we need / A. Soares, F. Fonseca, Pennsylvania State University, 2010. **2.** Захарова И.В. Способы автоматического построения онтологии для задач анализа текстов, Институт математики им. С.Л. Соболева СО РАН, 2011. **3.** Eivazzadeh S., Anderberg P., Larsson T.C., Fricker S.A., Berglund J. Evaluating Health Information Systems Using Ontologies, JMIR Med Inform, 2016. **4.** Копайгородский А.Н. Применение онтологий в семантических информационных системах, «Ontology of Designing» scientific journal, 2014, 78–89 с. **5.** Deshendran Moodley Ontology driven multi-agent system: an architecture for sensor web applications, University of KwaZulu-Natal, Durban, South Africa, December 2009. **6.** Скобелев П.О. Приминение онтологии в интеллектуальной системе распределенного управления группировкой малоразмерных космических аппаратов, Самарский государственный аэрокосмический университет имени академика С.П. Королева, 2015. **7.** Иващенко А.В. Мультиагентные системы для управления производством в реальном времени, Научно-производственная компания «Разумные решения», 2011.

Samuel Adegoke
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

On implementing CARMA overlay in VOIP/SIP

Theoretical Implementation of CARMA model in VoIP/SIP network to show a potential improvement mechanism for SIP connection technique using CARMA a newly developed metric system.

List of Abbreviations. SIP – Session Inititiation Protocol, VOIP – Voice Over IP, CARMA – Combined Affinity Reconnaissance Metric Architecture, P2P – Peer-to-peer.

Problem Definition. In building an efficient distributed network, topological distance estimation is key to managing network resources. Traditionally, node communication via message exchange is required and this typically has impact on resource consumption within the network.

VoIP (Voice over IP) is a method used for real-time delivery of voice communications and media sessions over public or private data network. In VoIP systems, the SIP is used for session initiations and it uses three primary address parts to locate an endpoint. Requests are sent to an SIP server via TCP to locate the client and registrations are subsequently set up following the success of the server to locate the node.

The unreliability of the server-client model gave birth to a new system,peer-to-peer, where nodes on the network form swarms and are collective contributors to the larger network as a whole. This system does not require a central server in most cases or if it'll ever need one, it'll merely be for the sake of registration on the network and the server the server has no real role to play in the activities that take place within the network.

A node on the p2p network that needs a resource sends a request for a file on the network, other nodes propagate the request for the file and available nodes with the file create a swarm then can they begin ex- change of the file. Each node in the swarm is technically referred to as a seed. The requesting node maintains connection with the swarm till all data needed have been successfully downloaded.

This essentially makes the network scalable and to an extent guarantees more security for the nodes on the network as multilayered security approach is employed.

The application of P2P systems can be found in some of the most notable overlay networks today, instant messaging, VoIP, video streaming e.t.c all use the peer-to-peer system to transmit packet data from endpoint to endpoint. SIP despite being the preferred method of communication initiation in VoIP does not cater for scalability in the network and suffers greatly per bandwidth available hence a need for more ways to improve VoIP quality of service.

Previous Research. The concept of building a virtual network over an existing network is an old concept and the Internet can in fact be considered an overlay network as it was originally built to overlay on telephone networks which was the most common at the time.

As the Internet grew and more overlays were built, research has shown that VoIP which is an overlay on top of the Internet started getting prominent due the low cost of setup and maintenance.

Researchers claim Upon the receipt of an analog voice (standard voice) by telephone, Voice Gateway first digitizes the signal and compresses the new digital signal in the form of standard data blocks, known as IP packets [3, 6]. They say the packets are sent over the Internet to the entrance of a Voice Gateway, where the process is reversed and with this technology researchers claim it is possible to make three different types of calls: PC to PC, PC to phone, and phone to phone.

VoIP consists of three essential components: CODEC (Coder/Decoder), packetizer and playout buffer. At the sender side, analog voice signals are converted into digital signals, compressed and then encoded into a predetermined format using voice codecs. There are various voice codecs developed and standardized by International Telecommunication Union.

Telecommunication (ITU-T) such as G.711, G.729, and G.723 etc. After this, packetization

process is performed and this is when the encoded voice is fragmented into equal size of packets. Each packet has protocol header from the different layers that are attached to the voice bit that has been encoded. These headers are added to the packets using RTP (Real Time Protocol), UDP (User Datagram Protocol) and the IP (Internet Protocol) as well as DLL (Data Link Layer) header [5]. Additionally researchers claim, RTP and Real Time Control Protocol (RTCP) were designed at the application layer to support real-time applications. Although TCP transport protocol is commonly used on the internet, UDP protocol is preferred in VoIP and other delay sensitive real-time applications. TCP protocol is suitable for less delay-sensitive data packets and not for delay-sensitive packet due to the acknowledgement (ACK) scheme that TCP applies. This scheme introduces delay as receiver has to notify the sender for each received packet by sending an ACK. On the other hand, UDP does not apply this scheme and thus, it is more suitable for VoIP applications [6]. The packets are then sent out over IP network to its destination where the re- verse process of decoding and depacketizing of the received packets is carried out. During the transmission process, time variations of packet delivery (jitter) may occur. Hence, a play out buffer is used at the receiver end to smoothen the playout by mitigating the incurred jitter. Packets are queued at the playout buffer for a playout time before being played. However, packets arriving later than the playout time are discarded.

Besides, there are signaling protocols of VoIP namely Session Initiation Protocol (SIP) and H.323 [4, 6]. These signaling protocols are required at the very beginning to establish VoIP calls and at the end to close the media streams between the clients.

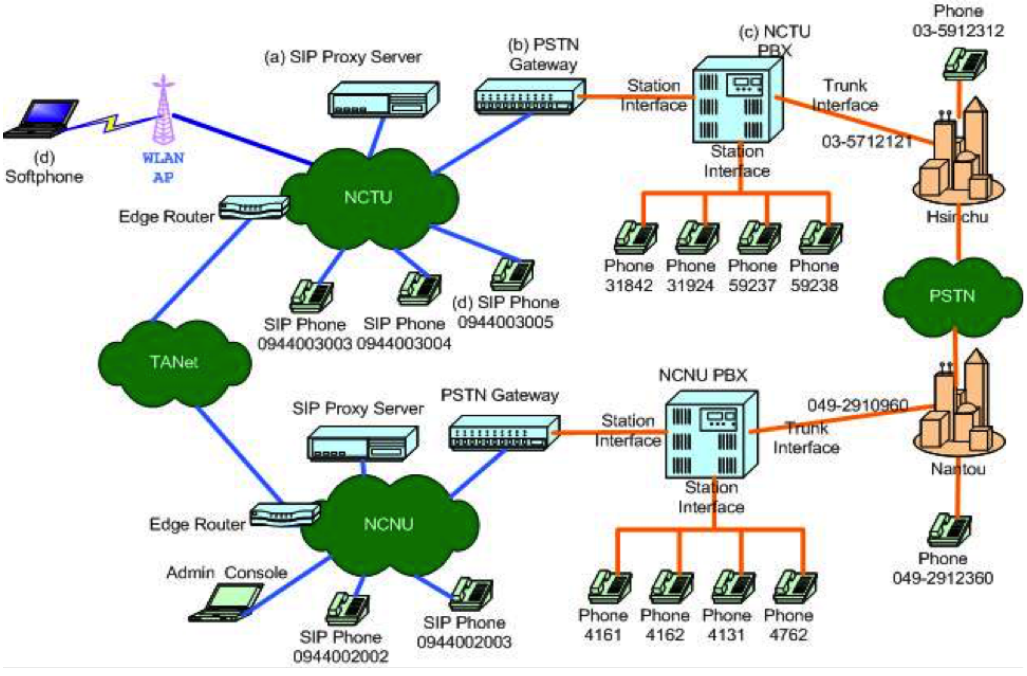


Figure 1. Showing the structure of SIP Architecture

Previous research shows that SIP uses DNS procedures to allow a client to resolve a SIP URI into the IP address, port, and transport protocol of the next hop to contact and that it also uses DNS to allow a server to send a response to a backup client if the primary client has failed. It also uses the DNS to address the issues of server, intermediate proxy location as well as location of the backup client in the failure of the failure of the primary client. First when an SIP connection is started, the SIP end system typically referred to as the user agent or client forwards the call request to a proxy server within it's own domain, this proxy server then locates the proxy server in

the other user agent's domain and forwards the request to the server who uses the URI to locate the other agent.

As part of this call flow, proxy 1 needs to determine a SIP server for domain B. To do this, proxy 1 makes use of DNS procedures, using both SRV and NAPTR records which are basically specifications of data in the DNS that defines the location and port number of servers and some other devices [2]. A diagrammatic representation of this usually referred to as the SIP trapezoid is shown in the figure below

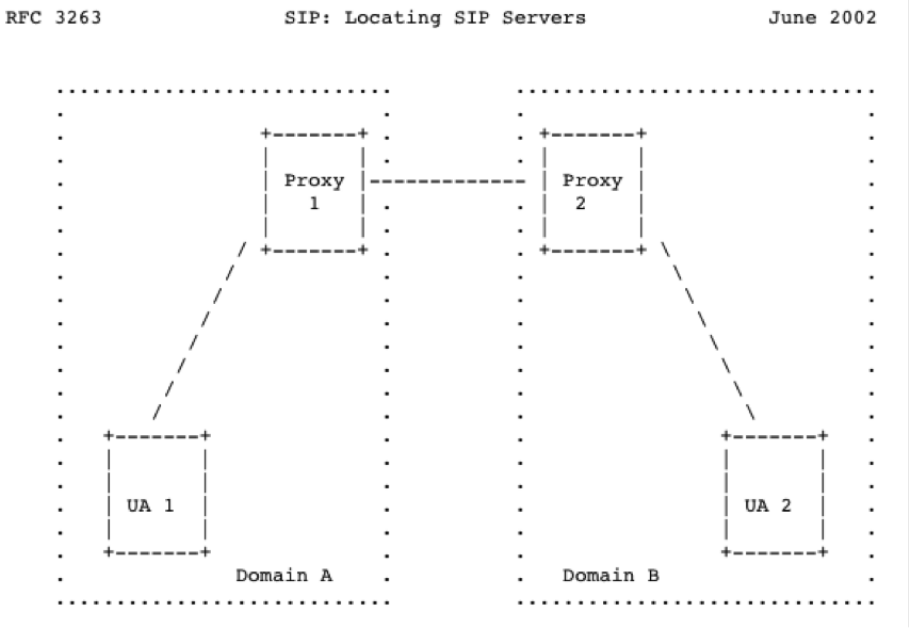


Figure 2. Showing SIP Trapezoid

It is important to note that DNS lookups can be used multiple times throughout the processing of a call. In general, an element that wishes to send a request may need to perform a DNS processing to determine the IP address, port and transport protocol of the next hop element (mostly the server) Such processing in principle could occur a number of times/number hops [5, 8]. A common resolve most past researches have is that since SIP is used for connection initiations, the time required to setup a call ideally should be minimal and the margin for error should follow suit.

2.1 VOIP/SIP Problems Despite the relative low cost of hardware and software maintenance involved in the use of VoIP, there are still a few key problems associated with it's use. During VoIP calls, packet loss is a common phenomenon and this often caused by jitters in the network and expired TTL in the packets. A network with high latency will often take longer to form the connection between the SIP clients and subsequent packet encoding/decoding will also suffer thereby leading to dropped calls, echos etc. [4, 5].

Typically, it is advisable to have 80 to 90 kbps upload and download speed for a two-way conversation using the G711 codec. Other applications running on the softphone may also try to use part of the bandwidth available thereby causing dead spots or creating choppy calls.

To achieve optimum call quality on VoIP, the use of network equipments that support packet prioritization along the link is required. This is not necessarily a good option as the balance of bandwidth usage is affected therefore there is a need for a solution that is less expensive to both hard and software components of the network and also provides quality service without affecting the throughput of the network.

CARMA. To achieve a more stable SIP process and ultimately a higher QoS, a method of connecting clients that will be less reliant on bandwidth and coordinating server is required and this is where CARMA comes in [6,8]. CARMA is a network model that approximates the Internet as a mesh of interconnected Autonomous Systems.

An AS is typically divided into IP ranges and subranges and may link to other ASes through a third subordinate AS. The mutual linkage of these ASes form an exchange point which allows to define the metric of how many AS junctions a packet has to travel between arbitrary nodes. But this falls short because the metric only allows for 7 locality flavors and practically, real hop counts that go beyond 7 are common which consequently leads to dropped packets and therefore inability to communicate with nodes farther away. CARMA addresses this problem by locally computing the metric and it does this by downloading AS IP structures from the publicly available FTP server of RIR [5,6].

After downloading the file, CARMA converts the information into a format that can be indexed and then uses an in-built algorithm to estimate the metric between the arbitrary nodes.

Judging by result accuracy, CARMA has been compared with other distance estimation techniques like PING and TRACEROUT and was even implemented in a p2p system, it was shown to have solved the problem of packets timing out before getting to destination and also the inaccuracy that occurs due to network servers using packet-shapers to block ICMP/TCP packets.

CARMA showed an improvement of 2.7% during implementation in the P2P environment. This of course is what the first version of CARMA was able to achieve on a broad scale p2p network [3,5].

Bringing the local distance estimation technique of CARMA into clients in a VoIP network, where if each client were to be fully aware of the network and knows what is the shortest path to the client to which it intends to connect, this will decentralize the workload on the SIP server and make the system more scalable as well, during the encoding and decoding of the voice packets during the phone call, voice packets can be routed to and fro communicating clients through other clients within the network. This is basically creating another overlay network on top of the existing VoIP network and also unleashing the full potential of CARMA-P2P in VoIP.

CONCLUSION. Since CARMA does local computation of network awareness, implementation of CARMA in SIP to replace client location tasks currently handled by SIP servers will eliminate latency that might occur during connection sessions when a call is being initiated.

While this may sound as simple as they come, a broader research work on the possibility of this will definitely pay off.

References. **1.** Node Placement Analysis for Overlay Networks in IoT Applications International Journal of Distributed Sensor Networks Volume 2014 (2014), Article ID 427496. **2.** https://datatracker.ietf.org/doc/rfc2848/. **3.** https://www.packetizer.com/ipmc/papers/understanding_voip/how_voip_works.html. **4.** Poryev G. Multi-Tier Locality Awareness in Distributed Networks // Інформаційні технології та комп'ютерна інженерія. 2009. № 3(16). С. 13–17. **5.** Poryev G. CARMA based MST Approximation for Multicast Provision in P2P Networks / Poryev G., Schloss H., Oechsle R. // In Proceedings of the Sixth International Conference on Networking and Services, 7–13 March 2010 p., Cancun: Proceedings Cancun, Mexico: IARIA, 2010. – P. 123–128. **6.** H. Schloss, R. Oechsle, J. Botev, M. Esch, A. Hohfeld, and I. Scholtes, "HiOPS overlay efficient provision of multicast in peer-to-peer systems", in 16th IEEE International Conference on Networks (ICON 2008), New Delhi, India, 2008, pp. 1–6. **7.** B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira, "Avoiding traceroute anomalies with Paris traceroute", in IMC '06: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. New York, NY, USA: ACM, 2006, pp. 153–158. **8.** S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection", in INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 3, 2002, pp. 1190–1199.

Trofimov A.S., Zenkovich M.V.
SeVirt, Moscow, Russia

Mobile secure runtime

This paper gives an overview of the current state of mobile security area, summarizes main areas of possible groundbreaking improvements and finally proposes the complex of steps to build an efficient secure mobile software stack. A special accent is given to a hardware virtualization capability utilization and isolation of running application instances. The paper splits software into two big groups where the first one is considered as an untrusted software and the second one is considered as a trusted software. The distribution of software over fast virtual machine(s) is isolating security critical applications from potentially malicious software. The paper also aligns software groups with common use cases such as mobile banking and enterprise operations.

Introduction. The area of embedded systems year by year shows quite intensive grow. The mobile devices are the significantly end-user negotiation specific kind of embedded systems. All nowadays mobile devices run under some operating system control. There are about three mainstream operating systems available on the market now. The Android OS as the mostly used one should be put at the first place, the overall amount of market is around 85.2%. The second place of market share is occupied by Apple iOS with the result of 14.4%. The third place with only 0.3% is taken by Microsoft Windows OS [1].

The mobile operating system is common target nowadays for various malicious software, viruses and hackers. The overall number of smartphones sold every year over the world reached 1495358000 in 2016 [1] which is really feasible number. Also it is necessary to highlight still limited set of security measures to protect end-users from the attack. The software platforms themselves as a fact cannot protect against attacks. Moreover even the most of available antivirus solutions cannot protect too.

There are multiple causes of this. The first and quite feasible is related to software distribution logic. Screening through the list of Application publishing news shows [2] high probability of getting a malicious software (SW) directly from official distribution stores (such as Apple Store, Google Play etc). The second is strictly tight to the software development methodology and quality management processes used at the software developer side. The main issue here is overall application code quality and presence of bugs in the code. The third and still quite feasible is related to the fact that there are a lot of functionality of the platform which is locked by a platform vendor and in the end end-users want to utilize rooting or jailbraking approaches [3] to be able to unlock specific features of a particular hardware.

This paper is intended to depict possible efficient approach of suppression of all three causes of attack vectors. It is designed through the utilization of virtualization capabilities of nowadays mobile hardware. The virtualization approach is combined with security policy enforcement model and instant activity monitoring. The name of the project implementing the architecture defined below is SeVirt [4] (abbreviation from Secure Virtualization).

Secure mobile system runtime architecture. Figure 1 defines the basic architecture structure. It considers that underneath hardware have support of virtualization feature acceleration (at least CPU, IC, Timers, MMU and SMMU). The secure boot logic is kept in place from the secure bootloader until the start of hypervisor.

Hypervisor (both L1 and L2 approaches are possible, in the following text the L1 hypervisor setup flow is considered) does initial initialization, applies default security policies and finally starts the control and driver virtual machine (VM) instance (CTL&DRV) which is a representation of platform and vendor specific code.

CTL&DRV sequentially starts and does the following steps: loads and setups all the platform hardware (HW) drivers, allocates and setups of all paravirtual server side drivers, finally it asks hypervisor to load precompiled trusted and untrusted VMs (TVM and UTVM) images and policies. TVM and UTVM utilize pure virtualized hardware without any direct access to the underneath real hardware.

CTL&DRV VM have runtime management components which are responsible for the following: 1. Software Manager (SM) is responsible for SW installation and uninstallation over all running VMs. 2. Window Manager (WM) is responsible for the consistent over the system windowing stack. 3. Runtime Manager (RM) is responsible for tracking the state of all running SW over the system. 4. Synchronization Manager (SYM) is responsible for keeping consistent state of system parameters quantity (such as audio volume, media storage, notifications etc) over VMs. 5. Security Model Description (SMD) is combining the functionality of keeping every particular application under security policy and tracking possible trials of policy avoiding.

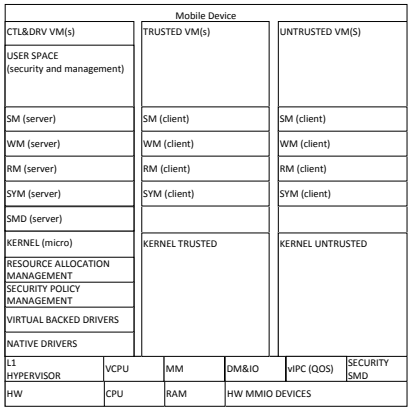


Figure 1. Architecture overview.

TVM contains single or multiple applications from one Software vendor with fairly strict policy enforcement, the policy operation mode is strictly enforced. UTVM contains multiple applications from various vendors with weak policy enforcement and possibility to restore initial state, the policy operation mode is permissive. Quality of service (QOS) is applied for all VMs over the system.

Application approaches. There are multiple areas of application of the SeVirt SW stack [4]. Two possible applications described below.

Enterprise use case. In this use case it is considered that work process related SW set (such as corporate mail client, file storage access, encryption, intranet, VPN) would be provided by an Enterprise customer. It is expected that SMD description would be also design in collaboration to achieve maximum security enforcement. From the runtime point of view all SW stays in corresponding TVM instance with strictly enforced SMD.

There are multiple evident benefits for Enterprise customers. The first is isolation of the Enterprise SW from the other end-user SW which is potentially malicious. The second is direct administrator rights of that particular VM instance. It basically means that all necessary updates might be installed in time, all user activity might be trackable and finally if the device is stolen it might be easily cleaned by the Enterprise IT department. Another one important benefit for the end-user is that the Enterprise IT department owns only corresponding VM and do not have permissions to wipe all end-user data which is not belongs to the Enterprise customer.

Mobile and on-line banking use case. In this use case it is considered that on-line banking SW is isolated from the other SW on the system. It is achieved by moving of banking SW with vendor provided SMD to the independent TVM and without vendor provided SMD to a single TVM. All other SW would be installed in a single UTVM instance.

There are multiple evident benefits for end-users and for banks. The first is isolation of on-line banking SW from the other potentially malicious SW. It may significantly reduce the probability of money steal due to three causes mentioned in introduction section. The second is possibility for banks to track avoidance of SMD policies during their SW usage.

References. **1.** Gartner, "Says Worldwide Sales of Smartphones Grew 7 Percent in the Fourth Quarter of 2016", February 15, 2017, http://www.gartner.com/newsroom/id/3609817. **2.** The Sun, "HummingWhale virus that can KILL smartphones infects millions of people's devices", January 24, 2017, www.thesun.co.uk/news/2695328/hummingwhale-virus-that-can-kill-smartphones-infects-millions-of-peoples-devices/. **3.** The Next Web, "Jailbreak attack reportedly stole more than 225,000 Apple logins", August 31, 2015, www.thenextweb.com/insider/2015/08/31/jailbreak-attack-reportedly-stole-more-than-225000-apple-logins/.tnw$_SOZfUVnu$. **4.** SeVirt project, www.sevirt.com.

Антонюк С.А.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Віртуальна реальність: новий погляд на web

Сьогодення є саме тим часом, коли сфера віртуальної реальності знаходиться на перехідному етапі. З одного боку, вже наявні цілком працездатні продукти як апаратного, так і програмного забезпечення, які відлагоджено та узгоджено взаємодіють між собою. Але в той же час, ці розробки носять скоріше тестово-прототипний характер для демонстрації можливостей та перспектив віртуальної реальності, а сама область випробовується лише одиницями. Причиною даної ситуації є те, що сучасна спільнота не бажає витрачати час та матеріальні ресурси на вивчення нової розробки, яка може здатися недосяжною в найближчому майбутньому та не дає доказів того, що стане невід'ємною частиною повсякденного користування більшості. Що ж, насправді, це не так.

Основною сферою розвитку віртуальної реальності стабільно залишається entertainment [1]. На цю нішу виділяються великі кошти гігантами світової розважальної індустрії, серед яких фігурують Google, Sony, Samsung, HTC, Acer, Facebook [2]. Розробка потужних ігрових застосунків дійсно вимагає великих ресурсів, так само, як і можливість для користувачів випробувати їх на собі, оскільки потрібно придбати як мінімум шолом віртуальної реальності, спеціалізовані маніпулятори та власне платформу роботи застосунка.

Ситуація докорінно змінюється зі створенням та введенням на ринок сумісної розробки Samsung та Oculus – окулярів віртуальної реальності Samsung Gear VR. Цей пристрій дозволяє використовувати в якості робочої платформи смартфон і працює в парі з ним, не маючи власного дисплею, а лише надаючи корпус з лінзами для кріплення смартфону [3]. Дана розробка запатентована компанією Samsung, але на теперішній час існує багато бюджетних реплік на ринку. В результаті, доступність віртуальної реальності для звичайного користувача ще більше зростає, оскільки необхідністю залишається лише смартфон з датчиком гіроскопу, який забезпечує зворотній зв'язок у VR-застосунках.

Крім описаного вище hardware, необхідним є також спеціалізоване програмне забезпечення. У сучасному розумінні, подорож веб-простором більшою мірою визначається роботою з вкладками браузера та елементами двовимірного веб-інтерфейсу користувача. У такому представленні останній реалізує лише доповнену реальність, оскільки, на відміну від цього, віртуальна реальність передбачає повне занурення у штучний світ [4]. Таким чином, перед проектувальниками та розробниками застосунків постає проблема створення, з одного боку, знайомого та зручного способу взаємодії, а з іншого – такого, який створюватиме ефект просторового оточення. На сьогодні, в цьому напрямку розвиваються спеціалізовані VR-браузери, серед яких найяскравішими представниками є Samsung Internet for Gear VR, Steam VR Web Browser, MozVR, JanusVR, Oculus Web Browser, FullDrive VR [5, 6]. Але досі більша частина інтерфейсу зводиться до представлення двовимірних об'єктів як тривимірних [4].

Для представлення у режимі віртуальної реальності найкраще підходить саме мультимедійна частина контенту: статичні VR-зображення, відео, ігри та тематичні застосунки. Так, ресурс відеообміну YouTube надає можливість перегляду VR-відео як у звичайному форматі, так і у режимі віртуальної реальності. Більш того, дедалі більше зростають можливості для створення VR-контенту користувачами: зробити таку фото- або відео-зйомку під силу будь-кому із застосуванням спеціалізованої камери 360°, а Google уже опублікували застосунок для створення VR-фотографій на платформі Android без необхідності використання спеціалізованого апаратного забезпечення [7].

Також, незважаючи на те, що сфера віртуальної реальності йде в авангарді як наукових розробок, так і споживацьких вподобань, вона ще деякий час не буде широко розповсюдженою. Крім незвичності інтерфейсу взаємодії у VR, споживачів та розробників програмного забезпечення відлякує ціна на спеціалізоване апаратне забезпечення. Також постає проблема створення дійсно якісного зображення у додатках, особливо у тих, що працюють під мобільними опера-

ційними системами, оскільки на даний момент смартфони з якісними екранами з роздільною здатністю FullHD та вище також відносяться до розряду флагманських та мають відповідну ціну. Не менш важливою задачею є зменшення часу між поворотом голови та оновленням зображення, оскільки, крім зменшення "ефекту занурення", дане явище спроможне викликати рябіння в очах, тошноту та морську хворобу. Для рішення цих задач розробники шоломів віртуальної реальності використовують технології покращення як апаратної (використання сенсору з більшою частотою дискретизації, заміна LCD панелей на OLED з низьким періодом відображення), так і програмні (передбачення руху, "time warp") [8].

Щодо професійного створення VR-застосунків, то для цього уже підготовано велику кількість бібліотечних інструментів. Зокрема, веб-розробникам обов'язково слід поглянути на такі напрацювання, як MozVR, Three.js, WebVR, Babylon.js, Blend4web, A-Frame, X3DOM [1, 4].

Фахівці зазначають, що область віртуальної реальності сумісно з доповненою реальністю є тими, що розвиваються найстрімкіше та мають найбільший потенціал. За прогнозами Goldman Sachs Global Investment Research, до 2025 року ринок VR та AR досягне позначок в 16,1 млрд долларів США від публічного сектору та підприємств та 18,9 млрд долларів з боку споживчого ринку [1]. За даними DigiCapital, ринок лише віртуальної реальності вже до 2020 року буде складати близько 30 млрд долларів США, а інвестиції в цю нішу продовжуватимуть як мінімум подвоюватися кожен рік [9].

Не дивлячись на те, що VR-досвід у всесвітній павутині є ще досить фрагментарним та примітивним, він є початком для створення масштабних застосунків для різних сфер. Так, крім онлайн відео та ігор, віртуальна реальність уже знаходить своє застосування у навчанні (електронні VR-енциклопедії, різноманітні симулятори), воєнній галузі та сфері охорони здоров'я (також симулятори), інженерному проектуванні (тривимірне моделювання), а також у рекламі та торгівлі [10]. Уже незабаром в онлайн-покупців буде можливість перед замовленням одягу через інтернет приміряти його у VR-примірочній, перед покупкою авто – провести тест-драйв у віртуальній реальності, а перед бронюванням номеру в готелі – оцінити його інтер'єр та вигляд з вікна, і це все у режимі віртуальної реальності та лише за допомогою її засобів. Це може звучати фантастично, але двадцять років тому покупки через інтернет теж здавалися фантастикою!

Підсумовуючи, слід сказати, що, незважаючи на те, що віртуальна реальність загалом і конкретно у веб-застосуванні знаходиться на початку шляху свого розвитку та використання, ця сфера є майбутнім інформаційних технологій і однозначно потребує детального ознайомлення та вивчення.

Література. **1.** 3dspace. (2016). VR for Web Developers. Retrieved from 3dspace.com/2016/04/vr-for-web-developers/. **2.** Hi-News. (2016). Крупнейшие VR-компании сформировали глобальную ассоциацию виртуальной реальности. Retrieved from hi-news.ru/technology/krupnejshie-vr- kompanii-sformirovali-globalnuyu-associaciyu-virtualnoj-realnosti.html. **3.** wikipedia. (2016). Samsung Gear VR. Retrieved from uk.wikipedia.org/wiki/Samsung_Gear_VR. **4.** designmodo. (2016). Virtual Reality and Web Design: Everything You Need to Know. Retrieved from designmodo.com/virtual-reality-design/. **5.** techradar. (2016). Best VR web browser: how to browse the internet in virtual reality. Retrieved from www.techradar.com/news/internet/web/best-vr-web-browser-how-to-browse-the-internet-in-virtual-reality-1317762. **6.** vrsource. (2016). Best VR browsers – immerse yourself into the web. Retrieved from vrsource.com/best-vr-browsers-immerse-web-5538/. **7.** holographica. (2015). Google опубликовала приложение для создания VR-фотографий. Retrieved from holographica.space/news/google-cardboard-camera-android-2511. **8.** wikipedia. (2017). Oculus Rift. Retrieved from ru.wikipedia.org/wiki/Oculus_Rift. **9.** habrahabr. (2016). Рынок виртуальной и дополненной реальности: перспективы для стартапов с точки зрения инвестора. Retrieved from habrahabr.ru/company/friifond/blog/322230/. **10.** habrahabr. (2016). Анализ рынка виртуальной реальности 2016. Retrieved from habrahabr.ru/post/318868/.

***Арчвадзе Н.Н.*[1]*, Пховелишвили М.Г.*[2]*, Шецирули Л.Д.*[3]**

[1] Тбилисский государственный университет им. И.Джавахишвили, Тбилиси, Грузия; [2] Институт вычислительной математики им. Н.Мусхелишвили Грузинского технического университета, Тбилиси, Грузия; [3] Батумский государственный университет им. Ш.Руставели, Батуми, Грузия

Об одном алгоритме параллельной сортировки

Алгоритм сортировки – это алгоритм для упорядочивания элементов в списке. Задача сортировки породила большое количество разнообразнейших решений, но некий "универсальный", наилучший алгоритм не существует. Однако, имея приблизительные характеристики входных данных, можно подобрать метод, работающий оптимальным образом. Для того, чтобы обоснованно сделать такой выбор, рассматривают параметры, по которым будет производиться оценка алгоритмов. Эти параметры: Время сортировки – основной параметр, характеризующий быстродействие алгоритма; Память – ряд алгоритмов требует выделения дополнительной памяти под временное хранение данных; Устойчивость – устойчивая сортировка не меняет взаимного расположения равных элементов; Естественность поведения – эффективность метода при обработке уже отсортированных, или частично отсортированных данных.

Основные методы сортировки делятся на два группы: 1. Квадратичные и субквадратичные алгоритмы – Сортировка выбором (SelectSort), Сортировка пузырьком (BubbleSort) и ее улучшения, Сортировка простыми вставками (InsertSort), Сортировка Шелла (ShellSort). 2. Логарифмические и линейные алгоритмы – Пирамидальная сортировка (HeapSort), Быстрая сортировка (QuickSort), Поразрядная сортировка (Radix sort) [1].

В данной статье рассматривается новый эффективный алгоритм сортировки с использованием параллельного программирования, которая упоминается как алгоритм с запозданием ("small delay") [2]. Деление исходного массива выполняется по принципу "распределения карт". Первый элемент массива передается на первое ядро, второй элемент – на второе ядро, и т.д. до n элемента (n является количеством ядер) на n-ое ядро. $n+1$ элемент не только передается на первое ядро, но и выполняется операция сортировки с элементами, расположенными там до него. $n+2$ элемент передается аналогично на второе ядро и выполняется сортировка с элементами, расположенными там. Процесс продолжается, пока все элементы массива не будут разделены на отсортированных отдельно n подмассивов и так далее, пока все элементы массива не будут разложены на n упорядоченных подмассивов.

После этого выполняется традиционная операция слияния. Основной выигрыш времени происходит потому, что сортировка массивов начинается не после распада массивов на элементы, а начинается параллельно с ним. Каждое ядро начинает операцию сортировки параллельно с "небольшой задержкой" в сравнении с предыдущим ядром, сортируя полученными каждый новый элемент с расположенными элементами уже отсортированного массива.

В представленном алгоритме фактически используются следующие алгоритмы:

- деление по принципу карты;
- поиск места по принципу поиска абонентов в телефонной книге;
- деление в месте слиянием;
- параллельное разложение паркета.

Главная новизна данного алгоритма состоит в том, что методы разложения и сортировки массива происходят в параллельном режиме, и это возможно из-за того, что обработка на ядре начинается с задержкой по сравнению с предыдущим ядром. Выигрыш времени обусловлен тем, что не происходит задержка при разложении массива на подмассивы.

Литература. **1.** http://algolist.manual.ru. **2.** N. Archvadze, M. Pkhovelishvili, L. Shetsiruli, O. Ioseliani. The Algorithm of Parallel Programming Using "Small Delay". SCCTW'2016 South-Caucasus Computing and Technology Workshop. 2–7/10/2016. indico.cern.ch/event/572800/contributions/2319259/attachments/1347569/2032637/WOR1.pdf.

Бакун С.А., Литвинюк А.А.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Прогнозування врожайності сільськогосподарських культур з використанням технологій Machine Learning

Вступ. У сучасному світі одним з найбільш важливих стратегічних завдань будь-якої країни є забезпечення продовольчої безпеки. В умовах ринкових відносин і самостійності господарюючих суб'єктів задача прогнозування обсягів врожайності набуває ще більшої актуальності [1]. Оцінювання умов вегетації, стану посівів та прогнозування врожайності сільськогосподарських культур є досить важливим, але використання загальноприйнятих методів не дозволяє задовольнити потреби в аналізі гібридів сільськогосподарських рослин для знаходження найбільш врожайного. Обробка великої кількості параметрів, перерахунок прогнозів та пошук оптимальних комбінацій є досить трудомісткою задачею.

Вхідні дані. База даних за 2013–2015 роки, що складається з 3365 записів по врожайності гібридів кукурудзи. Кожен запис містить 33 параметра, що описують цільову змінну (гібрид кукурудзи, місце розташування, площа, агрокліматичні та агрохімічні параметри).

Постановка задачі. Опрацювання технологій та моделей для прогнозування врожайності сільськогосподарських культур (на прикладі кукурудзи зернової та її гібридів) для умов України. Знаходження оптимальних гібридів з використанням побудованої прогностичної моделі.

Метод вирішення. Розрахунки проведено на основі статистичного ряду середньообласної врожайності кукурудзи та значень агрокліматичних параметрів з використання можливостей Azure Machine Learning. Для побудови прогностичної моделі (рис. 1) було використано наступні елементи Azure ML:

1. Edit Metadata – для визначення типів параметрів (числові, категоріальні) та цільової змінної;
2. Clean Missing Data – для обробки пропущених даних (в залежності від того чи іншого параметру);
3. Filter Based Feature Selection – для вибору найбільш значущих параметрів (на основі кореляції Пірсона);
4. Split Data – для розбиття вибірки на навчальну та тестову;
5. Train Model – для відновлення та знаходження зв'язків між параметрами та врожайністю (побудова прогностичної моделі);
6. Score Model – для прогнозування врожайності з використанням побудованої моделі;
7. Evaluate Model – для оцінки якості побудованої моделі на основі статистичних характеристик.

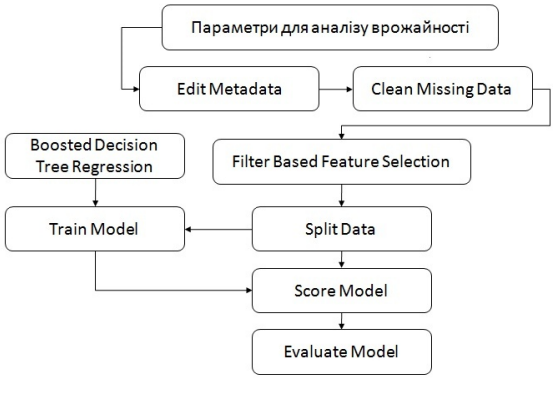


Рис. 1. Схема побудованої прогностичної моделі в Azure ML

Результати дослідження. У результаті роботи було побудовану модель прогнозування врожайності з використанням методу Boosted Decision Tree Regression. Після проведеного кореляційного аналізу в структуру моделі ввійшли наступні параметри: випаровування вологи в травні, випаровування вологи в червні, випаровування вологи в серпні, випаровування вологи в вегетаційний період, опади в травні, опади в вегетаційний період, баланс вологи в вегетаційний період, сума активних температур градусів більше 10 градусів по Цельсію (з травня по вересень), тривалість днів з температурою більше 30 градусів по Цельсію безперервно протягом вегетаційного періоду, середня вологість грунту на глибині 0–7 см (серпень), середня вологість грунту на глибині 7–28 см (серпень), середня вологість грунту на глибині 28–100 см (серпень), норма добрив типу N (азот) кг/га, норма добрив типу N+P+K кг/га, попередник, тип грунту (рис. 2).

Опади ; мм (травень)	Випаровування; мм (червень)	Середня Вологість грунту на глибині 0-7 см (серпень)	Код попередника	КОД Типу грунту	Scored Labels
82	109.85	0.43	22	11	9.643968
57.4	108	0.41	7	8	7.206676
66	110	0.48	16	11	9.692732
113	89	0.66	16	11	9.779696

Рис. 2. Результати прогнозування

Характеристики побудованої моделі говорять про її хорошу прогностичну здібність (табл. 1).

Табл. 1. Результати оцінювання якості побудованої моделі

Статистичні характеристики	Значення
Mean Absolute Error	0.8372
Root Mean Squared Error	1.1386
Relative Absolute Error	0.3311
Relative Squared Error	0.1558
Coefficient of Determination	0.8441

Для пошуку найбільш врожайного гібриду кукурудзи було проведено серію експериментів з комбінацією гібридів та агрокліматичних параметрів. Аналіз результатів показав, що вибір гібридів кукурудзи для посіву з використанням Azure Machine Learning може призвести до збільшення врожайності в середньому на 10%.

Висновки. Прогнози врожайності зернових культур, що складаються в період вегетації рослин, ґрунтуються на стані посівів і агрокліматичних умовах, що склалися в період їх росту та розвитку [2]. Проведене дослідження демонструє, що моделювання і прогнозування врожайності сільськогосподарських культур з використанням технологій Machine Learning на різних рівнях виробництва може використовуватись з досить високою точністю. А пошук оптимального гібриду для посіву в тих чи інших умовах значно спрощується та стає більш точним.

Література. **1.** Куссуль Н.Н. Регрессионные модели прогнозирования урожайности зерновых в Украине по спутниковым данным различной природы / Н.Н. Куссуль, А.В. Колотий, С.В. Яцков, Т.В. Олейник // Наукові праці Донецького національного технічного університету. Сер.: Інформатика, кібернетика та обчислювальна техніка. – 2013. – № 1. – С. 94–101. **2.** Сердюченко Н.М. Прогнозування врожайності кукурудзи з використанням даних ДЗЗ / Н.М. Сердюченко, Р.В. Сайдак // Меліорація і водне господарство. – 2014. – Вип. 101. – С. 41–50.

Бакурова А.В., Ропало А.М.
Запорожский национальный технический университет, Запорожье, Украина

Анализ существующих информационных систем в управлении аптечными сетями

Аптечный рынок Украины непрерывно развивается. Это связано с наличием у больших игроков финансовых ресурсов и их желанием увеличить свое присутствие на рынке. Поэтому количество торговых точек увеличивается от года к году. Количество аптек в 2016 году в Украине составило 19 799 (это на 685 больше, чем в 2015 году). Более половины торговых точек – это аптеки небольших аптечных предприятий. Но развитие больших сетей приводит к уменьшению веса малых сетей на рынке, их доля сократилась с 55,2% (в 2014 году) до 53% (в 2016 году). При этом доля больших сетей увеличилась с 24,1% (в 2014 году) до 27,7% (в 2016 году) [1].

Для розничной торговли аптечные сети используют специальное программное обеспечение (ПО), обеспечивающее не только отпуск лекарственных препаратов на точке, но и сбор, хранение, анализ данных о продажах, а также множество других важных функций.

Как правило, покупаются уже готовые решения, поскольку разработка подобного ПО своими силами является продолжительным и затратным проектом. Обслуживание установленной программы происходит удаленно. Основными требованиями к аптечному ПО становятся функциональные цели аптечной сети. При этом опыт использования подобных программ большими игроками рынка показывает, что, в первую очередь, аптечное ПО должно осуществлять вспомогательную функцию продажи товаров с возможностью оперативного анализа продаж в режиме реального времени. Также качественная программа должна иметь интеграцию с информационными системами автоматизации фармзаказов МОРИОН и LIKIS, иметь гибкую политику ценообразования (учет цен производителя, контроль максимальной наценки, фиксированные цены).

Сейчас в Украине на рынке аптечного ПО чаще всего используются системы: «1С-Аптека», «Аптека Морион», «АНР», «Парацельс», «Скарб». Крупные торговые сети предпочитают программы «Аптека Морион» и «АНР». Эти программы – лидеры рынка, при этом на данном этапе развития рынка система «АНР» является лидером, как за счет ценовой политики компании-разработчика, так и в силу своей функциональности.

Каждая из систем имеет свои достоинства и недостатки. Хорошим решением для аптек является программа «1С Аптека». Программа может охватить большое количество аптек, но базовой оболочки недостаточно для ее моментального использования, как и большинство программ 1С она требует доработки под конкретное предприятие.

Самой дешевой программой для аптеки является программа «Парацельс». Но, оптимальный охват аптек программой – до 20, что ограничивает ее применение для аптечных сетей. Также, за счет низкой стоимости среди аптечных сетей часто представлена информационная база «Скарб».

Проанализируем функциональность существующего аптечного ПО.

Система «1С Аптека» была разработана для обслуживания специальных операций (учет и продажа лекарств, а также товаров медицинского назначения) и управления бизнес-процессами (учет сроков годности лекарственных средств, формирования специальной системы ценообразования) аптек. Решение поддерживает контроль жизненно важных препаратов, ассортимента для формирования минимального остатка, формирует группы товаров по действующим веществам и по типам заболевания. Информационная база «1С Аптека» состоит из распределенных баз с разделением подзадач по торговым точкам, консолидируя при этом информацию по всем

аптекам [2]. Программа имеет базовые возможности оперативного формирования отчетов: «Отчет по продажам», «Отчет по возвратам», «Товарный отчет», «Отчет по минимальным остаткам», «Показатели эффективности». Но для эффективной работы аптечной сети подобных базовых отчетов недостаточно, поэтому при интеграции данного ПО, как правило, компанией-заказчиком оплачивается доработка оболочки. И чем больше компания, тем больше программа требует дополнительных программных модулей. То есть, это не готовое решение, а решение с открытым программным кодом, который может доработать как продавец, так и заказчик силами своих специалистов.

Система «Аптека Морион» имеет один из самых больших справочников лекарственных препаратов. Программа позволяет сравнивать цены заказа медикаментов у поставщиков, включает в себя описание многих препаратов. Информационная база «Аптека Морион» имеет заполненную базу аналогов на препараты, автоматически отображает их. Программа имеет множество основных форм анализа, позволяет строить различные типы анализа в зависимости от индивидуальных процессов у клиента. Программа является интеллектуальной, так как помимо возможности анализа продаж в режиме реального времени позволяет обмениваться документами (в т.ч. прайсами) с поставщиками в online режиме [3]. Недостатком программы является закрытая среда программирования, что может создать сложности при обслуживании отдаленных точек, а также при доработке.

Информационная база «АНР-Аптека» – система автоматизированного управления процессами аптечных предприятий. Система «АНР-Аптека» с 2011 года доступна на платформе «1С: Предприятие 8.2». Программа поддерживает работу основных бизнес-процессов аптеки: оприходование лекарственных препаратов, перемещения их между складами (складской учет), переоценка товара, дисконтирование скидки, продажа лекарств, закупка товаров у поставщиков, экспорт данных о продажах в учетную систему компании и импорт ее обратно в аптеки. Программа позволяет формировать первичные отчеты по продажам с их категоризацией и детализацией. Помимо отчетности управлять торговыми точками в режиме реального времени помогают дополнительные опции программы: «Накладные» (фиксирует оприходование препаратов складом посредством получения электронных накладных); «Запасы» (проводит ABC и XYZ анализ, анализирует плановые показатели оборачиваемости, автоматически формирует дефектуру, оптимизирует наличие товаров внутри сети, позволяет перераспределять товары внутри сети между торговыми точками, контролирует складские запасы – определяет неликвидные препараты); «Заявки от аптек» (формирует и передает информацию по дефектуре); «Заказы» (импортирует прайсы от поставщиков, анализирует предложения от них, формирует заказы); «Контроль» (контролирует заказы от поставщиков); «Дисконтная программа»; «Экспорт отчета». Помимо данных о продажах система владеет информацией о прайсах поставщиков и может автоматически формировать и заказывать товары [4].

«Парацельс» – готовая программа, которую можно приобрести единоразово (например, в отличие от системы «Аптека Морион»). В «Парацельс» представлена достаточная для розницы база лекарственных средств (32 000 препаратов). Также «Парацельс» содержит базу аналогов, систему оперативной обработки прайсов-листов поставщиков для формирования ежедневного заказа (аналог LIKIS), разнообразную отчетность, возможность проводить первичную аналитику продаж, генератор отчетов. Программа «Парацельс» – довольно простое решение, основным преимуществом которой является ее низкая стоимость [5].

Особенностью программы «Скарб» является возможность формирования различных отчетных форм, с которыми смогут работать руководители или сотрудники, которым эта

информация важна в режиме реального времени. Программа позволяет формировать отчетности для контроля и принятия качественных управленческих решений, статистику продаж, отчетности для поставщиков, отчетности для маркетинговых служб [6].

