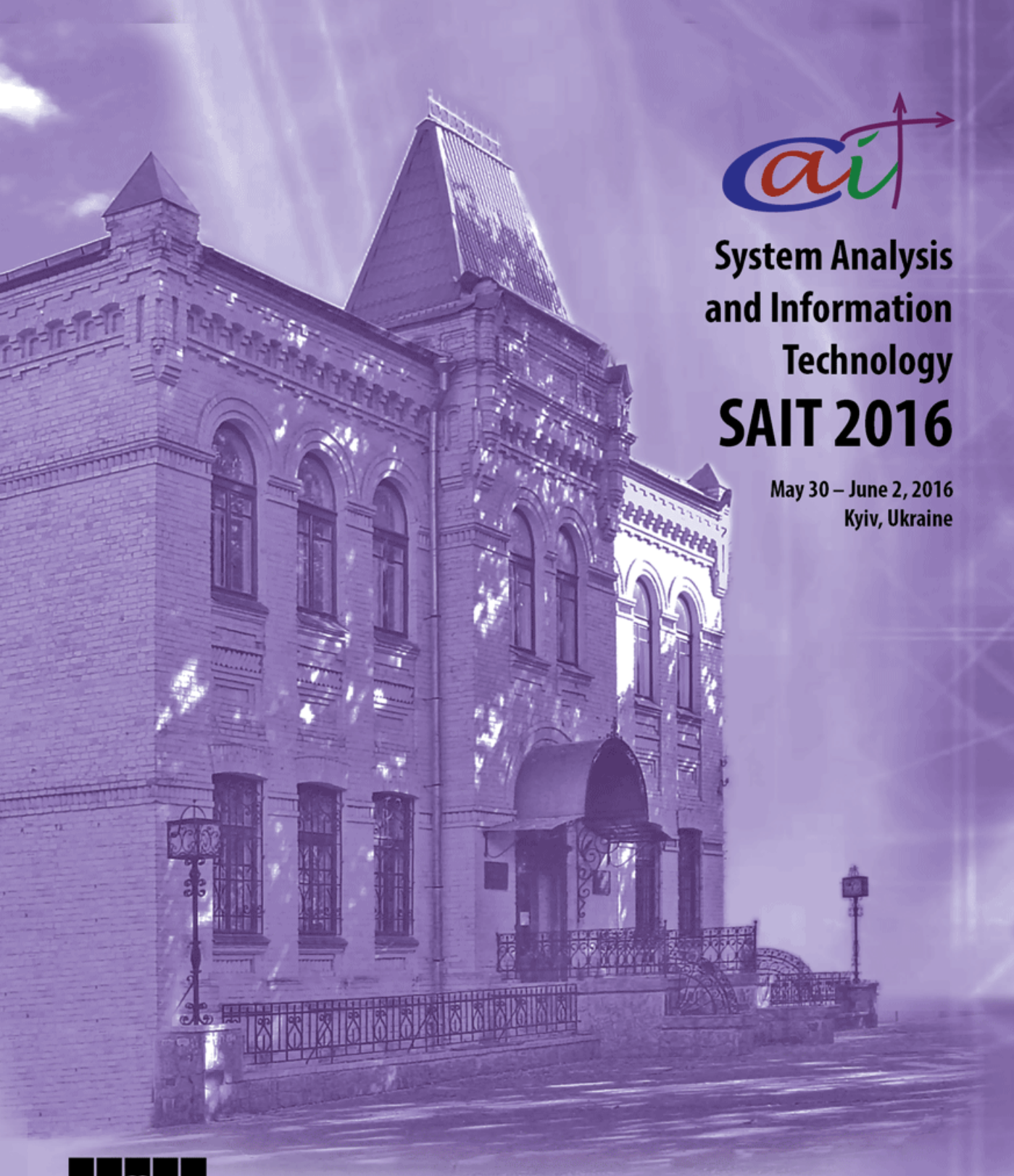


System Analysis and Information Technology

SAIT 2016

May 30 – June 2, 2016
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 "Kyiv Polytechnic Institute"

Nataliya D. Pankratova (Ed.)

System Analysis and Information Technologies

18-th International Conference SAIT 2016
Kyiv, Ukraine, May 30 – June 2, 2016

Proceedings

IASA

Institute for Applied System Analysis
of National Technical University of Ukraine
"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

Design and typesetting:
Mykhailo P. Makukha

System analysis and information technology: 18-th International conference SAIT 2016, Kyiv, Ukraine, May 30 – June 2, 2016. Proceedings. – ESC "IASA" NTUU "KPI", 2016. – 440 p.

С40 **Системный анализ и информационные технологии**: материалы 18-й Международной научно-технической конференции SAIT 2016, Киев, 30 мая – 2 июня 2016 г. / УНК "ИПСА" НТУУ "КПИ". – К.: УНК "ИПСА" НТУУ "КПИ", 2016. – 440 с. – Текст: укр., рус., англ.

С40 **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. – 440 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 NTUU "KPI", 2016
ISBN 978-966-2748-83-3 (print)
ISBN 978-966-2748-84-0 (ebook)
http://sait.kpi.ua

Table of contents · Содержание · Зміст

Plenary talks · Пленарные доклады · Пленарні доповіді **13**
Дорофеев Ю.И., Любчик Л.М. Робастное управление запасами в сложных сетях поставок 14
Згуровський М.З. Комплекс робіт з Форсайту та побудови стратегії соціо-економічного розвитку України на середньотерміновому (до 2020 року) і довготерміновому (до 2030 року) часових горизонтах 16
Петренко А.І. Електронна охорона здоров'я (eHealth) 17
Романенко В.Д., Мілявський Ю.Л. Нові методи управління імпульсними процесами у когнітивних картах 20

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 **22**
Bilushchak Y.I., Chernukha O.Y., Davydok A.Y., Goncharuk V.Y. Mathematical model for mass flow in a strip with randomly disposed sublayer of stochastical thickness 24
Chenarani A., Druzhinin E.A. Analysis and selecting the appropriate technique for identifying risks of designing complex product systems 26
Cojocaru S., Petic M., Coman L., Horos G. Determining "controlled vocabulary" (lexicon of markers) of a text 28
Fomin O.O., Pavlenko V.D. Feature space building for nonlinear systems diagnostics based on wavelet transforms of integral models 30
Gorban N.V., Paliichuk L.S. Dynamics of solutions for non-autonomous reaction-diffusion equations with Carathéodory's nonlinearity 32
Herasymiv I.S., Kovalenko A.E. System diagnosis models for distributed systems 33
Lulu K.M. Analysis of the effect of news related to war on the exchange rate in Ukraine . . 34
Mieshkov O.Y., Novikov O.O. Automated system for identification and human condition diagnostics based on its voice signal analysis 35
Pankratova N.D., Dozirtsiv V.O. Application of methods for text analysis of the emotional tone to identify social disasters 38
Parkhomchuk D.M., Tymoshenko Yu.O. Regularization of second order dynamic method . . 39
Savastiyanov V.V. Extraction of expert sentiments from interview context 40
Shcherbachenko L.V. Methods for improving work with Public Procurement systems 41
Sobko V.O., Bidyuk P.I., Terentiev O.M., Dzhygyrey I.M. Scenario-based forecasting and analysis of air pollution for Ukraine: outlook to 2020 42
Teodorescu H.N., Bolea C. Analysis of probabilities of specified words' occurrences in SN messages related to catastrophes 44
Titchiev I.I. Modelling and verification of evacuation system using Time Petri nets in case of disaster 46
Zorin Yu.M. An Improved Cuckoo Search Algorithm 48
Андреев Н.В., Статкевич В.М. О задаче управления неоднородным процессом рождения и гибели 50

Андреева М.О., Левко Л.Р. Метод для побудови генератора псевдовипадкових рівноважових послідовностей заданої довжини і заданої ваги . . . 51
Андрущенко В.Б., Ланде Д.В. Нові наукометричні сервіси на базі Google Scholar Citations 52
Барановский А.Н., Кириченко Л.О., Кобицкая Ю.А. Рекуррентный анализ фрактальных временных рядов . . . 53
Бардадим І.В. Алгоритм оцінювання параметрів регресійної з випадковим шумом там можливими сезонними коливаннями . . . 54
Буценко Ю.П., Лабжинський В.А. Удосконалений алгоритм голосування в складних технічних системах . . . 56
Васильев В.И., Виштальь Д.М., Любашенко Н.Д. Об оценке доступности сетевой услуги в классе бинарных стохастических моделей . . . 57
Вероцький Д.С. Застосування поєднаних марківських ланцюгів для оцінювання кредитних ризиків . . . 59
Волкова В.Н., Черный Ю.Ю. Системный анализ информационных основ развития социума 60
Гайвась Б.І., Дмитрук В.А., Торський А.Р., Дмитрук А.А. Модель імпульсного осушення шару зерна з врахуванням неоднорідної структури зернини . . . 63
Гамоцька С.Л. Аналіз особливостей моніторингу проектних ризиків в IT . . . 65
Головатюк В.О. Багатокритеріальне оцінювання якості роботи співробітників технічної підтримки компанії з використанням методів підтримки прийняття рішень . . . 67
Гончар В.О. Складність задачі синтезу мережі . . . 69
Горбань І.М., Хоменко О.В. Застосування схеми спійманих вихорів до керування потоком навколо квадратної призми . . . 70
Горелова Г.В. Критерии оценки эффективности процессов в сложных системах в условиях неопределенности, когнитивный аспект . . . 72
Городецкий В.Г., Осадчук Н.П. Моделирование активности нейрона по временным рядам 74
Губарев В.Ф., Дяденко О.Н. Анализ наблюдаемости бортового измерительного космического аппарата . . . 76
Даник Ю.Г., Писарчук А.А., Дюков И.Н., Тимчук С.В. Самоорганизующаяся математическая модель ситуационного управления структурой и параметрами систем мониторинга . . . 77
Даник Ю.Г., Писарчук О.О., Дюков І.М., Тимчук С.В. Математичне забезпечення автоматизованої системи збору та обробки інформації від технічних засобів моніторингу 79
Денисенко О.І. Тривимірне моделювання елементів теплообінних приладів з турбулізаторами . . . 81
Дмитрієнко В.Д., Заковоротний О.Ю., Мезенцев М.В. Оптимізація тягового електроприводу за допомогою геометричної теорії керування . . . 82
Донець Н.П., Ясинський В.К. Існування сильного розв'язку нелінійного стохастичного диференціально-функціонального рівняння з зовнішніми випадковими збуреннями 83
Донченко В.С., Зінько Т.П., Скотаренко Ф.М. Базові структури евклідових просторів в прикладних задачах: кортежні оператори для матричних кортежів . . . 84
Дяченко О.М. Математичне моделювання поведінки гравців в нечітких умовах за неповної інформації . . . 85
Зражевська Н.Г. Оцінювання динамічних мір ризику VaR і CVaR для фондових індексів на основі гетероскедастичної моделі часового ряду . . . 86
Калюжний М.С. Використання інтелектуального аналізу тексту для структурування знань з різних текстових джерел . . . 88
Каніовська І.Ю., Зубарева П.К. Імовірнісні моделі Страффіна для виміру впливу голосу у простих іграх голосування . . . 89
Качаловський А.С. Дослідження методів прогнозування лінійних та нелінійних нестаціонарних процесів . . . 91

Киселева Е.М., Гарт Л.Л., Селищев В.Л., Матяш В.В. Анализ сложности проекционно-итерационных алгоритмов решения линейных интегральных уравнений 92
Киселева Е.М., Притоманова О.М., Журавель С.В. О решении задачи «location-allocation» методами оптимального разбиения множеств 94
Козуля Т.В., Білова М.О. Інформаційно-методична основа визначення системи оцінки екологічності складних об'єктів 96
Козуля Т.В., Ємельянова Д.І. Визначення інформаційно-алгоритмічної бази для комплексної оцінки екологічності системних об'єктів 98
Колбасин В.А. Категоризация большого объема текстовых документов с использованием ElasticSearch и Spark ML 100
Комлєв О.О., Довніч О.С. Термодинамічна система Землі: механізми її самоорганізації . 101
Кораб К.О., Каніовська І.Ю. Тестування скорингової моделі за допомогою коефіцієнту Джині 102
Кравченко А.М. Аналіз та моделювання процесів перехідної економіки 104
Кузенков О.О., Падалко В.Г. Біфуркаційні властивості у вироджених диференціальних моделях природничих процесів 105
Кузнєцова Н.В. Скорингові карти як інструмент аналізу ризиків 106
Левко Л.Р., Андреєва М.О. Про один метод побудови та модифікації графо-логічної моделі багатопроцесорної відмовостійкої системи, підсистеми якої мають спільні процесори 108
Ліщук К.І. Оцінка якості програмних продуктів з використанням експертних методів . 110
Лопатин А.К., Черненко О.Б. Индикаторы общеэкономической динамики в задачах предсказания кризисов 111
Мазурік О.Ю. Порівняльний аналіз моделей оцінювання в рекомендаційних системах . 113
Марченко О.І., Молчанов О.А. Імітаційне моделювання TCP потоків: еволюційно-ігровий підхід 114
Маслянко П.П., Прудський Ю.Ю. Компонентна модель уніфікованої системи оптимізації складу сумішей для ентерального харчування 117
Мигаль Г.В., Мигаль В.П. Концептуальные основания системного подхода к исследованию и анализу безопасности сложных технических систем 119
Мосійчук М.Д., Олефір О.С. Використання поліноміального хаосу для моделювання біологічних процесів та хімічних реакцій клітини 121
Недашковская Н.И. Согласованность и неопределенность экспертных оценок при принятии решений 123
Омельчук А.А., Рудакова А.В., Поливода О.В. Использование алгоритма стохастической аппроксимации для идентификации параметров модели перемещения судна на слипе 124
Орєхов О.А., Орєхова Н.А. Методи верифікації паралельних програм за допомогою мереж Петрі 126
Павлюк Ю.Р., Подладчіков В.М. Розробка методів статистичного аналізу для визначення швидкості корональних хвиль 127
Панкратов В.А. Развитие будущего металлургической промышленной Украины 129
Панкратова Н.Д., Головань М.В. Системна стратегія функціонування складної технічної системи в умовах багатофакторних ризиків 131
Панкратова Н.Д., Кондратова Л.П. Модель гарантированного предотвращения нештатного режима в реальных условиях функционирования сложной технической системы 132
Панкратова Н.Д., Сльота М.Р. Класифікація і розпізнавання рівня небезпеки критичних ситуацій в задачах інформаційного аналізу 134
Панкратова Н.Д., Тарасюк Т.С. Розпізнавання ситуацій за умов неповноти і нечіткості інформації 135
Панкратова Н.Д., Шелест А.В. Исследование развития ИКТ кластера Украины на основе когнитивного моделирования 136

Печурин Н.К., Кондратова Л.П., Печурин С.Н. Модель распределения вычислительных ресурсов в беспроводной компьютерной сети . . . 137
Пікоож А.В. Основа для автоматизованої інтелектуальної системи керування квадрокоптером . . . 138
Погребной Е.А. Моделирование и анализ взаимодействия биологических популяций сложной структуры . . . 140
Попова Ю.А., Подладчиков В.Н. Идентификация неизвестных статистических характеристик шумов для нелинейной модели . . . 141
Потебня А.В., Погорілий С.Д. Дослідження процесів самоорганізації у тематично близькій групі наукових публікацій . . . 142
Пудло І.В., Бідюк П.І. Прогнозування дохідності банківських продуктів з використанням скорингового підходу . . . 145
Путренко В.В. Статистичний аналіз даних землекористування для цілей регіонального управління в Україні . . . 147
Рубин Э.Е., Годлевский М.Д., Голоскокова А.А. Технология планирования улучшения качества процесса разработки программного обеспечения . . . 149
Савченко І.О., Іваненко Є.В. Моделювання воєнних конфліктів за допомогою модифікованого методу морфологічного аналізу . . . 151
Сальніков М.М., Вергун К.В. Відновлення просторової картини еволюції хвильових процесів у сферичному шарі за результатами вимірів вздовж лінії . . . 152
Сащук І.М., Романов О.М. Застосування системного підходу для синтезу структури апаратно-програмних комплексів технічного аналізу сигналів на основі інформаційно-сигнатурних технологій . . . 153
Сирик С.В., Цверчкова А.А. Коррекция и Фурье-анализ конечноэлементных стабилизированных вычислительных схем при использовании технологии mass lumping . . . 155
Слухай Я.О. Алгоритми логічного виведення фактів в OWL-онтологіях . . . 157
Соколовський Я.І., Левкович М.В., Яркун В.І., Ликтей Ю.Т., Стельмах Р.П. Чисельне моделювання двовимірного тепломасоперенесення на підставі похідних дробового порядку . . . 158
Соловйов В.М., Чабаненко Д.М. Мережні міри складності у передбаченні кризових явищ 162
Стасюк А.В. Логічна парадигма подання даних . . . 164
Степаненко А.І., Чкалов О.В. Алгоритми планування дій гравців у віртуальному футболі роботів . . . 165
Стефанишин Д.В. Про ситуаційно-індуктивне моделювання в задачах екстраполяційного прогнозування за даними моніторингу . . . 167
Стоянова А. Анализ преступности в Украине с использованием регрессионных моделей в SAS Enterprise System . . . 168
Супруненко О.О., Лісун А.Г. Засоби імітаційного моделювання детермінованого та недетермінованого вибору в алгоритмічних конструкціях . . . 169
Харчун В.М. Оцінювання якості життя населення України з використанням модифікованих показників МІУР . . . 171
Ходневич Я.В., Стефанишин Д.В., Щодро О.Є., Корбутяк В.М. Про системний підхід до прогнозування розмивів берегів рік на основі імітаційного моделювання з врахуванням геопросторових даних . . . 173
Цегелик Г.Г., Краснюк Р.П. Математичне моделювання оптимального розподілу реплікаційних баз даних у розподілених інформаційних системах . . . 174
Чертов О.Р., Сахаров С.Ю., Юрченко Д.В. Виділення ключових значень з показів CGM для прогнозування нічної гіпоглікемії у хворих на цукровий діабет 1-го типу . . . 176
Шаронова Н.В., Козуля М.М. Знання-орієнтоване інформаційне забезпечення досліджень складних систем . . . 178

Шевченко Д.С. Нестаціонарна задача комівояжера в умовах авіатрансферу з додатковими обмеженнями . . . 180
Яковлева А.П., Мухін В.Є., Корнага Я.І., Курдуп І.О. Про одну оптимізаційну задачу розташування об'єкта на конкретній території . . . 182

Section 2. Intelligent systems for decision-making
Секция 2. Интеллектуальные системы принятия решений
Секція 2. Інтелектуальні системи прийняття рішень **184**
Arkhipov D.V. The rational approach for working with heterogeneous knowledge bases . . . 186
Azzuz I.D. Analysis of decision-making methods in computer poker . . . 187
Bidyuk P.I., Vertiletskij V.G. Decision support system for adaptive forecasting of financial and economic processes . . . 188
Levin D.K. Neo-fuzzy neural networks in forecasting stock prices . . . 190
Markova V.M. Application of Neural Networks for E-shop Goods Demand Forecasting . . . 191
Reznik I.M., Shatalov V.M. Data mining of Wikipedia reveals historical changes of human priorities . . . 192
Romanko O.R., Radchenko O.Yu., Livinskiy O.V. Fractal nature of financial time series: R/S analysis, V-statistic and predictive modelling . . . 194
Shaptala R.V., Kyselova A.G. Location as a context parameter in energy management systems 195
Zaychenko Yu.P., Guskova V.G. Application of fuzzy neural network nefclass for recognition of medical images in diagnostics . . . 196
Zaychenko Yu.P., Ovi Nafas Aghaei Agh Ghamish Bankruptcy Risk Forecasting under Uncertainty Application of Fuzzy Neural Network NEFClass . . . 198
Алексеєв М.О., Безлюднов Д.Є. Визначення аномальних станів у роботі пристроїв IoT 199
Аникин В.Ю. Комплексная оценка риска банкротства путем сравнительного анализа с придприятиями-конкурентами . . . 201
Бідюк П.І., Липчак І.Р. Моделювання гетероскедастичних процесів та побудова функцій прогнозування на основі моделей реальних фінансово-економічних процесів . . . 202
Богун К.В., Бідюк П.І. Побудова моделі ВВП в економіці перехідного преріоду на прикладі України . . . 203
Бойко О.М. Моделювання кредитоспроможності фізичних осіб . . . 205
Ганчукова Д.В., Подладчіков В.М. Узагальнення методу апроксимації експериментальних даних за критерієм гладкості на гетероскедастичні процеси . . . 207
Глоба Л.С., Коваль О.В., Новогрудська Р.Л., Сенченко В.Р. Створення сценаріїв обробки даних на основі онтології . . . 209
Гусак О.В., Алексеєв М.О. Балансування навантаження розподілених гетерогенних інформаційних мереж з допомогою підходів штучного інтелекту . . . 211
Декрет М.В. Підвищення ефективності якості прогнозування часових рядів інструментів фінансових ринків за рахунок агрегації різних довжин тайм-фреймів . . . 213
Дуда О.М., Кунанець Н.Е., Липак Г.І., Мацюк О.В., Небесний Р.М., Пасічник В.В. Консолідація інформаційних ресурсів соціокомунікаційного середовища в проектах "Розумне місто" . . . 214
Дуда О.М., Кунанець Н.Е., Мацюк О.В., Пасічник В.В. Системні комплекси інформаційних технологій у проектах "Розумне місто" . . . 215
Зайченко Ю.П., Четырбок П.В. Распознавание объектов с использованием модульных нейронных сетей . . . 217
Іванущак Н.М. Інформаційні технології моделювання процесів розвитку комп'ютерних мереж в умовах дестабілізаційних загроз . . . 218
Каковська Т.І., Олефір О.С. Використання генетичних алгоритмів для формування персонального раціону харчування людини . . . 220
Караюз І.В., Бідюк П.І. Системний підхід до прогнозування макроекономічних процесів 222

Квитко М.В. Распознавание речи с помощью глубоких рекуррентных нейронных сетей 223
Кондратенко Н.Р., Снігур О.О. Використання інтервальних нечітких множин типу 2 в умовах недовизначеності вхідних даних . . . 224
Кондратьєв І.І. Трекінг швидких об'єктів на відео . . . 226
Кравець І.О., Козловська В.О., Тумко В.С. Система підтримки прийняття рішень для трейдера на базі розпізнавання фігур майбутнього розвороту тренду та створення бази правил для нечіткої нейронної мережі . . . 227
Курінний І.В. Композиційні методи класифікації та регресії . . . 228
Мейтус В.Ю. Знания в 1-интеллектуальных системах . . . 229
Мейтус В.Ю., Цепкова Д.П. База знаний с неполной информацией представления знаний 231
Миселюк А.Ю. Глибинні нейронні мережі в медицині . . . 233
Мурга Н.А. Принятие решение об открытии позиции по финансовому инструменту и её структуре на основе прогнозирования и вычисления статистических характеристик временного ряда . . . 234
Новицкая Е.Г., Железко Б.А. Интеллектуальная система поддержки принятия решений как элемент регионального электронного правительства . . . 236
Огурцов М.И., Ходзинский А.Н. Про гибридный алгоритм комбинаторной оптимизации решения транспортных задач повышенной сложности . . . 238
Перепеличный Я.И., Янковский И.П., Дидковская М.В. Интеллектуальная система построения контекстной рекламы . . . 240
Пилипенко І.В. Прогнозування тривалості процесу тестування програмного забезпечення 241
Піпіч А.А. Виявлення та усунення технічних конфліктів в контекстно-залежних системах 243
Поворознюк А.И., Поворознюк О.А. Комплексная оценка этапов диагностически-лечебных мероприятий в системах поддержки принятия решений в медицине . . . 244
Поршнєва К.О. Дослідження та розробка мультиагентної системи керування базами даних 246
Ржепішевський А.Л. Архітектура системи доповненої реальності у відеочатах . . . 247
Романенков Ю.А., Вартанян В.М. Задачи прогностического обеспечения систем поддержки принятия стратегических решений . . . 249
Саржинська М.С., Кисельов Г.Д. Виявлення прихованих взаємозалежностей між даними на прикладі аналізу тональності відгуків користувачів інтернет . . . 250
Скірко І.О. Система розпізнавання текстур на зображеннях з БПЛА . . . 252
Слободюк А.Ю. Построение модели для расчета лучшего маркетингового предложения для клиентов банка . . . 253
Таранян А.Р., Алфимцев А.Н. Применение ковариационных матриц для отслеживания людей в видеопотоке . . . 254
Тертишний А.А. Розподілене навчання нейронних мереж у браузерах . . . 256
Филатова А.Е. Исследование альтернативных диагностических признаков в интеллектуальных компьютерных кардиологических системах поддержки принятия решений 258
Штельмах С.О., Олефір О.С. Дослідження методів очистки мовленнєвого сигналу від шуму . . . 260
Яковенко М.М. Моделі і методи інтелектуального аналізу даних для оцінювання кредитоспроможності фізичних осіб . . . 261
Яковлєва О.Є., Бідюк П.І. Прогнозування динаміки процесів у фінансових системах . 262

Section 3. High Performance Computing and Microsystems Engineering
Секция 3. Высокопроизводительные вычисления и разработка микросистем
Секція 3. Високопродуктивні обчислення та розробка мікросистем 264
Antonyuk S.A., Artyukhov V.G. Digital Signal Processing and Applications Design with the TMS320C6713 DSK . . . 266
Kharchenko K.V. An Architecture and Test Implementation of Data Flow Virtual Machine 268

Kovalenko A.E. Structure decomposition for system-level diagnosis models in distributed fault-tolerant systems . . . 269
Бритов О.А., Тарасенко О.Г. Програмований кодер циклічного коду . . . 271
Булах Б.В., Крамар О.В. Мікросервісна архітектура з семантичною складовою для комплексів інженерних розрахунків . . . 273
Денисов С.В. Взаємодія агентної моделі та мікросервісів . . . 274
Капшук О.О. Аутентифікація користувачів в інформаційних системах і веб-сервісах з використанням двохфакторних і біометричних технологій . . . 276
Кирюша Б.А. Використання масиву MEMs акселерометрів, розташованих на одному кристалі IC, для зменшення похибки вимірювань . . . 278
Литвиненко А.М. Розподілена обробка великих об'ємів зображень за допомогою технологій Hadoop . . . 279
Масліков В. Використання графічних процесорів для розв'язання задачі фолдингу білків 280
Надточій Т.В. Реалізація функції імітації присутності в системі "розумний дім" на базі Arduino Nano 3.0 з обробкою даних на хмарному сервері . . . 282
Петренко О.О. Пошук семантичних сервісів у репозитарії сервісів . . . 283
Попович Є.С., Безносик О.Ю., Фіногенов О.Д. Використання розподіленої архітектури зберігання даних при вирішенні задачі скорочення розмірності великих щільних графів . . . 285
Романов В.В., Корначевский Я.И. Выбор параметров моделей компонентов интегральных микросхем по критерию чувствительности . . . 287
Свістунов С.Я., Юрченко В.В. Досвід впровадження технологій хмарних обчислень в Інституті теоретичної фізики ім. М.М.Боголюбова НАН України для обробки даних по експерименту ALICE в CERN . . . 290
Тесленко О.К., Шаповал І.І. Нероздільні блочні коригуючі коди . . . 292

Section 4. Progressive information technologies
Секция 4. Прогрессивные информационные технологии
Секція 4. Прогресивні інформаційні технології **294**

Ademola Samuel Adegoke Network locality awareness comparison and method of improvement of existing models . . . 296
Bakun S.A., Lytvyniuk A.A., Terentiev O.M. Using SAS Enterprise Miner to build scoring card to evaluation solvency of individuals . . . 299
Basyuk T.M. Web site promotion on Google . . . 301
Didenko D.G. Fuzzy recovery dictionary by the id and the name into database . . . 303
Ivanova Y.V., Terentiev O.N., Korshevnyuk L.O., Prosyankina-Zharova T.I. Using modified logistic regression to increase customer response rate to marketing campaigns . . . 304
Kavrelishvili A.G. Hotel Management System, Development stages and Processing . . . 306
Malinovskyi A.D., Nikolaiev S.S. The Architecture of a mobile system for heart health monitoring . . . 307
Markov D.K. Improved frustum culling . . . 310
Povisenko S.V. Using of Saga pattern in Asynchronous Messaging Architecture for Big Data Processing . . . 312
Radovanovic M., Strbac D. The Standardization of Geospatial Data in Serbia - Experiences and Problems . . . 314
Sidletskiy V.M. The automated operational control system for the food industry . . . 316
Terentiev O.M., Makukha M.P., Savastynov V.V., Oparina E.L. Analysis and modeling the dynamics changing of registered crimes taking into account the macroeconomic and political situation in Ukraine . . . 318
Trofimov A.S., Zenkovich M.V. Virtualization approaches in the area of fault-tolerant embedded systems . . . 320

Tsymbal M.D. Remote electronic identification methods . . . 322
Арчвадзе Н.Н., Пховелишвили М.Г., Шецирули Л.Д. Особенности параллельного программирования на языке Haskell . . . 324
Балашевич Р.В., Гіоргізова-Гай В.Ш. Забезпечення якості обслуговування трафіку VoIP у мережах кабельного телебачення . . . 325
Бондаренко А.С. Использование мобильных устройств для мониторинга состояния человека с диабетом 1-го типа . . . 326
Боярінова Ю.Є., Каліновський Я.О., Сукало А.С. Застосування узагальнених кватерніонів у цифровому підпису . . . 327
Бугаєва Л.М., Безносик Ю.О. Програмні засоби обчислювальної гідродинаміки для розв'язання задач хімічної інженерії . . . 328
Будник С.А. Який вибір ми зробимо: HTML 5 чи Flash? . . . 330
Буценко Ю.П., Савченко Ю.Г., Семенов С.В. Алгоритми формування послідовностей псевдовипадкових чисел . . . 332
Виклюк Я.І., Гусак О.М. Методи і засоби підвищення інформативності даних дистанційного зондування Землі . . . 334
Галіцька В.В. Алгоритм штучної бджолиної колонії для задачі комівояжера . . . 335
Галушко М.О. Бібліотека наукових публікацій з використанням RDF-сховищ . . . 337
Ганиева С.А., Танырвердиев Ч.Г., Мутталибова Ш.Ф. Метод создания электронной кадастровой информационной базы недвижимости и цифровых кадастровых карт на примере города Сумгаит . . . 338
Дичко В.М. Система співставлення супутникових зображень . . . 341
Діденко Д.Г., Журке Б.С. Імітаційна модель для прогнозування фармацевтичного ринку України . . . 342
Діденко Д.Г., Климковська К.Д. Розподілене обчислення і прогнозування важливості брендів в Інтернеті . . . 343
Діденко Д.Г., Конєв Д.В. Прискорення пошуку напівдублікатів текстових файлів за допомогою алгоритму MapReduce . . . 345
Діденко Д.Г., Крюков О.О. Модель адаптивного навчання на базі алгоритму кластеризації категорійних даних CLOPE . . . 346
Євтухов С.К. Стеганографічний метод ін'єкціювання інформації в зображення динамічного кодування . . . 348
Журавлёв П.В. Выбор архитектуры корпоративных серверных приложений . . . 350
Забелин С.И. Контрмеры против Arp-Spoofing в SDN . . . 352
Загороднюк А.А. Использование ORM при разработке корпоративных приложений . . . 353
Ивченко Д.А. Использование мобильных устройств для мониторинга состояния человека с астмой . . . 354
Калінін В.В., Боярінова Ю.Є. Огляд існуючих нейропакетів та перспективи їх застосування . . . 356
Калінін В.В., Черненко П.Р. Метадані у алгоритмах пошуку . . . 357
Капустін В.В., Маслянко П.П. Система мультипоточного тестування веб застосунків на основі алгоритмів паралельних обчислень . . . 358
Карпенко О.В., Голубовський О.І. Методи забезпечення якості обслуговування в мережах IP . . . 360
Кириленко Е.Г., Стрельчук Е.А. Методический инструментарий для оценки удобства использования пользовательского интерфейса информационного веб-сайта . . . 361
Киричек Г.Г., Котов Р.О. Автоматичне розгортання мережних сервісів . . . 362
Кириченко В.Е., Терентьев О.М., Связінська Н.О. Застосування наївного та дереводоповненого байєсівський класифікаторів для прогнозування кредитоспроможності фізичних осіб . . . 364

Кислий Р.В., Письменний І.О. Мікросервісна архітектура в побудові подійно-орієнтованої системи моніторингу та аналізу стану здоров'я пацієнта . . . 366
Козленко А.С., Подладчиков В.Н. Минимизация смещения последовательных изображений на основе метода наименьших квадратов . . . 367
Копичко С.М., Бобрик І.М. Аналіз моделей управління запасами для СППР . . . 368
Корепанова Н.Л., Лебедева М.А. Помехоустойчивая облачная платформа на базе ОС LINUX . . . 369
Кошкин Е.Г. Гибкая методология управления проектами как инструмент для качественного контроля разработки программного обеспечения . . . 371
Кравчук Є.С. TensorFlow, як нове слово в області систем машинного навчання . . . 372
Крижанівський В.В. Система оптичного 3D сканування . . . 374
Круковский М.Ю. Дискретное моделирование рабочих групп . . . 375
Круш І.В. Особливості побудови веб систем на базі мікросервісів . . . 376
Кузнецова Ю.А. Решение оптимизационной задачи визуализации управляющих алгоритмов методами программирования в ограничениях . . . 377
Лавренюк А.М., Лавренюк С.І. Аспекти використання C++ AMP для моделювання та обробки даних . . . 379
Ланде Д.В., Сулема О.К. Ієрархізація моделі предметної галузі шляхом визначення центральних вузлів . . . 381
Литвинов В.А., Майстренко С.Я., Хурцилава К.В. Классификация основных правил и методов проверки целостности информационных ресурсов (ИР) ГИС . . . 383
Лютенко И.В., Лиликович С.А. Измерение субъектной диверсности в условиях многоверсионной разработки программного обеспечения . . . 385
Магас В.В. Інтелектуальний аналіз просторових даних . . . 386
Марковський О.П., Гура Е.С. Метод побудови ортогональних систем нелінійних булевих SAC–функцій . . . 388
Марченко О.І., Погорелов В.В. Використання розширеної форми статичного однократного присвоєння для трансляції програм . . . 389
Медведський А.М. Застосування WhitestormJS для оптимізації 3Д графіки в онлайн іграх 390
Михалько В.Г. Використання асинхронної моделі програмування для розробки високонавантажених веб-серверів . . . 392
Михняк О.О. Комп'ютерна система обробки зображень на сітківці ока . . . 394
Морозов К.В., Потапова Е.Р. О самотестировании многопроцессорных систем . . . 396
Осадчий Д.Ю. Auto-Layout как способ построения графических и мультимедийных интерфейсов для приложений операционной системы iOS . . . 397
Пашаев Н.М., Рагимов Р.М., Пашаева М.М., Мутталибова Ш.Ф. Аэрокосмический мониторинг динамики развития элементов наземной инфраструктуры Апшеронского региона на основе космических изображений и технологии ГИС . . . 398
Поворознюк Н.І., Гридчук Д.Т., Шибін О.В. Частотно-часовий аналіз сигналів при аускультації . . . 402
Прогонов Д.О. Стегодетектор на основі мультифрактального аналізу цифрових зображень 403
Роганов Е.В. Сравнительный анализ эффективности алгоритмов трекинга . . . 404
Романкевич В.А., Дейнеко А.Л., Олейник В.В. О диагностировании многопроцессорных систем с ограничением на число допустимых неисправностей . . . 405
Романкевич В.А., Малышева М.О., Примак И.К. Об одном способе формирования реберных функций GL-модели . . . 406
Романовський О.В., Сало А.М. Технології захисту платіжних транзакцій у кіберфізичних системах . . . 408
Рубець А.В. Моделювання систем високої доступності . . . 410
Савчук В.В., Лозицький О.А., Пасічник В.В. Динамічне формування персоніфікованого мультимедійного контенту туристичного путівника міських екскурсійних маршрутів 411

Сергеєв Д.С. Методика оцінки якості роботи природно-мовних пошукових систем . . . 413
Сергеєв А.В. Роль децентралізованих мереж у побудові некомерційних систем доставки контенту . . . 416
Скукіс О.Є. Підхід до побудови програмних систем для розв'язання задач транспортної логістики . . . 417
Слободянюк Б.Р., Павловський В.І. Засоби інтеграції успадкованих систем з новітніми системами предачі даних . . . 418
Сутула О.В. Порівняння методів розробки для мобільної платформи Android . . . 420
Сухомлин І.Б. Технологія тестування мобільних додатків для контролю здоров'я, що підключаються до зовнішніх приладів . . . 421
Сухоручкина О.Н., Прогонный Н.В., Воронов М.А. Технология использования измерений сенсоров-дальномеров в задачах управления автономным мобильным роботом 422
Томашевський В.М., Дифучин А.Ю. Візуальне середовище моделювання дискретно-подійних систем . . . 424
Ухина А.В., Ситников В.С., Ситникова В.А. Управление процессом предварительной обработки информации в экспериментальных медико-биологических исследованиях при использовании специализированных компьютерных систем . . . 426
Шевчик М.О. Особливості діагностики нерозпізнанного інтернет трафіку . . . 428
Шеренковський А.О. Програмне забезпечення медіа-серверу Інтернет радіо . . . 429
Шишков О.Н. Распознавание рукописного ввода с помощью рекуррентный нейронных сетей . . . 431
Шліхтенко Н.М., Ярмак Д.О. Паралельні алгоритми розв'язування систем лінійних алгебраїчних рівнянь та алгебраїчної проблеми власних значень . . . 433
Шпортько О.В., Шпортько Л.В. Особливості реалізації розбиття даних зображень на однорідні блоки перед застосуванням арифметичного кодування . . . 434

Authors · Авторы · Автори **436**

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

Дорофеев Ю.И., Любчик Л.М.
Национальный технический университет "ХПИ", Харьков, Украина

Робастное управление запасами в сложных сетях поставок

Введение. Среди современных проблем управления сложными технико-экономическими системами важное место занимают проблемы управления сетями поставок, а именно совокупностью взаимосвязанных объектов, осуществляющих добычу сырья, производство, хранение, транспортировку и распределение материальных ресурсов с целью удовлетворения потребительского спроса и получения прибыли. Примерами служат производственные, транспортные системы, системы распределения ресурсов (воды, электроэнергии) и т.п.

Каждый узел сети поставок в реальном времени принимает заказы на поставку ресурсов от других узлов, а также от внешних потребителей, и формирует заказы узлам, которые являются для него поставщиками ресурсов. Управление запасами заключается в определении размеров заказов на их пополнение и моментов времени формирования заказов. Определение оптимальных уровней страховых запасов ресурсов в узлах сети поставок, которые рассматриваются в качестве задающих воздействий соответствующих многомерных регуляторов, и разработка методов синтеза стратегии робастного управления запасами в сетях поставок в условиях неопределенности является актуальной научной и практической проблемой.

С точки зрения теории управления объемы спроса, поступающие из внешней среды, целесообразно рассматривать в качестве внешних возмущающих воздействий. Выбор стратегии управления запасами определяется характером внешнего спроса. В условиях неопределенности спроса используется концепция «неизвестных, но ограниченных» воздействий [1], при этом соответствующая модель спроса характеризуется интервальной неопределенностью. Другим источником неопределенности является наличие транспортных запаздываний. Предполагается, что номинальные значения длительности транспортировки и переработки ресурсов в узлах сети известны. Однако, в процессе функционирования эти параметры могут отличаться от своих номинальных значений. В результате возникает необходимость обеспечения робастности системы управления запасами относительно вариаций указанных параметров.

Особенностью рассматриваемой задачи является наличие ограничений на объемы хранилищ и размеры заказов. В теории управления традиционно рассматривают ограничения на векторы переменных задачи, заданные в какой-либо норме. Однако, для задач управления запасами характерно требование неотрицательности значений переменных, что приводит к необходимости учета несимметричных ограничений на значения состояний и управлений.

Для математического описания управляемой сети поставок предложена [2] дискретная модель с запаздыванием в пространстве состояний, уравнения которой описывают изменение уровня запасов каждого вида ресурсов с течением времени. В качестве переменных состояний рассматриваются наличные уровни запаса ресурсов. Управляющими воздействиями являются размеры заказов на поставку ресурсов, формируемые узлами в текущем периоде. На основе расширения пространства состояний построена модель без запаздывания. Для учета неопределенности интервалов задержки пополнения запасов обоснована необходимость использования матрицы динамики с параметрической неопределенностью аффинного типа. В результате построена расширенная модель сети поставок с параметрической неопределенностью, у которой область неопределенности представляет собой выпуклое многогранное множество.

Синтез ограниченного робастного управления запасами. Закон управления строится на основе периодической проверки уровней запаса ресурсов в виде линейной динамической обратной связи по сигналу рассогласования между наличными и страховыми уровнями запаса ресурсов, которые вычисляются с помощью продуктивной модели Леонтьева на основании верхних граничных значений внешнего спроса с учетом запаздываний.

Для подавления влияния внешнего спроса на уровни запаса ресурсов, одновременно с обеспечением робастной устойчивости замкнутой системы применен метод инвариантных эллипсоидов, который позволяет свести задачу синтеза оптимального регулятора к поиску

наименьшего инвариантного эллипсоида замкнутой системы. С помощью техники линейных матричных неравенств задача синтеза регулятора сведена к последовательности задач полуопределенного программирования, которые решаются численно в реальном времени. Основным недостатком указанного подхода является консерватизм полученных результатов [3], который проявляется в том, что с практической точки зрения границы робастности оказываются неоправданно заниженными. Для преодоления указанного недостатка построена параметризованная модифицированная квадратичная функция Ляпунова замкнутой системы. Использование дескрипторного подхода [4] позволило добиться уменьшения степени консерватизма результатов управления за счет того, что при вычислении матрицы регулятора не используются матрицы, зависящие от вектора неопределенных параметров модели.

Синтез децентрализованной системы управления запасами. Большинство процедур, разработанных для анализа и синтеза систем управления запасами в последние десятилетия, используют централизованный подход, когда вся информация о текущем состоянии системы передается на единый регулятор, формирующий управляющие воздействия для всех узлов. Однако, для сетей поставок перспективным является децентрализованный подход, при котором исходная оптимизационная задача заменяется набором локальных задач меньшей размерности, решаемых параллельно и независимо. Решение задачи синтеза децентрализованной стратегии управления запасами получено на основе робастной стабилизации подсистем. Решена задача анализа устойчивости децентрализованной системы управления запасами, которая с помощью метода векторных функций Ляпунова и метода сравнения сведена к задаче проверки условия нильпотентности нестационарной матрицы динамики позитивной системы сравнения.

Анализ условий разрешимости задачи синтеза управления. Необходимыми и достаточными условиями существования допустимого управления для рассматриваемой задачи являются:

1) условие управляемости модели сети поставок;

2) условие достаточности ресурсов управления, которое допускает следующую интерпретацию в терминах теории множеств: выпуклый многогранник, описывающий влияние внешних возмущений, должен находиться строго внутри выпуклого многогранника, описывающего ограничения на ресурсы управления.

На практике к регуляторам и переходным процессам в замкнутых системах предъявляются различные инженерные требования. Одним из наиболее значимых является учет ограниченности ресурса управления. Спецификой задач управления запасами является неотрицательность значений переменных, поэтому применяется линейное управление с несимметричными ограничениями, которое определено в ограниченной области фазового пространства. Задача оценивания допустимой области в пространстве управляющих воздействий сформулирована в терминах разрешимости системы билинейных матричных неравенств, для решения которой предложен итерационный алгоритм [5].

Литература. **1.** Bertsekas D.P. Recursive state estimation for a set-membership description of uncertainty / D.P. Bertsekas, I. Rhodes // IEEE Trans. Automat. Control. – 1971. – Vol. 16. – P. 117-128. **2.** Дорофеев Ю.И. Построение математических моделей управляемых сетей поставок с учетом запаздываний потоков / Ю.И. Дорофеев, А.А. Никульченко // Системні дослідження та інформаційні технології. – 2013. – № 1. – С. 16-27. **3.** Щербаков П.С. Приближенные методы в параметрической робастности линейных систем управления : дис. ... д-ра физ.-мат. наук : 05.13.01 / Щербаков Павел Сергеевич. – М., 2004. – 215 с. **4.** Дорофеев Ю.И. Дескрипторный подход к синтезу ограниченного робастного гарантирующего управления запасами с использованием параметризованной функции Ляпунова / Ю.И. Дорофеев // Радіоелектроніка, інформатика, управління. – 2015. – № 4(35). – С. 87–95. **5.** Дорофеев Ю.И. Анализ условий разрешимости задачи синтеза ограниченного стабилизирующего управления запасами с помощью билинейных матричных неравенств / Ю.И. Дорофеев // Інформатика та математичні методи в моделюванні. – 2015. – Том 5. – № 2. – С. 177-190.

Згуровський М.З.
НТУУ "КПІ", Київ, Україна

Комплекс робіт з Форсайту та побудови стратегії соціо-економічного розвитку України на середньотерміновому (до 2020 року) і довготерміновому (до 2030 року) часових горизонтах

Побудова сценаріїв соціо-економічного розвитку України (Форсайт 2016) ґрунтується на дослідженнях, виконаних в 2008 і 2015 роках (Форсайт 2008 і Форсайт 2015) Світовим центром даних з геоінформатики і сталого розвитку і являє собою подальший розвиток і логічне продовження цих досліджень, http://wdc.org.ua/sites/default/files/WDC-IASA-FORSIGHT-UA.pdf.

При цьому під *сценарієм* будемо розуміти гіпотетичну послідовність взаємопов'язаних подій, яка показує, як за відомих початкових умов, ймовірної дії сукупності різноманітних факторів в майбутньому, може послідовно змінюватися стан досліджуваного об'єкта.

У дослідженні Форсайт 2008 розглядалося три часових відрізки розвитку держави, для кожного з яких було проведено SWOT-аналіз і побудовано відповідні сценарії розвитку:

- короткостроковий (до 2010 року) – сценарій "Старі політики";
- середньостроковий (2010–2020 роки) – сценарій "Нова генерація";
- довгостроковий (2020–2050 роки) – сценарій "Швидкий розвиток".

Аналіз подій, що відбулися в Україні на часовому відрізку з 2008 року по 2016 рік показав, що сценарій "Старі політики" з великим ступенем збіжності відбувся, тобто "Революція гідності" 2013–2014 років залишила в минулому покоління старих політиків. Після цього розпочався якісно новий етап розвитку країни, який характеризується приходом до влади нової генерації політиків, що дає країні великі шанси для подальшого прискореного розвитку, але й приносить небачені раніше нові випробування і виклики.

З урахуванням цих нових реалій в дослідженні Форсайт 2015, на основі методології сценарного планування та SWOT-аналізу, було побудовано групу з восьми сценаріїв розвитку майбутньої економіки України: чотири на середньостроковому (до 2020 року) та чотири на довгостроковому (до 2030 року) часових горизонтах. Ці сценарії призначені для використання людьми, що приймають рішення на рівні держави, інституціями громадянського суспільства й міжнародними організаціями з метою розроблення раціональної політики та конструктивних планів соціально-економічного розвитку України на зазначених часових горизонтах.

Враховуючи дуже високу соціальну динаміку, яка має місце після "Революції гідності" на середньостроковому часовому горизонті (до 2020 року), доцільно продовжити побудову групи сценаріїв соціо-економічного розвитку України і здійснювати періодичну (один раз на рік) корекцію цих сценаріїв. Цій меті і присв'ячується наступне дослідження Форсайт 2016.

Для побудови сценаріїв використовувалися наступні методи: методи Делфі та SWOT-аналізу, метод сценарного планування, методика "Двох вісей", методика "Гілки", методика "Конусу правдоподібності".

Петренко А.І.
Інститут прикладного системного аналізу НТУУ «КПІ», Київ, Україна

Електронна охорона здоров'я (eHealth)

Розглядаються сучасні дослідження з галузі удосконалення охорони здоров'я в Європі.

Вступ. Комп'ютеризація у сфері охорони здоров'я на базі ІКТ технологій проходить через три етапи управління даними: збір даних, спільне використання даних і аналіз даних. Перший етап характеризується широким застосуванням електронних персональних медичних карт (ЕПМК) для збору і фіксації медичної інформації про пацієнта. На другому етапі необхідність обміну даними між всіма учасниками процесу (пацієнтом, лікарем, госпіталем, швидкою допомогою) стає очевидною і реалізується мережевими засобами. Охорона здоров'я в даний час вступає в третій етап аналізу даних, який пов'язаний зі створенням розподілених сховищ даних, або Enterprise Data Warehouses (EDW), коли дані збираються і спільно використовуються для аналізу різних аспектів процесу медичного обслуговування, що відображаються в характері агрегованих даних.

Задача сьогодення полягає в тому, щоб знати про пацієнта якомога більше і протягом як можна більш тривалого часу , щоб було можливо розпізнати тривожні ознаки серйозної хвороби на ранній стадії, в результаті чого лікування стає набагато простішим (і менш дорогим) в порівнянні з випадком виявлення хвороби на більш пізньому етапі. Мета полягає в тому, щоб зібрати дані з різних джерел (наприклад, медичних і страхових записів, портативних датчиків, лабораторних обстежень та медичних мобільних додатків, генетичних даних та медичних зображень, і навіть соціальних медіа), щоб скласти повну картину пацієнта як індивіда і адаптувати для нього пакет послуг охорони здоров'я, що пропонується. Це потребує в'язання даних і отримування потенційно цінної інформацію з неструктурованих даних автоматизованим економічно ефективним способом із застосуванням мобільних додатків, додатків управління лікуванням захворювання, платформ соціальної інтеграції, широкого діапазону зв'язку, Інтернет речей, хмарних обчислень.

Європейські проекти. Хоч організація системи охорони здоров'я є національною або регіональною компетентністю, постало питання про можливість координаційних дій на рівні ЄС для створення транс- європейської системи охорони здоров'я, що не знає кордонів. З цією ціллю в 2008-2009 рр. Європейською Комісією започатковано спеціальне дослідження ***SMART 2007/0059*** для створення правової основи для сумісних електронних персональних медичних послуг в Європі, для дослідження стану і особливостей систем охорони здоров'я кожної держави-члена ЄС. Це сприяло порівнянню і потенційному об'єднанню всіх національних законодавств в державах-членах ЄС і наданню детальної інформації про 10 прикладів передового досвіду (скажімо, про те, як Барселонська телемедична клініка надає послуги шведським і датським лікарням).

Проект SMART 2007/0059 показав, що рівень використання ІКТ в секторі охорони здоров'я значно відрізняється від однієї держави-члена до іншої. Був зроблений висновок, що сектор охорони здоров'я в Європі обов'язково потрібно розглядати як дуже фрагментований ландшафт і що ця фрагментація не буде усунута найближчим часом. Європейський нормативна база в області електронної охорони здоров'я повинна не тільки сприймати це різноманітність до уваги, але й і принципово спиратися на неї. Ще однією перешкодою для надання транскордонних послуг електронної охорони здоров'я є розбіжність в рівні зобов'язань. Якщо пацієнт бажає використовувати транскордонні послуги системи електронної охорони здоров'я, що надаються його провайдером, яка встановлена в іншій державі-члені, він буде негайно стикається з невизначеністю щодо діючих в ній прав для пацієнта. Надання транскордонних послуг електронної охорони здоров'я може бути полегшено, якщо провайдери не обов'язково повинні зареєструватися або звернутися за ліцензією в кожній державі-члені окремо.

Інший проект ***e-SENS*** з європейських мережевих послуг тривалістю 36 місяців було започатковано Європейською Комісією в квітня 2013 року. Завдяки розробці технічних компонентів для безшовної електронного зв'язку, цей проект сприяє наданню транскордонних цифрових послуг в ряді галузей: електронне правосуддя, бізнес, електронної ідентифікації ID, електронної закупівлі та електронної охорони здоров'я. Він заснований на проекті EC SMART 2011/0074 з аналізу потреб в транскордонних послугах та оцінки організаційних, правових, технічних і семантичних бар'єрів, які необхідно провести і подолати, щоб розгорнути ключові транскордонні громадські послуги з високим рівнем впливу. Проект вказує на потенційні вигоди від реалізації прикордонних комунальних послуг і забезпечує сценарії для обговорення, якщо і як ці послуги можна було б реалізувати до 2015 року.

Проект e-SENS буде розвивати цифрову інфраструктуру для підвищення якості громадських послуг в ЄС. Це сприятиме реалізації європейської політики, зокрема, програми Digital Agenda для Європи, в якій вказується, що "держави-члени погодити спільний перелік ключових транскордонних комунальних послуг", який також включає транскордонні послуги в області електронної охорони здоров'я, зокрема, в таких напрямках, як електронні медичні записи, телемедицина та електронний рецепт.

У 2013 році Європейська комісія також ініціювала проект ***EIF*** з розробки загальної сумісної платформи електронної охорони здоров'я (eHealth EU Interoperability Framework), яка може допомогти розширити цифровий єдиний ринок для охорони здоров'я, відкриваючи конкуренцію і знижуючи витрати для майбутніх реінвестицій в оновлення системи охорони здоров'я. Проект визначає бачення сумісної платформи електронної охорони здоров'я на чотирьох рівнях: технічному, семантичному, організаційному та правовому. Він також оцінює деякі технічні характеристики проекту ***epSOS*** (European Patients - Smart open Services) та розробок двох організацій Continua Health Alliance і Integrating the Healthcare Enterprise (IHE) на відповідність критеріям ідентифікації додатку II Положення про європейську стандартизацію (Regulation on European Standardization).

Зокрема, проект epSOS (2012-2014) був спрямований на проектування, будівництво і оцінювання сервісної інфраструктури, яка би демонструвала транскордонну сумісність електронних персональних медичних карт в Європі. Тому EpSOS було зосереджено реалізації обміну рецептами і резюме електронних медичних карт пацієнтів серед країн-партнерів ЄС.

Трьох річний проект ***SUSTAINS*** спрямований на розробку і розгортання кошика послуг в 11 європейських регіонах, забезпечуючи доступ пацієнтів до електронних персональних медичних карт. Хоч кожна з цих послуг має певну мету, усі послуги сприяють досягненню нової парадигми в галузі охорони здоров'я, в якому пацієнт більше не є пасивним суб'єктом, а стає активним гравцем в управлінні його власного здоров'я.

Проект ***United4Health*** (UNIversal solutions in TElemedicine Deployment for European HEALTH care) спрямований на використання і розгортання інноваційних варіантів застосування телемедицини, раніше випробуваних в рамках проекту European Renewing Health при лікуванні пацієнтів з цукровим діабетом, хронічними обструктивними або серцево-судинних захворюваннями. Ці рішення дозволять пацієнтам управляти лікуванням їх захворювань шляхом регулювання вибору і дози ліків, сприяють дотриманню лікування, допоможуть фахівцям при виявленні ранніх ознак погіршення хвороби і тим самим сприятимуть стійкості системи охорони здоров'я. Нарешті, проект ***SemanticHealthNet (SHN)***, що виконувався в 2012-2015 рр., присвячено поліпшенню семантичної сумісності змісту електронних персональних медичних карт шляхом внесення в них інтегрованого семантичного резюме. Цей підхід був продемонстрований на прикладах хронічної серцевої недостатності і серцево-судинної профілактики.

Результати проектів e-SENS, EIF, epSOS і SHN будуть використовуватися для розгортання транскордонних послуг електронної охорони здоров'я в рамках ***Connection Europe Facility***

(CEF), але вони також потенційно можуть бути використані для національного або регіонального розгортання. Були деякі спроби створити єдине вікно для медичної інформації. Таке завдання дуже важке і навіть ***Google Health***, який був однією з перших спроб інтеграції, не впорався, тому що завантаження даними було клопітною справою. Microsoft була більш успішним з своїм ***HealthVault***, що являє собою безпечне місце для збирання, зберігання, використання і поширювання інформації онлайн, яку пацієнти завантажують самі. IBM недавно оголосила, що збирається зосередити свою систему ***Watson*** на галузь охорони здоров'я (для підтримки діагностики і лікування).

Висновки. Варто відзначити, що Європа ще не отримати максимальну вигоду від сумісності систем електронної охорони здоров'я та використання технологій Big Data. Недоліки в нормотворчій діяльності, державної закупівлі та координації між державними органами перешкоджають цифровим послугам і пристроям, що використовуються європейцями, працювати разом і так успішно, як вони можуть. Необхідні проекти для підтримки верхніх рівнів аналітики в охороні здоров'я, які включають: персоналізовану медицину і розпорядчу аналітику (рівень 8), ризики клінічного втручання і прогностичну аналітику (рівень 7), управління здоров'ям населення і сугестивну аналітику (рівень 6), які передують попередні рівні зменшення мінливості послуг, зовнішньої звітності, автоматизованої внутрішньої звітності, стандартного словника пацієнта і сховищ даних.

Проект Горизонт-2020 надає можливість розпочати такі проекти і об'єднати сумісні зусилля науковців різних країн на подальший прогрес удосконалення охорони здоров'я на базі ІКТ.

Література. **1.** SMART 2007/0059: Study on the Legal Framework for Interoperable eHealth in Europe. Final Report: www.ec.europa.eu/digital-agenda/en/news/legal-framework-interoperable-ehealth-europe/. **2.** Project e-SENS: www.esens.eu/home/. **3.** SMART 2011/0074: Inventory of cross-border eGovernment services & Analysis of existing and future needs and demand for cross-border eGovernment services: www.ec.europa.eu/digital-agenda/en/news/inventory-cross-border-egovernment-services-analysis-existing-and-future-needs-and-demand-cross **4.** eHealth European Interoperability Framework: www.ehealthnews.eu/images/stories/pdf/ehealth-interoper-ability-framework-study-assessmentframework.pdf. **5.** Integrating the Healthcare Enterprise (IHE): www.ihe.net/About_IHE/. **6.** Continua Health Alliance: www.healthcareitnews.com/directory/continua-health-alliance. **7.** EpSOS project: www.epsos.eu/home/about-epsos.html. **8.** Regulation on European Standardization: www.bsigroup.com/en-GB/about-bsi/uk-national-standards-body/European-standardization-regulation/. **9.** Support USers To Access INformationand Services: www.eu-patient.eu/globalassets/projects/sustains/sustains_user-requirement-recommendations.pdf. **10.** United4Health: www.united4health.eu/. **11.** SemanticHealthNet: www.semantichealthnet.eu/index.cfm/deliverables/. **12.** Connection Europe Facility (CEF): www.ec.europa.eu/inea/en/connecting-europe-facility. **13.** Google Health: www.google.com/intl/en_us/health/about/. **14.** Microsoft HealthVault: www.msdn.microsoft.com/en-us/healthvault/healthvault-introduction.aspx. **15.** IBM Health Integration Framework: www.www-01.ibm.com/software/bd/industry/healthcare/. **16.** Healthcare Analytics Adoption Model: A Framework and Roadmap (white paper): www.healthcatalyst.com/white-paper/healthcare-analytics-adoption-model/2/.

Романенко В.Д., Мілявський Ю.Л.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Нові методи управління імпульсними процесами у когнітивних картах

Когнітивна карта (КК) — це зважений орієнтований граф, що застосовується для опису взаємозвязків між різними елементами складної системи [1]. На сьогодні розроблено багато методів побудови, аналізу та моделювання КК, що широко застосовуються до систем різної природи — економічних, соціальних, фінансових, екологічних, політичних тощо. Автори доповіді розробили ряд методів управління імпульсними (перехідними) процесами у КК на основі застосування підходів з теорії автоматичного керування ([2,3] та ін.). За основу для усіх методів прийнято модель керованого імпульсного процесу

$$\Delta Y_i(k+1) = \sum_{j=1}^{n} a_{ij}\Delta Y_j(k) + b_i\Delta u_i(k), \tag{1}$$

де $\Delta Y_i(k) = Y_i(k) - Y_i(k-1), i = 1,2,...,n; a_{ij}$ – ваги ребер КК, що з'єднують j-у вершину з i-ою; n – кількість вершин КК, $\Delta u_i(k) = u_i(k) - u_i(k-1)$ – прирости зовнішніх управлінь.

У даній доповіді розглядається ще два нових методи управління складними динамічними системами, представленими у вигляді КК. Перший метод стосується різнотемпового представлення системи та керування нею із використанням частих і рідких управлінь. Другий метод полягає в управлінні системою не на основі варіювання наявних ресурсів, як усі попередні методи, а на основі варіювання вагових коефіцієнтів КК, тобто змінюючи взаємозвязки між елементами предметної області.

Управління імпульсними процесами КК з різнотемповою дискретизацією координат вершин. Розглянемо модель імпульсного процесу КК складної системи, у якій частина координат вершин вимірюється з періодом T_0, а частина – з періодом $h = mT_0, m > 1$. Тоді запишемо аналог (1) у вигляді системи:

$$\Delta Y_i\left[\left[\frac{k}{m}\right]h + (l+1)T_0\right] = \sum_{j=1}^{p} a_{ij}\Delta Y_j\left[\left[\frac{k}{m}\right]h + lT_0\right] + \\ + \sum_{j=p+1}^{n} a_{ij}\Delta\tilde{Y}_j\left[\left[\frac{k}{m}\right]h\right] + b_i\Delta u_i\left[\left[\frac{k}{m}\right]h + lT_0\right], \tag{2}$$

$$\Delta Y_s\left[\left(\left[\frac{k}{m}\right]+1\right)h\right] = \sum_{j=1}^{p} a_{sj}\Delta\tilde{Y}_j\left[\left[\frac{k}{m}\right]h + lT_0\right] + \\ + \sum_{j=p+1}^{n} a_{sj}\Delta Y_j\left[\left[\frac{k}{m}\right]h\right] + b_s\Delta u_s\left[\left[\frac{k}{m}\right]h\right], \tag{3}$$

де $i = 1,2,...,p; s = p+1,...,n; l = 0,1,...,m-1;$

$$\Delta\tilde{Y}_j\left[\left[\frac{k}{m}\right]h\right] = \begin{cases} Y_j\left[\left[\frac{k}{m}\right]h\right] - Y_j\left[\left(\left[\frac{k}{m}\right]-1\right)h\right], l = 0, \\ 0, l \neq 0 \end{cases} \tag{4}$$

при $j = p+1,...,n$.

Координати $\Delta\tilde{Y}_j\left[\left[\frac{k}{m}\right]h + lT_0\right]$ при $j = 1,...,p$ обчислюються за формулою [3]

$$\Delta\tilde{\overline{Y}}_1\left[\left[\frac{k}{m}\right]h + lT_0\right] = \overline{Y}_1\left[\left[\frac{k}{m}\right]h + (m-1)T_0\right] - \overline{Y}_1\left[\left[\frac{k}{m}\right]h - T_0\right]. \tag{5}$$

У векторно-матричній формі рівняння (2), (3) можна записати так:

$$\Delta\overline{Y}_1\left[\left[\frac{k}{m}\right]h + (l+1)T_0\right] = A_{11}\Delta\overline{Y}_1\left[\left[\frac{k}{m}\right]h + lT_0\right] + \\ + A_{12}\Delta\tilde{\overline{Y}}_2\left[\left[\frac{k}{m}\right]h\right] + B_{11}\Delta\overline{u}_1\left[\left[\frac{k}{m}\right]h + lT_0\right], \tag{6}$$

$$\Delta\overline{Y}_2\left[\left(\left[\frac{k}{m}\right]+1\right)h\right]=A_{21}\Delta\tilde{\overline{Y}}_1\left[\left[\frac{k}{m}\right]h+lT_0\right]+$$
$$+A_{22}\Delta\overline{Y}_2\left[\left[\frac{k}{m}\right]h\right]+B_{22}\Delta\overline{u}_2\left[\left[\frac{k}{m}\right]h\right]. \quad (7)$$

Застосувавши до системи (6), (7) квадратичні критерії оптимальності та мінімізувавши їх по векторах керування, отримуємо такі закони керування з різнотемповою дискретизацією:

$$\overline{u}_1\left[\left[\frac{k}{m}\right]h+lT_0\right]=\overline{u}_1\left[\left[\frac{k}{m}\right]h+(l-1)T_0\right]-$$
$$-(B_{11}^TB_{11}+R_1)^{-1}B_{11}\{(I+A_{11}-A_{11}q_1^{-1})\times$$
$$\times\overline{Y}_1\left[\left[\frac{k}{m}\right]h+lT_0\right]+A_{12}\Delta\tilde{\overline{Y}}_2\left[\left[\frac{k}{m}\right]h\right]-\overline{G}_1\Big\}, \quad (8)$$

$$\overline{u}_2\left[\left[\frac{k}{m}\right]h\right]=\overline{u}_2\left[\left(\left[\frac{k}{m}\right]-1\right)h\right]-$$
$$-(B_{22}^TB_{22}+R_2)^{-1}B_{22}\{(I+A_{22}-A_{22}q_2^{-1})\times$$
$$\times\overline{Y}_2\left[\left[\frac{k}{m}\right]h\right]+A_{21}\Delta\tilde{\overline{Y}}_1\left[\left[\frac{k}{m}\right]h+lT_0\right]-\overline{G}_2\Big\}. \quad (9)$$

Управління КК на основі варіювання вагових коефіцієнтів ребер. Варіювання вагового коефіцієнта можливо тоді, коли можна змінювати ступінь чутливості від впливу однієї вершини КК на іншу. Особа, що приймає рішення, може реалізувати цей принцип шляхом зміни коефіцієнтів передачі адміністративних, наукових, фінансових, політичних, освітніх, інформаційних взаємодій на координати складної системи, представлені вершинами КК. При управлінні імпульсним процесом КК шляхом варіювання вагових коефіцієнтів a_{ij} змінюється ступінь впливу на координату ΔY_i інших координат ΔY_j. При цьому величина керування u_i формується не за рахунок зміни ресурсів Y_j, що безпосередньо діють на вершину Y_i, а за рахунок зміни впливу інших координат Y_j на вершину Y_i.

Запишемо аналог (1) у повних значеннях координат вершин КК і зі змінними вагами замість звичайних керувань:

$$\overline{Y}(k+1)=(I+A-Aq^{-1})\overline{Y}(k)+B(k)\Delta\overline{a}(k)+\overline{\xi}(k), \quad (10)$$

де $\Delta\overline{a}(k)$ – вектор приростів вагових коефіцієнтів матриці A (по одному в кожному рядку матриці), q^{-1} – оператор зворотного зсуву на один період дискретизації, а $B(k)$ складається з тих координат вектора $\overline{Y}(k)$, які через ребра зі змінними коефіцієнтами діють на координати $Y_i(k), i=1,2,...,n$, $\overline{\xi}$ – вектор незалежних випадкових збурень з нульовим середнім.

Введемо квадратичний критерій оптимальності:

$$J(k+1)=E[[\overline{Y}(k+1)-\overline{G}]^T[\overline{Y}(k+1)-\overline{G}]+\Delta\overline{a}^T(k)R\Delta\overline{a}(k)], \quad (11)$$

де $\overline{G}$ – вектор задавальних діянь, R – задана матриця, E – оператор математичного сподівання.

Мінімізувавши (11) по вектору приростів вагових коефіцієнтів $\Delta\overline{a}(k)$, отримаємо закон керування:

$$\Delta\overline{a}(k)=-(B^T(k)B(k)+R)^{-1}B^T(k)[(I+A-Aq^{-1})\overline{Y}(k)-\overline{G}]. \quad (12)$$

Література. **1.** Roberts F. Discrete Mathematical Models with Applications to Social, Biological, and Environmental Problems. – Englewood Cliffs, Prentice-Hall, 1976. — 559 p. **2.** Романенко В.Д., Милявский Ю.Л. Обеспечение устойчивости импульсных процессов в когнитивных картах на основе моделей в пространстве состояний // Системні дослідження та інформаційні технології. — 2014. — № 1. — С. 26—42. **3.** М.З. Згуровский, В.Д. Романенко, Ю.Л. Милявский. Принципы и методы управления импульсными процесами в когнитивных картах сложных систем. Часть 2 // Проблемы управления и информатики. – 2016. (Прийнято до друку)

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.

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.

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.

***Bilushchak Y.I.*[1], *Chernukha O.Y.*[1], *Davydok A.Y.*[1], *Goncharuk V.Y.*[2]**
[1]*Centre of Mathematical Modelling of Y.S.Pidstryhach Institute of Applied Problems of Mechanics and Mathematics of NASU, Lviv, Ukraine;* [2]*National University "Lviv Polytechnic", Lviv, Ukraine*

Mathematical model for mass flow in a strip with randomly disposed sublayer of stochastical thickness

Diffusion flows play an important value on studying parameters of membranes and filters, diagnosing materials (diffusive-structural analysis), determining the structure of metal growth during hardening alloys etc.

On modeling transfer processes in multiphase stratified systems that we consider as randomly nonhomogeneous structures than the unknown coordinates of inclusion dispositions it can be no information on thicknesses of several inclusions. So the thickness of inclusions should also be considered as a random variable. This work is devoted to stating the model and solving the initial-boundary problem of diffusion formulated at once for the stochastically admixture diffusion flow [1] in a strip contained a randomly disposed sublayer of the stochastic thickness with triangular or uniform distributions [2] on the given interval $[h_{min}; h_{max}]$ (fig. 1), where $0 < h_{min} < h_{max} < z_0$.

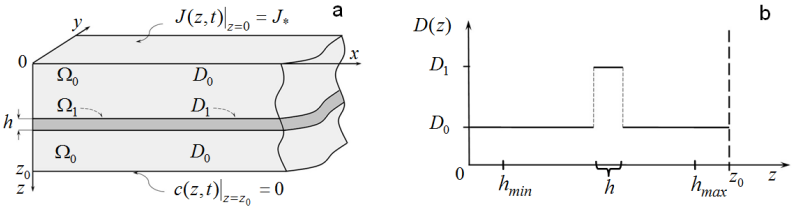


Figure 1. Possible realization of both a strip structure with sublayer of random thickness (fig.a) and corresponding random diffusion coefficient (fig.b)

Let the processes of diffusion of admixture substance run in a strip of thickness z_0. Coordinates of sublayer locations are unknown. The volume fraction of the basic phase v_0 (matrix) is much greater than the volume fraction of inclusion v_1. And diffusion coefficients are constant within the scope of each phase. In one-dimensional in spatial coordinate case the equation of admixture diffusion formulated for function of mass flow in a multiphase body is as follows [1]

$$\frac{\partial J(z,t)}{\partial t} = D(z)\frac{\partial^2 J(z,t)}{\partial z^2}. \tag{1}$$

Here $J(z,t)$ is the random mass flow; $D(z) = \{D_j, z \in \Omega_j\}$ is the stochastically coefficient of admixture diffusion, Ω_j is the domain of the j-th phase $(j = 0; 1)$.

Let in the initial moment of time the diffusion flow be absent in the body. The constant diffusion flow is supported on the boundary $z = 0$ and the admixture concentration $c(z,t)$ equals zero on the lower surface of the strip $z = z_0$. Namely, the following initial and boundary conditions are given

$$J(z,t)\big|_{t=0} = 0; \quad J(z,t)\big|_{z=0} = J_* \equiv const, \quad c(z,t)\big|_{z=z_0} = 0. \tag{2}$$

Herewith the diffusion flow on the lower boundary is certain function of time $F(t)$ that is to determine additionally

$$J(z,t)\big|_{z=z_0} = F(t). \tag{3}$$

We find the function $F(t)$ from the corresponding initial-boundary value problem for the concentration of migrating substance under the initial condition $c(z,t)\big|_{t=0} = c_* \equiv const$.

By using the theory of generalized functions the initial-boundary value problem (1)-(3) is reduced to the equivalent integro-differentual equation, which is solved by the method of successive iterations in the form of absolutely and uniformly convergent Neumann series [3]. We restrict ourselves by two terms of the series and average it as follows

$$\Big\langle \langle J(z,t)\rangle_{conf} \Big\rangle_h = J_0(z,t) + \left\langle \int_0^t \int_0^{z_0} G(z,z',t,t') \langle L_s(z')\rangle_{conf} J(z',t') dz' dt' \right\rangle_h .$$

Here $J_0(z,t)$ is the admixture flow in the homogeneous strip with matrix characteristics, $G(z,z',t,t')$ is deterministic Green function [3], $L_s(z) = (D_1 - D_0)\eta_1(z)\partial^2/\partial z^2$ is the perturbed diffusion operator, $\eta_1(z)$ is the random function of structure [3].

Procedures of averaging are carried out over the ensemble of phase configurations and over the random inclusion thickness. Calculated formulae for the averaged mass flow are obtained. For the uniform distribution of inclusions the calculated formula takes the form

$$\frac{1}{J_*}\left\langle \langle J(z,t)\rangle_{conf}\right\rangle_h = 1 - \frac{2}{z_0}\sum_{n=1}^{\infty} e^{-D_0\xi_n^2 t}\left(\frac{1}{\xi_n} + (-1)^n D_0\frac{c_*}{J_*}\right)\sin(\xi_n z) + \frac{2(D_1-D_0)}{D_0 z_0^3}\times$$

$$\times\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\frac{\xi_n}{y_k^2-\xi_n^2}A^1_{kn}\left(1+(-1)^n D_0\xi_n\frac{c_*}{J_*}\right)\left[e^{-D_0\xi_n^2 t} - e^{-D_0 y_k^2 t}\right]\sin(y_k z), \qquad (4)$$

where $\xi_n = \frac{\pi(2n-1)}{2z_0}$, $y_k = k\frac{\pi}{z_0}$; $b^{\pm}_{kn} = y_k \pm \xi_n$, $A^1_{kn} = \frac{1}{\Delta h}\left[\frac{\sin(b^-_{kn}h_{max}) - \sin(b^-_{kn}h_{min})}{(b^-_{kn})^3} - \right.$

$$\left. -\frac{\sin(b^+_{kn}h_{max}) - \sin(b^+_{kn}h_{min})}{(b^+_{kn})^3} - \frac{4y_k\xi_n\Delta h}{\left(y_k^2-\xi_n^2\right)^2} + (-1)^{k+n}y_k\frac{h^2_{max}-h^1_{min}}{y_k^2-\xi_n^2}\right],\ \Delta h = h_{max} - h_{min}.$$

On the basis of the obtained expressions the software modules are designed for dimensionless variables $\zeta = z/z_0$, $\tau = D_0 t/z_0^2$ and dependence of the averaged diffusion flow on input data is analyzed. Fig.2 illustrates graphs of the flows averaged over random inclusion thickness for different intervals $[h_{min}; h_{max}]$ of uniform distribution calculated by the relation (4) in fig.a and triangular one in fig.b. Here curves 1 and 2 are built for the intervals [0.005;0.205], [0.1;0.3] in the case of D_1/D_0=0.01, curves 3-6 for [0.005;0.205], [0.05;0.25], [0.1;0.3], [0.15;0.35] in the case D_1/D_0 =5. Curves a correspond the value c_*/J_*=0.1, curves b – c_*/J_*=0.4. Dash lines mark the flows in the homogeneous layer with the matrix characteristics.

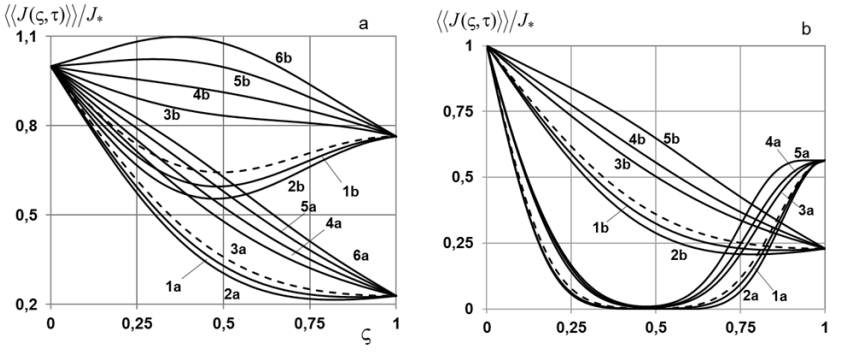


Figure 2. Averaged mass flows for uniform (fig.a) and triangular (fig.b) distribution of inclusion thickness on the interval $[h_{min}; h_{max}]$

Account of stochasticity of the inclusion thickness for both its uniform and triangular distributions affects substantially on values and/or behaviour of the averaged mass flow. So in the case of larger values of diffusion coefficient of admixture in the inclusion leads to increase of the diffusion flow in the whole body (curves 3-6 in fig.2). If the diffusion coefficient in the inclusion is smaller than in the matrix then increase of Δh under the constant value of h_{min} causes decreasing of the averaged flow (curves 1, 2 in fig.2). Growth of the initial concentration c_* can lead to formation of a local minimum of the function $\langle\langle J(z,t)\rangle\rangle/J_*$ in the middle of the strip (curves 1b, 2b in fig.2).

References. **1.** Chaplya, Y., Chernukha, O., Davydok, A., Pabyrivskyi, V. (2015) Diffusion equation for mass flux and problem of initial and boundary conditions. 17-th International conference SAIT 2015, Kyiv, Ukraine, June 22-25, 2015. Proceedings, 25-26. **2.** Korolyuk, V., Portenko, N., Skorokhod, A., Turbin, A. (1985) Reference Work on a Probability Theory and Mathematical Statistics, Nauka, Moscow, 640 p. **3.** Chaplya, Y., Chernukha, O., Davydok, A. (2015) Mathematical modeling of random diffusion flows in two-phase multilayered stochastically nonhomogeneous bodies. Task Quarterly, 19(3), 297-320.

Chenarani A., Druzhinin E.A.
Department of Information Technology Design, National aerospace university named after N.E. Zhukovsky "KhAI", Kharkiv, Ukraine

Analysis and selecting the appropriate technique for identifying risks of designing complex product systems

Today, the Complex Product Systems (CoPS) are a special class of products, which have critical importance for the economy and the industrial future of the modern world. Some of the most important characteristics of CoPS are significant economic and political value; consisting of many interconnected parts and components performing multiple and important functions; involving high degree of technological novelty and innovation; high level of coordination and collaboration during design and implementation; wide breadth of knowledge and skills and a high level of system integration. Examples include aircraft, high-speed trains, naval vessels, aircraft engines, etc.) In the process of developing CoPS, after defining the requirements by the customer, the first phase is the design phase through which the full set of required documents for creating the product is produced. This phase usually undergoes three stages of evolution - conceptual, preliminary and detailed design. Early design phases are critically important in complex product's life cycle because this is when the main decisions are made which in turn determine the product performance and development project's cost, time and quality. The complex systems design phase is highly iterative as it needs the satisfaction of different conditions and requirements and considering diverse technical and economical features. The inherent complexity and lack of data at this phase are the major source of risk and uncertainty that if not identified and managed properly can lead to increasing the time and cost or even failure of development project. Considering above-mentioned factors, it can be concluded that proper risk identification in designing phase of CoPS can lead to eliminate their undesirable consequences effectively, which in turn results in better achievement of assigned project goals. Our goal at present work is analyzing and comparing different risk identification techniques and selecting the most appropriate one for our field of interest. Our research object is the process of designing a new gas turbine engine for aviation purposes. Large number of risk identification techniques are available. For example, [1] mentioned 42 techniques. According to [2] based on their origin, these techniques can be classified in three categories. These categories with some examples are presented as follows [1–3]:

1. Historical Review (past experience): techniques which are based on experiences of either current project, or similar projects in past. Some techniques include: 1-1) *Checklists*: this technique is quick to use, and provides useful guides for areas in which there is a depth of experience, particularly for projects that are standard or routine in nature. It is possible to structure the risk identification checklist around a risk breakdown structure. 1-2) *Experiences of similar projects*: The risk assessments from previous projects may be an ideal guide but this technique has some limitations. 1-3) *Industry knowledge base*: It is a special case of a checklist and is used similarly. 1-4) *Post-Project Reviews/Lessons Learned/Historical Information*: by reviewing databases of risks of previous similar situations, information about risks are obtained. 1-5) *Questionnaire*: It can be presented as a special form of checklist but the result depends on the quality of the questions, and is limited to the topics covered by the questions. 1-6) *Risk Breakdown Structure (RBS)*: Is a hierarchical framework of potential sources of risks and ensures coverage of all types of risk. A comprehensive RBS is required, often tailored to the project.
2. Current Assessments (present): these techniques analyze current project characteristics. Some techniques include: 2-1) *Failure Modes and Effects Analysis (FMEA) or Fault Tree Analysis*: it is a systems engineering method for representing the logical combinations of the system states and possible causes that can contribute to system failure. This technique requires statistically accurate data on fault probabilities of events. 2-2) *Root-Cause Analysis*: Seeks to identify basic causes of risks that may be visible symptoms of more fundamental forces. Allows identification

of additional, dependent risks but can oversimplify and hide existence of other potential causes. 2-3) *Assumptions & Constraints Analysis*: Simple structured approach, which generates project specific risks, but Implicit/hidden assumptions or constraints are often missed. 2-4) *Cause and Effect (Ishikawa) Diagrams*: Presents in diagrammatic form the causes which contribute to a given outcome. The diagram identifies risks as those uncertain events, which could result in occurrence of the impact, but it, can quickly become over-complex for complex projects. 2-5) *SWOT analysis*: acronym stands for Strengths, Weaknesses, Opportunities and Threats and the method considers risk from internal and external environment. The idea is conducted in a structured manner and is particularly useful for identifying internally-generated risks arising from within the organization. 2-6) *Document review*: Careful review of current project documentation is done in order to identify risks. 2-7) *Interviews*: Risk identification interviews are conducted by an independent skilled interviewer, a risk breakdown structure, checklist or prompt list can be used as a framework for risk interviews. 2-8) *Influence Diagrams*: Is a diagrammatic representation of a project situation. It can identify risks when combined with sensitivity analysis or Monte Carlo simulation. 2-9) *System Dynamics*: The SD model represents entities and information flows within a project, and analysis of the model can reveal feed-back and feed-forward loops which lead to uncertainty or instability. 2-10) *WBS Review*: this technique provides for risks related to different levels of detail (from high-level to those related to individual work packages) but excludes external risks or those not specifically related to WBS elements.

3. Creativity Techniques (future): the stakeholders use their imagination and creativity to find new probable risks which might affect the project. The effectiveness depend on the ability of participants. Some of these techniques are: 3-1) *Delphi technique*: An information-gathering technique in which experts on the subject participate anonymously. The responses to the questionnaire are summarized and re-circulated to the experts for further comment. This process may be repeated few rounds to reduce bias in the data. It is limited to technical risks. 3-2) *Brainstorming*: Is an interactive, team-based approach where the risks are identified in sessions by project team and other experts, without making judgments about their importance. 3-3) *Crawford Slip method*: In the meeting, each participant writes the risks on sticky pads during 10 minutes. Each risk should be stated in a complete sentence, which states the cause of the risk, the risk, and its effect. 3-4) *Nominal Group Technique*: This is an adaptation of brainstorming, where participants share and discuss all issues before evaluation. Each technique has its strengths and weaknesses, and no single technique can be considered suitable, so it is recommended that identification process should use a combination of techniques, perhaps one from each category.

As in our case, main features and characteristics (designing the aviation engine)are:

- high degree of technical and technological complexity;
- high uncertainty and lack of knowledge about system in initial development phases;
- presence of good experience about similar projects;
- high required expertise level;
- relatively low level of economical and administrative features;

As the result, with considering these features in mind, for our research goals it is decided to use a combination of checklist, Delphi and interview techniques, which are assisted by Risk Breakdown Structure (RBS).

References. **1.** Raz T. A., Hillson D. (2005). Comparative Review of Risk Management Standards. *Risk Manag. An Int. J., 7(4), 53–66.* **2.** Project Management Institute Inc. (2009). *Practice standard for project risk management* **3.** Cooper D. F., Grey S., Raymond G., Walker P. (2005). *Project Risk Management Guidelines, Managing Risk in Large Projects and Complex Procurements*, John Wiley & Sons Ltd

***Cojocaru S.*[1] *Petic M.*[1,2] *Coman L.*[1] *Horos G.*[1]**
[1]*Institute of Mathematics and Computer Science, Chisinau, Republic of Moldova;* [2]*Alecu Russo Balti State University, Balti, Republic of Moldova*

Determining "controlled vocabulary" (lexicon of markers) of a text

We propose in this study to identify some data source, which can be used to determine "controlled vocabulary" (lexicon of markers) in order to develop an engine for extracting of articles referring to a given topic.

In our research 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 detecting, monitoring, modelling and mitigating social disasters caused by actions of different nature. We will identify some of the useful data sources, accumulate and process a limited number of texts in order to create a lexicon of relevant words (markers) for disaster domains. Also, our aim is to develop methods of automatic enriching of this lexicon. Thereafter the formed lexicon of markers will serve as a basis for development of tools to extract texts from the identified sources, to classify them and to provide sentiment analysis.

The development of resources and tools for sentiment and subjectivity analysis often starts with the construction of a lexicon, consisting of words and phrases annotated for sentiment or subjectivity [1].

The analysis of data in order to extract relevant structured information can be performed on different categories of data and for different purposes [2].

Public safety and emergency tools, in the context of the high popularity of social applications and the abundance of individual broadcasting messages issued inside them, must be reviewed and adapted to the new possibilities of possible social threat detection.

By its nature, social media data can contain a large portion of noisy data. Facebook and other social networks are the place where we can find different promotional campaigns having commercial character, but not only. For this data, we notice that blindly noise removing can worsen the problem stated in the big data paradox because the removal can also eliminate valuable information [3].

Therefore the process of information collecting from social networks has advantages and disadvantages. On the one hand, we can easily identify a chain of information that relates to a particular topic by following a certain hash-tag, on the other – it is difficult to identify automatically the marker itself, because the literal sense of hash tags might not refer directly to the disaster topics. For example, #jesuischarlie (I am Charlie) contains no words related directly to terrorism, but everyone knows that it is a slogan to express support of freedom of the press after the 7 January 2015 massacre at the French satirical weekly newspaper Charlie Hebdo.

We collected manually 616 news articles from 10 categories of disasters (railway, air and cars accidents, fire, earthquake, hurricanes, radioactive substances, attacks and diseases), consisting of 142840 words, from several Moldovan news sites.

On the other hand in order to automate the process of text collecting referring to disasters, we started with Natural kit for NodeJS. It consists of different natural language tools, including two classifiers, Naive Bayes and logistic regression. This mechanisms give us the possibility to use them in order to collect the articles from an established topic.

Then we made some experiments.

One of the experiment consisting of analysing the site www.noi.md from which we collected news articles referring to the topic of social disasters. The collected articles are referred to the topic of social disasters, and classified into three categories: catastrophes, epidemics and disasters.

NATO.NUKR.SFPP 984877 "Modeling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism".

For the beginning and approving the hypothesis of research we collected 26 articles containing 8450 words. The collection of texts is annotated at sentence and word levels, providing morpho-lexical information using UAIC Romanian Part of Speech Tagger (Simionescu, 2011).

The obtained collection of Romanian texts about disasters permits to extract and annotate a lexicon of markers concerning the topic on disasters. We annotated manualy those words and/or word expressions that express the meaning of social disasters. The most frequent words are "catastrofă" (catastrophe), "accident" (accident), "mori" (dead) and "incendiu" (fire). As we established 3 categories of articles of social disasters, each of them has its own lexical markers. Together with the words selected from social networks we obtain 117 markers. 78 words refer to the catastrophes, the epidemics are described by 24 words, and the fire - by 15 [3].

In another collection of texts consisting of 45 news articles, consisting of 11257 words, referring to railway, air and car accidents. The same procedure was applied: annotation at sentence and word levels, providing morpho-lexical information using UAIC Romanian Part of Speech Tagger (Simionescu, 2011). Based on the obtained results, we got 2659 unique lemmas. In addition, extracting only those which have the frequency more than one, and part of speech noun, verb, adjective and adverb, we obtained 1093 different lemmas. So, the procedure showed how to reduce the number of susceptible words for markers.

The obtained collection and markers were extracted manually. They serve as a basis for elaboration of an engine which inspects the given list of sites and extracts the articles containig markers. The corresponding software is under development. In its elaboration the date of publishing is taken into account (in order to avoid old posts), a special action is included to delete eventual noise (non-relevant publicity etc.)

Usually we operate with word-markers in the form of lemma. To extract relevant texts it is necessary to use instruments with rules of inflection and derivation. If in the case of inflection, that does not change the meaning, the problem is solved automatically [4], in the process of derivation things are more complicated.

The particularities of the derivational morphology mechanisms help in lexical resources extension without any semantic information. Moreover, there are similar processing mechanisms for different languages spoken in Europe, namely English, French, Spanish, Russian and Romanian. The approaches and mechanisms presented in the paper have been studied on the examples from Romanian, but, in most of the cases, they can be more or less applicable to other languages [3].

References. **1.** Banea, C., Mihalcea, R., and Wiebe J. (2011) Multilingual Sentiment and Subjectivity, In: Multilingual Natural Language Processing, editors Imed Zitouni and Dan Bikel, Prentice Hall. **2.** Bolea C.(2015), Vocabulary, Synonyms and Sentiments of Hazard-related Posts on Social Networks. An analysis for Romanian messages, Proc. IEEE Conf. SPED 2015, Bucharest, Oct. 2015 **3.** Petic M., Cojocaru Sv., Gîsca V. Exploring list of markers in unstructured text automatic processing, In:Proceedings of the 11th International Conference "Linguistic Resources and Tools for Processing the Romanian Language", 26-27 November 2015, pp. 125-136 . **4.** Petic, M. and Cojocaru, S. (2015) Vocabulary enriching for text analysis. In System analysis and information technology: 17-th International conference SAIT 2015, Kyiv, Ukraine, June 22–25, 2015. Proceedings. pp. 37-38.

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

Feature space building for nonlinear systems diagnostics based on wavelet transforms of integral models

The technology of intelligent diagnostic systems building used for improving the reliability of nonlinear dynamic objects fault diagnosis is presented. It is proposed to improve the method of model-based diagnostics based on nonparametric systems identification and diagnostic models parametrization. The effective algorithms of diagnostic models parametrization based on wavelet transforms are offered. The efficiency of the proposed diagnostic model is analyzed using a switched reluctance motor example.

Introduction. Nowadays the direction in technical diagnostics, based on diagnosed object models reconstruction is widely developed. Methods of nonparametric identifications used to build the object models based on "input/output" experimental data are applied when the faults change not only object features, but it's structure. In this work the model diagnostics based on nonparametric identification of objects using models in a form of Volterra series. In this case, the diagnostic procedure consists in obtaining of informational model of an object in the form of Volterra kernels based on "input/output" experimental data [1, 2] in time or frequency domain. The secondary diagnostic model is built using received Volterra kernels (dimension reduction of diagnostic feature space). The decision functions are built in diagnostic feature space.

The aim of this work is to improve the quality and reliability of diagnosis of the nonlinear dynamic object state using a model-based diagnostic nonparametric identification of objects in a form of Volterra kernels and wavelet transform procedure for secondary diagnostic model building. The switched reluctance motor is considered as real-nature dynamic object with significant nonlinearities. This direction is fast developing scientific and technical one. The electric motors are widely used in machine-tool construction, robotics, automated production lines, transportation, aerospace engineering and etc.

Feature space building. Using of the proposed diagnostics method leads to a need of Volterra kernels functions parametrization (the system of secondary feature space building). Selection of a feature space has a decisive influence on an accuracy of the diagnostic model and on the reliability of the object state diagnosis.

Features space diagnostic value is determined as a maximum of true recognition probability (TRP) criteria P , applied on a subset $\mathbf{X}'$ of a features space $\mathbf{X}$ ($\mathbf{X}' \subset \mathbf{X}$). It estimates by results of a classification problem solving on examination sample of data.

Classifier builds using decision rules based on discriminant functions constructed during learning process. TRP is calculated for each decision rule. Then it is searched the maximum value of TRP. Feature spaces having low quality of recognition are discarded. In summary, there are selected feature spaces for which the adding of any new feature doesn't increase diagnostic value of ones.

So, the most valuable spaces of two, three, etc. features are determined during the exhaustive search procedure for considered diagnostic features.

Common methods of Volterra kernels parametrization use samples of these functions of 1^{st} order $w_1(t)$ and diagonal section of these ones of 2^{nd} order $w_2(t,t)$ with a specified discreteness ($\mathbf{V}_k$). Also it's used moments of Volterra kernels of different orders r, $r{=}\overline{0,3}$, ($\mathbf{M}_k$)).

It's offered a method of Volterra kernels parametrization based on coefficients of direct continuous wavelet transform [3] of main sections of these functions of the 1^{st} order (feature set$\mathbf{W}_1$) and diagonal sections of these ones of the 2^{nd} order (feature set $\mathbf{W}_2$) along with common methods of function parametrization.

Experiment. Offered method of feature space building is analyzed on example of the switched reluctance motor diagnostics task. This electric motor is a real-nature dynamic object with significant nonlinearities The results of numerical experiments (feature space building) will allow to make a conclusion about efficiency of wavelet transforms Volterra kernels in problem of nonlinear

dynamic object diagnostics.

Preliminary evaluation shows that nonparametric dynamic models based on Volterra kernels sections of 2nd order give more diagnostic information about object than models of 1st order. So, further experiments are performed using models based on Volterra kernels sections of 2nd order.

It compares a diagnostic value of an offered feature space based on wavelet transforms and common feature spaces based on samples and moments of Volterra kernels sections. Discrete wavelet *coiflet* of different orders uses as a mother wavelet.

It's found that feature spaces based on wavelet transforms of Volterra kernels are most informative (fig. 1).

Further it's analyzed a fault tolerance of feature spaces diagnostic value. Errors in estimates of Volterra kernels are additive noise with zero expected value and variance depending of Volterra kernels extremum. Four training samples formed on base of noisy Volterra kernels 1st order and diagonal sections of Volterra kernels 2nd order. The noise level of training samples corresponds to 1%, 3%, 5%, 10% of Volterra kernels extremum.

The best results of fault tolerance analysis are shown in fig. 1. Feature spaces based on wavelet transforms of Volterra kernels are most fault tolerance among considered feature spaces. Feature space $\mathbf{W}_2$ save a fault tolerance both at low and at high noise levels unlike feature spaces $\mathbf{V}_2$, $\mathbf{M}_2$.

So, the method of feature spaces building based on wavelet transforms using *coiflet* of 1st and 4th orders provides the efficient dimension reduction of model with small losses of information.

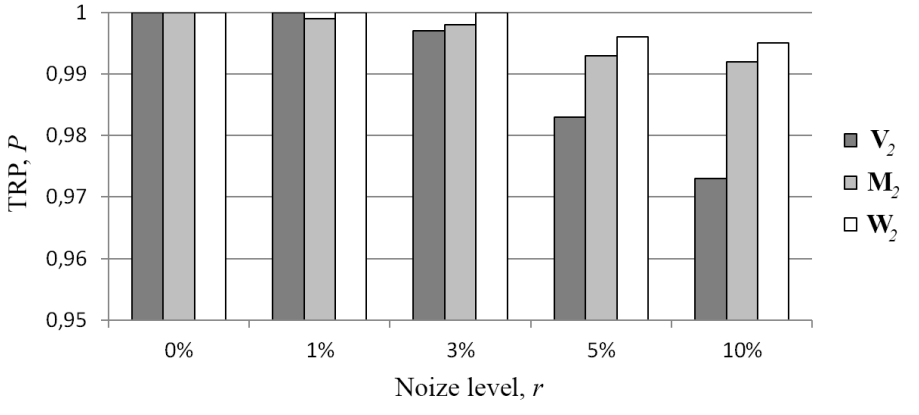


Figure 1. Diagnostic values for feature spaces $\mathbf{V}_2$, $\mathbf{M}_2$, $\mathbf{W}_2$ when the noise effects on Volterra kernels estimations.

Conclusion. In this work the technology of intelligent diagnostic systems building used for improving the reliability of nonlinear dynamic objects fault diagnosis is presented. It founds on the method of model diagnostics based on nonparametric identification of objects using models in a form of Volterra series. The effective algorithms of diagnostic models parametrization based on wavelet transforms are offered. The efficiency of the proposed diagnostic model is analyzed using a switched reluctance motor example.

References. **1.** Fomin, A., Pavlenko, V., Construction of diagnostic features space using Volterra kernels moments. Methods and Models in Automation and Robotics (MMAR), 2015 20th International Conference: 24-27 Aug. 2015 Miedzyzdroje, Poland. P: 1022 – 1027. **2.** Pavlenko, V. & Fomin, A., Methods For Black–Box Diagnostics Using Volterra Kernels. Proc. of the 2nd Int. Conf. on Inductive Modelling (ICIM'2008), Kyiv, Ukraine, Sept. 15-19, pp.104–107, 2008. **3.** Raghuveer M. Rao and Ajit S. Bopardikar. 1998. Wavelet Transforms: Introduction to Theory and Applications. Eddison-Wesley.

Gorban N.V., Paliichuk L.S.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Dynamics of solutions for non-autonomous reaction-diffusion equations with Carathéodory's nonlinearity

We consider non-autonomous reaction-diffusion system with Caratheodory's nonlinearity in a bounded domain $\Omega \subset \mathbb{R}^N$ with sufficiently smooth boundary $\partial\Omega$:

$$\begin{cases} u_t = a\Delta u - f(x,t,u), \quad x \in \Omega,\ t > 0, \\ u|_{\partial\Omega} = 0, \end{cases} \tag{1}$$

where $N, M = 1, 2, \ldots$, $u = u(x,t) = (u^{(1)}(x,t), ..., u^{(M)}(x,t))$ is unknown vector-function, a is real $M \times M$ matrix, $f = f(x,t,u) = (f^{(1)}(x,t,u), ..., f^{(M)}(x,t,u))$ is given interaction function.

We assume the following main conditions on parameters of Problem (1).

Assumption (A) There exists a positive constant d such that $\frac{1}{2}(a + a^*) \geqslant dI$, where I is identity $M \times M$ matrix, a^* is a transposed matrix for a.

Assumption (B) The interaction function $f = (f^{(1)}, ..., f^{(M)}) : \Omega \times \mathbb{R}_+ \times \mathbb{R}^M \to \mathbb{R}^M$ satisfies the standard Carathéodory's conditions, i.e. the mapping $(x,t,y) \to f(x,t,y)$ is continuous in $y \in \mathbb{R}^M$ for a.e. $(x,t) \in \Omega \times \mathbb{R}_+$, and it is measurable in $(x,t) \in \Omega \times \mathbb{R}_+$, for any $y \in \mathbb{R}^M$.

Assumption (C) There exist a translation uniform integrable in $L_1^{\mathrm{loc}}(\mathbb{R}_+; L_1(\Omega))$ function $c_1 : \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$ and a constant $c_2 > 0$ such that

$$\sum_{i=1}^{M} \left| f^{(i)}(x,t,y) \right|^{q_i} \leqslant c_1(x,t) + c_2 \sum_{i=1}^{M} \left| y^{(i)} \right|^{p_i},$$

for any $y = (y^{(1)}, ..., y^{(M)}) \in \mathbb{R}^M$ and a.e. $(x,t) \in \Omega \times \mathbb{R}_+$, where $p_i \geqslant 2$ and $q_i > 1$ are such that $\frac{1}{p_i} + \frac{1}{q_i} = 1$, for any $i = 1, 2, \ldots, M$.

Assumption (D) There exist a constant $\alpha > 0$ and a translation uniform integrable in $L_1^{\mathrm{loc}}(\mathbb{R}_+; L_1(\Omega))$ function $\beta : \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\sum_{i=1}^{M} f^{(i)}(x,t,y) y^{(i)} \geqslant \alpha \sum_{i=1}^{M} \left| y^{(i)} \right|^{p_i} - \beta(x,t),$$

for any $y = (y^{(1)}, ..., y^{(M)}) \in \mathbb{R}^M$ and a.e. $(x,t) \in \Omega \times \mathbb{R}_+$.

We investigate the long-time dynamics of all globally defined weak solutions as $t \to \infty$ under assumptions (A)–(D). We prove the existence of uniform global attractor for multivalued sewiflow generated by considered problem. Moreover, we obtain new topological properties of solutions, in particular flattening property, i.e. we prove that dynamics of all weak solutions of studied problem is finitedimensional within a small parameter.

As applications we may consider FitzHugh-Nagumo system (signal transmission across axons), complex Ginzburg-Landau equation (theory of superconductivity), Lotka-Volterra system with diffusion (ecology models), Belousov-Zhabotinsky system (chemical dynamics) and many others [1], whose dynamics are well studied in autonomous case [2,3], and in non-autonomous case, when all coefficients are uniformly continuous on time variable (see [2,4] and references therein).

References. **1.** Smoller, J.: Shock Waves and Reaction-Diffusion Equations. Springer-Verlag, New York (1983) **2.** Chepyzhov, V.V., Vishik, M.I.: Attractors for equations of mathematical physics. American Mathematical Society, Providence RI (2002) **3.** Temam, R.: Infinite-Dimensional Dynamical Systems in Mechanics and Physics. Appl. Math. Sci. Springer-Verlag, New York (1988) **4.** 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)

This work was partially supported by the Ukrainian State Fund for Fundamental Researches under grants GP/F61/017, and by the National Academy of Sciences of Ukraine under grant 2284/15.

Herasymiv I.S., Kovalenko A.E.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

System diagnosis models for distributed systems

Productive identification of system faults remains a common problem in today's large decentralized distributed systems. During last years, the complexity of distributed systems has grown in extremely large proportions, so it creates insistent need for development of automatic distributed system diagnosis for system faults.

System faults categorization is important part of system diagnosis process. There a lot of different approaches to divide system faults into types, but the common idea that each system unit can be in one of two states in concrete period of time: faulty or fault-free. There are three main approaches: based on duration, how it behaves after failure and occurrence of fault during diagnosis session [1].

Based on the Duration:

1. Transient fault: it appears in a network for short time.
2. Intermittent fault: we cannot predicts appearance and disappearance in the network.
3. Permanent fault: it remains until it removed and repaired by some external administrator.

Based on the Behavior:

1. Soft Fault: units can communicate with neighbors but with unexpected behaviors.
2. Hard fault: units cannot communicate with neighbors.

Based on the Occurrence:

1. Static fault: fault-free node cannot become faulty during diagnosis session.
2. Dynamic fault: fault-free node may become faulty during diagnosis session [1].

Chien, Metze and Preparata proposed the first approach to diagnose distributed systems in 1967 [2]. In their model, system is divided into nodes and these nodes should be atomic. Atomic node cannot be further decomposed for diagnosis. Each node diagnose a subset of nodes, and correct diagnosis can be reached for any type of diagnosis. Authors propose methods for instantaneous and sequential diagnosis procedures.

In our days bio-inspired methods become more popular in solving distributed systems diagnosis problems. Methods such as artificial immune systems, genetic algorithms and ant colonies have wide applications in distributed system diagnosis. In 2011 Falcon, Almeida Amiya Nayak proposed fault identification model with binary adaptive fireflies in parallel and distributed systems [3]. Authors describe a method, which use a swarm of artificial fireflies to find possible sets of faulty nodes under comparison and invalidation models. This model uses a binary encoding of the fireflies, an adaptive light absorption coefficient to handle impossible solutions. The empirical analysis confirms that the proposed algorithm exceed existing approaches in time and memory usages.

Conclusions. There are many different approaches to diagnose distributed systems. This research area is very important because unhandled faults in large distributed systems, such as mobile or other telecommunication networks, may cause a huge financial loses. In our days bio-inspired and artificial intelligence methods become more popular, they can highly improve existing methods of diagnosis. In future work it is important to create diagnosis model with ability to find system faults, which appear during diagnosis session, with acceptable time and memory characteristics.

References. **1.** Madhu Chouhan Dynamic Fault Diagnosis in Mobile Ad Hoc Networks Thesis for the degree of Master of Technology in Computer Science and Engineering / Madhu Chouhan, Department of Computer Science and Engineering National Institute of Technology Rourkela.- India Rourkela: Odisha, 2011.- 64 p. **2.** F. Preparata, G. Metze, and R. T. Chien. On the connection assignment problem of diagnosable systems. IEEE Transactions on Computers, 1967, pp. 848-854. **3.** R. Falcon, M. Almeida, A. Nayak, Fault identification with binary adaptive fireflies in parallel and distributed systems, in: 2011 IEEE Congress on Evolutionary Computation (CEC), 2011, pp. 1359-13.

Lulu K.M.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Analysis of the effect of news related to war on the exchange rate in Ukraine

In the period 2014 - 2015, news related to the war in the East of Ukraine dominated the Ukrainian mass media. In the same period, also the dynamics of the exchange rate were widely discussed. Intuitively, there should be a link between news related to the war and the exchange rate, especially between war related news and the black market rates as such news affects the confidence in the local currency. A decrease in confidence is likely to lead to a higher substitution between local and foreign currencies and to a higher demand for more stable assets, especially for US Dollars. This hypothesis about the effect of violence and the exchange market's fluctuations is the main topic of this paper.

Why news related factor is incorporated. A substantial volume of the papers demonstrated that describing exchange rate fluctuations only by fundamentals is unaffordable to some extent [1]. Meanwhile, changes can be better understood through the channel of socio – political determinants, for example politics, civil wars, sanctions etc., when time series specification accounts for such factors. Cornerstone of studies which concern about non – economic factors is that the most of them considered well – developed countries where financial instruments are well – known and widely used altogether with healthy market system. However, imperfections related to the financial system and relatively weaker capital markets in developing countries where a violence plays not minor role (case of Ukraine) can induce dependencies of the economic indicators on socio – political factors [2].

Black currency market as a mirror of the NBU's restrictions. Official foreign exchange market had been experiencing severe capital controls imposed by National Bank of Ukraine in 2014 - 2015 which caused substantial increase in the volume of the currency black market (Fig. 1). For that reason, black market exchange rate is subject of interest. At the same time, strong depreciation of the Russian ruble during 2014 and heavy tie to Russian economy might explain to some extent successive reduction in the value of currencies of the most CIS countries. Due to that reason sanctions imposed on Russia from Western countries and vice versa and oil prices are the subject of matter in explaining evolution of the foreign exchange rate in Ukraine.

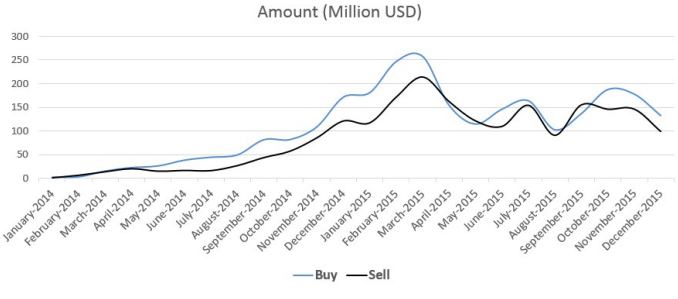


Figure 1. Volume of the currency black market

In order to investigate strength of the link between black market exchange rate and news related to war Vector Error Correction Auto Regression (VEC - VAR) specification is applied. Using VEC model has allowed us to obtain equilibrium equation which confirms strong dependency of the exchange rate generating process on the news related factor. Including socio - political factor has decreased error rate for prediction and increased predictive power of the model.

References. **1.** E. Charles, D. Kenneth Exchange Rates and Fundamentals http://www.ssc.wisc.edu/~cengel/PublishedPapers/EngelWestJPE.pdf **2.** F. Odhuno Does Civil War Violence Contribute to Exchange Rate Fluctuations? Evidence from Uganda http://nzae.org.nz/wp-content/uploads/2011/08/Does_Civil_War_Violence_Contribute_to_Exchange_Rate_Fluctuations-NZAE09.pdf

***Mieshkov O.Y.*[1,2] *Novikov O.O.*[1]**
[1]*Kherson national technical university, Kherson, Ukraine;* [2]*Kherson physical and technical lyceum by KNTU and DNU, Kherson, Ukraine*

Automated system for identification and human condition diagnostics based on its voice signal analysis

A two-steps system for human condition analysis based on human voice signal analysis is proposed. System extracts specified characteristics of human voice signal and on this basis identifies the person among other speakers. Then it compares characteristics of incoming signal with the database signals to define the presence of certain deviations in human organism condition. The developed system is able to perform analysis on the basis of both individual and group standards of human voice signals.

Introdution. Nowadays there are a great amount of methods and systems that allow to analyze human physical condition, i.e. to perform its diagnostics. For this purpose definition of a number of indicators that characterize human condition and their assessment with respect to the optimal level is performed. By the indicators deviation value the decision on the presence or absence of deviations in human organism condition in general is made.

One of the methods of human condition analysis that is actively used now is human voice signal analysis. It is known that human voice is definitely influenced with all systems of human organism. That is why different deviations that take place in human organism will be automatically displayed on its voice condition. Technologies for human state diagnostics are based on this aspect today [1,2].

Human voice signal is individual, i.e. for every certain person it differs. A number of technologies for personal identification in access systems are based on this fact.

The purpose of research. The purpose of the research is to develop a system that identify the person among other speakers by the incoming voice signal and perform the analysis of this voice signal to determine the presence of deviations in human physical condition.

Realization of a settled purpose caused the following tasks:

1. Defying of voice signal characteristics that can be used for identification and human condition analysis;
2. Development of two-steps human voice analysis system ('identification-diagnostics').

Human voice signal and its main characteristics. Human voice is a complex signal that is formed by the human vocal apparatus. This apparatus consists of different organs of respiratory, musculoskeletal and other systems of human organism. The whole voice formation process is coordinated by nervous system.

Human voice signal can be approximately divided into a number of quasi-periodic sounds. Every such sound has its own characteristics (amplitude, frequency, duration etc.). These characteristics are used in human voice analyzing systems in different ways. In most cases analysis is performed in frequency domain, where signal is transferred by means of the Fourier transform. The main characteristics for this domain are amplitude-frequency and phase-frequency spectrum, formant frequencies, dynamic spectrograms. While analyzing voice signal in time domain the main characteristics of human voice signal must be fundamental frequency and its dynamics, signal amplitude time distribution, duration of phonemes.

Voice signal characteristics extraction. In the first stage the developed system extract voice signal time domain characteristics which will be used for voice analysis further. We referred to these characteristics fundamental frequency and amplitude curve structure. Analysis is performed on recordings of vowel sound /A/ in Ukrainian pronunciation. Recordings are performed in *wav format (Sample rate – 22050 Hz, Bit depth – 16 bit, Channels – mono) for both male and female speakers. All operations, described below, are composed in script-files for applied mathematics program package SciLab 5.5.2.

Fundamental frequency definition is based on the cepstral transformation and searching the local maximum in the voice frequency range. In current research it was found that for men this range is about 40-220 Hz and for women – 90-300 Hz, that slightly differs with data in references [1].

Every voice fragment can be divided into a sequence of single quasiperiodic oscillations called frames. Dividing of voice signal into separate frames and their scaling algorithm have been described in research [3]. However, in specified research fragments with 20 consecutive frames have been extracted manually. Current research results shows that single vowel of human voice signal does not always contain 20 frames. Therefore, the automatic limitation by signal amplitude has been introduced to the developed system. As a result our system automatically choose only those frames of voice signal, which are suitable for analysis. The number of such frames is different for different speakers. After extraction each frame is scaled to clearly defined duration.

Described procedures are performed for all speakers, which are available in the database. Results – fundamental frequency of each frame and amplitude time distribution of scaled frames – are recorded in the special database file.

Persons' identification by the extracted characteristics. In the second stage the developed system identify the person by the incoming voice signal. An obligatory condition is the presence of persons' voice signal in the database. Incoming signal is pre-processed by the same algorithm the basic signals are. In this case, the average value of fundamental frequency and average amplitude time distribution structure are defined.

After that two-dimensional space of speakers is built on the basis of specified characteristics. Than the developed system calculates the distance from the incoming voice signal to each signal available in the base. The distance between signals is defined as Euclidean distance, taken with the weight coefficients:

$$d_{x,i} = \sqrt{\omega_F (F_{0,x} - F_{0,i})^2 + \omega_K K_{x,i}^2}, \tag{1}$$

where $d_{x,i}$ - distance between incoming signal (x) and i-th database signal;
$F_{0,x}, F_{0,i}$ – normalized fundamental frequency values of incoming and database signal;
ω_F, ω_K – characteristics' weight coefficients;
$K_{x,i}$ – normalized signal structure standard deviation coefficient, calculated by the formula:

$$K_{x,i} = \frac{\sqrt{\frac{1}{T}\sum_{j=1}^{n}(Y_{x,j} - Y_{i,j})^2 \Delta t}}{\sqrt{\frac{1}{T}\sum_{j=1}^{n} Y_{i,j}^2 \Delta t}} \tag{2}$$

where $Y_{x,j}, Y_{i,j}$ – amplitudes of j-th count of incoming and database signal;
T - frame period (duration);
Δt - time intervals between counts [4].

Database signal that will have a minimal distance to incoming signal is considered to belong to incoming speaker. If the distance exceeds the minimal threshold d_max, the system decides the speaker is out of database.

Experimental research. During the pilot study for base speakers set (60 female, 70 male speakers) weight coefficients values ($\omega_F = 0,67; \omega_K = 0,33$) and threshold value ($d_{max} = 0,2$)) were determined. Developed system was also tested using the test speakers set (30 female, 30 male speakers) and showed absolute accuracy in speakers' identification for the short-time period (one week). However, research on long-time period has shown that some basic speakers' signals ceases to be relevant, so it must be updated.

The analysis of human condition by voice signal. However, achieved results can testify that changes that have occurred in person's voice can be caused by certain deviations in the organism condition. In this case, these changes can determine the overall human condition. Taking into

account that voice signal is dynamic, all records of each speaker are saved in a personal database directory, so they can be used for the purpose of diagnostics of persons' condition.

In research [3] authors have developed a mathematical model of fundamental frequency and amplitude time distribution structure. This model takes into account individual features of each speaker and is able to simulate the speakers' voice signal dynamically:

$$F_0 = \alpha_F A^{a_F} M^{m_F} H^{h_F} \tag{3}$$

$$Y_i = \alpha_{Y,i} A^{a_{Y,i}} M^{m_{Y,i}} H^{h_{Y,i}}, \tag{4}$$

where F_0 - fundamental frequency;
Y_i - i-th count signal amplitude;
A - persons' age, complete years;
M - persons' weight, kg;
H - persons' height, cm;
α_F, a_F, m_F, h_F - frequency factor coefficients;
$\alpha_{Y,i}, a_{Y,i}, m_{Y,i}, h_{Y,i}$ - amplitude factor coefficients.

Amplitude and frequency factor coefficients are calculated every time when speakers database is updated using multiple regression method.

At this stage of the study the adequacy of using this model for every certain speaker as an individual voice signal standard is confirmed. Also in developed system there is an opportunity to construct human voice signal standard that will take into account features of different speakers' voice signals that are close to the needed signal. This type of voice signal standard is based on the analogical mathematical model and database clustering using KNN-graph technology.

Conclusions. In this study an automated system for human voice analysis is developed. This system is based on authors' algorithms and mathematical models of human voice signal, that take into account individual features of this type of signals. In the first stage, the developed system identifies the person among other database speakers by incoming voice signal. In the second stage, the system analyzes persons' physical condition using a number of voice signal characteristics based on persons' individual data or clustering the speakers' database. Experimental study of the developed system showed expediency of its use for solving the tasks of research.

References. **1.** Chorna L.B. Stochastic model of human voice signal for the task of human heart rhythm diagnostics, Ph.D. dissertation, Dept. Math. Model., TNTU after I.Pyliyi, Ternopil, Ukraine, 1999. **2.** Speech treatment for Parkinson's disease // Lorraine O Ramig, Cynthia Fox and Shimon Sapir // Expert Rev. Neurotherapeutics 8(2), 2008. – P. 299–311. **3.** Mieshkov O.Y, Novikov O.O. Development of Universal Program Complex for Human Condition Analysis, Based on the Analysis of Human Voice // Theoretical and Applied Aspects of Cybernetics. Proceedings of the 4th International Scientific Conference of Students and Young Scientists – Kyiv: Bukrek, 2014. – P. 294-305. **4.** Ponyzov A.G. Arrangement and formation methodology for acoustic test signals equivalent to pitchfork for hearing quality assessment, Ph.D. dissertation, Dept. Calc.Tech., TUSUR, Tomsk, Russia, 2012.

Pankratova N.D., Dozirtsiv V.O.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Application of methods for text analysis of the emotional tone to identify social disasters

The main source of information for instrumentality of prevention and mitigation of social disasters are huge volumes of unstructured data available on the global information networks, mass media, social networks, blogs, etc. Currently, development of information technology, computer networks and communications caused rapid rate of generation of new information, thus formed available huge amount of information. This information is very difficult to work because it largely is unstructured but contains markers for social processes that may be useful for monitoring and early identification of these processes, especially during social unrest and disasters, particularly caused by terrorism, because people and media creates a large amount of relevant information.

To solve problems related to the extraction and subsequent analysis of emotive phrases and phrases in the text, use methods "analysis of the emotional tone of the text" (Sentiment Analysis). Analysis of the emotional tone of the text is the part of the field of computational linguistics tasks and subtasks of receiving and processing information (Information Retrieval) [1]. Nowadays, not many researchers engaged this problem. Recently, the problem Sentiment Analysis pay attention mainly because people are not able to process this information. There are systems that analyze text emotion, but they all have some drawbacks. Therefore, the problem of determining the emotional elements of the text remains the important task of automated processing of textual information [2].

Large companies use Sentiment Analysis according to the significant proliferation of Internet, as business efficiency depends on the quality of using CRM-systems, which need to track references brand and analyze the text on the emotional satisfaction of the product. In a public policy the major using of analysis methods of emotional intelligence is the state tints, detection and prediction sentiments of the people and different social groups. In the public sector using of assessment emotional coloring is appropriate in the analysis of protocols and policy documents presented in the form of unstructured data, including Parliament, for checking for authenticity expressed views and reliability of opinions intentions of politicians in their election programs and in the public service.

Nowadays, there are several methods of analysis of emotional coloring the text of which is a method of two stages: analysis of the relationship between words, the analysis of emotive words. In the analysis of relationships between words exude categories and the weight of term in it. On the terms build set of terms, which among ourselves can strengthen or weaken the effect terms. Thus formed multilevel structure of sets of terms that, consequently, strengthen or weaken the effect of the terms of the previous set. Weakening or strengthening emotional coloring of term reflected in the modification of its weight. In the second stage of analysis need to calculate value of the degree of identity document to each category using the weight of each of the terms. Classification is usually used Bayesian model, the assumption of independence of statistical terms.

In solving the problem of detecting and mitigating on social calamities regarding disasters and attacks, implemented the project of NATO, proposed a new method based on Statistical Analysis Software (SAS), which significantly speeds up the processing of large data sets. In method occurs analyzing of emotional coloration in several stages: extraction and obtain documents, extract the theme of the document to the appropriate category and analysis on it, analysis of such categories.

References. **1.** Bo Pang and Lillian Lee. Foundations and Trends in Information Retrieval Vol. 2, No 1-2 - Pre-publication version – 2008 – 135 p.

2. О. В. Каніщева, А. В. Медведська, А. О. Панчул. Визначення типів емоційного мовного висловлювання у додатках автоматичного опрацювання текстів - Вісник Національного університету "Львівська політехніка". Інформаційні системи та мережі. - 2014. - № 805. - С. 289-297.

This research was sponsored in part by the NATO Science for Peace and Security Programme under grant G4877

Parkhomchuk D.M., Tymoshenko Yu.O.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Regularization of second order dynamic method

Introduction. Inverse problems with approximation of continual differential and integral equations by finite discrete algebraic systems or local linearization of non-linear systems reduced to the solution of system of linear equations

$$Ax = b, A \in R^{m\times n}, x \in R^n, b \in R^m, \tag{1}$$

where matrix A is usually singular, so such systems usually do not have mathematically rigorous solutions, and it is only possible to find approximate solutions [1]. In engineering practice matrix A and vector b are usually known only approximately with with some errors ς, ε i.e.:

$$||A - \tilde{A}|| < \varsigma, ||b - \tilde{b}|| < \varepsilon$$

Second order dynamic method. This paper analyzes the dynamic method for solving ill-posed problems, according to which, the solution of the following Cauchy problem is accepted as an approximate solution of the system 1 [2]:

$$\xi_1 * x'' + \xi_2 * x' + \tilde{A}^T\tilde{A}x = \tilde{A}^T b, x(0) = x_0, x'(0) = x_1 \tag{2}$$

In order to represent the solution of the problem 2 we will introduce the following notation:

x^+ - pseudo solution of the system 1

$k-$ dimension of $\tilde{A}^T\tilde{A}$ kernel,

$P(x)$ - projection of the vector x on the kernel of the matrix $\tilde{A}^T\tilde{A}$

a_i - i-th column of the matrix $\tilde{A}^T\tilde{A}$, $\delta(a_1, a_2, ..., a_n)$ - determinant of matrix, which consists of columns $a_1, a_2, ..., a_n$,

λ_i - i-th eigenvalue of the matrix $\tilde{A}^T\tilde{A}$,

$\triangle_i^j = \delta(a_1, ..., a_{i-1}, \tilde{A}^T b, a_{i+1}, ..., a_{j-1}, e_j, a_{j+1}, ..., a_n)$,

$\nabla_i^j = \delta(a_1, ..., a_{i-1}, \xi_2/\xi_1 * x_0 + x_1, a_{i+1}, ..., a_{j-1}, e_j, a_{j+1}, ..., a_n)$,

$L_i(t)-$ some linear combination of $e^{(-\lambda_j)}$, where $\lambda_j > 0$

so i - th coordinate of the solution 2 will be:

$$x_i(t) = \frac{\sum\limits_{1<j_1<j_2<\ldots<j_k<n} \triangle_i^{j_1,j_2,\ldots,j_k} + \sum\limits_{1<j_1<j_2<\ldots<j_{k-1}<n} \nabla_i^{j_1,j_2,\ldots,j_{k-1}}}{\prod_{j=1}^{n-k}\lambda_j} + L_i(t)$$

As a result of this solution analysis, two fundamental theorems were proved:

Theorem 1. If the matrix $\tilde{A}$ and vector $\tilde{b}$ in 1 defined precisely then

$$\lim_{t\to\infty} x(t) = x_n$$

where $x_n = x^+ + P(\xi_2/\xi_1 * x_0 + x_1)$.

Theorem 2. If matrix A and vector b in 1 defined approximately, i.e.

$$||A - \tilde{A}|| < \varsigma, ||b - \tilde{b}|| < \varepsilon$$

then we could choose some value $t^*(\varsigma, \varepsilon)$, that

$$\lim_{\substack{\varsigma\to 0\\ \varepsilon\to 0}} x(t^*) = x_n$$

where $x_n = x^+ + P(\xi_2/\xi_1 * x_0 + x_1)$.

Conclusions. This paper shows, that second order dynamic method can be used for solving ill-posed problems. According to theorems 1 and 2 we can see, that second order dynamic method produces the regularization family of operators with time t^* as a regularization parameter.

References. **1.** Фадеев Д.К., Фадеева В.Н. Вычислительные методы линейной алгебры -М. -Л.: Физматгиз, 1969. -654 с. **2.** Гутенмахер Л. И.,Тимошенко Ю. А,Тихончук С. Т. О динамическом методе решения некорректных задач: Докл. АН СССР. — 1977. — Т. 237. — № 4. — С. 776—778.

Savastiyanov V. V.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Extraction of expert sentiments from interview context

Most vital tasks in decision support systems are based on processing of experts' opinion. An example of such systems could be Foresight platform [1] [Pankratova and Savastiyanov, 2014], scenario analysis systems for mitigation of social disaster factors caused by catastrophes, terrorism, war or natural disaster [2] [Pankratova et al, 2014], and many others. Given approach could be used to extract of expert sentiments from interview context according to subjects, objects and systems which expert is talking about.

On the input the interview is captured in free form of dialog between interviewer and expert. All dialog is recorded as text. The algorithm is using sentiments vocabulary in first stage to estimate distances from extracted subjects, objects and systems to emotional words. There are several steps to estimate experts key objects in every emotional context:

1. Split text into paragraphs which begin from interviewer question and expert answer;
2. Filter out or join to next/previous all paragraphs with small length;
3. Set emotion vector for every paragraph if it could be matched with words from emotions vocabulary;
4. Extract objects and subjects from every paragraph;
5. If there is a lack of objects/subjects, the algorithm extracts objects/subjects from the question of interviewer;
6. Calculate distances between detected emotion words and extracted objects and subjects;
7. Calculate total score of emotional context [3] [Savastiyanov, 2015]
8. Output top5 objects/subjects for every emotion listed.

As an example lets look on questions were raised regarding the socio-political position of expert who is CEO of several large companies in energy market. After running the algorythm showed next results:

1. The biggest experts' concerns were caused by the following topics: effective leadership, their future, political future, etc.
2. Most experts' discontent call the following topics: mutual claims, coal contract, self minded persons, etc.
3. Most of all expert welcomes the following facilities: good education, capable people, the presence of their own business and so on.

The results that were produced in the analysis can be used in the methods of qualitative analysis as key factors, objects, subjects and events that are important to consider expert consideration under drawing up future scenarios.

The disadvantage of the approach is that experts is talking not only about objective information, but also about personal. The result is depending on interviewer skills to ask question about objective information.

References. **1.** Pankratova N., Savastiyanov V. Foresight Process Based on Text Analytics // International Journal «Information Content and Processing». — 2014. — 1, No 1, ITHEA. — P. 54–65. **2.** N. D. Pankratova, P. I. Bidyuk, Y. M. Selin, I. O. Savchenko, L. Y. Malafeeva, M. P. Makukha, V. V. Savastiyanov Foresight and Forecast for Prevention, Mitigation and Recovering after Social, Technical and Environmental Disasters // Improving Disaster Resilience and Mitigation - IT Means and Tools, // Springer. — 2014. — P. 119-134, ISBN 978-94-017-9136-6 **3.** Savastiyanov V.V. Discovering of potential positive and negative factors of social disaster using sentiment analysis // System analysis and information technology: 17-th International conference SAIT 2015, Kyiv, Ukraine, June 22–25, 2015. Proceedings. – ESC "IASA" NTUU "KPI", 2015. – 304 p.

This research was sponsored by the NATO Science for Peace and Security Programme under grant SPS G4877

Shcherbachenko L.V.
Faculty of Applied Mathematics, NTUU "KPI", Kyiv, Ukraine

Methods for improving work with Public Procurement systems

Every year Ukrainian government pays big amount of money (nearly half of budget) on public procurement. Therefore, this area of interaction between government and business was highly important and needed for innovations. Public procurement data became available for every user in May, 2011. After this many marketplaces were created for helping business find tenders, e.g. e-tender, dz.prom.ua, publicbid, newtend, smarttender.

Key stakeholders in public procurement. For public procurement area can be selected next stakeholders:

Definition 1. Customer – legal person or organization who is in demand of a product or service [1] (e.g. government);

Definition 2. Candidate – legal person or organization who provides a product or service [1] (e.g. business).

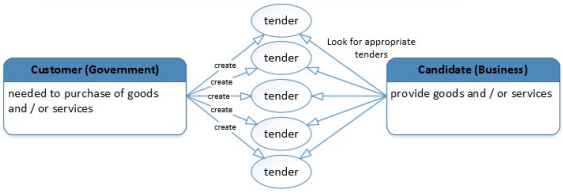


Figure 1. Schema of interaction between stakeholders

Studies results. Now process of finding appropriate tender was simplified only with search engine and process of finding required tender was not efficient and require time. Such process was shown on figure 1. Many electronic public procurement systems are presented on market, but studies have shown that they don't have next features:

1. Processing data of Ukrainian public procurement using data mining and machine learning techniques.
2. Expert system that:
   a) recommend tenders for candidates. This feature was based on candidates viewing history and on their finished tenders;
   b) resent candidates who accomplish similar tenders for customers.
3. Trust evaluation of candidates and customers based on their history of sales, comments, some official ratings. Provide statistical information about procurement participants.

Also, electronic public procurement system in Ukraine has problems with controlling changing states of the tenders (creation of interesting tender, who won tender or when was it paid for). I think that this problems can be solved by making a newsletter about tenders that contains some keywords or changing of tender state.

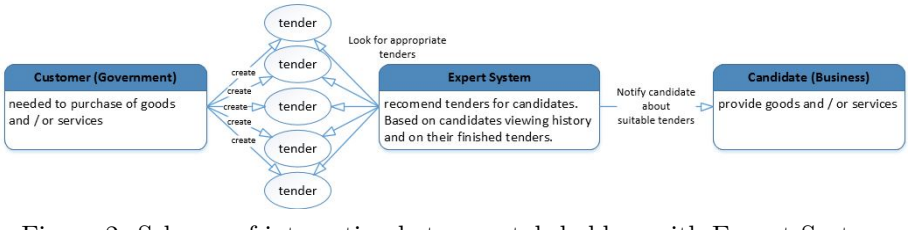


Figure 2. Schema of interaction between stakeholders with Expert System

Conclusions. Figure 2 show model of interaction between stakeholders with expert system. Such expert system make work with public procurement more effective for participates from Small and Medium Enterprise. Because this participates seek for big amount of small projects and features provided above can help them to find required tenders and evaluate them in shorter period of time.

References. **1.** Business Interoperability Interfaces for Public procurement in Europe - Use of profiles in the tendering process, 2012.

***Sobko V.O.*[1], *Bidyuk P.I.*[1], *Terentiev O.M.*[1], *Dzhygyrey I.M.*[2]**
[1]*Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine;* [2]*World Data Center for Geoinformatics and Sustainable Development, Kyiv, Ukraine*

Scenario-based forecasting and analysis of air pollution for Ukraine: outlook to 2020

Air pollution is among the most important environmental issues in modern world. Global climate change, human mortality and morbidity – is not the whole list of consequences of air pollution.

In order to prevent further climate change and reduce risks of human mortality, the global community should unite their efforts and adopt common agreement on environmental management. The actions should be directed towards reducing the emissions, and following the common climate policy [1]. That is why, according to [2], forecasting the emissions and analyzing their influencing factors are essential in order to adjust to policies which have a goal to mitigate the climate change, and overall environmental degradation.

In 2010, Ukraine started active work on the strategic goals of The United Nations Framework Convention on Climate Change (UNFCCC). In the context of UNFCCC, Ukraine promised to keep its emissions 20% below 1990 levels and declared a long-term goal of a 50% reduction by 2050 from 1990 levels. In addition, Ukraine has ratified the Kyoto protocol and is now on the negotiation process for the new global climate agreement [1].

Whereas many efforts in environmental monitoring and control were already done, the recent events happening in Ukraine were not encountered in any adequate study. The emphasizing characteristic of this paper is a new and fresh view on the emissions targeting, while encountering the recent events which take place in Ukraine.

Air pollution is analyzed in this paper in terms of its *"intensity"*, i.e. what amount of pollutants is emitted for a unit of economic output. This concept has a reasonable application, as it allows to compare countries, regions regardless of their size or population. The environmental impact is judged by emissions of air pollutants into the air, and carbon emissions separately, because in Ukraine data on these emissions is gathered separately. Economic output of regions of Ukraine will be judged by Gross Regional Product (GRP), and economic output of the whole country will be judged by Gross Domestic Product (GDP), which is a sum of regional GRPs. The choice of such variables is based on the fact that absolute values of emissions do not take into consideration the sizes of population in the regions [3].

The influence of conflict in eastern part of Ukraine on studied variables requires making scenario-based forecasts to 2020. *Scenario 1 – Integrity* assumes that all currently occupied regions, namely Crimean peninsula, Donetsk and Luhansk regions are get back to Ukraine. In other words, under this scenario the conflict in Ukraine is not taken into account. The forecasts are made as if Ukraine consists of its all 25 regions, and 2 administrative centers, as before 2014. *Scenario 2 – Occupation* assumes that apart from Crimean peninsula Ukraine loses Donetsk and Luhansk regions as well. The emissions of air pollutants and GRPs of these regions are not taken into account starting from 2014. It is assumed, that the development of other regions of Ukraine is not influenced by the conflict. For the forecasts to 2020 under *Scenario 3 – Protracted conflict* it is assumed, that the drop of emissions in air pollutants and GRP in Donetsk and Luhansk regions in 2014 is one-third compared to the levels of 2013.

The general forecasting workflow can be presented as the following procedure [4–6]: 1) data preparation (ETL); 2) forecasting methods selection; 3) model building; 4) results visualization; 5) results quality assessment.

The annual data on Gross Regional Product of Ukraine (2004 - 2013), carbon dioxide emissions (2008 - 2013), as well as annual data on emissions of air pollutants into the air from stationary and movable pollution sources (2000 - 2013) were taken from the official website of State Statistics Service of Ukraine.

Forecasting was done in SAS Enterprise Miner using the Time Series Exponential Smoothing

(TS Exponential Smoothing) node. Using this tool the best-fitting model is selected automatically based on the statistic criteria Mean Square Error (MSE).

The loss of Crimean peninsula (with the city of Sevastopol), as well as the Donetsk and Luhansk regions under Scenario 2 (Occupation) compared to Scenario 1 (Integrity) over the period from 2014 to 2020 will result in:

- 33% reduction in emissions of air pollutants;
- 39% reduction in carbon dioxide emissions;
- 19% reduction in GDP;
- 17% reduction in air pollution intensity;
- 24% reduction in carbon intensity;

The loss of Crimean peninsula (with the city of Sevastopol), as well as one-third of the Donetsk and Luhansk regions under Scenario 3 (Protracted conflict) compared to Scenario 1 (Integrity) over the period from 2014 to 2020 will result in:

- 12% reduction in emissions of air pollutants;
- 12% reduction in carbon dioxide emissions;
- 15% reduction in GDP;
- 3% increase in air pollution intensity;
- 3% increase in carbon intensity;

The tendency of air pollution intensity as well as carbon intensity is to decrease over time. This fact corresponds to the expectations from literature [2, 7] that GDP increases faster than emissions over time. However, the fact, that under Scenario 3 both intensities tend to be higher than under Scenario 1 emphasizes the high intensity levels of the Donetsk and Luhansk regions. Massive power plants, geographically concentrated mainly in the eastern part of Ukraine are characterized by high levels of energy inefficiency. In addition, as they rely on non-renewable fossil fuels, this results in increased emissions.

Validity of the results can be proved with the values of validation metrics. In forecasts of GRP exponential smoothing method shows strong forecasting performance with MAPE less than 9% on average, and R-squared = 0,96. In forecasts of emissions of air pollutants MAPE less than 6% on average, and R-squared = 0,7. In forecasts of carbon emissions MAPE is less than 7% on average, and R-squared = 0,6.

The overall results show that Ukraine easily fulfils its emission reduction targets by 2020. Also Ukraine has a great potential to further minimize intensity levels. As a main source of hazardous emissions in Ukraine comes from the industry sector, in order to improve intensity indicators, the implementation of improved energy policies is needed, as well as modernization of industrial equipment, so that the binding effect of economic growth and emissions growth will be avoided.

References. **1.** Dzhygyrey, I. 2014. Carbon intensity assessment of macroregions of Ukraine. Journal for Sustainable Development Data Analysis, 1, 40-48. **2.** Wu, L., Liu, S., Liu, D., Fang, Z. & Xu, H. 2015. Modelling and forecasting CO2 emissions in the BRICS (Brazil, Russia, India, China, and South Africa) countries using a novel multi-variable grey model. Elsevier Energy, 79, 489-495. **3.** Kokovskyy, L. 2014. Regional analysis of decoupling economic growth and pollutant emmissions into the air in Ukraine. Revista de Geografia e Ordenamento do Território (GOT), 143-160. **4.** Croker, S. T., Hornibook, S. L. & Ishikawa, T. 2007. Automated Real-Time Forecasting of Stream Conditions with SAS. http://analytics.ncsu.edu/sesug/2007/SA06.pdf. **5.** Waal, A. D. 2011. An Introduction to SAS®. Enterprise Miner and SAS®. Forecast Server. Waal, A. D. 2011. An Introduction to SAS®. Enterprise Miner and SAS®. Forecast Server. http://spears.okstate.edu/cepd/files/osu_analyticsdadewaal.pdf **6.** Rey, T., Kordon, A. & Wells, C. 2012. Applied Data Mining for Forecasting Using SAS®, SAS Institute Inc., Cary, NC, USA. **7.** Sun, J. W. 2006. An alternative viewpoint on the forecasting of energy-related CO2 emissions in the OECD countries. Elsevier Energy Policy, 34, 377-382.

***Teodorescu H.N.*[1], *Bolea C.*[2]**
[1]*Romanian Academy, Iasi Branch, Iasi, Romania;* [2]*Institute of Computer Science, Romanian Academy, Iasi Branch, Iasi, Romania*

Analysis of probabilities of specified words' occurrences in SN messages related to catastrophes

Recently, social networks (SNs) have become an important source of information for social sciences, linguistics, marketing, and even for technology and for combating the effects of catastrophic events [1–5]. This interest in SNs requires improvements in the methods and tools for data acquisition on SNs [3–5]. Making decisions and forecasts based on data collected from SNs asks for high precision in relevant data collection and specific data discovery. Redundant data reduce the efficiency of information extraction, while missing data or data misinterpretation may compromise the decision.

Various types of catastrophic events, such as earthquakes, floods, aviation crashes and terrorist acts may produce huge increases of the traffic on SNs [3,4]. Consequently, several research groups proposed methods for making use of this information [1–5]. While several authors have attempted the analysis of the vocabulary and the statistics of the words in messages related to such events for English and a few other major languages, we are not aware of any extended research on the statistics of the language (language model) specific to SNs related to catastrophic conditions, except a few very recent papers [6–8]. Moreover, such studies are almost completely missing for 'smaller' languages such as Romanian.

In previous papers, e.g., [7], we dealt with various issues related to the vocabulary of the Internet and social networks, including frequently used terms on specified SNs, in relation with natural disasters (earthquakes in [7], earthquakes and flood in other papers). Also, this research complements those results and the study of the jargon of Social Networks and in general on Internet; the latter study produced an *Internet Slang Annotated Dictionary* (ISAD) which is freely available at the address http://iit.academiaromana-is.ro/isad/.

The examination of the vocabularies of messages related to disasters on various SNs has several reasons. In the first place, it helps selecting the best keywords and search logical conditions; second, it provides clues on the main issues perceived by the public in relation with disasters, moreover on the sentiments and attitudes prevailing in these conditions. Third, from a linguistic point of view, it allows us study the specificity of the idiomatic expressions and jargon used on SNs in general, on individual SNs, and in relation with specified disaster conditions.

In the current investigation, we are specifically interested in the following:

- probabilities of the stopwords in specific categories of SN messages, i.e., in SNs messages that relate to a specific type of disaster, $p(w \in Stop | w \in m, m \to D)$, where $p(|)$ denotes conditional probability, w is a word, $Stop$ denotes the set of stopwords, m denotes a message, and $m \to D$ means that the message refers to a specific type of disaster D;
- probabilities of the most frequent words that are not stopwords,
- ratios of the probabilities of specified couples of synonyms, e.g., *earthquake* and *seism*, or *flood* and *inundation*, in messages on SNs;
- probabilities of co-occurrences of specific words, $p(w_1 \wedge w_2)$, for example $w_1 =$ *earthquake* and $w_2 =$ *collapsed*;
- probabilities of co-occurrences of pairs of words that comprise one geographic name, for example earthquake and *Vrancea*.

The present study refers only to earthquakes and to messages and posts (on social media) in Romanian related to seisms. Beyond the probabilities of synonyms, we are also interested in the analysis of the degree of synonymy and of the contexts that determine synonymy.

The authors acknowledge the partial support of the NATO Science for Peace and Security Programme under grant G4877.SPS grant G4877 (*Modeling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism.*) Authors' contribution: HNT conceived the research and methods, and wrote the abstract; CB gathered the data and performed the analysis with the help of HNT.

Synonyms and synonymic expressions are important because they are frequently used. Messages with synonyms must be interpreted accordingly, without loosing information and without erroneous interpretation. Their use increases the difficulty of the information extraction task. When pairs of words are synonyms only in specific contexts, the complexity of their role in the analysis increases significantly; this is because of the need of added context analysis.

Among others, the analysis of synonyms on SNs is aimed to determine the context where two words are truly synonyms, respectively the contexts making them different. In a previous research (HNT, unpublished, Apollonia Congress, Iaşi, March 4, 2016), the first author has shown that synonymy and related meanings obey context-dependent rules, providing as a typical example the Romanian words *cutremur* and *seism* which are synonyms only when the first is a noun. Although the word *seism* obeys S. Marcus' rule that the literary and scientific languages tend to converge, it seems that it enters the common language only in some restricted contexts. The statistics performed tends to support this hypothesis. In another direction of research, also of interest was the differential use of synonyms under normal circumstances and under surges (spikes) of SNs' activity (see [9]), due to disasters.

For example, after a destructive earthquake, finding where are located the large collapsed buildings and where are located victims is crucial. But such information can be expressed with a large number of almost synonym words. In fact, one can say about a building that "s-a dărâmat" (collapsed), "s-a prăbuit" (downfall), "s-a năruit" (fell), "s-a ruinat" (ruined). Similarly, casualties may be described with many words, such as casualty victim decease fatality injured person wounded person loss dead person (in Romanian, 'victime, decedati, raniti, pierderi, morti etc.).

Supplementary materials, including bags of words, can be found on the SRoL site described in [10] and at the above mentioned web address for ISAD.

References. **1.** Pankratova N.D., Modelling and mitigation of social disasters of various nature. Int. Conf. on System Analysis and Information Technology SAIT 2015, Kyiv, Ukraine, June 22–25, 2015, ISBN 978-966-2748-69-7, pp. 13-14. **2.** Pankratova N.D., Pyvovar V.V., Big Data Overview. Int. Conf. on System Analysis and Information Technology SAIT 2015, Kyiv, Ukraine, June 22–25, 2015, ISBN 978-966-2748-69-7, pp. 35. **3.** H.N.L. Teodorescu, On the Responses of Social Networks to External Events. Proc. ECAI 2015 7th IEEE Int. Conf. on Electronics, Computers and Artificial Intelligence, 25-27 June, 2015, Bucharest, Romania, pp. 13-18, DOI: 10.1109/ECAI.2015.7301138. **4.** H.N. Teodorescu, Using analytics and social media for monitoring and mitigation of social disasters. Procedia Engineering Vol. 107C, (2015) pp. 325-334, DOI 10.1016j.proeng.2015.06.088, Reference: PROENG17426. **5.** Pankratova N.D., Savastiyanov V., Foresight Process based on Text Analytics. Int. J. "Information Content and Processing", Vol. 1, Number 1, 2014, pp. 54-65. **6.** Petic M., Cojocaru S., Vocabulary enriching for text analysis. 17-th Int Conf. on System Analysis and Information Technology SAIT 2015, Kyiv, Ukraine, June 22–25, 2015, pp. 37-38. **7.** S.C. Bolea, Vocabulary, synonyms and sentiments of hazard-related posts on social networks. SPED 2015, 8th IEEE Int. Conf. on Speech Technology and Human-Computer Dialogue, Oct 14-17, 2015, Bucharest, Romania. **8.** M Pirnau Tool for monitoring Web sites for emergency-related posts and post analysis. SPED 2015, 8th IEEE Int. Conf. on Speech Technology and Human-Computer Dialogue, Oct 14-17, 2015, Bucharest, Romania. **9.** H.N.L. Teodorescu, Emergency-Related, Social Network Time Series – Description and Analysis. In: I. Rojas, H. Pomares (Eds.), Advances in Time Series Analysis. Series Contributions to Statistics, Springer, 2016. **10.** SM Feraru, HN Teodorescu, MD Zbancioc, SRoL-Web-based resources for languages and language technology e-Learning. International Journal of Computers Communications and Control 5 (3), pp. 301-313.

Titchiev I.I.
Institute of Mathematics and Computer Science, Chisinau, Republic of Moldova; Tiraspol State University, Chisinau, Republic of Moldova

Modelling and verification of evacuation system using Time Petri nets in case of disaster

The aim of these article is to addressing issues related to the extension of an evacuation system by adding time and verification of properties mostly related to evacuation time in case of disaster.

Introduction. The Petri net formalism is applied in both theoretical and practical aspects. The classical *Petri net* has been extended with the notion of time in order to surprise as close as possible modeled real systems. Petri nets are a powerful modeling technique because they can be use to model complex systems and to verify if the modeled systems satisfy some criteria.

The classical Petri net is a directed bipartite graph. The nodes are called places and transitions, and they are connected by arcs. Places are graphically represented by circles, transitions by rectangles. Places can store tokens, represented by black tokens. A distribution of tokens on the places of a net is called a marking, and corresponds to the "state" of the Petri net.

A transition of a net is enabled at a marking if all its input places (the places from which some edge leads to it) contain at least as many tokens as is the weight of arcs connecting them. An enabled transition can fire: it removes tokens from each of the input places, and adds tokens to each of its output places and exactly so much tokens how is the weight of arcs connecting them. This is called the *firing rule*.

Time in Petri nets. There are many ways to incorporate time in Petri net [4,5]. For many practical applications, the addition of time is a necessity. Time can be associated with places, transitions, or arcs.

For evacuation system [1] it is more natural to associate time to transitions: transitions represent activities (moving from the rooms to the doors or from the doors into the rooms) and activities take time. Evacuation time depends on walking ability of each inhabitant, for different categories of persons time is given in Table 1.

Table 1. Standard walking speed [2] of inhabitants

Age(year)	Walking speed (m/s)
0-3	Same speed like parents
4-6, 65-	0.8
7-12, 50-64	1
13-15, 40-49	1.2
16-18, 26-39	1.3
19-25	1.5

Also for a more accurate modeling we consider the density of people in rooms that influence the walking speed as shown in Table 2.

To determine the time of evacuation of people from rooms we will take into account the speed of one or more flows of people:

$$t_{ev} = \sum_{i=1}^{n} t_i \tag{1}$$

NUKR.SFPP 984877 "Modeling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism".

n-number of rooms, t_i - the travel time of the first flow, which depends on length of room and walking speed of people flow.

The transfer function is used, which take into account the time spent in the queue (moving time of human in the room) and the density and flow rate. The walking speed on the first flow depends on density of people (Table 2.) in rooms. The maximum density of inhabitants in every room is assumed to be 6.5 people/m^2.

$$D_1 = \frac{N_1}{(l_1 + w_1)} \tag{2}$$

N_1 - the number of people in the first room, l_1 - length of the room, w_1- width of the room.

The maximum number of persons who can enter in each room is:

$$N_{max} = l * w * 6.5 \tag{3}$$

l - length of the room, w- width of the room.

Walking speed on the next rooms depends on flow rate, which depends on flow rates of previous flows.

$$r_i = \frac{\sum r_{i-1} * w_{i-1}}{w_i} \tag{4}$$

r_i, r_{i-1} - flow rate in $i, i-1$ rooms , w_{i-1}, w_i - width of the $i-1, i$ rooms.

Table 2. Correlation of density and walking speed [2] of inhabitants

Density of inhabitants (people/m^2)	Walking speed (m/s)
0-1.5	Standard walking speed (Table 1)
1.6-3.8	0.5
3.9-6.5	0.2
6.5-	No approach

In correspondence with the specified plane in [1], it is necessary to translate evacuation plan in terms of Petri nets, considering the respecting of the conditions specified above. Using analysis modules of PIPE [3]: *DNAmaca interface* - calculates how long it takes for a system to go from the initial state to the target state,

GSPN analysis - calculates the average number of tokens on a place, the token probability density and the throughput of timed transitions by exploring the state space of the Petri net and determining the steady state solution of the model.

References. **1.** TITCHIEV I., *Petri nets to model disaster prevention*, Proceedings of the Workshop on Foundations of Informatics, August 24-29, Chisinau, Republic of Moldova, pp. 445-449, 2015. **2.** TAKASHI M., YOSHIFUMI N., YASUHIRO F.a and ATSUSHI M., *Development of Tsunami refuge PETRI-NET simulation system utilizable in independence disaster prevention organization*, The 14^th World Conference on Earthquake Engineering October 12-17, Beijing, China, 2008. **3.** AKHARWARE, N. PIPE2: Platform Independent Petri Net Editor.http://pipe2.sourceforge.net/documents/PIPE2-Report-20050814.pdf, 2005. **4.** BERTHOMIEU B. and DIAZ M.,*Modelling and verification of time dependent systems using Time Petri Nets*. IEEE Transactions on Software Engineering, 17(3):259-273, March 1991. **5.** AALST van der W.M.P., *Interval Timed Coloured Petri Nets and their Analysis*. In M. Ajmone Marsan, editor, Application and Theory of Petri Nets, v.691 of Lecture Notes in Computer Science, pages 453-472, Springer-Verlag, Berlin, 1993.

Zorin Yu.M.
Faculty of Applied Mathematics of NTUU "KPI", Kyiv, Ukraine

An Improved Cuckoo Search Algorithm

The paper presents a novel metaheuristic algorithm for multimodal functions optimization based on idea of the Cuckoo Search Algorithm. Due to introduction of two novelties into randomization component of a cuckoo movement, the algorithm performance has been substantially improved. Numerical experiments have shown that proposed algorithm substantially outperforms standard Cuckoo Search Algorithm in regard of the conversion speed and accuracy.

The Cuckoo Search [1] uses the following three idealized rules. 1) Each cuckoo lays one egg at a time, and dumps its egg in randomly chosen nest. 2) The best nests with high quality of eggs will carry over to the next generations. 3) The number of available host nests is fixed, and the egg laid by a cuckoo is discovered by the host bird with a probability $p_\alpha \in [0,1]$. In this case, the host bird can either throw the egg away or abandon the nest, and build a completely new nest. For simplicity, this last assumption can be approximated by the fraction p_α of the n nests are replaced by new nests (with new random solutions).

When generating new solution x_i^{t+1} for a cuckoo i the Lévy Flight is performed as

$$x_i^{t+1} = x_i^t + \alpha_t \oplus Levy(\lambda),$$

where $\oplus$ is dimension-wise sum.

In the standard Cuckoo Search all the cuckoos perform Levy Flight with $1 < \lambda \leqslant 3$ while movement step length α_t is fixed at value 1. So, all the cuckoo move with a fixed step length in all iterations. Due to the fixed step length, the algorithm might miss better local search capabilities and sometimes it traps into several local optimums. In order to increase the quality of cuckoo flock in searching the problem space efficiently and find the optimal solution we will use the Levy Flight with controlled step length [2].

The Levy Flight [3] is a random walk in which the step lengths have a stable probability distribution. When defined as a walk in a space of dimension greater than one, the steps made are in isotropic random directions. As it known, the step length of the Levy Flight depends on the value of λ . The less the value of λ , the bigger is the step length. Let's rank all the cuckoos in descending order accordingly to the value of their fitness function. To get a balance between explorations and exploitations capabilities of cuckoo flock in searching the problem space efficiently we assume that "worse" cuckoos should perform the Levy Flight with bigger swing (less λ value) to explore remote parts of the search space. On the other hand, "better" cuckoos 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 is bigger. It may be achieved if λ belongs to the interval $[\lambda\min; \lambda_{\max}]$ and the dependence of λ value on the cuckoo's rank is as follows

$$\lambda = \lambda_{\max} - i(\lambda_{\max} - \lambda_{\min})/(m-1),$$

where m is the flock size, i - rank of i-th cuckoo.

The other way to increase the thoroughness of the perspective areas exploitation is reduction of the α_t value form $\alpha_{\max}$ to $\alpha_{\min}$ during the iterations, so that it varies with the iteration counter t as follows

$$\alpha_t = \alpha_0 \delta^t,$$

where $0 < \delta < 1$ and is calculated as

$$\delta = \sqrt[k]{\frac{\alpha_{\min}}{\alpha_{\max}}},$$

where k is maximum number of iterations, $\alpha_0 = \alpha_{\max}$.

Parameter Settings. From the test phase simulations it has been found out that the best results can be obtained with the following values of algorithm's parameters: $\lambda_{\min} = 0.5$, $\lambda_{\max} = 1.9$, $\alpha_{\min} = 10^{-4}$ and $\alpha_{\max} = 0.5$. The flock size m has been fixed at 50 for all given problems.

Computational simulation results. Computational simulations have been aimed to show that the developed metaheuristic algorithm has substantial advantages in comparison with the standard Cuckoo Search algorithm. The capability of this algorithm has been investigated on well known test problems [4]. The proposed algorithm (MCS) and the standard Cuckoo Search Algorithm (SCS) have been run 50 times on each test problem.

Table 1. Comparative analysis of improved and standard Firefly algorithms.

	Problem size		
Algorithm	n = 30	n = 40	n = 50
	F*unction - Sphere*		
SCS	1.**94e-6 (3.13 e-7)**	9.**11e-6(5.37e-7)**	3.**32e-5(2.61e-5)**
MCS	1.**34e-12 (2.11e-14)**	2.**82e-12(3.1e-13)**	2.**25e-11 (5.1e-12)**
	F*unction - Rosenbrock*		
SCS	2 **.14e-4(2.12e-4)**	8.**67e-3(4.13e-3)**	1.**27e-2(4.51e-4)**
MCS	2.**13e-7(2.35e-8)**	9.**31e-7(4.34e-8)**	2.**37e-6(2.05e-10)**
	F*unction - Rastrigin*		
SCS	8.**08-1(4.06e-2)**	5.**41e-1(1.12e-1)**	0.**27(4.13e-1)**
MCS	1.**54-10(5.93e-10)**	3.**37e-9(3.22e-11)**	7.**83e-8(1.48e-10)**
	F*unction - Griewank*		
SCS	1.**07e-3(2.66e-3)**	4.**25e-3(5.57e-4)**	1.**71e-2(1.67e-3)**
MCS	3.**71e-5(1.27e-6)**	1.**15e-4(1.08e-5)**	1.**11e-3 (2.31e-4)**
	F*unction - Ackley*		
SCS	1.**62e-3(4.1e-4)**	6.**29e-2(4.01e-3)**	9.**49e-2(4.43e-3)**
MCS	7.**36 e-9(3.287e-11)**	7.**79e-8(1.12e-9)**	3.**31e-6(4.28e-8)**

Table 1 shows mean result obtained by the algorithm, and in the brackets the standard deviation for the results obtained upon 50 runs.

Conclusions. The statistical results obtained from the benchmark functions demonstrate the superiority of the proposed MCS algorithm to the standard Cuckoo Search Algorithm in regard of results accuracy and convergence speed. The values of the result standard deviation demonstrate more stable behavior of the proposed algorithm. These results were obtained due to the introduction of the two novelties – iteration depending value of λ and variable step length of the Levy Flight. It should be noted that the contribution into performance improvement of both novelties have been approximately the same. Further improvement of proposed algorithm might be obtained by the study of the influence of the parameter p_α on the algorithm performance.

References. **1.** X.-S. Yang. *Nature-Inspired Metaheuristic Algorithms*, 2nd Edition, Luniver Press, 2010. **2.** Yu.Zorin. "A metaheuristic algorithm for multimodal functions optimization," Международная научная конференция имени Т. А. Таран Интеллектуальный анализ информации ИАИ2015, Сборник трудов, с. 88-92, 2015. **3.** A. M. Reynolds, C. J. Rhodes. "The Lévy flight paradigm: random search patterns and mechanisms," Ecology, vol. 90, pp. 877-887, 2009. **4.** Virtual Library of Simulation Experiments: Test Functions and Datasets, Available: http://www.sfu.ca/~ssurjano/index.html

Андреев Н.В., Статкевич В.М.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

О задаче управления неоднородным процессом рождения и гибели

Рассмотрим неоднородный во времени процесс рождения и гибели [1–4]. Такой процесс является математической моделью ядерной реакции, эволюции биологических популяций и т.д. Это марковский процесс $\xi(t)$, $t > 0$ с состояниями $\{0, 1, 2, \ldots\}$, для которого из состояния $\xi(t) = n > 0$ возможны лишь переходы в состояния $n - 1$ и $n + 1$, причём

$$P\{\xi(t + \Delta t) = n + 1/\xi(t) = n\} = n\lambda(t)\Delta t + o(\Delta t),$$
$$P\{\xi(t + \Delta t) = n - 1/\xi(t) = n\} = n\mu(t)\Delta t + o(\Delta t),$$
$$P\{\xi(t + \Delta t) = n/\xi(t) = n\} = 1 - n(\lambda(t) + \mu(t))\Delta t + o(\Delta t).$$

Здесь $\lambda(t)$ и $\mu(t)$ – положительные непрерывные на $(0; +\infty)$ функции, называемые интенсивностями рождения и гибели соответственно. В начальный момент времени $\xi(0) = 1$. Рассматривается следующий случай: возможность самозарождения исключается (переход из состояния 0 в состояние 1 невозможен), а отношение $\frac{\mu(t)}{\lambda(t)} = c > 0$ постоянное для всех $t > 0$. Обозначим $P_x(t) = P\{\xi(t) = x\}$, $x = 0, 1, 2, \ldots$. В [3] доказаны следующие формулы:

$$P_0(t) = \left(1 + \left(\int_0^t \lambda(\tau)d\tau\right)^{-1}\right)^{-1} \quad (c = 1); \quad P_0(t) = \frac{c(1 - e^{\rho(t)})}{1 - ce^{\rho(t)}} \quad (c \neq 1); \tag{1}$$

$$P_x(t) = \frac{e^{\rho(t)}}{c^{x-1}} P_0^{x-1}(t)(1 - P_0(t))^2 \quad x = 1, 2, \ldots; \quad \rho(t) = (c - 1)\int_0^t \lambda(\tau)d\tau. \tag{2}$$

Из указанных соотношений следует, что $P_0(t) < c$, а функция $P_x(t)$ при фиксированном t непрерывна по c.

Пусть задано число $N > 0$. Рассмотрим задачу управления: минимизировать вероятность выхода процесса на $t \to +\infty$ из полосы $[1; N - 1]$, управляя параметром c, т.е.

$$P(c) = \lim_{t\to+\infty} P\{\xi(t) = 0\} +$$
$$+ \lim_{t\to+\infty} P\{\xi(t) \geqslant N\} \to \min_c.$$

Подставив в первое слагаемое критерия формулы (1), а во втором вычислив сумму бесконечной геометрической прогрессии согласно (2), получаем непрерывную по c функцию

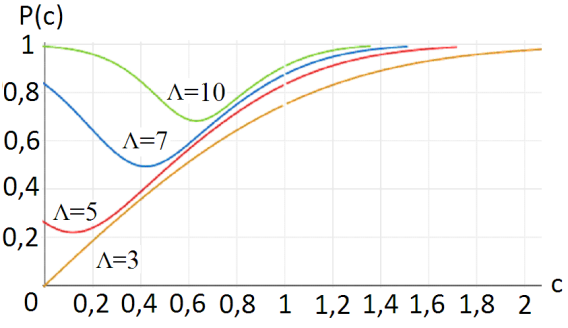


Рис. 1. Графики функции $P(c)$ при $N = 200$ и $\Lambda = 3, 5, 7, 10$

$$P(c) = \frac{\Lambda}{1 + \Lambda} + \frac{\Lambda^{N-1}}{(1 + \Lambda)^N} \quad (c = 1); \quad P(c) = \frac{c(1 - e^{(c-1)\Lambda})}{1 - ce^{(c-1)\Lambda}} + \frac{(1 - c)(1 - e^{(c-1)\Lambda})^{N-1}}{(1 - ce^{(c-1)\Lambda})^N} \quad (c \neq 1),$$

где $\Lambda = \int_0^{+\infty} \lambda(\tau)d\tau < +\infty$ – интегральная интенсивность рождения. Для случая $\Lambda = +\infty$ выполнено равенство $P(c) = 1$. Функция $P(c)$ при $\Lambda < +\infty$ имеет единственную точку минимума c^*, её график зависит от соотношения N и Λ (см. рис. 1). Для функций на рис. 1 численный подсчёт (например, методом золотого сечения) даёт соответствующие точки минимума $c^* = 0$, $c^* \approx 0,117263$, $c^* \approx 0,420581$, $c^* \approx 0,632043$.

Литература. **1.** Kendall D.G. On the Generalized "Birth-and-Death" Process, Ann. Math. Statistics, Vol. 19, No. 1 (1948), P. 1–15. **2.** Бартлетт М.С. Введение в теорию случайных процессов. – М.: ИЛ, 1958. – 384 с. **3.** Horalek V. Nonhomogeneous birth-death processes with constant ratio of rates, Aplikace matematiky, Vol. 11 (1966), Issue 4, P. 296–302. **4.** Андреев Н.В., Штатланд Э.С. О некоторых задачах управления неоднородными процессами рождения и гибели//Труды семинара "Теория оптимальных решений". – Вып. 4. – Киев, 1968. – С. 42–46.

Андреєва М.О., Левко Л.Р.
Факультет прикладної математики НТУУ "КПІ", Київ, Україна

Метод для побудови генератора псевдовипадкових рівновагових послідовностей заданої довжини і заданої ваги

В останні десятиліття все більша кількість систем управління різними об'єктами реалізуються на основі мікропроцесорів. Однак, для деяких систем, особливо, так званих, систем критичного застосування, тобто таких, неправильна робота яких загрожує життю або здоров'ю людей, призводить до значних матеріальних втрат, завдає серйозної шкоди екології і т. д., ставляться підвищені вимоги до надійності.

Розрахунок надійності відмовостійких багатопроцесорних систем часто є нетривіальною задачею. Існує багато методів розрахунку імовірності безвідмовної роботи [1]. Одним з таких методів є метод статистичних експериментів з моделями, які відображають реакцію системи на появу відмов певних її елементів.

Для проведення статистичних експериментів на вхід системи, робота якої моделюється подається двійковий вектор довжини n, який містить задану кількість елементів, що відмовили k. Для формування таких векторів використовуються генератори псевдовипадкових послідовностей з можливістю генерування послідовностей сталої ваги і такого, що не повторює свої стани. Кількість станів такого генератора дорівнює кількості комбінацій C_n^k.

Запропонований генератор складається з трьох частин: генератора М-послідовності, схеми зміщення та відсіювання, схема перетворення отриманих чисел у вихідний вектор [2]. Вихід кожної частини генератора використовується як вхід для наступної.

Генератор М-послідовності – регістр зсуву з лінійним зворотнім зв'язком, який генерує двійкову послідовність, значення якої лежать в множині $\{1, ..., 2^P - 1\}$, де p – розрядність регістру. Значення p обирається так, щоб $2^P - 1$ було мінімально більшим (або рівним) за кількість станів генератора, тобто: $2^P - 1 \geqslant C_n^k$.

В основі схеми перетворення отриманих чисел у вихідний вектор лежить використання комбінаторної системи чисел, де кожному числу із множини чисел $\{0, ..., C_n^k\}$ відповідає один і тільки один вектор з множини векторів, що мають довжину n і вагу k [3,4]. В $\binom{n}{k}$ комбінаторній системі чисел ціле число $N < \binom{n}{k}$ представляється як $N = c_k c_{k-1} \dots c_1$, де:

$$N = \binom{c_k}{k} + \binom{c_{k-1}}{k-1} + + \binom{c_1}{1}, \tag{1}$$

таких, що $c_k > c_{k-1} > ... > c_1 \geqslant 0$

Результуючий вектор X матиме вигляд:

$$X = 2^{C_k} + 2^{C_{k-1}} + + 2^{C_1}, \tag{2}$$

де C_k - номер розряду двійкового вектору в якому буде встановлено 1.

Оскільки потужності виходу генератора М-послідовності і схеми перетворення отриманих чисел у вихідний вектор можуть відрізнятися, то нами запропоновано відсіювати значення М-послідовності, які не належать множині $\{0, ..., C_n^k\}$. При відсіюванні числа вихідне значення генератора в поточному такті вважається недійсним або пропускається.

Наступним етапом у розробці даного методу генератора буде конвеєризація схеми перетворення отриманих чисел у вихідний вектор, оскільки вона вносить найбільші затримки і обмежує максимальну частоту роботи генератора.

Література. **1.** Мораведж Сейед Милад. Методы и средства оценки технической безопасности и расчёта параметров эффективности тестирования и реконфигурации многопроцессорных систем управления : дисс. ... канд. техн. наук. - 05.13.05 **2.** Jon T. Butlery and Tsutomu Sasaoz (2011). High-Speed Constant Weight Codeword Generators. Monterey, CA, USA: Naval Postgraduate School. pp. 3-8. **3.** Knuth, D. E. (2005). "Generating All Combinations and Partitions", The Art of Computer Programming, 4, Fascicle 3, Addison-Wesley, pp. 5–6 **4.** Ed. E. F. Beckenbach (1964). "Applied Combinatorial Mathematics", pp. 27-30.

Андрущенко В.Б., Ланде Д.В.
Інститут проблем реєстрації інформації, Київ, Україна

Нові наукометричні сервіси на базі Google Scholar Citations

Розвиток наукового знання, поширення наукової інформації сьогодні вимагає додаткових засобів для оцінки, структуризації та візуалізації.

Наукометричні показники – є одним із вагомих інструментарієв, що використовуються для оцінки стану, перспектив розвитку та визначення пріоритетних напрямків наукових досліджень, колективу, колаборації, стану науки в окремій країні та у світі загалом.

На сьогоднішній день в багатьох країнах світу наукометричні показники стають визначальними параметрами не тільки для оцінки діяльності наукової організації, окремого вченого, а і вирішальними при наданні грантової підтримки для проведення наукового дослідження. Адже публікаційна активність вченого чи колективу є однією з вадливиг вимог при поданні запитів на гранти для підтримки наукового досілдження.

Нажаль, серед інсуючих ресурсів, що надають інформацію про наукометричні показники того чи іншого вченого, інститутції, публікаційну активність є передплачуваними, що обмежує доступ до них. І сприяння у доступі до таких ресурсів у будь-якому вигляді є відсутнім.

В той же час в мережі Інтернет присутні ресурси відкритого доступу, такі як Google Scholar Citations, ArXiv та вітчизняний ресурс, що базується на даних сиситеми Google Scholar - «Бібілометрика української науки», і відображає рейтинги виключно українських вчених в різних галузях знань. Користування вищезазначеними ресурсами передбачає вилучення інформації про окремих вчених, їх публкаційну активність і деякі наукометричні показники, зокрема індекс Гірша. Але на сьогодні всі можливості зазначених ресурсів не є реалізованими повноцінно.

Авторами було досліджено та реалізовано алгоритми побудови мережі співавторів за даними сервісу Google Scholar Citations [1]. В той же час вивчення стану проблеми показало, що є необхідність реалізації більш розширених алгоритмів, що передбачають крім побудови мережі співавторів ще реалізацію побудови моделей предметних областей[2] та формування додаткових структурованих масивів даних.

Запропонований алгоритм передбачає й вилучення додаткової інформації із подальшою її структуризацію. Алгоритм був реалізований на прикладі використання тегу “text mining” сервісу Google Scholar Citations.

1. При введенні даного тегу у рядку пошуку ресурсу, отримуємо перелік профілів вчених, відсортований за кількістю цитувань, в сфері зацікавленості яких “text mining” було зазначено, як один із інтересів.

2. До розгляду приймалися автори, чия загальна кількість цитувань є не менше 500.

3. За рахунок покрокового аналізу ключових слів авторів із переліку було сформовано бази понять, що відносяться до різних предметних областей (в даному випадку - 4), що визначає точки перетину тієї чи іншої предметної області в рамках міждісциплінарності.

4. Відповідно до зазначених вченими афіліацій було створено перелік організацій, напрямки яких передбачають дослідження проблем глибинного аналізу тексту.

5. Шляхом почергового вивчення профілів вчених, було відібрано та складено переліки: - найбільш цитованих публікацій за тегом; - публікацій за останні 3 роки; - видань, в яких було розміщено результати досліджень за тегом.

Наразі алгоритм вимагає додаткової апробації із використанням інших тегів, пошуку закономірностей, переваг та недоліків його використання, розширення масивів інформації шляхом звернення до допоміжних ресурсів. Також є необхідність реалізації подібних задач за даними інших відкритих наукометричних та реферативних ресурсів.

Література. 1. Д.В. Ландэ, И.В. Балагура, В.Б. Андрущенко, Постороение сетей соавторства по данным сервиса Google Scholar Citations //Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2016): материалы VI Междунар. научн.-техн. конф. (Минск, 18-20 февраля 2016 года) - МинскЖ БГУИР, 2016. - 596с.

2. Д.В. Ландэ, Построение модели предметной области путем зондирования сервиса Google Scholar Citations // Онтология проектирования, 2015. - N 3(17). - С. 328-335.

***Барановский А.Н.*[1], *Кириченко Л.О.*[2], *Кобицкая Ю.А.*[2]**
[1]*НТУУ "КПИ", Киев, Украина;* [2]*ХНУРЭ, Харьков, Украина*

Рекуррентный анализ фрактальных временных рядов

Большинство динамических систем являются сложными системами, что подразумевает многозвенную структуру с нелинейной обратной связью. К ним относятся процессы присущие человеческому организму и природе: информационные, физические, технические, социальные и т.д. Многие публикации свидетельствуют о том, что большинство природных процессов обладают фрактальными свойствами и долгосрочной зависимостью. Выбор и построение математических моделей, которые соответствуют естественным процессам, требуют новых инструментов исследования, основанных как на методах нелинейной динамики, так и методах статистического анализа. В то же время открытым остается вопрос взаимосвязи меры фрактальности (показателя Херста) и нелинейно-информационных характеристик современных методов анализа, в частности метода анализа рекуррентных диаграмм (РД).

Метод РД является одним из методов построения математических моделей естественных процессов, основанный на использовании теоремы Таккенса-Паккарда и восстановлении псевдофазовой траектории [1]. Рекуррентная диаграмма является проекцией m-мерного псевдофазового пространства на плоскость. Пусть точка x_i соответствует точке фазовой траектории $x(t)$, описывающей динамическую систему в m-мерном пространстве в момент времени $t = i$, для $i = 1, \ldots, N$, тогда рекуррентная диаграмма RP есть массив точек, где ненулевой элемент с координатами (i, j) соответствует случаю, когда расстояние между x_i и x_j меньше ε:

$$RP_{j,j} = \Theta(\epsilon - ||x_i - x_j||), \qquad x_i, x_j \in R^m, \qquad i, j = 1, ..., N$$

где ε – размер окрестности точки x_i, $||x_i - x_j||$ – расстояние между точками, $\Theta(.)$ – функция Хэвисайда.

Анализ топологии диаграммы позволяет классифицировать наблюдаемые процессы: определять однородные процессы с независимыми случайными значениями; процессы с медленно меняющими параметрами; периодические или осциллирующие процессы, соответствующие нелинейным системам и т.д. Численный анализ РД позволяет вычислять меры сложности структур РД, такие как мера рекуррентности, мера детерминизма, мера энтропии и др., которые в свою очередь характеризуют определенные параметры поведения исходного процесса.

В работе были получены зависимости качественных и количественных характеристик рекуррентных структур от степени самоподобия и мультифрактальности. На рисунке 1 представлены рекуррентные диаграммы фрактального броуновского движения для разных показателей Херста и взаимосвязь меры детерминизма и показателя Херста. Основываясь на результатах проведенных исследований предложены рекомендации выбора параметров рекуррентных диаграмм в зависимости от характеристик исследуемых самоподобных временных рядов.

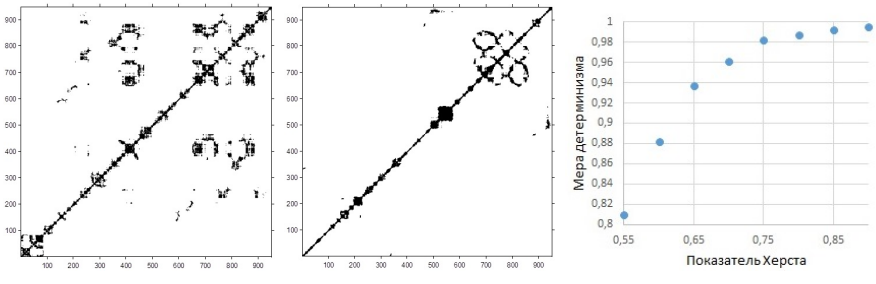


Рис. 1. Рекуррентные диаграммы для H=0,6 и H=0,8; зависимость меры детерминизма от H

Литература. 1. Marwan N. Recurrence-plots-based measures of complexity and application to heart-rate-variability data / N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths // Physical Review, 2002. – E 66 (2).

Бардадим І.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Алгоритм оцінювання параметрів регресійної з випадковим шумом там можливими сезонними коливаннями

Серед великого різноманіття математичних моделей статистики випадкових процесів одне з провідних місць займає широко використовувана модель "сигнал плюс шум". Математичні моделі такого типу мають велику сферу застосувань в різних галузях природничих та соціальних наук. Для сучасного етапу розвитку цієї теорії властиве ускладнення математичних моделей випадкового шуму, що маскує функцію регресії (корисний сигнал) для процесів, що мають сезонний характер. У роботі в якості об'єкта досліджень пропонуються, так звані, М-оцінки, які узагальнюють класичні оцінки. Вони викликають великий науковий інтерес протягом останніх десятиріч завдяки своїй стійкості (робастності) до аномальних спостережень породжених шумами. [1].

Концепція алгориитму робастної апроксимації. Головним недоліком методу найменших квадратів є висока залежність до початкових даних та помилок. Якщо серед вхідних даних координати вхідних даних мають сильне відхилення від інших – результат апроксимації може вийти досить неточним.

Щоб уникнути подібного ефекту, вводиться наступні функції втрат вигляду:

1. Модифікована оцінка найманших квадратівequation:
$$\rho(x,c) = \frac{x^2}{x^2+c^2}, \tag{1}$$
де x - відстань від будь-якої точнки до прямої,
c - масштабуюча константа.
2. Оцінка Ендрюсаequation:
$$\rho(x) = \begin{cases} 1-\cos(x), |x| \leqslant \pi \\ 2, |x| > \pi, \end{cases} \tag{2}$$
де x - відстань від будь-якої точнки до прямої.

Для мінімізації функції помилки складається наступна система нелінійних рівнянь
$$\begin{cases} \frac{\partial \widetilde{Q_N}(\tau)}{\partial \tau_l} = \sum\limits_{j=1}^{N} \rho' \left(X_j - \sum\limits_{i=1}^{q} \tau_i a_{i,j} \right) a_{l,j} = 0 \\ l = \overrightarrow{1 \ldots q} \end{cases} \tag{3}$$

де N – об'єм вибірки;
q – розмірність вектора оцінених параметрів;
$\vec{X} = \{X_j\}_{j=1}^{N}$ – реалізація вектора-регресора;
$\vec{\tau} = \{\tau_i\}_{i=1}^{q}$ – вектор оцінюваних параметрів,
$a_{i,j}$ – відповідний елемент регресивної матриці.

Введення подібної модифікації суттєво знижує вплив помилкових точок набору даних на результат апроксимації.

Загальний алгоритм робастної апроксимації можна розбити на п'ять основних кроків:

1. класичним МНК отримати початкове лінійне приближення набору точок прямою. Отримати початкове наближення вектор параметрів прямої $\vec{\tau}_0$;
2. для $\vec{\tau}_0$ перерахувати початкове наближення усіх додаткових параметрів функції втрат;
3. доки вектор параметрів істотно відрізняється: $\vec{\tau}_n - \vec{\tau}_{n-1} > \varepsilon$;
4. апроксимувати набір точок, використовуючи попередні модельні параметри функції втрат. Отримати вектор параметрів $\vec{\tau}_n$;
5. повторно перерахувати модельні параметри функції втрат [2]

Приклад моделювання агрегованого цінового споживчого індексу України. Кількісні методи теорії перехідної економіки, як і будь-якої іншої теорії, спираються на дані про об'єкт дослідження, отримані в результаті вимірювань.

У відносно слабо дослідженій моделі перехідної економіки України роль точності вимірювання економічних показників зростає ще більше. Даний аналіз присвячується одному з аспектів моделювання української економіки – точності агрегованого показника макроекономічної динаміки. Будь-який опис макроекономічної динаміки передбачає використання агрегованого індексу цін, що описує динаміку інфляційних процесів та агрегованого індексу виробництва, що описує економічний зріст чи спад. [3]

Основні проблеми визначення рівня росту цін української перехідної економіки проявляються в різному масштабі похибки для сильно рознесених у часі періодів, а тому можуть впливати на результати відносно довгострокових спостережень. [4]

Економіка України має суттєву специфіку. Для її трендового складника характерні високі темпи спаду та росту. Масштаб подібного еволюційного процесу вирізняє Україну на тлі інших країн світу.

Досліджується залежність агрегованого індексу споживчих цін від наступних економічних факторів:

1. індекс цін споживчої продукції,
2. курс долара США,
3. номінальна середньомісячна заробітня плата, грн.

Період фіксації модельної вибірки становить 1 місяць. Усі данні опубліковано Державною службою статистики України на офіційному сайті

Для зменшення рівня похибки. обумовленої високою інтенсивністю еволюції слід проводити підвищення адаптивності існуючих методів моделювання макроекономічних показників.

Відповідно до результатів моделювання отримано наступний результат:

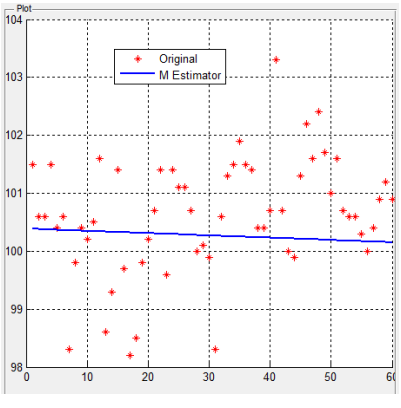


Рис. 1. Графічний результат моделювання

Відзначається стійкість результатів моделювання до деяких спостережень 2002, 2003 та 2004 років, що суттєво відрізняються. Наявність середньоквадратичної похибки рівня: $SSE = 1.1293$, характеризує високу степінь адекватності моделі.

Література. **1.** Adenstedt R.K. On Large Sample Estimation of the Mean of a Stationary Random Sequence // Ann. Statist. – 1974. – Vol. 2. – P. 1095-1107. **2.** Дегтярева А, Вержневец В. Line fitting, или методы аппроксимации набора точек прямой [Электронный ресурс].- Режим доступа cgm.computergraphics.ru/content/view/41 **3.** Бессонов В.А. О точности сводных показателей экономической динамики в переходной экономике. – М., 2002. – 15 с **4.** Дрейпер Н., Смит Г. Прикладной регрессионный анализ. – М.: Финансы и статистика, 1986. – Кн.1. – 336 С., Кн.2. – 351 С.

Буценко Ю.П., Лабжинський В.А.
НТУУ "КПІ", Київ, Україна

Удосконалений алгоритм голосування в складних технічних системах

Традиційне застосування процедури голосування для підвищення надійності складних технічних систем пов'язане зі збільшенням вірогідності їх діагностування за рахунок використання структурної надлишковості (резервування) або алгоритмічної надлишковості (багатоверсійне програмування) [1]. Використання такого підходу у випадку розподіленої системи датчиків, від яких надходить інформація про стан системи (при цьому можливі ситуації їх розташування на окремих компонентах системи), є відносно новим, проте важливим та перспективним. Запровадження такого роду алгоритмів контролю та керування пов'язане з двома важливими їх особливостями. По-перше, у багатьох випадках існують внутрішні зв'язки між різними датчиками розподіленої системи, що вимагає попереднього "внутрішньокластерного аналізу", для якого, знову ж таки, доцільною є процедура голосування. Зрозуміло, що питання про вагу кластера вимагає детального аналізу в кожній конкретній ситуації. По-друге, у вказаному специфічному випадку доречними виявляються наявні удосконалення процедур голосування і, перш за все, зважене голосування з урахуванням наявної передісторії. Точніше, враховуючи те, що на датчики та канали зв'язку системи впливають потоки збоїв і відмов, встановлюють:

- мультиплікативні коефіцієнти "знецінення" показників адекватності результатів попередніх голосувань окремо для випадків різних дійсних станів системи [2];
- показники "накопиченої ваги адекватності" для датчиків, що визначають необхідність їх заміни або технічного обслуговування;
- рівні для результатів голосування, за яких система вважається такою, що знаходиться в одному зі скінченної множини своїх станів ("нормальному", "аварійному" тощо).

Вказана методологія підвищення надійності функціонування системи вимагає:

- проведення класифікації станів системи, яких може виявитись достатньо велика кількість, що має місце, наприклад, у природно-технічних системах;
- виявлення та формалізації у вигляді математичних моделей зв'язків між елементами системи, що дозволяє представити її у вигляді сукупності кластерів, до кожного з яких належать взаємозалежні елементи, причому структура кластерів є динамічною (пов'язаною, наприклад, з природними факторами, а також з характеристиками системи);
- визначення сукупності контрольованих показників, які повною мірою визначають стан системи, та діапазонів їх значень, характерних для її знаходження у кожному зі станів;
- опису процедур "внутрішньокластерного" та "загальносистемного" голосування.

При цьому:

- голосування всередині кластера пов'язане з визначенням вагомості показів кожного з датчиків по кожному з параметрів, які вимірюють, на основі наявної інформації про відповідність його показів реальній ситуації (з визначеного вище спектру), обчислення ваги для j-го датчика здійснюють за формулою $w_{\text{j}} = a_{\text{k}}q^{\text{k}} + a_{\text{k-1}}q^{\text{k-1}} + ... + a_1 q$, де a_{i} - показник відповідності отриманої від датчика у i-му вимірюванні інформації дійсному стану системи (може бути як додатним, так і від'ємним), $q \in (0; 1)$ - коефіцієнт "знецінення" інформації з часом, i - номер попередньої процедури голосування у зворотному напрямку;
- врахування результатів "внутрішньокластерного" голосування при "загальносистемному" може здійснюватись на основі "рівноважності" кластерів або надання їм знову ж таки певних ваг аж до права "вирішального голосу".

Література. **1.** Дедков В. К. Основные вопросы эксплуатации сложных систем / В. К. Дедков, Н. А. Северцев. - М.: Высшая школа, 1976. - 406 с. **2.** 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., Виштальд Д.М., Любашенко Н.Д.
Национальный технический университет Украины "Киевский политехнический институт", Украина

Об оценке доступности сетевой услуги в классе бинарных стохастических моделей

Рассматривается сеть связи, структура которой представлена размеченным связным неорграфом $G(Y;X)$ без петель, где: $Y=[y_1,..,y_m]$ - множество вершин (узлов сети); $X=[x_1,\ldots,x_n]$ – множество ребер (дуг).

Эволюция элементов и сети в целом моделируется соответствующими альтернирующими процессами восстановления с конечными средними значениями. Предполагается, что зависимостью между отдельными элементами сети в процессе функционирования, по крайней мере в заданном режиме сетевой нагрузки, можно пренебречь.

Пусть структура сети представлена матрицей смежности узлов.

Доступность сетевой услуги трактуется как вероятность предоставления связи для любого пользователя в любой момент времени. Можно дать несколько формальных определений понятия доступности сетевой услуги:

- сетевая услуга доступна, если в любой момент времени между любой парой узлов существует по крайней мере один простой путь;
- сетевая услуга доступна, если в любой момент времени существует по крайней мере одно остовное дерево сети;
- сетевая услуга доступна, если в каждом из разрезающих множеств сети существует по крайней мере один доступный элемент.

Эти определения эквивалентны, но по-разному представляют отношения доступности на множестве узлов сети. Отношения доступности сетевой услуги обладают свойствами рефлексивности, симметричности и транзитивности [1–3]. В силу этого пространство состояний сети разбивается на два класса эквивалентности. Во избежание излишней громоздкости будем считать, что узлы сети абсолютно доступны и в представлении структуры сети играют вспомогательную роль при представлении отношений доступности сетевой услуги. С учетом этого замечания будем считать, что доступность (недоступность) сетевой услуги однозначно определяется состоянием каналов (ребер) сети в данный момент времени [4]. Обозначим $X(t)=x_1(t)\ldots x_n(t)$ - вектор состояния каналов (ребер) в момент времени t

$$X_i(t)=\begin{cases}1, & \text{если в момент времени } t \text{ } i\text{-й элемент будет доступен для выполнения сетевой услуги}\\ 0, & \text{в противном случае, } i=1,...,n.\end{cases}$$

Тогда пространство состояний сети можно рассматривать в зависимости от заданных операций на множестве векторов и как векторное пространство по $mod(2)$ и как булеву векторную решетку. Это дает возможность представлять отношение доступности сетевой услуги соответствующей монотонной функцией алгебры логики (ФАЛ)

$$\phi(t)=\begin{cases}1, & \text{если в момент времени } t \text{ существует по крайней мере одна минимальная конфигурация}\\ & \text{элементов, обеспечивающая доступность сетевой услуги,}\\ 0, & \text{в противном случае}\end{cases}$$

Воспользуемся следующими свойствами монотонной ФАЛ:

1. Если $\phi(x_1(t),\ldots,x_n(t))$ - монотонная функция, то ее минимальная дизъюнктивная нормальная форма (МДНФ) представима суммой простых импликант, то есть $\phi(x_1(t),\ldots,x_n(t))=\vee_{i=1}^{|I|}\alpha_i(x_1,\ldots,x_n)$, где I – полный список простых импликант функции $\phi(x_1(t),\ldots,x_n(t))$, $\alpha_i(x_1,\ldots,x_n)$ - i-я простая импликанта $\phi(x_1(t),\ldots,x_n(t))$.
2. Если $\phi(x_1(t),\ldots,x_n(t))$ – монотонная функция, то и двойственная по отношению к ней функция $\phi^*(x_1(t),\ldots,x_n(t))$ монотонна. Так как $\phi^*(x_1(t),\ldots,x_n(t))$ монотонна, то ее

МДНФ представима в виде суммы простых импликант, то есть $\phi^*(x_1(t),\dots,x_n(t)) = \vee_{j=1}^{|J|}\beta_j(x_1,\dots,x_n)$ где J – полный список простых импликант функции ϕ^*, $\beta_j(x_1,\dots,x_n)$ - j-я простая импликанта функции ϕ^*.

Учитывая свойство дуальных автоморфизмов булевой алгебры $(\phi^*)^* \equiv \phi$, получаем простое, но важное утверждение: $\vee_{i=1}^{|I|}\alpha_i(x_1,\dots,x_n) \equiv \phi(x_1(t),\dots,x_n(t)) \equiv \wedge_{j=1}^{|J|}\beta_j(x_1,\dots,x_n)$, то есть любая монотонная ФАЛ имеет однозначное МДНФ- и МКНФ-представление. Утверждение остается верным не только для дискретных, но и для непрерывных неотрицательных значений переменных.

Введем следующие обозначения: $\alpha_i(x_1,\dots,x_n) = \wedge_{k_i\in\{\alpha_i\}}x_{k_i}$; $\beta_j(x_1,\dots,x_n) = \wedge_{r_i\in\{\beta_i\}}x_{r_i}$. $\{\alpha_i\},\{\beta_i\}$ - множество индексов существенных переменных соответствующих простых импликант. С учетом этого замечания рассмотренное выше утверждение можно записать в виде $\vee_{i=1}^{|I|}(\wedge_{k_i\in\{\alpha_i\}}x_{k_i}) \equiv \phi(x_1,\dots,x_n) \equiv \wedge_{j=1}^{|J|}(\vee_{r_i\in\{\beta_i\}}x_{r_j})$

Для нахождения наиболее компактного представления доступности сетевой услуги используются алгоритмы поиска путей, поиска циклов, поиска остовных деревьев, поиска разрезающих множеств сети. Для нахождения разрезающих множеств, циклов, остовных деревьев, пространство состояний элементов сети рассматривается и как векторное пространство по $mod(2)$ и как булева векторная решетка. Процесс функционирования сети может рассматриваться как процесс случайного блуждания по булевой векторной решетке (гиперкубу). Интерес представляет стационарная вероятность пребывания сети в подмножестве состояний, когда сетевая услуга доступна (недоступна), а также среднее стационарное время пребывания сети в подмножестве состояний доступности (недоступности) сетевой услуги. Векторное пространство по $mod(2)$ содержит два ортогональных подпространства: подпространство разрезающих множеств, размерность которого равна рангу графа $\rho(G) = |Y| - 1$ и подпространство циклов, размерность которого равна цикломатическому числу графа $\mu(G) = |X| - |Y| + 1$.

Вопрос о том, какому из представлений МДНФ (МКНФ) отдать предпочтение, зависит как от размерности соответствующих подпространств так и от специфики рассматриваемой задачи. Пусть найдено наиболее компактное (по числу символов) представление функции $\phi(x_1,\dots,x_n)$ в виде МДНФ или МКНФ. Для нахождения вероятности доступности сетевой услуги необходимо ортогонализовать выбранное представление, используя обобщенную формулу де Моргана. После чего переменным найденного представления без отрицания, например x_i, присвоить значения $P_i = \frac{M\Theta_i}{M\Theta_i+M\Xi_i}$, а переменным $\overline{x_j}$ присвоить значения $q_j = \frac{M\Xi_i}{M\Theta_i+M\Xi_i}$ где $M\Theta_i$, $M\Theta_j$ - средние длины интервалов времени доступности элементов x_i и $x_j, i, j = 1,\dots,n$; $M\Xi_i$, $M\Xi_j$ - средние длины интервалов времени недоступности элементов x_i и x_j; $(M\Theta_i + M\Xi_j)$, $(M\Theta_j + M\Xi_j)$ - средние стационарные длины циклов соответствующих альтернирующих процессов восстановления. Подставив в ортогонализованную функцию на место переменных без отрицания значения вероятности доступности элементов, а на место переменных с отрицанием вероятности недоступности элементов и проведя соответствующие вычисления, получим точную оценку вероятности доступности сетевой услуги.

Задачи оценки доступности сетевой услуги относятся к классу трудновычислимых задач комбинаторной логики свойств и классов. По этой причине получение точной оценки вероятности доступности не всегда возможно и приходится ограничиваться оценками “снизу” и “сверху”. Использование изоморфизма булевых алгебр свойств и классов, а также использование дуальных изоморфизмов булевых алгебр позволяет на несколько порядков поднять порог вычислимости при точной оценке вероятности доступности сетевой услуги, а в случае сетей очень большой размерности конструктивным образом получать оценки “снизу” и “сверху”.

Литература. **1.** Г. Биркгоф. Теория решеток - М.: Наука, 1984. **2.** Г. Биркгоф, Т. Барти «Современная прикладная алгебра» М.: «Мир» 1976 г. **3.** Р. Сикорский. Булевы алгебры. - М.: Мир - 1969.- 375 с. **4.** Р. Барлоу, Ф. Прошан «Статистическая теория надежности и испытания на безотказность» М., «Наука», 1984 г. – 327 с.

Вероцький Д.С.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Застосування поєднаних марківських ланцюгів для оцінювання кредитних ризиків

Банківські регулятори використовують кредитний рейтинг як значимий параметр при обчисленні достатності капіталу [1]. Залежність між переходами рейтингів різних компаній моделюється за допомогою поєднання відповідних марківських ланцюгів, вважається, що одиночні зміни рейтингу є марківськими процесами. Модель не залежить від припущення про сумісний розподіл вартості активів компанії.

Спільні рейтингові переходи компаній в різних класах, що належать до різних промислових секторів моделюються наступним чином:

- Переходи компаній з однаковими кредитними рейтингами є залежними.
- Змінення компаній за кредитним рейтингом також є залежними.
- Кожен одиночний перехід визначається марківською матрицею, яка однакова для всіх компаній.

Розвиток портфелю моделюється за допомогою багатовимірного випадкового процессу $X^t = (X_1^t, \dots, X_N^t)$. Через X_n^t позначено рейтинг компанії n на кінець періоду t, де період t є проміжком часу між та $t-1$ та t. Процессы X_n^t моделюються як залежні дискретні марківські ланцюги з простором стану $\{1, \dots, M+1\}$, матрицею перехідних ймовірностей P та поглинаючим станом $M+1$. На основі [2] будуємо модель наступним чином. Переміщення рейтингу компанії n в момент часу t розділене на ринкову η_n^t та залишкову частину δ_n^t. Компоненти приймають значення $\{1, \dots, M+1\}$, їх перехідні ймовірності залежать від X_n^{t-1}.

$$X_n^t = \delta_n^t \xi_n^t + (1 - \delta_n^t)\eta_n^t, \tag{1}$$

де δ_n^t — бернулівські величини, що залежать від рейтингового класу компанії n в момент часу $t-1$, від X_n^{t-1} та промислового сектору $s(n)$ компанії n.

Через $\chi = (\chi_1, \dots, \chi_M)$ позначимо вектор бернулівських величин, де χ_M визначає, чи відбувається крок без погіршення для всіх компаній n та $X_n^{t-1} = m$ в момент часу t. Для набору спостережуваних значень $x = (x^1, \dots, x^T) \in \mathbb{R}^{N \times T}$ процессу X^t та періоду в T кроків оцінюються [3] методом максимальної правдоподібності параметри моделі $Q = (q_{m,s})$, де $q_{m,s} = \mathbb{P}(\delta_n^t = 1 | X_n^{t-1} = m, s(n) = s)$, та щільність спільного розподілу P_χ для $\chi = (\chi_1, \dots, \chi_M)$. Функція максимальної правдоподібності має вигляд:

$$L(x, Q, P_\chi) = \prod_{t=2}^{T} \sum_{\overline{\chi} \in \{0,1\}^M} P_\chi(\chi^t = \overline{\chi}) \mathbb{P}(X^t = x^t | \chi^t = \overline{\chi}, X^{t-1} = x^{t-1}) \tag{2}$$

Метод максимальної правдоподібності приводить до необхідності розв'язання неопуклої задачі оптимізації, рішення якої дає можливість отримати оцінки параметрів моделі. Розглянута модель порівнювалась з моделлю, що описана в роботі [4]. Було показано, що на згенерованих сценаріях для спільних рейтингових переходів тестових компаній існує більш сильна залежність між переміщенням окремих боржників. Це в свою чергу призводить до більш консервативних портфельних рішень.

Література. **1.** Basel Committee on Banking and Supervision. International convergence of capital measurement and capital standards: A revised framework. Technical report, 2004. **2.** Y. M. Kaniovski and G. Ch. Pflug. Risk assessment for credit portfolios: A coupled markov chain model. Journal of Banking & Finance, 31(8):2303–2323, 2007. **3.** Wozabal, D., & Hochreiter, R. (2012). A coupled markov chain approach to credit risk modeling. Journal of Economic Dynamics and Control, 36(3), 403–415. **4.** A.J. McNeil and J.P. Wendin. Bayesian inference for generalized linear mixed models of portfolio credit risk. Journal of Empirical Finance, 14(2):131–149, 2007.

Волкова В.Н.[1]; ***Черный Ю.Ю.***[2]
[1] *Санкт-Петербургский государственный университет Петра Великого, г. Санкт-Петербург, Российская Федерация;* [2] *Институт научной информации по общественным наукам РАН, г.Москва, Российская Федерация*

Системный анализ информационных основ развития социума

Одна из основных принципиальных закономерностей сложных систем – способность противостоять энтропийным (разрушающим систему) тенденциям и проявлять негэнтропийные тенденции. Именно эта закономерность принципиально отличает развивающиеся открытые системы с активными элементами от закрытых систем, для которых справедлив закон физики – стремление системы занять минимальное энергетическое состояние, т.е. к возрастанию энтропии («второе начало термодинамики» или просто «второе начало»).

При применении этого закона для осмысления процессов в живых и социально-экономических системах процесс возрастания энтропии в закрытых системах, изолирован-ных от других систем, объясняют таким образом: "*...В технике это называется аморти-зацией, в химии – деструкцией, а в биологии – старением*" [4, с.5]. Такие процессы наблюдаются на практике и в социально-экономических системах.

По ***Л. фон Берталанфи*** [1] способность противостоять энтропийным (разрушающим систему) тенденциям и проявлять негэнтропийные тенденции обусловлено наличием *активных элементов*. Эти элементы стимулируют обмен материальными, энергетическими и информационными продуктами со средой и проявляют собственные "инициативы". Благодаря действию активных элементов в системах нарушается закономерность возрастания энтропии (аналогичная "второму началу") и наблюдаются *негэнтропийные* тенденции, т.е. *самоорганизация и развитие*.

Энтропийно-негэнтропийные процессы сложно трактовать в упрощенных терминах.

С одной стороны, существует представление об энтропии как о хаосе, мере неупорядоченности системы, а о негэнтропии как мере упорядоченности. Но, с другой стороны, именно негэнтропийные тенденции являются основой развития (травинка пробивает асфальт благодаря этой тенденции), любых инноваций, а они дестабилизируют систему, вносят нестабильность, т.е. неупорядоченность ("креативное разрушение" по ***В. Зомбарту*** и ***Й. Шумпетеру***). В то же время энтропийные тенденции, напротив, стабилизируют состояние системы, поскольку минимальное энергетическое состояние, к которому они приводят, является самым устойчивым.

Анализ истории появления представлений об источниках *негэнтропийных тенденций* позволяет осознать, что их источником является информация. Становление этого понятия связано с рядом предшествующих понятий и философских концепций, близких терминов и определений, вводимых философами античности, средних веков и последующих периодов развития науки: *эйдос, идея, форма* (*In-form* – в форме), *мысль, сведения, знание, разъяснение* (лат. *informatio*) [8, 10], которые использовались для обозначения явлений и процессов, связанных с негэнтропийными тенденциями. Эти формы до появления собственно термина "информация" можно назвать *протоинформацией* [11, с.234].

Диалектический подход предполагает, что для развития необходимы парные категории, что подтверждается развитием представлений о материи и сопутствующих категорий – от эйдоса до информации. В диалектике ***Г.Гегеля*** информации соответствует "абсолютная идея", с которой он начинает процесс эволюции, что вызвало критику в материалистической философии. Ф. Энгельс попытался убрать абсолютную идею и представить процесс развития без нее ("Диалектика природы"), что не получилось. Вместо абсолютной идеи в теории ***К.Маркса*** – ***Ф.Энгельса*** предлагается понятие "самодвижение целостности", т.е. наличие неких активных начал, которые обеспечивают развитие. В современной атрибутивной теории информация

предполагается неотъемлемым свойством (атрибутом) материи, и поэтому она может проявлять себя во всех объектах, процессах и явлениях как живой, так и неживой природы [9]. Таким образом, развитие понятия "информация" привело ученых к выводу, что ее можно рассматривать как субстанцию, определяющую единство и развитие мира, категорию парную материи.

Это позволяет глубже осознать роль *институтов социальной памяти* и *научных коммуникаций* как основы создания эмерджентных свойств системы, т.е. появления у нее принципиально новых, неожиданных свойств, способствующих развитию системы. Негэнтропийные тенденции обеспечиваются передачей накопленной новой информации от предыдущего поколения к последующему, что осуществляется различными способами, основные из которых приведены на рис. 1.

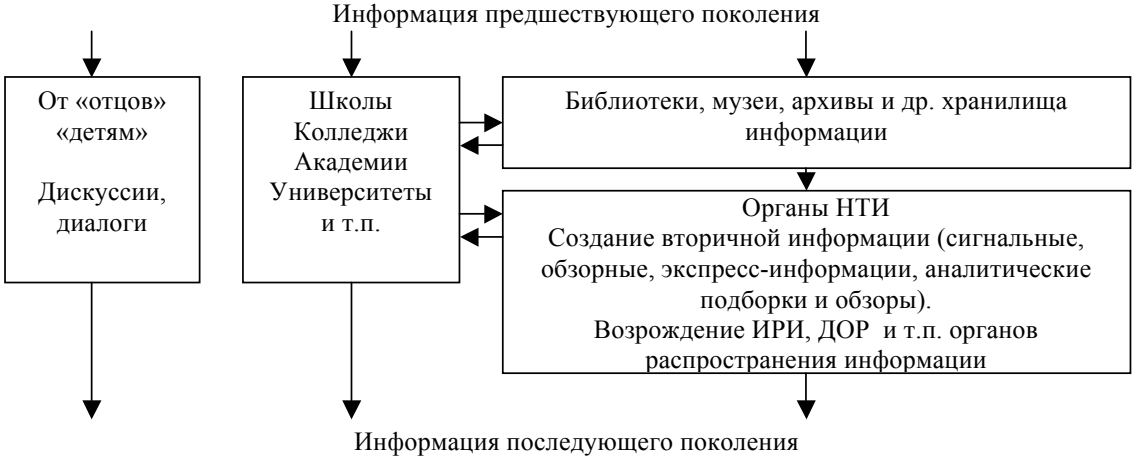


Рис. 1. Способы передачи информации от предшествующего поколения к последующему

Библиотеки, музеи, архивы и др. хранилища информации играют важнейшую роль. Они накапливают информацию и передают информацию от предшествующего поколения к последующему, сохраняя ее в течении длительных периодов с момента возникновения.

В то же время при оформлении информации в виде документов, создании хранилищ, архивов она в определенной мере "*омертвляется*", и необходимы меры по ее "оживлению", т.е. обеспечивающие активизацию ее использования. В музеях эту роль выполняют гиды; в библиотеках – библиографы, информационные работники справочно-информационных фондов. Однако при росте источников информации, при усложнении общественных и информационно-коммуникационных процессов нужны и специальные формы, обеспечивающие активизацию использования информации.

Расширить возможности библиотек, являющихся основными хранилищами культурного наследия любого общества, можно путем создания специальных органов научно-технической информации (НТИ), выполняющих функции обработки (реферирование, подготовка обзорных, экспресс-информаций, ориентированных на развитие науки в различных отраслях, аналитических обзоров и т.п.), избирательного распространения научно-технической информации, издания сигнальных информаций).

Научно-техническая *информации* является основой развития науки, которая, в свою очередь, является основой негэнторопийных тенденций в социальной системе, основой развития всех сфер общества. В большинстве развитых стран из разрозненных информационных служб были созданы *центры* НТИ, а в ряде стран – и национальные системы НТИ. В нашей стране такая система формировалась централизованно как *Государственная система научно-технической информации* (ГСНТИ) [6], В тот же период 1960–1970-х гг. была осознана значимость научных

коммуникаций как одного "*из важнейших компонентов сложнейшего социального явления, называемого наукой*" [7].

Для углубленного исследования информационно-коммуникационных процессов и роли различных институтов, обеспечивающих переработку информации, полезно использовать закономерности теории систем.

Заключение. Таким образом, осмысление роли энтропийно-негэнтропийных процессов в развитии общества позволяет глубже осознать роль институтов социальной памяти и научных коммуникаций как основы создания эмерджентных свойств системы, способствующих развитию системы. В то же время при оформлении информации в виде документов, создании хранилищ, архивов она в определенной мере "омертвляется", и необходимы меры по ее "оживлению", обеспечивающие активизацию ее использования, что возможно на основе реализации функций переработки информации, обеспечиваемых органами системы научно-технической информации.

Исследование энтропийно-негэнтропийных и коммуникационных процессов с точки зрения закономерностей теории систем позволяет сделать вывод о том, что для развития общества необходимо создание *единого информационного пространства*, объединяющего все информационные ресурсы общества и системы социальной памяти, включая библиотеки, органы НТИ, музеи, архивы, локальные базы данных, хранилища информации, ресурсы интернет. При этом в настоящее время осознается возрастающее влияние новых технологий [5] на информационно-коммуникационные процессы, увеличение степени целостности за счет увеличения взаимосвязей в социальных сетях, в большой степени за счет компьютеризации, что иногда квалифицируют как "коммуникативный тоталитаризм". Следовательно, необходимо искать формы управление энтропийно-негэнтропийными процессами.

Литература. **1.** Берталанфи Л. фон. Общая теория систем: критический обзор / Л. Фон Берталанфи // Исследования по общей теории систем. – М.: Прогресс, 1969. – С.23–82. **2.** Волкова В.Н. Теория систем и системный анализ: учебник / В.Н.Волкова, А.А.Денисов. – М.: Юрайт, 2014. – 616 с. **3.** Волкова В.Н. К 50-летию появления термина «информатика» в отечественной научной литературе: Памяти Ф.Е.Темникова / В.Н.Волкова, Ю.Ю.Черный // Научно-практич. журнал: Прикладная информатика. № 4(46)2009. – М.: Market, 2013. – С.129–133. **4.** Граве П.С. Кибернетика и психика / П.С.Граве, Л.А.Растригин. – Рига: Зинатне, 1973. – 96 с. **5.** Классификация информационных технологий / А.Ю.Васильев, А.А.Ефремов, В.Н.Юрьев, Н.Б.Паклин // Научно-практич. журнал: Прикладная информатика. Том 10.№ 5 (591)2015. М.: Market, 2054. – С.127–147. **6.** Михайлов А.И. Основы информатики / А.И.Михайлов, А.И.Черный, Р.С.Гиляревский. – М.: Наука, 1976. – М.: Наука, 1968. – 756 с. **7.** Михайлов А.И. Научные коммуникации и информатика / А.И.Михайлов, А.И.Черный, Р.С.Гиляревский. – М.: Наука, 1976. – 435 с. **8.** Новая философская энциклопедия. В 4 т. Т.3. – М.: Мысль, 2001. **9.** Урсул А.Д. Проблема информации в современной науке / А.Д.Урсул. – М.: Наука, 1975. **10.** Философский энциклопедический словарь / Ред.-сост. Е.Ф.Губский и др, 2003. **11.** Черный Ю.Ю. Послесловие к 7-му заседанию семинара "Методологические проблемы наук об информации" / Ю.Ю.Черный / Теория и практика общественной научной информации: сб. науч. тр. // РАН, ИНИОН. – М. 2014. – Вып.22. – 234 с.

Гайвась Б.І. [1], ***Дмитрук В.А.*** [1,2], ***Торський А.Р.*** [1], ***Дмитрук А.А.*** [3]
[1] *Центр математичного моделювання ІППММ ім. Я.С.Підстригача НАН України, Львів;* [2] *Національний університет «Львівська політехніка», Україна;* [3] *Львівський національний університет імені Івана Франка, Україна*

Модель імпульсного осушення шару зерна з врахуванням неоднорідної структури зернини

Технологічні властивості зернових культур можна змінювати в процесі їхньої переробки за допомогою відповідної гідротермічної обробки. На усіх її етапах контроль за показниками технологічних процесів вимагає досконалої вартісної техніки. В умовах конкурентної ринкової економіки високу якість вихідної продукції можуть забезпечити якісні зміни в організації виробництва, зокрема шляхом використання комп'ютерного моделювання усього технологічного процесу та корекції його окремих ланок.

Експериментально доведено, що для рівномірного нагріву шару зерна під час його осушення необхідно забезпечити розділення окремих зерен [1]. Це досягається за допомогою імпульсного режиму сушильного агента, коли при нагріванні (охолодженні) зернової маси теплоносієм зерно "бере участь" у загальному пульсаційному русі суміші. При визначенні гранично допустимої температури нагріву зерна основним фактором є його цільове призначення. Для посівного матеріалу важливо не зруйнувати шкірку зернини, яка зберігає алейроновий пласт, який оберігає клітини ядра. Різниця в будові оболонки і ядра впливає на термодинамічні характеристики вологопереносу в цих матеріалах, тому при математичному моделюванні важливо окремо описувати процеси, які відбуваються в шкірці і в ядрі зернини.

Розглянуто задачу про осушення в імпульсному режимі агента сушіння зернового шару товщиною L. Шар віднесено до прямокутної декартової системи координат так, що вісь Oz є перпендикулярною до поверхонь шару. Зернини шару насичені вологою та мають форму сферичних кульок радіусу R. Межа розділу ядра (область 1) та оболонки (область 2) зернини $r = r*$. Початкові температури шару зерна та i-ї області зернини T_z, $T^{(i)}$відповідно. Шар кульок продувається сухим теплим повітрям зі швидкістю v конвективного переносу в міжзерновому середовищі. Приймається гіпотеза ідеального витіснення. Основне співвідношення для моделі ідеального витіснення витікає з умови неперервного потоку $v = const$ [2]. На вході приймаються збурення за концентрацією та температурою A_c, A_T відповідно.

Вологоперенос у зерні носить дифузійний характер [3]. Повна система рівнянь для зернини складається з рівнянь механодифузії та теплопровідності, які у сферичній системі координат для i-ї області зернини (надалі позначатимуться верхніми індексами в дужках) мають вигляд:

$$\frac{\partial}{\partial r}\left[\frac{1}{r^2}\frac{\partial\left(r^2u_r^{(i)}\right)}{\partial r}\right]-\xi^{(i)}\beta^{(i)}\frac{\partial c^{(i)}}{\partial r}=0,\ \frac{\partial\left(rc^{(i)}\right)}{\partial\tau}=D^{(i)}\frac{\partial^2\left(rc^{(i)}\right)}{\partial r^2}-D_\varepsilon^{(i)}\frac{\partial^2\left(r\varepsilon^{(i)}\right)}{\partial r^2},$$

$$\frac{\partial\left(rT^{(i)}\right)}{\partial\,\tau}=a_T^{(i)}\Delta\left(rT^{(i)}\right)\quad(i=1;\ 2)\,, \tag{1}$$

де $\xi^{(i)}=3K^{(i)}\big/\left(3K^{(i)}+4G^{(i)}\right)$ – механічна стала, $\varepsilon^{(i)}=\partial u_r^{(i)}\big/\partial r+2u_r^{(i)}\big/r$ – об'ємна деформація, $c^{(i)}$ – концентрація вологи, τ – час, $u_r^{(i)}$ – радіальні переміщення, $\beta^{(i)}$ – концентраційний коефіцієнт об'ємного розширення, $K^{(i)}$ – модуль всестороннього об'ємного стиску, $G^{(i)}$ – модуль зсуву, $D^{(i)}$ – коефіцієнт дифузії, $D_\varepsilon^{(i)}$ – коефіцієнт впливу градієнта поля об'ємної деформації на масовий потік вологи за рахунок колоїдності матеріалу зерна, T – температура, $a_T^{(i)}$ – коефіцієнт температуропровідності в i-ій області.

У початковий момент часу ($\tau = 0$) задано переміщення $u_r = 0$ та концентрації $c = c_0 = const$.

На зовнішній поверхні зернини $r = R$ дифузійний потік пропорційний до різниці концентрацій вологи на поверхні зернини і міжзернового середовища

$$c_z\ :\ -\tilde{D}^{(2)}\partial c^{(2)}\Big/\partial r\Big|_{r=R}=k(\,c^{(2)}\Big|_{r=R}-c_z),\quad \tilde{D}^{(i)}=D^{(i)}-\xi^{(i)}\beta^{(i)}D_\varepsilon^{(i)},$$

де k– коефіцієнт масовіддачі. На межі контакту області 1 та області 2 зернини $r = r_*$ справджується рівність потоків пароповітряної суміші при випаровуванні:

$$\left(D^{(1)}\frac{\partial\vartheta^{(1)}}{\partial r}+D_{\varepsilon}^{(1)}\frac{\partial\varepsilon^{(1)}}{\partial r}\right)_{r=r*}=\left(D^{(2)}\frac{\partial\vartheta^{(2)}}{\partial r}+D_{\bar{\varepsilon}}^{(2)}\frac{\partial\varepsilon^{(2)}}{\partial r}\right)_{r=r*}. \qquad (2)$$

Тут $\vartheta^{(i)}(r,\bar{z},\tau)=c_r^{(i)}(r,\bar{z},\tau)-c_{z0}(\bar{z},0)$, $c_{z0}(\bar{z},0)$ – початкова умова в міжзерновому середовищі шару, $\varepsilon=\varepsilon_{\cdot\alpha}^{\alpha\cdot}$ – перший інваріант тензора деформації.

При осушенні шару зерна розрізняється тепломасоперенос в зернині і в середовищі, що її оточує. Рівняння дифузії в міжзерновому середовищі аналогічне рівнянню конвективної теплопровідності :

$$\frac{\partial c_z}{\partial\tau}+\upsilon_z\frac{\partial c_z}{\partial z}=D_z\frac{\partial^2 c_z}{\partial z^2}+J_r(r,z,\tau),\quad \frac{\partial T_z}{\partial\tau}+\upsilon_z\frac{\partial T_z}{\partial z}=a_{T_z}\frac{\partial^2 T_z}{\partial z^2}+J_rT(r,z,\tau), \qquad (3)$$

де a_{Tz} – температуропровідність міжзернового середовища, D_z – коефіцієнт дифузії пароповітряної суміші міжзернового середовища, $J_r(r,z,\tau)$, $J_{rT}(r,z,\tau)$ – інтенсивність локального джерела вологи та тепла в результаті випаровування з окремих зернин; z – координата за товщиною шару. Для задовільнення граничних умов на поверхні зернини розв'язок для $\vartheta^{(2)}$ шукаємо у вигляді $\vartheta^{(2)}=\vartheta_1^{(2)}+c_2^{(2)}$, де $\vartheta_1^{(2)}$ задовольняє початковим та граничним умовам та на границі областей (1), (2), а $c_2^{(2)}$ вибирається так, щоб сумарно задовільнились умови конвективного нестаціонарного масообміну на поверхні зернини при імпульсному осушенні.

Визначивши власні значення відповідної задачі Штурма-Ліувіля, отримано спочатку власні функції та фундаментальні розв'язки задачі теплопереносу, а потім, враховуючи зв'язки цих розв'язків з розв'язками задачі вологопереносу, визначено фундаментальні розв'язки задачі окремо для області ядра та шкірки зернини:

$$T^{(i)}(r)-T_{z0}=\frac{1}{r}\sum_{j=1}^{\infty}e^{-\mu_{jT}^2\tau}\left[\left(A_{1jT}^{(i)}+A_{2jT}^{(i)}\right)\cos\left(\lambda_{jT}^{(i)}r\right)+\left(B_{1jT}^{(i)}+B_{2jT}^{(i)}\right)\sin\left(\lambda_{jT}^{(i)}r\right)\right],$$

$$c^{(i)}(r)-c_{z0}=\frac{1}{r}\sum_{j=1}^{\infty}e^{-\mu_j^2\tau}\left[\left(A_{1j}^{(i)}+A_{2j}^{(i)}\right)\cos\left(\lambda_j^{(i)}r\right)+\left(B_{1T}^{(i)}+B_{2T}^{(i)}\right)\sin\left(\lambda_j^{(i)}r\right)\right],$$

де $A_{1jT}^{(i)}$, $A_{2jT}^{(i)}$, $B_{1jT}^{(i)}$, $B_{21jT}^{(i)}$ – коефіцієнти залежності температури та потоку тепла за товщиною зернини від температури збурень зовнішнього середовища в залежності від часу.

Для ілюстрації на рис. 1 наведено 3D розподіли температурного поля $T^{(i)}(r,\tau)$ при нагріванні зерна пшениці (радіус зернини $R=2.00086\cdot10^{-3}$м, вологовміст 12%, температура T_z=248 ° K) в потоці теплого повітря (амплітуда збурення температури A_{1T}=80 ° K) зі швидкістю υ=1.2 м/с за час τ=10 с.

Зерно швидко нагрівається за рахунок тепла сушильного агента. Проте зміна температури в різних частинах зернини відбувається неоднаково. В оболонці різниця температур менша, ніж в ядрі зернини. Тому при сушінні великих об'ємів в агентах імпульсного режиму необхідно проводити перерви між етапами сушіння для перерозподілу вологи з внутрішніх частин зернівки до поверхні.

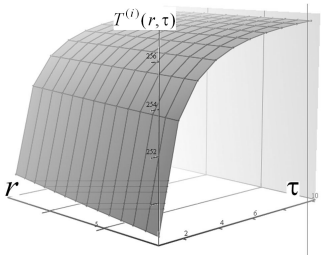


Рис. 1. 3D графік зміни температури в процесі сушіння зерна

Література. **1.** Зимин Е. М. Движение влаги в зернове при сушке / Е. Зимин, В. Крутов // Механизация и электрофикация сел. хоз.-ва. – 2001. – № 4. – С. 11-13. **2.** Сажин Б.С. Основы техники сушки. Москва : Химия, 1984. – 319 с. **3.** Chaplya Y., O. Chernukha, V. Dmytruk. Advective-diffusive mass transfer in binary regular structures in the steady-state regime // Applied Math. Modelling. – 2013. – V. 37. – P. 6191-6211.

Гамоцька С.Л.
Київський національний університет імені Тараса Шевченка, Київ, Україна

Аналіз особливостей моніторингу проектних ризиків в IT

Проблема та її аналіз. Реалізація проекту здебільшого відбувається в умовах невизначеності й ризику, що викликає необхідність виявляти та ідентифікувати ризики, вибирати методи управління, розробляти та вживати заходів для зниження цих ризиків, контролювати й оцінювати результати впроваджуваних заходів.

Необхідність моніторингу проектних ризиків обумовлена відмінностями між проектною та операційною діяльністю. Дані види діяльності розрізняються головним чином тим, що операційна діяльність - це тривалий в часі і повторюваний процес, в той час як проектна діяльність - є тимчасовою і спрямованою на створення унікальних послуг, продуктів або результатів [1].

Моніторинг і контроль – це процедури спостереження за ідентифікацією ризиків, забезпечення виконання плану ризиків і оцінки його ефективності з урахуванням зниження ризику. Показники ризиків, пов'язані зі здійсненням умов виконання плану, фіксуються. Моніторинг проектних ризиків супроводжує весь процес впровадження проекту в життя.

Моніторинг проектних ризиків. Моніторинг ризиків проекту забезпечує поточний аналіз рівня діагностованих ризиків і його відповідності рівню прийнятного ризику, а також розробку контрольних процедур, спрямованих на підвищення ефективності інтегрованого управління ризиками проекту.

В процесі здійснення діяльності проектні ризики змінюються, прогнозовані ризики можуть зникати і замість них виникають нові ризики. Добре поставлені процеси моніторингу та контролю дають можливість отримати інформацію, яка допомагає приймати ефективні рішення ще до того, як виникають нові ризики. Мета моніторингу проектного ризику - визначити, що:

- реакцію на ризики впроваджено відповідно до плану й необхідності змін;
- противоризикові заходи були прийняті так, як це планувалося;
- противоризикові дії настільки ефективні, наскільки очікувалося;
- загроза ризику змінилася з моменту, коли вона була виявлена;
- є сигнали, що свідчать про можливість настання ризику;
- виникли ризики, не виявлені раніше.

Для аналізу та впорядкування ідентифікованих проектних ризиків з урахуванням пріоритетів проекту використовується якісний аналіз ризиків. Отримані результати використовуються згодом в ході кількісного аналізу ризиків. Основне завдання кількісного підходу полягає в чисельному вимірі впливу факторів проектного ризику на поведінку критеріїв ефективності IT-проекту.

Якісний контроль виконання проекту надає можливість отримати інформацію для прийняття ефективних рішень по запобіганню виникнення ризиків, забезпечує ефективний вибір альтернативних стратегій, прийняття коректив, перепланування проекту для досягнення базового плану. Щоб отримати повну та достовірну інформацію про проектні ризики, необхідна постійна взаємодія між усіма менеджерами проекту і розробниками. Для цього регулярно формуються звіти про виконання проекту й системи ризиків.

При моніторингу проектних ризиків застосовуються такі інструменти та методи [1, 2]:

- переоцінка ризиків;
- аудит ризиків;
- аналіз відхилень і тенденцій;
- технічне вимірювання виконання;
- аналіз резервів;
- наради щодо поточного стану.

Переоцінка ризиків відбувається постійно протягом всього життєвого циклу проекту. Обсяг і ступінь деталізації повторень переоцінки залежать від ходу виконання проекту по відношенню

до поставлених цілей.

Аудит ризиків передбачає вивчення і надання в документальному вигляді результатів оцінки ефективності заходів з реагування на ризики, вивчення основних причин їх виникнення, а також оцінку ефективності процесу управління ризиками.

Для *аналізу відхилень* і *тенденцій* необхідні дані, зібрані в процесі виконання проекту. На їх підставі можна прогнозувати потенційні відхилення проекту на момент його завершення за показниками вартості і розкладу. Відхилення від базового плану можуть вказувати на наслідки, викликані загрозами або сприятливими можливостями.

При *технічному вимірі виконання* порівнюються одержувані в процесі реалізації проекту технічні результати із запланованими.

При *аналізі резервів* для визначення адекватності залишку резерву проводиться порівняння обсягу залишилися резервів на непередбачені обставини з кількістю залишилися ризиків станом на будь-який момент часу процесу виконання проекту.

Управління проектними ризиками може бути одним з пунктів порядку денного періодичних *нарад про поточний стан* проекту. Чим частіше застосовується управління ризиками, тим легше воно відбувається, а часті обговорення питань, пов'язаних з ризиками, роблять розмови про ризики, особливо про загрози, більш легкими і точними.

Моніторинг та управління ризиками являє собою процес застосування планів реагування на ризики, стеження за виявленими ризиками, контролю залишкових ризиків, ідентифікації нових ризиків і оцінки ефективності процесу управління ризиками протягом усього життєвого циклу проекту. В результаті його виконання процесу моніторингу на виході будуть отримані такі артефакти:

- Оновлення реєстру ризиків включає в себе результати аудиту та періодичної перевірки ризиків, фактичні результати ризиків проектів і результати реагування на ризики. Реєстр ризиків є частиною документації щодо закриття проекту;
- Запити на зміну виникають в результаті необхідності зміни плану управління проектом у відповідь на ризик. Схвалені запити на зміни оформляються документально;
- Рекомендовані коригувальні дії - це роботи, внесені в плани на непередбачені обставини;
- Рекомендовані запобіжні дії, що використовуються для приведення проекту у відповідність до плану управління проектом;
- Оновлення плану управління проектом виникає, якщо схвалені запити на зміни стосуються процесів управління ризиками;
- Оновлення документів проекту. Результати управління ризиками проекту можуть бути використані в майбутніх проектах і повинні увійти до складу активів організаційного процесу.

Висновки. Таким чином, причиною виникнення ризику є невизначеність, характерна для всіх проектів. Моніторинг проектних ризиків дозволяє оцінити і відслідкувати їх та запланувати витрати і відстежити рентабельність проекту на кожному з його етапів. Використовуючи цю інформацію, проектна команда уникне відставання від графіка та підвищить рентабельність проекту.

На жаль, існуюча на сьогоднішній день загальна методологія моніторингу проектних ризиків, наведена в більшості літературних джерел, носить описовий характер. Наявної в них інформації недостатньо для визначення обсягу коштів, які необхідно закласти в бюджет проекту для компенсування збільшення цін на проектні роботи. У зв'язку з цим вкрай актуальними є науково-прикладні дослідження, спрямовані на розвиток методів моніторингу проектних ризиків в умовах високої невизначеності.

Література. **1.** PMBOK. Руководство к Своду знаний по управлению проектами, 5-е изд./ Project Management Institute. PMI, 2013. 589 с. **2.** Кошелівський І.С. Огляд методів управління проектними ризиками [Текст] // Проблеми сучасної економіки: матеріали II міжнар.наук.конф. (Челябінськ, жовтень 2012 року). - Челябінськ: Два комсомольця, 2012. С. 164-166.

Головатюк В.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Багатокритеріальне оцінювання якості роботи співробітників технічної підтримки компанії з використанням методів підтримки прийняття рішень

Одним із важливих факторів успішності програмних продуктів в умовах сучасної ринкової економіки є надання якісної технічної підтримки. Часто може виникати ситуація, коли товар чи послуга відмінної якості не приносить бажаного результату через те, що клієнти вчасно не отримують необхідної для них інформації, або допомоги зі сторони творців цього продукту.

Задача оцінювання якості роботи працівників технічної підтримки є актуальною, як з точки зору управління персоналом компанії, так і для підвищення привабливості програмного продукту загалом. Адже допомагає звернути увагу на проблемні ділянки в роботі з клієнтами, і вжити заходи спрямовані на покращення обслуговування.

В організаційному контексті працівник є важливим активом організації для досягнення її цілей та мети. У поставленій здачі саме від працівників, в першу чергу, залежить наскільки продукт буде популярним і привабливим для клієнтів. Атестація співробітників є важливим аспектом управління людськими ресурсами, і допомагає оцінити внесок кожного працівника відділу компанії. В даному дослідженні якість роботи співробітників технічної підтримки пропонується розглядати як комплексний показник, що враховує кількість оброблюваних заявок за певний період часу, якість спілкування з клієнтом і особисті знання та навички спрямовані на прокращення обслуговування.

Процес агрегування показників є завданням для систем підтримки прийняття рішень. Для досліджуваної задачі пропонується використати модифікований метод аналізу ієрархій (ММАІ)[1], оскільки він дозволяє поєднати якісні і кількісні показники, а також взяти до уваги уподобання особи, що приймає рішення. Важливою особливістю цього методу є обробка нечітких експертних оцінок. Адже в умовах невизначеності людина експерт не в змозі дати оцінку у вигляді скалярного значення, вона може давати її тільки з певним суб'єктивним ступенем впевненості.

Згідно з МАІ виконується декомпозиція проблеми в ієрархічну структуру, вершиною якої є сама проблема (головна ціль), на останньому рівні розміщується набір альтернатив - співробітники діяльність і компетентність яких мають бути оцінені. Результатом МАІ є глобальні ваги – інтегральні показники якості роботи кожного з працівників. Підготовка вхідної інформації для МАІ відбувається в декілька етапів: 1) формування ОПР/ експертом загальної структури ієрархії, що містить критеріїв і зв'язки між ними; 2) формування оцінок парних порівнянь критеріїв у фундаментальній шкалі відносної важливості та оцінок парних порівнянь альтернатив за якісними критеріями останнього рівня ієрархії критеріїв; 3) формування бази статистичних даних альтернатив за кількісними критеріями[1].

Модифікацією класичного методу МАІ є можливість обробляти нечіткі експертні оцінки. Для цього вводиться поняття нечіткої матриці парних порівнять НМПП $A^{fuz} = \left\{ \left(a_{ij}^{fuz} | i,j = \overline{1,n} \right) \right\}$, для якої $a_{ij}^{fuz} = (x, \mu_{ij}(x))$ – нормальна нечітка множина і відображає результат парного порівняння об'єктів O_i та O_j, $x \in R$, де R – множина дійсних чисел. Значення функції приналежності $\mu_{ij}(x)$ нечіткої множини a_{ij}^{fuz} є ступенем виконання переваги $O_i \succeq O_j$, причому результат порівняння об'єкта з самим собою дорівнює одиниці, тобто $a_{ij}^{fuz} = 1$.

Необхідно:

- знайти вектор нечітких глобальних ваг $W^{fazglob} = \left\{ W_i^{fazglob} | i = \overline{1,N} \right\}$, який відображає переваги з множини НМПП $A^{faz} = \left\{ A_k^{faz} | k = \overline{1,K} \right\}$;
- оцінити узгодженість нечітких експертних оцінок;
- визначити ранжування нечітких глобальних ваг W_i^{faz} ;

• оцінити ступінь довіри до отриманого ранжування.

Під час дослідження розроблено ієрархію критеріїв (рис. 1), які відображають складові якості технічної підтримки програмного продукту [2].

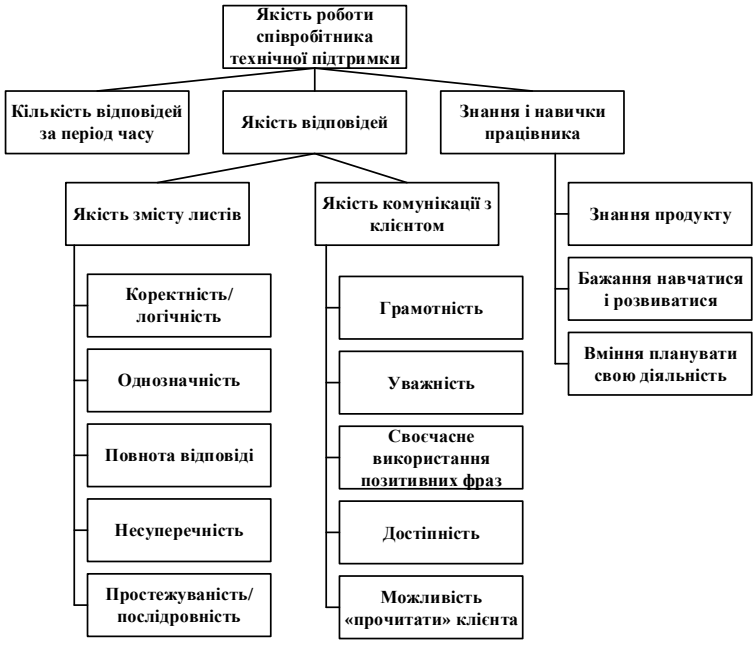


Рис. 1. Ієрархія критеріїв, за якими оцінюється якість роботи працівників технічної підтримки

Для кожного із співробітників ведеться облік кількості діалогів із клієнтами за певний період часу. Ця інформація врахована у критерії кількість відповідей за період часу, і відображає такі показники як швидкість і продуктивність. Якість відповідей клієнтам в свою чергу оцінюється за двома аспектами. Перший з них відображає якість змісту листів, другий – навички комунікації. Не менш важливим є також рівень знання продукту, бажання навчатися і розвиватися, та вміння планувати свою діяльність.

Таким чином ОПР отримує ієрархію критеріїв, що тою чи іншою мірою впливають на якість надання технічної підтримки. Метод аналізу ієрархії дозволяє врахувати важливість кожного із них у вигляді ваг під час формування інтегральної оцінки. На основі отриманої ієрархії розробляється комп'ютерна система прийняття рішень з використанням MAI яка дозволить:

- Оцінити важливість критеріїв;
- Розрахувати ваги кожного із критеріїв методами парних порівнянь;
- Оцінити узгодженість оцінок експертів;
- Розрахувати глобальні ваги альтернатив.

Подальше застосування отриманих результатів дозволяє полегшити і прискорити процес прийняття кадрових рішень у відділі технічної підтримки компанії.

Література. **1.** Панкратова Н.Д., Недашківська Н.І. Моделі і методи аналізу ієрархій: Теорія. Застосування: Навчальний посібник. – К: ВЦ .Видавництво .Політехніка., 2010. – 371с. **2.** Ciotti G., 15 Customer Service Skills that Every Employee Needs [Електронний резурс]. - Режим доступу: https://www.helpscout.net/blog/customer-service-skills/

Гончар В.О.
Національний технічний університет України «КПІ», Київ, Україна

Складність задачі синтезу мережі

Задачі синтезу мереж є математичними моделями чисельних економіко-математичних проблем. Одним з важливих для практики класів задач є задачі оптимального синтезу надійних мереж [1]. На відміну від звичайних задач синтезу мереж в них передбачається, що елементи системи можуть час від часу виходити з ладу, порушуючи нормальне функціонування системи. Припускається, що елементи системи виходять з ладу в певний момент часу, але в той же час не може вийти з ладу не більше одного елемента.

Далі буде розглянуто задачу синтезу надійної мережі як штейнерову версію задачі проектування 2-зв'язної мережі мінімальної вартості, так звану задачу Штейнера про 2-зв'язний граф [2].

Теорема 1. Штейнерова задача на графі є NP-повною.

Доведення було проведено в праці Гері та Джонсона [3].

Розглянемо неорієнтований граф $G = (V, E)$, де V – множина його вершин, а E – множина ребер. Для кожної пари вершин s, t з множини V шляхом між цими (зв'язаними) вершинами називається послідовність вершин та ребер $s = v_0, e_1, v_1, \ldots, v_{l-1}, e_l, v_l = t$, де кожне ребро e_i з'єднує вершини $v_{i-1}, v_i, i = 1, \ldots l$ і жодні вершини чи ребро не зустрічаються в цій послідовності більше, ніж один раз. Коли говорять, що граф $G' = (V', E')$ є підграфом G, це означає, що $V' \subset V$, а множина E' складається з ребер, що належать E, причому їх кінцеві вершини знаходяться в V'. $N_0 = \frac{N}{s}\ \Gamma(i)$ - множина ребер з однією вершиною i на графі, що утворений з графа G шляхом видалення ребер підграфа Г.

Припустимо, що задано множину термінальних вузлів $N (N \subset V)$ та невід'ємні вартості се для кожного з ребер е множини E. Задано також неорієнтований граф $H = (V_h, E_h)$, де V_h та E_h позначено відповідно множини вершин та ребер графа H. Іноді термінальні вузли з N називають просто вузлами. Вартість будь-якого підграфа $G' = (V', E')$ графа G дорівнює сумі вартостей ребер у множині E'. Тоді задача Штейнера про 2-зв'язний граф має вигляд:

$$min \sum_{e \in E} c_e x_e, \tag{1}$$

$$\sum_{j \in \Gamma(i)} x_{ij}^r - \sum_{j \in \Gamma(i)} x_{ji}^r = \begin{cases} 2, & i = s, \\ 0, & i \neq s, r, \\ -2, & i = r, i \in V, r \in N_0 \end{cases}, \tag{2}$$

$$0 \leqslant x_{ij}^r \leqslant x_{ij} \leqslant 1, r \in N_0, e = (i, j) \in E, \tag{3}$$

$$x_e = 0 \cup 1, e \in E. \tag{4}$$

В реальних задачах розглянутого типу можуть зустрічатись графи G з декількома сотнями і навіть тисячами вершин. Розв'язання таких задач точними методами практично неможливо. Тому застосовуються наближені алгоритми. Наприклад, в мережі з 10 вузлів існує 2^{45} варіантів розміщення ліній зв'язку, включаючи множину тривіальних випадків. Якщо припустити, що час необхідний для аналізу кожного варіанту буде становити 1 секунду, то на дослідження всіх варіантів знадобиться більш ніж $9 * 10^8$ років. Навіть для малих мереж з кількістю вузлів 6 та 7 знадобиться 8,8 годин та 24,2 доби відповідно.

Література. **1.** Михалевич В.С., Трубин В.А., Шор Н.З. Оптимизационные задачи производственно транспортного планирования. Модели, методы, алгоритмы. –М.: Наука, 1986. – 260 с. **2.** Шор Н. З., Сергієнко І. В. та інші. Задачі оптимального проектування надійних мереж. — К.: Наук. думка, 2005. — 230 с. **3.** Michael R. Garey and David S. Johnson.(1979) Computers and Intractability: A Guide to the Theory of NP-Completeness. W.H. Freeman and Company.

***Горбань І.М.*[1], *Хоменко О.В.*[2]**
[1]*Інститут гідромеханіки НАН України, Київ, Україна;* [2]*Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна*

Застосування схеми спійманих вихорів до керування потоком навколо квадратної призми

В роботі розглянуто проблему зменшення гідродинамічних навантажень на квадратну призму. Її актуальність зумовлена тим, що призма належить до фундаментальних погано обтічних конфігурацій, що використовується у багатьох технічних системах. Розвинута стратегія управління ґрунтується на концепції спійманих вихорів Рінглеба [1], яка спрямована на створення поблизу тіла нової топології течії зі стійким вихором. В запропонованій керуючій схемі стійку систему вихорів передбачається генерувати за допомогою двох симетричних пластин, приєднаних до передньої грані призми (рис. 1а).

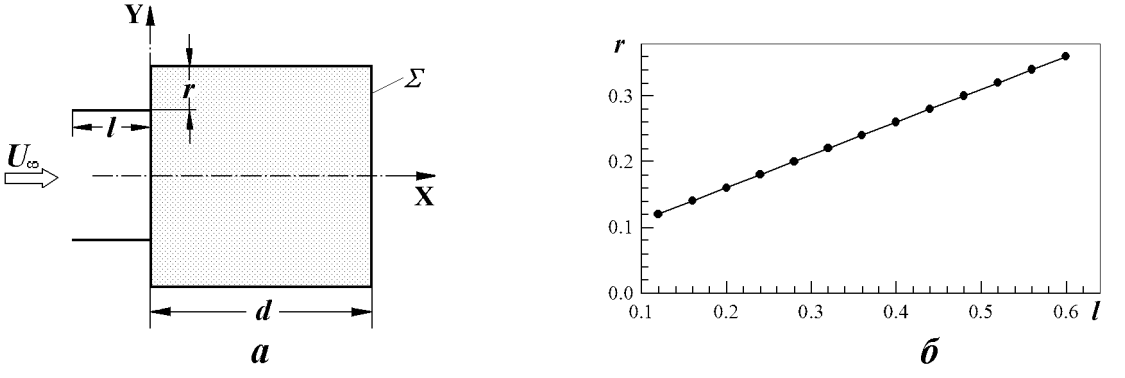


Рис. 1. а)Конфігурація течії навколо призми з пластинами; б)оптимальна залежність положення пластини r від її довжини l.

Оптимальні параметри керуючої схеми, при яких забезпечується стоячий (спійманий) вихор між пластиною і кутом призми, знаходяться на основі теорії критичних точок, яка застосовується до вивчення топології течії, що генерується при обтіканні призми рівномірним потоком швидкості U_∞. Задача керування при цьому зводиться до системи нелінійних диференційних рівнянь, які випливають з фізичних умов нерухомості вихора і безвідривного обтікання крайок призми і пластини [2]. Розрахунки виявили, що стійкий стоячий вихор в даній конфігурації існує лише при певних співвідношеннях довжини пластини l до її положення r відносно кута призми. Отримана в розрахунках залежність $r(l)$ представлена на рис. 1б (зазначимо, що всі лінійні параметри віднесені до довжини ребра призми d).

Для оцінки ефективності запропонованої керуючої схеми виконане чисельне моделювання в'язкої течії навколо призми з малими передніми пластинами (при $l = 0.2$, $r = 0.16$) в діапазоні помірних чисел Рейнольдса: $Re = 100 \div 500$, де $Re = U_\infty d/\nu$. У розрахунках рівняння Нав'є-Стокса в $2D$ просторі інтегрувалися у змінних "завихреність-швидкість" за схемою, яка докладно описана в роботі [3]. Її апробація на прикладі обтікання квадратної призми показала відповідність одержаних динамічних коефіцієнтів експериментальним даним, а також розрахункам, виконаним із застосуванням кінцево-різницевих схем високого порядку.

Виконане чисельне моделювання виявило істотну регуляризацію сліду за призмою з пластинами (рис. 2), що проявляється у зменшенні розмірів циркуляційної зони, а також інтенсивності та масштабу вихорів, які формують доріжку Кармана.

Регуляризація сліду призводить до зменшення динамічних навантажень на призму, причому, вплив пластин стає більш істотним, коли число Рейнольдса зростає. Так, зменшення коефіцієнту сили опору змінюється від 20% при $Re = 100$ до 35% при $Re = 500$ (рис. 3 а). Ще

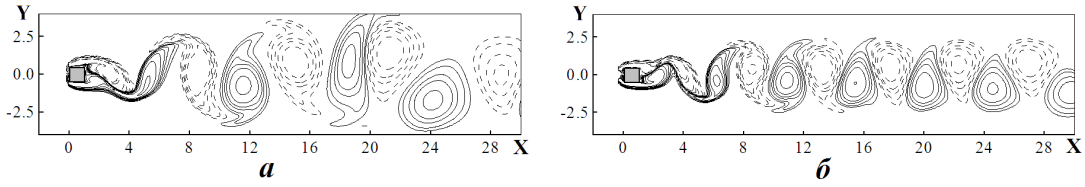


Рис. 2. Контури завихреності за квадратною призмою при $Re = 500$: а)без керування, б)з оптимальним керуванням (– – –– додатна циркуляція, — — — від'ємна циркуляція).

більш значним є падіння амплітуди під'ємної сили, яке, в залежності від числа Рейнольдса, знаходиться в діапазоні 50% - 70%. Число Струхаля, яке характеризує частоту відриву вихорів, у керованому потоці збільшується (рис. 3 б).

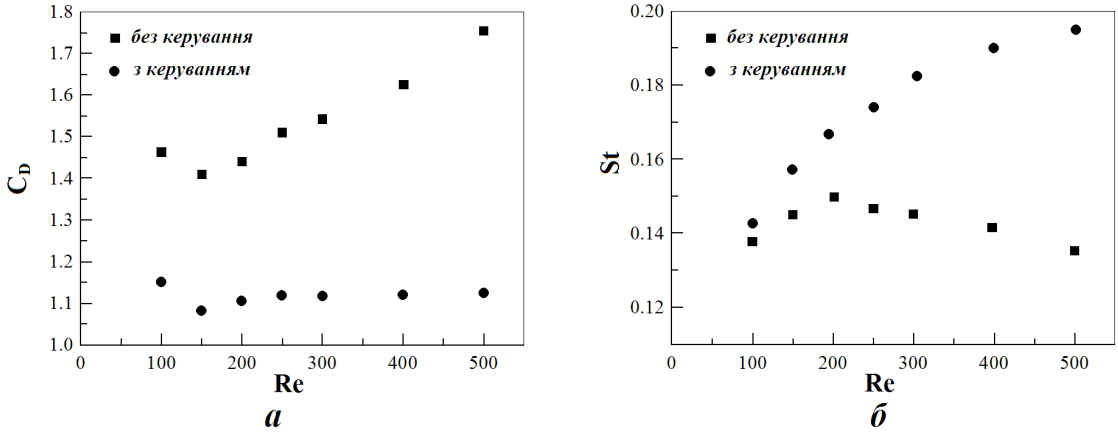


Рис. 3. Порівняння характеристик квадратної призми без керування та з керуванням при різних числах Рейнольдса: а) середній коефіцієнт сили опору, б)число Струхаля.

Виконане чисельне моделювання підтверджує, що вплив управляючих пластин на структуру течії і гідродинамічні характеристики призми є найбільш ефективним при оптимальних співвідношеннях довжини та положення пластин, які отримані на основі спрощеної моделі стоячого вихора (рис. 1 б). З цього випливає головний висновок проведеного дослідження про те, що нові успішні способи керування потоками рідини мають будуватися із залученням динамічних моделей, які спрямовані на зміну топології течії відповідно до мети керування.

Література. **1.** Ringleb F.O. Two-Dimensional Flow with Standing Vortex in Ducts and Diffusers//ASME J. Basic Eng. – 1960. – 82, №4. – P. 921–927. **2.** Gorban I.M., Homenko O.V. Dynamics of vortices in near-wall flows with irregular boundaries// Continuous and Distributed Systems: Theory and Applications. Solid Mechanics and Its Applications / M.Z. Zgurovsky, V.A. Sadovnichiy (Eds.). — 2014. — 211. — P.115-128. **3.** Горбань В.О., Горбань І.М. Вихрова структура потоку при обтікання квадратної призми: числова модель та алгоритми управління// Прикладна гідромеханіка. – 2005. – 7. №2. – С. 8–26.

Горелова Г.В.
Институт управления в экономических, экологических и социальных системах ИТА ЮФУ, Таганрог, Россия

Критерии оценки эффективности процессов в сложных системах в условиях неопределенности, когнитивный аспект

В данной работе рассматриваются в качестве объекта исследования социально-экономические системы (СЭС). СЭС, как известно, относятся к классу сложных систем, процессы в которых происходят под воздействием явных и латентных факторов, цепочек внутренних и внешних событий. Это затрудняет как прогнозирование развития процессов и оценку их эффективности или опасности, так и принятие решений по управлению ими или адаптации к ним. Исследование поведения сложных систем включает моделирование их структуры и поведения, поскольку эксперимент над реальной сложной системой чаще всего нежелателен, поскольку может быть дорогим, длительным, физически невозможным, опасным и т.п. Классы моделей сложных систем, процессов в них, критериев и правил принятия решений чрезвычайно многообразны в настоящее время. Остановимся на классе когнитивных моделей [1–6] и вероятностных задачах принятия решений [7], инструментарий которых развивается в наших работах.

Разработанная методология когнитивного моделирования сложных систем согласуется с работами сотрудников ИПУ РАН в этом направлении (когнитивный анализ и управление ситуациями, работы Н.А.Абрамовой, В.В. Кульбы, В.И.Максимова и многих других), и в известной мере опирается на работы Р.Аксельрода, Дж.Касти, Ф.Робертса, Р.Эткина.

Разработанная методология опирается на теорию сложных систем, системный и когнитивный подходы, является системой моделей и методов решения задач идентификации объекта в виде когнитивных моделей разного уровня сложности, анализа структурных и поведенческих свойств системы на когнитивной модели, разработки возможных сценариев развития системы, выбора желаемых решений на основе оценки результатов развития процессов.

Если для конкретного объекта возможна интерпретация части происходящих в нем процессов как вероятностных, то для оценки его эффективности для лица, принимающего решения, может быть использована обобщенная функция эффективности оптимума номинала, развившая идеи Д.В. Свечарника [7]. В своё время моделирование задач об оптимуме номинала проводилось для решения задач оптимизации по полезности (чаще – экономической) различных производственных процессов (в металлургической промышленности, машиностроении, электроэнергетике, нефте – и газопереработке, обогащении руд цветных металлов и др.).

В простейшем случае модель задачи об оптимуме номинала имеет вид

$$\varphi(m_y) = \sum_{i=1}^{s} \int c_i f(y, m_y, \sigma_y) dy \qquad (1)$$

где $\varphi(m_y)$ – оценочная функция – функция эффективности оптимума номинала от математического ожидания случайной величины Y (показателя, фактора), $f(y, m_y, \sigma_y)$ – плотность распределения Y; m_y, σ_y – математическое ожидание и среднее квадратичное отклонение соответственно, c_i – полезность i–го интервала $[y_{\text{iH}}, y_{\text{iB}}]$ значений Y, $\int_{Y_{\text{iH}}}^{Y_{\text{iB}}} f(y, m_y, \sigma_y)$ – вероятность попадания y в i–й интервал значений.

Для производственного процесса $m_y = m_{уном}$ – номинальное значение целевого показателя Y. Задача об оптимуме номинала – это задача определения такой оптимальной для данного процесса в конкретных производственных условиях величины m_y^o, которая приведет, например, к минимизации производственных затрат или к максимизации прибыли и другого показателя,

т.е. к максимизации (минимизации) функции эффективности оптимума номинала

$$\max_{my} \varphi(m_y) = \varphi^o \tag{2}$$

заметим, что принятие решения по правилу (2) является достаточно широко распространенным правилом рационального выбора решения по критерию максимизации математического ожидания полезности.

В общем случае при непрерывной функции полезности $C(Y)$, многих показателях Y и зависимости параметров от времени t наилучшее решение выбирается по правилу максимизации обобщенной функции эффективности оптимума номинала

$$\varphi^o \max_{M_h} \varphi(M_h, t) = \iint\limits_S \ldots \int C(y) f(y, M_h, t) dcdydt \tag{3}$$

при наличии ограничений на обобщенную функцию эффективности в виде условий управления моментами M_h распределения $f(y, M_h)$

$$M_h = \varphi(X, t), h = 1, 2, 3, 4 \tag{4}$$

и ограничений на управляющие факторы $X \in X_{\text{доп}}$ и результативные $Y \in Y_{\text{доп}}$.

Использование когнитивных моделей для представления сложных систем и определение возможного их поведения путем импульсного моделирования на когнитивных моделях [1,3,4] позволяет получить достаточное количество реализаций импульсных процессов. Наличие таких реализаций позволяет поставить и решить задачу оптимума номинала по выбору "лучшего" импульсного процесса, который в дальнейшем может быть принят в качестве желаемого сценария развития исследуемого объекта.

Литература. **1.** Максимов В.И. Когнитивные технологии – от незнания к пониманию. // Сб. трудов 1-й Международной конференции "Когнитивный анализ и управление развитием ситуаций", (CASC'2001) – М.: ИПУ РАН, 2001. – т.1, – С.4–18. **2.** Кульба В.В., Кононов Д.А., Ковалевский С.С., Косяченко С.А., Нижегородцев Р.М., Чернов И.В. Сценарный анализ динамики поведения социально-экономических систем (Научное издание). – М.: ИПУ РАН, 2002. – 122 с. **3.** Горелова Г.В., Захарова Е.Н., Радченко С.Н. Исследование слабоструктурированных проблем социально-экономических систем: когнитивный подход. – Ростов н/Дону: Изд-во РГУ, 2006. – 332 с. **4.** Инновационное развитие социоэкономических систем на основе методологий предвидения и когнитивного моделирования. / Под ред. Г.В.Гореловой, Панкратовой Н.Д. – Киев: Изд-во "Наукова Думка", 2015. – 464 с. **5.** Gorelova G.V. Cognitive modeling as the instrument in course of knowledge of big system // International Journal "Information Theories and Applications", Bulgaria, Vol.18, Number 2, 2011. – P.172–182. **6.** Горелова Г.В., Верба В.А., Захарова Е.Н. Когнитивные модели в интеллектуальных системах поддержки управленческих решений. // Известия ТРТУ, тематический выпуск "Проектирование и моделирование интеллектуальных систем". Таганрог: изд-во ТРТУ, 2004, – №5. – С.35–39. **7.** Горелова Г.В., Свечарник Д.В., Здор В.В.Метод оптимума номинала и его применения М.: Изд-во "Энергия", 1970.

Городецкий В.Г., Осадчук Н.П.
Национальный технический университет Украины "Киевский политехнический институт", Киев, Украина

Моделирование активности нейрона по временным рядам

В наше время известны модели активности нейрона в виде системы обыкновенных дифференциальных уравнений (ОДУ), которые способны воспроизводить как периодические [1,2], так и хаотические [3] колебания. Поскольку во многих случаях при исследовании нейрона для измерения и записи изменений во времени доступна только одна физическая величина — электрический потенциал, то целесообразно определять параметры модели нейрона используя только одну наблюдаемую переменную.

В качестве источника данных была рассмотрена предложенная в [3] система ОДУ третьего порядка, имеющая вид

$$\begin{cases} \dot{x}_1 = a_0 + a_2 x_2 + a_3 x_3 + a_4 x_1^2 + a_{10} x_1^3, \\ \dot{x}_2 = b_0 + b_2 x_2 + b_4 x_1^2, \\ \dot{x}_3 = c_0 + c_1 x_1 + c_3 x_3, \end{cases} \tag{1}$$

где $a_0 = 3$, $a_2 = 1$, $a_3 = -1$, $a_4 = 3$, $a_{10} = -1$, $b_0 = 1$, $b_2 = -1$, $b_4 = -5$, $c_0 = 0.096$, $c_1 = 0.06$, $c_3 = 0.015$. Система (1) была проинтегрирована методом Рунге-Кутта 4 порядка с шагом дискретизации 0.025 с, и для дальнейшего исследования использовался временной ряд переменной x_1, содержащий 20000 точек.

Для определения параметров модели нейрона был использован подход, предложенный в [4]. Вместо того, чтобы из временного ряда напрямую определять коэффициенты оригинальной системы (ОС) (1), можно определить коэффициенты стандартной системы (СС), одна из переменных которой совпадает с наблюдаемой переменной ОС, и коэффициенты которой могут быть выражены через коэффициенты ОС. После этого, используя числовые значения коэффициентов СС, можно аналитически найти числовые значения коэффициентов ОС. СС, соответствующая ОС (1), имеет вид

$$\begin{cases} \dot{y}_1 = y_2, \\ \dot{y}_2 = y_3, \\ \dot{y}_3 = N_0 + N_1 y_1 + N_2 y_2 + N_3 y_3 + N_4 y_1^2 + N_5 y_1 y_2 + N_6 y_1 y_3 + N_7 y_2^2 + \\ \quad + N_{10} y_1^3 + N_{11} y_1^2 y_2 + N_{12} y_1^2 y_3 + N_{13} y_1 y_2^2, \end{cases} \tag{2}$$

где $y_1(t) \equiv x_1(t)$. Первые два уравнения СС (2) позволяют заменить недостающие переменные производными наблюдаемой переменной [5]. Последнее уравнение СС содержит полиномиальную стандартную функцию (СФ), коэффициенты которой могут быть определены численными методами [6,7] из временного ряда.

В результате применения численного метода были получены значения коэффициентов СС: $N_0 = -0.036$, $N_1 = -0.06$, $N_2 = -0.075$, $N_3 = -1.015$, $N_4 = -0.03$, $N_5 = -3.91$, $N_6 = N_7 = 6$, $N_{10} = -1.015$, $N_{11} = -3.045$, $N_{12} = -3$, $N_{13} = -6$. На основе числовых значений коэффициентов СС и аналитических соотношений между коэффициентами ОС и СС были однозначно определены значения коэффициентов ОС a_4, a_{10}, b_2, c_3. Также были определены произведения коэффициентов $a_2 b_4$ и $a_3 c_1$. Если в каждом произведении один из коэффициентов считать известным (например, равным единице), то можно определить еще два коэффициента ОС. Кроме того оказалось достаточным, чтобы только один из коэффициентов a_0, b_0, c_0 был ненулевым, а остальные два могут равняться нулю. При этом числовые значения коэффициентов СС не изменяются, следовательно такая упрощенная система ОДУ способна воспроизводить исследуемый временной ряд.

Используя рассмотренный выше подход можно найти модели активности нейрона, уравнения которых не являются частными случаями уравнений модели (1). Для этого рассмотрим систему

ОДУ

$$\begin{cases} \dot{x}_1 = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_1^2 + a_5x_1x_2 + a_6x_1x_3 + a_{10}x_1^3 + a_{11}x_1^2x_2 + a_{12}x_1^2x_3, \\ \dot{x}_2 = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1^2 + b_5x_1x_2 + b_6x_1x_3 + b_{10}x_1^3 + b_{11}x_1^2x_2 + b_{12}x_1^2x_3, \\ \dot{x}_3 = c_0 + c_1x_1 + c_2x_2 + c_3x_3 + c_4x_1^2 + c_5x_1x_2 + c_6x_1x_3 + c_{10}x_1^3 + c_{11}x_1^2x_2 + c_{12}x_1^2x_3, \end{cases} \quad (3)$$

где x_1 — наблюдаемая переменная, и некоторые из коэффициентов a_i, b_i, c_i могут быть равны нулю. Аналитически было установлено, что системе (3) соответствует СС с дробно-рациональной СФ, в числителе которой содержатся 63 коэффициента, а в знаменателе — 10 коэффициентов. Если в (3) обнулить часть коэффициентов, то ей будет соответствовать СС (2) и наблюдаемая переменная будет совпадать с наблюдаемой (1). Таким образом были получены три системы ОДУ, содержащие от 17 до 19 коэффициентов в правых частях уравнений, причем одна из систем включает в себя (1) как частный случай (в двух других системах отсутствует коэффициент a_3). Дальнейшее упрощение альтернативных систем производилось с использованием аналитических соотношений и числовых значений коэффициентов СС (2). Так были найдены примеры систем ОДУ, способных воспроизводить исследуемый временной ряд, но отличающиеся по структуре от (1):

$$\begin{cases} \dot{x}_1 = a_0 + a_2x_2, \\ \dot{x}_2 = b_1x_1 + b_3x_3 + b_4x_1^2 + b_5x_1x_2 + b_{11}x_1^2x_2, \\ \dot{x}_3 = c_2x_2 + c_3x_3 + c_5x_1x_2 + c_{10}x_1^3, \end{cases} \quad (4)$$

$$\begin{cases} \dot{x}_1 = a_0 + a_2x_2 + a_4x_1^2 + a_{10}x_1^3, \\ \dot{x}_2 = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1^2, \\ \dot{x}_3 = c_1x_1 + c_3x_3. \end{cases} \quad (5)$$

Далее для систем (4), (5) были вычислены значения коэффициентов при которых временной ряд переменной x_1 совпадает с исследуемым временным рядом. При этом для системы (4) существуют два разных набора значений коэффициентов, позволяющих получить один и тот же временной ряд.

Рассмотренные примеры определения коэффициентов и поиска альтернативных систем ОДУ для модели нейрона являются частным случаем задачи реконструкции системы ОДУ по одной наблюдаемой переменной. Предложенный авторами подход сочетает аналитические и численные методы и позволяет не только реконструировать систему ОДУ, но и упростить ее или получить альтернативные системы ОДУ. В качестве исходных данных используются значения коэффициентов СС, которые могут быть получены численно (если для исследования доступен только временной ряд) или аналитически (если известна ОС исследуемого процесса). В последнем случае возможно, что уравнения ОС неточно отражают физику исследуемого процесса, а альтернативные системы ОДУ могут служить основанием для правильного понимания протекания процесса.

Литература. **1.** R. Fitzhugh, Impulses and physiological states in models of nerve membrane, Biophys. J. 1 (1961) 445–466. **2.** J. S. Nagumo, S. Arimoto, S. Yoshisava, An active pulse transmission line simulating a nerve axon, Proc. IRE 50 (1962) 2061–2070. **3.** J. L. Hindmarsh, R. M. Rose, A model of neuronal bursting using three coupled first order differential equations. Proc. R. Soc. Lond. B 221 (1984) 87–102. **4.** G. Gouesbet, Reconstruction of standard and inverse vector fields equivalent to the Rössler system, Phys. Rev. A 44 (1991) 6264–6280. **5.** F. Takens, Detecting strange attractors in turbulence, in: D.A. Rand, L.S. Young (Eds.), Dynamical System and Turbulence, volume 898 of Lecture Notes in Mathematics, Springer, New York, 1981, pp. 366–381. **6.** G. Gouesbet, Reconstruction of the vector fields of continuous dynamical systems from numerical scalar time series, Phys. Rev. A. 43 (1991) 5321–5331. **7.** В. Г. Городецкий, Н. П. Осадчук, Алгоритм реконструкции динамических систем по одной наблюдаемой переменной, Кибернетика и вычислительная техника. 179 (2015) 56–69.

***Губарев В.Ф.*[1], *Дяденко О.Н.*[2]**
[1]*Института космических исследований НАН Украины, Киев, Украина;* [2]*Институт прикладного системного анализа НТУУ "КПИ"*

Анализ наблюдаемости бортового измерительного космического аппарата

При построении прецизионных систем управления ориентацией космических аппаратов (КА) очень важно выбрать такой комплект измерительных приборов, который с одной стороны не был бы избыточным, а с другой – позволял бы оценивать все параметры ориентации с требуемой точностью.

В представленной работе проведен анализ наблюдаемости наиболее распространенных бортовых измерительных приборов, а именно - магнитометра, построителя местной вертикали, астродатчика и датчика угловых скоростей с целью определения минимально необходимого набора измерительных приборов, который обеспечивал бы наблюдаемость космического аппарата.Оценка пригодности и качество различных по составу измерительных комплексов проводилась на основе существующих методов теории наблюдаемости динамических систем с учетом уравнений наблюдения и динамики углового движения КА.

Рассматривается ориентация связанной с осями КА и началом в его центре масс системы координат (ССК) относительно орбитальной системы координат (ОСК связана с орбитальным движением КА) [1]. Под параметрами ориентации КА подрузмеваются позиционные параметры - компоненты нормированных кватернионов $\Lambda = (\lambda_0, \lambda)^T$, где $\lambda = (\lambda_1, \lambda_2, \lambda_3)^T$, а также скоростные $\omega = (\omega_1, \omega_2, \omega_3)^T$ - проекции абсолютной угловой скорости КА на оси ССК. Тогда уравнения углового движения КА принимают вид:

$$\Lambda' = \frac{1}{2} A(\omega, \omega_*)\Lambda, \quad A(\omega, \omega_*) = \begin{bmatrix} 0 & -(\omega - \omega_*)^T \\ (\omega - \omega_*) & -\Omega(\omega + \omega_*) \end{bmatrix}$$

$$J\omega' = M - \Omega(\omega)J\omega, \quad \Omega(z) = \begin{bmatrix} 0 & -z_3 & z_2 \\ z_3 & 0 & -z_1 \\ -z_2 & z_1 & 0 \end{bmatrix}$$

где $\omega_* = (0, 0, v)^T$ - вектор угловой скорости, который определяет вращательное движение ССК, v - текущее значение истинной аномалии КА, J - симметричная положительно определенная матрица момента инерции обьекта, M - суммарная матрица управляющего и возмущающего моментов.

Описанные выше уравнения углового движения КА и некоторые уравнения измерительных приборов являются нелинейными. Как известно, получение общих глобальных условий наблюдаемости таких систем представляет сложную практически не решенную на сегодняшний день задачу. Поэтому в данной работе реализован подход на основе приближенного анализа с использованием процедур линеаризации и локальных условий наблюдаемости при вычислении производных по времени [2].

Исследования показали, что наиболее эффективная наблюдаемость обеспечивается астроизмерительной системой. Для неё характерно близкое к абсолютному минимуму (единица) число обусловленности. Более того, на её основе может быть построен простейший оцениватель состояния, который описан в работе.

Литература. **1.** Волосов В.В., Хлебников М.В., Шевченко В.Н. Алгоритм прецизионного управления ориентацией космического аппарата при действии неконтролируемого возмущения // Проблемы управления и информатики. – 2011. – №2. – с.114-121. **2.** Справочник по теории автоматического управления // Под ред. А.А. Красовского. – М.:Наука, 1987. – 712 с.

Даник Ю.Г., Писарчук А.А., Дюков И.Н., Тимчук С.В.
Житомирский военный институт имени С.П. Королёва, Житомир, Украина

Самоорганизующаяся математическая модель ситуационного управления структурой и параметрами систем мониторинга

Развитие информационных технологий обусловило создания систем мониторинга (СМ), функционирующих в едином информационном пространстве и обеспечивающих выявление, локализацию и ликвидацию кризисных ситуаций (КС) различной прикладной направленности. Единая информационная среда породила актуальную проблематику эффективного функционирования СМ в условиях значительной динамики и плотности потока КС. Традиционные подходы в архитектуре и функционировании СМ опираются на ее статичность с обеспечением открытости, модульности и унификации. При этом возрастание динамики изменения и плотности потока КС компенсируется включением новых элементов в систему. Такой путь является малоэффективным, поскольку приводит к усложнению системы, повышает ее стоимость, инертность, снижая уровень отказоустойчивости и т.д.

В докладе рассматривается самоорганизующаяся [1] математическая модель ситуационного управления структурой и параметрами СМ, обеспечивающая эффективное ее функционирования в условиях значительной динамики и плотности потока КС.

Модель формируется как решение многокритериальной задачи фоомирования информационно-управляющего кластера (ИУК) с индивидуальными особенностями под конкретную КС из элементов известной избыточной структуры СМ.

Формирование ИУК реализуется по восьми группам критериев эффективности.

1. *Множество критериев технических характеристик* подсистем СМ:
$$TX(N_{KS},N)=\begin{cases}TS_{i_{TS}}(N_{KS},N)\to max;SNO_{i_{SNO}}(N_{KS},N)\to max;\\ STO_{i_{STO}}(N_{KS},N)\to max;STK_{i_{STK}}(N_{KS},N)\to max.\end{cases}\quad(1)$$
$TS_{i_{TS}}(N_{KS},N), i_{TS}=1...I_{TS}; SNO_{i_{SNO}}(N_{KS},N), i_{SNO}=1...I_{SNO}; STO_{i_{STO}}(N_{KS},N)$, $i_{STO}=1...I_{STO}; STK_{i_{STK}}(N_{KS},N), i_{STK}=1...I_{STK}$ — технические характеристики элементов технических средств мониторинга (ТСМ), подсистемы накопления и обработки информации мониторинга (СНОИМ), специального технического оборудования (СТО), подсистемы телекоммуникации (СТК); N_{KS} — классификационный признак КС, N — количество i-х элементов подсистем ИУК.
2. *Критерии функциональной направленности* подсистем СМ:
$$FN(N_{KS},N)=\begin{cases}TS_{i_{TS}^{FN}}(N_{KS},N)\to max;SNO_{i_{SNO}^{FN}}(N_{KS},N)\to max;\\ STO_{i_{STO}^{FN}}(N_{KS},N)\to max;STK_{i_{STK}^{FN}}(N_{KS},N)\to max.\end{cases}\quad(2)$$
$TS_{i_{TS}^{FN}}(N_{KS},N), SNO_{i_{SNO}^{FN}}(N_{KS},N), STO_{i_{STO}^{FN}}(N_{KS},N), STK_{i_{STK}^{FN}}(N_{KS},N)$ — кодовые комбинации, характеризующие функциональную направленность подсистем СМ.
3. *Структурные критерии функциональной устойчивости* с избыточностью элементов:
$$FN_N(t,N)=\begin{cases}FN_{TS}(t,N)\to max;FN_{SNO}(t,N)\to max;\\ FN_{STO}(t,N)\to max;FN_{STK}(t,N)\to max.\end{cases}\quad(3)$$
t – временной параметр, характеризующий возникновение, модификацию и развитие КС.
4. *Вероятностные критерии функциональной устойчивости*, отражающие наличие межэлементных горизонтально-вертикальных связей в ИУК:
$$P_N(t,N)=\begin{cases}P_{TS,SNO}(t,N)\to max;P_{TS,STO}(t,N)\to max;\\ P_{TS,STK}(t,N)\to max;P_{SNO,STO}(t,N)\to max;\\ P_{SNO,STK}(t,N)\to max;P_{STO,STK}(t,N)\to max.\end{cases}\quad(4)$$
5. *Критерий эффективности эргатической составляющей* СМ:
$$P_{OPi}(t,N)=K_{OPi}(t,N)P_{\Phi i}(t,N)P_{\psi i}(t,N)\to max.\quad(5)$$

$P_{OPi}(t,N)$ — вероятность безотказной работы i -го оператора в течении времени t, $K_{OPi}(t,N)$ — коэффициент готовности оператора, $P_{\Phi i}(t,N), P_{\psi i}(t,N)$ — его биологическая и психофизиологическая надежности.

6. *Критерии, характеризующие условия среды работы оператора*:

$$SR(t,N) = \left\{ SR_{PF}(t,N) \to min; SR_{VS}(t,N) \to min; SR_{KC}(t,N) \to min. \right. \quad (6)$$

$SR_{PF}(t,N)$ — оценка негативного действия природных факторов; $SR_{VS}(t,N)$ — уровень напряженности внешней психофизической ситуации; $SR_{KC}(t,N)$ — уровень напряженности, отрабатываемой КС.

7. *Критерии автоматизации работы оператора*:

$$AR(t,N) = \left(\frac{N_{SH}(t,N)}{SN_{SH}(t,N)} + \frac{N_{PG}(t,N)}{SN_{PG}(t,N)} + \frac{N_{KR}(t,N)}{SN_{KR}(t,N)} + \frac{N_{PK}(t,N)}{SN_{PK}(t,N)} \right) \to min, \quad (7)$$

характеризуются отношением количества функций, выполняемых оператора в штатном $N_{SH}(t,N)$, повышенной готовности $N_{PG}(t,N)$, кризисном $N_{KR}(t,N)$ и посткризисном $N_{PK}(t,N)$ режимах функционирования СМ, соотнесенных к общему количеству функций, необходимых для выполнения системой — $SN_{SH}(t,N)$, $SN_{PG}(t,N)$, $SN_{KR}(t,N)$, $SN_{PK}(t,N)$.

8. *Пропускная способность ИУК* с ее минимизацией (максимизация для СМ $PSp(N)$):

$$PSp_{iUIK}(N) = 1 - PSp(N)^{-1} \to min. \quad (8)$$

Самоорганизующаяся математическая модели формируется путем агрегации частных критериев (1) – (8) с использованием свертки по нелинейной компромиссов проф. Воронина А.Н. [2] с применением технология вложенных от элемента до ИУК. С учетом изложенного, получим критерии для подсистем ИУК: ТСМ — $F_{iTS0}(t,N_{KS},N)$, СНОИМ — $F_{iSNO0}(t,N_{KS},N)$, СТО — $F_{iSTO0}(t,N_{KS},N)$, СТК – $F_{iSTK0}(t,N_{KS},N)$:

$$\begin{gathered} F_i(t,N_{K,S},N) = \gamma_{i10}(1 - TX_{i0}(N_{KS},N))^{-1} + \gamma_{i20}(1 - FN_{i0}(N_{KS},N))^{-1} + \\ + \gamma_{i30}(1 - FN_{iN0}(t,N))^{-1} + \gamma_{i40}(1 - P_{iOP0}(t,N))^{-1} + \gamma_{i50}(1 - IPO_{i0}(t,N))^{-1} \to min, \\ IPO_i(t,N) = \gamma_{i10}(1 - SR_{iPF0}(t,N))^{-1} + \gamma_{i20}(1 - SR_{iVS0}(t,N))^{-1} + \\ + \gamma_{i30}(1 - SR_{iKS0}(t,N))^{-1} + \gamma_{i40}(1 - AR^{S}_{iOP0}(t,N))^{-1} \to min. \end{gathered} \quad (9)$$

Из (9) формируется интегрированный критерий эффективности ситуационного синтеза:

$$\begin{gathered} \Omega_{iIUK}(t,N_{K,S},N) = \gamma_{i10}(1 - F_{iTS0}(t,N_{KS},N))^{-1} + \gamma_{i20}(1 - F_{iSNO0}(t,N_{KS},N))^{-1} + \\ + \gamma_{i30}(1 - F_{iSTO0}(t,N_{KS},N))^{-1} + \gamma_{i40}(1 - F_{iSTK0}(t,N_{KS},N))^{-1} + \\ + \gamma_{i50}(1 - P_{iN0}(t,N))^{-1} + \gamma_{i60}(1 - PSp_{iUIK}(N))^{-1} \to min. \end{gathered} \quad (10)$$

Самоорганизация модели (10) проявляется во времени и заключается в последовательном приближении структуры ИУК к аттрактору – известному образу кластера для конкретной КС, что реализуется с контролем выполнения условия:

$$\frac{\Omega_{iIUK}(t,N_{KS},N)}{\Omega^{KSi}_{IUK}(t,N_{KS},N)} \to min. \quad (11)$$

Сформированная таким образом оптимизационная модель реализует ситуационный принцип управления структурой и параметрами СМ с использованием принципов самоорганизации. В модели учтены технические характеристики и функциональные возможности, свойства эргатической составляющей и распределенная структура СМ.

Литература. **1.** Ивахненко А.Г. Принятие решений на основе самоорганизации / А.Г. Ивахненко, Ю.П. Зайченко, В.Д. Дмитриев. — М. : Сов. радио, 1979. — 280 с. **2.** Воронин А.Н. Сложные технические и эргатические системы: метод использования / А.Н. Воронин, Ю.К. Зиатдинов, А.В. Харченко, В.В. Осташевский. — Харьков : Факт, 1997. — 240с.

Даник Ю.Г., Писарчук О.О., Дюков І.М., Тимчук С.В.
Житомирський військовий інститут імені С.П. Корольова, Житомир, Україна

Математичне забезпечення автоматизованої системи збору та обробки інформації від технічних засобів моніторингу

На теперішній час у будь-якій сфері діяльності людини широкого поширення набули сучасні автоматизовані системи управління (АСУ) різного призначення. Одним із ключових компонентів будь-якої АСУ є система збору і обробки інформації (СЗОІ) моніторингу. СЗОІ моніторингу забезпечує пункти управління різного рівня попередньо обробленою узагальненою та адаптованою інформацією про поточну обстановку через спостереження за інформаційними джерелами (ІД), тим самим надаючи основу для прийняття рішення та подання відповідних команд на об'єкти управління(ОУ).

Метою роботи є вироблення напрямів щодо ситуативного синтезу структури втоматизованої СЗОІ моніторингу в умовах значної інформаційної надмірності, високої щільності потоку конфліктних ситуацій (КС) та динаміки зміни поточної обстановки.

Першопричиною формування інформаційної моделі та ситуаційного конфігурування СЗОІ є КС, опис якої реалізується певною множиною (формуляром). Формуляр формується наступною трійкою: P_{ks} — ознака КС з унікальною для кожного її типу літерно-цифровою комбінацією; T_{ks} — множина частинних завдань системи з усунення КС, що виникла (формується, виходячи із завдань системи в цілому й завдань окремих складових); I_{ks} — множина інформаційних потреб системи з усунення КС. Тоді формуляр КС задаємо множиною

$$KS_i = \left\{P_{ksi}, T^{KS}_{ksj}, I^{KS}_{ksf}\right\}, i = 1...I, j = 1...J, f = 1...F. \tag{1}$$

У свою чергу, множина частинних завдань з усунення КС і перелік інформаційних потреб для i-ї КС можна подати у вигляді підмножин:

$$\begin{aligned} T^{KS}_{ksj} &= \{T_{ksi1}, T_{ksi2}, T_{ksi3}, T_{ksil}, ..., T_{ksiL_i}\}, l = 1...L_i; \\ I^{KS}_{ksf} &= \{I_{ksi1}, I_{ksi2}, I_{ksi3}, I_{ksik}, ..., I_{ksiK_i}\}, k = 1...K_i. \end{aligned} \tag{2}$$

Опис системи обробки інформації реалізується для кожного її окремого елемента (автоматизованого робочого місця (АРМ) або технічного засобу моніторингу (ТЗМ)). Кожне АРМ описується формуляром, що містить перелік завдань та інформаційних потреб:

$$ES_j = \left\{T^{ES}_{ksj}, I^{ES}_{ksj}(I^{ID}_{ksf})\right\}. \tag{3}$$

Функціональне позначення $I^{ES}_{ksj}(I^{ID}_{ksf})$ характеризує взаємозв'язок інформаційних потреб j-го АРМ, забезпечуваних інформаційними можливостями f-го інформаційного джерела (ІД).

Опис ТЗМ характеризується множиною пар (формуляром ТЗМ): I^{ID}_{ks} — перелік інформаційних можливостей ТЗМ; TX — перелік технічних характеристик (ТХ) ТЗМ —

$$ID_f = \left\{I^{ID}_{ksf}, TX_f\right\}. \tag{4}$$

Таким чином, на понятійному рівні здійснено формальний математичний опис складових СЗОІ.

Побудова адекватної моделі функціонування СЗОІ вимагає опису не лише її структури, а й врахування взаємодії системи із зовнішнім середовищем (ІД та користувачами інформації). Для опису взаємодії СЗОІ з ІД та користувачами інформації пропонуємо використати багатошаровий граф [1]. Багатошарова фрактально-графова структура з урахуванням результатів формування передфракталу, що описує СЗОІ, складатиметься з трьох шарів (підграфів) (рис. 1): G^1 — підграф множини ІД на об'єкті моніторингу (ОМ), G^2 — підграф множини ТЗМ та множини АРМ системи обробки інформації, G^3 — підграф споживачів інфор-

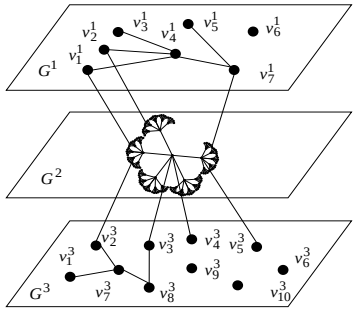


Рис. 1. Багатошарова фрактально-графова модель СЗОІ

мації. Результатом є багатошарова фрактально-графова інформаційна модель. Для графа G^1 вершинами v_k^1 є ІД, які об'єднані ребрами r_k^1 та утворюють дерева, що характеризують ОМ, їх ієрархічну структуру. У графі G^2 вершини – ТЗМ та наявні АРМ системи обробки інформації, а інформаційні потоки між ними показують ребра графа. Вершинами графа G^3 є користувачі інформації, які з'єднуються ребрами відповідно до їх ієрархічних зв'язків. Ребра, що сполучають вершини з різних шарів, відображають внутрішні та зовнішні інформаційні зв'язки СЗОІ через ребра відповідних графів. Застосування фрактально-графових операцій для отримання багатошарової математичної інформаційної моделі СЗОІ в запропонований спосіб та опис процесів, що відбуваються в системі, відповідними моделями дадуть фрактальну графодинамічну модель для опису динаміки зміни процесів у системі.