Однако следует отметить, что аптечный рынок – высококонкурентный, где эффективная управляющая программа приведет компанию к усилению своей позиции на рынке. Поэтому в последнее время большие аптечные корпорации задумываются о внедрении в свою информационную базу статистической аналитической системы, которая бы отвечала следующим требованиям:

1. Интеграция данных и соблюдение требований. Этому требованию условно соответствуют только 3 системы: «Аптека Морион», «АНР», «Скарб».
2. Инновационные статистические методы. Это требование возможно реализовать с помощью программ, у которых открытый программный код. Это программы «1С-Аптека» и «АНР».
3. Анализ, формирование отчетов, визуализация информации. Полноценная система вспомогательных инструментов одного разработчика значительно уменьшает расходы на внедрение, поддержку и обучение, обеспечивает актуальность информации и простой доступ к ней всех участников процесса. Этому критерию соответствуют все исследуемые программы, поскольку аптечное ПО – узкоспециализированный скомпилированный целостный продукт.
4. Простое создание графиков и отчетов. Этому требованию не соответствует ни одна программа. Только 3 программы являются условно интеллектуальными. Это «Аптека Морион», «АНР» и «Скарб». Но, ни одна из них не осуществляет ассоциативный поиск и качественную визуализацию обработанных вычислений в оперативной памяти.

К подобным системам нового поколения относятся программные решения, разработанные на базе технологий SAS. В основе построения системы интеллектуального управления аптеками компания ставит следующие постулаты:

1. лица, участвующие в анализе результатов продаж должны иметь возможность работы непосредственно с готовыми данными, а не на преобразовании их в логический вид;
2. функциональность процессов должна постоянно развиваться и дополняться новейшими решениями;
3. данные должны быть представлены в единой форме;
4. информация должна быть целостной и доступной.

Сравнение систем с точки зрения критериев их функциональности и цены представлено в таблице.

Согласно критерию «Интеллектуальная компонента» баллы проставлялись тем программам, которые позволяют лицам, принимающим решения получать необходимую информацию в режиме реального времени для оперативного вмешательства в деятельность торговой точки.

Ключевым фактором при выборе аптечного ПО также является критерий открытости среды программирования для возможности его самостоятельной доработки.

Системы, а именно, «Аптека Морион», «АНР» имеют мощные аналитические модули, использование которых позволит аналитической службе сосредоточиться на подготовке решений, а не на приведение информации в удобную аналитику форму.

К важному критерию также отнесем стоимость программного обеспечения. Так для установки на 1 аптеку программ «Парацельс», «Скарб» покупатель заплатит единоразово менее 5 000 грн., при этом стоимость 1 часа доработки ПО составляет менее 200 грн. Например, оболочка программы «1С-Аптека» изначально требует программного дополнения и стоимость

подобных работ повлияют на ее окончательную стоимость.

Как правило, большая торговая сеть приобретет программу, способную обслуживать большое количество аптек. Такому критерию соответствуют «Аптека Морион», «АНР», «Скарб».

Лидерами по простоте обучения являются информационные системы для аптек «Аптека Морион», «Парацельс». «Аптека Морион» – за счет длительного присутствия компании на рынке и опыта внедрения, «Парацельс» – за счет понятной логики интерфейса. Программы написанные на платформе 1С (такие как «АНР аптека» и «1С-Аптека») достаточно сложны для восприятия человеком, не работавшим в базах 1С.

Критерий\ ПО	«1С-Аптека»	«Аптека Морион»	«АНР»	«Парацельс»	«Скарб»	Возможности ПО на базе SAS
Интеллектуальная компонента		+	+			+
Открытая среда программирования	+		+			+
Аналитика		+				+
Цена, в т. ч. стоимость одного часа доработки				+	+	+
Управление большим количеством аптек		+	+		+	+
Легкость обучения		+		+		+
Распространенность		+	+	+		
Удаленное обслуживание	+	+	+	+	+	+
Единоразовое приобретение	+		+	+	+	+
Количество баллов	**3**	**6**	**6**	**5**	**4**	**8**

Согласно приведенной таблице, можно сделать вывод, что наиболее оптимальным для аптечной сети является продукт компании SAS (8 баллов). При этом программа не получила баллы только за критерий «Распространенность», что можно отнести к ее недостаткам условно.

Литература. **1.** Кирсанов Д. Как развивается розничный сегмент? Инфраструктура аптечного рынка Украины. / Еженедельник Аптека, №3(1024). **2.** www.rarus.com.ua/roznica/1c8-apteka-ukraina/. **3.** www.pharmbase.com.ua/projects/avtomatizaciya-aptek/#x19. **4.** www.anr.ua/system/functionality/. **5.** www.farmprofi.net/paracels/. **6.** www.skarb.ua/uchot.html.

Бидниченко А.А., Безносик А.Ю.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Использование объектной модели в ядрах современных Unix-подобных операционных систем

Целесообразность использования объектной парадигмы при проектировании современных приложений неоднократно доказывалась теоретическими исследованиями и практическими результатами ее применения. Можно сказать, что практически все программные продукты, если и не всецело, то частично, прибегают к этой парадигме.

В соответствии с объектной парадигмой, задачи реализации нового функционала, понимание структуры существующего кода возводятся в рамки функциональных сущностей – «объектов», которые, по существу, являются абстракцией повседневной жизни. Как безусловное преимущество объектной парадигмы стоит отметить пониженный «порог вхождения» в продукт, а также трудозатраты на устранение дефектов и некорректного поведения.

Учитывая вышеупомянутые преимущества, возникает вполне обоснованное желание использовать объектную парадигму в одном из самых сложных программных продуктов – ядре операционной системы. Но такого рода задача отнюдь не тривиальна в силу ряда причин: временные затраты на создание такой модели в рамках существующей ОС могу привести к потере рынков использования последней, более того, изменения такого рода однозначно приведут к нестабильности в работе системы, что поставит под угрозу не один миллион пользователей. Вариант написания «новой ОС» так же маловероятен в силу времени и стабильности работы [1].

macOS. Самое продвинутое ядро в вопросах абстрагирования сущностей до уровня объектов имеет операционная система macOS фирмы Apple. После десятилетий попыток создания собственной ОС от первой до последней строчки кода, Apple представила свою Mac OS X, которая, в сущности, объединила исходные коды операционных систем FreeBSD, NeXTSTEP с уникальным интерфейсом компании Apple. В составе данной операционной системы есть модуль, который носит название IOKit (рис. 1). Данный модуль представляет собой иерархию классов-устройств, что позволяет свести написание расширений ядра к реализации интерфейсов данной иерархии на языке программирования C++. Реализация определенного интерфейса дает четкое понимание типа устройства, драйвер которого мы пишем, и снижает трудозатраты и время на исправление ошибок [2]. Как видно из вышеприведенного рисунка, IOKit не покрывает всех элементов системы. Его цель – предоставить интерфейсы для написания драйвера определенного класса устройств. При возникновении необходимости модификации поведения VM (виртуальной памяти) либо VFS (виртуальной файловой системы), либо при желании написать собственный сетевой стек, разработчик будет вынужден вернуться к классическому функциональному программированию расширений ядра.

Linux. В то же время, самая популярная Unix-подобная ОС Linux содержит в себе два типа кода: «ассемблерные вставки» и код на языке C. Оба языка не имеют никаких механизмов для создания объектных абстракций. Тем не менее, такая логичная с точки зрения объектной модели иерархия классов-устройств не могла не остаться незамеченной разработчиками данного ядра. И, как следствие, в Linux появились Kobject (рис. 2), изначально как система классов для устранения проблем утечки памяти путем подсчета ссылок. Но к текущей версии ядра они разрослись до полноценных объектов. Однако, при написании драйверов использование Kobject не является очевидным. Реализация собственного Kobject необходима, только если класс устройств, драйвер которого мы собираемся написать, отсутствует в Linux, а это, как показывает практика, – крайне редкий случай. В остальных случаях написание драйверов устройств сводится к реализации колбеков (ссылок на функции), предоставляемых определенным Kobject-ом [3].

Дальнейшие исследования в данном направлении предполагают оценивание рисков, вре-

менных затрат и коммерческой выгоды при использовании объектной модели для создания расширений ядра на примере реализации драйвера сетевого устройства в Linux и macOS. Также запланировано провести аналогичные исследования существования объектных моделей в других Unix-подобных операционных системах: HP AIX, Oracle SunOS.

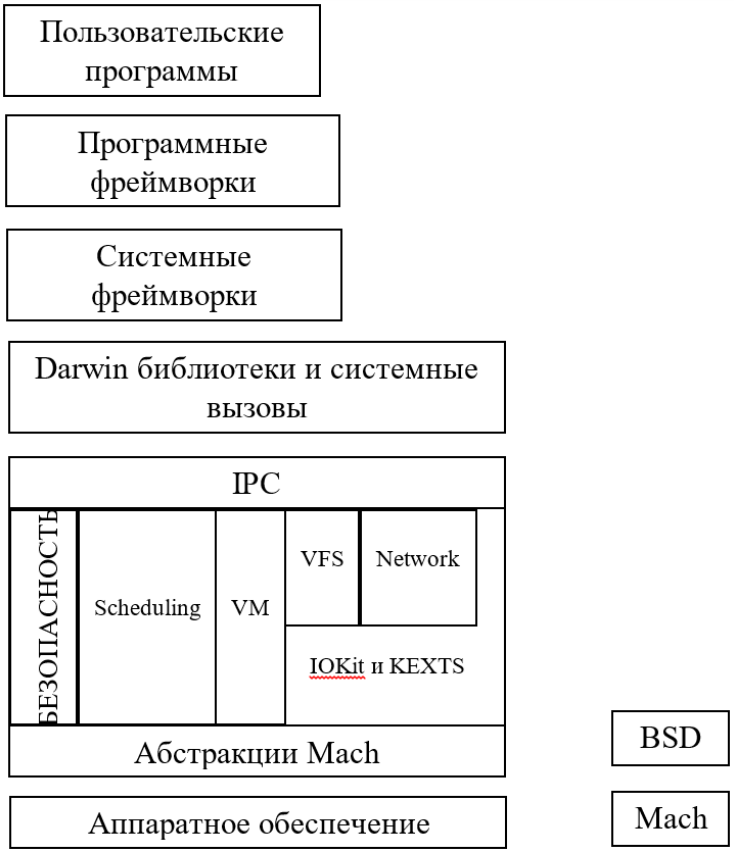


Рис. 1. I/O Kit в общей архитектуре macOS

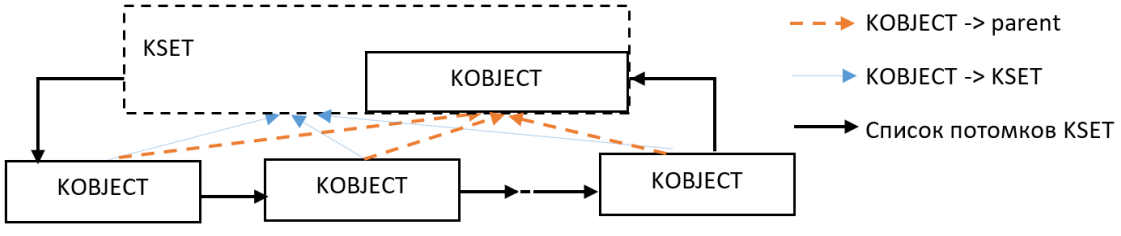


Рис. 2. Общая организация классов для KOBJECT объектов

Литература. **1.** Таненбаум Э. Современные операционные системы / Таненбаум Э., Бос Х. – СПб.: Питер, 2015. – С. 22–106. **2.** Levin J. Mac OS X and iOS Internals: To the Apple's Core / Levin J. – Wrox, 2012. – P. 711–773. **3.** Corbet J. Linux device drivers / Corbet J., Rubini A., Kroah-Hartman G. – O'Reilly, 2005. – P. 362–407.

Бідюк П.І., Терентьєв О.М., Просянкіна-Жарова Т.І., Савастьянов В.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Застосування інструментів SAS BASE для дослідження ефективності методів обробки пропусків у вибірках даних з метою підвищення якості прогнозування показників продовольчої безпеки країни

Вступ. В умовах суспільно-політичних перетворень проблема забезпечення населення продовольством в достатній кількості і за доступною ціною є не лише економічним чинником, а й передумовою соціальної стабільності. Саме тому у даному дослідженні, виконаному в рамках проекту NUKR.SFPP G4877 «Modeling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism» [1], як одні із пріоритетних, розглядаються питання розробки та впровадження ефективних інструментів аналізу та прогнозування кількісних параметрів продовольчої безпеки. Однією із практично значущих проблем при розв'язуванні таких задач є проблема розкриття невизначеностей різного роду, зокрема інформаційної, зумовленої пропусками даних.

Мета дослідження. Аналіз існуючих методів заповнення пропусків та розроблення рекомендацій щодо вибору оптимального підходу до оброблення неповних вибірок даних з метою підвищення якості прогнозного моделювання показників продовольчої безпеки.

Викладення основного матеріалу. У роботах багатьох вітчизняних та закордонних вчених, пов'язаних із прогнозним моделюванням соціально-економічних процесів, значна увага приділена питанням, пов'язаним із заповненням пропусків у відповідних часових рядах. Очевидно, що вибір методики залежить від типу змінних, що містять пропуски, особливостей предметної області дослідження, виду аналізу, що в подальшому буде використаний для побудови прогнозних моделей [2]. Враховуючи специфіку часових рядів, що описують досліджувані процеси, для заповнення пропусків доцільно використовувати адаптивні моделі з трендом та сезонною компонентою, зокрема, моделі із затухаючим трендом, адитивні та мультиплікативні моделі Хольта-Вінтерса, авторегресійні моделі тощо. Зазначимо, що методи, які використовують заміну пропусків сталими значеннями є взагалі неприйнятними. Саме тому, особливе значення має ефективна організація робіт з накопичення необхідної інформації, попередньої обробки даних для їх подальшого опрацювання у системі підтримки прийняття рішень (рис. 1).

Рис. 1. Послідовність обробки інформації на етапі попереднього аналізу даних для прогнозування продовольчої безпеки

Не менш складною проблемою є вибір комп'ютерної аналітичної системи, яка б дозволяла реалізувати якнайширший спектр методів дослідження часових рядів та інтелектуального аналізу даних. Тому у даному дослідженні використано середовище SAS Base [3], зокрема, модулі SAS/ETS (Econometrics and Time Series Analysis) та SAS/STAT (Statistical analysis).

В рамках дослідження було виконано декілька обчислювальних експериментів із використанням вибірок різних статистичних даних, що описують продовольчу безпеку. Зокрема, були досліджені часові ряди показників, вимірюваних у натуральних одиницях (дані щодо валового збору основних сільськогосподарських культур в Україні за період 1940–2014 рр.) та у грошових одиницях (дані про виробництво валового регіонального продукту однієї з

областей України за 2003–2014 рр.) [4]. Досліджувана вибірка даних щодо валового збору основних сільськогосподарських культур України за 1940–2014 рр. у ряду «фрукти та ягоди» містить пропуски даних за окремі роки, додатково 10 відсотків вхідних даних були заповнені пропусками у довільному порядку. Заміна пропусків даних здійснювалась на: середнє, медіанне значення, попереднє заповнене значення, середнє сусідніх значень, значення, отримане в результаті побудови моделі з трендом. Крім того, запропоновано власний метод, який передбачає заповнення пропусків усередненим середнім, розрахованим на основі попередніх п'яти відомих значень ряду із використанням вагових коефіцієнтів. Вагові коефіцієнти розраховувалися за тим же принципом, що використовується у методі експоненційного згладжування. Прогнози найкращої якості одержані за методом заміни пропусків на попереднє заповнене значення (MSE=172,94) та за методом заміни пропусків на усереднене за попередні п'ять років значення (MSE=160,55).

Враховуючи особливості формування вибірок даних фінансово-економічних показників продовольчої безпеки, а саме, неможливість одержання досить довгого часового ряду співставних статистичних даних за наявності значної кількості показників (використано 26 показників за 2003–2014 рр. на мезорівні та 32 показники за 2003–2013 рр. на макрорівні), для відновлення пропущених значень були застосовані методи імпутування із використанням регресійних моделей, дерев рішень, зокрема методу Кааса і методу Stepwise. Для прогнозування показників продовольчої безпеки на мезорівні побудовані такі моделі: нейронна мережа зі структурою багатошарового персептрона з трьома прихованими вершинами, з критерієм оптимізації моделі – мінімізація середньої похибки моделі (MSE – 9589,711, SSR – 86307,4); нейронна мережа із структурою багатошарового персептрона з трьома прихованими вершинами, використанням кластерного аналізу та дерев рішень і критерієм оптимізації – мінімум середньої похибки моделі нейронної мережі (AIC – 114,6902, MSE – 112672,7, BSC – 115,676, SSR – 1014054); регресійна модель, у якій кластери, побудовані для попередньої моделі, використано у якості вхідних змінних (AIC – 23,67, MSE – 31,925, BSC – 31,03, SSR – 287,322); пропущені значення часового ряду в окремих кластерах доповнені за методом дерев рішень. Для прогнозних моделей, побудованих для макрорівня (відсутні значення заповнювалися у часових рядах чотирьох показників із 32, що використані у прогнозному моделюванні), прийнятні результати заповнення пропусків у всіх випадках одержані з використанням регресійних моделей, значення критеріїв якості знаходяться в межах: AIC: 5,67–20,68, MSE: 0,13–0,91, BSC: 5,71–20,73 (відповідно). Для перевірки якості прогнозування використано порівняння прогнозних показників із відповідними статистичними показниками за 2012–2014 рр. Розбіжність значень реальних та прогнозованих показників складає від 3 до 5%, що вказує на високу якість прогнозу.

Висновки. Як показало виконане дослідження, для заповнення пропусків цілком прийнятні результати можна одержати використовуючи не лише ітеративні алгоритми, а й навіть такі методи, як просте усереднення (для випадків, коли відсутня лише незначна частина спостережень), регресійне моделювання, заповнення пропусків шляхом генерації прогнозованих значень на основі наявних, генерації відсутніх (втрачених) даних на основі форми розподілу та параметрів, що визначені використовуючи наявну частину даних, оптимізаційні методи, зокрема EM-алгоритми, експоненційне згладжування, тощо. Однак, найкращі результати можна одержати лише за умови якісної попередньої обробки вхідної інформації.

Література. **1.** Офіційний сайт «Україна-НАТО». (2016) Програма НАТО «Наука заради миру та безпеки» [Електронний ресурс]. Режим доступу: http://ukraine-nato.mfa.gov.ua/ua/about-nato/science-for-peace. **2.** N. V. Kusnietsova and P. I. Bidyuk, "Business intelligence techniques for missing data imputation", Наукові вісті НТУУ «КПІ», No. 5, 2015, pp. 47–56. **3.** Multiple imputation in SAS / Institute for digital research and education. (2016) [Електронний ресурс]. Режим доступу: http://stats.idre.ucla.edu/sas/seminars/multiple-imputation-in-sas/mi_new_1. **4.** Офіційний сайт Державної служби статистики України. (2016) Комплексні статистичні публікації [Електронний ресурс]. Режим доступу: http://www.ukrstat.gov.ua.

Біла А.І., Бідюк П.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Побудова моделей гетероскедастичних процесів

Гетероскедастичні процеси (ГСП) – це процеси із змінною у часі дисперсією, які зустрічаються в економіці, фінансах, технологічних процесах та багатьох інших галузях діяльності. Особливо часто необхідно моделювати динаміку дисперсії процесу (або відповідного стандартного відхилення), яка необхідна для аналізу стаціонарності, обчислення ринкових фінансових ризиків, прийняття рішень при виконанні торгових операцій з фінансовими активами, для формування портфелів фінансових активів та побудови систем підтримки прийняття рішень для керування фінансовими потоками [1–3]. Стандартне відхилення в аналізі фінансових процесів називають волатильністю. Її можна визначити як міру мінливості (варіабельності) випадкової складової часового ряду, який характеризує поведінку фінансової змінної. Волатильність – це неспостережуваний параметр, а тому її необхідно коректно оцінювати на основі спостережень основної змінної. Простим популярним підходом до оцінювання волатильності при аналізі можливих втрат є використання квадратів інновацій для щоденних значень фінансового процесу (наприклад, доходності).

Однією з ключових моделей, які використовують для опису умовної дисперсії, є узагальнена авторегресія із умовною гетероскедастичністю (АРУГ), що має структуру авторегресії з ковзним середнім [3, 4]. Недоліком цієї моделі є те, що згідно з таким описом волатильність експоненційно затухає в часі, що не завжди відповідає дійсності. Численні дослідження свідчать, що зміни волатильності можна краще описати так званою "довгою пам'яттю", яка передбачає гіперболічне спадання цінових шоків, тобто спадання з меншою швидкістю. З іншого боку, така поведінка волатильності спостерігається не завжди. При дослідженні фінансових процесів також важливо встановити наявність структурних змін у відповідних часових рядах, що дає можливість відслідковувати довгострокові ефекти, включаючи волатильність. Також з основних моделей пропонується розглянути такі: модель авторегресії з умовною гетероскедастичністю – УАРУГ, подальшим розвитком моделі гетероскедастичних процесів є експоненціальна модель УАРУГ (ЕУАРУГ), ще одна модель – стохастична модель волатильності характеризується мінливою структурою дисперсії, яка описується стохастичним, наприклад, нормально розподіленим процесом.

Постановка задачі. Метою роботи є:

1. аналіз і вибір для практичного використання сучасних моделей для прогнозування волатильності;
2. модифікація існуючих моделей умовної дисперсії з метою підвищення якості оцінок прогнозів;
3. застосування вибраних і запропонованих моделей до оцінювання прогнозів волатильності, виконання порівняльного аналізу.

Методика побудови моделі ГСП. Відома методика побудови моделі ГСП, яка складається з трьох кроків:

Крок 1: Побудувати просту модель низького порядку АРКС (або АР) для основної досліджуваної послідовності і обчислити додатковий ряд із квадратів похибок. Обчислити вибіркову (на основі значень елементів ряду) дисперсію залишків.

Крок 2: Обчислити вибіркову автокореляційну функцію квадратів залишків за виразом

$$\rho(s) = \frac{\sum_{k=s+1}^{N} \left[\hat{\varepsilon}^2(k) - \hat{\sigma}^2\right] \left[\hat{\varepsilon}^2(k-s) - \hat{\sigma}^2\right]}{\sum_{k=1}^{N} \left[\hat{\varepsilon}^2(k) - \hat{\sigma}^2\right]}, s = 0, 1, ..., N/3. \quad (1)$$

Крок 3: Для вибірок досить великої довжини стандартне відхилення для змінної $\rho(s)$ приблизно прирівнюють значенню $N^{1/2}$. Індивідуальні значення змінної $\rho(s)$, які суттєво відрізняються від нуля, вказують на наявність процесу УАРУГ. Для того щоб визначити, наскільки значення вибіркової АКФ є відмінними від нуля (тобто чи вони значимі в статистичному смислі), можна застосувати Q^- статистику Люнга-Бокса.

Якщо значення $\hat{\varepsilon}^2$ некорельовані між собою, то Q^- статистика має в асимптотиці розподіл χ^2 з n ступенями свободи. За нульову гіпотезу можна прийняти те, що значення $\hat{\varepsilon}^2$ некорельовані між собою. Відмова від цієї нульової гіпотези еквівалентна тому, що процес АРУГ чи УАРУГ в конкретному випадку наявний.

Виконання обчислюваних експериментів. Розглянемо конкретний практичний приклад моделювання гетероскедастичних процесів. У якості статистичних даних розглянемо ряд значень, які характеризують об'єм авіаперевезень. Результати побудови моделі та однокрокових прогнозів представлені у таблиці 1.

Табл. 1. Результати моделювання і оцінювання прогнозу для даних стосовно авіаперевезень

Тип моделі	Якість моделі			Якість прогнозу		
	R^2	SSE	DW	СКП	САПП	U
АРУГ(1)	0,269	1976,80	1,944	4,850	43106,25	0,507
АРУГ(4)	0,234	2057,91	1,990	5,315	44019,99	0,582
УАРУГ(6,2)_м	0,998	4,684	1,667	2,450	179,61	0,138
УАРУГ(6,2)_д	0,985	0,455	1,932	0,429	28,47	0,120
УАРУГ(8,10)_м	0,999	2,878	1,981	1,279	50,93	0,068
УАРУГ(8,10)_д	0,987	0,371	2,009	0,319	20,29	0,091
ЕУАРУГ(1,6,10)	0,988	0,120	1,895	0,063	34,54	0,065

Примітка: моделі, що позначаються буквою «м» (УАРУГ(6,2)_м), побудовані на основі умовного математичного сподівання; моделі, що позначаються буквою «д» (УАРУГ(6,2)_д), побудовані на основі умовної дисперсії процесу.

Таким чином, найкращий прогноз дає модель УАРУГ(8,10), побудована за допомогою умовної дисперсії процесу.

Висновки. У статті розглянуто деякі популярні моделі гетероскедастичних процесів, а саме показана основна структура моделей АР, АРУГ, УАРУГ та ЕУАРУГ. Також зроблено огляд інших відомих моделей ГСП. Наводиться методика побудови моделей ГСП. Проведені експерименти та аналіз вибраних моделей ГСП за основними показниками, виконано порівняння якості прогнозу. У першому випадку, де проводилось моделювання даних стосовно авіаперевезень, модель УАРУГ(8,10), побудована з урахуванням умовної дисперсії процесу, показала найкращий результат. А в другому випадку дані описують ціноутворення на біржі, найкращий прогноз дає модель ЕУАРУГ(1,6,5). Таким чином, можна сказати, що кожний окремий випадок потребує детального вивчення та індивідуального підбору моделі залежно від походження даних.

Література. **1.** Xekalaki E., Degiannakis S. ARCH models for financial applications. – New York: John Wiey & Sons Ltd., 2010. – 550 p. **2.** Tsay R.S. Analysis of financial time series. New York: John Wiley & Sons, Inc., 2010, 715 p. **3.** Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів. – К.: Політехніка, 2012. – 520 с. **4.** Engle F.R. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", Econometrica, vol. 50, no. 4, pp. 987–1007, 1982.

Бойко Р.С.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Комп'ютерне бачення: метод розпізнавання джерел вогню та ознаки появи вогню в відеокадрі

Пожежа була і залишається одним з найстрашніших лих в житті людини. На сьогоднішній день, навіть у найрозвиненіших країнах досить важко попередити виникнення пожеж, тому розробка системи детектування вогню є важливою, навіть необхідною, для запобігання пожежі. Фокус даної роботи направлено на охоронну діяльність, на вивчення і розробку способів виявлення полум'я в цифровому відеопотоці, який дозволив би швидко, точно і з низькими витратами обчислювальних ресурсів визначати наявність вогню на зображенні.

Для досягнення цієї мети були сформульовані наступні задачі:

- Проаналізувати існуючі алгоритми, підходи та методи для розпізнавання вогню в кадрі відеоряду.
- Створити алгоритм для розв'язку задачі розпізнавання джерел вогню.

Основним фактором вогню є його кольорові ознаки. Оскільки контрастність полум'я досить велика і його зображення є яскравим – колір відіграє принципово важливу роль при розпізнаванні. Отже в нашому алгоритмі використаємо комбінацію підходів, що дозволяють знайти маску вогню.

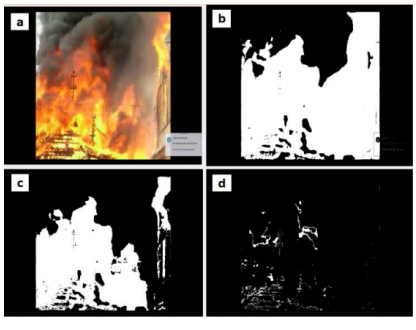


Рис. 1. Знахождення маски вогню:

Поступово накладуємо правила (1) на (2)–(4) і комбінуємо з (5) або (6)–(8) та отримуємо максимально відфільтровану маску вогню, де мінімізовано помилкове детектування схожих на полум'я обєктів. Процес знаходження маски вогню різними комбінаціями методів показано на Рис. 1, де: (a) вихідне зображення, (b) використання правил (1), (c) використання правил (1) разом з (2)–(4), (d) використання правил (5)–(8) [1,2].

$$R(x,y) > R_{mean} = \frac{1}{N}\sum_{i=1}^{N} R(x_i,y_i) \quad R(x,y) > G(x,y) > B(x,y) \tag{1}$$

$$0.25 <= \frac{G(x,y)}{R(x,y)+1} <= 0.65 \tag{2}$$

$$0.05 <= \frac{B(x,y)}{R(x,y)+1} <= 0.45 \tag{3}$$

$$0.20 <= \frac{B(x,y)}{G(x,y)+1} <= 0.60, \tag{4}$$

$$Y(x,y) > Cb(x,y) \quad Cr(x,y) > Cb(x,y) \tag{5}$$

$$Y(x,y) > Y_{mean} = \frac{1}{N}\sum_{i=1}^{N} Y(x_i,y_i) \tag{6}$$

$$Cb(x,y) < Cb_{mean} = \frac{1}{N}\sum_{i=1}^{N} Cb(x_i,y_i) \tag{7}$$

$$Cr(x,y) > Cr_{mean} = \frac{1}{N}\sum_{i=1}^{N} Cr(x_i,y_i) \tag{8}$$

Наступним кроком є використання часових вікон при обробці кольорів полум'я. Робимо перевірку частоти появи вогню в послідовності кадрів методом усереднення за часом ймовірності того, що даний піксель належить області вогню на зображенні за формулою (9). Потім за умовою (10) залишаємо лише пікселі, що перевищили поріг значень $\delta_1 = 0,6$.

$$P(x,y) = \frac{\sum_{t=1}^{N} C(I_t(x,y))}{n}, \tag{9}$$

$$Color(x,y) = \begin{cases} 1, & P(x,y) > \delta_1, \\ 0, & P(x,y) \leqslant \delta_1. \end{cases} \tag{10}$$

Для коректного розпізнавання подальшим кроком є використання ознак, заснованих на динамічних змінах сцени в поєднанні з ознаками, що визначають колір полум'я. Основна концепція в тому, що інтенсивність пікселів вогню поступово змінюється в часі, тому між двома послідовними кадрами буде простежуватись певна різниця інтенсивності. Для цього використовуємо усереднення модуля різниці інтенсивності фонових пікселів за формулою (11) [3].

$$D(x,y) = \frac{\sum_{t=1}^{N} |J(I_t(x,y) - I_{t-1}(x,y))|}{n-1}, \quad N(x,y) = \frac{\sum_{x,y,Color(x,y)=0} D(x,y)}{M}. \tag{11}$$

Внаслідок фільтрування ми втратили велику частину пікселів, що насправді є пікселями вогню. Проте це було необхідним кроком, щоб уникнути помилкового розпізнавання обєктів, що лише схожі на вогонь, наприклад мають колір вогню та рухаються. Для відновлення втрачених елементів використаємо процес дилатації. За знайденою маскою, що є результатом фільтрування всіма методами, рекурсивно будемо розглядати сусідні від пікселів вогню пікселі, і використовуючи лише їх перевірку на колір, доповнимо маску новими пікселями, що також є частиною полум'я.

Рис. 2. Остаточний варіант роботи алгоритму після процесу дилатації

Висновки. Після проведення огляду різноманітних методів, були обрані підходи в порядку спадання передбачуваної ефективності. Запропоновано алгоритм розпізнавання вогню, який використовує підходи засновані на аналізі кольорових, динамічних, просторових та часових характеристик для знаходження маски вогню. Точність роботи алгоритму складає 93–95%.

Література. **1.** Yu-Chiang Li, & Wei-Cheng Wu. (2012). Sequential Pattern Technology for Visual Fire Detection. Journal of electronic science and technology, 10, 276–280. http://ccs.asia.edu.tw/ezfiles/2/1002/img/370/1203-14.pdf. **2.** Peiyi Shen, Liang Zhang, Juan Song, Hu Xu, Lingfei Zhou, Wei Wei, Wenzheng Zhang. (2013). DM642-Based Fire Detection in Video Sequences Using Statistical Color Model. International Journal of Digital Content Technology and its Applications, 7, 669–678. http://globalcis.org/jdcta/ppl/JDCTA2553PPL.pdf. **3.** Che-Bin Liu. (2006). Linear models of dynamic 2d shape and appearance. Globally-coordinated locally-linear modeling of multi-dimensional data (pp. 80–84). Urbana, Illinois: National Taiwan University. http://vision.ai.illinois.edu/publications/vision%20based%20fire%20detection.pdf.

Вельбицкий И.В.
Фонд Глушкова, Киев, Украина

Программирование без языков программирования (новая графическая полиглот-концепция)

Впервые после 1947 года предлагается использовать более простую и математически строгую концепцию программирования графами, нагруженными только по горизонтальным дугам символами, выражениями и функциями из элементарной математики. Такой граф является полиглотом, имеет стандарт ISO 8631/1989 и только один эффективно используется на всем жизненном цикле процесса разработки и эксплуатации программ. Традиционные языки программирования не используются, но сохраняются их библиотеки. Любая программа из библиотеки может быть автоматически переведена в единую графическую оболочку как эталон. Новая концепция имеет гораздо лучшие характеристики по наглядности, простоте, компактности до 100+ раз больше и скорости ввода программ в компьютер до 100+ раз быстрее. Для ввода используется только существующая клавиатура и/или мышь. Существенно упрощаются, улучшаются и ускоряются процессы безошибочного проектирования алгоритмов, программ, данных, сетевых графиков, доказательство их правильности, самодокументирование и документирование мотивации принимаемых решений. Получаемые в результате программы более эффективны. Чем больше и логически сложнее программный проект, тем больше эффект применения новой концепции, которая настолько проста, что позволяет программировать ВСЕМ, а не только программистам. Это делает программирование элементом всеобщей грамотности и культуры общества. Нам не известен аналог новой концепции и ее серьезная критика. В докладе приведена *история [1–3] возникновения и проверки [3–8]* новой концепции программирования, описание ее *сути, преимуществ, реализованной* графической среды и *перспективы* ее использования.

В существующей, неизменной с 1947года, концепции программирования все традиционные машинно-ориентированные и НЕестественные для человека операторы типа: **if-then, else, for, goto ...,** метки, блочные скобки типа **begin-end,...,** отступы (лесенки), большинство знаков препинания и т.д. **все! *исключены*** из новой концепции. Они устарели. Для человека их слишком много – *больше половины (70–80% мусора)* в существуюших текстах программ. Они сложны, эмпиричны (строго не определены), маломощны и обеспечивают только примитивно-ремесленническую технологию работы. Для *нейтрализации* их недостатков человек тратит много усилий, создавая огромное число новых языков, методов и сред программирования, которые “якобы упрощают”, а на самом деле делают программирование еще более *сложным, запутанным и недоступным* для всех.

В новой концепции *единые* и очень *простые* графические принципы используются для записи и алгоритмов, и программ, и данных, и сетевых графиков *на всем жизненном цикле* их разработки и эксплуатации. В основе новых принципов лежит только ОДНА АКСИОМА – горизонтальная дуга со стрелкой влево или вправо, у которой сверху записано Условие, а снизу – *Действия*, которые выполняются, если *Условие* “истинно”. После выполнения всех Действий осуществляется переход по стрелке к следующей дуге. Из вершины графа может исходить любое число дуг влево и/или вправо. Ввод каждой горизонтальной дуги графа делается одним нажатием мыши или клавиатуры или пальца по сенсорному экрану. При одном нажатии может быть введено сразу *несколько* новых горизонтальных дуг. Для записи Условий и Действий используются любые *естественные* языки: английский, русский, китайский и т.д., символы, выражения и функции *элементарной* математики, соответствующие *фрагменты* из любых существующих языков программирования для записи только Условий и Действий и функции обращения к их Библиотекам. Такой граф имеет стандарт ISO 8631/1989 (в отличие от всех известных языков программирования), является полиглотом – понятным всем без больших (до больше 1000 страниц для, например, C++) определений и сливается с известными, *строгими* и *краткими* математическими принципами *представления* и *обработки* информации, которые

приближаются к принципам работы *мозга*. Граф имеет *имя* с параметрами или без них, что соответствует имени некоторой графической функции в математике. *Программа* в новой концепции задается любым числом взаимосвязанных по имени таких графов и существенно компактнее записи ее в традиционных языках программирования. Сейчас известна и приведена в докладе графическая программа, которая компактнее в 400 раз этой же программы в C++ и вводится в компьютер в 578 раз быстрее (содержит меньше мусора, чем C++). Новая концепция *универсальна*. Она позволяет записать любой алгоритм, любую программу, любым известным методом от структурного до глубокого интеллектуального анализа, обеспечивая кибербезопасность, и с использованием уже готовых библиотек, которые могут быть объединены в единую графическую *гипербиблиотеку*. Новая концепция впервые вводит в программирование *доказательный* и *безошибочный* стили, *самодокументирование* и документирование *мотивации принимаемых решений, 3D* и *многомерное* программирование. Позволяет **каждому** человеку легко, быстро и наглядно нарисовать логическую схему любой своей проблемы (из любой области) и решить ее самому или на компьютере. Принципы новой концепции *так просты* и *естественны* для человека, что программирование впервые *доступно всем* (а не только программистам) и становится элементом *всеобщей грамотности* и *культуры* общества. Приведенного выше описания, например, достаточно, чтобы начать программировать в графах новой концепции. Новая концепция прошла всестороннее обсуждение и экспертизу на многих международных конференциях [4–7], а новая, реализованная (лабораторный вариант) графическая среда программирования опробована многими пользователями [8]. Ниже (от неизвестных нам авторов) приведены их *независимые мнения*. 1) Интересная, сильная полиглот-концепция программирования, которая проще существующей и может быть использована многими людьми. 2) Новый математический инструмент эффективный для анализа, синтеза и обучения программированию. 3) Увеличивает долголетие и эволюционное использование программ применительно к условиям их эксплуатации. 4) Повышенная защищенность от ошибок и высокая отлаженность готовых графических программ. 5) После двух недель рисования R-схем в новой среде программирования я решил задачу, которую без них не решил бы никогда... 6) Новая концепция – это полезный и прекрасный инструмент не только в программировании, но и ВЕЗДЕ, где надо найти решение любой логически сложной (запутанной) проблемы. По сравнению с известными Flow-chart, UML и др. R-схемы имеют заметные преимущества. Они выполнимы на компьютере на всем жизненном цикле любых работ, более наглядны, компактны и не загромождены лишними деталями (фигурами). Подходят для визуализации нелинейных алгоритмов (например, классов в C++) или структур данных. Поэтому мне совершенно не понятно, почему этот инструмент не пользуется *бешеной* популярностью. 7) Совершенно шикарная система, которая позволила мне, не будучи программистом, использовать её для составления различных диаграмм и решения технологических, управленческих задач. Всё интуитивно понятно и просто в использовании. 8) Я в восторге от использования R-схем для преподавания на третьем курсе спецкурса по “Элементам структурной технологии программирования”. Неожиданно даже для себя убедился в замечательной возможности R-схем для структурированного хранения больших объемов данных. **Каждый может попробовать R-схемы и *убедить себя* в сказанном другими**.

Литература. **1.** М. Глушков, И.В. Вельбицкий, “Технология программирования и проблемы ее автоматизации”, УСИМ, №6, С. 75–93, 1976. **2.** И.В. Вельбицкий, “Технология программирования”. Техника, Украина, 279 с., 1984. **3.** Сергеев В.Г. – “Главный конструктор систем управления ракет и космических комплексов”, Украина, Харьков, С. 285–292, 2014. **4.** McHenry, W.K. Technology: A soviet visual programming. Journal of Visual Languages and Computing, 1990; V.1, № 2. **5.** Velbitskiy, I.V. Graphical Programming and Program Correctness Proof, IEEE: 10.1109/ CSIT-13.6710368, 2013; P. 85–89. **6.** Velbitskiy, I.V. New Graphic Poliglot Concept of Programming, IEEE 40th Annual COMPSAC 2016, USA, P. 139–144, **7.** Velbitskiy, I.V. Ukrainian Programming Concept without Programming languages. Английская академия наук, в печати, 2017, 15 с. **8.** Письма Вельбицкому в интернете, 2015–2017.

Виклюк Я.І., Гусак О.М.
ПВНЗ "Буковинський університет", Чернівці, Україна

Інформаційна технологія раннього виявлення лісових пожеж

У зв'язку з загостренням регіональних і глобальних екологічних проблем виникла необхідність інформаційного забезпечення сучасних екологічних програм, зокрема тих, в яких основна увага приділяється ранньому оперативному виявленню епіцентрів лісових пожеж. Актуальність досліджень в цьому напрямі обумовлюється тим, що не дивлячись на зусилля вчених-пірологів усього світу, лісові пожежі залишаються одним з найвпливовіших факторів негативного впливу на довкілля і продовжують завдавати величезних збитків. Тому створення нових та підвищення ефективності існуючих інформаційних технологій раннього виявлення епіцентрів лісових пожеж є вельми актуальною задачею.

Основною проблемою прийняття рішень на основі даних космічних та аерофотознімків є те, що часто неможливо безпомилково виокремити джерело запалювання, визначити масштаб пожежі, оцінити площу вигорілих ділянок [1]. Це пов'язано з присутністю шумів і завад, обумовлених складними умовами отримання та недосконалістю систем передачі фото- та відеоінформації.

Одним з можливих рішень вказаної проблеми може стати технологія, заснована на використанні можливостей сучасних безпілотних літальних апаратів та мобільних пристроїв зв'язку. На квадрокоптери покладаються завдання повітряної розвідки епіцентрів лісових пожеж та перевірки стану термоточок (теплових аномалій) визначених заздалегідь під час планового обстеження лісу. При підвищенні рівня пожежної небезпеки відбувається запуск та організація польоту за пожежонебезпечним маршрутом. Камери квадрокоптера знімають відео, яке надходить на вхід інформаційної системи. Результат обробки відображається на екрані мобільного пристрою. Достовірність інформації забезпечується оптимальною комбінацією методів обробки цифрових відео зображень. Суть цих методів полягає в наступному:

1. Вхідний відео потік проходить попередню обробку, в ході якої послідовно відбуваються налаштування контрасту та вейвлет перетворення.
2. Формування навчальної вибірки, з подальшим навчанням нейронної мережі.
3. Кластеризація, яка дозволяє розділити об'єкти на класи природнього та чужорідного походження [2].

Запропонована інформаційна технологія дозволяє в автоматичному режимі виокремлювати чужорідні кластери, проводити донавчання нейронної мережі та формувати відеопотік в режимі реального часу за рахунок технології паралельних обчислень. Результатом є створення підсумкового відеопотоку даних, що підвищує ймовірність виявлення епіцентру пожежі на фоні гілок і листової підстилки.

Висновки. Розроблена інформаційна технологія надає можливість:

1. Забезпечити мобільність та оперативність пошуку епіцентрів лісових пожеж, а також достовірність інформації про наявність епіцентрів лісових пожеж на відеосюжетах, відзнятих за допомогою автоматичних безпілотних літальних апаратів.
2. Удосконалити метод обробки відеоінформації, з врахуванням специфіки пошуку епіцентрів лісових пожеж та можливостей сучасних засобів зв'язку та безпілотних літальних апаратів.
3. Забезпечити підтримку прийняття рішень при гасінні лісових пожеж, оскільки дозволяє здійснювати ефективне протипожежне патрулювання, вчасне визначення теплових аномалій, ймовірність пожежі в яких найбільша, можливість вчасного реагування лісопожежних сил та корегування дій в ході боротьби з вогнем.

Література. **1.** Стругайло В.В. Обзор методов фильтрации и сегментации цифровых изображений. / Наука и образование // МГТУ им. Баумана. – №77. – С. 270–281. **2.** Гонсалес, Р. Цифровая обработка изображений (Digital Image Processing) [Текст] / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2012. – 1104 с.

Ворона М.В., Лазаренко И.С.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Механизмы повышения успешности управления проектами по созданию WEB-сайтов

На современном этапе при создании WEB-сайтов, как и для программного обеспечения в целом, применяются подходы, основанные на использовании методов управления проектами и системного анализа. Успешным считается программный проект, который уложился в установленный бюджет, был выполнен в срок и вышел на рынок с надлежащим уровнем качества. Статистика успешности программных продуктов по данным The Standish Group International [1] показывает, что успешность малых ИТ-проеков (срок реализации которых не более, чем полгода, бюджет не превышает 1 млн долл.) составляет 76%, в то время как больших – всего лишь порядка 10% (под большими проектами понимаются проекты, которые реализуются от полугода и бюджет которых превышает 10 млн долл). Такая тенденция достаточно стабильна и сохраняется уже в течении 7 последних лет. Сравнительно с малыми проектами, в больших проектах в 10 раз чаще наступают рисковые события, влияющие на срок разработки, бюджет и качество [1]. Таким образом, разработка механизмов управления проектами по созданию WEB-сайтов, которые позволят уменьшить процент рисковых событий, влияющих на сроки создания проекта, а также повысить вероятность успешности такого программного проекта является актуальной.

Изначально программные проекты плохо поддаются планированию и оценке, что обусловлено влиянием многочисленных факторов: решаемой задачей, составом и квалификацией команды проекта, взаимодействием с заказчиком, выбранной методологией проектирования программного продукта и пр. При этом влияние факторов чаще всего носит динамический характер. В таком случае классические методы управления проектами перестают работать.

В качестве нового подхода к процессу управления разработки WEB-сайта можно применить модель конвейерного производства товаров в промышленности. В своем классическом представлении конвейерное производство применяется к востребованным товарам массового использования, сегодня таковыми являются и WEB-сайты в информационном секторе.

Сама концепция данного подхода может применятся только к определенной нише данного продукта, которая должна быть определена при стадии разработки организации работы компании. Для максимальной эффективности работы целесообразно выбирать среднеценовой сегмент. Также должна быть четко разработана юридическая документация с жестко прописанной ответственностью заказчика для минимизации рисков срыва процесса работы всей системы и возникновения задержек, не связанных с технологическим процессом.

Сам процесс разработки продукта делится на 8 этапов: формирование задания на проект, создание прототипа сайта, производство (распараллеливаем на три линии – дизайн сайта, наполнение сайта, программирование), согласование, тестирование, запуск проекта, повторное тестирование, сдача проекта. Всем этим этапам предшествует процесс продажи сайта клиенту, подходящему под критерии данной модели.

На этапе формирования задания менеджером проекта составляется техническое задание, заказчиком предоставляются все необходимые материалы (текстовые и графические), после чего менеджером и дизайнером составляется прототип сайта. Прототип сайта представляет из себя визуальный скелет сайта с окончательно расположенными всеми функциональными элементами, который утверждается заказчиком и не может быть изменен на последующих стадиях.