Побудована таким чином інформаційна фрактальна графодинамічна модель забезпечує формування надмірної структури СЗОІ для заданого рівня ієрархії. В отриманій структурі реалізується в подальшому ситуативний динамічний синтез системи – процес ситуативного управління структурою і параметрами СЗОІ.

Ситуативний структурно-параметричний синтез СЗОІ відповідно до КС реалізується із визначенням кількісної та якісної структури системи. Кількісний склад системи визначається за оптимізаційною моделлю, отриманою шляхом зведення вектора суперечливих критеріїв до єдиного функціонала за нелінійною схемою компромісів як розв'язок багатокритеріальної оптимізаційної задачі:

$$F(N_{ARM}) = (1 - t_{ks0}(N_{ARM}))^{-1} + (1 - D_{ks0}(N_{ARM}))^{-1} + (1 - ID_{ks0}(N_{ARM}))^{-1} \to min, \quad (5)$$

де залежними від кількості АРМ N_{ARM}, нормованими до мінімізованої форми, частинними критеріями є:

$t_{ks0}(N_{ARM})$ — час, витрачений на усунення КС;

$D_{ks0}(N_{ARM})$ — достовірність рішень на усунення КС;

$ID_{ks0}(N_{ARM})$ — характеристика інформаційної надмірності синтезованої системи.

Кількість АРМ визначає також кількість інших складових СЗОІ через міжрівневі зв'язки її графодинамічної фрактальної інформаційної моделі. Якісний склад системи встановлюють, виходячи з вимог найкращого відображення в її структурі КС, що виникла. При цьому оптимальною буде структура, яка забезпечує:

- виконання найбільшої кількості функцій АРМ із заданих формуляром КС $T_{Sj} \to max$;
- найбільшу кількість використовуваних для усунення КС ІД $I_{Sj} \to max$ з найкращими ТХ $TX_{Sj} \to max$.

Операцію конфігурування реалізовують відповідно до оптимізаційної моделі

$$\Psi_j = GT_{j0}(1 - T_{Sj0})^{-1} + GI_{j0}(1 - I_{Sj0})^{-1} + GTX_{j0}(1 - TX_{Sj0})^{-1} \to min. \quad (6)$$

За вагові коефіцієнти використовують значення GT_j, GI_j, GTX_j, що відображають значущість АРМ за параметрами T_{Sj}, I_{Sj}, TX_{Sj}.

Для виконання завдань обробки інформації слід використовувати сукупність фактів, даних, суджень про ОМ, потік неформалізованої інформації про поточну ситуацію. Завдання обробки інформації у СЗОІ моніторингу зводяться до трьох етапів:

- спільна траєкторна обробка вимірів від різнотипних ТЗМ;
- ідентифікація КС та ОМ;
- формування аналітичного висновку про поточну ситуацію та прийняття рішення на формування команд управління виконавчим елементам.

Зазначені підходи забезпечують реалізацію процесів ситуативного синтезу структури та обробки даних у СЗОІ моніторингу в умовах значної інформаційної надмірності.

Література. 1. Агеев Д.В. Методика описания структуры современных телекоммуникационных систем с использованием многослойных графов / Д.В. Агеев // Восточно-Европейский журнал передовых технологий. — 2010. — № 6/4 (48). — С. 56–59.

Денисенко О.І.
Запорізький національний технічний університет, Запоріжжя, Україна

Тривимірне моделювання елементів теплообінних приладів з турбулізаторами

Проблеми енергозбереження набули особливої актуальності останнім часом у зв'язку з дефіцитом та подорожчанням енергоресурсів. Ефективність використання енергоресурсів для опалювання приміщень у великій мірі залежить від конструктивних особливостей опалювальних котлів.

Зазвичай, конструкція опалювального котла являє собою ємність з теплоносієм та каналами, через які походять продукти горіння з високою температурою. З метою підвищення ККД котлів в каналах монтуються турбулізатори у вигляді системи пластин або спіральних пластинчатих вставок. Існуючі балансові, двовимірні або осесиметричні моделі таких каналів не дозволяють дослідити складну структуру потоку та оцінити ефективність теплообміну.

В роботі побудована тривимірна математична модель процесу теплообміну в каналах теплообмінника з пластинчастими спіральними вставками. Для моделювання руху газів використовувалася система диференціальних рівнянь Навьє-Стокса. Для опису температурного режиму металевих каналів і вставок використовувалося рівняння теплопровідності. Розглядалася сполучена постановка задачі. Температура теплоносія зовні каналу підтримувалася постіною. Проходження продуктів горіння через канали з турбулізаторами відбувалось природнім шляхом без примусової витяжки. Розрахунок здійснювався методом кінцевих елементів за допомогою програмного комплексу COMSOL. Проведена серія чисельних експериментів для різної геометрії вставок та кількості витків спіралі. Розрахунки проведені для широкого діапазону можливих значень перепаду тиску в системі газовідводу. Ускладнення геометрії турбулізаторів призводить до підвищення ефективності теплообміну, але, з іншого боку, призводить до підвищення аеродинамічного опору в системі газовідводу. Визначені умови, коли надлишкова щільність турбулізаторів шкодить евакуації продуктів горіння, що може привести до отруєння чадним газом.

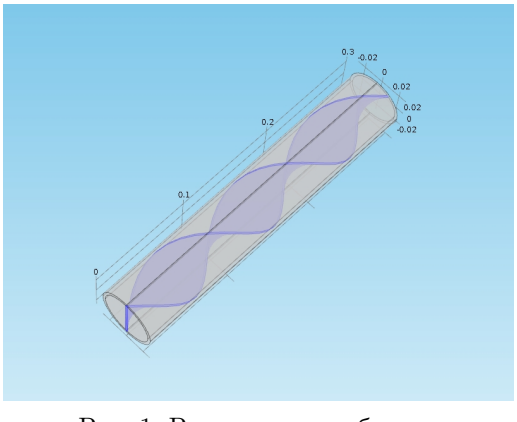


Рис. 1. Розрахункова область

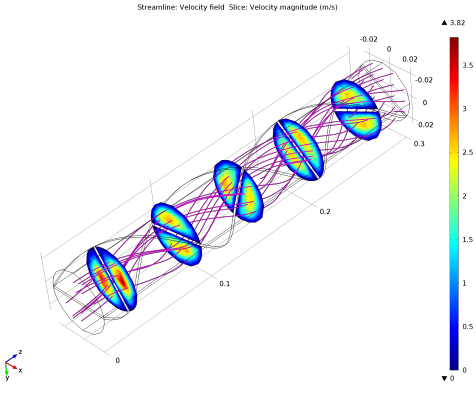


Рис. 2. Розподіл швидкостей та лінії тока

На рис.1 зображена геометрія розрахункової області. На рис.2 наведено приклад розрахунку розподілу швидкостей та ліній току в каналі зі скрученою пластиною. Спостерігається трансформація профіля швидкостей вздовж канала. Так, в кінцевій частині канала поле швидкостей в перетині, за рахунок турбулентного перемішування, має більш рівномірний розподіл, ніж на початку.

Результати роботи дозволяють прогнозувати поведінку теплообмінних елементів в залежності від геометричних параметрів та зовнішніх умов, на стадії проектування визначати ефективність опалювального обладнання та обирати оптимальні конструктивні параметри.

Дмитрієнко В.Д., Заковоротний О.Ю., Мезенцев М.В.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Оптимізація тягового електроприводу за допомогою геометричної теорії керування

Одним з стратегічних напрямів державної політики України в галузі розвитку сучасних видів транспорту є залучення до їх виробництва вітчизняних виробників, що забезпечує зменшення імпортної залежності Укрзалізниці від закордонних поставок рухомого складу. При цьому особливу увагу приділено, з одного боку, впровадженню сучасних енергозберігаючих технологій, а з іншого боку, раціональному використанню вже наявних енергоресурсів. Останнє вирішується, як шляхом вдосконалення вже існуючих систем автоматичного керування рухомим складом в напрямку оптимізації режимів ведення поїздів по залізничним перегонам при яких дотримується заданий графік руху й мінімальні витрати енергоресурсів, так й оптимізацією роботи окремих складових частин рухомого складу. Для України ці питання мають істотне значення, оскільки витрата енергоносіїв на перевезення пасажирів та вантажів в нашій країні вище, ніж в більшості розвинених країн світу.

В наш час питаннями створення систем керування залізничним транспортом займається безліч фахівців, роботи яких засновані на теорії керування та теорії оптимальних систем керування, а також методів функцій Ляпунова, принципу максимуму Понтрягіна, класичного варіаційного числення, термінальних керувань і т.д. Дані методи дозволяють виконувати синтез регуляторів для нелінійних об'єктів, однак вони мають й істотні недоліки. Їх використання для синтезу оптимальних систем керування тяговим рухомим складом ускладнено, особливо якщо мова йде про управління приводом змінного струму. Труднощі синтезу систем керування тяговими асинхронними приводами, які зазвичай описуються системами нелінійних диференціальних рівнянь вище п'ятого-шостого порядку з декількома керуваннями, привели до пошуку нових методів синтезу систем керування приводом, зокрема, єдиної геометричної теорії керування, перевагою якої є можливість перетворень нелінійних систем високого порядку з декількома керуваннями до еквівалентних лінійних, за рахунок розбиття вихідних моделей на ряд підсистем меншою розмірності, в кожну з яких входить тільки одне керування. Такі перетворення відкривають можливості для використання при вирішенні задач розробки нелінійних систем керування методів і засобів теорії лінійних систем. При цьому лінеаризація нелінійної системи виконується на основі використання лінійного зворотного зв'язку в просторі "вхід – стан", та включає в себе таку послідовність дій: формування векторних полів об'єктів по їх моделям; перевірка послідовності розподілів на виконання умов інволютивності; визначення індексу керованості для заданої системи керування, використовуючи теорію про лінійний еквівалент для нелінійної афінної системи з векторним керуванням й форми лінійного еквівалента; побудова системи диференціальних рівнянь, з якої шляхом послідовного диференціювання, уздовж відповідних векторних полів, визначаються функції переходу до канонічної формі Бруновського, пошук оптимальних керувань за допомогою систем лінійних рівнянь, та перехід до початкової нелінійної системи.

У доповіді розглянутий приклад лінеаризації математичної моделі руху дизель-поїзда на основі інволютивних розподілів геометричної теорії керування у просторі "вхід – стан", що описується системою нелінійних диференціальних рівнянь двадцятого порядку з чотирма керуваннями та враховує, з одного боку, основні види коливань вагонів рухомого складу та розподіл сил взаємодії між ними, а з іншого боку, паралельну роботу еквівалентних тягових асинхронних двигунів двох обмоторених вагонів. Приведено опис спеціально розроблених функцій для пакета Matlab, що дозволяють перетворювати нелінійні моделі об'єктів високого порядку у еквівалентні лінійні моделі за допомогою геометричної теорії керування. Приведені результати моделювання та оптимізації у пакеті Matlab нелінійної та лінійної моделей руху дизель-поїзда в різних режимах роботи, які підтверджують перспективність застосування геометричної теорії керування для синтезу систем керування тяговим приводом.

Донець Н.П., Ясинський В.К.
Чернівецький національний університет ім. Юрія Федьковича, Чернівці, Україна

Існування сильного розв'язку нелінійного стохастичного диференціально-функціонального рівняння з зовнішніми випадковими збуреннями

Постановка задачі. Нехай на імовірнісному базисі $(\Omega, \mathcal{F}, \mathbb{F} \equiv \{\mathcal{F}_t, T \geqslant t \geqslant 0\}, \mathbb{P})$ задано випадкові процеси $x(t) \equiv x(t,\omega) : [0,T] \times \Omega \to \mathbb{R}^n$; $x_t(\theta) \equiv x(t+\theta) \in D(\mathbb{R}^n)$; $\varphi(\theta) \in D(\mathbb{R}^n)$ [1], $\theta \in [-\tau, 0]$; вінерів процес $w(t) := w(t,\omega) : [0,T] \times \Omega \to \mathbb{R}$; $\widehat{v}(dt, du, \omega) :\equiv v(dt, du) - E\{v\}$ - центрована пуассонова міра [2]; $f_j(t,\omega) : [0,T] \times \mathbb{R} \to \mathbb{R}$ [3], $j = \overline{1,3}$, - неперервні обмежені випадкові процеси, що не залежать від подій $N_0(dw) \cup N_0(d\widehat{v}) \cup N_0(f_j)$, $j = 1,3$. $D \equiv D(\mathbb{R}^n)$ - простір Скорохода неперервних справа функцій, які мають лівосторонні границі. Для $\forall t \in [0,T]$ має місце інтегральне стохастичне диференціально-функціональне рівняння з зовнішніми випадковими збуреннями (СДФР ВЗ)

$$x(t) = \varphi(0) + \int_0^t f_1(s,\omega) a(s, x_s(\theta)) ds + \int_0^t f_2(s,\omega) b(s, x_s(\theta)) dw(s) + \tag{1}$$

$$+ \iint_0^t f_3(s,\omega) c(s, x_s(\theta), u) \widehat{v}(ds, du)$$

за початковою умовою

$$x(t)\,|_{t=0} = \varphi(\theta) \equiv \varphi(\theta,\omega) \in D(\mathbb{R}^n), \forall t \in [-\tau, T] \tag{2}$$

$a, b, c : [0,T] \times D \to \mathbb{R}^n$ - вимірні (неперервні) за сукупністю змінних функціонали; $x_t(\theta) = \{x(t+\theta), \theta \in [-\tau, 0]\} \subset D(\mathbb{R}^n)$.

Теорема. Нехай:

1. коефіцієнти рівняння (1) $a(t,\alpha)$, $b(t,\alpha)$ та $c(t,\alpha,u)$ неперервні за сукупністю змінних для будь-яких $\alpha \in D$, $t \in [0,T]$;
2. a, b, c, f_j задовольняють умову Ліпшиця

$$|a(t,\varphi) - a(t,\psi)|^2 + |b(t,\varphi) - b(t,\psi)|^2 + \int_U |c(t,\varphi,u) - c(t,\psi,u)|^2 \mathbf{\Pi}(du) \leqslant K \|\varphi - \psi\|^2,$$

$$|f_j(t,\varphi) - f_j(t,\psi)|^2 \leqslant K \|\varphi - \psi\|^2$$

для будь-яких $\varphi, \psi \in D$, $j = \overline{1,3}$.

3. зовнішні випадкові збурення $f_j(t,\omega)$, $j = \overline{1,3}$ неперервні за сукупністю змінних;
4. існують математичні сподівання для $\forall t \in [0,T] : E\{\|f_j(t,\omega)\|^4\} < \infty$, $j = \overline{1,3}$;
5. випадкові функції $a(t,\xi)$, $b(t,\xi)$, $c(t,\xi,u)$ не залежать від $f_j(t,\omega)$, $j = \overline{1,3}$, де $\xi \equiv \xi(\omega) \in \mathbb{R}^n$ - довільна випадкова величина з імовірнісного базису;
6. існують четверті моменти для початкових процесів

$$E\|\psi\|^4 < \infty.$$

Тоді для будь-якого $T > 0$ існує сильний розв'язок з точністю до стохастичної еквівалентності задачі (1), (2) на $[0,T]$.

Доведена теорема існування та єдиності сильного розв'язку СДФР з зовнішніми випадковими збуреннями методом стислих відображень.

Література. **1.** Хейл Дж. Теория функционально-дифференциальных уравнений. – М.: Мир, 1984. – 421 с. **2.** Ясинський В.К., Ясинський Є.В., Юрченко І.В. Стабілізація у динамічних системах випадкової структури. – Чернівці: Золоті литаври, 2011. – 738 с. **3.** Царьков Е. Ф. Случайные возмущения дифференциально-функциональных уравнений. – Рига: Зинатне, 1989. – 421 с.

Донченко В.С., Зінько Т.П., Скотаренко Ф.М.
Київський національний університет імені Тараса Шевченка, Київ, Україна

Базові структури евклідових просторів в прикладних задачах: кортежні оператори для матричних кортежів

Розглядається концепція кортежних операторів, яка реалізована для варіанту, коли кортежем є набір матриць фіксованої розмірності. Запропонована концепція дозволяє перенести конструктивні засоби опису основних лінійних та нелінійних структур евклідового простору числових векторів на матричні евклідові простори із збереженням основних змістів та інтерпретацій у запропонованій на основі „кортежної концепції“ формалізму.

Евклідові простори числових векторів R^n та їхнє використання в прикладних дослідженнях визначається з одного боку — багатством структур в рамках цих просторів, породжених операціями та скалярним добутком, з іншого — багатством конструктивних можливостей в оперуванні таким структурами. До перших: структур в евклідових просторах, які можна називати базовими, — можна, наприклад віднести лінійні підпростори та гіперплощини (зсунені підпростори). Це ті структури, які можна позначити як множинні, як і еліпсоїди чи еліпсоїдальні циліндри. З іншого боку до базових структур можні віднести матриці лінійних операторів та невід'ємно визначені квадратичні форми — синглові структури. В евклідових просторах R^n множинні та синглові структури пов'язані між собою. Так з матрицею лінійного оператора природним чином пов'язуються два підпростори: множина можливих значень та ядро оператора, а множина можливих розв'язків чи псевдорозв'язків СЛАР (системи лінійних алгебраїчних рівнянь) є гіперплощиною. Так само пов'язані між собою еліпсоїдальні циліндри та матриці невід'ємно визначених квадратичних форм, що їх задають. Ефективність та конструктивного використання базових структур значною мірою визначається засобами оперування з ними, зокрема, можливістю переходу від синглових до множинних і навпаки. Ці можливості реалізуються через SVD та ПдО (псевдообернення) матриці слушного лінійного оператора чи системи лінійних рівнянь (СЛАР). Так само SVD та ПдО дозволяють побудувати ортогональні проектори на підпростори та гіперплощини, породжені наборами векторів: побудувати синглові структури за множинними. Зважаючи, що навчальні виборки в прикладних задачах представляються саме такими наборами, такий перехід реалізує можливість ефективного обчислення відстаней від множинних утворень (підпросторів, гіперплощин), породжених наборами векторів. Так само, SVD та ПдО реалізують можливості ефективного опису матриці квадратичної форми, що задають еліпсоїди, які в певному сенсі оптимальним чином „накривають“ заданий набір векторів. Згаданий зв'язок між множинними та сингловими структурами, можливості його конструктивного опису є важливими у прикладних задачах, які можна позначити терміном „задачі групування інформації“ (GIP): задачі класифікації-кластеризації та задачі відновлення функції, представлених своїми спостереженнями. Одначе GIP-задачі є важливими не тільки в ситуаціях, коли об'єктами групування є числові вектори. Типовим прикладом GIP-задач такого типу є задачі розпізнавання аудіо сигналів представлених матрицями спектрограм, а також задачі обробки зображень. Використання матричних об'єктів як „узагальнених векторів“ стримується недостатньою розвиненістю техніки оперування в матричних евклідових просторах, а отже — розвиток такої техніки для матричних об'єктів є нагальною та важливою задачею для прикладних досліджень. Цю задачу, як показує досвід, можна реалізувати у цьому — тай не тільки в цьому — випадку засобами побудови слушних операторів, які автори називають „кортежними“ (див., наприклад, роботу авторів [1]). Рівень складності розв'язання задач зазначеного підходу зводиться до класичного варіанту.

Література. 1. Донченко В.С., Зінько Т.П., Скотаренко Ф.М. Концепція кортежності для лінійних операторів та її реалізація для матричних кортежів // Журнал обчислювальної та прикладної математики: Серія „Прикладна математика“. №3. 2015. —С. 127–140.

Дяченко О.М.
НТУУ "КПІ" ННК "ІПСА", Київ, Україна

Математичне моделювання поведінки гравців в нечітких умовах за неповної інформації

Теорія ігор та її застосування. Теорія ігор - це теорія математичних моделей прийняття оптимальних рішень в умовах конфлікту. З іншої точки зору — це розділ прикладної математики, який використовується в соціальних науках (найбільше в економіці), біології, політичних науках, комп'ютерних науках (головним чином для штучного інтелекту) і філософії [1–3]. Теорія ігор намагається математично зафіксувати поведінку в стратегічних ситуаціях, в яких успіх суб'єкта, що робить вибір, залежить від вибору інших учасників. Якщо спочатку розвивався аналіз ігор, в яких один із супротивників виграє за рахунок інших (так звані ігри з нульовою сумою), то згодом почали розглядати широкий клас взаємодій, які були класифіковані за певними критеріями.

Основні розбіжності теорії з реальністю. Базовим припущенням теорії ігор є те, що усі учасники конфлікту вибирають найбільш раціональну стратегію. В такому випадку розв'язання задачі полягає в пошуці так званої "точки Неша" для змішаних стратегій гравців. Проте на практиці теорія не відображає дійсний стан речей відносно поведінки гравців, оскільки ключовою відмінністю є її нераціональність в більшій частині випадків [1]. Якщо ж керуватись в такій ситуації методологією теорії ігор, то "раціональний" гравець може отримати гірший результат, ніж він очікує отримати в точці рівноваги раціональних стратегій.

Психологічні основи поведінки в умовах конфлікту. Зважаючи на можливу мінливість поведінки противників та неоптимальність їх дій виникає питання, як потрібно реагувати в таких ситуаціях. Однією з базових причин даної проблеми є те, що середньостатистичний гравець використовує від 0 до 3 рівнів логіки, в той час, як теорія ігор відштовхується від припущення, що кількість цих рівнів безкінечна [4]. Крім цього, значна частина сфер її застосування має настільки складні для багаторівневого аналізу моделі, що обчислення оптимальної змішаної стратегії стає фізично неможливим [2].

Модифікований підхід до теорії ігор. Для вирішення проблеми неоптимальної поведінки виникає потреба відходу від вибору загальної стратегії до алгоритму моментального вибору в нечітких умовах. Для досягнення даної цілі крім введення нечітких характеристик в модель, змішані стратегії замінюються еквівалентними стратегіями, які базуються на поведінці гравців (behavioral strategies). Такий підхід разом з використанням оцінки психологічних портретів гравців дає можливість оцінити ймовірності їх дій в кожній ситуації і діяти відповідно. Математична модель в даному випадку приймає вигляд оціночної функції, яка розраховує суб'єктивну корисність кожної позиції, використовуючи коефіцієнти відношень між гравцями (які можуть приймати як додатні, так і від'ємні значення) та психологічну схильність до кожної з можливих стратегій, які можуть включати в себе як базові, так і провокативну, стратегію шантажу, хаотичну, тощо.

Висновок. Використовуючи психологічні моделі поведінки гравців в умовах конфлікту була створена модифікована модель теорії ігор. Використувуючи дану модель, були проведені дослідження поведінки гравців, які підтвердили вищевказані відхилення від немодифікованої моделі.

Література. **1.** Myerson, Roger B. (1991), *Game Theory: Analysis of Conflict*, Harvard University Press, p.1. **2.** Shoham, Yoav (2008), *Computer Science and Game Theory*, Communications of the ACM 51 (8): 75–79. **3.** Aumann R., Hart S. (1994), *Handbook of Game Theory with Economic Applications*, v. 2, ch. 30-31. **4.** Drew Fudenberg (2006), *Advancing Beyond Advances in Behavioral Economics*, Journal of Economic Literature, pp. 694–711.

Зражевська Н.Г.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Оцінювання динамічних мір ризику VaR і CVaR для фондових індексів на основі гетероскедастичної моделі часового ряду

Метою дослідження є вибір методу оцінювання динамічних мір ризику *VaR* і *CVaR* для часового ряду логарифмічної доходності на денній основі фондового композитного індексу, обчисленого на основі 300 стоків, що торгуються на шанхайській стоковій біржі - *CSI300*.

Вважається, що часовий ряд $\{X_t, t \in Z\}$ підкоряється стохастичному процесу, що має вигляд: $X_t = \mu_t + \varepsilon_t Z_t = \mu_t + \sigma_t Z_t$, де μ_t і σ_t - умовні середнє і варіація, визначені на інформаційному просторі Ψ_t , $\{Z_t\} \sim F_t(0,1)$. Тоді [1]: $VaR_\alpha{}^t = \mu_t + F^{-1}(\alpha)\sigma_t = \mu_t + VaR_\alpha(Z)\sigma_t$, $CVaR_\alpha{}^t = \mu_t + CVaR_\alpha(Z)\sigma_t$.

Для оцінювання волатильності ряду була використана модель *GARCH*, оцінка параметрів здійснювалась за методом максимальної правдоподібності, для перевірки залишків моделі на iid застосовувався BDS тест та критерій відношення дисперсій [2,3,4]. Для знаходження мір ризику для залишків моделі $(VaR_\alpha(Z), CVaR_\alpha(Z))$, проведено аналіз основних методів оцінювання статичних мір ризику. Методи були зведені до спільних економетричних позначень, був проведений їх порівняльний аналіз. Результат проведеної класифікації представлений у вигляді структорно-ієрархічної схеми, що спрощує процедуру вибору конкретного метода (Рис.1).

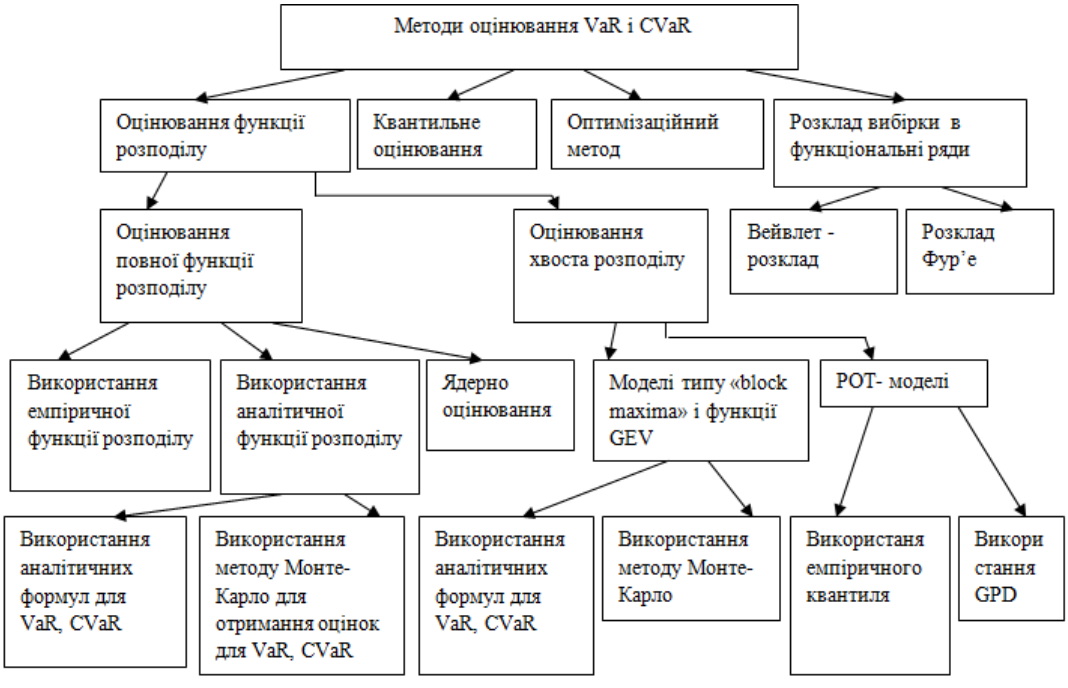


Рис. 1. Структурно-ієрархічна схема вибору методу оцінювання *VaR* і *CVaR*

Для знаходження мір ризику для залишків моделі були застосовані 10 різних методів, що, відповідно з Рис.1, представляють різні підходи: метод історичного моделювання, дискретний метод Рокофеллера-Урясьєва, оцінки в припущенні нормальності розподілу (оцінки параметрів отримано по методу максимальної правдоподібності ММП) по явним формулам та із застосуванням методу Монте-Карло, метод Ричардсона, оцінки з використанням функцій *GEV* та *GPD*, параметри яких оцінювались по ММП - по явним формулам та із застосуванням методу Монте-Карло, емпіричний *POT* - метод [1,2,5]. Для аналізу якості використовувалися тести Купеца і Кристофферсена для оцінок *VaR* та V тест для оцінок *CVaR* [3].

Для верифікації моделі оцінювання динамічних *VaR* і *CVaR* були змодельовані дані по моделі $GARCH(1,1)$ з коефіцієнтами $a_0 = 1$, $a_1 = 0.1$, $b_1 = 0.8$ із нормально розподіленими залишками. ММП дав наступні оцінки фактичних параметрів: $\widehat{a_0} = 0.9$, $\widehat{a_1} = 0.09$, $\widehat{b_1} = 0.81$. Проведені BDS-тест для $m = 4$ і критерій відношення дисперсій показали, що залишки моделі- *iid*. Аналіз похибок оцінок статичних *VaR* і *CVaR*, отриманих із застосуванням десяти вище згаданих методів, свідчить про можливість застосування цих методів для оцінювання. Результати тестів на якість оцінок динамічних мір ризику свідчать про високу якість оцінок.

В якості моделі для часового ряду *CSI300* була використана модель $GARCH(5,0)$ з коефіцієнтами $a_0 = 0.00014$, $a_1 = 0.0713$, $a_2 = 0.15300$, $a_3 = 0.12165$, $a_4 = 0.18222$, $a_5 = 0.10322$. Застосування десяти методів для оцінювання мір ризику для залишків моделі показало, що критерію якості задовольняють оцінки мір ризику, що були отримані з використанням теорії екстремальних значень, а саме оцінки по явним формулам з використанням функцій *GEV* та *GPD*, параметри яких оцінювались по ММП. Це свідчить про те, що часовий ряд залишків моделі демонструє товсті хвости розподілів, отже, для ряду не можна застосовувати методи, що базуються на припущенні нормальності розподілів, а використовувати моделі на основі *GEV* та *GPD*. На рис. 2 представлені змодельовані ${VaR^t}_{0.9}$, ${CVaR^t}_{0.9}$ для ряду *CSI300*, де для моделювання мір ризику для залишків був застосований метод на основі *GPD*.

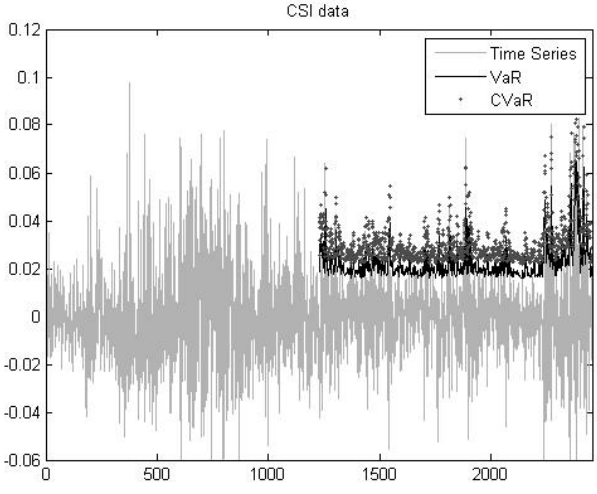


Рис. 2. ${VaR^t}_{0.9}$ і ${CVaR^t}_{0.9}$ для часового ряду *CSI300*

Таким чином, запропонована у роботі класифікаційна схема дозволяє спростити процедуру вибору методу оцінювання статичних мір ризику і визначити оптимальний метод оцінювання динамічних мір ризику на базі гетероскедастичної моделі. Наведений приклад продемонстрував ефективність запропонованої схеми. Схема може бути використана при розробці систем прийняття рішень для формулювання математичних моделей поведінки волатильних фінансових інструментів.

Література. **1.** Tae-Hwy Lee, Y. Bao and B. Saltoglu, "Evaluating predictive performance of value-at-risk models in emerging markets: a reality check", Journal of Forecasting - J FORECASTING , vol. 25, no. 2, 2006, pp. 101-128. **2.** R.S. Tsay, "Analysis of Financial Time Series" (third edition).Hoboken: John Wiley & Sons, Inc. 2010. **3.** B. Kjellson, "Forecasting Expected Shortfall. An Extreme Value Approach.", Bachelor's thesis in Mathematical Sciences, 2013:K7, 43.p. **4.** Зражевська Н.Г. "Метод згладженої автокореляційної функції для прогнозування варіації гетероскедастичних часових рядів". **5.** S. Nadarajah, Bo Zhang and S. Chan, "Estimation methods for expected shortfall", Quantitative Finance, vol. 14, issue 2, 2014, pp. 271-291.

Калюжний М.С.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Використання інтелектуального аналізу тексту для структурування знань з різних текстових джерел

На даному етапі розвитку технологій була спродукована велика кількість інформації. Цей обсяг зростає експоненціально. За даними міжнародної дослідницької і консалтингової компанії International Data Corporation ця кількість у 2007 році становила приблизно 161 ексабайт (161 * 1018 байт або 161 * 1010 Гігабайт). Станом на 2010 рік це число зросло в 6.2 рази, тобто досягло відмітки у 998 ексабайт.

Головною проблемою у максимально повному використанні таких обсягів даних є те, що лише 10% даних є структурованими, а решта, головним чином текст – є неструктурованими. Споживачеві інформаційного продукту доводиться довго шукати корисну і значущу інформацію, методом ручного перебору. Раніше цю проблему вирішували ручним способом структуризації та класифікації, як, наприклад, Карл Лінней у 18 столітті, але, враховуючи сьогоденний обсяг інформації та динаміку його росту, постала необхідність у розробці технологій, які виконували б вищезазначені процеси автоматично.

Технології інтелектуального аналізу тексту (Text Mining) мають своєю метою створити інформаційну систему, яка б автоматично проводила семантичний аналіз тексту, виділяючи ключові слова, а на основі результатів аналізу, будувала б залежності, використовуючи алгоритми класифікації та кластеризації та їх модифікації. Після того як система отримала документ на опрацювання, для ефективного подальшого аналізу, необхідно виконати попередню обробку тексту. По-перше вона включає в себе видалення слів, які є допоміжними і не несуть інформації про зміст документа. Далі, для зручності подальшої обробки, текст приводиться до нижнього регістру. Після цього текст проходить процедуру стреммінгу, тобто приведення слів до нормальної морфологічної форми. Ці алгоритми дуже залежать від мови документа.

Автоматичний пошук ключових слів – виділення слів, або сталих висловів природної мови, які важливі для опису змісту документа та має істотне смислове навантаження. Це можуть бути наукові терміни, відомості про географічні місця, історичні події, персон, витвори мистецтва, тощо. Вони часто служить запитом під час користування пошуковими системами. Цей пошук поділяється на 2 стадії: локальний: виділення лексем, дат, понять, власних назв та інтеграцію і виведення понять.

Наступним кроком аналітичного та інтелектуального аналізу тексту є структурування та знаходження залежностей між знайденими ключовими словами. Для вирішення даної задачі застосовують алгоритми класифікації (наприклад метод Native Bayes), кластеризації (ієрархічний, бінарний, нечіткий), автоматичне анотування. При чому ці алгоритми працюють ефективніше, коли також використовують дані з Інтернет-енциклопедій. Таким чином, із сукупності текстів, у яких описується одне спільне ключове слово, формується поняття, яке автоматично доповнюється даними с інших статей і так розширюється. Далі поняття класифікують для більш зручної роботи.

Таким чином, система аналітичного та інтелектуального аналізу тексту вирішує проблему неструктурованості інформації. Це досягається завдяки визначенню смислового наповнення тексту та об'єднання результатів в єдине поняття структуру з урахуванням усіх залежностей. Вона включає в себе інструменти для попередньої обробки тесту, виділення ключових слів, а також їх подальше структурування.

Література. **1.** 1. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009 **2.** 2. A vector space model for automatic indexing / G. Salton, A. Wong, C. Yang // Commun. – ACM, 1995. – Vol. 18, no. 11. – P. 613–620. **3.** 3. Нога Р. Аналітичний огляд методів та засобів опрацювання текстової інформації / Р. Нога, Н.Б. Шаховська // Вісник Національного університету "Львівська політехніка". – 2011. – № 715 : Інформаційні системи та мережі. – С. 323–332.

Каніовська І.Ю., Зубарева П.К.
НТУУ "КПІ" ІПСА, Київ, Україна

Імовірнісні моделі Страффіна для виміру впливу голосу у простих іграх голосування

Проста гра голосування являє собою набір W кінцевих множин, що задовольняють таким властивостям:

1. W містить найбільшу множину $N = \{1, 2, \ldots, n\}$, що включає в собі всі елементи W. Таким чином, $N \in W$. Тоді $\forall\ S \in W$, $S \subseteq N$;
2. $\varnothing \notin W$
3. (Монотонність) Якщо $S \subseteq T \subseteq N$ і $S \in W$, то тоді $T \in W$ також;

N - об'єднання усіх членів W. Елементи, що містяться в N, будемо називати виборцями у W. Множина виборців буде називатись коаліцією. Кількість виборців у коаліції S буде позначатись $|S|$.

Вплив голосу виборця вимірює можливість гравця мати вирішальне значення для результату голосування. У простій грі голосування є тільки два варіанти, в яких гравець може мати вирішальне значення. Він може бути критичним, якщо його голос перетворить невиграшну коаліцію у виграшну, і навпаки - з виграшної коаліції зробить програшну.

Візьмемо програшну коаліцію T, що включає гравця i, $T \subset N$: $i \notin T$ та $w(T)$=0. Якщо i – критичний гравець, тоді $w(T \cup \{\mathrm{i}\})$-$w(T)$=1-0=1. Якщо i – не критичний гравець, тоді $w(T \cup \{\mathrm{i}\})$-$w(T)$=0-0=0. Бачимо, що вираз $w(T \cup \{\mathrm{i}\})$-$w(T)$ поводить себе як функція-індикатор, показуючи нам, коли i є критичним гравцем для програшної коаліції. Міра можливості гравця i бути критичним для програшної коаліції задається формулою (1):

$$\sum_{T \subset N, i \notin T} \pi_T [w(T \cup \{i\}) - w(T)], \tag{1}$$

де π_T – імовірність появи коаліції T.

А зараз візьмемо виграшну коаліцію S, що включає гравця i, $S \subseteq N$: $i \in S$ та $w(S)$=1. Якщо i – критичний гравець, тоді $w(S)$-$w(S \setminus \{i\})$=1-0=1. Якщо i – не критичний гравець, тоді $w(S)$-$w(S \setminus \{i\})$=1-1=0. Тоді вираз $w(S)$-$w(S \setminus \{i\})$ індикує нам ті випадки, в яких гравець i є критичним для виграшної коаліції. Міра можливості гравця i бути критичним для виграшної коаліції задається формулою (2):

$$\sum_{S \subset N, i \in S} \pi_S [w(S) - w(S \setminus \{i\})], \tag{2}$$

де π_S – імовірність появи коаліції S.

Страффін стверждував, що сила голосу має бути визначена здатністю гравця бути критичним у коаліції при голосування, будь-то виграшна або програшна, та вимірюватись сумою двох попередніх виразів (3):

$$\sum_{T \subset N, i \notin T} \pi_T [w(T \cup \{i\}) - w(T)] + \sum_{S \subseteq N, i \in S} \pi_S [w(S) - w(S \setminus \{i\})], \tag{3}$$

Слід зазначити, що суми містять рівне число доданків та $T = S \setminus \{i\}$

Нехай p_i – імовірність того, що гравець i проголосує «Так». «Умова Незалежності»: імовірність p_i вибирається незалежно з рівномірного розподілу на $[0, 1]$. Тобто всі гравці голосують незалежно один від одного, та імовірність голосу «Так» рівна $1/2$ для кожного з них. «Умова Однорідності»: спочатку генеруємо значення величини p, що рівномірно розподілена на відрізку $[0, 1]$, а потім імовірність голосу «Так» обирається рівною p для кожного гравця. $p_i = p\ \forall\ i$.

Умова Незалежності. Нехай P_i – імовірність того, що гравець i проголосує «Так» і нехай $(P_1, \ldots, P_n)$ вектор з n незалежних і рівномірно розподілених випадкових величин. Беручи p_i як реалізацію P_i, імовірність профіля голосування $v = (v_1, ..., v_n)$, де $v_j = 1$, коли гравець j

голосує «Так», та v_j=0, коли гравець j голосує «Ні» отримуємо:

$$\pi_v(p_1,\dots,p_n)=\prod_{j=1}^{n}p_j^{v_j}(1-p_j)^{1-v_j} \tag{4}$$

При умові незалежності P_i:

$$\pi_v=E\prod_{j=1}^{n}P_j^{v_j}(1-P_j)^{1-v_j}=\prod_{j=1}^{n}EP_j^{v_j}(1-P_j)^{1-v_j} \tag{5}$$

З того, що

$$EP_j^{v_j}(1-P_j)^{1-v_j}=\begin{cases}EP_j=1/2, \text{якщо } v_j=1,\\ E(1-P_j)=1/2, \text{якщо } v_j=0,\end{cases}$$

Ми отримуємо, що

$$\pi_v=\frac{1}{2^n} \tag{6}$$

Підставляючи ці імовірності в рівняння (3) міра Пенроуза – Банзафа набуває вигляду:

$$\beta(i)=\sum_{T\subset N,i\notin T}\frac{1}{2^{|N|}}[w(T\cup\{i\})-w(T)]+\sum_{S\subseteq N,i\in S}\frac{1}{2^{|N|}}[w(S)-w(S\setminus\{i\})], \tag{7}$$

Замінюючи $T=S\setminus\{i\}$, отримуємо:

$$\beta(i)=\frac{1}{2^{|N|-1}}\sum_{S\subseteq N,i\in S}\frac{1}{2^{|N|}}[w(S)-w(S\setminus\{i\})], \tag{8}$$

що також відома як абсолютна міра Банзафа. Відносна міра Банзафа - відома як індекс Банзафа, отримується шляхом нормалізації абсолютної.

Умова Однорідності грунтується на тому, що будь-які два голоси є незалежними та обумовлені реалізацією p. Отже, умовна імовірність того, що з n гравців s проголосують «Так» рівна $\pi_s(p)=p^s(1-p)^{n-s}$, де $s=0,\dots,n$. Безумовна імовірність спостереження профілю голосування з s голосами «Так» задається:

$$\pi_s=E(\pi_s|P=p)\int_0^1 p^s(1-p)^{n-s}\,dp=B(s+1;n-s+1)=\frac{\Gamma(s+1)\Gamma(n-s+1)}{\Gamma(n+2)}=\frac{s!(n-s)!}{(n+1)!}, \tag{9}$$

В термінах простих ігор голосування для $S\subseteq N$:

$$\pi_s=\frac{|S|!(|N|-|S|)!}{(|N|+1)!}, \tag{10}$$

Підставляючи ці імовірності у вираз (3) отримуємо визначення індексу Шеплі-Шубіка:

$$\phi_(i)=\frac{(|S|-1)!(|N|-|S|)!}{|N|!}\sum_{S\subset N,i\in S}[w(S)-w(S\setminus\{i\})], \tag{11}$$

Література. **1.** Straffin, P. D.: 1977, Homogeneity, independence and power indices, Public Choice 30, 107–118 **2.** Straffin, P. D.: 1978, Probability models for power indices, in P. Ordeshook (ed.), Game Theory and Political Science, New York University Press, pp. 477–510. **3.** Kaniovski, S.: 2008, The exact bias of the Banzhaf measure of power when votes are neither equiprobable norindependent, Social Choice and Welfare 31, 281–300. **4.** Felsenthal, D. S. and Machover, M.: 1998, The measurement of voting power: Theory and practice, problems and paradoxes, Edward Elgar

Качаловський А.С.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Дослідження методів прогнозування лінійних та нелінійних нестаціонарних процесів

В сучасних умовах загальної економічної нестабільності, коли у світі наступають фінансово-економічні кризи, виникає необхідність удосконалення світової фінансової системи, а також удосконалення методів і технологій управління фінансовими ризиками. Актуальною задачею є виконання поглиблених наукових досліджень в напрямі математичного моделювання і прогнозування фінансових процесів. Це стосується, безумовно, і українського фінансового ринку, на якому фінансовий ризик-менеджмент знаходиться на ранній стадії свого становлення. Фінансові ринки – це великі та складні ієрархічні системи з великою кількістю змінних та факторів впливу, а також функціональних зв'язків між їх підсистемами. Вони потребують застосування системного підходу та достатньо складних математичних методів для виконання аналізу їх функціонування, методів статистичної обробки даних та належних комп'ютерних обчислювальних засобів. Багато економічних задач можна проаналізувати і розв'язати за допомогою часових рядів даних. При цьому важливим напрямом досліджень є моделювання динаміки доходності фондового та валютного ринків з використанням волатильності, тобто міри їх мінливості, яка визначається умовною дисперсією процесу. Метою моделювання волатильності є оцінювання її прогнозу для вивчення різних аспектів ринкової доходності та використання цього параметра при створенні правил торгівлі і в менеджменті ризиків. [1] Для цього треба розглянути деякі особливості сучасної економіки:

Нестаціонарність процесів - це процеси для яких властиво наявність тренду або змінної дисперсії (гетероскедастичні процеси). Тренд - це загальний довгостроковий напрям розвитку процесу. Фактично він співпадає з поточним середнім значенням. Існує два типи тренду: детермінований (регулярний) та стохастичний (нерегулярний). Визначення типу тренду відбувається в – побудові моделі. Процеси зі змінною дисперсією називають гетероскедастичними. Дисперсія може змінюватися за досить складним законом, який і потрібно знайти при створенні моделі процесу.

Нелінійність. Однією з проблем визначення структури моделі є встановлення факту нелінійностей в досліджуваному процесі та їх типу. Економічна динаміка являє собою нелінійне явище. Нелінійність - це можливість неочікуваних змін напрямів розвитку процесів. В динамічних економічних системах нелінійність виявляється як підвищена реакція на зміну одних факторів і повною нечутливістю до інших.

Моделювання. Оскільки розвиток нестаціонарних процесів представляють, як правило, часовими рядами, основну увагу приділено проблемам моделювання і прогнозування фактичних часових рядів за допомогою моделей авторегресії з інтегрованим ковзним середнім (АРІКС) [2], а також моделі гетероскедастичних процесів різних модифікацій. Такі моделі застосовують в автоматизованих інформаційно-аналітичних системах та системах підтримки прийняття рішень для прогнозування динаміки технічних систем і технологічних процесів, вартості акцій та інших біржових активів, курсів валют, рівня інфляції, цін на товари та ін. Моделі такого класу можна застосовувати для аналізу сезонності змін статистичних характеристик, в тому числі для прогнозування процесів на макрорівні. Узагальнюючи, можна сказати, що вищеозначені моделі ефективно застосовують там, де основною інформацією для аналізу процесу та прогнозування його розвитку є часовий ряд.

Література. **1.** Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів. – Київ: НТУУ КПІ, Політехніка, 2013. – 607с. **2.** Демківський Е. О. Інформаційні технології аналізу і прогнозування нестаціонарних процесів.– Київ: - 2007.

Киселева Е.М., Гарт Л.Л., Селищев В.Л., Матяш В.В.
Днепропетровский национальный университет имени Олеся Гончара, Днепропетровск, Украина

Анализ сложности проекционно-итерационных алгоритмов решения линейных интегральных уравнений

Исследование многих сложных процессов и явлений, возникающих во многих областях естествознания, техники и управления, требует построения и изучения соответствующих адекватных математических моделей, значительное место среди которых принадлежит интегральным уравнениям. Интегральные уравнения естественным образом возникают в процессе математического моделирования в геофизике, астрофизике, аэродинамике, радиотехнике, компьютерной томографии, при обработке и интерпретации данных физических экспериментов, в анализе информации, экономике, медицине. Аппарат интегральных уравнений прочно вошел в теорию управления, теорию надежности аппаратов, теорию распространения волн и др.

Интенсивное развитие теории интегральных уравнений, в том числе некорректных, во многом обусловлено появлением в последние десятилетия высокопроизводительной вычислительной техники. На сегодняшний день разработаны и представлены в многочисленной литературе в доступной для практического применения форме методы приближенного решения основных классов интегральных уравнений: традиционные (метод квадратур, итераций, резольвенты, вырожденных ядер, моментов, коллокации, преобразования Фурье, собственных функций и др.), эффективные, главным образом, при решении уравнений второго рода, а также регулярные, или робастные (методы Тихонова, Лаврентьева, Иванова, Фридмана, Бакушинского и др.), обеспечивающие устойчивость решения уравнений первого рода [1]. В тоже время актуальными остаются вопросы разработки новых, более эффективных методов решения интегральных уравнений, усовершенствования уже существующих, а также доведения методов до практических алгоритмов, особенно до компьютерных программ.

Целью данной работы является анализ сложности проекционно-итерационных вычислительных схем, основанных на методе квадратур как методе проекционного (аппроксимационного) типа и методах итераций, для приближенного решения линейных интегральных уравнений с постоянными пределами интегрирования вида

$$Ay \equiv g(t)\,y(t) - \alpha \int_a^b K(t,s)\,y(s)\,ds = f(t), \qquad a \leqslant t \leqslant b, \tag{1}$$

где $K(t,s)$ – ядро интегрального уравнения, $f(t)$ – правая часть уравнения, α – параметр уравнения, $g(t)$ – заданная функция на отрезке $[a,b]$, $y(t)$ – искомая функция на $[a,b]$. При $g(t) \equiv 0$ уравнение (1) есть уравнение первого рода, записываемое в виде

$$\int_a^b K(t,s)\,y(s)\,ds = f(t), \qquad a \leqslant t \leqslant b. \tag{2}$$

Если $g(t) \neq 0$ на $[a,b]$, то уравнение (1) допускает деление на $g(t)$ и в этом случае можно рассматривать уравнение второго рода

$$y(t) - \alpha \int_a^b K(t,s)\,y(s)\,ds = f(t), \qquad a \leqslant t \leqslant b. \tag{3}$$

Если же $g(t) = 0$ для некоторых, но не для всех $t \in [a,b]$, то (1) есть уравнение третьего рода, слабо исследованное, но все же встречающееся в приложениях.

Линейные уравнения (1)-(3) можно записывать также в операторном виде и использовать для их исследования общую теорию операторных уравнений в некоторых функциональных пространствах.

При численном решении интегральных уравнений любыми методами неизбежно приходится заменять входящие в них интегралы конечными суммами. Метод квадратур (конечных

сумм, квадратурных формул), как известно [1], состоит в сведении интегрального уравнения к конечной системе алгебраических уравнений относительно приближенных значений искомого решения в точках разбиения отрезка $[a, b]$ путем замены интегральных операторов конечными суммами на основе применения различных квадратурных формул. Полученная система сеточных уравнений решается впоследствии любым известным методом, как правило, итерационным.

Суть проекционно-итерационного подхода к решению линейных интегральных уравнений состоит в том, что исходное уравнение заменяется последовательностью аппроксимирующих его сеточных задач (систем линейных алгебраических уравнений) на совокупности измельчающихся сеток. При этом для каждой из таких "приближенных" задач строится с помощью некоторого итерационного метода лишь несколько приближений к решению, последнее из которых с использованием кусочно-линейной интерполяции принимается в качестве начального приближения в итерационном процессе для следующей "приближенной" задачи. Последовательность линейных интерполянтов построенных приближенных решений выступает в роли последовательности приближений к решению исходного интегрального уравнения.

Упомянутый прием сведения интегрального уравнения к его сеточному аналогу укладывается в общую схему проекционных методов, предложенных в [2] для решения линейных операторних уравнений в нормированных пространствах. Описанная же выше реализация метода квадратур в сочетании с методом итераций на последовательности сеток может быть отнесена к группе проекционно-итерационных методов, теоретически обоснованных в [3] и [4] для линейных операторних уравнений в корректной и некорректной постановках соответственно.

Разработан программный продукт, реализующий алгоритмы классического метода квадратур и его проекционно-итерационного варианта на последовательности сеток с использованием различных итерационных процедур (простой итерации, скорейшего спуска, минимальных невязок, минимальных ошибок), позволяющий сопоставить результаты расчетов по точности получаемых приближенных решений и количеству вычислительных затрат на их нахождение. Проведен сравнительный анализ названных алгоритмов, использующих правило остановки итерационного процесса по невязке, показана их практическая сходимость на примере решения конкретных задач.

Выполнена оценка ресурсов, необходимых для проведения вычислений с помощью разработанных алгоритмов, а именно оценка вычислительной сложности алгоритмов (процессорного времени, количества выполняемых операций) и оценка сложности алгоритмов по памяти [5].

Результаты численных экспериментов свидетельствуют о том, что проекционно-итерационные вычислительные схемы, основанные на методе квадратур и методах итераций, имеют определенные преимущества в сравнении с классическим алгоритмом метода квадратур не только по качеству получаемого приближенного решения, но и по количеству вычислительных затрат, большая часть которых приходится на решение "приближенных" задач невысокой размерности. Наряду с тем, что программная реализация алгоритмов требует от компьютера примерно одинаковой вычислительной мощности (поскольку модифицированный алгоритм лишь многократно использует классический), оценка количества необходимых арифметических операций и время счета подтверждают эффективность проекционно-итерационного подхода.

Литература. **1.** Верлань А.Ф., Сизиков В.С. Интегральные уравнения: методы, алгоритмы, программы. – К.: Наукова думка, 1986. – 544 с. **2.** Габдулхаев Б.Г. Теория приближенных методов решения операторных уравнений. – Казань: Изд-во Казанского гос. ун-та, 2006. – 112 с. **3.** Балашова С.Д. Приближенные методы решения операторных уравнений. – Д.: ДГУ, 1980. – 112 с. **4.** Гарт Л.Л. Явный проекционно-итерационный метод решения некорректных операторных уравнений// Питання прикладної математики і математичного моделювання. – Д.: РВВ ДНУ, 2015. – Вип. 15. – С. 33-47. **5.** Абрамов С.А. Лекции о сложности алгоритмов. – М.: МЦНМО, 2012. – 248 с.

Киселева Е.М., Притоманова О.М., Журавель С.В.
Днепропетровский национальный университет имени Олеся Гончара, Днепропетровск, Украина

О решении задачи «location-allocation» методами оптимального разбиения множеств

Типичным представителем непрерывных задач оптимального разбиения множеств является бесконечномерная транспортная задача, или более общая – бесконечномерная задача размещения предприятий с одновременным разбиением данного региона, непрерывно заполненного потребителями, на области потребителей, каждая из которых обслуживается одним предприятием, с целью минимизации транспортных и производственных затрат. В зарубежной научной литературе такие задачи известны как «location-allocation problems» (задачи размещения-разбиения) и рассматривается как в непрерывной так и в дискретной постановках. В роли потребителей здесь могут выступать также телефонные абоненты, школьники, избиратели, точки орошаемой территории, пациенты для диагностики заболеваний и т. д. Особый интерес вызывают непрерывные задачи размещения-разбиения, в которых потребительский спрос в области размещения представляется непрерывной функцией плотности.

Сформулируем математически бесконечномерную задачу размещения производителей. Пусть потребитель некоторой однородной продукции равномерно распределен в области $\Omega \subset \mathrm{E}_2$. Конечное число N производителей, расположенных в изолированных точках $\tau_i = (\tau_i^1, \tau_i^2)$, $i = 1, ..., N$, области Ω, образует систему точек τ, причем координаты некоторых из них, или даже всех, могут быть заранее не известны. Считаются известными: $\rho(x)$ – спрос в продукции потребителя с координатами $x = (x^1, x^2)$; $c(x, \tau_i)$ – стоимость транспортировки единицы продукции от производителя $\tau_i = (\tau_i^1, \tau_i^2)$ потребителю с координатами $x = (x^1, x^2)$.

Предполагается, что прибыль производителя зависит только от его расходов, которые являются суммой производственных затрат и транспортных расходов. Для каждого i-го производителя задана функция $\varphi_i(Y_i)$, описывающая зависимость стоимости производства от его мощности Y_i, определяемой по формуле

$$\mathrm{Y_i} = \int_{\Omega_\mathrm{i}} \rho\,(\mathrm{x})\,\mathrm{dx},$$

и приведенные капитальные затраты на реконструкцию i-го производителя с целью увеличения его мощности от существующей до проектной Y_i.

Множество потребителей Ω можно разбивать на зоны обслуживания Ω_i потребителей i-м производителем так, чтобы

$$\bigcup_{i=1}^{N} \Omega_i = \Omega, \quad \Omega_i \cap \Omega_k = \varnothing, \quad i \neq k, \quad i, k = 1, ..., N, \tag{1}$$

причем мощность i-го производителя определяется суммарным спросом потребителей, принадлежащих Ω_i, и не должна превышать заданных объемов:

$$\int_{\Omega_i} \rho(x)dx \leqslant \mathrm{b_i}, i = 1, ..., N; \tag{2}$$

не исключается, что некоторые из подмножеств Ω_i окажутся пустыми. Требуется разбить множество потребителей Ω на зоны обслуживания их N производителями, т. е. на подмножества Ω_i , $i = 1, \ldots, N$, и разместить этих производителей в Ω так, чтобы минимизировать функционал суммарных затрат на производство продукции и доставку ее к потребителю:

$$F(\{\Omega_1, ..., \Omega_N\}, \{\tau_1, ..., \tau_N\}) = \sum_{i=1}^{N} \{ \int_{\Omega_i} c(x, \tau_i)\rho(x)dx + \varphi_i(\int_{\Omega_i} \rho(x)dx)\} \tag{3}$$

при условиях (1), (2).

Получено решение задачи (1)-(3), следуя [1,2], в следующем виде:

$$\lambda_i^*(x) = \begin{cases} 1, & c(x,\tau_i) + \psi_i + \varphi'_{iY_i}\left(\int\limits_\Omega \rho(x)\lambda_i^*(x)dx\right) \leqslant \\ & \leqslant c(x,\tau_k) + \psi_k + \varphi'_{kY_k}\left(\int\limits_\Omega \rho(x)\lambda_k^*(x)dx\right), \\ & i \neq k \text{ п.в. для } x \in \Omega \text{ (другими словами} \\ & i = k \text{ только в точках границы между} \\ & \text{подмножествами } \Omega_i \text{ и } \Omega_k) \; i,k = 1,...,N, \\ 0, & \text{в остальных случаях.} \end{cases}$$

Здесь $\lambda_i(x)$ – характеристическая функция подмножества Ω_i, i = 1,...,N, имеющая вид

$$\lambda_i(\mathrm{x}) = \begin{cases} 1, \; \mathrm{x} \in \Omega_\mathrm{i}, \\ 0, \mathrm{x} \in \Omega \backslash \Omega_\mathrm{i}, \; \mathrm{i} = 1,...,\mathrm{N}, \end{cases}$$

а в качестве $\mathrm{Y}_1^*,...,\mathrm{Y}_\mathrm{N}^*$, $\tau_1^*,...,\tau_N^*$, $\psi_1^*,...,\psi_N^*$ выбирается оптимальное решение следующей задачи

$$G(\psi) = \min_{\tau\in\Omega^N}\max_{Y\in U} G_2(Y,\tau,\psi) =$$

$$= \min_{\tau\in\Omega^N}\max_{Y\in U}\left\{-\sum_{i=1}^N \psi_i b_i + \sum_{i=1}^N \left[\left(\varphi_i\left(Y_i\right) - \varphi'_{iY_i}(Y_i)\cdot Y_i\right) + \right.\right.$$

$$\left.\left. + \int\limits_\Omega \min_{k=\overline{1,N}}\left(c(x,\tau_k) + \psi_k + \varphi'_{kY_k}(Y_k)\right)\rho(x)dx\right]\right\} \to \max,$$

при условиях

$$\psi_i \geqslant 0, \; i = 1,...,N.$$

Здесь

$$\Omega^N = \underbrace{\Omega \times ... \times \Omega}_{N},$$

$$U = \left\{Y = (Y_1,...,Y_N) \in E_N : 0 \leqslant Y_i \leqslant b_i, \; i = 1,...,N\right\}.$$

Литература. 1. Киселева Е. М. Непрерывные задачи оптимального разбиения множеств: теория, алгоритмы, приложения: монография / Е. М. Киселева, Н. З. Шор. – К.: Наукова думка, 2005. – 564 с. **2.** Киселева Е. М. Модели и методы решения непрерывных задач оптимального разбиения множеств: линейные, нелинейные, динамические задачи: монография / Е. М. Киселева, Л.С. Коряшкина. – К.: Наукова думка, 2013. – 606 с.

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

Інформаційно-методична основа визначення системи оцінки екологічності складних об'єктів

Актуальність дослідження. визначається необхідністю розробки інформаційного забезпечення системного аналізу складних об'єктів з оцінки їх безпечності та екологічності з метою впровадження управлінського рішення. Вихідна інформація щодо стану та функціонування системних утворень характеризується надзвичайно великою кількістю параметрів, які взаємодіють між собою та трансформуються у системі в умовах змін, пов'язаних з внутрішніми процесами і взаємодією самої системи з навколишнім середовищем. На цей час за даною тематикою в теоретично-практичному плані з урахуванням методології позитивних результатів досліджень науковців запропоновані удосконалення з інформаційного забезпечення ідентифікації якості складних систем [1–3]. У методичному плані для кінцевої об'єктивної оцінки стану системи запропонована попередня класифікація та визначення головних компонент інформаційного потоку [3]. Така обробка вихідних даних для вирішення завдань оцінки сталого розвитку складних систем була визначена недостатньою для отримання інформації з забезпечення комплексної ідентифікації якості.

Таким чином, з метою обґрунтування достатності і цілісності вихідних даних для компараторної ідентифікації у роботі пропонується застосування додатково методики когнітивного моделювання у загальній системі оцінювання складних об'єктів.

Методика дослідження і аналіз результатів. У даній роботі для моделювання взаємозв'язків між параметрами, що характеризують стан регіону, пропонується використовувати знакові когнітивні карти [4]. Такі графи успішно використовуються для визначення збалансованості у системах різної природи, для вирішення завдань моделювання взаємовпливів факторів, дослідження стійкості систем і виявлення вагомих чинників змін у поведінці системи.

У рамках запропонованої системи комплексного оцінювання рівня екологічної безпеки досліджувані параметри стану регіональних об'єктів було розділено на три групи відповідно до екологічного, економічного чи соціального змісту [3]. Розподіл чинників за інформативністю у визначених групах встановлюється за методом головних компонент. Для обґрунтування об'єктивності отриманих таким чином результатів пропонується побудова когнітивної карти, елементами якої є визначені за групами параметри стану системи. Вузлові точки графу розглядаються як якісні характеристики системи та визначаються факторами стійкості системи. У даній роботі такий розподіл достатньо умовний – дослідження передбачає зосередження уваги на екологічних завданнях при урахуванні впливу економічних і соціальних параметрів. Відповідно до імпульсного аналізу змін в системі підтверджуються вихідні дані факторного аналізу і в остаточному варіанті формується інформаційна основа для компараторної ідентифікації рівня екологічності складних об'єктів (рис. 1).

За наданими пропозиціями щодо комплексування методів аналізу отримано таку послідовність ідентифікації якості складних систем (рис. 2).

Висновки. Результатом даної роботи є запровадження додаткової методики когнітивного моделювання у загальній системі оцінювання об'єктів, яка дозволила доформувати комплексне методичне забезпечення системи аналізу якості складних утворень.

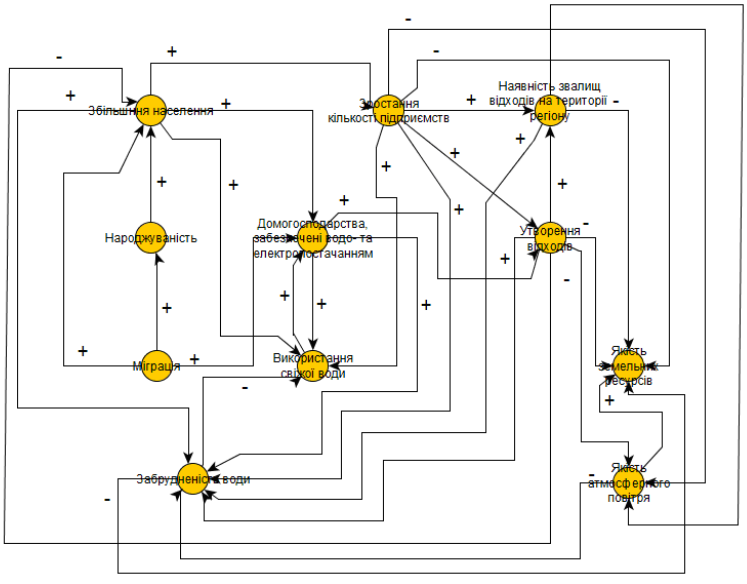


Рис. 1. Приклад виділення взаємозв'язків характеристик безпечності регіону на когнітивній карті

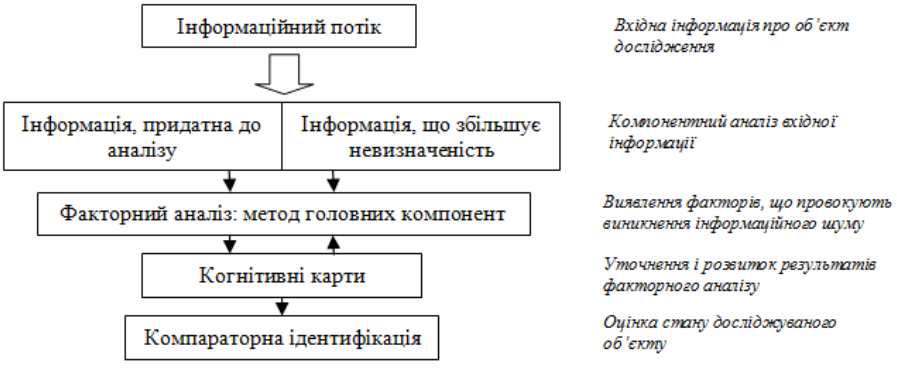


Рис. 2. Приклад виділення взаємозв'язків характеристик безпечності регіону на когнітивній карті

Література. **1.** Згуровский М. З. Глобальное моделирование процессов устойчивого развития в контексте качества и безопасности жизни людей / М. З. Згуровский, А. Д. Гвишиани. – К.: Політехніка, 2008. – 331 с. **2.** Статюха, Г. О. До питання оцінки безпечності промислових об'єктів в аспекті сталого розвитку [Текст] / Г. О. Статюха, Т. В. Бойко, В. І. Бендюг // Екологія довкілля та безпека життєдіяльності. – 2003. – № 3. – С. 57–61. **3.** Козуля Т. В., Білова М. О. Удосконалення комплексної методики оцінки екологічності складних систем методом компараторної ідентифікації // Системный анализ и информационные технологии: материалы 17-й Международной научно-технической конференции SAIT 2015, Киев, 2015 / УНК «ИПСА» НТУУ «КПИ». – К.: УНК «ИПСА» НТУУ «КПИ», 2015. – С.74–76. **4.** Робертс Ф. С. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам: Пер. с англ. – М.: Наука, 1986. – 496 с.

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

Визначення інформаційно-алгоритмічної бази для комплексної оцінки екологічності системних об'єктів

Актуальність і мета дослідження. Важливість розробки комплексних підходів формування методичного забезпечення системного аналізу та оцінки якості об'єктів дослідження обумовлена складністю інформаційного простору функціонування різнорідних структурованих утворень. Особливо це вагомо при визначенні основ інформаційної підтримки прийняття рішень щодо сталого розвитку системних об'єднань.

Метою дослідження є формування інформаційно-алгоритмічної бази щодо удосконалення теоретичних та методичних підходів з комплексної оцінки екологічності природно техногенних системних об'єктів (ПТСО) в контексті сталого розвитку.

Реалізація мети пов'язана з вирішенням таких завдань: 1) визначення на комплексній основі системи інформаційно-методичного забезпечення аналізу стану ПТСО і оцінки їх рівня якості і безпечності; 2) розробка алгоритму комплексного аналізу складних об'єктів відповідно до запровадженої аналітичної системи обробки інформації «стан$_1$ об'єкта – процес – стан$_2$ об'єкта».

Методика дослідження і аналіз результатів. На основі узгодження MIPS і ризик-аналізу запропоновано комплексне методичне забезпечення оцінки якості і безпечності системних об'єктів з метою отримання контрольованих результатів. В межах інформаційного забезпечення роботи методів аналізу передбачено співставлення отриманих за попередніми етапами результатів оцінювання для підтвердження їх достовірності.

Загальні результати оцінки за MIPS-аналізом екологічності природних складових об'єкта дослідження обґрунтовуються розрахунками загальної оцінки ризику щодо стану систем і «позитивності» вихідної дії процесів в них [1].

Деталізація досліджень з оцінки безпечності ґрунтується на методиці комплексного ризик-аналізу «стан$_1$ об'єкта – процес – стан$_2$ об'єкта». Такий підхід дозволяє встановити вагомі механізми регулювання сталого розвитку, що забезпечує інформаційну основу прийняття управлінського рішення.

Комплексність інформаційно-методичного забезпечення дозволяє отримати узгоджені результати різнорівневих досліджень складних систем у такій послідовності: 1) визначення загального рівня екологічної безпечності за MIPS-аналізом і підтвердження цієї оцінки результатами сумарного ризик-аналізу; 2) виявлення точок екологічної небезпеки за пофакторним ризик-аналізом станів ПТСО і процесів в них при високому значенні коефіцієнта кореляції між однаковими ризик-факторами; 3) інтерпретація результатів ризик-аналізу відповідно до загальної оцінки стану «система-навколишнє середовище» і детальної оцінки змін в системі «стан – процес – кінцевий стан» [2].

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

Література. **1.** Козуля Т. В. Інформаційно-методичне забезпечення комплексної оцінки екологічності системних об'єктів / Т. В. Козуля, Н. В. Шаронова, Д. І. Ємельянова, М.М. Козуля // Системні дослідження та інформаційні технології. – 2014. – №3. – С. 25 34. **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.

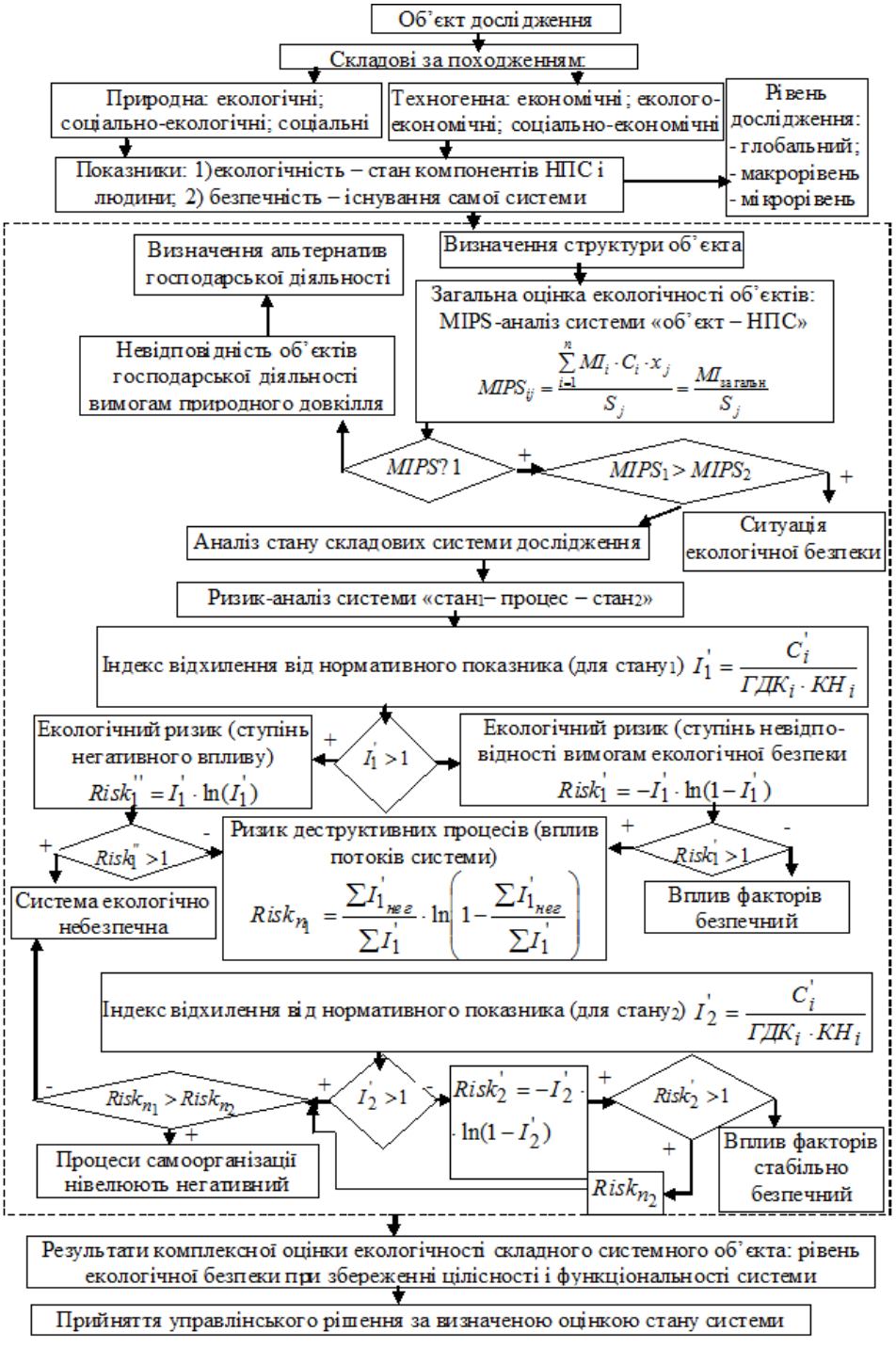


Рис. 1. Алгоритм комплексної оцінки екологічності системних об'єктів

Колбасин В.А.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Категоризация большого объема текстовых документов с использованием ElasticSearch и Spark ML

Задача категоризации текстов. В задачах интеллектуального анализа текстов часто применяется операция категоризации (topic mining) набора (корпуса) текстов при помощи алгоритма LDA (Latent Dirichlet Allocation) [1]. Модель категорий, полученная в результате этой операции, является порождающей вероятностной моделью, задающей вероятность появления терма в документе соответствующей категории. При этом предполагается, что для генерации документа могли использоваться все категории с некой фиксированной для каждого конкретного документа вероятностью. Сам документ при этом описывается набором вероятностей применения различных категорий для его генерации и моделью категорий. Результаты категоризации могут быть использованы как сами по себе для анализа содержимого корпуса, так и для выделения численных характеристик документов для их последующего сравнения или кластеризации.

Промышленные системы сбора и обработки текстовых данных. Промышленные системы сбора и обработки текстовых данных чаще всего строятся на базе систем полнотекстового поиска таких как ElasticSearch [2] и Apache Solr [3]. Такие системы обеспечивают быструю индексацию, поиск и извлечение документов из индекса, но не поддерживают "из коробки"операций интеллектуальной обработки текстов. Применение же реализаций LDA из большинства распространенных библиотек машинного обучения сопряжено с проблемами нехватки памяти при обработке достаточно больших корпусов данных. К счастью, реализация LDA входит в состав зрелого и успешно используемого продукта – библиотеки машинного обучения Spark ML системы распределенных вычислений Apache Spark [4]. Интеграции этих двух технологий и посвящена данная работа.

Категоризация в системах обработки текстовых данных. Поскольку алгоритм LDA использует модель представления текста "мешок слов"(Bag-of-words), он требует в качестве входных данных набор частот встречаемости каждого терма в документе (term-frequency vector, TFV). Традиционно для получения TFV требуется выполнить разбор текста, однако при использовании ElasticSearch можно извлечь TFV непосредственно из индекса средствами поисковой системы. При этом ElasticSearch, как и Spark, позволяет производить параллельную обработку на кластере. Для обеспечения принципа локальности данных можно использовать ключи шардирования (sharding/routing keys) ElasticSearch. Тогда каждый модуль выполнения Spark будет выбирать локальные для него TFV из индекса ElasticSearch и обрабатывать их в параллельном режиме с другими модулями выполнения, при необходимости обмениваясь между ними данными.

Выводы. Практическая реализация данного подхода позволяет добавлять методы интеллектуального анализа текста, основывающиеся на построении модели категорий корпуса текстов, к существующим системам обработки текстовых данных без внесения существенных изменений в них. Также данный подход позволяет варьировать соотношением времени построения модели категорий к количеству машин в кластере и соответственно стоимости такой обработки.

Литература. **1.** D.M. Blei, A.Y. Ng, M.I. Jordan, "Latent Dirichlet Allocation,"J. of Machine Learning Research no. 3, pp. 993–1022, 2003. **2.** R. Kuc, M. Rogozinski, Mastering ElasticSearch, Packt, 2013. **3.** A. Gazzarini, Apache Solr Essentials, Packt, 2015. **4.** N. Pentreath, Machine Learning with Spark, Packt, 2015. **5.** Bag of words model. [Online]. Available: http://en.wikipedia.org/wiki/Bag-of-words_model.

Комлєв О.О., Довніч О.С.
Київський національний університет імені Тараса Шевченка, Київ, Україна

Термодинамічна система Землі: механізми її самоорганізації

Наша планета складно побудований космічний об'єкт – відкрита термодинамічна система, що почала формуватися 4,5-4,65 млрд. р. тому внаслідок *акреції* (нарощування) хаотично захопленої космічної речовини (пилу) до її певної критичної маси. В перші 100 млн. р. Земля пройшла планетний етап розвитку (гадей), зміст якого визначали: взаємодія частинок пилу на рівні атомів і *виділення* значної теплової енергії, *розігрів* і *розплавлення* цієї маси, її швидке *охолодження* і *щільнісне* (гравітаційне) розділення. Виникли первинні *внутрішні* оболонки (кора, мантія, ядро) і *зовнішні* сфери (атмогідросфера, із зниженням температури якої виділились атмосфера і гідросфера) і Земля набула фундаментальної якості багатьох планет - *шаруватості*, яка відображує структуру її гравітаційного поля. Акреція триває постійно, але космічна речовина все більше *розпорошується* і згорає в атмосфері, поповнює загальну енергетику планети. На наш погляд, щільнісне розділення речовини планети в гадеї - це *перше* і *визначальне* для подальшої еволюції термодинамічної системи Землі проявлення її *самоорганізації (синергетики)*. В подальшому відбувалися диференціація утворених компонентних оболонок, поглиблення їх *взаємодії*, і *інтеграція* в єдиній надсистемі планети. Так, відбулися: *заміна первинної «базальтової» протокори* (до базальтового шару додались гранітно-метаморфічний і осадочний); *розділення мантії на нижню і верхню* завдяки появі між ними рідко-пластичної *астеносфери Гутенберга; розділення ядра на рідке зовнішнє і тверде внутрішнє; зміна* атмосфер - *першої* воднево-гелієвої *другою* – вуглецевою, далі *третьою* і *четвертою* азотно-кисневими (відмінність їх у складі ізотопів кисню – у четвертій значно виріс відсоток «екзогенного» O18 і зменшився «ендогенного» O16); в гідросфері змінювались температура, мінералізація, хімічний склад води, відбувалося і її поповнення ювенільними водами. Все ці події добре пояснюється саморозвитком відкритих термодинамічних систем, зокрема їх здатністю до *негентропії* (появі зустрічного порядку), яка проявляється в них як *спряжений* процес, що відбувається на тлі основного - дисипації теплової енергії в навколишнє середовище і супроводжується ростом їх ентропії, а з позиції термодинамічної системи Землі є актами її планетарної самоорганізації. На наш погляд, ними були створені і *умови* для появи біосфери і розвитку її відомих форм життя, антропогенезу та формуванню ефективних *механізмів* їх захисту, в основі яких багаторівнева взаємопов'язана система *колообігів* речовини, енергії, інформації, ентропії. До таких механізмів відносяться: *перешарування* у внутрішній будові планети *ослаблених* пластичних *астеносфер* (Гутенберга та інших) і *міцних* твердих ділянок збільшує стійкість планетного тіла, дозволяє «гасити» як внутрішні його збурення, так і протистояти зовнішнім космічним впливам (перед усім гравітаційним); стійкий планетарний *термостат* (сприятливі для життя температури), який існує тут впродовж мільярдів років; рідке зовнішнє ядро Землі генерує *геомагнітне поле*, що блокує шкідливі для біосфери і людини короткохвильові випромінювання космосу (Сонця); *хімічний склад нинішньої атмосфери*, зокрема здатність кисню стратосфери до нейтралізації жорстких космічних променів, що веде до виділення теплоти і утворення *озоносфери*. Планетарний гомеостаз забезпечує і *морфосистема Землі (геоморфосистема)*, зокрема функціонування колообігових метаболічних систем речовини, енергії, інформації, ентропії на денній поверхні. Роль її в планетарній еволюції постійно зростає, особливо із входженням в *автоколивальну мегасистему*, що інтегрує системи зовнішніх і внутрішніх оболонок Землі. Причинно-наслідкові їх відношення значно прискорили загальні темпи метаболізму. Ріст перепаду висот, площі, експозиційних контрастів денної поверхні, збільшили і накопичення нею екзогенної енергії. Це веде до урізноманітнення геосистем, зменшення тривалості циклів їх розвитку, зокрема і геоморфосистеми. Звідси важливий висновок: *термодинамічна система Землі вичерпує власну (ендогенну) енергію, збільшує ефективність її використання і все більше накопичує зовнішню екзогенну.*

Кораб К.О., Каніовська І.Ю.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Тестування скорингової моделі за допомогою коефіцієнту Джині

Об'єкт дослідження. Скорингова модель, побудована на основі логістичної регресії.

Мета роботи. Розробка скорингової моделі для оцінювання кредитоспроможності позичальника на основі логістичної регресії та її тестування з використанням коефіцієнту Джині.

Методи дослідження. Логістична регресія, коефіцієнт Джині.

Результати. Скоринг - це методика оцінки кредитного ризику, яка дозволяє, оцінивши набір ознак, що характеризують позичальника, визначити, чи варто надавати йому кредит. В основі скорингової системи лежить припущення, що люди зі схожими соціальними показниками поводяться однаково, наприклад, за статистикою люди з вищою освітою повернуть кредит з більшою ймовірністю, ніж люди з початковою освітою [1]. Суть скорингу полягає в тому, що кожному параметру, що характеризує позичальника, надається реальна оцінка в балах. Чим вищий бал, тим більш надійним позичальником є людина.

У роботі було розглянуто скорингову модель у вигляді зваженої суми визначених характеристик позичальника: вік, сімейний стан, місце роботи, дохід, наявність телефонів та адресів та багато інших факторів. На основі анкетних даних позичальників банку будується статистика, що визначає відповідні коефіцієнти моделі. Ці коефіцієнти вказують, на скільки даний фактор впливає на загальну характеристику позичальника.

Для дослідження використовується логістична регресія як найпоширеніша статистична модель для побудови скорингових карт. Математично модель логістичної регресії виражає залежність логарифма шансу (логіт) від лінійної комбінації незалежних змінних [2].

$$\ln \frac{p_i}{1-pi} = b_0 + b_1 x_i^1 + b_2 x_i^2 + .. + \epsilon^i. \quad (1)$$

Рівняння (1) відображає лінійну залежність ймовірності настання прострочення по кредиту в залежності від значень незалежних змінних. Константа в моделі відображає природний рівень ризику настання модельованої події за рівності всіх незалежних змінних нулю. Значення коефіцієнтів при незалежних змінних, що відображають ступінь їх впливу на шанс дефолту в логарифмічною шкалою, використовуються для побудови скорингової карти.

Значення константи в моделі логістичної регресії залежить від розподілу в даних за категоріями залежної змінної. У разі перезважування вибірки для зміни цього розподілу для більш адекватної подальшої оцінки якості отриманої моделі константу було скоректовано і отримано наступну модель логістичної регресії:

$$\ln \frac{p_i^*}{1-p_i^*} = \ln \frac{p_1 \pi_0}{p_0 \pi_1} + b_0 + b_1 x_i^1 + b_2 x_i^2 + .. + \epsilon^i. \quad (2)$$

де $p_i{}^*$ - відкоригована апріорна ймовірність; p_0 і p_1 - частки «хороших» і поганих» позичальників у вибірці; π_0 і π_1 - частки «хороших» і «поганих» позичальників в генеральної сукупності.

При включенні в модель логістичної регресії неперервних кількісних змінних коефіцієнти при них будуть показувати, на скільки в середньому змінитися логарифм шансу настання прострочення по кредиту при зміні незалежної змінної на одиницю свого виміру при незмінності інших змінних.

Основне завдання скорингу полягає не лише в тому, щоб з'ясувати, чи спроможний клієнт повернути кредит та виплатити відсотки чи ні, але і ступінь надійності потенційного позичальника [3]. Так після побудови моделі було досліджено залежність скорингового балу, отриманого на основі моделі, та ймовірності платежу на прикладі історії платежів банку за 2015 рік (рис. 1).

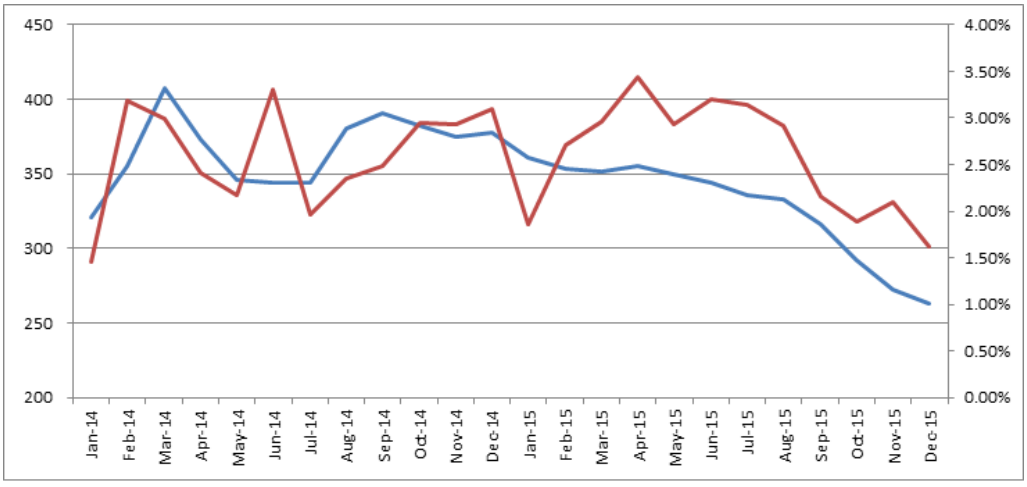


Рис. 1. Ймовірність платежу та середній скоринговий бал кредитного портфелю банку

За рисунком бачимо, що скорингова модель досить точно описує динаміку залежності скорингового балу та ймовірності платежу. Менший бал призводить до зменшення ймовірності платежу та навпаки. Важливою характеристикою будь-якої-регресійної моделі є достовірність. Достовірність моделі логістичної регресії характеризується її здатністю відрізняти «хороших» позичальників від «поганих». Для оцінки якості класифікації моделі побудуємо ROC-криву, яка показує залежність кількості вірно класифікованих позитивних результатів від кількості невірно класифікованих негативних наслідків [4]. Для порівняння двох і більше моделей між собою порівнюються площі під ROC-кривими - цей показник називається AUC і вимірюється в діапазоні від 0.5 до 1.

Чим більше значення площі, тим краща модель. Зазвичай вважають, що значення площі від 0.9 до 1 відповідає відмінній якості моделі, від 0.8-0.9 - дуже хорошому, 0.7-0.8 - хорошому, 0.6-0.7 - середнього, 0.5-0.6 - незадовільного. За значенням площі під ROC-кривої можна обчислити показник індекс Джіні. Цей показник переводить значення площі під кривою в діапазон від 0 до 1. Для побудованої моделі було обчислено коефіцієнт Джині. Він складає 0.68 - модель відповідає середньому рівню якості. Для дослідження варіантів вдосконалення моделі було застосовано процедуру виключення боржників з максимальним та з мінімальним балом. Таким чином, можна досліджувати, які сегменти виборки модель гірше прогнозує, виключити їх та досягти збільшення коефіцієнту Джині. В побудованій моделі виключення 10 % найгірших позичальників призвело до збільшення коефіцієнту до 0.75. Це означає, що модель краще відрізняє боржників з середнім та високим балом. Даний алгоритм можна застосовувати для вдосконалення та подальшого дослідження параметрів моделі.

Маючи остаточно скореговану модель, можна швидко настроїти процес видачі кредитів. На основі персональних даних позичальника визначаємо скоринговий бал, який вказує, чи запропоноувати людині більш вигідні пропозиції кредитування, бо ризик неповернення досить низький, чи одразу відмовити у видачі кредиту.

Література. **1.** Камінський А. С. Експертна модель кредитного скорингу позичальника банку / А. С. Камінський // Банківська справа. – 2006. - № 1. – С.75 - 81. **2.** С. В. Шкляр // Теорія ймовірностей та математична статистика. - 2011. - Вип. 85. - С. 150-160. **3.** Івченко І. Ю. Економічні ризики: Навчальний посібник / І. Ю. Івченко. – Київ: «Центр навчальної культури», 2004. – 304 с. **4.** Lyn C Thomas // Measuring the Discrimination Quality of Suites of Scorecards: ROCs Ginis, Bounds and Segmentation // Quantitative Financial Risk Management Centre, University of Southampton UK// August 2007.

Кравченко А.М.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Аналіз та моделювання процесів перехідної економіки

Перехідні процеси – це найтяжчі і найскладніші періоди існування для будь-яких систем – економічних, технічних чи біологічних. Тому управління такими процесами потребує набагато більших зусиль та інтелекту ніж управління усталеними процесами.

Протікання сучасних фінансово – економічних процесів характеризується великою кількістю специфічних особливостей, які треба враховувати при побудовах адекватних моделей та їх прогнозування. Можна виділити розглянуті нижче особливості.

Висока нестаціонарність процесів. Під нестаціонарними процесами будемо розуміти такі процеси для яких властиво наявність тренду або змінної дисперсії (гетероскедастичні процеси). Під трендом будемо розуміти загальний довгостроковий напрям розвитку процесу. Фактично він співпадає з поточним середнім значенням. Існує два типи тренду: детермінований (регулярний) та стохастичний (нерегулярний). Визначення типу тренду відбувається в подальшому аналізу – побудові моделі. Процеси зі змінною дисперсією (змінюється з часом) називають гетероскедастичними. Дисперсія може змінюватися за досить складним законом, який і потрібно знайти при створенні моделі процесу. Процеси з трендами або зі змінною дисперсією особливо характерні для нестійкої шводкоростущої перехідної економіки.

Наявність нелінійностей. Однією з проблем визначення структури моделі є встановлення факту нелінійностей в досліджуваному процесі та їх типу. Економічна динаміка являє собою нелінійне явище. Нелінійність означає можливість неочікуваних змін напрямів розвитку процесів. В динамічних економічних системах нелінійність виявляється як підвищена реакція на зміну одних факторів і повною нечутливістю до інших.

Вплив сильних економік на слабкі. Однією із невід'ємних складових перехідного процесу є вплив на економічні процеси держави зі слабкою економікою держав з вищим рівнем економічного розвитку. Слабкість ринку (як складової частини економіки) проявляється неможливістю ринковими механізмами розв'язувати деякі економічні задачі взагалі або найкращим способом. Більш розвинені економіки напряму впливають на ситуацію розвитку фінансово – економічних процесів. Встановлення ціни на природні ресурси, залучення інвестицій, підтримання експортно-імпортних відносин, надання кредитів відіграє значний вплив на формування процесів, тому при їх моделюванні потрібно враховувати фактори впливу цих країн. Наприклад Росія має великий вплив на макроекономіку країн колишнього Радянського Союзу, а рівень промислового розвитку Сполучених Штатів Америки має прямий вплив на рівень промислового розвитку Канади та Великобританії.

Висока динаміка змін економічних процесів. На різних ділянках часового ряду може спостерігатись різна динаміка, тобто різна швидкість падіння або зростання значень змінної. Це особливо стосується інфляції, об'єму промислового та сільськогосподарського виробництва, податкових надходжень.

Моделі для опису перехідної економіки. Основними моделями опису процесів перехідної економіки являються модель АРУГ [1], основана на теорії залежності дисперсії від історичних данних процесу, та модель УРАУГ [1], яка враховує вплив попередніх значень дисперсії на її значення в поточний момент часу.

Література. 1. Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів. – Київ: НТУУ КПІ, Політехніка, 2013. – 607с.

Кузенков О.О., Падалко В.Г.
Дніпропетровський національний університет імені Олеся Гончара, Дніпропетровськ, Україна

Біфуркаційні властивості у вироджених диференціальних моделях природничих процесів

Актуальність роботи обумовлена важливістю дослідження складних систем, які становлять основу функціонування більшості біологічних, економічних та соціальних процесів, а також необхідністю розробки ефективних методів керування цими системами для досягнення відповідного економічного або соціального ефекту. Досить часто дослідження в цій галузі обмежуються використанням простих (лінійних) моделей і методів, які можуть бути корисними як перший крок, що дозволяє дати принциповий опис системи та встановити загальні взаємозалежності між її параметрами. У той же час зрозуміло, що дослідження будь-яких складних систем не може обмежуватися лінійним описом, без урахування нелінійних ефектів в їх функціонуванні. Зокрема, це стосується дослідження їхніх біфуркаційних властивостей, тобто можливості й характеру якісної зміни загальної поведінки системи при неперервній зміні значень її параметрів. Особливо цікавим є таке дослідження для вироджених систем, тобто систем, деякі параметри яких, за ознакою системи, мають критичні (граничні) значення. І якщо для невироджених (робастних) систем біфуркаційні дослідження інколи проводяться, то для вироджених систем відповідні методики практично не розроблялися. Дана робота присвячена саме розробці такої методики для одного класу вироджених моделей. Визначення критичних станів екологічних або соціальних систем, що складаються із декількох груп, а також умов, за яких у системі можливі біфуркації (катастрофи), дає можливість вирішувати суто практичні проблеми, зокрема, дозволяє розробляти рекомендації щодо запобігання стрибкоподібних змін у функціонуванні реальних систем.

Серед авторів класичних наукових праць, що покладені в основу запропонованої методики дослідження біфуркацій, назвемо такі імена, як H. Poincare, В.І. Арнольд, О.О. Андронов, О.М. Ляпунов, Д. Ероусміт, серед українських авторів – це А.М. Самойленко [1], Ю.С. Іляшенко, М.О. Перестюк, І.О Парасюк та ін. Із сучасних авторів-дослідників у даній царині варто згадати В.В. Калініна, М.В. Шамоліна, Т.Г. Незбайла, Г.Н. Яковенка, а також зарубіжних учених K.Morris, M.Hirsch, S.Smale, E.A.Coddington, N.Levinson, S.Lefshets, Н.О. Магніцького, С.В. Сидорова та інщих. Запропоновані моделі дають можливість передбачати стан складних систем із певною виродженою структурою. Якісні методи дозволяють визначити загальні динамічні тенденції системи без проведення натурних або комп'ютерних експериментів.

В роботі були отримані такі основні результати:

-розглянуто нелінійні моделі міжгрупової взаємодії другого та третього порядку з урахуванням властивостей гомогенних груп;

-застосовані основні етапи методики дослідження нелінійних моделей міжгрупової динаміки;

-досліджено локальну та глобальну динаміку моделей, а також класифіковані всі можливі нееквівалентні топології їхнього фазового простору [2];

-розроблено загальну класифікацію біфуркацій, які є можливими в розглянутих моделях міжгрупової динаміки [3];

-проведено порівняльний аналіз результатів моделювання з відомими теоретичними результатами та результатами біологічних й соціологічних досліджень;

Література. 1 Самойленко А. М. Диференціальні рівняння: підручн. / А.М. Самойленко, М.О. Перестюк, І.О. Парасюк. – 2-ге вид., перероб. і доп. – К.: Либідь, 2003. – 600с. **1.** Кузенков О.О. Класифікаційна система точок рівноваги лінійної системи диференціальних рівнянь третього порядку/ О.О. Кузенков // Вісн. Київ. нац. ун-ту імені Тараса Шевченка. Сер. "фіз.-мат. науки". – К., 2011. – №3. С.61-69. **2.** Kuzenkov O.O. Biffurcation effects in degenerate differential models of subpopulation dynamics/ O.O. Kuzenkov, S.V. Chernyshenko // ECMS2013, 27-th european conference on modelling and simulation. – Norway, 2013.

Кузнєцова Н.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Скорингові карти як інструмент аналізу ризиків

Ризик – це міра можливих відхилень від цілі внаслідок впливу внутрішніх або зовнішніх факторів, які у випадку реалізації призводять до втрат. Основними параметрами оцінки ступеня ризиків є ймовірність втрат (чим вона більша, тим більший ризик) та розмір втрат (чим він більший, тим небезпечніший ризик). Появу ризику необхідно розглядати у контексті проблеми невизначеності (неточності, недостовірності, ненадійності) знань про умови та процеси, які відбуваються в об'єкті та зовнішньому середовищі, з ймовірнісним характером виникнення небажаних подій. Тому дослідження ризиків пов'язане з аналізом невизначеності, а ефективні шляхи для попередження небажаних подій, зменшення рівня ризиків при прийнятті рішень пов'язані із цілеспрямованою мінімізацією невизначеності [1,2]. Ризик прояву зовнішніх факторів в такому випадку буде визначатися як ймовірність попадання даної характеристики (фактору) в очікуваний інтервал $P(R) = P(x_i \in [a_i, b_i])$, де x_i – i-й фактор, а a_i та b_i - відповідна ліва та права межа інтервалу, в який оцінюється попадання фактору. Оскільки ймовірність оцінюється в межах від 0 до 1, то тоді ймовірність попадання фактору x_i в інтервал $1 \leqslant j \leqslant N$ оцінюється як $\frac{1}{a_{j+1}-a_j}$, де $[a_j, a_{j+1}]$ – межі інтервалу для фактору x_i з розбиттям на N однакової ширини інтервалів (наприклад ризик 0.1; 0.2; 0.3 і т.д.) [1].

Типи скорингових карт. В залежності від галузі підприємства, для якого необхідно аналізувати ризики, розроблюються наступні типи скорингових карт:

- для оцінки ризиків;
- CRM або маркетингові скорингові карти;
- комбіновані.

Найбільш поширеними є скорингові карти для оцінки кредитних ризиків [3]. Це пов'язано з великою кількістю задач, які постають перед банками і потребують періодичного вирішення, та великою кількістю клієнтів та продуктів, для яких необхідна періодична перевірка та відслідковування. Прикладами таких задач є:

- оцінка кредитоспроможності претендентів на отримання нових кредитів (application scoring),
- оцінки ймовірності повернення вже виданих кредитів (behavioral scoring),
- відслідковування можливості повернення кредиту клієнтом при порушенні ним строків та умов погашення кредиту (collection scoring),
- оцінка ймовірності шахрайства з боку існуючого клієнта банку (fraud scoring).

Розробка скорингових карт. Ефективність та якість розробленої методології побудови скорингових карт суттєво залежить від наступних факторів:

- якість доступних даних;
- тип очікуваного результату, тобто двійковий («хороший»/«поганий») або неперервний (прибуток/збитки в доларах США);
- розмір доступних вибірок;
- платформа для впровадження (тобто чи можливе впровадження певної скорингової карти в системі обробки заявок);
- можливість інтерпретації результатів, наприклад зручність застосування бальної системи оцінки скорингових карт, побудованих на основі регресії;
- відповідність методології правовим нормам; зазвичай, наглядові органи вимагають, щоб методи, що використовуються, були прозорими та просто пояснювались;
- можливість відслідковувати і діагностувати ефективність скорингових карт.

Розробка скорингових карт складається з наступних етапів [4]: обробка та очистка даних, початковий аналіз характеристик, побудова попередньої скорингової карти, аналіз відхилених заявок, початковий аналіз характеристик, побудова фінальної скорингової карти (з виконанням

масштабування та оцінки) та виконання валідації.

Для вибору найбільш значущих вхідних змінних в моделі використовуються коефіцієнти Gini та Information Value. Для формування груп значень (групування) в якості критеріїв розбиття діапазону значень на групи використовуються коефіцієнти Weight of Evidence (WOE).

Скорингова карта є візуальним представленням роботи логічних алгоритмів побудови скорингових моделей. В якості методів для побудови скорингових карт використовуються: логістична регресія, нейронні мережі, дерева рішень, мережі Байєса, нечіткі підходи тощо [3, 4]. Скорингова карта є модифікацією експертного підходу, де в якості експертних оцінок використовуються скорингові бали, обчислені за допомогою скорингових моделей наступним чином:

$$\begin{aligned} \text{score} = \log(\text{odds}) \cdot \text{factor} + \text{offset} = -(\sum_{j,i=1}^{k,n} (\text{woe}_j \cdot \beta_i) + a)) \cdot \text{factor} + \text{offset} = \\ = -(\sum_{j,i=1}^{k,n} (\text{woe}_j \cdot \beta_i) + \frac{a}{n})) \cdot \text{factor} + \text{offset} = -(\sum_{j,i=1}^{k,n} (\text{woe}_j \cdot \beta_i) + \frac{a}{n}) \cdot \text{factor} + \frac{\text{offset}}{n}, \end{aligned} \quad (1)$$

де β_i – коефіцієнти регресії;
woe_j – оцінки WOE для атрибутів моделі;
n – кількість характеристик в моделі;
a – відрізок, що відсікається на осі логістичною регресією;
k – кількість груп (атрибутів) для кожної характеристики.

За результатами таких оцінювань отримується таблиця скорингових балів, які присвоюються за кожною характеристикою клієнту, і обраховується сумарний скоринговий бал.

На етапі аналізу відхилених заявок при побудові скорингових карт (Reject Inference) існуюча навчальна вибірка доповнюється за рахунок даних по претендентам, яким було відмовлено у видачі кредиту, при цьому автоматично проводиться розмежування претендентів на позитивні та негативні випадки.

Для оцінки якості побудованої скорингової карти використовуються наступні характеристики: помилки першого та другого роду, ROC-крива або крива Лоренца, вартість неправильної класифікації(втрати внаслідок видачі кредитів "поганим" клієнтам та втрачені можливості внаслідок невидачі кредитів "гарним" клієнтам) тощо. Суттєвим на етапі розробки та у подальшому застосуванні скорингової карти є правильний вибір порогу відсікання моделі, тобто такої оцінки, яка сумарно розраховується по всіх характеристиках клієнта.

Етап валідації скорингової карти передбачає порівняння статистик для розроблених версій скорингових карт та порівняння розподілів хороших та поганих клієнтів для розроблених версій, обрання кращої з них для подальшого використання.

Слід зазначити, що навіть після впровадження скорингової карти процес моніторингу ефективності скорингових карт продовжується, оскільки в процесі її роботи можуть зазнати змін фактори та характеристики, що були включені в скорингову модель, змінитися законодавство, або ключові фактори оцінки. На практиці використання скорингових карт є ефективним інструментом для ризик-менеджерів, оскільки дає наочне представлення щодо рівня ризикованості того чи іншого клієнта чи продукту.

Література. **1.** Кузнєцова Н.В. Практичні підходи до визначення та урахування невизначеностей, що формують фінансові ризики / Н.В. Кузнєцова // Тр. Одес. политехн. ун-та. – Одесса, 2014. – Вып. 2(44). – С. 160–170. **2.** Кузнєцова Н.В. Системний підхід до аналізу кредитних ризиків з використанням мереж Байєса / Н.В. Кузнєцова, П.І. Бідюк // Наукові вісті НТУУ «КПІ». — 2008. — № 3. — С. 11 — 24. **3.** Kiss F. Credit scoring processes from a knowledge Management perspective / F.Kiss // Periodica Polytechnica Ser. Soc .Man. Scl. – 2003. – vol. 11, №1. – P. 95–110. **4.** Siddiqi N. Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring. – 2005. – 196 p.

Левко Л.Р., Андреєва М.О.
Факультет прикладної математики НТУУ "КПІ",; кафедра системного програмування і спеціалізованих комп'ютерних систем, Київ, Україна

Про один метод побудови та модифікації графо-логічної моделі багатопроцесорної відмовостійкої системи, підсистеми якої мають спільні процесори

Системи реального часу можуть бути класифіковані як системи "жорсткого" та "м'якого" реального часу. В системах "жорсткого" реального часу час реакції має бути мінімальним. Порушення реакції на будь-які події в заданий момент часу викликає відмову і зводить нанівець вирішення поставленого завдання. У таких системах важливо, щоб завдання було завершено протягом відведеного часу, навіть у присутності збою. Прикладами систем "жорсткого" реального часу є системи управління космічними станціями, система-автопілот, а також системи моніторингу для пацієнтів у критичному стані. У системах "м'якого" реального часу наслідки є відносно прийнятні. У таких системах реального часу важливіше виявити збій якомога швидше, аніж замаскувати помилку. Тому основне завдання для елементів системи полягає у можливості виявлення помилки швидко та за найменших витрат. Передбачається, що відповідна дія може бути виконана відразу після виявлення несправності. Прикладами систем "м'якого" реального часу є всі види обробки транзакцій онлайн, таких як бронювання авіаквитків, банківські послуги і додатки електронної комерції. Іноді вихід процесору з ладу не є катастрофічним, але все ж таки цей недолік повинно бути усунено якомого скоріше. Наприклад, значий час простою сервера може призвести до зменшення прибутку компаній, які використовують цей сервер. Потрібно відзначити, що в той час як вартість системи зменшується, вартість обслуговування та втрати за час простою збільшуються навпаки. Вихід із ладу одного процесора не є критичним, якщо несправність виявлено виправлено вчасно з мінімальними затратами.

Розрахунок рівня надійності відмовостійких багатопроцесорних систем (ВБС) часто виявляється складною задачею, оскільки системи управління можуть мати складну і неоднорідну структуру за рахунок великої кількості процесорів. Для розрахунку надійності ВБС використовують багато методів, кожен з яких має свої недоліки та переваги [1]. Один з таких методів був запропонований в [2] - це побудова графо-логічної моделі поведінки m-відмовостійкої системи K(m,n).

Графо-логічна модель (GL-модель) – це математична модель, яка описує залежність стану системи (роботоздатний або відмова) від станів елементів системи. GL-модель представляє собою неорієнтований граф, ребрам якого ставляться у відповідність деякі булеві функції (реберні функції). Реберні функції залежать від змінних, які відображають стани елементів системи. Вектор стану системи складається з набору булевих змінних, які відповідають кожному з процесорів системи і приймають значення 1, якщо процесор роботоздатний, і 0 - якщо він вийшов з ладу. В останньому випадку процесор не є ативним, тому відповідне ребро на графі видаляється. Вектор стану системи подається на вхід графо-логічної моделі і визначає, чи залишиться в такій ситуації система працездатною або вийде з ладу. Зв'язність графу зазначає стабільну роботу всієї системи, в іншому випадку - система виходить з ладу.

В [3] розроблені ієрархічні графо-логічні моделі і відповідна методика розрахунку надійності, згідно якої розрахунок здійснюється по всім рівням ієрархії для кожної підсистеми окремо. Основною перевагою «ієрархічної» методики порівняно з «неієрархічною» є значне зменшення необхідної кількості статистичних експериментів. Методика розрахунку надійності, запропонована в [3], враховує окремий випадок ієрархічних GL-моделей, коли системи не мають спільних процесорів.

Розглянемо реконфігуровану відмовостійку багатопроцесорну систему (РВБС), що містить n процесорів.

Введемо наступні позначення:

y – булева змінна, що відображає стан системи (y=0, якщо система відмовила, y=1, якщо система працездатна).

$x_1, x_2, \ldots, x_n$ – булеві змінні, які відображають стани процесорів системи ($x_i = 0$, якщо i-й процесор роботоздатний, i = 1,. . . ,n).

C – множина індексів процесорів системи, тобто C = $\{1, 2, \ldots, n\}$.

X – вектор стану системи, $X = (x_1, x_2, \ldots, x_n)$.

φ – графо-логічна модель поведінки системи:

$y = \varphi(X) = \varphi(x_1, x_2, \ldots, x_n)$.

Графо-логічна модель поведінки системи є ієрархічною, якщо існують такі підмножини $C_1, C_2, \ldots, C_w$ множини C і GL-моделі β, $\varphi_1, \varphi_2, \ldots, \varphi_w$, що

$$\varphi(X) = \beta(\varphi_1(X(C_1)), \varphi_2(X(C_2)), \ldots, \varphi_w(X(C_w)), \quad (1)$$
$$C = C_1 \cup C_2 \cup \ldots \cup C_w,$$

де $X(C_j) = (x_i | i \in C_j)$ - вектор, який складається з булевих змінних x_i, індекси яких належать множині $C_j (i \leqslant j \leqslant w)$.

Тобто ієрархічна GL-модель може бути представлена у вигляді суперпозиції інших GL-моделей. Зазначимо, що GL-моделі φ_j у свою чергу можуть бути ієрархічними, тоді має місце багаторівнева ієрархія. Практично ієрархічна GL-модель деякої системи може бути побудована, якщо система складається з конкретних підсистем, відомі GL-моделі φ_j підсистем і GL-модель β, яка відображає залежність стану системи від станів підсистем. Варто зазначити, що обернена задача, тобто задача пошуку GL-моделей β, φ_1,φ_2,. . . ,φ_w, при відомій GL-моделі φ, у загальному випадку є дуже складною.

Для ієрархічних GL-моделей можна виділити дві цікаві особливості:

1)ієрархічна GL-модель може бути побудована для системи, структура якої не є ієрархічною;

2)для однієї системи може бути побудовано кілька різних ієрархічних GL-моделей у залежності від розбиття множини C на підмножини $C_1, C_2, \ldots, C_w$.

Такі особливості ієрархічних графо-логічних моделей обумовлюють необхідність конкретизації в межах даної роботи поняття підсистеми.

У відмовостійкій базовій системі, для якої побудована ієрархічна GL-модель у вигляді (1), підсистемою будемо називати підмножину процесорів системи, індекси яких належать множині C_j, і яка має GL-модель φ_j.

В доповіді нами запропонована ієрархічна GL-модель, що представлена у вигляді суперпозиції інших графо-логічних моделей φ_j, кожна з яких відповідає базовій багатопроцесорній відмовостійкій системі і описується як K(m,n). Оскільки ці підсистеми мають спільні процесори, для них також існує деяка графо-логічна модель φ_p. У випадку якщо y(φ_p) приймає нульове значення - це може вплинути на стан інших систем. Фактично при y(φ_p) = 0 підсистема φ_p, що складається зі спільних процесорів підсистем, втрачає працездатність, що впливає на працездатність кожної з підсистем φ_j, до складу кожної з яких входять процесори з даної підсистеми φ_p. Отже, подалі у роботі нами буде проведено статичні експерименти, що надасть змогу визначити імовірність виходу з ладу таких систем φ_p, та відповідний вплив на кожну із систем φ_j.

Література. **1.** Мораведж Сейед Милад. Методы и средства оценки технической безопасности и расчёта параметров эффективности тестирования и реконфигурации многопроцессорных систем управления : дисс. ... канд. техн. наук. - 05.13.05 **2.** Романкевич В.О. Про розрахунок надійності відмовостійних багатопроцесорних систем, підсистеми яких мають спільні процесори / В.О. Романкевич, А.П. Фесенюк // Радіоелектронні і комп'ютерні системи. – 2010. - No 3. – С. 62-67. **3.** Романкевич В.А. Объединение моделей подсистем в рамках графо-логической модели / В.А. Романкевич, А.В. Темноход // Збірник тез 2 Всеукраїнської конференції "Людина і космос". – Дніпропетровськ, 2000. - С. 273. **4.** В.А.Романкевич. Об одном способе построение GL-моделей відмовостійних багатопроцесорних систем / В.А.Романкевич // Радіоелектронні і комп'ютерні системи. – 2006. - No 7. – С. 47-51.

Ліщук К.І.
Національний технічний університет України "Київський політехнічний інститут", Київ, Україна

Оцінка якості програмних продуктів з використанням експертних методів

Розробка програмного забезпечення на сьогодні є однією з найбільш прибуткових діяльностей. Будь-які порушення при розробці (помилки, котрі невиявлені під час тестування, неефективні інтерфейси, тощо) можуть призвести до небажаних результатів, котрі призведуть до вибору покупцем більш якісного програмного забезпечення конкурентів. Тому, одним з найбільш актуальних питань сучасної індустрії програмного забезпечення є оцінка якості.

Визначення якості програмного забезпечення є достатньо складною задачею, котра може трактуватися по-різному в залежності від ситуації та стадії життєвого циклу програмного забезпечення, не дивлячись на наявність стандартів, котрі описують оцінювання якості на різних стадіях життєвого циклу програмного забезпечення. Розробка програмного забезпечення (ПЗ) потребує контроль за якістю функціонування на кожному із запланованих етапів розробки. Стеження за показниками якості на протязі всього процесу розробки і супроводу ПЗ надає можливість контролювати і покращувати кінцевий продукт. Методи забезпечення контролю якості ПЗ на всіх етапах розробки, націлені на підвищення ступеня задоволення клієнтів від покращення якості і зручності застосування на виробництві. Професійний підхід до управління якістю базується на тому, що якість це чітко визначена величина, яку можна виміряти і проконтролювати, вона піддається управлінню і поліпшенню [2]. Існує безліч різних моделей для оцінки якості програмного забезпечення (МакКола [2], Боема [3], Гилба [4],ISO-9126 [5],Маркова [6], тощо) та їх модифікацій, однак результати більшості з них можуть трактуватися по-різному в залежності від ситуації [7].

Запропонований метод та програмна система підтримки прийняття рішень, котра його реалізує, використовується на етапі оцінки якості програмного забезпечення під час його експлуатації. При оцінці якості експертам пропонується виконати оцінку якості програмного забезпечення за певним набором критеріїв, котрі найкращим чином зображають за різними критеріями. Крім того, набір критеріїв, котрі використовуються при оцінці якості так само визначаються експертами. Для аналізу та оцінки характеристик якості ПП по сукупності критеріїв пропонується використовувати модифікації методу аналізу ієрархій (MAI) Т.Сааті. Необхідно відмітити, що кількість можливих характеристик та критеріїв, за якими виконується оцінка, достатньо велика, тому використання класичного методу аналізу ієрархій (MAI) є неможливим. Тому для розв'язання задачі пропонується використовувати одну з модифікації MAI, запропоновані Павловим О.А., Ліщук К.І. [8].

Література. **1.** Kelemen,Z.D.,Kusters,R. and Trienekens,J.(2012),Identifying criteria for multi-model software process improvement solutions – based on a review of current problems and initiatives. J. Softw. Evol. and Proc., 24: 895–909. doi: 10.1002/smr.549. **2.** McCall J.,Richards P.,Walters G. Factors in Software Quality. Three volumes:NTIS AD-A049-014, AD-A049-015, AD-A049-055, 1977. **3.** Боэм Б. и др. Характеристики качества программного обеспечения. – М.: Мир, 1981. – 208 с. **4.** Gilb T. Principles of Software Engineering Management. – Reading MA: Addison Wesley, 1988. – 464 p. **5.** International Standard ISO/IEC 9126. Information technology – Software product evaluation – Quality characteristics and guidelines for their use. International Organization for Standardization, International Electrotechnical Commission, Geneva, 1991. **6.** Марков А.С.,Годердзишвили Г.М. Модель оценки степени отлаженности программного обеспечения по результатам испытаний // Методы и средства управления и контроля. Л.:МО СССР,1987.С.51-52. **7.** Кривуля Г.Ф.,Шкиль А.С., Кучеренко Д.Е.(2013).Экспертное оценивание качества программного обеспечения. Радіоелектронні і комп'ютерні системи, (5), 282-286. **8.** Павлов О.А. Модифікований метод аналізу ієрархій (версія 1,2)/О.А. Павлов, К.І. Ліщук, О.С.Штанькевич, Г.А.Іванова, О.П. Федотов//Вісник НТУУ "КПІ".–К.:"ВЕК+", 2009 – №51.

Лопатин А.К.[1,2] ***Черненко О.Б.***[1]
[1] *Национальная академия управления;* [2] *УНК «ИПСА» НТТУ «КПИ», Киев, Украина*

Индикаторы общеэкономической динамики в задачах предсказания кризисов

В работе [1] выделены статистические циклы в эталонном диффузионном индексе PMI (Еврозона) методом нормированного размаха Херста. В работе [2] построены вероятностные модели оценки фазы финансовых рынков. В настоящей работе разработаны индикаторы общеэкономической динамики в задачах предсказания кризисов на примере мирового финансового кризиса 2007-2009 годов.

Решены следующие задачи:

Задача 1. Разработаны алгоритимы выделения статистических циклов в числовых рядах, которые соответствуют структурным перестройкам системы.

Задача 2. Разработаны идикаторы общеэкономической динамики раннего обнаружения кризиса в экономической системе.

В качестве объекта исследования выбран индекс PMI (Еврозона) на период мирового финансового кризиса 2007-2009 годов. Особенность-малое количество данных ($<$120).

Задача 1. Строится кривая нормированного размаха для индекса PMI_EURO. Кривую нормированного размаха можно представить как кусочно-линейную функцию. Каждому линейному отрезку соответствует локальный показатель Херста (тангенс угла наклона этой линии). Суть алгоритма построения границ циков: строится кривая средних значений нормировнного размаха (mu_herst_EURO). По средним значениям нормировнного размаха строится статистика tan. Локальные экстремумы построенной функции ответствуют началу или концу статистического цикла. На рисунке приведен окончательный результат.

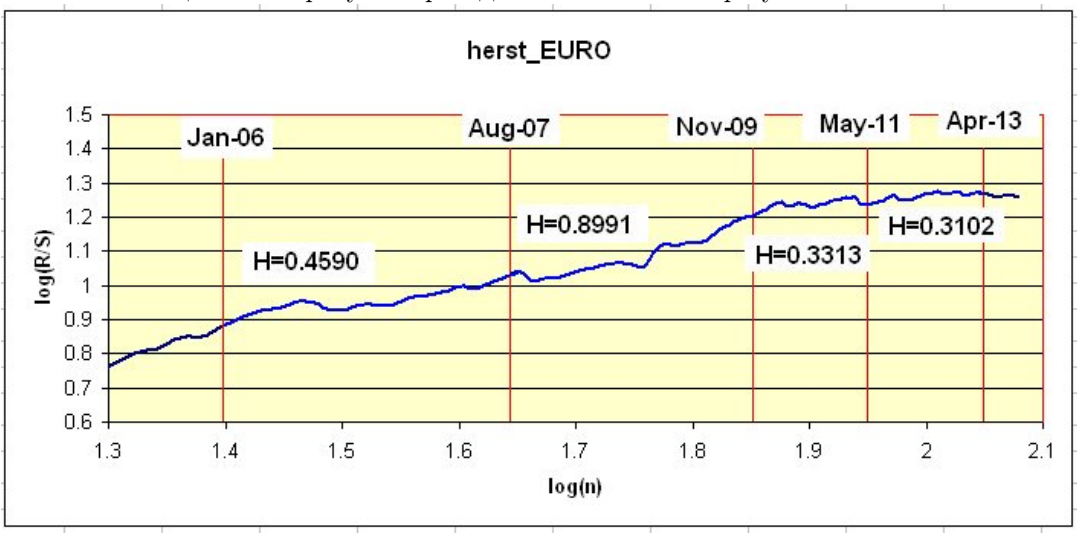


Рисунок 1. Показатель Херса на статистических циклах диффузионного индекса EURO

Задача 2. Экономичекая интерпретация показателя Херста ($0 < H < 1$). Показатель Херста описывает эффективность или неэффективность финансового рынка. Выделяется по меньшей мере 4 различные области [3].

1 область: $H = 1/2$. Это так называемый белый шум. Рынок, который ведет себя в соответствии с этим равенством называется математически эффективным рынком (характеризуется математической теорией мартингалов).

2 область: : H принимает значения в области $H = 1/2 + \varepsilon$, где величина $\varepsilon > 0$ мала и обычно не превосходит $0,03$. Это ведет к исчезновению тенденций и к возвращению рынка в состояние мартингальности. Такой рынок называется экономически эффективным рынком.

Модель экономически эффективного рынка дает лучшее приближение к реалиям рынка по сравнению с математически эффективным рынком.

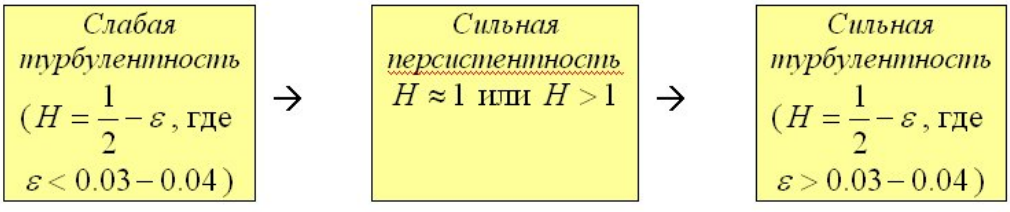
3 область: $H < 1/2 - \varepsilon, \varepsilon > 0$ Случай соответствует турбулентному рынку. В финансовом плане турбулентность означает невозможность статистически обоснованных предсказаний.В зависимости от места нахождения актива его цена на турбулентном рынке может принимать перемежающиеся случайные значения в весьма широком диапазоне.

4 область: H принимает значения в области $H = 1/2 + \varepsilon > 0$, где величина $\varepsilon > 0.03$. Для этой области справедливы сильные последействия и память.

В заключении приведем итоговую таблицу индикаторов общеэкономической динамики раннего обнаружения кризиса в экономической системе.

Цикл 1	Январь_2006-Август_2007	0.4590	*Слабо турбулентный* (перемежающиеся случайные значения в низком диапазоне)
Цикл 2	Август_2007-Ноябрь_2009	0.8991	*Сильно персистентный*(сильные последействия и память)
Цикл 3	Ноябрь_2009-Май_2011	0.3313	*Сильно турбулентный* (перемежающиеся случайные значения в широком диапазоне)
Цикл 4	Май_2011-Апрель_2013	0.3102	*Сильно турбулентный* (перемежающиеся случайные значения в широком диапазоне)

Наличие финансового кризисна характеризуется такой последовательностью статистических циклов



Как видно из приведенной таблицы мировой финансовый кризис 2007-2009 годов имеет предварительную (скрытую) фазу: появление слабой турбулентности H=0.4590. Основная фаза кризиса на втором цикле характеризуется сильной персистентнотью (H=0.8991). Это является результатом поведения «толпы» на финансовом рынке. Неожиданный результат дают третий и четвертый циклы: здесь имеет место сильная турбулентность ($H = 0.3313$ и $H = 0.3102$ на третьем и четвертом циклах соответственно). По-видимому, перестройка мирового финансового рынка после кризиса достаточна сложна и имеет относительно долгий период. Сделанные выводы касаются конкретного события—мирового финансового рынка 2007 2009 годов.

Вывод: Для рассмотренного конкретного события приближение финансового кризиса можно спрогнозировать за 0.5-1 года по обнаружению слабой турбулентности рынка.

Литература. **1.** Лопатин А.К., Черненко О.Б., Построение статистических циклов на основе метода нормированного размаха Херста, Матеріали 13 Міжнародної науково-технічної конференції «Системний аналіз та інформаційні технології», 27-31 травня 2013 року, Київ, 127-128. **2.** Лопатин А.К., Черненко О.Б., Лопатин А.К., Черненко О.Б. , Вероятностные модели оценки фазы финансовых рынков на основе методов нелинейной динамики, Матеріали 17 Міжнародної науково-технічної конференції «Системний аналіз та інформаційні технології», 26-30 травня 2015 року, Київ, 82-83. **3.** Костюк В. Н., Два вида финансовых рисков, Труды Института системного анализа РАН, Москва, 2008, 36, No.1, 56-60.

Мазурік О.Ю.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Порівняльний аналіз моделей оцінювання в рекомендаційних системах

На сучасному ринку IT існує безліч проектів як комерційних, так і відкритих. Але майже всі ці проекти об'єднує одна мета: залучити якомога більше користувачів до нього та змусити їх користуватися послугами якомога частіше. Цю проблему вирішують рекомендаційни системи. Їх призначення: надати користувачу якомога більше релевантної інформації.

Для оцінювання товарів в рекомендаційних системах використовують різні типи систем оцінювання. Найпоширеніші з них: 5-бальна, 10-бальна, бінарна (сподобалось/не сподобалось).

Вибір способу оцінок товару впливає на подальший розвиток проекту та прибуток від нього. Також він суттєво впливає на роботу алгоритму рекомендаційної системи. Отже, необхідно проаналізувати усі типи систем оцінювання та обрати найкращий.

Рекомендаційні системи та алгоритми у свій час започаткувала та активно розвивала компанія Netflix. На свому сайті вони надали інформацію, як саме користуватися такою системою [1]:

★☆☆☆☆ Зовсім не сподобався
★★☆☆☆ Не сподобався
★★★☆☆ Сподобався
★★★★☆ Дійсно сподобався
★★★★★ В захваті

Такий підхід перейняли багато інших відомих компаній, які надають послуги: Amazon, eBay та ін. Але він дуже суб'єктивний. Що може означати "дійсно сподобався" фільм? Чому інтервали між різними операціями нерівні (немає опції "дійсно не сподобався")? Тож, можна зазначити що навіть текст, який несе на своїй меті допомогти виставити рейтинг товару, містить в собі суб'єктивність.

Незважаючи на росповсюдженість п'ятибальної шкали користувачи несвідомо зводять свій вибір до оцінювання бінарним підходом. В 2009 році Youtube (відео-платформа від Google) поширила статистику оцінок користувачів, серед яких найпопулярнішими оцінками стали "5" (∼70%) та "1" (∼25%). Це є прекрасним прикладом того, що бінарне оцінювання вже працює. Адже користувачам значно зручніше та швидше визначити: чи сподобався їм товар, чи ні.

Але не все так добре, як можна помітити в статистичних матеріалах. Бінарний підхід також містить декілька недоліків. Якщо користувачу не сподобався товар чи послуга, то існує велика вірогідність, що він зовсім не буде ставити оцінку. Така система - це система типу "все або нічого". Але користувачі, які все таки ставлять оцінки дуже серйозно відносяться до своїх рекомендацій та дуже зацікавлені в тому, щоб система пропонувала їм релевантні результати.

Деякі проекти (в основному - соціальні мережі) долають проблему відмови від голосування (такі як VK, Facebook). Якщо опустити деякі деталі, то якийсь товар буде оцінено в +1, якщо ви проголосували позитивно, та -1, якщо ви не звернули увагу на нього. Це є різновид бінарного голосування, який теж гарно працює в сучасних умовах (в основному його використовують СМІ) [2].

Отож, незважаючи на велику кількість способів оцінити товар чи послугу в проекті, краще це робити за допомогою бінарного методу, оскільки він не містить недоліків та суттєво знижує суб'єктивність оцінки. Також завдяки бінарному підходу відпадає проблема "холодного старту" (рекомендації новим користувачам) та значно покращується робота алгоритму (менше даних для обробки - більша швидкість).

Література. **1.** David Celis. (2012, Feb. 2). *Collaborative Filtering With Likes and Dislikes* [Online]. Available: https://davidcel.is/posts/collaborative-filtering-with-likes-and-dislikes/ **2.** David Celis. (2012, Feb. 7). *Why I Hate Five-Star Ratings* [Online]. Available: https://davidcel.is/posts/why-i-hate-five-star-ratings/

Марченко О.І., Молчанов О.А.
Факультет прикладної математики НТУУ "КПІ", Київ, Україна

Імітаційне моделювання TCP потоків: еволюційно-ігровий підхід

Розвиток комп'ютерних мереж тісно пов'язаний із конкурентним розвитком різних версій мережевих протоколів. Найбільш вживаним протоколом на сьогодні є протокол TCP. Ядро протоколу базується на двох алгоритмах: *керування потоком (flow control)* та *керування переповненням (congestion control)*. У той час як перший регулює передачу даних через мережу в залежності від стану готовності приймача, другий контролює передачу пакетів в залежності від стану самої мережі. Поточний *стан мережі*, за умови її спільного використання багатьма *користувачами (кінцевими користувачами, end users)*, визначається завантаженістю її *вузьких місць (bottleneck)*, які ці користувачі спільно використовують. Ідея керування переповненням досить проста: збільшити швидкість, якщо мережа недозавантажена, або зменшити швидкість, якщо перевантажена. Момент зміни швидкості визначається на основі інформації, яку користувач може отримати про те, чи дійшов його пакет (за допомогою пакетів підтвердження ACK), або про те, що мережа перевантажена (якщо він не отримає підтвердження протягом певного часу).

Прямими реалізаціями основної ідеї керування переповненням є такі алгоритми, як *адитивне збільшення/мультиплікативне зменшення (АЗМЗ, additive increase/multiplicative decrease - AIMD)* та *мультиплікативне збільшення/мультиплікативне зменшення (МЗМЗ, multiplicative increase/multiplicative decrease - MIMD)*. Однак, окрім цих основних алгоритмів, в реальних системах також використовують допоміжні алгоритми, такі як *швидке відновлення (fast recovery)*.

Концепція еволюційних ігор є частиною теорії ігор, яка зосереджена на вивченні взаємодії між популяціями, а не окремими гравцями. Однією з ключових концепцій, що вивчають еволюційні ігри, є *еволюційно-стабільна стратегія (ECC, ESS - Evolutionary Stable Strategy)*, яку визначив біолог Мейнард Сміт [3]. Вона характеризується властивістю стійкості до загарбників (мутацій, змін). Є два критерії стійкості стратегії, якщо вона досягнута:

1) пропорції популяцій не змінюються у часі
2) популяції захищені від нападу інших малих популяцій

Якщо описувати взаємодію користувачів мережі у термінах еволюційної гри, то можемо дати наступні визначення. *Грою* називається взаємодія користувачів мережі, при якій вони обирають певну реалізацію протоколу TCP. Кожен користувач мережі — це *гравець*. Він може обрати реалізацію протоколу TCP, яку він використовуватиме, тобто вибрати певну *стратегію*. *Виграшем* у грі називатимемо дотримання такої стратегії, за якої будь-яка дія гравця вже не може покращити значення його функції корисності. Прийнявши такі визначення, можна розглянути модель еволюційної гри між різними реалізаціями протоколу TCP.

У роботах [1,4] розглянута модель еволюційної гри між двома користувачами мережі, що використовують різні реалізації протоколу TCP (популяції). Було показано, що така взаємодія формує модель гри, відомої як гра Hawk-Dove (HD-гра), а динамічне рівняння еволюційної гри має стабільне рішення, тобто ECC (яка визначається згідно з [3]).

Можливі значення функції корисності для першого користувача в HD-грі у загальному вигляді можуть бути представлені матрицею, показаною на рис. 1. Для другого користувача ці значення є симетричними.

		Гравець 2 H	Гравець 2 D
Гравець 1	H	A11	A12
	D	A21	A22

Рис. 1. Загальне представлення HD-гри у матричній формі для першого користувача

HD-гра, представлена у загальному вигляді, характеризується співвідношеннями A11<A22<A12 і A21<A22, причому можливі дві ситуації:

1. Якщо A11>A21, то чиста стратегія H є єдиною ЕСС.
2. Якщо A11<A21, то унікальна ЕСС $p = (p_D, p_H)$ є змішаною стратегією, що визначається виразом (1).

$$p = (p_D = \frac{v}{v+u}, p_H = \frac{u}{v+u}), \tag{1}$$

де $u = A12 - A22$, $v = A21 - A11$, $Aij = J(i, j)$, $i, j \in H, D$, $J(i, j)$ — функція корисності, H, D – стратегії гравців (“агресивна” і “мирна” відповідно).

Значення функції корисності для користувачів мережі може бути визначене за формулою (2).

$$J_i = K \cdot (Thp_i(t) - \lambda\frac{1}{T_i}), \tag{2}$$

де K — коефіцієнт збільшення значення корисності; Thp_i — швидкість надсилання пакетів для даної стратегії на момент часу t; $\lambda\frac{1}{T_i}$ — швидкість втрати пакетів; T_i — інтервал часу між моментами зменшення швидкості (при виявленні перевантаження мережі); λ — коефіцієнт чутливості до втрат пакетів.

Розглянута модель має обмеження, яке полягає в тому, що за час Ti втрачається лише один пакет, після чого виконується зменшення *вікна переповнення (congestion window)*, яке регулює швидкість посилання пакетів, і відбувається повторне надсилання пакетів, починаючи з втраченого. Зазначимо також, що в даному випадку розглядається симетрична втрата пакетів (тобто, відкидання пакетів маршрутизатором трапляється одночасно для обох користувачів), тому час T_i є однаковим для всіх i.

Для перевірки моделі гри була розроблена імітаційна модель мережі, що складається з чотирьох вузлів: двох користувачів, які надсилають дані на сервер, сервера, що приймає пакети, і маршрутизатора, що з’єднує користувачів з сервером. Канал між маршрутизатором і сервером є саме тим вузьким місцем мережі, ресурс якого користувачі поділяють між собою (рис. 2).

Рис. 2. Модель мережі

Користувачі можуть обрати одну з двох реалізацій протоколу, що є модернізованими варіантами TCP Tahoe [2]. На відміну від оригіналу, в даних реалізаціях параметри α і β є змінними в алгоритмі АЗМЗ, де α відповідає кількості пакетів, на яку збільшується вікно переповнення після вдалої пересилки пакетів попереднього вікна, а β — це величина, у які зменшується вікно після втрати пакету внаслідок відкидання цього пакету маршрутизатором. Перший варіант має параметри $(\alpha_1, \beta_1) = (1, 0.5)$, а другий — $(\alpha_2, \beta_2) = (2, 0.6)$, що відповідає “мирній” та “агресивній” стратегіям.

На рис. 3 зображені показники статистики проходу пакетів користувачів через спільний канал, що були отримані в результаті моделювання.

Отриманий розподіл пакетів користувачів у спільному каналі відповідає умовам HD-гри. Коли обидва користувачі використовують “мирну” стратегію, вони порівну поділяють канал

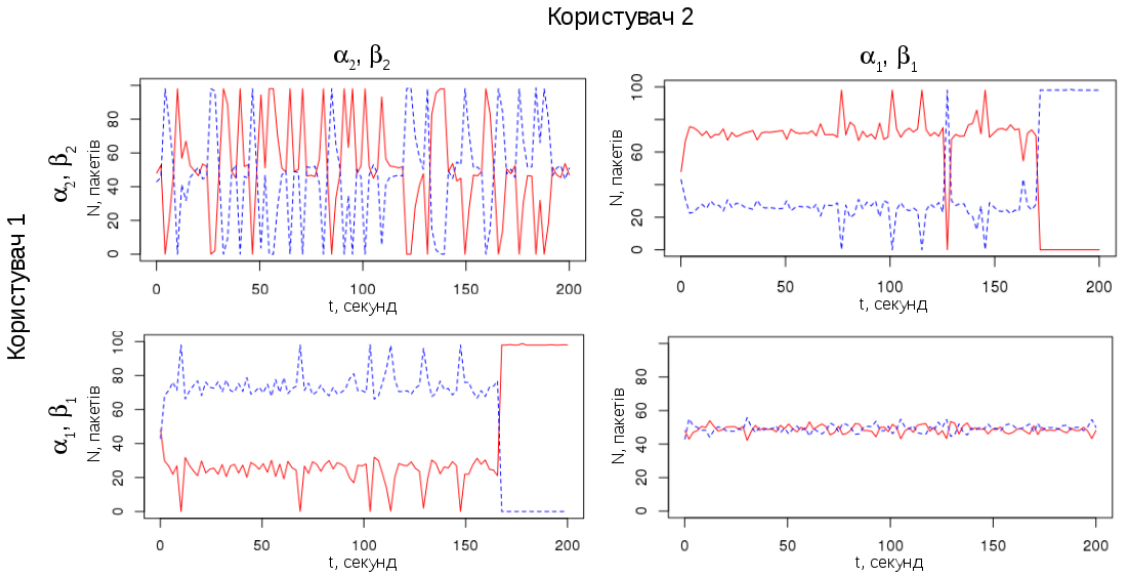


Рис. 3. Статистика проходу пакетів через вузьке місце мережі (суцільна лінія — перший користувач, пунктирна лінія — другий користувач)

між собою, коли "агресивну" — поділ відбувається також порівну, але нестабільно. Коли обидва користувачі використовують різні стратегії, виграш отримує "агресивніша". У даному випадку спостерігається співіснування обох стратегій з перевагою "агресивнішої".

Під час моделювання також були отримані показники функції корисності, наведені у табл. 1.

Табл. 1. Значення функцій корисності отримані моделюванням моделі мережі

		Користувач 2	
		$\alpha_2,\ \beta_2$	$\alpha_1,\ \beta_1$
Користувач 1	$\alpha_2,\ \beta_2$	$J_1 = 194.02$ $J_2 = 205.78$	$J_1 = 402.84$ $J_2 = 159.35$
	$\alpha_1,\ \beta_1$	$J_1 = 163.84$ $J_2 = 395.11$	$J_1 = 268.87$ $J_2 = 269.59$

Отримані результати дають змогу пересвідчитись, що взаємодія у мережі кінцевих користувачів, які використовують різні реалізації протоколу TCP, відповідає моделі HD-гри. Це дозволяє використовувати дану модель для представлення еволюційної гри в мережі, а саме для розгляду ситуацій динамічної зміни реалізацій TCP під час взаємодії, в залежності від корисності, яку принесе така зміна.