На следующем этапе задачи распараллеливаются для экономии временных ресурсов. На этапе программирования разработчик формирует всю функциональную часть сайта, контент-менеджер организовывает наполнение сайта согласно задания, как результат, создается три варианта дизайна. Процесс разработки дизайна имеет свою специфику, в специальную систему

выкладывается прототип сайта, задание и стоимость работы дизайнера по данному проекту; первые 3 дизайнера, которые взли этот заказ берутся за его исполнение. После выдачи готовых решений, заказчик имеет возможность выбрать один из трех предложенных дизайнов. Исполнитель выбранной работы получает 70% от стоимости работ, остальные два – по 15%.

После выбора дизайна происходит его верстка и интеграция в функциональную часть, что делается программистом. Далее проект тестируется и запускается в работу, после чего происходит его закрытие (сдача заказчику) с учетом всех формальных процедур.

Конвейерное производство позволяет исключить риск возникновения спорных ситуаций и откат проекта на предыдущие стадии без денежных компенсаций, поскольку каждая стадия оплачивается клиентом отдельно. При необходимости внесения изменений клиентом на предыдущей стадии проекта она оплачивается еще раз.

Такой подход к организации процесса представляет собой симбиоз системы управления “Водопад”, так как все этапы четко последовательны и распределены, и гибкой методологии управления проектами. Это выражается в том, что в системе работы компании постоянно ведется мониторинг и визуализация всех стадий всех проектов, а в случае попадания одного проекта в зону риска задержки сроков – он ставится в приоритет и на него перебрасываются ресурсы других менее “горящих” проектов.

Следует отметить, что системы управления проектами по разработке ПО (WEB-сайтов) относятся к динамическим системам с нечетко выраженными состояниями, где в силу воздействие внутренних или внешних факторов информация об их состоянии (этапах выполнения, задачах и подзадачах), представленная значениями атрибутов, является нечетко выраженной. В этих системах информация о состоянии объекта управления оказывается размытой из-за субъективности оценок состояния [2].

Согласно [2] существует решение, которое основывается на возможности корректировки значений атрибутов с учетом природы источника искажений и избыточной информации о состоянии системы. При этом механизм управления динамической системой с нечетко выраженными состояниями будет иметь вид, представленный на рисунке 1.

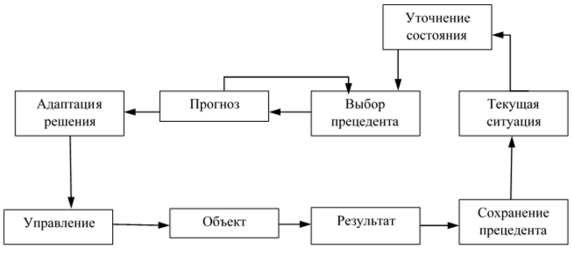


Рис. 1. Механизм адаптивного управления по прецедентам с корректировкой состояния системы

Данный подход к управлению проектами по разработке WEB-сайтов, создаваемых на основе конвейерной модели, является механизмом повышения эффективности их управления, позволяющим учитывать и корректировать ход выполнения проекта с учетом возможных отклонений объемов работ проектов и ресурсов исполнителей от плановых, которые могут возникнуть из-за возможного источника искажений и избыточной информации о состоянии системы.

Литература. **1.** Николаенко В.С. Разработка принципов управления ИТ-проектом [Электронный ресурс] // Научная электронная библиотека «Киберленинка» [сайт]. URL: http://cyberleninka.ru/article/n/razrabotka-printsipov-upravleniya-it-proektom-1. **2.** Гайда А.Ю. Механизмы эффективного управления проектами в организационных системах с нечетко выраженными состояниями / А.Ю. Гайда, Т.А. Фарионова, М.В. Ворона // Управління розвитком складних систем. – 2016. – № 28. – С. 116–122.

Гнатенко В.Ю., Ситников В.С., Ступень П.В.
Одесский национальный политехнический университет, Одесса, Украина

Метод получения полиномов Фаулхабера для интерполяции сигналов

При описании сигнала по его промежуточным значениям необходимо определить функцию с необходимой степенью точности.

В информационных системах обработки сигналов подобные функции реализуются на основе интерполяционных полиномов. Существующие методы рассчета интерполяционных полиномов позволяют получить либо приближенное решение, либо являются сложными в реализации, поэтому разработка методов интерполяции является актуальной задачей.

Рассмотрим график $y(x) = x^2$ – полупараболу, ограниченную на участке $x \in [0, n]$ на Рис. 1. Если рассматривать $y_1(x) = x$ как сумму нечетной на участке $[0, n]$ функции $y_{11}(x) = x - \frac{n}{2}$ и постоянной составляющей $y_{10}(x) = n$, $y_1(x) = y_{10}(x) + y_{11}(x)$, то можно увидеть в $p_1(x)$ результат интегрирования $y_{10}(x)$, а в $p_2(x)$ (полная парабола) – результат интегрирования $y_{11}(x)$.

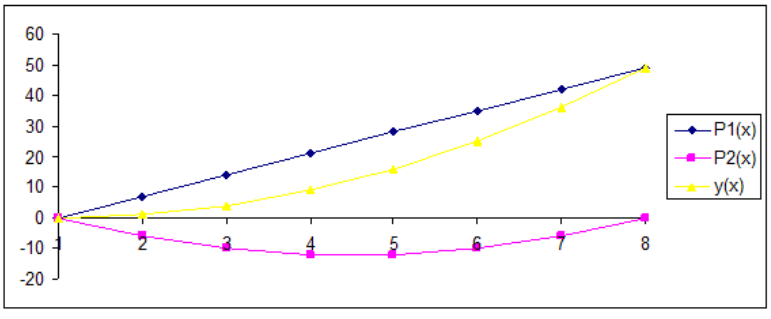


Рис. 1

Рассмотренный пример можно обобщить для любого монома k-й степени $x^k, x \in [0, n]$, который можно представить суммой полиномов, назовем их полиномами Фаулхабера, причем вклад в график $x^k, x \in [0, n]$ каждого из них уменьшается с увеличением его порядка.

Получение полиномов Фаулхабера. $p_i(x)$ – полином Фаулхабера i-го порядка.

$$\begin{gathered} \mathrm{p}_0(x) = 1, \\ \mathrm{p}_1(x) = x - 0, \\ \mathrm{p}_2(x) = (x-0)(x-n), \\ \mathrm{p}_3(x) = (x-0)(x-n)(x-\tfrac{n}{2}),\ p_i(x) = p_{i-2}(x)(x-x_1)(x-x_2), \\ |\mathrm{p}_{i-2}(x_1)| = |p_{i-2}(x_1)| = |p_{i-2}(x_2)| = max(|p_{i-2}(x)|), x \in [0, n]. \end{gathered} \tag{1}$$

Каждый полином Фаулхабера k-го порядка получается из суммы ряда значений полинома Фаулхабера $k-1$-го порядка по формуле Фаулхабера [1] (Формула $\int x^k dx = \frac{x^{k+1}}{k+1} + C$ из таблицы интегралов получена из формулы Фаулхабера), путем вычитания из него полинов Фаулхабера меньшего порядка с соответствующими коэффициентами.

После получения суммы ряда значений полинома Фаулхабера $k-1$-го порядка $\sum_{i=0}^{x} p_{k-1}(i)$ вычитаем из нее $\frac{1}{2}p_{k-1}(x)$, если $k - 1 > 2$.

Каждый полином Фаулхабера меньшего порядка, исключая уже вычтенный, начиная с полинома Фаулхабера минимального порядка, вычитается из k-го с соответствующим коэффициентом. Для получения коэффициента находим максимум вычитаемого полинома Фаулхабера, соответствующее ему значение x, и делим на максимум соответствующее значение k-го полином Фаулхабера (полученный коэффициент – это частное от деления полиномов, которые делятся без остатка), чтобы полностью исключить вычитаемый полином. По сути, находимые

максимумы – это корни k-го полинома Фаулхабера. Для четных и нечетных полиномов Фаулхабера, кроме краев отрезка $[0, n]$, корни не повторяются. Полученный в итоге результат соответствует правилу получения полиномов Фаулхабера (1).

С учетом симметрии полиномов Фаулхабера коэффициенты для полиномов можно получить из двух систем линейных уравнений с матрицами треугольного вида и полиномы меньшего порядка вычисляются независимо от полиномов более высокого порядка на меньшем количестве точек графика.

Рассмотрим получение коэффициентов полиномов Фаулхабера для ряда данных из 8 элементов $Y = \begin{pmatrix} y_0 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 \end{pmatrix}$. Y можно представить в виде суммы четной и нечетной частей $Y = Y_0 + Y_1$.

$Y_0 = \begin{pmatrix} \frac{y_0+y_7}{2} & \frac{y_1+y_6}{2} & \frac{y_2+y_5}{2} & \frac{y_3+y_4}{2} \end{pmatrix}$

$Y_1 = \begin{pmatrix} \frac{y_0-y_7}{2} & \frac{y_1-y_6}{2} & \frac{y_2-y_5}{2} & \frac{y_3-y_4}{2} \end{pmatrix}$

Матрицы строим с учетом содержания в $p_1(x)$ четной и нечетной частей.

$p_1(x) = p_{01}(x) + p_{11}(x)$

$p_{11}(x) = x - \frac{n}{2}, p_{10}(x) = \frac{n}{2}$.

Линейные уравнения для вычисления четных и нечетных полиномов Фаулхабера – (2), (3). После решения этих линейных уравнений легко получить коээфициенты для $p_1(x)$ и $p_0(x)$.

$$\begin{pmatrix} 0 & 0 & -12 & 1 \\ 80 & -20 & -10 & 1 \\ 0 & -36 & -6 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_6(x) \\ p_4(x) \\ p_2(x) \\ p_0 + p_{10} \end{pmatrix} = \begin{pmatrix} \frac{y_3+y_4}{2} \\ \frac{y_2+y_5}{2} \\ \frac{y_1+y_6}{2} \\ \frac{y_0+y_7}{2} \end{pmatrix}. \tag{2}$$

$$\begin{pmatrix} 72 & -36 & 6 & -0.5 \\ 0 & -60 & 15 & -1.5 \\ 0 & 0 & 15 & -2.5 \\ 0 & 0 & 0 & -3.5 \end{pmatrix} \begin{pmatrix} p_7(x) \\ p_5(x) \\ p_3(x) \\ p_{11}(x) \end{pmatrix} = \begin{pmatrix} \frac{y_3-y_4}{2} \\ \frac{y_2-y_5}{2} \\ \frac{y_1-y_6}{2} \\ \frac{y_0-y_7}{2} \end{pmatrix}. \tag{3}$$

Чтобы получить нормальный треугольный вид матрицы, нужно поменять местами строки в уравнении (2). Получим уравнение (4).

$$\begin{pmatrix} 80 & -20 & -10 & 1 \\ 0 & -36 & -6 & 1 \\ 0 & 0 & -12 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_6(x) \\ p_4(x) \\ p_2(x) \\ p_0 + p_{10} \end{pmatrix} = \begin{pmatrix} \frac{y_2+y_5}{2} \\ \frac{y_1+y_6}{2} \\ \frac{y_3+y_4}{2} \\ \frac{y_0+y_7}{2} \end{pmatrix}. \tag{4}$$

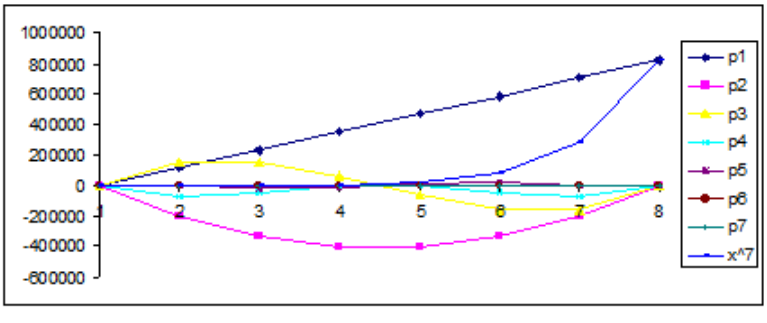


Рис. 2. График монома $x^7, x \in [0, n]$ и составляющих его полиномов Фаулхабера

Применение. Полиномы Фаулхабера неортогональны, симметричны относительно середины отрезка $\in [0, n]$.

Применимы для интерполяционных целей, выбора узлов интерполяции.

Литература. **1.** APA, en.wikipedia.org/wiki/Faulhaber's_formula.

Губар Б.Д., Олійник Ю.О.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Архітектура рекомендаційної системи з вибору продуктів харчування

Рекомендаційні системи (РС) – аналізують інтереси користувачів і намагаються передбачити, що саме найбільш зацікавить користувача в даний момент часу. РС вже зараз широко розповсюджені в електронній комерції, а компанії, що продають товари через інтернет, мають від 25 до майже 75% прибутку, за рахунок продажу рекомендованих товарів [1]. Переваги РС можна також використовувати і в офлайн рітейлі для продажу продуктів харчування, але точність рекомендацій в звичайних магазинах, менше, оскільки для аналізу користувачів є тільки дані про продажі.

На відміну від електронної комерції, де продаються вже готові товари, продукти харчування можуть не використовуватись самі по собі, а складатися у певні страви. Тому при обробці покупок необхідно враховувати не лише певний окремий продукт, а й інші продукти в чеку. Тому для використання РС для рекомендації продуктів необхідно додати в архітектуру РС додаткові модулі (рис. 1): модуль розбиття чеку, який виконує розбиття продуктів з чеку на страви, порівнює продукти, куплені користувачем з базою страв і визначає, які страви може приготувати користувач.

На відміну від загальних схем онлайн продажу, РС проводить аналіз закупок користувачів на основі страв, а не продуктів, що дозволяє формувати рекомендації до набору продуктів, а не з певного окремого продукту.

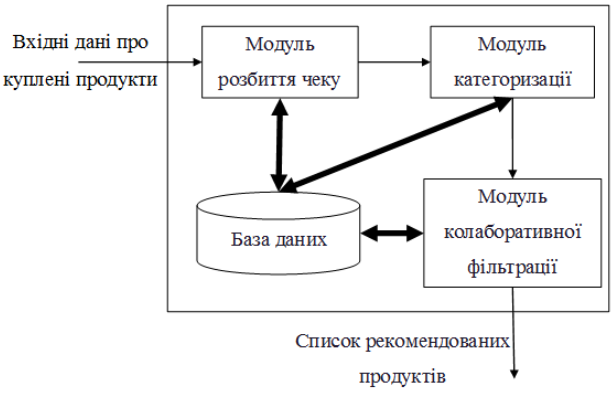


Рис. 1. Схема структурна РС з вибору продуктів харчування

Модуль категоризації встановлює відповідність між користувачами та категоріями страв, що вони готують. Категоризація в РС відомостей про товар відноситься до розповсюдженого методу побудови прогнозу – фільтрацією за змістом.

Наступним етапом є генерація рекомендацій методом колаборативної фільтрації, який використовує відомі переваги (оцінки) групи користувачів для прогнозування невідомих переваг іншого користувача. Основне припущення колаборативної фільтрації полягає в наступному: ті, хто однаково оцінювали будь-які предмети в минулому, схильні давати близькі оцінки інших предметів і в майбутньому [2]. Комбінація методів колаборативної фільтрації та фільтрації за змістом дозволяє вирішити одразу декілька проблем: проблему «холодного старту», проблему «розрідженості даних» та підвищити точність рекомендацій. При додаванні нового товару в систему, його одразу можна віднести до певної категорії і включати в рекомендації. З іншого боку категорії зменшують кількість об'єктів для пошуку рекомендацій, що зменшує навантаження на систему.

Висновки. Запропонована РС дозволяє надавати рекомендації з вибору продуктів харчування, враховуючи контекст покупок, збільшити продажі офлайн рітейлерів, а також підвищити якість обслуговування покупців

Література. **1.** Ian MacKenzie, Chris Meyer and Steve Noble, How retailers can keep up with consumers, 2013 www.mckinsey.com/industries/retail/our-insights/ how-retailers-can-keep-up-with-consumers. **2.** Мазурік О.Ю., Покращення результатів роботи рекомендаційних систем за допомогою алгоритму SVD, 2015.

***Губарєв В.Ф.*[1], *Комісаренко В.І.*[2]**
[1]*Інститут космічних досліджень НАН і ДКА України, Київ, Україна;* [2]*КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна*

Визначення параметрів відносного руху космічних апаратів на основі аналізу зображень за допомогою інформаційних ознак

Вступ. Для задачі зближення і стикування із некооперуючим космічним апаратом (НКА) необхідне використання систем технічного зору для визначення параметрів взаємного положення і орієнтації. Система технічного зору у реальному часі надсилає зображення на вхід алгоритму, на виході якого повинні отримати параметри орієнтації і положення, які потім використовуються для керування космічним апаратом.

Основні припущення. Припускається, що алгоритм розпочинає роботу, коли НКА знаходиться на достатньо близькій відстані і в такому положенні, що його видно на зображеннях. Також відома трьохвимірна графічна модель НКА.

Основна частина. Шляхом моделювання формується вибірка із штучних зображень, які покривають усі можливі значення параметрів положення і орієнтації із певним кроком дискретизації. Крок дискретизації визначається із узгодження точності алгоритму, об'єму пам'яті і обчислювальних ресурсів. Підхід полягає у виборі найближчого за певним критерієм зображення із вибірки і ототожнення його параметрів орієнтації і положення із шуканими параметрами. Кожне зображення будемо характеризувати набором інформаційних ознак – певних властивостей, що характеризують зображення [1]. Усі ці властивості можна представити у вигляді чисел. Зберігання замість матриці пікселей лише деякого набору чисел дозволяє зекономити значну частину ресурсів. Ефективно обрана структура даних дозволяє виконувати прохід по вибірці за логарифмічний від розміру вибірки час, що дозволяє алгоритму функціонувати у режимі реального часу. Головні вимоги, які ставляться перед інформаційними ознаками, наступні: між зображенням і значеннями ознак має бути взаємно однозначна відповідність, інформаційні ознаки мають змінюватись при різних положеннях орієнтації і положення, при чому при невеликих змінах параметрів значення ознак також має змінюватись незначним чином. Для даної задачі важливо, щоб інформаційні ознаки були інваріантними до зовнішніх умов, до яких перш за все відноситься освітлення [2]. Тому більшість відомих ознак, пов'язаних із кольором зображення, не можна застосовувати до цієї задачі. Окремо варто виділити ознаки, пов'язані із конструктивними особливостями графічної моделі. Зазвичай вони є найбільш інформативними, однак без знання моделі, у загальному випадку, їх виділити не можливо. Виділяються інформаційні ознаки із усіх зображень набору. Можливі неточності відбору єдиного зображення набору пов'язані із похибками дискретизації. Задля цього завчасно обчислюються максимально можливі відхилення значень інформаційних ознак зображень, які внаслідок кроку дискретизації вважаться однаковими, тобто із рівними параметрами положення і орієнтації. Тоді кожне зображення із набору характеризується деяким околом значень інформаційних ознак, куди мають потрапляти ті зображення, які мають ті самі, із точністю до кроку дискретизації, параметри орієнтації і положення. Вибір інформаційних ознак проводиться таким чином, щоб у описаний окіл потрапляли лише зображення із тими ж параметрами орієнтації і положення. Останній факт перевіряється експериментальним шляхом і отримується після значної кількості спроб комбінування різнотипних інформаційних ознак.

Література. **1.** Васильева Н.С. Методы поиска изображений по содержанию. / Программирование. 2009. №3. С. 51–80. **2.** Губарев В.Ф. и др. Использование систем технического зрения для определения параметров относительного движения космических аппаратов / Боюн В.П., Мельничук С.В., Сальников Н.Н., Симаков В.А., Комисаренко В.И., Годунок Л.А., Деркач С.В., Добровольский В.Ю., Матвиенко С.А. // Проблемы управления и информатики, 2016, № 6, С. 103–119.

Джамалов А.Т., Рагимов Р.М.
Национальное аэрокосмическое агентство Азербайджана, Баку, Азербайджан

Аэрокосмический мониторинг обеспечения экологической безопасности транспортировки энергоносителей по магистральным трубопроводам

In the given article we give the methodical basis of using the various types of space imaging for solving the problems of diagnostics of the objects of geo-technical systems for providing the ecologically safe transportation of energy carriers and timely revealing of potentially dangerous areas.

When carrying out diagnostics of objects of geotechnical systems, the best results are obtained only by combining the land based complex diagnostics with the diagnostics based on interpretation of the space images of meter and sub-meter spatial resolution. It is shown that the high authenticity during the evaluation of condition of geotechnical systems is reached on the basis of information obtained from the satellites Ikonos-2, Quick Bird-2 and Resource ФІМ.

We also give the analysis of possibilities of space diagnostics of the condition of geotechnical systems with the usage sensors of high and extra-high spatial resolution OSA, BHRC-60, KFA-1000.

In the result of theoretical researches and practical usage it was determined that the space diagnostics allows solving of the following issues:

- Determining the main pipe line routes and their true location;
- Revealing of the objects of geotechnical systems, their characteristic elements and the closest surrounding, which gives the possibility of correct evaluation of technical condition of the investigated object;
- Revealing and monitoring of the development of negative processes at the main pipeline laying route, caused by impact of surface and ground waters;
- Analysis of the landscape of investigated object's location, which allows determining the relative location of the objects, evaluate their mutual interrelation and analyze the influence of the difference in elevation at the object of diagnostics;
- Evaluation and diagnostics of the shape, geometry, and the size of the objects of geotechnical system;
- Revealing of the potentially dangerous parts of the investigated objects.

Finally it is shown that the space diagnostics of geo-technical systems based on interpretation of high resolution space images allows managing of important tasks, but there is a variety of unsolved problems, including the limitations of spectral ranges of the used sensors, which leads to limitations of the possibility of interpretation of the objects (for example, underground main pipelines).

Важную роль в обеспечении безопасности эксплуатации магистральных трубопроводов для транспортировки энергоносителей играет совершенствование систем диагностики и применение новых информационных технологий основанных на дешифровании космических снимков высокого разрешения.

В данной работе анализируется различные методы космической съемки для решения задач диагностики магистральных нефтегазопроводов с целью оценки их состояния и своевременного обнаружения потенциально опасных участков.

Совершенствование аппаратуры наблюдения установленной на спутниках ДЗ, развитие цифровых технологий, появление космических снимков высокого и сверхвысокого пространственного разрешения метрового и субметрового диапазонов позволяет решать задачи не только обнаружения и классификации объектов, но и для оценки и прогнозирования их состояния. На сегодняшний день космические снимки самого высокого пространственного разрешения – это Ikonos-2, разрешение до 0,82 м, QuickBird-2 разрешение до 0,61 м, Worldview с пространственным разрешением до 0,47 м с расширенным спектральным диапазоном за счет добавления новых каналов (423–453, 590–640, 700–730 нм, 900–1500 нм).

При диагностике объектов Геотехнических систем наиболее перспективной является методика, сочетающая проведение наземной комплексной диагностики с космической диагностикой. Наилучших результатов космической диагностики удается достичь на основе информации

полученной с космических систем Ikonos-2, Quick Bird-2, Ресурс–Ф1М. Применяемые на этих космических системах сенсоры и камеры по пространственному разрешению условно можно разбить на две группы: к первой относятся сенсоры и камеры OSA на Ikonos-2, BHRC – на Quick Bird-2, KFA – 1000 на Ресурс-Ф1М. Сенсор ETM+ устанавливаемый на спутнике Landsat-7, можно отнести ко второй группе. В соответствии с пространственным разрешением сенсоры OSA и BHRC-60 могут быть использованы для решения многих задач геотехнической диагностики, в частности, выявить такие объекты как магистральные трубопроводы (МТ), трассы коммуникаций, природные и техногенные нарушения которые являются потенциальными источниками возникновения дефектов МТ. Недостатками этих сенсоров является узкий спектральный диапазон (0,45–0,90 мкм.), что затрудняет поиск и выявление объектов геотехнических систем, особенно, если это касается подземных объектов. В сравнении с ними сенсор ETM + более эффективен по своим спектральным характеристикам, а наличие красных каналов, например, термического, позволяет улучшить процесс распознавания объектов на космических снимках. Однако, в связи с низким пространственным разрешением использование ETM+ сенсора актуально при решении крупных Проектов, например в задачах оценки состояния магистральных нефтегазопроводов. При проведении аэрокосмического мониторинга состояния природных и инженерных объектов распознаваемых в пределах технического коридора прокладки МТ, в качестве исходных данных используются материалы воздушной и космической съемки, преимущественно видимого и инфракрасного диапазонов. Обработка аэрокосмической информации включает следующие этапы:

– техническая обработка материалов съемки;
– дешифрирование снимков;
– аналитическая обработка результатов дешифрирования.

Возможности космической диагностики объектов геотехнической системы с использованием сенсоров высокого и сверхвысокого пространственного разрешения, включают:

1. *Определение маршрута прокладки магистральных трубопроводов.*

При решении этой проблемы на основе космических снимков, полученных с систем Ikonos-2 и Quick Bird-2 оказалось, что наиболее информативными является красный канал №3 и ближний ИК №4. Использование этих каналов позволяет выявить открытые участки МТ и четко обнаружить трассу прокладки подземных трубопроводов (рис. 1).

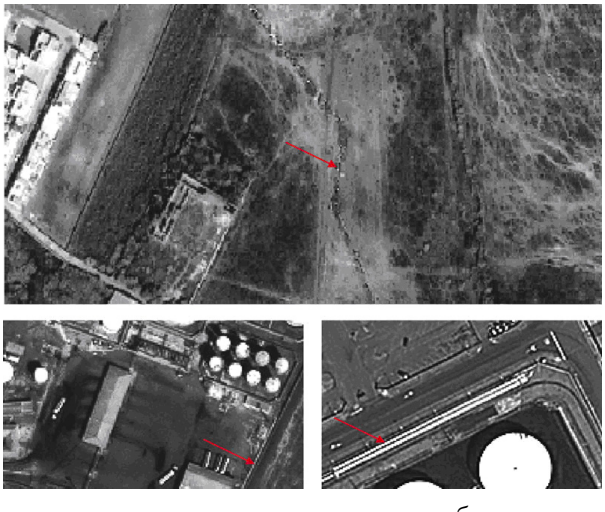


Рис. 1. Выявление трассы прокладки магистрального трубопровода по данным сенсоров с систем Ikonos-2 и Quick Bird-2

На верхнем фрагменте снимка (рис. 1) в красном канале №3 виден след подземного участка трассы МТ, на нижних фрагментах выявляются в ближнем ИК канале №4 открытые участки

МТ.

При проведении космической диагностики состояния МТ необходимо классифицировать не только сам объект, но и элементы его окружения, что позволяет дать полноценную оценку технологического состояния исследуемого объекта.

Все объекты Геотехнической системы можно разделить на две группы. К первой группе относятся объекты искусственного происхождения: различные виды магистральных трубопроводов, резервуары разной формы и назначения, мосты, линии электропередач, элементы инфраструктуры, а также другие объекты промышленного назначения. Ко второй группе следует отнести объекты природного происхождения находящиеся в пределах технического коридора прокладки МТ-это реки, открытый грунт, растительный покров.

Дешифрирование исходной информации с сенсоров OSA и BHRC-60 космических систем Ikonos-2 и Quick Bird-2 показало, что наилучших результатов при распознавании промышленных объектов, находящихся в пределах трассы прокладки МТ, удается достичь в диапазоне 0,6–0,9 мкм.

Современные методы космической диагностики позволяют также дать оценку влажности открытых грунтов в области трассы прокладки МТ, оценить воздействие подземных вод, поверхностных вод, а также смещение исследуемого объекта, вызванное действием гравитационных и других сил.

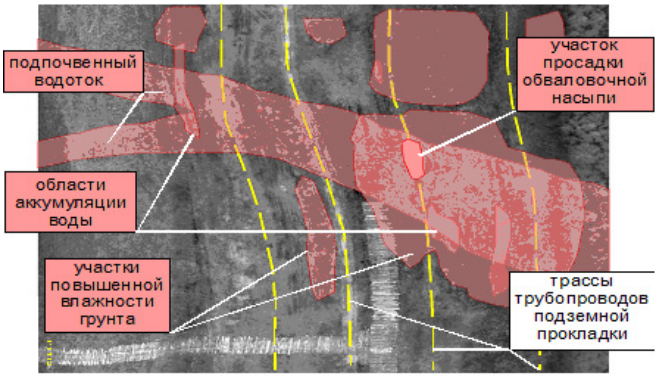


Рис. 2

На рис. 2 в результате дешифрирования фрагмента изображения магистрального газопровода, полученного со спутника Ikonos-2, удалось выявить:

- подпочвенный поток;
- несколько областей аккумуляции воды;
- участки с повышенной влажностью грунта, а также определить их фактическое расположение относительно трасс четырех ниток газопровода.

В качестве другого примера рассмотрим результаты работы по оценке влажности земель по данным, полученным с космической системы Quick Bird-2. Сравнительный анализ информативности спектральных каналов космического снимка показал, что лучше всего влажные участки почвы и границы между сухим и влажным грунтом отображаются в зеленом спектральном канале (0,52–0,61 мкм). В процессе обработки в программной среде ERDAS IMAGINE исходное изображение было классифицировано с целью выделения участков с растительным покровом, а также сухих открытых и влажных участков почвы без растительного покрова. В конечном итоге после обработки космического снимка была получена тематическая карта с нанесением следующих информационных слоев: открытая водная поверхность; густой растительный покров; сухая песчаная поверхность; сухая почва; влажная почва (рис. 3).

В результате были выявлены области, представляющие опасность из-за повышенной влажности, и области, покрытые густой растительностью, которая может служить индикатором наличия воды. По результатам оценки влажности грунтов в коридоре прокладки МТ были

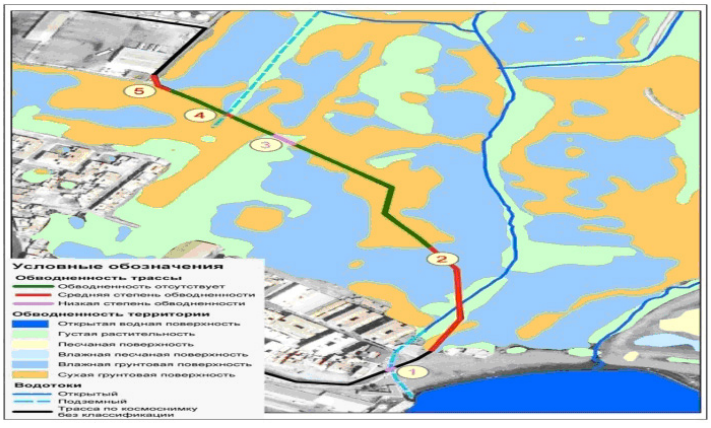


Рис. 3. Тематическая карта влажности грунтов вокруг прокладки магистрального газопровода

выделены места пересечения газопровода с участками территории повышенной влажности (на рис. 3 обозначены цифрами). При наличии космических снимков исследуемой территории за разные периоды времени можно получить данные по изменению увлажненности почвы, а полученные результаты позволяют провести оценку динамики подтопления территории и составить прогноз тенденции дальнейшего изменения влажности почвы. Важной проблемой является анализ рельефа территории расположения диагностируемого объекта, оценка рельефа территории позволяет определить взаимное расположение объектов, оценить их взаимосвязь и проанализировать влияние высоты рельефа на объект диагностики. Современные ГИС технологии позволяют получать для космической диагностики 3D модели с заданной точностью. При создании современной 3D модели исследуемого объекта необходимо совместно использовать технологии обработки данных ДЗ, ГИС-технологий, технологий трехмерного проектирования и визуализации. Информационная основа для проектирования 3D модели показана на рис. 4. Здесь отчетливо виден фрагмент ЦМР территории, прилегающей к техническому коридору прокладки магистрального трубопровода с наложенным космическим снимком Quick Bird.

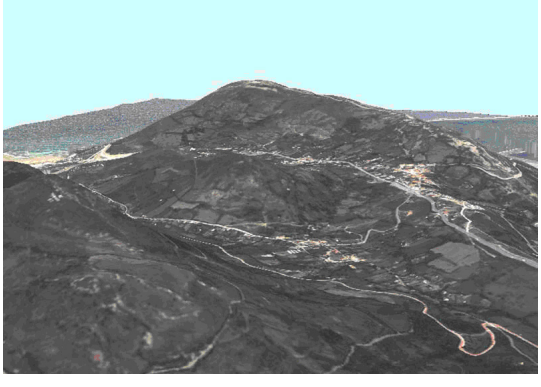


Рис. 4. 3D–основа для анализа инфраструктуры диагностируемых объектов

Пространственное разрешение современных сенсоров сверхвысокого разрешения позволяет не только выявлять объекты геотехнической системы, но и проводить оценку их внешнего состояния. Анализируя информацию сенсоров OSA, BHRC-60 и KBP-100, можно обнаружить нарушение внешнего покрытия трубопроводов, изменение и нарушение формы или целостности окружающих МТ стен, изменение геометрии и формы сферических и цилиндрических резервуаров и т.д.

Ранее мы отмечали, что основным назначением метода геотехнической диагностики МТ по космическим многоспектральным снимкам сверхвысокого пространственного разрешения является определение потенциально опасных участков на маршруте его прокладки. Такими потенциально опасными областями могут являться, например, участки территории, подверженные водным скоплениям. В таких областях возможна высокая коррозионная активность грунтов, что может привести к дальнейшему развитию дефектов МТ. Кроме того, систематизация и объединение всех данных, полученных в ходе дешифрирования и аналитической обработки космических снимков, позволяет выявить участки происхождения трассы МТ под автомобильными дорогами, под железнодорожными путями, участки изгибов трубопровода вызванных геодинамическими процессами и т.д. Анализ рельефа также является источником информации об опасных областях. На рис. 5 и 6 приведены примеры диагностирования потенциально опасных участков, требующих постоянного аэрокосмического мониторинга.

Рис. 5. На снимке обнаруживается, что часть МТ проходит под построенным позже промышленным предприятием с подъездными и железнодорожными путями

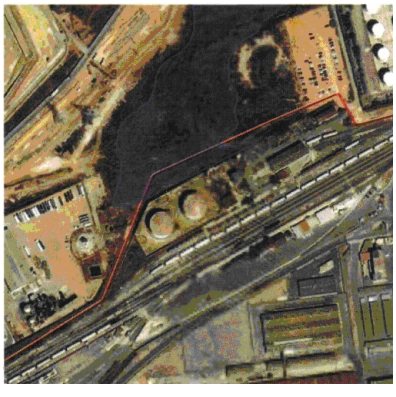


Рис. 6. На снимке выявлено, что часть МТ проходит в зоне повышенного увлажнения, возникшего в результате утечек из водопроводной насосной системы

Достоверность результатов геотехнической диагностики потенциально опасных участков технического коридора прокладки МТ по космическим снимкам высокого пространственного разрешения была проверена путем сравнения с данными наземной диагностики. В выбранной области прокладки МТ дефекты были выявлены в 4 из 5 участков, определенных по космической диагностике.

Заключение

Технология космической диагностики состояния объектов геотехнических систем позволяет существенно улучшить характеристики комплексной системы диагностики за счет использования ряда преимуществ, таких как:

- информативность (за счет учета негативного влияния природных и техногенных факторов и прогноза технического состояния объектов);
- оперативность (обусловлена высокой скоростью обработки информации);
- обзорность, обеспеченная большой полосой захвата космической съемки;
- картографичность (возможность предоставления информации в виде электронных тематических карт с координатной привязкой).

Однако приходится признать, что некоторые проблемы аэрокосмической диагностики геотехнических объектов остаются нерешенными. В первую очередь сюда следует отнести ограниченность спектральных диапазонов применяемых сенсоров, что существенно ограничивает возможность дешифрирования объектов Геотехнических Систем (в частности подземных участков магистральных трубопроводов).

***Залевська О.В.*[1], *Котолупов О.В.*[2]**
[1]*КПІ ім. Ігоря Сікорського, ФМФ, Київ, Україна;* [2]*КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна*

Дослідження властивостей тривимірних клітинних автоматів

Клітинні автомати вже десятки років активно досліджуються через їх унікальні характеристики. Деякі клітинні автомати мають такі властивості, як хаотична непередбачувана реакція на незначні зміни в конфігурації та можливість формування стійких структур, здатних зберігати свою форму протягом безліч ітерацій [1]. Особливо дослідженими є двовимірні автомати, такі як мураха Ленгтона [2] та гра "Життя" [3]. Останній, наприклад, є Тьюринг-повною системою, що дозволяє реалізовувати в ньому вирішення майже будь-якої обчислювальної задачі [4].

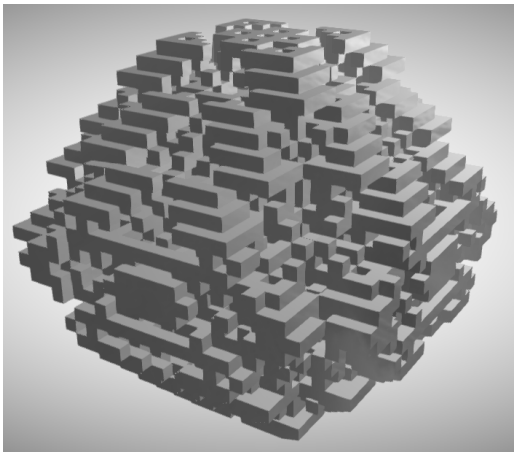


Рис. 1. Приклад еволюції тривимірного автомату

Рис. 2. Приклад еволюції двовимірного автомату гра "Життя"

Більшість ПЗ орієнтована на роботу з двовимірними клітинними автоматами, зокрема існує велика кількість реалізацій клітинного автомату гра "Життя".

Недоліками існуючого ПЗ для роботи з клітинними автоматами є:

- обмеженість можливостей задання початкового стану автомату;
- відсутність перевірки переходу автомату до тривіального стану;
- неможливість зберігання та відновлення поточного стану клітинного автомату.

На відміну від двовимірних тривимірні клітинні автомати мають ряд особливостей, серед яких:

- збільшення кількості сусідів кожної клітини;
- ускладнення процесу задання початкового стану автомату;
- проблематичність відображення внутрішньої структури клітинного автомату.

Робота з тривимірними автоматами вбачається особливо перспективною, адже їх можна використовувати для моделювання складних природних процесів в трьох вимірах, але існуюче ПЗ має вкрай обмежені можливості.

Література. **1.** Эткинс П. Порядок и беспорядок в природе [Текст]: М.: Мир, 1987. – 224 с. **2.** David Darling. The Universal Book of Mathematics: From Abracadabra to Zeno's Paradoxes. – John Wiley & Sons, 2004. – P. 180–181. **3.** Martin Gardner The fantastic combinations of John Conway's new solitaire game "life" // Scientific American. – № 4 (October 1970). **4.** Andrew Adamatzky. Game of Life Cellular Automata [Текст]. – Springer-Verlag London, 2010.

Замекула О.І., Заболотня Т.М.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Розроблення програмної системи для створення та аналізу результатів онлайн-опитувань

На сьогоднішній день непроста політично-економічна ситуація у світі піднесла на новий рівень важливість соціологічних досліджень. На сторінках ЗМІ ми часто бачимо результати соцопитувань, що стосуються підтримки виборцями політичних сил чи відношення людей до певних подій. Разом з тим опитування населення, які мають маркетинговий характер, щодо побутових питань також лишаються актуальним інструментом вивчення суспільної думки.

Сучасні інформаційні технології, зокрема Інтернет, дозволили підвищити ефективність досліджень у галузі соціології, в першу чергу, за рахунок можливості опитування ширших кіл респондентів. Через це онлайн-опитування сьогодні є найбільш розповсюдженою формою соцопитування. Користувачі можуть навіть самостійно створювати такі опитування за допомогою безкоштовних конструкторів. Але в своїй більшості існуючі механізми онлайн-опитувань характеризуються рядом недоліків:

- в них закладено проведення обмеженого статистичного аналізу отриманих результатів;
- всі механізми мають обмеження щодо кількості створюваних опитувань, кількості питань в кожному опитуванні, а також кількості респондентів;
- всі онлайн-опитування мають лінійну структуру, хоча можливість їх розгалуження дозволила б динамічно реагувати на відповіді респондента, одразу відносити його до певної соціальної групи та задавати йому більш релевантні запитання;
- не всі конструктори надають користувачеві можливість вбудовування онлайн-опитування в інший сайт, що погано впливає на швидкість поширення опитування.

Таким чином, розроблення ПЗ для онлайн-опитувань, яке підтримувало б створення анкет різного рівня складності, проведення розширеного статистичного аналізу отриманих результатів за різними показниками, а також оброблення значної кількості наборів даних для виявлення латентних закономірностей, є актуальною задачею і її вирішенню присвячена дана доповідь.

Авторами були розглянуті існуючі методи обробки великих обсягів пов'язаних між собою даних у реальному часі – OLAP, досліджено різні реалізації даної технології (ROLAP, MOLAP, HOLAP) та обрано модель HOLAP, тому що вона поєднує переваги MOLAP (висока швидкість запитів) та ROLAP (масштабування). Такі можливості необхідні, оскільки онлайн-опитування пов'язані з великими обсягами даних, з якими MOLAP працює погано, та великою кількістю запитів, для яких критична швидкість виконання (тому недоцільно використовувати ROLAP). Також розглянуто методи побудови розгалуженої структури опитувань та збереження їх у вигляді реляційної бази даних для подальшого аналізу у багатовимірних структурах даних та встановлення зв'язку з основною системою. Крім того, у доповіді розглядається можливість застосування ненормалізованої реляційної БД для збереження інформації про опитування в одній базі, оскільки у іншому випадку необхідно створити 2 БД: для збереження сконструйованих опитувань та для результатів. Такий підхід також забезпечує вищу швидкість обробки даних у OLAP-кубі та дозволяє збільшити складність опитувань за рахунок незначної втрати швидкості у виконанні запитів до БД (до 10%). Основні осі кубу в такому випадку містять спільні дані для більшості опитувань: вік, стать, професію, регіон проживання респондента та дату проходження опитування.

Також в доповіді розглядаються переваги мов ООП для уніфікації алгоритмів обробки правил переходу між питаннями різного типу в залежності від відповідей на попередні питання. Проведено порівняння засобів створення динамічних веб-сторінок (jQuery, JavaScript, Angular та React). Для розробки додатку обрано ReactJS, адже він більше орієнтований на відображення даних, ніж на побудову правильної архітектури додатка. Досліджено переваги та недоліки авторизації через сторонні сервіси за допомогою використання бібліотеки OWIN.

Згуровський М.З., Павлов О.А.
КПІ ім. Ігоря Сікорського, Київ, Україна

Методологічні основи побудови ПДС-алгоритмів для важкорозв'язуваних задач комбінаторної оптимізації

Задачі комбінаторної оптимізації відносяться до широкого кругу задач, що мають застосування в організації виробництва, в бізнесі, в оптимізації комп'ютерних обчислень, в індустрії сервісу та ін. Більшість важкорозв'язуваних задач комбінаторної оптимізації (ВЗКО) є NP-складними. Пошук для них точних алгоритмів розв'язання, час роботи яких обмежений поліномом від розмірності задачі, в даний час є безперспективним. Навіть найкращі алгоритми переборного типу як вітчизняних, так и іноземних вчених засновані на методах гілок і меж, динамічного програмування і т.п., мають експоненціальну трудомісткість і вимагають значних обчислювальних витрат навіть при розв'язанні прикладів середньої розмірності, що робить неможливим розв'язання реальних практичних задач. Тому одним з важливих напрямів досліджень є створення і аналіз ПДС-алгоритмів на основі теорії та методології їх побудови.

Теорія ПДС-алгоритмів (алгоритмів з поліноміальною складовою та достатніми ознаками оптимальності) [1–4] є новим окремим класом теорії алгоритмів та перспективним підходом для точного розв'язання ВЗКО великої розмірності. ПДС-алгоритм включає поліноміальну складову, яка строго реалізує оптимальний розв'язок, і точний експоненціальний підалгоритм або поліноміальну апроксимацію точного алгоритму (наближений або евристичний алгоритм). Поліноміальна складова ПДС-алгоритму породжується теоретично обґрунтованими логіко-аналітичними умовами (p-умовами) – достатніми ознаками оптимальності допустимого розв'язку. Поліноміальна складова може бути:

- частиною точної експоненціальної складової ПДС-алгоритму або поліноміальної апроксимації точного алгоритму;
- збігатися повністю з поліноміальною апроксимацією точного алгоритму;
- окремим поліноміальним алгоритмом, що не збігається з точною експоненціальною складовою або поліноміальною апроксимацією точного алгоритму.

ПДС-алгоритми можна розділити на два класи. ПДС-алгоритми першого класу включають поліноміальну складову і експоненціальну складову, представлену точним експоненціальним алгоритмом, побудованим як на теоретично обґрунтованих спрямованих перестановках, так і на ефективних відсіканнях. ПДС-алгоритми другого класу містять поліноміальну апроксимацію, яка може бути або евристичним поліноміальним алгоритмом, якщо отримана оцінка відхилення розв'язку від оптимального, або наближеним алгоритмом, якщо такої оцінки немає. ПДС-алгоритми суттєво розширюють границі розв'язуваних задач комбінаторної оптимізації, дозволяючи точно і ефективно розв'язувати NP-складні задачі суттєво більшої розмірності, ніж існуючі методи, а при невиконанні ознак оптимальності отримувати ефективний точний розв'язок експоненціальним підалгоритмом або евристичний чи наближений розв'язок з оцінкою відхилення від оптимуму для індивідуальних задач великої розмірності.

Необхідною і достатньою умовою побудови ПДС-алгоритмів є дослідження властивостей розв'язуваної задачі і знаходження теоретично обґрунтованих достатніх ознак оптимальності допустимого розв'язку, перевірка виконання яких, на відміну від відомих результатів, може бути реалізована лише в процесі побудови допустимого розв'язку поліноміальною складовою ПДС-алгоритму.

Ефективність отримання точного розв'язку поліноміальною складовою ПДС-алгоритму визначається двома факторами:

1. Статистично значимий обчислювальний експеримент, який повинен складатися з двох частин:
   - моделюються індивідуальні задачі, кожна з яких має допустимий розв'язок, який задовольняє p-умовам; поліноміальна складова повинна практично завжди знаходити

такий допустимий розв'язок, в іншому випадку вона є неефективною;
- моделюється довільна множина індивідуальних задач і визначається відсоток отримання оптимальних розв'язків, який повинен бути статистично значущим.

2. Обчислювальна простота поліноміальної складової ПДС-алгоритму.

Якщо в процесі реалізації поліноміальної складової ПДС-алгоритму p-умови не виконувалися, то розв'язання задачі здійснюється або точним експоненціальним алгоритмом, або поліноміальною апроксимацією точного алгоритму (наближеним або евристичним підалгоритмом). Отже, інтегральна ефективність ПДС-алгоритму визначається як ефективністю його поліноміальної складової, що реалізує перевірку достатніх p-умов оптимальності в процесі побудови допустимого розв'язку, так і ефективністю в традиційному сенсі його наближеного або евристичного підалгоритму або експоненціального за складністю точного підалгоритму.

Перевагою ПДС-алгоритмів перед всіма існуючими методами є те, що на їх основі можна побудувати для задач великої розмірності як точні, так і ефективні наближені алгоритми з оцінками якості отриманих розв'язків або високоефективні евристичні алгоритми. Побудова ПДС-алгоритму вимагає:

- виявлення властивостей оптимальних розкладів, тобто визначення та дослідження структурних властивостей задачі. В процесі розв'язання задачі досліджується структура отриманого поточного розв'язку, яка визначає подальшу стратегію пошуку оптимального розв'язку;
- виявлення достатніх ознак оптимальності допустимого розв'язку, на основі яких формалізується поліноміальна складова та розробляються ефективні поліноміальні алгоритми, які статистично гарантовано знаходять оптимальний розв'язок, якщо він задовольняє одному з достатніх ознак оптимальності, у цьому випадку поліноміальна складова ПДС-алгоритму реалізується;
- побудову ефективних експоненціальних точних алгоритмів та ефективної поліноміальної апроксимації точного алгоритму, побудованих на основі дослідження властивостей задачі. Якщо поліноміальна складова ПДС-алгоритму не реалізувалась, виконується його експоненціальна складова, наближений чи евристичний підалгоритм. Усі підалгоритми будуються на основі ПДС-алгоритму, застосовуючи його властивості та можливості.

Незважаючи на однозначність визначення ПДС-алгоритму і наявність спільних рис у методології його побудови для різних ВЗКО, якісно різні теоретичні властивості задач призводять до різноманіття модифікацій ПДС-алгоритмів в рамках одного визначення. Тому створення ПДС-алгоритму є творчим процесом для кожної задачі, що включає дослідження властивостей задач, обґрунтування достатніх умов оптимальності розв'язків, розробку та обґрунтування відповідних оптимізаційних процедур. Теорія та методологія побудови ПДС-алгоритмів для ВЗКО є конструктивним апаратом в руках спеціалістів, що дасть можливість будувати нові ПДС-алгоритми для актуальних практичних задач комбінаторної оптимізації.

В рамках теорії ПДС-алгоритмів вже побудовані високоефективні ПДС-алгоритми розв'язання 11-ти ВЗКО, що входять до складу математичного забезпечення створеної авторами чотирьохрівневої моделі календарного та оперативного планування в системах з мережним представленням технологічних процесів та обмеженими ресурсами [4].

Література. **1.** Павлов А.А. Конструктивные полиномиальные алгоритмы решения индивидуальных задач из класса NP [Текст] / А.А. Павлов, А.Б. Литвин, Е.Б. Мисюра и др., под ред. А.А. Павлова. – К.: Техника, 1993. – 126 с. **2.** Pavlov A. PDC-algorithms for intractable combinatorial problems. Theory and methodology of design [Текст] / A. Pavlov, L. Pavlova. – Uzhhorod: «Karpatskij region» shelf, №15, 1998. – 320 pp. **3.** Згуровский М.З. Принятие решений в сетевых системах с ограниченными ресурсами: Монография [Текст] / М.З. Згуровский, А.А. Павлов. – К.: Наук. думка, 2010. – 573 с. **4.** Згуровский М.З. Труднорешаемые задачи комбинаторной оптимизации в планировании и принятии решений: Монография [Текст] / М.З. Згуровский, А.А. Павлов. – К.: Наук. думка, 2016. – 716 с.

Землянський Д.І., Ліщук К.І.
КПІ ім. Ігоря Сікорського, Київ, Україна

Спосіб побудови структури бази даних для динамічно змінюваних об'єктів

З кожним роком кількість даних, структура яких з часом може змінюватись, потрібно опрацювати та зберігати, безперервно зростає. При необхідності змінити структуру БД, ми бачимо два можливих сценарії розвитку подій. Перший – це кожного разу вносити зміни в структуру, що зазвичай тягнуть за собою внесення істотних змін у відповідне програмне забезпечення. Другий спосіб – це спроектувати структуру БД таким чином, щоб при зміні об'єктів, не потрібно було вносити зміни до самої структури таблиць, а лише відкоригувати вже раніше збережені дані. З нашої точки зору, другий спосіб є найбільш прийнятним, так як при необхідності масштабування систем він дозволяє достатньо швидко здійснити зміни в структуру об'єктів БД, при цьому не зупиняючи роботу користувачів.

На сьогодні існують два основних підходи для розв'язання цієї задачі: використання патерну сутність-атрибут-значення (Entity–attribute–value model EAV) та популярний на сьогодні NoSQL. В обох підходах є свої переваг і недоліки. Однією з основних проблем патерну EAV є те, що його використання унеможливлює використання деяких можливостей стандартних реляційних баз даних, а при невдалому проектуванні ще й сильно уповільнює роботу. Але ми вважаємо і підтвердили цей факт практично, що взявши його за основу можливо створити структуру БД, яка буде гнучкою, і не сильно поступатись в швидкості [1]. Використання NoSQL має дві значні переваги: велику швидкість і гнучкість. Однак швидкість роботи досягається лише за певних умов. Наприклад, виконання певних запитів з метою збору статистики та аналізу здійснюється набагато повільніше, ніж при використанні сучасних реляційних БД [2]. NoSQL, як правило, застосовуються в тих випадках, коли потрібен швидкий доступ до даних для подальшої обробки на стороні серверу.

З нашої точки зору, для розв'язання задач, в котрих попередньо передбачити зміну структури даних, які зберігаються, неможливо, доцільніше при проектуванні БД використовувати EAV-патерн. При використанні цього підходу коректна робота EAV забезпечується за рахунок метаданих, які зберігаються в допоміжних таблицях. В них описуються додаткові правила валідації, типи даних, додаткові перевірки і багато іншого. За рахунок цього, також організовується доступ і для роботи з попередніми версіями даних. Однак, це призводить до більш складного вигляду простих на перший погляд запитів [3].

Для доведення ефективності запропонованого рішення було розроблено структуру БД, котра поєднує EAV та стандартний реляційний підхід. В якості прикладу візьмемо задачу зберігання даних користувачів та прив'язаних до них ролей. Подібний функціонал існує майже в усіх сучасних автоматизованих системах. Використання запропонованого підходу при необхідності масштабування дозволили збільшити продуктивність та зменшити навантаження на сервер бази даних. Для спрощення побудови складних запитів користувачами, базові типи запитів були реалізовані у вигляді збережених процедур, які набагато краще оптимізовані та дозволяють використовувати БД менш досвідченим користувачам. Крім того, в програмний код застосунку, котрий використовує розроблену БД, ніякі зміни структури схеми не вносились.

Використання запропонованого підходу до створення структури бази даних дозволяє динамічно змінювати об'єкти бази даних (додавати нові сутності, атрибути) з мінімальними змінами в програмному коді. Наведений підхід дозволяє виконувати масштабування програмних продуктів із мінімальною шкодою для користувачів.

Література. **1.** Entity–attribute–value model, en.wikipedia.org/wiki/Entity-attribute-value_model. **2.** NoSQL, en.wikipedia.org/wiki/NoSQL. **3.** David C. HayData model patterns: a metadata map, Morgan Kaufmann, 2006, 432 p.

Зубок В.В.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Виявлення мережевих атак на комп'ютерні системи за допомогою нейронних мереж

Вступ. В ході інтенсивного розвитку інформаційної сфери і поступової міграції різноманітних служб з різних сфер людської діяльності у Інтернет стрімко зростає необхідність у надійних методах захисту інформації від кіберзлочинних дій. Особливо гостро ця проблема стоїть перед ресурсами, які реалізують банківські операції, або оперують персональними даними користувачів. Найчастіше такі ресурси є об'єктами мережевих атак. Засоби безпеки комп'ютерних систем передбачають використання програмного або апаратного компоненту, який аналізує процеси і дані, які протікають у комп'ютерній системі, фіксує аномальні події і сповіщає про це адміністратора. Такий "сигналізуючий" компонент називається системою виявлення вторгнень (IDS – Intrusion Detection System). Для розробки IDS використовуються різноманітні підходи, в тому числі й такі, що засновані на нейронних мережах.

IDS, що використовують багатошарові персептрони. Одним зі найпростіших і найпоширеніших способів реалізації IDS за допомогою нейронної мережі є використання багатошарового персептрону (MLP, Multi-layer Merceptron). В роботі [1] досліджена система виявлення DoS-атак, що реалізована за допомогою багатошарового персептрону. В якості вхідного вектору використовується набір параметрів мережевого з'єднання, вихідний шар складається з одного нейрону, стан якого сигналізує про факт атаки. Для навчання нейронної мережі використовується алгоритм зворотнього поширення помилки (backpropagation). Навчальна вибірка береться з бази *KDD '99 Cup* [2]. IDS продемонструвала найкращі результати роботи за наявності в структурі персептрона двох прихованих шарів з 28 і 14 нейронами – загальна частка помилкових рішень склала 0,298%. Також багатошаровий песпетрон використовується в роботі [3]. Автори пропонують структуру з одним прихованим шаром на 35 нейронів. Особливістю запропонованої IDS є можливість не лише виявити факт мережевої атаки, а й класифікувати її. Класифікація проводиться на основі даних бази KDD '99 Cup, що в свою чергу обумовлює структуру вхідного і вихідного шарів персептрона – вхідний шар складається з 41 нейронів, а вихідний шар містить 4 нейрони, що відповідає 4 класам мережевих атак (DoS, U2R, R2L, Probe). Для навчання нейронної мережі використовується *trainscg* – прискорена модифікація backpropagation-алгоритму. Також був застосований метод ранньої зупинки навчання, який дозволяє уникнути ефекту, коли на тестовій вибірці нейронна мережа дає значно гірші результати, ніж на навчальній. Запропонована IDS продемонструвала високу ефективність виявлення атак всіх типів. Найкращий результат зафіксований для виявлення Probe-атак – 0,9% хибних спрацювань.

У роботі [4] пропонується модифікація базового алгоритму навчання методом зворотнього поширення помилки. Автори використовують еврістичний метод рою часток (PSO – Particle Swarm Optimization), який виконує оптимізацію нових вагових коефіцієнтів для нейронів персептрона, отриманих в результаті роботи backpropagation-алгоритму. Результати дослідження показали, що IDS з модифікованим алгоритмом навчання виявляє атаки з вищою точністю (відсоток помилок склав 0,439%, на відміну від 0,670% за базового алгоритму зворотнього поширення помилки), що в свою чергу зменшує час навчання нейронної мережі.

IDS, що використовують гібридні нейронні мережі. Автори статті [5] розглядають декілька архітектурних рішень для побудови нейромережевої IDS, які передбачають використання деякої рециркуляційної нейронної мережі (RNN) спільно з багатошаровим персептроном. Завданням RNN є стиснення вхідного 41-розмірного вектора в 12-розмірний вихідний вектор. Багатошаровий персептрон здійснює обробку стислого простору вхідних образів (головних компонент) з метою розпізнавання класу атаки. В роботі [6] описана IDS, яка виконує виявлення мережевих атак за допомогою гібридної нейронної мережі, що складається з мережі Кохонена

і багатошарового персептрону. Мережа Кохонена (SOM – Self-organizing Map), отримуючи на вході групу зі 150 мережевих пакетів, виконує кластеризацію і передає на вхід багатошарового персептрона вектори, компонентами яких є номер кластера і кількість пакетів, що належать йому. Частка помилок першого роду на тестовій вибірці склала 1,4%, помилки другого роду не були зафіксовані. Аналогічний підхід до створення IDS застосований у роботі [7]. Тут Гаусіана використовується в якості функції сусідства SOM, а в результаті кластеризації утворюється 7 кластерів. Ефективність такої системи виявлення атак виявилася 96,5%.

Інші реалізації нейромережевих IDS. Серед інших способів реалізації можна виділити роботу [8], де для побудови класифікатора використовуються засоби нечіткої логіки та лінгвістичні змінні. Утворена нечітка нейронна мережа за своєю будовою нагадує персептрон, але її характерною особливістю є наявність різних класів нейронів, які мають різні функції активації. Автор не наводить результати роботи запропонованої IDS, тому не можна оцінити ефективність цього методу. Ще один альтернативний метод запропонований у роботі [9]. Тут в якості класифікатора використовується мережа Хемінга спільно з MAXNET. Такий вибір обумовлюється тим, що мережа Хемінга не потребує тривалого навчання. Також перевагою такої мережі є невелика кількість з'єднань і її лінійна залежність від кількості вхідних параметрів. Середньостатистична ефективність класифікатора на основі нейронної мережі Хемінга склала приблизно 70%, що суттєво менше, ніж у описаних вище реалізаціях.

Запропонована IDS. В результаті досліджень запропонована IDS на основі гібридної нейронної мережі, що складається з мережі Хемінга та багатошарового персептрону. Така IDS потребує менше ресурсів і часу для навчання (42 епохи) за рахунок використання мережі Хемінга, а багатошаровий персептрон дозволяє підвищити якість розпізнавання атак до 99% і знизити відсоток помилок типу false-negative, властивий мережі Хемінга, до 1,4%.

Висновки. Нейронні мережі є ефективним інструментом для побудови систем виявлення мережевих атак. Перевагами нейромережевих IDS є їхня адаптивність і відносно невелика ресурсоємність. Існує багато методів реалізації, найбільш поширеним з яких є використання багатошарового персептрону. Однак, найвищу ефективність виявлення мережевих атак продемонстрували IDS на основі гібридних нейронних мереж. Запропонована IDS на основі гібридної нейронної мережі, що складається з Хемінга і багатошарового персептрону. Її ефективність виявилась вищою, ніж у інших розглянутих рішеннях.

Література. **1.** Частикова В.А., Картамышев Д.А., Власов К.А. (2015). Нейросетевой метод защиты информации от DDoS-атак. Современные проблемы науки и образования, № 1–1 / 2015. **2.** KDD Cup 1999 Data. Режим доступу http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html **3.** Haddadi F., Khanchi S., Shetabi M., Derhami V. (2010). Intrusion Detection and Attack Classification Using Feed-Forward Neural Network. Second International Conference on Computer and Network Technology, 262–266. DOI: 10.1109/ICCNT.2010.28. **4.** Частикова В.А., Власов К.А., Картамышев Д.А. (2014). Обнаружение DDoS-атак на основе нейронных сетей с применением метода роя частиц в качестве алгоритма обучения. Фундаментальные исследования №8, 2014, 829–832. **5.** Пархоменко І.І., Пасько О.З. (2012). Штучні біологічні системи захисту комп'ютерних мереж. Науково-практичний журнал "Захист інформації" №3, 2012, 76–81. **6.** Абрамов Е.С., Сидоров И.Д. (2009). Метод обнаружения распределенных информационных воздействий на основе гибридной нейронной сети. Известия ЮФУ. Технические науки, №11 (100), 154–164. **7.** Sodiya A.S., Ojesanmi O.A., Akinola O.C., Aborisade O. (2014). Neural Network based Intrusion Detection Systems. International Journal of Computer Applications (0975 – 8887) Volume 106 №18, 19–24. **8.** Слеповичев И.И., Ирматов П.В., Комарова М.С., Бежин А.А. (2009). Обнаружение DDoS атак нечеткой нейронной сетью. Известия Саратовского университета. 2009. Т.9. Сер. Математика. Механика. Информатика, вып. 3, 84–89. **9.** De Sa Silva L., Dos Santos A.C.F., Da Silva J.D.S. Montes A. A Neural Network Application for Attack Detection in Computer Networks. Режим доступу http://mtc-m16.sid.inpe.br/col/sid.inpe.br/marciana/2004/12.03.10.53/doc/Lilia%201572_ijcnn2004.PDF.

Калинин В.В., Бояринова Ю.Е.
КПИ им. Игоря Сикорского, Киев, Украина

Клеточные автоматы

Клеточные автоматы – дискретные детерминированные системы, поведение которых полностью определяется в терминах локальных взаимодействий [1].

Основное направление исследования клеточных автоматов – алгоритмическая разрешимость тех или иных задач. Также рассматриваются вопросы построения начальных состояний, при которых клеточный автомат будет решать заданную задачу.

Клеточные автоматы были использованы Джоном фон Нейманом для исследования самовоспроизведения. Однако они применяются и для совершенно других целей [2].

Отличительной особенностью клеточных автоматов является то, что они имеют структуру, объединяющую вычислительную компоненту и данные, с которыми автомат оперирует [2]. Эта особенность хорошо сочетается с принципами объектно-ориентированного программирования, поскольку класс обладает такими же свойствами.

Существует специальный класс клеточных автоматов, называемых тоталистичными. На каждом шаге эволюции клеточного автомата значение ячейки равно какому-либо целому числу (обычно выбираемого из конечного множества), а новое состояние клетки определяется суммой значений клеток-соседей и, возможно, предыдущим состоянием клетки. Если состояние клетки на новом шаге зависит от её предыдущего состояния, то такой клеточный автомат называется внешним тоталистичным.

Понятие окрестности клетки задается как свойство клеточного автомата, причем для каждой из клеток это понятие одинаково. Например, для двумерного клеточного автомата с тетрагональной решеткой чаще всего в качестве соседей рассматривают либо четыре клетки, соприкасающиеся с данной стороны, либо восемь клеток, соприкасающиеся с данной хотя бы углами.

Клеточные автоматы обычно обладают следующими свойствами:

- решетка однородна;
- взаимодействия локальны;
- множество состояний клетки конечно;
- изменения значений всех клеток происходят одновременно.

Как правило, рассматривают одномерные либо двумерные клеточные автоматы. Хотя количество измерений ничем не ограничено, трехмерные и выше клеточные автоматы сами по себе практически не исследуются, либо имеют очень схожие с двумерными свойства [3]. Отдельного внимания заслуживают одномерные клеточные автоматы. В частности, существует такой одномерный клеточный автомат, для которого доказана универсальность – с его помощью реализуется любой вычислительный процесс.

Клеточный автомат – это однородная «решетка», в каждой клетке которой находится конечный автомат. В двумерном автомате наиболее часто используется тетрагональная решетка [4]. Возможны также клеточные автоматы с треугольной или гексагональной решеткой. Отличительная особенность конечных автоматов, являющихся элементами клеточного автомата: в них не используются входные и выходные воздействия, а на каждом шаге новое состояние каждого конечного автомата определяется его собственным состоянием и состоянием его соседей (элементов окрестности).

Литература. **1.** Столбов С.А. Разработка метода реализации игр на тетрагональном поле, в которых преобладают локальные взаимодействия http://is.ifmo.ru/papers/gamecell/ **2.** Тоффоли Т., Марголус Н. Машины клеточных автоматов. М.: Мир, 1991. **3.** Фон Нейман Дж. Теория самовоспроизводящихся автоматов. М.: Мир, 1971. **4.** Шалыто А., Туккель Н. От тьюрингова программирования к автоматному // Мир ПК. 2002, № 2, с. 144–149. http://is.ifmo.ru/works/turing/.

Касумова Т.А., Пашаева М.М., Джамалова А.Г., Аралис С.В.
Научно-исследовательский институт аэрокосмической информатики НАКА, Баку, Азербайджан

Роль информационной культуры в создании и развитии информационного космического общества

В связи с появлением современных информационных технологий, развитием рыночной экономики и формированием новых рыночных отношений в культурных, экономических и социальных областях общества происходят коренные перемены. Сегодня страны мира живут в переходный период от промышленного общества к информационному.

Благодаря созданию информационного общества, формированию экономического правительства, развитию интеллектуального потенциала, область информации и коммуникационных технологии (ИКТ) считается приоритетной после нефтяного сектора и в последние годы сильно влияет на социально экономическую жизнь нашей страны. Современные технологии информации и связи, всестороннее развитие, повышение интеллектуального уровня населения и активное привлечение республики в интеграционные процессы открывают новые возможности формирования информационного общества [1].

Основными компонентами информационного общества являются: информационные технологии, электронное государство, электронная наука, интернет, информационная культура, электронная экономика, информационная безопасность, информационные права, информационная инфраструктура. К основным свойствам развития информационного общества Азербайджана относится создание: глобальной информационной системы, новых форм (дистанционное получение образования, электронная торговля, теле работа, электронная демократия, электронное правительство и т.д.) социально-экономической деятельности, рынка информации и знаний и т.д. Таким образом, создание информационного общества дополнительно способствует полной демократизации жизни общества.

Благодаря информационно-коммуникационным технологиям космическая промышленность в своем развитии имеет высокие достижения. Развитие космической промышленности и вывод на орбиту телекоммуникационных и многоцелевых спутников устраняет информационную зависимость от других стран и обеспечивает информационную безопасность. Начиная с президентского указа “Создание Азербайджанской космической промышленности и вывода на орбиту телекоммуникационных спутников” от 2008 года, в 2009 году были продолжены работы государственной программы по созданию и развитию космической промышленности Азербайджана, что сделало возможным сотрудничество с 10-ю развитыми странами.

Азербайджан является членом космического клуба и одной из первых стран СНГ, имеющей телекоммуникационные спутники в Кавказском регионе. Главной целью государственной программы является, используя достижения космической промышленности, обеспечение потребностей государственных структур в спутниках связи, обеспечение потребностей населения регионов в радиовещании, увеличение каналов международной связи страны и полезное использование космического пространства, развитие социальных, экономических, культурных, научных и других областей, а также безопасности страны [2]. Программа предоставляет широкие перспективы для развития международных отношений в космической области, увеличение потенциала космической промышленности в республике, развитие техники в космической промышленности, предоставление дополнительных услуг связи, телерадиовещание, дистанционные зондирование Земли, гидрометеорология, экологический мониторинг, контроль чрезвычайных ситуаций, космические исследования, программы поиска и спасении и т.д.

8-го февраля 2013-го года в Азербайджане был выведен на орбиту первый телекоммуникационный спутник “Azərspace 1”. В 2017 планируется вывести спутник “Azərspace 2” и в 2020–2021-ом году еще несколько дополнительных спутников. После запуска первого спутника “Azercosmos” приобрел опыт в этом направлении. “Azercosmos” обменивается опытом с компаниями и странами, имеющими опыт в этой области. Наша страна тесно сотрудничает с США, Канадой, Малайзией, Францией, Турцией.

В настоящее время развитие многих областей общества связано с ростом информационных ресурсов и возможностей компьютеров и интернета. Будущее любой страны определяется не только богатством природных ресурсов, а также людскими ресурсами, имеющими высокий научный-культурный и информационно-интеллектуальный потенциал.

Человечество должно быть подготовлено к тому, чтобы пользоваться возможностями информационного общества для удовлетворения своих потребностей и улучшения благосостояния. Для достижения этих целей очень важно реализовать процесс формировании информационной культуры у людей. Основы информационной культуры человечества состоят из получения и освоения информации, запоминания, обработки, представления, знания и умения обеспечить безопасность. Одновременно, важными компонентами информационной культуры являются соблюдение законов и норм этикета, умение уберечься от вредной информации. В настоящее время в глобальном информационном пространстве вредной информацией, являющейся опасной для морально-психологического здоровья человечества, являются проблемы киберопасности и интернет-зависимости. Решение этих проблем состоит в формировании информационной культуры у людей [3].

Сегодня в направлении развития ИКТ были сделаны важные шаги. Информация и знания являются стратегическими ресурсами общества и составляют основу социально– экономической, технологической, культурной областей страны. При развитии информационного общества меняется отношение к людским ресурсам, информационным услугам. Для работы с информацией, предоставленной ИКТ–информационных услуг, от людей требуются новые формы культуры. Сегодня граждане информационного общества должны стать носителями информационной, технологической культуры и стратегически мыслить для обнаружения и решения проблем. Одной из главных задач является подготовка кадров, имеющих информационную культуру для создания и развития космической промышленности. Для благополучного исполнения государственной программы имеет особое значение подготовка кадров по эксплуатации и управлению спутником, производству космического оборудования. получению и обработки космической информации. Имея в виду, что космическая промышленность является новой областью, необходимо за короткий срок подготовить отечественные кадры по этим направлениям [4].

Быстрое развитие ИКТ и телекоммуникаций в мире требует принять меры в направлении повышения конкурентной способности спутниковых сетей. Кроме этого, создание супер многофункционального компьютерного центра по принятию и обработки космической информации, требует подготовку специалистов в области космической промышленности и спутниковых систем. Экономика государства и вооруженные силы требуют расширения фундаментальных научных знаний. Освоение космического пространства является одним из определяющих признаков уровня развития экономической и национальной безопасности.

Динамика развития ИКТ в Азербайджане, повышение доверия к нашей стране в результате благополучной реализации нефтяной стратегии, образцы благополучного сотрудничества с международными организациями, поставленные задачи по формированию экономики, основанной на высоких технологиях, оказание большой заботы и внимания к области ИКТ, благодаря высоким технологиям, гарантирует в ближайшем будущем нашей стране высокие достижения в области создания и развития космической промышленности [5].

Литература. **1.** T.İ.Süleymanov, T.A.Qasımova, M.M.Paşayeva. “İnformasiya cəmiyyətində informasiya mədəniyyətinin formalaşması məsələləri”. İnformasiyan təhlükəsizliyinin multidissplinar problemləri üzrə II respublika elmi-praktik konfrans, Bakı – 2015, S. 95–98. **2.** www.google.az **3.** T.A.Qasımova, S.İ.Hüseynova, M.M.Paşayeva. “Azərbaycan kosmik sənayesinin yaradılmasında və inkişafında informasiya mədəniyyətinin rolu”. Memarlıq, İnşaat və Nəqliyyat sahələrində proqressiv texnoloqiyalar mövzusunda elmi-praktik konrans, Bakı-2016, S. 161–163. **4.** R.M.Əliquliyev, R.Ş.Mahmudova. İnformasiya mədəniyyəti, formalaşması problemləri və həlli yolları., Bakı Universitetinin xəbərləri. Sosial-siyasi elimlər seriyası, 2008, №1, 137 s. **5.** www.ict.az.

Киричек Г.Г., Овчар А.О.
Запорізький національний технічний університет, Запоріжжя, Україна

Захист передачі мережних даних

Приводиться аналіз сучасних методів захисту передачі даних, їх недоліків та переваг і пропонується рішення, у вигляді системи опосередкованої передачі, яке виключає безпосередню взаємодію між клієнтами.

Вступ. Проведення фінансових операцій в мережі, використання кредитних карток, доступ до закритих інформаційних ресурсів і передача телефонних розмов вимагають відповідного рівня безпеки [1]. Існує безліч сучасних криптографічних алгоритмів, які відповідають певним вимогам безпеки передачі даних [2]. Їх стійкість ґрунтується на складності обчислення зворотної функції до функції шифрування [3]. Як видно з успішних атак на алгоритми шифрування, є необхідність в розробці ефективних методів реалізації безпечного обміну даними, що є актуальним в умовах мережної передачі даних.

Постановка завдання. Основними завданнями роботи є розробка методу опосередкованої передачі даних в мережі, розробка алгоритму випадкового визначення засобів передачі (серверів) та розробка гнучкої зміни шифрувального алгоритму в системі обміну, а також проведення дослідження ефективності розробленого методу, визначення ймовірності несанкціонованого доступу до інформації, яка передається та тестування загальної криптостійкості системи.

Результати досліджень показали, що спеціалізовані програмні засоби захисту інформації від несанкціонованого доступу мають кращі можливості і характеристики, ніж вбудовані засоби мережних ОС [2]. Також визначено, що момент передачі даних в мережі є найбільш слабким місцем в електронному обміні інформацією, тому для забезпечення захищеної передачі, в процесі реалізації методу, пропонується внесення додаткових вузлів, через які передається повідомлення між кінцевими вузлами.

Реалізація системи. Виходячи з основної мети проведення досліджень та спираючись на загальні завдання розробки, перед початком реалізації системи наведемо модель функціонування в мережному середовищі (рис. 1). При реалізації основних завдань дослідження, розроблено алгоритми та методи за допомого яких досягається непередбачуваність шляху передачі повідомлення в мережі, неповторність алгоритмів шифрування з випадковим вибором кількості перехідних вузлів та їх адрес.

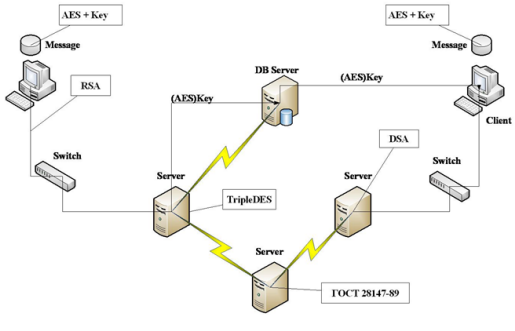


Рис. 1. Модель системи

Зашифроване повідомлення клієнта та ключ для розшифрування передаються на попередньо випадково визначений сервер. Передача при цьому шифрується іншим алгоритмом, визначеним на сервері методом визначення алгоритму шифрування. Сервер розшифровує пакет, який містить зашифроване повідомлення та ключ, після чого відправляє ключ на сервер баз даних, відомий тільки серверам даної системі. Потім визначається наступний сервер з відповідним алгоритмом шифрування на даний момент часу та повторно шифрується і пересилається повідомлення. Клієнти, як і сервери, знають усі можливі алгоритми шифрування, що змінюються через визначений час. В процесі передачі число серверів в системі визначається попередньо, на відміну від числа проміжних вузлів, яке генерується при передачі за допомогою генератора псевдо-випадкових чисел. Кожен з проміжних серверів має унікальний ідентифікатор, який дописується в кінець повідомлення, задля запобігання петель, при визначенні наступного серверу. Останній сервер, перед передачею повідомлення клієнтові, дістає ключ першого алгоритму з сервера баз даних, додає його до повідомлення, зашифровує своїм алгоритмом, а потім передає пакет клієнтові. Клієнт розшифрує пакет відповідно до

серверного алгоритму та дістанеться останнього (клієнтського) алгоритму з відповідним ключем. Основна мета методу полягає у випадковості вибору серверів та неможливості визначити, яким з алгоритмів зашифроване повідомлення в даний проміжок часу.

Задля запобігання можливості здійснення DDoS-атак, доступ до серверів в системі обмежуємо шляхом настроювання списків доступу так, щоб вони не виконували команду пінг із зовні. Вибір криптоалгоритмів шифрування даних здійснювався виходячи з можливостей бібліотек System.Security.Cryptography каркаса .NET Framework. Вони включені в аналіз реалізації симетричних алгоритмів шифрування Rijndael, DES, 3DES, RC2 і реалізацію несиметричного алгоритму шифрування RSA. Для симетричних алгоритмів проведено аналіз їх швидкодії з точки зору шифрування документів різного об'єму і зроблена порівняльна характеристика різних симетричних алгоритмів [4]. Для несиметричних алгоритмів шифрування проведено аналіз швидкодії алгоритму RSA при обробці документів різного об'єму при фіксованій довжині модуля і з точки зору обробки документів фіксованого обсягу при різних значеннях модуля. У каркасі .NET Framework симетричний криптоалгоритм Rijndael представлений managed-реалізацією RigndaelManaged. В проведеному досліджені розглядалися реалізації криптоалгоритмів з розмірами ключів 128, 192, 256 бітів [4]. Симетричний алгоритм RC2 представлений типом RC2CryptoService Provider. При проведенні аналізу використовувались екземпляри класу провайдера з ключами 40, 64, 96 і 128 бітів. Симетричні алгоритми DES і 3DES (TripleDES) представлені типами DESCryptoServiceProvider і TripleDESCryptoServiceProvider відповідно. Дані типи також є оболонками і делегують відповідність реалізацій алгоритмів бібліотеки CryptoAPI. Реалізація алгоритму DES підтримує довжину ключа 64 біта, тоді як реалізація 3DES дозволяє задавати ключі довжиною 128 біт (2 різних ключа) і 192 біта (3 різних ключа). У таблиці 1 наведено результати порівняльного аналізу симетричних алгоритмів і несиметричного алгоритму RSA.

Табл. 1. Результати порівняльного аналізу роботи алгоритмів

Об'єм Кбайтах	DES 64 bit	3DES 192 bit	RC2 128 bit	Rijndael 128 bit	RSA 512 bit	RSA 1024 bit	RSA 2048 bit
20000	$1,10 * 10^{-3}$	$1,85 * 10^{-3}$	$1,35 * 10^{-3}$	$2,12 * 10^{-3}$	$1,25 * 10^{-1}$	$2,5 * 10^{-1}$	$6,5 * 10^{-1}$
50000	$2,03 * 10^{-3}$	$4,24 * 10^{-3}$	$3,60 * 10^{-3}$	$4,46 * 10^{-3}$	$3,08 * 10^{-1}$	$6,3 * 10^{-1}$	$1,8$
100000	$3,80 * 10^{-3}$	$8,50 * 10^{-3}$	$4,72 * 10^{-3}$	$8,54 * 10^{-3}$	$6,14 * 10^{-1}$	$1,26$	$3,61$

В таблиці наведено середній час в секундах, який потрібен вказаному криптоалгоритму, щоб перетворити документ певного розміру, який вказано у відповідному стовбці таблиці. Як видно з таблиці, алгоритм RSA з довжиною модуля в 512 біт працює в середньому на два порядки довше, ніж в середньому реалізації симетричних алгоритмів. При цьому збільшення довжини модуля в 2 і 4 рази (1024 і 2048 біт) призводить до збільшення часу криптоперетворень приблизно в 2 і в 4 рази відповідно.

Висновки. Результатом проведених досліджень є підвищення показника надійності при обміні даними за рахунок збільшення проміжних вузлів, випадкового вибору наступного вузла, шифрування даних різними алгоритмами, неповторність крипто-алгоритмів та їх зміна на вузлах відповідно до певного показника часу. В ході подальших досліджень планується реалізація алгоритму зі збільшення швидкодії системи при передачі даних в мережі.

Література. **1.** Баричев С.Г. Основы современной криптографии / С.Г. Баричев, Р.Е. Серов // Учебное пособие. – М.: Телеком, 2002. – 152 с. **2.** Шнайер Б. Прикладная криптография. Протоколы, алгоритмы и исходные тексты на языке С / Б. Шнайер. – М., 2012. – 610 с. **3.** Мао В. Современая криптография. Теория и практика / В. Мао. – М.: Вильямс, 2005. – 768 с. **4.** Проценко А.Г. Исследование быстродействия алгоритмов шифрования данных на базе технологии .Net Framework / А.Г. Проценко, И.В. Лысенко. – Харків, 2011. – Вип. 4. – С. 179–184.

Кіріленко О.Г., Труш Г.О.
Національний аерокосмічний університет ім. М.Є. Жуковського "Харківський авіаційний інститут", Харків, Україна

Систематичний огляд екологічного програмного забезпечення

Програмне забезпечення (ПЗ) відіграє важливу роль у всіх сферах життєдіяльності людини. Кількість, обсяг та області застосування ПЗ різко збільшилися. В результаті індустрія програмного забезпечення стала витрачати значні ресурси на розроблення, супровід і використання ПЗ, що негативно впливає на оточуюче середовище. Зазначена проблема сприяла посиленню уваги до реалізації концепції сталого розвитку, в контексті якої ПЗ розглядають як інструмент для її реалізації та як об'єкт для забезпечення такої сталості. В рамках цієї концепції ПЗ, що відповідає вимогам енергоефективності, екологічності, безпеки має назву екологічне («зелене») програмне забезпечення.

На даний час аналіз науково-технічної літератури переконливо показує, що результати досліджень стосовно екології програмного забезпечення сегментарні, суперечливі, а доступ до інформації, що стосується вказаної тематики, здебільш закрито. Саме тому науковці та практики мало ознайомлені з новою областю знань – екологією програмного забезпечення. Даний огляд має на меті узагальнення емпіричних наукових даних про використання екологічного підходу до процесу розроблення ПЗ, виявлення прогалин в дослідженнях, а також визначення нових напрямків науково-дослідницької діяльності.

Для синтезу результатів досліджень у певній галузі знань у науковій практиці використовують систематичні огляди. Такі огляди передбачають всебічний пошук публікацій з певного питання; використання точних, відтворюваних критеріїв відбору статей для огляду, а також оцінку структури та особливостей оригінального дослідження; узагальнення даних та інтерпретацію отриманих результатів.

Детальний план проведення систематичного огляду, описано далі у вигляді протоколу огляду. Цей протокол розроблено на основі процесу огляду та керівних принципів щодо здійснення систематичного огляду літератури, запропонованих B. Kitchenham і S. Charters [1]. Протокол огляду включає такі компоненти.

1. Питання дослідження:
   a) «Як впливає ПЗ та його розробка на навколишнє середовище?»,
   b) «Яким чином можна зробити ПЗ «зеленим»?».
2. Електронні бази даних IEEE [2], ACM [3], ScienceDirect [4], SpringerLink [5] і ABI/ Inform [6], вибір яких обумовлений тим, що вони є найбільшими в світі електронними базами даних в області програмної інженерії та комп'ютерних наук. При цьому дві останні включають найбільші в світі онлайн-колекції науково-технічних і медичних журналів, книг, довідників.
3. Стратегію пошуку, що було розроблено відповідно до рекомендацій в [1].
4. Критерії відбору досліджень, в ході реалізації стратегії пошуку: а) критерії включення: джерела, опубліковані в період 2010–2016 рр. на англійській мові; первинні джерела, які пов'язані з науковими розробками; наукові статті або журнали; дослідження, проведені в промисловості та в академічному середовищі у контексті досліджуваних питань, а також доступність повного тексту; б) критерії виключення: дублікат наукового дослідження; статті, опубліковані на веб-сайтах компаній або студентські роботи; глави книг.
5. Вилучення даних з досліджуваних питань.
6. Синтез масиву даних за наслідками реалізації попередньої компоненти плану.

У ході дослідження було знайдено та розглянуто 856 публікацій, з яких було відібрано

46 джерел, що охоплюють різні аспекти дослідження: емпіричні дослідження, звіти, що свідчать про досвід в галузі програмної інженерії, тематичні дослідження та ін. Кожне дослідження було розглянуто на основі аналізу контексту дослідження, питань дослідження й емпіричного підтвердження результату. Для короткого викладу змісту відібраних для дослідження публікацій були визначені блоки, які служать основою для аналізу (назва джерела, анотація, ключові слова, методи дослідження, висновки).

Систематичний огляд дозволив узагальнити науково-технічні дослідження в контексті питань, що розглядаються. Встановлено, що більшість публікацій (74%) присвячена вирішенню проблеми зменшення шкідливих факторів впливу на навколишнє середовище процесу розроблення ПЗ та його використання в різноманітних технічних пристроях. Найменша кількість публікацій присвячена вирішенню проблеми визначення метрик «відтінків зеленого» на стадіях розробки ПЗ (11%). Визначено, що екологічне використання ПЗ, технічне обслуговування, утилізація та його розроблення засновані на зменшенні: споживання різноманітних ресурсів (обсяг оперативної пам'яті комп'ютерів, тактова частота та швидкодія центрального процесора, обсягу споживаної енергії, матеріальних, людських, природних ресурсів та ін.); а також викиду шкідливих речовин в атмосферу; на забезпеченні надійності та безпеки.

Встановлено, що розроблення «зеленого» ПЗ спрямовано на зниження впливу на навколишнє середовище, викликаного самою програмою. Доведено, що «зеленість» (англійський термін «greenest») є показником якості ПЗ, що визначає здатність програми мінімально впливати на навколишнє середовище за рахунок ефективного використання ресурсів. «Зеленість ПЗ» (ЗПЗ) досягається за рахунок застосування «зелених» принципів і процесів розроблення, використанням методів, що дозволяють визначити ступінь «зеленості» на всіх етапах розроблення ЗПЗ. Показано, що тематичні дослідження, які проводяться авторами, допомагають підвищити обізнаність компаній та користувачів щодо ефективності ЗПЗ та мотивувати його використовувати. До перспективних напрямків досліджень в області розробки та використання ЗПЗ відносять: спеціалізовані методи та засоби розроблення екологічного ПЗ; розроблення моделей якості ПЗ з урахуванням метрик «відтінків зеленого»; методи та інструменти для визначення метрик «відтінків зеленого» на етапах: аналізу вимог, проектування, кодування; розроблення «зелених» предметно-орієнтованих мов програмування, а також компіляторів для них; метрики, методи та засоби генерації тестових сценаріїв для оцінки «зеленості» ПЗ; середовища розроблення ЗПЗ; використання методів рефакторинга для підвищення «зеленості» ПЗ; методи та інструменти для вимірювання «зеленості» ПЗ; методи й інструменти для моніторингу «відтінків зеленого».

Література. **1.** B. Kitchenham and S. Charters (2007). Guidelines for performing Systematic Literature Reviews in software engineering. Keele University and Durham University joint report. Durham, UK. **2.** IEEE Xplore Digital Library. Retrieved from http://ieeexplore.ieee.org/Xplore/guesthome.jsp. **3.** ACM. Retrieved from http://dl.acm.org. **4.** ScienceDirect. Retrieved from http://www.sciencedirect.com. **5.** SpringerLink. Retrieved from http://link.springer.com. **6.** Inform. Retrieved from http://www.proquest.com.

Копп А.М., Орловский Д.Л.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Континуум предприятия как составная часть знание-ориентированного подхода к управлению предприятием

Архитектура предприятия. Современное предприятие представляет собой сложную систему, состоящую из нескольких областей, влияющих друг на друга. Архитектура предприятия описывает компоненты, их отношения и принципы, лежащие в основе данной системы [1].

Архитектурный стандарт. В качестве архитектурного стандарта консорциумом The Open Group была разработана методология TOGAF (The Open Group Architecture Framework), предоставляющая методы и инструменты, предназначенные для поддержки принятия, разработки, использования и поддержки архитектуры предприятия. Данная методология является одним из ведущих архитектурных решений, благодаря углубленному рассмотрению вопросов использования и управления архитектурой предприятия в ее последних версиях [2].

Согласно методологии TOGAF, архитектура предприятия включает следующие области [3]:

1. Бизнес-архитектура – описывает организационную структуру, бизнес-процессы и деятельность организации с учетом установленных показателей эффективности.
2. Архитектура информации – определяет информационные потоки процессов организации, методов сбора, обработки, предоставления и хранения информации.
3. Архитектура приложений – определяет, какие приложения используются и должны использоваться для управления данными и поддержки бизнес-функций, и реализует объекты информационной архитектуры в прикладных системах.
4. Техническая архитектура – определяет, какие аппаратные и программные средства необходимы для обеспечения работоспособности прикладных систем в соответствии с операционными требованиями (надежность, производительность и т.п.).

Центральным элементом TOGAF является метод разработки архитектуры предприятия ADM (Architecture Development Method). Он совмещает в себе элементы онтологии – определение основных элементов структуры и их взаимодействия – и методологии – последовательность разработки, внедрения и поддержания архитектуры в актуальном состоянии [4].

Континуум предприятия. В состав методологии TOGAF входит понятие континуума предприятия (Enterprise Continuum) – совокупности всех архитектурных решений, существующих как внутри предприятия, так и в отрасли в целом (рис. 1) [5]. Континуум предприятия состоит из набора готовых моделей, шаблонов и архитектурных описаний, которые максимально обобщены и сгруппированы по уникальным или специальным направлениям [3].

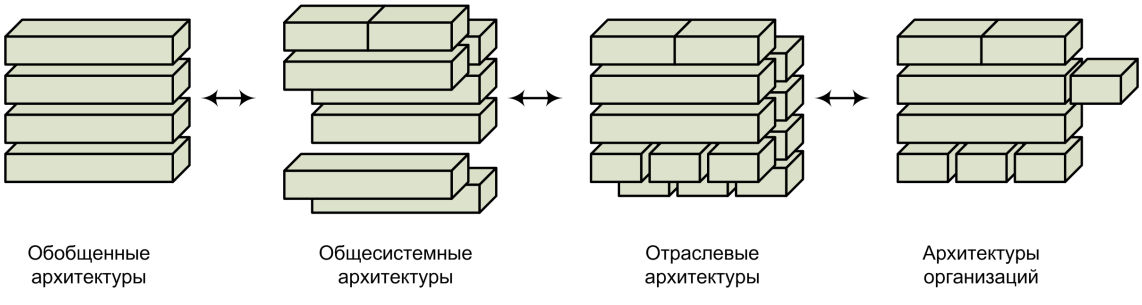


Рис. 1. Континуум предприятия

На уровне обобщенных архитектур определяется ряд баз знаний – техническая эталонная модель TRM (Technical Reference Model) и информационная база стандартов SIB (Standards Information Base). Техническая эталонная модель TRM представляет собой рекомендуемое описание архитектуры предприятия. Информационная база стандартов SIB является набором

стандартов, рекомендуемых консорциумом The Open Group при построении архитектуры предприятия. Данные принципы построения архитектуры могут использоваться практически любым предприятием. На уровне общесистемных архитектур определяются общие принципы, прослеживаемые во многих предприятиях. Принципы, определяемые отраслевыми архитектурами, являются характерными для предприятий, занятых в одной сфере деятельности. К самому высокому уровню специализации – архитектурам организаций – относятся архитектуры конкретных предприятий [3].

Архитектурный репозиторий. В методологии TOGAF понятие континуума предприятия тесно связано с понятием архитектурного репозитория (Architecture Repository), который представляет своего рода накопитель объектов архитектуры – компоновочных блоков (Building blocks). Компоновочные блоки, в свою очередь, состоят из архитектурных описаний (потенциально многократного использования), представленных в виде каталогов, матриц и диаграмм [6]. Данные компоновочные блоки позволяют разработчикам создавать готовые архитектурные решения на протяжении всего процесса разработки, внедрения и поддержания архитектуры предприятия в актуальном состоянии.

Для формализации взаимосвязи архитектурного репозитория с решениями, создаваемыми в континууме предприятия, была разработана следующая UML диаграмма классов (рис. 2).