Література. **1.** E. Altman, R. El-Azouzi, Y. Hayel, H. Tembine, "The evolution of transport protocols: an evolutionary game perspective", Computer Networks Journal, 2009, pp. 1751–1759. **2.** A. Afanasyev, N. Tilley, P. Reiher, L. Kleinrock, "Host-to-Host Congestion Control for TCP", IEEE Communications Surveys & Tutorials 12, 3 (2012), pp. 304–342. **3.** J. Maynard Smith, "Game theory and the evolution of fighting", John Maynard Smith (Ed.), On Evolution, Edinburgh University Press, 1972, pp. 8–28. **4.** O. Ignatenko and O. Synetskyi, "Evolutionary Game of N Competing AIMD Connections", Information and Communication Technologies in Education, Research, and Industrial Applications, Springer International Publishing, 2014, pp. 325–342.

Маслянко П.П., Прудський Ю.Ю.
Кафедра пикладної математики НТУУ "КПІ", Київ, Україна

Компонентна модель уніфікованої системи оптимізації складу сумішей для ентерального харчування

Вступ. Протягом багатьох років спеціальне харчування широко використовується в стаціонарах як засіб нутритивної підтримки і терапії різних захворювань. Але навіть після виписки з лікарні людина, що одужує, потребує збалансованого лікувального харчування. Потреби організму не обмежуються тільки білком або будь-якими вітамінами - це завжди комплекс речовин в певній пропорції. Для вирішення цієї проблеми використовуються ентеральні суміші [1].

Ентеральне харчування — тип лікувального або додаткового харчування спеціальними сумішами, при якому всмоктування їжі (при її надходженні через рот, через зонд в шлунку або кишечнику) здійснюється фізіологічно адекватним шляхом, тобто через слизову оболонку шлунково-кишкового тракту [2]. Можливість скласти збалансовану за складом ентеральну суміш ускладнюється кількістю факторів, що впливають на вимоги до цієї суміші. Серед таких факторів: антропометричні показники організму людини, рівень та тип її фізичної активності, тип захворювання, алергічні реакції чи індивідуальна непереносимість певних речовин. Також важливим є те, що засвоювання одних мікронутрієнтів впливає (позитивно чи негативно) на засвоювання інших. На даний час немає системи, що робила б усі розрахунки автоматично, з урахуванням вищенаведених факторів.

Постановка задачі. Об'єктом дослідження є система процесів метаболізму в організмі людини, моделі визначення денної норми калорій, особливості засвоєння вітамінів та мікроелементів, особливості впливу віку людини та типу фізичної активності на процеси метаболізму, методи оптимізації складу ентеральних сумішей, моделі дієтичного харчування при захворюваннях шлунково-кишкового тракту (ШКТ), графічні засоби мови UML для представлення структурних і динамічних моделей, існуючі рішення для генерації рецептів поживної суміші, існуючі рішення для визначення денної норми поживних речовин.

Предметом дослідження є уніфікована модель розрахунку оптимальних параметрів лікувальної поживної суміші для лікування таких захворювань ШКТ, як панкреатит, холецистит та гастрит з використанням генетичних алгоритмів, з урахуванням антропометричних показників організму людини та рівня її фізичної активності.

Метою дослідження є забезпечення одужання людини шляхом моделювання та розрахунку складу ентеральних сумішей для лікування захворювань ШКТ з урахуванням індивідуальних параметрів людини.

У результаті дослідження встановлено, що для забезпечення урахування усіх індивідуальних особливостей людини при її лікуванні, актуальною э задача створення системи оптимізації складу ентеральної суміші. Пропонується використовувати уніфіковану систему, структурна схема якої представлена на рис. 1.

Опис компонентів системи. Аналіз даних користувача – пакет компонентів, загальною метою якоих є визначення параметрів уніфікованої моделі метаболізму користувача, яка буде використовуватись для визначення денної норми поживних речовин.

Система управління - компонент, що відповідає за взаємозв'язок між іншими компонентами системи та забезпечує роботу з базою даних.

База даних - компонент, що відповідає за зберігання інформації про калорійність, макронутрієнтний (білки, жири, вуглеводи) та мікронутрієнтний (вітаміни і мінерали) склад продуктів.

Інтерфейс користувача - компонент, що дозволяє користувачу системи вводити особисті дані, які передаються на обробку до пакету аналіза даних користувача за допомогою системи управління.

Визначення денної норми поживних речовин - сукупність двох компонентів: розрахунок норми мікронутрієнтів та розрахунок норми КБЖВ (калорій, білків, жирів та вуглеводів). Ці

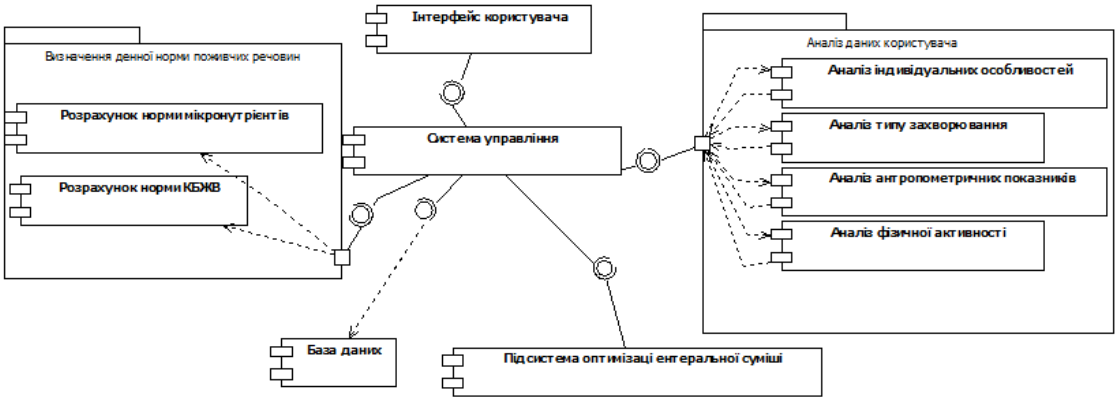


Рис. 1. Компонентна модель уніфікованої системи оптимізації складу ентеральних сумішей

величини розраховуються на основі розробленої моделі та з використанням інформації про особливості засвоєння вітамінів та мінералів.

Підсистема оптимізації складу ентеральної суміші – компонент, що відповідає за створення оптимального рецепту ентеральної суміші. Алгоритмом оптимізації є генетичний алгоритм, з наступними особливостями:

1. Початкова популяція складається зі 100 особин.
2. Селекція проводиться за методом «рулетки»:

$$p_i = \frac{f_i}{\sum_{i=1}^{N} f_i}, \text{ де} \tag{1}$$

p_i - ймовірність вибору i особи, f_i - значення функції пристосування для i особи, N - кількість особин у популяції [3]

3. Початкова популяція складається зі 100 особин.
4. Розмноження реалізується за допомогою оператора схрещування.
5. Різноманітність в особах досягається мутаціями.

Висновки. Було оглянуто основні проблеми галузі лікувальної дієтології та запропонована модель уніфікованої системи, що їх вирішує. Розроблена система відрізняється від існуючих урахуванням особливостей взаємодії мікроелементів при сумісному вживанні та урахуванням найважливіших індивідуальних особливостей організму людини: вид захворювання, антропометричні показники, індивідуальна непереносимість та алергії, рівень та тип фізичної активності.

Було встановлено, що уніфікована система оптимізації складу сумішей для ентерального харчування є зручним і ефективним рішенням проблеми лікувальних дієт. Похибка разрахунків при оптимізації раціону не перевищує 2%, тоді як кращий результат серед існуючих систем – 4%. Можливі подальші шляхи розробки:

- Збільшення спектру хвороб, до яких може бути застосована створена система.
- Охоплення більшої кількості вікових категорій.
- Врахування виду діяльності людини.

Література. **1.** Згржебловская Л.В. Роль подготовки гастро-интестинального тракта в раннем энтеральном питании у пострадавших с тяжелой политравмой // Проблеми військової охорони здоров'я. – Київ: «Янтар», 2002. – С. 682-691. **2.** Г.П. Арутюнов Методические рекомендации: Энтеральное питание в лечении хирургических и терапевтических больных // МЗСР РФ 08.12.2006, №6530-РХ, С. 5. **3.** Емельянов В. В., Курейчик В. В., Курейчик В. М. Теория и практика эволюционного моделирования. — М: Физматлит, 2003. — С. 432.

Мигаль Г.В., Мигаль В.П.
Национальный аэрокосмический университет им. Н.Е.Жуковского "ХАИ Харьков, Украина

Концептуальные основания системного подхода к исследованию и анализу безопасности сложных технических систем

Ключевой проблемой безопасности сложных технических систем (СТС) в экстремальных условиях эксплуатации является динамическая индивидуальность (ДИ). Она приводит к локальным нарушениям совместимости подсистем и, как следствие, не согласованности их функционирования, что является одной из причин аварий и техногенных катастроф. ДИ свойственна как сенсорам разной природы, так и системам организма человека [1]. Она является системообразующим фактором, в котором скрыты тонкие (информационные) особенности сигналов функционирования, а также в функциональных характеристиках технических и биологических подсистем. Поэтому информативными оказались представления физиологических сигналов (ЭКГ и др.) и откликов сенсоров технических систем в виде графических образов (фазовых портретов, диаграмм комплексной плоскости [2] и др.). Однако многообразие методов обработки информации и способов ее представления, разнообразие параметров, показателей и моделей породили новые проблемы – однозначности и др. При сглаживании сигналов и их усреднении теряется важная информация, вследствие чего анализ комплексных исследований безопасности СТС неоднозначен и противоречив. Поэтому целью работы являлся поиск концептуальных оснований системного исследования разных по природе сигналов функционирования подсистем СТС.

Динамическая индивидуальность наиболее проявляется в фазовых портретах, диаграммах и др. графических образах сигналов функционирования. Следовательно, посредством геометризации динамики функционирования подсистем СТС основные принципы системного анализа можно естественным образом совместить с фундаментальными законами и принципами динамики. Следуя методологическому принципу единства и борьбы противоположностей, можно предположить, что индивидуальность формируется в переходном процессе от динамического подобия к динамической индивидуальности. Этот процесс можно обобщенно рассматривать как переход от динамического порядка к беспорядку, который сопровождается потерей энергетического баланса. Взаимосвязь этих процессов определяет динамику функционирования подсистем, скрытая структура которых технологически или генетически унаследованы. В структуре сигнала функционирования скрыто влияние двух взаимодополняющих принципов. Первый из них – принцип механического детерминизма, определяющий причинно-следственные связи в сигнале, как в последовательности динамических событий. Второй – принцип структурной детерминации, отображающий связь структуры и функций. Взаимосвязь этих принципов, а, следовательно, и структура цикла управления, скрыты во взаимных корреляциях динамических параметров сигнала функционирования X(t), его скорости dX(t)/dt и ускорения d2X(t)/dt2. В пространстве (состояние – скорость – ускорение) каждое событие можно отобразить точкой с координатами (X, dX/d, d2X/dt2). Поэтому сигнал функционирования X(t) любой подсистемы можно представить в виде сложной траектории причинно-следственно связанных между собой динамических событий [3]. Траектории разных сигналов состоят из криволинейных участков, которые отличаются длиной и линейной плотностью динамических событий. При этом ДИ пространственно-временных корреляций динамических событий на этих участках наиболее проявляется в трех проекциях этой траектории на ортогональные плоскости. Все конфигурации сигнатур представляют собой замкнутые последовательности из геометрически упорядоченных участков постоянной крутизны или кривизны. Это отображает естественную декомпозицию сигнала на составляющие, физический смысл которых определяется порядком сигнатуры. Так, проекция на плоскость (состояние–скорость) является сигнатурой сигнала 1-го порядка X–dX/dt, в конфигурации которой отображаются все фазы функционирования подсистемы. При этом, площадь сигнатуры можно представить в виде мощности подмножества динамических микросостояний |W|, а натуральный лога-

рифм |W| является энтропией H [4]. Проекция на плоскость (состояние–ускорение), является сигнатурой 2-го порядка X–d2X/dt2, в которой проявляются особенности энергетического баланса противофазных процессов в сигнале. Наиболее информативна проекция на плоскость (скорость–ускорение), которая является сигнатурой 2-го порядка dX/dt – d2X/dt2. Ее конфигурация расположена в 4 квадрантах плоскости, а охватываемые площади отображают мощности основных фаз структуры цикла управления функционированием подсистемы. Поэтому эту сигнатуру можно рассматривать как геометрическую модель структуры цикла управления [5]. Важно, что конфигурации трех сигнатур сигнала любой природы можно анализировать с помощью универсальных дифференциально-геометрических параметров составляющих и интегративных показателей их упорядоченности и сбалансированности.

Наполним конкретным содержанием основные принципы системного подхода (абсолютного приоритета, единства, связанности, модульности, иерархии, функциональности и развития). Так, абсолютным приоритетом подхода является учет ДИ, которая присуща конфигурациям сигнатур, их площадям и составляющим. Так, все сигнатуры состоят из замкнутой последовательности геометрически упорядоченных участков, распределение которых в конфигурации сигнатур индивидуально. Это позволяет в соответствии с принципом единства анализировать подсистему в целом и как совокупность составляющих. Все сигнатуры и их составляющие взаимосвязаны, на что ориентирует принцип связанности. Конфигурации сигнатур располагается в квадрантах плоскости, которые выделяют основные фазы сигнала, отображая тем самым принцип модульности. Поэтому в рамках подхода принципы единства, связанности, модульности естественным образом взаимосвязаны и составляют ядро концепции. Другие принципы (функциональности, развития и неопределенности) приобретают конкретное содержание при пакетном представлении сигнатур последовательности сигналов. Так, принцип развития проявляется в характере перестройки цикла функционирования системы [6]. Изменения конфигурации и площади сигнатур взаимосвязаны, что позволяет выявлять влияние на безопасность различных стресс-факторов среды и деятельности. Характер изменения площади сигнатуры X–dX/dt в экстремальных условиях можно анализировать посредством операций над подмножествами микросостояний. Поэтому в характере перестройки конфигурации и площади сигнатур естественным образом отображается связь структуры и функций, которая проявляется в зависимостях энтропии от времени H(t)) и ее производной dH(t)/dt. Так, согласованное изменение конфигурации сигнатур и энтропии H свойственно биологическим системам. Для их исследования впервые предложены H-сигнатуры H(t) – dH/dt, в пакетах которых естественным образом проявляются переходы порядок-беспорядок. Унаследованная связь структуры и функций скрыта в перестройке геометрической модели структуры управления, характер которой проявляется в пакете F-сигнатур F(t) – dF/dt. Их применение упрощает структурно-функциональный анализ функционирования СТС.

Таким образом, принципы системного анализа естественным образом проявляются в сигнатурах 1-го и 2-го порядков. Это позволяет системно исследовать безопасность функционирования СТС в экстремальных условиях.

Литература. **1.** Mygal, V.P., Klimenko, I.A., Mygal, G.V. et al. Application of parametric and wavelet-signatures for sensor diagnostics. Radioelectronic and computer systems, vol. 36, pp. 143–148, 2009. **2.** Komar, V.K., Migal, V.P. et al. Investigation of localized stated in cadmium zinc telluride crystals by scanning photodielectric spectroscopy. Appl. Phys. Lett., vol. 81, pp. 4195–4197, 2002. **3.** But, A.V., Migal, V.P. et al. Structure of a time variable photoresponse from semiconductor sensors. Technical Physics vol. 57, pp. 575–577, 2012. **4.** Mygal, V.P., Phomin, A.S. et al. Geometrization of the dynamic structure of the transient photoresponse from zinc chalcogenides. Semiconductors, vol. 49, pp. 634–637, 2015. **5.** But, A.V., Mygal, V.P., Phomin, A.S. Photoelectric signatures of CdZnTe crystals. Semiconductors vol. 43, pp. 1217–1220, 2009. **6.** Мигаль В.П., Мигаль Г.В. Сигнатурний подход к анализу и обеспечению безопасности системы «человек-машина» / Открытые информационные и компьютерные интегрированные технологии: сб. науч. трудов. – Х.: Нац. аэрокосм. ун-т «ХАИ». – 2014. Вып. 65. – С. 152-159.

Мосійчук М.Д., Олефір О.С.
Національний технічний університет України "КПІ", факультет прикладної математики, Київ, Україна

Використання поліноміального хаосу для моделювання біологічних процесів та хімічних реакцій клітини

Вивчення генетичних процесів всередині клітин неможливе без моделювання цих процесів та явищ. Зокрема, цікавим є питання зміни концентрації та властивостей активних речовин протягом цих процесів. Моделювання задачі ускладнюється тим, що зазвичай коливання концентрацій цих речовин є випадковою величиною та може бути оцінене доволі приблизно. Тому актуальною є задача розробки математичного та програмного забезпечення для моделювання таких систем за допомогою випадкових диференційних рівнянь та їх розв'язок.

Зазвичай хімічні кінетичні реакції можуть бути представлені за допомогою системи поліноміальних диференційних рівнянь.

$$y(t) = (y_1(t), y_2(t), ..., y_n(t)), \tag{1}$$

де $y_i(t)$ - є концентрацією хімічних регенератів. Загальний вигляд рівняння реакції:

$$\dot{y}(t) = \sum_{j=1}^{q} k_j f_j(y(t))(v_j^+ - v_j^-), \tag{2}$$

де q – число реакцій, k_j – коєфіцієнти рівняння, v_j^+ - стехіометричний коефіцієнт добутку,v_j^- стехіометричний коефіцієнт субстрату, $f_j(y(t))$ - змінна добутку у вигляді:

$$f_j(y(t)) = \prod_{l=1}^{n} (y_l(t))^{v_j^- l} \tag{3}$$

Початковими умовами системи є $y(0) = (y_1(0), ..., y_n(0))$. Зазвичай існують тільки добутки, що складаються з двох елементів і змінні мають квадрати в рівняннях. Проте, більш високі порядки теж допускаються. Основна проблема полягає в тому, що, як правило, коефіцієнти і/або початкові умови невідомі і можуть бути знайдені тільки з деякою ймовірністю. Таким чином, ці коефіцієнти розглядаються як випадкові величини з деякою функцією розподілу ймовірностей.

Одним з підходів до вирішення такого роду систем є використання поліноміального хаосу Вінера-Аскі [1, 2]. Основна ідея полягає в тому, що, згідно з [2], функція f при точному часі τ може бути апроксимована за допомогою ортогональної проекції.

Хаос Вінера-Аскі може бути використаний для представлення кожного елементу y у вигляді суми базисних поліномів.Таке ж представлення є можливим і для випадкових коефіцієнтів та початкових умов. Визначимо $Z = (k_1, ..., k_N)$, де N - є кількість випадкових коефіцієнтів, в свою чергу $/Phi(Z)$ - є поліноміальним базисом.

$$y_l(t) = \sum_{m=0}^{k} y_{lm}(t)\Phi_m(Z) \tag{4}$$

$$y_l(0) = \sum_{m=0}^{k} y_{lm}(0)\Phi_m(Z) \tag{5}$$

$$k_j(t) = \sum_{m=0}^{k} k_{jm}(t)\Phi_m(Z) \tag{6}$$

Розподіл випадкових величин і їх поліноміальних базисів може бути різним. Зазвичай це зручно обирати базуючись на ваговії функції в якості випадкової величини розподілу, яка дає так зване сильне наближення. Наприклад, якщо змінна має нормальний розподіл, то

зручно використовувати поліноми Ерміта, якщо вона має бета розподіл - поліноми Якобі і так далі. Справді, якщо змінна, що знаходиться форматі функції розподілу ймовірностей, відома, то відносно легко побудувати ортогональну систему, яка відповідає розподілу за допомогою, наприклад, методу вагової функції моментів.

Єдине правило полягає у використанні функції розподілу ймовірностей в ролі вагової функції, що означає, що для кожного $\Phi_m(Z)$ $\Phi_l(Z)$ поліноми наступні:

$$E[\Phi_m(Z)\Phi_l(Z)] = \int_S \Phi_m(Z)\Phi_l(Z)\omega_m(Z)dZ = \gamma_m\delta_{ml}, \tag{7}$$

де $\omega(x)$ - функція розподілу ймовірностей, δ_{ml} - дельта Кронекера, $\gamma_m = E[\Phi_n^2(Z)]$. Таким чином, систему стохастичних диференціальних рівнянь можна представити у вигляді:

$$\sum_{m=0}^{k} \dot{y_m}(t)\Phi_m(Z) = \sum_{j=1}^{q}\sum_{m=0}^{k} k_{jm}\Phi_m(Z)g_j(\tilde{y}(t))(v_j^+ - v_j^-), \tag{8}$$

де

$$g_j(\widetilde{y}(t)) = \Pi_{l=1}^{n}(\sum_{d=0}^{k} y_{ld}(t)\Phi_d(Z))^{v_{jl}^-} \tag{9}$$

У цій формі всі стохастичні компоненти розташовуються усередині поліноміальних базисних елементів і всі коефіцієнти є детермінованими.

Наступним кроком є використання проекції Гальоркіна, щоб позбутися стохастичних компонентів і надати можливість розрахувати коефіцієнти. Основна ідея полягає в тому, що ми повинні знайти рішення так, що похибка загальної системи є ортогональною до поліноміального базису.

Після проекції з'являється система з nk диференційних рівнянь, де n - число змінних в y, k - загальний порядок рядів. Кожен поліноміальний коефіцієнт для кожного елемента рішення тепер має вигляд:

$$\dot{y_p}(t) = 1/\gamma_p \sum_{j=1}^{q}\sum_{m=0}^{k} k_{jm}\pi_{jm}(\tilde{y}(t))(v_j^+ - v_j^-) \tag{10}$$

Остаточний вигляд рівняння:

$$\dot{\tilde{y}}(t) = 1/\gamma_p \sum_{j=1}^{q}\sum_{m=1}^{q} k_j m \sum_{l=1}^{q_1}(\prod_{d=0}^{q_2}(y_l^{a_j ld}d)\sum_{j_1=0}^{p_1} m_j i E[k_{i_1}^{j_1}]....\sum_{j_N=0}^{p_N} m_l N i E[k_{i_N}^{j_N}](v_j^+ - v_j^-) \tag{11}$$

де $\sum_{d=0}^{q=2} a_{jld} = v_{jl}^-$.

Розроблено та протестовано модель, яка дозволяє вирішити задачу моделювання зміни концентрації речовин практично для будь-яких реальних біологічних процесів.

Результатом роботи даної моделі є розв'язок складної системи диференційних рівнянь, який показує очікувану модель зміни концентрації речовин протягом трансмутації генів.

Тестування показало, що алгоритм демонструє необхідну точність обчислень для коректного моделювання поведінки біологічних систем.

Література. **1.** Dongbin Xiu. Numerical Methods for Stochastic Computations. Princeton University Press, 2010. **2.** Dongbin Xiu and George Em Karniadakis. The Wiener–Askey polynomial chaos for stochastic differential equations. SIAM journal on scientisfic computing, 24(2):619– 644, 2002.

Недашковская Н.И.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Согласованность и неопределенность экспертных оценок при принятии решений

Рассматривается задача вычисления весов альтернатив решений на базе экспертных оценок парных сравнений альтернатив. Для ее решения обычно применяются методы главного собственного вектора, геометрической средней, а также модели оптимизации [1]. В этих методах качество экспертных оценок парных сравнений проверяется с помощью показателей согласованности. Если показатели превышают свои пороговые значения, то такие оценки считаются противоречивыми, недопустимо несогласованными и не могут применяться для вычисления весов. С другой стороны, множество полностью согласованных оценок считается истинным, т.е. таким, что отображает истинные веса альтернатив по исследуемой их характеристике. На основании таких оценок известные методы главного собственного вектора, геометрической средней и ряд остальных [1] приводят к четким (вещественным) результирующим весам. Однако, утверждение об истинности полностью согласованных экспертных оценок не всегда является верным, поскольку эксперты, выполняя оценивание в шкале, могут поставить в соответствие реальным весам разные множества полностью согласованных оценок парных сравнений [2]. Поэтому необходим метод, результатом работы которого будут нечеткие (интервальные) веса, независимо от уровня несогласованности экспертных оценок.

В последнее время предлагаются модификации методов парных сравнений с использованием теории нечетких множеств [1]. Экспертные оценки формализуются с помощью функций принадлежности, причем вид и параметры этих функций выбираются произвольно. Используя разные функции принадлежности, на основании одного и того же множества экспертных оценок в результате получаем разные нечеткие (интервальные) веса. Результирующие веса характеризуются большей шириной, если исходные экспертные оценки формализованы нечеткими множествами с большей степенью нечеткости. Такой субъективизм с выбором вида и параметров функций принадлежности ставит под сомнение достоверность результатов, получаемых известными нечеткими методами парных сравнений.

В [2, 3] разработан новый метод, в соответствии с которым на основании экспертных оценок парных сравнений альтернатив решений вычисляются доверительные интервалы для весов альтернатив. В основу положено утверждение, что оценки только в некоторой степени отражают реальные отношения весов альтернатив и содержат неопределенность, независимо от уровня их согласованности. Предполагается, что неопределенность экспертных оценок парных сравнений обусловлена шкалой, в которой эксперт выполняет оценивание, и такими личными его качествами, как пессимизм и оптимизм. Метод использует аппарат теории доверия (свидетельств) и результаты компьютерного моделирования суждений экспертов. Компьютерное моделирование позволило при некоторых условиях получить количественные оценки неопределенности суждений эксперта-реалиста, пессимиста и оптимиста в задаче вычисления весов методом парных сравнений в шкале Саати [2, 3]. Предложен показатель неопределенности экспертных оценок рассматриваемой задачи вычисления весов. На основании решения большого количества тестовых задач разной размерности построены таблицы оценок параметров метода в зависимости от количества сравниваемых альтернатив решений.

Результаты моделирования показывают, что предложенные доверительные интервалы [2, 3] более достоверно отображают реальные веса по сравнению с точечными весами, получаемыми известным методом главного собственного вектора анализа иерархий, а также нечеткими весами по методу нечеткой геометрической средней.

Литература. **1.** Панкратова Н.Д., Недашківська Н.І. Моделі і методи аналізу ієрархій: Теорія. Застосування: Навчальний посібник.–К:ІВЦ «Видавництво«Політехніка»,2010.–371 с. **2.** Недашковская Н.И. Метод оценивания неопределенности экспертных оценок парных сравнений при вычислении весов альтернатив решений//Проблемы управления и информатики.–2015.-No5.–С.1–13. **3.** Недашковская Н.И. Построение доверительных интервалов для весов альтернатив решений на основе экспертных парных сравнений//Системні дослідження та інформаційні технології.–2015.-No3.- С.121 - 130.

Омельчук А.А., Рудакова А.В., Поливода О.В.
Херсонский национальный технический университет, Херсон, Украина

Использование алгоритма стохастической аппроксимации для идентификации параметров модели перемещения судна на слипе

Процесс перемещения судна на слипе осуществляется в нестационарных условиях при воздействии внешних факторов, которые существенно изменяются на протяжении всего пути [1], что зачастую приводит к возникновению нештатных ситуаций. Обеспечение должной безопасности и повышения надежности функционирования многоприводной спускоподъемной системы возможно за счет дальнейшей автоматизации судоподъемного сооружения путем внедрения компьютеризированной системы управления и использования для оперативного управления процессом перемещения методов адаптивного управления с прогнозом.

Работа адаптивной системы основывается на своевременной оценке текущего состояния объекта и прогнозе его возможного состояния на ближайший период времени, что обуславливает необходимость наличия адекватной математической модели объекта.

В зависимости от положения, перемещаемого на слипе судна, должны формироваться соответствующие сигналы управления движением тележек, связанных общим объектом (корпусом судна). На рис. 1 рассмотрены силы, действующие на распределенный, движущийся по направлению оси x, объект «судно-тележки».

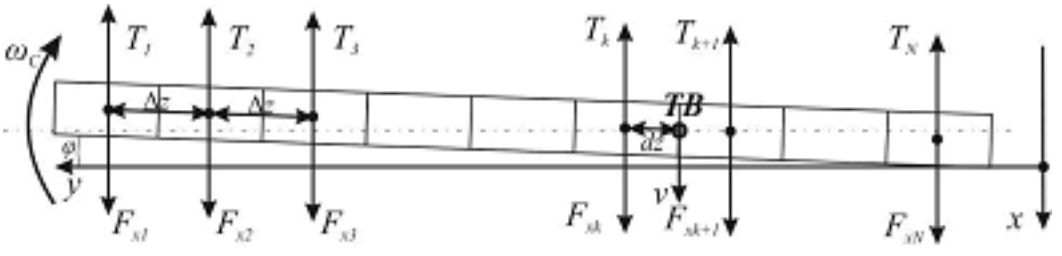


Рис. 1. Действие сил на объект «судно-тележки»

Сложный процесс перемещения можно разделить на два вида движения: поступательное со скоростью v и ускорением a, вращательное с угловой скоростью ω и углом поворота φ . На основе анализа уравнений движения можно записать уравнения состояния модели объекта «судно-тележки» в пространстве состояний, как показано в [2], в виде

$$\dot{x}_1 = x_2, \;\; \dot{x}_2 = \frac{1}{m}\sum_{i=1}^{N} F_{xi} - \frac{T_m}{m}\sum_{i=1}^{N} u_i, \tag{1}$$

$$\dot{x}_3 = x_4, \;\; \dot{x}_4 = \frac{1}{J}\sum_{i=1}^{N}(T_m u_i - F_{xi}) \cdot [(k-i)z_0 + dz] \cdot cos(x_3), \tag{2}$$

где $x_1 = l$ – перемещение центра масс судна вдоль оси x; $x_2 = v$ – скорость поступательного движения центра масс вдоль оси x; $x_3 = \varphi$ – угол поворота судна; $x_4 = \omega$ – скорость вращения судна; $\vec{u} = (u_1, u_2, \ldots, u_N)^T$ – вектор управления, $u_i = T_i/T_m$, $i = \overline{1,N}$; T_i – сила натяжения троса i-й тележки; T_m – предельно допустимое натяжение троса; $F_{xi}(l_i)$ – величина нагрузки на i-й привод, обусловленная силами, влияющими на тележку в точке траектории движения l_i; $m = m_c + N \cdot m_T$ – масса объекта «судно-тележка», m_c – масса судна, N – количество тележек, m_T – масса тележки; J – момент инерции судна; z_0 – расстояние между центрами соседних тележек; dz – расстояние от точки вращения (центра масс объекта) к центру k-й тележки.

Уравнения выхода модели формируются как

$$y_j = x_1 - [(k-j)z_0 + dz] \cdot x_3, \;\; j = \overline{1,r}, \tag{3}$$

где r – количество выходов, необходимых для идентификации параметров движения объекта.

Разработанная в пространстве состояний модель перемещения судна на слипе, которая учитывает все существенные внешние факторы, является нелинейной, 4-го порядка. В результате линеаризации (1)-(3) и с учетом малых допустимых углов поворота $x_3 < 2^\circ$ (можно принять $sin(x_3) \approx x_3$, $cos(x_3) \approx 1$) была получена математическая модель, пригодная для синтеза оптимального управления с использованием методов аналитического конструирования оптимальных регуляторов [3]. Линеаризованные уравнения в пространстве состояний записываются в векторно-матричной форме как

$$\dot{\widehat{\mathbf{x}}} = \mathbf{A}\widehat{\mathbf{x}} + \mathbf{B}\widehat{\mathbf{u}}, \quad \widehat{\mathbf{y}} = \mathbf{C}\widehat{\mathbf{x}}, \tag{4}$$

где матрицы $\mathbf{A}$, $\mathbf{B}$ и $\mathbf{C}$ в (4), имеют постоянную структуру, но их компоненты изменяются в зависимости от условий функционирования системы, и имеют вид

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ a_{21} & 0 & a_{23} & 0 \\ 0 & 0 & 0 & 1 \\ a_{41} & 0 & a_{43} & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ K & K & \dots & K \\ 0 & 0 & \dots & 0 \\ b_{41} & b_{42} & \dots & b_{4N} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 & c_{13} & 0 \\ 1 & 0 & c_{23} & 0 \end{pmatrix}. \tag{5}$$

Элементы матриц (5) определяются как $a_{i1} = f_{i1}(x_{1s}, x_{3s})$, $a_{i3} = f_{i3}(x_{1s})$, где x_{1s}, x_{3s} – элементы вектора установившегося состояния объекта $\vec{x}_s = (x_{1s}, \dots, x_{4s})^T$; $K = -T_m/m$, $b_{4j} = (T_m/J)[(k-j)z_0 + dz]$; элементы матрицы $\mathbf{C}$ соответствуют структуре системы измерения при $r = 2$, $c_{13} = (k-1)z_0 + dz$, $c_{23} = (k-N)z_0 + dz$.

Параметры линеаризованной модели перемещения судна в зависимости от точки положения на слипе принимают разные значения вследствие изменчивости возмущающих факторов. В связи с этим возникает потребность в своевременной оценке компонентов матрицы $\mathbf{A}$.

Для оценивания параметров в реальном масштабе времени обычно используют рекуррентные методы [4], которые относятся к многошаговым процедурам и вычисляют токующую оценку на основе предшествующей по мере поступления информации в дискретные моменты времени. С точки зрения простоты технической реализации оценивание (идентификацию) параметров модели объекта целесообразно проводить рекуррентным алгоритмом стохастической аппроксимации [4].

В связи со свойством достаточно быстрой сходимости псевдоградиентного метода стохастической аппроксимации процедуры идентификации параметров модели, необходимо активизировать заново при существенном изменении внешних факторов. Условие окончания этапа идентификации $|a_{ij}(k+1) - a_{ij}(k)| < \delta$ для всех i, j. Условие активации нового этапа идентификации $|y(t) - y_m(t)| > \epsilon$.

Исследование эффективности процедуры идентификации параметров модели и выбор оптимальных значений коэффициентов настройки алгоритма можно осуществить в результате моделирования работы адаптивной системы управления процессом перемещения судна на слипе.

Литература. **1.** Омельчук А.А. Проблеми узгодженого керування складною електромеханічною системою / А.А. Омельчук, Ю.О. Лебеденко, Г.В. Рудакова // Наукові праці Національного університету харчових технологій. – К.: НУХТ, 2013. – №49. – С. 19 – 23. **2.** Омельчук А.А. Моделирование пространственных параметров процесса спуска судна на воду с помощью слипа / А.А. Омельчук, А.В. Рудакова // Системні технології. Регіональний міжвузівський збірник наукових праць. – Дніпропетровськ, 2015. – №4(99). – С. 31 – 39. **3.** Omelchuk A.A. Optimal control of the ship's motion at the cross slipway / A.A. Omelchuk, G.V. Rudakova, O.V. Polivoda // Праці Одеського політехнічного університету. – 2015. – Вип. 3(47). – С. 75 – 84. **4.** Киричков В.Н. Автоматика и управление в технических системах. В 11-ти кн. Кн.2. Идентификация объектов систем управления технологическими процессами / В.Н. Киричков; под ред. А.А. Краснопрошиной. – К.: Выща шк. 1990. – 263 с.

***Орєхов О.А.*[1], *Орєхова Н.А.*[2]**
[1] *Фізико-технічний інститут НТУУ "КПІ", Київ, Україна;* [2] *Інститут кібернетики ім.В.М.Глушкова НАН України, Київ, Україна*

Методи верифікації паралельних програм за допомогою мереж Петрі

В доповіді надається опис методу верифікації комп'ютерної програми шляхом перетворення її на мережу Петрі спеціального виду та пошуку Т-інваріантів для отриманої мережі Петрі. У залежності від наявності та кількості Т-інваріантів робляться висновки про поведінку асинхронного процесу, що описується цією мережею Петрі.

Мережі Петрі спільно з алгебрами процесів є одним із основних інструментів моделювання складних розподілених асинхронних обчислювальних систем [1,2]. Паралельна мова програмування описується відповідною алгеброю процесів з переліком бажаних властивостей. На наступному кроці для артефактів побудованої алгебри процесів будуються мережі Петрі (звичайні, кольорові, стохастичні). Для того, щоб модель, яку побудовано в вигляді мережі Петрі, демонструвала строго циклічну поведінку, потрібно, щоб ця мережа Петрі належала до класу "fair"("справедливих") [3,4]. І навпаки, можна передбачати суттєво ациклічну поведінку асинхронної моделі, якщо відповідна мережа Петрі не є "справедливою"навіть за наявності двох чи більшої кількості Т-інваріантів.

Для визначення "справедливості"мережі Петрі доведено можливість побудови поліноміально трудомісткого алгоритму від розміру матриці переходів [5,6], але для вирішення практичних задач великої розмірності цей метод вживати недоцільно. У зв'язку з цим було проведено перевірку можливості використання бібліотеки scipy.optimize [7] (мова програмування Python 3.5) пошуку інваріантів мережі Петрі шляхом редукції до задачі лінійної оптимізації спеціального типу та її вирішення за допомогою симплекс-методу. Сама модель мережі Петрі будувалась за допомогою бібліотеки snakes [8] мови програмування Python 3.5.

Доповідь містить результати обчислень, що свідчать про можливість практичного застосування описаного підходу для вирішення задачі побудови моделей асинхронних розподілених систем з передбачуваною поведінкою. Виконано пошук Т-інваріантів для мереж Петрі, що утворюються внаслідок імплементації алгоритмів взаємного виключення Петерсона, Деккера та Морріса [2].

На підставі отриманих результатів також було проаналізовано можливість пошуку Т-інваріантів мереж Петрі для типових завдань стосовно "відомих"моделей (known models) 2016р. міжнародного змагання Model Checking Contest @ Petri Nets [9]. Знайдено Т-інваріанти для всіх "відомих"моделей 2011 - 2015 років та зроблено висновки щодо їхньої можливої "справедливої"поведінки на підставі результатів обчислень. Такий підхід дозволяє розширити перелік сучасних методів верифікації складних асинхронних обчислювальних систем.

Література. **1.** http://ls4-www.cs.tu-dortmund.de/cms/de/home/bause/bause_kritzinger_spn_book_screen.pdf

2. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.52.3685&rep=rep1&type=pdf

3. https://inst.eecs.berkeley.edu/~ee249/fa07/discussions/PetriNets-Murata.pdf

4. http://link.springer.com/chapter/10.1007/978-1-4612-3462-3_61#page-1

5. Орехов А.А. Полиномиальный алгоритм нахождения Т-инвариантов сети Петри. "Кибернетика", 1988, № 3.

6. A.Orekhov. Petri Nets: combinatorial approach and modelling flexible manufacturing systems. - In: IMACS'91: Proceedings of the 13th World Congress on Computation and Applied Mathematics. Trinity College, Dublin, 1991

7. http://docs.scipy.org/doc/scipy/reference/optimize.html

8. https://www.ibisc.univ-evry.fr/~fpommereau/SNAKES/

9. http://mcc.lip6.fr

Павлюк Ю.Р., Подладчіков В.М.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Розробка методів статистичного аналізу для визначення швидкості корональних хвиль

EUV (Extreme Ultraviolet) хвилі виникають під час ерупцій в активних зонах, тому вивчення їх розповсюдження допоможе пояснити питання, пов'язані з процесом формуванням корональних викидів мас і виникненням сонячних спалахів. Важливість цих питань визначається їх прикладним характером, зокрема геоефективними наслідками.

Розглянуто корональні викиди мас, які відбулись 7 квітня та 12 травня 1997 року. На основі послідовності зображень з SOHO-EIT було сформовано фіксовані EUVI різницеві зображення, що враховують особливості обертання Сонця [1]. Активні зони під час ерупції знаходяться у відносній близкості до центрального меридіану, що зменшує негативні ефекти від роботи з проекціями. Під час обох подій EUV хвиля мала майже циркулярну форму.

Для обробки супутникових даних використовується метод кільцевого аналізу. Розроблений метод фільтрації для вилучення хвильового фронту враховує те, що EUV хвилі мають вигляд дифузійної кільцевої структури, яка поширюється від центру виверження [2], [3]. Динаміка EUV хвилі залежить від відстані до еруптивного центру, тобто довжини радіус-вектора $\overline{r}$ та його напряму (кут ϕ). Саме тому для обчислень формується полярна система координат з початком в центрі ерупції.

Метод фільтрації зображень базується на обробці множини кілець на поверхні Сонця з центром в центрі виверження. Досліджується динаміка сумарної інтенсивності кілець $S(m)$, на основі інтенсивностей всіх пікселів в кільці m, під час проходження хвильового фронту. Потрібно зауважити, що варіації інтенсивностей кожного окремого пікселя можуть бути викликані артефактами, іншими подіями, не пов'язаними з корональними викидами мас, можливими помликами спостережень, фрагментальним характером хвильового фронту, діммінгами, обмеженими певними кутовими секторами. Проте загальна інтегральна інтенсивність уздовж кільця менше буде залежати від локальних відхилень, ніж інтенсивність окремих пікселів [4].

Для спостереження за поширенням фронту EUV хвилі і визначення його радіальної швидкості використовується часова послідовність загальної інтенсивності в концентричних кільцях $S(m)$. Для дослідження напрямку розповсюдження хвилі, розіб'ємо проміжок $0 \leqslant \phi < 2\pi$ на рівні сектори. Інтегральна інтенсивність m кільця визначається наступним чином: $S(m) = \sum_k S(m,k)$, де $S(m,k)$ - загальна інтенсивність для k сектора в m кільці.

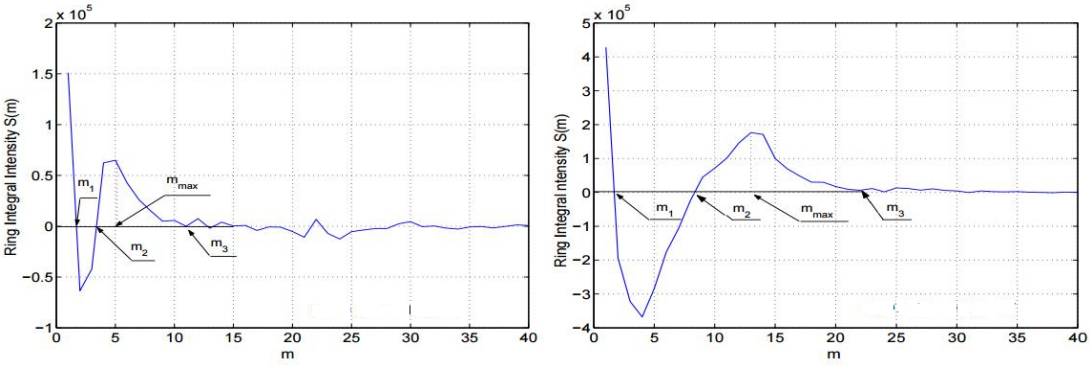


Рис. 1. Залежність інтегральної інтенсивності від номеру кільця.

Межі фронту хвилі визначаються нулями (m_2;m_3) функції $S(m)$, що відповідають кордонам зони діммінгу. В кожний момент часу визначається точка m_{max}, де інтенсивність хвильового фронту досягає свого максимального значення. Кільце хвильового фронту обмежується радіусами $\overline{r_{m_2}}, \overline{r_{m_3}}$, що відповідають кордонам фронту хвилі. На графіках (Рис. 1) видно, що значення m_1, m_2, m_3 монотонно зростають від одного зображення до іншого, що пояснюється розширенням активної області під час розповсюдження хвильового фронту.

Чисельна залежність відстані до фронту хвилі від часу може бути використана для оцінки радіальної швидкості хвилі. Для автоматичної ідентифікації хвильового фронту, ми визначаємо точку максимуму гребеня хвилі m_{max} для зображень на інтервалі від $\overline{r_{m_2}}$ до $\overline{r_{m_3}}$ в різні моменти часу.

На основі розробленого методу було оцінено швидкості для хвиль, які відбулись 7 квітня та 12 травня 1997 року. EUV хвиля на зображеннях від 12 травня має майже правильну кільцеву форму, проте можна виділити північно-східний і південно-західний напрями розповсюдження фронту хвилі. На зображеннях від 7 квітня 1997 року EUV хвиля також розповсюджується в двох напрямках - південно-західному та північно-східному. За допомогою методів кільцевого та секторного аналізу було отримано оцінки радіальноїV_r і тангенціальної складової швидкості ω та повне значення швидкості V (Табл.1, Табл. 2).

Табл. 1. Оцінки кутової і радіальної швидкості для EUV хвилі 12 травня 1997 року

Time		Південно-східний напрям		Північно-західний напрям	
	V_r, km/s	ω, rad/s	V, km/s	ω, rad/s	V, km/s
04:50 - 05:07 UT	259,65	1,27e-05	224,87	3,97e-03	207,31
05:07 - 05:24 UT	226,74	1,07e-05	205,66	2,12e-04	207,80
05:24 - 05:41 UT	227,64			-2,19e-05	95,78

Табл. 2. Оцінки кутової і радіальної швидкості для EUV хвилі 7 квітня 1997 року

Time	Південно-західний напрям			Північно-східний напрям		
	ω, rad/s	V_r, km/s	V, km/s	ω, rad/s	V_r, km/s	V, km/s
14:00 - 14:12 UT	2,70e-04	52,03053648	233,30	2,30e-04	295,96	223,80
14:12 - 14:21 UT	1,24e-04	461,7610544	470,26	7,40e-04	386,24	577,11
14:21 - 14:35 UT	1,32e-04	390,8053615	172,22	3,91e-05	89,41	137,17

Література. **1.** Eit waves and eruptive dimmings catalog [Online]. Available: http://sidc.oma.be/nemo/catalog/. **2.** Moses. D. et al.,: 1997. ElT observations of the extreme ultraviolet Sun. Solar Phys. 175(2), 571 -599. ADS:1997SoPh.175.571M. doi:10 1023/A; 1004902913117. Thompson et al., 1998). **3.** Thompson. B.J.. Plunkett. S.P.. Gurman. J.B.. Newmark, J.S.. St Cyr. O.C.. Michels. D.J.: 1998. SOHO/EIT observations of an Earth-directed coronal mass ejection on May 12, 1997. Geophys. Res. ten. 25(1). 2465 - 2468. doi:1O1029/98GL50429. **4.** Podladchikova. O,. Bcrghmans, D,: 2005, Automated detection of EIT waves and dimmings. Solar Phys. 228. 265-284. doi:l0.1007.S11207-005-5373 z.

Панкратов В.А.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Развитие будущего металлургической промышленной Украины

После потери из-под контроля Украины в 2013–2014 годах газовых месторождений Черноморского шельфа и туристического комплекса Крыма и значительной части горнодобывающей, металлургической и химической промышленности Донбасса очевидным стал факт, что будущая экономика страны уже не будет иметь предыдущей структуры и характера низкотехнологического, сырьевого, перерабатывающего развития. В связи с этим в международной кооперации труда экономика Украины должна найти для себя другие сферы по тем кластерами экономики, по которым она имеет географические, ресурсные, человеческие и другие преимущества. Безальтернативной для нее является переориентация на высокотехнологичный путь дальнейшего развития. При этом государство должно сделать ставку исключительно на качественный человеческий капитал, включить в основные драйверы экономикиконкурентоспособную науку, передовое образование и высокотехнологичный бизнес. Поэтому безальтернативным является переход экономики на высокотехнологичные рельсы. Было обнаружено и описано области, которые могут стать главными драйверами или локомотивами такого развития [1]. Одни из них – это области создания новых веществ и материалов, нанотехнологии, высокотехнологичное машиностроение и другие – всего десять позиций.

В связи с нынешним чрезвычайно острым российско-украинским конфликтом в газовой сфере особое значение для энергетической и экономической независимости нашего государства приобретает состояние отечественной угольной промышленности. Впрочем, эскалация боевых действий на Донбассе нанесла значительный ущерб этой отрасли украинской экономики. Уже сегодня кризис в отечественной угольной промышленности негативно сказывается на производстве электроэнергии. Аналогичная ситуация, хотя в несравненно меньших масштабах, наблюдается и в отечественной черной металлургии. Из-за эскалации непосредственно поддерживаемых Россией военных действий на Донбассе в этом регионе прекратили или сократили работу некоторые металлургические предприятия. Сложности в работе, прежде всего из-за проблем с логистикой, возникли и в некоторых отечественных металлургических предприятиях, расположенных вне зоны непосредственных боевых действий. В результате Украина по итогам работы металлургических предприятий в 2014 сократила производство стали на 37% по сравнению с 2013, опустившись с 10 на 11 место в рейтинге 65 стран – основных мировых производителей этой продукции. Однако, по итогам 2014, в топ-10 мировых производителей стали вошли Китай, Япония, США, Индия, Россия, Южная Корея, Германия, Турция, Бразилия и Украина. В то же время ослабление рыночных позиций украинских металлургов на мировом рынке в определенной степени идет на пользу их российским конкурентам, которые таким образом косвенно получают выгоду от военного конфликта на Донбассе.

Промышленная деятельность Украины – сложная многофакторная система, в которой присутствует множество проблемных направлений, поэтому важно отобрать те, которые являются наиболее важными, не потеряв при этом информацию и не внеся ошибку на данном этапе вычисления.

На основе предварительного исследования, выполненного с привлечением метода SWOT анализа, выявлены сильные и слабые стороны, а также возможности и угрозы в промышленной деятельности Украины. *Самые опасные угрозы*: отток инвестиционных капиталов из Украины; медленное восстановление объектов промышленности на Востоке Украины; экономическая нестабильность, высокая вероятность дефолта страны. *Наиболее вероятные для реализации возможности*: максимальная переориентация на собственное производство во всех сферах промышленности; создание замкнутых циклов производства в различных отраслях промышленности; привлечение инвестиций и развитие бизнеса с зарубежным капиталом; переориентация старых предприятий, модернизация технологических процессов; высокий спрос на продукцию из композиционных материалов. *Слабые характеристики*: политическая

и социальная нестабильность; устаревшая материальная база; высокий уровень коррупции; низкое качество экономической инфраструктуры; девальвация национальной валюты; низкая инвестиционная привлекательность страны в связи с социально-политическим положением. *Сильные характеристики*: большой запас природных ресурсов для различных отраслей; возрождение оборонной промышленности и ее совершенствования; выгодное торговое географическое положение; разнообразная отраслевая структура машиностроения; значительные научно-исследовательские и конструкторские достижения металлургического производства; высокое качество и соответствие международным стандартам украинских разработок в прогрессивной сфере композиционных материалов.

Таким образом, выполненны еисследования показали, что одним из перспективных направлений развития металлургической отрасли в Украине является развитие рынка композиционных материалов (КМ). Исследование направлений развития рынка КМ до 2025 года выполнялось с помощью метода Делфи. Рассмотрены 4 главные области применения КМ: промышленность; транспорт; жилищное строительство; военный сектор.Эксперты в составе 16 квалифицированных профессионалов предоставили оценки каждой области применения по критериям: объем производства; объем спроса; эффективность применения; степень отношение к области. Для выявления приоритетности направлений развития рынка КМ необходимо было проранжировать области применения по выбранным критериям.

Для решения поставленной задачи выполнено следующее: проведен анализ согласованности экспертных оценок для каждого показателя каждой области применения с учетом компетентности и уверенности экспертов и априорно заданным радиусом доверительного интервала; определены согласованныеэкспертныеоценки для каждого показателя каждой области применения; определена приоритетность областей применения по варианту и проранжавано их, учитывая веса показателей; на основе проведенного анализа предложено объективную информацию по оценке областей применения и даны соответствующие рекомендации.

Итак, заказчику исследуемого направления можно дать некоторые советы и рекомендации по отраслям применения композиционных материалов по их приоритетом. На основе проанализированной ситуации представленные отрасли можно проранжировать по приоритету эффективности следующим образом: военный сектор, транспорт, жилищное строительство, промышленность.

Разработку указанных стратегий и планов предполагается осуществлять путем масштабного ситуационного моделирования построенных сценариев на различных временных горизонтах на базе создаваемого в ИПСА ситуационно-аналитического центра (САЦ) в рамке Программы “Наука для мира и безопасности” (грант НАТО, 2015–2016 гг.) с последующим предоставление результатов исследований лицам, принимающим решения на разных уровнях управления государством. Решение задач предвидения по разработке стратегий социально-экономического развития [2] в среде САЦ позволит в режиме реального времени проводить мониторинг и моделирование социальных, экономических, экологических процессов различной природы и на этой основе анализировать и строить стратегии желаемого будущего на уровне крупных предприятий, мегаполисов, регионов и государства в целом.

Литература. **1.** Згуровький М.З., Панкратова Н.Д. Форсайт економіки України на середньострокову і довгострокову перспективу // XIII міжнародна науково-практична конференція “Математичне та програмне забезпечення інтелектуальних систем”, Матеріали конференції, 18–20 листопада 2015 р., Дніпропетровськ, Україна. – 2015. – С.75–77. **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

Панкратова Н.Д., Головань М.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Системна стратегія функціонування складної технічної системи в умовах багатофакторних ризиків

Актуальність та зацікавленість світової наукової спільноти проблемою забезпечення функціонування СТС в останнє десятиліття зумовлене усвідомленням кардинальної проблеми безпеки та стійкості систем після їх впровадження. Системна стратегія функціонування СТС спрямована на забезпечення її живучості з одночасним забезпеченням безпеки, а саме вона базується на запропонованому принципі своєчасного виявлення причин, оперативного запобігання переходу від штатних ситуацій в нештатні, аварійні чи надзвичайні, виявлення факторів ризику, прогнозування основних показників живучості об'єкта протягом заданого періоду його експлуатації як основи забезпечення гарантованої безпеки в динаміці функціонування системи, а також усунення причин можливих переходів робочих станів об'єкта в неробочий стан на основі системного аналізу багатофакторних ризиків нештатних ситуацій [1]. Розв'язок такого роду завдань для будь-яких систем, зокрема технічних, допоможе уникнути значних витрат на ремонт систем, шляхом недопущення аварійних та катастрофічних ситуацій, а також уникнути значних людських жертв за допомогою достовірного прогнозування стану досліджуваної системи.

Для складних систем різної природи актуальне достовірне і своєчасне передбачення, прогнозування і запобігання критичним, надзвичайним та іншим небажаним ситуаціям і впливам, що можуть призвести до позаштатного режиму, аварії, катастрофи або суттєво вплинути на працездатність, живучість, безпеку, ефективність та інші властивості таких об'єктів. Можливість появи і наслідки від впливу таких ситуацій, умов і факторів зумовлені випадковими і хаотичними процесами, які за механізмами впливу можна охарактеризувати як ризики. Розглянуто поняття «безпеки» системи – як властивості своєчасно запобігати послідовному переходу із штатного режиму системи в аварію чи катастрофу на основі своєчасного виявлення факторів суттєвих ризиків і усунення можливостей їх переходу у фактори катастрофічного ризику [2].

Інформацію про фактори ризику можна доповнити й уточнити у процесі формування рішення, зорієнтованого на запобігання небажаним впливам досліджуваних факторів ризику. Найважливішою особливістю задачі, що розглядається, є якісно новий вид обмеження, а саме порогове обмеження часу формування рішення, крім того множини факторів ризику і множини ситуацій є принципово необмеженими. Реалізація стратегії формується на основі наступних принципів: системна узгодженість по цілям,задачам, ресурсам і очікуваному результату за засобами забезпечення безпеки функціонування складної системи; своєчасне визначення, гарантоване розпізнання і системна діагностика факторів та ситуацій ризиків; оперативне прогнозування, достовірна оцінка нештатних і критичних ситуацій; формування і реалізація раціонального рішення за практично прийнятний час в рамках неусувного часового обмеження.

Особлива увага призначається задачі визначення і забезпечення раціонального узгодження ресурсів допустимого ризику позаштатних, критичних, надзвичайних, аварійних і катастрофічних ситуацій в процесі функціонування складної системи на її життєвому циклі. Для ситуацій, в яких параметри виходять за штатні режими функціонування, своєчасне прийняття рішення щодо зміни режиму функціонування, або штучного корегування деяких параметрів дозволяє забезпечити гарантоване функціонування технічної системи в реальному часі.

В роботі представлена реалізація функціонування системи оборотного водопостачання в штатному та позаштатному режимах.

Література. **1.** Pankratova N.D. System strategy for guaranteed safety of complex engineering systems //Cybernetics and Systems Analysis: Volume 46, Issue 2 (2010), Page 243-251. **2.** Pankratova N. System strategy of survivability and safety complex engineering objects operation // Computer science journal of Moldova. Vol. 22, no. 3 (66), (2014), Page 1-15.

Панкратова Н.Д., Кондратова Л.П.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Модель гарантированного предотвращения нештатного режима в реальных условиях функционирования сложной технической системы

Обеспечение безопасности, живучести, эффективности функционирования технического объекта с учетом целей, сроков, ресурсов, ожидаемых результатов представляет основную задачу технической диагностики. Различают средства функциональной и тестовой диагностики. Средства функциональной диагностики представляют типичные системы контроля, применяемые для оценки общего технического состояния объекта. Средства тестового диагностирования применяются для оценки технического состояния отдельных элементов, узлов, деталей со специально организуемыми тестовыми воздействиями. В работе [1] используется прием диагностики отказов, заключающийся в определении типа отказа, места и времени его обнаружения с использованием аналитических и экспертных оценок. В [2] используется подход к проектированию системы контроля с одновременным обнаружением отказов элементов технологического оборудования, осуществляя переход объекта управления в одно из безопасных состояний. Для установления работоспособности элементов диагностируемой технической системы в [3] рассматривается подход, основанный на построении оптимальной тестовой последовательности на заданных наборах входных данных.

Для сложных систем различной природы актуально своевременное предотвращение ситуаций нештатного режима при воздействии дестабилизирующих факторов на работоспособность, безопасность, живучесть, эффективность в процессе функционирования, которые выявляются с использованием показателей степени риска, уровня риска, ресурса допустимого риска. Ресурс допустимого риска характеризует ресурс, определяемый состоянием работоспособности, безопасности, живучести в течение определенного периода с учетом реальных условий эксплуатации.

Предлагается модель гарантированного предотвращения ситуаций нештатного режима при воздействии дестабилизирующих факторов и различного типа неопределенностей в реальном времени функционирования сложной технической системы (СТС) на основе оценивания и рационального согласования ресурсов допустимого риска. В основе реализации данной модели используется принцип своевременного обнаружения и устранения причин возможного перехода СТС в неработоспособное состояние, учитывая интегральный показатель информированности о ситуациях как комбинацию частных показателей полноты, достоверности, своевременности информированности и неустранимое пороговое ограничение времени в условиях воздействия дестабилизирующих факторов [4].

Ресурс допустимого риска оценивается с использованием функциональных зависимостей $\widetilde{y}_i = \phi_i\left(\{x_j, j=\overline{1,n}\}\right), i=\overline{1,m}$, восстановленных с приемлемой погрешностью в заданном интервале времени по выборке значений реальных показателей функционирования $y_i, i=\overline{1,m}$ и параметров $x_j, j=\overline{1,n}$ СТС с учетом прогноза на некоторый период. С использованием ограничений для степени и уровня риска по функциональным зависимостям $\widetilde{y}_i$ выявлены ситуации нештатного режима с интервалами времени формирования и реализации решений. Для учета различного типа неопределенностей вводится рациональный уровень информированности, определяемый с использованием средних значений интегрального показателя информированности в интервалах времени формирования и реализации решения относительно ситуаций для $\widetilde{y}_i$.

Ресурс допустимого риска для $\widetilde{y}_i$ реализуется в интервалах времени поиска дополнительной информации по принятию решения ЛПР относительно ситуаций, выявленных на основе функциональных зависимостей $\widetilde{y}_i$, с целью предотвращения ситуаций определенного нештатного режима. Согласованное значение ресурсов допустимого риска определется длиной интерва-

ла времени с минимальными значениями для нижней и верхней границ моментов времени реализации ресурсов допустимого риска для $\widetilde{y}_i$.

Модель гарантированного предотвращения определенного нештатного режима для функциональных зависимостей $\widetilde{y}_i$ описывает оценивание рационального согласования ресурсов допустимого риска в виде

$$R_c^{\text{доп}} = \frac{|\left[\min_{i=\overline{1,m}}\left\{t^i_{a_1r}\right\}, \min_{i=\overline{1,m}}\left\{t^i_{s_2r}\right\}\right]|}{|\left[\min_{i=\overline{1,m}}\left\{t^i_{s_1r}\right\}, \min_{i=\overline{1,m}}\left\{t^i_{s_2r}\right\}\right]|}, r \in [1, n_r]; \quad (1)$$

$$t^i_{a_1r} = \min_{t^i_k \in \left[t^i_{s_1r}, t^i_{s_2r}\right]} \left\{t^i_k | t^i_k < t^i_{s_2r}\right\}, r \in [1, n_r], i = \overline{1,m}; \quad (2)$$

$$\eta_i\left(\widetilde{y}_i\left(t^i_k\right)\right) < \eta_i\left(\widetilde{y}_i\left(t^i_{s_2r}\right)\right), W_i\left(\widetilde{y}_i\left(t^i_k\right)\right) < W_i\left(\widetilde{y}_i\left(t^i_{s_2r}\right)\right), r \in [1, n_r], i = \overline{1,m}; \quad (3)$$

$$U_0^i\left(t^i_{s_2r}\right) \geqslant U_0^-, r \in [1, n_r], i = \overline{1,m}. \quad (4)$$

В (1), (2) $t^i_{a_1r}$ характеризует нижнюю границу интервала времени реализации ресурса допустимого риска r-го нештатного режима относительно $\widetilde{y}_i$; n_r - количество видов нештатных режимов, возникающих в процессе функционирования СТС. Параметр $t^i_{a_1r}$ оценивается в интервале $\left[t^i_{s_1r}, t^i_{s_2r}\right]$ времени поиска дополнительной информации по принятию решения ЛПР с учетом значений степени $\eta_i\left(\widetilde{y}_i\left(t^i_{s_2r}\right)\right)$ и уровня $W_i\left(\widetilde{y}_i\left(t^i_{s_2r}\right)\right)$ риска в конечный момент времени $t^i_{s_2r}$ интервала времени поиска дополнительной информации по принятию решения ЛПР, максимизируя продолжительность реализации ресурса допустимого риска. Верхняя граница интервала времени реализации ресурса допустимого риска относительно $\widetilde{y}_i$ определяется с учетом значения U_0^- рационального уровня для интегрального показателя информированности в момент времени $t^i_{s_2r}$.

Исследование описанной модели проведено на примере функционирования системы водоснабжения с показателями уровня воды y_1 в резервуаре и напора воды y_2 на входе технологического объекта, значения которых характеризуют состояние системы в каждый момент времени заданного периода эксплуатации. Согласованный для указанных показателей функционирования системы водоснабжения ресурс допустимого риска, полученный в результате расчетов на основе предложенной модели, реализуется в интервале времени, который может быть использован для гарантированного предотвращения нештатной ситуации для $\widetilde{y}_1$ и аварийной ситуации для $\widetilde{y}_2$.

Литература. **1.** Fagarasan I. Fault detection and diagnosis of technical systems / Fagarasan I., St Iliescu S. // Proceedings of the 9th WSEAS International Conference on Automation and Information, Bucharest, Romania, Jun, 24-26, 2008. – 2008. – P. 446-453. **2.** Потехин А.И. Отказобезопасные дискретно-событийные системы // Технические и программные средства систем управления, контроля и измерения: труды конференции. - М., 2010.- С.943-958. **3.** Губернаторов В.П. Модификация эволюционно-генетического алгоритма для построения оптимальных тестовых последовательностей // Вестн. Нижегородского ун-та им. Н.И.Лобачевского. - 2012. - №3. - С.179-183. **4.** Pankratova N.D. Safety operations of the complex engineering objects // International Journal. “Information technologies & knowledge". – ITHEA: SOFIA. – V.5, 2. – 2011. –P.152–167.

Панкратова Н.Д., Сльота М.Р.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Класифікація і розпізнавання рівня небезпеки критичних ситуацій в задачах інформаційного аналізу

На практиці часто виникає необхідність дослідження системних задач розпізнавання і запобігання позаштатним, критичним і катастрофічним ситуаціям у складних системах різного призначення. Адже від вчасного виявлення та класифікації небезпечних ситуацій, залежить не тільки надійсність функціонування окремих функціональних елементів, а складної системи в цілому. Катастрофічні ситуації згубно впливають на населення і навколишнє середовище в регіональному чи навіть у глобальному масштабі. Упродовж останніх десятиліть аварії і катастрофи з багатомільйонними втратами мали місце майже в усіх промислово розвинених країнах. Аналіз цих катастроф свідчить, що масштаби втрат можна суттєво зменшити, якщо в позаштатній ситуації вчасно сформувати і реалізувати раціональні рішення ОПР.

У процесі функціонування складної системи під впливом множини $\phi = \{\phi_j | j = 1, m\}$ неконтрольованих факторів ризику ϕ_j штатна ситуація S_i може перейти в критичну, надзвичайну або катастрофічну. Такий перехід може відбуватися упродовж деякого періоду часу, тривалість якого апріорі невідома, і який залежить від кількості факторів ризику і тривалості їх впливу. Потрібно визначити такий допустимий період часу T_0 на формування і реалізацію рішення, для якого ймовірність переходу ситуації S_i в небезпечну, не перевищуватиме задане значення η. Імовірність переходу ситуації S_i під впливом фактора $\phi_j \in \phi$ у небезпечну ситуацію залежить від зміни в часі якісних показників повноти $I_{\text{П}}^{ij}$, достовірності $I_{\text{Д}}^{ij}$ і своєчасності I_T^{ij} інформованості ОПР. Імовірність η_{ij} такої події з урахування інтегрованого показника інформованості визначає співвідношення $\eta_{ij} = 1 - lg(1 + \alpha_{ij} \cdot I_{\text{П}}^{ij}(t) \cdot I_{\text{Д}}^{ij}(t) \cdot I_{\text{Т}}^{ij}(t))$. На цьому етапі визначається тривалість допустимого періоду $T_0^{ij} = [T_1^{ij}; T_2^{ij}]$ на формування рішення і приходимо до задачі класифікації ситуації [1]. В залежності від предметної області виділяємо множину критичних ситуацій $A = \{A_k | k = 1..K\}$. Кожна ситуація характеризується множиною проміжків часу T_0^{ij}, необхідних для прийняття рішення. Кожна критична ситуація має певний критерій, по якому можна її класифікувати. Критерій може бути сформований експертно у вигляді нечіткої оцінки [2]. З такої нечіткої оцінки формується функція належності для кожного A_k заданої на X, де X це універсальна множина $X = [0, T_{max}]$.Основна мета подальшої задачі – маючи нечітко задані межі для критичних ситуацій A_k визначити належність конкретної ситуації, що характеризується проміжками T_0^{ij}, до одного з класів критичних ситуацій. Таку задачу нечіткої класифікації вирішуємо за три етапи.

1. Етап фазифікації проміжків T_0^{ij}. Знаходження ступеня належності кожного проміжку, що характеризує ситуацію при дії певного фактора ризику, для кожного A_k.
2. Етап внутрішньофакторної нечіткої композиції.
3. Етап дефазифікації вихідних змінних.

Таким чином, можемо класифікувати ситуацію, враховуючи фактори ризику і нечіткість критерію відбору. Описана модель класифікації включає множину методів, оскільки можна застосувати різні методи внутрішньофакторної композиції фазифікованих змінних, використовуючи операції над нечіткими множинами. Розглянута проблема класифікації і розпізнавання рівня небезпеки критичних ситуацій має велике значення для формування необхідного і технологічно можливого рівня інформаційного забезпечення достовірності та обґрунтованості розв'язання прикладних системних задач. Наведений прийом розв'язання задачі та запропонована математична модель інтегрована в загальну методологію інформаційного аналізу[3].

Література. **1.** Згуровський М.З., Панкратова Н.Д. Основи системного аналізу.//Кібернетика і системний аналіз, 2007-С.263. **2.** Зайченко Ю.П. Нечіткі моделі і методи в інтелектуальних системах – К.:Видавництво «Слово», 2008 – 335с. **3.** Zgurovsky M.Z., Pankratova N.D. System analysis: Theory and Applications. Springer.-2007. - 475 p.

Панкратова Н.Д., Тарасюк Т.С.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Розпізнавання ситуацій за умов неповноти і нечіткості інформації

Актуальність дослідження. Необхідність прийняття рішення супроводжує людину у всіх сферах її діяльності. Своєчасність і правильність вибору забезпечує досягнення поставлених цілей. Тому необхідно забезпечити людину, що приймає рішення необхідним і технологічно можливим рівнем інформаційного забезпечення достовірності та обґрунтованості вибору для прийняття рішення. Створення методологічного і математичного інструментарію, що допоможе прийняти рішення у визначені терміни є основною проблемою інформаційного аналізу. Оскільки під час формалізації і розв'язання реальних задач важливе значення мають якісні характеристики інформації, важливою для вирішення постає задача розпізнавання ситуації за умов неповноти і нечіткості інформації. Розв'язання даної задачі допоможе людині, що приймає рішення, зробити оптимальний вибір, тим самим досягти поставлених перед нею цілей, незалежно від сфери її діяльності.

Постановка задачі. Для заданої множини S_0 ситуацій $S_k \in S_0$ існує поділ на скінченну кількість класів Ω_r, r=$\overline{1, R_0}$. При цьому поділ визначений не повністю, задано лише деяку неповну нечітку інформацію I_r про класи Ω_r, $r = \overline{1, R_0}$. Опис I_r кожного класу Ω_r є набором функцій належності $\mu_{\Omega_r}(\widetilde{I}_{\Omega_r})$, де $\widetilde{I}_{\Omega_r}$ - вектор часткових показників інформованості про елементи класу Ω_r, $\widetilde{I}_{\Omega_r} = \{I_{rq} | q \in Q_c\}$. Кожен досліджуваний об'єкт $S_k \in S_0$ описує нечітка інформація, що є набором функцій належності $\mu_k(\widetilde{I}_k)$, де $\widetilde{I}_k = \{I_{kq} | q \in Q_c\}$ - вектор часткових показників інформованості про об'єкт S_k. Потрібно за наявною інформацією $\widetilde{I}_k$ про ситуації $S_k \in S_0$ визначити до якого класу Ω_r належить ця ситуація. [1].

Для розв'язання даної задачі формалізуються наступні процедури:

- процедура класифікації множини S_0 з нечітко заданою інформацією про часткові показники інформованості
- процедура задання нечіткого відношення, що дасть змогу однозначно зіставляти нечіткий опис класу Ω_r і нечіткий опис досліджуваної ситуації S_k за кожним показником із системи часткових показників інформованості
- процедура композиції нечітких відношень, що дозволить однозначно зіставляти ситуації S_k і класи Ω_r за системою часткових показників [1]

Задамо функцію належності нечіткого відношення, що буде характеризувати ступінь належності ситуації S_k до класу Ω_r по характеристиці q у наступному вигляді:

$$\mu^q_{k\Omega_r}(I_{kq}, I_{rq}) = min(\mu_k(I_{kq}), \mu_r(I_{rq})), q \epsilon Q_c.$$

Для того аби однозначно визначити до якого класу належить ситуація S_k за заданою системою характеристик, представимо композицію нечітких відношень наступним чином:

$$\mu^*_{k\Omega_r}(\widetilde{I}_k, \widetilde{I}_r) = \prod_q C_q \mu^q_{k\Omega_r}(I_{kq}, I_{rq}), \text{де} q \epsilon Q_C, \text{а} C_q \epsilon [0, 1].$$

Саме таке представлення функції належності дозволить класифікувати ситуацію S_k до класу Ω_r за заданою системою характеристик з допомогою ваг C_q, які задаються при розв'язанні конкретної задачі. Також варто врахувати те, що при збільшенні кількості характеристик, значення $\mu^*_{k\Omega_r}$ швидко спадають. Тому варто врахувати, що коректність вибору рівня α залежить від кількості характеристик, що розглядаються.

Висновок. Таким чином, у роботі розглянуто формалізацію процедур класифікації та розпізнавання ситуацій за умови нечіткості інформації. Виявлено неоднозначність розпізнавання належності ситуацій до введених класів. Запропоновано способи уникнення неоднозначності класифікації за системою часткових показників інформованості.

Література. **1.** Згуровський М.З., Панкратова Н.Д. Основи системного аналізу. –К.: BHV, 2007. – 544 с.

Панкратова Н.Д., Шелест А.В.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Исследование развития ИКТ кластера Украины на основе когнитивного моделирования

В последние годы когнитивное моделирование занимает всё более уверенные позиции в междисциплинарных науках. Спектр применения методологии когнитивного моделирования достаточно широк и включает такие разделы наук как: когнитивная психология, когнитивная социология, коммуникативно-когнитивная прагматика, когнитивное моделирование социальных и экономических систем [1].

Главной задачей когнитивного моделирования является более широкое, более точное и комплексное восприятие некоторого события, явления, процесса. Когнитивная модель позволяет исследователю систематизировать знания об изучаемом объекте, выделить основные факторы, влияющие на его поведение, определить основные закономерности процессов, происходящих с его участием. Когнитивный подход предлагает удобный способ изучения сложных процессов и систем, позволяя исследователю фокусироваться на самых важных аспектах решаемой задачи, подключая второстепенные факторы лишь по мере необходимости. Такое свойство когнитивного моделирования обуславливает успешное примение данного научного подхода в исследовании экономических систем, которые характеризуются высоким уровнем сложности процессов, происходящих в них, и чрезвычайно широким разнообразием факторов, определяющих поведение системы в целом.

Проведенный сценарный анализ экономики Украины на среднесрочную и долгосрочную перспективы позволил выявить главные драйверы или локомотивы экономического развития [2]. Одни из них - это области создания новых веществ и материалов, нанотехнологии, ИКТ, высокотехнологичное машиностроение и другие - всего десять позиций.

В данной работе когнитивное моделирование используется для исследования развития ИКТ кластера Украины. ИКТ сфера является одной из наиболее динамично развивающихся и наиболее перспективных в Украине. ИКТ отрасль дает возможность Украине конкурировать на мировом рынке, повышая добавленную стоимость внутреннего производства страны и при этом не нанося ущерба окружающей среде и не расходуя природных богатств. Несмотря на то, что по сравнению с другими отраслями ИКТ демонстрирует уверенный рост, данный кластер экономики является весьма неустойчивым ввиду высокой мобильности его участников, непрекращающейся утечки квалифицированных кадров, нестабильности геополитической ситуации и непрекращающегося экономического кризиса.

В работе предлагается построить сценарий развития ИКТ кластера при помощи когнитивного моделирования, достигая устойчивости когнитивной карты кластера с применением формальной теории устойчивости. Из результатов проведённого анализа делается вывод относительно устойчивости системы к возмущениям, устойчивости к заданным начальным значениям, а так же структурной устойчивости.

Далее планируется использовать методику инновационного развития ИКТ кластера Украины для определения рационального сценария его развития с использованием мощностей создаваемого в ИПСА ситуационно-аналитического центра (САЦ) в рамках Программы «Наука для мира и безопасности» (грант НАТО, 2015-2016 гг.).

Литература. **1.** Горелова Г.В., Панкратова Н.Д. Инновационное развитие социально-экономических систем на основе методологии предвидения и когнитивного моделирования //Наук. Думка, Киев, 2015.

2. Згуровский М.З., Панкратова Н.Д. Форсайт економіки України на середньострокову і довгострокову перспективу //XIII международная научно-практическая конференция «Математичне та програмне забезпечення інтелектуальних систем», Материалы конференции, 18-20 ноября 2015 г., Днепропетровск, Украина.- 2015. – С.75- 77.

***Печурин Н.К.*[1], *Кондратова Л.П.*[2], *Печурин С.Н.*[1]**
[1] *Институт компьютерных информационных технологий НАУ, Киев, Украина;* [2] *Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина*

Модель распределения вычислительных ресурсов в беспроводной компьютерной сети

При обмене вычислительными ресурсами в компьютерных сетях, использующих в качестве среды передачи информации радиоэфир, особое место занимает функция расчёта эффективного плана распределения локальных вычислительных ресурсов для решения общей прикладной задачи. В существующих беспроводных компьютерных сетях, обладающих принципиально ограниченными телекоммуникационными ресурсами, эта функция реализуется средствами системы управления сетью NMS. Традиционная (эталонная) модель NMS предполагает возможность управления потоками данных в сети, причём увеличение эффективности процедур управления достигается за счёт следующих составляющих (сервисов): управление производительностью, управление конфигурацией, управление учётными записями, управление отказоустойчивостью, управление безотказностью [1]. Здесь в качестве критериев используются показатели, «прописанные» на нижних уровнях семиуровневой эталонной модели взаимодействия открытых систем. Например, для предотвращения уменьшения эффективности из-за перегрузок сети может использоваться инструментарий буферизации, принудительного снижения скорости передачи данных, введение окон; соответствующими критериями эффективности могут служить максимальная пропускная способность составляющих сети, минимальное значение среднего времени реакции сети на запрос и подобное [2]. Эти критерии лишь опосредованно отражают истинную эффективность плана распределения вычислительных ресурсов, необходимых для решения общей прикладной задачи.

Для уменьшения этого (естественного для многоуровневой иерархической модели) критериального разрыва предлагается комплексная модель распределения вычислительных ресурсов в беспроводной компьютерной сети, основанная на классических формализациях распределения потоков в (абстрактных) сетях. Рассматривается (абстрактная) сеть, представленная ориентированным двунаправленным графом $G = (V, E)$ с множеством узлов V локализации базовых вычислительных ресурсов. Множество V дополняют рабочие станции, представляющие пары $(s_i, t_i) \in V$ источников и стоков, а также рабочая станция, представляющая общий источник s_0 для передачи вычислительных ресурсов в узлы s_i. Базовые и рабочие станции сети связаны дугами множества E с весами, характеризующими параметры буферизации и оконной технологии управления потоками данных, значения скорости передачи данных в канале, пропускные способности канала, стоимости транспортировки протокольной единицы данных. Задача заключается в расчёте плана распределения вычислительных ресурсов (определение потока в абстрактной сети) между узлами $s_0, s_i \in V$ при различных критериях, в частности, - минимизации стоимости потока, максимизации потока в сети. Способом решения потоковых задач выбрано сведение их к задачам линейного и целочисленного программирования. В линейном случае использован инструментарий блочного программирования [3], для целочисленной модели – переборные алгоритмы [4]. Планирование распределения вычислительных ресурсов на основе универсального инструментария математического программирования увеличивает гибкость представляемой в докладе комплексной модели.

Литература. **1.** Руководство по технологиям объединённых сетей: Пер. с англ.–М.: Издательский дом «Вильямс», 2005.–1040с. **2.** Рошан П., Лиэри Д. Основы построения локальных беспроводных сетей стандарта 802.11.: Пер. с англ. – М.: Издательский дом «Вильямс», 2004.–304с. **3.** Жариков А.В. Исследование скорости сходимости некоторых алгоритмов блочного линейного программирования // Управление, вычислительная техника и информатика.–2011.– С.100-105. **4.** Назарова И.А. Анализ переборных алгоритмов для задачи оценки уязвимости многопродуктовых сетей. – М.:ВЦ РАН, 2006.–72 с.

Пікоэс А.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Основа для автоматизованої інтелектуальної системи керування квадрокоптером

В роботі розглянуто і проаналізовано базу для автономної інтелектуальної системи керування квадрокоптером.

Комплементарний фільтр. Гіроскоп забезпечує хорошу швидку кутову оцінку для малих інтервалів часу, а акселерометр дає точні значення кута без дрейфу на довгих інтервалах часу. Ми повинні об'єднувати покази цих датчиків разом, щоб скористатися перевагами кожного з них.

Щоб поєднати покази датчиків разом так, щоб отримати найкращу оцінку кута, використовується спеціальний цифровий метод обробки даних – комплементарний фільтр. Комплементарний фільтр поєднує в собі цифровий фільтр верхніх частот з цифровим фільтром нижніх частот для отримання точної оцінки кута [1, 4].

Низькочастотний фільтр пропускає довгострокові зміни, фільтруючи короткострокові коливання. Одна з реалізацій такого фільтру:

$$\theta_{n+1} = a\theta_n + (1-a)\theta_{accel} \tag{1}$$

Якщо кут з датчика не змінюється, то оцінка кута буде наближатися до цього кута, поки не досягне його значення. Час, необхідний для досягнення незмінного кута, залежить як від коефіцієнта фільтра, так і від часових інтервалів між ітераціями.

Фільтр високих частот повністю протилежний фільтру низьких частот – він пропускає короткострокові зміни сигналу (високі частоти) і фільтрує мало змінні сигнали (низькі частоти). Він використовується для того, щоб подолати дрейф гіроскопа.

Повний фільтр виглядає так:

$$\theta_{n+1} = A(\theta_n + \theta'_{gyro}t) + (1-A)\theta_{accel} \tag{2}$$

Де A – коефіцієнт фільтру, який визначає часовий інтервал для довіри гіроскопу і фільтрації акселерометра.

ПІД регулятор. Після того, як програмою управління були отримані кути крену і тангажу, необхідно застосувати поправку до потужностей двигунів, для усунення можливого відхилення від потрібних кутів крену та тангажу. Але із законів динаміки обертового руху випливає, що безпосередньо управляти можна лише другою похідною кута. Тобто, контролюючи швидкість обертання двигунів, ми насправді контролюємо силу, з якою ротори тягнуть квадрокоптер. Для вирішення цієї задачі використовується ПІД регулятор.

На вхід регулятора подається узагальнена координата (в даному випадку кут), на виході отримуємо момент сил (друга похідна кута). Кожен ПІД регулятор стабілізує значення однієї узагальненої координати. Так як стабілізація по крену і тангажу – дві незалежні задачі, то потрібно використати два окремих ПІД з постійними коефіцієнтами.

Визначимо нев'язку – різницю між необхідним і реальним значенням кута:

$$e(t) = \alpha(t) - \alpha_0(t), \tag{3}$$

де α_0 – потрібне значення кута, α – поточне значення кута.

Запишемо ПІД регулятор для даного кута [2]:

$$u(t) = P + I + D = -\left(K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d \frac{d}{dt}e(t)\right) \tag{4}$$

Знак мінус говорить про те, що при позитивних K_d, K_i, K_p вплив спрямований проти відхилення квадрокоптера.

Оскільки похідною від кута є кутова швидкість, то у диференціальній складовій даного ПІД регулятора можна використати покази з гіроскопа. Це позбавить необхідності проводити

апроксимоване диференціювання і додасть точності у розрахунки.

$$u(t) = -\left(K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d e'_{gyro}(t)\right) \tag{5}$$

Для дискретного часу ця формула приймає наступний вигляд:

$$u_n = -\left(K_p e_n + K_i \sum_{j=1}^{n} e_j t_j + K_d e_{gyro_n}\right), \tag{6}$$

де $\Delta t_j = t_j - t_(j-1)$ – інтервал часу між $(j-1)$-ою і j-ою ітерацією.

Застосувавши такі ПІД регулятори до кутів тангажу і крену, можна вирішити проблему стабілізації.

Розрахунок потужностей двигунів. ПІД регулятори розраховують тільки поправки до потужностей двигунів, на загальну потужність вони не впливають. Загальна потужність контролюється окремо.

Акселерометр і гіроскоп розташовуються на платі так, щоб їх осі були співнапрямлені з власним осям квадрокоптера. Для конфігурації «плюс» електродвигуни з несучими пропелерами знаходяться на цих осях, а в конфігурації «ікс» вони повернуті на 45 градусів відносно осі перпендикулярної до площини квадрокоптера. Внесемо поправку на відповідні мотори для конфігурації «плюс».

Нехай $M(t)$ – загальна потужність двигунів в момент часу t, $u_\theta(t)$ – поправка потужності по куту тангажу, $u_\gamma(t)$ – поправка потужності по куту крену, $u_\psi(t)$ – поправка потужності по куту рискання, $P_1(t)$, $P_2(t)$, $P_3(t)$, $P_4(t)$ – потужності двигунів.

Щоб квадрокоптер не втрачав висоту, сума тяг гвинтів повинна зберігатися при довільному нахилі. З закону аеродинаміки, підйомна сила гвинта пропорційна квадрату кутової швидкості обертання [3, 4]. Тоді потужності двигунів розраховуються за такими формулами:

$$P_1(t) = \sqrt{M^2(t)/4 + u_\theta(t) + u_\psi(t)} \tag{7}$$

$$P_3(t) = \sqrt{M^2(t)/4 - u_\theta(t) + u_\psi(t)} \tag{8}$$

$$P_2(t) = \sqrt{M^2(t)/4 - u_\gamma(t) - u_\psi(t)} \tag{9}$$

$$P_4(t) = \sqrt{M^2(t)/4 + u_\gamma(t) - u_\psi(t)} \tag{10}$$

Можна перевірити, що при такому виборі потужностей підйомна сила зберігається:

$$P_1^2(t) + P_2^2(t) + P_3^2(t) + P_4^2(t) = \frac{M^2(t)}{4} + u_\theta(t) + u_\psi(t) + \frac{M^2(t)}{4} - u_\gamma(t) - u_\psi(t) + \tag{11}$$

$$+\frac{M^2(t)}{4} - u_\theta(t) + u_\psi(t) + \frac{M^2(t)}{4} + u_\gamma(t) - u_\psi(t) = M^2(t) \tag{12}$$

Література. **1.** Матвеев В.В. Основы построения бесплатформенных инерциальных навигационных систем / Матвеев В.В., Распопов В.Я. — СПб. : ГНЦ РФ ОАО "Кон-церн "ЦНИИ "Электроприбор", 2009. — С. 118–157. **2.** Денисенко В.В. ПИД-регуляторы: принципы построения и модификации / Денисенко В.В. — Современные технологии автоматизации. 2006. № 4. С. 66 – 74; 2007. № 1. С. 78 – 88. **3.** Tommaso Bresciani Modeling, identification and control of a quadrotor helicopter / Tommaso Bresciani – Master thesis, Lund University. 2008. **4.** Pounds P. "Modelling and Control of a Quad-Rotor Robot"/ Pounds, P.; Ma-hony, R., Corke, P. (December 2006) — In the Proceedings of the Australasian Conference on Robotics and Automation. Auckland, New Zealand.

Погребной Е.А.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Моделирование и анализ взаимодействия биологических популяций сложной структуры

Математическое моделирование развития биологических популяций это основной метод контроля за сохранением биологического разнообразия на планете Земля, одной из ключевых составляющих устойчивого развития. Этот метод позволяет прогнозировать численность популяций. Исходя из полученых прогнозов устанавливаются квоты отлова, идентифицируются виды которым в ближайшем будующем угрожает вымирание.

Используемые модели отражают лишь изменения общего количества популяции, принимая во внимание лишь некоторые наиболее общие факторы развития популяций в изолированой среде. Такие модели называются точечными моделями изолированых популяций. Как правило они записываются и анализируются в виде системы дифференциальных уравнений, в которой функциями выступают количества популяций. Наиболее известными среди них являются:

- Линейная модель Мальтуса развития одной популяции
- Логистическая модель развития одной популяции
- Модель Лотки-Вольтерра
- Модель Холлинга-Теннера

Эти и другие существующие модели рассматривают либо развитие одной изолированой популяции, либо взаимодействие двух изолированых видов. Одновременное взаимодействие большего числа видов на данный момент остаётся не исследованым.

Автор рассматривает обобщение модели Лотки-Вольтерра для большего количества популяций основываясь на следующих её принципах:

- Размер биологической среды достаточно велик
- Изменение численности каждого вида не зависит от времени и точки в пространстве
- Изменение количества популяции каждого вида пропорционально её текущей численности
- Непосредственное влияние численности одного вида на численность другого пропорционально численности каждого из них
- Непосредственное влияние одного вида на другой не зависит от остальных видов

Исходя из этих принципов автором была построена модель развития этих популяций в виде системы n дифференциальных уравнений вида

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = N_i \left(a_i + \sum_{j=1}^{n} b_{i,j} N_j \right), \quad i = \overline{1,n},\ b_{i,i} = 0$$

Эта модель была подробно рассмотрена для взаимодействия трёх видов, проанализированы её качественные особенности, найдены стационарные точки. Получены следующие её свойства:

- Всегда существуют две стационарные точки, соответствующие ситуациям когда один из видов вымирает.
- Аналитически найдены условия устойчивости [1] этих точек.
- Может существовать третья стационарная точка соответствующая сосуществованию всех трёх видов.

Аналитически определить условия существования и устойчивости третьей точки достаточно сложно, однако автору удалось численно подобрать значения коэффициентов модели, при которых эта точка существует и первые две стационарных точки устойчивы. Также был проведён анализ поведения модели при полученых коэфициентах, построены фазовые портреты [1] в окрестностях стационарных точек.

Литература. **1.** Леонтович Е.А. Методи і прийоми якісного дослідження динамічних систем на площині. — Москва, «Наука», 1990.

Попова Ю.А., Подладчиков В.Н.
УНК «ИПСА» НТУУ «КПИ», Киев, Украина

Идентификация неизвестных статистических характеристик шумов для нелинейной модели

Не секрет, что практически все процессы в природе, равно как и в экономике, подвержены колебаниям с определённой периодичностью и имеют некоторый тренд.Так же стоит учесть, что фрагмент колебания на одном уровне может быть трендом для колебаний на другом уровне. Именно так связанны между собой колебания экономических параметров. В доказательство этому можно привести теорию волн Кондратьева [4]. Учитывая этот факт, можно утверждать, что данная работа будет иметь большое прикладное значение, так как имеет необычайно широкую сферу применения.

Объект исследования. Цикличные экономические процессы на примере пассажирских перевозок международными авиалиниями [1].

Цель работы. Описать детерминированную основу циклического процесса гармоническими колебаниями с переменной амплитудой. На этой основе построить нелинейную стохастическую модель в пространстве состояний. Построить адаптивный фильтр Калмана [2] для решения задачи прогнозирования. Разработать алгоритм идентификации дисперсии шумов состояния и применить его к разработанной модели. Скорректировать работу фильтра с учётом полученных значений. Построить модель прогнозирования.

Методы исследования. Используется механизм перехода в пространство состояний путём записи основной формулы модели в рекуррентном виде, МНК [3], аппарат фильтра Калмана.

Результаты.

- Разработана методика построения стохастических моделей в пространстве состояний с использованием фильтров Калмана для анализа и прогнозирования циклических процессов с трендом.
- Сформирована стохастическая модель с неизвестными статистическими характеристиками шума: математическим ожиданием и дисперсией.
- Выполнена идентификация неизвестных статистических характеристик в предположении о пропорциональности дисперсии возмущений изменяющейся амплитуде колебаний.

Предложенный подход позволил количественно оценить уровень флюктуаций по полученным данным и обеспечить надежный прогноз , более точный по сравнению с известными методами.

Значимость работы. Предложенный подход позволяет построить модель интересующего нас экономического процесса, если он является циклическим с трендом, для предоставления возможности прогнозирования. Очевидно, что прогнозирование рынка необходимо при планировании работы любого предприятия: ведь оно терпит убытки как от перевозок в холостую или переизбытка изготовленной продукции, так и недополучает прибыль, если спрос на услуги/товары этого предприятия будет выше чем то, сколько они могут их предоставить в заданный промежуток времени.

Новизна в работе. это принципиально новый подход – описывать периодичный процесс в пространстве состояний и использовать для прогноза фильтры Калмана. Также научным достижением является расширение алгоритма идентификации неизвестных статистических характеристик шумов на нелинейный нестационарный случай.

Литература. **1.** . Дж.Бокс, Г.Дженкинс Анализ временных рядов прогноз и управление. – Москва «Мир», 1974. – 406с. **2.** Згуровский М.З. Подладчиков В.Н. Аналитические методы калмановской фильтрации для систем с априорной неопределённостью. – Киев «Наукова думка», 1995. – 283с. **3.** МНК, http://i.kpi.ua/podladchikov/index.htm?q=podladchikov/index.htm. **4.** Волны Кондратьева, .http://www.br.com.ua/referats/Econom_teoriya/2777-2.html.

Потебня А.В., Погорілий С.Д.
Київський національний університет імені Тараса Шевченка, Київ, Україна

Дослідження процесів самоорганізації у тематично близькій групі наукових публікацій

У статті виконано аналіз процесів формування прихованих структур у наукометричних базах даних та встановлено їх здатність до самоорганізації з проявом емерджентних властивостей. Виконано експериментальне дослідження фрагменту системи *IEEE Xplore*.

Вступ. Відомо, що останні десятиріччя пов'язані зі стрімким впровадженням у багатьох сферах наукових досліджень методів синергетики, орієнтованих на пошук універсальних законів еволюції та самоорганізації складних систем. Зокрема, проявом самоорганізації реальних систем є їх схильність до утворення прихованих структур та неявних спільнот з особливим режимом функціонування. У зв'язку з цим, динаміка розвитку складних систем визначається не лише зовнішніми впливами, але й перебігом внутрішніх процесів, неконтрольована активізація яких може призвести до руйнування системи з розпадом її на окремі нежиттєздатні компоненти [1]. Яскравим прикладом може слугувати явище турбулентності, що супроводжується виникненням низки вихорів з якісною зміною загальних властивостей потоку та сповільненням його швидкості. Виникнення подібних процесів у наукометричних базах даних може виявитися згубним для сучасного наукового розвитку, що обумовлює актуальність запропонованого дослідження. Слід зазначити, що розмірність реальних систем вимірюється мільйонами вузлів, тобто задачі їх аналізу належать до напряму *Big Data*.

Процеси самоорганізації у складних системах. Відповідно до уявлень синергетики, в процесі еволюції складної системи найбільш важливими є точки біфуркації, пов'язані з її переходом до альтернативних напрямків розвитку (рис. 1). Їх формування пов'язано з розвитком внутрішніх неврівноважених динамічних процесів, викликаних нарощенням флуктуацій, загасити які система самостійно не спроможна. При цьому тривале відхилення від стану рівноваги свідчить про виникнення дисипативності, що може призвести до катастрофи, пов'язаної з різкою зміною системи у напрямку притягнення її атрактором [2]. Прикладом такого процесу може слугувати стрімкий занепад соціальних мереж *Friendster* та *MySpace*, якому передував тривалий процес накопичення в них структурних недоліків. Як наслідок, після досягнення критичного рівня ураження архітектури, система, опинившись у точці біфуркації, починає розвиватися відповідно до вибраного атрактора, яким у цьому випадку виявляється остаточне руйнування архітектури. При цьому система може вийти за межі гомеостатичного стану, в якому вона здатна до справного виконання необхідних макрофункцій.

Таким чином, для повноцінної роботи та підтримки мінімального набору критично важливих функцій інформаційна система повинна мати деякий запас стійкості. Його забезпечення

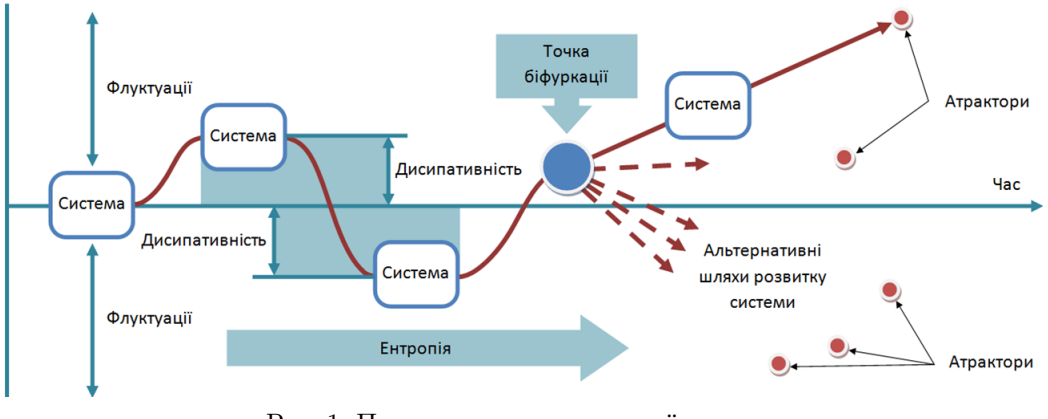


Рис. 1. Процес розвитку складної системи

потребує дотримання умов високої активності вузлів, структурної цілісності та значної живучості. Важливо, що за втрати цих компонентів система зазнає незворотних уражень архітектури, які призводять до дезорганізації управління, зниження гнучкості та втрати важливих макрофункцій. Разом з тим, синергетика гомеостазу інформаційної системи визначається балансом ентропії зовнішнього середовища та запасом живучості складових елементів [1, 2].

Структурна цілісність системи пов'язана з первинністю її макровластивостей відносно характеристик окремих компонентів, здатністю до самозбереження, еволюційного розвитку та підтримки стійкого стану. Живучі системи виявляють здатність до регенерації своєї структури після її пошкодження, що забезпечує відновлення втраченої архітектурної цілісності. З використанням структурного підходу інформаційні системи можуть бути представлені у вигляді динамічних мереж. При цьому поточний стан системи подається ациклічним графом $G_m = (V, E)$, де V – набір компонентів системи, а E – сукупність зв'язків між ними. Для встановлення наявності неявних спільнот у системі застосовується функція модулярності Ньюмана-Гірвана Q, що описує співвідношення густини внутрішніх та міжгрупових зв'язків за обраного розподілу вузлів. Експериментально показано, що значення модулярності, які перевищують 0.3 є свідченням реальної наявності спільнот у системі [3].

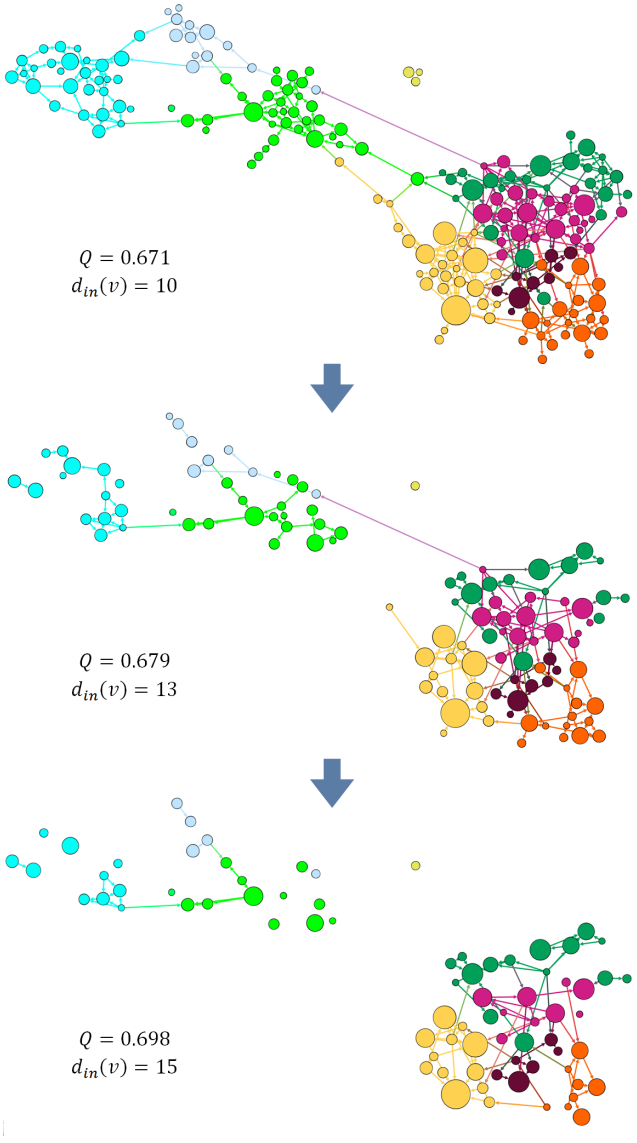


Рис. 2. Процес руйнування мостових з'єднань та збільшення значення модулярності при вилученні вузлів з середнім рівнем цитування

Експериментальні дослідження фрагмента системи IEEE Xplore. У цій роботі для дослідження обрано фрагмент наукометричної бази *IEEE Xplore*, що охоплює 2592 тематично близькі наукові публікації, присвячені візуалізації інформації, між якими встановлено 8984 цитувань. Зокрема, до складу вибірки входять праці конференцій *VIS*, *InfoVis*, *VAST*, а також наукові статті з журналу *IEEE Transactions on Visualization and Computer Graphics* [4]. Зв'язки з публікаціями, що виходять за межі відібраного фрагмента, були вилучені зі сформованого графа.

Шляхом оптимального розбиття такого графа було отримано значення $Q = 0.639$, що є підтвердженням високої активності процесів самоорганізації та свідчить про достатньо автономну діяльність спільнот у системі. При цьому лише дві з них охоплюють більше 10% вузлів і дев'ять – більше 5% статей. Природно, що значення модулярності для тематично близької групи публікацій виявилося на 12% нижчим, ніж значення Q, отримане для цілісної бази

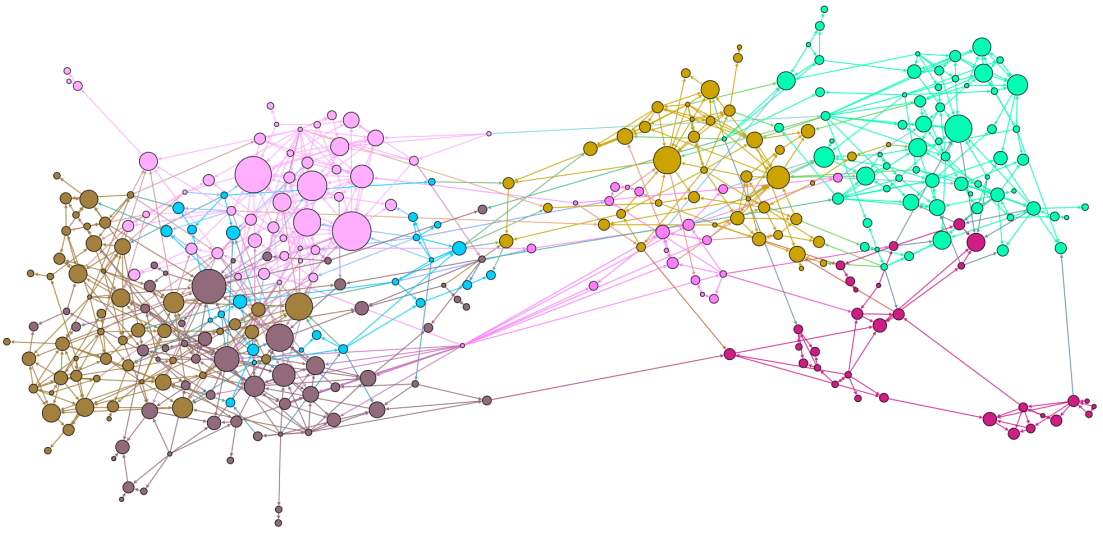


Рис. 3. Структура фрагменту наукометричної бази *IEEE Xplore* для публікацій, що мають більше 5 посилань. Розмір вузлів вказано пропорційно до індексу їх цитування

ArXiv [5]. Однак, за цілісністю структури ця група суттєво поступається его-графам соціальних мереж. Разом з тим, встановлено, що в тематично близькій групі публікацій, як і в цілісній наукометричній базі, основну зв'язувальну роль виконують статті з середнім рівнем цитування. Зокрема на рис. 3 показано процес руйнування мостових з'єднань між окремими спільнотами, спричинений збільшенням порогу цитування $d_{in}(v)$, за яким відбувалася фільтрація вузлів у графі. Загальна структура відібраного фрагменту наукометричної бази *IEEE Xplore* для статей, цитованих більше 5 разів, міститься на рис. 2.

Висновки. Формування неявних спільнот, що є проявом здатності складної системи до самоорганізації, забезпечує виділення найбільш актуальних сфер дослідження та задає загальний напрямок наукового розвитку. Натомість, публікації з середнім рівнем цитування беруть найбільш активну участь у процесах інформаційного обміну та мережевої мобілізації. Надзвичайно важливим для підтримки гомеостатичного стану сучасних систем є аналіз внутрішніх процесів динамічного розвитку неявних спільнот та мостових з'єднань між ними, оскільки повільні зміни архітектури можуть призвести до формування точки біфуркації, в якій поведінка системи стає непередбачуваною.

Література. **1.** C.S. Wagner, L. Leydesdorff, Network structure, self-organization, and the growth of international collaboration in science // Research Policy. – 2005. – Vol. 34 (10). – P. 1608 – 1618. **2.** В.А. Колодийчук, Синергетический эффект логистических систем // Журнал научных публикаций аспирантов и докторантов. – 2014. – №12 (102). – С. 28 – 31. **3.** А.В. Потебня, С.Д. Погорілий, Дослідження процесів поширення інформації у наукометричних базах даних // Системний аналіз та інформаційні технології: матеріали 17-ї Міжнародної науково-технічної конференції SAIT 2015 (Україна, Київ, 22 – 25 червня, 2015). – С. 97 – 102. **4.** Visualization Publication Dataset [Електронний ресурс]. – Режим доступу: URL: http://www.vispubdata.org. **5.** А.В. Потебня, С.Д. Погорелый, Анализ процессов информационного обмена в наукометрических базах данных // Открытые семантические технологии проектирования интеллектуальных систем: материалы 6-й международной научно-технической конференции OSTIS 2016 (Беларусь, Минск, 18 – 20 февраля, 2016). – С. 247 – 252.

Пудло І.В., Бідюк П.І.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Прогнозування дохідності банківських продуктів з використанням скорингового підходу

Кредитування - це найбільш прибутковий і одночасно ризикований вид банківської діяльності. Стратегія розвитку банку включає в себе управління кредитним та іншими типами фінансових ризиків, наприклад, ринковим та операційним. Побудова скорингової моделі є одним із основних методів зниження рівня втрат, починаючи з можливого неповернення тіла кредиту (дефолту) до недоотримання відсотків.

Револьверні картки з грейсовим періодом користуються найбільшою популярністю через можливість відстрочки платежу по кредиту [1]. Саме тому було вирішено побудувати аплікаційну модель для даного типу продукту. Ознакою дефолту такої моделі є частка недоотриманих відсотків по кредиту. Основне ускладнення полягає у тому, що ця модель є неперервною, оскільки прогнозоване значення це не просто 0 або 1 ("гарний" або "поганий" клієнт), а значення від 0 до 1, тобто це частка можливих втрат.

Оскільки сучасна історія не є найкращим представленням для найближчого майбутнього (кризовий період, коливання національної валюти, падіння ВВП тощо), то аналіз та побудова скорингової моделі базується на банківських даних за часовий проміжок з 01.07.2012 по 31.12.2012 рр. Основним мотивом у такому випадку є максимальне наближення популяції, на якій будується модель, до популяції, на якій вона буде застосовуватись. Після побудови модель була розглянута в динаміці до настання кризового періоду, тобто до 28.02.2014р.

В результаті проведення дослідження визначено такі параметри, як найбільш предикативні для прогнозування дефолту клієнта:

- APPBIRTHDATE - вік клієнта у заявці.
- APPFAMILYSTATUS - сімейний стан.
- APPPOSTCATEG - займана позиція.
- APPSEX APPBIRTHDATE - схрещений параметр статі та віку.
- APPSEX APPEDUCATION - схрещений параметр статі та освіти
- DEP NUMACT DEP SUM - схрещений параметр кількості активних депозитів та їх суми.
- APPWPERIODG - робочий стаж.
- CHAR10 - кількість місяців до планованої дати закриття.
- COMPPO - форма влаштування.
- CHR4 - кількість активних депозитів за 2 роки.
- FLAT - наявність житла.

Серед наведених вище параметрів є 3 схрещені параметри (APPSEX APPBIRTHDATE, APPSEX APPEDUCATION, DEP NUMACT DEP SUM). Суть схрещування полягає у збільшені предикативної сили цих параметрів. Таким чином, схрестивши два слабких параметри, ми отримуємо один, який є значно сильнішим за суму цих двох параметрів. Це суттєво підвищує адекватність математичної моделі, що будується на основі статистичних даних стосовно клієнтів (позичальників кредитів).

У результаті виконання обчислювальних експериментів побудована математична модель з індексом GINI 24% [2]. Динаміка індексу GINI коливається в межах допустимого рівня протягом всього розглянутого періоду (рис. 1).

Згідно зі створеною моделлю, граничні значення скорингового балу можуть коливатися від -92 до 205 (рис. 2). Концентрація популяції у скоринговому балі (розподіл з кроком 20) наведена також на рис. 2.

Саме такий підхід до аналізу даних позичальників кредитів застосовується у формуванні бізнес-стратегії.

Таким чином, у роботі запропонована методика аналізу можливих фінансових втрат внаслідок недоотримання відсотків по кредиту (фізичних осіб), яка ґрунтується на використанні

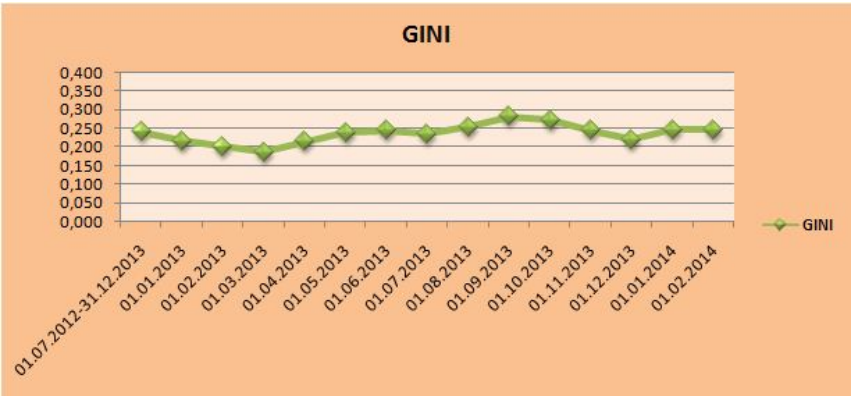


Рис. 1. Динаміка індексу GINI на вибраному часовому періоді

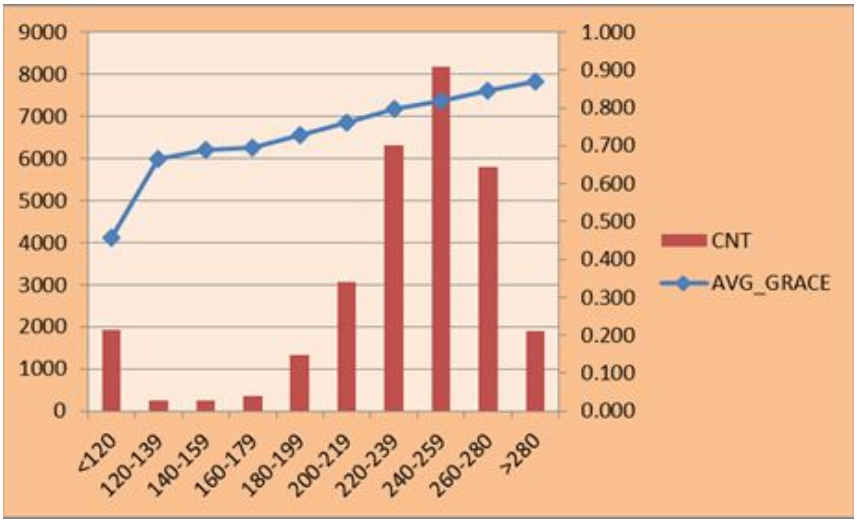


Рис. 2. Концентрація популяції у скоринговому балі

револьверних карток з грейсовим періодом, що користуються найбільшою популярністю у клієнтів завдяки наявності можливості відстрочки платежу за кредитом. Результатом дослідження стала побудована аплікаційна модель для даного типу продукту, яка має цілком прийнятні характеристики за точністю прогнозування можливих втрат. Для побудованої моделі індекс GINI мав значення 24%. Динаміка індексу GINI коливається в межах допустимого рівня протягом всього розглянутого періоду, що свідчить про високу стійкість побудованої моделі до змін у даних.

Література. **1.** Міщенко В. І. Банківські операції: Підручник /В.І. Міщенко, Н.Г. Слав 'янська, О.Г. Коренєва. — 2-е вид., переробл. і доп. — К.: Знання, 2007. — С. 280 — 283. **2.** Бідюк П.І., Матрос Є.О. Моделі оцінки ризиків кредитування фізичних осіб // Кібернетика та обчислювальна техніка. – 2007. – № 153. – С. 87-95.

Путренко В.В.
Світовий центр даних з геоінформатики та сталого розвитку НТУУ «КПІ», Київ, Україна

Статистичний аналіз даних землекористування для цілей регіонального управління в Україні

Розвиток технологій дистанційного зондування Землі дозволив сформувати тематичні продукти класифікації, які лягли в основу покриття землекористування. Дані землекористування є основою для управління природними ресурсами і регіонального планування. Європейська програма CORINE створила єдине покриття для країн Європи на основі затверджених стандартів. Україна до цього списку не потрапила, тому тривалий час точні дані землекористування не були доступні для території України. Це положення справ змінилося з виходом китайського глобального продукту, який дозволяє оцінити розподіл видів землекористування на території України і суміжних країн, а також провести паралелі з даними CORINE [1].

З метою підтримки досліджень глобальних змін і розвитку системних моделей функціонування Землі Міністерство науки і технологій Китаю запустило Національну високотехнологічну програма досліджень «Картографування за допомогою даних ДЗЗ і дослідження ключові технології побудови глобального рослинного покриву» у 2010 році. Одним з результатів проекту є картографічний продукт GlobeLand30-2010 глобального покриву з 30-метровий просторовим розрізненням, отриманий із зображень дистанційного зондування в 2010 році.

Для генерації даних земного покриття використовувалися класифіковані зображення із супутникових даних мультиспектральних зображень GlobeLand30-2010 з роздільною здатністю 30m, в тому числі Landsat TM і ETM+ і мультиспектральні зображення китайської системи Environmental Disaster Alleviation Satellite (HJ-1) [2].

Формування програм управління природними ресурсами і землекористуванням в Україні має чотири рівні організації: національний, регіональний, районний, місцевий. На кожному з рівнів відбувається коригування програм і вхідних даних, які використовуються для прийняття рішень в сфері землекористування. Дані землекористування є загальною інформаційною основою для прийняття управлінських рішень на цих рівнях. Однак на різних організаційних рівнях виникають протиріччя, основними причинами яких є диспропорції в розподіл різних видів землекористування на різних територіальних рівнях управління, неузгодженість цілей окремих суб'єктів управління, необхідність в узагальнення інформації на більш високих рівнях управління.

Для вирішення цих завдань була проведена обробка даних землекористування для територіальних одиниць кожного рівня управління. Нами були завантажені тайлові дані, які покривають територію України. Далі було проведено побудову тайловой мозаїки і її маскування межами кордону України. Отримані дані були оброблені за допомогою інструментів ArcTollbox ArcGis. В якості основних видів аналізу використовувалася група інструментів зонального аналізу, яка дозволяє отримати узагальнену статистику для декількох наборів даних. Зонування проводилося з використанням інструменту Tabulate Area, який обчислює перехресну таблицю площ між двома наборами даних і надає вихідну таблицю. В якості вхідних даних використовувався растровий набір даних землекористування і векторні набори адміністративних одиниць. На місцевому рівні використовувалися полігони сільських рад, на районному рівні - полігони міст і районів, на регіональному рівні - полігони областей України. Отримані результати зонування оцінювалися щодо статистичних показників по Україні, областях і районах. Основними статистичними характеристиками, які описують розподілу стали середні значення, амплітуда, дисперсія показників, а також такі показники як:

Ентропія даних:

$$E = -\sum_{i=1}^{n}\sum_{j=1}^{n} p(i,j) lg^{p(i,j)},$$

Однорідність:

$$O = -\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{p(i,j)}{1+(i-j)^2},$$

Несхожість:

$$E = \sum_{i=1}^{n}\sum_{j=1}^{n}|i-j|p(i,j),$$

де кожний $p(i,j)$ є значенням комірки в позиції (i,j) растру покриття.

Також велике значення мають методи класифікації, які використовуються при побудові картограм. Такий підхід дозволяє враховувати і погоджувати цілі раціонального землекористування і територіального планування на різних територіальних рівнях, тому що дає можливість порівняння показників як між адміністративними одиницями одного рівня, так і між регіоном і його складовими.

Табл. 1. Мінімальні та максимальні показники розподілу типів землекористування за регіонами та районами України у порівнянні

Показник	С/г землі,%	Ліси,%	Населені пункти,%	Луки,%
Min регіон	23,21	0,84	3,44	0,31
Min район	2,54	0,02	0,5	0
Max регіон	84,14	57,89	40,26	22,93
Max район	97,55	79,24	71,55	49,45

Розподіл амплітуди показників типів землекористування за регіонами та районами України у порівнянні показав, що для адміністративних одиниць меншого рангу за чотирма типами землекористування розподіл значень має більшу амплітуду, ніж для регіонів. Це вказує на ускладнення планування та прогнозування при переході від регіональних до районних та локальних показників.

Для всіх одиниць також було зроблено розрахунок індексу ентропії Шеннона, який вказує на різноманітність форм землекористування, що допомагає враховувати адміністративні одиниці з низьким рівнем різноманітності, пов'язані з антропогенним впливом. Низьке значення індексу Шеннона характерно для районів східної України, де воно пов'язане з високим рівнем урбанізації та південних районів, де рівень сільськогосподарських земель може досягати 80-90

Другий напрямок дослідження пов'язане з необхідністю проведення генералізації даних для прискорення розрахунків на регіональному та національному рівні. В даному випадку необхідно проведення збільшення розміру комірки растра, що позначається на точності розрахунків і їх перерозподіл між класами. Це вимагає вибору оптимального розміру комірки растра, при якому співвідношення між точністю результату і швидкістю обробки буде оптимальним.

Висновки. Сучасні дані землекористування є одним з базових продуктів класифікації даних дистанційного зондування Землі. Вони використовуються як основа для управління територіальним розвитком на різних рівнях регіонального управління. Необхідно, щоб українські дослідження проходили на системній основі і були інтегровані в європейський науковий простір. Дані землекористування створюють основу для отримання зональної статистики в вертикальному і горизонтальному розрізі для всієї території країни, що має бути використано при розробці планів використання території і збереження природного різноманіття.

Література. **1.** Попов М. А., Станкевич С. А., Сахацкий А. И., Козлова А. А. Использование полного набора нормализованных межканальных индексов многоспектральных космических изображений при классификации покрытий ландшафта // Уч. зап. Таврического нац. ун-та им. В. И. Вернадского. — 2007. — 20 (59), № 1. — С. 175–182. **2.** National Geomatics Center of China, 2014 National Geomatics Center of China 30 M Global Land Cover Data Product (GlobeLand30) (2014).

***Рубин Э.Е.*[1], *Годлевский М.Д.*[2], *Голоскокова А.А.*[2]**
[1] *Харьковский национальный университет радиоэлектроники, Харьков, Украина;* [2] *Национальный технический университет «Харьковский Политехнический Институт», Харьков, Украина*

Технология планирования улучшения качества процесса разработки программного обеспечения

В работе рассматривается технология улучшения качества процесса разработки программного обеспечения (ПР ПО), суть которой заключается в предварительном решении задачи стратегического планирования на некотором плановом подпериоде [1, T], результаты которго являются основой для решения задачи текущего планировнания на первом подпериоде. В дальнейшем задачу стратегического планирования будем решать на основе синтеза динамической модели, а задачу текущего планирований – на основе синтеза статической модели. Построения двух типов этих моделей базируется на вербальном описании модели зрелости CMMI [1].

В структуре модели CMMI, представленной в виде иерархии понятий, на самом нижнем уровне находятся практики, каждая из которых характеризуется определённым уровнем возможности. Каждая ключевая цель ПР ПО этой модели реализуется на основе ряда практик, которые позволяют снять или существенно уменьшить соответствующие проектные риски. Все цели сгруппированы в фокусные области. Каждая фокусная область вносит свой вклад в достижение некоторого уровня зрелости ПР ПО организации. Модель CMMI предполагает пять уровней зрелости. Формализация вербального описания этой модели реализована в [2, 3] путём введения переменных, на основе которых определяются уровни возможности и зрелости отдельных практик, фокусных областей и всего ПР ПО в целом.

Для динамической постановки задачи были построены модели, оптимальные значения переменных которых определяются на основе метода последовательного анализа вариантов. В качестве базового алгоритма, модификация которого была реализована применительно к поставленной задаче, был выбран алгоритм «Киевский веник» [4]. Входной информацией для решения задачи является исходное состояние ПР ПО IT компании в виде значений уровней возможностей отдельных практик и фокусных областей, целевого профайла по этим параметрам на конечном подпериоде планирования, а так же ограничения по финансовым и временным ресурсам в течении рассматриваемого периода планирования. Результатом решения задачи является значение уровней возможностей, практик и фокусных областей по подпериодам планирования.

В результате решения динамической задачи делается оценка необходимых ресурсов для достижения поставленной цели как суммарных, так и по подпериодам планирования. Результаты, полученные для первого подпериода, являются целевым профайлом для решения задачи текущего планирования, для которой была применена теория полезности, которая предполагает построение функции полезности степени достижения целевого профайла, а так же функции полезности ресурсного обеспечения. Это финансовые ресурсы и ресурсы, связанные со временем сотрудников компании, которое необходимо затратить для продвижения ПР ПО к более высокому уровню зрелости. Первый ресурсный критерий базируется на двух составляющих:

- Необходимые затраты на дополнительное техническое, информационное, программное и методическое обеспечение.
- Финансовые затраты на оплату труда сотрудников компании, которые овладевают новыми знаниями.

Обычно оплата труда базируется на затраченном времени и состоит из следующих трёх составляющих:

- Время на формулирование практики и её проработку.
- Необходимое время на обучение новой практике внутри компании.
- Время необходимое на контроль применения практики преодоление сопротивления и

притирку сотрудников.

Второй критерий фактически связан со второй составляющей первого критерия и определяется временем, которое необходимо потратить сотрудникам компании на освоение нового уровня возможности практик. Однако он имеет совершенно другой смысл. Дело в том, что это время приводит к отвлечению сотрудников компании от основного вида деятельности, связанного с разработкой ПО, а это может привести к напряженным графикам выполнения проектов, необходимости коррекции времени их выполнения и функциональности, многим другим негативным явлениям. В результате была построена модель многокритериальной оптимизации, состоящая из двух групп критериев:

1. Степень достижения целевого профайла на уровне всего ПР ПО.
2. Степень использования ресурсного обеспечения.

В качестве переменных модели используются дискретные целочисленные переменные, определяющие уровень возможности отдельной практики, а так же булевы переменные, определяющие вариант вклада ресурсов.

Рассматриваются различные варианты решения многокритериальной задачи.

1. Построение обобщенной функции полезности на основе свертки критериев с учетом весовых коэффициентов важности отдельных критериев на основе теоремы Карлина.
2. Нахождение некоторого компромиссного решения задачи на основе минимизации одинакового отклонения от оптимума всех критериев с учетом их весовых коэффициентов [5]. Данная задача решается методом ограничений.

Решение динамической задачи, на основе которого определяется целевой профайл для первого подпериода планирования, является целью при рассмотрении статической задачи. В свою очередь результаты, полученные в ходе решения статической задачи, уточняют план развития ПР ПО и являются исходной информацией для дальнейшего решения динамической задачи с последующей сдвижкой планового периода на один подпериод.

Результаты решения данной задачи предназначены для руководства IT компаний при управлении качеством ПР ПО, которые позволяют синтезировать оптимальную стратегию продвижения организации-разработчика программного обеспечения к определенному уровню зрелости в условиях ограниченных ресурсов.

Целью дальнейших исследований является разработка архитектуры и соответствующей информационной технологии скользящего планирования улучшения качества ПР ПО, проверка её работоспособности на реальной информации IT компаний, а так же разработка различных модификаций алгоритмов решения задач текущего и стратегического планирования с целью уменьшения времени решения задач, которые являются NP-сложными.

Литература. **1.** Ahern, D.M. CMMI Distilled: A Practical Introduction to Integrated Process Improvement, Third Edition / D.M. Ahern, A. Clouse, R. Turner. – Addison-Wesley, 2008. – 288 p. **2.** Годлевский М.Д. Синтез статических моделей планирования улучшения качества процесса разработки программного обеспечения / М.Д. Годлевский, А.А. Голоскокова // Східно–Європейський журнал передових технологій. – 2015. – № 3/2 (75). – С. 23–29. **3.** Годлевский М.Д. Динамическая модель планирования улучшения качества процесса разработки программного обеспечения / М.Д. Годлевский, Э.Е. Рубин, А.А. Голоскокова // Вісник Національного технічного університету "ХПІ". Збірник наукових праць. Серія: Системний аналіз, управління та інформаційні технології. – Х. : НТУ «ХПІ» – 2015. – № 58 (1167). – С. 3–6. **4.** Годлевский М.Д. Улучшение качества процесса разработки программного обеспечения на основе методов последовательного анализа вариантов и локальной оптимизации / М.Д. Годлевский, Э.Е. Рубин, А.А. Голоскокова // Вісник Національного технічного університету "ХПІ". Збірник наукових праць. Серія: Системний аналіз, управління та інформаційні технології. – Х. : НТУ «ХПІ» - 2014. - № 55 (1097). - С. 5-10. **5.** Михалевич В.С. Вычислительные методы исследования и проектирования систем / В.С. Михалевич, В.Л. Волкович. – М.: Наука, 1982. – 287 с.

Савченко І.О., Іваненко Є.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Моделювання воєнних конфліктів за допомогою модифікованого методу морфологічного аналізу

З кінця XX століття новітні технології набувають визначального значення. На сьогоднішній день можливість передбачити, яким буде розвиток тієї чи іншої галузі, є визначальним фактором досягнення успіху [1]. В даному дослідженні увага приділяється прогнозуванню наслідків воєнних конфліктів, оскільки воєнні зіткнення були і є на сьогодні доволі поширеним явищем: від невеликих локальних сутичок до широкомасштабних протистоянь. Таким чином, здатність адекватно змоделювати стан, перебіг та можливі наслідки воєнних конфліктів є важливим аспектом для здобуття перемоги та успішного подолання наслідків, який би характер вони не мали.

На сьогодні у багатьох країнах проводяться подібні дослідження. Зокрема, Шведське морфологічне товариство проводило аналіз щодо використання системи наземного розпізнавання цілей для вибору типу зброї. В результаті було отримано кілька моделей, використання яких є ефективним засобом боротьби проти ряду наземних ворожих об'єктів [2].

Для дослідження було використано модифікований метод морфологічного аналізу, як один із потужних методів технологічного передбачення, що дозволяє розглянути одночасно велику кількість різних конфігурацій об'єкта [3]. Даний метод має у своєму складі засоби для прогнозування розвитку досліджуваної системи, а також дозволяє враховувати чинники навколишнього світу, які впливають на систему і, як наслідок, планувати стратегію поведінки у майбутньому.

На першому етапі була побудована морфологічна таблиця для опису воєнних конфліктів, яка містила дев'ять характеристичних параметрів:

- тип конфлікту: *внутрішній конфлікт/громадянська війна, регіональний конфлікт, воєнний конфлікт між територіально розподіленими державами, тероризм, вітчизняна війна, "холодна війна", світова війна*;
- тип використовуваної зброї: *стрілецька/гранати, вибухівка/міни, хімічна, ядерна, бронетехніка, авіація, артилерія, запалювальна*;
- що призвело до конфлікту: *напруженість у стосунках між державами, замах/усунення керівної особи, порушення територіальних кордонів*;
- основний сектор враження: *сільське господарство, інфраструктура, області проживання, виробництво, видобування ресурсів та виробництво енергії*;
- тривалість конфлікту: *до 0.5 року, 1-2 роки, до 5 років, до 10 років*;
- умови проживання: *забруднення, військові дії, руйнування, недостатнє постачання води/їжі*;
- тощо (жертви; сторони конфлікту; причини виникнення конфлікту);

Представлення воєнного конфлікту у вигляді морфологічної таблиці дозволяє на основі експертного оцінювання побудувати його морфологічну модель і оцінити ймовірності появи альтернатив параметрів [3]. Це дає змогу на другому етапі знаходити пріоритети альтернатив щодо можливих засобів запобігання або усунення наслідків, які є потенційно найбільш результативними в умовах різноманіття воєнних конфліктів, описаних морфологічною таблицею. Також на основі морфологічної моделі можуть оцінюватись ризики при воєнних конфліктах (наприклад, жертви серед мирного населення, мародерство).

Література. **1.** Згуровський М. З. Геоекономічні сценарії розвитку і Україна / М. З. Згуровський, Ю. М. Пахомов, А. С. Філіпенко. – К. "Академія", 2010. – 323 с. **2.** Ritchey, T. Developing Scenario Laboratories with Computer-Aided Morphological Analysis / Presented at the 14th International Command and Control Research and Technology Symposium Washington DC – June 15-17, 2009. **3.** Панкратова Н.Д., Савченко І.О. Морфологічний аналіз. Теорія, проблеми, застосування. Навчальний посібник. – Київ: Вид-во «Наукова думка». – 2015. – 244 с.

***Сальніков М.М.*[1], *Вергун К.В.*[2]**
[1]*Інститут космічних досліджень НАНУ та ДКАУ, Київ, Україна;* [2]*НТУУ "КПІ", факультет прикладної математики Київ, Україна*

Відновлення просторової картини еволюції хвильових процесів у сферичному шарі за результатами вимірів вздовж лінії

Задачі оцінювання стану динамічних систем виникають у різноманітних сферах науки і техніки. Окремо виділяються задачі оцінювання стану систем з розподіленими параметрами, які потребують застосування особливих методів оцінювання [1].

У даній роботі розглянута модельна задача оцінювання стану хвильового процесу в області сферичного шару з використанням вимірів оцінюваної величини, що здійснюються вздовж кільцевих ліній, які знаходяться у площинах, що проходять через центр сфери. Ця задача моделює дослідження хвильових процесів, що протікають в верхніх шарах атмосфери або магнітосфери Землі на базі вимірювань, які здійснюються супутниками вздовж траєкторії польоту. Мета роботи полягає у побудові та вивченні властивостей алгоритму оцінювання, який буде придатний для застосування щодо моделей, що більш точно відображають специфіку реальних фізичних процесів в атмосфері та магнітосфері.

Постановка задачі. У якості моделі хвильового процесу обрана змішана крайова задача для хвильового рівняння у сферичному шарі, з заданими однорідними крайовими умовами першого роду на внутрішній межі шару та другого роду на зовнішній межі:

$$\Delta u = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}, \tag{1}$$

$$u|_{\Gamma_1} = 0, \quad \left.\frac{\partial u}{\partial \vec{n}}\right|_{\Gamma_2} = 0, \tag{2}$$

де Γ_1 — внутрішня поверхня області, а Γ_2 — зовнішня поверхня.

Враховуючи специфіку області протікання процесу, доцільним є перехід у сферичну систему координат. При цьому форма рівняння (1) суттєво ускладнюються, однак крайові умови (2) можуть бути виражені простими співвідношеннями:

$$\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial u}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 u}{\partial \phi^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}, \tag{3}$$

$$u|_{r=r_0} = 0, \quad \left.\frac{\partial u}{\partial r}\right|_{r=r_1} = 0, \tag{4}$$

де $r \in [r_0, r_1], \phi \in [0, 2\pi), \theta \in [0, \pi]$.

Метод побудови алгоритму. Початкові умови для цього рівняння невідомі. Задача полягає у відновлені стану системи з використанням вимірів, що здійснюються вздовж кільцевих ліній $r = const$ з невідомими, але обмеженими похибками. Стохастичні характеристики оцінюваних величин задачі вважаються невідомими, більш того, невідомо чи вони мають схотастичну природу. Тому для розв'язку задачі можна використовувати лише так звані нестохастичні методи оцінювання, які не використовують інформацію про ймовірнісні характеристики невідомих величин.

У роботі [2] продемонстровано застосування методу еліпсоїдів до оцінювання стану круглої мембрани, що коливається. Автори поточного дослідження ставлять за мету розширити застосування методу еліпсоїдів на аналогічну тривимірну задачу щодо сферичного шару.

Література. **1.** Егоров А. И. Основы теории управления. — А. И. Егоров. — М.: Физматлит, 2004 — 504 с. **2.** Сальников Н. Н. , А. Н. Сальникова Эллипсоидальная фильтрация процессов с распределенными параметрами с использованием конечномерных моделей и локальных измерений. — Проблемы управления и информатики, 2015, № 3, С.75-90.

Сащук І.М., Романов О.М.
Науково-дослідний інститут Міністерства оборони, Київ, Україна

Застосування системного підходу для синтезу структури апаратно-програмних комплексів технічного аналізу сигналів на основі інформаційно-сигнатурних технологій

Подальший розвиток суспільства нерозривно пов'язаний з широким впровадженням у всі сфери людської діяльності комунікаційних та інформаційних технологій, їх інтеграцією і, як наслідок, створенням складних інформаційно-телекомунікаційних систем. Їх велике розмаїття потребує, з одного боку, застосування універсальних методів аналізу та синтезу, а з іншого – індивідуального підходу до визначення оптимальної структури конкретної складної системи залежно від призначення та сфери застосування.

Так, широкого розповсюдження набули системи радіо- та телевізійного цифрового мовлення, геофізичного моніторингу, вимірювально-діагностичні тощо, у яких використовуються різноманітні методи та протоколи передачі даних. Доступ до інформації, що циркулює в таких системах, здійснюється на основі технічного аналізу сигналів (ТАС) за допомогою апаратно-програмних комплексів (АПК). До АПК ТАС пред'являються такі основні вимоги:

- проведення ТАС в умовах суттєвої апріорної невизначеності про структуру та параметри складних сигналів, що використовуються в системах телекомунікацій;
- оперативність аналізу, що полягає у досягненні мети технічного аналізу у визначений термін;
- повнота аналізу, що полягає у забезпеченні енергетичного, структурного та інформаційного доступу до систем телекомунікацій.

Очевидно, що забезпечення цих, до певної міри суперечливих, вимог можливе, якщо при синтезі АПК використовуються основні принципи системного підходу: системність та ієрархічність дослідження [1]. Принцип системності передбачає, що АПК ТАС слід розглядати як складну систему, в якій зв'язки між апаратними та програмними засобами комплексу визначають його структуру, а системний ефект (повнота доступу до інформації при збереженні оперативності) досягається лише при спільному їх застосуванні. Ієрархічність вимагає врахування трьох рівнів: по перше, АПК ТАС є елементом більш складної системи інформаційного забезпечення прийняття рішень; по друге, АПК ТАС, відповідно до принципу системності, слід розглядати як окрему систему; по третє, елементи комплексу – апаратні і програмні засоби аналізу сигналів – хоча і мають самостійне значення, але розробляються лише у взаємозв'язку та розвитку.

Врахування часткової апріорної інформації про об'єкт ТАС при синтезі АПК досягається на основі впровадження інформаційно-сигнатурних технологій (ІСТ) – прийомів, способів та методів автоматизованого збору, обробки, накопичення, зберігання та передачі інформації про об'єкт (подію, явище) на основі виділення та аналізу його сигнатури (множини частинних сигнатур). Сигнатура об'єкта – сукупність характерних системних ознак (показників, параметрів, властивостей, характеристик), притаманних безпосередньо саме досліджуваному об'єкту. Складається із частинних сигнатур.

Інформаційна система, що реалізує ІСТ, включає в себе датчики, що можуть бути побудовані на різних фізичних принципах чи інші джерела інформації, засоби (апаратні та програмні) первинної обробки інформації, засоби (як правило, програмні) вторинної обробки інформації, базу даних сигнатур об'єктів, органи прийняття рішення.

Застосування принципів системного підходу на основі ІСТ розглянемо на прикладі розробки АПК ТАС телекомунікаційних систем, що використовують передачу даних із застосуванням мережевої моделі взаємодії відкритих систем Open systems interconnection basic reference model (OSI) в умовах часткової апріорної невизначеності (рис. 1).

АПК ТАС, що вирішують таку задачу, можна віднести до складних інтелектуальних систем, оскільки на кожному рівні моделі може застосовуватись широкий перелік протоколів. Ланцюг

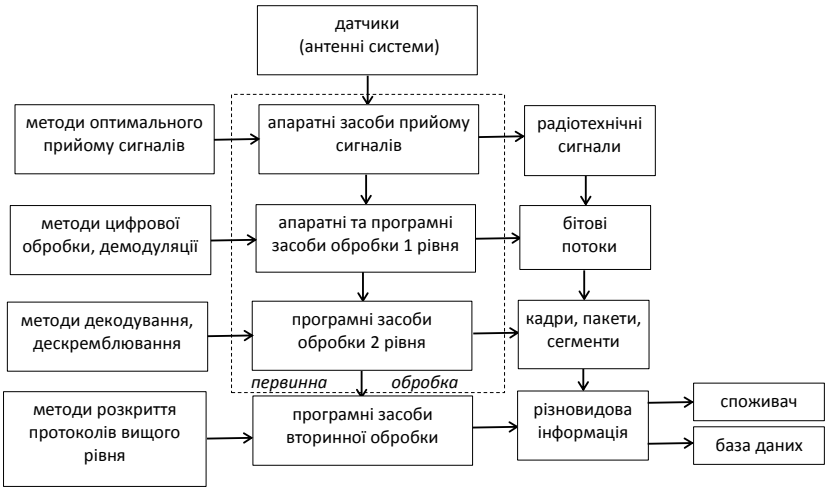


Рис. 1. Структура АПК ТАС на основі ІСТ

протоколів від верхнього до нижнього рівнів моделі залежить від топології системи, виробника апаратури, що в ній використовується, та налаштувань апаратури і є практично унікальним.