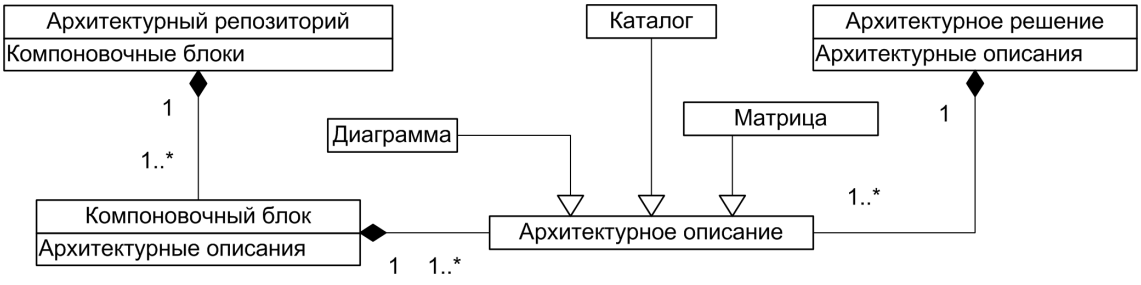


Рис. 2. Взаимосвязь архитектурного репозитория с создаваемыми решениями

Выводы. Концепция континуума предприятия, поддерживаемая архитектурным репозиторием, может быть использована для накопления и повторного использования знаний об архитектуре предприятия, представляемых в виде архитектурных описаний. Повторное использование накопленных знаний позволит повысить эффективность процесса ADM при переходе от обобщенной архитектуры, определяемой в методологии TOGAF, к архитектуре конкретной организации. Вследствие высокого уровня абстракции и обобщения в методологии TOGAF, актуальными остаются вопросы, связанные с информационными технологиями, которые могут быть использованы для реализации архитектурного репозитория.

Литература. **1.** Minoli, D. (2008). Enterprise Architecture A to Z: Frameworks, Business Process Modeling, SOA, and Infrastructure Technology. Boca Raton: Auerbach Publications. 10.1201/9781420013702. **2.** Iacob, M.I., Jonkers, H. (2006). Quantitative analysis of enterprise architectures, Interoperability of Enterprise Software and Applications (pp. 239–252). London: Springer. 10.1007/1-84628-152-0_22. **3.** Логиновский, О.В., Нестеров М.И., Шестаков А.Л. (2013). Применение методов архитектурного подхода в развитии информационной системы крупного ВУЗа. Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника, 4, 123–128. **4.** Липунцов, Ю.П. (2016). Прикладные программные продукты для экономистов. Основы информационного моделирования. Учебное пособие. Москва: Изд-во Проспект. **5.** Кудрявцев, Д.В., Арзуманян М.Ю., Григорьев, Л.Ю. (2014). Технологии бизнес-инжиниринга. Учебное пособие. СПб: Изд-во Политехн. ун-та. **6.** The Open Group. (2011). Part IV – Architecture Content Framework. Retrieved from http://pubs.opengroup.org/architecture/togaf9-doc/arch/toc-pt4.html.

Корнієнко О.О.
КПІ ім. Ігоря Сікорського, ФТІ, Київ, Україна

Алгоритм апроксимації вейвлет-пакетного дерева декомпозиції

Вступ. Важливим етапом усіх систем обробки мовних сигналів (МС) є виділення ознак. Одними з найбільш розповсюджених параметрів МС є кепстральні коефіцієнти лінійного передбачення та мел-частотні кепстральні коефіцієнти. Ці ознаки МС використовують перетворення Фур'є, яке не дозволяє описати МС, як сукупність квазістаціонарних сегментів (звуків голосних та приголосних фонем, міжфонемних переходів, переходів мова-тиша). Тому для усунення цих недоліків застосовують вейвлет-пакетне перетворення (ВПП) [1]. Структуру вейвлет-пакетного дерева декомпозиції (ВПДД) МС адаптують до шкали мелів [2] або барків [3], що уже є неактуальними [4] та застосовують для обробки сигналів з частотою дискретизації 8 кГц. Це обмежує придатність таких ознак до застосування у сучасних системах обробки МС. Тому актуальною задачею є розробка загального алгоритму апроксимації вейвлет-пакетного дерева декомпозиції МС до сучасної психоакустичної моделі [4], що є адаптивним до параметрів сигналу.

Алгоритм формування вейвлет-пакетного дерева. Дискретне ВПП є узагальненням класичного дискретного вейвлет перетворення та виконується шляхом каскадного застосування пари низько- та високочастотних фільтрів. Структуру ВПП зображають бінарним деревом [4], де вузол відповідає вейвлет-пакетним коефіцієнтам розкладу піддіапазону, рух по лівій гілці – розрахунку низькочастотних компонент сигналу, по правій – високочастотних. Шлях по дереву до вузла може бути кодованим бінарною послідовністю довжини L (глибина декомпозиції сигналу). Старший біт визначає дію на рівні декомпозиції 0, молодший – на максимальному L_{max}. Біт 0 послідовності відповідає низькочастотному розкладанню сигналу, 1 – високочастотному. Кількість частотних піддіапазонів рівня розкладання L_{max} ВПП становить $2^(L_{max}+1)$, а частотна роздільна здатність дерева рівна $f_s/2^(L_{max}+2)$. На рис. 1 зображена структурна схема алгоритму формування ВПДД МС, що апроксимує функцію $B(f)$ залежності ширини критичної смуги психоакустичної моделі від її центральної частоти. Вхідними параметрами алгоритму є: максимальна кількість рівнів розкладання L_{max}, частота дискретизації сигналу f_s та функція $B(f)$. Результатом роботи алгоритму є масив бінарних послідовностей шляхів ВПДД МС, що апроксимують психоакустичну модель.

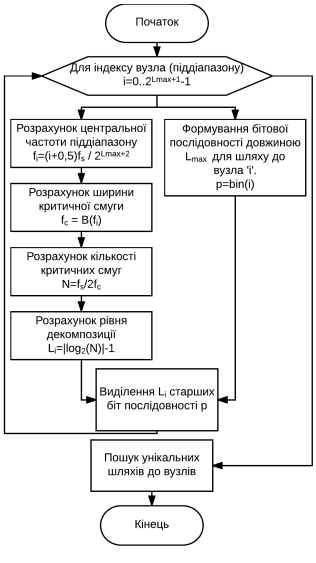


Рис. 1. Блок-схема алгоритму формування апроксимації ВПДД МС

Висновок. Запропонований алгоритм використовується для апроксимації еквівалентних прямокутних смуг пропускання (ERB, Equivalent Rectangular Bandwidth) при виділенні ознак МС у задачах розпізнавання мовця, мови та емоцій.

Література. **1.** Farooq O. Wavelet based robust sub-band features for phoneme recognition / O. Farooq, S. Datta // IEE Proceedings – Vision, Image and Signal Processing. – 2004. – Vol. 151, no. 3. – pp. 187–193. **2.** Sarikaya R. Wavelet packet transform features with application to speaker identification / R. Sarikaya, B.L. Pellom, J.H.L. Hansen // Proceedings of the IEEE nordic signal processing symposium: (NORSIG'98). – 1998. – pp. 81–84. **3.** Deshpande M. Speaker Identification Using Admissible Wavelet Packet Based Decomposition / M. Deshpande, R. Holambe // International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering. – 2010. – Vol. 4, no. 1. – pp. 83–86. **4.** Siafarikas M. Wavelet packet approximation of critical bands for speaker verification / M. Siafarikas, T. Ganchev, N. Fakotakis // International Journal of Speech Technology. – 2007. – Vol. 10, Issue 4. – pp. 197–218.

Куцаченко Н.Г.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Оценка эффективности реализаций алгоритмов трекинга движения средствами OpenCV

В системах компьютерного зрения и обработки изображений часто возникает задача определения смещения объектов в трехмерном пространстве с помощью камеры. Цель оценки оптического потока – вычислить аппроксимацию к области движения в зависимости от изменяющейся во времени интенсивности изображения.

Наиболее известными алгоритмами для решения задач трекинга движения, которые реализованы в библиотеке OpenCV являются: Лукаса–Канаде, Farneback, SimpleFlow.

Как показал анализ источников, наиболее часто используемым является алгоритм Лукаса–Канаде [1]. При реализации этого алгоритма может возникнуть такая проблема, как aperture problem, вследствие которой на выходе получаются слишком большие значения сдвига, которые не соответствуют действительности. При использовании приема multi-scaling алгоритм успешно и стабильно обнаруживает как малые, так и большие сдвиги, довольно устойчив к шумам.

Farneback разработал быстрый алгоритм для вычисления плотного оптического потока (dense-потока) [2]. Он предложил аппроксимировать изменение интенсивности в окрестности с помощью квадратичной формы с симметричной матрицей A. Этот алгоритм менее стабилен, не очень хорошо определяет большие сдвиги, но имеет большее быстродействие.

Метод SimpleFlow заключается в том, что для всех точек в окне находится функция «энергии», отвечающая за вероятность перехода исходной точки в эту точку [3]. Далее, считается свертка этой энергии с гауссовым окном и находятся значения точки, минимизирующие эту функцию. Алгоритм работает значительно медленнее других. Результат работы более точный, чем Farneback, лучше определяется сдвиг на равномерных изображениях, заметно лучше срабатывает с большими сдвигами.

Алгоритм	Достоинства	Недостатки
Лукаса-Канаде	Хорошо работает со sparse-потоком, достаточно точен и быстр	Плохо определяет сдвиги на равномерных изображениях
Farneback	Быстрая скорость работы алгоритма; хорошо работает с dense-потоком	Не всегда удовлетворительная точность результатов
SimpleFlow	Хорошо определяет сдвиги на равномерных изображениях; хорошо работает с большими сдвигами	Медленная скорость работы

Рис. 1. Достоинства и недостатки алгоритмов

Выводы: при разработке программного обеспечения для реализации трекинга движения с использованием средств OpenCV необходимо учитывать следующие факторы: требуемое быстродействие, плотность потока, который необходимо определять, точность результатов работы алгоритма. Целесообразным является использование алгоритма Лукаса-Канаде, если достаточно посчитать сдвиг некоторого набора точек (sparse-поток). Для вычисления dense-потока лучше использовать алгоритм SimpleFlow, при условии, что быстродействие не является критичным. Если же быстродействие является необходимым фактором, то следует использовать алгоритм Farneback, но необходимо удостовериться, что точность его работы является удовлетворительной.

Литература. **1.** B.D. Lucas and T. Kanade. An iterative image registration technique with an application in stereo vision. In Seventh International Joint Conference on Artificial Intelligence, pages 674–679, Vancouver, 1981. **2.** G. Farnebäck. Two-Frame Motion Estimation Based on Polynomial Expansion. In SCIA'03 Proceedings of the 13th Scandinavian conference on Image analysis, pages 363–370, Halmstad, 2003. **3.** Michael W. Tao, Jiamin Bai, Pushmeet Kohli, and Sylvain Paris. SimpleFlow: A Non-iterative, Sublinear Optical Flow Algorithm. Computer Graphics Forum (Eurographics 2012), 2012.

***Лавренюк А.М.*[1], *Лавренюк С.І.*[2], *Тульчинський П.Г.*[2]**
[1]*КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна;* [2]*Інститут кібернетики ім. В.М.Глушкова НАН України, Київ, Україна*

Оптимізація архітектури програмного забезпечення для аналізу великих даних базі мови програмування python

На прикладі аналізу даних для рішення задачі відтоку клієнтів (churn predict) телекомунікаційної компанії показано створення системи прийняття рішень з гнучкою модульною структурою та з оптимізацією часових та ресурсних витрат на розробку програмного забезпечення.

Вступ. При рішенні задачі аналізу великих даних для прогнозування відтоку клієнтів необхідно вести аналіз дуже великого об'єму даних, тому що сучасні телекомунікаційні компанії мають клієнтську базу від 1000000 і більше. Дані необхідно брати за останні три місяці. Для побудови моделей та тестування прогнозування на першому етапі достатні дані для телекомунікаційних компаній, що містяться в відкритих джерелах [1,2]. При розробці програмного забезпечення та його впровадженні необхідно збалансовувати три компоненти, що значним чином впливають на результат: T – час, C – ресурси (люди, гроші), Q – якість. При цьому замовник часто має потребу отримати програмне забезпечення, яке масштабоване, мультиплатформне та яке не вимагає великої заміни наявної інфраструктури та комп'ютерної техніки замовника.

Однією із сучасних вимог є необхідність роботи програмного забезпечення на різних основних програмних платформах (операційних системах), та на різних пристроях. Інтерфейс програмних систем повинен бути розрахований на роботу різних категорій співробітників. Для однієї групи користувачів, наприклад, аналітиків необхідна більша інтерактивність, та можливість проводити самостійно різне представлення результатів, розрахунок додаткових показників. Для другої групи, наприклад, керівництва необхідно швидко отримувати точні результати розрахунків основних показників для прийняття управлінських рішень. У цьому випадку інтерфейс повинен містити тільки необхідні дані, працювати швидко і на різних пристроях та системах.

Якщо розробляти різні модулі на різні операційні системи, для різних пристроїв (робочі станції, смартфони, планшети...) та для різних груп користувачів будуть значно збільшені T, C при $Q = const$:

$$T = T0 + deltaT, C = C0 + deltaC, \tag{1}$$

де $deltaT > 1$, $deltaC > 1$.

Запропонований підхід. Виходячи з наявного досвіду побудови програмного забезпечення для прогнозування даних [3] та для обробки великих об'ємів даних [4], також провівши ряд експериментів, запропоновано для рішення задачі аналізу великих даних використати мову програмування python, та бібліотеки, що необхідні для машинного навчання та прогнозування: pandas, numpy, sklearn та інші.

Запропоновано використати триярусну архітектуру: сервер баз даних, сервер додатків та робочі станції користувачів. Головну задачу по обробці даних, машинному навчанню та прогнозуванню візьме на себе сервер додатків. Так як для машинного навчання в python на базі бібліотеки sklearn використовується не просто бібліотека numpy, а numpy+mkl, то буде забезпечена висока швидкодія.

Python+django дає можливість побудувати веб-додаток (web-application) з високою швидкодією, гнучкістю та мультиплатформністю. Веб-додаток проводить аналіз даних, будує інтерфейс користувача для другої групи користувачів, та правильно працює на різних пристроях. Jupiter notebook веб-додаток для створення інтерфейсу для першої групи користувачів (рис. 1).

Ядро системи – обробка та аналіз даних, також спільні процедури доступу до даних

Рис. 1. Архітектура програмного забезпечення

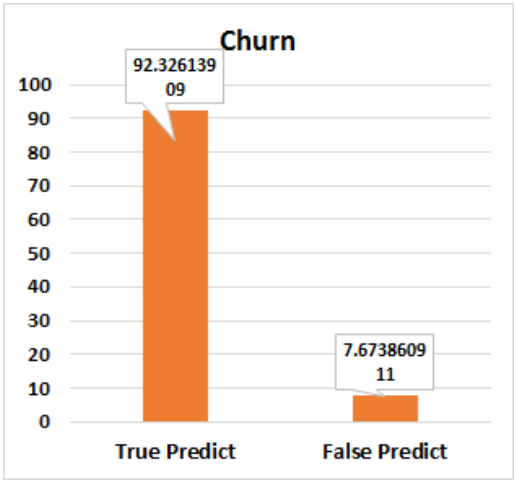


Рис. 2. Результат прогнозування

винесені в окремі модулі та є спільними для двох систем. Таким чином, вірно виділивши спільні процедури, функції, та з'єднавши їх із першою та другою системами, ми отримуємо:

$$T = T0 + deltaT1, C = C0 + deltaC1, deltaT1 < deltaT, deltaC1 < deltaC, \quad (2)$$

де $Q = const, deltaT1 => 0, deltaC1 => 0$.

Слід зазначити, що технології глибинного навчання та методи машинного навчання успішно використовуються для великих об'ємів даних при проведенні аерокосмічних спостережень [5,6]. Тому запропонований авторами підхід можна буде успішно використати в інших галузях.

Проведені експерименти показали високий результат якості прогнозування та швидкості обробки даних (рис. 2).

Висновок.

- мова програмування python та наявні компоненти дають можливість побудови оптимальних програмних систем для аналізу великих об'ємів даних, для машинного навчання, та для побудови мультиплатформних веб-додатків або додатків робочого столу (desktop app), не критичних до великої швидкодії;
- запропонований підхід до побудови програмного забезпечення для аналізу великих даних на базі мови програмування python зменшує витрати на створення програмного забезпечення та його супровід в подальшому.

Література. **1.** Telco Customer Churn data set, https://www.ibm.com/communities/analytics/watson-analytics-blog/predictive-insights-in-the-telco-customer-churn-data-set/. **2.** Machine Learning Repository, http://archive.ics.uci.edu/ml/?cm_mc_uid=15134185542514827482720&cm_mc_sid_50200000=1485174656. **3.** Автоматическое построение числовых прогнозов / [Лавренюк С.И., Перевозчикова О.Л., Тульчинский В.Г., Харченко А.В.] // Компьютерная математика. – 2010. – № 2. – С. 52–61. **4.** Лавренюк А.Н., Лавренюк С.И., Назаренко Е.В. Подход к оптимизации программы обработки больших объемов сейсмических данных на примере 3D миграции дуплексных волн // Компьютерная математика., сборн. научн. трудов – 2014. №1. – Р. 69–75. **5.** Лавренюк М., Новіков О. Технологія глибинного навчання для великих об'ємів даних // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки, GEO-UA 2016. – 2016. – Р. 46–47. **6.** Новіков О., Шелестов А., Лавренюк М. Розробка методів машинного навчання в проекті Google Earth Engine // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки, GEO-UA 2016. – 2016. – Р. 20–21.

Лисенко А.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Порівняльний аналіз методів трекінгу об'єктів на відеоряді

В даний час відбувається швидкий розвиток систем відеоспостереження і повсюдне впровадження їх в усі сфери діяльності на транспорті, на військових і промислових об'єктах. Системи відеоспостереження додатково до традиційного спостереження за об'єктами володіють потенційними технічними можливостями для вирішення цілої низки завдань. Актуальними завданнями, які повинні вирішувати системи відеоаналітики в найближчому майбутньому, є виявлення, ідентифікація об'єктів і визначення траєкторій їх руху, вимірювання швидкості руху об'єктів, впровадження систем відеоспостереження в автоматизовані системи управління технологічними процесами, що працюють в режимі реального часу, а також ефективне використання каналів передачі даних за рахунок попередньої обробки сигналів і передачі технічної (аналітичної) інформації замість передачі відеозображення. Наразі існує невелика кількість методів розв'язку задачі трекінгу. Зокрема у роботах [1, 2] розглядають інформацію про колірні характеристики та контур об'єкта для розпізнавання дорожньої розмітки і об'єктів на дорозі. Також у роботі [3] розглядають методи виділення характерних точок об'єкта та слідкування за їх зміщенням.

Методи. *Базований на SIFT* Ідею використання особливих точок реалізує алгоритм SIFT (Scale Invariant Feature Transform) [4]. На основі піраміди гауссіанів і піраміди різниць гауссіанів знаходяться ключові точки зображення, які, після уточнення їх положення, описуються за допомогою SIFT-дескриптора. Для кожної особливої точки визначається її напрямок, що обчислюється на основі напрямків градієнтів точок, сусідніх з нею. Дескриптором в методі SIFT є вектор, що обчислюється, виходячи з градієнтів в деякому вікні ключової точки. Отримані дескриптори інваріантні щодо зміщення і повороту зображення так само забезпечують часткову інваріантність щодо збільшення або зменшення зображення, зміни яскравості і положення камери.

Базований на SURF Метод SURF [5] шукає особливі точки за допомогою матриці Гессе. Детермінант матриці Гессе – гессіан – досягає екстремуму в точках максимального зміни градієнта яскравості, добре детектуючи плями, кути і краю ліній. Гессіан інваріантний щодо обертання але не щодо масштабу, тому використовують різномасштабні фільтри для знаходження гессіанів. Для кожної ключової точки вважається напрямок максимального зміни яскравості (градієнт) і масштаб, взятий з масштабного коефіцієнта матриці Гессе. Для забезпечення інваріантності обчислення дескрипторів особливої точки щодо обертання, потрібно визначити переважну орієнтацію перепадів яскравості в особливій точці. До афінних перетворень SURF НЕ інваріантний, в чому явно програє менш чутливому до них SIFT. З іншого боку, алгоритм відрізняється підвищеною щодо SIFT швидкодією.

PREDATOR Підхід, який використаний в даному алгоритмі, називається Супровід-Моделювання-Виявлення (Tracking-Modeling-Detection (TMD)), він поєднує адаптивний супровід об'єкта з навчанням детектора об'єкта в процесі розпізнавання [6]. Після того як об'єкт був захоплений за допомогою будь-якого алгоритму захоплення, траєкторія об'єкта починає спостерігатися двома процесами (що розширює і звужує події). Вони будують детектор об'єкта на льоту. Обидва процеси роблять помилки, стабільність системи досягається скасуванням подій. Навчання на льоту і класифікація виробляються за допомогою рандомізованого лісу.

Порівняльні метрики. Одними із основних критеріїв є кількості похибок першого та другого роду. Похибкою першого роду в даній галузі є нерозпізнавання явно присутнього в кадрі об'єкта, помилкою другого роду є розпізнавання об'єкта, що явно на ньому відсутній. Для цієї системи вищеописані критерії можна вимірювати в такий спосіб:

- оцінкою за кількістю помилок першого роду буде кількість кадрів з даного відеоряду, на яких не було знайдено шуканий об'єкт за умови його присутності в кадрі;
- оцінкою за кількістю помилок другого роду буде кількість кадрів з даного відеоряду, на

яких було знайдено шуканий об'єкт за умови його відсутності в кадрі.

Крім того, ще вводять оцінку швидкості алгоритму, таким чином визначаючи швидкість обробки кадрів. Для оцінки якості алгоритму потрібно сформувати вибірки з відеофайлів різних типів та визначити послідовності кадрів, на яких присутні шукані об'єкти, та запустити алгоритм на кожному і обчислити кількість не розпізнаних та хибно розпізнаних кадрів, а також середній час обробки кадру. Для зручності порівняння якості роботи методів за першими двома критеріями введемо таку оцінку:

$$Q = \frac{N_s d}{N_s t} \times \left(1 - \frac{N_n d}{N_n t}\right) \times 100\% \quad (1)$$

де $N_s d$ – кількість кадрів, на яких було розпізнано об'єкт, $N_s t$ – загальна кількість кадрів, що містять об'єкт, обчислена раніше, $N_n d$ – кількість кадрів, на яких було розпізнано об'єкт попри його відсутність, $N_n t$ – загальна кількість кадрів, що не містять об'єкт, обчислена раніше.

Порівняльний аналіз. В результаті проведеного порівняння було отримано такі оцінки для розглянутих методів:

- SIFT – 91.3%, 16.5 fps
- SURF – 87.4%, 41.8 fps
- PREDATOR – 94.5%, 20 fps

SIFT Недолік полягає в тому, що метод є математично складним і в обчислювальному відношенні важким. SIFT заснований на гістограмі градієнтів, тобто градієнти кожного пікселя у кластері повинні бути обчислені і ці обчислення потребують часу. Це не є ефективним для малопотужних пристроїв. Крім того, метод погано реагує на зміну освітлення об'єкта та розпливчастість зображення. З іншого боку, метод обирає велику кількість стійких характерних точок, що гарантує високу його точність.

SURF Метод за своєю природою подібний до SIFT, але він є більш сфокусованим на швидкості знаходження характерних точок за рахунок зниження обчислювальної складності алгоритму. За рахунок цього тим дещо втрачається стійкість точок що призводить до зростання кількості помилок.

PREDATOR Основною перевагою методу є висока стійкість до зміни кута та точки огляду за рахунок моделювання супроводжуваного об'єкта. З іншого боку основним недоліком є необхідність у попередньому тренуванні класифікатора для розпізнавання об'єкта, що призводить до великої кількості помилок при трекінгу незнайомого класифікатору об'єкта.

Висновки. Можна зробити висновок, що базований на SURF метод трекінгу показав хорошу пару оцінок порівняно з іншими методами. Тим не менш слід зауважити, що кожен з методів має свою галузь застосування, зокрема PREDATOR куди краще підходить для задач, що потребують високої точності та в яких є можливість певний час дослідити потрібний об'єкт під час трекінгу. В той же час базований на SURF метод трекінгу дозволяє з трохи гіршою точністю переслідувати об'єкт безпосередньо після його вибору/виявлення.

Література. **1.** Belcher C, Du Y. Region-based SIFT approach to iris recognition. Opt Lasers Eng 47 (1): 139–147, (2009). **2.** Lin M.X., Xu X.H. Multiple object visual tracking from a moving object. In: Proc. the sixth international conference on intelligent systems design and applications, Jinan, China, vol.2, pp. 373– 378, (2006). **3.** Conaire C.O., O'Connor N.E., Smeaton A.F. An improved spatiogram similarity measure for robust object localisation //Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on. – IEEE, 2007. – T. 1. – C. I-1069-I-1072. **4.** Lowe D.G. Distinctive image features from scale-invariant keypoints // International journal of computer vision. – 2004. – T. 60. – №. 2. – C. 91–110. **5.** Bay H. et al. Speeded-up robust features (SURF) // Computer vision and image understanding. – 2008. – T. 110. – №. 3. – C. 346–359. **6.** Online learning of robust object detectors during unstable tracking Zdenek Kalal, Jiri Matas, Krystian Mikolajczyk – 3rd On-line Learning for Computer Vision Workshop 2009, Kyoto, Japan, IEEE CS.

Лілікович С.О., Лютенко І.В.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Суб'єктно-орієнтований підхід до забезпечення диверсності N-версійного програмного проекту

В рамках загальної схеми розробки N-версійного програмного забезпечення пропонується формальний підхід щодо забезпечення командної диверсності для підвищення якості ПЗ. Розроблено концептуальну схему підходу, позначені подальші роботи та перспективи.

Одним з загальновживаних підходів щодо підвищення надійності програмного забезпечення є його диверсифікація, яка може полягати в підготовці $N \geqslant 2$ версій: вони відповідають однаковим вимогам, проте розробку ведуть різні, ізольовані одна від одної, команди.

N-версійність забезпечується множиною засобів, яка називається Design Diversity Dimensions (DDD). Зазвичай в цю множину входить різноманіття інструментів реалізації проекту, різноманіття варіантів специфікації ПЗ, тощо. Різноманіття виконавців можна також розглядати як частину множини DDD, адже виконавці безпосередньо впливають на реалізацію ПЗ.

Концептуальна схема суб'єктно-орієнтованого підходу до оцінки командної диверсності. Враховуючи одну з добре знаних схем N-версійної розробки [1], автори пропонують доповнити її процедурою цілеспрямованого формування диверсних команд розробників. Модифіковану схему в нотації діаграми діяльності UML наведено на рисунку 1.

Запропонований процес передбачає наступні дії:

1. Налаштування початкової інфраструктури N-версійного проекту, що полягає у виборі відповідних технологічних параметрів для визначення DDD.
2. Формування диверсних команд розробників: вони розглядаються як додаткове вимірювання в просторі DDD.
3. Якщо узгоджено всі атрибути проекту, виконуються загальні дії щодо реалізації N-версійного ПЗ.
4. Після завершення виконання проекту визначається рівень диверсифікації, що був досягнутий з даними технологічними параметрами проекту.
5. Якщо досягнутий рівень диверсифікації є прийнятним – продукт передається замовнику. Інакше повторюються дії з кроку 3.
6. У випадку якщо на кроці 2 виникає необхідність підвищення рівня командної диверсності проекту, то виконується відповідна процедура оцінки персональних якостей розробників з подальшим формуванням нових диверсних команд. Оцінку персональних якостей можна виконувати за допомогою таких методів оцінки особистості як Myers-Briggs Type Indicator (MBTI), Minnesota Multiphasic Personality Inventory (MMPI), тощо [2].

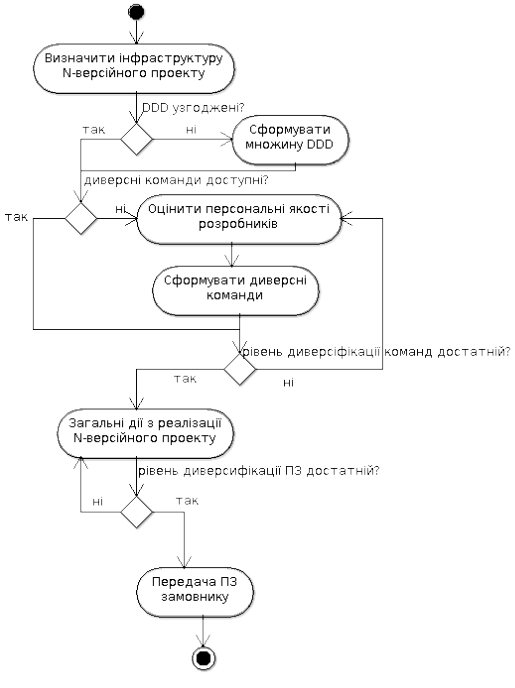


Рис. 1. Модифікована схема розробки N-версійного ПЗ

Процес формування диверсних команд ґрунтується на гіпотезі про те, що певні особисті якості розробників (стать, вік, соціальний статус, рівень освіти, тощо) впливають на рівень диверсифікації проекту в цілому. Також відомо, що такі особливості розробника впливають на

програмний код, що він створює, і оцінка цього програмного коду за допомогою відповідних метрик може допомогти визначити особистий стиль розробника.

Деякі питання реалізації підходу. Персональні якості розробників пропонується оцінювати за допомогою методу MBTI, який не є медичною процедурою і не потребує для використання ані кваліфікованого спеціаліста, ані придбання стимульного матеріалу. MBTI надає оцінки за 8 показниками та поділяє респондентів на 16 груп.

Для формування команд розробників необхідно досягти мінімального рівня диверсифікації всередині команд та максимального рівня між ними. Така задача може бути зведена, наприклад, до задачі лінійного програмування.

Рівень диверсифікації програмного коду пропонується оцінювати за допомогою відповідних метрик визначення стилю програмування. В роботі [3] пропонується 50 таких метрик, що успішно використовуються для встановлення авторства програмного коду. Деякі, найбільш важливі з них, наведено в таблиці 1, де значення N залежить від конкретної платформи розробки.

Табл. 1. Деякі метрики визначення стилю програмування

Назва	Опис	Діапазон значень
STY1c	Кількість відкриваючих дужок, які є останнім символом в рядку	$[0; +\infty]$
STY2a	Відношення кількості рядків, що містять тільки коментарі, до загальної кількості рядків з коментарями	$[0; 1]$
STY5	Середня кількість пробілів праворуч від операторів	$[0; N]$
PRO2b	Середня довжина назви методів	$[0; N]$

Оцінку сформованих команд можна виконати через обчислення показника надійності розробленого на даній ітерації програмного забезпечення. Для цього можна використовувати показник Distinct Failure Diversity (DFD) [4]:

$$DFD = \sum_{n=1}^{N} \frac{N-n}{N-1} t_n.$$

Тут N позначає загальну кількість версій, а t_n характеризує відношення кількості провальних запусків у випадку коли n версій показали ідентичний дефект до загальної кількості унікально провальних запусків. Таким чином показник DFD є нормованою величиною, яка лежить в діапазоні $[0; 1]$.

Подальші роботи та перспективи. В даній публікації представлено суб'єктно-орієнтований підхід до забезпечення диверсності N-версійного програмного проекту, в основу якого покладено гіпотезу про наявність взаємозв'язку між певною множиною персональних якостей виконавців проекту та їх стилем програмування. В подальшому планується: проведення більш детального аналізу альтернативних методів оцінки особистих якостей розробників та методів оцінки програмного коду, мотивований вибір методу визначення кореляції між зазначеними показниками, розробка процедури формування диверсних команд із заданим рівнем диверсності, проектування та реалізація відповідного CASE-засобу автоматизації основних етапів запропонованого підходу.

Література. **1.** Lyu Michael R., He Y.-T. Improving the n-version programming process through the evolution of a design paradigm // IEEE Transactions on Reliability. – 1993. – Vol. 42, no. 2. – P. 179–189. **2.** Anne Marie Wheeler Burt Bertram. Assessment in Counseling: A Guide to the Use of Psychological Assessment Procedures. – 5 edition. – Amer Counseling Assn, 2012. **3.** Ding Haibiao, Samadzadeh Mansur H. Extraction of Java program fingerprints for software authorship identification. – Vol. 72, no. 1. – P. 49–57. **4.** Partridge Derek, Krzanowski Wojtek. Distinct failure diversity in multiversion software // Res. Rep. – 1997. – Vol. 348. – P. 24.

Ляхевич О.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Порівняльний аналіз методів відновлення зображення

Відновлення спотворених зображень – одна з найцікавіших та найважливіших проблем у задачах обробки зображень – як з теоретичної, так і з практичної точок зору. Окремими випадками є розмиття через неправильне фокусування та розмиття внаслідок руху (змаз). Ці дефекти, які достатньо часто зустрічаються в житті кожної людини, дуже важкі у виправленні.

Основною проблемою під час відновлення розмитих зображень є наявність шумів у вихідному зображенні. Окрім того, що наявність шумів заважає використовувати алгоритм без додаткових модифікацій, кількість шуму багатократно збільшується у відновленому зображенні. Це пов'язано з тим, що кожна точка відновленого зображення може утворюватися з сотень точок початкового.

Відновлення розфокусованих зображень може застосовуватись для відновлення інформації з розмитої фотографії. Наприклад, це може допомогти відновити автомобільний номер зі змазаної фотографії чи зображення з камери спостереження. Це може допомогти, наприклад, у криміналістичній справі.

Цифровий шум – дефект зображення, внесений фотосенсором і електронікою пристроїв, які їх використовують внаслідок недосконалості технології. Причини шуму в цифрових сенсорах можуть бути найрізноманітнішими, але основні – це теплові коливання і темнові струми. На величину шуму також впливає ряд факторів, таких як значення ISO, тип матриці, розмір пікселя, температура, електромагнітні наведення тощо.

В більшості випадків шум розподілений за гауссівським законом розподілу, а також є адитивним, не корелює з зображенням та не залежить від координат точки. Саме з цих припущень і будемо виходити при побудові математичної моделі шуму.

Метод фільтрації Вінера поєднує в собі врахування властивостей спотворювальної функції та статистичних властивостей шуму в процесі відновлення. Метод заснований на розгляді зображення та шуму як випадкових процесів, і задача постає наступним чином: знайти таку оцінку для неспотвореного зображення, щоб середньоквадратичне відхилення цих величин одна від одної було мінімальним [1].

Метод регуляризації Тихонова полягає в регуляризації задачі, яка досягається заміною вихідної задачі на задачу знаходження мінімуму певного згладжуючого функціоналу. В якості такого функціоналу можна використовувати квадрат норми лапласіана, з додатковим обмеженням (зв'язком) [2].

Був проведений аналіз результатів роботи методів. В якості оцінки була використана дисперсія градієнтів зображення – у чіткого зображення вона менша, ніж у розмитого. Також було проведене опитування серед експертів.

Табл. 1. Якість відновлення зображень

Метод	Оцінка
Інверсна фільтрація	1.1
Фільтрація Вінера	4.2
Регуляризація Тихонова	3.8

Висновки. Порівняння відновленого різними методами зображення показало перевагу методу фільтрації Вінера над методом регуляризації Тихонова. Маючи однаковий час обробки зображення, метод фільтрації Вінера показав кращий результат, що показують математичний критерій та експертні оцінки.

Література. **1.** R.C. Gonzalez, R.E. Woods, „Digital Image Processing," 3rd edition, Prentice-Hall, 2008. **2.** Tikhonov, A.N. (1977). Solution of Ill-posed Problems. Washington: Winston & Sons.

Максим К.Є.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Дослідження сучасних методів стеганоаналізу

На сьогодні існує загроза використання інформаційних потоків у соціальних мережах, службах обміну повідомленнями, сервісах зберігання мультимедійних даних зловмисниками для формування прихованих (стеганографічних) каналів несанкціонованої передачі даних. Поява новітніх стеганографічних методів приховання повідомлень (стегоданих) у файли-контейнери, зокрема цифрові зображення, заснованих на комплексному використанні декількох перетворень контейнера та стегоданих, суттєво ускладнює виявлення сформованих стеганограм.

Для детектування наявності прихованих повідомлень у зображеннях-контейнерах було запропоновано ряд методів пасивного стегоаналізу, зокрема алгоритмів статистичного стегоаналізу. Для ускладнення виявлення несанкціонованої передачі конфіденційних даних державних установ та приватних підприємств широко використовуються методи цифрової стеганографії. Особливістю даних методів є використання існуючих інформаційних потоків в інформаційно-комунікаційних системах для формування прихованих каналів зв'язку, що дозволяє долати системи моніторингу та контролю [1].

В якості файлів-контейнерів для вбудовування стегоданих можуть бути використані мультимедійні дані, наприклад, аудіо- та відеодані, зображення, текстові файли, виконувані файли тощо [2]. Згідно з дослідженнями провідних аналітичних компаній [3], переважна кількість стеганографічних методів, що перебувають у відкритому доступі, заснована на прихованні повідомлень у цифрові зображення.

Відомі методи вбудовування стегоданих у зображення-контейнери можливо розділити на три групи [2, 3] – засновані на зміні стандартних методів обробки цифрового зображення, приховуваанні повідомлень в просторовій (LSB-методи) області зображення та в області перетворення зображення-контейнера. Вагомою перевагою перших двох груп є приховання значних об'ємів стегоданих, а також відносна простота алгоритмів вбудовування та вилучення повідомлень. Проте висока чутливість отримуваних стеганограм до будь-яких змін в процесі передачі відкритими каналами зв'язку, наприклад, стиснення із втратами, застосування методів підвищення візуальної якості цифрового зображення, обмежує практичне застосування даних стеганографічних методів [2]. Для подолання зазначених обмежень використовуються методи вбудовування стегоданих в області перетворення контейнера шляхом використання особливостей поширених, наприклад двовимірного дискретного вейвлет-перетворення, та спеціальних, зокрема сингулярного розкладу матриць яскравості пікселів зображення, методів обробки цифрових зображень.

Одним із поширених підходів до проведення пасивного стегоаналізу цифрових зображень є дослідження зміни їх статистичних характеристик внаслідок вбудовування даних [3]. Обмеженням даного підходу є використання окремих характеристик зображення-контейнера при проведенні стегоаналізу, що дає можливість виявляти стеганограми, сформовані лише із використанням «класичних» LSB-методів, наприклад (±1)-алгоритм, MMx-алгоритм.

Для підвищення ймовірності розпізнавання стеганограм, сформованих згідно з адаптивними чи багатоетапними методами, перспективним є використання комплексних статистичних моделей цифрового зображення, отриманих шляхом об'єднання декількох «простих» моделей, зокрема моделей в просторовій (модель SPAM) та в частотній (модель CC-PEV) області зображення.

Література. **1.** FireEye Threat Intelligence (9 July 2015). "HAMMERTOSS: Stealthy Tactics Define a Russian Cyber Threat Group". Retrieved 7 August 2015. **2.** Конахович Г.Ф. Компьютерная стеганография / Г.Ф. Конахович, А.Ю. Пузыренко. – К.: «МК-Пресс», 2006. – 288 с. **3.** Holub, V. Low Complexity Features for JPEG Steganalysis Using Undecimated DCT / V. Holub, J. Fridrich // IEEE Transactions on Information Forensics and Security. – 2015. – vol. 10 (2). – P. 219–228.

Малишев А.І.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Синхронізація анімації обличчя віртуального персонажа із голосом в реальному часі

Вступ. По своїй природі людська мова є бімодальним процесом. Тобто сприйняття мовлення людиною залежить не тільки від акустичних сигналів, але і від візуального їх подання, такого як, наприклад, рух губ (артикулярне подання) та вирази обличчя (емоції). Саме за допомогою мовлення повинна відбуватись передача інформації від віртуального персонажа до людини. Тому візуальне подання звуку є дуже важливою складовою, яка дозволяє зробити віртуальний персонаж наближеним до реального життя. Якщо віртуальна розмова не буде відбуватися плавно, а рухи губ не будуть синхронізовані із голосом людини, – це буде досить важким для сприйняття та виглядатиме незграбно.

Основні положення. Розмова зазвичай представляється як послідовність фонем. Кожна фонема має бути пов'язана з відповідною віземою (візуальне подання руху губ, щелепи, язика та щік при формуванні звуку). Кожна літера у словах впливає на формування певної фонеми, але не всі із них повинні бути видимі. Тому різні фонеми можуть бути пов'язані із однаковими віземами, які попередньо створюються і, власне, формують анімацію мовлення. Пізніше вони об'єднуються в потік і переходи між ними згладжуються блендінгом або певними інтерполяційними схемами, що візуально надає віртуальному персонажу живості мовлення. Головною проблемою цього підходу є синхронізація тривалості звукового потоку та тривалостей фонем разом із відповідними їм віземами [1, 2].

Фонема – це найменша одиниця мовлення, яка може бути протиставлена вислову. В фонетичній мові поєднання фонем, а не власне літер, створюють слова. Для прикладу візьмемо англійське слово "foot". Поєднання літер "oo" повинні бути представлені однією фонемою "UH". Повне фонетичне представлення даного слова матиме вигляд "F-UH-T". При вимові фонем рот змінює змінює свою форму, щоб сформувати їх правильне звучання. Наприклад, слова "bye" та "pie" мають зовсім різне значення і відповідно промовляються як "baі" та "paі" і є дуже схожими. Тут всього дві фонеми, які відрізняються [2].

Взаємозв'язок фонем та візем. Віземи – це візуальні фонеми, які описують положення точок або рухи обличчя під час артикуляції. Вони у анімації обличчя мають залежність "один-до-багатьох" по відношенню до фонем. У наступній таблиці приведено співставлення візем та відповідних їм фонем. За допомогою такого набору можливо реалізувати мовлення віртуального персонажу, яке буде виглядати природно [3].

Табл. 1. Віземи та відповідні їм фонеми

Візема	Фонеми
O	AA, AH, ER, OY
E	AE, EH, EY, IH, IY, OW
A	AH, AW, AY
UQ	AW, OW, UH, UW
FV	AY, EY, F, OY, V, Y
MBP	B, M, P
TS	D, DH, G, JH, S, T, TH, Z, ZH
LN	L, N, NG
WR	R, W
Close	SIL(тиша)
nothing	CH, HH, K, SH(не відображаються)

Алгоритм синхронізації. На діаграмі, яка зображена на рисунку (рис. 1) показано загальний алгоритм синхронізації голосу із анімацією обличчя віртуального персонажу.

Рис. 1. Алгоритм синхронізації анімації обличчя із голосом

Для застосування даного алгоритму на практиці потрібно використати бібліотеки, які надають можливіть в реальному часі розпізнавати звуковий потік та виділяти з нього фонеми та їх тривалості. Такий функціонал надає CMU Sphinx та спрощене рішення для мобільних платформ PocketSphinx [4]. Також заздалегідь повинні бути підготовані анімації візем для віртуального персонажу з відповідними переходами, яких повинно бути 11.

Висновки. При підготовці тез було розглянуто можливі підходи до рішення проблеми синхронізації голосу з анімацією обличчя віртуального персонажа. В результаті було отримано взаємозв'язок фонем англійської мови з віземами, яких всього 11. Запропоновано рішення для розпізнавання голосу. Також представлено спрощену версію алгоритму синхронізації, який, при певних допрацюваннях, може вирішити дану проблему.

Література. **1.** Y. Cao, P. Faloutsos, E. Kohler, and F. Pighin. Real-time speech motion synthesis from recorded motions. In ACM SIGGRAPH/Eurographics symposium on Computer animation, p. 345–353, 2004. **2.** Yuyu Xu, Andrew W. Feng, Stacy C. Marsella, Ari Shapiro. A Practical and Configurable Lip Sync Method for Games, In ACM SIGGRAPH Motion in Games, 2013. **3.** Junho Park, Hanseok Ko. Real-Time Continuous Phoneme Recognition System Using Class-Dependent Tied-Mixture HMM With HBT Structure for Speech-Driven Lip-Sync. IEEE Trans. Mult., 10(7): 1299–1306, 2008. **4.** CMU Sphinx, http://cmusphinx.sourceforge.net/.

***Мегрелишвили Р.П.*[1], *Шенгелия С.В.*[2]**
[1] *Тбилисский государственный университет им. И.Джавахишвили, Тбилиси, Грузия;* [2] *Сухумский государственный университет, Тбилиси, Грузия*

Исследование и синтез матричных структур

Главное найти обобщенное решение задачи. Задача состоит в том, чтобы априори знать для быстродействующего матричного алгоритма необходимое и достаточное условия для составления исходных матриц. Для всех матриц порядок e равно числу мерсенна или максимальному значению (размерность матрицы равно: n, максимальное значение $e = 2^n - 1$). Иследования были проводимы для новых фрактальных матричных структур. Для каждой размерности $n > 1$ исходная $n \times n$ матрица должна генерировать либо максимальное число матриц $(2^n - 1)$, либо это число должно быть числом Мерсена, т.е. $2^j - 1$, где $j < n$; [1,2].

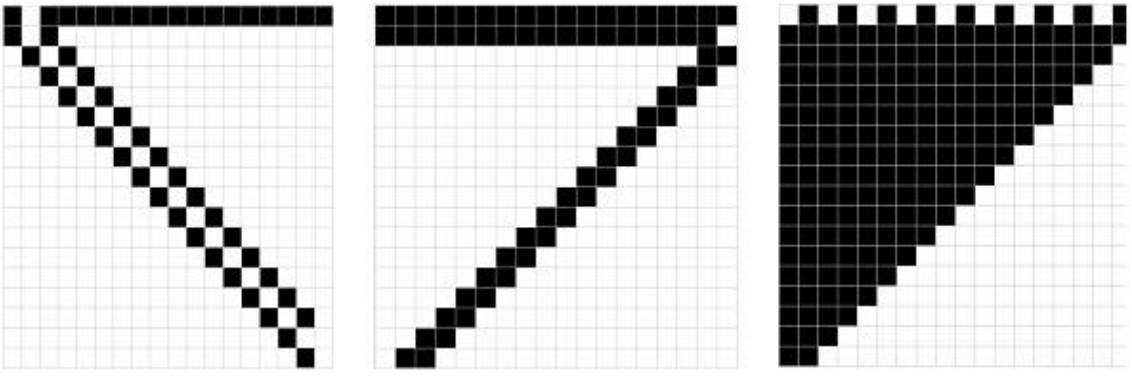


Рис. 1. Новые фрактальные структуры

N_1 Матричные структуры:

$$n = 3, A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}; \quad n = 4, A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix};$$

$$n = 5, A = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

N_2 Матричные структуры:

$$n = 3, A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}; \quad n = 4, A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix};$$

$$n = 5, A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

N_3 Матричные структуры:

$$n = 3, A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}; \quad n = 4, A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix};$$

$$n = 5, A = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Результаты вычисления порядков примитивных $n \times n$ матриц представлены в таблице. Из анализа данных, представленных в таблице для быстродействующего матричного алгоритма образовали матричное множество, который состоит из высокого порядка.