В ході проведення технічного аналізу, починаючи з нижнього рівня моделі OSI, сигнал з фізичного середовища перетворюється у двійкові дані. При цьому застосовуються традиційні методи цифрової обробки сигналів – використання віконних функцій, перетворення Фур'є, методів визначення модуляційних параметрів сигналів, їх цифрової фільтрації та децимації з подальшою демодуляцією. Реалізація первинної обробки здійснюється із застосуванням SDR [2], FPGA, CUDA та інших технологій.

На наступному рівні роботи АПК ТАС здійснюється визначення параметрів перешкодостійкого кодування, декодування, дескремблювання, демультиплексування. Обробці підлягають двійкові дані з використанням математичного апарату теорії матриць, полів Галуа, вейвлет-аналізу.

Для визначення протоколів при послідовному перетворенні сигналу з двійкового виду в кадри (frame), пакети (packet), сегменти (segment) і дейтаграми (datagram) використовуються алгоритми пошуку в отриманих реалізаціях відомих структур протоколів на основі апріорних відомостей з бази даних та визначення протоколів наступного рівня.

На верхніх рівнях моделі OSI визначаються протоколи відображення даних споживачу інформації: текстова, графічна, голосова та відеоінформація.

На більшості рівнів моделі можливе широке застосування теорії розпізнавання отриманих реалізацій [3].

Загальне функціонування АПК в залежності від протоколу, визначеного на попередньому рівні, може реалізовуватись також під управлінням штучного інтелекту.

Швидкість визначення протоколів та кількість протоколів, які можуть паралельно оброблятися в масштабі реального часу, визначається повнотою бази даних відомих протоколів, наявними обчислювальними ресурсами, застосовуваним математичним апаратом та розробленим програмним забезпеченням.

Література. **1.** Романов В.Н. Системный анализ для радиоинженеров. – СПБ: СГЗТУ – 2006. – 186 с. **2.** Силин Алексей. Технология Software Defined Radio. Теория, принципы и примеры аппаратных платформ // Беспроводные технологии. – 2007. – №2. – С. 22–27. **3.** Ревуцкий В.А. Устойчивые к мешающим факторам алгоритмы распознавания вида помехоустойчивых кодов в радиотехнических системах: автореф. дис. на соиск. уч. степ. канд. техн. наук: 05.12.04 / Вадим Андреевич Ревуцкий. — Рязань, 2013. — 16, [3] с.

Сирик С.В., Цверчкова А.А.
Факультет прикладной математики НТУУ "КПИ", Киев, Украина

Коррекция и Фурье-анализ конечноэлементных стабилизированных вычислительных схем при использовании технологии mass lumping

Известно (см. [1–4]), что применение сосредоточения (mass lumping) в методе Петрова-Галеркина (МПГ) [3–5] может приводить к дисперсионным и диссипативным ошибкам, и, соответственно, к значительным погрешностям численного решения. Это является существенным недостатком приема сосредоточения, изучению и преодолению которого посвящены работы [2–4]. Подход работы [3] основан на том, чтобы с помощью специального выбора стабилизирующих параметров сделать решение МПГ с использованием сосредоточения близким (в смысле локальной погрешности) к решению исходной (несосредоточенной) аппроксимации МПГ. Для этого сравниваются локальные погрешности исходной и сосредоточенной аппроксимаций МПГ на решении дифференциального уравнения. Отметим, что в работе [3] рассматривалось применение МПГ к достаточно общему уравнению конвекции-диффузии (КД) с переменными коэффициентами. Подход работы [2] основан на использовании матричных рядов для аппроксимации матрицы M^{-1} (обратная к матрице масс M МПГ): для этого M представляется в виде $M = \bar{M}(I - A)$ (где I — единичная матрица, $A \equiv \bar{M}^{-1}(\bar{M} - M)$, $\bar{M}$ — диагональная матрица, элементы диагонали которой равны суммам элементов соответствующих строк M), откуда $M^{-1} = (I + A + A^2 + \cdots)\bar{M}^{-1}$ (ряд Неймана). В [2] авторы отмечают, что уже использование $(I + A)\bar{M}^{-1}$ вместо M^{-1} способно значительно улучшить точность численного решения. Заметим, что (в отличии от [3]) в работе [2] рассматривалось только уравнение переноса с постоянными коэффициентами, а основное внимание уделено полудискретным аппроксимациям Галеркина (и сходимость соответствующих матричных рядов Неймана доказана только для классического МКЭ Галеркина с линейными элементами). В работе [4] подход работы [3] получил дальнейшее развитие (а также произведено сравнение указанных подходов работ [2] и [3]), результаты статьи [3] обобщены и показано, что в некоторых случаях подходы работ [2] и [3] приводят к одинаковым вычислительным схемам (а для тех случаев, где схемы не совпадают, в [4] были приведены примеры, в которых схемы [3, 4] обеспечивают более высокую точность численного решения по сравнению со схемами из [2]).

Рассмотрим одномерное нестационарное уравнение КД [1–4]

$$\frac{\partial u}{\partial t} + \lambda \frac{\partial u}{\partial x} - \kappa \frac{\partial^2 u}{\partial x^2} = 0,$$

где $\lambda = \lambda(t)$ и $\kappa = \kappa(t)$ заданы, а $u = u(x,t)$ — неизвестное решение. Для численного решения используем МПГ с кусочно-линейными базисными $\{N_i(x)\}$ и кусочно-квадратичными весовыми функциями [3, 4] $\{W_i(x)\}$ (с соответствующими стабилизирующими параметрами α_i) на равномерной сетке с шагом h. Введем сокращенные обозначения для выражения разностных соотношений: $u_{\bar{x},i} \equiv (u_i - u_{i-1})/h$, $u_{x,i} \equiv (u_{i+1} - u_i)/h$, $u_{\mathring{x},i} \equiv (u_{i+1} - u_{i-1})/(2h)$, $u_{\bar{x}x,i} \equiv (u_{\bar{x}})_{x,i} = (u_{i+1} - 2u_i + u_{i-1})/h^2$. Используя сосредоточение и опираясь на [3, 4], можно получить следующие полудискретные аппроксимации для нахождения $\{a_i(t)\}$:

$$\dot{a}_i + \lambda a_{\mathring{x},i} = \kappa a_{\bar{x}x,i} + \theta_3 a_{\bar{x}\mathring{x}x,i} + \mu_4 a_{\bar{x}x\bar{x}x,i} + \theta_5 a_{\bar{x}x\mathring{x}\bar{x}x,i} + \mu_6 a_{\bar{x}x\bar{x}x\bar{x}x,i} + \theta_7 a_{x\bar{x}x\mathring{x}\bar{x}x\bar{x},i} + \mu_8 a_{\bar{x}x\bar{x}x\bar{x}x\bar{x}x,i}, \quad (1)$$

где коэффициенты θ_3, θ_5, θ_7, μ_4, μ_6, μ_8 определяются следующими выражениями:

$$\theta_3 \equiv \frac{\kappa\alpha_i h}{2} + \frac{\lambda h^2}{6}, \quad \theta_5 \equiv -\left(\frac{\alpha_i \kappa h^3}{24} + \frac{\lambda h^4}{36}\right), \quad \theta_7 \equiv \frac{\kappa\alpha_i h^5}{180} + \frac{\lambda h^6}{180},$$

$$\mu_4 \equiv -\left(\frac{\alpha_i \lambda h^3}{24} + \frac{\kappa h^2}{6}\right), \quad \mu_6 \equiv \frac{\kappa h^4}{72} + \frac{\alpha_i \lambda h^5}{240}, \quad \mu_8 \equiv -\frac{\kappa h^6}{540}.$$

Частные случаи соотношения (1) (получающиеся при отбрасывании членов величин $O(h^3)$ и

$O(h^5)$ соответственно), уравнения

$$\dot{a}_i + \lambda a_{\mathring{x},i} = \kappa a_{\bar{x}x,i} + \theta_3 a_{\bar{x}\mathring{x}x,i} - \left(\kappa h^2/6\right) a_{\bar{x}x\bar{x}x,i}, \tag{2}$$

$$\dot{a}_i + \lambda a_{\mathring{x},i} = \kappa a_{\bar{x}x,i} + \theta_3 a_{\bar{x}\mathring{x}x,i} + \mu_4 a_{\bar{x}x\bar{x}x,i} + \theta_5 a_{\bar{x}x\mathring{x}\bar{x}x,i} + \left(\kappa h^4/72\right) a_{\bar{x}x\bar{x}x\bar{x}x,i}, \tag{3}$$

были выведены в работах [3] и [4].

При использовании подхода [2] (в этом случае все $\alpha_i = 0$) и использовании $(I + A)\bar{M}^{-1}$ вместо M^{-1} получим схему (2) (доказательство приведено в [4]). При использовании $(I + A + A^2)\bar{M}^{-1}$ вместо M^{-1} получим уравнение (3), в котором θ_3, θ_5, μ_4 равны прежним значениям, а коэффициент перед $a_{\bar{x}x\bar{x}x\bar{x}x,i}$ равен $\kappa h^4/36$ (доказательство также приведено в [4]). При использовании выражения $(I + A + A^2 + A^3)\bar{M}^{-1}$ для аппроксимации M^{-1} получим уравнение $\dot{a}_i + \sum\limits_{j=i-3}^{i+3} \tilde{d}_j \left\{\lambda(a_{j+1} - a_{j-1})/2 - \kappa(a_{j+1} - 2a_j + a_{j-1})/h\right\}/h = 0$ (где $\tilde{d}_i = 43/27$, $\tilde{d}_{i\pm1} = -25/72$, $\tilde{d}_{i\pm2} = 1/18$, $\tilde{d}_{i\pm3} = -1/216$), приводящееся к (1), в котором θ_3, θ_5, μ_4 равны прежним значениям, а коэффициенты $\mu_6 = \kappa h^4/36$, $\theta_7 = \lambda h^6/216$, $\mu_8 = -\kappa h^6/216$.

Проведен Фурье-анализ систем (1)-(3) (изучение поведения допускаемых системами плоских монохроматических гармоник вида $a_i(t) = e^{\omega t} e^{jihp}$, где $j = \sqrt{-1}$, h — шаг сетки, $p \in \mathbb{R}$ — (пространственная) частота гармоники, $\omega \in \mathbb{C}$ — величина, определяющая диссипативные и дисперсионные характеристики систем). В частности, показано, что для (1) имеет место

$$\mathrm{Re}(\omega) = -\frac{4\kappa}{h^2}\sin^2\frac{ph}{2} - \left(\frac{8\kappa}{3h^2} + \frac{2\alpha\lambda}{3h}\right)\sin^4\frac{ph}{2} - \left(\frac{8\kappa}{9h^2} + \frac{4\alpha\lambda}{15h}\right)\sin^6\frac{ph}{2} - \frac{64\kappa}{135h^2}\sin^8\frac{ph}{2},$$

$$\mathrm{Im}(\omega) = -\left(\frac{\lambda}{h} + \left(\frac{2\lambda}{3h} + \frac{2\alpha\kappa}{h^2}\right)\sin^2\frac{ph}{2} + \left(\frac{4\lambda}{9h} + \frac{2\alpha\kappa}{3h^2}\right)\sin^4\frac{ph}{2} + \left(\frac{16\lambda}{45h} + \frac{16\alpha\kappa}{45h^2}\right)\sin^6\frac{ph}{2}\right)\sin(ph).$$

Соответствующие выражения ω_2 и ω_3 для (3) и (2) имеют вид

$$\mathrm{Re}(\omega_2) = -\frac{4\kappa}{h^2}\sin^2\frac{ph}{2} - \left(\frac{8\kappa}{3h^2} + \frac{2\alpha\lambda}{3h}\right)\sin^4\frac{ph}{2} - \frac{8\kappa}{9h^2}\sin^6\frac{ph}{2},$$

$$\mathrm{Im}(\omega_2) = -\left(\frac{\lambda}{h} + \left(\frac{2\lambda}{3h} + \frac{2\alpha\kappa}{h^2}\right)\sin^2\frac{ph}{2} + \left(\frac{4\lambda}{9h} + \frac{2\alpha\kappa}{3h^2}\right)\sin^4\frac{ph}{2}\right)\sin(ph),$$

$$\mathrm{Re}(\omega_3) = -\frac{4\kappa}{h^2}\sin^2\frac{ph}{2} - \frac{8\kappa}{3h^2}\sin^4\frac{ph}{2}, \qquad \mathrm{Im}(\omega_3) = -\left(\frac{\lambda}{h} + \left(\frac{2\lambda}{3h} + \frac{2\alpha\kappa}{h^2}\right)\sin^2\frac{ph}{2}\right)\sin(ph).$$

Фурье-анализ показал, что оригинальные уравнения (1)-(3) в сравнении с уравнениями, получающимися после применения подхода [2], дают более точные численные решения (в смысле невязки между численным и точным решением задачи), однако, менее близкие к решению МПГ (по сравнению со схемами на основе подхода [2]). Это также подтверждается проведенными численными экспериментами.

По аналогии с [3,4], от (1) можно перейти к семействам разностных схем с весами и показать их среднеквадратическую устойчивость в сеточной норме L_2.

Литература. **1.** *Wendland E., Schulz H.E.* Numerical experiments on mass lumping for the advection-diffusion equation // Revista Minerva – 2005. – Vol. 2, № 2. – P. 227-233. **2.** *Guermond J.-L., Pasquetti R.* A correction technique for the dispersive effects of mass lumping for transport problems // Comput. Methods Appl. Mech. Engrg. – 2013. – Vol. 253. – P. 186-198. **3.** *Сирик С.В.* Анализ применения сосредоточенных аппроксимаций в методе конечных элементов при решении задач конвекции-диффузии // Кибернетика и системный анализ – 2013. – № 5. – С. 152-163. **4.** *Сирик С.В., Сальников Н.Н.* О применении сосредоточения в методе конечных элементов Петрова-Галеркина при решении задач конвекции-диффузии // Доповіді НАН України – 2014. – № 5. – С. 39-44. **5.** *Сирик С.В.* Оценки точности конечноэлементного метода Петрова-Галеркина при интегрировании одномерного стационарного уравнения конвекции-диффузии-реакции // Украинский математический журнал – 2015. – Т.67, № 7. – С. 937-961.

Слухай Я.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Алгоритми логічного виведення фактів в OWL-онтологіях

Сьогоднішні виклики в IT-галузі полягають, зокрема, в організації автоматичного пошуку, збору і обробки інформації, об'єми якої інтенсивно зростають. Зокрема, окремим завданням цього напрямку, яке покликані вирішити технології Semantic Web, є автоматизація роботи з різноманіттям даних у Всесвітній мережі. Серед останніх значних досягнень там - специфікація мови опису онтологій OWL.

Під онтологією тут слід розуміти представлення знань, що побудоване за допомогою певної описової мови, яке можна розділити на елементи (сутності, категорії), та пов'язати їх між собою через відношення. Така модель забезпечує здатність виводити нові логічні висновки зі старих, знаходити закономірності в базах знань. Загалом, автоматичні висновки, навіть дуже прості, забезпечують ефективнішу навігацію і пошук.

OWL (Web Ontology Language) розширює RDFS (Resource Description Framework) з можливістю висловити додаткові обмеження. Твердження формулюються у вигляді «суб'єкт – предикат – об'єкт», як і у RDFS. Але у OWL з'явилась можливість описати нові типи відношень, що у RDFS описати було неможливо. Та головною особливістю OWL є створення нових класів з існуючих, що дозволяє виразити більш складні відношення між об'єктами та суб'єктами бази знань [1].

Перші алгоритми виведення нових фактів у OWL оперували численням предикатів: класи виступають у ролі унарних предикатів, властивості – бінарних, а відношення – це логічні формули над цими предикатами. Але виявилось, що неможливо знайти єдиного правила, яке б вирішувало загальну задачу для багатьох наборів даних. Тому звернули увагу на методи описової логіки, яка в залежності від вхідних наборів даних давала бажаний результат. Описова мова складається з двох частин - термінологічних аксіом (відповідностей між ролями і поняттями) і тверджень. Разом вони складають конструктор описової мови. Зазвичай у описовій мові є три компонента:

1. Словник, що містить множини концептів, атомарних ролей та
2. Множина конструкторів, що використовуються для побудови індивідуальних об'єктів складних концепцій і ролей з атомарних понять і ролей.
3. Мова аксіом, які можуть бути вказані для обмеження словникового запасу, щоб вказати обмеження пошукового інтервалу.

Отже, щоб формувати алгоритми формування логічних висновків у OWL, використання засад описової логіки є необхідністю для автоматичного отримання відношень для заданих комбінацій конструкторів OWL (напр. алгоритми доказів на основі "семантичного табло"). У найгіршому випадку, алгоритм пошуку фактів може мати експоненційну складність у часі в зв'язку з використанням повного набору відношень (повне заперечення, кон'юнкція, функціональні відношення і т.д.) [2]. Існуючі описові мови (FaCT, RACER, Pellet) мають прийнятну швидкодію, адже на практиці онтології зазвичай не являють собою надскладні випадки, які б могли досягти теоретично найвищої складності. Втім, задача оптимізації відомих алгоримтів виведення може бути поставлена і з урахуванням специфіки даних і їх взаємозв'язків з певної предментої області (соціальні мережі, бібліотеки програмних компонентів тощо). Дослідження способів таких алгоритмічних удосконалень є нині досить актуальним, оскільки алгоритми виведення логічних висновків є одним з ключових елементів прикладних семантичних систем.

Література. **1.** Hayes, P., Horrocks, I., Patel-Schneider, P.F.: OWL Web Ontology Language Semantics and Abstract Syntax. W3C Recommendation 10 February 2004 (2004) **2.** Schlobach, S., Cornet, R.: Non-standard reasoning services for the debugging of description logic terminologies. In Gottlob, G., Walsh, T., eds.: IJCAI, Morgan Kaufmann (2003) 355– 362

Соколовський Я.І., Левкович М.В., Яркун В.І., Ликтей Ю.Т., Стельмах Р.П.
Національний лісотехнічний університет України, Львів, Україна

Чисельне моделювання двовимірного тепломасоперенесення на підставі похідних дробового порядку

Синтезовані двовимірні математичні моделі неізотермічного вологоперенесення у середовищах з фрактальною структурою з врахуванням ефектів пам'яті та просторової кореляції. Побудовані явні та неявні різницеві схеми для зв'язаних рівнянь тепломасоперенесення у двовимірній області з граничними умовами третього роду. Проведена алгоритмічна реалізація різницевих рівнянь та отримано умови стійкості різницевих схем.

Математична модель процесу тепло- та вологоперенесення у середовищах з фрактальною структурою описується системою диференціальних рівнянь у частинних похідних із дробовим порядком за часом τ та просторовими координатами x_1 та x_2

$$c\rho\frac{\partial^{\alpha}T(\tau,x_1,x_2)}{\partial\tau^{\alpha}}=\lambda_1\frac{\partial^{\beta}T(\tau,x_1,x_2)}{\partial x_1^{\beta}}+\lambda_2\frac{\partial^{\beta}T(\tau,x_1,x_2)}{\partial x_2^{\beta}}+\epsilon\rho_0 r\frac{\partial^{\alpha}U(\tau,x_1,x_2)}{\partial\tau^{\alpha}},\tag{1}$$

$$\frac{\partial^{\alpha}U(\tau,x_1,x_2)}{\partial\tau^{\alpha}}=a_1\frac{\partial^{\beta}U(\tau,x_1,x_2)}{\partial x_1^{\beta}}+a_2\frac{\partial^{\beta}U(\tau,x_1,x_2)}{\partial x_2^{\beta}}+a_1\delta\frac{\partial^{\beta}T(\tau,x_1,x_2)}{\partial x_1^{\beta}}+a_2\delta\frac{\partial^{\beta}T(\tau,x_1,x_2)}{\partial x_2^{\beta}},\tag{2}$$

з початковими умовами

$$T\mid_{\tau=0}=T_0(x_1,x_2),U\mid_{\tau=0}=U_0(x_1,x_2),\tag{3}$$

та граничними умовами третього роду

$$\lambda_i\frac{\partial^{\gamma}T}{\partial x_i^{\gamma}}\mid_{x_i=l_i}+\rho_0(1-\epsilon)\beta_i(U\mid_{x_i=l_i}-U_{pi})=\alpha_i(T\mid_{x_i=l_i}-t_{ci}),\tag{4}$$

$$a_i\delta\frac{\partial^{\gamma}T}{\partial x_i^{\gamma}}\mid_{x_i=l_i}+a_i\frac{\partial^{\gamma}U}{\partial x_i^{\gamma}}\mid_{x_i=l_i}=\beta_i(U_{pi}-U\mid_{x_i=l_i}),\tag{5}$$

$$\lambda_i\frac{\partial^{\gamma}T}{\partial x_i^{\gamma}}\mid_{x_i=0}+\rho_0(1-\epsilon)\beta_i^{*}\left(U\mid_{x_i=0}-U_{pi}^{*}\right)=\alpha_i^{*}(T\mid_{x_i=0}-t_{ci}^{*}),\tag{6}$$

$$a_i\delta\frac{\partial^{\gamma}T}{\partial x_i^{\gamma}}\mid_{x_i=0}+a_i\frac{\partial^{\gamma}U}{\partial x_i^{\gamma}}\mid_{x_i=0}=\beta_i^{*}\left(U_{pi}^{*}-U\mid_{x_i=0}\right),\tag{7}$$

де $(\tau,x_1,x_2)\in D, D=[0,t]\times[0,l_1]\times[0,l_2]$; T,U - шукані функції, де T- температура, U - вологовміст, c - питома теплоємність, ρ - густина, ρ_0 - базисна густина, ϵ - коефіцієнт фазового переходу, r - питома теплота пароутворення, $\lambda_i\,(i=1,2)$ - коефіцієнти теплопровідності, $a_i\,(i=1,2)$ - коефіцієнти вологопровідності, δ - термоградієнтний коефіцієнт, $t_{ci},t_{ci}^{*}\,(i=1,2)$ - значення температури середовища, $U_{pi},U_{pi}^{*}\,(i=1,2)$ - значення відносної вологості зовнішнього середовища, $\alpha_i,\alpha_i^{*}\,(i=1,2)$ - коефіцієнти теплообміну, $\beta_i,\beta_i^{*}\,(i=1,2)$ - коефіцієнти вологообміну, $l_i\,(i=1,2)$ - геометричні розміри, α - дробовий порядок похідної за часом $(0<\alpha\leqslant 1)$, β,γ - дробові показники похідної за просторовими координатами $(1<\beta\leqslant 2),(0<\gamma\leqslant 1)$.

В області D введемо просторово-часову сітку

$$\omega_{\triangle\tau,h_1,h_2}=\langle\left(\tau^k,x_{1(n)},x_{2(m)}\right):x_{1(n)}=(n-1)h_1,x_{2(m)}=(m-1)h_2,$$
$$\tau^k=k\triangle\tau,n=\overline{1,N},h_1=\frac{l_1}{N-1},m=\overline{1,M},h_2=\frac{l_2}{M-1},k=\overline{0,K},\triangle\tau=\frac{t}{K}\rangle.$$

Використавши формулу Рімана-Ліувілля [1]

$$\frac{\partial^{\alpha}f(\tau)}{\partial\tau^{\alpha}}\mid_{\tau^k}=\frac{1}{\Gamma(1-\alpha)}\left(\frac{f\left(\tau^k\right)}{\left(\tau^{k+1}-\tau^k\right)^{\alpha}}+\int_{\tau^k}^{\tau^{k+1}}\frac{f'(\xi)}{\left(\tau^{k+1}-\xi\right)^{\alpha}}d\xi\right)$$

різницеву апроксимацію дробової похідної $\alpha\,(0<\alpha\leqslant 1)$ на відрізку $\left[\tau^k,\tau^{k+1}\right]$ можна записати

таким чином [4]:

$$\frac{\partial^\alpha u}{\partial \tau^\alpha} \mid_{\tau^k} \approx \frac{u^{k+1} - \alpha u^k}{\Gamma(2-\alpha)\triangle\tau^\alpha}, \triangle\tau = \tau^{k+1} - \tau^k, \tag{8}$$

де $\Gamma(\alpha)$ - гамма-функція.

Для визначення дробової похідної $\beta\,(1 < \beta \leqslant 2)$ за координатами x_1, x_2 скористаємось формулою Грюнвальда-Летнікова [2]

$$\frac{\partial^\beta f(\tau)}{\partial x_i^\beta} = \lim_{h_i \to 0} \frac{1}{h_i^\beta} \sum_{j=0}^{\left[\frac{\tau}{h_i}\right]} (-1)^j \frac{\Gamma(\beta+1)}{\Gamma(j+1)\Gamma(\beta-j+1)} f(x_i - j + 1),$$

де $h_i = x_{i(n+1)} - x_{i(n)}, [x_i]$ - ціла частина x_i, $i = 1, 2$.

Тоді різницева апроксимація дробової похідної β за координатами x_1, x_2 матиме вигляд [3]:

$$\frac{\partial^\beta u}{\partial x_1^\beta} \mid_{x_{1(n)}} \approx \frac{1}{h_1^\beta} \sum_{j=0}^{n} q_j u_{n-j+1}, \frac{\partial^\beta u}{\partial x_2^\beta} \mid_{x_{2(m)}} \approx \frac{1}{h_2^\beta} \sum_{j=0}^{m} q_j u_{m-j+1}, \tag{9}$$

де $q_0 = 1, q_j = (-1)^j \frac{\beta(\beta-1)\ldots(\beta-j+1)}{\Gamma(j+1)}$.

Враховуючи (8), (9) отримаємо неявну різницеву схему для чисельної реалізації системи диференціальних рівнянь (1), (2):

$$c\rho \frac{T_{n,m}^{k+1} - \alpha T_{n,m}^k}{\Gamma(2-\alpha)\triangle\tau^\alpha} = \frac{\lambda_1}{h_1^\beta} \sum_{j=0}^{n} q_j T_{n-j+1,m}^{k+1} + \frac{\lambda_2}{h_2^\beta} \sum_{j=0}^{m} q_j T_{n,m-j+1}^{k+1} + \epsilon\rho_0 r \frac{U_{n,m}^{k+1} - \alpha U_{n,m}^k}{\Gamma(2-\alpha)\triangle\tau^\alpha}, \tag{10}$$

$$\begin{aligned} \frac{U_{n,m}^{k+1} - \alpha U_{n,m}^k}{\Gamma(2-\alpha)\triangle\tau^\alpha} = {} & \frac{a_1}{h_1^\beta} \sum_{j=0}^{n} q_j U_{n-j+1,m}^{k+1} + \frac{a_2}{h_2^\beta} \sum_{j=0}^{m} q_j U_{n,m-j+1}^{k+1} + \frac{a_1\delta}{h_1^\beta} \sum_{j=0}^{n} q_j T_{n-j+1,m}^{k+1} + \\ & + \frac{a_2\delta}{h_2^\beta} \sum_{j=0}^{m} q_j T_{n,m-j+1}^{k+1}. \end{aligned} \tag{11}$$

Явна різницева схема буде така:

$$c\rho \frac{T_{n,m}^{k+1} - \alpha T_{n,m}^k}{\Gamma(2-\alpha)\triangle\tau^\alpha} = \frac{\lambda_1}{h_1^\beta} \sum_{j=0}^{n} q_j T_{n-j+1,m}^{k} + \frac{\lambda_2}{h_2^\beta} \sum_{j=0}^{m} q_j T_{n,m-j+1}^{k} + \epsilon\rho_0 r \frac{U_{n,m}^{k+1} - \alpha U_{n,m}^k}{\Gamma(2-\alpha)\triangle\tau^\alpha}, \tag{12}$$

$$\begin{aligned} \frac{U_{n,m}^{k+1} - \alpha U_{n,m}^k}{\Gamma(2-\alpha)\triangle\tau^\alpha} = {} & \frac{a_1}{h_1^\beta} \sum_{j=0}^{n} q_j U_{n-j+1,m}^{k} + \frac{a_2}{h_2^\beta} \sum_{j=0}^{m} q_j U_{n,m-j+1}^{k} + \frac{a_1\delta}{h_1^\beta} \sum_{j=0}^{n} q_j T_{n-j+1,m}^{k} + \\ & + \frac{a_2\delta}{h_2^\beta} \sum_{j=0}^{m} q_j T_{n,m-j+1}^{k}. \end{aligned} \tag{13}$$

Скінченно-різницевий аналог граничних умов має вигляд:

$$\lambda_i \frac{T_R^{k+1} - \gamma T_{R-1}^{k+1}}{\Gamma(2-\gamma)h_i^\gamma} + \rho_0(1-\epsilon)\beta_i \left(U_R^{k+1} - U_{pi}\right) = \alpha_i \left(T_R^{k+1} - t_{ci}\right), \tag{14}$$

$$a_i\delta \frac{T_R^{k+1} - \gamma T_{R-1}^{k+1}}{\Gamma(2-\gamma)h_i^\gamma} + a_i \frac{U_R^{k+1} - \gamma U_{R-1}^{k+1}}{\Gamma(2-\gamma)h_i^\gamma} = \beta_i \left(U_{pi} - U_R^{k+1}\right), \tag{15}$$

$$\lambda_i \frac{T_2^{k+1} - \gamma T_1^{k+1}}{\Gamma(2-\gamma)h_i^\gamma} + \rho_0(1-\epsilon)\beta_i^* \left(U_1^{k+1} - U_{pi}^*\right) = \alpha_i^* \left(T_1^{k+1} - t_{ci}^*\right), \tag{16}$$

$$a_i\delta \frac{T_2^{k+1} - \gamma T_1^{k+1}}{\Gamma(2-\gamma)h_i^\gamma} + a_i \frac{U_2^{k+1} - \gamma U_1^{k+1}}{\Gamma(2-\gamma)h_i^\gamma} = \beta_i^* \left(U_{pi}^* - U_1^{k+1}\right). \tag{17}$$

де $i = 1, 2$, (якщо $i = 1$, то $R = N$ та якщо $i = 2$, то $R = M$).

Для знаходження чисельних розв'язків отриманої системи різницевих рівнянь використаємо метод предиктор-коректор [5]. Предиктор реалізований за допомогою неявної різницевої схеми

(10), (11), а коректор - явної різницевої схеми (12), (13).

На часовому кроці $k=0$ знайдемо значення функцій $T^0_{n,m}, U^0_{n,m}$ із початкових умов:

$$T^0_{n,m}=T_0\left(x_{1(n)},x_{2(m)}\right), U^0_{n,m}=U_0\left(x_{1(n)},x_{2(m)}\right). \tag{18}$$

На першому півкроці інтервалу $\frac{\triangle\tau}{2}$ запишемо неявну різницеву схему, у якій врахуємо тільки похідну дробового порядку β за координатою x_1:

$$c\rho\frac{T^{k+\frac14}_{n,m}-\alpha T^k_{n,m}}{\Gamma(2-\alpha)\left(\frac{\triangle\tau}{2}\right)^\alpha}=\frac{\lambda_1}{h_1^\beta}\sum_{j=0}^{n}q_jT^{k+\frac14}_{n-j+1,m}+\epsilon\rho_0 r\frac{U^{k+\frac14}_{n,m}-\alpha U^k_{n,m}}{\Gamma(2-\alpha)\left(\frac{\triangle\tau}{2}\right)^\alpha}, \tag{19}$$

$$\frac{U^{k+\frac14}_{n,m}-\alpha U^k_{n,m}}{\Gamma(2-\alpha)\left(\frac{\triangle\tau}{2}\right)^\alpha}=\frac{a_1}{h_1^\beta}\sum_{j=0}^{n}q_jU^{k+\frac14}_{n-j+1,m}+\frac{a_1\delta}{h_1^\beta}\sum_{j=0}^{n}q_jT^{k+\frac14}_{n-j+1,m}. \tag{20}$$

Запишемо у матричному вигляді рівняння (19) та граничні умови (14), (16), що відповідають йому:

$$U_n^{k+\frac14}=AT_n^{k+\frac14}+\alpha U_n^k-\frac{\alpha c\rho}{\epsilon\rho_0 r}T_n^k+\Psi_1, \tag{21}$$

де $U_n^{k+\frac14}=\left[U^{k+\frac14}_{1,m},...,U^{k+\frac14}_{N,m}\right]^T, T_n^{k+\frac14}=\left[T^{k+\frac14}_{1,m},...,T^{k+\frac14}_{N,m}\right]^T, U_n^k=\left[0,U^k_{2,m},...,U^k_{N-1,m},0\right]^T,$

$T_n^k=\left[0,T^k_{2,m},...,T^k_{N-1,m},0\right]^T, \Psi_1=\left[U^*_{p1}-\frac{\alpha_1^*t^*_{c1}}{\rho_0(1-\epsilon)\beta_1^*},0,...,0,U_{p1}-\frac{\alpha_1t_{c1}}{\rho_0(1-\epsilon)\beta_1}\right]^T.$

Компоненти $a_{ij},\left(i,j=\overline{1,N}\right)$ матриці A визначаються за виразами:

$$a_{ij}=\begin{cases}0, j\geqslant i+2,\\ 0, i=N, 1\leqslant j\leqslant N-2,\\ \left(\frac{c\rho}{\epsilon\rho_0 r}-A_1q_1\right), i=j\neq1\neq N,\\ \left(\frac{\alpha_1^*}{\rho_0(1-\epsilon)\beta_1^*}+B_1^*\gamma\right), i=j=1,\end{cases}\qquad \begin{cases}\left(\frac{\alpha_1}{\rho_0\left(1-\epsilon\right)\beta_1}-B_1\right), i=j=N,\\ -B_1^*, i=1, j=2,\\ B_1\gamma, i=N, j=N-1,\\ -A_1q_{i-j+1},\end{cases}$$

в інших випадках.

де $A_1=\frac{\lambda_1\Gamma(2-\alpha)\triangle\tau^\alpha}{\epsilon\rho_0 r2^\alpha h_1^\beta}$, $B_1=\frac{\lambda_1}{\rho_0(1-\epsilon)\beta_1\Gamma(2-\gamma)h_1^\gamma}, B_1^*=\frac{\lambda_1}{\rho_0(1-\epsilon)\beta_1^*\Gamma(2-\gamma)h_1^\gamma}$.

Аналогічним чином можна записати у матричному вигляді рівняння (20) та граничні умови (15), (17), що відповідають йому:

$$BT_n^{k+\frac14}+CU_n^{k+\frac14}+\Psi_2+\alpha U_n^k=0, \tag{22}$$

$C=(c_{ij}), B=(b_{ij}), \left(i,j=\overline{1,N}\right),$

$$c_{ij}=\begin{cases}0, j\geqslant i+2,\\ 0, i=N, 1\leqslant j\leqslant N-2,\\ \left(Z_1q_1-1\right), i=j\neq1\neq N,\\ \left(a_1\gamma-\beta_1^*\Gamma\left(2-\gamma\right)h_1^\gamma\right), i=j=1,\\ -\left(a_1+\beta_1\Gamma\left(2-\gamma\right)h_1^\gamma\right), i=j=N,\\ -a_1, i=1, j=2,\\ a_1\gamma, i=N, j=N-1\\ Z_1q_{i-j+1},\end{cases}\qquad \text{інше.}$$

$$b_{ij}=\begin{cases}0, j\geqslant i+2,\\ 0, i=N, 1\leqslant j\leqslant N-2,\\ Z_1q_1\delta, i=j\neq1\neq N,\\ a_1\delta\gamma, i=j=1; i=N, j=N-1,\\ -a_1\delta, i=j=N; i=1, j=2,\\ Z_1q_{i-j+1}\delta,\end{cases}\qquad \text{інше.}$$

$Z_1=\frac{a_1\Gamma(2-\alpha)\triangle\tau^\alpha}{2^\alpha h_1^\beta}, \Psi_2=\left[\beta_1^*\Gamma\left(2-\gamma\right)h_1^\gamma U^*_{p1},0,...,0,\beta_1\Gamma\left(2-\gamma\right)h_1^\gamma U_{p1}\right]^T.$

Підставляємо (21) в (22) і отримаємо систему рівнянь, яку розв'язуємо відносно функції T:

$$\left(B+CA\right)T_n^{k+\frac14}-\frac{\alpha c\rho}{\epsilon\rho_0 r}CT_n^k+\left(\alpha C+\alpha\right)U_n^k+\Psi_1+\Psi_2=0 \tag{23}$$

Знайшовши із (23) множину розв'язків - $T^{k+\frac14}_{1,m}, T^{k+\frac14}_{2,m},...,T^{k+\frac14}_{N,m},\left(k=\overline{0,K-1}\right)$, шукаємо із

(21) множину розв'язків - $U_{1,m}^{k+\frac{1}{4}}, U_{2,m}^{k+\frac{1}{4}}, ..., U_{N,m}^{k+\frac{1}{4}}, \left(k=\overline{0,K-1}\right)$.

На часовому півкроці інтервалу $\frac{\triangle\tau}{2}$ запишемо неявну різницеву схему, у якій врахуємо тільки похідну дробового порядку β за координатою x_2

$$c\rho\frac{T_{n,m}^{k+\frac{1}{2}}-\alpha T_{n,m}^{k+\frac{1}{4}}}{\Gamma(2-\alpha)\left(\frac{\triangle\tau}{2}\right)^{\alpha}}=\frac{\lambda_2}{h_2^{\beta}}\sum_{j=0}^{m}q_jT_{n,m-j+1}^{k+\frac{1}{2}}+\epsilon\rho_0 r\frac{U_{n,m}^{k+\frac{1}{2}}-\alpha U_{n,m}^{k+\frac{1}{4}}}{\Gamma(2-\alpha)\left(\frac{\triangle\tau}{2}\right)^{\alpha}}, \tag{24}$$

$$\frac{U_{n,m}^{k+\frac{1}{2}}-\alpha U_{n,m}^{k+\frac{1}{4}}}{\Gamma(2-\alpha)\left(\frac{\triangle\tau}{2}\right)^{\alpha}}=\frac{a_2}{h_2^{\beta}}\sum_{j=0}^{m}q_jU_{n,m-j+1}^{k+\frac{1}{2}}+\frac{a_2\delta}{h_2^{\beta}}\sum_{j=0}^{m}q_jT_{n,m-j+1}^{k+\frac{1}{2}}. \tag{25}$$

Аналогічно як і на першому півкроці, на другому півкроці отримаємо дві системи. У результаті їх розв'язання яких отримаємо множину розв'язку відносно функції T - $T_{n,1}^{k+\frac{1}{2}}, T_{n,2}^{k+\frac{1}{2}}, ..., T_{n,M}^{k+\frac{1}{2}}$, $\left(k=\overline{0,K-1}\right)$, та множину розв'язку відносно функції U - $U_{n,1}^{k+\frac{1}{2}}, U_{n,2}^{k+\frac{1}{2}}, ..., U_{n,M}^{k+\frac{1}{2}}, \left(k=\overline{0,K-1}\right)$.

Для знаходження розв'язків на усьому інтервалі $\triangle\tau$ використаємо коректор, який реалізовується на явній різницевій схемі:

$$c\rho\frac{T_{n,m}^{k+1}-\alpha T_{n,m}^{k}}{\Gamma(2-\alpha)\triangle\tau^{\alpha}}=\frac{\lambda_1}{h_1^{\beta}}\sum_{j=0}^{n}q_jT_{n-j+1,m}^{k+\frac{1}{2}}+\frac{\lambda_2}{h_2^{\beta}}\sum_{j=0}^{m}q_jT_{n,m-j+1}^{k+\frac{1}{2}}+\epsilon\rho_0 r\frac{U_{n,m}^{k+1}-\alpha U_{n,m}^{k}}{\Gamma(2-\alpha)\triangle\tau^{\alpha}}, \tag{26}$$

$$\begin{aligned}\frac{U_{n,m}^{k+1}-\alpha U_{n,m}^{k}}{\Gamma(2-\alpha)\triangle\tau^{\alpha}}=&\frac{a_1}{h_1^{\beta}}\sum_{j=0}^{n}q_jU_{n-j+1,m}^{k+\frac{1}{2}}+\frac{a_2}{h_2^{\beta}}\sum_{j=0}^{m}q_jU_{n,m-j+1}^{k+\frac{1}{2}}+\frac{a_1\delta}{h_1^{\beta}}\sum_{j=0}^{n}q_jT_{n-j+1,m}^{k+\frac{1}{2}}+\\&+\frac{a_2\delta}{h_2^{\beta}}\sum_{j=0}^{m}q_jT_{n,m-j+1}^{k+\frac{1}{2}}.\end{aligned} \tag{27}$$

Отже, із (27) знаходимо множину розв'язку - $\left\{U_{n,m}^{k+1}: k=\overline{0,K-1}, n=\overline{1,N}, m=\overline{1,M}\right\}$, а із (26) отримаємо множину розв'язку - $\left\{T_{n,m}^{k+1}: k=\overline{0,K-1}, n=\overline{1,N}, m=\overline{1,M}\right\}$.

Для визначення умов стійкості отриманих різницевих рівнянь зв'язаного тепломасоперенесення було використано метод умовного задання деяких відомих функцій системи [5]. Такий підхід дозволяє звести стійкість системи рівнянь до дослідження стійкості її окремих рівнянь. Оскільки предиктор реалізований на неявній схемі, а вона у свою чергу є абсолютно стійкою, то питання стійкості різницевих рівнянь для чисельного розв'язку задачі (1)-(7) залежить від коректора, що реалізований на явній різницевій схемі. Отже, загальною умовою стійкості для рівнянь(26), (27) буде:

$$\triangle\tau^{\alpha}\left(\frac{C_1}{h_x^{\beta}}+\frac{C_2}{h_y^{\beta}}\right)\leqslant\frac{(\alpha+1)C_3}{(2+\beta)\Gamma(2-\alpha)},$$

де $C_1=\lambda_1,a_1; C_2=\lambda_2,a_2; C_3=(c\rho-\epsilon\rho_0 r),(1+\delta)^{-1}$.

Література. **1.** Podlubny I. Fractional Differential Equations / I. Podlubny, vol. 198 of Mathematics in Science and Engineering, Academic Press, San Diego, Calif, USA, 1999. – 340s. **2.** Васильев В.В. Дробное исчисление и аппроксимационные методи в моделировании динамических систем. Научное издание / В.В. Васильев, Л.А. Симак. – Киев, НАН Украины, 2008. – 256с. **3.** Соколовський Я. І. Числовий метод дослідження теплообміну на підставі похідних дробового порядку / Соколовський Я. І., Москвітіна М. В. // журнал «Технічні вісті», 2015/1(41), 2(42). – С. 12-15. **4.** Соколовський Я. І. Математична модель теплового перенесення та напружено-деформівного стану у капілярно-пористих матеріалах з фрактальною структурою / Соколовський Я. І., Шиманський В. М. // Фізико-математичне моделювання та інформаційні технології, 2012. – Вип.16. – С.133-141. **5.** Никитенко Н. И. Теория тепломассопереноса / Н. И. Никитенко. - Киев: Наук. думка, 1983. – 352с.

Соловйов В.М., Чабаненко Д.М.
Черкаський національний університет імені Богдана Хмельницького, Черкаси, Украина

Мережні міри складності у передбаченні кризових явищ

У роботі обговорюються основні методологічні і методичні особливості нового мережного підходу до побудови індикаторів кризових явищ. На сучасному етапі розвитку складних систем домінує так звана мережна парадигма складності, в основу якої покладена ключова ідея щодо можливості представлення будь-якої складної системи у вигляді мережі (і відповідного їй графа). Подальший аналіз зміни з часом віконних мережних мір складності – графодинаміка – дає можливість встановити універсальні топологічні і спектральні властивості складних систем в різні періоди функціонування, зокрема, і в періоди криз на фінансових ринках.

В останні 10-15 років все більшою популярністю міждисциплінарний напрям дослідження складних систем, який отримав назву теорії складних мереж (complex networks) [1]. Він вивчає характеристики мереж, враховуючи не тільки їх топологію, але й статистичні властивості, розподіл ваг окремих вузлів і ребер, ефекти розповсюдження інформації, стійкість (robustness) і т.п. До складних мереж відносяться електричні, транспортні, інформаційні, соціальні, економічні, біологічні, нейронні та інші мережі.

Мережна парадигма стала домінуючою при дослідженні складних систем оскільки дозволяє ввести не існуючі для часового ряду нові кількісні міри складності.

Раніше нами було введено різні кількісні міри складності для окремих часових рядів, зокрема: алгоритмічні, фрактальні, хаос-динамічні, рекурентні, неекстенсивні, нереверсивні та ін. [2].

Суттєвою перевагою введених мір є їх динамічність, тобто можливість відстежувати у часі зміну обраної міри та порівнювати з відповідною динамікою вихідного часового ряду. Це дозволило нам співставити критичні зміни динаміки системи, що описується часовим рядом, з характерними змінами конкретних мір складності [3]. Виявилось, що кількісні міри складності реагують на критичні зміни в динаміці складної системи, що дозволяє використовувати їх в процесі діагностики та прогнозування майбутніх змін.

У даній роботі ми введемо мережні міри складності і адаптуємо їх з метою дослідження системної динаміки. Найбільш вживаними методами перетворення часових послідовностей у відповідні мережі є рекурентні [4], графи видимості [5] та кореляційні [6].

Технологія рекурентних діаграм для візуалізації рекурентностей у фазовому просторі заснована на ідеї Анрі Пуанкаре щодо рекурентності фазового простору динамічних систем.

Згідно з теоремою Такенса, еквівалентна фазова траєкторія, що зберігає структуру оригінальної фазової траєкторії, може бути відновлена з одного спостереження або часового ряду методом часових затримок: $\hat{x}(t) = \left(u_i, u_{i+\tau}, \ldots, u_{i+(m-1)\tau}\right)$, де m – розмірність вкладення, τ – часова затримка (реальна часова затримка визначається як $\tau \cdot \Delta$). Рекурентна ж діаграма відображає наявні повторюваності у формі бінарної матриці R, де $R_{i,j} = 1$, якщо x_i є сусіднім до стану x_j, які потрапляють в m-вимірну околицю радіусом ϵ і центром x_i. Зрозуміло, що параметри m, τ та ϵ є ключовими при проведенні рекурентного аналізу. Рекурентна діаграма легко трансформується у матрицю суміжності, за якою розраховуються спектральні і топологічні характеристики графа [3].

Алгоритм графа видимості реалізується у наступний алгоритм. Візьмемо часовий ряд $Y(t) = [y_1, y_2, \ldots, y_n]$ довжини N. Кожну точку даних часового ряду можна розглядати як вершину в асоційованій мережі, а ребро буде з'єднувати дві вершини, якщо дві відповідні точки даних можуть "бачити" один одного з відповідної точки часового ряду. Формально два значення ряду y_a (на момент часу t_a) і y_b (на момент часу t_b) пов'язані, якщо для будь-якого іншого значення (y_c, t_c), яке розміщене між ними (тобто, $t_a \leqslant t_c \leqslant t_b$), задовольняється умова: $y_c < y_a + (y_b - y_a)\,\frac{t_c - t_a}{t_b - t_a}$.

Для побудови і аналізу властивостей кореляційного графа слід сформувати з кореляційної матриці матрицю суміжності. Для цього треба ввести величину, яка для поля кореляцій буде

слугувати відстанню між корелюваними агентами. Такою відстанню може слугувати залежна від коефіцієнта взаємної кореляції C_{ij} величина $x(i,j) = \sqrt{2(1-C_{ij})}$.

Так, якщо коефіцієнт кореляції між двома активами помітний, відстань між ними є малою, і, починаючи з деякої критичної величини x_{cr}.

Для матриці суміжності це означає, що вони є суміжними на графі. В противному випадку активи не є суміжними. При цьому обов'язковою умовою є зв'язність графа.

Розрахунки проводились у такий спосіб. Обирався часовий проміжок (вікно), наприклад, два роки (приблизно 500 торгівельних днів), для нього будувались відповідні графи та розраховувались його спектральні і топологічні властивості. До спектральних, зокрема, відносяться максимальне власне значення матриці суміжності, спектральний розрив, енергія графа, а до топологічних – ступінь вершини, щільність зв'язків у графі, найкоротший шлях, коефіцієнт кластеризації тощо. Далі вікно зміщувалось з кроком, наприклад, один тиждень (5 торгівельних днів) і процедура повторювалась до вичерпання часових рядів.

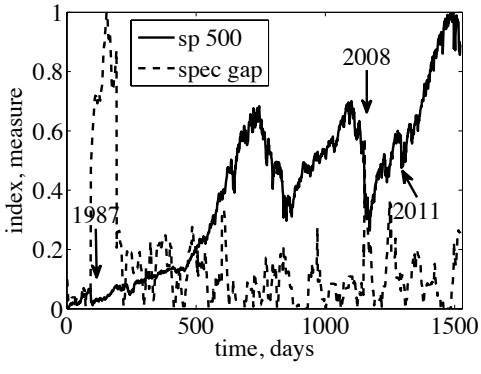


Рис. 1

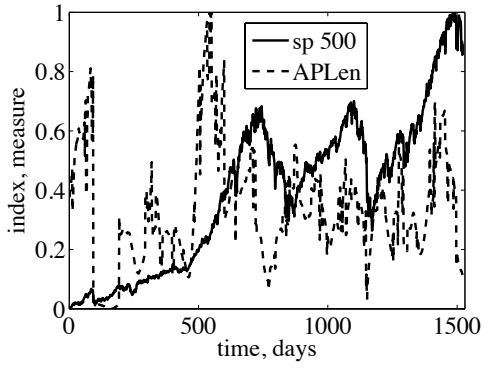


Рис. 2

Динаміка індексу sp 500 і спектральної міри складності: спектральний розрив (spec gap) (рис. 1) та топологічної: середня довжина шляху (APLen) (рис. 2) для рекурентної мережі, отриманої для індексу sp 500 за період з 1984 по 2016рр.

На прикладах відомих фінансових криз показано, що деякі з мережних мір можуть слугувати індикаторами-передвісниками кризових явищ.

Література. **1.** Halvin S., Cohen R. Complex networks. Structure, robustness and function / Halvin S., Cohen R. // Cambridge University Press, 2010. – 238 p. **2.** Соловйов В.М., Чабаненко Д.М. Методи кількісної оцінки складності. Системний аналіз таінформаційні технології: матеріали 15-ї Міжнародної науково-технічної конференції SAIT 2013, Київ, 27-31 травня 2013 р. / ННК "ІПСА" НТУУ "КПІ". – К.: ННК "ІПСА" НТУУ "КПІ", 2013. – С.192-193. **3.** Цветкович Д. Спектры графов. Теория и применение / Цветкович Д., Дуб М., Захс Х. - К.: Наукова думка, 1984. – 384 с. **4.** 4. Donner R.V. Recurrence-based time series analysis by means of complex network methods / R.V. Donner, M. Small, J.F. Donges, N. Marwan et.al. // [Електронний ресурс] – Режим доступу: arXiv:1010.6032v1 [nlin.CD] 25 Oct 2010. **5.** Lacasa L. From time series to complex networks: The visibility graph / L. Lacasa, B. Luque, F. Ballesteros et.al. // PNAS. – 2008. – V. 105, No 13. – P. 4972-4975. **6.** Лук'янчук О.С. Фолксономія соціально-економічних об'єктів в складних мережах засобами CorrRank / О.С. Лук'янчук, В.М. Соловйов //Моделювання та інформаційні технології в економіці: монографія / За заг. ред. Соловйова В.М. – Черкаси: Брама-Україна, 2014. – с. 140-151.

Стасюк А.В.
Факультет прикладної математики НТУУ "КПІ", Київ, Україна

Логічна парадигма подання даних

За декілька десятиліть розвитку комп'ютерних наук було запропоновано багато способів подання даних для їх тривалого зберігання та обробки. На сьогодні домінуючою технологією вирішення даної задачі є реляційні бази даних, що зберігають інформацію у вигляді багатьох зв'язаних таблиць. Таке рішення має значні недоліки, хоча роки наукових досліджень реляційних баз даних зробили існуючі реалізації досить ефективними [1].

Запропонована логічна парадигма опирається на філософські ідеї, які дещо відрізняються від описаних Аристотелем (саме аристотелівське бачення світу є основою реляційних баз даних). Тому розглянемо далі поняття, якими оперує логічна парадигма.

Основною категорією логічної парадигми, є *об'єкт*. Під цим поняттям розуміють довільну підмножину чотиривимірного часу-простору. В аналітичній практиці достатньо розуміння того, що об'єкт є абстрактним носієм унікальності (тому в контексті технічної реалізації для існування об'єкта досить існування унікального ідентифікатора цього об'єкта). Виділення об'єктів із реального світу диктується *виключно конкретною предметною областю*. Логічна парадигма також визначає три особливі типи об'єктів - *класи*, *кортежі* та *події* [2].

Класами в логічній парадигмі називають множини об'єктів (в тому числі у ролі елементів можуть виступати інші класи, кортежі або події), що мають деяку спільну ознаку. Прицьому така ознака може бути обрана довільно, залежно від потреб даної прикладної задачі. Наприклад, замість того, щоб сказати, що автомобіль має атрибут *колір* із значенням *червоний* (як це прийнято в реляційній парадигмі), можна просто включити об'єкт автомобіля у *клас* червоних речей. Оскільки класи є математичними множинами, у логічній парадигмі допускаються множинна класифікація (включення до багатьох класів) та перекласифікація [2].

Логічна парадигма не змінює загально прийняте значення поняття *кортеж*. Тобто кортежем є впорядкована послідовність об'єктів, що має фіксовану довжину. Однак так само, як і класи, кортежі в логічній парадигмі є об'єктами і повинні мати здатність відрізнятися від інших об'єктів в рамках даної системи. Оскільки елементами кортежів також є об'єкти, то кортежі можуть включати класи, події та інші кортежі, при цьому в будь-яких комбінаціях.

Подією в логічній парадигмі є об'єкт, що представляє часову миттєвість (або ж часовий діапазон, шириною якого в даному випадку можна знехтувати). З філософського погляду, подією є тривимірний зріз *усього* чотиривимірного часу-простору в деякий конкретний момент. Окрім цього, в практичних цілях виділяють три спеціальних події - *нескінченно віддалене минуле*, *поточний момент часу* та *нескінченно віддалене майбутнє*. Ці спеціальні події дозволяють моделювати, наприклад, напіввідкриті часові інтервали тощо [2].

Розглянемо коротко основні пропозиції щодо моделювання в рамках логічної парадигми. Сутності (які в реляційних базах даних є таблицями) найпростіше моделювати класами. Відношення між сутностями можна подавати як кортежі, що включаються у відповідний клас (наприклад, можна визначити клас кортежів для відношення *батько-дитина* або *студент-факультет* і включати у них кортежі, які пов'язують конкретні об'єкти). Спеціалізація (наслідування) просто є одним із типів звичайного відношення з додатковим обмеженням, що учасниками відношення можуть бути тільки класи. Моделювати атрибути можна 2 способами: або на основі відношення *об'єкт має атрибут*, або як включення відповідного об'єкта в клас (наприклад, в клас об'єктів, атрибут *прізвище* яких має значення *Петренко*).

Наступним етапом дослідження логічної парадигми буде відображення її моделі, на реляційну парадигму (з метою уможливлення використання реляційних баз даних).

Література. **1.** Begg, C. E. (2005). "Database Systems. A Practical Approach to Design, Implementation, and Management Addison-Wesley, pp. 1300 **2.** Patridge C. (2000). "Business Objects: Re-Engineering for Re-Use CRC Press LLC, pp. 554

Степаненко А.І., Чкалов О.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Алгоритми планування дій гравців у віртуальному футболі роботів

Вступ. На даний момент задача групового керування роботами розглядається як одна з найактуальніших проблем робототехніки [1]. Оскільки розробка узагальнених алгоритмів і принципів групового керування поки перебуває на початковій стадії, науковою спільнотою було сформульовано ряд модельних задач, на яких рекомендується відпрацьовувати алгоритми. Однією з основних модельних задач є так званий віртуальний футбол роботів.

В даній роботі розглядаються не детерміновані алгоритми гри роботів-футболістів, що базуються на плануванні дій у часі. Відомо, що не детерміновані алгоритми краще вирішують задачі управління в неповній мірі відомих на етапі проектування умов, а також в умовах динамічної протидії. Тому в деяких випадках детерміновані алгоритми можуть бути витіснені, наприклад, нейронними системами [1, 2].

В задачах віртуального футболу найкращі результати було досягнуто алгоритмами задач пошуку шляху у просторі та на графах. Різні техніки планування дій активно використовуються в алгоритмах RoboCup Simulation Leage. Але планування дій як правило виконується за допомогою перебору на невелику глибину, або ж за допомогою уточнень знайденої послідовності дій ітеративним методом. Ці алгоритми потребують попередньої побудови карт пошуку для ефективної реалізації. Питання побудови подібних карт добре вивчені для випадків вимірів пошуку, які моделюються полігональними поверхнями в тривимірному просторі. Для вирішення подібних задач існують як окремі комерційні системи, так и некомерційні бібліотеки. Однак через велику глибину та ширину перебору на кожному кроці використання методів планування дій на велику глибину потребує створення дещо інших алгоритмів планування.

Для того, щоб представити задачу віртуального футболу як задачу пошуку шляху в дискретному просторі станів, слід ввести дискретизацію для можливих положень футболістів і ввести квантизацію часу, розбивши його на послідовні такти. Для того, щоб послідовність дій, що буде отримана в ході пошуку, виявилась ефективною, крок дискретизації як часу, так і просторових координат повинен бути достатньо малим. Тому довжина шуканої послідовності виявляється досить великою (ігрова ситуація суттєво змінюється за час порядку декількох сотень кроків), а простір станів об'єкту керування має великий коефіцієнт розгалуження (порядку 30 для одного футболіста і порядку $2 \cdot 10^7$ для команди з 5 гравців. З цієї причини пряме використання алгоритмів планування, що широко використовуються в антагоністичних і неантагоністичних іграх, веде до величезних обчислювальних витрат, що практично унеможливлює можливість побудови керування в реальному часі. В роботі розглядаються різні методи планування дій, створені з урахуванням вказаної особливості, зокрема, ієрархічні методи пошуку шляхів та метод прорахунку стратегій в глибину.

Ієрархічна реалізація пошуку оптимального шляху робота-футболіста. Опишемо коротко запропоновану в роботі ієрархічну реалізацію пошуку шляху для одного футболіста. Метою пошуку є переведення робота за мінімальний час з поточної позиції до цільової. На відміну від детермінованих алгоритмів, керування будується не аналітично, а в ході процедури перебору. Це, зокрема дає можливість врахувати наявність стінок ігрового поля, штанг, можливих зіткнень з іншими гравцями.

На початковому етапі створюється високорівневе представлення простору станів робота-футболіста. Воно представляється як орієнтований граф, вершини якого відповідають певному розбиттю множини можливих позицій робота-футболіста на підмножини-кластери, що не перетинаються. Кожен такий кластер складається з близьких позицій. Ця задача вирішується описаним в роботі алгоритмом «розфарбовування в три проходи». Ця процедура виконується однократно до початку власне гри.

Пошук шляху складається з двох етапів. На першому за допомогою відомого алгоритму A^* відшукується шлях на високорівневому представленні простору станів. На другому за

допомогою модифікованого алгоритму Дийкстри, що використовує фіксовану пам'ять, ведеться безпосередній пошук шляху. При цьому вважається, що якщо шлях на високорівневому представленні не знайдено, то його не існує і в базовому просторі позицій. Якщо ж високорівневий шлях знайдено як послідовність кластерів, подальша його оптимізація виконується лише локально в межах кожного кластера та його найближчих сусідів.

Асимптотичні оцінки обчислювальної складності такого пошуку показують переваги даного підходу перед не ієрархічним. Практичне використання алгоритму в програмі віртуального футболу підтверджують такі оцінки.

Метод прорахунку в глибину для групового керування. Вказаний алгоритм успішно вирішує задачу тактичного керування одним футболістом, але не вирішує задачі групового керування. Для вирішення цієї задачі пропонується метод прорахунку стратегій в глибину. Його застосування передбачає, що існує деякий набір стратегій управління S, кожна з яких є оптимальною для частини можливих початкових ситуацій, наприклад, додаткових обмежень чи припущень, що зроблені для спрощення. Як такий набір можуть бути використані, наприклад, прості стратегії управління без перебору, або це може бути набір стратегій, що їх в аналогічних ситуаціях обирала людина. Для задачі керування командою роботів такий набір будується на основі детермінованих алгоритмів.

Метод прорахунку в глибину полягає в тому, що замість задачі планування у вихідному просторі станів команди розглядається пошук на орієнтованому графі станів, що визначається на наступною рекурсивною процедурою. Для кожного стану A, що включений до графу, будується набір станів, що виникають внаслідок застосування кожної стратегії з набору S протягом певної кількості кроків, що може бути як фіксованою, так і змінюватись в залежності від стратегії, що обрана. Отримані в такий спосіб вершини також включаються до графу і процедура повторюється рекурсивно до певної глибини. Далі серед термінальних для даної процедури вершин знаходиться та, що забезпечує найкращу оцінку, і за шукану послідовність дій приймається та, що веде до цього стану.

В тому випадку, коли число s виявляється більшим за довжину оптимального шляху, прорахунок в глибину вироджується в простий перебір стратегій з набору S. Шлях (складена стратегія), отримана за допомогою прорахунку в глибину, може бути знайдений суттєво швидше в порівнянні з іншими відомими алгоритмами пошуку, в тому числі й інформованого (таких як A^*). Змінюючи число кроків s і набір стратегій S, можна встановлювати баланс між точністю знайденого розв'язку (його близкістю до оптимального) та обчислювальною складністю процедури пошуку.

На практиці використання методу прорахунку в глибину не вимагає побудови графу в явному вигляді та збереження його в пам'яті. Цікавою особливістю процедури стратегічного пошуку в глибину є те, що вона дозволяє прозорим чином поєднувати в одній програмі гетерогенні (тобто створені на основі різних принципів) стратегії із збереженням сильних сторін кожної з них.

Література. **1.** Д. Е. Охоцимский, В. Е. Павловский, А. Г. Плахов, А. Н. Туганов. Система моделирования игры роботов-футболистов. // Тр. Конференции «Мобильные роботы и мехатронные системы», МГУ, 5-6 декабря 2000 - Под ред. д.ф.-м.н. А. М. Формальского и к.ф.-м.н. В. М. Буданова. М.: Издательство МГУ, 2001, с.192 - 203. **2.** А. В. Малолетов, Р. В. Лещенко, К. Ю. Лепетухин, Е. В. Ветошкина «Моделирование стратегии управления группой роботехнических объектов» Искусственный интеллект, N 4, 2002, с. 574-579

Стефанишин Д.В.
Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна

Про ситуаційно-індуктивне моделювання в задачах екстраполяційного прогнозування за даними моніторингу

Моніторинг передбачає цілеспрямовані і систематичні спостереження за визначеними компонентами або параметрами об'єкта досліджень, спеціальним чином організовані в просторі і в часі, а також комплекс методів обробки накопичених даних та прогнозування поведінки досліджуваної системи, процесу, явища за отриманими даними спостережень на основі математичних моделей з метою прийняття відповідних рішень [1]. На разі в більшості випадків прогнозування за даними моніторингу здійснюється у формі екстраполяцій — з виявленням усталених тенденцій в розвитку системи та їх перенесенням на майбутнє [2]. Серед основних математичних моделей, що використовуються при цьому, виділяються різного роду статистично-імовірнісні моделі: тренди, регресії, функції розподілу ймовірності [2,3]. Такий підхід до прогнозування поведінки динамічних систем в умовах усталених, еволюційних, режимів розвитку, за обмежень, що не допускають різких змін в поведінці системи, біфуркацій й катастрофічних сценаріїв, може вважатися цілком виправданим.

Однак традиційні методи побудови математичних моделей за даними спостережень, що ґрунтуються на принципі оптимізації, не завжди відповідають граничним обмеженням, які накладаються на модель. Практика показує, що зі збільшенням кількості накопичених даних спостережень проблеми, пов'язані з рішенням оптимізаційної задачі можуть виникати навіть у випадках використання простих моделей-екстраполяцій. При цьому ускладнення екстраполяційної моделі за рахунок врахування додаткових факторів та параметрів, нелінійних ефектів тощо можуть погіршувати її якість як прогностичної моделі.

В [4] нами було запропоновано підхід до прогнозування за даними моніторингу, що ґрунтується на ідеї ситуаційно-індуктивного моделювання. Згідно з цим підходом підставою для побудови більш складних індуктивних моделей є результати більш простого ситуаційного моделювання на окремих часових інтервалах функціонування системи, які характеризуються однорідним прогнозним фоном і обмежуються монотонними або квазістаціонарними динамічними рядами даних. При цьому за допомогою ситуаційних моделей, в межах обмежених інтервалів часу, де такі моделі можуть вважатися адекватними, здійснюється оперативне (короткотермінове) прогнозування. Строкове (довготермінове) прогнозування здійснюється на основі індуктивних моделей, за допомогою яких відслідковується еволюція ситуаційних моделей як фазових портретів минулих станів динамічної системи та відтворюються ситуаційні моделі її станів для майбутніх періодів.

Розглянуто практичні приклади екстраполяційного прогнозування з використанням запропонованих ситуативно-індуктивних моделей. Одна з задач стосується прогнозування ризиків збитків від руйнівних повеней в Закарпатті в басейні р. Тиса за рядами динаміки збитків та максимальних витрат води р. Тиса, що визначалися на гідрометричному посту Вилок. Інші приклади мають відношення до прогнозування значень діагностичних параметрів напірних гідротехнічних споруд за даними регулярних інструментальних спостережень.

Література. **1.** Израэль Ю. А. Экология и контроль состояния природной среды / Израэль Ю. А. — Л.: Гидрометеоиздат, 1979. — 376 с. **2.** Єріна А. М. Статистичне моделювання та прогнозування / А. М. Єріна. — К.: КНЕУ, 2001. — 170 с. **3.** Стефанишина-Гаврилюк Ю. Д. Використання нечіткої міри для подолання невизначеності довгострокових прогнозів на основі екстраполяцій / Ю. Д. Стефанишина-Гаврилюк, Д. В. Стефанишин // Системні дослідження та інформаційні технології. — 2013. — № 4. — С. 99-110. **4.** Stefanyshyn D.V. A Method of Forecasting of Indexes of Dynamic System that evolves slowly, based on Time Series Analysis / D.V. Stefanyshyn // ICIM 2013. Proc. of 4th Int. Conf. on Inductive Modelling. Kyiv, Ukraine, September 16-20, 2013. — P.P. 221-224.

Стоянова А.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Анализ преступности в Украине с использованием регрессионных моделей в SAS Enterprise System

Прогнозирование на основе моделей, построенных по экспериментальным (статистическим) данным - один из самых популярных подходов к прогнозированию динамики процессов в социально- экономических, финансовых, технических и других системах, коротко и среднесрочного прогнозирования [1].

Преступность — исторически переходное, переменное, социально-правовое явление, которое являет собой совокупность всех содеенных в государстве преступлений за некоторый период времени. Общие закономерности развития общества распространяются и на преступность как социальное явление. Вместе с этим преступности присущий элемент стихийности и стойкости. Сезонные колебания свойственны абсолютному большинству юридически значимых явлений. Некоторые сезонные волны имеют различные сдвиги по фазе. В автономных системах деятельности это не имеет особого значения, а в зависимых системах различия в сезонных колебаниях могут иметь отрицательные последствия, это имеет прямое отношение к системе «преступность - борьба с преступностью[2]. Проблемы связанные с искоренением преступности усиленно расматривается правовыми органами всего мира. На данном этапе в мире существует множество баз данных, каждая их которых имеет специфические возможности. Главным заданием которых является то, что со всех разнообразных данных отобрать такие фрагменты, которые складывались бы в согласованную целую картину реальных событий. Очень часто на составлении подобной головоломки могут уходить года аналитической работы. Поэтому главной проблемой является оптимизация программ, которая ускорит данный процесс. Общество стремится к такому состоянию общественного развития, когда такие общественно опасные деяния, которые определяются как преступление, членами общества не будут совершаться.

Постановка задачи. Предоставлена МВД информация, которая является миллионом записей совершенных преступлений по отношению к жителям городов. Эта статистическая информация была представлена с использованием ETL (Extract, Transform, Load) процедур[3], на основе языка SAS Base, более 50000 временных рядов на ежемесячной основе, для всех регионов страны (25 регионов) и типов преступности (учитывая все категории и подкатегории), согласно с Уголовно-процессуальным Кодексом. Каждый временной ряд должен быть проанализирован на наличие тренда, сезонности и аномалий и на основе этого построить модель прогноза на основе информации о регрессорах.

Решение. Для решения поставленной задачи используем SAS Enterprise Guide 6.1 для того, чтобы использовать автоматическую обработку временных рядов. SAS Text Analytics инструментарий для определения существующих трендов в сети Интернет. SAS Visual Analytics для построения отчетности, с возможностью отображения на мобильных устройствах. SAS Enterprise Guide 6.1 является мощным инструментом для создания математических моделей на основе временных рядов, такие регрессионные модели: нелинейной регрессии, экспоненциального сглаживания, авторегрессии, авторегрессии с трендом и скользящим средним.

Выводы. SAS аналитические инструменты позволяют создать модели автоматически или с помощью экспертов [4]. Далее на базе технологии SAS планируется реализовать систему автоматически быстрого предсказания(прогнозирования) количество преступлений по всем типам и подтипам.

Литература. **1.** Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів (навчальний посібник) - Київ: Політехніка, 2010. - 317 с. **2.** Лунеев В.В. Юридическая статистика (Учебник.) — М.: Юристъ, 1999. - 400с. - ISBN 5-7975-0156-2 **3.** Терентьев А.Н., Домрачев В.Н., Костецкий Р.И. SAS BASE: Основы программирования (научное издание). – К: Эдельвейс, 2014. – 304 с. – ISBN 978-966- 2748-49-9. **4.** Lora D. Delwiche, Susan J. Slaughter The Little SAS Book. - SAS Institute Inc., Cary, NC, USA, 2008. - 358 с. - ISBN 978-1-59994-725-9

Супруненко О.О., Лісун А.Г.
Черкаський національний університет імені Богдана Хмельницького, Черкаси, Україна

Засоби імітаційного моделювання детермінованого та недетермінованого вибору в алгоритмічних конструкціях

При моделюванні процесів обробки даних у програмному додатку виникають випадки, коли треба проводити вибір між альтернативними гілками, причому такий вибір може бути як детермінованим так і недетермінованим [1]. При детермінованому виборі результат залежить від вихідних параметрів виконання деяких попередніх алгоритмів, при недетермінованому виборі перевага може надаватися будь якій альтернативі та вибирається переважно за певним довільним критерієм. При проектуванні процесів, що мають функціонувати у додатку, виникають випадки, коли детермінований вибір замінюють недетермінованим, що може перетворитися на приховану помилку. Розв'язання задачі виявлення вищевказаного типу помилок проілюструємо на прикладі моделювання алгоритмів засобами безпечних мереж Петрі, які дозволяють здійснювати опис детермінованого і недетермінованого вибору в алгоритмах [2] та ідентифікувати обидва варіанти реалізації.

Розглянемо варіант, який проілюструє детермінований та недетермінований вибір на прикладі моделі інваріанту. В даній роботі інваріант (invariant) розглядається, як деяка логічна умова (істинна чи хибна), яка має виконуватись завжди. В об'єктній моделі це поняття застосовується до об'єктів і означає умову, якій повинен відповідати стан об'єкту.

Наприклад, є клас Employee, який представляє працівника. Він містить поле *age* – вік працівника, а також два методи: метод *setAge*(*integerage*) за допомогою якого можна задати вік працівника та метод *getAge*() за допомогою якого можна отримати його вік. Припустимо, що головною умовою роботи працівника в деякій компанії є його вік. Він має бути від 18 до 65 років. Отже ця умова і буде інваріантом класу *Employee*, який необхідно зберігати.

На рис. 1, *a* за допомогою мережі Петрі описаний процес збереження інваріанту класу *Employee*, який не є детермінованим, тобто неможливо однозначно визначити коли інваріант буде виконуватися, а коли навпаки.

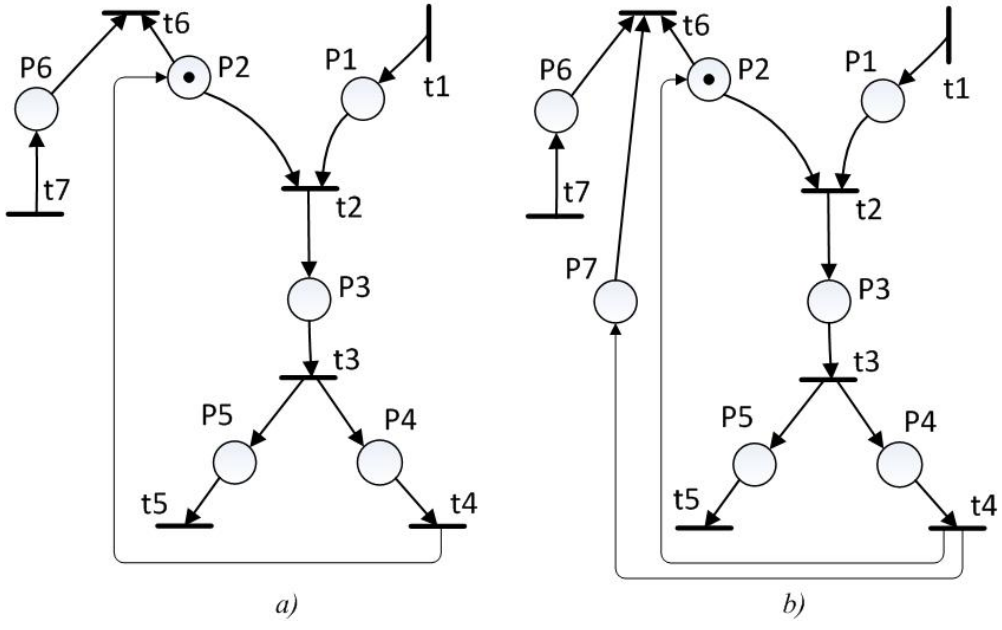


Рис. 1. Модель *a)* можливого порушення та *b)* збереження інваріанту класу Employee

Виклик *setAge*(*integerage*) відбувається в переході $t1$ після чого мітка переходить у вершину

місця $P1$. Це робить можливим спрацювання переходу $t2$, який в свою чергу забирає мітку із вершини-місця $P2$, що унеможливлює спрацювання переходу $t6$. З переходу $t2$ виконуються дії по збереженню інваріанта. В $t3$ відбувається перевірка умови $(age \geqslant 18)$ *and* $(age \leqslant 65)$. Якщо умова виконується, то у $t4$ змінній *age* присвоюється нове значення і мітка повертається в вершину-місце $P2$. Тепер перехід $t6$ може спрацювати, а отже це робить можливим виклик функції $getAge()$. Якщо умова не виконується, то мітка попадає у $t5$ та відбувається виведення повідомлення про помилку.

При детальному розгляді моделі, представленої на рис. 1, можна помітити, що вибір альтернатив з вершини місця $P2$ відбувається недетерміновано, тобто може відбутися передача мітки для активізації як вершини $t2$, так і вершини $t6$, якщо умови активізації вершини переходу $t6$ будуть виконані раніше, ніж умови запуску переходу $t2$. Це не дає однозначної гарантії збереження інваріанту. Для однозначного опису даного алгоритму застосуємо додаткову умову від вершини переходу $t4$ та вершини місця $P7$, яка при використанні у вершині $t6$ забезпечить непорушність інваріанту (рис. 1, *b*). Крім того, потрібно повернути мітку у вершину місця $P2$, що дозволить даному процесу виконуватися багаторазово і забезпечувати живість [3] моделі алгоритму. Остання обставина є показником працездатності реалізуємого алгоритму.

Однозначно виявити подібну приховану помилку реалізації алгоритму допоможе конструкція моделі (у термінах мережі Петрі), у якій мають бути відображені наступні умови:

1. у розгалуженні, яке не містить умови, мають бути альтернативи, що залежать від сукупності різних умов, які не можуть активізувати альтернативні гілки алгоритму одночасно;
2. циклічні конструкції, які побудовані відносно «дозволяючої» вершини місця (у прикладі – вершини місця $P2$) мають повертати початкову розмітку у «дозволяючу» вершину (як у моделі на рис. 1, *a*) для можливості подальшої роботи,
3. гілка від «дозволяючої» вершини, яка активізується першочергово, повинна мати зв‘язок з іншою, альтернативною гілкою вибору, яка моделюється вихідною відносно «дозволяючої» вершиною переходу ($t4 \to P7 \to t6$ у моделі на рис. 1, *b*).

Висновок. Таким чином, використаний при побудові розглянутої моделі інструментарій дає можливість описати набір ідентифікаторів, який дозволяє при верифікації реалізованого алгоритму з детермінованим вибором виявити приховану помилку заміни детермінованого вибору на недетермінований.

Мережа $PN = (P, T, F)$ відображає інваріант у випадку коли

1. ${}^{\bullet(k,s)}t_i \neq {}^{\bullet(v,u)}t_j$ та ${}^{\bullet(i)}t_i = {}^{\bullet(i)}t_j$,
2. $m^0(p_i) = 1 \xrightarrow{1} m^1(p_i) = 0 \ldots \xrightarrow{k} m^k(p_i) = 1$,
3. існує шлях $\alpha(C)$ з вершини t_i до вершини t_j.

При аналізі алгоритмів, що мають розгалуження та паралельні ділянки обробки даних, існує кілька критичних ситуацій, які приводять до втрати працездатності алгоритму. Їх можливо однозначно описати в термінах безпечних мереж Петрі та застосовувати при верифікації реалізованих програмних модулів, що дає можливість позбавлятися більшої частини прихованих помилок у реалізації програмних додатків.

Література. **1.** W.M.P. van der Aalst. Three good reasons for using a Petri-net-based workflow management system. / Eindhoven University of Technology. Available at: http://wwwis.win.tue.nl/~wvdaalst/publications/p52.pdf (accessed 09 January 2015). **2.** Кузьмук В.В. Модифицированные сети Петри для формирования параллельных процессов в системах управления / В.В. Кузьмук, О.О. Супруненко // Восточно-европейский журнал передовых технологий. – 2010. – № 6/8 (48). – С. 50-53. **3.** Супруненко О.О. Засоби автоматизованого аналізу паралельних алгоритмів на основі модифікацій мереж Петрі. / О.О. Супруненко, А.М. Парнюк // Восточно-европейский журнал передовых технологий. – 2010. – № 4/3 (46). – С. 66-71.

Харчун В.М.
ННК "Інститут прикладного системного аналізу" НТУУ "КПІ", Київ, Україна

Оцінювання якості життя населення України з використанням модифікованих показників МІУР

Світовою спільнотою поступово сформувалась нова філософія суспільного розвитку, основу якої складають концепція сталого розвитку та концепція якості життя. Ці дві парадигми тісно пов'язані між собою, багато в чому тотожні, хоча й відрізняються в певних акцентах.

Сьогодні багато наукових центрів світу розробляє різні підходи до переліку показників, що складають основу первинних даних, та застосовуються для визначення єдиного індексу.

Аналіз підходів до поняття «якість життя» свідчить, що на сьогодні немає єдиного усталеного його визначення. Дана категорія ще досліджується, наповнюється новим змістом. Перелік компонентів якості, обґрунтування їх важливості, часто залежить від рівня соціально-економічного розвитку країни, де здійснюється дослідження. У високорозвинених країнах все частіше на перший план виходять суб'єктивні параметри задоволеності життям, тоді як у слаборозвинених країнах переважають дослідження рівня життя. Все це породжує певні проблеми, коли мова йде про вимірювання якості життя. Відповідно до належності дослідника до тих чи інших наукових галузей, шкіл і країн пропонуються різні набори показників, індикаторів, індексів її оцінки. Тому, на нашу думку, вибору показників якості життя має передувати чітко окреслена мета проведення такої оцінки та виявлення основних тенденцій розвитку країни – об'єкту дослідження.

Нове тисячоліття в Україні розпочалося з економічного відновлення після тривалої кризи, пов'язаної з розривом усталених господарських зв'язків у 90 рр. XX ст. У подальшому світова фінансова криза 2008 - 2009 рр. спричинила нову економічну кризу в країні. Посткризове відновлення 2010 – 2011 років змінилось у 2012–2013 роках стагнацією. У 2014 році Україна стикнулася з найгострішими викликами, включно з економічною кризою, військовим конфліктом на Сході країни та анексією АР Крим. Падіння внутрішнього попиту та слабкий зовнішній спричинили падіння реального ВВП на 6,8% та стрімке знецінення гривні.

Реальний ВВП за III квартал 2015 р. порівняно з III кварталом 2014р. (у постійних цінах 2010 р.) склав 92,8%, падіння промисловості у січні-грудні 2015 р. порівняно з відповідним періодом 2014 р. склало 13,4%, зменшення обсягів виробництва у сільському господарстві – 6,8%, скорочення будівництва – 15,9%, споживчі ціни за 2015 р. зросли на 43,3%.

За показником ВВП в розрахунку на одну особу (8 178 дол. США за ПКС 2011 року) Україна посідає 106 місце у 2015 році.

Внаслідок геополітичного конфлікту у 2014–2015 роках перепонами на шляху розвитку економіки України стали:

- руйнування виробничих потужностей та транспортної інфраструктури на території військового конфлікту;
- втрата міжгалузевих та логістичних зв'язків у регіональному та зовнішньо-економічному просторі;
- ускладнення міжнародних відносин із країною – колишнім основним торговельним партнером (Російською Федерацією);
- нестача енергетичної сировини (вугілля);
- зростання інвестиційних ризиків та негативних очікувань населення.

Негативний вплив на економіку України у 2015 році справляють і накопичені системні диспропорції, наслідками яких стали девальваційні та інфляційні шоки.

Найвищий рівень безробіття населення віком 20–64 роки спостерігався у 2001 році (10,6%), найнижчий – у 2008 році (6,2%). Під впливом економічної кризи 2009 року цей показник підвищився до 8,8%, а у 2015 р. рівень безробіття (за методологією МОП) серед економічно активного населення віком 15–70 років становив 9,0%, а працездатного віку – 9,4%. Демографічна ситуація в Україні в період 2000–2015 рр. характеризувалася такими особливостями:

низька народжуваність; висока смертність (особливо серед чоловіків працездатного віку, а воєнний конфлікт ще посилив цю тенденцію) і відповідно низька середня очікувана тривалість життя з вираженою гендерною диференціацією. З 2000 року Україні вдалося досягти певних успіхів у подоланні бідності, однак з 2014 р. показники знову стали погіршуватись. [1]

Це свідчить, що в Україні більш актуальним є дослідження якості життя за об'єктивними, а не суб'єктивними параметрами, висновки якого можна було б використати для вироблення відповідної державної політики.

Об'єктивна сторона якості життя населення визначається комбінацією різних нормативних, статистичних характеристик, за допомогою яких можна судити про ступінь задоволеності науково обґрунтованих потреб й інтересів людей.

Згідно підходу, запропонованого Інститутом прикладного системного аналізу, компонента «якості життя» є інтегрованою оцінкою, яка враховує сумісно усі три виміри сталого розвитку, і тим самим відображає суттєві відносини і зв'язки між трьома нероздільними сферами розвитку суспільства: економічною, екологічною та соціальною. Це передає основний зміст процесів збалансованого соціально-економічного та екологічного розвитку.Економічна складова якості життя оцінюється за двома індексами (глобальна конкурентоспроможність та економічна свобода) і 32 індикаторами, екологічна - за двома індексами (екологічне здоров'я та життєздатність екосистем) і 24 індикаторами, соціальна – за двома індексами (людський розвиток та якість життя) і 13 індикаторами. [2]

Такий підхід дозволяє на основі інверсного аналізу сформувати «простір критичних показників», поліпшення яких вимагає прийняття відповідних управлінських рішень. Одночасно необхідно відзначити і проблемні моменти методу, пов'язані з тим, що використання метрики істотно залежить від якості вихідних даних. Крім того, окремі індикатори та набори даних або базуються на експертних оцінках, або вимагають адаптації до існуючої в країні статистичної бази, що ускладнює розрахунок якості життя. Також є складність при агрегуванні інформації в індекси, що полягає у визначенні ваги вихідних показників без втрати їх значимості та запобігаючи суб'єктивізму при вирішенні цього питання. Із збільшенням рівня агрегування інформації складність зважування непорівнюваних величин істотно зростає.

Тому пропонується модифікація показників МІУР шляхом регресійного аналізу індексів з метою виділення тих, що мають вирішальний вплив на якість життя в Україні. [3]

Таким чином базуючись на показниках за моделлю МІУР в результаті було отримано нову модель оцінки якості життя, яка є більш легкою для використання, адже кількість індексів, що використовується менша, ніж в МІУР, що спрощує процес збору статистичної інформації. При цьому зберігається та сама точність вимірювання та чутливість до змін вхідної інформації. Це пов'язано з тим, що під час регресійного аналізу досліджувався вплив змін кожного з показників на загальний результат вимірювання і, в разі не перевищення мінімального порогового значення зміни індексу якості життя, відповідний показник відкидався.

Література. **1.** Національна парадигма сталого розвитку України / за заг. ред. академіка НАН України, д.т.н., проф., засл. діяча науки і техніки України Б. Є. Патона. – К.: Державна установа "Інститут економіки природокористування та сталого розвитку Національної академії наук України 2012. – 72 с. **2.** Згуровський М.З. Основы устойчивого развития общества: курс лекций в 2 ч. / М. З. Згуровський, Г. О. Статюха. – К.: НТУУ «КПИ», 2010. – Ч.1. – 464 с. **3.** Аналіз сталого розвитку – глобальний і регіональний контексти : монографія / Міжнар. рада з науки (I CSU) та ін.; наук. кер. проекту М. З. Згуровський. – К. : НТУУ КПІ , 2013. – Ч. 1. Глобальний аналіз якості та безпеки життя людей (2011 – 2012). – 328 с.

***Ходневич Я.В.*[1], *Стефанишин Д.В.*[1], *Щодро О.Є.*[2], *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 системний підхід до прогнозування розмивів берегів рік на основі імітаційного моделювання з врахуванням геопросторових даних

Однією з причин виникнення надзвичайних ситуацій на прирічкових територіях є ерозійно-акумулятивна діяльність рік, яка інтенсифікується на ділянках незавершеного і обмеженого меандрування. Наприклад, серйозною проблемою у Волинській області є руслова ерозія на Західному Бузі, в результаті якої відбувається постійна зміна державного кордону України шляхом відторгнення її території. Натурні спостереження за русловим процесом і на інших ріках, зокрема в Карпатському регіоні, показують, що найбільш часто місцеві розмиви їх дна і берегів при паводках спостерігаються в місцях розташування в руслах різного роду перешкод, що формуються природним чином, і відхиляються від берега вниз за течією: дерев, що падають в русла, гряд наносів тощо. При цьому при проходженні паводків інтенсифікуються процеси місцевих руслових переформувань і водна стихія приносить найбільші збитки.

В [1], за результатами лабораторних досліджень, висунуто гіпотезу, що на глибину місцевого розмиву можуть суттєво впливати кут відхилення перешкоди в руслі від берега та її геометричні параметри. Для перевірки цієї гіпотези було проведено імітаційне математичне моделювання умов, при яких за донною перешкодою, що відхиляється вниз за течією, турбулентна течія, направлена від берега, підсилюється. Було підтверджено існування залежності між кінематичними характеристиками потоку за перешкодою та її кутом розташування і відносною висотою [2]. Було встановлено, що місцевий розмив інтенсифікується біля перешкод, що мають висоту 1/3 глибини потоку та кут розташування до набігаючого потоку в 35-40 °.

Отримані результати не тільки підтвердили дані натурних і лабораторних досліджень [1], а й вказали на можливість адекватного прогнозування місцевих розмивів дна і берегів рік, що мають ділянки незавершеного і обмеженого мандрування, на основі імітаційного математичного моделювання кінематичних характеристик водних потоків з виявленням відповідних умов для розмиву. Однак для проведення такого моделювання необхідно володіти мінімальним об'ємом геопросторових даних щодо особливостей перебігу руслового процесу, наявності в руслі річки донних перешкод в різні моменти часу та їх геометрії. Особливу цінність при цьому мають дані дистанційного зондування Землі (ДЗЗ) [3]. Вони дають можливість оцінювати зміни руслових елементів за порівняно короткий проміжок часу. При цьому для складання ситуаційних прогнозів достатньо, щоб знімки забезпечували ідентифікацію гребенів руслових перешкод у руслі, що є можливим при просторовій та радіометричній роздільності принаймні 2 м. Дані ДЗЗ можуть також бути використані з метою попередньої оцінки шорсткості та гідравлічного опору ділянок, для яких проводяться розрахунки елементів гідравлічної структури потоку. Остаточна ідентифікація шорсткості здійснюється в рамках прийнятої математичної моделі.

Література. **1.** Щодро О.Є. Порівняння розрахункових та експериментальних даних про кінематичну структуру, поле тисків, наносотранспортуючу та розмивну спроможності потоку в зоні просторової відривної течії за косо розташованими грядами та донними уступами / О.Є. Щодро // Вісник НУВГП. Вип. 4 (36). – Рівне: 2006. – Ч. 1. – С. 173-179. **2.** 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 // Zeszyty Naukowe Inżynieria Lądowa i Wodna w Kształtowaniu Środowiska. Nr 10. – Kalisz, 2014. – S. 7-18. **3.** Ходневич Я.В. Обґрунтування вибору просторових даних для математичного моделювання руслових потоків в зонах місцевого розмиву /Я.В. Ходневич, В.М. Корбутяк, Д.В. Стефанишин // Зб. наук. праць. 14 Міжнародна науково-практ. конф. «Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях». (5-9 жовтня 2015 р.), м. Київ, Пуща-Водиця. – К.: 2015. – С. 241-247.

Цегелик Г.Г., Краснюк Р.П.
Львівський національний університет імені Івана Франка, Львів, Україна

Математичне моделювання оптимального розподілу реплікаційних баз даних у розподілених інформаційних системах

В роботі розглянуто питання оптимізації розподілу реплікаційних баз даних при моделюванні інформаційних систем за умов мінімізації експлуатаційних витрат.

Вступ. Сучасні тенденції розвитку інформаційних систем обумовлюють перехід від централізованих обчислювальних систем до розподілених. Стратегії розподілу даних за вузлами мережі визначаються як управлінськими, так і виробничими завданнями конкретних організацій та підприємств. Однак, одним з основних недоліків фізичного розподілу даних є жорсткі вимоги щодо швидкодії та надійності каналів зв'язку. Якщо база даних розподілена у декількох вузлах, які територіально є віддаленими та поєднаними повільними і ненадійними каналами зв'язку, а число одночасно працюючих користувачів вимірюється десятками, то ймовірність фіксації розподіленої трансакції невпинно зменшується. У таких випадках обробка розподілених даних суттєво ускладнюється. Однак, у практиці використання розподілених баз даних не для всіх задач необхідним є забезпечення ідентичності баз даних на різних вузлах в довільний час. Достатнім є підтримка ідентичності даних лише у певні критичні моменти часу. Тому, можна накопичувати зміни в даних у вигляді трансакцій в одному вузлі та періодично копіювати ці зміни у інші вузли мережі. Цей процес має назву тиражування даних (*Data Replication – DR*) – асинхронний процес переносу змін об'єктів вихідної бази даних (*source database*) у бази даних, які належать різним вузлам розподіленої системи. Функції DR виконує спеціальний модуль СКБД – сервер тиражування даних, шо носить назву реплікатора (*replicator*). Його завдання – підтримання ідентичності даних у приймаючих базах даних (*target databases*) до даних у вихідній БД.

Принциповою характеристикою тиражування даних є відмова від фізичного розподілення даних. Значення *DR* полягає в тому, що довільна база даних (як для СКБД, так і для працюючих з нею користувачів) завжди є локальною (розміщується на комп'ютері користувача при монопольному доступі користувача чи у локальній мережі установи для забезпечення доступу до бази даних її співробітників); дані розміщуються локально на тому вузлі даних, де вони обробляються; усі трансакції у системі завершуються локально.

Технологія реплікації необхідна для інформаційних систем, що використовуються у галузях в високими вимогами щодо гарантій отримання, своєчасності та цілісності даних, які передаються. Це є системи міжнародних безготівкових банківських платежів, системи обробки даних розподілених технологічних процесів реального часу, білінгові системи, а також спеціалізовані корпоративні системи.

Однак, використання технології реплікації даних висуває низку задач оптимізації, як то оптимального розподілу реклікаційних баз у розподіленій інформаційній системі за умов мінімізації витрат на обслуговування чи обмежень на обчислювальні ресурси вузлів. Окрім того, у практиці застосування інформаційних систем має місце розташування масиву даних у декількох базах даних за певними логічними чи технологічними ознаками. Наприклад, останнім часом популярним є поєднання реляційної бази даних (*Microsoft SQL Server*), яка виступає сховищем бізнес-даних та *NoSQL* бази даних (наприклад, *MongoDB*) як сховища службової інформації (наприклад, лог файлів) [1]. Тому, може мати місце розмішення в одному вузлі розподіленої інформаційної системи декількох баз даних. Як наслідок, дослідження питань оптимального розподілу реплікаційних баз даних у інформаційних системах є актуальними і тому формулювання відповідних математичних моделей та розв'язок відповідних задач оптимізації і є предметом цього дослідження.

Математичне формулювання задачі оптимального розподілу реплікаційних баз даних у вузлах мережі за умови мінімізації витрат на обслуговування. Нехай m баз даних необхідно розподілити у n вузлах мережі ($m < n$). Витрати на обслуговування i-ї бази у j-му вузлі

складають c_{ij} одиниць. Умовою оптимальності у цьому випадку є мінімізація загальних витрат L на обслуговування баз даних інформаційної системи. Математична модель задачі у цьому випадку набуває вигляду:

$$L = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to min, \tag{1}$$

$$\sum_{j=1}^{n} x_{ij} = 1, x_{ij} \in \{0 \cup 1\}, i = 1, 2, ..., m; j = 1, 2, ..., n. \tag{2}$$

При математичному формулюванні цієї оптимізаційної задачі накладається обмеження щодо розміщення тільки однієї бази даних у вузлі (2). Якщо ж позбутися цієї умови, то природньо виникає обмеження щодо обсягів ресурсів пам'яті у вузлі розподіленої інформаційної системи.

Математичне формулювання задачі оптимального розподілу реплікаційних баз даних у вузлах мережі за умови мінімізації витрат на обслуговування та обмеженнях на ресурси пам'яті. Нехай, як і раніше, необхідно розподілити m баз даних, які можуть займати відповідно a_i одиниць пам'яті у n вузлах мережі, загальні об'єми ресурсів пам'яті яких складають b_j одиниць відповідно. Витрати на обслуговування i-ї бази у j-му вузлі складають c_{ij} одиниць. Умовою оптимальності, як і у попередньому випадку, є мінімізація загальних витрат L на обслуговування баз даних інформаційної системи. Тоді, математична модель задачі формулюється наступною системою залежностей:

$$L = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to min, \tag{3}$$

$$\sum_{j=1}^{n} x_{ij} \leqslant m, \sum_{i=1}^{m} a_i x_{ij} \leqslant b_j, \tag{4}$$

$$x_{ij} \in \{0 \cup 1\}, i = 1, 2, ..., m; j = 1, 2, ..., n. \tag{5}$$

Формулювання обчислювальних алгоритмів побудови розв'язків задач оптимізації. Ефективним підходом до побудови розв'язків зформульованих оптимізаційних задає є використання методу динамічного програмування [2]. Пропустивши проміжні викладки, наведемо тільки остаточні результати: рекурентні рівняння Белмана. Так, для задачі (1)–(2):

$$L_1(\{x_{ij}\}) = g_1(\{x_{1j}\}), L_k(\{x_{ij}\}) = min\{g_k(x_{kj}) + L_{k-1}(\{x_{ij}\} \setminus \{x_{kj}\})\},$$
$$g_k(x_{kj}) = \min_j \{c_{kj}\}, j = 1, 2, ..., n; k = 2, 3, ..., m.$$

Для задачі ж (3)–(5) відповідні рекурентні співвідношення набувають вигляду:

$$L_1(\{x_{ij}\}) = g_1(\{x_{1j}\}),$$
$$L_k(\{x_{ij}\}) = min\left\{g_k(x_{kj}) + L_{k-1}(\{x_{ij}\})_{\{b_1, b_2, ..., b_j - a_k, ..., b_n\}}\right\},$$
$$g_k(x_{kj}) = \min_{j, a_k \leqslant b_j} \{c_{kj}\}, j = 1, 2, ..., n; k = 2, 3, ..., m.$$

З використанням отриманих співвідношень неважко побудувати розв'язки задач оптимізації за допомогою таблиць чи графічним методом.

Висновки. Здійснено математичну постановку задач оптимального розподілу реплікаційних баз даних у розподілених інформаційних системах та сформульовано рекурентні рівняння Белмана розв'язку задач методом динамічного програмування. Результати роботи мають як самостійне значення, так і можуть бути використані при тестуванні наближених методів розв'язку задач оптимізації з використанням «жадібних» або генетичних алгоритмів.

Література. **1.** Sadalage P.J. and Fowler M., NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence. – NY: Addison-Wesley, 2012. **2.** Цегелик Г.Г. Математичне програмування: навч. пос. – Львів, ЛНУ ім. Івана Франка, 2011.

Чертов О.Р., Сахаров С.Ю., Юрченко Д.В.
Кафедра прикладної математики, Національний технічний університет України "Київський політехнічний інститут", Київ, Україна

Виділення ключових значень з показів CGM для прогнозування нічної гіпоглікемії у хворих на цукровий діабет 1-го типу

Вступ. Цукровий діабет – це хронічне захворювання, яке потребує постійного медичного догляду і нагляду з боку хворого, щоб попередити можливі ускладнення та зменшити ризик довгострокових ускладнень [1]. Згідно з даними International Diabetes Federation (IDF), в світі нараховується більше 415 мільйонів хворих на діабет людей [2]. Гіпоглікемія є нагальною проблемою для хворих на діабет першого типу (тобто таких, організм котрих не може самостійно виробляти інсулін). Відповідно з [1], хворі на діабет першого типу мають в середньому два приступи симптоматичної гіпоглікемії кожного тижня і один тяжкий приступ гіпоглікемії раз на рік.

Як правило, вважається, що має місце приступ нічної гіпоглікемії у хворого, якщо рівень глюкози в плазмі його крові падає нижче 70 мг/децилітр у нічний час доби [3]. В такому випадку функціонування організму хворого порушується, що може призвести навіть до його смерті.

У роботі пропонується методика виділення ключових значень з часового ряду рівня глюкози, який отримується шляхом зняття показів з пристроїв CGM (Continuous Glucose Monitoring) у хворих на цукровий діабет 1-го типу [4], з метою побудови в подальшому прогнозу щодо можливого приступу нічної гіпоглікемії у хворого на основі виділених значень.

Методика до виділення ключових значень з часового ряду рівня глюкози. Пропонована методика складається із таких послідовних етапів:

1) фільтрація даних;
2) корекція даних;
3) попередня оцінка часових проміжків основних прийомів їжі;
4) фіксація значень базальної глікемії в динаміці постпрандіальної глікемії.

Фільтрація даних проводиться з метою відкидання із часового ряду нульових показань, що з'явилися у зв'язку з некоректною роботою пристрою CGM.

Корекція даних полягає в обробці пилоподібних стрибків часового ряду. Було виявлено, що показання в такі періоди з'являються декілька разів на 10 хвилин, а не один раз, як це передбачено в налаштуваннях пристроїв CGM. Тоді обирається одне з таких значень, а саме — те, що отримується з кратністю через 10 хвилин від показань, які були б за нормальної роботи апаратури. Наприклад, якщо перед проміжком з пилоподібними стрибками виміри робились в моменти часу гг:09, гг:19, гг:29, гг:39, гг:49, гг:59, гг:09, то при збоях обладнання будуть братися ті значення, які є продовженням даного ряду, а не "альтернативного"типу гг:04, гг:14, гг:24, гг:34, гг:44, гг:54, гг:04.

Попередня оцінка часу основних прийомів їжі виконується окремо для кожного пацієнта: для кожного дня визначаються моменти часу, коли спостерігаються значні підйоми рівня глюкози. Далі відбувається підрахунок кількості таких підйомів та типові для даного пацієнта періоди часу, коли вони мають місце.

На основі цих даних відбираються для подальшого розгляду значущі підйоми рівня глюкози. Значущим вважається такий підйом, максимальне значення якого є найбільшим серед підйомів протягом типового для даного пацієнта часового проміжку основного прийому їжі.

Фіксація значень базальної глікемії в динаміці постпрандіальної глікемії здійснюється для кожного значущого підйому рівня глюкози. Фіксуємо 3 показання: $M_1(t_1; g_1)$ — локальний мінімум перед підйомом, $M_2(t_2; g_2)$ — максимум підйому, $M_3(t_3; g_3)$ — показання через 2 год

Дана робота була виконана в рамках проекту AMMODIT (Approximation Methods for Molecular Modelling and Diagnosis Tools) за Європейською програмою Horizon2020, грантова угода 645672-AMMODIT.

після показання $(t_1; g_1)$. Знаходимо допоміжну точку $M_4(t_4; g_4)$ — локальний мінімум після підйому, й обчислюємо швидкості підйому та спаду рівня глюкози:

$$v_{up} = \frac{g_2 - g_1}{t_2 - t_1}, \quad v_{down} = \frac{g_2 - g_4}{t_2 - t_4}. \tag{1}$$

А також фіксується найменший показник за нічний період (00:00-06:00).

Приклад застосування методики. Використаємо часовий ряд (рис. 1) рівня глюкози одного з пацієнтів за добовий період. Результати занесемо у таблицю (табл. 1).

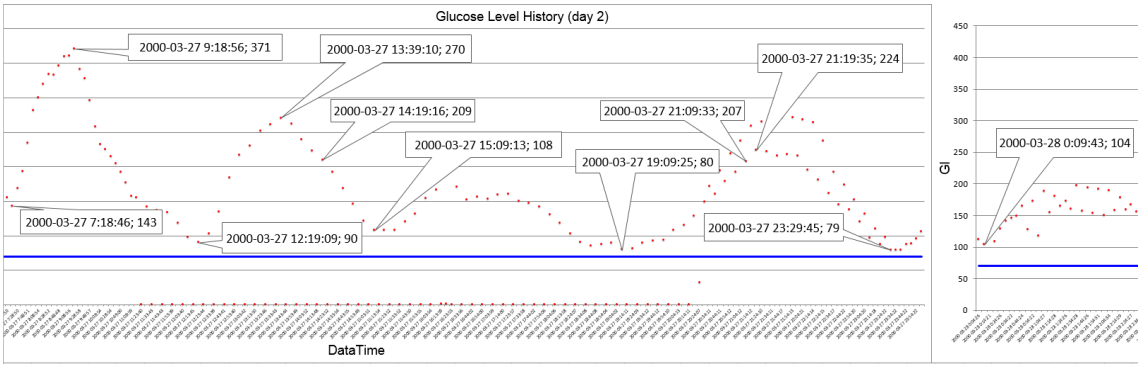


Рис. 1. Часовий ряд рівня глюкози за добовий період з поміченими ключовими точками

Дата	27.03.2000						
Глікемія	базальна		постпрандіальна				
	t_1	g_1, мг/дл	t_2	g_2, мг/дл	g_3, мг/дл	v_{up}, мг/дл/год	v_{down}, мг/дл/год
1	7:18	143	9:18	371	371	114	94
2	12:19	90	13:39	270	209	138	65
3	19:09	80	21:19	224	207	65	67
ніч (мін.)	00:09	104					

Табл. 1. Виділені ключові значення

Висновки. У роботі запропоновано методику виділення ключових значень з часового ряду рівня глюкози з метою побудови в подальшому прогнозу щодо можливого приступу нічної гіпоглікемії у хворих на цукровий діабет 1-го типу.

Перевагою методики є її простота та детермінованість. Незалежні застосування методики до однакових вхідних даних приведуть до однакових результатів. Це дозволяє адекватно порівнювати ефективність предикторів, яким на вхід подаються результати застосування методики.

Запропонована методика може бути покращена за рахунок врахування додаткової інформації про раціон пацієнта та схему його лікування. В подальшому планується враховувати дози та тип швидкого інсуліну, який вводиться пацієнту перед регулярними прийомами їжі.

Література. **1.** The importance of hypoglycemia in diabetic patients / G. Shafiee, M. Mohajeri-Tehrani, M. Pajouhi, B. Larijani // Journal of Diabetes & Metabolic Disorders. — 2012. — Vol. 11 (1). — pp.1-7. **2.** International Diabetes Federation. IDF Diabetes Atlas. — Brussels, Belgium: International Diabetes Federation, 2015. — 7 ed. — 144 p. **3.** Defining and reporting hypoglycemia in diabetes / American Diabetes Association Workgroup On Hypoglycemia // Diabetes Care. — 2005. — Vol. 28 (5). — pp.1245-1249. **4.** A Pilot Study to Evaluate the Navigator Continuous Glucose Sensor in the Management of Type 1 Diabetes in Children // Diabetes Research in Children Network (DirecNet) [Online] Available: http://direcnet.jaeb.org/Studies.aspx?RecID=166

Шаронова Н.В., Козуля М.М.
Національний технічний університет "Харківський політехнічний інститут", Харків, Україна

Знання-орієнтоване інформаційне забезпечення досліджень складних систем

Актуальність і мета роботи. За останні десятиліття розроблені два напрямки розв'язання аналітичних завдань: застосування формальних методів пошуку рішення без попередніх процедур, зведення до задачі математичного програмування; логіко-математичний підхід з експертними процедурами з включенням людини (особи з прийняття рішення – ОПР). Дані підходи мають суттєві недоліки при наявності елементів невизначеності: для першого напрямку – не розв'язуються різного виду невизначеності й ігноруються неформалізовані фактори, пов'язані з досягненням мети (прийняття рішення); у другому випадку суттєво зростає суб'єктивна складова при встановленні остаточного рішення, що упускає об'єктивні фактори багатокритеріальної оцінки стану, процесів і ситуації за умови досягнення цілі [1–4].

Виходячи з вище сказаного, мета роботи спрямована на підвищення ефективності інформаційних технологій інтелектуальної обробки даних на основі ідентифікації знань і подання результатів для прийняття рішень на об'єктивній основі стосовно управління складними об'єктами соціально-еколого-економічного призначення.

Методика дослідження і аналіз результатів. Взаємодія складних систем за структурою (складом і зв'язками), властивостями, функціональністю потребує нової за обсягом і характером вихідної інформації, яка складатиме додаткове системне утворення – інформаційну систему, яка визначатиме раціональне об'єднання сучасних об'єктивних знань, суперечливих даних, неповноту, неточність, нечіткість характеристик, сукупність теоретично-практичного інструментарію, засобів, методів досягнення рішення, цілей або не вирішення задач для об'єкта – умови ризику [3–5]. Основою для таких системних комплексів є інформаційна компонента, яка враховує особливості окремих складових, їх відношення і функціональність в межах об'єкта і умови взаємодії з навколишнім середовищем.

Підвищення ефективності управління будь-якого рівня складності системного об'єкта в умовах невизначеності інформації і знань про неконтрольовані ситуації в організаційно-технічних системах пов'язано з побудовою нечіткої системи управління на феноменологічній базі – досвід, знання, використання знання-орієнтованих систем. Для формування математичної бази обробки даних використовується ентропійна ризик-оцінка складних систем в умовах невизначеності. Розв'язання задач з оцінки якості системних об'єктів передбачає змістовність управління як необхідність встановлення початкової рівноваги на будь-якому рівні організації системи. За таких умов необхідним є досягнення точки відповідності, що гарантується при мінімальному ризику прийняття рішення: не втручання, а контроль самовільної адаптації за рахунок довільних процесів; регулювання і управління до максимального упорядкування і самоорганізації об'єкта.

З цією метою запропоновано інформаційно-аналітичну систему виду «0-стан об'єкта (системи як складові) – процес (система процесів, оператор змін стану або його підтримки) – і-стан об'єкта (системи)» для розв'язання завдань сталого розвитку складного об'єкту соціально-еколого-економічного змісту. Така організація об'єкта системного аналізу виходить за межі класифікаційного поняття складності (великі системи тощо), стає доцільним відображати його як систем-системне утворення і визначати як корпоративну систему з кооперативними зв'язками, які становлять основу її самоорганізації для досягнення стабільності і рівноважного розвитку (синергетика) [5, 6].

У такому складному об'єкті дослідження виділяють цільовий комплекс «(система $\overset{\text{процес}}{\longleftrightarrow}$ зовнішнє середовище)$\overset{\text{стан}}{\longleftrightarrow}$ система $\overset{\text{процес}}{\longleftrightarrow}$ зміни системи – процес – (стан системи)' – (система' $\overset{\text{процес}}{\longleftrightarrow}$ зовнішнє середовище)», в якому розглядають як елементи системи, зовнішнє відповідно до системи середовище, що утворює внутрішній простір об'єкта, де самоорганізуються зв'язки і формується структура об'єкта, взаємодіючого з оточуючим середовищем. Кооперативна дія

внутрішніх і зовнішніх процесів стабілізує складну систему, що і відповідає уявленням синергетики («Synergia» – корпоративна дія). На практиці застосування запропонованого підходу формування знання-орієнтованого інформаційного забезпечення застосовано для розв'язання завдань в системі екологічного моніторингу. Аналіз предметної області дозволив виділити й деталізувати такі основні бізнес-процеси (роботи) (рис. 1). 1. Система нагляду: 1.1. Вибір території. 1.2. Методика виміру параметрів навколишнього середовища. 1.3. Характеристика факторів дестабілізації досліджених систем. 1.4. Заповнення вихідної інформації на основі спостережень. 1.5. Класифікація отриманої інформації. 2. Система аналізу: 2.1. Аналіз інформації. 2.2. Вибір моделі розрахунку. 2.3. Розрахунок фактичних параметрів стану системи і оцінка процесів у ній. 2.4. Вибір бази порівняння. 2.5. Порівняння фактичних і розрахункових результатів з параметрами природнього стану системи: $\Delta S \to 0, S_1 \to \min \Delta S > 0, S_2 \to \max$. 2.6. Аналіз отриманих результатів порівняння. 3. Система оцінки: 3.1. Вивчення тенденцій змін стану за попередніми дослідженнями. 3.2. Визначення механізмів самоорганізації з регулювання стійкого розвитку системи. 3.3. Встановлення управляючої дії при необхідності стабілізації стану системи.

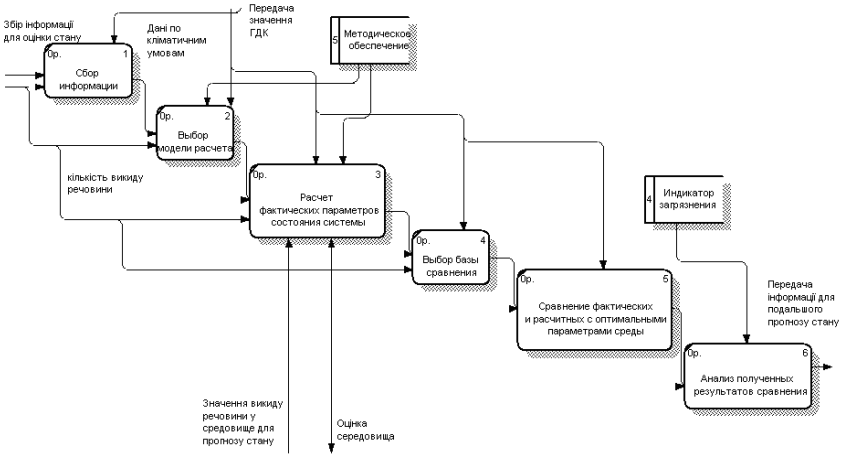


Рис. 1. Діаграма потоків даних

Висновки. У роботі розглянуто доцільність запровадження знання-орієнтованих систем інформаційного забезпечення дослідження складних об'єктів відповідно до аналітичної системи на базі синергетичного підходу з розв'язання завдань прийняття рішень. Надано практичність реалізації запропонованого методичного забезпечення для розв'язання комплексних завдань з обробки моніторингової інформації.

Література. **1.** Huang G.H. Perspectives of environmental informatics and system analysis /G.H. Huang, N.B. Chang // Journal of Environmental Informatics – 2003. – №1(1) – p. 1–6. **2.** Chang N.B. Optimal planning of the reservoir watershed by grey fuzzy multi-objective programming (I): theory / N.B. Chang, C.G. Wen, Y.L. Chen, Y.C. Yong // Water Res. – 1996 – 30(10) – p. 2329–2334. **3.** Мельник Л. Г. Методология развития: Монография/ Л. Г. Мельник. – Сумы: ИТД «Университетская книга», 2005. – 602 с. **4.** Згуровський М. З. Основи системного аналізу / М. З.Згуровський, Н. Д. Панкратова. – К.: Видавнича група BHV, 2007. –544 с. **5.** Козуля Т. В. Теоретико-практические основы методологии комплексной оценки экологичности территориальных и объектовых систем. Монография. / Т. В. Козуля – Saarbrücken: Palmarium Academic Publishing, 2014. – 298 с. **6.** Козуля Т.В. Розробка оцінки екологічності техногенних об'єктів на основі методу компараторної ідентифікації / Т.В. Козуля. М.О. Білова. М.М. Козуля // Східно-Європейський журнал передових технологій. – 2015. – №5/10(77) – с. 27-33.

Шевченко Д.С.
Національний технічний університет України «КПІ», Київ, Україна

Нестаціонарна задача комівояжера в умовах авіатрансферу з додатковими обмеженнями

Розвиток та популяризація інформаційних і комунікаційних технологій дозволяє збирати інформацію і планувати індивідуальні екскурсійні маршрути самостійно. Існуючі сервіси пошуку подорожей орієнтовані на пошук оптимального шляху за певним критерієм лише між двома пунктами. У ряді випадків користувачу бажано поєднати відвідування кількох міст у певний період часу і оптимізувати свій маршрут за вартістю і особистими вподобаннями.

Розглянуто задачу планування авіамаршруту, де відомі бажані міста відвідування і додаткові побажання мандрівника, такі як часовий проміжок подорожі, мінімальна тривалість перебування в місті, максимальна вартість авіатрасферу. Така задача є узагальненим випадком класичної задачі комівояжера і відрізняється від неї наявністю більше ніж одного ребра між двома вершинами, що обумовлено наявністю множини рейсів між двома містами [1,2]. Вартість рейсів різна і залежить від часу відправлення та авіакомпанії.

Розглянемо математичну постановку нестаціонарної задачі комівояжера. Дано зважений орієнтований мультиграф $G = (V, F)$ де $V = \{v_1, v_2, ..., v_n\}$ – множина n вершин, які являють собою n міст; F – множина дуг, які відображають наявні авіасполучення. Між двома вершинами може існувати декілька дуг, що відображає наявність декількох рейсів між містами від різних авіакомпаній, які можуть відрізнятися за вартістю і часом. Планування маршруту по містам має вкладатися в часові межі. Нехай t – поточний час, $t \in [t_{min}, t_{max}]$, де t_{min} – бажаний час початку авіатуру, а t_{max} – бажаний час повернення у початкове місто. Тоді тривалість подорожі буде дорівнювати $T = t_{max} - t_{min}$.

Позначимо F множину рейсів $(v_i v_j)^t$ з міста i в місто j в момент часу t, а f_{ij}^t – конкретну дугу, яка відповідає рейсу. Нехай ціна рейсу задається функцією $cost(f_{ij}^t)$ – невід'ємна залежна від часу функція, яка визначає загальну вартість перельоту з міста i в місто j у час t. Необхідно знайти оптимальний за вартістю шлях з початкової вершини $s \in V$, що охоплює всі бажані міста та повертається знову у початкову вершину у заданий проміжок часу $[t_{min}, t_{max}]$ чи повідомити що неможливо побудувати авіамаршрут у заданий інтервал. Шляхом x будемо називати упорядковану за часом множину рейсів $(f_{i_1 i_2}^{t_1}, f_{i_2 i_3}^{t_2}, \ldots, f_{i_n i_{n+1}}^{t_n})$, для якої $i_1 = s$, $i_{n+1} = s$. Тоді вартість шляху x визначається як $c(x) = \sum_{k=1}^{n} cost(f_{i_k i_{k+1}}^{t_k})$.

Метою є знаходження оптимального за вартістю шляху x^*. Додатково може накладатися умова про тривалість перебування комівояжера в місті d. Нехай $x_{ij}^t = \{0, 1\}$ – ідентифікатор наявності рейсу в маршруті і приймає значення 1, якщо маршрут містить дугу f_{ij}^t, і 0 в протилежному випадку. Математична постановка нестаціонарної задачі комівояжера в умовах авіатрансферу з додатковими умовами задається так:

$$c(x) = \sum_{k=1}^{n} cost(f_{i_k i_{k+1}}^{t_k}) \to min, \tag{1}$$

$$\sum_{t=t_1}^{t_n} \sum_{i=1}^{n} x_{ij}^t = 1, \forall j = \overline{1, n}, \tag{2}$$

$$\sum_{t=t_1}^{t_n} \sum_{j=1}^{n} x_{ij}^t = 1, \forall i = \overline{1, n}, \tag{3}$$

$$t_k - t_{k-1} \geqslant d, \forall k = \overline{1, n}, \tag{4}$$

$$t_n - t_1 \leqslant T. \tag{5}$$

Метою оптимізаційної задачі є мінімізація загальної вартості авіамаршруту, яка подана цільовою функцією (1). Умови (2) та (3) гарантують, що комівояжер прилітає і вилітає з

міста лише один раз. Умова (4) гарантує, що часовий проміжок між вилітами не менший ніж заданий, а умова (5) – що тривалість подорожі не перевищує заданий поріг.

Нині не існує прийнятних на практиці алгоритмів роз'язання нестаціонарної задачі комівояжера. Результатом пошуку ефективних методів розв'язання такої задачі стало використання біонічних алгоритмів, у тому числі, еволюційних та генетичних. Далі пропонується алгоритм, розроблений на основі методу імітації відпалу [3,4].

Наведемо пояснення ключових понять методу імітації відпалу таких як: стан, сусідній стан, прийняття імовірностей. Станом s будемо називати послідовність вершин графу задачі чи нулів у випадку, коли комівояжер не здійснює переліт. Часовою одиницею вважатимемо один день. Кількість елементів послідовності дорівнює бажаній тривалості подорожі у днях. Номер елементу в послідовності відповідає дню подорожі, а елементи безпосередньо – номеру міста призначення рейсу в поточний день або 0 у випадку відсутності перельоту. Наприклад, можливим станом для планування маршруту тривалістю в 9 днів через 3 міста може бути наступна послідновність: $(0, 3, 0, 0, 1, 0, 0, 0, 2)$.

Сусідні стани утворюються перестановкою двох елементів в послідовності. Імовірність переходу від поточного стану s в новий стан s' визначається ймовірнісною функцію прийняття $P(e, e', T)$, що залежить від енергії двох станів: $e = E(s)$ та $e' = E(s')$; та від глобального параметру T, який змінюється в часі та називається «температура»: $P = e^{-\frac{e'-e}{T}}$. Така стратегія переходу в сусіній стан запобігає від застрявання в локальному мінімуму, який гірший, ніж глобальний. Коли $T \to 0$, ймовірність $P(e, e', T)$ прямує до нуля, якщо $e' > e$. Тобто з кожним пониженням температури гірше рішення обирається з меншою ймовірністю.

До основних кроків методу імітації відпалу додано додатковий крок, так званий крок балансування. Він полягає у пошуці дешевшого рейсу в найближчому околі. Якщо такий рейс знайдено, то розглядуваний сусідній стан буде модифікований заміною рейсу. Наприклад, для міста номер 3 будуть проаналізовані рейси з міста 2 у місто 3 у перший та третій день:

$$(\overline{0, 3, 0}, 0, 1, 0, 0, 0, 2) \to \begin{bmatrix} (\overline{3, 0, 0}, 0, 1, 0, 0, 0, 2) \\ (\overline{0, 0, 3}, 0, 1, 0, 0, 0, 2) \end{bmatrix}$$

Загальну схему алгоритм розв'язання нестаціонарної задачі комівояжера можна подати так.

1. Початковий стан генерується випадковим чином. Поточний стан дорівнює початковому. Вибір початкової температури.
2. Оцінка поточного стану.
3. Основний крок алгоритму.
   - 3.1. Генерування сусіднього стану.
   - 3.2. Балансування сусіднього стану.
   - 3.3. Оцінка сусіднього стану.
   - 3.4. Критерій зміни поточного стану.
4. Зменшення температури. Якщо температура більша за визначений поріг – перехід до кроку 3, інакше - до кроку 5.
5. Завершення алгоритму. Поточний стан є розв'язком задачі.

Отже, сформульовано математичну постановку нестаціонарної задачі комівояжера в умовах авіатрансферу з додатковими обмеженнями та запропоновано алгоритм розв'язання на основі методу імітації відпалу.

Література. **1.** Fox, K., Gavish, B., Graves, S.: An n-constraint formulation of the (time dependent) traveling salesman problem. Oper. Res. 28, 101–102 (1980) **2.** Bigras, L.-Ph., Gamache, M., Savard, G.: The time-dependent traveling salesman problem and single machine scheduling problems with sequence dependent setup time. Discrete Optim. 5, 685–699 (2008) **3.** Kirkpatrick S., Gelatt C. D., Vecchi M. P. Optimization by simulated annealing. Science. v220 (1983), pp 671–680. **4.** Aarts E. H. L., Korst J. H. M. Simulated annealing and Boltzmann machines. Chichester: Wiley, 1989.

***Яковлева А.П.*[1], *Мухін В.Є.*[2], *Корнага Я.І.*[2], *Курдуп І.О.*[1]**
[1] *Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна;* [2] *Факультет інформатики та обчислювальної техніки НТУУ "КПІ", Київ, Україна*

Про одну оптимізаційну задачу розташування об'єкта на конкретній території

Шукається місце розташування об'єкта на де-якій території. Ключовим для роботи об'єкта є розташування його цільових клієнтів. Розглядається ряд математичних постановок.

1. Детермінована задача безкмовної оптимізації такого розміщення об'єкта, щоб сумарна відстань до заданих точок (розташування клієнтів) була найменьшою.
   Нехай $x^i, i = 1, 2, ...m$, - координати клієнтів; $x^i = (x_1^i, x_2^i), \overline{1,m} \in R^2$. Потрібно знайти точку $x^0 \in R^2$, що мінімізує наступну функцію:
   $f(x) = \sum_{i=1}^{n} ||x^0 - x^i||^2$, де $||x|| = \sqrt{x_1^2 + x_2^2}$ - евклідова норма вектора.
   Аналітичним розв'язком цієї задачі буде точка $x^0 = \frac{1}{m}\sum_{i=1}^{m} x^i$, оскільки система $f'(x) = 0$ набуде вигляду $2\sum_{i=1}^{m} |x - x^i|$, або $mx - \sum_{i=1}^{m} x^i$.
   Це буде точкою строгого глобального мінімуму в силу опуклості цільової функції.
2. Задача з розташуванням точок на еліпсоїді.
   Розглядається наступна задача умовної оптимізації: в просторі R^2 задано m точок $x^i, i = \overline{1,m}$. Необхідно знайти точку $x^0 \in D$

$$D = (x \in R^2 : \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} \leqslant 1, a_i \in R) \tag{1}$$

   таку, щоб мінімізувати функцію $f(x) = \sum_{i=1}^{n} ||x^0 - x^i||^2$
   У випадку, коли $D = (x \in R^2 : x_1^2 + x_2^2 = 1)$, тобто $a_1 = a_2 = 1$, задача має наступний аналітичний розв'язок [1]
   Задача формулюється так:

$$f(x) \to \min, g(x) = ||x||^2 - 1 = 0$$

   Функція Лагранжа дає умову стаціонарності, з якої випливає, що $x^0 = \frac{1}{m}\sum_{i=1}^{m} x^i$, $||x^0||^2 = 1$.
   Таким чином, точка мінімуму буде $y^0 = \frac{x^0}{||x^0||}$, де $||x^0|| = \sqrt{x_1^0 + x_2^0}$
   Задача з обмеженням (1) може бути розв'язана за допомогою чисельних методів (метод проекції градієнту або метод штрафних функцій)
3. Стохастична залача з випадковим значенням точок.
   В просторі R^2 задано m точок $x^i = (x_1^i, x_2^i), \overline{1,m}$. Необхідно знайти точку $x \in D$:

$$D = (x \in R^2 : \sum \frac{x_i^2}{a_i^2} \leqslant 1, a_i \in R)$$

   таку, щоб мінімізувати функціонал $f(x) = \sum_{i=1}^{m} ||x^- x^i||^2$. Але при цьому кожна з точок x^i рівномірно розподілена у сфері радіусу r з центром в точці x^i.
   Задано щільність розподілу. Необхідно знайти мінімум стохастичної функції:

$$F(x) = Mf(x, \Theta) = M\sum_{i=1}^{m} ||x - \Theta_i||^2$$

   Розв'язується за допомогою стохастичного квазіграфічного метода.

Як результат планується створити програмний продукт, що реалізує розв'язок задачі оптимального розташування об'єктів для різних постановок.

Література. 1. Васильев Ф. П. Чисельные методы решения экстремальных задач. - М.: Наука, 1998.

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.).

Arkhipov D.V.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

The rational approach for working with heterogeneous knowledge bases

Nowadays the power of information is an axiomatic fact that could be illustrated by Ronald Reagan words: "Information is the oxygen of the modern age". And it is hard to argue with that, because even if we take into account digital storage only, the total amount of possible information is estimated at 295 exabytes for 2007 [1] and that value doubles every three years roughly. However, the usage of these huge amounts of structured and unstructured data required the development of "knowledge base" concept.

There is a variety of ways for organizing knowledge bases, which leads to significant differences in their structures. In addition, the absence of knowledge types unambiguous classification and forms representation causes problems with sharing knowledge base data with external agents, as well as with other knowledge bases. Thus, taking into account, that each data storage uses its own conceptual "language", usage of multiple distributed sources for run-time data mining, sharing and reusing of knowledge becomes very complex. Therefore, it is necessary to investigate the approaches for correct work with different data sources.

There are currently several approaches for working with data distributed among different sources.

The simplest ones allow agents to extract data in the most common formats, such as CSV or XLS. While this makes possible for various applications to access the data without any special converters, these applications have to "know" about knowledge base structure beforehand. Other obvious disadvantages include the inability to keep data up to date without regular synchronization with the source, as well as possible loss of information due to restrictions of the format being used for data export.

Another approach assumes that data mining agents use information on a particular knowledge base structure to develop data conversion plugins. In this case, all the data can be preserved, and run-time analysis becomes possible.

Unfortunately, every new data source connection setup is going to be time-consuming, because it requires new converter implementation. The development of Knowledge Query and Manipulation Language (KQML) [2], which was later superseded by FIPA Agent Communication Language [3], was an ambitious attempt to solve the aforementioned problem. However, despite the open-source language design and partnership with large software companies (such as HP, IBM, and Sun Microsystems), the FIPA standards were never widely adopted. Probably, that failure was a result of a solution complexity and, as a consequence, high implementation costs for commercial software.

The traditional solution of the problem was based on acceptance of new standard by the majority of software companies. However, the increased availability of relatively cheap "cloud" computing has created another solution to this issue. It is possible to develop a cloud-based SaaS platform comprising independent, open source converters, that could communicate with various data sources and provide output in a standardized form. Similar approaches already exist in other fields - for instance, video decoding. Currently, there are several actively used video codecs. However, you only need FFMpeg [4] decoder framework to be able to "read" the majority of them.

Therefore, the basic functionality of the proposed solution could be provided for free, while paid services might include bigger bandwidth, unlimited data processing, professional support, etc. The open source platform could guarantee the relevance whereas scalable computing architecture would increase maintainability and simplify requirements for data mining agents.

References. **1.** Hilbert, M., and Lopez, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60-65. Retrieved from http://science.sciencemag.org/content/332/6025/60 **2.** Finin, T., Fritzson, R., Mckay, D., & Mcentire, R. (1994). KQML as an agent communication language. *Proceedings of the Third International Conference on Information and Knowledge Management - CIKM '94*. Retrieved from https://www.aaai.org/Papers/Workshops/1994/WS-94-02/WS94-02-007.pdf. **3.** Poslad, S. (2007). Specifying protocols for multi-agent systems interaction. *ACM Transactions on Autonomous and Adaptive Systems TAAS, 2(4)*. Retrieved from http://www.fipa.org/subgroups/ROFS-SG-docs/2007-TAAS-specifying-MAS.pdf **4.** About FFmpeg. Retrieved from https://www.ffmpeg.org.

Azzuz I.D.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Analysis of decision-making methods in computer poker

Introduction. A significant development in Artificial Intelligence is tightly connected to the optimal strategies in games. When trying to solve perfect information games, we obtain notable results. It is hard for a human mind to predict all the possible actions in chess, and always loses to Artificial Intelligence because it can compute all the precise decisions.[1] However, the brute force approach is not enough to determine optimal strategy in such a game as GO. Go is a hard game for computers to play: it has a high branching factor, a deep tree, and lacks any known reliable heuristic value function for non-terminal board positions.[2] Computing approximation of Nash equilibrium is a part of solving games with imperfect information. This research is aiming at analyzing algorithms that apply to the most popular card game, poker. There has been a lot of research made to solve the problem of a poker game in the second half of the 20th century. Nevertheless, Heads-up limit hold'em poker, which is one of the variants of poker, was solved by applying Counterfactual Regret Minimization algorithm to it.[3]

Methods. The usage of CFR algorithm and building full tree for poker with more than 2 players is very space-efficient. There has been a research that was examining the application of Monte Carlo tree search method to poker.[4] That MCTS method explores more promising branches, and thus, finds more optimal actions to win. The tree search builds according to these steps until run out of computational time.

1. *Selection*: Starting at the root node, a child selection policy is recursively applied to descend through the tree until the most urgent expandable node is reached.
2. *Exploration*: If it is not a terminal node then add a child node to expand a tree
3. *Simulation*: A simulation is run from the expanding node until a result is achieved.
4. *Backpropagation*: The simulation result is "backed up" (i.e. backpropagated) through the selected nodes to update their statistics.[2]

The more iterations are repeated, the better. Choosing the best action is based on the estimated value, which is stored in each node. Using deep learning neural networks succeeded in improving AI results. For instance, Google combined MCTS and NN algorithms for a new approach in computer GO, AlphaGO. A *supervised* learning policy network was trained by millions of games which were performed by experts and professionals. Afterwards, the *reinforcement* learning policy was trained by self-play, which helped improve the SL policy. Then, a *value* network was trained to predict the winner of games played by the RL policy network against itself. Value and policy networks significantly improve MCTS rollout.[5] Applying MCTS to reduce tree search and using deep learning neural networks to compensate psychological factor such as bluffing for poker game can produce a significant result.

Conclusion. We understand that a poker is a complicated system with imperfect information. As shown in this paper, a combination of different algorithms can prove itself as a successful tool for achieving the best optimal strategy. That is why such an approach to this kind of games is a rich domain that can offer a vast horizon for researchers of AI.

References. **1.** M. Campbell, A. Hoane and F. Hsu, "Deep Blue Artificial Intelligence, vol. 134, no. 1-2, pp. 57-83, 2002. **2.** C. Browne, E. Powley, D. Whitehouse, S. Lucas, P. Cowling et al, "A Survey of Monte Carlo Tree Search Methods IEEE Trans. Comput. Intell. AI Games, vol. 4, no. 1, pp. 1-43, 2012. **3.** M. Bowling, N. Burch, M. Johanson and O. Tammelin, "Heads-up limit hold'em poker is solved Science, vol. 347, no. 6218, pp. 145-149, 2015. **4.** J. Rubin and I. Watson, "Computer poker: A review Artificial Intelligence, vol. 175, no. 5-6, pp. 958-987, 2011. **5.** D. Silver, A. Huang, C. Maddison, A. Guez, L. Sifre et al, "Mastering the game of Go with deep neural networks and tree search Nature, vol. 529, no. 7587, pp. 484-489, 2016.

Bidyuk P.I., Vertiletskij V.G.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Decision support system for adaptive forecasting of financial and economic processes

Introduction. Today exists a growing necessity for development of modern decision support systems (DSS) that would help decision makers (DM) in processing data and expert estimates, adequate model constructing, estimating of quality forecasts for a given horizon and generating decision alternatives.

Here DSS provides wide possibilities for computing high quality forecasts with the most different techniques and combine the estimates generated by different methods.

Another important task that could be performed is identification and taking into consideration possible uncertainties, creating difficulties with mathematical modeling of selected processes and computing forecasts as well as generating alternative decisions.

Thus, modern DSS are considered as highly developed (user friendly) computer systems that are capable to perform multiple complicated computations directed towards providing maximum help for a DM.

Also adaptive features are usually implemented in such system aiming to constructing model with adaptive structures and parameters.

This way we can maintain model adequacy in conditions of non-stationary process characteristics.

Fighting the uncertainties. Among possible uncertainties taking place in modeling and forecasting the following ones should be mentioned: – data uncertainties; – model structure and parameters uncertainties; and uncertainties related to computational process, for example, systematic computing and methodology errors.

The data uncertainties are caused by the following reasons: incomplete measurements, short samples and missing data; influence of external random disturbances on functioning of a system under study; and measurement errors. Certainly there is an appropriate statistical instrumentation for 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.

A negative influence of external random disturbances and measurement errors could be partially eliminated by application of Kalman filter, elliptical filtering, exponential smoothing or other appropriate method.

The Kalman filtering techniques are among the most often applicable due to availability of different useful possibilities of the filter 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 usually is refined thanks to implementation of appropriate adaptation scheme, application of 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 parameter estimates, say they may lead to higher variance and/or to bias of the estimates.

To reduce influence of the uncertainties we introduce adaptive estimation schemes for both model structure and parameters.

Also 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 then we apply appropriate Markov chain Monte Carlo algorithm adjusted to specific distribution of statistical data [1].

DSS architecture and functional layout. The DSS proposed has an architecture consisting of 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 includes the following computational operations: preliminary data processing; statistical analysis of data aiming to preparing the data for model structure and parameters estimation, computing of short- and medium-term of forecasts using constructed forecasting functions, and quality analysis subsystem that is based on three sets of statistical quality criteria.

The criteria help to estimate quality of data, model adequacy, and quality of forecasts. In each case we hire separate sets of statistical parameters (statistics) directed towards solving specific monitoring problems for the computational processes [2,3].

Such systemic approach is directed towards quality improvement for all computations in the system. We also take into consideration the necessity of hiring different kinds of criteria when linear or non-linear models are analyzed.

For example, non-linear classification models require for application index Gini, common accuracy, area under curve (AUC) etc.

Modeling and forecasting techniques. 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, and the model could have a very high dimensionality, say, several thousand variables.

The theory of Bayesian networks is quickly developing these days. Another possibility for solving the forecasting problem provide such methods as Kalman filtering techniques, hierarchical models, nonparametric and Bayesian regression

The processes studied and results. Among the processes studied are the following: Gross Domestic Product (GDP) for Ukraine, USA and some other selected countries, consumer and industrial price indexes, human development index, stock prices, and many others.

To select the best models the following statistics were hired: determination coefficient, sum of squared residuals, and Durbin-Watson statistics.

The tests for stationarity and nonlinearity were applied for identifying time series structures.

The best forecasts are selected with mean absolute percentage error (MAPE), mean squared errors, and Theil coefficient [3,4].

In a case of studying nonlinear processes some other criteria were applied. The results achieved for short- and medium-term forecasting are quite acceptable or high quality, say with MAPE changing in the range between 0.5% and 7.5%.

Future developments and studies. 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 other specific uncertainties. We are also aiming to automating all basic functions of the DSS.

References. **1.** Murray I. Advances in Markov chain Monte Carlo methods. – London: University of London, 2007. – 176 p. **2.** Bidyuk P.I., Gozhyj O.P., Korshevnyuk L.O. Design of Decision Support Systems. – Mykolaiv: Black Sea State University named after Petro Mohyla, 2012. – 380 p. . **3.** Tsay R.S. Analysis of Financial Time Series. – Hoboken (New Jersey): John Wiley & Sons Inc., 2010. – 715 p. **4.** Aralbayev T.Z., Aralbayeva G.G. Information Systems in Economy (reading and instructions for students). – Orenburg: Orenburg State Unversity, 2004. – 43 p.

Levin D.K.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Neo-fuzzy neural networks in forecasting stock prices

Stock price forecasting. The widely adopted efficient-market hypothesis states, that the asset price fully reflects all available information. This and the assumption that new information on the market is independent of asset's past values implies, that stock prices follow the random walk model. The famous example of the random walk process is Gaussian random walk:

$$X_t = X_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma^2) \tag{1}$$

The most important implication for forecasting is that this time series increments $\delta_t = X_t - X_{t-1} = \eta_t$ follow the i.i.d condition and have zero mean, therefore the best forecast of the increments is $\widehat{\delta_t} = 0$. As a result, the best forecast of the price in this random walk model is $\widehat{X_t} = X_{t-1}$.