Таблица 1. Результаты вычисления порядков примитивных $n \times n$ матриц

n	e	n	e	n	e	n	e	n	e
3	2^3-1	26	$2^{26}-1$	50	$2^{50}-1$	83	$2^{83}-1$	105	$2^{105}-1$
5	2^5-1	29	$2^{29}-1$	51	$2^{51}-1$	86	$2^{86}-1$	113	$2^{113}-1$
6	2^6-1	30	$2^{30}-1$	53	$2^{53}-1$	89	$2^{89}-1$	119	$2^{119}-1$
9	2^9-1	33	$2^{33}-1$	65	$2^{65}-1$	90	$2^{90}-1$	131	$2^{131}-1$
11	$2^{11}-1$	35	$2^{35}-1$	69	$2^{69}-1$	95	$2^{95}-1$	134	$2^{134}-1$
14	$2^{14}-1$	39	$2^{39}-1$	74	$2^{74}-1$	98	$2^{98}-1$	135	$2^{135}-1$
23	$2^{23}-1$	41	$2^{41}-1$	81	$2^{81}-1$	99	$2^{99}-1$	146	$2^{146}-1$

Изучили новые, разные фрактальные матричные структуры. Они состоят из примитивных матриц, а также они свободны от внутри рекуррентных матриц. Мы создали базу из матриц для безопасности быстродействующего, матричного криптографического алгоритма.

Литература. **1.** Megrelishvili R.P. Analysis of the matrix one-way function and two variants of its implementation, IJMRAE, Vol.5, No.IV, 2013, P. 99–105. **2.** Мегрелишвили Р., Шенгелия С. “Оригинальная матричная однонаправленная функция и вопросы ее реализации” / System Analysis and information technologies. Conference SAIT 2013, Ukraine, Kyiv, P. 465–466, 2013. **3.** Мегрелишвили Р., Шенгелия С. “Фрактальные матрицы в криптографии” / System Analysis and information technologies. Conference SAIT 2014, Ukraine, Kyiv, P. 418–419, 2014. **4.** R. Megrelishvili, S. Shengelia. “Open Channel Key Exchange Algorithm and Fractal Structures Research” / 2015 Tbilisi International Conference on Computer Science and Applied Mathematics, March 21–23, P. 93–97, 2015.

Москаленко В.В., Миронюк И.А.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Построение иерархии ключевых показателей деятельности предприятия в рамках информационной системы управления эффективностью

Наиболее сложными проблемами в управлении предприятием остаются решение задач стратегического уровня. Для их решения используют концепцию управления эффективностью бизнеса. Она охватывает весь спектр задач в области стратегического и финансового управления предприятием, включает в себя такие управленческие технологии как моделирование стратегии, стратегические карты, системы сбалансированных показателей, процессно-ориентированное планирование и функционально-стоимостной анализ, бюджетирование и бизнес-планирование, консолидированную управленческую отчётность и анализ. Эта концепция реализуется в информационных системах класса EPM (Enterprise Performance Management). Основное преимущество EPM-систем – это связанность всех показателей планов в единую иерархию, что позволяет отследить взаимосвязь между стратегическими целями предприятия и ключевыми плановыми показателями. Внедрение такой системы на предприятии позволяет разработать систему планов стратегического, тактического и текущего (оперативного) уровней. В рамках EPM-системы осуществляется мониторинг выполнения планов и оценка эффективности реализации стратегий. Актуальной остаётся проблема каскадирования стратегических показателей до показателей бюджетов. Трудность заключается в больших отличиях стратегического и текущего (годового) планирования. Стратегическое планирование ориентируется на агрегированные показатели и на достижение их плановых значений через несколько лет. А в основе годового планирования и учёта лежит деление деятельности предприятия на основную (операционную), инвестиционную и финансовую на принципах учёта затрат и доходов на коротких промежутках времени – месяц, квартал, год. Предлагается при разработке механизма каскадирования учитывать эту специфику. В работах [1, 2] предложена технология каскадирования стратегических показателей. Предприятие рассматривается как четырёхуровневая система: Центр (руководство предприятия) – бизнес-процессы – структурные подразделения – персонал. Такое разделение по уровням соответствует процессному подходу в управлении. Система управления рассматривается как совокупность управления бизнес-процессами: 1) управленческие бизнес-процессы, которые образуют верхний уровень, здесь формируются стратегические цели, строится стратегическая карта, определяются ключевые показатели на стратегический период (при рассмотрении холдинговой структуры эти показатели определяются ГК); 2) основные и вспомогательные бизнес-процессы, которые охватывают текущую деятельность предприятия. Основные бизнес-процессы осуществляются в рамках операционной деятельности, вспомогательные – обеспечивают непрерывность основных процессов, в качестве обеспечивающих процессов будем рассматривать бизнес-процессы инвестиционной и финансовой деятельности предприятия; 3) структурные подразделения, реализующие бизнес-процессы; 4) исполнители процедур бизнес-процессов – персонал структурных подразделений. Строится система ключевых показателей, которая охватывает все эти уровни [2]. Технология каскадирования включается как отдельный модуль в EPM-систему (рис. 1) [2]. В рамках EPM–системы реализуются основные этапы стратегического планирования и процесс каскадирования. В результате стратегического планирования формируется стратегическая карта и система стратегических показателей деятельности, устанавливаются их целевые значения на конец планового периода. Систему бюджетирования предлагается рассматривать как трёхуровневую систему: уровень стратегических бюджетов, уровень тактических бюджетов, уровень годовых бюджетов. Стратегический уровень бюджетов охватывает последние года стратегического периода, здесь определяются бюджетные показатели за последние года стра-

тегического периода с учетом плановых значений KPI, для KPI и бюджетных показателей строятся функции принадлежности. Тактический уровень содержит значения бюджетных показателей, которые представлены в виде интервальных оценок. Оперативный уровень – это система бюджетирования, которая позволяет построить бюджет предприятия на 1 год – первый год стратегического периода на основе прогнозных данных и на основе плановых значений стратегических KPI.

Рис. 1. Структура EPM–системы

Литература. **1.** Москаленко В.В. Технология формирования показателей эффективности как основа EPM-системы предприятия. / В.В. Москаленко, Н.Г. Фонта // Информационные технологии в управлении, образовании, науке и промышленности: монография / под ред. В.С. Пономаренко. – Х.: Издатель Рожко С.Г., 2016. – С. 476–490. **2.** Москаленко В.В. Структура системы Enterprise Performance Management с учётом технологии каскадирования ключевых показателей деятельности / В.В. Москаленко, Н.Г. Фонта // Вісник Національного технічного університету «ХПІ». Серія «Системний аналіз, управління та інформаційні технології». – Харків: НТУ «ХПІ». – 2016. – № 45. – С. 34–40.

Муха И.П., Вольский А.В.
КПИ им. Игоря Сикорского, Киев, Украина

Анализ и оптимизация структуры сайта

Веб-сайты являются как визитной карточкой компаний и организаций, так и основным источником информации для огромного числа пользователей сети Internet. Поскольку подавляющее большинство сайтов нацелены на ознакомление посетителей с размещенной на них информацией, а иногда и на использование дополнительных функциональных возможностей, таких как покупки, то остро стоят вопросы о привлечении посетителей (целевой аудитории) на сайт и их удержании (переведении посетителей в постоянных пользователей), особенно при наличии конкуренции со стороны других компаний или аналогичных ресурсов.

На выбор пользователя влияют несколько факторов, основные из которых: *уникальность контента, время доступа к сайту и поиска нужной информации, понимание структуры и функциональных действий* и другие.

Уже существует набор требований и рекомендаций по оптимизации сайтов, как рекомендации Google для веб-мастеров [1], или сервисы, которые проводят диагностику сайта и указывают на существующие ошибки (например, Netpeak Spider [2]). Однако рекомендации обычно даются в общем виде и не учитывают специфики сайта. Использование подробной классификации сайтов (например, табл. 1) дает возможность сделать рекомендации по улучшению структуры более точными, однако препятствием является отсутствие единой классификации [3, 4].

Таблица 1. Типы сайтов

№	Тип сайта	№	Тип сайта
1	Сайт-визитка	8	Промо-сайт
2	Корпоративный информационный веб-сайт	9	Корпоративный имиджевый веб-сайт
3	Сайт-форум	10	Блог
4	Интернет-магазин	11	Дорвей
5	Информационный сайт	12	Каталог
6	Игровой портал	13	Сателлит
7	Персональный проект	14	MFA,MFC-сайт

Для создания набора целевых рекомендаций по изменению структуры сайта на начальном этапе можно воспользоваться рекомендациями веб-студий, например «Антула» [5] (табл. 2) с дальнейшим анализом типа сайта и поведения пользователей при изменении структуры.

Таблица 2. Рекомендации по структуре сайта в зависимости от числа его страниц

№	Число страниц	Рекомендованная структура
1	до 40	Одноуровневая
2	40–700	Двухуровневая, допускается 3-й уровень вложенности для 10–15% страниц
3	700+	Трехуровневая, допускается 4-й уровень вложенности для 10–15% страниц

Литература. **1.** Рекомендации Google для опытных веб-мастеров [Электронный ресурс]. URL: https://support.google.com/webmasters/answer/35769?hl=ru (дата обращения: 07.03.2017). **2.** Netpeak Spider: Your personal SEO crawler [Электронный ресурс].URL: https://netpeaksoftware.com/spider (дата обращения: 07.03.2017). **3.** Классификация сайтов в Интернет-маркетинге [Электронный ресурс]. www.seonews.ru/analytics/klassifikatsiya-saytov-v-internet-marketinge/ (дата обращения: 07.03.2017). **4.** Идеология каталога Яндекса [Электронный ресурс]. URL: https://yandex.ru/support/catalogue/basic-info/ideology.html (дата обращения: 07.03.2017). **5.** Сайт веб-студии “Антула” [Электронный ресурс].URL: http://www.antula.ru/index.htm.

***Мухін В.Є.*[1], *Корнага Я.І.*[1], *Яковлева А.П.*[2], *Герасименко О.Ю.*[3], *Козерацький М.С.*[4]**

[1]*КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна;* [2]*КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна;* [3]*Київський національний університет ім. Тараса Шевченка, факультет інформаційних технологій, Київ, Україна;* [4]*КПІ ім. Ігоря Сікорського, ММІ, Київ, Україна*

Комплексний аналіз часу обробки запитів в розподілених базах даних

У СУБД обробка запитів заснована на визначенні кількості записів, які зчитуються або записуються на жорсткий диск та обробляються в оперативній пам'яті. Можна зробити припущення, про те, що системна пам'ять не є достатньо об'ємною для корпоративних баз даних, щоб зберігати весь об'єм даних, який обробляється, тому домінуючими операціями, від яких залежить час виконання запитів, є дискові операції вводу/виводу. У даній роботі ми розглядаємо задачу витрат часу СУБД на обробку запитів.

Для визначення варіантів часу обробки запитів введемо позначення: T_{ni} – час обробки i запиту в БД; t_{1i} – час обробки i запиту при наявності індексу; t_{2i} – час обробки i запиту в оперативній пам'яті (ОП); t_{3i} – час завантаження даних в (вивантаження даних з) ОП з (на) жорсткий диск; t_{4i} – час обробки i запиту без індексу. Відповідно: f – функція використання режиму обробки запиту.

Для запису та пошуку час обробки запитів залежить від різних змінних, тому для пошуку: $t_{s1}=\sum_{j=1}^{k} m_j\ (t_{dj}+t_{vj})$, де m_j – висота дерева пошуку, t_{dj} – час пошуку у вузла дерева, t_{vj} – час пошуку у вузлі дерева; $t_{s2}=\sum_{j=1}^{k} t_{opj}$, де t_{opj} – час пошуку в ОП однієї таблиці з даними; $t_{s3}=\sum_{j=1}^{k} (t_{gj}+t_{cj})$, де t_{gj} – час вибору головки жорсткого диску, t_{cj} – час очікування відповідного сектору; $t_{s4}=\sum_{j=1}^{k} u_j\ t_{pj}$, де u_j – кількість записів, t_{pj} – час перебору запису в таблиці.

Для запису: $t_{r1}=\sum_{j=1}^{k} s_j\ (t_{zj}+t_{pj})$, де s_j – кількість записів в одну таблицю, t_{zj} – час запису, t_{pj}- час перебудови індексу; $t_{r2}=\sum_{j=1}^{k} t_{opj}$, де t_{opj} – час пошуку в ОП однієї таблиці з даними; $t_{r3}=\sum_{j=1}^{k} (t_{gj}+t_{cj})$, де t_{gj} – час вибору головки жорсткого диску, t_{cj} – час очікування відповідного сектору; $t_{r4}=\sum_{j=1}^{k} s_j\ t_{pj}$, де s_j – кількість записів в одну таблицю, t_{pj} – час одного запису в таблиці.

Тоді виникає чотири загальних типи визначення часу обробки запитів:

$$\begin{cases} T_{n1}=t_1+t_2; \\ T_{n2}=t_1+t_2+t_3; \\ T_{n3}=t_2+t_4; \\ T_{n4}=t_2+t_3+t_4. \end{cases}$$ і через функцію: $$\begin{cases} T_{n1}=f(t_1,t_2) \text{ при } t_{s1}<t_{r1}; \\ T_{n2}=f(t_1,t_2,t_3) \text{ при } t_{s1}<t_{r1}; \\ T_{n3}=f(t_2,t_4) \text{ при } t_{s4}>t_{r4}; \\ T_{n4}=f(t_2,t_3,t_4) \text{ при } t_{s4}>t_{r4}. \end{cases}$$

Тому $f(T_{n1},T_{n2},T_{n3},T_{n4})\rightarrow\min$

Дана оптимізаційна задача вирішується багатьма існуючими методами рішення оптимізаційних задач.

Таким чином, запропонований варіант аналіз часу обробки запитів в розподілених базах даних на основі комплексного врахування факторів впливу дозволяє визначити оптимальний режим функціонування механізму пошуку в базах даних.

Література. **1.** Корнага Я.І. Методи підвищення швидкості запису та пошуку даних у базах даних / Я.І. Корнага // Вісник Чернігівського державного технологічного університету серія «Технічні науки». – Чернігів. – 4(61). – 2012. **2.** Корнага Я.І. Порівняльні оцінки застосування методів підвищення швидкості пошуку та запису даних в базах даних / Я.І. Корнага // Міжвідомчий науково-технічний збірник «Адаптивні системи автоматичного управління». – Київ. – 1(22). – 2013. **3.** Мухін В.Є. Механізми пошуку в розподілених базах даних на основі К-дерев / В.Є. Мухін, Я.І. Корнага // HPC-UA 2012, 2 international conference on high performance computing "Proceedings". – Kiev. – 2012.

Негода К.В., Чертов О.Р.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Система ідентифікації психологічного стану дитини за її малюнками

У зв'язку з подіями в Україні багато дітей, що були вивезені з зони АТО, потребують психологічної допомоги. Але первинна діагностика психологічного стану дітей вимагає досить багато часу. Якщо цей процес автоматизувати, то у психолога буде більше вільного часу для роботи з конкретною дитиною, у якої був виявлений незадовільний психологічний стан.

Існують наступні методики та підходи для оцінки психологічного стану дитини.

Проективні методики [1,2]: «Дім – дерево – людина», «Неіснуюча тварина», «Дитина малює родину» тощо. За малюнком можна судити про особистість, її потреби, рівні психологічного розвитку тощо.

Тестологічний підхід [2]. Тест – це система спеціально підібраних завдань, які пропонують дітям у суворо визначених умовах. За виконання кожного завдання дитина отримує оцінку в балах. Правила виставлення балів зазвичай дуже прості і однозначні. Оцінка не повинна залежати від особистого враження людини, яка проводить випробування.

Формуючий експеримент [2], в якому дослідник навмисно створює і видозмінює умови, в яких протікає діяльність дитини, ставить перед нею певні завдання і по тому, як ці завдання вирішуються, робить висновок про психологічні особливості дитини.

Найбільш цікавим для дитини і тому більш релевантним з точки зору отримуваних результатів способом проведення психологічної діагностики є малюнок. На думку класика американської психології Джона Бака (John N. Buck), автора методики «Дім – дерево – людина», кожен малюнок – це своєрідний автопортрет, деталі якого мають особистісне значення. У зв'язку з суб'єктивністю та неоднозначністю оцінок об'єктів на малюнку психологом (великий, середній чи малий розмір об'єкта на малюнку, нижня, верхня частина листа тощо), виникає необхідність об'єктивної оцінки. У даній роботі запропоновано перейти до точних оцінок за допомогою системи нечіткого виведення.

Функції належності для цієї системи у даній роботі побудовані для наступних лінгвістичних змінних [3] (характеристик малюнку та об'єктів, зображених на ньому): розмір об'єкту, розташування об'єкту, стан об'єкта (скажімо, вікно відчинено чи закрито), колір. Функції належності складені на основі експертних оцінок спеціалістів у галузі психології. Приклад однієї з таких функцій, що використовуються у роботі:

$$\mu_{\text{нижня частина листа}}(x) = \begin{cases} 1, & x \leqslant 0,3, \\ \dfrac{0,5-x}{0,2}, & 0,3 < x \leqslant 0,5, \\ 0, & x > 0,5. \end{cases}$$

Правила для системи нечіткого виведення у даній роботі складені на основі інструкцій з інтерпретації кожного об'єкта на малюнку. Наприклад:

ЯКЩО об'єкт у нижній частині листа ТА дуже малий, ТОДІ відчуття незахищеності;

ЯКЩО об'єкт у верхній частині листа, ТОДІ мрійливість;

ЯКЩО об'єкт точно у центрі листа, ТОДІ відчуття незахищеності, прямолінійність.

Розроблену систему можна впровадити для первинної діагностики психологічного стану дітей в ігровій формі на основі інтерпретації їхніх малюнків.

Література. **1.** Шапарь В.Б., Шапарь О.В. Практическая психология. Проективные методики / В.Б. Шапарь, О.В. Шапарь. – Ростов н/Д: Феникс, 2006. – 480 с. **2.** Gromova C.R., Zakirova V.G. Psychological and pedagogical conditions of primary schoolchildren civil identity development (results of complex research). Life Sci J 2014;11(10s): 518–521. **3.** Tsumoto S., Hirano S. (2016) Linguistic Variables Construction in Information Table. In: Flores V. et al. (eds) Rough Sets. IJCRS 2016. Lecture Notes in Computer Science, vol 9920. Springer.

Нестюк О.М., Ліщук К.І.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Технічна підтримка продукту з використанням автоматизованого пошуку відповідей за змістом та чат-ботів

Наведено опис методу реалізації першої лінії технічної підтримки продукту з використанням питально-відповідальної системи та інтелектуальних чат-ботів.

Сучасні темпи розвитку все більше вимагають максимальної автоматизації тих частин роботи підприємства, які взагалі можуть бути автоматизовані. Це дозволяє звільнити людські ресурси від монотонної та простої праці і перекласти ці обов'язки на машину. Необхідність такої автоматизації вимагає створення нових інструментів, які могли б автоматизувати не лише ті процеси розробки, які мають чітку послідовність дій та умов їх виконання. Одним з таких процесів є технічна підтримка продукту.

Опис проблеми. Для реалізації першої лінії технічної підтримки на сьогодні, як правило, використовуються оператори, які взаємодіють з клієнтом у вигляді діалогу через телефон, чат, систему підтримки чи інше. Робота таких операторів у більшій частині зводиться до транслювання користувачеві відповідей на питання зі спеціально підготовленої інструкції [1]. Враховуючи таку ситуацію, можна зробити висновок, що перший рівень технічної підтримки може бути ефективно спрощеним за рахунок інтеграції програмного забезпечення для автоматичного пошуку відповідей за змістом. Таке програмне забезпечення може стати своєрідним «рівнем 0» технічної підтримки.

Відомі рішення. Відомими рішеннями для часткової автоматизації процесу надання інформації по роботі системи, є так звані FAQ (Frequently Asked Questions) – часто поставлені, поширені питання, підбірка питань, які часто озвучують клієнти та відповіді на них [2]. Такі FAQ розміщуються на сайтах продукту та дають можливість користувачеві самостійно знайти відповідь на питання, що його хвилює. Відомими системами, що можуть ділитися інформацією з користувачем в режимі діалогу є голосові помічники, як Google Assistant, Apple Siri чи Microsoft Cortana. Але такі системи мають значну ширшу спеціалізацію та закриту архітектуру, що не дозволяє використати їх для задач автоматизації технічної підтримки.

Пропоноване рішення. Для зменшення навантаження на першу лінію технічної підтримки пропонується розробити механізм автоматизованого пошуку відповідей на питання користувачів в вигляді комбінації питально відповідальної системи [3] та чат-боту, що задовольнить користувачів як якістю отримуваної інформації, так і способом комунікації. Даний механізм повинен відповідати наступним критеріям:

1. сфера знань обмежена продуктом, для якого виконується технічна підтримка;
2. база даних відповідей не є кінцевою та незмінюваною, дані в ній можуть змінюватись та поповнюватись;
3. система повинна мати можливість повідомити про відсутність необхідної відповіді та з'єднати з технічним спеціалістом, оскільки даних в базі може бути недостатньо для вирішення абсолютно всіх можливих ситуацій;
4. для покращення роботи та підвищення якості відповідей необхідний зворотний зв'язок, система повинна знати, чи була конкретна відповідь на конкретне питання правильною для користувача;
5. комунікація з користувачем повинна бути в режимі діалогу та з використанням природної мови.

Література. **1.** Phillip J. Windley, Delivering High Availability Services Using a Multi-Tiered Support Model. **2.** FAQ, https://uk.wikipedia.org/wiki/FAQ. **3.** Abdelghani Bouziane, Djelloul Bouchiha, Noureddine Doumi, Mimoun Malki, Question Answering Systems: Survey and Trends. Procedia Computer Science 73 (2015) 366–375.

***Орєхов О.А.*[1], *Орєхова Н.А.*[2]**
[1]*КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна;* [2]*Інститут кібернетики ім. В.М.Глушкова НАН України, Київ, Україна*

Дослідження властивостей бібліотеки алгоритмів глибокого навчання Keras на прикладі задачі компрометації нейронної мережі

Досліджуються властивості нової бібліотеки алгоритмів глибокого навчання Keras (надбудови над відомими та все більш популярними бібліотеками tensorflow та Theano) та розглядається приклад її застосування для дослідження сценаріїв компрометації нейронної мережі при незначному спотворенні зразків, що надаються для її навчання. Робиться висновок про певне спрощення процесу навчання більш стійкої до атаки на зразки навчання нейронної мережі.

Бібліотека Keras [1, 2] набуває швидкої популярності як засіб, що спрощує роботу з класом задач глибокого навчання нейронних мереж за рахунок надання інтуітивно зрозумілого інтерфейсу до інших бібліотек, щр також були запропоновані в останні роки і вже набули широкої популярності: deeplearning4j [3], tensorflow [4] та Theano [5]. Слід відзначити, що власне бібліотека не скорочує час розрахунків, але спрощує процес прийняття рішень на етапах адаптивної модифікації проміжних результатів. Також слід зазначити, що бібліотеку створено в CPU- та GPU-версіях для основних відомих операційних систем.

Бібліотеку Keras було використано для часткового відтворення результатів досліджень, описаних в [6] та [7]. Змістом оригінальних досліджень було навчання за стандартною процедурою глибоких нейронних мереж (DNN – Deep Neural Networks) та надання певним чином змінених зразків (adversarial samples – "зразків атаки") оракулу [8] з фіксацією наступної помилкової класифікації. Метою було виявлення закономірностей настання випадків помилкової класифікації. Також вивчався ефект передачі "зразків атаки" від одного оракула іншому без втручання у внутрішню структуру, тільки за рахунок вивчення відгуків, отриманих при навчанні.

Основні висновки роботи:

- бібліотека Keras надає всі можливості для зазначених обчислювальних експериментів;
- обмеження на наявні обчислювальні ресурси не дало змоги підтвердити статистику компроментації у 84 відсотки, але було достатнім для відтворення факту компрометації оракула;
- порівняльний аналіз результатів застосування алгоритму Гудфеллоу (Goodfellow) та алгоритму Паперно (Papernot) в цілому збігається з висновком [8] стосовно того, що перший є більш зручним для генерації великої кількості "зразків атаки", але може досить швидко привернути увагу захистника, в той час, як другий вимагає більш складних обчислень, одночасно зменшує спотворення, потрібні для побудови результативного "зразку атаки";
- застосування двох вищезгаданих алгоритмів у комбінації може покращити підсумкові результати у порявнянні з кожним з них, що застосовується окремо.

Література. **1.** F. Chollet Deep Learning with Python MEAP V.2 Manning Publications. **2.** https://keras.io. **3.** https://deeplearning4j.org. **4.** https://tensorflow.org. **5.** http://deeplearning.net/software/theano/. **6.** Papernot, N., McDaniel, P., Jha, S., Fredrikson, M., Celik, Z.B., and Swami, A. The limitations of deep learning in adversarial settings. In Proceedings of the 1st IEEE European Symposium on Security and Privacy (2016), arXiv preprint arXiv:1511.07528. **7.** Papernot, N., McDaniel, P., Godfellow, I., Jha, S., Berkay Celik, Z., Swami. A., 'Practical Black-Box Attacks against Deep Learning Systems using Adversarial Examples' (2016) arXiv preprint arXiv:1602.02697v2. **8.** https://metamind.io.

Осиюк Д.В., Шубенкова И.А.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Сквозная аналитика для проектов сферы электронной коммерции на основе Google Analytics

Торговля товарами через интернет – это наиболее успешно развивающаяся на сегодняшний день отрасль в Украине. Люди все чаще покупают через интернет и интернет-магазины. Если за 2015 год товарооборот увеличился на 32% – с 19 до 25 млрд грн, то в 2016 году он вырос уже на 52% – до 38,5 млрд грн. [1].

Используемая Интернет-реклама является одним из самых дешевых, но не всегда эффективным рекламным инструментом продвижения, поскольку не используются все возможности современных рекламных систем и систем веб-аналитики.

Рекламные системы, такие как Google AdWords, Facebook Ads, Реклама VK.com, ADMIXER помогают сегментировать аудиторию интернета по различным признакам, таким как:

- пол/возраст;
- интересы пользователя;
- местоположение;
- типы используемых устройств;
- операционные системы, браузеры;
- источники трафика и множество других.

Признаки определяются с помощью технологий машинного обучения и нейронных сетей, а также на основе тех данных, которые собирает сама рекламная система. Например, Google AdWords может определять характеристики пользователей на основании поисковых запросов в системе Google, а Facebook умеет распознавать предпочтения и интересы на основании данных о том, на каких публикациях задерживается пользователь, что он комментирует и отмечает, как понравившееся. Все это позволяет качественно таргетировать рекламу в интернете на целевую аудиторию. Кроме этого, важно учитывать и поведение пользователя на сайте, анализировать его движение по воронке продаж.

Google Analytics бесплатный сервис по сбору детальной статистики посетителей веб-сайтов. Статистика собирается на сервере Google, пользователь только размещает JS-код на страницах своего сайта. Код отслеживания срабатывает, когда пользователь открывает страницу в своем веб-браузере. Google Analytics позволяет узнать, о том по каким ссылкам переходят на ваш сайт посетители, как долго они остаются на сайте, где они находятся географически и какие действия совершают на сайте [2].

Когда на сайте есть возможность оплатить заказ онлайн – можно настроить отслеживание таких транзакций через блок расширенной электронной торговли в Google Analytics. Кроме самих транзакций, это позволит отслеживать:

- Product Impressions / Просмотры товаров в каталоге.
- Product Clicks / Клики по товарам.
- Views of Product Details / Просмотры карточек товаров.
- Adding a Product to a Product Cart / Добавление товаров в корзину.
- Removing a Product from a Product Cart / Удаление товара из корзины.
- Checkout Steps / Шаги оформления заказа.
- Checkout Options / Варианты оформления заказа.
- Purchases / Совершенные покупки.
- Refunds / Отмены транзакций.
- Promotion Impressions / Показы внутренней рекламы.
- Promotion Clicks / Клики по внутренней рекламе.

Большинство проектов в сфере электронной торговли еще не смогли перенести все бизнес-

процесы в онлайн. В таких случаях, обычно заказ оформляется на сайте, но дальше его статус не известен в системах веб-аналитики. Такие заказы могут быть оплачены, изменены, либо отменены по какой-либо причине. Как тогда оценивать эффективность рекламных кампаний и как эффективно перераспределить бюджет для увеличения прибыли? Как отслеживать реальный доход от рекламы и считать ROI для отдельных рекламных каналов?

Сейчас подобные отчеты создаются путем сведения данных изо всех систем (рекламные сервисы, системы веб-аналитики, CRM/ERP) в Excel или другую среду обратотки информации. Это достаточно долгий и трудоемкий итерационный процесс, который необходимо автоматизировать. Это можно сделать путем интеграции в единую систему рекламных сервисов, CRM систему и систему веб-аналитики чтобы получить хранилище данных, которое максимально точно описывает весь путь пользователя от клика по рекламному объявлению до получения товара (оплаты заказа). Это позволит компаниям перейти на существенно новый уровень эффективной бизнес-аналитики.

В новой версии Universal Analytics реализован протокол передачи статистических данных Google Analytics (Measurement Protocol) позволяет отправлять необработанные данные напрямую на серверы Google Analytics посредством HTTP-запросов практически в любой среде. С помощью этого протокола разработчики могут решать следующие задачи:

- отслеживать, как пользователи взаимодействуют с компанией, в новых средах (например, в платежных терминалах, мобильных приложениях и т.д.);
- связывать онлайн- и офлайн-действия, чтобы получить полную картину поведения клиентов;
- отправлять данные как с веб-сайтов, так и с сервера;
- отслеживать телефонные звонки в компанию по номеру на сайте.

Гибкость настройки Google Analytics позволяет собирать данные из разных систем в единой структуре. Такой результат достигается за счет:

- API рекламных сервисов;
- Трекера Google Analytics;
- Measurement Protocol.

Таким образом на основе Google Analytics можно реализовать систему сквозной аналитики, которая позволяет динамически управлять маркетинговыми коммуникациями в интернете для максимизации дохода и повышения окупаемости инвестиций в рекламу (ROI).

Такая система позволит тестировать изменения в интерфейсе сайта, бизнес-процесах в привязке к реальному доходу и количеству состоявшихся сделок. Принятие решений в бизнесе можно будет осуществлять, опираясь на полные данные.

Результаты анализа рекламных кампаний и контроль KPI по проекту можно предоставлять в бесплатном BI-инструменте Google Data Studio. Этот инструмент позволяет создавать интерактивные автоматические отчеты, а широкий выбор элементов визуализации позволяет визуализировать данные в удобном виде для лиц, принимающий решения на любом уровне управления проектом.

Разрабатываемая система призвана решить все выше описанные проблемы. Ее настройка позволит сэкономить время на подготовку отчетов по эффективности использования рекламы в интернете за счет автоматизации передачи данных в единую систему, что в свою очередь позволит быстро принимать решения по оптимизации рекламных кампаний и перераспределять рекламные бюджеты для максимизации прибыли. Эта система позволит эффективно использовать сайт для повышения конверсии и увеличения продаж.

Литература. **1.** Что ожидает украинский рынок e-commerce в 2017 году – прогнозы EVO, ain.ua/2016/12/16/e-commerce-v-2017-evo. **2.** Google Analytics, ru.wikipedia.org/wiki/Google_Analytics.

***Павлов О.А.*[1], *Місюра О.Б.*[1], *Мельников О.В.*[1], *Лисецький Т.М.*[2]**
[1]*КПІ ім. Ігоря Сікорського, Київ, Україна;* [2]*ПАТ "КП ОТІ", Київ, Україна*

Система календарного та оперативного планування і прийняття рішень для підприємств дискретного типу

В [1–3] розглядається проблема реалізації ефективного календарного та оперативного планування в системах, що мають мережеве представлення технологічних процесів і обмежені ресурси. Такі системи складають 80% від всіх виробничих систем, до яких відносяться, зокрема, виробництва дискретного типу (у тому числі літакобудування та суднобудування) та, що найбільш важливо в сучасних умовах, виробничі та конструкторські підприємства військово-промислового комплексу України.

Авторами створена система планування (включаючи оперативне) та прийняття рішень для підприємств дискретного типу у вигляді інтегрованого пакету програм, до складу якого входять інформаційна система підтримки прийняття рішень та універсальна ієрархічна система календарного та оперативного планування виробництв дискретного типу. Система реалізована як програмне забезпечення, налаштоване для впровадження без суттєвих змін у різних прикладних областях.

У створеній системі поставлена проблема вирішується через побудову узгоджених планів виробництва для всіх підрозділів підприємства і пооперaційного плану, що реалізує потенційний портфель замовлень і є оптимальним за критерієм, заданим замовниками.

В основу системи покладена інноваційна чотирьохрівнева модель календарного та оперативного планування в мережевих системах [3].

Постановка задачі. Маємо мережу, що відображає реальну технологію виконання множини комплексів взаємозв'язаних робіт (виробів); по кожній роботі задається питома тривалість виконання для одного виробу; по кожному виробу задається кількість виробів у портфелі замовлень, вага, директивний строк (або припустимий інтервал часу); для кожного ресурсу вказується його тип і число паралельних приладів, продуктивності приладів. Вироби в мережі групуються за критеріями оптимізації, між такими групами виробів не повинно бути виробничих зв'язків. Потрібно побудувати пооперaційний план виконання робіт з розподілом їх по ресурсах з урахуванням одного з 31 критеріїв оптимальності (п'яти базових і 26 синтетичних критеріїв, що є всіма можливими лінійними комбінаціями базових).

Загальним критерієм оптимальності для всіх рівнів моделі планування є максимізація прибутку підприємства, що є головною метою планування в умовах ринку. Базові критерії:

1. Максимізація сумарного прибутку підприємства у випадку відсутності директивних строків, що еквівалентно до мінімізації сумарного зваженого моменту закінчення виконання виробів (МЗМ) при відношенні порядку, заданому орієнтованим ациклічним графом, побудованим на критичних шляхах виробів.
2. Максимізація сумарного прибутку підприємства за умови: для всіх виробів введені директивні строки, які не можуть бути порушені (планування «точно в строк»).
3. Максимізація сумарного прибутку підприємства за умови: для всіх виробів введені директивні строки, необхідно мінімізувати сумарний штраф за запізнювання виконання виробів відносно директивних строків.
4. Максимізація сумарного прибутку підприємства за умови: для всіх виробів введені директивні строки, для кожного виробу зазначена величина прибутку від виконання виробу у випадку його виконання без запізнювання відносно директивного строку.
5. Для всіх виробів задані директивні строки. Необхідно мінімізувати сумарний штраф підприємства як за випередження, так і за запізнювання відносно директивних строків.

У чотирьохрівневій моделі планування (включаючи оперативне) і прийняття рішень [2, 3] формалізуються як процедура одержання календарного пооперaційного плану, так і формальні процедури його оперативного коректування у випадку його порушення в процесі реалізації. Ієрархічність моделі планування дозволила ефективно використовувати ряд NP-складних

задач календарного планування, ефективне розв'язання яких було обґрунтовано авторами в рамках оригінальної теорії ПДС-алгоритмів, що розвивається ними [1, 2]. ПДС-алгоритми (алгоритми з поліноміальною складовою та достатніми ознаками оптимальності) дозволяють точно розв'язувати NP-складні задачі комбінаторної оптимізації суттєво більшої розмірності, ніж існуючі методи, а при невиконанні ознак оптимальності – отримувати ефективний точний розв'язок експоненціальним підалгоритмом або евристичний чи наближений розв'язок з оцінкою відхилення від оптимуму для індивідуальних задач великої розмірності. Модель також включає ефективну формальну процедуру розв'язання слабко формалізованої задачі вибору найкращого плану із множини планів, побудованих за різними критеріями оптимальності. Алгоритмічне забезпечення моделі складається з п'яти блоків.

Блок 1. Блок прийняття рішень (підсистема ERP- або MRP II-системи підприємства або зв'язок з ERP-системою) має функції: формування портфеля замовлень; узгодження із замовником при необхідності всіх змін плану і нового портфеля замовлень; вибір найкращого плану за критерієм мінімізації ризику неотримання розрахункового прибутку; визначення моменту коректування поопераційного плану та ін.

Блок 2. Попереднє планування. На вхідній мережній моделі виконується дворівнева агрегація: перший рівень – об'єднання робіт і ресурсів; другий рівень – пошук критичних шляхів виробів і побудова мережі на критичних шляхах виробів із спільними вершинами. Далі для кожної базової складової синтетичного функціоналу розв'язується задача МЗМ, у результаті одержуємо оптимальні послідовності виконання агрегованих робіт з кожної групи виробів.

Блок 3. Узгоджене планування з урахуванням заданого критерію оптимальності. Узгоджені плани виконання агрегованих робіт будуються для кожної можливої черговості виконання груп виробів, при цьому агреговані роботи призначаються на ресурси агрегованої мережі. В блоці прийняття рішень з побудованих планів вибираються найкращі за значенням функціонала і за критерієм мінімізації ризику неотримання розрахункового прибутку.

Блок 4. Точне планування. Виконується дезагрегація агрегованих робіт, будується багатоетапна мережна задача календарного планування з критерієм максимізації моменту початку виконання виробів при умовах мінімізації сумарного випередження і відсутності запізнення відносно кінцевих директивних строків. Розв'язання цієї задачі полягає у визначенні директивних строків кожної роботи та розв'язанні одноетапних задач календарного планування на елементах мережі за допомогою ПДС-алгоритмів, описаних в [2], при цьому одержуємо моменти початку і закінчення виконання всіх робіт мережі (поопераційний план).

Блок 5. Оперативне планування. У випадку часткового невиконання плану за допомогою оригінальних методів будується новий поопераційний план.

У разі впровадження створених методів та технологій економічний ефект буде досягнутим завдяки оптимізації планів виробництва з урахуванням критеріїв максимізації прибутку підприємства та мінімізації штрафів за затримку поставок готових виробів. Тому застосування нових ефективних моделей і методів планування, що засновані на нових ПДС-алгоритмах розв'язання задач календарного та оперативного планування, сприятиме підвищенню конкурентоспроможності України як на внутрішньому, так і на зовнішньому ринках та її обороноздатності.

Література. **1.** Згуровский М.З. Принятие решений в сетевых системах с ограниченными ресурсами: Монография [Текст] / М.З. Згуровский, А.А. Павлов. – К.: Наук. думка, 2010. – 573 с. **2.** Згуровский М.З. Труднорешаемые задачи комбинаторной оптимизации в планировании и принятии решений: Монография [Текст] / М.З. Згуровский, А.А. Павлов. – К.: Наук. думка, 2016. – 716 с. **3.** Згуровский М.З. Методология построения четырехуровневой модели планирования, принятия решений и оперативного планирования в сетевых системах с ограниченными ресурсами [Текст] / М.З. Згуровский, А.А. Павлов, Е.Б. Мисюра та ін. // Вісник НТУУ "КПІ". Серія «Інформатика, управління та обчислювальна техніка». – К.: "ВЕК+", 2014. – №61. – С. 60–84.

Погорелов В.В., Марченко О.І.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Використання графа залежностей значень і станів для трансляції програм

У галузі проектування та розробки сучасних компіляторів можна виокремити наступні тенденції розвитку внутрішніх форм представлення програми (ВФПП) [1]:

1. Інформація про оператори присвоєння набуває другорядної ролі.
2. Потік управління відображається непрямим способом.
3. Залежності між даними стають чітко вираженими.

На практиці у якості ВФПП найчастіше виступає граф потоку управління (ГПУ) [2]. ГПУ – це орієнтований граф, вершини якого є одиничними операторами чи базовими блоками. Базовий блок – це послідовність операторів з однією точкою входу та однією точкою виходу, що не містить операторів умовних та безумовних переходів. Дуги графа відображають можливі шляхи виконання програми. ГПУ, по суті, є відображенням структури потоку виконання операторів у програмі. Тобто залежності між даними виражені неявним чином. Це, у свою чергу, ускладнює проведення оптимізацій пов'язаних з аналізом потоку даних.

Недоліки ГПУ можна усунути, перетворивши цей граф у граф залежностей значень і станів (ГЗЗС) [3]. ГЗЗС – це дводольний граф, що може бути визначений наступним чином: $G = (N, E_V, E_S, L, N_0, N_\infty)$. Складовими частинами цього графа є:

1. N – множина вершин графа.
2. E_V – множина дуг графа ($E_V \subseteq N \times N$), що відповідають залежностям значень.
3. E_S – множина дуг графа ($E_S \subseteq N \times N$), що відповідають залежностям стану.
4. L – множина функцій, що асоціюють кожну вершину графа з відповідним оператором.
5. N_0 – початкова вершина графа.
6. N_∞ – кінцева вершина графа.

Цикли та умовні оператори задають у ГЗЗС за допомогою вершин μ та γ.

ГЗЗС є паралельним представленням, у якому відношення порядку між операціями задається частково і базується лише на залежностях за даними та станами. Тобто, на відміну від ГПУ, ГЗЗС передбачає обчислення значень лише у разі потреби. Звідси слідує, що обчислення ГЗЗС може бути таким, що завершилось, навіть якщо виконання вхідної програми буде нескінченним. Ця особливість графу пояснюється тим, що значення, обчислення яких відбувається під час виконання програми але які не потрібні для побудови ГЗЗС, не обчислюються.

У ГЗЗС непрямим чином забезпечується властивість однократного присвоєння, оскільки існує однозначна відповідність між операторними вершинами та вершинами-значеннями. Кожній вершині-значенню можна присвоїти ім'я, що буде унікальним і відповідатиме змінній програми у формі статичного однократного присвоєння.

У порівнянні з ГПУ ГЗЗС надає переваги у наступних аспектах:

1. Відсутні побічні ефекти.
2. Залежності між місцями використання змінних та значеннями змінних в цих місцях виражені однозначно.

Література. **1.** О.І. Марченко, В.В. Погорелов. Використання розширеної форми статичного однократного присвоєння для трансляції програм. – 18-th International conference on System Analysis and Information Technology SAIT 2016, May 30 – June 2, 2016 Institute for Applied System Analysis of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. **2.** В.В. Погорелов, О.І. Марченко. Огляд внутрішніх форм представлення програми для трансляції з процедурних мов програмування у функціональні мови. – Комп'ютерно-інтегровані технології: освіта, наука, виробництво. – Фаховий науковий журнал Луцького НТУ, Випуск 23. – 2016. **3.** Lawrence A.C. Optimizing compilation with the Value State Dependence Graph. – 2007. – 183 p.

Попеляев Д.П.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Подходы к распознаванию огня на изображении. Сверточные нейронные сети

В наши дни, огонь по-прежнему является серьезной угрозой для безопасности человеческой жизни. В сочетании с объективными условиями глобального потепления, эксперты прогнозируют, что лесные пожары станут серьезной угрозой для экосистемы в будущем. Не меньшую же опасность несут пожары в жилых домах и приусадебных участках. Поэтому задача раннего выявления огня остается важной, так как ее решение имеет прямое отношение к сохранности здоровья и материальной собственности.

Задача раннего обнаружения огня. В большинстве случаев задача раннего обнаружения огня делегирована специальным датчикам, таким как датчикам-сигнализаторам угарного газа, датчикам изменения температуры, инфракрасным датчикам пламени, датчикам задымленности, светочувствительным датчикам и даже специальным датчикам звука. Тем не менее большая часть этих специальных средств подвержены серьезным проблемам. Прежде всего, для срабатывания они требуют непосредственной близости к источнику возгорания. Вдобавок, они не всегда есть надежными, так как они не всегда реагируют на возгорание, их триггерами являются непосредственно продукты горения, такие как температура, угарный газ, излучаемый свет. А также большинство их них ограничены применением лишь внутри помещения. С развитием технологий, стали появляться и другие методы распознавания открытого огня. Современные подходы к распознаванию объектов по большей части связаны с методами machine learning (машинного обучения). В рамках этого подхода уже разрабатывались системы для обнаружения огня. Основным принципом их действия был анализ данных, которые получались с упомянутых ранее датчиков. Данные напрямую поступали на входной слой некоторой искусственной нейронной сети, которая имела один скрытый слой а на выходе сигнализировала о признаках наличия огня в помещении. Этот подход является более прогрессивным так как достаточно гибко комбинирует данные с нескольких датчиков. Но ему присущи те же недостатки что и самим датчикам.

Распознавание огня с графического изображения. С набирающей популярность инсталяцией систем видеонаблюдения, задача обнаружения огня получила еще одно возможное решение – распознавание огня с графического изображение и видеоряда. Существует хорошо описанный и изученный подход – нахождение на изображении с помощью специальных фильтров и алгоритмов так называемых fire pixels, а затем анализ их расположения, количества, концентрации [1]. Так же были работы, связаны с распознаванием огня с помощью классического перцептрона, который являет собой полносвязную нейронную сеть, каждый нейрон связан со всеми нейронами предыдущего слоя, причем каждая связь имеет свой персональный весовой коэффициент.

Сверточные нейронные сети (convolutional neural network, CNN). Более современным решением для классификации объектов на изображении является использование свёрточных искусственных нейронных сетей (convolutional neural network, CNN). В свёрточной нейронной сети, в отличии от перцептрона, в операции свёртки используется лишь ограниченная матрица весов небольшого размера, которую «двигают» по всему обрабатываемому слою (в самом начале – непосредственно по входному изображению), формируя после каждого сдвига сигнал активации для нейрона следующего слоя с аналогичной позицией. То есть для различных нейронов выходного слоя используются одна и та же матрица весов, которую также называют ядром свёртки. Обычно сеть состоит из множества сенсорных элементов (входных узлов), которые образуют входной слой, одного или нескольких скрытых слоев вычислительных нейронов и одного выходного слоя нейронов. Идея свёрточных нейронных сетей заключается в чередовании свёрточных слоев (convolution layers) и субдискретизирующих слоев (subsampling layers).

Структура сети – однонаправленная (без обратных связей), принципиально многослойная. Для обучения используются стандартные методы, чаще всего метод обратного распространения ошибки. Работа свёрточной нейронной сети обычно интерпретируется как переход от конкретных особенностей изображения к более абстрактным деталям, и далее к ещё более абстрактным деталям вплоть до выделения понятий высокого уровня. При этом сеть самонастраивается и вырабатывает сама необходимую иерархию абстрактных признаков (последовательности карт признаков), фильтруя маловажные детали и выделяя существенное [2]. Основной идеей является решение задачи классификации огня с помощью этого более современного подхода с его уникальными свойствами.