But in [1] authors argue the random walk hypothesis and empirically show presence of a low short-term autocorrelation of 0.05 in stock price series. Because of this forecasting models performance needs to be compared relative to the random walk model.

Neo-fuzzy neural networks. A neo-fuzzy neuron (NFN) model first was proposed in [2]. It is described by the function:

$$y = \sum_{i=1}^{n} f(x_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \mu_{ij}(x_i) \tag{2}$$

Here x_i are input signals; $f(x_i)$ is a non-linear synapse on signal x_i; $\mu_{ij}(x_i)$ are complementary triangular membership functions; w_{ij} are tunable weights.

As follows from (3), NFN cannot accurately restore functions with multiplicative relations between input variables. This makes it necessary to build cascades and multilayer networks from neo-fuzzy neurons [3] to address this problem, which dramatically increases the number of tunable weights and therefore the complexity of a learning process.

To address the aforementioned problem and keep the single NFN unit as flexible, as possible and thus allow to create smaller networks, a new neo-fuzzy neuron model is proposed:

$$y = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \sum_{j=1}^{m} w_{ij} \mu_{ij}(x_i) \tag{3}$$

This multiplicative NFN (mNFN) is capable of approximating larger class of functions and allows to create simpler forecasting models with the same accuracy compared to original NFN, which is shown by experiments.

Experimental result were achieved on the data over the last year on NASDAQ exchange. In the Table 1 model sizes (number of rules m) are shown. All calculations were done for $n = 7$; absolute percentage error (APE) is computed by cross-validation with 10 folds relative to the random walk model prediction and kept equal for both models. Multiplicative NFN achieved the same accuracy as original NFN but with much smaller number of weights.

Table 1. Experimental results

Asset	NFN	mNFN	APE
GOOG	70	20	+2.7%
MSFT	50	12	+3.2%
AAPL	70	25	+2.8%

References. **1.** A. W. Lo, A. C. MacKinlay "Stock market prices do not follow random walks: evidence from a simple specification test". Review of Financial Studies, 1988, pp. 41–66. **2.** T. Yamakawa, E. Uchino et al. A neo fuzzy neuron and its applications to system identification and predictions to system behavior, Proceedings of the International Conference on Fuzzy Logic and Neural Networks, vol. 1, 1992, pp. 477–484. **3.** Ye. Bodyanskiy, Yu. Zaychenko, et al. The neo-fuzzy neural network structure optimization using the GMDH for solving forecasting and classification problems, Proceedings of the IWIM workshow, 2009, pp. 77-89.

Markova V.M.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Application of Neural Networks for E-shop Goods Demand Forecasting

Motivation. Nowadays there are lots of e-shops. Each entrepreneur wants to make his e-shop the most profitable. The forecasted demand can be used to solve different problems, such as "How much do we need spend for different adverts?", "What price is better to put in the next period?" or "How much do we need for the next period to minimize "frozen" money?". The work is dedicated to the last one and solves it by using neural networks. The problem of "frozen money" means minimizing cash that is not used in specified period. The forecast demand will get an information to understand how does the cash flow changes and how much cash the entrepreneur need to have in each time period to make purchases from suppliers. There is useless to have so more much money that we really need because of money valuable downtime.

Goal. To forecast demand for getting an information about cash quantity which is better to have for each time period using Multilayer Back Propagation Neural Network.

Methods. Application is written on a Java programming language. Neuroph Framework is used to design Multilayer Back Propagation Neural Network.

Description. The goods demand time series will be investigated and, basing on these results, determined effective structure of forecasted models. According to these models structures of Multilayer Back Propagation and Radial basis neural networks will be created and their comparative analysis with AR, MA, ARMA and ARIMA models will be processed. At first stage investigation will be reviewed just 3 exactly different goods from 1 category. Each good is studying by using appropriate Neural Network. Discovered Back Propagation Neural Network has the next structure. For input use old prices for chosen period (it depends of model) is used. Network weights are adjusting by using information about last week's cash flow. Each Network training will minimize difference between guess output value and desired output value according to information about real last week cash flow given from a practice. It's computing a steering force as the difference between our desired velocity and our current velocity. The current velocity acts as a guess and the error (the steering force) tells how to adjust the velocity in the right direction. Adjusting the vehicle's velocity to follow a target is just like adjusting the weights of a neural network to arrive at the right answer.In the case of the perceptron, the output has only two possible values: +1 or -1. This means there are only three possible errors. If the perceptron guesses the correct answer, then the guess equals the desired output and the error is 0. If the correct answer is -1 and we've guessed +1, then the error is -2. If the correct answer is +1 and we've guessed -1, then the error is +2.At first Network will be trained using real data that we have. After few trainings it will have enough information to forecast new values and then work independent on its own.

Summary. Application is developing for real existing e-shop, which has a 5 years trade experience and now wants to improve trading system with artificial intelligence in a goal of increasing their gain. So the problem is solving on real data basis. Now it's on developing stage, but predictably in a few months the result will be.

References. **1.** Methods of Back Propagation Trainee - http://robocraft.ru/blog/algorithm/560.html. **2.** Hajkin S. Nejronnye seti: Polnyj kurs - //Neural Networks. **3.** Quick explanation of the backpropagation algorithm, http://www.tek271.com/documents/others/into-to-neural-networks. **4.** A Gentle Introduction to Backpropagation - An intuitive tutorial by Shashi Sathyanarayana The article contains pseudocode ("Training Wheels for Training Neural Networks") for implementing the algorithm., http://numericinsight.com/uploads/A_Gentle_Introduction_to_Backpropagation.pdf. **5.** Bishop, C. M. (1995) Neural Networks for Pattern Recognition, Oxford University Press. **6.** Weitzenfeld, A., Arbib M. A., Alexander A.(2002)The Neural Simulation Language: A System for Brain Modeling, The MIT Press, https://mitpress.mit.edu. **7.** Shiffman D., (2002) The Nature of code, http://natureofcode.com/book/chapter-10-neural-networks/.

Reznik I.M., Shatalov V.M.
National Technical University of Ukraine "Kiev Polytechnic Institute", Slavutych branch, Ukraine

Data mining of Wikipedia reveals historical changes of human priorities

Wikipedia, without exaggeration, can be considered "contemporary Bible". However, unlike the Bible, it is written simultaneously and independently by many people. This process is not controlled by anyone and some features of knowledge and facts, imprinted in Wikipedia, are objective evidences of the history of humans. The opinion of the authors of biographical pages has no effect on dates of birth and death of famous people. Thus, these data form the reliable statistical dataset for quantitative studies of the history [1]. This paper data mines over 600000 Wikipedia's biographies, focusing on individuals' longevity, profession and activity years. The number of cross-citations Wikipedia pages looks like an independent "expert" assessment of the significance of the cited person.

The most cited person of the 1st-century is Jesus, 16th-century – William Shakespeare, 18th-century – George Washington, 19th-century – Abraham Lincoln, 20th-century – George W. Bush, 21st-century – Roger Federer.

Some historical cataclysm dates may be detected in peculiarities of lifespan distribution. For example, the dips in the lifespan (Fig. 2, top) around 1915 and 1945 are obviously caused by the World War I, the 1918 flu pandemic, and the World War II. The less intense dip at 1865 possibly corresponds to the American Civil War (but also to the Austrian-Prussian and the Prussian-French wars around the same time); the weak dip at 1969 (it is also statistically significant) — to the Vietnam War. We guess the less pronounced features might also be related to wars: at 1810 — the Napoleonic Wars, at 1795 — the French Revolution and so on. The top and bottom (standard error of the mean value) parts of Fig. 2 demonstrate strong anti-correlation. To study this, mortality rates were calculated for the two war periods and the two relatively peaceful ones preceding them. The wartime histograms appear to be significantly distorted and about two times wider than "peaceful" ones. Broadening of the distribution leads to an increase in the standard deviation that appears as a peak of statistical error.

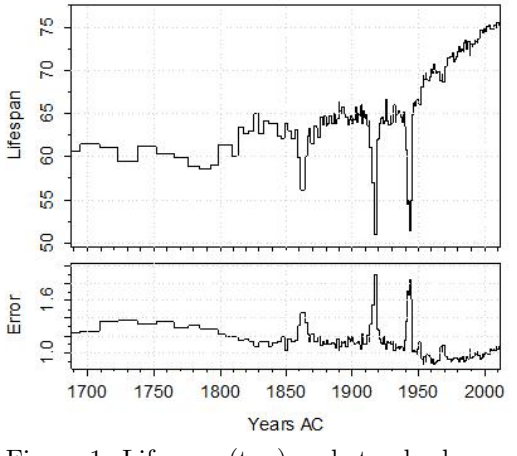


Figure 1. Lifespan (top) and standard error of the mean value (bottom) in years

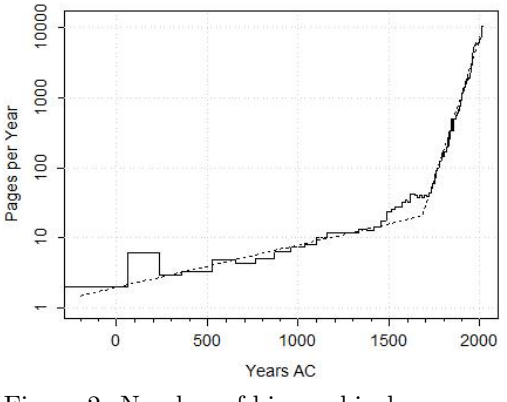


Figure 2. Number of biographical pages per year (solid line) and its trends (dashes lines)

New mass media generations cause changes in trends of number of biographies with time. A year of activity was assigned to every biographical page. Fig. 3 presents a histogram of the number

of biographical pages versus the year of activity. The shape of the resulting distribution is quite surprising. It has a broken linear trend that shows a sudden change of the slope near 1700 AC. Such a behavior was associated with the advent of the newspaper era. The hump after 1500 AC was considered to correspond to the advent of the book printing era that took place in the 15th century.

Areas of personal activities (and, accordingly, of human culture) were categorized in just a few broad and easily understandable terms. A keyword classification was used to plot time dependence of categories that reveals an evolution of human priorities (Fig. 3). Thus, sport rises only in the 20th century (particularly from the 1990s), politics surpassed religion in the 13th century, until it was surpassed by sport, and so on. Then we divided the biographies into public (politicians, businessmen, religion) and private (artists, scientists and sportspeople) and noted that it was only in the last few decades that the second group started to significantly outnumber the first; we conclude that this represents a major shift in societal values, which we refer to as "hidden revolution in human priorities".

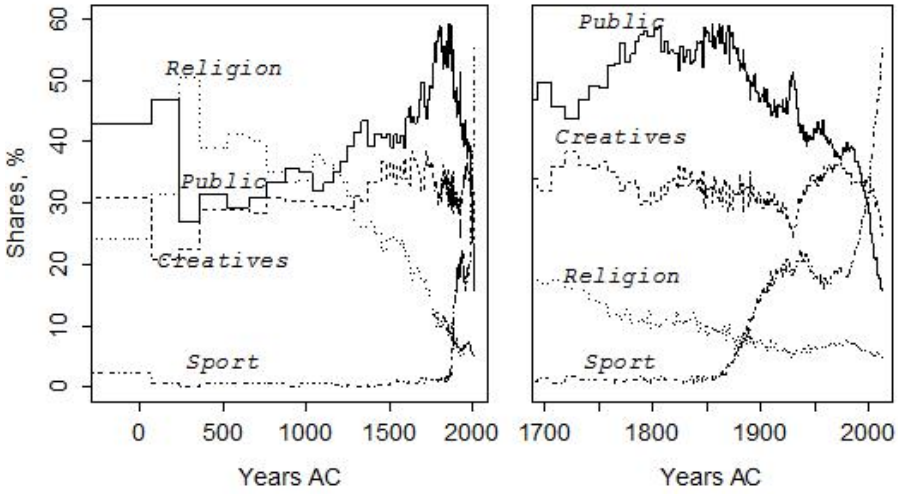


Figure 3. Shares of main groups of categories from 200 BC to 2000 AC (left) and detailed results for 1700-2000 AC (right)

A new index of human priorities, namely the Personal to Public Ratio, was introduced. Being almost a constant over centuries it increases eight times in the last decades due to the Sport and Art group growth. Now the majority of Wikipedia biographies are about sportspeople (half of them soccer players), followed by artists and politicians.

Thus, Wikipedia biographical pages can be used as a unique self-consistent source of information for studies in historical sociology. Wikipedia data mining may reveal changes over time in the human perception of the world, and may also serve as an independent reliable quantitative method of investigation of historical events.

References. **1.** Reznik I.M., Shatalov V.M. Hidden revolution of human priorities: An analysis of biographical data from Wikipedia. *Journal of Informetrics* 10(1): 124–131. doi:10.1016/j.joi.2015.12.002. ISSN 1751-1577.

Romanko O.R., Radchenko O.Yu., Livinskiy O.V.
Institute of applied system analisys NTUU "KPI", Kiev, Ukraine

Fractal nature of financial time series: R/S analysis, V-statistic and predictive modelling

Fractal Time Series Analysis - a new methodology using modern vision of the financial market and fractal mathematics tools for analyzing and forecasting financial and economic processes. There are descriptions and algorithms of a new approach to the analysis of time series - methodology of R/S analysis and V-statistic.

Misconseptions. The emergence of a new fractal approach is due to the discovery of some of the misconceptions about the nature and behavior of many financial processes. There is a widespread use of the fundamental theorem of statistics - the central limit theorem. This theorem states that the sum of a large number of independent identically distributed random variables has a normal (Gaussian) distribution. Based on the assumption of a normal process, they often use the well-known three sigma rule – with a probability of 99.7% the values of the random variable fall into the gap defined by three standard deviations. But usually time values of a financial process are not independent. Therefore, the central theorem may fail. The phenomena of "heavy"tails and high peaks are quite common, but there is no holistic approach to the study of processes with similar properties. Heavy tails are often caused by values of processes, that are not independent and identically distributed(IID). Therefore, there are different dependencies between values of the process that are separated by the time. And there is a big question, how strong is correlation between them, does it decreases slightly with time and what is the amount of values, that will significantly influence future ones.

Memory of the process. The proposed approach allows, most importantly, measuring such important component of the process, as memory. The Hurst exponent, named in honor of the discoverer of the phenomenon of memory in the study of the indicators of the Nile inundation, gives us information about the time dependence and tendencies of change in the process. When $H \in (0.5, 1]$ the process is persistent, it maintains the existing tendency, i.e. the increase in the past are more likely to lead to an increase in the future, and vice versa. When $H \in [0, 0.5)$ the process is called antipersistent, any tendency tends to be replaced by its counterpart. Process is making fluctuations in some corridor, formed with maximum and minimum values. If $H = 0.5$ it is a completely random process, there is no clear trends. It is worth noting that natural processes have $H = 0.72 - 0.73$. H. E. Hurst discovered a nonparametric methodology that allows us to distinguish between random and nonrandom systems, permanence of trends and duration of cycles if there are any [1]. The idea of the method is that "standardized range"is calculated, which shows the distance S by which to move the system in a discrete time n. It is known that random walk (e.g. random Brownian motion of molecules) passes the distance $S = cn^{0.5}$ where c is an arbitrary constant) in the time n. Hurst generalized this equation for systems that are not «IID»: $s = cn^H$,where $H \in (0, 1)$. Therefore, we can perform a simple linear regression on $log(n)$ as an independent variable and on $log(R/S)$ as a dependent variable.

Conclusion. There are two types of the memory of processes – "long" and "short". In terms of statistics, there are processes with correlation between distant values of time series and processes with correlation between distant values at the level of zero [2]. Such Models as fBm and ARFIMA are used to model processes with long memory. ARIMA model is often used to describe processes with short memory. Measuring the value of memory is highly required to build a correct model and to make accurate predictions.

Литература. **1.** Peters, Edgar E. Fractal market analysis: applying chaos theory to investment and economics. Vol. 24. John Wiley & Sons, 1994. **2.** The forecast of the time series by approximating the fractal Brownian motion / Bondarenko V.V. // System Research and Information Technologies. — 2013. — № 4.

Shaptala R.V., Kyselova A.G.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Location as a context parameter in energy management systems

Introduction. In recent years, energy requirement has been growing and context-aware energy management systems have been refined to suit this demand. Prediction of context parameters is a vital task for increasing the productivity of such systems.

Context-aware energy management system. Context-aware energy management system(CAEMS) is a smart information system, that combines the data from diverse heterogeneous sources like alternative power sources, sensors, which are defined by different kinds of physical data and loads, that solve crucial user problems in several zones or areas (structures or groups of structures). The fundamental aim of the CAEMS is to introduce a control system which can foresee and adapt intelligently to the activity of every connected electrical equipment in the cooperative information setting. The integration of such a system allows to manage energy resources taking into account user requirements [1].

Indoor location prediction. Context parameters are management system factors, influencing the decision-making [2]. Such parameters include location, time, environment, infrastructure and user information. Location prediction tries to deduce spatial context by looking at the context history, a list of contextual data points which tell the system how users are changing their positions around a definite environment. The contextual information is given by several kinds of sensors like GPS, RFID, or wireless equipment that can provide data about users' locations or changes in physical status. As the context-aware energy management system becomes more and more aware of user location, it can predict the parameter to adopt the most effective strategies to successfully work on the given objectives such as user comfort and energy efficiency. Indoor location prediction works only within a single structural unit such as a building or an office. Given some amount of previous visits data points, the goal of indoor location prediction is to infer the most likely location for the next point in time. The classic testing ground for this problem is Augsburg Indoor Location Tracking Benchmark [3], which is also used in our research.

Algorithms and evaluation results. We tested some of the most popular and proven location prediction techniques: Neural Networks, Naive Bayes(NB), Decision Trees, Support Vector Machines(SVM), Hidden Markov Models, Random Forests(RF), Gradient Boosting, Ensemble of several models. The evaluation procedure implied 10-fold cross-validation and was rerun in 2 different settings: using a single previous point, using 4 previous points for predicting the next. Moreover there are 4 different behavior settings in the dataset corresponding to 4 different people. The results reveal that the installation of additional previous location data points increases the accuracy of every model except neural networks which seem to overfit to the training data. The most accurate standalone models are Support Vector Machines and Random Forests, which outperform every other model. However voting procedure between SVM, RF and NB improved their prediction capabilities in several testing circumstances, which implies that further parameter tuning can be beneficial.

Conclusion. SVMs and RFs introduced themselves as the most accurate models for indoor location prediction. Such algorithms can be profitably implemented as a component of context-aware energy management systems.

References. **1.** Zhuikov, V.; Kyselova, A., "Integration of context-aware control system in microgrid,"in Electronics and Nanotechnology (ELNANO), 2013 IEEE XXXIII International Scientific Conference , vol., no., pp.386-390, 16-19 April 2013. **2.** Hong, J., Suh, E., Kim, S.: Context-aware Systems: A Literature Review and Classification. Expert Systems with Applications 36(4), 8509–8522 (2008). **3.** J. Petzold. Augsburg Indoor Location Tracking Benchmarks. Technical Report 2004-9, Institute of Computer Science, University of Augsburg, Germany, February 2004.

Zaychenko Yu.P., Guskova V.G.
Institute for applied system analysis, Kyiv, Ukraine

Application of fuzzy neural network nefclass for recognition of medical images in diagnostics

An important application sphere of pattern recognition systems is the problem of classification of optical medical images and diagnostics in medicine.

One of such tasks is cervix epithelium state analysis and diagnostics using optical images obtained with colposcope (a method of survey of a mucous membrane of part of a neck of a uterus in the conditions of additional lighting and optical increase with the help of a colposcope). The problem of classification cervix epithelium state using images obtained with colposcope was considered in [1] where for its solution was suggested the application of crisp neural networks Back propagation, neural networks with radial basis functions (RBFNN) and cascade RBFNN.

The goal of this paper is the investigation of fuzzy neural network NEFClass [2] for recognition of state of cervix epithelium in medical diagnostics and comparison of its efficiency with conventional RBF network.

Stages of cervix epithelium states recognition.

1. Work with data. Construct a database of examples, characteristic for this task. Split all data set into two sets: training and test in the following ratio: training 50%, test 50%;training 60%, test 40%; training 70%, test 30%; training 80%, test 20%; training 90%, test 10%.
2. Preliminary processing. Choose system of features, characteristic for this task, and transform data appropriately that is to be fed into network inputs. As input data for medical images of 5 benign processes, are used namely: inflammatory processes in the form of branching of vessels; cervical erosion; traumatic deformation; large cervical ectropion, small cervical ectropion.
3. Designing, training and assessment of a network work quality. At this stage the number of rules, quantity of fuzzy sets and percentage ratio of training and testing samples are determined.
4. Choosing algorithm of a network training. As a training algorithm the gradient method was used. At this stage it is necessary to specify the accuracy, the steps size for all variables and a number of iterations.
5. Application and diagnosing. At the last stage we receive result of application of the neural NefClass network to a problem of medical diagnostics. The amount of misclassifications and an average error on sample are determined. Sample size is 70 elements.

The experimental results of classification after training are presented in the table1.

The next step in experiments was calculating of results change due to variation of the rules number. For each number of fuzzy sets (3, 6, 7, and 11) training/test sample ratio was used. It should be noted there is a number of rules, after which there is no change in the classification of samples and in the mean square error.

The work of fuzzy neural network NefClass was compared with the work of the neural network RBF. The results of RBF are shown in Table 2.

Presented curves indicate that the fuzzy neural network NefClass shows the better results than non-fuzzy RBF network.

Conclusions.

1. The problem of recognition of objects on medical images in medical diagnostics is considered. The investigations were performed on the cervix uterus images obtained using colposcope. 70

Table 1. Performance results of NEFClass

# of sets	Ratio training/ test sample %	misclassified patterns training	misclassified patterns testing	MSE training	MSE testing	% Misclassification
3	50-50	13	16	0,559	0,588	45.12%
	60-40	17	12	0,588	0,551	42.1%
	70-30	20	9	0,570	0,549	42.85%
	80-20	23	6	0,562	0,540	42.857%
	90-10	25	4	0,551	0,562	57.14%
6	50-50	2	9	0,1697	0,336	25.71%
	60-40	2	9	0,1699	0,330	32.14%
	70-30	3	6	0,167	0,306	28.57%
	80-20	2	2	0,1495	0,254	14.28%
	90-10	2	0	0,154	0,197	0%
11	50-50	3	11	0,091	0,440	31.42%
	60-40	1	7	0,055	0,466	25%
	70-30	1	8	0,0434	0,550	38.09%
	80-20	2	4	0,054	0,377	28.57%
	90-10	1	1	0,064	0,221	14.28%

Table 2. Performance results of RBF network

	50-50	60-40	70-30	80-20	90-10
number of coincidences	20	16	17	10	6
number of non coincidences	15	12	3	4	1
% Misclassification	42.9%	42.9%	19%	28.6%	14.3%

images were selected which contained 5 classifications of diseases.

2. Fuzzy neural network NefClass and non fuzzy neural network RBF were used for classification. Experiments were carried out on training / test samples in the ratios: 50/50, 60/40, 70/30, 80/20 and 90/10. In process of experiment with NefClass number of fuzzy sets varied 3, 6, 7 and 11, the number of rules – 50, for each sample the value of MSE (training and testing) was calculated. The best result was obtained for samples ratio 90-10.
3. The experiments with non-fuzzy RBF neural network had shown the best result was obtained for training/test sample ratio 90-10, with an error of classification 14.3%. The results of the fuzzy neural network proved to be much better than the RBFN.

References. **1.** K.Malyshevska The analysis of neural networks' performance for medical image classification / K. Malyshevska // International Journal "Information Content and Processing Volume 1, Number 2, 2014. – C.194-199.

2. М.З. Згуровский, Ю.П. Зайченко. Основы вычислительного интеллекта. - Киев: Изд. Наукова Думка, 2013.- 406 стр.

Zaychenko Yu.P., Ovi Nafas Aghaei Agh Ghamish
NTUU "KPI", Kyiv, Ukraine

Bankruptcy Risk Forecasting under Uncertainty Application of Fuzzy Neural Network NEFClass

The problem of classifying data is currently one of the most pressing areas of applications of artificial intelligence . To solve this problem have been proposed various approaches and directions, including the most popular acquired solutions that combine neural networks and fuzzy inference system.

The task of financial analysis and forecasting of risk of bankruptcy of corporations is one of the most important. The timely identification of signs of possible bankruptcy allows corporate management to make operational decisions to improve the financial condition and prevent a possible bankruptcy.

Traditionally , this problem was solved with the use of classical methods of multivariate discriminant analysis (MDA),beginning with the pioneering work of Professor E. Altman in 1968. Further, this method was developed by Lis, Davydova - Belikova and other authors.

A specificity of this problem in terms of Ukraine's economy is the presence of considerable uncertainty caused by the desire of corporations top managers to hide the true state of sponsored enterprises to obtain loans.

This calls for the development and application of new methods and algorithms, allowing to work in a fuzzy and uncertain information. These methods are based on a system of fuzzy logic and fuzzy neural networks.

In our previous studies to the problem of predicting the risk of bankruptcy have been proposed and investigated fuzzy neural network (FNN) and Mamdani Tsukamoto, as well as cascading neo-fuzzy neural networks.

The experimental studies of their effectiveness in relation to the economy of Ukraine and the comparative analysis with classical methods were performed.

The aim of this work is the development of the proposed methods and research of FNN NEFClass for problem of predicting the risk of corporations bankruptcy.

The structure of FNN NEFClass-M is presented and the learning algorithms for fuzzy rules base and membership functions are described in the paper.

The experimental investigations of FNN NEFClass for bankruptcy risk forecasting for Ukrainian corporations were carried out and their efficiency was estimated and compared with fuzzy neural networks with algorithms of Mamdani, Tsukamoto and classical statistical methods.

Алексеєв М.О., Безлюднов Д.Є.
Інститут телекомунікаційних систем НТУУ "КПІ", Київ, Україна

Визначення аномальних станів у роботі пристроїв IoT

Вступ. Інтернет речей - концепція комунікаційної мережі фізичних або віртуальних об'єктів, які мають технології для взаємодії між собою та з оточуючим середовищем, а також можуть виконувати певні дії без втручання людини. Дана концепція дозволяє змінити якість життя.

Кількість розумних пристроїв та протоколів їх взаємодії збільшується щодня, але всі ці пристрої працюють на різних платформах, мають різні служби та взаємодіють на основі різних протоколів та ін. Відсутні чітко прописані стандарти у цій сфері, що є причиною безлічі проблем, пов'язаних із сумісністю та коректною взаємодією даних пристроїв. Існує безліч пристроїв, які через відсутність стандартів та наявності великої конкуренції у цій сфері взаємодіють лише з такими самими пристроями або пристроями, що виготовлені тим самим виробником.

Різноманітність протоколів зв'язку, невеликі потужності у каналі зв'зку, вплив та нестабільність навколишнього середовища ускладнюють можливість детектування позаштатних ситуацій в режимах роботи пристроїв IoT. Також, негативно впливають на можливість детектування позаштатних ситуацій, великі об'єми даних та високі швидкості надходження нових даних.

Проблема. Існує потік даних X(t), заданих на дискретній часовій осі. Даний потік даних є стаціонарним (або квазістаціонарним). В кожен часовий інтервал tk кожна частина системи генерує вектор даних vi , де i=1..N. Аномалією в наборі даних називається дані, які не сумісні із іншою частиною набору даних. Метою є знайти викиди O ⊂ X в наборі даних.

Існуючі методи пошуку аномальних станів роботи пристроїв. Існуючі методи визначення аномальних станів технічних систем є малоефективними для IoT. Розглянемо основні існуючи методи пошуку аномалій у сенсорних мережах (табл. 1).

Табл. 1. Порівняльна таблиця методів детектування аномалій

Назва	Режим реального часу	Працює із будь-якими даними	Складність алгоритму	Примітка
SSA	+	-	+	не адаптований під велику кількість даних
авто-регресія	+	-	-	низька точність
метод Фур'є	-	-	+	
параметрична регресія	+	-	+	низька точність
лінійні фільтри	+	-	+	не визначає деякі види аномалій
модифікований SSA [1]	+	-	+	не адаптований під вимоги IoT
Кластерний алгоритм [2]	+	+	+	висока алгоритмічна складність (O(N2))

Основними недоліками існуючих методів по відношенню до існуючих є:

- Комплексна складність алгоритмів (алгоритмічна складність алгоритму; складність реалізації алгоритмів; складній математичний апарат, який складно реалізувати на пристроях IoT);
- Орієнтованість тільки на числові данні (всі алгоритми не можуть працювати із текстовими даними, або із многовимірними даними);
- Не можуть працювати із потоком даних (частина алгоритмів є offline'овими, тобто тими,

що працюють із готовим набором даних, та не можуть працювати із потоками даних);
- Недостатня точність (алгоритми, такі, як авто-регрессія та лінійні фильтри, не працюють із аномаліями середнього значення та можуть помилково детектувати аномалії за їх відсутності [1]).

Розроблений метод. На рис. 1 зображена нейронна мережа Елмана із блоком пам'яті [4], яка є перцептроном із оберненім зв'зком. Особливостями слоїв даної модифікованої мережі Елмана є:

- На вхідному рівені перцептрону, кількість нейронів рівна розмірності вхідних даних (N);
- На проміжному рівені перцептрону, кількість нейронів не менша за 2N (N входів та N обернених зв'язків);
- На вихідному рівень нейронної мережі, кількість нейронів може змінюватись від 1 до N;
- Рівень керування - допоміжний прихований рівень нейронів. Кількість нейронів рівна N;
- Блок пам'яті в найпростішому рішенні являє собую чергу довжиною T.

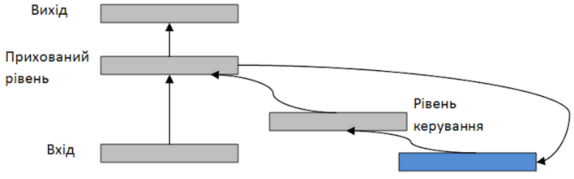


Рис. 1. Структурна схема нейронної мережі Елмана із блоком пам'яті

Додавання блока пам'яті дає можливість детектувати зміни у порівнянні із попереднім значенням. Для цього раціонально використати від'ємні значення вагів на входах нейронах із рівня керування. Від'ємні значення, затримані на T, підсумовуючись із виходами із вхідного слою дозволяють отримати дельту змін у потоці даних. В якості передаточних функцій використовуються сигмодальні функції, які дозволяють перейти від бінарніх рішень до більш гнучких - відсоткових. Дані від передаточних функцій на вихідному рівні, які лежать у межах від 0 до 1 включно, накладені на певну шкалу можуть визначити із заданою достовірністю, наявність аномалій, змін у потоці даних або їх відсутність. Запропонований метод детектування аномальних режимів роботи пристроїв IoT вирішує подані проблеми наступним шляхом:

- Режим роботи у реальному часі витікає у постійному відстежуванні надходжуванних даних, та зберіганні невеликої кількості "історії" даних;
- Метод працює із будь-якими даними через те, що для порівняння використовуються функції предикати, які можна визначити майже на будь-яких даних (числові, знакові, текстові, багатовимірні).
- Метод базується на нейронній мережі Елмана, яка складається із порогових суматорів, зв'язків та блоку пам'яті, як зображено на рис. 1, які являють собою найпростіші математичні операції, які можна виконувати над будь-якими даними.

Висновки. Проблема пошуку аномальних станів пристроїв досліджується вже давно. Однак розроблені методи погано застосовуються у сфері IoT через велику кількість даних різних типів. Розроблений метод детектування аномальних станів пристроїв IoT є ефективнішим у сфері IoT, ніж інші існуючи через його застосовуваність для будь-яких даних, режим роботи в реальному часі та простішу реалізацію алгоритму, ніж у SSA.

Література. **1.** Y. Yaoa et al., Online Anomaly Detection for Sensor Systems: a Simple and Efficient Approach, Performance Evaluation, vol 00, pp. 1-24, 2010. **2.** S. Rajasegarar et al., "Distributed anomaly detection in wireless sensor networks in conf., ICCS, Florida, USA 2006. **3.** B. Abhishek et al., Sensor Faults: Detection Methods and Prevalence in Real-World Datasetss, Trans. on Sensor Networks 6(3), 2010. **4.** John A. (2015) Recurrent Neural Networks [Online]. Available: https://en.wikipedia.org/wiki/Recurrent_neural_network.

Аникин В.Ю.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Комплексная оценка риска банкротства путем сравнительного анализа с придприятиями-конкурентами

Понятие банкротства характеризует несостоятельность предприятия удовлетворить требованиям кредиторов относительно оплаты товаров, услуг, а также обеспечить обязательные платежи в бюджет и внебюджетные фонды. [1] Закон Украины "О возобновлении платежеспособности должника или признания его банкротом"под банкротством понимает признанную судом несостоятельность должника возобновить свою платежеспособность и удовлетворить признанные судом требования кредиторов, не иначе как через применение ликвидационной процедуры. [2] Задача определения степени риска банкротства актуальна как для собственников предприятия, так и для его кредиторов.

Большинство методов предсказания банкротства дают точечную оценку устойчивости предприятия в разрезе финансовых показателей самого предприятия. Однако, устойчивость не является единственным возможным критерием оценки деятельности фирмы. Инвесторы заинтересованы не только в бесконечно долгом существовании фирмы, но и в получении достаточно весомой прибыли. Помимо состояние фирмы на настоящий момент, важными являются характеристики существующих тенденций. [3] Более точно сильные и слабые стороны предприятия можно оценить в разрезе набора подобных ему предприятий-конкурентов. При такой формулировке проблемы целесообразно делать комплексную оценку финансового состояния предприятия путем систематического сравнения его показателей с показателями остальных участников данного рынка.

В данной работе для этой цели была использована самоорганизующаяся карта Кохонена, которая позволяет распознать кластеры данных. В качестве входных данных были испольованы финансовые показатели из годовых балансов и отчетов о прибыли 80 банков Украины за 2013 и 2014 годы. Каждый банк описывался 20 финансовыми показателями – отношением балансовых статей к общей сумме активов банка. Благодаря такой нормализации были сглажены различия между крупными и мелкими банками, составляющие несколько порядков величин. Из данных 20 параметров конструировались две обобщенные координаты для каждого предприятия.

В процессе обучения сети информация о банкротстве не была учтена, а была изображена уже на готовой карте в качестве индикатора области параметров с повышенным риском банкротсва. С помощью этой особенности были выбраны области риска по отсносительно небольшому числу примеров. Используя различные финансовые показатели (такие как «Размер активов», «Итого доходов», «Всего обязательств»), был получен набор раскрасок, с помощью которого была проведена оценка положения предприятий среди конкурентов.

Также, для сравнения были проведены оценки с помощью многомерного дискриминантного анализа. Нейросетевое моделирование обеспечивает лучшую точность предсказания банкротств: порядка 95%, по сравнению с 75%-85% для использованных статистических методик.

Выводы:

1. В результате исследований были определены условия для успешной оценки риска банкротства.
2. Разработана система, позволяющая классифицировать предприятия Украины относительно риска банкротства на основе количественных показателей с точностью до 95%.

Литература. **1.** Зайченко, Ю.П. Нечеткие модели и методы в интеллектуальных системах [Текст]: учеб. для вузов / Ю.П. Зайченко. - К.: Слово, 2008. -344 с. **2.** Закон Украины "О возобновлении платежеспособности должника или признания его банкротом" // Ведомости Верховной Рады. – 2012, № 32-33, ст.413. **3.** Ежов А.А., Шумский С.А. Нейрокомпьютинг и его применения в экономике и бизнесе. М., МИФИ, 1998.

Бідюк П.І., Липчак І.Р.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Моделювання гетероскедастичних процесів та побудова функцій прогнозування на основі моделей реальних фінансово-економічних процесів

У даному дослідженні розглянуті методи, що дають можливість моделювати нестаціонарні фінансово-економічні процеси та забезпечують можливості оцінювання коротко- та середньострокових прогнозів економічно-фінансових процесів та їх волатильності (дисперсії).

На основі розроблених математичних моделей, методів оцінювання параметрів використаних рівнянь та трьох множин критеріїв реалізовано інформаційну систему підтримки прийняття рішень (ІСППР), за допомогою якої можна вибирати кращі моделі для відповідних типів нелінійних нестаціонарний процесів з мінімальними витратами обчислювальних ресурсів.

У результаті досліджено такі методи: побудова різнотипних регресійних моделей (за методикою регресійного аналізу даних), метод експоненційного згладжування, метод комбінування прогнозів та деякі інші. На основі цього дослідження здійснено порівняльний аналіз методів між собою.

Розроблену ІСППР протестовано на 3 рядах і отримані результати порівняно із аналогічними обчисленнями у системі Eviews 3.0. В результаті порівняння очевидно, що з точки зору якості обчислень реалізований програмний продукт не поступається уже існуючому комерційному.

Наведені методи практично застосовано для моделювання різних фінансово-економічних процесів, наприклад, цін на валютні пари та на акції вибраних підприємств. Критеріями якості побудованих моделей були такі: середньо-квадратична похибка, середня абсолютна похибка, середня абсолютна похибка у процентах, коефіцієнт Тейла. Отримані результати засвідчили високу якість побудованих моделей та можливість їхнього практичного застосування в рамках розробленого програмного продукту. Водночас варто зазначити, що для кожної конкретної ситуації неможливо заздалегідь передбачити, який з методів дасть найкращий прогноз.

Отже, загальний підхід до прогнозування фінансово-економічних процесів має бути комплексним та базуватися на сучасних принципах системного аналізу. При цьому побудований прогноз стає підґрунтям для прийняття відповідних управлінських рішень.

У подальших дослідженнях планується використання методів моделювання і прогнозування, які логічно доповнюють вже використані у роботі. Зокрема, будуть використані дерева рішень, байєсівські мережі (статистичні і динамічні), метод опорних векторів, метод групового урахування аргументів та інші методи інтелектуального аналізу даних.

Література. **1.** Бідюк П. І. Аналіз часових рядів (навчальний посібник) / Бідюк П. І., Романенко В. Д., Тимощук О. Л. – К.: Політехніка, 2010. – 317с. **2.** Бідюк П.І. Моделювання і прогнозування гетероскедастичних процесів [Текст] / П.І. Бідюк // Системні дослідження та інформаційні технології : Міжнародний науково-технічний журнал. - 2004. - № 1. - с. 115-134. **3.** Бідюк П.І. Проектування комп'ютерних інформаційних систем підтримки прийняття рішень / П.І. Бідюк, Л.О.Коршевнюк. – К.: - 2010 – 340 с. **4.** Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования/ Ю.П. Лукашин – М.: Финансы и статистика, 2003. – 414 с. **5.** Матвєєва Т. Ю. Введення в макроекономіку. / Т. Ю. Матвєєва — "Видавничий дім ГУ ВШЕ": 2007. — 511 с. **6.** Молчанов И. Н. Компьютерный практикум по начальному курсу эконометрики (реализация на Eviews): Практикум / Ростовский государственный экономический университет. / И. Н. Молчанов, И. А. Герасимова И. А. – Ростов-н/Д.: –2001. – 58 с. **7.** Єріна А. М. Статистичне моделювання та прогнозування: Навч. посібник. / Єріна А. М. – К.: КНЕУ, 2001. – 170с. **8.** Бідюк П.І., Зворигіна Т.А. Структурний аналіз методики побудови регресійних моделей за спостереженнями часового ряду // Управляющие системы и машины. – 2003. – № 3.

Богун К.В., Бідюк П.І.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Побудова моделі ВВП в економіці перехідного преріоду на прикладі України

Вступ. Економіка України більше двадцяти років перебуває у перехідному періоді, за цей час не були вирішенні першочергові завдання, що мали привести країну до ринкової економіки. Більшість програм, законів і проектів були непослідовними і не достатньо науково обґрунтованими. Для подальшого розвитку потрібне чітке розуміння ситуації, вміння правильно моделювати і прогнозувати процеси в економічному секторі. Це дасть можливість приймати обгрунтовані рішення.

Для переходу від командно-адміністративної системи до ринкової економіки існують такі перевірені шляхи: радикальний та еволюційний. Перший шлях передбачає одночасну лібералізацію практично усіх ринків і цін, причому зростання цін обмежене жорсткою монетарною та фіскальною політикою. Радикальним шляхом пішли багато країн Центрально-Східної Європи (класичний приклад – Польща). Другий шлях – поступове запровадження ринкових відносин упродовж 10-15 років і більше. Еволюційний шлях поступового створення ринкових інститутів випробуваний у Китаї. [1]

Існують декілька математичних моделей, які можуть застосовуватися для формального опису процесів перехідного періоду. Ми спираємося на модель переходу від економіки централізованого планування (ЦП) до ринкової економіки, де ринкова економіка розглядається у відповідності із традиційним представленням по Ремсі-Кассу (РК), а перехід до нових ринкових відносин – як перехід із сталого стану економіки централізованого планування до сталого стану економіки РК. В запропонованій моделі розглядається також вплив розбіжностей в облікових ставках держави до домашнього господарства на параметри перехідного процесу. Такий вплив називають ефектом зміни облікової ставки (або ефектом переважаючих змін). Якщо перехід відбувається успішно, то економіка повинна досягати уставленого стану, що відповідає моделі РК [2].

Для моделювання і прогнозування економіки перехідного періоду надзвичайно важливо правильно розрахувати показник ВВП, оскільки він характеризує можливості економіки щодо забезпечення добробуту суспільства в трьох аспектах: ефективності розподілу виробництва на продукти проміжного та кінцевого призначення; структури кінцевого попиту (споживання, нагромадження, зовнішньоторговельне сальдо); первинних доходів. ВВП є також основою вимірювання структури виробництва, міжнародних порівнянь якості життя, оцінки дефіциту державного бюджету, грошової емісії та ін. Головною особливістю ВВП є те, що він охоплює вартість тільки тих товарів і послуг, які вироблені на території даної країни [3].

При побудові математичної моделі у вигляді різницевих рівнянь необхідно вибрати кращу модель з моделей-кандидатів. Оскільки процес оцінювання параметрів рівняння грунтується на даних з випадковою складовою, то для підвищення надійності правильного вибору моделі необхідно використовувати, по можливості, більше статистичних параметрів, що характеризують адекватність моделі, а саме:

- коефіцієнт множинної детермінації;
- сума квадратів похибок;
- статистика Дарбіна-Уотсона.

А також параметри для оцінювання якості прогнозу:

- середня абсолютна похибка в процентах (САПП);
- коефіцієнт Тейла [4].

Для побудови моделі дані ВВП, індексу споживчих цін та грошовий агрегат були взятіз елктронного джерела Державної служби статистики України та Світового банку.

Результати моделювання на три кроки, наведені вище, свідчать про те, що всі моделі є придатними для моделювання, оскільки коефіцієнт Тейла вимірюється сотими долями.

Тип моделі	Характеристики моделі				Характеристики прогнозу		
	R^2	Akaike	$\sum e^e(k)$	DW	СКП	САПП	Коеф-т Тейла
ARMAX+ICЦ	0.861030	9.631469	7377.469	2.2419	9,803565	6,423132	0,031532
AR (6)	0.867350	9.584922	7041.939	2.114848	6,258251	4,147031	0,020394
AR (7)	0.863419	9.316714	6915.097	2.179179	9,201781	7,944902	0,038579
ARIMA(1, 3)	0.867178	8.953026	8020.643	1.956997	12,00931	1,198774	0,008543
ARIMA(1, 6)	0.966695	8.004518	2011.191	1.660721	4,1711349	2,667609	0,013199
ARIMA(3,3)	0.919992	8.928393	4761.756	2.052934	7,309842	4,885511	0,023379
ARIMA(6,1)	0.910896	9.298124	4730.235	2.074529	1.308944	0.756722	0.004171

Рис. 1. Зведені характеристики математичних моделей, побудовані для прогнозу ВВП України

Але в деяких моделях САПП більше 5, що свідчить про те, що у таких випадках потрібно покращувати структуру моделі.

Порівнюючі побудовані моделі, найкращі характеристи в моделях ARIMA(1,6) та ARIMA(3,3), так як у першої моделі кращі кофефіцієнти множинної детермінації = 0,966695,коєфіцієнтк Акайке = 8.004518 і сума квадратів похибок = 2011.191, що є у два рази краще від моделі ARIMA(3,3), але статистика Дарбіна-Уотсона рівна 1.660721 .

Найкраща модель для оцінювання прогнозу ARIMA(6,1), оскільки середня абсолютна похибка в процентах дорівнює 1,308944, а коєфіцієнт Тейла 0,004171.

Висновок. Зважаючи на кризові явища в українській економіці, побудова національних та багатонаціональних макроекономічних моделей дозволяє прогнозувати певні тенденції в економіці та досліджувати реальний стан економічного розвитку на певний момент. А прогноз показника ВВП характеризує кількісні параметри економічного розвитку країни та його складові за етапами економічного обігу, оскільки відображає вартість кінцевих результатів економічної діяльності всіх інституційних секторів та галузей економіки за певний проміжок часу. Отже, ВВП визначає ефективність економіки в даній країні та забезпечення добробуту суспільства.

Подальше удосконалення моделі ВВП можливе за рахунок введення додаткових незалежних змінних, а також застосування методів інтелектуального аналізу даних.

Література. **1.** Економіка перехідного періоду [Електронний ресурс]. – Режим доступу: http://lnu.edu.ua/faculty/ekonom/Economics/publish/Ek_Ukr_10/R_1.pdf. **2.** Бідюк П.І. Аналіз та математичне моделювання економічних процесів перехідного періоду/Бідюк П.І., Половцев О.В. –К:ПЛАБ-75, 1999. – 209с. **3.** Фатюха Н.Г., Аналіз динаміки ВВП України/ Н.Г.Фатюха, Т.П.Макушина//Ефективна економіка [Електронний ресурс]. – Режим доступу:http://www.economy.nayka.com.ua/pdf/12_2015/30.pdf **4.** 11) Бідюк П.І. Аналіз часовиз рядів/ Бідюк П.І., Романенко В.Д., Тимощук О.Л. – К: Політехніка, 2010. – 317 с.

Бойко О.М.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Моделювання кредитоспроможності фізичних осіб

Сьогодні існує множина моделей для оцінювання кредитоспроможності на основі ринкових показників, наприклад: модель Блека-Шоулза і Мертона [1, 2]. Перевагами таких моделей є висока прогнозуюча спроможність, абсорбуюча інформація про позичальника, доступна всім інвесторам, чинним на ринку. Недоліками є те, що інформація про позичальника буде показовою лише за умови ефективності ринку, також необхідно мати великий масив даних.

Підходи до розробки моделей оцінки кредитоспроможності на основі фундаментальних показників базуються на макроекономічних показників, фінансових показників та на даних рейтингових агентств. Особливості таких підходів є те, що вони враховують циклічність економіки, дають довгострокову оцінку, підходять для крос-аналізу. Основні переваги: доступність інформації, простота підрахунків і точність прогнозу. Проте присутні недоліки: важко визначити періодичність циклів економіки та оцінити ймовірність дефолту конкретного позичальника; не завжди надані дані достовірні; бухгалтерська звітність показує результати постфактум, що володіє недостатньою передбачувальною силою щодо майбутніх перспектив; переоцінка рейтингу відбувається з тимчасовим лагом.

Для мінімізації кредитного ризику потрібно мати інструмент для управління, що представляє собою процес виявлення і оцінювання ризиків, а також необхідно коректно вибрати методи для його аналізу. Традиційно кредитний ризик розглядається в розрізі кожного конкретного позичальника. Існує велика кількість моделей, що використовують складний математичний апарат для оцінки кредитного ризику. Ключовою задачею побудови математичної моделі кредитного ризику є оцінювання розподілу збитків всього агрегованого кредитного портфеля; це так звані системи скорингу [3].

На практиці банки управляють кредитними ризиками, керуючись власними методиками кредитного аналізу та відбору позичальників. Основний акцент в кредитному аналізі робиться на готовності і можливості позичальника виплатити кредит, для оцінки яких ретельно вивчається характер діяльності позичальника, його кредитна історія, поточний фінансовий стан , його можливості та потенціал.

Дослідження присвячене аналізу можливості застосування лінійної і нелінійної регресії, а також дерев рішень; виконанню та аналізу результатів обчислювальних експериментів з оцінювання кредитоспроможності клієнтів за заданими статистичними даними; порівняння результатів застосованих методів стосовно розв'язання задачі оцінювання кредитів.

Встановлено, що існує потреба у створенні нових сучасних систем для оцінювання ризиків з метою їх мінімізації та розв'язання задач ризик-менеджменту. Для створення таких систем потрібно використовувати сучасні математичні моделі, такі як байєсівські мережі, дерева рішень, лінійна та нелінійна регресія, які показали можливість досягнення точних результатів вимірювання кредитного ризику. Нелінійна регресія має значно кращі якісні показники, ніж лінійна регресія за рядом критеріїв. Встановлено, що кращі результати класифікації клієнтів у даному випадку отримано за допомогою дерев рішень. Показано, що дерева рішень дають можливість отримати прийнятний за якістю результат класифікації. Для виконання обчислювальних експериментів використано вибірку даних для 15000 клієнтів банку. Вибрано множину статистичних критеріїв для аналізу якості класифікації клієнтів.

Узагальнена порівняльна характеристика лінійної моделі та моделі логіт показана в таблиці1:

Отже, нелінійна регресія за всіма показниками якості прогнозу дає кращі результати ніж лінійна регресія. Побудована модель нелінійної регресії є кращою за показниками якості. Нижче наведено результати ризику неправильної оцінки та стандартне відхилення для трьох вибірок, використовуючи дерева рішень(таблиця 2).

Табл. 1. Результати Лінійної і нелінійної регресії

Варіант розрахунків	Коефіцієнт детермінації	Сума квадратів залишків	Статистика Дарбіна -Уотсона	Середня квадратична похибка	Середня абсолютна похибка	Коефіцієнт Тейла
Лінійна регресія						
All_1	0,004825	709,0625	1,854974	0,217419	9,4637	0,112956
BASE 743	0.384027	114.4168	0.271820	0.392419	34.0567	0.302704
NK 588	0.280168	105.6350	0.537552	0.414596	35.4870	0.288926
Нелінійна регресія						
All_1	-	708,0545	-	0,217264	9,4463	0,112876
BASE 743	-	106.1273	-	0.377937	28.4419	0.289203
NK 588	-	104.7132	-	0.410257	33.7454	0.285487

Табл. 2. Оцінка ризику та стандартна похибка дерева рішень трьох статистичних даних

Вибірки	Ризик неправильної оцінки	Стандартне відхилення
1) ALL 1	0.043	0.001
2) BASE 743	0.073	0.006
3) NK 588	0.188	0.007

Отже, побудовані дерева рішень дали досить точні результати, оскільки стандартні відхилення моделей невеликі.

Кращі результати класифікації клієнтів на дві групи отримано за допомогою дерев рішень (загальна точність склала 4,3%, що можна пояснити можливістю досягнення високої якості класифікації на дві групи за допомогою цього методу.)

Для оцінювання кредитоспроможності клієнтів фінансової установи доцільно побудувати спеціалізовану систему підтримки прийняття рішень на основі методів регресійного та інтелектуального аналізу даних. Також, для підвищення достовірності результатів, доцільно використати дані, отримані з альтернативних джерел.

Література. **1.** Бідюк П.І. Аналіз часових рядів / П.І. Бідюк. — К.: НТК ІПСА, 2006. — 188 с. **2.** Бідюк П.І. Системний підхід до прогнозування на основі моделей часових рядів / П.І. Бідюк // Системні дослідження та інформаційні технології. — 2003. — №3. — С. 88 – 110. **3.** Ковалев, В. В. Финансовый анализ: методы и процедуры: учебник. [Текст]/ В.В. Ковалев. — М.: Финансы и статистика, 2001. — С. 560.

Ганчукова Д.В., Подладчіков В.М.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Узагальнення методу апроксимації експериментальних даних за критерієм гладкості на гетероскедастичні процеси

Об'єкт дослідження. Процеси сонячної активності.

Мета роботи. Розробка та вдосконалення методу аналізу гетероскедастичних процесів із змінною дисперсією за комплексним критерієм; порівняння запропонованого алгоритму з традиційно використовуваними алгоритмами в умовах невизначеності – 13-місячним згладжуванням, Гауссівським фільтром, – для відновлення динаміки 11-річних сонячних циклів.

Методи дослідження. 13-місячне згладжування, Гауссівський фільтр, метод апроксимації експериментальних даних за критерієм гладкості.

Результати. У попередній роботі [2] було запропоновано алгоритм згладжування, що є оптимальним за комплексним критерієм, але його застосування обмежувалося гомоскедастичністю процесів. В цій роботі пропонується використовувати метод апроксимації експериментальних даних для гетероскедастичних процесів, тобто тих процесів, у яких вимірювання мають різну дисперсію. Цей метод не потребує побудови математичної моделі та є оптимальним за комплексним критерієм, що включає вимоги гладкості шуканої кривої та її близькості до експериментальних даних. Виконується порівняння запропонованого алгоритму з традиційно використовуваними алгоритмами в умовах невизначеності – 13-місячним згладжуванням, Гауссівським фільтром, – для відновлення динаміки 11-річних сонячних циклів. Демонструється перевага запропонованої методики для згладжування гетероскедастичних процесів, що виявляється у більшій гладкості кривої та близькості до експериментальних даних.

Так як на практиці спостереження неоднорідні, що виявляється у змінній дисперсії, то було запропоновано вдосконалений критерій методу апроксимації, який враховує змінні дисперсії та заснований на алгоритмі згладжування, що враховує гладкість шуканої кривої та її близькість до експериментальних даних. Основна ідея алгоритму базується на перетворенні вихідних даних вимірювань у нормований вектор вимірювань. У комплексному критерії, на відміну від попередньої роботи, формується зважена сума квадратів відхилень, в якій вагові коефіцієнти обернено пропорційні дисперсіям помилок вимірювань:

$$J = \frac{\rho}{\sigma^2}\sum_i (p_i - g_i)^2 + \sum_j (g_j - 2g_{j+1} + g_{j+2})^2,\ i = \overline{1,N},\ j = \overline{1,N-2} \qquad (1)$$

де ρ – коефіцієнт згладжування, який обирається виходячи з цілей та умови задачі, σ^2 – змінна дисперсія, p_i – значення деякої функції, яку необхідно згладити, і заданої на множині рівновіддалених точок: $i = \overline{1,N}$, g_i – значення згладженої функції, які необхідно визначити. Перший доданок у (1) визначає внесок у критерій вимогу близькості згладженої кривої до експериментальних даних. Другий доданок враховує вимогу до її гладкості. Аналогічно видозмінюємо матрицю в алгоритмі згладжування, тобто на головній діагоналі коефіцієнт згладжування ділимо на дисперсію. Таким чином, отримуємо критерій, що є функцією від змінної дисперсії.

Спочатку було проведено імітаційне моделювання з використанням вдосконаленого методу апроксимації експериментальних даних для відновлення моделі гармонічного ряду. Дисперсія була задана таким чином, що із ростом часу величина дисперсії збільшувалася. Після застосування вдсконаленого алгоритму згладжування відновлена крива майже співпадала з заданим гармонічним рядом (рис. 1).

Після імітаційного моделювання запропонований алгоритм було застосовано до реальних даних – середньомісячних чисел Вольфа, наданих Центром аналізу даних по впливу Сонця (Бельгія) [4]. Вдосконалений алгоритм значно перевершує традиційно застосовувані алгоритми

Тут можна вказати відомості про грант, в рамках якого виконано дослідження.

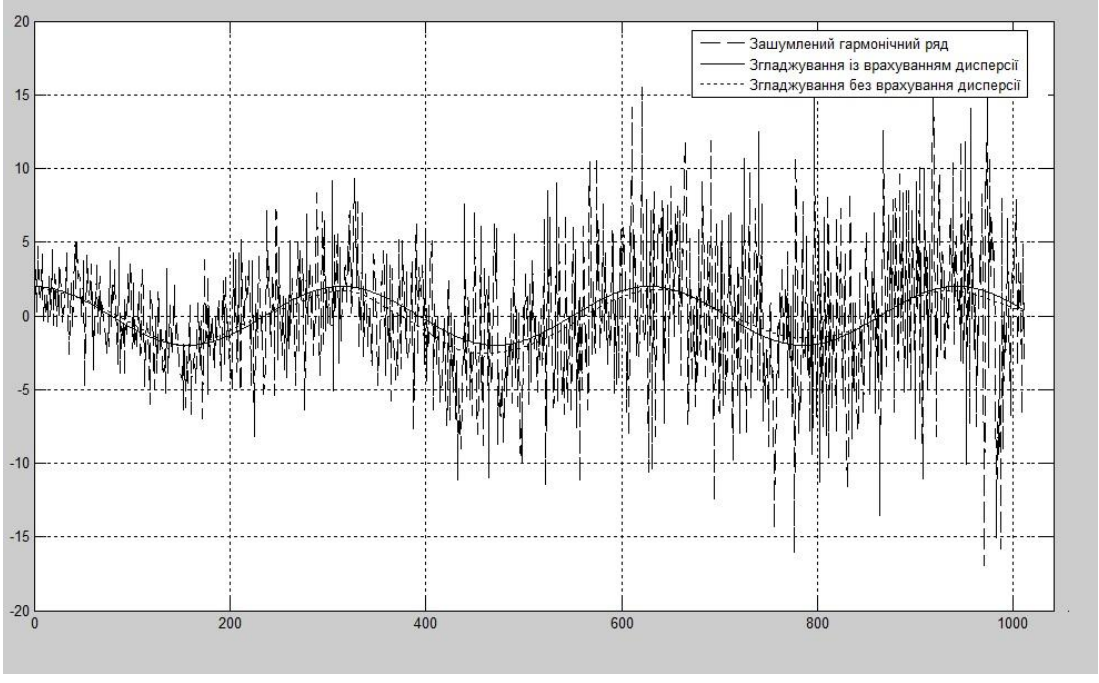


Рис. 1. Імітаційне моделювання

(13-місячне згладжування, Гауссівський фільтр) та алгоритм згладжування, що був запропонований у попередній роботі [2] як за близькістю до експериментальних даних, так і за гладкістю відновлюваної кривої.

Запропонований вдосконалений метод апроксимації експериментальних даних може бути використаний для виявлення закономірностей широкого класу гетероскедастичних процесів природничо-наукової, економічної, соціальної природи за відсутності будь-якої іншої апріорної інформації.

Література. **1.** Глухов Г. Н. Алгоритм цифрового сглаживания // Научно-технический и рекламно-коммерческий периодический журнал ФГУП ННИПИ «Кварц» Нижний Новгород. – 2004 – С. 43–46. **2.** Ганчукова Д. В., Подладчіков В. М. Пошук закономірностей динаміки артеріального тиску на основі апроксимації експериментальних даних за критерієм гладкості // Системні дослідження та інформаційні технології. – 2016 – № 1. **3.** David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann. A Synthesis Of Solar Cycle Prediction Techniques // Journal Of Geophysical Research, Vol. 104, No. A10, Pages 22375–22388, October 1, 1999. **4.** Джерело експериментальних даних: WDC-SILSO, Royal Observatory of Belgium, Brussels [Електронний ресурс] - Доступний з: http://www.sidc.be/silso/datafiles.

***Глоба Л.С.*[1] *Коваль О.В.*[1] *Новогрудська Р.Л.*[1] *Сенченко В.Р.*[2]**
[1] *Національний технічний університет України «Київський політехнічний інститут», Київ, Україна;* [2] *Інститут проблем реєстрації інформації НАН України, Київ, Україна*

Створення сценаріїв обробки даних на основі онтології

Однією з найважливіших тенденцій розвитку систем обробки великих обсягів даних є перехід до підтримки парадигмі ДАНІ→ ІНФОРМАЦІЯ→ЗНАННЯ [1]. В цьому контексті дані розглядаються як масиви незв'язних різнорідних територіально-розподілених даних фактичного характеру, що описують об'єкти, процеси або явища предметної області (ПрО), а також їхні властивості, за якими можна лише виконувати елементарний пошук. Інформація розглядається як попередньо оброблені та узагальнені дані, класифіковані за певною ознакою, що явно становлять інтерес для фахівців при проведенні аналітичної діяльності у визначеній ПрО. Отримання інформації з даних виконується із застосуванням методів аналітичної обробки, узагальнення, групування, математичних моделей і методів, методів класифікації, структуризації і систематизації, аналітичної обробки даних, узагальнення, системного та сценарного аналізу. Під знаннями, в контексті цієї парадигмі, розуміється необхідна для фахівців сукупність даних, використовувана за певними правилами і процедурами з урахуванням професійного ставлення фахівців до цієї інформації. Для отримування знань з інформації використовується широкий спектр методів, включаючи експертні оцінки, інтелектуальної обробки інформації аналізу даних, а також онтології [2–4].

Об'єктивними труднощами при створенні ефективних моделей і методів обробки великих обсягів інформації є: недостатня систематизація інформації, неоднозначність метаописів в онтологіях для багатьох наукових та інженерних порталів особливо, якщо інформація розподілена в Web-ресурсах; неможливість автоматизованого створення workflow для здобуття знань з цих Web-ресурсів. З метою підвищення продуктивності обробки великих обсягів даних пропонується об'єднати три важливі групи математичних моделей і методів, а саме: онтологічні моделі як засіб опису наборів даних у мережі Internet; методи побудови метаграфів для наочного відображення об'єктів онтології ПрО; нечітку логіку як засіб обробки великих обсягів інформації для функціонування систем, заснованих на знаннях.

Процес обробки великих обсягів інформації, згідно запропонованого підходу, включає: по-перше, створення онтології, яка прив'язує інформаційні ресурси під конкретні задачі на підставі моделі онтологій; по-друге, формування пошукових запитів «на лету» (з урахуванням неточності метаописів) в процесі обробки на підставі адаптуємих сценаріїв обробки інформації.

Для здобуття знань використовують сценарії роботи із завданнями користувача (workflow), які дозволяють звертатися до тих web-сервісів, що семантично найбільш відповідають запитам користувачів в розподілених інформаційних сховищах порталів знань. Для такої реалізації workflow і подальшої обробки отриманих даних формують семантичні відносини між процесами (web-сервісами), щоб відобразити подібності та відмінності між запитами кінцевих користувачів і моделями, що описують web-сервіси.

Проблема неоднозначності мовних конструкцій вирішується за рахунок того, що web-служби пошуку використовують метаописи онтологічних понять ПрО, написані за допомогою такого інструментарію, як: мова DAML (DARPA Agent Markup Language), у вигляді RDF-графів (Resource Description Framework), мова LARKS (Language for Advertisement and Request for Knowledge Sharing), або LSD (Learning Source Descriptions). Таки метаописи онтологічних понять ПрО засновані на розподілі ймовірностей примірників різних заздалегідь визначених онтологічних моделей. Незважаючи на значне число наявних рішень, залишається істотним вплив людського фактору і, крім того, не підтримуються опис для входів і виходів web-сервісів. Для поліпшення релевантності виявлення web-сервісів в інформаційному просторі пропонується використовувати семантичну близькість [4], яка враховує входи і виходи операцій, наприклад, безліч типів параметрів і «шаблони сервісів».

Підхід до пошуку web-сервісів в онтологічних середовищах складається з трьох послідовних етапів:

1. Створення «шаблону сервісу», виходячи з ймовірних запитів користувача.
2. Порівняння «шаблону сервісу» з web-сервісами, які були визначені як сервіси-кандидати.
3. Повернення web-сервісів, які задовольняють вимогам мінімальної подібності, виходячи з ймовірних запитів користувача.

Складний сценарій користувача можна ідентифікувати із складним інженерним або науковим розрахунком, який реалізується в програмному середовищі у вигляді конкретного workflow, а кожен частковий елемент сценарію ідентифікується з вирішення конкретного завдання, реалізованим у вигляді окремого web-сервісу. Множина часткових сценаріїв - T зберігається в сховищі, яке не має зв'язків зі сховищем складних сценаріїв - чСе sSe. Складний сценарій чСе sSe формується із часткового сценарію чСэ рСе на базі порівняння значений параметрів, які використовуються в Сэ. Це дозволяє автоматизувати процес формування складного сценарію та встановлювати порядок його на «лету». В цьому випадку складний сценарій чСе sSe являє упорядковане дерево, порядок обходу вершин якого відповідає порядку обчислень за допомогою знайдених web-сервісів. Отриманий складний сценарій для певного інженерного, наукового розрахунку або для процесу аналізу великих обсягів даних може бути представлений як шаблон для подальшого використання – як типовий сценарій. Докладний опис методу формування складного сценарію представлено у роботі [6].

Для візуального аналізу отриманого на «льоту» сценарію пропонується використовувати метаграф. Мета застосування метаграфу полягає в можливості моделювання бізнес-процесів з використанням діаграм виконання робіт (workflow), візуальному представленні та аналізі аномалій в нечітких наборах даних. При нечіткості метаописів web-сервісів, вершини можна представити у вигляді метаописів тих web-сервісів, при пошуку яких всі критерії пошуку виконані, а метавершіни - це метаописания таких web-сервісів, критерії пошуку яких задоволені частково або коли для вирішення однієї задачі знайдено достатню кількість сервісів - претендентів. В цьому випадку візуальний аналіз аномалій у метаграфі дозволяє уникнути помилкових рішень при опису понять ПрО.

Розглянутий підхід до автоматизації побудови сценаріїв обробки даних відрізняється використанням онтологічних моделей для опису інформаційних і обчислювальних ресурсів, а застосування метаграфів (для візуалізації та аналізу отриманих сценаріїв) є особливо ефективним в таких складних сферах діяльності як інженерія та наукові дослідження. Використання онтологій в комплексі з методами семантичного пошуку дозволяє сформувати сценарії рішення складних наукомістких завдань в середовищі Internet, наприклад, ідентифікація об'єктів методами гідроакустики.

Література. **1.** On Data, Information, and Knowledge beaststwo.org/inforant/index.shtml. **2.** Basu A., Blanning R. Metagraphs and their applications / A. Basu, R. Blanning // Springer Science+Business Media, LLC, 2008, 173 p. **3.** Коваль А.В. Верификация компьютерной модели системы информационного управления / А.В. Коваль, К.А. Зайцева // Вестник НТУУ «КПІ» Информатика, управление и вычислительная техника. - 2014. - №61, С. 43-48. **4.** Globa L. Increasing web services discovery relevancy in the multi-ontological environment / L. Globa, M. Kovalskyi, A. Stryzhak // Advances in Intelligent Systems and Computing, ISBN 978-3-319-15147-2 (eBook): Springer International Publishing Switzerland, 2015, p. 335-344. **5.** Новогрудска Р.Л. Інформаційна технологія створення та підтримки порталів інженерних знань // Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Національний технічний університет України «Київський політехнічний інститут», Київ, 2015. **6.** Globa L. Based on Force-Directed Algorithms Method for Metagraph Visualization / L. Globa, M. Ternovoy, O. Shtogrina, O. Kryvenko // Advances in Intelligent Systems and Computing, ISBN 978-3-319-15147-2 (eBook): Springer International Publishing Switzerland, 2015 – p. .359-370.

Гусак О.В., Алєксєєв М.О.
Інститут телекомунікаційних систем НТУУ "КПІ", Київ, Україна

Балансування навантаження розподілених гетерогенних інформаційних мереж з допомогою підходів штучного інтелекту

Вступ. Розподілені розрахунки з використанням портативних пристроїв набувають все більшої популярності. Це викликано різким їх поширенням, збільшенням потужностей, часу автономної роботи. За рахунок оптимального балансування навантаження і коригування такої мережі можливо отримати великі обчислювальні потужності. При реалізації розподілених розрахунків за допомогою гетерогенних, невідчужуваних, некластеризованих обчислювальних ресурсів має ряд проблем основою яких є ефективний розподіл завдань між вузлами. Оскільки вузли мережі можуть довільно від'єднуватись і обчислювальні потужності таких пристроїв заздалегідь не відомі, необхідно застосувати такий підхід до розподілу завдань, який враховував би топологію мережі, пропускну здатність до вузлів, їх потужність і прогнозував час роботи. Необхідно знайти новий підхід для реалізації балансування навантаження в розподілених гетерогенних мережах. Метою даної роботи є огляд існуючих алгоритмічних рішень, їх актуальність. Впровадження рішень основаних на штучному інтелекті, застосування або вдосконалення якого дозволило б наблизити ефективність розподілених розрахункових мереж до кластерних, знизити негативний ефект динамічних і гетерогенних складових, тим самим розширивши можливий діапазон задач.

Існуючі алгоритми балансування розділяють на дві основні групи: статичні і динамічні. В обох групах є алгоритми, які застосовують методи штучного інтелекту(ШІ) для реалізації складної логіки організації роботи. Враховуючи переваги кожного з підходів можна підбирати комбінацію алгоритмів для поставлених цілей (адаптивні алгоритми). Оскільки ми маємо справу з динамічною гетерогенною архітектурою, статичні алгоритми балансування не підходять за визначенням. В даній групі алгоритмів рішення про розподілення навантаження приймається заздалегідь і не змінюється в процесі роботи. Динамічне балансування передбачає перерозподіл розрахункової потужності на вузлах в ході виконання програми. Завдяки тому, що динамічні алгоритми відслідковують і аналізують такі параметри як завантаження розрахункових вузлів, пропускну здатність лінії, частоту обміну повідомленнями між логічними процесами, це призводить до збільшення накладних розрахунків на балансування в порівнянні із статичними алгоритмами. Адаптивні ж алгоритми використовують поєднання декількох алгоритмів. В залежності від умов і поточних параметрів розрахункової системи вибирається найбільш доречний алгоритм. Прикладом є система [1] з підтримкою поколінь планування (The introduction of the Generational Scheduling) з реплікацією завдань. Цей алгоритм адаптується до змін в продуктивності з реструктизацією завдань. Хоч даний підхід має низку переваг, він періодично потребує втручання людини для корегування або внесення змін при зміни топології мережі. Даний недолік значно знижується в підходах з використанням штучного інтелекту. В таких рішеннях алгоритм адаптується до змін в топології, неоднорідностей в мережі, пропускної здатності з'єднань, об'єму поставленої задачі. Недоліком такого підходу є складність реалізації яка залежить від реалізації самого штучного інтелекту.

Класифікація ШІ.

- *Класичний ШІ.* Цим методом вперше були спроби змоделювати комп'ютерні програми які б могли розпізнавати людську мову, або виявляти об'єкти на зображеннях. Як правило, для досягнення поставлених цілей необхідно мати доступ до великих об'ємів даних. Однім з прикладів сучасних систем що реалізує даний підхід є IBM Watson [2]. Вона націлена на створення складної бази даних по заданому запитанню. Хоч Watson і не залежить від кодованих правил, він все рівно потребує експерта з певної області для надання інформації і оцінки ефективності. Класичний підхід націлений на розв'язок чітко поставлених задач. Він не може вчитися на своїх помилках і для розв'язку нових проблем необхідно створювати нові рішення. Такий інтелект має мало спільно з людським

інтелектом і потребує значних затрат на побудову і підтримку.

- *Прості нейронні мережі.* Коли межі класичного підходу були досліджені, виникло припущення що штучний інтелект повинен бути змодельований шляхом наслідування структури нервових тканин. Саме так реалізована більшість підходів до розпізнавання образів. Розв'язуються задачі нелінійної оптимізації по навчанню нейронної мережі. Тривалий час нейроні мережі є основним способом реалізації штучного інтелекту з допомогою комп'ютерних алгоритмів і є частиною великої категорії «machine learning» і підкатегорії «deep learning» з допомогою яких вдалося наблизитися реалізованими розрахунковими потужностями до швидкісних комп'ютерів. Перші прототипи були створені в 1950-х рр. і незважаючи на те що зараз про роботу мозку і нервових клітин стало відомо набагато більше, архітектура нейронних мереж майже не змінилася. Незважаючи на назву, дизайн нейронних мереж має мало спільного з реальними прототипом. Нейронні мережі погано працюють якщо обмежено данні для навчання і закономірності вхідної інформації постійно змінюються.
- *Біологічні нейронні мережі.* В даному напрямку відштовхувалися від того що єдиною інтелектуальною системою є людський мозок і правильним напрямком в створені штучного інтелекту це симуляція роботи тієї частини мозку яка відповідає за інтелект. Відомо що мозок обробляє інформацію за допомогою розподіленої памяті SDR (Sparse Distributed Representations) [3] яка необхідна для творчості і семантичного узагальнення. По справжньому інтелектуальну мережу можна зробити використовуючи SDR як основу для всієї архітектури. Відомо що різниця між біологічним нейроном і нейроном в простих нейронних мережах досить значна і не можна досягти бажаного результатом шляхом дорощування, ускладнення існуючої нейронної мережі - необхідно почати побудову заново з більш реалістичним біологічним підходом.

Найкращим прикладом біологічних нейронних мереж є HTM (Hierarchical Temporal Memory) [4]. На сьогодні HTM здатний робити передбачення, розпізнавати залежності у вхідній інформації і працювати з структурою даних в потоковій інформації [5]. Він постійно навчається реалізуючи людський підхід в аналізі даних, уяві, поведінці і короткочасній пам'яті. Це відрізняє HTM від інших аналогів, роблячи його кращим алгоритмом для використання в розподілених гетерогенних динамічних мережах.

Висновок. В ході дослідження було розглянуто основні типи алгоритмів і зроблена класифікація підходів до побудови штучного інтелекту, які можуть бути застосовані для організації розподілених розрахунків з використанням гетерогенних невідчужуваних некластерізованих обчислювальних ресурсів. Наведено аргументи доречності вибору HTM в якості основи для алгоритму балансування навантаження. Перевагою HTM є можливість працювати з потоком даних, самостійний пошук інформаційних залежностей, детектування аномалій. В подальшому досліджені планується реалізація підходу розподілу завдань за допомогою алгоритму HTM.

Література. **1.** Nissimov A., Feitelson D.G.(2007) *Probabilistic Backfilling.// Job Scheduling Strategies for Parallel Proc.* LNCS. Springer-Verlag. Vol. 4942. **2.** Dr. John E. Kelly III(2015) *Computing, cognition and the future of knowing* [Online]. Available: http://www.research.ibm.com/software/IBMResearch/multimedia/Computing_Cognition_WhitePaper.pdf. **3.** Geoffrey E. Hinton(1997) *Generative models for discovering sparse distributed representations* [Online]. Available: http://rstb.royalsocietypublishing.org/content/352/1358/1177#related-urls. **4.** ©Numenta, Inc. (2011) *Hierarchical Temporal Memory including HTM Cortical Learning Algorithms* [Online]. Available: http://numenta.com/assets/pdf/whitepapers/hierarchical-temporal-memory-cortical-learning-algorithm-0.2.1-en.pdf. **5.** ©Numenta, Inc.(2015) *The Numenta Anomaly Benchmark The first temporal benchmark designed for anomaly detection in streaming data* [Online]. Available: https://numenta.com/assets/pdf/numenta-anomaly-benchmark/NAB-Business-Paper.pdf.

Декрет М.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Підвищення ефективності якості прогнозування часових рядів інструментів фінансових ринків за рахунок агрегації різних довжин тайм-фреймів

Для ведення ефективної торгівлі на фінансових ринках необхідним є якісне прогнозування значень цін фінансових інструментів, з якими виконуються операції. Класичний підхід прогнозування використовує аналіз часового ряду на одному тайм-фреймі. Однак очевидною є можливість збільшення ефективності якості прогнозування, завдяки агрегації оцінок одного і того ж прогнозованого значення на різних довжинах тайм-фреймів. Для підтвердження даної теорії в доповіді будуть представлені результати порівняльного аналізу різних підходів.

Об'єкт дослідження. Часові ряди інструментів фінансових ринків.

Предмет дослідження. Моделі, методи та технології прогнозування часових рядів інструментів фінансових ринків.

Мета роботи. Розробити оптимальну модель для прогнозування значень тайм-фреймів та порівняти ефективність реалізації за допомогою нейронних мереж *Back Propagation*, авторегресійної моделі прогнозування та методу головних компонент.

Дослідження. Для досягнення поставленої цілі були проаналізовані прогнози на різних часових інтервалах. В якості результату представлені значення критеріїв *RMSE*, *MAE*, *MAPE* та *maxAPE*. В ході експерименту розглянули пару фунт/долар за січень-жовтень 2015 рік.

Табл. 1. Порівняння оцінок для різних періодів прогнозу

	Години	Дні	Тижні
MAPE	0.0011	0.0073	0.0872
RMSE	1.0000	0.9970	0.9681
maxAPE	4.0570	7.5505	14.4762
MAE	6.3200	11.9086	22.5046

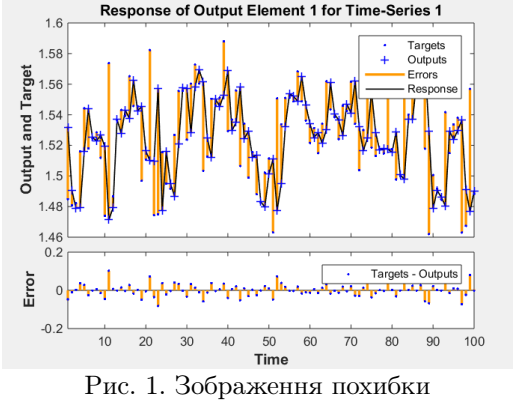


Рис. 1. Зображення похибки прогнозування по дням

Проаналізувавши кожну вибірку по критеріям, що представлені в таблиці 1, побудували модель, результати якої наведені в таблиці 2. Отримана модель, маючи інформацію про більш глобальну поведінку котирування фінансового інструменту, з огляду на *maxAPE* та *MAE*, краще реагує на нестабільні значення, що актуально для ринку Forex і матиме широке застосування для прогнозування значень фінансових інструментів на різних довжинах тайм-фреймів.

Табл. 2. Порівняння звичайного прогнозу і моделі

	MAPE	RMSE	maxAPE	MAE
Прогноз	0.0073	0.9970	7.5505	11.9086
Модель	0.0779	0.6340	0.9320	1.4365

Література. **1.** Хайкин Саймон.- Нейронные сети: полный курс, 2-е издание. : Пер. с англ. – М. : Издательский дом «Вильямс», 2006. – 1104 с. **2.** Бідюк П.І., Романенко В.Д., Тимощук О.Л. - Аналіз часових рядів. - Київ: Політехніка, 2010. - 317с. **3.** Зайченко Ю.П. - Основи проектування інтелектуальних систем. - Київ: Видавничий Дім «Слово», 2004. - 352с.

***Дуда О.М.*[1], *Кунанець Н.Е.*[2], *Липак Г.І.*[1], *Мацюк О.В.*[1], *Небесний Р.М.*[1], *Пасічник В.В.*[2]**

[1] *Тернопільський національний технічний університет імені Івана Пулюя;* [2] *Інститут комп'ютерних наук та інформаційних технологій, Національного університету "Львівська політехніка"*

Консолідація інформаційних ресурсів соціокомунікаційного середовища в проектах "Розумне місто"

Стрімкий розвиток інформаційних технологій, зорієнтованих на формування соціокомунікаційного середовища міста, генерує потреби щодо нових інформаційних ресурсів та суттєво підвищує вимоги до методів, засобів та способів задоволення інформаційних запитів громадян. З цією метою і формується комфортне інформативно насичене міське соціокомунікаційне середовище, яке отримало назву "Розумне місто". Розвиток соціальних комунікацій, як системи суспільної взаємодії, потребує удосконалення інформаційно-технологічного базису, що забезпечує комфортне міське середовище «Розумного міста», яке є унікальним і в кожному випадку специфічним. Формування такого якісно оновленого соціокомунікаційного середовища неможливе без застосування широкого спектру технологій та методів консолідації інформації.

Мета даного дослідження – реалізація дослідницького проекту із розроблення методологічних підходів та інформаційних технологій, що дозволяють формувати платформи та макети соціокомунікаційного середовища різних за масштабами та профілями міст. Однією із компонент в таких макетах є консолідований інформаційний ресурс соціальних інститутів (бібліотек, архівів, музеїв, місцевих ЗМІ та влади).

Проект, що реалізовується у віртуальній науково-дослідній лабораторії Тернопільського національного технічного університету імені І.Пулюя, спрямований на вирішення важливої проблеми - побудови методологічних засад та інформаційно-технологічного прототипу соціокомунікаційного середовища сучасного міста з метою задоволення інформаційних потреб мешканців та туристів шляхом надання їм зручних інформаційно-технологічних інструментів.

Вирішення проблеми формування якісного соціокомунікаційного середовища сучасних міст з використанням технологій консолідації інформаційних ресурсів базується на спеціальних технологіях опрацювання фондів місцевих бібліотек, музеїв та архівів, що відіграють роль соціальної пам'яті міської громади, уособлюють їх культурну та інтелектуальну спадщину.

У проекті "Розумне місто", в якому прототипом було обрано місто Тернопіль, передбачається моделювання та макетування процесів формування соціокомунікаційного середовища міста, зокрема, здійснюватиметься:

- розроблення методологічних засад формування комплексу інформаційних технологій для системи ефективного та комфортного забезпечення інформаційних потреб міської спільноти та гостей міста;
- побудова комплексу моделей, прототипів та макетів процесів формування та розвитку соціокомунікаційного середовища сучасного міста;
- створення консолідованого інформаційного ресурсу на основі фондів міських бібліотек, музеїв та архіві з метою забезпечення інформаційних потреб міської спільноти, популяризації потенційних туристичних маршрутів та історичних пам'яток регіону.

Реалізація проекту щодо розроблення макетів та прототипу соціокомунікаційного середовища на прикладі міста Тернопіль передбачає поглиблений аналіз результатів та напрацювань як українських, так і зарубіжних дослідників, формування методологічних засад побудови перспективних процесів консолідації різнорідних метаданих щодо інформаційних ресурсів бібліотек, музеїв та архівів міста, реалізації інформаційних технологій швидкого та ефективного доступу до них.

***Дуда О.М.*[1], *Кунанець Н.Е.*[2], *Мацюк О.В.*[1], *Пасічник В.В.*[2]**
[1] *Тернопільський національний технічний університет імені Івана Пулюя;* [2] *Інститут комп'ютерних наук та інформаційних технологій, Національного університету "Львівська політехніка"*

Системні комплекси інформаційних технологій у проектах "Розумне місто"

Запропоновано застосовувати комплексний системний підхід до розроблення методологічних засад та застосування сучасних інформаційних технологій для формування інформаційно-технологічного базису реалізації проектів "Розумне місто". Авторами проаналізовано використання програмно-алгоритмічних застосунків, які реалізують процеси побудови відповідних онтологій, інтерпретують методи вирішення задач пошуку інформації, системи формування профільних рекомендацій, адаптивні процедури вибору моделей і критеріїв оцінки управлінських рішень.

"Smart city" ("Розумне місто") – місто, в якому для підвищення якості життя міської спільноти активно використовуються сучасні інформаційні технології. Інформаційні технології "Розумного міста" "вбудовуються" у виробничі процеси міських структур з метою якіснішого надання послуг, зменшення вартостей споживаних ресурсів, поліпшення комунікації та взаємодії влади з містянами, розширення спектру послуг надаваних туристам.

Успішно впроваджують концепцію "Розумного міста" в Європейських країнах, США та Канаді. В країнах європейського союзу зазначена системна концепція [1] включає наступні напрямки: мобільність інфраструктури, управління навколишнім середовищем [2], управління міською спільнотою, управління міським господарством, розвиток та підвищення ефективності економіки.

В переважній більшості відомих на даний час проектів "Smart city" чітко не окреслено вичерпний перелік та профільність використання інформаційних технологій. Діючі проекти в основному спрямовані на вирішення проблем енергоефективності, оптимізації трафіку міського транспорту, прикладних застосунках на базі геоінформаційних технологій і т. ін. Це зумовлено, зокрема, необхідністю вирішенням біжучих проблем функціонування міста як складної багатопрофільної та багатокомпонентної системи. Інформаційні технології при цьому покликані зіграти роль інтегруючих ланок між різними інфраструктурними рівнями та підсистемами і забезпечувати надійний та якісний обмін інформацією між ними.

В подібній ситуації, яка потребує суттєвого поліпшення, знаходиться переважна більшість українських міст, які декларують необхідність реалізації перспективних проектів "Розумних міст". Розробляються окремі елементи та ланки інформаційних технологій "Розумного міста". В обласних центрах та великих містах реально забезпечується вільний доступ до WiFi, реалізується моніторинг руху громадського транспорту, функціонують комплекси мереж "розумних" світлофорів, підтримуються on-line сервіси з оплати комунальних послуг і т. ін. На даний час вичерпно не опрацьовані системні комплексні підходи до формування цілісного інформаційно-технологічного базису, необхідного для якісної реалізації проектів "Розумного міста".

Для кваліфікованого вирішення проблем, пов'язаних з формуванням класів IT для реалізації проектів "Розумне місто", зокрема, до розробки методологічних засад та способів застосування сучасних інформаційних технологій, що дозволяло б враховувати як загальнодержавні, так і місцеві інтереси, забезпечувати раціональне використання людських, технологічних, природних, матеріальних та фінансових ресурсів, без сумніву слід застосовувати комплексний системний підхід. Дослідження таких складних комплексів інформаційних технологій пов'язане з об'єктивними складнощами, які зумовлені зокрема наступними причинами:

- класи інформаційних технологій, що використовуються у сучасних містах, належать до класу складних систем, які характеризуються великими множинами кількісних і якісних параметрів;

- між зазначеними технологіями існують різнотипові зв'язки різної природи та інтенсивності, котрі суттєво можуть змінюватись в часі;
- комплексні системи управління сучасних міст мають слабку організаційну структурованість, є доволі нестійкими та достатньо хаотичними; явища та параметри, які характеризують системи, у багатьох випадках не мають стійких імовірнісних розподілів;
- складна система, якою є сучасне місто, визначається параметрами, що динамічно змінюються і відображають рівень знання щодо таких систем, за якого неможливо однозначне визначення їхніх морфологічних і функціональних особливостей.

Стан комплексної адміністративно-господарської та інформаційно-технологічної складної системи, якою є система класу "Розумне місто", визначається її функціональним наповненням, характеристиками внутрішніх процесів, процедурами її взаємодії із зовнішнім середовищем, державними, муніципальними та громадськими чинниками.

Можна виділити базові напрямки використання інформаційних технологій в проектах "Розумне місто":

- використання великих масивів даних та аналітичного їх опрацювання з метою підвищення якості інформаційних послуг;
- мобільне збирання даних та оперативне реагування на проблемні ситуації;
- віртуальна взаємодія між різнорідними установами, структурами та мешканцями міста;
- активне використання соціокомунікаційних технологій, що покращують взаємодію місцевих органів влади з громадянами.

Для підготовки та прийняття управлінських рішень, які будуть сприяти досягненню мети, інформаційні технології в проектах "Розумного міста" повинні реалізовувати сучасні інтелектуальні методи та підходи до розв'язання широкого спектру задач, опиратись на гіпервеликі бази даних.

Одним з базових напрямків подолання проблем інтеграції комплексу різнопланових інформаційних технологій у проектах "Розумне місто" є використання програмно-алгоритмічних застосунків, які реалізують процеси побудови відповідних онтологій, інтерпретують методи вирішення задач пошуку інформації, системи формування профільних рекомендацій, адаптивні процедури вибору моделей і критеріїв оцінки управлінських рішень. Такий підхід дає змогу формувати повний та якісний інтегрований простір даних та знань "Розумного міста".

Використання принципів відкритих систем, комплексного запровадження аналізу широкого спектру сучасних інформаційно-технологічних концептів, таких як сховища та простори даних, Великі дані, хмарні обчислення, геоінформаційні технології, GRID-системи, мобільний комп'ютинг, видобування даних та знань та реалізації при цьому інформаційного користувацького гасла "ВСЕ! ТУТ! НЕГАЙНО! ЗРУЧНО! КРАСИВО!" дозволяє реально втілювати такого роду проекти.

Окреслені базові характеристики системних класів інформаційних технологій закладаються в основу масштабного інноваційного проекту "Розумне місто", над втіленням якого активно працюють автори доповіді. Прототипування та макетування базових ланок проекту здійснюється розробниками на прикладі міста Тернопіль.

Література. **1.** European smart cities 4.0 (2015). The smart city model [Electronic resource] – Mode of access: WWW.URL: http://www.smart-cities.eu/?cid=2&ver=4. – Last access: 23.02.2016. – Title from the screen. **2.** Deakin M. Sustainable urban development: Use of the environmental assessment methods. / M Deakin, A Reid // Sustainable Cities and Society, 2014

Зайченко Ю.П., Четырбок П.В.
Учебно-научный комплекс "Институт прикладного системного анализа" Национального технического университета Украины "Киевский политехнический институт", Украина

Распознавание объектов с использованием модульных нейронных сетей

Постановка проблемы. Построить модульную нейронную сеть для отображения множества распознаваемых образов (векторов параметров образов) на множество векторов ошибок распознавания образов нейронною сетью, которое позволяет связать классификацию образов с анализом векторов в пространстве ошибок.

Концепция модульности нейронных сетей. Для описания сложных систем распознавания объектов применяются тысячи показателей. Число экспериментальных наблюдений по каждому показателю составляет до несколько тысяч. Это выдвигает проблемы перед технологиями реконструктивного анализа сложных систем. Современная методология реконструктивного анализа основана на конструктивной процедуре индукции, обеспечивающей развитие первоначально отобранных простых структур в конечные результа-тивные структуры большой сложности. Модульность нейронной сети позволяет выполнить иерархическую декомпозицию сложной задачи в ряд более простых подзадач, а соответствующая структура сети может быть оптимизирована под конкретную задачу, то есть нейронная сеть может служить средством моделирования структуры данных. Структурный анализ модульной сети инвариантен внутреннему устройству модулей. Вместо нейронного модуля может быть звено системы автоматического регулирования или узел коммутационной системы. Техническая реализация больших и сверхбольших нейрон-ных сетей на современных универсальных компьютерах имеет модульную распределенную структуру, поскольку технически невоз-можно реализовать сверхбольшую нейронную сеть на одном процессоре. При программной реализации быстродействие модульных сетей зависит от выбранной структуры. Это качество делает перспективным использование модульных нейронных сетей в системах реального времени, которой и является система распознавания объектов. Исполь-зование технологии сверхбольших интегральных схем (СБИС) для реализации нейронных сетей приводит к необходимости разделения нейронных сетей на однородные фрагменты. Размер фрагмента определяется площадью кристалла, энергопотреблением, числом внешних выводов. Проблема выводов, возникшая при реализации кристалла, продолжается и при построении схем соединений между ними. Таким образом возникает задача разбиения нейронной сети на модули с ограниченным числом исходящих связей. Данную задачу решает по-строение модулей нейронной сети, обученных с использованием кри-терия близости образов в пространстве ошибок.

Выводы. Каждому образу, распознаваемому многослойным персептроном в многофакторном пространстве ошибок соответствует свой вектор ошибок. В докладе построено решающее правило для классификации образов в виде утверждения: каждому образу, распознаваемому мно-гослойным персептроном в многофакторном пространстве ошибок будет соответствовать свой вектор ошибок и образ ближе к эталону, чем больше $cos(\lambda)$:

$$\cos(\lambda) = \frac{(\overline{E},\overline{X})}{||\overline{E}||_c,||\overline{X}||_c},$$

где $\overline{E}$- вектор ошибок в пространстве ошибок, полученный при распознавании нейронной сетью входного образа, $\overline{X}$ - вектор весовых коэфициентов выходного слоя, полученный при обучении сети. Предложенное решающее правило (критерий для распознавания образов) позволяет создать структуру модульной нейронной сети.

Литература. **1.** Хайкин Саймон. Нейронные сети. – М. : Издательский дом «Вильямс», 2006. – 1104 с. **2.** Зайченко Ю.П. Нечеткие модели и методы в интеллектуальных системах. –К.: Издательский дом «Слово», 2008. – 344 с. **3.** Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей. - СПб.: «Политехника», 2014. - 328 с. **4.** Четырбок П.В. Построение решающего правила для классификации образов на основе векторов ошибок / П.В. Четырбок // САИТ. – Київ, 2013. – №2. – С. 114 – 120.

Іванущак Н.М.
Чернівецький національний університет імені Юрія Федьковича, Чернівці, Україна

Інформаційні технології моделювання процесів розвитку комп'ютерних мереж в умовах дестабілізаційних загроз

Метою роботи є дослідження і обґрунтування підходів, які дозволяють ідентифікувати структуру і параметри моделей локальних комп'ютерних мереж на основі даних спостережень; розроблення методів та засобів математичного моделювання процесів розвитку структур комп'ютерних мереж.

Комп'ютерна мережа подається у вигляді графа G, який визначається як сукупність (V, E) кінцевої множини вершин V, $dim(V) = N$, і множини ребер E. Кожна вершина характеризується своїм ступенем, тобто кількістю інцидентних їй ребер. Інтегральною характеристикою комп'ютерної мережі є закон розподілу ступенів p_k, який задає ймовірність того, що випадково вибрана вершина має ступінь приєднання k.

У багатьох реальних мережах емпіричний розподіл ступенів вершин інтерпретується в термінах ступеневого розподілу $p_k = k^{-\gamma}$. Для здійснення процесу моделювання використовуються характеристики реальних мереж Інтернет-провайдерів в м. Чернівцях: «BW-Star & FoxNet», «КТМ» та «DSS-Group». Ступінь вузла k задає кількість ребер інцидентних конкретній вершині, а n_k – кількість вершин у графі із заданим k. За цими даними побудований розподіл ступенів вершин p_k (рис.1).

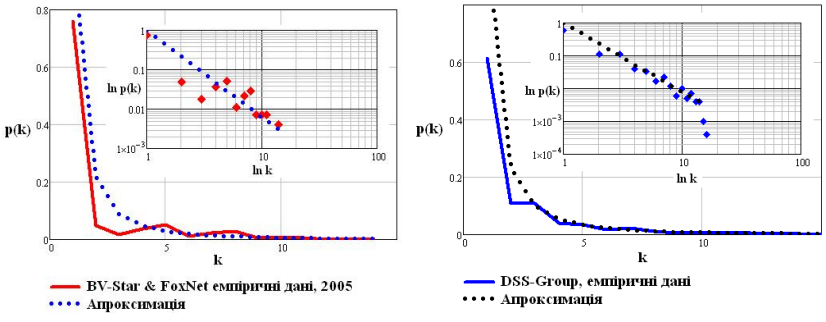


Рис. 1. Апроксимація «хвостів» розподілів ступенів вершин досліджуваних мереж.

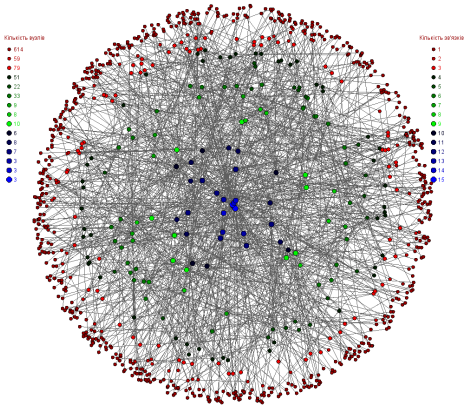


Рис. 2. Графічне зображення топології мережі «BW-Star & Fox Net»: $p_k = k^{-\gamma}$: $\gamma = 2,4$, N=915

Для побудови моделі розвитку комп'ютерної мережі, параметри якої близькі до реальних, пропонується програмна реалізація мережі у вигляді стохастичного графа з відомою кількістю вершин і заданим розподілом ймовірностей їх приєднання p_k.

Здійснивши апроксимацію «хвостів» розподілів ймовірностей приєднання у мережу вершин з різними ступенями k, що проілюстровано на рис.1, та визначивши тим самим показники γ для різних локальних комп'ютерних мереж, змодельовано ці мережі за відомим показниковим $p_k = k^{-\gamma}$ розподілом ймовірностей приєднання користувачів у мережу з відповідними γ. На рис. 2 проілюстровано приклад візуалізації стохастичного графу, який відображає властивості однієї із досліджуваних комп'ютерних мереж.

Запропонована вище абстрактна модель комп'ютерної мережі дає змогу здійснювати прогностичні

оцінки атакувальних дій різних категорій порушників з реалізації найпоширеніших сценаріїв загроз безпеки мережі. Основними компонентами описової конструкції мережі є ***модель загроз*** та ***модель безпеки*** [1].

Сформований стохастичний граф, що відображає структуру реальної мережі, подається на вхід ***моделі загроз***. Імітаційне моделювання дестабілізаційних загроз полягає у проведенні над графом двох сценаріїв атак: «природної» ліквідації вершин і ребер, тобто випадкової R-атаки, та ліквідації вершин з найбільшими ступенями, тобто I-атаки. На рис.3 подано результати зміни відносного розміру максимального кластера S залежно від кількості атакованих вузлів безмасштабних комп'ютерних мереж «BW-Star & FoxNet», «KTM» та «DSS-Group» в м. Чернівці.

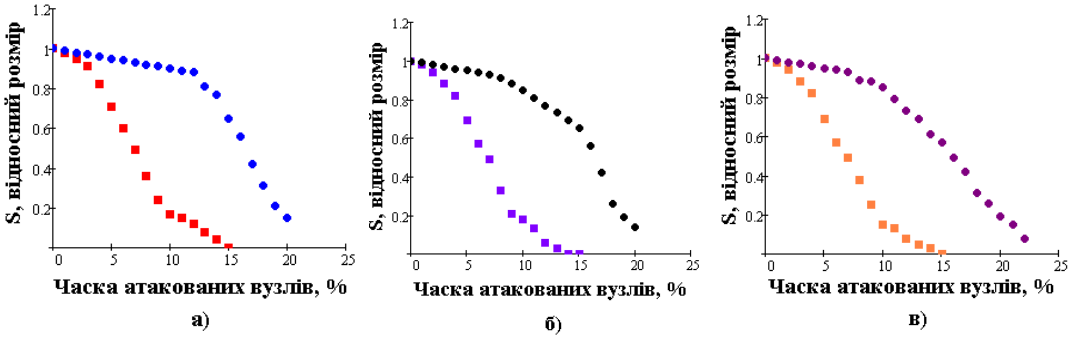


Рис. 3. Залежність відносного розміру максимального кластера в мережах «BW-Star & FoxNet» (а), «KTM» (б) та «DSS-Group» (в) від кількості атакованих вузлів, що піддаються атакам у послідовності: I+R (■) та R+I (•).

Модель безпеки враховує, що основною мірою безпеки є ризик: $R = \rho_{th} \cdot \rho_{\nu} \cdot Pl$, де R – ризик, ρ_{th} і ρ_{ν} – відповідні ймовірності загрози та уразливості, Pl –ціна втрати. На відміну від багатьох відомих методів, де наслідком атаки завжди є знищення вузла чи зв'язка, цей підхід дає можливість для кожного вузла ввести параметр уразливості (або захищеності). Елементам мережі приписують величини їх уразливості, в загальному випадку, з $\rho_{\nu}(i)$ відмінними від 1, які визначаються об'ємом фінансових вкладів F_i в безпеку i-го елемента: $\rho_{\nu}(i) = f(F_i)$.

В якості комп'ютерного експерименту проводилися дослідження реакції відповідних змодельованих стохастичних графів, що відображають структури реальних комп'ютерних мереж, на цілеспрямовані загрози захищених елементів – вузлів.

На основі проведених досліджень обчислено інтегральні ймовірнісні характеристики реальних комп'ютерних мереж у різні часові проміжки, апроксимовано їх задля моделювання.

Вперше розроблено нову математичну модель генерування структури локальних комп'ютерних мереж із заданою функцією щільності розподілу ступенів вузлів. Відтворена в результаті моделювання мережа як стохастичний граф володіє параметрами, які близькі до реальної мережі, що уможливлює вивчення природи формування і розвитку мереж, виявлення нових закономірностей в них, моделювання задач зараження та імунізації, протидії мережевим атакам.

Запропоновані алгоритми моделювання сценаріїв атак використано для оцінювання уразливості комп'ютерних мереж та розв'язання задачі про їхню стійкість до випадкових та спрямованих атак. Розроблено узагальнену модель еволюції мережевого ансамблю в умовах дестабілізаційних загроз.

Література. **1.** Іванущак Н.М. Узагальнена модель еволюції мережевого ансамблю в умовах дестабілізаційних загроз / Н.М. Іванущак, В.В. Пасічник // Вісник Національного університету «Львівська політехніка». – 2015. - №814: Інформаційні системи та мережі. – с. 457-467.

Каковська Т.І., Олефір О.С.
Національний технічний університет України "КПІ", факультет прикладної математики, Київ, Україна

Використання генетичних алгоритмів для формування персонального раціону харчування людини

Однією з головних причин ожиріння виступає незбалансоване та нездорове харчування. Найважливішим завданням збалансованого харчування є правильне поєднання всіх харчових речовин у добовому раціоні людини, притому вони повинні бути введені в кількостях, які повністю покривають всі витрати організму як в кількісному, так і в якісному відношенні. Актуальною задачею є розробка математичного та програмного забезпечення системи створення збалансованого раціону харчування людини.

До процесу складання персональної дієти входить [1]:

1. Анкетування користувача з метою отримання особистих даних.
2. Складання математичної моделі вимог користувача до раціону харчування.
3. Складання персональних дієт на основі вимог користувача до раціону та даних про страви.
4. Схвалення та відсіювання сформованих раціонів харчування користувачем.

Складність становить процес побудови раціону харчування з переліку страв, що відповідає вимогам користувача. Тому формальними критеріями вибору методу розв'язання задачі є:

- необов'язковість отримання точного розв'язку;
- отримання значної кількості наближених розв'язків за малих час;
- врахування додаткових обмежень, зокрема поєднуваності певних продуктів;
- можливість повторного використання результатів обчислень, здійснених на попередніх етапах;
- можливість введення нових обмежень в задачу (наприклад, мікронутрієнтів).

Задача оптимізації полягає у виборі раціонів, що відповідають таким вимогам [2]:

- сумарна добова енергетична цінність (DailyKcalRequirements);
- сумарний добовий склад за нутрієнтами (білки (p), жири (f), вуглеводи (c));
- склад за нутрієнтами щодо кожного прийому їжі (x_{ij});
- поєднуваність страв в прийомі їжі ($food_{incompatibility}$).

Формалізоване представлення задачі:

$$x_1 > DailyKcalRequirements > x_2, \tag{1}$$

$$\sum_{i=1}^{n} p_i = p,\ p_{min} < p < p_{max}, \tag{2}$$

$$\sum_{i=1}^{n} f_i = f,\ f_{min} < f < f_{max}, \tag{3}$$

$$\sum_{i=1}^{n} c_i = c,\ c_{min} < c < c_{max}, \tag{4}$$

$$\sum_{j=1}^{m} x_{ij} = k_p \cdot p_i + k_f \cdot f_i + k_c \cdot c_i, \tag{5}$$

$$\bigcap_{i=1}^{a} food_{incompatibility} = \varnothing \tag{6}$$

Дану задачу можна розглядати як задачу цілочисельного лінійного програмування (ЗЦЛП). Ідея – складання алгоритмів пошуку на основі біологічної моделі механізмів природного добору. Базовими поняттями є: популяція, відбір, мутація, схрещування.

Для вирішення задачі побудови раціону харчування відповідно до вимог користувача обрано модифікований генетичний алгоритм, який відповідає формальним критеріям.

Вхідними даними для алгоритму є [3]:

1. Вимоги щодо добової кількості макронутрієнтів (білків, жирів, вуглеводів).
2. Вимоги щодо кількості прийомів їжі.
3. Вимоги щодо макронутрієнтів відносно кожного прийому їжі.
4. Дані про страви – кількість макронутрієнтів та поєднуваність продуктів.

Кроки алгоритму:

1. Підготовка даних – побудова списків страв за кожним з нутрієнтів.
2. Визначення коефіцієнту зміни розміру порції відносно еталону. Доки не буде отримано необхідну кількість в прийоми їжі (відповідно до вимог щодо макронутрієнтів відносно кожного прийому їжі).
3. Формування початкової популяції прийомів їжі, що базується на генераторі випадкових величин. Першою до прийому їжі додається страва зі списку, що відповідає тому макронутрієнту, до якого висунуто найбільші, щодо кількості, вимоги.
4. Схрещення отриманих прийомів їжі.
5. Мутація – навмисне витіснення страв з прийому їжі, зокрема за такими ознаками, як поєднуваність між собою.
6. Відбір – знищення тих кандидатів, що показують гіршу відповідність вимогам. Доки не буде отримано необхідну кількість добових раціонів:
7. Формування початкової популяції добового раціону з отриманих на попередньому етапі прийомів їжі.
8. Схрещення кандидатів у добові раціони.
9. Мутація – заміна одних прийомів їжі на інші.
10. Відбір – знищення тих кандидатів, що мають недостатню відповідність вимогам до добового раціону харчування [4].

Результатом роботи розробленого алгоритму є набір добових раціонів. Кожен раціон складається з кількох прийомів їжі, в якому міститься декілька страв. Про кожну страву вказано розмір порції та кількість макронутрієнтів в ній.

Розроблено та протестовано модель, що надає такі можливості:

- побудова моделі користувача до раціону харчування на основі таких особистих даних: вік, стать, зріст, вага, рівень фізичної активності, мета харчування;
- побудова персонального раціону харчування на основі списку вимог користувача.

Тестування показало, що алгоритм демонструє необхідну точність та сформовані меню відповідають висунутим вимогам до раціону харчування людини.

Література. **1.** Calorie Distribution Throughout Day [Електронний ресурс]. – Режим доступу: http://forum.bodybuilding.com/showthread. **2.** Calorie Know-How: Get The Equation Right To Get Results! [Електронний ресурс]. – Режим доступу: http://www.bodybuilding.com. **3.** Calorie Distribution Throughout Day [Електронний ресурс]. – Режим доступу: http://forum.bodybuilding.com. **4.** Принцип оптимальности Беллмана для диннамического программирования [Електронний ресурс]. – Режим доступу: http://www.ngpedia.ru/id319149p1.html

Караюз І.В., Бідюк П.І.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Системний підхід до прогнозування макроекономічних процесів

Сьогодні існує необхідність виконання системного (поглибленого) аналізу фінансово-економічних процесів з метою побудови адекватних математичних моделей (ММ) для розв'язання задач прогнозування і прийняття управлінських рішень. При побудові ММ, як правило, зустрічаються невизначеності, зокрема у такому вигляді:

- невизначеності даних: неповнота даних (коротка вибірка); пропуски значень; недостовірні дані; похибки вимірів; вплив зовнішніх випадкових збурень;
- невизначеності структур і параметрів моделей: зміни структури представлення даних (дані збирає та оприлюднює Держстат [1] згідно з кодами класифікацій видів економічної діяльності (КВЕД), структура яких протягом останніх 20 років змінювалась тричі [2–5]); невизначеності ймовірнісного типу (стани досліджуваних процесів змінюються випадково); наявність та глибина причинно-наслідкових зв'язків відомі з деякою ймовірністю; невизначеності даних і структурні невизначеності спричиняють появу невизначеностей параметричного типу (зміщеність оцінок параметрів та зростання їх дисперсії);
- невизначеності оцінок прогнозів, зумовлені згаданими вище причинами та методичними похибками обчислень.

Побудова адекватних моделей ускладнюється також внаслідок наявності перехідних процесів в економіці (наприклад, змінами у законодавстві, зокрема у Податковому кодексі України [6]). Деякі засоби боротьби з невизначеностями: фільтр Калмана; мережі Байєса; ієрархічні байєсівські моделі; інтегровані моделі, які поєднують у собі моделі ідеологічно різних типів (рис. 1). Використання інтегрованих моделей надає можливість суттєво зменшити вплив випадкових факторів і покращити якість остаточних результатів: адекватність математичних моделей і якість оцінок прогнозів та рішень, які на них ґрунтуються.

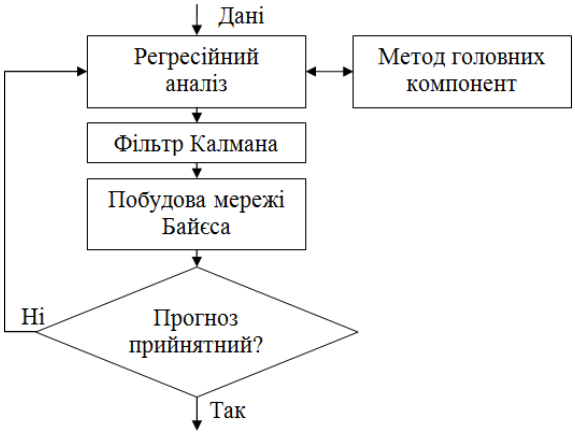


Рис. 1. Функціональна схема інтегрованої моделі

Література. **1.** Державна служба статистики України – Режим доступу: http://www.ukrstat.gov.ua/. **2.** Наказ Державного комітету України по стандартизації, метрології та сертифікації від 22.10.1996 № 441 "Про затвердження Державного класифікатора України". **3.** Наказ Державного комітету України з питань технічного регулювання та споживчої політики від 26.12.2005 р. №375 "Про затвердження національних, класифікаторів, зміни до державного класифікатора та скасування чинності державних класифікаторів". **4.** Наказ Державного комітету України з питань технічного регулювання та споживчої політики від 11.10.2010 р. № 457 національний класифікатор України "Класифікація видів економічної діяльності". **5.** Наказ Державного комітету статистики України від 30.08.2002 р. № 320 "Про затвердження Номенклатури продукції промисловості", зі змінами, затвердженими наказами від 29.12.2005р. №439, від 29.12.2007р. №477, від 25.12.2008р. №523, від 30.12.2009р. №503, від 01.12.2010р. №482 та від 23.12.2011р. №376. **6.** Податковий кодекс України : прийнятий Верховною Радою України 02.12.2010 р. № 2755-VI – Режим доступу: http://zakon0.rada.gov.ua/laws/show/2755-17/page.

Квитко М.В.
Институт прикладного системного анализа НТУУ "КПИ Киев, Украина

Распознавание речи с помощью глубоких рекуррентных нейронных сетей

В сегодняшнем мире технологий люди все чаще управляют электронными устройствами с помощью голоса. Начиная с команд вкл/выкл свет и заканчивая набором текста и работой с голосовыми помощниками на мобильных устройствах. В данной работе мы исследуем алгоритм распознавания речи, основанный на комплексной работе модифицированных рекуррентных нейронных сетей [2]. Для анализа и распознавания акустических моделей в основу современных рекурентных нейронных сетей (англ.: recurrent neural networks - RNN) легла нейросетевая темпоральная классификация (англ.: Connectionist Temporal Classification — CTC [1]) и дискриминантный анализ для последовательностей, адаптированный для обучения подобных структур. Данные RNN намного точнее, особенно в условиях посторонних шумов, а главное — они работают быстрее, чем все предыдущие модели распознавания. На сегоднешний день переход на технологию Глубоких Нейронных Сетей (англ.: Deep Neural Networks — DNN) осуществил настоящий прорыв, они лучше других методов подходят для распознавания отдельных звуков, произносимых пользователем, благодаря чему точность распознавания речи увеличилась [2].

Из чего состоит механизм распознавания. Улучшенная акустическая модель основана на Рекуррентных Нейронных Сетях (RNN). Их преимущество заключается в том, что они имеют циклы обратной связи в своей топологии, позволяющие им моделировать временные зависимости: когда пользователь произносит /u/ в слове «museum» (его фонетическая запись имеет вид: /m j u z i @ m/), то его речевой аппарат одновременно выходит из процесса произношения предыдущих звуков /j/ и /m/. Если попробовать сказать в слух — «museum», то можно заметить, что слово выходит на одном выдохе, и RNN могут это распознать.

RNN бывают разных типов, и для распознавания речи были использованы специальные RNN с «длинной кратковременной памятью» (англ.: Long Short-Term Memory — LSTM). Эти ячейки памяти и сложный механизм гейтов (Рис. 1) позволяют LSTM RNN лучше других нейронных сетей запоминать информацию [2].

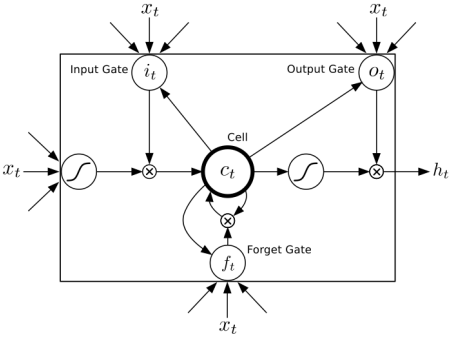


Рис. 1. Один блок LSTM

Данный подход был применен для двунаправленных рекуррентных сетей с несколькими внутренними слоями. Но основная проблема использования LSTM RNN несовпадение количества параметров на входе и выходе сети. Для решения этой проблемы была использована нейросетевая темпоральная классификация (CTC) [1].

С Нейросетевой темпоральной классификацией (CTC) модели научились выводить своеобразные «пики», которые и отображают последовательность различных звуков в звуковой волне. Они могут выделять различные фонемы в правильной с точки зрения языка последовательности звуков.

Результатом построения архитектуры, состоящей из акустической и лингвистической модели стали две RNN, одна из которых строит результаты на базе символов, а другая на базе звуков, при чем вторая модель использует результаты первой. Использование данных моделей позволило значительно улучшить результат распознавания речи [2].

Литература. **1.** A.Graves, S. Fern´andez, F.Gomez Connectionist Temporal Classification ftp://ftp.idsia.ch/pub/juergen/icml2006.pdf **2.** A.Graves, G.Hinton Speech recognition with deep recurrent neural networks, 22 May 2013. http://arxiv.org/pdf/1303.5778.pdf.

***Кондратенко Н.Р.*[1], *Снігур О.О.*[2]**
[1] *Вінницький національний технічний університет, Вінниця, Україна;* [2] *ТОВ "Дельфи Вінниця, Україна*

Використання інтервальних нечітких множин типу 2 в умовах недовизначеності вхідних даних

Проблеми раціонального природокористування та збалансованого використання природних ресурсів усе частіше розглядаються як такі, що вимагають вирішення на науковій основі, що включає пізнання об'єктивних закономірностей, які існують у природних системах та зумовлюють динаміку процесів, що протікають у цих системах. Як наслідок, виникає практична необхідність розробки математичних моделей, здатних із практично прийнятною достовірністю характеризувати природні процеси різних видів за певні періоди та інтервали часу. На практиці задача виявлення природних закономірностей являє собою системну задачу, що важко піддається формалізації [1]. Пропонується розглянути її на прикладі оцінювання стану артезіанської свердловини в довільний момент часу в межах від початку гідрогеологічної розвідки до її повного завершення. Дана задача 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 3 місяців до кількох років. Тому особливістю природної системи є недовизначеність, наявність пропусків у векторі даних, що її характеризує в кожен конкретний момент часу.

Недовизначеність є одним із НЕ-факторів [2], що визначають якість вихідної експериментальної вибірки. В ролі НЕ-факторів виступають неточність вимірювань, відсутність можливостей для спостереження компонент вхідного та вихідного векторів, статистична неоднорідність експериментальної послідовності, неповнота та неоднозначність знань про предметну галузь та розв'язувану задачу тощо [3].

Крім того, можливі різні порушення процесів збору інформації, що можуть стати причиною втрати даних, що суттєво впливають на результати аналізу [1]. Не слід також виключати ефекту неврахованих (прихованих) параметрів, що може вплинути на однозначність моделювання системи [4]. Статистичні моделі, що з успіхом застосовуються в таких задачах за умов визначеності вхідних даних, на недовизначених даних є непрацездатними. Особливо небезпечне застосування статистичних моделей на невеликих вибірках експериментальних даних, оскільки отримані на них закони розподілу можуть бути нестійкими [4]. Апарат інтервальних нечітких множин дає змогу будувати математичні моделі складних систем та процесів, здатні оперувати вхідними даними, що містять пропуски [5]. Важливою їх особливістю є відсутність обмежень на складність моделі, оскільки моделюється відгук системи на вектор вхідних даних, без необхідності побудови фізичної моделі внутрішніх процесів, що відбуваються в системі, та причинно-наслідкових зв'язків [4]. Як правило, для вивчення поведінки складних систем будується єдина модель та єдиний критерій, що характеризує невідповідність моделі та спостережень. В [6] показано, що такий підхід працює лише тоді, коли між входами та виходами системи є функціональна залежність, та спостереження проводяться точно. Якщо хоча б одна з цих умов не виконується, рекомендується побудова системи моделей та критеріїв відповідності. Природні системи та процеси належать саме до таких, в силу важкодоступності даних спостережень та неможливості гарантувати точність вимірювання кількісних показників. При цьому, відповідно до [6], чим складніша система та чим менша достовірність та/або доступність даних спостережень, тим більш відмінними будуть моделі, обрані за різними критеріями. Цей фактор також повною мірою стосується природних систем. З іншого боку, система моделей придатна для довгострокових кількісних оцінок, коли природні системи, що розглядаються, є стійкими в інтервалі спостережень. Гідрогеологічні системи володіють цією властивістю в зв'язку зі специфікою геологічних процесів: усталений комплекс умов формується протягом мільйонів років, існує тривалий час та не схильний до різких змін своїх параметрів [6]. Тому

має сенс побудова агрегованої моделі, що являє собою систему математичних моделей, кожна з яких має на меті виділення закономірностей природних процесів за даними дискретних спостережень за весь доступний період (рис. 1).

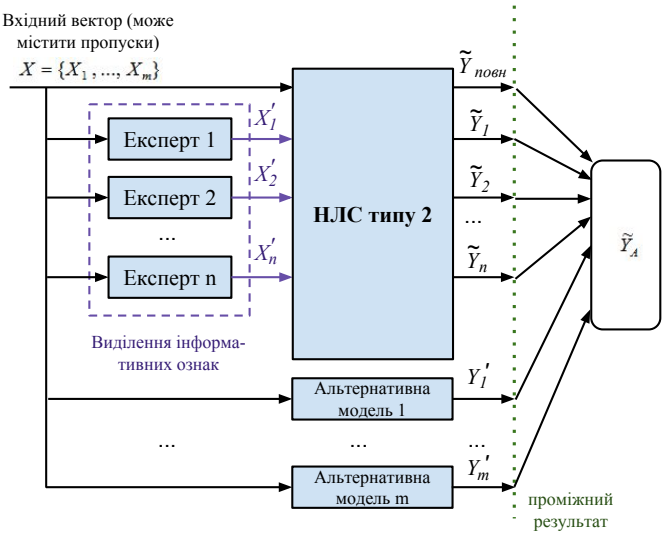


Рис. 1. Агрегована модель прийняття рішень в умовах недовизначеності вхідних даних

Реалізація єдиної агрегованої моделі прийняття рішення про стан природної системи має такі переваги порівняно з іншими існуючими підходами:

- Узагальнення та повторне використання обчислювальних методів, що використовуються для побудови виходу окремих моделей
- Скорочення проміжку часу, необхідного для прийняття практичних рішень у сфері природокористування, раціонального освоєння та охорони природних ресурсів
- Забезпечення своєчасного прийняття рішень у динаміці природного процесу, та достатньої достовірності прийнятого рішення.

Література. **1.** 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. **2.** Нариньяни А.С. Неопределенность в системе представления и обработки знаний // Изв. АН СССР. Техническая кибернетика. – 1986 – №5. – С. 8-11. **3.** 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. **4.** Минаев Ю.Н. Методы и алгоритмы решения задач идентификации и прогнозирования в условиях неопределенности в нейросетевом логическом базисе. – М.: Горячая Линия-Телеком. – 2003. – 205 С. **5.** Кондратенко Н.Р. Діагностика гіпотиреозу на основі нечіткої логіки з використанням інтервальних функцій належності / Н.Р. Кондратенко, Н.Б. Зелінська, С.М. Куземко // Наукові вісті національного технічного університету України „Київський політехнічний інститут". – 2003. – №4. **6.** Ивахненко А.Г. Моделирование сложных систем по экспериментальным данным. – М.: Радо и связь, 1987. – 120 С.

Кондратьєв І.І.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Трекінг швидких об'єктів на відео

Комп'ютерний зір є відносно новою областю сучасної науки, стрімкий розвиток якої почався всього лиш два десятиліття назад. Розвиток даної галузі пов'язаний, в першу чергу, зі збільшенням доступності систем відеоспостереження, що призвело до їх масового використання у повсякденному житті, а також зі зростанням потужностей обчислювальних пристроїв, що дозволило втілити алгоритми розпізнавання та супроводження на практиці.

Важливою задачею в області комп'ютерного зору є задача супроводження об'єктів у відеопотоці. В загальному вигляді задача полягає в знаходженні положення цільового об'єкта на послідовності відеокадрів при відомому положенні на одному або декількох перших кадрах послідовності. Положення може визначатись різними способами, наприклад:

- точка - центр об'єкта,
- контури об'єкта,
- обмежуючий прямокутник або еліпс [1].

Найбільшу практичну цінність мають алгоритми, що забезпечують швидкість роботи в реальному часі. Завдяки величезному розмаїттю об'єктів реального світу, загальний розв'язок даної задачі є дуже складним, тому на практиці зазвичай використовують алгоритми, націлені на супроводження конкретного класу об'єктів. В даній роботі розглядається задача трекінгу в контексті класу швидких об'єктів. Об'єкт вважатимемо швидким, коли для нього виконуються такі умови:

- між двома послідовними кадрами об'єкт може переміститись на відстань, більшу за власні лінійні розміри,
- об'єкт може зникати з поля зору супроводжуючої системи.

Застосування алгоритму, що розв'язує дану задачу може бути дуже різноманітним, починаючи від спортивних змагань (автоматичне стеження за футбольним чи тенісним м'ячем для допомоги арбітрам), продовжуючи застосуванням на дронах чи коптерах, які можуть використовуватись, наприклад, для супроводження автомобілів чи очищення злітних смуг від птахів, і закінчуючи застосуванням у авіаційній галузі для стеження за літаками.

В роботі розглядається алгоритм, що базується на підході Tracking-Learning-Detection і відомий під назвою "Predator" [2]. Сильними сторонами даного алгоритму є здатність до тривалого стеження, адаптація до зміни розмірів та можливість пошуку об'єкта у кадрі після втрати з поля зору. Однак, на практиці даний алгоритм досить часто показує недостатню швидкість роботи, що не дає можливості повноцінно використовувати його у системах стеження в реальному часі. Розглядаючи лише клас швидких об'єктів, до алгоритму можна застосувати певні модифікації, що дозволяють покращити показники якості роботи.

Для локалізації областей, що можуть містити об'єкт спостереження на поточному кадрі, використовується алгоритм розрахунку оптичного потоку, розроблений Гуннаром Фарнебаком [3]. Завдяки даному кроку, можна зменшити навантаження на систему розпізнавання.

Для пришвидшення пошуку об'єкта на поточному кадрі також можна застосувати структуру даних R-Tree, що дозволяє визначати покриття локалізованих областей прямокутниками швидше, ніж у випадку повного перебору всіх варіантів, за умови, що область переднього плану має відносно невеликі розміри (іншими словами, відеоряд має статичний фон).

Література. **1.** Лукьяница А. А. Цифровая обработка видеоизображений / А. А. Лукьяница, А. Г. Шишкин. – М.: Ай-Эс-Эс Пресс, 2009. – 518 с. **2.** Kalal Z. Tracking-Learning-Detection / Z. Kalal, K. Mikolajczyk, J. Matas. // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2012. – №7. – С. 1409–1422. **3.** Farnebäck G. Two-frame motion estimation based on polynomial expansion / Gunnar Farnebäck. // SCIA'03 Proceedings of the 13th Scandinavian conference on Image analysis. – 2003. – С. 363–370.

Кравець І.О., Козловська В.О., Тумко В.С.
Чорноморський державний університет ім. П.Могили, Миколаїв, Україна

Система підтримки прийняття рішень для трейдера на базі розпізнавання фігур майбутнього розвороту тренду та створення бази правил для нечіткої нейронної мережі

Дослідження, що проводилися в останні десятиліття, показали практичну цінність застосування математичних методів прогнозування, розподілу капіталу та аналізу ризиків для аналізу ринкової ситуації. Поряд з активним використанням фундаментального аналізу трейдери потребують програмних продуктів, здатних розв'язувати ці завдання. Більшість відомих методів аналізу вже є вбудованими в окремі торгівельні термінали, а інші реалізуються в окремі програмні додатки. Одначе, існуюче програмне забезпечення (додатки до торгових платформ) дають запізнення у прогнозі і не завжди сприяють прийняттю вірного рішення. Поряд з технічними проблемами та вимогами існують й інші, з якими стикаються трейдери. Це – дія психологічних факторів на процес трейдингу, брак часу для постійного слідкування за ринковою ситуацією, висока волатильність короткострокових ринків. Тому дуже доцільне використання методів штучного інтелекту у процесі біржового трейдінгу [1].

В роботі запропоновано систему прийняття рішень для біржового трейдера на базі розпізнавання фігур майбутнього розвороту тренду на графіках котирувань, фрактального аналізу, та нечітких нейронних мереж. Застосування фрактального анлізу для фінансових часових рядів дає можливість визначити зони турбулентності та персистентності тренду і передбачити зміну тренду [2]. Розпізнавання фігур майбутнього розвороту тренду на графіках біржових котирувань та побудова бази правил також дають можливість розпізнати поворот в напрямку тренду швидше ніж адаптивні методи, які завжди дають затримку на один часовий крок. Адаптивні методи прогнозування визначають загальний напрямок тренду, який зберігається протягом останніх часових кроків. Ці дані (напрямок тренду, показник Херсту, наявність фігур розвороту тренду) є вхідними для бази правил, яка вказує на можливість майбутнього розвороту тренду и використовується у нечіткої нейронної мережі, яка формує рішення о майбутньої зміні напрямку тренду. Це допомагає трейдеру своєчасно прийняти рішення щодо відкриття і закриття позиції..

На підставі наведеного вище у запропонованої СППР:.

- для задачі розпізнавання фігур майбутнього розвороту тренду на графіках біржових котирувань використано чітку нейронну мережу багатошаровий персептрон;
- розроблено алгоритм підрахунку показника Херста ;
- створена база правил для нечіткої нейронної мережі NEFCLASS, яка використовує показник Херста та розпізнавання фігур майбутнього розвороту тренду на графіках біржових котирувань у якості вхідних даних
- розроблена нечітка нейронна мережа NEFCLASS, яка використовує базу правил для індикації зміни напрямку тренда;
- розроблено інтерфейс під'єднання вхідних та вихідних даних нечіткої нейронної мережі до програмного блоку розпізнавання фігур біржових котирувань;
- розроблено програмне забезпечення, що дозволяє передбачити зміну напрямку тренду.

Література. **1.** Зайченко Ю.П. Нечеткие модели и методи в интеллектуальних системах.-К., „Слово", 2008.- 333 с. **2.** Петерс Э. Фрактальный анализ финансовых рынков. Применение хаоса в инвестициях и экономике. – М.: Интернет-трейдинг, 2004 **3.** 22. Программное обеспечение для трейдеров: [Электронный ресурс]. https://freshforex.org/encyclopedia-forex/Software-for-the-currency-market/.

Курінний І.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Композиційні методи класифікації та регресії

Композиційні методи класифікації та регресії демонструють гарні результати при роботі з дуже складними за структурою та великими за об'ємом даними. Дані алгоритми мають високу точність прогнозування та є стійкими до перенавчання. Ідея методів навчання ансамбля полягає в тому, що із простору гіпотез потрібно вибрати цілу колекцію, або так званий ансамбль гіпотез, і в подальшому комбінувати передбачення, отримані за допомогою гіпотез цього ансамбля. Наприклад, може бути сформовано сто різних дерев рішень із однієї навчальної множини, після чого проведено голосування для визначення найкращої класифікації нового прикладу. В основі прагнення використовувати навчання ансамбля лежить проста причина. Розглянемо ансамбль із $M = 5$ гіпотез і припустимо, що їх передбачення комбінуються з використанням нескладного мажоритарного голосування. Для того, щоб цей ансамбль неправильно класифікував новий приклад, його повинні неправильно класифікувати щонайменше три з п'яти гіпотез. Проте можна розраховувати на те, що дана ситуація є набагато менш ймовірною у порівнянні з помилковою класифікацією при використання єдиної гіпотези. Нехай кожна з гіпотез h_i в ансамблі допускає помилку з ймовірністю p. Крім того припустимо, що помилки, допущені із застосуванням кожної гіпотези, є незалежними. В такому разі, якщо ймовірність p є малою, то ймовірність одночасної появи великої кількості помилок стає мікроскопічною. Простий розрахунок показує що використання ансамбля із п'яти гіпотез дозволяє скоротити частоту помилок від величини $\frac{1}{10}$ до величини менше ніж $\frac{1}{100}$. Тим не менше очевидно, що припущення щодо незалежності гіпотез є неоправданим, оскільки у всіх гіпотезах, скоріш за все, будуть виникати нові спотворення, викликані одними і тими ж спотворюючими їх аспектами однакових навчальних даних. Якщо ж гіпотези відрізняються між собою, то це призводить до зменшення кореляції між їх помилками, то навчання ансамбля може виявитися дуже корисним [1].

Найбільш широко використовується метод формування ансамбля, який називається підсилення. Він базується на ідеї зваженої навчальної множини. В такій навчальній множині з кожним прикладом пов'язана вага $w_i \geqslant 0$. Чим більша вага прикладу, тим вище важливість, яка присвоєна йому процесі вивчення якоїсь гіпотези. Процедура підсилення починається із задання w_i для всіх прикладів, тобто зі звичайної навчальної множини. На основі цієї множини створюється перша гіпотеза h_1, яка класифікує одні навчальні приклади правильно, а інші – неправильно. Бажано, щоб наступна гіпотеза краще справлялася з неправильно класифікованими прикладами, тому ваги останніх збільшуються, а ваги правильно класифікованих прикладів зменшуються. По цій навчальній множині із перерахованими вагами створюється гіпотеза h_2. Описаний процес продовжується таким чином до тих пір, поки не буде створено M гіпотез, де M стає вхідним параметром для алгоритму підсилення. Остаточна гіпотеза-ансамбль представляє собою зважену мажоритарну комбінацію із всіх M гіпотез, кожній з яких присвоєна вага, яка відповідає тому, на скільки високу продуктивність вона показала при опрацюванні навчальної множини. По мірі збільшення розміру ансамбля M виявляється дуже цікава властивість композиційних методів. Вона полягає в тому, що при введені додаткових гіпотез в уже сформований на попередніх кроках ансамбль помилка на тестовій множині продовжує зменшуватися впродовж тривалого часу, при тому як помилка на навчальній множині практично прямує до нуля. Для пояснення цього феномену було запропоновано декілька трактувань. Одне з можливих пояснення полягає в тому, що введення додаткових гіпотез дозволяє досягти того, що ансамбль проводить все більш чіткіші границі між класами, а ця властивість сприяє кращій класифікації нових прикладів [2].

Література. **1.** Рассел С. Искусственный интеллект: современный подход / С. Рассел, П. Норвиг, — М.: Издательский дом "Вильямс", 2006 — 1408 с.
2. Hastie T. The elements of statistical learning 2nd ed. / Hastie T., Tibshirani R., Friedman J. — Springer, 2009 — 745 p.

Мейтус В.Ю.
Международный научно-учебный центр информационных технологий и систем НАНУ, Киев, Украина

Знания в 1-интеллектуальных системах

Переход к созданию и использованию интеллектуальных машин и систем отражает новый этап в развитии современной цивилизации, и качественный скачек в истории человечества. А создание таких машин и систем определяет один из главных путей совершенствования технологий решения задач и организации сложного поведения, связанных с нашей цивилизацией. Поэтому такой актуальной является разработка методов построения систем, поведение которых определяется как интеллектуальное в той предметной области (ПО), в которой действует эта система.

Исследования по проблемам интеллекта показывают, что, не существует единого определения понятия «интеллект», хотя можно сказать, что **интеллект – это свойство высокоорганизованного субъекта, позволяющее ему моделировать ПО и решать задачи, относящиеся к этой области на уровне создаваемой модели. Тогда уровни интеллекта конкретного субъекта определяются тем, какие ПО и насколько адекватно он может моделировать, и тем, какие задачи в этих областях он решает.**

Интеллект нулевого уровня (0-Интеллект) – это способность решать задачи в заданной модели ПО. Сама модель предполагается заданной и неизменной. Такой интеллект определяется методами решения задач в ПО и способами их комбинирования. Это вид интеллекта достаточно продуктивен, и используется человеком в процессе его обучения и формировании его последующего поведения.

Более сложным является интеллект первого уровня (1-Интеллект). Приведем его определение из работы [1, 2]. 1-Интеллект – это неотъемлемое качество субъекта, ориентированного на взаимодействие с различными ПО, составляющими внешнюю среду. Это качество позволяет субъекту адекватно моделировать ПО и эффективно решать задачи, поставленные в этой области. Для этого субъект использует свой когнитивный опыт, сохраненный в виде системы знаний о свойствах ПО и способах решения задач в ней, расширяемый за счет применения алгоритмов анализа внешней и внутренней информации с последующим синтезом знаний о ПО на основе этой информации. Поскольку в дальнейшем рассматривается только интеллект этого уровня, то символ «1-» будем опускать.

1-Интеллект в отличие от 0-Интеллекта предполагает возможность внесения изменений в модель ПО, корректировку модели, модификацию ее составляющих, установление новых связей между элементами модели. Дальнейшие существующие уровни интеллекта пока не будем рассматривать.

И модель ПО, и существующие методы решения задач в ней задаются и интеллектуальных системах в виде знаний. Потому представление знаний является ключевым элементом в создании любой интеллектуальной системы

***Интеллектуальная система* (ИС) – это система, предназначенная для достижения в ПО некоторой цели (исполнения заданной программы действий, решения проблем в заданной области, сохранения некоторых параметров ПО), и выполняющая свои действия, опираясь на присущий ей интеллект. Поэтому можем говорить о 1-интеллектуальных системах, определяемых 1-интеллектом.**

В ИС знания сохраняются в базе знаний (БЗ), состоящей из банка знаний и системы управления базой знаний. Эти знания являются основой построения модели ПО. Изменения в модель вносятся на уровне БЗ. При решении задачи сначала из знаний воссоздается та часть ПО, в которой решается задача, а затем уже отыскиваются составляющие самого процесса решения задачи. В [1] было предложено задавать отдельные знания в виде базовой конструкции, состоящей из четырех компонентов: *экстенсионала, интенсионала, коннотации и типа.*

Под *интенсионалом* понимается класс, совокупность признаков или классов признаков, характеризующих данное знание. Под *экстенсионалом* понимается множество утверждений, примеров, выражений, в которых может быть использована данная форма знания. Соответственно, с интенсионалом связываются интенсиональное представ-ление знания, а с экстенсионалом – экстенсиональное представление. Компонент *коннотация* – это языковое задание подмножества признаков, которые, не включаясь в интенсионал, связаны со знанием в естественном языке, приписываются ему в силу различных ассоциаций – познавательных (знаний о ПО), эмоциональных, стилевых (по принадлежности к тому или иному стилю речи).

Для задания каждой компоненты (исключая тип) используется свой логический язык, создаваемый на основе модальной логики, и пополняемый дополнительными операторами. Этим языкам сопоставляется соответственно интенсиональная, экстенсиональная и коннотативная логики. На основе этих логик определяются способы формирования новых знаний, соответствующих решениям задач.

Экстенсиональная логика позволяет вычислять значение истинности некоторой фразы и совокупности фраз, представленных в виде логической формулы, по отношению к заданной модели, представляющей ПО. Интенсиональная логика определяет смысл, понимание каждой фразы, также по отношению к модели, построенной по ПО. Наконец, коннотационная логика определяет смысловую связь между различными элементами ПО, отражаемую в модели этой области. [3]

Использование базовой конструкции для каждого знания в общей схеме представления знаний задает аппарат оперирования со знаниями, связывающий знания как элементы представления модели ПО в единой форме, позволяющей решать задачи, относящиеся к этой области. Это своего рода метрика, вносимая в пространство,построенное из отдельных знаний. При этом решение согласуется конструктивными особенностями этой области и со смыслом соответствующих выражений. Одновременно, новые знания вписываются в корректируемую модель, что позволяет относить интеллект систем, основанную на таком представлении, к 1-Интеллекту.

Используемые при построении отдельных элементов знания логические языки рассматривают высказывания, каждое из которых представляется предикатом, аргументы которых являются индивидами константами, определяемыми в ПО. Соответственно высказывания характеризуют отношения между этими константами. Введение семантических отношений позволяет приблизить операции со знаниями к тем отношениям, которые использует человек, рассматривая соответствующие выводы. В этом случае можно ожидать, что интеллектуальные системы, база знаний которых будет основана на таком представлении, смогут оперировать с семантической информацией, представленной знаниями, повысив свой уровень интеллекта в общении с человеком.

Литература.

1. Мейтус В.Ю. Введение в теорию интеллектуальных систем. Основные представления. Palmarium academic publishing, 2015. -189 с. ISBN 978-3-659-60297-9

2. Мейтус В.Ю. Проблемы разработки интеллектуальных систем // 21-я Международная Крымская конференция “СВЧ-техника и телекоммуникационные технологии”, 12-16 сентября 2011 г. Том 1. - 2011. - С.43-46. ISBN 978-966-335-351-7

3. Логический подход к искусственному интеллекту: от модальной логики к логике баз данных. Пер.с франц. / Тейз А., Грибомон П., Юлен Г. и др. - М.: Мир, 1998. - 494 с. ISBN 5-03-002519-7

***Мейтус В.Ю.*[1], *Цепкова Д.П.*[2]**
[1]*Международный научно-учебный центр информационных технологий и систем НАНУ, Киев, Украина;* [2]*НУ «Киево-Могилянская академия, Киев, Украина*

База знаний с неполной информацией представления знаний

Проблема создания интеллектуальных систем включает в себя необходимость решения целого ряда промежуточных задач, которые являются составными частями этой проблемы. Общая схема построения такой системы, представленная в работе [1], основывается на том, что интеллектуальная система (ИС) всегда связана с одной или несколькими предметными областями (ПО), в которой система должна решать задачи, возникающие в этой области. А интеллект как свойство, определяющее возможность поиска решений, непосредственно используется с этой целью.

Каждая ПО в окружающем нас мире бесконечно разнообразна по своим свойствам и качествам. Поэтому с ней могут быть связаны разнообразные модели, каждая из которых отражает только некоторые черты ПО, достаточные для решения отдельных классов задач. Модель заменяет ПО, а связь между моделью и ПО задается интерпретацией модели в ПО. В основе модели лежит информация, воспринимаемая и используемая при построении модели. Эта информация задается в виде знаний, которые рассматриваются как составные элементы при построении модели ПО.

Любая задача в ПО представляется как пара точек (или областей) в пространстве, связанном с моделью и интерпретируемом в ПО. Одна из одна из этих точек представляет собой условие задачи, а вторая – результат ее решения. Само же решение рассматривается как путь в ПО, связывающий условие задачи с ее решением. Обычно это решение сначала находится в модели ПО, а затем переносится (рассматривается) как решение в ПО. Если модель была адекватна ПО, то и решение, полученное по модели, будет соответствовать этой области, будет решением задачи, относящейся к ПО.

Как **знания *будем рассматривать совокупность фактов, утверждений (одно или одновременно несколько), форм, представлений, схем, моделей, которые задают объекты ПО, образы, структуры и комбинации элементов в ПО, структурные связи (зависимости, отношения) между объектами и структурами, геометрию ПО*** .

Поскольку знания играют основополагающую роль при построении модели и поиске решения задач, т.е. при выполнении основных функций ИС, то необходимо их получать, сохранять, изменять и выдавать при формировании поведения системы. Для этого в ИС существует ***база знаний*** (БЗ), которая состоит из двух компонент – банка знаний (БкЗ) и системы управления базой знаний (СУБЗ). В банке знаний накапливаются сами знания, в СУБЗ выполняются все операции, связанные с управлением знаниями. В зависимости от того, как представлены знания, определяются и операции над ними, сам процесс использования знаний для решения задач ИС.

В работе [1]предлагается рассматривать знания в виде именованной четырехкомпонентной структуры, включающей для каждого знания, сохраняемого в БЗ, следующие компоненты представления знаний: ***тип, интенсионал, экстенсионал и коннотацию*** знания. Для задания типов применяется алгебра типов, а для каждой из остальных компонент представления знания используется свой логический язык. В основу этих языков положен язык высказываний L_0, в котором высказывания записываются в виде предикатов.

Для БЗ индивидные константы – это элементы словаря БЗ, интерпретируемые как составляющие объекты ПО. Кроме того, к индивидным константам относятся и обозначения

неопределенных величин ω_i $(i = 0, 1, \dots)$, которые используются в тех случаях, когда отсутствует полная информация, относящаяся к некоторому предикату. Наши знания о ПО связываются с теми наборами констант, для которых предикаты и построенные из них логические формулы оказываются истинными.

Отдельные высказывания характеризуют различные элементы общей ситуации, которая описывается в БЗ. Эти высказывания объединяются между собой с помощью обычных логических операций – конъюнкции и дизъюнкции, отрицания, импликации. Но поскольку используются индивидные константы, то необходимы и соответствующие переменные, которые связываются с помощью кванторов существования и всеобщности. Дополнительно вводятся модальные операции возможности и необходимости. А затем полученная формула оценивается с точки зрения семантики соответствующего выражения.

Выражения языка связываются с ПО, в которой рассматриваются модели, построенные из составляющих ПО и отношений между ними (структуры для некоторой части ПО). В такой модели построенные выражения соответствующего языка должны быть истинными. Иначе отсутствовала бы адекватность знаний и ПО. Тогда можно ввести определения формулы, истинной в модели, истинной в структуре, и общезначимой.

Отметим, что возможна не только модальная логика для анализа имеющихся знаний. Существуют и другие логики, например, логика веры и знаний, динамическая логика, временная логика. Каждая из них определяет свой подход к анализу ПО и знаний о ней.

Вычисление семантического значения формулы, соответствующей условию задачи, не всегда возможно, например, в тех случаях, когда в формулу входят кванторы существования, или если значение некоторого предиката оказывается неизвестным в данной ПО. Эта ситуация соответствует неполной информации для данной БЗ. Для преодоления этой проблемы возможны несколько вариантов. Во-первых, поиск дополнительной информации, используя коннотативную логику, с последующим преобразованием рассматриваемой модели. Во-вторых, включение нескольких моделей, в которых рассматриваются разные варианты пополнения информации. В-третьих, введение в семантику неопределенных констант и рассмотрение решения в этом виде. В последующем константы могут рассматриваться как некоторые гипотезы в семантике данного выражения.

Неполная информация может в некоторых случаях дополнительно анализироваться с использованием коннотативной логики, которая позволяет искать информацию по аналогии с другими похожими объектами ПО. Сталкиваясь в некоторой модели с неполной информацией на основе зафиксированных в коннотации связей с другими объектами ПО анализируется дополнительное ограничение на переменную, значение которой отсутствует в рассматриваемой ситуации. Происходит модификация знаний, их пополнение на основании гипотезы о схожести объектов в ПО.

Литература.

1. Мейтус В.Ю. Введение в теорию интеллектуальных систем. Основные представления. Palmarium academic publishing, 2015. -189 с. ISBN 978-3-659-60297-9

2. Логический подход к искусственному интеллекту: от модальной логики к логике баз данных. Пер.с франц. / Тейз А., Грибомон П., Юлен Г. и др. - М.: Мир, 1998. - 494 с. ISBN 5-03-002519-7

Миселюк А.Ю.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Глибинні нейронні мережі в медицині

Серцево-судинні захворювання (ССЗ) є однимb із найбільш поширених причини смерті у світі і становлять приблизно 30-40% смертей у всьому світі [1]. Зростання захворюваності ССЗ привернула увагу у всьому світі до наукових досліджень і розробки автоматизованих інструментів для виявлення порушень серцевого ритму з великою точністю. У даній роботі розглядається проблема діагностики ССЗ за допомогою аналізу електрокардіограм. Для вирішення даної проблеми розглядається модель на основі глибинного навчання. Аналізуються результати моделювання та описуються можливі варіанти покращення ефективності роботи моделі.

Для навчання моделі використано базу даних електрокардіограм PhysioNet. Очищення вхідних даних від шумів було виконане за допомогою адаптивного фільтру та фільтру низьких частот. Під час попередньої обробки електрокардіограм було виділено серцеві цикли, застосувавши алгоритм Pan-Tompkins [1].

Традиційним інструментом для аналізу сигналів є віконне перетворення Фур'є (ВПФ). Недоліком ВПФ є вибір сталого вікна, що може спричинити погіршення частотних або часових характеристик. Для уникнення даної проблеми було використано вейвлет перетворення. Вейвлет перетворення для безперервного сигналу $f(x)$ щодо вейвлет функції ψ визначається наступним чином:

$$[W_\psi f](a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \psi^* \left(\frac{x-b}{a} \right) f(x) dx. \quad (1)$$

де ψ^* означає комплексне спряження для ψ, параметр $b \in \mathbb{R}$ відповідає тимчасовому зміщенню, і називається параметром положення, параметр $a > 0$ задає масштабування і називається параметром розтягування.

Для побудови моделі на основі глибинних нейронних мереж, здатних виявляти захворювань на електрокардіограмах, використовуються згорткова нейронна мережа (англ.: Convolutional Neural Network, — CNN), багатоярусні автокодувальники (англ.: Stacked Autoencoders) і глибинні машини Больцмана (англ.: Deep Boltzmann Machine — DBM) [2,3]. Навчання таких моделей, як правило, здійснюється за допомогою двох підходів. У першому використовується навчання із вчителем. Серед різних архітектур, які використовують такий підхід, високу продуктивність мають CNN. У порівнянні зі стандартними нейронними мережами прямого поширення, вони мають згорткові шари, що складаються з невеликої кількості нейронів, які приймають на вхід обмежену область даних. Ці шари мають загальні ваги, таким чином, процес навчання еквівалентний процесу навчання фільтрів, які застосовуються для кожної області даних. Через невелику кількість з'єднань, вони навчаються більш ефективно, і в той же час ці фільтри автоматично виділяють ознаки з даних.

Друга концепція навчання базується на виявлення ознак без учителя (англ.: Unsupervised Representation Learning). В даний час в основному використовуються три основні підходи: МГК, автокодувальники і глибинні машини Больцмана. Автокодувальники та глибинні машини Больцмана мають багатошарову архітектуру, представлену "видимим"шаром, які відповідає вхідним даним, і «прихованих» шарів, які відповідають ознакам даних. Основна відмінність між цими підходами, що в машині Больцмана приховані блоки розглядаються як приховані випадкові величини, а в автокодувальника вони розглядаються як обчислювальні вузли.

Література. **1.** R. Martis , C. Chakraborty, A. Ray Wavelet-based Machine Learning Techniques for ECG Signal Analysis, 10 Dec 2013. **2.** A. Ignatov Deep Learning in information analysis of electrocardiogram signals for disease diagnostics **3.** X. Jia, K. Li, X. Li, A. Zhang A Novel Semi-supervised Deep Learning Framework for Affective State Recognition on EEG Signals http://www.cse.buffalo.edu/DBGROUP/bioinformatics/papers/cameraready_xwjiabibe.pdf

Мурга Н.А.
Учебно-научный комплекс "Институт прикладного системного анализа" Национального технического университета Украины "Киевский политехнический институт", Украина

Принятие решение об открытии позиции по финансовому инструменту и её структуре на основе прогнозирования и вычисления статистических характеристик временного ряда

Доклад посвящён описанию результатов применения подхода к принятию решений относительно открытия предварительно спрогнозированных прибыльных позиций по финансовым инструментам, который основан на расчёте статистических характеристик временных периодов, в течении которых происходит торговля.

Выделение прибыльных позиций с использованием прогнозирования. Использование авторегрессионных линейных моделей ([1]) и нейронных сетей ([2]) применительно к прогнозированию временных рядов финансовых инструментов, позволяет выделять их колебания, которые могут произойти в ближайших тайм-фреймах в будущем и являются прибыльными.

Однако, в следствие высокой волатильности большинства финансовых инструментов, сделанные прогнозы могут оказаться ошибочными и наносить серьёзные уроны капиталу трейдера.

Таким образом, кроме выделения прибыльных позиций, необзодимо производить оценку их возможности на основании некоторых показателей.

В данной работе разрабатывается и исследуется критерий, который позволяет численно оценить перспективность выделенной позиции и, как результат, автоматически принять решений по ней.

Критерий основывается на оценке отклонения текущего состояния торговли от типичного на торговом временном интервале. И если такое отклонение существенно, позиции предлагается не открывать.

Критерий принятия решения. Пусть выполнен прогноз значения финансового инструмента на основании регрессионной модели его временного ряда, который показал прибыльность возможной открытой позиции по нему. Для принятия решения по по возможности открытия и её структуре необходимо ответить на ряд вопросов:

1. Не было ли уже открыто чрезмерное количество позиций по финансовому инструменту за данный временной период, что делает маловероятным прогнозируемый скачёк цен?
2. Насколько правдободобна прогнозируемая ампитуда скачка?
3. Насколько правдоподобно, что скачёк цен будет происходить в течении спрогнозированного временного периода?
4. Как корректно выставить ордер Stop-Loss?

Очевидно, что ответы на эти вопросы должны основываться на стойких статистических характеристиках временного ряда финансового инструмента, полученных из его предыстории. В работе к рассмотрению предлагается следующее:

1. Максимальное, среднее и минимальное количество оптимальных позиций, которые можно было бы открыть в течении торгового дня.
2. Максимальное, среднее и минимальное количество оптимальных позиций, которые можно было бы открыть в течении торгового дня недели.
3. Максимальная, средняя и минимальная величина Take-Profit и Stop-Loss ордеров для оптимальных позиций, которые можно было бы открыть в течении торгового для.
4. Максимальная, средняя и минимальная величина Take-Profit и Stop-Loss ордеров для оптимальных позиций, которые можно было бы открыть в течении торгового для недели.
5. И аналогичные характеристи для подпериодов торогового дня (дня недели).

Имея значения данных характеристик, можно сделать заключени о том, не вступает ли полученный прогноз временного ряда в противоречие с ними и безопасно ли открывать позицию.

Струкрура исследований. Для проведения исследований были временные ряды валютных пар: AUD/USD, NZD/USD, EUR/USD, GBP/USD, USD/CHF на 15-ти, 30-ти, 45-ти минутных, часывых, 2-х, 3-х, 4-х часовых тамфреймах.

Для определения моментов времени открытия и закрытия оптимальных позиций по данных финансовым инструментам была поставленна задача оптимизации, которая позволяет их вычислять в течении расчётного периода (торгового дня):

$$max \sum_i \sum_j x_{ij} r_{ij}$$
$$\prod_{k=i+1}^{j-1} \prod_{l=k+1}^{j-1} x_{kl} + x_{ij} \leqslant 1$$
$$x_{ij} x_{kl} = 0, k - i > l$$
$$x_{ij} \in \{0, 1\}$$

, где x_{ij} - факт проведения, либо не проведения позиции по финансовому инструменту в период времени $[0, 1]$, r_{ij} - доходность, которая будет реализована по данной позиции. Первое ограничение устанавливает факт того, чтобы вложенные позиции для данной не реализовывались, а второе - позиции не пересекались во времени. Целевая функция определяет процес оптимизации в сторону максимизации доходности за временной период.

Эксперименты проводятся следующим образом. Выборка значений временного ряда отдельно взятого финансового инструмента делится на периоды длины: торговый день, первая и вторая половина торгового дня, четверти торгового дня, восьмые части. На основании значений из выделенных периодов, расчитываются r_{ij} для их подпериодов и решается задача оптимизации, благодаря которой выделяются моменты времени открытия и закрытия оптимальных позиций в течении выделенного периода.

Используя полученные значения моментов времени открытия и закрытия позиций по финансовому инструменту, расчитываются указанные выше статистические характеристики.

После расчёта данных харакретистик, берётся продолжение временного ряда, на котором выполняется прогнозирование его значений и, в случае выделения прибыльных позиций, выполняется оценка их перспективности на основании приведённого выше набора показателей. В качестве критерия отбора потенциально прибыльных взят следующий:

$$\prod_i \frac{p_i - \overline{p_i}}{p_i^{max} - p_i^{min}},$$

где p_i - реальное значение показателя, $\overline{p_i}$ - среднее значение, p_i^{max} - максимальное, p_i^{min} - минимальное.

Пороговое значение данного критерия (значение выше которого позиция не открывается) расчитывается на той же выборке данных, на которой выполнялся расчёт статистических характеристик.

Качество рассмотренного подхода численно оценивается как отношение дохода, полученного на продолжении временного ряда благодаря его применению, к оптимальному, максимально возможному, доходу на данной выборке значений.

Уровни Take-Profit и Stop-Loss. В качестве ордера Take-Profit предлагается брать прогнозное значение, однако если оно не больше максимального по доходности на предыстории цен финансового инструмента (в этом случае рекомендуется брать именно максимальное значение) - это можно назвать подходом оптимиста. Однако, кроме данных двух видов установки Take-Profit, естественно выделить ещё два: торговля в среднем (выставлять среднее значение Take-Profit, полученное из предыстории) и использование минимального значения - торговля пессимиста.

В качестве уровня Stop-Loss можно взять его среднестатистическое значение и при сильных колебаниях установить в максимальное, полученное из обучающей части временного ряда.

Литература. **1.** Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів (навчальний посібник) - К.: Політехніка, 2010. - 317 с. **2.** Згуровский М.З., Зайченко Ю.П. Основы вычислительного интеллекта. - К.: НВП «Видавництво "Наукова думка" НАН України», 2013. - 408 с.

***Новицкая Е.Г.*[1], *Железко Б.А.*[2]**
[1]*Гродненский государственный университет имени Янки Купалы, Гродно, Беларусь;* [2]*Белорусский государственный экономический университет, Минск, Беларусь*

Интеллектуальная система поддержки принятия решений как элемент регионального электронного правительства

Введение. Концепция электронного правительства стала активно распространяться с начала 2000-х годов. В постсоветских странах, этот процесс начался ближе к концу первого десятилетия XXI века. На сегодняшний день в Беларуси и соседних странах завершена стадия формирования соответствующей инфраструктуры. Это подтверждается высокими значениями специализированных международных рейтингов, таких как индекс развития ИКТ, индекс готовности к сетевому обществу, рейтинг готовности стран к электронному правительству. Перечисленные рейтинги не учитывают развития соответствующих технологий на уровне региона, а оценивание происходит только на национальном уровне. В результате разрабатываемые инструментальные средства в большинстве своем направлены именно на поддержку общегосударственного управления и на уровне регионов не применяются. Это обусловило актуальность данной работы, целью которой является разработка методологических положений инструментального обеспечения регионального электронного правительства. Исследование базировалось на работах В.М. Глушкова, И.А. Хасаншина, В.Е. Кириенко и других. В отличие от других авторов в данной работе основное внимание уделено созданию интеллектуальной системы поддержки принятия решений по управлению социально-экономическим развитием региона и его использованием не только представителями органов региональной власти, но и другими пользователями для оценки и анализа развития отдельных территориальных образований.

Используемые методы. В предлагаемом комплексе используются следующие математические методы: факторный анализ и метод главных компонент, кластерный анализ, нечеткая кластеризация, метод анализа вариации и анализ временных рядов. Основной методологией автоматизации является теория мультиагентных систем, а основным методом принятия решений является ситуационный анализ.

Программно-инструментальный комплекс «СМАПС». Предлагаемый комплекс включает следующие элементы:

- *база знаний* включает как базы теоретических положений - математические модели и онтологию исследуемой области, так и хранилище данных. Последнее выражено в виде OLAP-модели, где одно измерение представлено локальными регионами, второе – периодами, третье – анализируемыми показателями. Пополнение хранилища данных происходит на основе получения информации от системы сбора данных. Объектом исследования в данном случае является Гродненская область (Республика Беларусь), а хранилище данных содержит сведения о значении 28 статистических показателей социально-экономического развития районов Гродненской области за 2008-2014 годы;
- *блок анализа и прогнозирования* состоит из *расчетного модуля* (вычисляет значения расчетных показателей по требованию пользователя и производит нормировку данных), *модуля кластеризации* (служит для формирования кластеров, содержащих районы с близким уровнем социально-экономического развития, включает четкую и нечеткую кластеризацию), *экспертной системы* (позволяет при необходимости привлекать экспертов для решения задачи), *модуля прогнозирования* (может строить несколько видов прогнозов: на основе заранее заложенных моделей экономического роста, на базе экстраполяции как общих тенденций, так и значений отдельных факторов с учетом степени их влияния на общий результат);
- *блок отчетов и визуализации* позволяет строить графики, картограммы, формировать отчеты на основе отобранных показателей и/или параметров, готовить документы лицу, принимающему решения.

Структура комплекса «СМАПС» представлена на рисунке 1.

Рис. 1. Структура программно-инструментального комплекса «СМАПС»

Стоит отметить, что каждый модуль представляет собой независимую процедуру, активизирующуюся по мере необходимости в зависимости от выбранного сценария. Фактически это совокупность отдельных независимых агентов, образующих мультиагентную систему. Так как каждая из процедур достаточно проста, то реализации отдельных сценариев не требует большого количества ресурсов.

Предложенные программно-инструментальный комплекс «СМАПС» позволяет проводить оценку уровня социально-экономического развития региона и отдельных составляющих экономического роста, выделять главные факторы, влияющие на развитие региона, определять степень их влияния, строить прогнозы развития исходя из заданных параметров и оценивать ресурсы, необходимые для достижения заданных показателей. Особенность предложенной системы в том, что, основываясь на открытых статистических данных, она может быть размещена в свободном доступе и использована не только органами регионального управления, но и общественными организациями, инвесторами для сравнительной оценки развития территориальных образований.

Результаты работы. Предложенный комплекс может стать эффективным элементом регионального электронного правительства, которое представляет собой систему электронного документооборота местного управления и способ предоставления информации и оказания набора государственных услуг гражданам, бизнесу, другим ветвям государственной власти, при котором личное взаимодействие между структурами регионального управления и заявителем минимизировано и максимально возможно используются информационные технологии, а использование автоматизированных систем позволяет повысить эффективность регионального управления. Он позволяет учесть особенности развития административно-территориальных единиц при построении прогнозов, может быть интегрирован в систему информационного обеспечения органов регионального управления и способствует повышению функциональности сайтов органов регионального управления. Кроме того, применение такой системы в открытом доступе способствует расширению круга пользователей, влияющих на принятие решений в области регионального управления.

Огурцов М.И., Ходзинский А.Н.
Институт кибернетики имени В.М. Глушкова НАН Украины, Киев, Украина

Про гибридный алгоритм комбинаторной оптимизации решения транспортных задач повышенной сложности

Математические модели многих транспортных задач относятся к классу NP-сложных. Этим обуславливается актуальность проблем разработки новых параллельных методов и алгоритмов решения задач этого класса при использовании суперкомпьютеров. Одной из таких задач является задача маршрутизации транспортных средств (ТС) при наличии временных окон [1,2].

Задача маршрутизации транспорта. Задача маршрутизации транспорта (Vehicle Routing Problems, VRP) заключается в выявлении и оптимизации маршрутов к заданным точкам-потребителям для транспортных средств (ТС), находящихся в депо [1]. Для каждого потребителя задан объем товаров, которые ему надо доставить и промежуток времени – временное окно, внутри которого нужно эту доставку выполнить. Задана также и грузоподъемность каждого из ТС. Решением задачи является разбиение множества заказов по транспортным средствам и совокупность маршрутов этих ТС. Каждый такое решение характеризуется стоимостью, которая зависит от затрат на осуществление перевозок и возможных штрафов за просрочку временных окон.

Поиск решения поставленной задачи маршрутизации транспортных средств с временными окнами осуществлялся путем гибридизации двух подходов: конструктивного алгоритма и итерационной процедуры. На первом этапе применялся детерминированный *алгоритм начального приближения* (АНП). На втором этапе выполнялась проверка оптимальности полученного начального приближения и его оптимизация итерационной процедурой. Процедура решения задачи маршрутизации транспортных средств с временными окнами в виде псевдокода:

```
Best_Solution = NULL;
Sorted_Data = Sorting_Input(данные клиентов);
Solution = DeterminedConstruction (Sorted_Data);
UpdateSolution (Solution, Best_Solution);
for k = 1,2,..., MAX_ITERATIONS do
    Solution = LocalSearch (Solution);
    if (Solution is better than Best_Solution) then
        UpdateSolution (Solution, Best_Solution);
    endif;
endfor;
```

При использовании АНП выполнялась такая последовательность действий:

1. При добавлении каждого клиента к маршруту проверялось, поместится ли необходимое количество груза в это ТС. Если нет – клиент будет обслуживаться другим ТС.
2. На втором шаге происходила *жадная сортировка частично сформированного маршрута*. Для каждого добавленного узла маршрута по формуле для декартовых координат рассчитывалось расстояние от депо до него с учетом времени до начала временного окна каждого клиента. Далее выбирался наименьший результат, ТС передвигалось к нему и процесс повторялся, пока все клиенты не были добавлены к маршруту.

После завершения построения начального приближения выполнялась оптимизация полученного приближенного решения. Первым методом комбинаторной оптимизации, выбранным для поиска лучших решений, была транспозиция компонент вектора решения (в окрестности с радиусом 1), то есть был использован локальный поиск (Local search) *методом вектора спада*. Он работал путем перебора всех возможных вариантов в окрестности, пока не будет достигнут локальный оптимум либо не будут превышены заданные временные ограничения.

Вторым примененным алгоритмом для оптимизации начального приближения является *алгоритм имитационного отжига* (Simulated Annealing - SA). Переход в найденный в

окрестности новый минимум происходил как и при обычном локальном поиске. Если же полученный набор маршрутов не являлся лучшим, переход к нему происходил с определенной вероятностью, уменьшающейся с каждой выполненной итерацией имитационного отжига и зависящей также от того, насколько менее эффективным, чем текущий минимум, являлось найденное решение.

Для исследования эффективности разработанного алгоритма был сгенерирован набор тестовых задач размерности 1000, а также выбрана известная задача Соломона C201 [2] размерности 100. Был разработан программный комплекс на языке C, он выполнялся в параллельном режиме на кластере Института кибернетики имени В.М. Глушкова.

При начальной сортировке клиентов C201 по объему требуемого груза, количество ТС составило 6, общее расстояние –891.66, время построения начального приближения в среднем составляло 15 секунд. Результатом применения метода вектора спада стало значительное улучшение полученных начальных решений. Расстояние сократилось на 32%; количество транспортных средств – до 4; время выполнения локального поиска составляло от 3 до 6 секунд.

Результаты применения алгоритма имитационного отжига (АИО): расстояние сокращалась при разных запусках до значений от 783.90 до 656.22; общее время выполнения составляло в среднем 15-45 секунд. В процентном соотношении лучший из полученных равен 73,6% от результатов АНП. Результаты о тестирования АИО приведены в таблице.

Таблица 1. Результаты оптимизации маршрутов алгоритмом имитационного отжига

Длина маршрута после АНП	Длина маршрута после АИО	Процент от исходного маршрута, %	Количество транспортных средств	Затраченное время, с
2516.88	872.64	34.6	4	21
2770.34	1007.71	36.4	5	7
2880.90	640.04	22.2	4	45
2923.31	621.09	21.2	4	39
3007.45	902.32	30.0	4	25
3134.32	858.02	27.4	4	32
3220.65	1321.66	41.3	5	8
3242.71	991.42	30.6	5	40

Проведенные исследования предложенного гибридного алгоритма маршрутизации транспортных средств подтвердили его практическую применимость и эффективность при решении прикладных задач маршрутизации. Разработанные алгоритмы на основе конструктивного алгоритма формирования начального приближения и итерационных – алгоритмов вектора спада и имитационного отжига – показали высокие результаты по минимизации количества ТС, общего пути всех ТС и общего времени ожидания, необходимых для обслуживания клиентов. Потенциальными потребителями разработанного параллельного алгоритма маршрутизации ТС являются логистические компании и компании, занимающиеся доставкой грузов.

Направлением дальнейших исследований является развитие предложенных алгоритмов комбинаторной оптимизации путем применения популяционных алгоритмов и сравнения их эффективности с уже исследованными методами и алгоритмами.

Литература. **1.** Гуляницкий, Л. Ф., Самусь А. В. Решение Н-методом задачи оптимизации маршрутов транспортных средств с временными окнами // Компьютерная математика (2). – 2012. – с. 147-155. **2.** Solomon M.M. Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints // Operations Research (35). – 1987. – pp. 254-265.

Перепеличный Я.И., Янковский И.П., Дидковская М.В.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Интеллектуальная система построения контекстной рекламы

На сегодняшний день реклама является одной из основных движущих сил рынка, в связи с чем рекламная машина требует постоянного улучшения и «движения в ногу» с современными технологиями. Одним из новых веяний последних нескольких лет стала контекстная реклама в сети Интернет, которая позволяет не просто рекламировать товары и услуги пользователям, а делать это, основываясь на его предпочтениях (история запросов, анализ посещенных страниц), что позволяет повысить КПД рекламной кампании в разы.

Контекстная реклама – это текстовые объявления, которые показываются пользователям по запросам, если эти запросы рекламодатель добавил в настройки рекламной кампании. Объявления показываются пользователю именно в тот момент, когда он сам проявил интерес к товару или услуге и, возможно, готов к покупке.

Интеллектуализация подобной системы заключается в том, чтобы использовать Искусственные нейронные сети для анализа данных и определения предпочтений пользователя. Вместо прямолинейного data mining'а можно обучить сеть результатами первичного анализа временных файлов, хранящихся на стороне пользователя. Таким образом мы не только избавим себя от необходимости хранить постоянно накапливающиеся знания о пользователе, но и получим систему, способную с каждым посещением всё лучше определять предпочтения пользователя.

Для организации подобной системы необходимо получить как можно больше информации со стороны пользователя, что достигается при взаимодействии со специфическими компонентами браузера (localStorage, Cookies, Local Shared Objects, IndexedDB API и т.д.). На практике существует большое число различных способов для отслеживания пользователя. Какие-то из них являются плодом ошибок в реализации или упущений и теоретически могут быть исправлены. Другие практически невозможно искоренить без полного изменения принципов работы компьютерных сетей, веб-приложений, браузеров. Каким-то техникам можно противодействовать — чистить cache, cookie и прочие места, где могут храниться уникальные идентификаторы. Другие работают абсолютно незаметно для пользователя, и защититься от них вряд ли получится. Это всё говорит нам о том, что практически всегда будет возможность получить информацию о пользователе и идентифицировать его. А большое разнообразие способов получения этой информации можно решить расширяемым адаптером в приложении, в котором будет предусмотрено использование источников информации о пользователе доступных на текущий момент.

Для обучения в работе предлагается использовать искусственные нейронные сети на основе радиально-базисных функций, или перцептрон. Это искусственная нейронная сеть, которая использует радиальные базисные функции как функции активации. Выходом сети является линейная комбинация радиальных базисных функций входов и параметров нейрона. Перцептрон - математическая или компьютерная модель восприятия информации мозгом (кибернетическая модель мозга).

Вывод. Полученная в работе интеллектуальная система имеет неограниченный потенциал при использовании со стороны провайдера (или другого узла, пропускающего через себя большое количество пользователей), или же может быть реализована в виде встраиваемого модуля для коммерческих сайтов. Так же данная система может быть использована для анализа метаданных пользователей определенных регионов и возможной их дальнейшей кластеризации для увеличения скорости работы существующих систем: поисковых или генерации контекстной рекламы.

Литература. **1.** J. Moody and C. J. Darken, "Fast learning in networks of locally tuned processing units, "Neural Computation, 1, 281—294 (1989). **2.** Уоссермен, Ф. Нейрокомпьютерная техника: Теория и практика — М.: Мир, 1992. — 240 с.

Пилипенко І.В.
Вінницький національний технічний університет, Вінниця, Україна

Прогнозування тривалості процесу тестування програмного забезпечення

Актуальність. Відповідно до поставлених завдань тестування підходять різні практики імплементації тестування, що в основному впливає на якість самих тестів. Тому актуальним є визначення певної моделі розробки рішень тестування, яка б дозволяла прогнозувати як процес загалом, так і його певні етапи [1].

Мета. Метою роботи є покращення якості рішень при тестуванні програмного забезпечення.

Аналіз процесу тестування програмного забезпечення показав, що даний процес можна віднести до розгалужено-циклічних технологічних процесів, оскільки виконується циклічно з прийняттям рішень на контрольних операціях [2].

Важливою складовою кожної моделі тестування програмного продукту є критерій достатності процесу тестування, який дозволяє керівникам проектів приймати обґрунтовані рішення про завершення даного етапу розробки. Принципова характеристика процесу тестування програмного забезпечення полягає у циклічності процесу, а саме в тому, що в програмі за період виконання кожного циклу тестування виявляються і виправляються помилки, які не були знайдені раніше. В даний час у переважній більшості IT компаній такі показники носять формальний характер та затрати на тестування намагаються зменшити, що призводить до зниження якості продукту. Необхідно спрогнозувати затрати на процес тестування, враховуючи його циклічну повторюваність [3].

Застосована модель в [4] розгалужено-циклічного технологічного процесу до процесу тестування ПЗ дає змогу прогнозувати час на виконання процесу. Складність застосування визначається тим, що існують стани, котрі не можливо одночасно віднести до бажаних чи небажаних. Кожний модуль, що тестується має свої характеристики та потребує окремого прогнозування. Запропонована методика дозволяє виконувати прогнозування для кожного окремого модуля програмного продукту, що призводить для якіснішого прийняття рішень на кожній контрольованій операції всього підпроцесу.

Для реалізації даної методики розробимо імітаційну модель процесу тестування програмного продукту.

На рис. 1 показано концептуальну модель процесу тестування.

Рис. 1. Концептуальна модель процесу тестування за технологією Scrum

На основі показника якості та показника покриття тестувальник приймає рішення:

- програмний продут відправити на доробку;
- програмний продукт вводити в експлуатацію.

Якщо програмний продукт було відправлено на доробку, він виправляється програмістом, і наступним етапом знову надходить на перевірку тестувальнику. Тестувальник перевіряє модуль за новим набором тестів з перевіркою попередніх, які виконалися не успішно. Такий процес тестування відбувається доти, поки продукт не буде відповідати показнику якості та не буде повністю покритий тестами.

Імітаційну модель розроблено з використанням середовища Simulink. Фактично, процес тестування можна розглядати як СМО. Покажемо комплексну модель процесу тестування з урахуванням підсистем розрахунку показників пориття та якості (рис. 2).

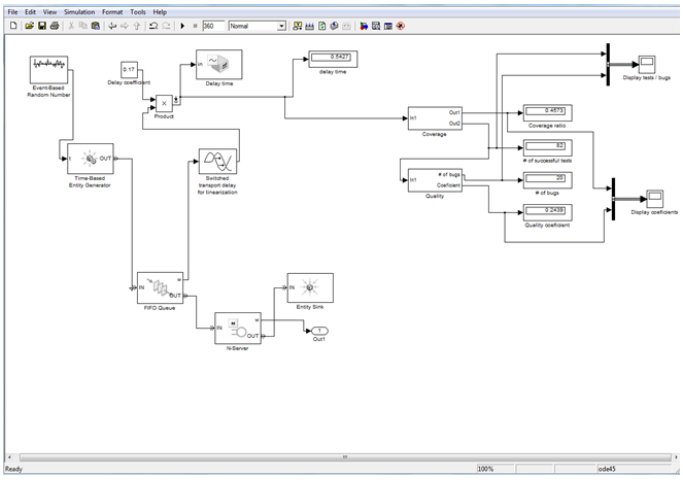


Рис. 2. Імітаційна модель процесу тестування з урахуванням підсистем розрахунку показників пориття та якості

Покажемо фрагмент випробовування імітаційної моделі на рис. 3.

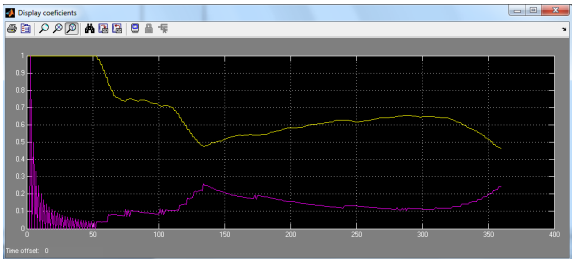


Рис. 3. Випробовування моделі процесу тестування для окремого модуля

Висновки. Здійснено прогнозування часу тестування програмного продукту за допомогою імітаційного моделювання в середовищі Simulink. Результати дослідження показали позитивні результати. Результати дослідження отримали практичну реалізацію у вигляді програмного забезпечення, яке впроваджене у ТОВ НВП «Спільна Справа» (м. Вінниця) в 2015 році на основі розроблених моделей, алгоритмів та методів.

Література. **1.** Ремінний О. А. Модель процесу автоматизованого тестування користувацьких інтерфейсів в умовах багатопродуктових компаній / О. А. Ремінний // Вимірювальна та обчислювальна техніка в технологічних процесах. - 2013. - № 4. - С. 106-113. **2.** Дубовой В.М., Пилипенко І.В. Прогнозування доцільної кількості повторень циклічного технологічного процесу / В.М. Дубовой, І.В. Пилипенко // Вісник ВПІ . - №1. - Вінниця: УНІВЕРСУМ Вінниця. - 2015. - С. 86-91 **3.** Дубовой В.М., Пилипенко І.В. Моделювання процесу тестування програмного забезпечення як розгалужено-циклічного технологічного процесу / В.М. Дубовой, І.В. Пилипенко // Автоматизація технологічних і бізнес-процесів. - №24 - Volume 7. - Issue 4. - 2015. - С. 55-64. **4.** Дубовой В.М., Пилипенко І.В., Стець Р.С. Застосування марковської моделі для аналізу впливу циклічності на управління розгалуженим техноло- гічним процесом / В.М. Дубовой, І.В. Пилипенко, Р.С. Стець // Наукові праці ВНТУ, 2014. - № 4. - [Електронний ресурс] - Режим доступу: http://praci.vntu.edu.ua/article/view/3827/5583.

Піпіч А.А.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Виявлення та усунення технічних конфліктів в контекстно-залежних системах

В даний час значна кількість систем в різноманітних галузях діяльності людини активно автоматизується - створюються механізми, що регулюють параметри елементів системи. Контекстом такої системи є середовище, що оточує механізм та несе певну інформацію про неї. В системах, що будуть розглядатися, вплив механізму на елементи залежить від контексту [1].

У випадку, коли аналіз контексту призводить до неоднозначності стану, в якому має перебувати механізм, виникає технічний конфлікт, який може призвести до зниження ефективності системи. Отже, виявлення та усунення конфліктів є важливою частиною керування системою.

Виявлення. Включає в себе реєстрацію конфлікту та його детермінування за одним з наступних критеріїв: джерело, учасники, час виявлення, вирішуваність.

Усунення. Під усуненням конфлікту слід розуміти мінімізацію його впливу на систему. Алгоритм, що буде для цього застосовано, залежить від типу конфлікту.

Таким чином передбачувані конфлікти можуть бути вирішені застосуванням заздалегідь запланованої послідовності дій.

Класифікація за учасниками, передбачає для кожного з них виконання аналізу наполегливості (міра, в якій учасник потребує виконання власних вимог) та ступені взаємодії (міра, в якій учасник допускає виконання вимог іншого учасника) [2]; в залежності від їх оцінки рішення може бути згенероване за таксономією Томаса-Кілмана [4] - уникнення, змагання, компроміс, взаємодія, пристосування.

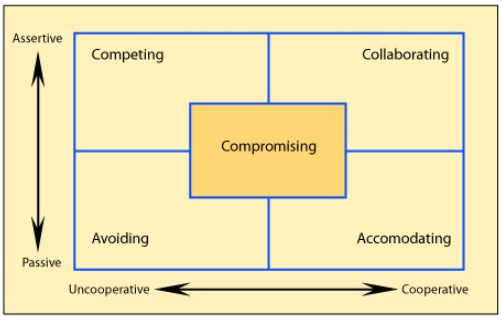


Рис. 1. Таксономія Томаса-Кілмана

Постановка задачі.

- Досліджуються шляхи усунення конфліктів, виконується їх порівняння;
- Розглядаються інструменти для вирішення конфліктів; можливість та доцільність їх застосування в системах різного типу;
- Застосовується Web Ontology Language (OWL) для визначення впливу конфлікту на систему;
- Моделюється поведінка системи при виникненні в ній конфлікту.

Література. **1.** Anind K. Dey. Understanding and Using Context, http://www.cc.gatech.edu/fce/ctk/pubs/PeTe5-1.pdf. **2.** Silvia Resendes, Paulo Carreira, Andre C. Santos. Conflict Detection and Resolution in Home and Building Automation Systems A literature review http://web.ist.utl.pt/paulo.carreira/downloadable/papers/resendes13overview.pdf **3.** John Herbert, John O'Donoghue, Xiang Chen. A Context-Sensitive Rule-based Architecture for a Smart Building Environment http://www.sersc.org/journals/IJSH/vol3_no1_2009/1.pdf **4.** Kenneth W. Thomas, Ralph H. Kilmann. Thomas-Kilmann Conflict Mode Instrument https://www.cpp.com/pdfs/smp248248.pdf

Поворознюк А.И., Поворознюк О.А.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Комплексная оценка этапов диагностически-лечебных мероприятий в системах поддержки принятия решений в медицине

Формализована задача комплексной оценки этапов лечебно-диагностического процесса с целью минимизации рисков врачебных ошибок. Разработаны методы кластеризации диагнозов в пространстве врачебных действий и коррекции порогов в диагностическом решающем правиле. Выполнена программная реализация системы и тестирование на реальных медицинских данных.

Ключевые слова: компьютерная система, диагностика, врачебное действие, врачебная ошибка.

Постановка проблемы. Процесс реабилитации пациентов состоит из двух взаимосвязанных этапов: диагностики заболеваний и лечения выявленных патологий, причем после диагностики и назначения лечебных процедур необходим мониторинг текущего состояния пациента с целью оценки эффективности процесса лечения и, при необходимости, его коррекции. Для лечения того или иного заболевания необходимо оказание определенных врачебных действий (ВД) на организм (хирургическое вмешательство, фармакологическое, лечебно-терапевтическое воздействие, реабилитационные мероприятия). На каждом из отмеченных этапов врач, как лицо принимающее решение (ЛПР), вырабатывает управленческое решение в условиях дефицита исходных данных и существенной априорной неопределенности, основываясь на своей квалификации, опыте и интуиции. При этом принятие неправильного решения (врачебная ошибка) как на этапе диагностики, так и на этапе лечения может иметь катастрофические последствия для здоровья пациента.

В настоящее время имеется широкий спектр компьютерных диагностических систем в разных предметных областях медицины [1], информатизация ВД ограничивается медицинскими справочниками, в том числе в виде информационно-поисковых систем [2].

В современных компьютерных системах поддержки принятия решений в медицине (СППРМ) задачи диагностики и ВД рассматриваются независимо друг от друга, при диагностике минимизируется риск неправильной постановки диагноза без учета ВД, поэтому актуальной является задача минимизации риска ВО при комплексной оценке диагностических и врачебных действий (ДВД).

Целью работы является разработка информационных технологий комплексной оценки этапов ДВД с целью повышения их эффективности и минимизации риска врачебных ошибок.

Формализация и информационная технология реализации этапов ДВД. В формализованном виде задачей диагностики является классификация состояния i-го пациента D_{ij} (i-й диагноз, j-я стадия тяжести заболевания) при анализе вектора диагностических признаков X_i. При этом выполняется синтез иерархической структуры диагностируемых состояний в заданной предметной области медицины (бинарное дерево решений S_D), в каждой вершине которого применяется решающее правило (РП), которое реализует метод последовательного анализа при диагностике между состояниями D_q и D_l. На каждом i-м этапе РП анализируется очередной признак x_i и вычисляется отношение правдоподобия $\Theta = \prod_i P(x_{ik}/D_q)/P(x_{ik}/D_l)$, которое сравнивается с порогами $\Theta > A$, $\Theta < B$. При выполнении одного из условий принимается решение о D_q или D_l соответственно и выполняется переход на более низкий уровень иерархии S_D с целью уточнения диагноза. При невыполнении обеих неравенств добавляется следующий $i+1$ признак и процедура повторяется.

ВД представляются моделью $M_a =< T_p, F, SI >$, где M_a – множество ВД; $T_p = \{t_{pi}\}$ – множество терапевтических действия (ТД), $F = \{f_i\}$ – множество фармакологических действии (ФД), $SI = \{si_i\}$ – множество видов хирургического вмешательства.

Назначение ВД при известном D_{ij} состоит из определение их типа (T_p, F, SI или их комбина-

ций) и перечня конкретных действий. Выбор типа ВД является задачей многокритериального выбора альтернатив, для решения которой используется метод анализа иерархий (МАИ). Для каждого из допустимых для данного диагноза D_{ij} типа ВД формируется подмножество необходимых ВД $t_{pD_{ij}} \in T_p$, $f_{D_{ij}} \in F$, $si_{D_{ij}} \in SI$, после чего определяется их реализация с учетом индивидуальных особенностей пациента, противопоказаний к отдельным ВД и многокритериальному выбору аналогов.

Для комплексной оценки этапов ДВД и минимизации риска ВО ищется зависимость между ошибкой диагностики (D_q вместо D_l) и ее последствий при реализации ВД. Так как для подавляющего множества патологий в различных областях медицины ВД реализуются медикаментозным путем, рассмотрим данный вариант реализации ВД, при котором каждому диагнозу D_i соответствует множество необходимых ФД f_{di}, на основании которого формируется комплекс лекарственных препаратов (КЛП) Y_i, который обеспечивает f_{di}, с учетом непереносимости j-го пациента к отдельным препаратам и многокритериального сравнения препаратов-аналогов [3].

Так как $Y_q \to f_{D_q}$, а $Y_l \to f_{D_l}$, то риск ВО определяется расхождением компонентов множеств f_{dq} и f_{dl}, и для него минимизации выполняется переход от традиционного пространства признаков X в пространство ФД F, компонентами которого $f_m \in F$ являются бинарные переменные (0 – отсутствует ФД, 1 – присутствует), а каждое состояние D_i представляется i-й вершиной гиперкуба. В качестве меры близости в пространстве F выбрано взвешенное расстояние Хэмминга: $r_{ij} = \sum_{k=1}^{g} w_{ij}|f_{ki} - f_{kj}|$, где $f_{ki}, f_{kj} \in [01]$ – k-е ФД i-го и j-го диагнозов соответственно; g – размерность пространства F; w_{ij} – коэффициент, который обеспечивает увеличения расстояния, в случае присутствия конфликтующих ФД.

Применение иерархической кластеризации по критерию минимума суммарной связи (минимальный разрез R) в пространстве F обеспечивает минимум риска принятия решения на этапе формирования КЛП при синтезе дерева решений S_D на этапе диагностики. Кроме того, в работе предлагается метод коррекции порогов A и B в РП, учитывая ошибки, которые возникают на этапе назначения КЛП. Получены зависимости между α, β и нормированным значением минимального разреза $\overline{R_i}$: $\alpha_n = 0,5(1 - \overline{R_{ql}})$, $\beta_n = 0,5(1 - \overline{R_{lq}})$. Определенные таким образом α и β задают пороги $A = (1-\beta)/\alpha$, $B = (1-\alpha)/\beta$ в РП, что обеспечивает учет рисков ВО при назначении КЛП в диагностическом РП.

Рассмотренная технология комплексной оценки ДВД с целью минимизации рисков ВО адаптируется к другим видам ВД, при этом выполняется переход из пространства признаков X в пространство соответствующих ВД (T_p или SI).

Выполнена программная реализация системы и ее тестирование на реальных медицинских данных с использованием обучающей выборки из 400 пациентов. В качестве основной платформы обоснован выбор Java. В архитектуре системы выделено три основных модуля: модуль взаимодействия с пользователем, базу данных, включающую в себя базу знаний и модуль построения знаний.

Выводы. Разработана система поддержки принятия решений при проведении ДВД на основе формализации этапов проведения ДВД при их комплексной оценке, что позволяет минимизировать риски врачебных ошибок, повысить достоверность и обоснованность решений. Архитектура программного обеспечения системы позволяет легко адаптироваться к различным предметным областям медицины.

Литература. **1.** Поворознюк А.И. Системы поддержки принятия решений в медицинской диагностике. Синтез структурированных моделей и решающих правил – Saarbrücken Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011. – 314 с. **2.** Компендиум 2015 – лекарственные препараты / Под ред. В.Н. Коваленко, А.П. Викторова. [Електронный ресурс] http://www.compendium.com.ua. **3.** Поворознюк О.А. Біотехнічна система призначення лікарських препаратів в дерматології. Автореферат дис. на здоб. вч. ступ. к. т. н.: спец. 05.11.17 "Біологічні та медичні прилади і системи"– Харків, 2010. – 21 с.

Поршнєва К.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Дослідження та розробка мультиагентної системи керування базами даних

Мотивація. Сучасні інформаційні системи характеризуються великим об'ємом розподілених по мережі даних. Часто виникають проблеми зі швидкістю доступу до даних, навантаженням мережі та зміною параметрів забору даних. Також трапляється така ситуація, що дані мають гетерогенний характер, тобто зберігаються у різних форматах, базах даних тощо. В такому випадку, доступ до них ускладнюється та сповільнюється ще більше.

Мета дослідження. Розробити та дослідити ефективність використання мультиагентної системи для керування розподіленими базами даних.

Методи дослідження. В якості мови програмування була використана мова - JavaScript на платформі NodeJS. В якості серверу був використаний звичайний персональний комп'ютер з unix-подібною операційною системою. Для обміну повідомленнями між агентами використовується з'єднання через веб-сокети. Для забезпечення комунікації мід агентами використовується мова Agent Communication Language (ACL).

Опис дослідження. Було визначено основні вимоги до керування даними в розподілених інформаційних системах. На підставі визначених вимог, було спроектовано та реалізовано мультиагентну систему керування базами даних. Було досліджено властивості розробленої системи та порівняно її ефективність з відомими системами керування розподіленими базами даних. Також, було проаналізовано ефективність її використання для упроавління гетерогенними даними.

В якості тестового прикладу була розглянута база даних на 1000000 записів. Було порівняно розроблену мультиагентну платформу та розподілені бази даних MongoDB та MySQL. При такому порівнянні зріст загальної ефективності виявився незначним.

Також було порівняно ефективність при роботі з залежними різноформатними даними. Для прикладу, було розглянено дані, що зберігаютья у трьох форматах: за допомогою бази даних MongoDB, в базі даних MySQL та у форматі JSON у вигляді звичайного файлу. Важливим було забезпечити відповідь на запит у єдиному форматі без залежності від початкового формату зберігання даних. В такому випадку, мультиагентний підхід виявився дуже зручним, так як за дапомогою створення агентів та налаштування процесу їх комунікації вдалося значно спростити логіку обробки запиту та абстрагувати роботу з різноманітними джерелами даних.Виявилося, що мультиагентний підхід є ефективним при налаштуванні роботи з даними, які зберігаються у різних форматах, та методи доступу до яких відрізняється.

Результат роботи. Перевагами мультиагентного підходу є гнучкість у налаштуванні виборок даних, швидкість доступу до розподілених гетерогенних систем та зменшенння навантаження на мережу.

Література. **1.** An Introduction to MultiAgent Systems - Second Edition / Michael Wooldridge **2.** G., Rzevski -Modelling large complex systems using multi-agent technology In Proc. of 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD2012), August 8-10, Kyoto, Japan, 2012. PP. 434-437. **3.** Амелина Н., Лада А., Майоров И., Скобелев П., Царев А. - Исследование моделей организации грузовых перевозок с применением мультиагентной системы адаптивного планирования грузовиков в реальном времени // Проблемы управления. 2011. No 6. **4.** В.А. Филатов, Е.Е. Цыбульник, Л.Э. Чалая - Модель мультиагентной системы автономного администрирования автономных систем и распределенных баз данных **5.** Поспелов Д.А. Многоагентные системы – настоящее и будущее // Информационные технологии и вычислительные системы. - 1998. - № 1. **6.** Wooldridge M., Jennings N. Intelligent agents: Theory and Practice // The Knowledge Engineering Review. - 1995. - № 10 (2). - P. 115-152.

Ржепішевський А.Л.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Архітектура системи доповненої реальності у відеочатах

Поява і розповсюдження засобів побудови доповненої реальності ставлять перед людством нові задачі обробки відео та ком'ютерного зору. Зокрема, використання звичайних інтерфейсів стає ускладненим, оскільки доповнена реальність передбачає, що доповнені частини не будуть заважати людській взаємодії з реальним світом. Використання системи доповнюваної реальності у відеочатах дозволить не тільки значно зменшити навантаження на мережу передачі даних та підвищити якість передачі зображення за рахунок відмови від передачі статичного фону та виділення тільки зображення людини, але й значно спростити спілкування у форматі доповнюваної реальності, роблячи його більш природнім та звичним.

Створення системи виділення контурів зображення людини на відео, що буде працювати зі швидкістю близькою до реального часу, результатом роботи якої буде вихідний потік відеоданих, що містить зображення людини без існуючого на вхідному потоці фону.

Знаходження контрастних контурів. Проблемі виділення контурів присвячена велика кількість літератури. Завдання обробки відеозображення в реальному часі накладає на обраний алгоритм істотні обмеження по продуктивності, тому з усіх алгоритмів слід в першу чергу вибирати найбільш швидкі, може бути, навіть на шкоду якості. На практиці добре себе зарекомендував метод. що полягає в послідовному проходженні бінарного зображення зверху вниз і зліва направо, знаходженні пов'язаних областей (далі компонент) і присвоєння їм міток. При цьому всі операції можна умовно розділити на чотири етапи.

Знаходиться перша точка зовнішнього контуру, яка на Рис. 8А позначена буквою А. після чого весь зовнішній контур проходиться цілком, поки знову не зустрінеться точка А. Їй і всім точкам зовнішнього контуру присвоюється одна і та ж мітка. Для всіх точок зовнішнього контуру, які перебираються зверху, виконується наступна операція. Починаючи з цієї точки, позначеної на Рис. 1б як А'. здійснюється сканування відповідного рядка пікселів, щоб визначити всі наступні чорні пікселі (якщо вони існують) шляхом присвоєння їм тієї ж мітки, що і у А '.

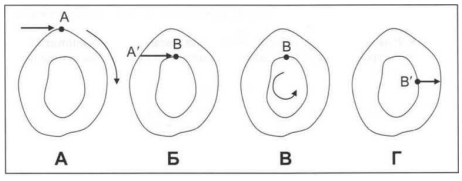


Рис. 1. Послідовність побудови контура

Коли в перший раз зустрічається точка внутрішнього контуру, їй присвоюється та ж мітка, що і у відповідній точки того ж ряду А зовнішнього контуру. Далі обходиться весь внутрішній контур, і всім його точкам присвоюється та ж мітка, що і у точки внутрішнього контуру. Коли зустрічається точка внутрішньогоконтуру, що вже має мітку, то продовжуємо йти вздовж цього ряду, щоб визначити всі наступні чорні пікселі (якщо такі існують) і присвоїти їм ту ж мітку, що і у В '.

Використання класифікаційних метрик і тимчасової узгодженості. Особливі труднощі при класифікації створює таку обставину: в поточному кадрі справжня природа об'єкта може бути спотворена внаслідок накладання декількох об'єктів. Вибір метричного оператора грунтується на тій обставині, що люди, як правило, мають менші розміри і більш складні контури, ніж транспортні засоби. Для обліку цієї закономірності досить ефективним є наступний метод. Введемо систему координат, в якій по одній осі будемо відкладати розмір об'єкта (загальне число пікселів), а по інший - величину D яка визначається за такою формулою:

$$D = p^2/a \qquad (1)$$

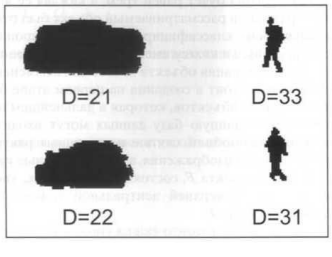


Рис. 2. Значення параметра D для деяких зображень автомобілів і людей

де p - це периметр об'єкта. Після цього на основі даних, отриманих під час навчання системи, на координатної площині можна виділити області, відповідні людям і транспортним засобам. Подальший поділ між цими двома класами проводиться шляхом розбиття масиву даних або лінійним чином, або за допомогою заходів Махаланобіса.

На Рис. 2 наведені зображення автомобілів і людей, під якими стоїть значення параметра D, округлене до цілого значення

Щоб уникнути невірної класифікації об'єктів під час їх перекриття або в тих випадках, коли об'єкт на зображенні з'явився тільки частково (при цьому автомобіль може виглядати як людина), можна використовувати метод максимальної правдоподібності. Для кожного переднепланового об'єкта слід побудувати класифікаційну гистограмму, у якій число осередків збігається з числом різних типів об'єктів. Наприклад, якщо вважати, що в відео потоці можуть бути присутніми автомобілі, групи людей і окремі люди, то розмір такої гістограми буде дорівнює трьом, а кожен її осередок буде відображати кількість разів, коли даний об'єкт був віднесений до конкретного класу.

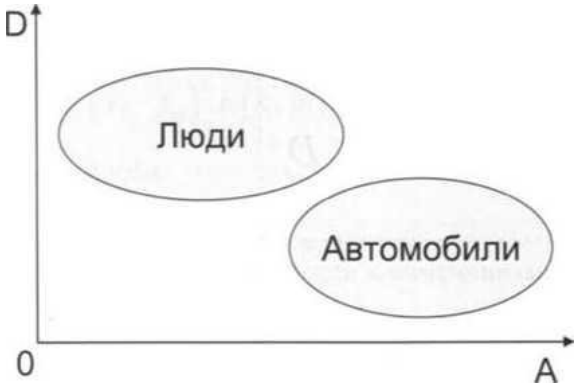


Рис. 3. Взаємне розташування областей, відповідних зображень людей і автомобілів, на площині (A, D)

Застосування описаних підходів до розробки архітектури системи. Предметна область передбачає високу якість виділення контурів, оскільки отримане зображення буде використовуватись як беспосередня частина графічного інтерфейсу користувача. В випадку з розпізнаванням людей за даними однієї камери, можуть виникнути ситуації, коли система не зможе виділити контур тіла людини через візуальну невідрізнимість тіла та фону. Також ускладнює задачу той факт, що фонове зображення може змінюватися із різною швидкістю в залежності від оточуючих умов. Шукане зображення буде використовуватися для відображення безпосередньо користувачу, тому це зображення не має містити ніякі частини фонового зображення. Враховуючи усе вищенаведене було вирішено використовувати послідовно декілька методів аналізу зображення та виділення контурів, щоб досягти необхідних показників якост отриманого зображення та швидкодії.

Література. **1.** Лукьяница. А. А. Цифровая обработка видеоизображений / А. А. Лукьяница, А. Г. Шишкин. - М.: "Ай-Эс-Эс Пресс". - 2009. - 518с. **2.** Методи компьютерной обработки изображений / Под ред. В.А.Сойфера. - 2- е изд., испр. - М.: - Физматлит, 2003. - 784 с. **3.** Балахонов Е.В. Алгоритми вьіделения контуров / Е.В.Балахонов / Научние работы белорусского государственного университета информатики и радиоэлектроники. - Минск, 2006. **4.** William E. Green. Canny Edge Detection Tutorial / William E. Green / Information resources and Technology of Mechanical Engineering and Mechanics Department, Drexel University. - Philadelphia, 2002. - Vol.8. **5.** Яне Б. Цифровая обработка изображений / Б.Яне - М.: Техносфера, 2007 - 584с.

Романенков Ю.А., Вартанян В.М.
Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков, Украина

Задачи прогностического обеспечения систем поддержки принятия стратегических решений

Прогнозирование является ключевым этапом процесса принятия управленческих решений, его результаты служат информационной основой для действий специалистов, а, следовательно, должны характеризоваться точностью, надежностью, устойчивостью [1].

Для моделирования процесса принятия стратегического решения необходимо формализовать условия внешней и внутренней среды объекта (например, организации). В некоторых случаях удается определить ряд количественных показателей, которые в полной мере определяют положение объекта в конкурентной технико-экономической среде. Они образуют фазовое пространство, траектория движения в котором определяет эффективность исследуемого объекта на протяжении собственного жизненного цикла.

Об эффективном стратегическом управлении сложными организационно-техническими объектами можно говорить лишь в том случае, когда кроме информации о текущем положении объекта в фазовом пространстве, субъект управления обладает прогнозными оценками относительно будущего состояния объекта.

Наглядным примером таких объектов могут служить современные авиакомпании. В Украине насчитывается около 30 авиакомпаний. Чтобы оценить финансовые потоки авиакомпаний, достаточно сказать, что стоимость часа полета на Як-42 составляет около 1000 евро. В год только одним бортом, при условии налета заявленного количества часов, можно получить доход в 840 000 евро [2].

Таким образом, мера риска при ошибочном планировании весьма существенна, что ужесточает требования к прогностическому обеспечению систем поддержки принятия стратегических решений, используемых в этой предметной области.

Научно-практической проблемой является прогностическое обеспечение стратегического управления организацией в условиях объективной неопределенности показателей её внутренней и внешней среды. Объектом исследования являются процессы принятия стратегических решений в организациях. Предметом исследования выступают принципы, модели и методы прогностического обеспечения систем поддержки принятия стратегических решений в организациях. Целью исследования является повышение качества прогностического обеспечения систем поддержки принятия стратегических решений в организациях путем разработки методов, моделей и инструментальных средств обработки информации.

Сформулируем основные задачи прогностического обеспечения систем поддержки принятия стратегических решений:

1) сбор, верификация и накопление статистической информации по ключевым координатам фазового пространства для объекта стратегического управления;

2) формирование группы основных методов прогнозирования нижнего уровня, пригодных для работы с временными рядами с учетом их особенностей (стационарность, пропуски, зашумленность и т.д.);

3) разработка алекватной модели адаптивного комплексирования прогнозных оценок из разных источников;

4) обеспечение интерактивного режима параметрической настройки как прогнозных моделей нижнего уровня, так и модели комплексирования, позволяющего обоснованное вмешательство субъекта управления.

Литература. **1.** Афанасьева, Т. В. Моделирование нечетких тенденций временных рядов / Т. В. Афанасьева. – Ульяновск : УлГТУ, 2013. – 215 с. **2.** А. Леденев. Инфографика: где зарабатывают самолеты с украинской пропиской http://forbes.net.ua/business/1411400-infografika-gde-zarabatyvayut-samolety-s-ukrainskoj-propiskoj.

Саржинська М.С., Кисельов Г.Д.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Виявлення прихованих взаємозалежностей між даними на прикладі аналізу тональності відгуків користувачів інтернет

Прийняття людиною рішень в будь якій сфері діяльності в значній мірі залежить від відгуків оточуючих. З поширенням Інтернету коло оточуючих значно збільшилось, користувачі Інтернету залишають свої відгуки на форумах, web-сторінках, в соціальних мережах. Щорічно кількість контенту в мережі Інтернет зростає на 30%. Близько третини блогерів регулярно публікують свої думки щодо поширених торгівельних марок та брендів. Весь цей контент несе в собі велику кількість інформації, яку доводиться регулярно отримувати, читати та аналізувати.

Опитування, проведене компанією Dimensional Research [1], показало, що 88% опитуваних схиляються до думки, що аналіз позитивних та негативних відгуків в Інтернеті впливає на їхнє рішення при покупці товарів та послуг. Власникам інформаційних ресурсів життєво необхідно знати відгуки користувачів – чи це оцінка нового продукту в інтернет-магазині, чи це реакція людей на свіжу інформацію на сайті новин. Щоб опрацювати цей колосальний об'єм даних, потрібно багато часу. Вирішенням цієї проблеми займається система аналізу тональності текстів.

В загальному вигляді задача текстової класифікації визначається наступним чином [2]. Нехай існує опис документу X, де X — векторний простір документів, і фіксований набір класів $C = \{c1, c2, \ldots, cm\}$. З навчальної вибірки (множини документів із заздалегідь відомими класами) D $=\{\langle d, c\rangle \mid \langle d, c\rangle \in X \times C\}$ за допомогою способу визначення тональності R необхідно отримати класифікуючу функцію (або класифікатор) R(D) $= \gamma$, яка відображає документи в класи $\gamma : X \to C$. В задачі визначення тональності множина C, наприклад, може складатися з трьох елементів {позитивний, негативний, нейтральний}.

Основні підходи [3] до способу визначення тональності можна розділити на наступні категорії:

1. підхід, оснований на правилах (rule-based approach);
2. підхід, оснований на використанні словників оціночної лексики (affective lexicons);
3. підходи, основані на навчанні з учителем (supervised learning);
4. підходи, основані на навчанні без учителя (unsupervised learning).

Для визначення тональності відгуків в якості предметної області була обрана тематика оцінок ресторану відвідувачами [4]. Навчальна вибірка складалася з 102544 об'єктів, тестова - 34194. Відгуки класифіковані за п'ятибальною шкалою. Дослідження роботи алгоритму здійснювалося за допомогою мови програмування Python, її бібліотеки NLTK та методів машинного навчання. Перед аналізом тональності текстів для зменшення розмірності та «зашумленості» даних застосовувалися такі способи нормалізації: 1) приведення всіх букв до нижнього регістру; 2) заміна смайликів на відповідні слова в залежності від тональності; 3) видалення стоп-слів (прийменників, сполучників та ін.), символів пунктуації та повторюваних букв в словах, які користувачі можуть писати в емоційному стані (наприклад, supppppperrr); 4) заміна всіх слів з часткою "не" на антоніми (тут виникає проблема вибору слів з множини антонімів з різним емоційним забарвленням); 5) використання стемерів та лематайзерів для виділення основ слів (стемер Портера, лематайзер WordNet). Отримані практичні результати класифікацій відгуків алгоритмом Naive Bayes в залежності від кількості найбільш часто зустрічаючих слів зображені на рис. 2. Порівняльна гістограма алгоритмів класифікації машинного навчання для кількості слів 1700, що найчастіше зустрічалися в тестовій вибірці, наведена на рис. 3.

Дані експериментальних досліджень показали, що кращим класифікатором аналізу тональності відгуків є алгоритм логістичної регресії, його достовірність складає майже 60% (для класифікованих за п'ятибальною шкалою відгуків). На підставі цих даних відвідувачі

ресторану можуть приймати рішення, чи варто їм відвідувати даний заклад, а власники ресторану можуть прогнозувати прибутки та популярність даного закладу.

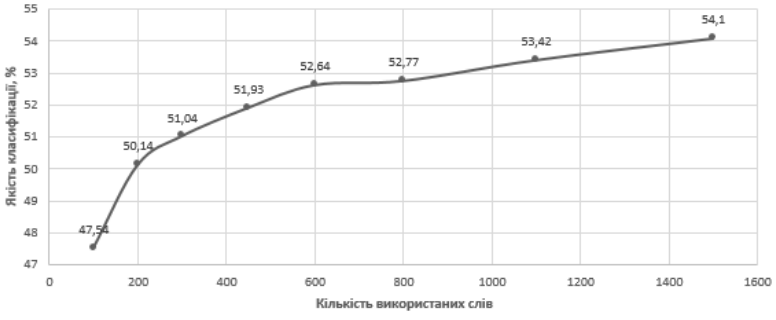


Рис. 1. Аналіз якості класифікації методом Naive Bayes в залежності від кількості слів

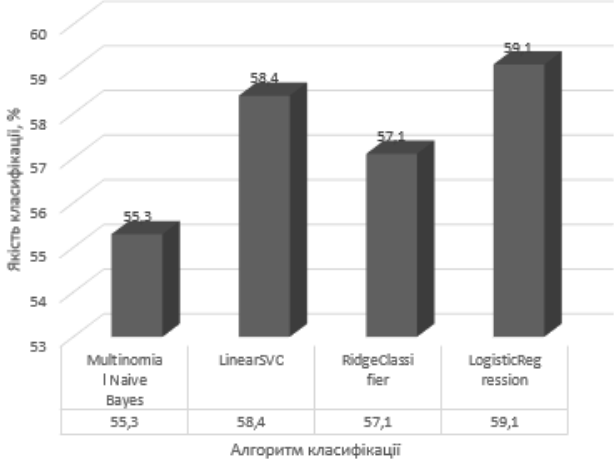


Рис. 2. Аналіз якості класифікації в залежності від методу класифікації

Висновок. Отримані результати свідчать про те, що створення системи аналізу відгуків для виявлення прихованих взаємозалежностей між даними та прийняття рішень на їх основі є складною задачею. Для її вирішення необхідні вхідні дані для навчання та завчасно визначена тематика аналізу. При використанні машинного навчання потрібно тестувати різні параметри, щоб підібрати ті, які краще працюють на тестових даних. Потрібно тестувати різні алгоритми класифікації (Naive Bayes, SVM, Random Forest, логістична регресія та інші), різні набори ознак (уніграми, біграми, N-грами слів) та функції зважування ознак. Практичне використання інструментарію аналізу відгуків доцільне при автоматизації експертних оцінок тендерних пропозицій.

Література. **1.** The Impact Of Customer Service On Customer Lifetime Value, www.zendesk.com/resources/customer-service-and-lifetime-customer-value. **2.** D. Manning, P. Raghavan, H. Sch¨utze,"Introduction to Information Retrieval"/ Cambridge University Press, 2008. **3.** Обучаем компьютер чувствам (sentiment analysis по-русски), habrahabr.ru/post/149605. **4.** Kaggle competition, inclass.kaggle.com/c/review-it/data.

Скірко І.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Система розпізнавання текстур на зображеннях з БПЛА

Системи комп'ютерного зору широко розповсюджені, та використовуються для автоматизації взаємодії комп'ютерів із зовнішнім світом - розпізнавання деяких об'єктів та реагування на них. В даній роботі пропонується використовувати текстурний аналіз для попередньої обробки великих масивів даних, зібраних за допомогою безпілотних літальних апаратів.

На сьогоднішній день у світі помітно тренд стрімкого розвитку безпілотних літальних апаратів, починаючи від маленьких любительських коптерів і закінчуючи великими військовими літаками, що обладнані різноманітними приладами для збору інформації про навколишнє середовище. Застосуванням безпілотних коптерів може бути доставка товарів від постачальника до споживачів, моніторинг різноманітних явищ та об'єктів (повені, популяції диких тварин, ситуації на дорогах, тощо), охорона великих територій (наприклад, охорона державного кордону). Всі ці застосування передбачають обробку великої кількості відео- та фотоматеріалу для виявлення об'єктів, що цікавлять серед великої кількості інших.

Для автоматизації попередньої обробки даних пропонується застосувати текстурний аналіз. За своєю природою кожна текстура є набором відносно простих підструктур (текстонів), що повторюються за деяким визначеним специфічним законом. Наприклад поверхня водойм є сукупністю хвиль, що йдуть одна за одною та в орієнтовані в одному напрямку. Таким чином поверхню води можна розглядати як текстури, де елементарними частинами є окремі хвилі, а законом є правила їх розповсюдження. Навчившись розпізнавати такі явища, можна автоматизувати пошук окремих текстур: наприклад, повідомити оператора, коли в полі зору з'явиться ліс; або окремих об'єктів попередньо видаливши "зайві"текстури згідно контексту.

Існує багато підходів до розпізнавання текстур. Наприклад, досліджувати мультифрактальний спектр (MFS) — дескриптор текстур, що має за основу фрактальний аналіз та надає ефективний фреймворк для роботи із сутністю структури текстур [1]. Або за допомогою нейронних мереж, попередньо визначивши метод формування вектору ознак [2]. Іншим варіантом є застосування до зображення множини фільтрів [3], вигляд яких є чітко визначеним для кожної текстури. На рисунку 1 зображено приклад обробки зображення набором фільтрів та кластеризації їх результатів.

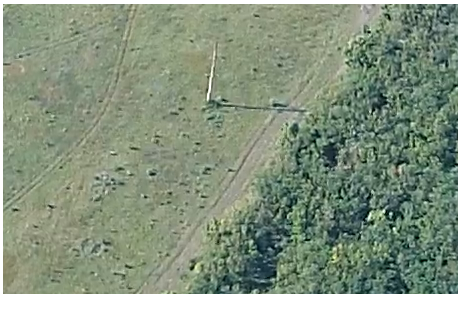


Рис. 1. Вихідне зображення

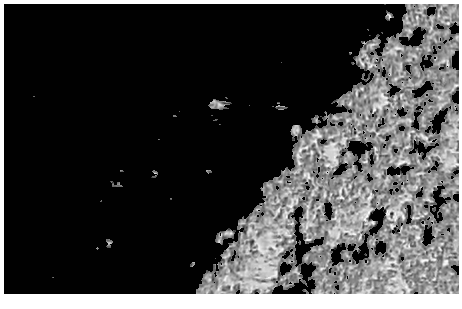


Рис. 2. Результат пошуку лісу

Пропонується отримані результати фільтрації поєднувати з методами контурного аналізу, іншими методами розпізнавання текстур для сегментації вихідного зображення та пошуку окремих об'єктів та текстур

Література. **1.** Y Xu, H Ji, C. Fermüller, "Viewpoint invariant texture description using fractal analysis". International Journal of Computer Vision, 83 (1), 85 - 100 (2009). **2.** Nidhal K. Al abbadi , Nizar Saadi Dahir, Zaid Abd Alkareem. "Skin Texture Recognition Using Neural Networks" **3.** Jitendra Malik, Serge Belongie, Thomas Leung, Jianbo Shi. "Contour and Texture Analysis for Image Segmentation". International Journal of Computer Vision, 43(1), 7-27, June 2001.

Слободюк А.Ю.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Построение модели для расчета лучшего маркетингового предложения для клиентов банка

Одним из основных направлений банковской деятельности является работа с внутренней базой клиентов, а именно перекрестные продажи. Основная задача эффективной кросс-продажи – это определение наилучшего маркетингового предложения для клиента. Под маркетинговым предложением будем понимать коммуникацию, в которой клиенту предлагается купить некий банковский продукт. Для расчета предложения необходимо принимать во внимание ряд поведенческих и социально-демографических характеристик клиента, таких как пол, возраст, доход клиента, его семейное положение, род деятельности и прочее. Также учитываются так называемые продуктовые характеристики, а именно наличие у клиента тех или иных продуктов банка, его транзакционная активность, сегмент клиента (определяется в зависимости от объёмов пассивов клиента и его доходов).

Пусть есть основная выборка клиентов, а также обучающая выборка. В роли обучающей выборки выступает пул клиентов, которые получали то или иное маркетинговое предложение. Кроме этого мы знаем результат коммуникации, т.е. есть информация о том, оформил ли клиент предлагаемый ему продукт.

Требуется построить математическую модель, которая бы описывала вероятность отклика клиента на маркетинговое предложение.

Одним из возможных методов решения этой задачи является метод логистической регрессии. Логит-регрессия – это статистическая модель, которая предсказывает апостериорные вероятности возникновения некоторого события путем подгонки данных к логистической кривой. Одним из преимуществ логистической регрессии является гибкость алгоритма, способного принимать входные данные любого рода и поддерживать несколько различных аналитических задач, поскольку в перспективе планируется использовать данный метод также и для портфельного анализа кредитных карт. [1]

Вторым этапом решения данной задачи выступает построение скоринговых карт для дальнейшего практического использования данной модели. Каждому клиенту в зависимости от его характеристик выставляются скоринговые баллы и клиент попадает в ту или иную скоринговую группу. [2] После этого экспертным путём выбирается порог отсечения и клиент получает рассчитанное маркетинговое предложение.

Внедрение данного исследования дало значительные результаты. Основным показателем эффективности кросс-продажи является конверсия контакта с клиентом в оформленную сделку. В среднем результаты продаж в наиболее эффективном канале «колл-центр» за 2015 год после внедрения увеличилась в 2 раза, что позволяет судить об эффективности модели. Такое увеличение конверсии было достигнуто за счет более таргетированных предложений и попадание в стартовую базу продаж только наиболее перспективных клиентов. А это влечет за собой уменьшение затрат.

В дальнейшем планируется работа над обогащением данных, что даст возможность построить более точную и эффективную поведенческую модель, а также применение данного метода в смежных областях, таких как портфельный анализ и активация портфеля кредитных карт.

Литература. **1.** Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: классификация и снижение размерности. — М.: Финансы и статистика, 1989. **2.** Усачёв С. А. Кредитный скоринг: решения desktop или enterprise. Журнал "Банки и технологии № 04,2008 год.

Таранян А.Р., Алфимцев А.Н.
МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Применение ковариационных матриц для отслеживания людей в видеопотоке

Автоматическое обнаружение и отслеживание людей в видеопотоке имеет целый ряд важных применений, среди которых можно выделить:

- Системы видеонаблюдения и системы безопасности [1].
- Роботы, взаимодействующие с людьми [2].
- Интеллектуальные мультимодальные интерфейсы [3].

В данной работе обсуждается задача отслеживания людей видеопотоке, которую можно рассматривать в виде одной из фундаментальных задач компьютерного зрения – классификации объектов, представленных в виде двумерных изображений.

Основными этапами большинства алгоритмов классификации объектов являются выделение дескрипторов, описывающих наиболее информативные признаки объектов с точки зрения их дальнейшей классификации, обучение классификаторов и непосредственно классификация объектов на основе их дескрипторов.

Существующие методы обнаружения и классификации объектов, которые могут быть применимы для решения задачи отслеживания людей в видеопотоке, можно разделить на две группы по принципу построения дескриптора.

Методы первой группы основаны на выделении репрезентативных областей объекта посредством алгоритмов нахождения ключевых точек.

В методах второй группы дескрипторы строятся на основе всей области локализованного объекта. Одним из наиболее популярных подходов данной группы является использование гистограммы области в качестве ее дескриптора [4]. Однако, отбрасывание пространственной информации при построении гистограмм, дающее с одной стороны преимущество, заключающееся в инвариантности дескриптора относительно изменения перспективы съемки объекта и его положения, с другой стороны, делает невозможным дифференциацию объектов схожих цветов с различной геометрией.

Основной причиной недостаточной инвариантности дескрипторов по отношению к изменению освещения, положения объекта и перспективы съемки является то, что они не в полной мере учитывают статистическую и пространственную информацию об объектах. Дескриптор на основании ковариационной матрицы, предложенный в работе [5], дает возможность захватывать и статистические, и пространственные свойства объекта, а также корреляцию между ними, при этом имея достаточно маленькую размерность и вычислительную сложность [6].

Разработанная модель. Учитывая особенности решаемой задачи, разработана модель, основанная на распознавании людей посредством классификации двумерных изображений головы и плеч. Область головы и плеч (далее – область интереса) выбрана в виду того, что она в наименьшей степени перекрывается иными объектами в кадре. Классификация областей производится путем сопоставления ковариационных матриц (дескрипторов), построенных для этих областей.

Поскольку прямоугольная область интереса, которую находит алгоритм Виолы-Джонса при поиске головы и плеч человека, содержит достаточно большие области, являющиеся, по сути, фоном, целесообразно строить ковариационную матрицу лишь для тех пикселей, которые принадлежат области головы и плеч человека. Для этого была предложена бинарная маска, изображенная на (рис. 1).

Ключевым вопросом при разработке данной модели являлся выбор функции признаков. Был разработан ряд новых функций признаков, эффективность применения которых была сравнена с эффективностью базовых функций признаков.

На (рис. 2) изображены графические представления ковариационных матриц, построенных для пяти различных обласей интереса с использованием одной из базовых функции признаков.

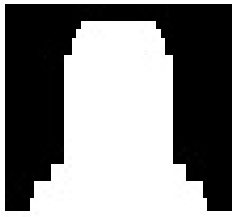


Рис. 1. Маска области интереса

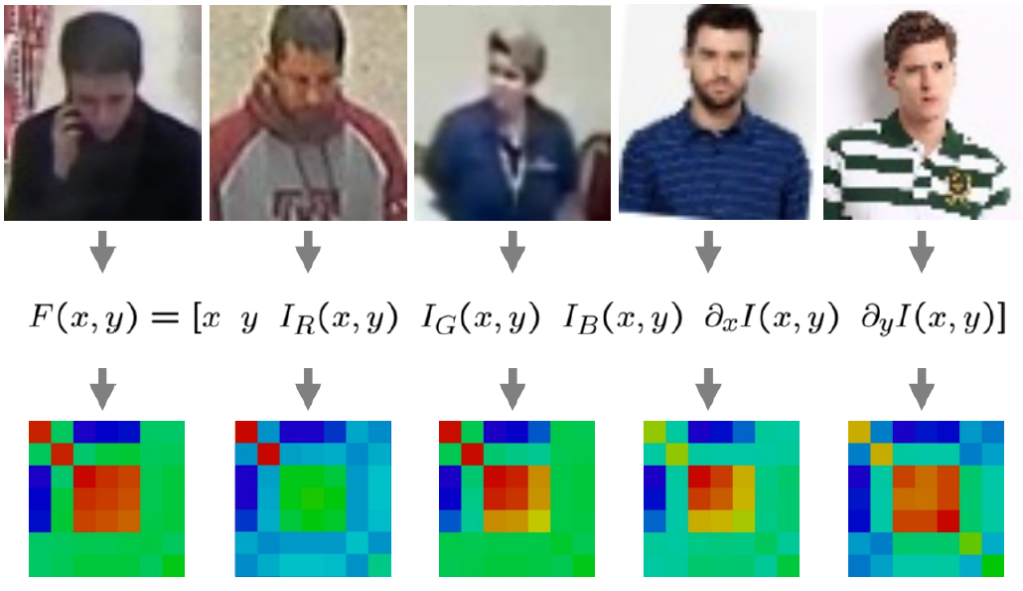


Рис. 2. Ковариационные матрицы для пяти изображений

Для определения «похожести» дескрипторов C1 и C2 двух областей интереса необходимо определить метрику для ковариационных матриц. В данной работе в качестве метрики была использована эвклидова мера.

Проведенный анализ на достаточно сложной тестовой выборке показал высокую эффективность представленной модели, которая сумела правильно классифицировать 94,6% образов.

Литература. **1.** Watada J., Musaand Z.B. Tracking human motions for security system // SICE Annual Conference, 2008. - P. 3344 – 3349. **2.** Bellotto N., Hu H. Multisensor-based human detection and tracking for mobile service robots // IEEE Trans. on Systems, Man, and Cybernetics, 2009. Vol. 39. N. 1. - P. 167-81. **3.** Алфимцев А.Н. Нечеткое агрегирование мультимодальной информации в интеллектуальном интерфейсе // Программные продукты и системы, 2011. №3. - С. 10. **4.** Comaniciu D., Ramesh V., Meer P. Kernel-based object tracking // IEEE TPAMI, 2003. Vol. 25. - P. 564-577. **5.** Tuzel O., Porikli F., Meer P. Region covariance: A fast descriptor for detection and classification // ECCV, 2006. - P. 589-600. **6.** Porikli F., Tuzel O., Meer P. Covariance tracking using model update based on Lie algebra // IEEE CVPR, 2006. Vol. 1. - P. 728-735.

Тертишний А.А.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Розподілене навчання нейронних мереж у браузерах

Робота присвячена написанню та дослідженню розподіленого навчання нейронних мереж на багатьох клієнтах з використанням браузера як обчислювального двигуна.

Ціль роботи. Дослідження можливостей використання браузерів для розподілення навчання нейронних мереж.

Методи. В якості основного обчислювального двигуна буде виступати браузер, що встановлений на персональному комп'ютері, телефоні, чи планшеті клієнта. В якості мови програмування буде використана мова - JavaScript, котру за замовчування підтримують всі сучасні браузери. В якості мастер серверу буде використаний звичайний персональний комп'ютер з Unix подібною операційною системою, сам сервер буде написаний на Javascript на платформи NodeJS. Для синхронізації клієнтів та мастер-серверу використовуються вебсокети.

Опис прототипу. В якості прототипу було створено мастер-сервер, що виступає виступає в ролі http-серверу. Після запуску мастер-серверу користувач може створити завдання на навчання нейронної мережі, налаштувати параметри мережі та інтервал синхронізації m, потім завантажити вибірки для навчання. Далі він отримує лінк який може розповсюдити серед клієнтів, кожен клієнт що відкриє цей лінк почне виконувати навчання, навчання на клієнті виконується в окремих потоках тому це не викликає блокування інтерфейсу чи підвисання системи. Також при використанні багатоядерного персонального комп'ютеру, мастер сервер буде використовувати вільні ядра в якості клієнтів. Кожні m мілісекунд мастер сервер буде збирати дані з клієнтів та починати наступну ітерацію доти доки помилка не досягне значення заданого користувачем, або не пройде максимальна кількість ітерацій встановлена користувачем.

Дослідження. В якості тестового прикладу була обрана нейронна мережа для класифікації зображень розміром 28*28 px, розмір вибірки було обрано - 16, кількість ітерацій - 100. Було запущене навчання на 0, 1, 2, 4, 8, 10, 12, 14, 16 клієнтах. Його результати наведено низже.

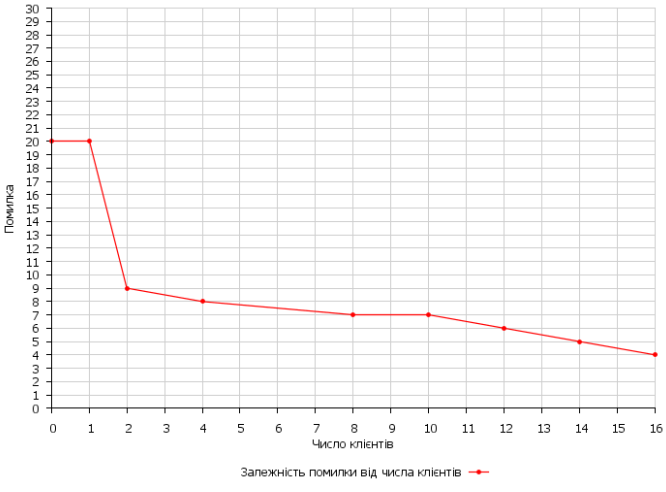


Рис. 1. Результати дослідження ефективності навчання

Результати. Для оцінки ефективності розподіленого навчання ми варіювали кількість клієнтів при 100 ітераціях як видно з графіку на рисунку 1 залежність між помилкою нейронної мережі та кількістю клієнтів - є близькою до лінійної, даний результат свідчить про ефективність даного підходу для навчання нейронних мереж. Для оцінки навантаження на мережу мастер-серверу ми варіювали кількість клієнтів як видно з рисунку 2 навантаження на мережу

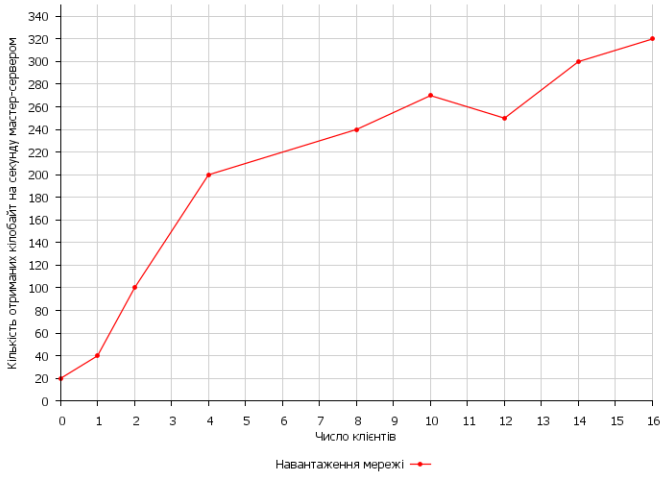


Рис. 2. Результати дослідження ефективності навчання

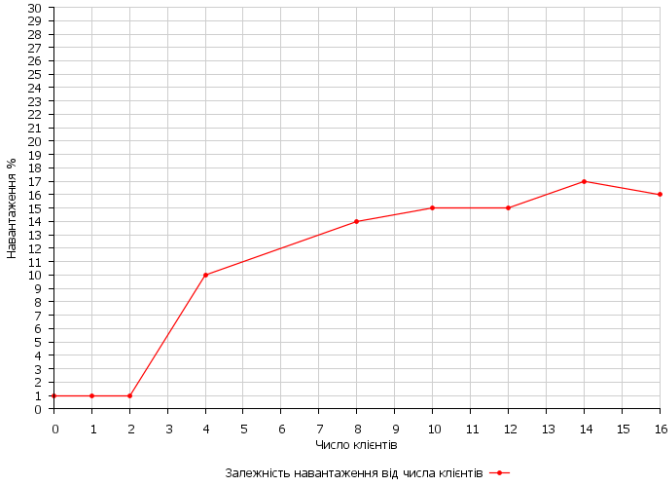


Рис. 3. Результати дослідження ефективності навчання

зростає лінійно і сильно залежить від конфігурації мережі, що свідчить, що при використанні звичайного інтернет з'єднаня користувач може бути обмежений в кількості клієнтів, навіть якщо обчисльовальної потужності достатньо. Також слід зазначити, що під час експериментів витрати оперативної пам'яті як на клієнтах так и на мастер сервері не перевищували 150 мб. Навантаження на процесори мастер серверу не значно зростало зі значним збільшенням клієнтів, це може можна бачити з рисунку 3. Середній час відповіді мастер серера залишався стабільним на всіх ітерацій. З часом значно збільшились енерговитрати мастер серверу. Під час експерименту на клієнтах не було помічено порушення стабільності чи відгукованості системи що свідчить про можливість використання даного підходу для навчання масово, адже це не буде заважати клієнтам і може виконуватися непомітно для них.

Література. **1.** Michael A. Nielsen, "Neural Networks and Deep Learning", Determination Press, 2015. **2.** Рутковская Д., Пилинський М., Рутковский Л.- Нейронные сети, генетические алгоритмы: Пер. с польск. И.Д. Рудинского.- М.:Горячая линя Телеком, 2006.-452 с.:ил. **3.** Хайкин Саймон.- Нейронные сети: полный курс, 2-е издание. : Пер. с англ. – М. : Издательский дом «Вильямс», 2006. – 1104 с. : ил. – Парал. тит. Англ. – ISBN 5-8459-0890-6.

Филатова А.Е.
Национальный технический университет "Харьковский политехнический институт", Харьков, Украина

Исследование альтернативных диагностических признаков в интеллектуальных компьютерных кардиологических системах поддержки принятия решений

Постановка проблемы и анализ литературы. С развитием производства отечественных цифровых электрокардиологических телеметрических комплексов, таких как транстелефонный цифровой 12-канальный электрокардиологический комплекс «Телекард», разработанный ООО «Компания TREDEX» (г. Харьков), актуальной является научно-практическая задача создания интеллектуальных компьютерных кардиологических систем поддержки принятия решений (ИККСППР) [1]. Создание ИКСППР в составе ЭКГ-телеметрических комплексов позволяет повысить качество консультативной помощи в условиях оказания экстренной медицинской помощи.

Традиционно при диагностике состояния сердечно-сосудистой системы по электрокардиограмме (ЭКГ) выполняется морфологический анализ, то есть выделение зубцов и комплексов, определение их амплитудно-временных характеристик, а также анализ формы выделенных структурных элементов. ЭКГ относится к биомедицинским сигналам (БМС) с локально сосредоточенными признаками (ЛСП). Для решения задачи морфологического анализа БМС с ЛСП в [2] автором предложен метод, основанный на построении многоканального нелинейного фильтра, задачей которого является обнаружение структурных элементов (СЭ) заданного вида, а также их локализация на рассматриваемом сигнале. При этом функция обнаружения i-го канала нелинейного фильтра имеет следующий вид:

$$\tilde{y}_i[t] = \frac{1}{1 + \alpha_i \sum\limits_j (y^{э}_{ij} - y^{t}_{ij})^2}, \tag{1}$$

где $\alpha_i > 0$ – коэффициент, отражающий чувствительность к изменениям СЭ одного класса за счет наложения помех и вариации параметров; $y^{э}_{ij} = f_{ij}(x_0[t], \vec{p})$, $y^{t}_{ij} = f_{ij}(x[t], \vec{p})$ – координаты эталона и текущего объекта соответственно; j – индекс координаты объекта; $f_{ij}(x[t], \vec{p})$ – функция преобразования сигнала $x[t]$ в пределах апертуры на основе модели полезного сигнала (МПС); $x_0[t]$ – эталон СЭ заданного класса; $\vec{p}$ – вектор параметров преобразования.

Для обнаружения СЭ по i-му каналу предложено частное решающее правило:

$$\tilde{x}_i[t] = \begin{cases} x[t] \ \forall t \in [t_{oij}; t_{oij} + T_0], & \text{если } \tilde{y}_i[t_{oij}] > Pd_i; \\ x^0, & \text{в остальных случаях,} \end{cases} \tag{2}$$

где Pd_i – пороговое значение ЧРПi; t_{oij} – точка локального максимума функции $\tilde{y}_i[t]$ такая, что $\tilde{y}_i[t_{oij}] \geqslant \tilde{y}_i[t] \ \forall t \in \dot{\mathbf{M}}(t_{0ij})$; $\tilde{y}_i[t]$ – функция обнаружения (1) i-го канала; $\dot{\mathbf{M}}(t_{0ij}) = \mathbf{M}(t_{0ij}) \setminus \{t_{0ij}\}$ – проколотая окрестность точки t_{0ij}; $\mathbf{M}(t_{0ij})$ – окрестность точки t_{0ij}; j – индекс локального максимума; $x^0 = const$ – константа, определяющая уровень сигнала, соответствующий отсутствию структурного элемента заданного типа на текущем фрагменте сигнала (например, уровень изолинии ЭКГ).

Цель работы. В работе предлагается наряду с традиционным представлением ЭКГ-сигнала предоставить врачу-кардиологу альтернативные диагностические признаки в удобном для восприятия графическом виде.

Альтернативное представление ЭКГ. Альтернативные диагностические признаки формируются в результате преобразования с помощью функций f_{ij} исходного пространства признаков, описывающих структурные элементы БМС с ЛСП . Для нормализации координат с учетом функции обнаружения (1) альтернативные диагностические признаки $\tilde{y}'^{t}_{ij}$ i-го канала

определяются по следующем выражению:

$$\tilde{y}'^{t}_{ij} = (y^{э}_{ij} - \tilde{y}^{t}_{ij}) \sqrt{\frac{\alpha_i}{1 + \alpha_i \sum\limits_{j} (y^{э}_{ij} - \tilde{y}^{t}_{ij})^2}}, \quad (3)$$

где $\tilde{y}^{t}_{ij}$ - координаты обнаруженного СЭ заданного класса с помощью решающего правила (2).

Рассмотрим пример обнаружения комплекса QRS в одном из каналов нелинейного фильтра по трем опорным точкам с использованием в качестве функций f_{ij} разделенных разностей 1-го порядка (далее индекс i опустим, т.к. будем рассматривать только один канал). Тогда каждый СЭ (комплекс QRS) в альтернативном признаковом пространстве описывается двумя признаками $\tilde{y}'^{t}_{j}$ $(j = 1, 2)$, рассчитанными по (3). Для ЭКГ с наличием желудочковых экстрасистол двумерная картина в альтернативном пространстве признаков (рис. 2) значительно отличается от аналогичного представления ЭКГ в норме (рис. 3).

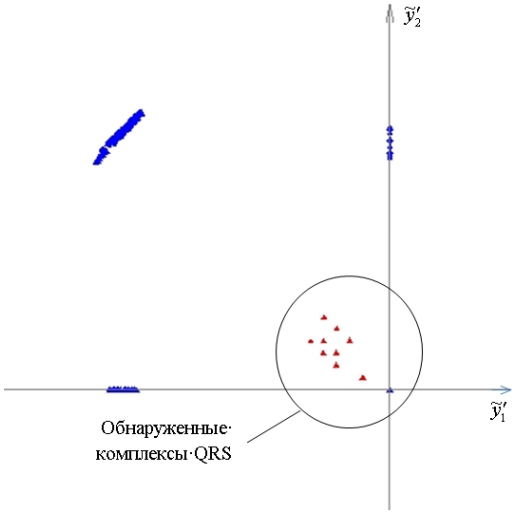


Рис. 1. Представление QRS-комплексов ЭКГ в норме

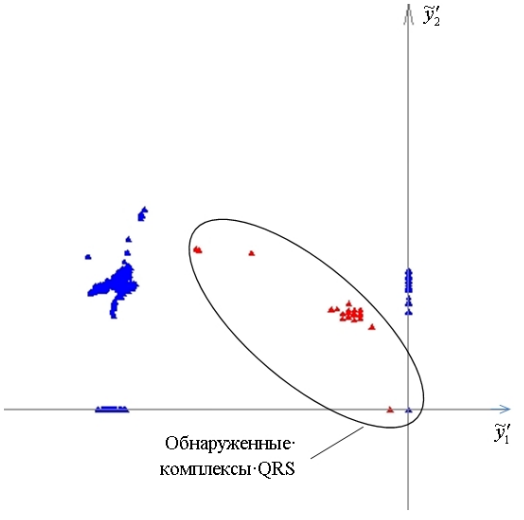


Рис. 2. Представление QRS-комплексов ЭКГ с экстрасистолами

Из рис. 2, 3 видно, что в случае обнаружения комплексов QRS на нормальной ЭКГ в альтернативном признаковом пространстве облако рассеяния найденных СЭ имеет сферическую форму, когда как в случае обнаружения комплексов QRS на ЭКГ с желудочковыми экстрасистолами конфигурация точек в альтернативном пространстве признаков имеет неправильную форму. Кроме того меняется конфигурация точек, которые соответствуют другим СЭ ЭКГ.

Выводы. В работе на основе разработанного метода обнаружения структурных элементов БМС с ЛСП предложено представление ЭКГ в альтернативном признаковом пространстве, в котором структурные элементы описываются адаптированными к модели полезного сигнала диагностическими признаками, максимально учитывающими свойства рассматриваемого биомедицинского сигнала с локально сосредоточенными признаками.

Дальнейшие исследования направлены на анализ информативности альтернативного пространства признаков и синтез диагностических правил в этом пространстве при проектировании интеллектуальных кардиологических систем поддержки принятия решений.

Литература. **1.** Здобникова Л.И. Телекардиология на догоспитальном этапе / Л.И. Здобникова [Електронный ресурс] http://tredex-company.com/ru/telekardiologiya-na-dogospitalnom-etape **2.** Поворознюк А.И. Проектирование нелинейного фильтра в задаче структурной идентификация биомедицинских сигналов с локально сосредоточенными признаками / А.И. Поворознюк, А.Е. Филатова // Системні дослідження та інформаційні технології. – 2014. – № 1. – С. 69-80.

Штельмах С.О., Олефір О.С.
Національний технічний університет України "КПІ", факультет прикладної математики, Київ, Україна

Дослідження методів очистки мовленнєвого сигналу від шуму

Кожен, хто працював з сигналами, стикався з шумом. На даний момент не існує універсального алгоритму для повного очищення сигналу від шуму. Розроблені методи можуть лише частково усунути цей дефект. Професіональна апаратура, унікальні алгоритми обробки звуку коштують тисячі доларів. Тому знаходження недорогого і ефективного методу для відокремлення корисної інформації з зашумленого сигналу залишається актуальною проблемою і сьогодні.

Модель вхідного сигналу. Вхідний сигнал y - N-вимірний вектор, який представляється в вигляді вибірки невідомої функції f на рівномірній часовій сітці t_i. ξ_i - незалежні випадкові величини, які належать одній і тій же функції нормального розподілу. Тоді

$$y_i = f(t_i) + \psi\xi_i, i = 0, \ldots, N-1 \tag{1}$$

де ψ - середньоквадратичне відхилення шуму. Задача полягає в тому, щоб отримати оцінку для функції f із зашумлених даних $y = (y_o, \ldots, y_{N-1})$, при невідомій параметричній моделі f. [1]

Пошук методів. В [2] зроблена класифікація і короткий аналіз методів для обробки зашумлених мовленнєвих сигналів. Тут автор виділяє два основні напрямки: підвищення якості і покращення розбірливості. Більшість методів побудовані таким чином, що покращуючи одну характеристику, вони погіршують іншу.

В даній роботі я досліджую методи для підвищення якості мовленнєвого сигналу. Основними критеріями для такого підходу є збереження натуральності голосу диктора і зниження втомленості аудитора. Методи, які базуються на оцінці спектральних характеристик шуму, добре зарекомендували себе в цьому напрямку. В [2] розглядається метод віднімання амплітудних спектрів шуму на базі перетворення Фур'є. В [1] замість перетворення Фур'є використовується вейвлет-перетворення. Такий підхід дає більше можливостей для модифікацій, він позбавлений недоліків, які притаманні методу Фур'є.

Схема очистки сигналу від шуму. Загальна схема очисти від шуму з використанням вейвлет перетворень включає в себе:

1. Пряме вейвлет-перетворення для вихідного сигналу $\{y_t : t = 0, \ldots, N-1\}$:
$$W = wy, \tag{2}$$
де W - вейвлет-коефіцієнти, ω - матриця перетворення, N - довжина вибірки.
2. Процедура усічення вейвлет-коефіцієнтів:
$$W_{thr} = T(W), \tag{3}$$
де T - оператор перетворення вейвлет-коефіцієнтів, яка залежить від способу порогової обробки і правила усічення.
3. Обернене вейвлет-перетворення, в результаті якого отримуємо оцінку істинної функції f:
$$f_{estim} = \omega^T W_{thr}, \tag{4}$$

Власне тут можна виділити два аспекти, над якими потрібно попрацювати: необхідно обрати базисну функцію для вейвлет перетворення; визначити правило, за яким буде виконано усічення.

Література. **1.** Орешко Н.И., Князева Т.Н. Вейвлет-технология анализа и очистки сигналов от шума [Текст] : УДК 519.218.82 Цифровая Обработка Сигналов. №3/2008 **2.** Чучупал В.Я., Чичагов А.С., Маковкин К.А. Цифровая фильтрация зашумленных речевых сигналов [Текст] : УДК 621.391;518.5 Вычислительный центр РАН Москва, 1998

Яковенко М.М.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Моделі і методи інтелектуального аналізу даних для оцінювання кредитоспроможності фізичних осіб

У роботі розглядаються основні існуючі підходи до оцінки кредитоспроможності фізичних осіб, побудовано моделі на основі нечіткої логіки, нейронної мережі та на основі рейтингового підходу.

Вступ. Для зменшення ризиків, які несе банк при кредитуванні фізичних осіб, дуже важливою задачею є визначення кредитоспроможності позичальника. Існує безліч визначень кредитоспроможності позичальника, найпоширенішим з яких є наступне: здатність особи повністю і в строк розрахуватися за своїми борговими зобов'язаннями. На жаль, на сьогодні не існує жодної ефективної методики оцінювання кредитоспроможності фізичних осіб, і банки застосовують різні способи, що не завжди вирішують поставлене завдання.

Для вирішення поставленої задачі існують різні моделі кредитного скорингу, що базуються на знаннях і даних стосовно клієнтів. Основним недоліком більшості моделей є те, що вони погано змінюються при надходженні нових даних, а також не завжди можуть правильно обробляти неточні та неповні дані про клієнтів. Тому на практиці зазвичай використовують сучасні моделі, що вміють адаптуватися під вхідні дані.

Результати роботи. У роботі побудовані наступні моделі:

1. На основі рейтингового підходу. Ця модель є найпростішим варіантом визначення кредитоспроможності. 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 кредитоспроможний, інакше – не кредитоспроможний. У такої моделі є два недоліки – суб'єктивність оцінок експертів, а також відсутність її адаптивності. До переваг моделі належить її простота. [1]
2. На основі нечіткої логіки. Системи нечіткого логічного висновку перетворюють значення вхідних змінних у вихідні з використанням нечітких правил. До недоліків такої системи також 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.
3. Нейронна мережа. Нейронна мережа самостійно визначає вклад кожного вхідного фактору у кінцевий результат. Недоліками цієї моделі є дуже висока складність її реалізації. До переваг можна віднести те, що вона може самостійно адаптуватися до нових даних, може враховувати взаємний вплив факторів, а також виявляти складні залежності.

У результаті роботи було визначено точність кожної побудованої моделі. Найточнішою виявилася нейронна мережа, за нею за рівнем точності виявилась модель на основі нечіткої логіки і на останньому місці – рейтингова модель. Також слід зазначити, що нейронна мережа виявилась дуже складною у реалізації. Тому можливим компромісом між складністю і точністю є модель на основі нечіткої логіки. Рейтингова модель виявилась дуже простою у реалізації, але вона дає дуже неточні результати порівняно з іншими методами.

Література. **1.** Скоринг как метод оценки кредитного риска [Електронний ресурс]. – Режим доступу: http://www.cfin.ru/finanalysis/banks/scoring.shtml

Яковлєва О.Є., Бідюк П.І.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Прогнозування динаміки процесів у фінансових системах

Задача прогнозування фінансових часових рядів – одна з високо актуальних задач економетричного аналізу і аналізу часових рядів, оскільки передбачування є необхідним елементом будь-якої інвестиційної діяльності, оскільки сама ідея інвестування – вкладення грошей з метою отримання прибутку у майбутньому – грунтується на ідеї прогнозування майбутніх значень досліджуваних процесів. Останнім часом, коли стали доступними потужні засоби збору, накопичення та обробки інформації, задача прогнозування фінансових часових рядів стала однією з найпопулярніших задач прикладної математики, що розв'язується різними, у тому числі комбінованими, методами.

Найбільш поширеними підходами до прогнозування динаміки процесів у фінансових системах є прогнозування на основі використання часового ряду однієї змінної, тобто на основі авторегресії (АР), авторегресії з ковзним середнім (АРКС) та авторегресії з інтегрованим ковзним середнім (АРІКС). При моделюванні та прогнозуванні будуються оптимальні за параметрами математичні моделі. Коефіцієнти моделі оптимізуються за допомогою відповідних критеріїв якості.

Аналіз та побудова моделей виконані за статистичними даними (ціна продажу акцій компаній ВАТ «Укртелеком», ВАТ «Київенерго» та ПАТ «Київський Річний Порт» [2]), всі дані у гривнях за часовий проміжок з 24.02.2014 по 19.01.2016 рр. Обчислювальні експерименти виконані за допомогою програми Eviews 8. Дані з 19.12.2015 по 19.02.2016 використано для оцінювання якості побудованих прогнозних моделей.

Наприклад для акцій компанії ВАТ «Укртелеком» рівняння моделей АР(1), АРКС(1,1), АРІКС(1,1,1) виглядають таким чином.

Модель АР(1):.

$$y(k) = 0,113846 + 0,668003y(k-1), \tag{1}$$

де y(k) – залежна змінна величина ціни купівлі акцій.

Модель АРКС(1,1):.

$$(y(k) = 0,113857 + 0,961049y(k-1) - 0,649627ma(k-1), \tag{2}$$

де y(k) – залежна змінна величина ціни купівлі акцій, а ma(k) – ковзне середнє.

Модель тренду 1-го порядку:.

$$y(k) = 0,116687 - 1,254192k, \tag{3}$$

де y(k) – залежна змінна величина ціни купівлі акцій, а k – змінна, що приймає значення від 1 до 20.

Модель АРІКС(1,1,1):.

$$y(k) = dy(k) + y(k-1), \tag{4}$$

де dy(k)=0.0978-0,030496y(k-1)-0,664271ma(k-1), dy(k) – перша різниця вихідного ряду, а ma(k) – ковзне середнє [1].

В результаті аналізу статистичних характеристик якості моделі та оцінок прогнозів визначено адекватні моделі досліджених процесів. Зроблено оцінку якості отриманих результатів прогнозування за множиною статистичних характеристик якості. За результатами статичного прогнозування ціни продажу акцій в середньому найкращою моделлю процесу виявилась АРІКС(1,1,1). При цьому отримане краще значення САПП =0,016195.

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.** Офіційний сайт інвестицій та фінансів [Електронний ресурс] – Доступний з:www.investfunds.ua.

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

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: Суперкомпьютерні, хмарні, Грід- та квантові обчислення, високопродуктивні системи та мережі..
2. Розробка та проектування мікросистем.

Antonyuk S.A., Artyukhov V.G.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Digital Signal Processing and Applications Design with the TMS320C6713 DSK

Introduction. Nowadays, digital signal processing plays a leading role in science, technology and industry, being used in different spheres, including communication equipment, audio and video channels encoding and compressing, signal filtration and commutation, data analysis and manipulation, spectral analysis, processes modelling and others, where rapid signal processing is required, including real-time handling. Progressive increase of digital signal processors usage is caused by necessity to obtain equivalent digital signal from the analog one by applying sampling and digitizing to analog signal and further manipulations with received digital signal.

TMS320C6713 DSP Characteristics. Among other digital signal processors, it is needed to highlight Texas Instruments TMS320 C6713 DSP distributed on board C6713 in TMS320C6713 DSP Starter Kit (DSK), which contains necessary hardware and software as well as information support for the correct work with the processor.

Such board and processor features give the opportunity to handle digital signals in the floating point format sampling the signals at 8-96 kHz sample rates with the means of advanced very long instruction word (VLIW) TMS320C67x DSP core operating at 225 MHz, organise audio data input and output through the widespread TRS 'mini-jack' socket with multichannel buffered serial ports utilizing and provide a user with the interactive feedback by LEDs and DIP switches, which are accessed by reading and writing to the CPLD registers.

Therefore, the avowed working processor frequency gives the opportunity to implement up to 1350 millions floating-point operations per second (MFLOPS), 1800 millions instructions per second (MIPS) and by means of floating or fixed point numeric dual multiplier - up to 450 millions multiply-accumulate operations per second (MMACS). But due to processor double-level cache architecture the executing speed of the program mostly depends on the operations to be carried out and can be enhanced through the system conveyor and calculations paralleling. So the actual program performing time can be determined by additional experiments and evaluations of each algorithm realization work duration.

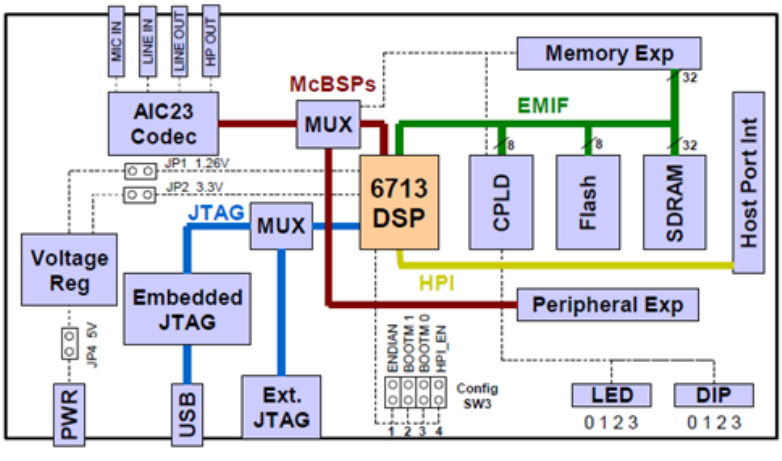


Figure 1. Block diagram C6713 DSK

Application design tools. Personal computer (PC) is necessary for applications design to carry out the instruction and data interchange with the board. There is dedicated software within the C6713DSK for specialized work - C6000 Code Composer Studio. This environment provides the capability to transmit data to and from the board by means of embedded JTAG emulator with a

USB host interface. The DSK can also be used with an external emulator through the external JTAG connector.

Methods of working with the package. It is needed to execute the board diagnostics prior to programming. For this purpose, there is a special utility incorporated in the program package, which implements tests of all board components as well as connection to the PC. A rapid board functionality test is also performed automatically right after the board is connected to the power supply. The successful result of Power-On Self Test (POST) is the blinking of all 4 LEDs for 3 times and their further remaining on.

The board must be connected to the PC via the USB-interface. Then, it is possible to start working in Code Composer Studio C6713 DSK.

After the program product launching is performed, it is compulsory to establish the connection via the menu bar, otherwise the project launched for execution will be carried out locally by the emulator, which has less possibilities and is used only for program code testing when there is no access to the board itself.

There is also a reference material available in the program package, which includes quick start guide, tutorial, user manual and other information materials about the DSK.

To see the system in operation, it is possible to upload one of the example projects distributed within the software to the environment. After this, the project must be built, and as the result the output project file is generated, specialized and ready for delivery to the board. When successful data uploading to the board is performed, the launch of project program can be made with the interface of Code Composer Studio. If the program is cyclic and implies the performance halt from the user, this must be done via the environment interface.

To verify if the board works correctly, the test example of 1 kHz sine signal generating can be launched. The result is driven to both the line out and fixed slots and can be heard via the headphones or seen on the oscilloscope.

If there is a new standalone projects design, then C source files must be included, and it's highly recommended to take standard project architecture as an example.

Conclusions and prospects. The main goal of work with the processor was experiments research stand development for further digital signal processing algorithms exploring and setting up work methods as well as processing power estimating for further experiments. Currently there is a developing project of real-time signals manipulating equalizer creation which is based on multi-bandpass IIR-filter (filters with infinite impulse response). MATLAB modelling environment and its Filter Design and Analysis Tool particularly are also considered to be used in this experiment.

Another important planned research direction is creation of heterogeneous system based on nearly 8 processors, which will give the opportunity to parallel signals processing procedures. Considering that operations on data transmitting take much less resources and fewer load the system comparing with the audio signals processing operations, such system holds out promises to achieve a great enhancement in digital signal processing algorithms work.

References. **1.** D. R. Brown III. (2007). *Digital Signal Processing and Applications with the TMS320C6713 DSK* [Online]. Available FTP: http://spinlab.wpi.edu Directory: courses/dspworkshop File: dspworkshop_part1_2007.pdf **2.** Spectrum Digital, Inc. (2003). *TMS320C6713 DSK Technical Reference* [Online]. Available FTP: http://www.etcs.ipfw.edu Directory: lin/ECET357/TI-TMS320C6713DSK File: dsk6713_TechRef.pdf **3.** R. Chassaing, *DSP Applications Using C and the TMS320C6x DSK*. New York: Wiley, 2002. **4.** R. Chassaing and D. S. Reay, *Digital Signal Processing and Applications with the TMS320C6713 and TMS320C6416 DSK*, 2nd ed. Hoboken: Wiley, 2008.

Kharchenko K.V.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

An Architecture and Test Implementation of Data Flow Virtual Machine

An architecture and principles of virtual machine for data flow computations are described. This virtual machine allows to run bytecode with description of data flow scheme of computing on various computer platforms.

Data flow paradigm for computational algorithms was proposed in early 1960 [1–4]. As well as command flow paradigm, it can be used for different areas of mathematical, logical and data processing applications. Most effective use of data flow paradigm can be digital signal processing, image processing, computer vision, control of robotic systems, control of complex business processes, etc. [5,6] There are many implementations of virtual machines for command flow paradigm and they are widely used in applied information systems, embedded devices, mobile systems and enterprise servers. A combination of these two well known concepts can be used for creation of the data flow virtual machine. This data flow virtual machine can input bytecode with description of netlist for connections of components, execute values of flows and produce its outputs. Advantages of such a kind of mixing concepts of the virtualization and data flow paradigm consist in universal tools that can be available on different operation systems, processor architectures, embedded systems with the power of the data flow paradigm, where many algorithms can be easily represented and implemented effectively. The theoretical part of the data flow based computations uses Khan process networks and Petri nets [7]. The experimental part for the data flow virtual machine consists of the following: module for loading a bytecode netlist from the input file, sorting the components in order to execute operations on the data flows, executable part. The implemented virtual machine supports integer type for the data flows, two inputs and one output for components, a dynamic type of the execution module for the data flows. The virtual machine is tested for different order of components in the netlist. To order components in netlist a simple sorting algorithm is used with complexity $O(n^2)$ and can be improved in the future releases. The test implementation of the data flow virtual machine is created on C++ with object oriented programming. For the multiplatform test of virtual machine Microsoft Windows 10 and Max OS X successfully used. C++ *gcc* compiler is used for cross platform support and the virtual machine is recompiled for each platform. On Windows platform *mingw* tools provides the necessary compilation support.

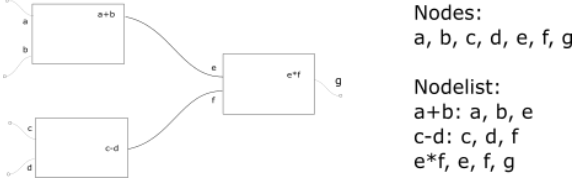


Figure 1. The test example for the data flow virtual machine

References. **1.** John L. Kelly Jr., Carol Lochbaum, V. A. Vyssotsky (1961). "A block diagram compiler". Bell System Tech. J. 40: 669–678. **2.** W.R. Sutherland (1966). "The On-line Graphical Specification of Computer Procedures". MIT. **3.** Jack B. Dennis. First version of a data flow procedure language. In "Lecture Notes in Computer Science 19: Programming Symposium pp 362-376. Springer-Verlag: Berlin, New York, 1974. **4.** J Paul Morrison, "Flow-Based Programming 2nd Edition: A New Approach to Application Development, CreateSpace, 2010. ISBN 1-451-54232-1 **5.** Labview, http://www.ni.com/labview/. **6.** NoFlo | Flow-Based Programming for JavaScript http://noflojs.org/. **7.** Kahn, G. (1974). Rosenfeld, Jack L., ed. The semantics of a simple language for parallel programming (PDF). Proc. IFIP Congress on Information Processing. North-Holland. ISBN 0-7204-2803-3

Kovalenko A.E.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Structure decomposition for system-level diagnosis models in distributed fault-tolerant systems

The technique and measures of the structure decomposition and appreciation of system syndrome are proposed based on splitting the diagnostic graph on the set of ordinary structures (trees, stars, chains). Some important properties of three-level cloud services classes (IaaS, SaaS, PaaS) for self-organizing diagnosable systems are discussed.

System-level diagnosis aims at the identification of faulty units in a self diagnosable distributed systems for elimination, repairing or recovering of these units. Identification is carried out by means of system analysis and/or data mining of diagnosed system (e.g. for syndromes, models of inter unit testing, system topology). Complexity of such analysis in a great extent depends on the preliminary appreciation of diagnosis model and evaluation of diagnosis processes characteristics.

Cloud computing and distributed systems . Artificial information systems with autonomic self-managing and high-level guidance from humans will be the dominating types of the future information systems generations.

Autonomic, or self-managing, systems rely on monitoring test probes and sensors, which indicate the technical states of system units (hardware and software), on an adaptation engine (autonomic manager) for computing based on monitoring data, and on control means to carry out changes on the system. Cloud computing inspire software technologies for data center automation, which may perform tasks such as: management of service levels of running applications (Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) [1, 2]; management of data center capacity; fault proactive disaster recovery; and automation of virtual machines (VM) provisioning. IaaS enables on-demand provisioning of servers running several choices of operating systems and a customized software stack.

Users are given privileges to perform numerous activities to the server, such as: starting and stopping it, customizing it by installing software packages, attaching virtual disks to it, and configuring access permissions and firewalls rules

Thus cloud computing gives an opportunity for creating the public, private, community, or hybrid clouds for self-organized distributed information systems with variety of virtualized structures, sensor networks and autonomic self-managing.

Fault-tolerant self-managing systems . The increasing complexity of distributed information systems has motivated researches to integrate the several technologies (cloud computing, NUMA), especially in hardware (virtualization, multi-core chips), Internet technologies (Web services, service-oriented architectures), distributed computing (clusters, grids), and systems management (autonomic computing, data center automation) [1,2]. Autonomic, fault-tolerant self-managing systems rely on the technical states of system units (hardware, software modules, services).

System-level diagnosis models . The most known system-level diagnosis models are PMC and BGM models, defined in the table.

Table 1. PMC, BGM system-level diagnosis models

Tester unit	Unit under test	PMC test result	BGM test result
fault-free	fault-free	0	0
fault-free	faulty	1	1
faulty	fault-free	0 or 1	0 or 1
faulty	faulty	0 or 1	1

System diagnosis in clouds. Let us define the structure of cloud system by means of diagnostic graph G(V,E), where V is the set of system autonomic units (elements, components, processes,

nodes, vertices) and E is the set of directed links $(v_i, v_j), v_i, v_j \in V$, between these units[3].

Every graph G(V,E) may be decomposed on some regular sub graphs G_j structures L_j in such a way that

$$G(V,E) = \bigcup_j G_j,\ G_j = (V_j, E_j), n_j = |V_j|$$

$$V_j \subseteq V, E_j \subseteq E, V = \bigcup_j V_j,\ E = \bigcup_j E_j.$$

For chain, star and tree structures $n_j - 1 = |E_j|$

Let us define the chain structure L_j as $L_j = L_j(G_j, A_j^i)$ where A_j^i is the system syndrome.

Lemma 1. For the connected graph G(V,E) covered by structures $L_j, j = 1, 2, \ldots, k$

$$n = \sum_{j=1}^{k} n_j - k + 1$$

Syndrome compatible sets . For every structureL_j there are syndrome compatible set (SCS) of unit states, not contradicted the given syndrome. Let us define the number of SCS for a given L_j and system syndrome A_j^i as $N(L_j)_i$. Then for the given A_j it may be evaluated as

$$N(L_j)_{max} = \max_i N(L_j)_i,$$

Such a structure L_j is called as maximum syndrome appreciation (MSA).

Lemma 2. The number of SCS for PMC 1 - chain structures with n elements is the MSA and is defined as

$$f_n = f_{n-1} + f_{n-2},\ n \geqslant 1, f_0 = 1; f_{-1} = 1.$$

This evaluation may be used for evaluation of the upper boundary measures in PMC structure set $L_1, L_2, \ldots, L_k$.

Lemma 3. Let $L_1, L_2, \ldots, L_k$ are the PMC chain structures of G=G(V,E) such as

$\forall$ l, m $\in$ 1, 2, ...,k , l $\neq$ m, $V_l \cap V_m = \varnothing$, $V_l, V_m \subseteq$ V. Then the number N of SCS for G(V,E) is satisfied the expression

$$N \leqslant 2^r \prod_{i=1}^{k} fn_i.$$

where $n_i = |V_i|$, r - $|V \setminus (V_1 \cup V_2 \cup ... \cup V_k)$ —

Problem solving methods . Some metaheuristic problem solving methods were considered for PMC models. Especially Fuzzy Bee Systems and Ant Systems approaches are enhanced on system diagnosis methods [3]. The PSO-based fault diagnosis may be useful by further enhancing of genetic algorithms, ant colonies and artificial immune system methods [4]. The swarm artificial models may be effective for quick and reliably navigation through the search space of all feasible sets of faulty units under the invalidation test models.

References. **1.** Cloud computing : principles and paradigms / edited by Rajkumar Buyya, James Broberg, Andrzej Goscinski.., Hoboken, New Jersey, 2011.-637 p. **2.** Kovalenko A.E. Distributed information systems / Kovalenko A.E..-Kyiv:NTUU "KPI", 2008.-244 p. (In Ukrainian). **3.** X.S. Yang, Nature-Inspired Metaheuristic Algorithms, Luniver Press, 2008. **4.** R. Falcon, M. Almeida and A. Nayak, "A Binary Particle Swarm Optimization Approach to Fault Diagnosis in Parallel and Distributed Systems", Proc. IEEE Congress on Evolutionary Computation (CEC), Barcelona, Spain, July 2010, pp. 41–48.

Бритов О.А., Тарасенко О.Г.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Програмований кодер циклічного коду

Пропонується опис апаратної реалізації кодера циклічних кодів, яка дозволяє динамічно змінювати параметри кода і утворюючий поліном.

Зберігання, пересилка та обробка інформації в інформаційних системах вимагають постійного контролю цілісності даних. Цей котроль найчастіше виконується за допомогою надлишкових кодів, які дозволяють виявляти та виправляти помилки. Програмне формування таких кодів вимагає досить великих витрат часу та пам'яті.

Проблема контролю цілісності інформації притаманна також вбудованим системам, які є невід'ємною частиною сучасних виробів різноманітного призначення. Зазвичай вбудовані системи є системами жорсткого реального часу, які повинні реагувати на зовнішні події з мінімальною затримкою.Сучасні вбудовані системи виконуються у вигляді систем на кристалі. Ці системи включають як програмовані, та і апаратурно реалізовані блоки, які виконують функції, необхідні для реалізації поведінки системи, заданої технічним завданням.

Виходячи з вимоги роботи кодека в реальному часі, кодування та декодування даних бажано реалізовувати апаратурно, у вигляді спеціалізованого блоку. На практиці для контролю рівня помилок часто використовують циклічні коди [1, 2]. Апаратурна реалізація кодерів циклічних кодів виявляється дуже компактною і ефективною.

Можливі два варіанти формування циклічного коду [1]. Перший полягає в обчисленні добутку інформаційного поліному на утворюючий поліном. Другий - в формуванні кодової комбінації у вигляді конкатенації інформаційної частини коду і залишку від ділення інформаційної частини, помноженої на 2^r, на утворюючий поліном. У другому варіанті іформаційні і контрольні біти коду розділені, завдяки чому інформаціна частина коду залишається незмінною. В термінах поліномів з двійковими коефіцієнтами поліном кодової комбінації утворюється згідно з рівнянням:

$$C(x) = U(x)x^r + R(x) \tag{1}$$

де C(X) - поліном отриманої кодової комбінації, U(x) - поліном інформаційної комбінації, R(x) - залишок від ділення, r - ступінь утворюючого полінома.

Цей варіант циклічного коду на практиці використовується частіше саме завдяки тому, що інформаційні біти залишаються незмінними.

Виявлення і виправлення помилок за допомогою циклічних кодів також основано на обчисленні залишку від ділення кодової комбінації на утворюючий поліном.

Таким чином, операція обчислення залишку від ділення є основною при кодуванні та декодування циклічних кодів. Від її реалізації залежить швидкісто роботи як кодера, так і декодера.

Ефективним пристроєм, який реалізує обчислення залишку від ділення, є регістр зсуву із зворотними зв'язками. Структура зворотних зв'язків визначається коєфіцієнтами утворюючого полінома (0 - зв'язку немає, 1 - є). При розробці апаратурного кодека поставлена задача побудувати його таким чином, щоб можна було змінювати параметри коду і утворюючий поліном в процесі роботи, не вносячи змін в схему пристрою.

Для розв'язання цієї задачі запропоновано побудувати регістр зсуву, який складається з комірок, схема яких наведена на Рис.1.

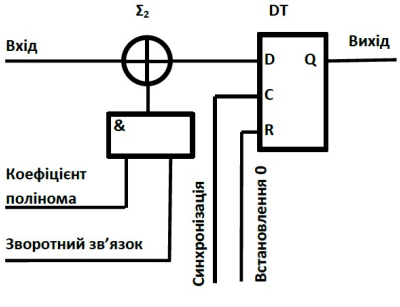


Рис. 1. Комірка регістра зсуву

Як видно зі схеми, комірка складається з однобітового суматора по модулю 2 Σ_2, тригера DT та логічного вентиля І.

Регістр зсуву реалізується послідовним з'єднанням однакових комірок. Вихід регістра зсуву подається на входи зворотного зв'язку комірок. Коєфіцієнти полінома подаються на входи комірок з окремого регістра, в який вони мають бути записані до початку кодування. В тих комірках, які видповідають нульовим коефіцієнтам полінома, виходом логічного вентиля є 0, і зворотний зв'язок відсутній. Відповідно, в комірках, які відповідають коефіцієнтам, що дорівнюють 1, сигнал зворотного зв'язку подається на суматор.

Таким чаном забезпечується можливість при потребі змінювати поліном, не змінюючи схему приладу. Зміна полінома і параметрів коду відбувається шляхом запису довжини кодової комбінації, порядку полінома та його коефіцієнтів в спеціально призначені для цього регістри.

Довжина кодової комбінації і порядок поліному записуються в регістр управління. В цьому ж регістрі знаходиться біт, який вказує на операцію, яка має бути виконана: кодвання чи декодування, а також біт запуску. Порядок поліному записується в регістр поліному.

Інформаційна частина кодової комбінації записується в регістр даних, доступний для читання та запису.

В регістрі стану формуються прапорці, які вказують на результат виконання операції. Зокрема, є біт готовності результата та біти помилок, які виявлені в процесі обробки даних.

Опис модуля розроблений на мові опису апаратури Verilog з використанням синтезованої підмножини мови. Завдяки цьому цей опис може бути використаний для апаратурної реалізації в проектах систем на кристалі без доопрацювання. Модулі пристрою параметризовані в максимально можливому ступені для розширення області можливих застосувань. Для параметризованої реалізації регістра зсуву на основі наведених вище комірок використаний опис на рівні вентилів у вигляді блока generate.

Тестування виконувалось за допомогою тестового модуля (testbench) в середовищі симулятора ModelSim. Результат тестування у вигляді часової діаграми наведений на Рис.2.

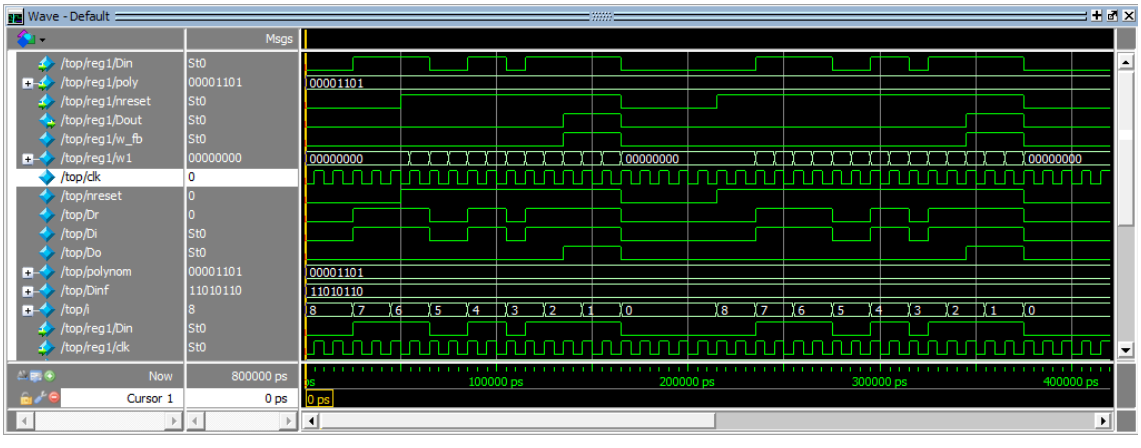


Рис. 2. Часові діаграми роботи кодера

На основі протестованого модуля розроблено і протестовано опис модуля кодека для роботи в складі системи на кристалі, побудованої на шині AMBA.

Пропонована схема придатна для реалізації кодеків циклічних кодів у вигляді автономних пристроїв або у вигляді ядер у складі систем на кристалі. При цьому конкретні параметри кода та вигляд утворюючого поліному задаються динамічно, в процесі роботи і можуь бути легко змінені в залежності від потреби.

Література. **1.** Дмитриев В.И. Прикладная теория информации.-М:Высшая школа, 1989.-320с. **2.** Жураковський Ю.П., Полторак В.П. Теорія інформації та кодування:Підручник.-К:Вища школа, 2001.-255с.

Булах Б.В., Крамар О.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Мікросервісна архітектура з семантичною складовою для комплексів інженерних розрахунків

Історично склалося, що програмна архітектура більшості систем інженерних розрахунків збудована за монолітним типом, коли усі складові тісно пов'язані у єдине програмне ядро. Оскільки такі системи за своїм призначенням мають охоплювати широке коло задач (і чим ширше — тим краще), то природним є бажання максимальної декомпозиції функціоналу системи на незалежні модулі, кожен з яким може розроблятись, розгортатись, підтримуватись окремо від інших. Це значно спростить підтримку усієї системи та нарощування її нового функціоналу без внесення змін в існуючий. На певному етапі вичерпним рішенням здавалася сервісно-орієнтована архітектура (СОА), однак, як виявилося, концепція екосистеми незалежних мережевих сервісів не завжди може бути успішно реалізавана на практиці. На зміну SOAP-сервісам, перевантаженим численними WS-* стандартами, прийшли REST-сервіси. Однак нині все більш популярною є нова "реінкарнація" ідеї СОА, що покликана ще більше мінімізувати зв'язки між сервісами — так звана архітектура мікросервісів. Мета даного дослідження - виявити переваги від застосування мікросервісів в програмній архітектурі систем інженерних розрахунків.

Концепція мікросервісів базується на наступних засадах [1]: чітка функціональна декомпозиція за принципом "одна задача — один мікросервіс"; зосередження логіки в кінцевих точках (сервісах), а не в каналах передачі даних (протиставляється СОА-патерну ESB – "шині" обміну повідомленнями між сервісами, відповідальній за узгоження численних інформаційних протоколів та форматів повідомлень, оркестрування сервісів, часто навантаженій й іншими функціями, відмінними від простої передачі даних від постачальників до споживачів); децентралізоване управління даними (жодних спільних централізованих БД); проектування "під відмову" (окремі мікросервіси мають зберігати працездатність в умовах імовірних відмов інших сервісів екосистеми); асинхронна взаємодія.

Наш попередній досвід реалізації комплексу схемотехнічного проектування з використанням SOAP-сервісів [2] виявив вузьке місце системи — оркеструвальник сервісів на основі стандарту WS-BPEL. Згадана вище ідея простих продуктивних каналів обміну повідомленнями може дати суттєву перевагу, дозволяючи уникнути централізованого маршрутизатора потоку робіт: кожний окремий мікросервіс тут діє автономно залежно від отриманих з каналу повідомлень, адресованих йому іншими сервісами. Такий підхід, відомий як хореографія сервісів, може бути також розвинутий залученням онтологій, роблячи сервіси ще менш зв'язаними.

Практична перевірка ідеї реалізації програмного забезпечення для інженерних розрахунків як екосистеми мікросервісів (з використанням прототипу системи на базі програмного забезпечення Apache Kafka – розподіленої вископродуктивної системи передач повідомлень з відкритим кодом) показала перспективність даного підходу: додавання нових сервісів та їх підтримка чинить мінімальний вплив на решту складових програмної системи, цей ефект посилюється з використанням онтологій у міжсервісній взаємодії. Недоліком даного підходу є накладні витрати на асинхронну міжсервісну мережеву взаємодію, що обмежують рівень можливої декомпозиції функціоналу системи. Дослідження будуть продовжені в рамках задачі функціональної декомпозиції на мікросервіси функціоналу програмного забезпечення для вирішення задач схемотехнічного проектування.

Література. **1.** Fowler M. Microservices: a definition of this new architectural term [Електронний ресурс]. – Режим доступу: http://martinfowler.com/articles/microservices.html (дата звернення 03.03.2016). **2.** Ladogubets V. Employing BPEL Engines for Engineering Calculations / Ladogubets V., Bulakh B., Chekaliuk V., Kramar O. // The Experience of Designing and Application of CAD Systems in Microelectronics : 12-th Intern. Conf. «CADSM'2013», 19-23 February 2013, Polyana-Svalyava (Zakarpattya), Ukraine : proc. – Lviv, 2013. – P. 427–430. – ISBN 978-617-607-393-2.

Денисов С.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Взаємодія агентної моделі та мікросервісів

З моменту свого заснування, перед областю комп'ютерних наук дуже часто поставала проблема організації складних систем. На ранніх етапах такі проблеми вирішувались правильним підбором структур даних чи виділенням прав для обмеження користувачів. З розвитком технологій, цього стало недостатнім і потрібно було продумувати структуру своїх систем – архітектуру. В свой більшості, розробники прагнули до організації більш логічного і читаємого коду, і не чіпали ідеї розміщення. Це дало те, що зазвичай, коли хтось говорить про клієнт-серверну архітектуру, слухач незважаючи на розмір системи, уявляє її архітектуру як монолітний сервер, які виконує запити користувача. Проте в останні роки, коли хмарні технології набувають все більшої і більшої популярності такий архітектурний підхід виявляється не дуже зручним – наприклад, при зміні невеличких кусків кода, треба передеплоювати цілу систему. [1] Саме тому доцільно замислитись про SOA – сервіс-орієнтовану архітектуру, яка передбачає модульність та слабку зв'язаність системи, забезпечену інтерфейсами для зв'язку.