Программная среда для разработки и обучения искусственной нейронной сети. Разработка программного решения для классификации открытого огня с помощью CNN подразумевает выбор фреймворка. По мере того как популярность CNN увеличилась, возникли различные реализации. Наиболее важными факторами фреймворков для глубокого обучения являются их производительность, удобство и простота использования, которая включает наличие интерфейсов API для различных языков программирования, качества документации и деятельности сообщества. Производительность является критически важным требованием, поскольку CNN являются сложными моделями со многими значащими параметрами.

Популярными фреймворками для глубокого обучения являются Torch, Theano и Caffe. Torch является фреймворком для научных вычислений, который реализует CNN, а также другие алгоритмы машинного обучения с использованием языков программирования C и Lua, где Lua используется в качестве языка сценариев. Он разработан и используется для Facebook AI, Google Search и Twitter DeepMind.

Caffe является фреймворком специализированным для глубокого обучения с использованием C++ для реализации и протокола Google Protobuf для определения сетевых архитектур и контроля обучения. Кроме того, Caffe предоставляет интерфейсы API для Python и Matlab. Оба фреймворка обеспечивают реализацию для обучения CNN на CPU и GPU.

Theano значительно отличается от Torch и Caffe. Это в основном библиотека Python и таким образом, также реализована на Python. Theano не является непосредственно библиотекой для реализацией только CNN, а более общей библиотекой, которая может быть легко использована для реализации модели CNN.

Выводы. В данной работе Caffe был выбран в качестве фреймворка для глубокого обучения, из-за его большой производительности, хорошей документации, простоты использования и развертывания а также модульности. В качестве модели нейронной сети была выбрана свёрточная искусственная нейронная сеть (convolutional neural network, CNN). Функция активации нейронов (передаточная функция) – сигмоидальная функция. Количество скрытых слоев и их очередность будет выплывать из конкретных результатов обучения сети. В качестве первоначальной конфигурации сети была взята классическая архитектура CNN: входной слой, слой свертки, блок линейной ректификации, субдискретизирующий слой, блок линейной ректификации, слой свертки, блок линейной ректификации, субдискретизирующий слой, выходной слой. Был взят публичный датасет из 5876 помеченных RGB изображений, который состоял из 2163 изображений с присутствием огня и 3713 нейтральных изображений. Было создано 3 поднабора: тренировочный (60%), валидиующий (20%) и тестовый (20%). Был использован метод стохастического градиентного спуска (SGD). Точность классификации которая была показана на тестовом наборе – 74.3%. В дальнейшем конфигурация сети и параметры обучения будут изменены с целью добиться лучших результатов.

Литература. **1.** Wen-Bing Horng, Jian-Wen Peng. A Fast Image-Based Fire Flame Detection Method Using Color Analysis. Tamkang Journal of Science and Engineering, Vol. 11, No. 3, pp. 273–285, 2008. www2.tku.edu.tw/~tkjse/11-3/11-3-6.pdf. **2.** Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25, pages 1097–1105. Curran Associates, Inc., 2012.

Прогонов Д.О.
КПІ ім. Ігоря Сікорського, ФТІ, Київ, Україна

Вплив невідповідності областей приховання повідомлень та проведення стегоаналізу на ефективність статистичних стегодетекторів

Вступ. Широке використання зловмисниками та терористами прихованих (стеганографічних) систем зв'язку для несанкціонованої передачі конфіденційних даних державних установ та приватних корпорацій обумовлює актуальність і важливість задачі створення високоточних методів виявлення прихованих повідомлень (стеганограм). Для вирішення даної задачі було запропоновано значну кількість алгоритмів статистичного [1,2] та структурного [3,4] стегоаналізу мультимедійних даних, зокрема цифрових зображень (ЦЗ). При розробці даних методів зазвичай використовується стандартне припущення щодо наявності апріорних даних відносно області вбудовування стегоданих у зображення-контейнер (ЗК), що дозволяє створювати цільові методи виявлення стеганограм. Проте у випадку використання стегодетекторів у реальних умовах, зокрема у системах моніторингу інформаційно-комунікаційних мереж, інформація щодо області приховання повідомлень є невідомою. Тому становить інтерес дослідження впливу невідповідності області приховання повідомлень та області проведення стегоаналізу ЦЗ на ефективність сучасних методів виявлення стеганограм, зокрема статистичних стегодетекторів.

Метою роботи є оцінка ефективності застосування стегодетектору на основі статистичної моделі зображення-контейнеру у частотній області (моделі DCTR) для виявлення стеганограм, сформованих згідно адаптивних методів вбудовування стегоданих у просторовій області ЗК.

Статистична модель DCTR цифрових зображень. Модель DCTR (Discrete Cosine Transform Residual) була запропонована у роботі [5] для скорочення тривалості налаштування стегодетектору при збереженні точності виявлення стеганограм, співставної з випадком використання потужних статистичних моделей ЦЗ, зокрема CC-JRM та J+SRM. Особливістю моделі DCTR є використання коефіцієнтів $\boldsymbol{U}^{(k,l)}, (0 \leqslant k,l \leqslant 7)$ блочного (undecimated) дискретного косинусного перетворення (ДКП) зображення-контейнеру $\boldsymbol{X} \in \mathbb{R}^{M\times N}$, що розраховуються згідно наступної формули:

$$\boldsymbol{U}^{(k,l)} = \boldsymbol{X} \star \boldsymbol{B}^{(k,l)}, \boldsymbol{B}^{(k,l)} = \left(B^{(k,l)}_{m,n}\right), 0 \leqslant m,n \leqslant 7,$$

$$B^{(k,l)}_{m,n} = \frac{\boldsymbol{w}_k \boldsymbol{w}_l}{4} cos\left(\frac{\pi k(2m+1)}{16}\right) cos\left(\frac{\pi l(2n+1)}{16}\right), \boldsymbol{w}_0 = \frac{1}{\sqrt{2}}, \boldsymbol{w}_{k|k>0} = 1,$$

де $\star$ – оператор згортки (конволюції); $B^{(k,l)}_{m,n}$ – ядра двовимірного дискретного косинусного перетворення; $\boldsymbol{w}_k$ – нормуючі множники. В якості параметрів моделі DCTR використовуються гістограми розподілу значень коефіцієнтів ДКП у блоках розбиття досліджуваного зображення. Загальна кількість параметрів моделі DCTR є рівною 8000.

Адаптивні методи приховання повідомлень в просторовій області зображення-контейнеру. В роботі були розглянуті сучасні адаптивні стеганографічні методи WOW [6] та MiPOD [7]. Особливістю даних методів є використання функції $F_{distortion}$ для оцінки спотворень зображення-контейнеру внаслідок приховання повідомлень.

Застосування методу WOW при формуванні стеганограм дозволяє мінімізувати спотворення статистичних характеристик ЗК при вбудовуванні повідомлень в частотній області з використанням двовимірного дискретного вейвлет перетворення. Значення функції оцінки спотворень $F^{WOW}_{distortion}$ у методі WOW розраховується для кожного пікселя зображення-контейнеру $\boldsymbol{X}$ розмірами $M \times N$ пікселів згідно наступної формули [6]:

$$F^{WOW}_{ij} = \Sigma^{n}_{k=1} |\xi^{(k)}_{ij}|^{-1},$$

$$\xi^{(k)}_{ij} = \Sigma_{(m,n)\in M\times N} |\boldsymbol{R}^{(k)}_{mn}| \cdot |\boldsymbol{R}^{(k)}_{mn} - \boldsymbol{R}^{(k)}_{[ij]mn}|,$$

де $\boldsymbol{R}^{(n)} = \boldsymbol{X} \star \boldsymbol{K}^{(n)}$ – коефіцієнти дискретного вейвлет перетворення зображення-контейнера, отримані шляхом згортки вихідного зображення з n-тою базисною функцією $\boldsymbol{K} = \{K^{(1)}, ..., K^{(n)}\}$.

Стеганографічний алгоритм MiPOD заснований на представленні зображення-контейнеру $\boldsymbol{X}$ розмірами $M \times N$ як D-вимірного вектору ($D = M \cdot N$) незалежних реалізацій D гаусових випадкових величин $Z_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$. Приховання окремих біт повідомлення проводиться шляхом мінімізації функції F^{MiPOD}:

$$F^{MiPOD} = \Sigma_{n=1}^{D} H\left(\beta_n\right),$$

де $\beta-$ величина зміни значення яскравості поточного пікселя зображення-контейнеру; $H(x) = -2x\log(x) - (1-2x)\log(1-2x)-$ тернарна ентропія. Значення параметру β розраховується шляхом мінімізації коефіцієнту відхилення розподілу значень яскравості пікселів ЗК від нормального (гаусового) розподілу:

$$\varrho^2 = 2\Sigma_{n=1}^{N} \sigma_n^{-4} \beta_n^2,$$

де σ_n^2- локальна оцінка дисперсії значень яскравості пікселів зображення-контейнеру.

Методика тестування стегодетектору. Дослідження точності виявлення стеганограм, сформованих згідно методів WOW та MiPOD, при використанні стегодетектору на основі моделі DCTR проводилося на тестовому пакеті зі 100,000 цифрових зображень зі стандартного пакету MIRFLICKR-1M [8]. Ступінь заповнення зображення-контейнеру стегоданими змінювалася від 5% до 25% з кроком 5% та від 25% до 95% з кроком 10%.

За результатами тестування стегодетектору були побудовані залежності значень стандартних метрик якості стегодетектору, зокрема площі під ROC-кривою (AUC), коефіцієнту кореляції Метьюса та інших, від ступеня заповнення ЗК стегоданими. Виявлено суттєве зниження точності виявлення стеганограм, сформованих згідно методів WOW та MiPOD, при використанні моделі DCTR навіть у випадку сильного заповнення ЗК стегоданими (більше 50%).

Висновки. За результатами проведеного дослідження встановлено, що невідповідність областей приховання повідомлень та проведення стегоаналізу призводить до суттєвого зниження точності виявлення стеганограм, особливо в області слабкого заповнення зображення-контейнеру стегоданими (менше 10%). Підвищення імовірності виявлення стеганограм незалежно від області вбудовування стегоданих потребує застосування комплексних моделей контейнеру, отриманих шляхом об'єднання декількох статистичних моделей ЦЗ.

Література. **1.** Fridrich J. Steganography in Digital Media: Principles, Algorithms, and Applications / Fridrich J. – Cambridge University Press, 2009. – 466 p. **2.** Böhme R. Advanced Statistical Steganalysis / Böhme R. – Information Security and Cryptography Series. – Springer-Verlag Berlin Heidelberg, 2010. – 288 p. **3.** Прогонов Д.О. Виявлення стеганограм з даними, прихованими в області перетворення цифрових зображень. / Прогонов Д.О., Кущ С.М. // Вісник НТУУ "КПІ". Серія – Радіотехніка. Радіоапаратобудування. – № 57, 2014. – С. 128–142. **4.** Прогонов Д.А. Мультифрактальный флуктуационный анализ стеганограмм. / Прогонов Д.А., Кущ С.Н. // Системные исследования и информационные технологии. – № 4, 2015. – С. 39–47. **5.** Holub V. Low-Complexity Features for JPEG Steganalysis Using Undecimated DCT / Holub V., Fridrich J. – IEEE Transactions on Information Forensics and Security. – Vol. 10, Iss. 2. – 2015. – pp. 219–228. **6.** Holub V. Designing steganographic distortion using directional filters / Holub V., Fridrich J. – IEEE International Workshop on Information Forensics and Security (WIFS). – 2012. **7.** Sedighi V. Content-Adaptive Steganography by Minimizing Statistical Detectability / Sedighi V., Cogranne R., Fridrich J. – IEEE Transactions on Information Forensics and Security. – Vol. 11, Iss. 2. – 2016. – pp. 221–234. **8.** Huiskes M.J. The MIR Flickr Retrieval Evaluation / Huiskes M.J., Lew M.S. – ACM International Conference on Multimedia Information Retrieval. – 2008.

Продан А.О.
КПІ ім. Ігоря Сікорського, ФПМ, Київ, Україна

Дослідження законів Ципфа з метою визначення унікальності автора за формулою

Закон Ципфа являє собою емпіричну закономірність розподілення частоти слів природньої мови: якщо всі слова мови впорядкувати за спадаючою частотою їх використання, то частота n-го слова у такому тексті виявиться приблизно обернено пропорційній його порядковому номеру [1].

Застосування концепції мета-книги, яка являє собою абстрактний нескінченний текст, в якому описаний світ очами одного чи іншого автора [2], з використанням законів Ципфа доводить, що закономірність залишається постійною в будь-яких текстах одного і того ж автора. На рис. 1 враховується залежність кількості слів, N, від загального числа слів, M. Дані представляють собою колекцію книг кожного автора.

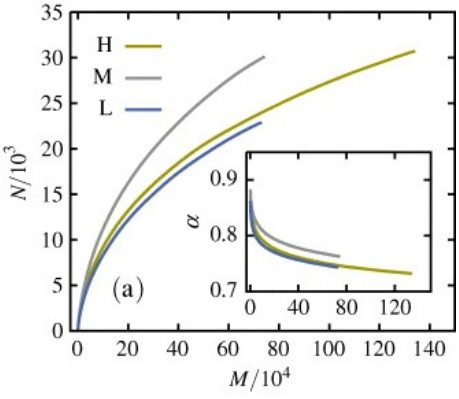


Рис. 1. Показник степені $\alpha = \frac{\ln N}{\ln M}$ виступає в ролі функції M для кожного автора.

Коли довжина тексту збільшується, кількість різноманітних слів також збільшується, проте середнє використання конкретного слова не є постійним, хоча і збільшується так само.

Для кожного розміру тексту, середнє входження слова розраховується як $k = \frac{N}{M}$. Крива $N(M)$ може бути перетворена в криву для середньої частоти в залежності від M.

Написання тексту можна порівняти з процесом, в якому автор витягує фрагмент тексту з великої мета книги [3], яка являє собою абстрактну нескінченну книгу, що дає уяву про частотну характеристику слів, які б автор використав у письмовій формі.

Знаходиться така властивість як частота найбільш поширених слів k_{max}. При діленні книги навпіл відбувається ділення k_{max} в два рази. Частота найбільш поширених слів, k_{max}, в залежності від розміру книги, M, для трьох авторів у подвійному логарифмічному масштабі показана на рис. 2.

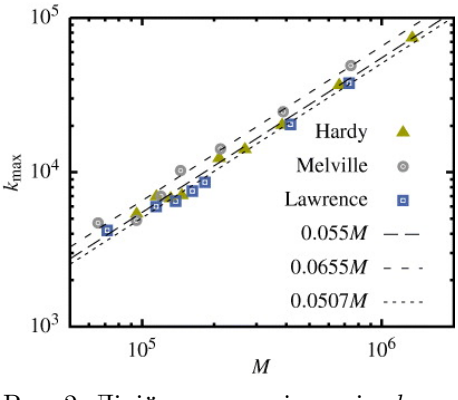


Рис. 2. Лінійна залежність між k_{max} та M.

Емпірично показано, що залежність функцій $N(M)$ та $k(M)$ має споріднену поведінку з розділенням великої книги. Процес випадкового перетворення книги може призвести до зменшення показника степені α. Концепція мета-книги, що являє собою абстрактну нескінченну книгу автора описує таку поведінку. Мета-книга має містити всі статистичні властивості реального тексту, пов'язані з конкретним стилем написання автора, які після цього переносяться в реальну книгу, коли будуть отримуватися з мета-книги. Мета-книга являється цілісною концепцією, завдяки якій будь-який текст неоднакової довжини, написаний автором буде містити інформацію про загальну про об'єм словарного запасу автора.

Література. **1.** Zipf's law, https://en.wikipedia.org/wiki/Zipf%27s_law. **2.** Авторство писателей можно узнать по специальной формуле, http://mignews.com.ua/science/nauka/2531956.html. **3.** The meta book and size-dependent properties of written language, http://iopscience.iop.org/article/10.1088/1367-2630/11/12/123015/meta.

Ревва Р.В.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Подготовка интерактивных отчётов с использованием визуальной аналитики для последующего упрощения процесса принятия решений в бизнесе

На сегодняшний день в бизнесе существует потребность в средствах, позволяющих максимально быстро провести необходимый анализ данных в интерактивном режиме и представить результаты анализа в визуализированном виде непосредственно в процессе принятия решений. Данная работа предназначена для демонстрации полного цикла процесса построения интерактивного бизнес-отчёта с использованием средств визуальной аналитики.

В качестве входных данных были использованы максимально приближённые к реальным демонстрационные данные отдела продаж производителя спортивных товаров, представленных на мировом рынке, собранные в период с 2011 по 2016 года. Главной задачей является построение бизнес-отчёта, позволяющего в доступной и, в то же время, максимально информативной форме, проанализировать продажи каждого вида продукции за всю историю компании, определить динамику доходности и продаваемости каждой группы товаров, а также, определить причины роста или падения продаж товаров каждой категории для дальнейшего принятия рациональных решений в процессе выбора стратегии продаж на следующий период.

В рамках работы будут рассмотрены такие этапы разработки отчёта как: подготовка данных; интеграция данных; построение отчёта; пост-анализ и принятие решений.

Решения визуальной аналитики, такие как SAS Visual Analytics [1], объединяют в себе высокопроизводительные средства анализа и визуализации данных, позволяющие анализировать большие объёмы данных вплоть до BigData и получать интерактивные, наглядные визуализации данных, консолидированные в виде доступных для понимания отчётов.

Подготовка данных играет одну из наиболее важных ролей в процессе анализа и включает в себя такие процедуры как очистка данных, заполнение пропусков, дополнение данных и др.

Цель интеграции данных – обеспечить возможность использования единой процедуры доступа к обработанным данным одновременно из разных приложений.

После того, как процесс интеграции завершён, можно приступить к исследованию данных и построению отчёта. Процесс построения отчёта начинается с подготовки концепции.

Для определения динамики необходимо рассматривать данные в контексте временных рядов с возможностью детализации выбранного периода. Анализ и выявление причин изменения продаж требует постепенного сужения охвата взаимосвязанных рассматриваемых данных с выделением наиболее значимых категорий. Кроме того, так как решения, принимаемые на основании этого отчёта, могут иметь стратегическое значение для компании, необходимо всегда понимать общую картину бизнеса по регионам. Таким образом, для решения задачи, поставленной в рамках данной работы, лучше всего подойдёт UpDown-подход в проектировании взаимодействий между результатами аналитических исследований.

Процесс принятия решений в современном большом бизнесе ставит всё больше и больше вызовов для advanced analytics технологий. Одним из таких вызовов является потребность интерактивности в бизнес-отчётах. Приводится один из вариантов использования интерактивного отчёта в процессе принятия решений по стратегическому планированию продаж и маркетинговых кампаний на следующий год. С помощью средств визуальной аналитики, в рамках работы, была рассмотрена динамика развития продаж компании. Результатом работы стало обнаружение сегмента клиентов рынка США, в возрасте 40–50 лет, которые не проявляют интерес к продуктам категории Tents в наиболее пиковый период с точки зрения увеличения продаж, в отличие от схожих сегментов клиентов в других странах.

Литература. **1.** SAS Institute Inc. 2014. SAS® Visual Analytics 7.1: User's Guide. Cary, NC: SAS Institute Inc., P. 3–7.

Сергеєв Д.С.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Виділення концептів у природно-мовному тексті як спосіб наповнення бази знань

Вступ. Обробка природної мови – це актуальний і перспективний напрям досліджень в галузі інформаційних технологій. Сучасні інформаційні природно-мовні технології часто використовують як свої складові різноманітні бази знань – спеціалізовані корпуси текстів, онтології, анотовані тексти, що дозволяють звести деякі складні задачі обробки мови до пошуку серед попередньо підготованих прикладів [1].

Аналіз проблеми. Наповнення бази знань, як правило, реалізується з використанням одного з двох основних підходів: наповнення в ручному режимі, що є відтворенням навчання дитини вчителем і вимагає великих витрат часу оператора, або наповнення з використанням стохастичних систем – аналогічно до навчання дитини на власних помилках, що вносить неминучі втрати точності та значно ускладнює аналіз та налаштування роботи системи.

Опис алгоритму. У даній роботі розглядається використання інтеграційного підходу [2] до моделювання мовленнєвої діяльності людини в якості основи для створення та наповнення природно-мовної бази знань. В рамках інтеграційного підходу визначається квант знань як структура, що описує у текстовому вигляді ситуацію зорового рівня і складається з об'єкту, руху, їх атрибутів та міри атрибутів. В рамках даної статті визначимо семантику кванту знань як його атрибутивне оточення, що не входить в сам квант знань.

Кожний з квантів знань, ідентифікованих у тексті, обробляється наступним чином. Фрагмент тексту визначається як контекст, початковий обсяг контексту визначається як один абзац. Оператор вводить в систему декілька концептів [3]. Кожний атрибут, що належить до семантики концепту, збільшує вагу цього концепту на 1; усі нові атрибути додаються до концепту з найбільшою вагою. Якщо неможливо визначити панівний концепт у контексті, контекст розширюється на попередній та наступний абзаци; якщо після обробки всього тексту встановлюється паритет двох або більше концептів, новий атрибут додається до усіх з них.

Приклад роботи. Для прикладу розглядаємо концепти (1) "навчальна програма", що має атрибути "міжнародна, дворівнева" та (2) "комп'ютерна програма", що має атрибути "системна, прикладна". Зазначимо, що вони можуть перетинатися ("навчальна комп'ютерна програма"), але розкриття таких випадків є окремою задачею.

В такому випадку атрибут "потужна" з фрагменту тексту "потужна системна програма" додається до семантики концепту (1), а атрибут "наукова" з фрагменту тексту "міжнародна наукова програма" – до концепту (2). Атрибут "спеціалізована" з фрагменту тексту "спеціалізована програма" додається до обох концептів, оскільки фрагмент тексту не містить елементів наявних семантик цих концептів.

Висновки. Запропонований підхід до наповнення бази знань моделює процес обробки природної мови людиною, а отже, потребує попереднього навчання на масиві простих текстів. Отримана в результаті база знань фактично представляє собою модель світу окремо взятої особистості. Це означає, що запропонована модель може бути використана зокрема для подальшого дослідження цього процесу, таких як формування моделі знань окремої особистості або моделювання процесу спілкування.

Література. **1.** Bernstein A. Evaluating the usability of natural language query languages and interfaces to Semantic Web knowledge bases / A. Bernstein, E. Kaufmann. – 2010. – №4. – С. 377–393. **2.** Kyslenko Y. Back to basics of speech activity / Yuri Kyslenko // Biologically Inspired Cognitive Architectures. – 2014. – №8. – С. 46–68. **3.** Laurence S. Concepts and Cognitive Science / S. Laurence, E. Margolis // Concepts: Core Readings / S. Laurence, E. Margolis., 1999. – (MIT Press). – С. 3–81.

Сергєєв А.В.
Київський національний університет ім. Тараса Шевченка, Київ, Україна

Особливості проектування гібридних систем доставки контенту

Розвиток інформаційних технологій у сучасному світі поставив перед комп'ютерними мережами ряд викликів, пов'язаних в першу чергу із різким збільшенням об'ємів контенту, що проходить через них. За таких обставин все більш розповсюдженими стають децентралізовані однорангові (p2p) мережі у випадку некомерційних організацій, а також так звані системи доставки контенту (CDN – Content Delivery Network) у протилежному випадку.

У випадку глобальної роззосередженості мережі (тобто її вузли значно географічно віддалені один від одного) у одноранговій мережі можуть з'явитися значні затримки при знаходженні та доставці інформації, які значно знижують комфортність користування мережею для кінцевого користувача.

У той же час, системи доставки контенту працюють за принципом віддзеркалення контенту із основного до периферійних серверів, що знаходяться максимально близько до клієнта шукача. До цього периферійного серверу і йде запит, і тільки у випадку відсутності потрібної інформації запит доходить до основного серверу. Таким чином й вирішується проблема затримок. Втім, постійна підтримка у актуальному стані як інформації на серверах, так і їх апаратної частини потребує значних фінансових ресурсів, які можуть бути недоступними некомерційним чи невеликим комерційним організаціям.

Можливим рішенням зазначених проблем є створення так званої гібридної системи доставки контенту, яка архітектурно поєднує як однорангові мережі, так і класичні системи доставки контенту.

Наразі існують дві можливі моделі побудови гібридних систем доставки контенту [1]:

1. Система, у якій основну роль відіграє саме класична CDN, при цьому ресурси одноранговой мережі застосовуються тільки за необхідності.
2. Зворотній варіант, при якому основні ресурси надаються одноранговою мережою, а CDN-серевери виконують роль резерву, запит до якого йде лише у випадку, коли необхідна інформація не була знайдена у p2p мережі.

Нажаль, на даний час, майже всі спроби реалізації гібридних систем доставки контенту обмежувалися лише теоретичним описом системи. Існує лише одна практична реалізація CoralCDN [2], яка була успішно введена в експлуатацію і працювала впродовж 2004–2012 рр. і була дуже простою у використанні для звичайних користувачів – для того, щоб інформація з сайту потрапила до цієї мережі, до його URL необхідно було просто додати суфікс .nyud.net. Але, незважаючи на простоту використання, деякі критичні недоліки [3, 4] не дозволили CoralCDN скласти якісну конкуренцію сучасним комп'ютерним мережам інших видів.

Всі інші спроби реалізації гібридних систем доставки контенту [5, 6] на даний час залишаються на рівні теоретичних розробок, в першу чергу через відсутність ефективних алгоритмів пошуку та доставки інформації і потребують подальших наукових досліджень.

Література. **1.** G. Urdaneta, G. Pierre, M. van Steen, A survey of DHT security techniques, ACM Computing Surveys 43 (2) 2011. **2.** M.J. Freedman, E. Freudenthal, and D. Mazieres. Democratizing content publication with Coral. In NSDI, Mar. 2004. **3.** P. Sun, M. Yu, M.J. Freedman, and J. Rexford. Identifying Performance Bottlenecks in CDNs through TCP-Level Monitoring. In SIGCOMM Workshop on Meas. Up the Stack, August 2011. **4.** P. Wendell and M.J. Freedman, "Going viral: flash crowds in an open CDN," in ACM Internet Measurement Conference (IMC), 2011, pp. 549–558. **5.** T.N. Kim, S. Jeon, and Y. Kim, "A CDN-P2P Hybrid Architecture with Content/Location Awareness for live Streaming Service Networks," Proc. IEEE 15th Int. Symposium on Consumer Electronics, pp. 438–441, Aug. 2011. **6.** Ratnasamy, S., Handley, M., Karp, R., Shenker, S., (2002), "Topologically-aware overlay construction and server selection," Proc. of IEEE INFOCOM, pp. 1190–1199.

Сибіряков П.А.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Система автоматизації генерування тестів

В наш час однією з найважливіших проблем галузі IT став високий рівень складності програмних систем, а також пов'язані з нею високі вартісні витрати на адміністрування, модифікацію та подальшу розробку. Зі зростанням складності систем зростав масштаб цієї проблеми. Для покращення надійності програмного продукту використовують тестування. Для заданих вхідних даних з кількома можливими значеннями, виконати для кожного параметра всебічне тестування всіх його можливих комбінацій практично неможливо. Виникає необхідність створення оптимального набору тестів, які будуть ефективно перевіряти ПЗ.

Тестування – це процес керованого експериментування з продуктом за допомогою тестів з метою виявлення в ньому помилок, тобто виявлення неточностей допущених розробниками програмних систем. Тест – контрольна задача для перевірки коректності функціонування програмних систем. Групу взаємозв'язаних тестів називають набором тестів [1].

Основні методи генерації наборів тестів можуть бути класифіковані в 3 основні групи:

- Математичні методи.
- Жадібні алгоритми.
- Мета-евристичні стратегії.

Математичні методи для створення наборів тестів, як правило, вимагають, щоб кожен фактор мав однакову кількість значень, це є суттєвим обмеженням. До цієї категорії належить метод ортогональної матриці [2]. Жадібні алгоритми починають з порожньої множини і додають один тест за крок до множини, поки не виконається умова покриття. Кінцеве значення множини і є розв'язком. Представники жадібних алгоритмів: автоматичний генератор ефективних тестів (Automatic Efficient Test Case Generator), алгоритм впорядкованих параметрів (In Parameter Order) [3]. Мета-евристичний підхід починається з одного чи декількох наборів тестів, у процесі змінювати набір тестів. Представники цього підходу: алгоритм сходження до гори (Hill Climbing), генетичний алгоритм.

Для генерації в роботі запропоновано алгоритм автоматизованого генерування тестів, що базується на використанні методу попарного тестування та алгоритму In Parameter Order. Алгоритм використовується для системи з двома або більше вхідними параметрами, процес розширення множини тестів складається з двох частин: горизонтального та вертикального росту. Нехай S – кандидат на розв'язок, який складається з декількох параметрів. S_{best} – найкращий поточний розв'язок. Алгоритм передбачає такі кроки:

1. Ініціалізуємо S та $S_{best} = S$.
2. Знайдемо локальну залежність між кожним параметром S.
3. Відсортуємо параметри S за їх локальною залежністю, найгірший буде змінений.
4. Обираємо S_q, як підмножина S вже з зміненим найгіршим параметром; $S = S_q$.
5. Якщо глобальна залежність між параметрами S_q краща, ніж S_{best}, то $S_{best} = S_q$.
6. Повторити кроки 2–5, поки не буде досягнута умова зупинки.

Висновок. У результаті даного дослідження розроблена автоматизована система генерації наборів тестів. Запропонована система може бути використана для аналізу існуючих наборів тестів та усунення надлишкових. В залежності від кількості вхідних параметрів та їх значень система зменшує набір тестів, що значно прискорює та зменшує витрати на процес тестування.

Література. **1.** Дідковська М.В. Тестування: Основні визначення, аксіоми та принципи. Текст лекцій. Частина І. / Дідковська М.В., Тимошенко Ю.О. www.mmsa.kpi.ua/disciplines/sac/14c. **2.** Brownlie R. Robust testing of AT&T PMX/StarMAIL using OATS / R. Brownlie, J. Prowse, and M.S. Padke. // AT&T Technical Journal.- May 1992. – vol. 71. – pp. 41–47. **3.** Bryce C. The density algorithm for pairwise interaction testing / Ren´ee C. Bryce and Charles J. Colbourn.- University of Nevada. – 2007.

Скукіс О.Є.
Інститут кібернетики ім. В.М.Глушкова НАН України, Київ, Україна

Застосування інформаційних технологій для розв'язання задач транспортних перевезень

Транспортна логістика – це переміщення необхідної кількості товару в потрібну точку, оптимальним маршрутом за необхідний час і з найменшими витратами [1–4]. Серед основних задач транспортної логістики, як правило, виділяють наступні: вибір виду та типу транспорту; вибір перевізника; визначення раціональних маршрутів доставки; оптимізація процесу транспортування товарів [5].

У роботі [6] запропоновано структуру та склад інформаційної системи (ІС), призначеної для автоматизації розв'язання типових задач із області транспортних перевезень. Одна з підсистем ІС призначена для побудови оптимального маршруту доставки грузів. Цей процес зводиться до розв'язання дискретних оптимізаційних задач які, як правило, мають складну природу – велику розмірність, багатоекстремальність, неточність інформації та ін. Враховуючи ці обставини та результати проведеного аналізу існуючих інформаційних технологій, призначених для розв'язання подібних задач, розроблено програмне забезпечення автоматизації обчислень, яке грунтується на наступних принципах: модульність програмного забезпечення; широкий вибір способів і інструментів обробки даних; використання простої проблемно-орієнтованої мови для формулювання завдань; автоматична організація процесу обробки даних; діалоговий режим роботи; сумісність з іншими програмними засобами.

Функціональні можливості програмного забезпечення реалізовано підсистемою, яка забезпечує формування і супровід паспортів задач. Останні описують моделі задач в термінах предметної області, визначають спосіб обчислень (розпаралелювання або послідовний варіант) та визначають список прикладних програмних модулів та порядок їх застосування. Дозволяє виконувати підготовку даних: введення і редагування вхідної інформації, підключення зовнішніх масивів вхідних даних. Реалізує автоматизацію процесу обробки даних із можливістю вибору оптимального алгоритму розв'язання задачі, послідовності виконання прикладних модулів, які реалізують обчислювальні алгоритми. Здійснення контролю часу виконання процедур обробки даних дозволяє прискорити процес отримання прийнятного розв'язку шляхом аналізу задачі та налаштування параметрів алгоритмів. Виконує збереження та візуалізацію розв'язків, перегляд і редагування вихідних даних, отриманих в результаті обчислень, побудову діаграм та графіків.

Розроблене програмне забезпечення забезпечило автоматизацію процесу побудови оптимального маршруту перевезень. Застосування об'єктно-орієнтованої технології дозволило використати її властивості спадкоємства та поліморфізму, надало можливість абстрагувати опис класів об'єктів, зокрема опис їх властивостей, методів обробки і аналізу подій, розробити структуроване програмне забезпечення як в частині предметного наповнення, так і в частині побудови інтерфейсу користувача [7]. Інструмент реалізації ІС – мова програмування С++.

Література. **1.** Бакаев О.О., Кутах О.П., Пономаренко Л.А. Теоретичні засади логістики: Підручник. У 2 т. – К.: Київ, ун-т економки і технологій транспорту. – Т.1, 2003. – 430 с; Т. 2, 2005. – 522 с. **2.** G. Dantzig and J. Ramser. The Truck Dispatching Problem. Management Science. – 1959. – №1(6). – P. 80–90. **3.** Parragh S., Doerner K., Hartl R. A survey on pickup and delivery problems. Part I: Transportations between customers and depot // Journal fur Betriebswirtschaft. – 2008. – №58. – P. 21–51. **4.** Parragh S., Doerner K., Hartl R. A survey on pickup and delivery problems. Part II: Transportations between customers and depot // Journal fur Betriebswirtschaft. – 2008. – №58. – P. 81–117. **5.** Скукис А.Е. Оптимизационные задачи в транспортной логистике // Теорія оптимальних рішень. 2015. – С. 106–113. **6.** Скукіс О.Є. Інформаційні технології та системи в транспортній логістиці // Теорія оптимальних рішень. 2016, – С. 149–153. **7.** Gruber T.R. A translation approach to portable ontologies. // Knowledge Acquistion. 1993. N 5 (2).

Сухоручкина О.Н., Прогонный Н.В.
Международный научно-учебный центр информационных технологий и систем НАН и МОН Украины, Киев, Украина

Информационные технологии интеллектуального управления мобильным роботом и их использование в учебных целях

Под интеллектуальным управлением мобильным роботом в данной работе понимается способность системы управления (СУ) робота обеспечивать его автономное целенаправленное функционирование в условиях неполноты информации о внешнем мире и неопределенностей в оценке собственных состояний [1]. Интеллектуальная СУ способна реализовывать допустимый для данной системы класс слабо формализованных задач управления, в отличие от адаптивных систем, нацеленных на один тип хорошо формализованной задачи [2].

Современный мобильный робот (МР) представляет собой техническую систему, объединяющую информационную, коммуникационную и исполнительную подсистемы, находящиеся в постоянном взаимодействии с базовым ядром СУ робота и с его внешним миром. Функционирование автономного МР сопровождается рядом информационных процессов, цель которых – сбор, анализ, интерпретация, хранение и использование информации, поступающей от различных аппаратных и программных модулей МР. Данные процессы асинхронные – каждый из них характеризуется своей частотой поступления данных в оперативную память СУ. Большинство из этих процессов должны протекать в реальном масштабе времени.

Фактически каждый из названных информационных процессов требует разработки соответствующей технологии его осуществления. Связующим фактором параллельно протекающих асинхронных процессов СУ МР является общность структур представления в оперативной памяти текущих состояний МР и внешнего мира, а также текущей цели каждого уровня управления, сведенных к единому такту управления на нижнем уровне.

Интеллектуальная СУ МР имеет три уровня иерархии – стратегический, тактический и исполнительный. В нашем случае на стратегическом уровне происходит определение текущей цели действий МР в виде значений параметров конкретного задания [3]. На тактическом уровне формируется способ достижения цели в соответствии с алгоритмом выбранного задания и конкретными его параметрами. На исполнительном уровне определяются параметры непосредственного управления исполнительными органами МР в соответствии с командами, участвующими в алгоритме выбранного задания.

Особую важность для интеллектуального управления МР имеют информационные процессы:

- связанные с восприятием внешнего мира в рамках тех его характеристик, которые доступны соответствующим программно-аппаратным компонентам информационной подсистемы МР;
- отвечающие за классификацию ситуаций и принятие решений об актуальной для текущего состояния МР задаче;
- обеспечивающие навигацию МР в нестационарной внешней среде.

Информационная технология восприятия внешнего мира и оценки состояния МР непосредственно связана с обработкой и интерпретацией сенсорной информации различных модальностей, при этом желательна унификация базовых компонентов такой технологии для применимости одного и того же программного обеспечения (ПО) в соответствующих задачах для различных типов сенсоров.

Аналогичная унификация методов и подходов желательна для информационной технологии классификации ситуаций и принятия решений на всех уровнях иерархии СУ. Автономное целенаправленное движение МР в не полностью детерминированной и изменяющейся во времени среде прежде всего требует корректной оценки системой собственного положения во

внешнем мире и постоянного обновления информации об окружающей робот обстановке в виде модели мира.

Нами реализована технология формирования пространственной модели мира по информации от видеокамеры и локационных сенсоров МР с соответствующими структурами данных, алгоритмами анализа и интерпретации информации в режиме реального времени, а также генерации изображений 3D-моделей мира и МР в нем в графическом интерфейсе пользователя робота [3–5].

Технология, обеспечивающая навигацию МР, в нашем случае включает в себя следующие методы, алгоритмы и соответствующие структуры данных для выполнения:

- первичной локализации и текущей коррекции оценки положения робота в пространстве по искусственным визуальным ориентирам;
- формализации представления множества положений МР, соответствующих некоторому целевому интересу активного задания СУ;
- планирования и динамического перепланирования траектории и параметров движения МР без угрозы столкновения с препятствиями;
- изменения параметров движения МР по реактивному принципу в экстренных ситуациях.

Также реализована информационная технология поддержки человеко-машинного взаимодействия с использованием устной речи на предметном подмножестве естественного языка. Главное окно графического интерфейса пользователя МР содержит реальные изображения от видеокамеры робота и динамически синтезируемые пространственные сцены среды функционирования робота в удобных пользователю ракурсах.

Помимо традиционных для мобильных роботов заданий, таких как «Подойди к объекту», «Исследуй объект», «Возьми объект», в репертуаре нашего МР имеются задания, связанные с его применением в учебных и исследовательских целях. Например, наш МР может удаленно поставлять входную информацию для лабораторных работ, целью которых является демонстрация студентами теоретических знаний и практических навыков в реализации программ анализа и интерпретации информации от видеокамеры робота либо его сенсоров-дальномеров. Кроме того, с помощью веб-технологии студенты могут проверить результат расчета параметров управления движением МР вдоль заданной траектории в виртуальной среде [6]. Ведется разработка дополнительных программно-аппаратных модулей и веб-средств для использования нашего МР в учебных целях.

Литература. **1.** Искусственный интеллект и интеллектуальные системы управления / И.М. Макаров, В.М. Лохин, С.В. Манько, М.П. Романов. – М.: Наука, 2006. – 333 с. **2.** Чечкин А.В. Активирующая подсистема – главная особенность интеллектуальной системы // Интеллектуальные системы. – 2001. – Т. 6, № 1–4. – С. 10–30. **3.** Сухоручкина О.Н. Структуры и информационные процессы интеллектуального управления мобильным роботом / Збірник наук. праць Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України. – Київ, 2012. – Вип. 62. – С. 93–101. **4.** Сухоручкина О.Н. Интеграция компонентов интерфейса пользователя и системы управления персонального мобильного робота: пример реализации // Искусственный интеллект. – 2008. – № 3. – С. 482–489. **5.** Сухоручкина О.Н., Прогонный Н.В., Воронов М.А. Интерпретация и использование измерений сенсора-дальномера в задачах управления // Управляющие машины и системы. – 2017. – № 1. – (в печати). **6.** Интернет ресурс Международного научно-учебного центра информационных технологий и систем НАН Украины. – http://www.mrs.irtc.org.ua/?module=TrainingProjects.

Ткаченко К.П., Шубенкова І.А.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Створення баз даних для медичних закладів у паліативній медицині

В роботі пропонується система керування базами даних для обліку історій хвороб [1] у картках пацієнта, карток лікарів та переліку карток фармакологічних препаратів.

Предметом дослідження є паліативна медицина і ті з її аспектів [2], що повинні враховуватись в системі керування базами даних медичних карток.

Створення системи управління базами даних медичних карток у галузі паліативної медицини надасть можливість створити необхідні передумови для поінформованості та участі пацієнтів в процесі лікування. А саме, знизити вплив втручання людського фактору на процес лікування, забезпечити надання хворому вчасної правдивої інформації стосовно його здоров'я та процесу лікування, можливість дистанційно записатись на прийом до фахівців, отримати перелік препаратів, які мають застосовуватись для конкретного визначеного захворювання та симптомів, при яких їх застосовують. Також система дозволить створювати самі картки пацієнтів та керувати ними. Передбачається і створення карток лікарів задля надання ресурсу більшої інформативності.

У ході дослідження було визначено суть такого явища як паліативна медицина, виділено головні слабкі місця у впровадженні такого напрямку медицини у сучасних реаліях в Україні [3]. Задля усунення визначених слабкостей було вирішено розробити проект для створення наборів баз даних, та системи керування, яка допоможе знівелювати більшість слабких місць при реалізації паліативної допомоги в Україні.

Вплив механізації та технологізації процесів впровадження паліативної медицини на реальну соціально-економічну складову паліативної медицини [4] відображено у результатах та підсумках по реалізації стартап-проекту зі створення системи керування базами даних медичних карток у галузі паліативної медицини.

Теоретична значимість дослідження полягає в розробці ефективної системи керування базами даних, що допоможе створити дієву модель розвитку та впровадження паліативної медицини як невід'ємної складової сучасної медицини у життя пересічних громадян України [5].

Практична значимість отриманих результатів полягає у створенні системи керування базами даних медичних карток у галузі паліативної медицини, спрощенні впровадження паліативної медицини у сучасні реалії України, нівелюванні негативних явищ, з якими вимушені зіткнутися працівники галузі паліативної медицини у ході своєї діяльності [6, 7].

Література. **1.** Вольф, О.О. Деякі міжнародні нормативно-правові акти щодо забезпечення реалізації соціальної політики та соціальної роботи у сфері паліативної та хоспісної допомоги [Текст] / О.О. Вольф // Реабілітація та паліативна медицина. – 2015. – N 2. – С. 49–53. **2.** Вольф, О.О. Порушення права на повагу до гідності, чуйне ставлення у сфері паліативної та хоспісної допомоги в Україні співвідносно до діяльності МСЕК [Текст] / О.О. Вольф // Главный врач. – 2014. – N 1. – С. 65–67; Головна медична сестра. – 2014. – N 1. – С. 23–24. **3.** Вороненко, Ю.В. Паліативна допомога, як інтегральна складова розвитку та реформування системи охорони здоров'я в Україні [Текст] / Ю.В. Вороненко // Главный врач. – 2014. – N 1. – С. 42–43; Головна медична сестра. – 2014. – N 1. – С. 12–13. **4.** Гойда, Н.Г. Соціально-медичні аспекти розвитку паліативної та хоспісної допомоги в Україні в умовах реформування системи охорони здоров'я [Текст] / Н.Г. Гойда, Ю.І. Губський, А.В. Царенко // Реабілітація та паліативна медицина. – 2015. – N 2. – С. 23–30. **5.** Губський, Ю.І. Розвиток паліативної та хоспісної допомоги в Україні: організаційні, юридичні та медичні аспекти [Текст] / Ю.І. Губський, Н.Г. Гойда, А.В. Царенко // Реабілітація та паліативна медицина. – 2015. – N 1. – С. 68–74. **6.** Рекомендації Європейської асоціації паліативної допомоги [Текст] // Главный врач. – 2013. – N 7. – С. 43–61. **7.** Стратегія розвитку паліативної допомоги в Україні на 2015–2025 роки [Текст] // Реабілітація та паліативна медицина. – 2015. – N 2. – С. 125–127.

Федь В.В.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Порівняльний аналіз криптографічних методів

Сучасне суспільство все більше стає інформаційно-обумовленим, успіх будь-якого виду діяльності все сильніше залежить від володіння певними відомостями і від відсутності їх у конкурентів. І чим сильніше проявляється зазначений ефект, тим більше потенційні збитки від зловживань в інформаційній сфері, і тим більше потреба в захисті інформації.

Широке застосування комп'ютерних технологій та постійне збільшення обсягу інформаційних потоків викликає постійне зростання інтересу до криптографії. Останнім часом збільшується роль програмних засобів захисту інформації, які не потребують великих фінансових витрат в порівнянні з апаратними криптосистемами.

Довгий час в Україні використовувався метод криптозахисту «ГОСТ 28147-89», розроблений ще у 1989 році. Незважаючи на те, що він все ще забезпечує практичну стійкість, для нього вже існують теоретичні методи криптоаналізу, що мають складність менше ніж повний перебір [1]. Цей стандарт вже не використовується в Білорусі та Росії. З точки зору продуктивності, «ГОСТ 28147-89» суттєво поступається сучасним аналогам, що призводить до удорожчання коштів на криптографічний захист.

З огляду на позитивний світовий досвід проведення криптографічних конкурсів, Державна служба спеціального зв'язку та захисту інформації України в 2007–2010 рр. успішно провела національний відкритий конкурс, в результаті якого був відзначений алгоритм «Калина», що став, після додаткових досліджень, основою для національного стандарту [2].

Новий національний стандарт підтримує розмір блоку і довжину ключа шифрування 128, 256 і 512 біт (довжина ключа дорівнює розміру блоку або в два рази перевищує його) [2]. Різні варіанти забезпечують гнучкість вибору параметрів для розробників систем криптографічного захисту, що дозволяє отримати як найвищий рівень швидкодії, так і найбільший запас стійкості перетворення.

Тестування продуктивності було орієнтоване на моделювання особливостей роботи засобів криптографічного захисту, які вимагають високої швидкодії перетворень (захист IP-трафіку і ін.). Порівняння проводилося для блочного шифру «Калина» (всі комбінації розміру блоку і довжини ключа), AES-128, AES-256, ГОСТ 28147-89, СТБ 34.101.31-2011 («BelT», національний стандарт Білорусі) і стандарту Росії (ГОСТ Р 34.12-2015). Результати наведені у таблиці 1.

Табл. 1. Продуктивність блочних шифрів

Блочний шифр	Продуктивність (Мбит/с)
Калина-128/128	1874.39
Калина-128/256	1295.55
Калина-256/256	1392.48
Калина-256/512	1088.88
Калина-512/512	1243.49
AES-128	1747.09
AES-256	1257.43
ГОСТ 28147-89	576.10
СТБ 34.101.31-2011	1080.02
ГОСТ Р 34.12-2015	1146.31

Висновки. Порівняння властивостей основних компонентів алгоритму з іншими регіональними та міжнародними шифрами показує перевагу саме національного стандарту України, що забезпечує більший запас стійкості до основних методів криптоаналізу. Продуктивність шифру порівнянна з AES або перевершує його на 64-бітових платформах. При програмній реалізації на сучасному обладнанні швидкодія національного стандарту України істотно вище, ніж у ГОСТ 28147-89 і стандартів Білорусі і Росії.

Література. **1.** Babenko, L., Maro, E. Algebraic Cryptanalysis of GOST Encryption Algorithm // Journal of Computer and Communications, 2014. 2 – P. 10–17. **2.** Roman Oliynykov, Ivan Gorbenko, Oleksandr Kazymyrov. A New Encryption Standard of Ukraine: The Kalyna Block Cipher // 15th Central European Conference on Cryptology, CECC 2015 – P. 18–20.

Хает Ф.И., Алфимцев А.Н., Девятков В.В.
Московский государственный технический университет им. Н.Э.Баумана, Москва, Россия

Гендерные поведенческие модели работы с пользовательскими интерфейсами различных типов устройств

Для качественного информационно-организационного обеспечения систем управления важной задачей является теоретическое исследование вопросов восприятия информации пользователем и оптимизации характеристик каналов передачи информации, а также процессов принятия решений пользователем и оптимизации путей и способов влияния на эти процессы. Указанные вопросы являются ключевыми для построения эффективных интерфейсов информационных систем самых различных классов.

Задача интеллектуальных дружественных интерфейсов состоит в том, чтобы, с одной стороны, обучить и проинструктировать пользователя так, чтобы он максимально использовал возможности информационной системы и соответствующего интерфейса, а с другой – максимально приспособить информационную систему и соответствующий интерфейс к способностям и возможностям пользователя.

Большое количество исследований в области человеко-машинного взаимодействия подтверждают важность учета личностных характеристик пользователей при построении интерфейсов информационных систем [1–3].

Из широкого круга задач персонализации пользовательского интерфейса в данной работе рассматривается задача определения гендерных различий в поведении пользователей для оптимизации параметров интерфейса.

В работе описано исследование влияния личностных характеристик пользователей (пол, возраст, уровень образования) на параметры интерфейсов информационных систем для устройств различного типа. Среди типов устройств рассматривались персональные компьютеры для личного и рабочего пользования, планшеты и смартфоны.

Для сбора и обработки экспериментальных данных было разработано специальное программное обеспечение распознавания образов на основе нечеткой конечно-автоматной модели [4].

Экспериментально подтверждена гипотеза о гендерных поведенческих особенностях восприятия информации и организации пользовательских интерфейсов информационных систем. Выявлено наличие «мужской» и «женской» поведенческих моделей работы с информационными системами различных классов на различных типах устройств.

На основе полученных результатов предложены рекомендации по построению эффективных интерфейсов информационных систем с учетом гендерных поведенческих моделей.

Отчётливо выявленное влияние культурных особенностей на указанные результаты обуславливает необходимость проведения исследований в среде пользователей соответствующей культуры.

Литература. **1.** Кочкин А.А., Калашников С.Н. Когнитивно-параметрическая оценка информационного потока пользовательского интерфейса // Электронный научный журнал. 2015. №3. С. 67–72. **2.** Редькина Н.С. Персонализированное обслуживание интрнет-пользователей библиотек // Научные и технические библиотеки. 2015. №7. С. 17–30. **3.** Фуртрат Ю.О. О влиянии адаптивных пользовательских интерфейсов на надежность и эффективность функционирования автоматизированных систем // Научно-технический вестник информационных технологий, механики и оптики. 2014. № 1. С. 71–76. **4.** Девятков В.В., Алфимцев А.Н. Нечеткая конечно-автоматная модель интеллектуального мультимодального интерфейса // Проблемы управления. 2011. № 2. С. 69–77.

Хіміч А.В.
КПІ ім. Ігоря Сікорського, ФІОТ, Київ, Україна

Моделювання процесу отримання знань із тексту

Усі загальнодоступні знання людства, зберігаються у текстових джерелах. І це зручно, адже можливо у будь-який момент звернутися до необхідної статті, підручника чи монографії і почерпнути звідти необхідну інформацію. Співставлення, порівняння та аналіз розрізнених фактів дозволяє отримати нове знання. Єдиний мінус, що присутній у такому способі роботи зі знаннями це їхня низька концентрованість. Часто для того, щоб отримати необхідну порцію інформації, доводиться прочитати об'ємний підручник замість однієї довідкової статті.

Аналіз проблеми. На шляху до вирішення цієї проблеми пропонувалися різні підходи, починаючи від анотування, реферування об'ємного джерела інформації і закінчуючи автоматичними системами категоризації знань на основі статистики вжитих термінів та даних корпусної лінгвістики. Важливо відмітити, що зв'язність тексту – основа для прискорення пошуку знань, адже у такому випадку семантичні зв'язки будуть сконцентровані навколо найбільш важливих сутностей.

Створення моделі. Продуктивним напрямом до вирішення проблеми отримання знань з тексту являється використання поняття базової семантико-синтаксичної структури (БССС) як фундаментальної мовної одиниці представлення знань [1]. На основі БССС проектується база знань (БЗ), яка у подальшому заповнюється інформацією, виділеною з тексту. Така БЗ моделює пам'ять людини на осмислену інформацію, повторюючи нейронні зв'язки головного мозку. Будучи заповненою, вона на отриманий формальний запит природної мови видасть інформацію, пов'язану з цим запитом, що зустрілася в опрацьованих раніше текстах. При побудові штучної бази знань слід прагнути в першу чергу до повноти видачі інформації за запитом і до її точності, тому кількість релевантних результатів буде вимірюватися великим числом. Перевага такої структури представлення знань перед підходом, що використовується у пошукових системах полягає у тому, що нові знання завжди будуть її доповнювати та розширювати без потреби постійно здійснювати переоцінку релевантності.

Для заповнення БЗ необхідна система аналізу природно-мовного тексту, який буде подаватися на вхід. Таким аналізатором виступає лінгвістичний процесор (ЛП). Його алгоритм є набором логічних правил опрацювання інформації на морфологічному, синтаксичному та семантичному рівнях. Лінгвістичний процесор як складова частина системи обробки та збереження знань, спрямований на аналіз та синтез мовних повідомлень. Практика показала, що поки не існує єдиного остаточно визначеного алгоритму розбору повідомлень для наповнення структури БССС. Найбільш імовірним вбачається поетапне формування бази правил, які будуть вирішувати окремі питання аналізу тексту, за аналогією з тим, як дитина поступово опановує мовний лад. Оскільки будь-яка природна мова містить скінченну кількість словоформ та правил їх вжитку, отримання робочого прототипу ЛП – це лише питання часу.

Головним призначенням штучного ЛП є відтворення та забезпечення процесу мовних активностей людини (як у режимі сприйняття, так і формування мовного повідомлення). З точки зору аналізу повідомлення ЛП повинен проводити декомпозицію мовного матеріалу за окремими складовими БССС, які виокремлюються як основа формування знань.

Структурний рівень організації ЛП відтворює головні розділи мови й охоплює такі питання:

- атрибутивний рівень опису складових предикатної конструкції;
- монопредикатна структура атомарного речення;
- особливості формування полі предикатного повідомлення.

Далі відбувається вихід на рівень аналізу структури цілісного тексту, де обов'язково повинні враховуватися:

- комунікативний рівень організації повідомлення;
- зв'язність текстової інформації.

- декомпозиція окремого повідомлення за складовими;
- аналіз окремих базових структур.

Що можна сказати про таку базу правил? Вона має бути ієрархічною, однозначною, повною та взаємоузгодженою із системою представлення знань на основі БССС (БЗ). Вимога до ієрархічності бази правил ЛП обумовлюється з одного боку необхідністю сталого процесу опрацювання будь-якого природно-мовного тексту, який подаватиметься на вхід цього ЛП, а з іншого – єдиної точки входу роботи системи. Однозначність – кожне правило з бази правил має виводити результат процесу опрацювання, який в контексті ЛП завжди трактується однозначно, тільки за такої умови можна буде стверджувати про те, що усі тексти, результати обробки яких внесено до БЗ, були опрацьовані за єдиним підходом. Повнота бази правил: ЛП має бути здатним опрацювати довільний природно-мовний текст для заданої мови, включаючи усі його синтаксичні, морфологічні особливості та семантичні зв'язки, тобто «білі плями» такого тексту по завершенні процесу аналізу мають бути відсутніми. Взаємоузгодженість з БЗ викликана потребою у постійній взаємодії: уже наявні у БЗ знання з одного боку використовуються ЛП для аналізу, а з іншого – БЗ повинна розширюватися на основі нових отриманих знань.

Тоді як БЗ являє собою структуру представлення знань на основі БССС, ЛП крім ієрархічної бази правил повинен містити ще численні допоміжні засоби та доповнення. Так, першим етапом його роботи є визначення типу речення. Одразу після цього відбувається виділення та виокремлення неінформативних складових. Таких як комунікаційні засоби, ставлення мовця до інформації, оцінки її вірогідності тощо. Ці складові не пов'язані з відтворенням конкретної ситуації, тобто не являються частиною БССС і не вносяться до БЗ, а тому мають бути помічені аналізатором як такі, що не потребують подальшого опрацювання. Наступним етапом є морфологічний аналіз з використанням морфологічних словників. Уся інформація, яку можливо отримати на основі морфологічного аналізу, може бути корисною при заповненні структури БССС.

Автоматизація процесу. Уже на цьому етапі можливі неоднозначності: одна і та ж словоформа може одночасно належати до різних частин мови, або являтися неологізмом, відсутнім у словнику. Зняти неоднозначності такого роду допоможе надбудова над словниками: морфологічний аналізатор, де, з одного боку, будуть використовуватися правила нечіткої логіки вигляду: «Якщо слово, відношення якого не вдалося однозначно ідентифікувати між прикметником та іменником стоїть у реченні біля іменника, то найбільш імовірно, що це – прикметник», з іншого боку, за типовими флективними складовими, слово, відсутнє у словнику, буде віднесено до тієї чи іншої частини мови. Звісно, остаточне подолання таких проблем буде носити імовірнісний характер доти, доки дану систему не проінспектує експертна система більш високого порядку. Після морфологічної обробки наступним етапом виступає виділення ситуаційних та предикативних складових. Тут за допомогою допоміжних словників системи відношень будуть окреслені елементи конструкцій, на основі яких буде зроблено висновок про заповнення структури БССС.

Дійсно, чим більш детально буде заповнено БЗ, тим більш релевантні результати будуть на виході. Запропонована система розрахована на автоматичне функціонування, однак, до факторів, що дозволять покращити процес взаємодії ЛП та БЗ, відносяться якісні морфологічні словники та деталізований алгоритм ЛП.

Висновки. Створення такої системи є важливим етапом на шляху до вирішення проблеми отримання та струкрутизації знань на основі природно-мвоних текстів. Разом з тим, складність створення високоточної системи пов'язана з деякими особливостями мови, такими як омонімія, коли однозначна інтерпретація тексту неможлива без визначення семантичної близькості слів.

Література. **1.** Kyslenko Y.I. Back to basics of speech activity / Yuriy Kyslenko // Biologically Inspired Cognitive Architectures. – 2014. – №8. – С. 47–69.

Хурцилава К.В., Литвинов В.А., Майстренко С.Я.
Інститут проблем математичних машин і систем НАН України, Київ, Україна

Про критерії вибору моделі представлення просторово-часових геоданих в ГІС лісового господарства

Останнім часом відбувається активний розвиток просторово-часових геоінформаційних систем, що оперують просторовими геоданими і їх змінами в часі. Концептуальною основою подання та обробки таких даних є просторово-часові моделі (ПВ-моделі). Залежно від типів властивостей об'єктів, що змінюються в часі (морфологічних, топологічних, атрибутивних) можливі різні сценарії поведінки ПВ-даних і різні критерії ефективності моделей опису цих сценаріїв. Аналітичний огляд основних існуючих ПВ-моделей [1], та [2], [4], проведений з метою виділення їх властивостей, істотних з точки зору особливостей можливих програмних додатків, показав, що жодна з розглянутих моделей не може розглядатися як універсальна, прийнятно ефективна для всіх поєднань типів змін і критеріїв. Однак, для конкретних предметних областей можливе виділення з безлічі моделей деякої предметно- і проблемно- орієнтованої підмножини, досить ефективної для подання відповідних даних і вирішення задач.

Однією з актуальних областей застосування технологій ГІС є управління лісовим господарством: відстеження кордонів і властивостей лісових об'єктів проводяться з метою вивчення динаміки природних процесів, а також виявлення синтаксичних і семантичних помилок в лісових геоданих [3], [5]. Процес ведення лісовпорядкування включає формування початкової тематичної і картографічної інформації для лісництв, збереження поточної інформації в локальних базах даних і формування вихідних паперових документів, накопичення картографічних і таксаційних даних на серверах баз даних по часовим, територіальним та іншим характеристикам, генерація різного роду вихідних документів (планшетів, планів, карт-схем, і ін.). В основі ієрархічної організації лісового фонду лежать таксаційні «виділи» – однорідні за таксаційними характеристиками та господарським функціональним призначенням ділянки лісового фонду, на всій площі яких при необхідності намічаються однакові господарські заходи. Крім зазначеної особливості організації лісових геоданих істотне значення для вибору альтернативних конкурентоспроможних моделей має обсяг даних, частота внесення змін, частота і складність виконання запитів, наявність (або відсутність) рухомих об'єктів та ін. З урахуванням зазначених критеріїв в якості згаданої проблемно-орієнтованої підмножини виділені три ПВ-моделі:

- модель часових шарів The Snapshot Model (SS-модель);
- модель найменших загальних геометрій Least Common Geometry model (LCG-модель);
- псевдо-об'єктна модель Psevdo-object model (PO-модель).

Для проведення порівняльного аналізу в якості основи «бенчмарка» побудовано фрагмент предметної області, що змінюється (див. Рис. 1), та розглянуто його представлення трьома виділеними моделями.

Оцінка кількісних показників, що характеризують властивості представлень, з урахуванням [4]) показала, що LCG-модель перевершує інші моделі за обсягом необхідної пам'яті і складності виконання запитів, пов'язаних з аналізом змін деякої області, що відслідковується за заданий інтервал часу. Крім того, LCG-модель простіше реалізується, ніж PO-модель (хоча і складніше, ніж SS-модель).

У свою чергу, PO-модель має переваги при великій кількості змін і складній неоднорідній ієрархії об'єктів (або її відсутності). Однак, ця особливість якраз не властива для чіткої ієрархічної організації лісових геоданих. Що стосується вимоги (з боку LCG-моделі) високої просторової і часової точності, то, можна вважати, що лісові геодані задовольняють їй прийнятною мірою.

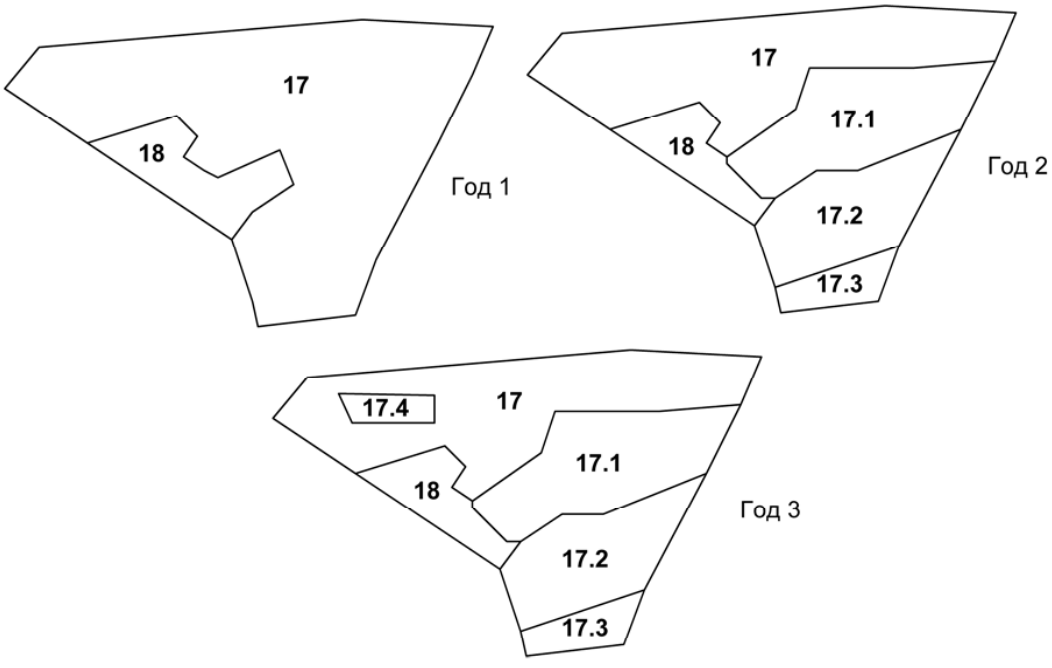


Рис. 1. Тестовий приклад. Зміна частини лісництва за 3 роки

В даний час в ГІС «ГІС-Ліспроект» [3], [5], для якої оновлення геоданих виконується з річним інтервалом, прийнята найбільш проста – SS-модель. Однак, у міру ускладнення задач аналізу динаміки лісового фонду, що визначається тенденцією переходу від опису і експлуатації до управління природними ресурсами, недоліки SS-моделі (зростання обсягу даних зі збільшенням кількості шарів, складність формування траєкторії зміни об'єктів та ін.) визначають необхідність застосування більш ефективних ПВ-моделей. Одним з найбільш прийнятних рішень для розвитку задач управління лісовим фондом вважається вибір LCG-моделі.

Література. **1.** Geographical Information Systems: Principles, Techniques, Applications and Management [Электронный ресурс] / P.A.Longley, M.F. Goodchild, D.J. Maguire, D.J. Rhind. // 2nd Edition Published by Wiley. – 1999. – Режим доступа: www.geos.ed.ac.uk/~gisteac/gis_book_abridged/. **2.** Майстренко С.Я. Пространственно-временные геоданные, их представление и обработка в задачах планово-картографического сопровождения лесоустройства // Математичні машини і системи. – 2015. – №2. С. 85–96. **3.** Литвинов В.А., Майстренко С.Я., Хурцилава К.В. Типовые ошибки в информационных ресурсах ГИС, правила и особенности их выявления в сфере лесного хозяйства (на примере системы «ГИС-Леспроект») // Математичні машини і системи. – 2016. – №2. – С. 78–90. **4.** Ott T. Time-Integrative Geographic Information Systems: Management and Analysis of Spatio- Temporal Data / T. Ott, F. Swiaczny. – Berlin; Heidelberg; New York; Barcelona; Hong Kong: Springer, 2001. – 234 с. **5.** Білецький Б.О., Литвинов В.А., Майстренко С.Я., Беспалов В.П., Загреба Т.О., Хурцилава К.В. Практичне застосування ГІС-технологій для планово-картографічного супроводження лісовпорядкування (на прикладі системи «ГІС-Ліспроект») // Математичні машини і системи. – 2013. – №3. – С. 76–86.

Шубенкова И.А., Клименко Ф.Т.
КПИ им. Игоря Сикорского, ИПСА, Киев, Украина

Проект управления созданием программного кода с использованием систем контроля версий

Задача любой системы управления версиями – записывать историю изменения файла или набора файлов, чтобы в будущем была возможность вернуться к конкретной версии. Данные системы позволяют возвращать в предыдущее состояние как отдельные файлы, так и целый проект, сравнивать сделанные изменения, смотреть, кто и когда последним редактировал файл, являющийся источником проблемы, и многое другое [1]. Хорошо реализованный контроль версий необходим для любого успешного IT-проекта.

Зачастую проектные менеджеры в сфере информационных технологий малоосведомлены о системах контроля версий. Как результат, управление кодовой базой остается неявным и нерегламентированным процессом, который целиком ложится на плечи разработчиков.

На раннем этапе выполнения IT-проекта процесс разработки зачастую выполняется мелкими итерациями, т.е. большое количество взаимосвязанных работ выполняются параллельно с непрерывным анализом полученных результатов и архитектуры приложения с последующей корректировкой предыдущих этапов работы. На данном этапе вопрос синхронизации изменений внесенных в общий программный код стоит наиболее остро.

Также, при выборе методологии управления созданием программного кода, необходимо учитывать, что в подавляющем числе современных IT-проектов процесс разработки выполняется согласно принципам Agile Manifesto, среди которых: частая поставка рабочего программного обеспечения, простота (не делать лишней работы) и постоянная адаптация к изменяющимся обстоятельствам и требованиям [2].

Рассмотрим методологию управления создания программного кода на примере наиболее популярной системы контроля версий Git [3].

Центральный репозиторий содержит две главные ветки, существующие всё время: *master* и *develop*. Мы считаем ветку *origin/master* главной. То есть, исходный код в ней должен находиться в состоянии *production-ready* в любой произвольный момент времени. Ветвь *origin/develop* мы считаем главной ветвью для разработки. Хранящийся в ней код в любой момент времени должен содержать самые последние изданные изменения, необходимые для следующего релиза. Когда исходный код в ветви разработки (*develop*) достигает стабильного состояния и готов к релизу, все изменения должны быть определённым способом влиты в главную ветвь (*master*) и помечены тегом с номером релиза. Каждый раз, когда изменения вливаются в главную ветвь (*master*), мы по определению получаем новый релиз.

Помимо главных ветвей *master* и *develop*, наша модель разработки содержит несколько типов вспомогательных ветвей, которые используются для распараллеливания разработки между членами команды, для упрощения внедрения нового функционала (*feature*), для подготовки релизов (*release*) и для быстрого исправления проблем в производственной версии приложения (*hotfix*). В отличие от главный ветвей, эти ветви всегда имеют ограниченный срок жизни.

Правила порождения и слияния вспомогательных ветвей:

1. ветки типа *feature* порождаются от ветки *develop* и в нее же вливаются;
2. ветки типа *release* порождаются от *develop* и вливаются в *master* и *develop*;
3. ветки типа *hotfix* порождаются от *master* и вливаются в *master* и *develop*.

Данная методология управления создания программного кода формирует простую мыслительную модель, которая позволяет проектной команде легко контролировать версионность и формирует совместное понимание процессов ветвления и слияния.

Литература. **1.** Чакон С., Штрауб Б. Git для профессионального программиста. – СПб.: Питер, 2016. – 496 с. **2.** A. Stellman, J. Greene. Learning Agile. – Sebastopol: O'Reilly Media, 2014. **3.** J. Loeliger, M. McCullough. Version Control with Git. – Sebastopol: O'Reilly Media, 2012.

Щука Р.В., Іванов С.С., Терентьєв О.М., Орешко В.А., Яворський Я.Б., Кузенко А.Ю.
КПІ ім. Ігоря Сікорського, ІПСА, Київ, Україна

Побудова предиктивної моделі прогнозування злочинності на основі методу подібних траєкторій в системі SAS

Вступ. Однією з наріжних задач розвитку соціальної сфери України, що тривалий термін перебуває під тиском несприятливих політичних та соціальних явищ, спричинених антитерористичною операцією в окремих районах Донецької та Луганської областей, залишається підвищення соціальної безпеки населення. У зв'язку із цим, дослідження феномену зростання злочинності сьогодні не тільки набувають великого значення у правовому житті суспільства, а й виступають одним із вирішальних чинників у справі успішного здійснення реформ на шляху розбудови якісно нової української держави і забезпечення безпеки її громадян. Успішне вирішення останніх уявляється можливим завдяки новітнім інформаційним технологіям та створенню на їх базі інформаційно-аналітичних систем.

Результати вивчення публікацій в галузі методологічних підходів до моделювання злочинності та злочинної поведінки у кримінологічній науці [1–6] дозволяють дійти висновку, що питанням створення відносно дешевих інформаційних систем для аналізу і прогнозування злочинності приділяється недостатньо уваги.

Запропонований підхід. У даній роботі в якості одного зі шляхів вирішення проблеми запропоновано використання розробленого К. Мортоном методу подібних траєкторій. Аналіз технологічного програмного забезпечення й додатків класу Business Intelligence показав, що згаданий метод не має програмної реалізації на платформі SAS. Це надало можливість висунути гіпотезу про те, що створення програмного продукту в SAS Enterprice Guide із використанням мови програмування SAS Base зможе посприяти підвищенню точності прогнозування наслідків соціально-економічних процесів, зокрема – рівня злочинності.

Метод подібних траєкторій. Нехай у ході соціального експерименту сформовано апріорний ряд спостережень $Y(t) = y(1), y(2), \ldots, y(i), \ldots, y(n)$, де $y(i), \ldots, y(n)$ – значення спостережуваної величини, що їх зафіксовано у ході експерименту з певним часовим кроком $\Delta t_i = const$, зображеним на рисунку 1:

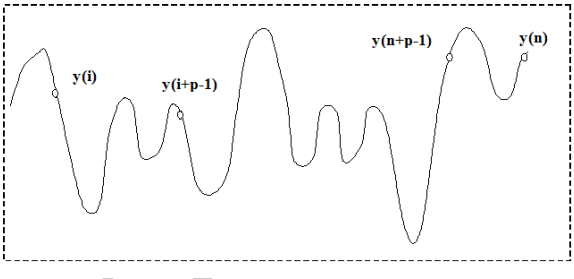


Рис. 1. Приклад спостережень

Обчислимо "перші різниці" $\Delta y(i) = y(i+1) - y(i)$, між сусідніми точками спостережень (різниці між сусідніми значеннями апріорного ряду). Для цього виберемо ділянку траєкторії, що передує прогнозованій точці на крок розміром $(p-1)$.

Сутність подальшого алгоритму полягатиме у наступному. Спочатку визначимо певну кількість символів алфавіту $B = b_k$, де k – довжина алфавіту.

Кількість символів в алфавіті є визначальною під час розбиття відсортованого ряду на часткові інтервали: кожному значенню ряду у відповідність ставиться власний символ. Таким чином числовий ряд спостережень $\Delta y(i)$ набуває символьних значень $b_1 =$ "A", $b_2 =$ "B", $b_3 =$ "C" і т.д.

Дане дослідження було здійснено в рамках проекту NUKR. SFPP G4877 "Modelling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism"

На наступному кроці визначаємо опорний ланцюжок $< e_i >$ – останні значення символьного ряду, які передують прогнозованій точці $y(n)$; весь ряд розбивається на часткові вектори спостережень $< e_i; e_{i+1} >$; вектор i є опорним. Далі підраховуємо кількість ланцюжків, що співпадають з вектором i та дивимось, яких значень набуває ряд на наступних кроках.

Частота появи того чи іншого ланцюжка відповідає ймовірності повторення відповідних їм послідовностей. Іншими словами, обчислюється умовна ймовірність $P(e_{i+1}|e_{i-n} \dots e_{i-1})$ появи символу e_{i+1} за умови, що йому передували символи: $e_{i-n}, e_{i-n+1}, \dots, e_{i-1}$.

За допомогою обчислених ймовірностей можемо зробити ймовірнісний прогноз виникнення якогось символу за умови що відомі ланцюжки попередніх символів. Після отримання прогнозу, робимо зворотній перехід від символьного ряду $B = b_k$ до числового $\Delta y(i)$.

Результати. Алгоритм, що описує метод подібних траєкторій, був реалізований в середовищі SAS Enterprise Guide. В результаті роботи програми були побудовані прогнозні моделі для тестових даних зі стандартної бібліотеки SAS AIR, що описує історію продажів авіабілетів на основі якої можна порівняти результати роботи нових методів з уже існуючими, та даних МВС України по кількості зареєстрованих правопорушень у місті Дніпро. Результати прогнозування наведені у таблиці 1:

Табл. 1. Статистики прогнозних моделей

	RMSE	MAPE	DW	R^2
Тестовий графік AIR $ql = 3, ww = 3$	9881	16,5016	0,8085	0,3459
Тестовий графік AIR $ql = 3, ww = 7$	3009	9,7480	0,9967	0,6732
Графік кількості правопорушень у м. Дніпро $ql = 5, ww = 3$	1935	114,1657	1,0323	0,7267
Графік кількості правопорушень у м. Дніпро $ql = 7, ww = 3$	960,6833	82,1263	0,9600	0,8642

де ww – це ширина опорного вектора, а ql – довжина алфавіту.

Висновки. З аналізу результатів таблиці 1 можна зробити наступні висновки:

- при фіксованій довжині опорного вектора значення RMSE та MAPE спадають за зростання алфавіту, тобто точність прогнозу покращується;
- при фіксованій довжині опорного вектора значення критерію R^2 наближається до одиниці при збільшенні кількості букв в алфавіті, тобто зростає точність прогнозу;
- за результатами роботи методу на даних числових рядах не уявляється можливим зробити остаточний висновок стосовно автокореляції за статистикою DW.

Результати дослідження у вигляді програмної розробки будуть використані при створені Ситуаційно-аналітичного центру при ННК ІПСА, для вирішення задач сценарного аналізу та моделювання.

Література. **1.** Бова А.А. Перспективи використання інтелектуального аналізу даних (DataMining) у кримінологічних дослідженнях [Текст] / А.А. Бова // Науковий вісник Національної академії внутрішніх справ України. – 2002. – № 5. – С. 32–35. **2.** Жеребкин В.Е. Логический анализ понятий права: монография [Текст]: / В.Е. Жеребкин. – К.: Издательское объединение «Вища школа», 1976. – 151 с. 8. **3.** Закалюк А.П. Курс сучасної української кримінології: теорія і практика: У 3-х кн. [Текст]: / А.П. Закалюк. – К.: Видавничий Дім «Ін Юре», 2008. – Кн. 3: Практична кримінологія. – 320 с. **4.** Кравчук С.Й. Економічна злочинність в Україні [Текст] / С.Й. Кравчук. – К.: Кондор, 2009. – 282 с. **5.** Кримінологія: Загальна та Особлива частини: підручник / І.М. Даньшин, В.В. Голіна, М.Ю. Валуйська та ін.; за заг. ред. В.В. Голіни. – 2-ге вид. перероб. і доп. [Текст]: – Х.: Право, 2009. – 288 с. **6.** Чорноус Ю.М. Теорія і практика криміналістичного забезпечення досудового слідства у справах про злочини міжнародного характеру [Текст]: монографія / Чорноус Ю.М.; Нац. акад. внутр. справ. – К.: Скіф, 2012. – 447 с.

Authors · Авторы · Автори

Albu Veaceslav, 20
Artuhov Vitalii G., 198
Baranovska Lesia Valeriivna, 18
Basyuk Taras M., 226
Bidyuk Peter I., 24, 233
Boyarinova Yuliya E., 231, 232
Brytov Oleksii A., 198
Chereda Hryhorii O., 228
Chugay Andrey M., 19
Cojocaru Svetlana, 20
Didenko Dmytro G., 230
Galuza Aleksei A., 12
Garev Yuriy V., 134
Gorban Nataliia V., 22
Gorodko Nataliya A., 231
Grinberg Galina L., 12
Hitsko Yana V., 232
Iievlieva Svitlana M., 134
Kachko Nikita A., 199
Kalinovsky Yakov A., 232
Kaniovska Irina Y., 23
Khomenko Olha V., 22
Kondratova Ljudmyla P., 26, 202
Kryvolap Dmytro V., 233
Kucher Vladyslav O., 201
Kuksa Mykola U., 136
Lazarenko Denys V., 137
Lyubchyk Leonid M., 12
Lyunda Volodymyr V., 137
Marchenko Oleksandr I., 140
Naderan Maryam, 139
Nadtochii Oleksandra V., 24
Naumenko Tetiana O., 234
Nikolaiev Sergii S., 228
Norkin Vladimir I., 136
Ovi Nafas Aghaei Agh Ghamish, 141
Paliichuk Liliia S., 22
Pankratova Nataliya D., 26
Pavlenko Vitaliy D., 28
Pechurin Nickolay K., 202
Pechurin Sergey N., 202
Petic Mircea, 20
Pikhota Sergey R., 134
Romanenko Yan V., 32
Salata Dmytro V., 28
Samuel Adegoke, 236
Solodrai Ihor I., 140
Soloviev Vladimir N., 32
Sukalo Alina S., 232
Titchiev Inga I., 34
Trofimov Alexander S., 240
Yaremenko Vadym S., 204
Zaitsev Andrii A., 23
Zaychenko Elena Yurievna, 141
Zaychenko Yuri Petrovich, 139, 141
Zenkovich Mikhail V., 240
Zorin Yuri M., 142
Алфимцев Александр Николаевич, 329
Антонюк Софія Андріївна, 242
Аралис Светлана Викторовна, 281
Архіпов Дмитро Володимирович, 144
Арчвадзе Натела Нодариевна, 244
Бакун Сабіна Антонівна, 245
Бакурова Анна Владимировна, 247
Безносик Александр Юрьевич, 251
Безсмертна Аліна Володимирівна, 36
Бидниченко Алексей Александрович, 251
Бігун Сергій Олександрович, 37
Бідюк Петро Іванович, 40, 76, 129, 146, 187, 195, 253, 255
Біла Анастасія Ігорівна, 255
Білова Марія Олексіївна, 83
Білущак Юрій Ігорович, 38
Бодянский Евгений Владимирович, 173
Бойко Ірина Олександрівна, 40
Бойко Роман Сергійович, 257
Болдак Андрій Олександрович, 41
Болдак Людмила Олексіївна, 42

Болжеларський Ярослав Володимирович, 43
Бондаренко Наталія Сергіївна, 45
Боровская Таиса Николаевна, 46, 116
Бояринова Юлия Евгеньевна, 280
Буткалюк Сергій Вікторович, 147
Буценко Юрій Павлович, 48
Васильев Владимир Иванович, 49
Васильєв Володимир Іванович, 149
Вельбицкий Игорь Вячеславович, 259
Виклюк Ярослав Ігорович, 261
Висоцька Вікторія Анатоліївна, 78
Вишталь Дмитрий Михайлович, 49
Віштальь Дмитро Михайлович, 149
Войтех Дмитро Володимирович, 51
Войтко Сергій Васильович, 127
Волкова Виолетта Николаевна, 53
Вольский Александр Владимирович, 305
Ворона Михаил Владиславович, 262
Гайвась Богдана Іванівна, 55
Галатенко Дмитро Володимирович, 151
Гамидов Галиб И., 160
Гера Богдан Васильович, 43
Герасименко Оксана Юріївна, 306
Гіоргізова-Гай Вікторія Шалвівна, 205
Главчева Юлія Миколаївна, 78
Гнатенко Валерий Юрьевич, 264
Гожий Віктор Вікторович, 152
Гончарук Володимир Євтихієвич, 38
Горбань Гліб Валентинович, 154
Горбань Наталія Володимирівна, 57
Горелова Галина Викторовна, 58
Городецкий Виктор Георгиевич, 61
Горох Олександр Сергійович, 42
Гритчук Денис Тарасович, 176
Гришин Дмитрий Игоревич, 116
Губар Богдан Дмитрович, 266
Губарєв В'ячеслав Федорович, 267
Гула Ігор Сергійович, 161
Гусак Олена Михайлівна, 261
Давидюк Олександр Юрійович, 156
Данилов Валерій Якович, 118
Девятков Владимир Валентинович, 329
Демьянчук Александр Зиновьевич, 99
Джамалов Алисейн Талыбович, 268
Джамалова Афаг Гараш, 281
Джигирей Ірина Миколаївна, 93
Дичко Аліна Олегівна, 158
Дідковський Олександр Ігорович, 62
Дмитрук Анатолій Андрійович, 55
Дмитрук Вероніка Анатоліївна, 43, 55
Донченко Володимир Степанович, 64–66
Дорофеев Юрий Иванович, 67
Ефремов Артем Александрович, 53
Ємельянова Дар'я Ігорівна, 68
Єремєєв Ігор Семенович, 70, 158
Єршов Сергій Володимирович, 207
Єщенко Олександр Іванович, 70
Жданова Олена Григорівна, 125
Журавель Сергей Вадимович, 109
Забелин Станислав Игоревич, 159
Заболотня Тетяна Миколаївна, 274
Заводник Вячеслав Владленович, 72
Зайченко Юрий Петрович, 160
Зайченко Юрій Петрович, 161, 191
Заковоротний Олександр Юрійович, 162
Залевська Ольга Валеріївна, 273
Замекула Олексій Ігорович, 274
Захарчук Наталія Геннадіївна, 74
Згуровський Михайло Захарович, 13, 275
Землянський Дмитро Ігорович, 277
Зінько Тарас Петрович, 65
Зражевська Наталія Григорівна, 163
Зубок Валерій Віталійович, 278
Іваненко Дмитро Олександрович, 37
Іванов Сергій Сергійович, 335
Калинин Владислав Валерьевич, 280
Кангун Ганна Іллівна, 76
Каніовська Ірина Юріївна, 80, 92
Каніщева Ольга Валеріївна, 78
Капшук Олег Олексійович, 209
Касумова Тахмина Агаджафар, 281
Киричек Галина Григорівна, 283
Кирюша Богдан Анатолійович, 211
Кир'яченко Світлана Петрівна, 185
Киселев Михаил Яковлевич, 109
Кіріленко Олена Георгієвна, 285
Клименко Федор Тарасович, 334
Кобера Валентина Василівна, 80
Коваль Лада Вікторівна, 120
Козерацький Михайло Сергійович, 306
Козуля Марія Михайлівна, 81
Козуля Тетяна Володимирівна, 83
Колесник Ирина Сергеевна, 116
Колінько Анжела Михайлівна, 211
Коляденко Андрій Владиславович, 164
Комісаренко Вячеслав Ігорович, 267

Кондратенко Наталія Романівна, 166
Копп Андрей Михайлович, 287
Корбутяк Василь Михайлович, 128
Корнага Ярослав Ігорович, 306
Корнієнко Олександр Олегович, 289
Коршевнюк Лев Олександрович, 85
Котолупов Олег Владиславович, 273
Кравець Ірина Олександрівна, 168
Крамаренко Юлія Юріївна, 146
Кузенко Андрій Юрійович, 335
Кулик Анатолий Степанович, 14
Кулиш Роман Дмитриевич, 61
Кульбачний Костянтин Андрійович, 205
Курач Ярослав Сергеевич, 88
Куца Карина Владимировна, 169
Куцаченко Никита Геннадиевич, 290
Куценко Александр Сергеевич, 87
Лабжинський Володимир Анатолійович, 48
Лавренюк Алла Миколаївна, 291
Лавренюк Сергій Іванович, 291
Лагодний Олександр Васильович, 103
Ладогубець Тетяна Сергіївна, 74
Лазаренко Ирина Сергеевна, 262
Лазаренко Ірина Сергіївна, 41
Лисенко Андрій Володимирович, 293
Лисецький Тарас Миколайович, 312
Литвинов Валерій Андронікович, 332
Литвинюк Антон Андрійович, 245
Лілікович Сергій Олександрович, 295
Ліщук Катерина Ігорівна, 222, 277, 308
Логинова Александра Викторовна, 53
Лопатин Алексей Константинович, 90
Лопатин Алексей Константинович, 88
Лубинець Юрій Миколайович, 92
Любашенко Наталія Дмитрівна, 149
Любашенко Наталья Дмитриевна, 49
Любчик Леонид Михайлович, 67
Лютенко Ірина Вікторівна, 295
Ляхевич Олександр Володимирович, 297
Мажара Ольга Олександрівна, 192
Майстренко Світлана Яківна, 332
Максим Катерина Євгенівна, 298
Малишев Андрій Ігорович, 299
Марченко Олександр Іванович, 213, 314
Марченко Олексій Олександрович, 213
Мегрелишвили Ричард Петрович, 301
Мельников Олег Валентинович, 312
Мерзляков Алексей Дмитриевич, 99
Минько Олексій Володимирович, 93
Миронюк Игорь Анатольевич, 303
Мілявський Юрій Леондіович, 111
Місюра Олена Борисівна, 312
Моїсєєв Віктов Федорович, 81
Мокін Віталій Борисович, 111
Морозов Константин Вячеславович, 94
Москаленко Валентина Владимировна, 303
Москаленко Юрій Володимирович, 192
Мурга Микола Олексійович, 170
Муха Ирина Павловна, 305
Мухін Вадим Євгенійович, 306
Мысник Богдан Викторович, 171
Назарага Інна Михайлівна, 66
Негода Катерина Володимирівна, 307
Недашківська Надія Іванівна, 95
Нестюк Олег Миколайович, 308
Никульченко Артем Александрович, 67
Ніколаєв Сергій Сергійович, 190
Новицький Євген Анатолійович, 96
Ноголь Ольга Святославовна, 72
Овчар Анастасія Олександрівна, 283
Олійник Юрій Олександрович, 266
Олокоба Олаийа Йусуф, 112
Опарина Елена Леонидовна, 97
Орешко Валерія Анатоліївна, 335
Орєхов Олександр Арсенійович, 309
Орєхова Наталія Артурівна, 309
Орлова Марія Миколаївна, 215
Орловский Дмитрий Леонидович, 287
Осиюк Дмитрий Васильевич, 310
Павлов Олександр Анатолійович, 275, 312
Палійчук Лілія Сергіївна, 57
Панкратов Владимир Андреевич, 98
Панкратова Наталья Дмитриевна, 99
Пархоменко Андрій Володимирович, 101
Пашаева Малахат Мухтар, 281
Пашинська Наталія Миколаївна, 102
Перова Ирина Геннадьевна, 173
Петренко Анатолій Іванович, 216
Писарчук Олексій Олександрович, 103
Пичугина Оксана Сергеевна, 105
Пишнограєв Іван Олександрович, 107
Піддубна Юлія Ігорівна, 107
Поворознюк Анатолий Иванович, 174
Поворознюк Назар Іванович, 176
Погорелов Володимир Володимирович, 314
Пономаренко Роман Миколайович, 207

Попеляев Денис Павлович, 315
Притоманова Ольга Михайловна, 109
Прогонный Николай Витальевич, 325
Прогонов Дмитро Олександрович, 317
Продан Анастасія Олегівна, 319
Просянкіна-Жарова Тетяна Іванівна, 253
Путренко Віктор Валентинович, 101, 102
Пховелишвили Мераб Гайозович, 244
Рагимов Рауф Махмудович, 268
Рак Олександр Олександрович, 110
Ревва Роман Владимирович, 320
Романенко Віктор Демидович, 111
Романкевич Виталий Алексеевич, 94
Романов Валерий Владимирович, 218
Романов Валерій Володимирович, 220
Ропало Анна Михайловна, 247
Рудакова Анна Владимировна, 112
Савастьянов Володимир Володимирович, 253
Савельев Юрий Дмитриевич, 177
Савченко Анатолій Вікторович, 113
Савченко Ілля Олександрович, 115
Сапсай Татьяна Григорьевна, 94
Сащук Ігор Миколайович, 180
Северилов Виктор Андреевич, 46
Северилов Павел Викторович, 46, 116
Селін Олександр Миколайович, 194
Селін Юрій Миколайович, 194
Сербул Роман Сергійович, 118
Сергеєв Данило Сергійович, 321
Сергеєв Єгор Ігорович, 220
Сергєєв Антон Валерійович, 322
Сибіряков Павло Анатолійович, 323
Ситников Валерий Степанович, 264
Скляр Андрей Владимирович, 182
Скукіс Олексій Євгенійович, 324
Смирнов Сергій Анатолійович, 120
Снитюк Виталий Евгеньевич, 171
Снігур Ольга Олексіївна, 166
Соловйов Сергій Олександрович, 185
Сперкач Майя Олегівна, 125
Статкевич Виталий Михайлович, 122
Стефанишин Дмитро Володимирович, 123, 124, 128
Стефанишина-Гаврилюк Юлія Дмитрівна, 124
Ступень Павел Вячеславович, 264
Суржан Антон Олегович, 187
Сухоручкина Ольга Николаевна, 325
Таранян Арам Робертович, 188
Тарасова Ольга Валеріївна, 66
Телелейко Інна Сергіївна, 215
Терентьєв Олександр Миколайович, 253, 335
Терновая Татьяна Ивановна, 112
Тимошенко Юрій Олександрович, 190
Ткаченко Віталій Юрійович, 125
Ткаченко Владимир Александрович, 218
Ткаченко Катерина Петрівна, 327
Ткаченко Наталія Володимирівна, 125
Товажнянский Владимир Игоревич, 87
Торський Адріан Романович, 55
Труш Галина Олександрівна, 285
Тульчинський Петро Григорович, 291
Федь Володимир Вадимович, 328
Федюков Владислав Дмитрович, 92
Филатова Анна Евгеньевна, 174
Фіногенов Олексій Дмитрович, 74
Хает Фаина Ильинична, 329
Харченко Артем Олександрович, 162
Харчук Владислав Васильович, 127
Хіміч Андрій Васильович, 330
Ходак Антон Романович, 190
Ходневич Ярослав Васильович, 128
Хоменко Ольга Володимирівна, 57
Хурцилава Костянтин Вікторович, 332
Чапалюк Богдан Володимирович, 191
Чепков Ігор Борисович, 180
Черненко Ольга Борисовна, 90
Чернуха Ольга Юріївна, 38
Чертов Олег Романович, 152, 307
Чучвара Анастасія Євгенівна, 38
Шаповалова Світлана Ігорівна, 192
Шевченко Євген Тарасович, 222
Шемеляк Роман Олегович, 129
Шенгелия София Вахтанговна, 301
Шецирули Лия Дмитриевна, 244
Шубенкова Ирина Анатольевна, 310, 334
Шубенкова Ірина Анатолівна, 327
Шубенкова Ірина Анатоліївна, 96
Шулькевич Тетяна Вікторівна, 194
Щеголь Юлія Миколаївна, 195
Щука Роман Володимирович, 335
Яворський Ярослав Борисович, 335
Якименко Ірина Василівна, 185
Яковлева Алла Петрівна, 306