Одним з передових принципів роботи за SOA – будувати додатки на мікросервісах. Архітектурний стиль мікросервісів – це підхід при якому один великий застосунок будується з набору невеликих сервісів, кожен з яких працює у власному процесі і спілкується з іншими сервісами за допомогою простих протоколів. [2]

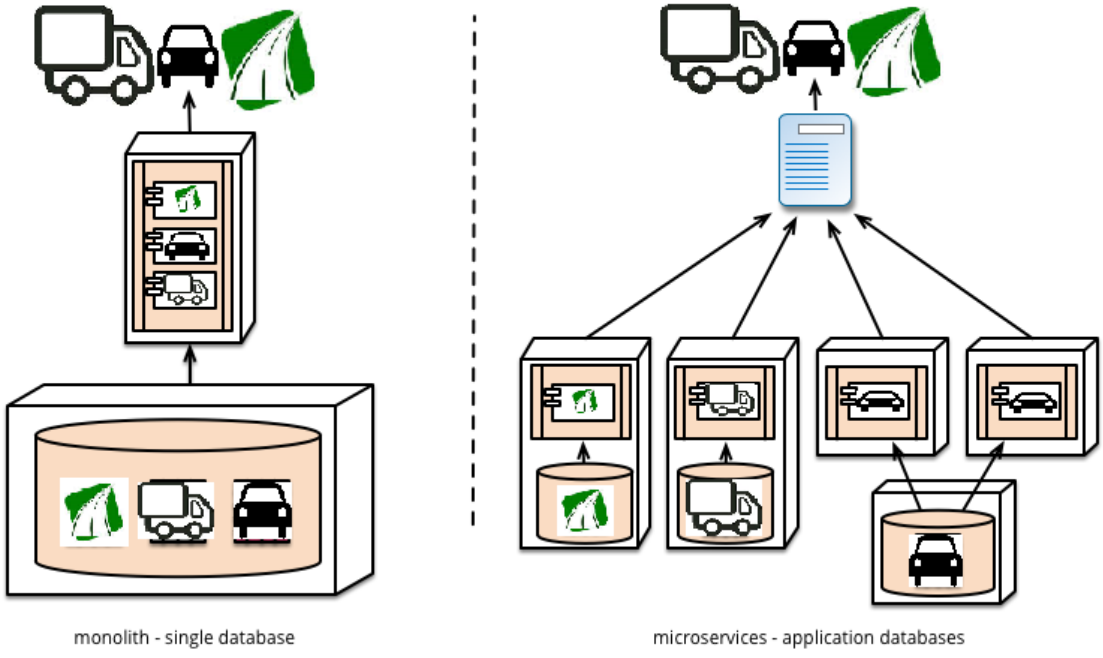


Рис. 1. Відмінність монолітної архітектури від архітектури на мікросервісах. [2]

Однією з найважливіших властивостей даної архітектури є децентралізація застосунку, що дає певні позитивні і негативні сторони:

- Якщо не працює якийсь сервіс, то не працює лише він;
- Якщо вносяться зміни в якийсь сервіс, то необхідно редеплоїти не весь додаток, а змінений сервіс і якщо змінився API, то і переважно всі сервіси які спілкуються зі зміненим;
- Для бізнес-використання дуже важливим є налаштування швидкої та безперебійної комунікації між сервісами, а також уникнення можливості передачі некоректних даних по сервісам;

- На відміну від монолітних додатків, де зазвичай для вирішення проблем не завжди виходить використовувати найкращі рішення через архітектурні обмеження, мікросервіси можуть вирішувати проблеми найбільш підходящим способом. [3]

В основі програмного забезпечення на мікросервісах можна розгледіти концепції, які використовуються також в сфері імітаційного моделювання, а саме - агентному моделюванні. Агентне моделювання передбачає зображення системи за допомогою опису поведінки і спілкування найпростіших елементів системи (агентів) та розміщення їх у певному середовищі. Простота агентної моделі полягає в тому, що системи описується лише простими правилами поведінки найменших елементів, що відкидає необхідність складної математичної моделі як в інших методах моделювання. [4, 5] В модель мікросервісів агентна модель впроваджується шляхом створення одного сервісу, який виконує роль середовища та необхідної кількості агентів-мікросервісів. Сервіс-середовища чекає та оброблює інформацію від мікросервісів-агентів, а також є єдиною точкою системи, яка контактує з користувачем. Мікросервіси-агенти під час моделювання посилають на сервіс-середовище інформацію про свій стан і очікують та оброблюють відповідь, на реакцію самого середовища на зміни, спричинені агентами. На відміну від десктопного варіанту моделювання, де при створенні великої кількості складних агентів, можна досягти межі працездатності системи. У випадку мікросервісного підходу до моделлювання, кожен мікросервіс окремо відповідає за поведінку моделі агента і сама модель може бути ускладнена наскільки того потребує експеримент. фівафівафвіафівафіва

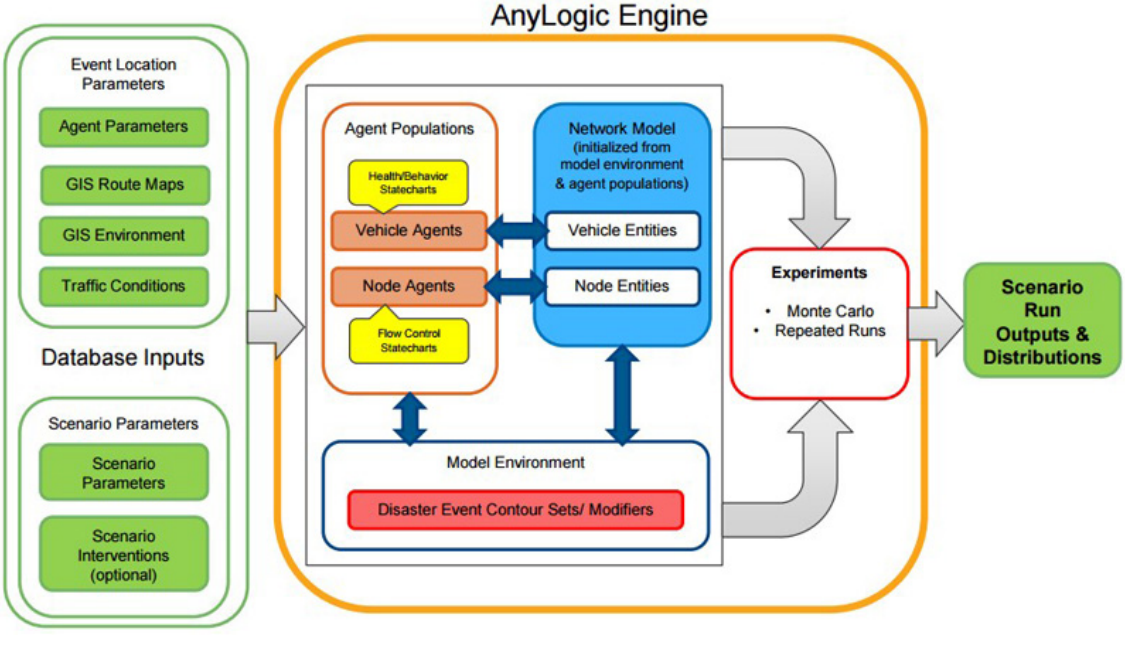


Рис. 2. Графічне зображення моделі руху транспорту, призначеного для реагування на надзвичайні ситуації, запропонованої компанією AnyLogic. [5]

Висновок. Хоч архітектура мікросервісів ще не досить поширеня і популяризована, але вже є певні галузі комп'ютерних наук, для яких впровадження мікросервісів може спростити чи покращити як процес розробки так і деякі функціональні можливості застосунків.

Література. **1.** Intuit. Архитектура приложений. http://www.intuit.ru/studies/courses/995/152/lecture/4232. **2.** Хабрахабр. Микросервисы. https://habrahabr.ru/post/249183/. **3.** Хабрахабр. Плюсы микросервисной архитектуры. https://habrahabr.ru/post/261237/. **4.** Хабрахабр. Имитационное моделирование. https://habrahabr.ru/post/246307/. **5.** AnyLogic. Имитационное моделирование. http://www.anylogic.ru/use-of-simulation.

Капшук О.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Аутентифікація користувачів в інформаційних системах і веб-сервісах з використанням двохфакторних і біометричних технологій

В роботі розглядаються існуючі технології аутентифікації користувачів в інформаційних системах і веб-сервисах, проблеми безпеки і напрями підвищення надійності аутентифікації за рахунок використання двохфакторних і біометричних технологій.

Багато компаній за кордоном і в Україні для організації праці своїх працівників починають використовувати модель BYOD (Bring Your Own Device (BYOD)), в якій передбачається можливість використовувати на роботі особисті пристрої. При цьому збільшується кількість мобільних пристроїв, за допомогою яких офісні і віддалені співробітники можуть працювати з конфіденційною корпоративною інформацією: електронною поштою, сховищами документів, різними бізнес-застосуваннями та ін. Протидія погрозам інформаційній безпеці (ІБ), що виникають при цьому, - одна з найбільш актуальних і складних проблем ІБ, яка вимагає впровадження надійних систем мобільної аутентифікації, що поєднують в собі зручність використання і посилений захист. До 2020 року на Землі очікується 26 млрд пристроїв з доступом до інтернету і 6,1 млрд. користувачів смартфонів. Це приведе до швидкого розвитку мобільного інтернету, до якого через п'ять років буде підключено 90% територій планети [1].

№ пп	Компанія	Програмний додаток/ пристрій	Технологія аутентифікації	Характеристики технології	Рік випуску/джерело
1.	Ratoc	SREX - FSU3	Біометрична за відбитками пальців,	3D оптичний сенсор, роздільна здатність 508 точок на дюйм, 256 градацій сірого, підключення до ПК через USB, сумісна як з вбудованими в ОС Microsoft Windows засобами біометричної ідентифікації користувача, так і з пакетом програмного забезпечення OmniPass. .	2016/[5]
2.	Nuance Enterprise	Dragon	Біометрична, голосовий пароль	Дистанційна аутентифікація за голосом з допомогою мобільних пристроїв	2015/[6]
3.	Microsoft	Смартфони Lumia 950 і Lumia 950 XL	Біометрична з використанням іридосканера	ОС Windows 10 Mobile з підтримкою сканування веселкової оболонки ока за допомогою засобів біометричної аутентифікації Windows Hello.	2015/[7]
4.	Intel Security	True Key™	Біометрична мультифакторна	Технологія надає доступ до сайтів, додатків і пристроїв шляхом розпізнавання рис обличчя, відбитків пальців або пристроїв, що належать власникові.	2015/[8]
5.	VisionLabs	VisionLabs LUNA	Біометрична з використанням технології рзпізнавання обличчя	Система розпізнавання обличчя клієнтів. Використовується для ідентифікації і верифікації осіб клієнтів банків, роздрібних магазинів. Біометричний модуль VisionLabs Luna інтегрований у новий сервіс по запобіганню кредитному шахрайству - Biometric Fraud Prevention Service	2014/[9]
6.	Sensory, Inc	TrulySecure™	Біометрична мультифакторна	Технологія дозволяє реалізувати технологію біометричної аутентифікації користувача за голосом і зображенням обличчя та інтегрувати її в мобільні пристрої.	-/[10]

Рис. 1. Приклади реалізації біометричних технологій аутентифікації

В даний час починається широке впровадження двохфакторної технології аутентифікації (2FA) з використанням мобільних телефонів, смартфонов, планшетів і інших мобільних пристроїв для отримання доступу до різних веб-сервисів. Відомі компанії Google, Yandex, Apple з 2015 року пропонують 2FA аутентифікацію [2–4].

В роботі проводиться аналіз двохфакторних і багатофакторних сервісів мобільної біометричної аутентифікації (МБА) з використанням технологій розпізнавання зображеня обличчя користувача, його голосу або відбитків пальців без використання додаткових апаратних засобів. Основні характеристики розглянутих технологій [5–10] наведені на рис.1. МБА приваблива для користувача ще і тим, що зазвичай вона достатньо проста у використанні, при цьому фактори аутентифікації не можна забути, втратити або поміняти, вони завжди залишаються при користувачеві. В якості прикладу реалізації і можливостей МБА можна навести розробки компанії Sensory, Inc., яка займається впровадженням технологій розпізнавания голосу і систем технічного зору у мільйони виробів від сотень провідних виробників споживчої електроніки .Компанія розробила технології:TrulyHandsfree™,TrulyNatural™ і TrulySecure™. TrulyHandsfree™-це технологія управління голосом для широкого кола безконтактних споживчих застосувань [10]. Технологія TrulyNatural™ надає розробникам можливість будувати системи розпізнавання голосу на базі нейронних мереж з різним розміром словників, а технологія TrulySecure™ - об'єднати технології розпізнавання голосу і зображення обличчя і, таким чином, реалізувати двохфакторну біометричну аутентифікацію користувача.

ВИСНОВКИ 1.Технологія мобільної багатофакторної біометричної аутентифікації знімає проблему використання великої кількості складних паролів і істотно підвищує рівень комфортності процесяу проходження аутентифікації.

2. Об'єднання засобів розпізнавання голосу і зображення обличчя (або веселкової оболонки ока) користувача дозволяє реалізувати повністю безконтактну технологію аутентифікації і відмовитись від необхідності вводу пароля.

3. Поява в мобільних пристроях (смартфони, планшети та ін.) вбудованих засобів отримання і попередньої обробки 2D і 3D зображень дозволяє реалізувати технології мобільної біометричної аутентифікації користувачів з використанням таких біометричних параметрів користувача як відбитки пальців, зображення обличчя і веселкової оболонки ока без додаткових апаратних засобів і забезпечити достатній рівень безпеки при дистанційному виконанні фінансових операцій.

Література. **1.** Количество пользователей смартфонов превысит 6 млрд.к 2020году [Електронний ресурс].– Режим доступу http:yamobi.rupostskolichestvo polzovateley smartfonov prevyisit 6 mlrd k 2020 godu.html : -Дата доступу 10.03.2016. – Назва з екрану. **2.** Google тестирует возможность аутентификации без пароля [Електронний ресурс]. – Режим доступу – http://www.3dnews.ru/925744/print/ : – Дата доступу: 10.03.2016. – Назва з екрану. **3.** Двухфакторная аутентификация пароля [Електронний ресурс]. – Режим доступу – https://www.yandex.ru/company/services news/2015/0203/ : – Дата доступу: 10.03.2016. – Назва з екрану. **4.** Apple выпустила iTunes 12.3 с поддержкой двухфакторной аутентификации [Електронний ресурс]. – Режим доступу – http://www.3dnews.ru /920406/: – Дата доступу: 10.03.2016. – Назва з екрану. **5.** Ratoc SREX-FSU3 наделит ПК средствами биометрической аутентификации [Електронний ресурс]. – Режим доступу:http://www.3dnews.ru/924553/ : – Дата доступу 10.03.2016. – Назва з екрану. **6.** Голландский банк предложил клиентам перейти на голосовую аутентификацию [Електронний ресурс]. – Режим доступу:http://www.3dnews.ru/917890/print/ : – Дата доступу: 10.03.2016. – Назва з екрану. **7.** Microsoft обещает больше устройств с иридосканером [Електронний ресурс]. – Режим доступу – http://www.3dnews.ru/923852/: – Дата доступу: 10.03.2016. – Назва з екрану. **8.** Технология True Key™ позволяет отказаться от необходимости запоминать пароли [Електронний ресурс].- Режим доступу-:http:www .prestigio.by/ news/company news ru/True key ru/:-Дата доступу:10.03.2016. **9.** VisionLabs LUNA – система распознавания лиц клиентов [Електронний ресурс].– Режим доступу-http:/www .visionlabs.ru/tech-section/: -Дата доступу: 10.03.2016. – Назва з екрану. **10.** Technologies for a Secure and Superior User Experience. Sensory's Truly Secure Technology [Електронний ресурс]. – Режим доступу – http://www.sensory.com/products/ technologies/truly secure /: – Дата доступу: 10.03.2016. – Назва з екрану.

Кирюша Б.А.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Використання масиву MEMs акселерометрів, розташованих на одному кристалі IC, для зменшення похибки вимірювань

Метою роботи є зменшення похибки та прискорення вимірювань без виконання розробки нового MEMs сенсору. Засобами досягнення мети обрано збільшення кількості сенсорів, їх початкових станів та способів розташування. Розраховані обмеження практичної вживаності такого підходу. Аналіз способів обробки даних з масиву акселерометрів виконано для одномірного лінійного сенсору.

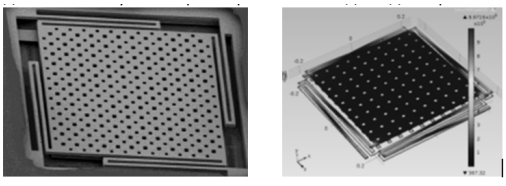


Рис. 1. MEMs акселерометр та його модель, розроблені для проекту EDUMEMS

Найбільший вклад до похибки вимірювань МЕМС давачем мають тепловий шум та ізотропні властивості матеріалів (ймовірнісна та дрейфова складові похибки). Оскільки розподілення похибки вимірювань майже рівномірне, то використання рухомого-середнього є нефективним засобом зменшення похибки. [1].

Розглянем роботу метода [2], який полягає у рішенні прямої і зворотньої задач кінематики на масиві сенсорів, які розташовани у діагональних напрямках. Це дає змогу: виділення лінійної складової із різних напрямків вимірювання, компенсувати асимметричність чутливості окремих сенсорів. Ще одним засобом розширення чутливості акселерометру є використання в кожній площині, принаймі, трьох сенсорів: один з нульовим зсувом, два - з максимальним зміщенням в позитивному і негативному напрямках, віповідно. Таким чином розширюється простір можливих значень та формується більш наближене до нормального розподілення похибки. Формули для прямої та інферсної кінематики мають вигляд: Пряма кінематика $a_i = G_i s$, де a - дані з сенсорів, d – зсув сенсорів, x, y, z – лінійне прискорення тіла, p, q, r – кутове прискорення тіла, $s = \begin{bmatrix} x" & y" & z" & p^2 & q^2 & r^2 & p' & q' & r' & pq & qr & rp \end{bmatrix}^T$

$$\mathbf{G}_i = \begin{bmatrix} 1 & 0 & 0 & 0 & -d_x & -d_x & 0 & d_z & -d_y & d_y & 0 & d_z \\ 0 & 1 & 0 & -d_y & 0 & -d_y & -d_z & 0 & d_x & d_x & d_z & 0 \\ 0 & 0 & 1 & -d_z & -d_z & 0 & d_y & -d_x & 0 & 0 & d_y & d_x \end{bmatrix}.$$

Зворотня кінематика для 4-х сенсорів: $s = \begin{bmatrix} G_1^T & G_2^T & G_3^T & G_4^T \end{bmatrix}^{T-1} \begin{bmatrix} a_1^T & a_2^T & a_3^T & a_4^T \end{bmatrix}^T$

Задача розміщення масиву сенсорів на одній IC. Використання лише однієї площини зменшує кількість сенсорів, які можуть бути враховані в моделі обчислень з інверсною кінематикою до 4-х: 2 сенсори з діагональною орієнтацією і два сенсори з різними початковими площинами розташування мас. Використання сенсорів, з однією загальною площиною розташування мас робить матриці, наведені вище, виродженими. Таким чином, координати сенсорів матимуть вигляд: 1 - d=[-dx 0 dz] r=0, 2 - d=[dx 0 -dz] r=0, 3 - d=[0 dy 0] r=-45o, 4 - d=[0 -dy 0] r=45o. Для подальшого моделювання та тестування запропонованого рішення планується розробка модифікованого алгоритму генерації топології на базі алгоритму DPLACE [3].

Література. **1.** Michalik P., An Approach to MEMS Smart Sensor Design using CMOS BEOL, Proc. MEMSWAVE 2013. **2.** Sebastian O.H. Measuring motion with kinematically redundant accelerometer arrays: Theory, simulation and implementation, Solid-State Sensors, Actuators and Microsystems 17, pages 22-25 – 2013. **3.** T.F. Chan, Multilevel generalized force-directed method for circuit placement. In Proc. Int. Symp. on Phys. Design, pages 185–192, 2005.

Литвиненко А.М.
Національний технічний університет України "КПІ", Київ, Україна

Розподілена обробка великих об'ємів зображень за допомогою технологій Hadoop

Обсяг інформації в світі росте за експоненційними законами, і його левова частка належить до неструктурованих даних. Проблема коректної інтерпретації інформаційних потоків стає все більш актуальною й одночасно складною.

Існує багато алгоритмів обробки зображень. У даній статті пропонується модифікація існуючих ітераційних алгоритмів, яка, шляхом застосування технології Hadoop MapReduce, дозволяє створювати масштабовані системи обробки великих об'ємів зображень.

Опис рішення. Програмна платформа Hadoop побудована на основі моделі MapReduce, яка накладає певні обмеження на застосування алгоритмів, вхідних та вихідних даних.

Програмна модель MapReduce працює з структурами даних типу ключ-значення [1].

Користувач MapReduce висловлює обчислення у вигляді двох Функції: Map і Reduce.

Функція Map створюється користувачем, приймає вхідну пару і продукує набір проміжних пар ключ-значення. Бібліотека MapReduce об'єднує всі проміжні значення, пов'язані з тим же самим проміжним ключем, і передає їх до функції Reduce. Функція Reduce також створюється користувачем. Вхідними даними функції є ключ і набір значень для цього ключа. Функція Reduce об'єднує ці значення з метою зменшити їх кількість для одного ключа. Часто на виході залишається тільки одне або не залишається жодного значення для кожного ключа [1].

Для застосування алгоритму в програмній моделі MapReduce потрібно привести структуру даних до вигляду пар ключ-значення і результатом роботи алгоритму також мають бути пари ключ-значення [1]. Результатами алгоритмів обробки зображень часто виступають певним чином модифіковані зображення або властивості вхідного зображення (наприклад, колір, середня інтенсивність і т.д.) [2]. У випадку, коли результатом алгоритму є зображення, то воно подається у вигляді пари ключ-значення, де ключ є ідентифікатором вихідного зображення, а значенням - саме зображення. Коли результатом алгоритму є властивість з обмеженого набору, то його подаємо в вигляді пари ключ-значення, де ключем буде ознака, а значенням - ідентифікатор вхідного зображення.

Таким чином, в тілі функції Map є можливість застосувати відомий алгоритм обробки зображення і, за допомогою моделі MapReduce [1], застосувати її для розподіленої обробки масиву зображень. На стадії Reduce результати групуються за ключами, що дозволяє по ідентифікатору записати вихідне зображення до файлової системи, чи отримати результати обробки, що згруповані за властивостями вхідного зображення.

Висновок. Запропонована модифікація алгоритму дозволяє застосовувати його разом з програмною моделлю MapReduce для розподіленої обробки масивів зображень.

Використання MapReduce дає можливість розпаралелювати обчислення на багатьох машинах та вирішувати задачі обробки великих обсягів зображень.

Однак, слід зазначити, що використання MapReduce доцільне в першу чергу при обробці великих об'ємів даних, які неможливо або економічно недоцільно тримати чи обробляти на одній машині [3].

Література. **1.** Dean, Jeffrey and Ghemawat, Sanjay "MapReduce: Simplified Data Processing on Large Clusters"in OSDI '04: 6th Symposium on Operating Systems Design and Implementation. — USENIX, 2004. P. 137-149. **2.** Потапов А. А., Пахомов А. А., Никитин С. А., Гуляев Ю. В., "Новейшие методы обработки изображений."— М.: Физматлит, 2008. **3.** Уайт, Том. "Hadoop. Подробное руководство = Hadoop: The Definitive Guide."— 2-е. — СПб.: Питер, 2013.

Масліков В.
Національний технічний університет України «КПІ», Київ, Україна

Використання графічних процесорів для розв'язання задачі фолдингу білків

Білки є одним з найважливіших класів біологічних макромолекул - вони виконують структурні функції у м'язах та кістках, каталітичні функції у всіх біохімічних реакціях, які формують метаболізм, і вони координують рух та трансдукцію сигналів [1]. Таким чином, білки є базовими одиницями життя, і гарне розуміння їхньої структури і функцій приведе до кращого розуміння процесів, що протікають у живому організмі. Як тільки, білок синтезований у вигляді лінійної послідовності амінокислот, він згортається протягом лічених секунд у стабільну тривимірну структуру - у цьому стані білок може виконувати свої функції в організмі. Визначення тривимірної структури білка на основі його амінокислотної послідовності є однією з головних задач біоінофрматики, і має багато застосувань у медицині (створення ліків, передбачення хвороб) та генній інженерії (моделювання клітин, модифікація та прокращення функцій окремих білків).

На відміну від інших біологічних макромолекул, білки мають складну структуру, яку складно передбачувати, тому застосовуються різноманітні наближені методи обчислення. У зв'язку з цим актуальною науковою задачею є розробка та реалізація ефективних алгоритмів розв'язання цієї задачі.

Графічні процесори (GPU), як потужні обчислювальні ресурси, використовуються у багатьох областях, у тому числі і для розв'язання задачі фолдингу білків [2]. Для розв'язання задачі фолдингу на GPU необхідно обрати модель та алгоритм, що дозволять ефективно розпаралелити обчислення з врахуванням особливостей архітектури процесору та пам'яті. У даному дослідженні запропоновано використання генетичного алгоритму для розв'язання задачі фолдингу на GPU в рамках "AB off-lattice" моделі.

В "AB off-lattice" моделі [3] послідовність амінокислот білка представлена лише двома видами мономерів: "A" - гідрофобними та "B" - гідрофільними. Відстань між мономерами (σ_i) прирівнюється одиниці, і кожен мономер σ_i приєднаний до наступного мономера σ_{i+1} через зв'язок який утворює кути θ та α відносно свого попередника. Загалом існує 20 стандартних амінокислот і кожну з них можна охарактеризувати як гідрофобну чи гідрофільну. Модель визначає значення енергії для мономерів: "A" має енергію 1, а "B" має енергію -1. Таким чином взаємодія різних типів мономерів (гідрофобних чи полярних) має різну енергію C. Додатні значення означають притягання, а від'ємні відштовхування. Функція енергії (цільова функція) може бути записана наступним чином [3]:

$$E = -k_1 \sum_{i=1}^{N-2} \widehat{b}_i \cdot \widehat{b}_{i+1} - k_2 \sum_{i=1}^{N-3} \widehat{b}_i \cdot \widehat{b}_{i+2} + \sum_{i=1}^{N-2} \sum_{j=i+2}^{N} E_{ij}(r_{ij}, \sigma_i, \sigma_j), \tag{1}$$

де $\widehat{b}_i$ - вектор зв'язку між мономерами σ_i та σ_{i+1} одиничної довжини. Оскільки $\widehat{b}_i \cdot \widehat{b}_{i+1} = \cos\theta_i$, $\widehat{b}_i \cdot \widehat{b}_{i+2} = \cos\alpha_i$, білок довжини N може бути представлений $N-2$ кутами зв'язку та $N-3$ двогранними (торсійними) кутами. r_{ij} - це Евклідова відстань між мономерами σ_i та σ_j, що залежить від кутів зв'язку та торсійних кутів. Енергія взаємодії між двома мономерами визначається через потенціал Леннард-Джонса (Lennard-Jones potential):

$$E_{ij}(r_{ij}, \sigma_i, \sigma_j) = 4C(\sigma_i, \sigma_j)\left(\frac{1}{r_{ij}^{12}} - \frac{1}{r_{ij}^{6}}\right). \tag{2}$$

Значення $C(\sigma_i, \sigma_j)$ обирається таким чином, щоб сприяти формуванню гідрофобного ядра з мономерів типу "A":

$$C(\sigma_i, \sigma_j) = \begin{cases} 1, & \sigma_i = \sigma_j = A, \\ 0.5 & \text{в інших випадках.} \end{cases} \tag{3}$$

Два параметри цільової функції k_1 та k_2 визначають силу локальної взаємодії мономерів що не залежить від їх типів. Отже, для будь-якої молекули, що має N мономерів, треба визначити $N-2$ кути зв'язку θ_i та $N-3$ торсійні кути α_i.

Генетичний алгоритм - це еволюційний метод пошуку, навіяний біологічним процесом еволюції. Він добре підходить для паралельних обчислень, оскільки операції схрещування, мутації та обчислення цільової функції можуть виконуватись паралельно для всіх елементів популяції. Далі наводяться генетичні оператори для задачі фолдингу білків в рамках "AB off-lattice" моделі.

Схрещування елементів популяції, що мають вигляд $h_i = \{\theta_2^i, \theta_3^i, ..., \theta_{n-1}^i, \alpha_3^i, \alpha_4^i, ..., \alpha_{n-1}^i\}$ та $h_j = \{\theta_2^j, \theta_3^j, ..., \theta_{n-1}^j, \alpha_3^j, \alpha_4^j, ..., \alpha_{n-1}^j\}$ відбувається за формулами:

$$\theta_m^{'} = r\theta_m^i + (1-r)\theta_m^j, \quad m \in \overline{2, n-1}, \tag{4}$$

$$\alpha_m^{'} = r\alpha_m^i + (1-r)\alpha_m^j, \quad m \in \overline{3, n-1}. \tag{5}$$

Після схрещування з певною ймовірністю відбувається мутація отриманого елементу. В процесі мутації всі кути $\theta_m^{'}$ та $\alpha_m^{'}$ змінюються на невелике значення r_m, тобто:

$$\theta_m^{''} = r_m\theta_m^{'}, \quad m \in \overline{2, n-1}, \tag{6}$$

$$\alpha_m^{''} = r_m\alpha_m^{'}, \quad m \in \overline{3, n-1}. \tag{7}$$

Після мутацій розраховується значення $\Delta E = E - E_{min}$, тобто різниця між фітнес-функцією отриманого значення і фітнес-функцією рекордного значення. Якщо $\Delta E < 0$, то конформація приймається, у іншому випадку конформація приймається з ймовірністю $e^{-(\Delta E\,/\,T)}$ (процедура аналогічна до симуляції відпалу). Значення T змінюється від максимального до 0 з певним інтервалом за формулою:

$$T_{i+1} = \lambda T_i \ (0 \leqslant \lambda \leqslant 1). \tag{8}$$

Ітерації алгоритму припиняються, коли значення T знизиться до 0, або якщо протягом певної кількості генерацій не знаходиться розв'язок кращий за рекордний.

Для ефективного розпаралелення даного алгоритму на графічному процесорі необхідно враховувати особливості роботи відеопам'яті та архітектури відеокарти загалом. Оскільки, відеокарта має невелику кількість більш швидкої колективної пам'яті (shared memory) на кожному мультипроцесорі, доцільно одночасно працювати з кількома невеликими популяціями. Таким чином, кожна популяція може бути розміщена у колективній пам'яті прив'язаній до одного з блоків мультипроцесору. При цьому кожен блок одночасно працює тільки над одним елементом популяції, внаслідок чого простіше і більш ефективно виконувати синхронізацію потоків, яка необхідна після обчислення нових кутів, перед мутацією та перед обчисленням фітнес-функції. З певним інтервалом блоки обмінюються найкращими елементами своїх популяцій через глобальну пам'ять (global memory). Кожен потік у блоці відповідає за обробку однієї амінокислоти.

Література. **1.** G. Czibula, M. Bocicor and I. Czibula, "A reinforcement learning model for solving the folding problem", International Journal of Computer Technology and Applications, vol. 2, pp. 171-182, 2011. **2.** Marlon H. Scalabrin, Rafael S. Parpinelli, Cesar M.V. Benitez and Heitor S. Lopes, "Population-based harmony search using GPU applied to protein structure prediction", International Journal of Computational Science and Engineering, vol. 9, pp. 106-118, 2014 **3.** Cesar M.V. Benitez and Heitor S. Lopes, "Molecular Dynamics for Simulating the Protein Folding Process Using the 3D AB Off-Lattice Model", in Advances in Bioinformatics and Computational Biology, 1st ed., Campo Grande, 2012, pp. 61-72.

Надточій Т.В.
Національтий технічний університет України "КПІ", факультет прикладної математики, Київ, Україна

Реалізація функції імітації присутності в системі "розумний дім" на базі Arduino Nano 3.0 з обробкою даних на хмарному сервері

В сучасному інформаційному світі системи "розумний дім", які дозволяють в автоматичному режимі керувати освітленням, кліматом, пожежними та охоронними системами, набувають активного розвитку. Вони використовують різні методи фіксації, передачі, обробки інформації та відповідного реагування системи.

Однією з функцій таких систем є забезпечення безпеки приміщення. Широкого використання набуває функція імітації присутності. Основне її завдання — імітувати присутність людини в будинку, в той час, коли там нікого немає довгий період.

Основними перевагами використання хмарних технологій [3] у системах "розумного дому" є: можливість користувачеві в будь-якому місці дізнатися обстановку в будинку, можливість більш гнучкої інтеграції нових пристроїв, можливість надання єдиного інтерфейсу управління системами "розумного дому" у випадку використання приладів різних виробників із різними протоколами передачі даних.

На сьогоднішній день мікроконтролер Arduino може в повному обсязі забезпечити користувача великим набором готових мікроконтролерних плат [1]. Використовуючи такі плати, певний набір датчиків та іншого обладнання можна достатньо просто реалізувати систему "розумний дім". Виконання таких систем на основі мікроконтролеру Arduino дозволяє значно зменшити вартість системи "розумний дім".

У даній роботі запропоновано систему на базі мікроконтролеру Arduino [2], яка фіксує окремі показники роботи "розумного дому", найбільш важливими з яких для імітації присутності є часові показники вмикання та вимикання світла. Через Ethernet або Wi-Fi модулі ці дані передаються на хмарний сервер, де на їх основі генерується сценарій вмикання та вимикання світла в "розумному домі". Він містить інформацію про номер зони освітлення, час вмикання світла в кожній зоні та час, через який світло вимкнеться. Згенерований сценарій у вигляді txt-файлу завантажується в пам'ять мікроконтролера і може приводитися у дію віддалено, через Інтернет.

Arduino Nano 3.0 – це пристрій, в основі якого лежить мікроконтролер ATmega328. Arduino Nano 3.0 має 14 цифрових входів/виходів, 6 з яких можуть використовуватись як ШІМ-виходи та 8 аналогових виходів. Живлення Arduino Nano 3.0 відбувається за допомогою підключення Mini-B USB, від нерегульованої напруги 6-20 В, або регульованої 5 В, зовнішнього джерела живлення. Мікроконтроллер ATmega328 має 32 кБ флеш-пам'яті для зберігання коду програми. Також за допомогою модуля micro SD/TF карт було розширенно об'єм пам'яті для зберігання розроблених сценаріїв. Для зв'язку мікроконтролера з локальним комп'ютером було обрано USB-UART (TTL/RS232) конвертор.

В запропонованій системі реалізовано два способи передачі інформації на хмарний сервіс: через Ethernet та через Wi-Fi. Для цього використані Ethernen модуль ENC28J60, який виконує комунікацію через LAN RJ45 Ethernet інтерфейс, та Wi-Fi модуль ESP-01, за допомогою якого можна реалізувати підключення до системи IP-камер, мережі сенсорів, гаджетів тощо.

Результатом виконаної роботи є система "розумний дім" на базі мікроконтролера Arduino Nano 3.0, яка виконує обробку інформації на хмарному сервері та має функцію імітації присутності. Реалізована система відповідає поставленій задачі, проста в реалізації та має невелику вартість порівняно з аналогічними системами.

Література. **1.** М. Э. Сопер. Практические советы и решения по созданию "Умного дома". - НТ Персс, 2007. - 432 с. **2.** Браян В. Еванс. Записник програміста Arduino. - Creative Commons, 2007. **3.** П.Л. Николаев. Архитектура интегрированной в облачную среду системы управления умным домом. - Программные продукты и системы 2015. - №2. - 110 с.

Петренко О.О.
Інститут прикладного системного аналізу НТУУ «КПІ», Київ, Україна

Пошук семантичних сервісів у репозитарії сервісів

Пошук, або відкриття (discovery), сервісів є однією з найбільш невід'ємних властивостей сервіс-орієнтованих систем і вважається однією з найважливіших задач у дослідженні веб-сервісів.. Пошук полягає в ідентифікації сервісу з безлічі описів сервісів, що відповідають вимогам запиту на пошук. Розглядаються автоматизовані методи відкриття сервісів, які можуть бути застосовані в різних бізнес-процесах, забезпечуючи при цьому приємний компроміс між виразністю опису сервісу і ефективністю його пошуку.

Вступ. Стрімке розповсюдження в усіх галузях сучасного суспільства індустрії сервісів, базованої на використанні сервіс-орієнтованої архітектури, передбачає створення найближчим часом розподілених сховищ з мільйонами сервісів, які можна було би знаходити і відкривати на основі вимог замовників прикладних додатків. Найкращим варіантом для реалізації таких сервісів є веб-сервіси завдяки їх стандартизації і прив'язку до Інтернет.

Веб-сервіси (WS). Універсальна система URL-адресів і технологія гіпертексту в поєднанні з пошуковими програмами утворили середовище, де інформація не тільки передається, а й інтелектуальне оброблюється. Під інтелектуальною обробкою мається на увазі, зокрема, автоматизоване збирання (композиція) об'єктів (наприклад, сервісів) в множини за певними ознаками, встановлення відносин між цими множинами і т.п. Веб-сервіси представляються за допомогою спеціальних мов опису сервісів (наприклад, WSDL) у формі, що робить їх доступними і для людини, і для комп'ютера. Це дозволяє організувати пошук і витяг веб-сервісів з універсального реєстру UDDI за їх синтаксичною інформацією при зіставленні ключових слів. Але все одно потрібна людина-програміст, який повинен розібратися в сервісах і зрозуміти їх зміст, або семантику, щоб вибрати потрібний сервіс з підмножини сервісів з подібними наборами ключових слів.

Семантичні веб-сервіси (SWS). Семантичні веб-сервіси усувають цей недолік, і це великий крок вперед в напрямку спрощеного доступу до сервісів через мережу. Вони мають свої мови опису (наприклад, OWL-S) і повністю незалежні від платформи реалізації, оскільки базуються на використанні онтології – свого роду словника понять предметної області і сукупності явним чином визначених припущень щодо змісту цих понять. У найпростішому випадку онтології використовуються для підвищення точності пошуку сервісу і відкриття такого сервісу, в якому згадується в точності шукане поняття, а не довільних сервісів, в текстах опису яких зустрілося задане ключове слово [1, 2].

Онтології і засоби їх побудови . Механізм онтології дозволяє формувати осмислені ієрархічні взаємозв'язки між сервісами, тобто реалізувати композицію сервісів спроможну задовольнити запит користувача, хоча в її описі може не бути деяких ключових слів з його вхідного запиту. Але для цього потрібно навчитися формувати семантичні анотації для опису веб-сервісів [3, 4].

Використання онтології для пошуку сервісів. Здатність ефективно знаходити сервіси залежить від того, як сервіси описані і як вони рекламуються, як можуть бути сформовані вимоги до сервісів і як ці вимоги перевіряються. Останнє спонукає до розроблення методів пошуку сервісів, які можуть бути застосовані в різних бізнес-процесах, забезпечуючи при цьому приємний компроміс між виразністю і ефективністю. Пошук сервісів заснований на описі сервісних компонентів, наприклад, інтерфейсів або операцій, але успішність застосування та використання веб-технологій у пошуку сервісів залежить від здатності постачальників сервісів описувати сервіси, здатності користувачів формувати запити і від порівнювача описів і вимог (matchmaker), який реалізує алгоритми, що враховують запит і знаходять найбільш підходящі

сервіси з набору сервісів компанії. Порівнювач фокусується на виборі необхідних елементів з опису сервісу, показниках подібності (similarity metrics) і комбінації отриманих значень подібності. Слід визнати, що веб-сервіси, описані з використанням мови WSDL, містять переважно синтаксичну інформацію і це затрудняє організацію автоматичного пошук веб-сервісу та виконання композиції сервісів. Тому було запропоновано використовувати також семантичну інформацію у веб-сервісах, внаслідок чого з'явилося поняття семантичних веб-сервісів (SWS) з своєю мовою опису OWL-S (Web Ontology Language for Web Services).

Методи пошуку семантичних веб-сервісів засновані на оцінках концептуальної близькості між елементами онтології сервісів (з репозитарію) і пошуковому запиті, заданому у вигляді множини триплетів [5]. Наступним кроком оцінки близькості між триплетами використовуються для визначення близькості мета-описів сервісів. Механізми пошуку можуть використовувати онтології для вибірки сервісів з синтаксично різними, але семантично однаковими словами. Онтології також можуть використовуватися для організації обміну даними та інтеграції програм.

Висновки. Після розгляду наявних методів оцінки семантичної подібності сервісів (семантичний, контексно-базований, кластерний, дедуктивний) [6] пропонується для реалізації гібридний підхід в експериментальному прототипі eHealth (сервісної медичної системи допомоги хронічно хворим пацієнтам), який передбачає пошук веб-сервісів в два етапи.

1. На першому етапі за допомогою синтаксичного порівняння сервісів звужується область пошуку шляхом відсіювання непричетних до запиту сервісів (застосуванням методу пошуку інформації ІК) або виділенням одного з кластерів, сформованих застосуванням кластерного методу.
2. На другому етапі проводиться уточнення попереднього вибору застосуванням семантичного порівняння описів відібраних сервісів із запитом, для чого складаються семантичні анотації для веб-сервісів або описи онтології для семантичних веб-сервісів.

Література. **1.** Banage T. G. S. Kumara, Incheon Paik, Koswatte R. C. Koswatte, Wuhui Chen. Improving web service clustering through post filtering to bootstrap the service discovery.-// International Journal of Services Computing (ISSN 2330-4472), Vol. 2, No. 3, July – Sept.2014 **2.** Yu Yue Du, Yong Jun Zhang and Xing Lin Zhang. A Semantic Approach of Service Clustering and Web Service Discovery.-// Information Technology Journal 12 (5): 967-974, 2013 ISSN 1812-5638, I DOI 10.3923/itj.2013.967.974 **3.** Jingyu Zhang, Xueli Yu, Peng Liu and Zhen Wang. Research on improving Performance of Semantic Search in UDDI.-// Proceedings of the WRI Global Congress on Intelligent Systems, Vol. 4, pp. 572-576, 2009. **4.** Richi Nayak and Bryan Lee. Web Service Discovery with additional Semantics and Clustering. - // Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, pp. 555-558, 2007. **5.** Крюков К.В., Панкова Л.А., Пронина В.А., Шипилина Л.Б. Меры семантической близости в онтологиях.- // Проблемы управления. – 2010. – №2. – С. 2–14. **6.** Rong, W., Liu, K. A Survey of Context Aware Web Service Discovery: From User's Perspective. - // Fifth IEEE International Symposium on Service Oriented System Engineering (2010)

***Попович Є.С.*[1], *Безносик О.Ю.*[2], *Фіногенов О.Д.*[3]**
[1] *ТОВ Глобал Лоджик, Київ, Україна;* [2] *Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна;* [3] *Факультет інформатики та обчислювальної техніки НТУУ "КПІ", Київ, Україна*

Використання розподіленої архітектури зберігання даних при вирішенні задачі скорочення розмірності великих щільних графів

Математична модель об'єкту складної фізичної природи на етапі схемотехнічного моделювання може складатися з десятків або навіть сотень тисяч рівнянь [1], а еквівалентна схема, відповідно, буде складається з десятків або навіть сотень тисяч компонентів. У роботі [2] розглянуто процес скорочення розмірності схеми з використанням Y-Δ перетворення, а у роботі [3] – паралельну версію відповідного алгоритму (рис. 2). Недоліком паралельного алгоритму є неефективність його повторного використання в разі великої розмірності скороченої схеми внаслідок її щільності.

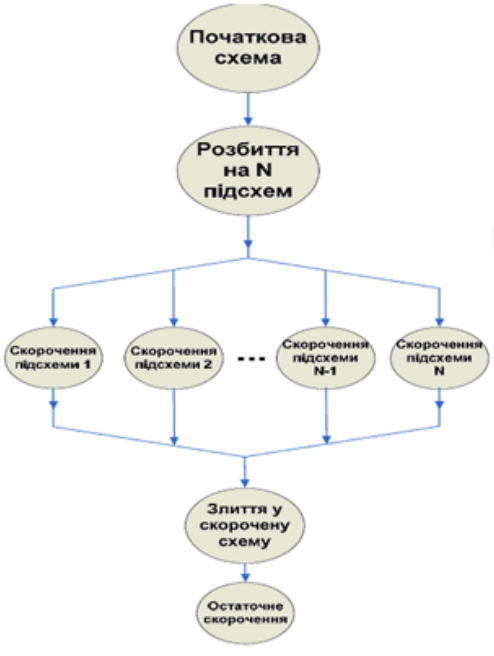


Рис. 1. Паралельний алгоритм скорочення

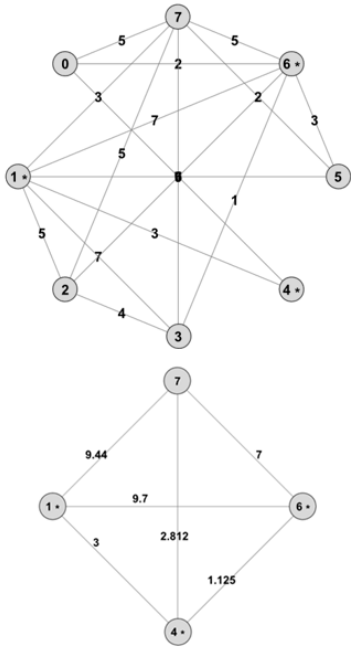


Рис. 2. Приклад початкового та скороченого графів

У свою чергу, зберігання та подальше скорочення великої за розмірами схеми на одному вузлі обчислювальної системи має ряд вад, пов'язаних з необхідним обсягом оперативної пам'яті та часом скорочення, що робить доцільним використання потужностей паралельно-розподілених систем. Абстрагуючись від деталей реалізації скорочення RLC-схем, задачу дослідження можна сформулювати як скорочення зваженого щільного графа за певними правилами, залишаючи семантику процесу скорочення RLC-схем незмінною (правила вибору елементу для скорочення, правила скорочення елементу, правила зупинки).

Вхідними даними модельної задачі є: простий неорієнтовний зв'язний зважений граф, що є умовним відображенням деякої RLC-схеми; ваги ребер графу відповідають значенням провідностей; деякі вершини графа не підлягають скороченню – вони відображають входи та виходи схеми; цільова кількість вершин скороченого графу як ознака припинення процесу скорочення. Методом скорочення для модельної задачі є метод скорочення на основі Y-Δ перетворення [2] із застосуванням деяких спрощень, зокрема, часова стала i-го вузла визначається як сума всіх провідностей, підключених до вузла i.

Вибір вершини, що буде скорочуватись, здійснюється за наступними правилами: першою скорочується вершина, що має найменшу часову сталу; якщо таких вершин декілька, то пріоритет має вершина з меншою кількістю сусідів (меншою степінню). У тому випадку, коли таких вершин декілька, пріоритет має вершина з меншим номером. Ітеративний процес скорочення продовжується до досягнення цільової кількості вершин. Результатом скорочення є граф зі зменшеною кількістю вершин (рис. 3).

Важливим є вибір структури даних, яка буде представляти граф у пам'яті ЕОМ. Цей вибір впливає на кількість використаної пам'яті для зберігання графу та ефективність застосованих алгоритмів. Для розподіленого зберігання вершини графу розподіляються порівну між наявними обчислювальними вузлами, при цьому обчислювальний вузол має повну інформацію про зв'язки тих вершин, що він зберігає. Для кожної вершини також зберігаються метадані, такі як степінь вершини, значення сталої часу та чи є вершина нескорочуваною. Степінь та значення сталої часу не перераховуються на кожній ітерації, а завжди підтримуються у актуальному стані. Такий спосіб розподілу графа має певні переваги: виконання операції підрахунку часової сталої вершини можливе без передачі даних між вузлами; при скороченні деякої вершини кожен вузол може перерахувати зв'язки тих вершин, що він зберігає, за наявності інформації про зв'язки скорочуваної вершини. Для зберігання зв'язків вершин графу (ребер) доцільно використовувати список вершин, представлений у вигляді бінарного дерева пошуку.

На відміну від послідовного алгоритму, алгоритм скорочення, що працює з розподіленою пам'яттю, включає в себе декілька додаткових етапів:

1. Вибір вершини, що буде скорочуватись:
   - кожен обчислювальний вузол вибирає вершину з найменшою часовою сталою серед вершин, що він зберігає;
   - потрібно вибрати вершину з найменшою часовою сталою серед усіх вузлів (операція оповіщення кожен-всім).
2. Для перерахунку ребер між усіма парами сусідніх вершин вузлам потрібна інформація про ребра скорочуваної вершини. Через це виконується поширення інформації про ребра скорочуваної вершини між усіма вузлами (як операція оповіщення один-до багатьох): вузол, що зберігає скорочувану вершину (кожен з вузлів на основі інформації, що поширюється на попередньому етапі, з'ясовує, чи є він таким, що містить скорочувану вершину), готує неперервний буфер (масив), в якому зберігає усі ребра скорочуваної вершини, після чого відправляє його іншим обчислювальним вузлам.

В якості інструментарію для розробки програмного забезпечення було обрано C++ у поєднанні з MPI. За результатами скорочення тестового графа за допомогою запропонованого підходу експериментально підтверджено, що з ростом кількості обчислювальних вузлів пропорційно зменшується час скорочення (табл. 1).

Табл. 1. Чисельні результати експериментів

Характеристики	Кількість вузлів		
	1	2	4
Кількість вершин графа	1000		
Щільність графу	0,6		
Час вирішення, с	52,8	27,8	14,0

Таким чином, запропонований підхід до скорочення розмірності великих щільних зважених графів може бути використано для вдосконалення алгоритму скорочення великих RLC-схем.

Література. **1.** Ладогубец В.В. Методы макромоделирования МЭМС / Ладогубец В.В., Безносик А.Ю., Крамар А.В., Финогенов А.Д. // Электроника и связь. Тематический выпуск «Проблемы электроники» ч.1. - 2008. - С. 244-248. **2.** Руденко Ю.А. Алгоритм уменьшения размерности RLC цепей / Руденко Ю.А., Ладогубец В.В., Ладогубец А.В. //Электроника и связь. – 2004. - №21. – с. 72-74. **3.** Петренко А.І. Паралельний алгоритм скорочення розмірності RLC-схем / Петренко А.І., Попов О.О., Кот Д.М. // Электроника и связь : тематический выпуск «Электроника и нанотехнологии». – 2010. – № 4. – С. 40–44.

Романов В.В., Корначевский Я.И.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Выбор параметров моделей компонентов интегральных микросхем по критерию чувствительности

Современные методы автоматизированного схемотехнического проектирования позволяют достичь высокой точности расчетов, т.е. существенно снизить алгоритмическую погрешность вычислений. Основной составляющей погрешности при этом остается погрешность, связанная с неточностью определения численных значений параметров моделей используемых компонентов схемы: диодов, транзисторов, интегральных операционных усилителей, логических элементов и т.п.

Стремление повысить точность проектирования электронных схем приводит к разработке оптимизационных процедур определения параметров модели. Такой подход позволяет снизить погрешность определения параметров до нескольких процентов, но при этом, естественно, повышаются затраты машинного времени, и процедура расчета параметров существенно усложняется.

Поэтому представляется целесообразным проанализировать связь между погрешностями измерения параметров моделей и погрешностью вычисления выходных характеристик схемы, и сформулировать требования к точности определения отдельных параметров используемых моделей компонентов электронных схем. Очевидно эти требования будут различными для разных типов моделируемых схем: переключательных, усилительных, генераторных и т.п. [3,4]

Относительную погрешность моделирования выходной характеристики схемы $y = f(q_1, \ldots, q_n)$ при известных относительных погрешностях измерения параметров моделей компонентов $\frac{\Delta q_i}{q_i}$ можно определить в первом приближении с помощью выражения:

$$\frac{\Delta y}{y} = \sum_{i=1}^{n} S_{q_i}^{y} \cdot \frac{\Delta q_i}{q_i} \tag{1}$$

где $S_{q_i}^{y} = \frac{q_i}{y} \cdot \frac{\delta y}{\delta q_i}$ – чувствительность выходной характеристики по отношению к параметру модели q_i.

В общем случае из известных методов, позволяющих определить чувствительность выходной характеристики физического объекта $S_{x_i}^{y}$ по отношению к параметру x_i наибольшее распространение получили расчетно-аналитический метод и экспериментально-статистические методы: метод малых приращений, корреляционный метод, метод планирования многофакторного эксперимента и регрессионного анализа. Расчетно-аналитический метод требует обязательного получения аналитического выражения выходного параметра как функции всех условных параметров объекта, что не всегда представляется возможным. В таких случаях используется часто экспериментальные методы, из которых наибольшими возможностями обладает метод статистического планирования многофакторного эксперимента, который в сочетании с регрессионным анализом опытных данных позволяет поставить и решить следующие основные задачи:

1. В значительной степени исключить интуитивный подход, заменить его научно обоснованной программой проведения эксперементального исследования, включающей объективную оценку результатов эксперимента, на всех последовательных этапах исследования.
2. Оптимизировать процесс исследования, т.е. найти совокупность варьируемых факторов q, при которых выбранная целевая функция принимает экстремальное значение. При этом возможно осуществление минимального числа опытов.
3. Путем направленного эксперимента можно определить наиболее "влияющие" факторы на выходные характеристики рассматриваемого объекта, скорректировать и уточнить теоретические представления об изучаемом явлении.

Используемый при решении выше перечисленных задач математический аппарат является

достаточно общим, что позволяет применять его как при исследовании реальных объектов, приборов, так и при исследовании их физических и математических моделей. Как показал предыдущий опыт разработчиков САПР, основная погрешность при математическом моделировании электронных схем обусловлена не методом вычислений, а, в основном, неточностью используемых моделей компонентов.

Ниже излагается методика определения чувствительностей выходных характеристик схемы к статическим и динамическим параметрам моделей компонентов, основанная на регрессионном анализе. Эта методика инварианта как к типу схемы и ее выходным характеристикам, так и к типам используемых компонентов и их моделей [1, 2].

Здесь, как и в общем случае, в основу построения плана проведения исследований может быть положена матрица полного факторного эксперимента и регулярных дробных реплик от него. Естественно, поэтому к создаваемой матрице предъявляются известные общий требования:

1. Симметричность – алгебраическая сумма элементов вектор-столбца каждого фактора должна быть равна нулю, т.е.

   $\sum_{N}^{i=1} x_{ji} = 0$, где N – число строк в матрице планирования, j – номер столбца.

2. Нормировка – сумма квадратов элементов каждого вектор-столбца должна быть равна числу опытов, а именно:

   $$\sum_{N}^{i=1} x_{ji}^2 = N.$$

3. Ортогональность – сумма произведений любых двух вектор-столбцов матрицы должно быть равна нулю, т.е.

   $$\sum_{N}^{i=1} x_{ji} \cdot x_{ui} = 0, \text{ при } j \neq u.$$

С помощью метода планирования многофакторного эксперимента вида «2^n» (n – число используемых факторов) по результатам вычислений выходной величины y при одновременном варьировании факторов по определенному плану определяем коэффициенты линейного уравнения регрессии:

$$y \approx b_0 + \sum_{n}^{i=1} b_i \cdot \tilde{q}_i \tag{2}$$

$$\text{где } \tilde{q}_i = \frac{q_i - q_{i0}}{\Delta q_i} \text{ – кодированное значение фактора} \tag{3}$$

q_{i0} – натуральное значение основного уровня фактора, выбор которого осуществляется на основе анализа априорной статистической измерительной информации; Δq_i – интервал изменения фактора; $q_i = q_{i0} \pm \Delta q_i$ – натуральное значение фактора.

При этих условиях параметр $\tilde{q}_i$ принимает значения $\pm I$, т.е. значения верхнего и нижнего уровней, формируемых относительно основного-центра эксперимента. Подстановка (3) в уравнение (2) дает:

$$y \approx b_0 + b_1 \cdot \frac{q_1}{\Delta q_1} + \cdots + b_i \cdot \frac{q_i}{\Delta q_i} + \cdots + b_n \cdot \frac{q_n}{\Delta q_n} - b_1 \cdot \frac{q_{10}}{\Delta q_1} - \cdots - b_i \cdot \frac{q_{i0}}{\Delta q_i} - \cdots - b_n \cdot \frac{q_{n0}}{\Delta q_n} \tag{4}$$

С помощью соотношения (1) можно получить выражение для определения искомых чувствительностей:

$$S_{q_i}^{y} = \frac{q_{i0}}{y} \cdot \frac{\delta y}{\delta q_i} = \frac{q_i 0}{b_0} \cdot \frac{b_i}{\Delta q_i}, \quad i = 1, 2, \ldots, n \tag{5}$$

Коэффициент b_0 представляет собой средне арифметическое значение выходной величины y т.е. значение y, вычисленное при основных уровнях варьируемых параметров.

Учитывая, что все вычисления производятся на математической модели схемы, отметим

следующие особенности применения метода планирования эксперимента для рассматриваемого случая:

- объект эксперимента (модель схемы) является управляемым, т.к. варьируемыми могут быть любые параметры модели;
- дисперсия воспроизводимости результатов равна нулю;
- факторы являются абсолютно управляемыми и независимыми, т.е. имеется принципиальная возможность установления любого уровня фактора вне зависимости от уровней других факторов.

Эти свойства обеспечивают удобство и эффективность использования метода планирования эксперимента для решения поставленной задачи.

Эксперименты, проведенные с помощью описанной выше методологии анализа выходных характеристик схемы инвертора и ряда других простых цифровых схем с использованием модели МОП транзистора BSIM3v3 подтвердили выводы, сделанные в работе [2,3].

При работе в субмикронном диапазоне выходные характеристики аналоговых схем оказались наиболее чувствительными к параметрам, связанных с короткоканальными эффектами. Для моделирования цифровых схем или схем, в которых короткоканальные эффекты проявляются в меньшей мере, можно существенно сократить число определяемых физических параметров, использовав при этом для остальных параметров стандартные значения [3,4]. При этом, в качестве выходных характеристик схем были приняты параметры: средняя задержка распространения сигнала ($t_{\text{з.ср}}$) и длительность выходного импульса ($t_{\text{имп}}$), которые являются качественными показателями функционирования переключательных схем. Средняя задержка распространения сигнала определялась как $t_{\text{з.ср}} = \dfrac{t_{\text{з.п}} + t_{\text{з.з}}}{2}$, где $t_{\text{з.п}}$ и $t_{\text{з.з}}$ времена задержек переднего и заднего фронтов выходного напряжения относительно соответственно переднего и заднего фронтов входного напряжения.

Интервал варьирования значений параметров модели МОП транзистора BSIMv3 (верхний уровень-нижний уровень) выбирался в пределах $\pm 1\%$ с целью получения более достоверных оценок коэффициентов регрессии. Расчет времени задержки и длительности выходного импульса производился с помощью комплекса программ ALLTED , разработанного на кафедре СП Киевского политехнического института.

Изложенная методика позволяет выявить наиболее существенные как статические, так и динамические параметры моделей с точки зрения их влияния на выходные характеристики анализируемого компонента или анализируемой схемы в целом, обеспечить получение количественных оценок такого влияния. Это, в свою очередь, приводит к повышению эффективности трудоемкого процесса информационного обеспечения систем схемотехнического проектирования, особенно в том случае, когда заранее требуется предсказать результат эксперимента в некоторой области с требуемой точностью.

Данная методика может быть применима также и для выбора наиболее эффективных способов учета различных физических эффектов в моделях активных компонентов электронных схем.

Литература. **1.** *Адлер Ю.П.* Планирование эксперимента при поиске оптимальных условий / Адлер Ю.П., Маркова Е.В., Грановский Ю.В. // М.:Издательство “Наука”. – 1976. – 279 с. **2.** *Романов В.В.* Сравнение инжекционной модели и модели переноса биполярного транзистора по критерию чувствительности /Бобин В.В., Романов В.В.// Известия Вузов. – Радиоэлектроника. – 1991. – №11 **3.** *Корначевський Я.І.* Порівняльне дослідження модифікацій моделей BSIM 3v3, вбудованих в пакети схемотехнічного проектування ALLTED та SPICE / Корначевський Я.І., Ладогубець В.В. // Электроника и связь. – 2004. – № 21. – С. 30–33. **4.** *Корначевський Я.І.* Особливості реалізації інверсного режиму в КМОН моделях на прикладі BSIM / Я.І. Корначевський, В.В. Ладогубець // Вісник національного університету “Львівська політехніка”. – 2003.

Свістунов С.Я., Юрченко В.В.
Інститут теоретичної фізики ім. М.М.Боголюбова НАН України, Київ, Україна

Досвід впровадження технологій хмарних обчислень в Інституті теоретичної фізики ім. М.М.Боголюбова НАН України для обробки даних по експерименту ALICE в CERN

Хмарні технології на даний час займають ключову роль у розвитку інформаційних інфраструктур для забезпечення наукових дослідженнях по всьому світу. В останні роки в CERN проводилися дослідні роботи з адаптації існуючої обчислювальної моделі обробки даних на LHC для використання в хмарній інфраструктурі. Ці роботи стосувалися як класичного проміжного програмного забезпечення gLite, так і прикладного програмного забезпечення, яке використовується в окремих експериментах для обробки даних. CERN інтенсивно займається розвитком хмарних технологій для обробки даних у фізиці високих енергій, створивши масштабну розподілену хмарну платформу для обробки первинної інформації, що побудована на базі програмного забезпечення OpenStack, на основі двох великих дата центрів, які знаходяться у м.Женева (Швейцарія) та м.Будапешт (Угорщина).

Стратегія використання хмарних обчислень полягає в тому, щоб перейти від класичних грід обчислювальних елементів (CE для gLite), до розгортання програмного забезпечення управління хмарою з відкритим вихідним кодом, щоб управляти обчислювальними вузлами на грід сайтах і надати можливість використовувати хмарні API. Це дає можливість забезпечити взаємодію між грід сайтами або між хмарними інфраструктурами і забезпечить однаковий підхід до підтримки програмного забезпечення. Така стратегія дозволяє ввести просту і уніфіковану систему забезпечення безпеки, обліку використання ресурсів та сервісів ідентифікації. Крім того, використання файлової інфраструктури CernVM FS (CernVM File System) дозволяє забезпечити загальний підхід до інсталяції та супроводу стандартного програмного забезпечення, скорочуючи кількість різних сервісів і забезпечуючи загальний, стійкий механізм підтримки обчислювальної моделі експериментів на LHC.

Коллаборація ALICE прагне повністю використовувати хмарні технології, щоб розширити свою розподілену обчислювальну інфраструктуру і використовувати, де це можливо, хмарні ресурси і програмні продукти підтримки експериментів, такі як CernVM FS і подібні йому інструменти.

В докладі представлені результати роботи по побудові в Інституті теоретичної фізики ім. М.М.Боголюбова НАН України (ІТФ) високонадійної хмарної інфраструктури з використанням програмного забезпечення OpenStack для участі в обробці даних по експерименту ALICE (Рис. 1).

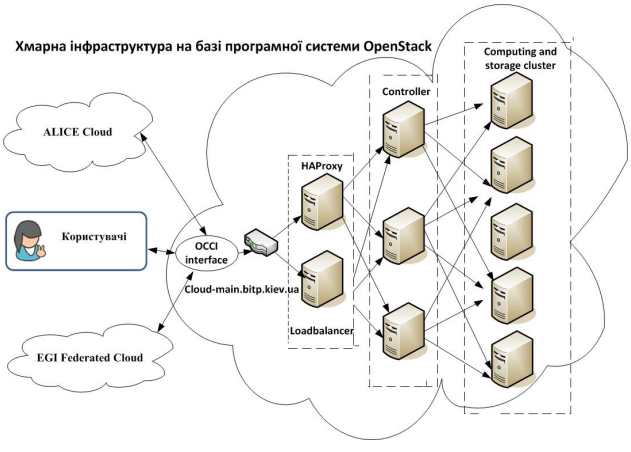


Рис. 1. "Активно-активна" схема хмарної інфраструктури ІТФ

Інсталяція виконувалась на 10 серверах з такими характеристиками: процесор Intel(R) Xeon(R) CPU E5335 2.00GHz (8 ядер); ОЗУ 8 ГБ, об'єм жорсткого диску 450 ГБ; операційна система CentOS Linux 7 (Core), архітектура x86_64. У результаті аналізу схем побудови високонадійних хмарних структур була обрана так звана "активно-активна" модель конфігурації в складі: двох мережевих серверів з системою балансування навантаження HAProxy; трьох контролерів, які забезпечують авторизацію та завантаження образів віртуальних серверів; п'ятьох обчислювальних серверів,

на яких відбувається запуск віртуальних машин, з встановленою розподіленою файловою системою Ceph. Таке рішення було прийняте, оскільки ця схема пропонує більше механізмів для забезпечення доступності та надійності сервісів у порівнянні з "активно-пасивною". Саме при цій конфігурації можливо використовувати систему балансування навантаження, що має додаткові переваги у вигляді зменшення навантаження на контролюючий та мережевий сервери. Для автоматизації процесу інсталяції OpenStack використовувалося програмне забезпечення Packstack. Для швидкої автоматичної установки та налаштування файлової системи Ceph був розроблено додатковий модуль для утиліти Packstack на мові Puppet. Цей модуль був інтегрований в структуру пакету openstack-packstack.

В докладі представлені результати роботи по інтеграції хмарної інфраструктури ІТФ з EGI Federated Cloud – об'єднанням хмарних та віртуалізованих ресурсів, яке побудоване на відкритих стандартах і сфокусоване на потребах наукових спільнот та результати роботи по забезпеченню сумісності з хмарною інфраструктурою CERN. Модель EGI Federated Cloud об'єднує хмарні ресурси шляхом визначення набору стандартних відкритих інтерфейсів та протоколів доступу до різних хмарних функцій – таких як публікація даних про ресурси, аутентифікація користувачів, обчислювальні сервіси та сервіси доступу до даних.

Така інтеграція різних хмарних інфраструктур виконана за допомогою стандартного інтерфейсу OCCI – протоколу і API, який поєднує в собі набір специфікацій, розроблених організацією Open Grid Forum. OCCI призначений для задач управління хмарними ресурсами різних провайдерів у різнорідних середовищах. Для підтримки інтерфейсу в хмарній інфраструктурі OpenStack був додатково встановлено модуль Keystone-VOMS.

Однак, сертифікація хмарної інфраструктури ІТФ та забезпечення виконання всіх тестів системи моніторингу EGI Federated Cloud (https://cloudmon.egi.eu/nagios/) є тільки першим обов'язковим кроком на шляху інтеграції до хмарної інфраструктури обробки даних для експерименту ALICE.

У 2015 році фахівці коллаборації ALICE почали тестування програмного пакету Vcycle для забезпечення обробки даних в хмарних інфраструктурах. За даними сайту http://ganglia-004.cern.ch/ п'ять європейських хмарних центрів вже задіяні в обробці даних по експерименту ALICE. Vcycle – це менеджер запуску віртуальних машин, який реалізує модель запуску задач з центральної черги ALICE на хмарні обчислювальні елементи. За архітектурним рішенням Vcycle отримує образ віртуальної машини з центрального репозіторію, і при інсталяції образу на хмарному ресурсі виконується контекстуалізація за допомогою скриптів, які розроблені спеціально для забезпечення обчислень з обробки даних в експериментах CERN. Для перевірки роботи програмного пакету Vcycle та дослідження реалізованого алгоритму взаємодії з хмарною інфраструктурою OpenStack в ІТФ було виконано інсталяцію програмного забезпечення Vcycle на окремому сервері (vcycle.bitp.kiev.ua). Проведені дослідження дозволили налаштувати хмарну інфраструктуру ІТФ для підключення до хмарної інфраструктури CERN.

Після тестування інтерфейсу взаємодії адміністраторами проекту ALICE було запущено 6 віртуальних машин на основі операційної системи µCernVM. Загальна кількість ресурсів, що використовуються: 12 віртуальних ядер; 24 ГБ оперативної пам'яті; 30 ГБ дискового простору. Віртуальні машини були запущені по протоколу OCCI за допомогою програмного забезпечення Vcycle, яке встановлене на стороні центральних сервісів ALICE. Для відображення кількості виконаних завдань в загальної хмарної інфраструктурі експерименту ALICE в системі моніторингу http://alimonitor.cern.ch/display було сконфігуровано фіктивний хмарний кластер Altaria.

На грудень 2015 року хмарна інфраструктура ІТФ є єдиним хмарним ресурсом в Україні, який відповідає всім вимогам EGI до хмарних інфраструктур і входить в EGI Federal Cloud. Також можна констатувати, що хмарна інфраструктура ІТФ є єдиним хмарним ресурсом в Україні, яка надає ресурси для виконання обчислювальних завдань по експерименту ALICE.

Тесленко О.К., Шаповал І.І.
Факультет прикладної математики НТУУ "КПІ", Київ, Україна

Нероздільні блочні коригуючі коди

ОГЛЯД ПРОБЛЕМИ. Зі зростанням застосування цифрових телекомунікаційних технологій все більш актуальними стають дослідження методів підвищення ефективності та надійності передачі цифрових даних. До них належать методи завадостійкого кодування, засновані на штучному введенні надлишковості.

Завадостійкий код відрізняється від звичайного тим, що у канал передаються не всі кодові комбінації N, котрі можна сформувати із наявного числа розрядів n, а тільки їх частина N_r, котра складає підмножину дозволених комбінацій. Це дозволяє на стороні адресата не тільки виявляти спотворення переданих даних, а і формувати правильні значення. Загальновідомо, що у випадку блочних n- розрядних кодів для виправлення t спотворень необхідно і достатньо, щоб $d >= 2t + 1$, де d – найменша відстань Хеммінга між будь якими кодами із множини N_r.

Найбільшого розвитку набула теорія та практика побудови блочних роздільних кодів, де існує чіткий розподіл між інформаційними та додатковими бітами. Прикладами таких кодів є коди Хеммінга, Боуза-Чоудхурі-Хоквінгема та ін. Перевагою таких кодів є аналітичне визначення дозволених комбінацій та порівняно не складна реалізація кодерів та декодерів. В той же час швидкість роздільних кодів може бути максимальною для достатньо обмежених значень n. Можна припустити, що для інших значень n швидкість можна збільшити при використанні нероздільних кодів.

У нероздільних блочних кодах надлишковість вводиться шляхом знаходження множини дозволених кодів, в якій між кожною парою кодових слів відстань за Хеммінгом не менша, ніж d. Для досягнення переваги при застосуванні нероздільних блочних кодів перед іншими альтернативами, необхідно визначити спосіб визначення множини кодів заданої розрядності зі швидкістю таких кодів, наближеною до максимальної.

У загальному випадку, досягнення вищезазначеної мети є нетривіальним завданням через високу складність обчислення всіх можливих кодових комбінацій заданої розрядності.

ОСНОВНА ЧАСТИНА. Для вирішення проблеми пошуку максимального коду нерозподільних блочних коригуючих кодів сформулюємо проблему у контексті теорії графів.

Елементи множини із $N = 2^n$ всіх двійкових кодів можна представити у вигляді вершин графа Хеммінга, де дві вершини суміжні, якщо відстань Хеммінга між відповідними кодами не менша, ніж d.

Очевидно, що максимальний код є максимальною клікою графа Хеммінга.

Відповідно до [1] задача пошуку максимальної кліки графа є NP-повною, а потенційно точне рішення цієї задачі повним перебором всіх можливих комбінацій має обчислювальну складність $O(k^2\binom{N}{k})$, де k-розмір кліки.

Необхідно зазначити, що, у разі існування точного рішення задачі пошуку максимального блочного коригуючого коду, його можна вважати рішенням задачі пошуку кліки графа у загальному випадку.

ЖАДІБНИЙ АЛГОРИТМ. Можливим вирішенням описаної проблеми є формулювання алгоритмів, заснованих на евристичних підходах. Одним з таких найбільш широко застосовуваних підходів є жадібні евристики.

Загалом, застосування концептуально простих підходів до формування жадібних евристик для пошуку клік графів показало свою ефективність при застосуванні до прикладних завдань, що виникають у процесі аналізу зв'язків у комп'ютерних мережах та «соціальних» зв'язків у соціальних мережах. Проте, такі підходи, як витікає з аналізу очікуваної обчислювальної складності, є прийнятними при застосуванні до графів малої та середньої щільності. Графи Хемінга ж при відповідних значеннях d можуть мати досить велику щільність

Виходячи з вищезазначеного, очевидною є необхідність сформулювати модифікацію існуючих підходів з метою вирішення завдання пошуку максимального коду нероздільних кори-

гуючих кодів з прийнятною обчислювальною складністю та з очікуваним результатом, що відповідає верхній межі аналітичної оцінки та критеріям практичної цінності.

Проблема невідповідності результатів на щільних графах, отриманих при застосуванні існуючих алгоритмів, заснованих на жадібних евристиках, полягає в тому, що базова жадібна евристика не враховує отримані раніше результати та не забезпечує близький до рівномірного розділ отриманих комбінацій вершин, що формують поточні кліки. При застосуванні такого підходу можливе отримання однакових або еквівалентних результатів для досить довгої послідовності виконань.

Для вирішення вищезазначених проблем та додаткової оптимізації жадібної стратегії модифікуємо її із застосуванням елементів стратегії «табу» - пошуку.

В якості елементів «табу» - пошуку введемо механізм запам'ятовування знайдених локальних рішень для відкидання їх елементів з множини пошуку на динамічно змінювану кількість ітерацій. Число таких ітерацій залежить від частоти занесення вершин графа до локальних рішень.

РЕЗУЛЬТАТИ. Експериментальні результати виконання імплементації зазначеної модифікації жадібного алгоритму довело перевагу такого підходу над можливими модифікаціями алгоритмів Програму було виконано для таких розрядностей та значень відстані за Хеммінгом:

n 4 6 8

d 3 3 3

Випадки 1) та 2) практичної цінності не мають, проте є важливими для виявлення закономірностей оцінки швидкодії імплементацій порівнюваних алгоритмів.

Отже, у випадках 1) та 2) імплементації алгоритмів виконали пошук клік за такий час:

Алгоритм Брона-Кербоша [2]: 4мс 1560мс 4 години 37 хвилин

Модифікований жадібний : 5мс 25мс 166мс

Для розрядності $n = 13$, відстані $d = 4$ із застосуванням модифікованого жадібного алгоритму було знайдено кліку розміром 256 елементів за 56 хвилин, тоді як кілька модифікацій алгоритму Брона-Кербоша не завершили виконання протягом 48-ми годин.

Для кодів розрядності $n = 8$ знайдено множину дозволених кодів, число яких 20, що більше, ніж 16 кодів для роздільних кодів Хеммінга такої ж розрядності.

Відносна швидкість коду $R = k/n$. Для нерозподільних кодів $R = log_2 N_k/n$, де N_k – кількість дозволених кодових слів. Порівняємо відповідні значення швидкості коду для роздільних та нерозподільних кодів у випадку кодів Хеммінга $R = 4/8 = 0.5$; у випадку нероздільних кодів $R = \log 220/8 = 0.54024$.

ВИСНОВКИ. Еквівалентність задачі пошуку максимальної кліки графа до задачі пошуку максимального коду нероздільних завадостійких блочних кодів теоретично обґрунтовує NP-повноту задачі нероздільного блочного кодування. Це вимагає розвитку евристик для знаходження максимальних нероздільних блочних кодів у загальному випадку та оригінальних вирішень при конкретних значеннях n та d.

Застосування запропонованої модифікації жадібного алгоритму із елементами табу-пошуку дозволяє суттєво (в десятки раз) підвищити швидкість знаходження максимальної кліки (максимального коду) в порівнянні з відомими, що підтверджено експериментально.

Нерозподільні коди при n=8, d=3 мають перевагу перед розподільними за показником кодової швидкості на ((0.54024 - 0.5)/0.5*100) = 8.048 %, що обґрунтовує перспективність подальших досліджень.

Література. 1. Томас Кормен. Алгоритмы: построение и анализ / Чарльз Лейзерсон, Рональд Ривест, Клиффорд Штайн, 2-е издание — М.: Издательский дом «Вильямс», 2009. — с. 442-481. **2.** E.A. Akkoyunlu, The enumeration of maximal cliques of large graphs, SIAM J. Comput., Vol. 2: 1-6, 1973.

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. Highly productive operating systems and networks, telecommunication technologies.
5. 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. Высокопродуктивные ОС и сети, телекоммуникационные технологии.
5. 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. Високопродуктивні ОС і мережі, телекомунікаційні технології.
5. Бази даних і знань як середовище інформаційної підтримки управління та проектування.

Ademola Samuel Adegoke
Taras Shevchenko National University of Kyiv "KNU", Kyiv, Ukraine

Network locality awareness comparison and method of improvement of existing models

This paper shows a comparison between current distance estimation techinques, how they differ from each other and proposes a more accurate means of locality awareness for network nodes.

Overlay networks have formed the basis of the most advanced network architectures of the recent times. They are essentially virtual networks built on the top of other networks such that they provide the services that are otherwise not present on the base networks [1]. The most successful overlay network to date presumptuously would be the Internet, which was initially built as an overlay on top of telephone networks [2], it has since become the backbone of information dissemination in the world. Over time, newer overlays have been built on top of the Internet and some of the most popular ones include the Voice Over IP service (VoIP), the peer-to-peer (P2P) networks amongst others. For a node to be said to be connected to the Internet, it should be able to communicate with another node on the Internet of which they have no physical links. The means and speed of communication as well as the physical resources required are very essential parts of the equation. The Internet is a complex structure of interconnected backbone networks and each of these network is down-divided into autonomous systems that are often defined by the IP address ranges assigned by the Internet Assigned Number Authority (IANA) and subsequently the Regional Internet Registry (RIR). Within a defined locality, depending on the size, there may be more than one Internet Service Provider (ISP) thus requiring a link between these ISPs for communication. Initially, ISPs formed individual connections among themselves, but this led to situations where a traffic from nodes physically closes to each other would have to travel long distances to get to their destination because the connections these ISPs formed were unstructured. Also, it had effects on the cost of Internet use for end-users as sometimes there had to be an unreliable third-party application involved in the forwarding of packets to the appropriate destinations. To help solve this problem, ISPs created Internet Exchange points (IXPs) [4][5] which essentially is a physical architecture for connecting Autonomous Systems, they aim at keeping local Internet traffic within local infrastructure. The consequence of this is that traffic within a single IXP arrives faster at the destination and it costs lesser for the end-user. Technically speaking, traffic passing through an IXP has a lower hop count because of the routing protocol configured on the IXP. It is assumed that the BGP (Border Gateway Protocol) is the main routing protocol running on a typical IXP architecture, which is a protocol that calculates multiple paths to a single destination and makes routing decisions based on the result.

Locally, applications that are network-based often require full knowledge of nodes on the network and the distance between them. This is typically managed by a central system that collects and computes these data then uses the resulting data for routing decisions and packet forwarding [4]. This however creates a problem of "single point failure" where if for some reason the central system where to be down, the whole system is affected. In a situation where there is no central system to use to get this information, nodes would have to resort to the traditional "ping" or "trace route" for hop count to get the metric required. The problem associated with this is that if nodes need this kind of information for a node that is not within their autonomous system or a node outside their IXP - a foreign node, they would have to explicitly request the information, which as a result causes delay and even sometimes dropped packets due to timeouts.

COMMON DISTANCE ESTIMATION TECHNIQUES. 1. PING (Packet InterNet Groper) This is a diagnostic tool that is used to test the accessibility of a node on the network (inter/intranet) as well as giving information on the round-trip time of packet from one node to another. It works by sending ICMP (Internet Control Message Protocol) echo to a target node and measures the

delivery time of the packets and also return time of the ICMP echo reply.

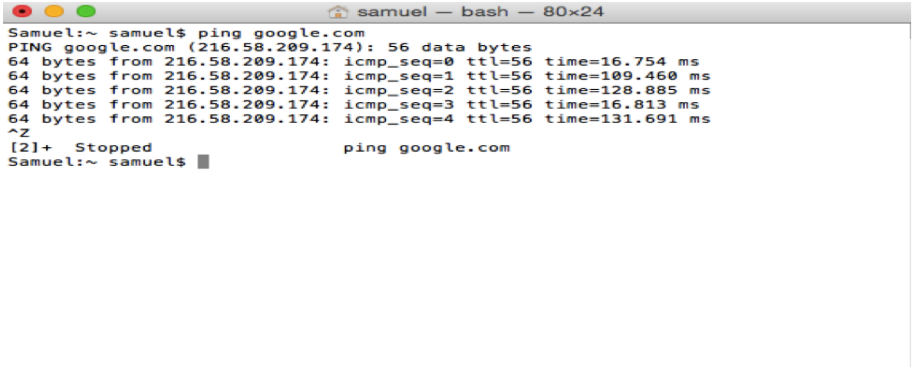


Figure 1. PING echo reply report.

The disadvantage of this is that some AS servers block ICMP packets from passing through because of the risks posed by malicious network users. A network can be exploited by simply flooding it with PING requests(DDoS). As a result of this, network servers use packet shapers [3][4], which limits the amount of ICMP/UDP packet allowed on the network. Some industrial systems also totally disable these packets on the network. All these are consequential on the result a PING request obtains, this means an "Unavailable host" report from PING may not necessarily be correct as the host may just be hidden behind a protected network, also considering that these hosts are not easily reachable due to security measures put in place, a packet that is allowed to reach a remote node may have an inaccurate round-trip time because of this. Hence, PING is not an accurate measuring scale of the proximity of nodes on a given network.

2. TRACERT. TRACECERT like PING uses ICMP echo requests; this is used to trace the path a packet takes from a given node to another. The problem with it is that it uses different protocols on different operating system environments. TRACECERT on windows uses ICMP while on UNIX based systems; it uses UDP thereby giving different results in the two environments. Considering that it uses the much dreaded ICMP/UDP packets, the results this protocols get is also subject to the security scrutiny employed on the network hence a need for a more accurate means of obtaining an arbitrary node information.

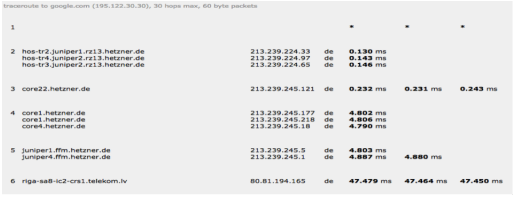


Figure 2. Traceroute on Windows.

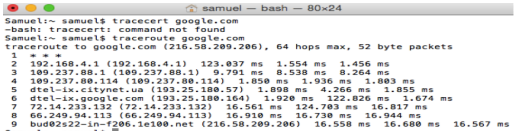


Figure 3. Tracecert on Unix.

3. CARMA (Combined Affinity Reconnaissance Metric Affinity). 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 sub ranges 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 structure from the publicly available FTP server of RIR. 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 [1][2][3].

Figure 4. RIR database file

Some of the advantages it has over the traditional method of distance estimation is that it doesn't have to probe traffic because it is a model that computes it values arithmetically locally. Basically, once the model is built, the rest of the process can be done offline, it can calculate the distance between arbitrary IPs by itself thereby saving data and bandwidth consumption. Also, it takes away the need to run estimation on a node whose locality is being estimated, one of it's highest points is that can for any arbitrary IP addresses. While all these are the obvious advantages it has, CARMA requires that the RIR database be downloaded whose size could be between 10 to 100mb and this may have to be done regularly so as to keep up with the changes that may have be made to the database. But since the RIR database is hardly updated due to the near exhaustion of the IPv4 then it may not be a big deal especially to nodes that don't need to do distance estimation often. The second disadvantage CARMA has is pertaining to irregular information supplied by RIRs. Not all RIRs provide detailed information about their ASes but for CARMA to work efficiently, it is needed that the data from RIR contains detailed, down-to-subrange and linkage information of ASes so this is an obvious limitation of efficiency. CARMA, though has been tested on a P2P overlay network and found to have a minimum 2

Considering the successful application of CARMA in P2P system, a probable way of improving network topology survey using CARMA model would be data extraction from P2P packets. As known, P2P systems transmit data using UDP/TCP/IP. The header of a TCP/IP packet contains a "reserved" block, which could be used to store subnet information of the node. Consequently, this could be another option for the improvement of the CARMA model, if CARMA were to optionally update it's database from CARMA-aware P2P systems that utilize the "reserved" block in the packet header alongside it's RIR update. This would ensure that CARMA has a larger database to get its data from. 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 NetworksVolume 2014 (2014), Article ID 427496, 2. https://www.euro-ix.net/documents/894-ixp-research-pdf? 3. 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 4. http://www.internetsociety.org/what-we-do/issues/internet-exchange-points-ixps

Bakun S.A., Lytvyniuk A.A., Terentiev O.M.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Using SAS Enterprise Miner to build scoring card to evaluation solvency of individuals

Introduction. Today, many banks faced with the problem of default issued by companies or individuals loans. Therefore, applying methods of evaluation solvency of individuals and risk of banks in lending in modern crisis of the financial sector by banking institutions is really necessary.

Incoming data. The database consists of two data sets. The first data set contains information about 5837 clients who were issued loans, and who has already ended period of loans, the second data set contain 4233 records of applicants who have been denied loan. Each entry in the two data sets contains 18 attributes that characterize the customers, moreover the first data set contains additional indicator (target variable) which reflects the result of the return credit of customer.

Problem. Need to prepare the sample data to build the scoring card, select the most important characteristics of the clients to be included in the model, develop a scoring card and evaluate the quality of developed model.

The method of solution. SAS Enterprise Miner includes specialized components for solving the problem of scoring. To build a scoring model that is shown in fig. 1 was applied components such as:

1. Data Partition - allows you to split our sample into training and testing;
2. Interactive Grouping - provides a selection of the most important characteristics and forms groups of values for the continuous input variables;
3. Reject Inference - allows to add information on applicants who has been denied loan with automatic markup customers in positive and negative to a training sample data;
4. Scorecard - automatically builds scoring card on the results of the regression model, based on data from the training set.

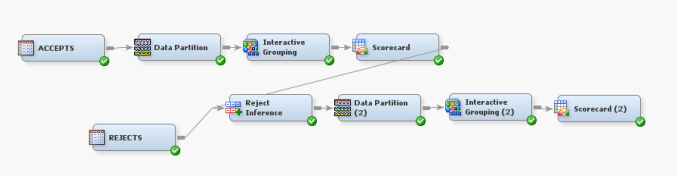


Figure 1. Scoring model built by SAS Enterprise Miner

Results. The results of system operation are displayed as built scoring card (fig. 2), which consists of parameters (attributes), ranges of values of each attribute and scoring mark for each of the ranges. To evaluate solvency of new applicant enough summarize marks for each indicator of scoring card. It also provides a number of reports with statistical indicators of quality built scoring card (fig. 3).

Conclusions. The risk of credit operations is one of the most important risks that may effect on the functioning of financial institutions. Credit operations are the most profitable type of bank's asset, but at the same time it is the most risky. Implementation of the scoring cards allow banking institutions to examine more comprehensive not only the financial capacity of the solvency of client, but also take into account other socio-demographic characteristics such as age, sex, education etc., which can negatively affect for ability of the client repaying the loan. Applying scoring evaluation will also reduce the cost and time that bank has spent for processing statements and deciding on the loan.

Scorecard

		Scorecard Points
age_oldest_tr	age_oldest_tr< 38	12
	38<= age_oldest_tr< 77	15
	77<= age_oldest_tr< 186	18
	186<= age_oldest_tr	25
	MISSING	12
bureau_score	_MISSING_	3
	bureau_score< 662	1
	662<= bureau_score< 717	22
	717<= bureau_score< 759	41
	759<= bureau_score	63

Figure 2. The fragment of scoring card built by SAS Enterprise Miner

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
bad		_AIC_	Информационный крите...	6579.409	.	
bad		_ASE_	Средний квадрат ошибки	0.141709	.	0.144467
bad		_AVERR_	Функция средней ошибки	0.438329	.	0.446885
bad		_DFE_	Число степеней свободы ...	7478.856	.	
bad		_DFM_	Число степеней свободы ...	8	.	
bad		_DFT_	Общее число степеней св...	7486.856	.	
bad		_DIV_	Делитель для асимптоти...	14973.71	.	1703.431
bad		_ERR_	Функция ошибки	6563.409	.	761.2373
bad		_FPE_	Итоговая ошибка прогноза	0.142012	.	
bad		_MAX_	Максимальная абс. ошиб...	0.9936	.	0.992683
bad		_MSE_	Средний квадрат ошибки	0.141861	.	0.144467
bad		_NOBS_	Сумма частот	7486.856	.	851.7156
bad		_NW_	Кол-во оценочных весов	8	.	
bad		_RASE_	Квадратный корень из су...	0.376443	.	0.380088
bad		_RFPE_	Корень из итоговой ошиб...	0.376845	.	
bad		_RMSE_	Стандартная ошибка	0.376644	.	0.380088
bad		_SBC_	Байесовский критерий Ш...	6634.777	.	
bad		_SSE_	Сумма квадратов ошибок	2121.915	.	246.0891
bad		_SUMW_	Сумма временных частот...	14973.71	.	1703.431
bad		_MISC_	Коэффициент неправиль...	0.200581	.	0.200697
bad		_AUR_	Area Under ROC	0.755207	.	0.743576
bad		_Gini_	Gini Coefficient	0.510414	.	0.487156
bad		_KS_	Kolmogorov-Smirnov Stati...	0.384649	.	0.383433
bad		_ARATIO_	Accuracy Ratio	0.510414	.	0.487156

Figure 3. The window with the results of statistical estimates of quality developed model

References. **1.** Anderson, Billie S., and R. Wayne Thompson. Developing Credit Scorecards Using SAS Credit Scoring for Enterprise Miner 5.3. - Cary: SAS Institute Inc, 2009. - 41 p. **2.** Siddiqi N. Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring - Cary: SAS Institute Inc, 2005. - 208 p. **3.** Терентьев А.Н. SAS BASE: Основы программирования (научное издание) / Терентьев А.Н., Домрачев В.Н., Костецкий Р.И. - К: Эдельвейс, 2014. - 304 с. - ISBN 978-966-2748-49-9

Basyuk T.M.
Lviv Polytechnic National University, Lviv, Ukraine

Web site promotion on Google

By analyzing known financial review [1] it can be concluded that the rating information company resource is comparable with the success rate of its activities. The success of any online resource depends on its popularity among search engines. Thus, it is important to make an emphasis on popular search engines before resource popularization in the region, since it is usually elected by their potential users. At the beginning of 2016 year Google is the most popular in the vast global network. It is the leader in all countries except Russia (Yandex), China (Baidu), Czech Republic (Seznam) and North Korea (Naver) [2]. Therefore, for the promotion of internet resources in Ukraine, it is necessary to make an emphasis on the search engine Google, as it can provide the greatest number of potential users in this market segment.

As it is shown [3] today, there is a set of methodologies for ranking online resource from local algorithms used by search engines (PR, TIC) to global, which are a set of parameters. Rating of internet resource, and thus its relative position is being determined by a set of factors that are actually unexplored due to the closeness of the search engine and copyright. Studies available for today [3,4], describe in the general form the features of the search engines and the possible factors that influence the popularizing of the resource, thus bypassing the necessary mechanisms of rankings that are used by the most popular search engine. Given that work on creation of the resource, in most cases, is divided into two parts. The first is the actual development of online resources in accordance with the user's criteria and location in a global network, the second is its popularizing (SEO) due to search engine. The above algorithm requires a significant amount of time (given the duration of the process) as financial costs (actually executed one and the same work, but with double salary).

The feature of conducted researches is that the determining of factors that influence the position of the online resource in the search results Google makes urgent task of system design evaluation and recommendations on creating websites which are optimized for the relevant user queries on the Internet. The study aims to determine the main factors ranging Internet resources in Google search engine. In order to achieve this goal it is necessary to solve the following main tasks: to analyze the known techniques and methods for evaluating online information and resources search engine Google; to identify factors that carry the greatest impact on the process of building high rating resource; to analyze the main methods of evaluation resource history, sanctions definition for each website, download speed from remote hosts and the access time, the structure of external links and keyword. Internet resource designed with the usage of these factors can pretend on top-end position in the ranking of Google search engine as well others who work on similar principles, which in turn will lead to its promotion in the global network. On the basis of research on the Internet resource position in the SERPs of Google it was managed to identify the following key factors of ranking (Figure 1).

The most important factor affecting the promotion of Internet resource is the domain name (the name which will post links and other resources that will be displayed in the search results). Thus, mostly it the impact was caused by age, history, keywords, address, subdomain and history of sanctions. Domain Age plays a significant role, as in the promotion of the new site (six months age) it is required to make multiple changes starting from the internal structure, completing by the external reference environment. Domain history can either help or influence the promotion, for example, if the subject of the resource did not change, it positively influences the ranking and vice versa. Given that set of SEO-optimizers looking for some thematic resources in order to develop websites. The history of any online resource can be found in the global archive ARCHIVE.ORG (http://archive.org/web/web.php) which are indexed websites WAN at different times.

Keywords in the domain - provide tangible result in the promotion, but with the introduction of the algorithm EMD (Exact Match Domains) domains that contain in the name the set of keywords

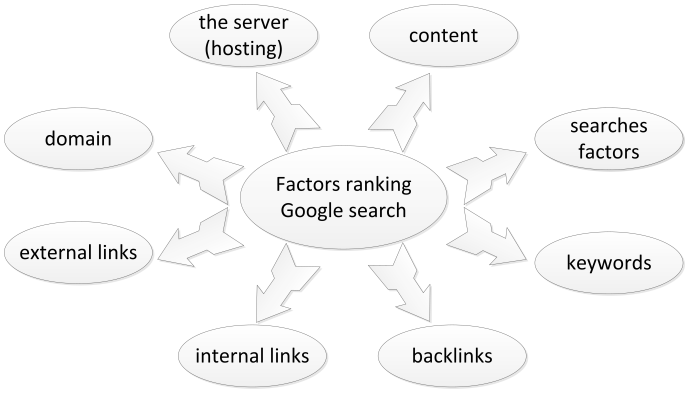


Figure 1. Ranking factors of Google search

or search query, are deliberately blocked.

Sanctions history of domain - causes a significant impact on the process of moving the resource, as well as search engine also evaluates past sanctions. Google uses multiple algorithms, including common - Google Panda and Google Penguin. Google Panda algorithm introduced on Feb. 23, 2011, whose main objective was to improve the quality of search results by evaluating material resources. Google Penguin entered in the search engine on 24 April 2012 and intended to fight online resources that use questionable backlinks. The credibility and weight of online resource is largely dependent on what web-sites are linked to it. Another factor that influences the ranking of resources is a server (hosting). The main criteria that can influence it are: long response of online resource, unavailability of Internet resources (especially when Google Bot tries to enter for indexing resource), Uptime-life (the time of continuous operation of a computer system or part of it which is measured from the moment of downloading until stop or completion of monitoring). This factor can be analyzed with the help of the Internet resource http://sitespeed.me/ showing information on the speed of loading web pages.

Content - the main element of the Internet resource that is subject to mandatory optimization to improve the popularization resource. All content must meet certain structure (nesting level, no duplicate content, availability microdata) and be informative for the user.

Searches factors - additional custom features that affect the ranking resource. The following are among the most common factors: the indicator and a different perception. Under the indicator of rejections the number of visitors to the resource is meant. Formally, the system Google Analytics considers the rejection to be a situation in which the process of visiting online resources, in addition to viewing the page, there were no other action. This indicator serves as a criterion of quality audiences page or internet resource.

Conclusion. The study proved that the Google search engine contains a significant amount factors of ranking. Thus, for successful implementation of the process of building high rating resources it is necessary to influence those that are known in order to obtain a meaningful result in a permanent increase of website traffic.

References. **1.** Стрикчиола Д. SEO. Искусство раскрутки сайтов / Джесси Стрикчиола, Рэнд Фишкин, Стефан Спенсер – СПб.:BHV, 2011. – 520с. **2.** Лоу Д. Google. Прошлое, настоящее, будущее / Джанет Лоу – М.: Манн, Иванов и Фербер, 2009. – 320с. **3.** Граппоне Д. Поисковая оптимизация сайтов. Исчерпывающее руководство / Дженнифер Граппоне, Градива Казн – М.:Эксмо, 2012. – 528с. **4.** Басюк Т.М. Принципи побудови системи аналізу та просування інтернет ресурсів / Т.М.Басюк // Комп'ютерні науки та інформаційні технології. – Львів : Нац. ун-т "Львівська політехніка". – 2012. – № 784. – С.43-48.

Didenko D.G.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Fuzzy recovery dictionary by the id and the name into database

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". Similarity duplicate determined by the distance between tuples using "editorial distance", which is more detail described in article [3].

Implementation. After the incoming import file as XML-file, 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 of lines.
2. Import incoming XML-file using JPA [4].
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".

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. A further development is the implementation of this approach in distributed systems for MapReduce (Hadoop for example) using the programming language Java. 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. + CD-ROM — (Учебная литература для вузов). **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.** 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. **4.** Schildt Herbert. Java 8. Ninth Edition. – New York. McGraw Hill Professional, 2014. - 1376 p.

***Ivanova Y.V.*[1], *Terentiev O.N.*[1], *Korshevnyuk L.O.*[1], *Prosyankina-Zharova T.I.*[2]**
[1]*Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine;* [2]*European University, Uman, Ukraine*

Using modified logistic regression to increase customer response rate to marketing campaigns

Introduction. While creating marketing campaigns, most companies use conventional predictive models, which target all customers who are likely to buy. This approach can lead to wasting money on customers who will buy regardless of the marketing contact [1]. Therefore, considered models which look for customers who are likely to buy or respond positively to marketing campaigns when they are targeted but are not likely to buy if they are not targeted.

Research methodology and result analysis. In incremental response modeling, one of the basic models is the difference score model, which measures the differences between predicted values from the model of the treatment group and from the model of the control group [1]. Suppose there are two data sets, D_T and D_C, which are the treatment group and the control group, respectively.

Denote a dependent variable y and explanatory variables x, and denote the number of observations in each group, n_t and n_c. Then,

$$D_T = \{x_i, y_i\}_{i=1}^{n_t} \tag{1}$$

$$D_C = \{x_i, y_i\}_{i=1}^{n_c} \tag{2}$$

$$D = D_T \cup D_C = \{x_i, y_i\}_{i=1}^{n}, n = n_t + n_c \tag{3}$$

Without loss of generality, a linear model is considered as follows:

$$Y = X\beta + \epsilon \tag{4}$$

Difference score from two separate models. For a more correct modeling two data sets are used – treatment and control. As a result, two models are built separately for each data set:

$$\hat{Y_T} = X_T\hat{\beta_T} \tag{5}$$

$$\hat{Y_C} = X_C\hat{\beta_C} \tag{6}$$

Then, both models are used to calculate predicted values from the entire data sets ($D = D_T \cup D_C$). The difference scores can be obtained from the predicted values as

$$\hat{DS}_i = (\hat{Y_T} - \hat{Y_C})_i \tag{7}$$

where $i = 1, 2, ...n$.

Customers who have a positive value of $\hat{DS}_i$ are initially considered as the incremental responders as a result of the promotion campaign. However, to determine the final set of incremental responders, further analyses must be performed with the ranked difference scores in decreasing order.

Difference score from a combined model. This model uses a new indicator variable, T_i: $T_i = 1$ for D_T and $T_i = 0$ for D_C and he fit the model to the entire, combined data ($D = D_T \cup D_C$) :

$$Y = X\beta + T\gamma + XT\varphi + \epsilon \tag{8}$$

Then, the difference scores are obtained from the estimated equation:

$$\hat{Y_T} = X\hat{\beta} + \hat{\gamma} + X\hat{\varphi} \tag{9}$$

$$\hat{Y_C} = X\hat{\beta} \tag{10}$$

$$\hat{DS}_i = (\hat{Y_T} - \hat{Y_C})_i = \hat{\gamma} + X\hat{\varphi} \tag{11}$$

where $i = 1, 2, ...n$.

This is a simplified expression of the difference scores in linear model, but in general they are calculated from the form of conditional expectation of Y:

$$\hat{DS}_i = (\hat{Y_T} - \hat{Y_C})_i = E(Y|X, T = 1) - E(Y|X, T = 0) \tag{12}$$

where $i = 1, 2, ...n$.

One advantage of this model is that it shows some influential variables directly to the incremental response through the significant parameter estimates [2].

Conclusions. A program has been created using Matlab to build incremental response models for different types of models. Customer data was taken from a database, which contains information on 9000 clients, including their previous records and traits [3]. The incremental response model diagnostics plot in fig. 2 shows both the predicted and observed incremental response rate by decile. The top decile has the highest incremental response rate, with a predicted increment of 21,2 % and an observed increment of 35,2 %. This predicted increment is approximately three times higher than the average incremental response rate of the data (8,3 %).

In fig. 3 response rate for a combined model is shown. The average increment in response rate equals 8,9%. This model also shows predicted increment of 44,4% , which is more than was achieved with previous model.

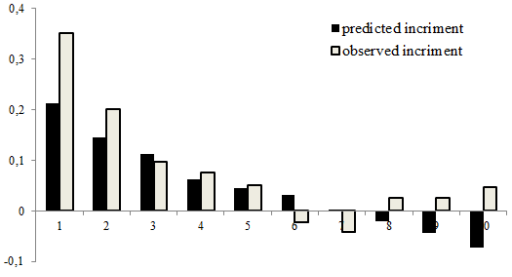


Figure 1. Incremental response model diagnostics plot for separate models. Y-axis shows deciles, while X-axis shows increment rate

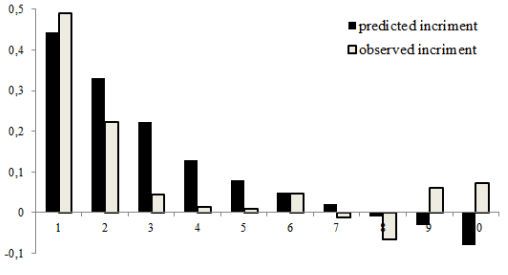


Figure 2. Incremental response model diagnostics plot for combined model. Y-axis shows deciles, while X-axis shows increment rate

Table 1. Example of wide table placed between text paragraphs

Model Name	SSE	Misc	Roc	Gini
Separate models	0.1878	0.2731	0.7106	0.4212
Combined model	0.1962	0.2837	0.6850	0.3701

For both models, all results are in the range of accepted values:

- *Roc* – the higher the value is, the better. $Roc = 1$ means perfect classifier, $Roc = 0,5$ means that the classifier is purely random.
- *Gini* - the higher the value is, the better. $Gini = 1$ means perfect classifier, $Gini = 0$ means that the classifier is purely random.
- *Misc* values range from 0 to 1, where 0 means a perfect classifier.

The first model shows better results in *Roc* and *Gini* as well as lower values of mean-square error and classification error rate. Combined model shows inferior results, which is caused by a large number of predictors which increase the model's complexity. This makes the model more suitable for use with stepwise regression. However, the combined model demonstrates a predicted increment rate much higher than average incremental response rate, which allows to find more potential customers.

References. **1.** Taiyeong L., Ruiwen Z., Xiangxiang M. and Laura R. Incremental Response Modeling Using SAS® Enterprise Miner // Conference SAS Global Forum 2013. - LA.: SAS Institute, 2013. - 13 p. **2.** Lo. V. The True Lift Model: A Novel Data Mining Approach to Response Modeling in Database Marketing, ACM SIGKDD Explorations Newsletter, 2002, 4. – 78–86 p. **3.** Data set and program https://sites.google.com/site/data4mining/datasets/sait_2016_thesis.

Kavrelishvili A.G.
Georgian Technical University, Georgia, Tbilisi

Hotel Management System, Development stages and Processing

Tourism is the world's largest industry and creator of jobs across national and regional economies. Travel and Tourism creates jobs and wealth and has tremendous potential to contribute to economically, environmentally and socially sustainable development in both developed countries and emerging nations. Tourism industry will continue to grow globally over the short to medium term. Cloud Based Hotel Management System is a system that helps hotel managers to manage all processes inside and outside of hotel. Our Software is designed to accommodate the needs of various types of properties such as the hotels, motels, resorts, small hotel franchisees, hostel and apartments. Software constitutes of Easy Check-in / Check-out, 2 Click guest reservation, Group Management, House Keeping, Night Audit, Travel Agent, Guest History and lot other. Our Hotel management system enables hotels of all sizes and types to manage day-to-day functions with ease. System offers customers to interfaces with many third party software and hardware solutions like Financial Accounting System, Payment Gateways and Call Accounting. Mobile and Tablet friendly, the online system can accessed from anywhere without any hassle. All the data is stored on secured servers with redundancy in place which makes sure that any hotel stay operational eternal. In the World everything is growing faster, technologies.

Why cloud based hotel management system is more effective than casual desktop software? There isn't one answer, because cloud based solution has a lots of benefits. It means that potential customer has ability to pay package fee and get unique username and password to access system and use it fluently anytime and everywhere with computer, tablet or mobile. This cloud software is designed in such a way that it programs automatically to accommodate any new updates or features. Moreover, hoteliers can customize the interface and choose the features that are pertinent to the property. There is no need to maintain disks of software or go through confusing instillation steps. Desktop based hotel software needs huge setup costs and regular maintenances to the server can cost a bomb to the hotelier. Cloud replaces the hardware and labor costs so the price drops automatically.

When a hotel receives a booking from a travel website, hotel website, email, phone, or walk-in, the system will automatically update revised availability. In addition, advance option such as stop-sale allows hoteliers to instantly stop distribution to avoid over booking. System has ability to save arbitrary error and after that it classify and get developers means to think about this error and careful planning fix way and implement it in system, when this stage finish, if error get again, system has ability to fix it automatically, without human. Consolidated reports show real-time data allowing hotels for live analysis and better strategic planning. The system collects all the relevant data and allows hotels to create reports on number of bookings, bookings received per channel, rates, etc. All the relevant information can be accessed anytime by identified user which has advanced permissions to see it. Report is presented in an easy to understand format. The software is unique in that it offers the flexibility and accessibility of a cloud-based solution while also providing the security and peace of mind from having the software installed locally. Software was designed to be extremely user-friendly, with an approachable, intuitive interface that helps accommodation managers gain control of their front desk, online booking engine, financials, maintenance, and more.

Cloud based hotel management system is a way to develop a hotel management, fix current problems, improve operation and increase the efficiency of hotel day to day activities and it's a chance to avoid waste of time. Researches shows that business always needs cheap and reliable software. It should be noted that, main for hotel business is all time accessibility and trustworthiness management system, which achieve cloud based solution and Sophisticated functionality

References. **1.** Larisa Dolikashvili (2010) Basis of Tourism: Hotel and Guest House management. P.149–340.

***Malinovskyi A.D.*[1], *Nikolaiev S.S.*[2]**
[1]*Faculty of informatics and computer science NTUU "KPI", Kyiv, Ukraine;* [2]*Institute for Applied System Analysis NTUU "KPI", Kyiv, Ukraine*

The Architecture of a mobile system for heart health monitoring

Abstract. This paper describes an architecture of a mobile system for daily noninvasive heart health monitoring using smartphone and cloud service. The major components of the system are described and functions of the sub-systems are explained. The client-server interaction model is discussed.

Relevance. Nowadays mobile e-health systems are becoming an important trend because new emerging technologies allow creating cheap and effective personal solutions for human biosignals monitoring. Portative heart monitoring systems have significant importance, because cardio vascular diseases are the major cause of death globally. CVD cause more than 4.3 million deaths in Europe (48 percent of all deaths in the world) every year [1]. There are established markets of different wearable devices (like fitness-trackers, pulsometers, personal electrocardiogram (ECG) analyzers), but most of them were created for entertainment. Some of the wearable devices were designed for personal health monitoring. These devices generate huge flows of data describing user organism state. For now they fully rely on the approach which involves real doctor in manual viewing and analyzing each user's record. This means that each user's data is viewed and analyzed by the real doctor. The share on the market of health devices is still small but is steadily growing. At the moment the described approach is viable but in the nearest future the biosignals data flows will overwhelm the ability of doctors to view through these amounts of records. Neither practitioners nor specially trained operators of the systems will have enough time to analyze the data. An automation and an approach called "healthy user records rejection" could be solutions to this problem. Application of machine and deep learning methods along with rule-based systems allow finding users at risk and automatically filtering out healthy people. To achieve better performance a set of models (like artificial neural networks, SVMs, RBMs, Random Forests, Fuzzy neural networks and others) are used [1]. Such personal systems for ECG monitoring and analysis allow to significantly reduce the load on doctors, that will increase their productivity and eliminate the probability of false diagnosis. These systems will be constructed using existing infrastructure, and will use the smartphone as a processing device and / or data storage. But taking into account limited memory of smartphones in tasks of constant monitoring the need in client-server model architecture arises. This will make it possible to connect smart wearables with sensors, mobile phones and powerful servers with huge storage for computing precise machine learning algorithms on the data into one distributed system [2], [3].

System overview. The system processes the ECG, and on the basis of the automatic analysis, gives the probability of pathologies and possible diseases. It also calculates quantitive metrics of health and shows them to the users. In the scope of the system we will describe in detail the architecture of lightweight mobile client that interacts with ECG sensors and a cloud service. Further the composite diagram of the light mobile client is shown and it's components are described(Fig. 1) [4], [5].

User Account Management Module. For personalized systems an important element is "User Account Management module" to track which user's data is collected and do we show results to their owners. Also this module is crucial to differentiate the owner of the device and guests who want just to try out the system. This module also processes and stores such personal information like height, weight, age, sex etc. It is necessary for data analysis that takes into account personal parameters of particular user.

Data Management. Module is responsible for planning and scheduling data collection from sensor device (to start recording at a particular time, when you connect device, etc.), upload the data to

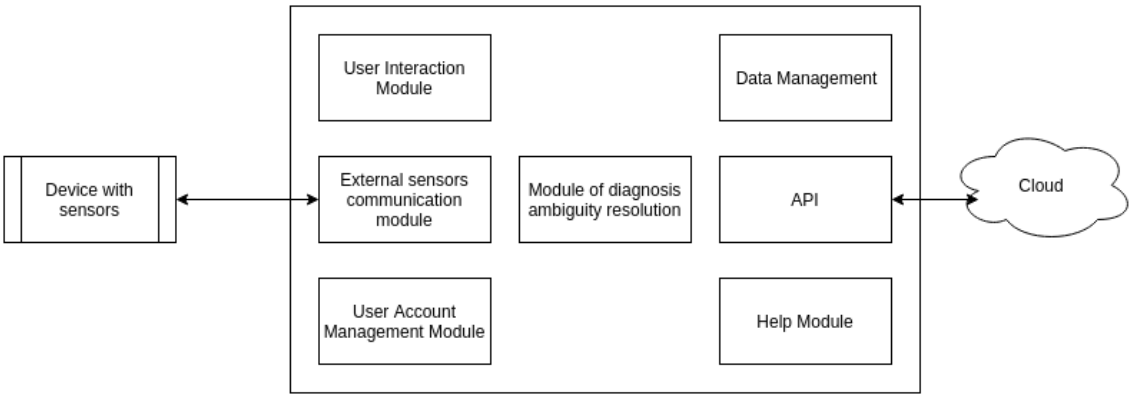


Figure 1. Mobile system architecture

your smartphone (on-demand, when your smartphone is connected to the charging, etc.). And also it is responsible for filtering downloaded data from the device with sensors (removal of all sorts of artifacts, noise).

External sensors communication module. This subsystem is responsible for interaction with the sensors. The interaction takes place via Bluetooth technology. Functionality includes the ability to search for devices with ECG sensors and connect them to the smartphone. Different data-acquisition modes, including manual recording, are switched on or when one of the rules is triggered(like alarm, connection to known device and so on). Also an option of displaying the data points in real-time is provided. It allows using this system for real-time analysis in medical institutions. The sensor device has the ability to record data. Then the user can download the records to his smartphone and send them to the doctor or to the cloud for analysis.

API. This subsystem is developed for interaction between cloud server and user's smartphone application via the Internet. For user data safety reasons, all analysis is performed and results are stored in the cloud, so the module is used to transmit user cardiograms, and receive the results of the data analysis. Storing data in the cloud also allows the system to increase accuracy by retraining it's models on the ever-growing dataset.

Module of diagnosis ambiguity resolution. In some cases the system cannot make a conclusion using only provided standard data about patient. For example when the user takes specific drugs his ECG might look as if the user has heart pathology. IN such cases the expert decision making system needs additional information. The purpose of this subsystem is to ask user relevant questions that will help to decrease the ambiguity. The user is interactively questioned according to the report received from the cloud and known rules of diagnosis ambiguity resolution. Depending on the answers to these questions, the diagnosis is fitted to current user personal condition.

User Interaction Module. The module is designed for the user to interact with data. The user can monitor their state using a calendar that displays the results of previous tests. User can share results in social networks to create social effect. He can view any recorded cardiogram that is stored on a smartphone. Access to older cardiograms may be useful in case of changing the doctor during clinical monitoring. The user can also manage the list of doctors with whom to share his data via the system. Also module supports comments displaying on top of provided cardiograms visualization.

Help Module. The module is a reference system element that consists of next elements:

- Help on the health indicator - a detailed description of all possible diagnoses and the heart health states, that may be shown on the indicator according to the results of analysis.

- Help on the diseases and pathologies that can and cannot be detected by the system. It is a necessary element of such systems, in order to enable patients to receive detailed information about the diseases which were diagnosed to them.
- Help on how to use the application.

Data Analysis Process. In the context of the system, the analysis of user's ECG looks like following (Fig. 2) :

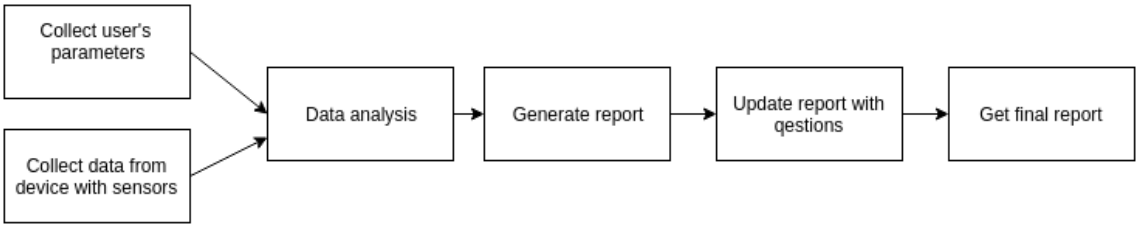


Figure 2. Data flow analysis

1. Parameters of the users and their ECG are sent to the cloud
2. Data analysis is carried out by machine learning methods and rule based systems
3. A preliminary report is sent to the user
4. If necessary, report is clarified by additional analysis results of issues detected on 2-nd stage
5. The user is given a personalized report on the state of health

Summary. The lightweight client architecture of the mobile system for heart health tracking was proposed and it's components were described. ECG analysis was described in context of the system. The proposed system allows gathering personalized ECG data in huge quantities and various situations of everyday life and it makes possible to automate analysis procedure with precise machine learning models for heart pathologies detection and prediction excluding doctors from the routine processes.

References. **1.** Sergii Nikolaiev : ARCHITECTURE OF AUTOMATIC MACHINE LEARNING SYSTEM FOR HEART PATHOLOGIES DETECTION USING ECGs // 3-rd International Conference on Computational Intelligence (ComInt 2015), pp. 159-160

2. S. Nikolaiev, Y. Tymoshenko : Reinvention of the Cardiovascular Diseases Prevention and Prediction Due to Ubiquitous Convergence of Mobile Apps and Machine Learning // The 2nd INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE "Information Technologies in Innovation Business"(ITIB) October, 7-9, 2015, Kharkiv, Ukraine

3. Fridman R.: Application Of Machine Learning Methods For Heart Pathologies Detection Using Ecgs. // Student Conference on Computer Science, Erfurt University of Applied Science (2015)

4. Rifat Shahriyar1, Md. Faizul Bari2, Gourab Kundu3, Sheikh Iqbal Ahamed4, Md. Mostofa Akbar5 : Intelligent Mobile Health Monitoring System (IMHMS) // International Journal of Control and Automation Vol.2, No.3, September 2009

5. Regin Joy Conejar, Haeng-Kon Kim : A Design of Mobile Convergence Architecture for U-healthcare // International Journal of Software Engineering and Its Applications Vol. 9, No. 1 (2015), pp. 253-260

Markov D.K.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Improved frustum culling

A typical game scene mostly contains a great variety of game objects. Every object may have a mesh, which consists of thousand of polygons. Therefore, if objects are not managed appropriately, they have major influence on performance and frame rate.

Of cause, libraries namely OpenGL and DirectX have a clipper (utility which deletes all polygons that are off the screen). However, as it is evident from the rendering pipeline (fig.1) , vertex shader and triangle assembly by the clipping stage already have processed vertex stream. Furthermore, all matrix computations should have been done by this moment. In this case, some useless work can be avoided using some high level technique, which will be deleting objects before the actual viewport clipping.

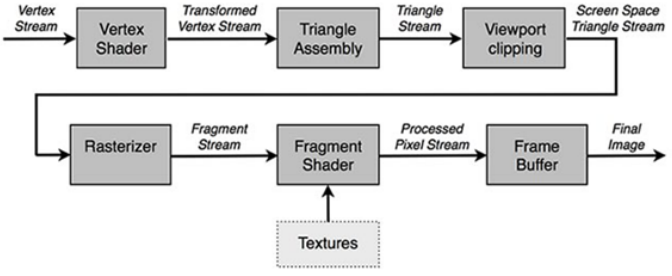


Figure 1. Rendering pipeline

One of these techniques is called view frustum culling. Frustum culling works by defining a volume that wraps the space currently visible from the given point of view. The volume is built by constructing a pyramid with the apex at the eye's position and the four triangular faces aligned with the screen borders. Two parallel planes (the near- and far- clipping planes) cut the pyramid. Nothing closer than the screen or further than some predefined distance will be included in the view volume. The resulting volume is known in mathematics as a frustum (fig.2). The objects that are completely outside of this volume have no chance to be seen by the camera and may be rejected on a per-object level.

Once the frustum is built, objects can be tested against it for visibility. Each object in the scene may be completely enclosed by a bounding primitive, such as a sphere, cube, cylinder, or cone, which used to test against the frustum rather than the objects itself. In such a way the overhead of duplicating the work already in the rendering pipeline is eschewed and a higher-level of efficiency in comparison to per-polygon culling is provided. If the bounding primitive is completely outside the frustum, there is no need to process and render this object.

The common representation of view frustum is a traditional six planes (up, down, left, right, near, far) approach. Every plane divides a space into two parts, so we can check in which part a point lies. However, there are some disadvantages of this approach. The first one is an expensive check operation, which should provided for every plane. Secondly, it is much cheaper to recreate the whole frustum than just to move it, because moving operation is followed by some difficulties including square root and divides.

The radar approach was designed for a two-dimensional system. The only intended purpose was to show the objects covered by a radar wave by leaving out those objects that were not covered. The approach proved to be very fast if some bounding primitives were inside the view, and the method was quickly extended to work in

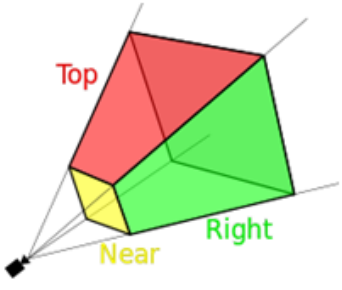


Figure 2. View frustum

three dimensions. Now, the algorithm has been used for frustum culling with incredible results. To describe how the radar approach works, the two-dimensional algorithm should be introduced first.

When working on a system with two dimensions, a frustum becomes a triangle, and as a frustum in a symmetric object, the triangle is known to be isosceles. The segment that starts in the apex and is perpendicular to the base is called in mathematics the triangle height. The handy part of the height is that as the frustum is symmetric; it splits the isosceles triangle into two equal halves. Coincidentally, the height happens to be directionally parallel to the forward vector of the camera. Using the radar approach, a frustum is defined to be as intuitive as the camera. Thus, the frustum construction is almost free. The two-dimensional radar only requires the field-of-view angle (FOV), the forward vector and the right vector. The same three parameters that the camera uses, can be extracted directly from the camera, which is very convenient. The radar approach does not use the FOV itself, but the tangent of this angle.

Therefore, to find out if the point P is in the view frustum in three-dimensional system, the following steps should be fulfilled:

```
bool Frustum::checkPointIn(const Vector3f &point) {
    // 1) Check z-coordinate against near and far planes
    if (point.z > Camera::far_plane || point.z < Camera::near_plane)
        return false;

    // Camera already has precomputed tangent of xFOV and yFOV (changes // only if FOV chan
    // 2) Check x-coordinate against left and right planes
    fp32 x_offset = point.z * Camera::x_halffov_tang;
    fp32 min_x_coord = Camera::position().x - x_offset;
    fp32 max_x_coord = Camera::position().x + x_offset;
    if (point.x > max_x_coord || point.x < min_x_coord)
        return false;

    // 3) Check y-coordinate against up and down planes
    fp32 y_offset = point.z * Camera::y_halffov_tang;
    fp32 min_y_coord = Camera::position().y - y_offset;
    fp32 max_y_coord = Camera::position().y + y_offset;
    if (point.y > max_y_coord || point.y < min_y_coord)
        return false;

    return true;
}
```

In work was presented improvement for frustum culling that uses precomputed FOV tangent instead of plane checking.

References. **1.** Culling explained, 2013. Available from: <http://docs.cryengine.com/display/SDKDOC4/Culling+Explained>. [Accessed: 28th October 2015].

2. Frustum Culling in OpenGL, 2000. Available from: <http://www.crownandcutlass.com/features/technicaldetails/frustum.html>. [Accessed: 28th October 2015].

3. Efficient View Frustum Culling, 2002. Available from: <http://www.cescg.org/CESCG-2002/DSykoraJJelinek/>. [Accessed: 28th October 2015].

4. View Frustum Culling, 2014. Available from: <http://www.lighthouse3d.com/tutorials/view-frustum-culling/>. [Accessed: 28th October 2015].

Povisenko S. V.
Faculty of Applied Mathematics of NTUU "KPI", Kyiv, Ukraine

Using of Saga pattern in Asynchronous Messaging Architecture for Big Data Processing

There is short description of CQRS (command-query responsibility segregation) pattern and its combination with Asynchronous Messaging Architecture. Shows the main advantages of their usage for processing large amounts of data.

Intro. Practice shows that the applications for processing large amounts of data need to be flexible and easily extensible. They should be able to continue work from the last successful status when an error occurs. As you build more advanced solutions, you may find the need to do collaborative processing that goes across multiple systems. Since failures are normal in these environments, some activities may fail while others succeed, which will yield inconsistent outcomes. There are mechanisms to handle these situations such as coordinated 2-phase commit (distributed transactions), but in the world of higher scale and cloud computing, these options are no longer suitable.

Saga pattern. The Saga interaction pattern was designed to handle these failures. A Saga is a distribution of long-living transactions where steps may interleave, each with associated compensating transactions providing a compensation path across databases in the occurrence of a fault that may or may not compensate the entire chain back to the originator.

A Saga is a set of rules for routing a job to multiple collaborating parties, and allowing these parties to backtrack and/or take corrective action in the case of failure. A Saga is a distribution of multiple workflows across multiple systems, each providing a path (fork) of compensating actions in the event that any of the steps in the workflow fails (fig. 2).

Sagas have been proposed as a transaction model for long lived activities. A saga is a set of relatively independent (component) transactions $T = T_1, T_2, ..., T_n$ which can interleave in any way with component transactions of other sagas. Component transactions within a saga execute in a predefined order which, in the simplest case, is either sequential or parallel (no order). Each component transaction $T_i, (0 \leqslant i \leqslant n)$ is associated with a compensating transaction CT_1. A compensating transaction CT_i undoes, from a semantic point of view, any effects of T_i, but does not necessarily restore the database to the state that existed when T_i began executing [1]. Further more there are other patterns that build from the Saga pattern such as Kangaroo Transactions [2]. Kangaroo transactions deal with transactions in mobile environments that hop from one base station to another as the mobile unit moves through cells. Similar to the Saga pattern a kangaroo transaction is designed for the purpose of compensating in the event of failure. Saga is similar to a transaction in the sense that it provides a shared context for an attempt to get a distributed consensus. Unlike a transaction which insures ACID (Atomicity, Consistency, Isolation, Durability) properties, Sagas are not. When a saga is aborted the only thing the coordinator can do is pass the status to the participants. Each of the services is responsible to do its best effort to handle the abort (either by rolling back, compensation or whatever).

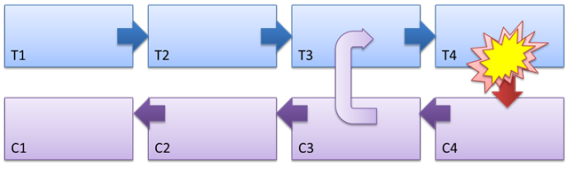


Figure 1. Saga compensating due to a failed transaction

Asynchronous Messaging Architecture. Suppose that we have a system that consists of different modules and they must interact in some way. On practice it turns out that the form and method of communication is a big problem, especially for processing large amounts of data applications. Asynchronous Messaging Architecture solve this problem.

In a messaging based system, both the input and output from services are defined as either commands or events. Each service subscribes to the events that it is interested in consuming, and then receives these events reliably via a mechanism such as a messaging queue/broker, when the

events are placed on the queue by other services.

Following this approach, the stock notification subsystem could now be remodelled as follows:

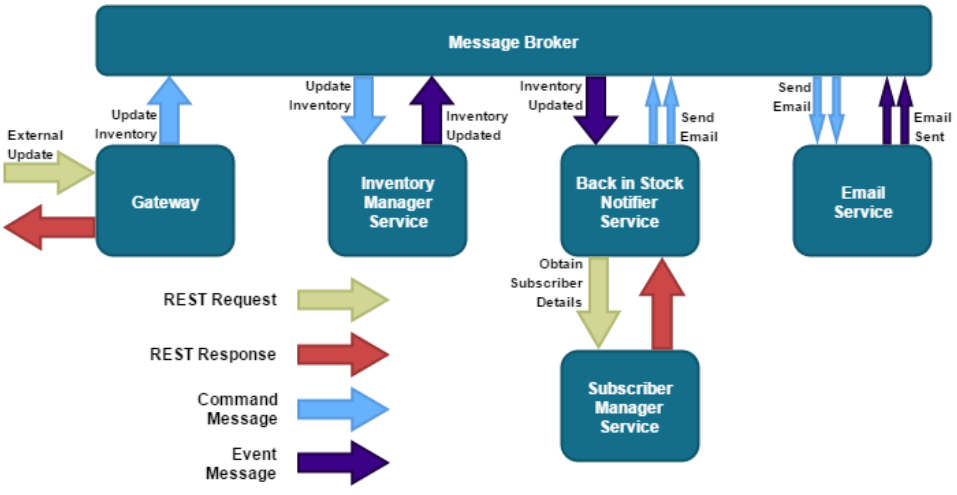


Figure 2. Asynchronous Messaging Architecture of hypothesis application

Cohesion is obtained via a shared knowledge of queue names, and a consistent and well known command/event format; an event or command fired by one service should be able to be consumed by the subscriber services. In this architecture, a great deal of flexibility, service isolation and autonomy is achieved.

The Inventory Manager for instance, has a single responsibility, updating the inventory, and is not concerned with any other services that are triggered once it has performed its task. Therefore, additional services can be added that consume Inventory Updated events without having to modify the Inventory Manager Service, or any pipeline orchestrator. Also, it really doesn't care (or have any knowledge of) if the Back in Stock Notifier has died a horrible death; the inventory has been updated so it's a job well done, as far as the Inventory Manager is concerned. This obliviousness of failure from the Inventory Manager service is actually a good thing; we MUST still have a strategy for dealing with the Back in Stock Notifier failure scenario, but as I have stated previously, it could be argued that this is not the responsibility of the Inventory Manager itself.

This ability to deal with change such as adding, removing or modifying services without affecting the operation or code of other services, along with gracefully handling stressors such as service failure, are two of the most important things to consider when designing a system for processing large amounts of data.

Conclusions. As we see Saga is a excellent candidate for Inventory Manage Service as a component which will be responsible for one business transaction. The main benefit of that is retryable conception, an ability of retry from last success state of Saga. So, we see, that Big Data processing can get very good benefit from using it, that allows us be free about unexpected system fails or fails of independent components such a extra Web Service, etc.

References. **1.** Panos K. Chrysanthis , Panos K. Chrysanthis , Krithi Ramamritham , Krithi Ramamritham (1992). ACTA: The SAGA Continues. California: ACM Digital Library. 100-104. **2.** A mobile transaction model that captures both the data and movement behavior. September 1997. Mobile Networks and Applications. Available from: http://link.springer.com/article/10.1023%2FA%3A1013672431080 [September 1997]. **3.** Sagas and Workflows Same thing with different names or not. By Arnon Rotem-gal-oz Jan. 19, 09. Integration Zone. Available from: https://dzone.com/articles/are-sagas-and-workflows-same-t [Jan. 19, 09]. **4.** Is REST best in a microservices architecture?. By Craig Williams. December 2015. Available from: http://capgemini.github.io/architecture/is-rest-best-microservices [December 2015].

Radovanovic M., Strbac D.
Geographical Institute Jovan Cvijić SASA, Belgrade, Serbia

The Standardization of Geospatial Data in Serbia - Experiences and Problems

About the INSPIRE directive. With the formation of the European Union, it was necessary to harmonize the standards for geospatial data. This was necessary because each member had its standards and legal regulation concerning the geospatial data. In order to solve this problem the INSPIRE (Infrastructure for Spatial Information in the European Community) Directive was made in 2007, which provided the basic principles of standardization of geospatial data through the process of establishing a geospatial data infrastructure. Serbia has accepted the INSPIRE Directive and in 2009 transposed certain provisions of the directive in the Law on State Survey and Cadastre, providing a legal framework for the establishment of the National Spatial Data Infrastructure (NSDI)). The Act defines the subjects, content and manner of use of the NSDI data, the metadata, the metadata sets and services, restrictions on access to metadata and geospatial data services, the establishment of the NSDI authorities and the formation of the geoportal. In this way, Serbia has begun the process of standardization of geospatial data.

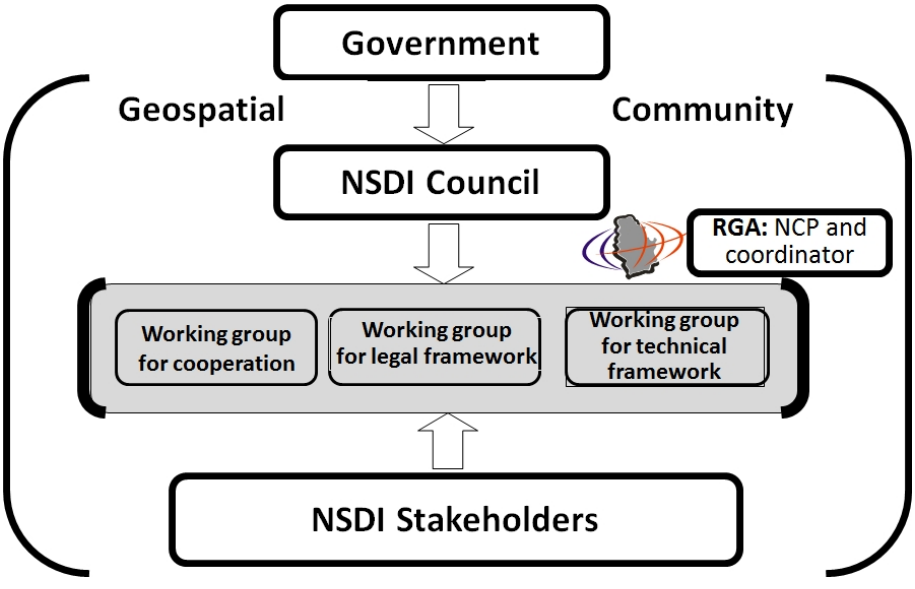


Figure 1. NSDI coordination structure

The bodies of the NSDI are the Council and working groups. The NSDI Council is appointed by the Government of Serbia and it is entrusted to create the institutional and legal framework for the establishment of geoinformation infrastructure on a national level. The role of the working groups is operationalization of topical issues for specific areas and their members are appointed by the Council (Figure 1).

In the part relating to the institutional framework, the following activities are developed:

- 2010: SDI Strategy of Serbia, period 2010 – 2012
- 2010; 2013: NSDI Council is appointed by The Serbian Government
- 2011: Decision on establishment of the NSDI Working groups
  - Working group for cooperation [19 members]
  - Working group for legal framework [10 members]
  - Working group for technical framework [23 members]

- 2011: Medium term program for NSDI, period 2011 – 2015
- 2012: Analysis of current situation of regulation in the geo-sector
- 2012: Geodata sharing policy: recommendations for cooperation
- 2012 – 2013: State of play: analysis of status within geo-sector
- 2014: Drafting on the Law on NSDI for full transposition

In the part which refers to the technical framework, the following activities are implemented:

- 2009: Initial Geoportal [www.geosrbija.rs] was launched
- 2010: Metadata – Proposal of metadata profile for geodata
- 2011: Metadata – Metadata Editor [INSPIRE + ISO 19115 + ISO 19139]
- 2009 – 2014: Geodata and WMS services provision for data publishing via the geoportal (ortophoto, cadastral parcels, addresses, administrative units, topographic, thematic maps, stakeholders/sector's spatial data, etc.)

Planned activities are related to:

- Translation of the INSPIRE Directive and the IRs under process of translation acquis communautaire into Serbian
- Drafting of NSDI Strategy for period up to 2020 (defined by NPAA)
- To define responsible public authority(s) for the INSPIRE themes prioritised per the Annexes [NSDI Council]
- Monitoring and Reporting in line with the IR (2016)
- Inclusion of the INSPIRE Implementing Rules in the national legislation by adoption of by-law regulations (up to a date of EU integration)

Experiences. Previous experiences in the implementation of the INSPIRE Directive are positive. During activity on its implementation awareness is raised of the need to establish NSDI for all users and makers of geospatial data. Their often opposed views were brought closer on the terms of use and exchange of data. An analysis of existing geo-data was done and determined their characteristics and authorities that are responsible for the collection and maintenance. A standard metadata creation was provided and raised awareness about their necessity. There is an initiative to change curricula in order to familiarize students with the characteristics of NSDI. A geoportal was formed with data from the competence of the institutions involved in the establishment of NSDI. Inquiries at the Geoportal are between six and seven million monthly. A regional cooperation has been realised with institutions from neighbouring countries, which led to a successful exchange of experiences. The transposition of the INSPIRE Directive has also been realised through legal regulation.

Problems. Among problems that have arisen in the implementation of the INSPIRE Directive, we will point out two. The first relates to the organization of an implementation unit. The implementation unit is located in the Republic Geodetic Authority which provides financial, technical and human resources. Employees of the implementation unit, in addition to activities relating to the establishment of NSDI perform tasks in the domain of home institution. In addition, the number of people employed in the implementation unit is small. This suggests that the scope of work that should be done has been underestimated. The second problem is related to the highly complex implementation in the technical part. Problems in this area significantly slow the implementation.

Sidletskiy V.M.
National University of Food Technologies, Kyiv, Ukraine

The automated operational control system for the food industry

Introduction. The use of computerized system technology is expected to continue to grow in the food industry as the cost of components decrease, equipment and manufacturing processes in an attempt to produce high quality, high value products and continue to update production facilities. Process food industries represents all measures in which mass, energy and information are transported, transformed or stored. Control system for a process processing industry, in general represents, control elements, sensors and actuators are assigned to the process, they form the base of observation, visualization, diagnosis of the process and the actuators possess the task of intervening into the process in order to achieve the desired result.Control system food industries that encompasses several types of control systems used in industrial production, including supervisory control and data acquisition systems, distributed control systems, operational control system and management and enterprise resource planning.

Problem statement. Modern automation systems require distributed structure because of the distributed nature of industrial plants in which the control instrumentation is widely spread throughout the plant. The efficiency of the system work as a whole depends on the efficient functioning of each subsystem on every level, this is why approach to the analysis and designing of the system not in terms of hierarchy, but as how a single mechanism in which every detail is important. Collection and prepossessing of sensors data requires distributed intelligence and an appropriate field communication system. The variety of plant automation functions to be executed and of decisions to be made at different automation levels require a system architecture that due to the hierarchical nature of the functions involved has also to be hierarchical. The efficiency of the system work as a whole depends on the efficient functioning of each subsystem on every level, this is why approach to the analysis and designing of the system not in terms of hierarchy, but as how a single mechanism in which every detail is important[1].This is due to complexity of the new manufacturing processes, the need to optimize costs, the increasing of the complexity in managing resources, the need to improve performances, the aspects tied to the information exchanges, the manufacture of complex products(Fig. 1). Then the Manufacturing Enterprise

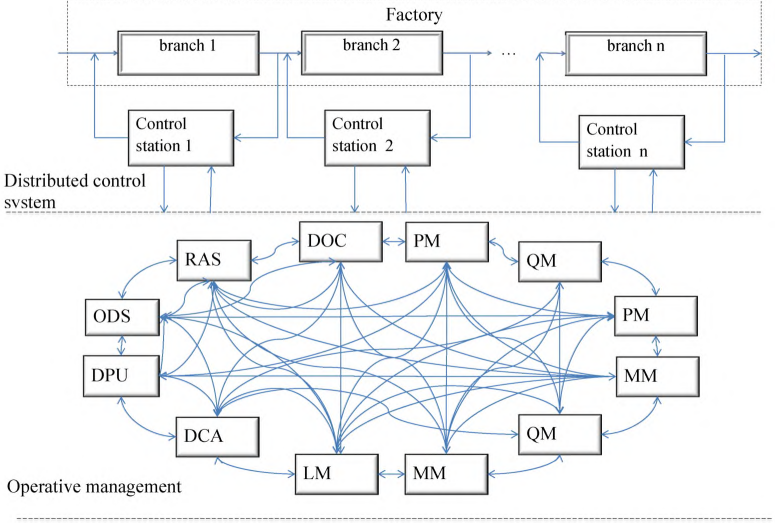


Figure 1. Block diagram of the operative management of the plant

Solutions Association (MESA) introduced some structure by defining 11 functions that set the scope of manufacturing execution systems (MES): RAS – Resource Allocation and Status, ODS –

Operations/Detail Scheduling, DPU – Dispatching Production Units, DOC – document Control, DCA – data Collection/Acquisition, LM – Labor Management, QM – Quality Management, PM – Process Management, MM – Maintenance Management, PTG – Product Tracking and Genealogy, PA – Performance Analysis[3]. This can be control and coordination of several interconnected unit operations that make up the total process, manufacturing cell or system level: control and co-ordination of groups of machines and supporting equipment working in co-ordination [2]. Operative control system and management,this is computerized systems used in manufacturing, to track and document the transformation of raw materials to finished goods. She can provide the right information at the right time and show the manufacturing decision maker "how the current conditions on the plant floor can be optimized to improve production output.

Solution. `wrapfigure` Parameters are interconnected through total process and twinning between operations and some other functions in process. A management is solved by a sequence of actions that can be represented with a tree structure (Fig.2). Each subtree associates a graph vertex with a set of tree nodes. To define this formally, we represent each tree node as the set of vertices associated with it. Graphs may be solved efficiently by dynamic programming for graphs of bounded treewidth, using the tree-decompositions of these graphs, by uses forward and backward recursion [4]. Such algorithms have a first step that approximates the treewidth, constructing a tree decomposition with this approximate width, and then a second step that performs dynamic programming in the approximate tree decomposition to compute the exact value of the treewidth. Let m_i, be the value for the variable of item i in the area operative management (Fig. 1) and define r_i and w_i as the revenue and of the cost of funds per unit for the variable of item i. Problem has the form.

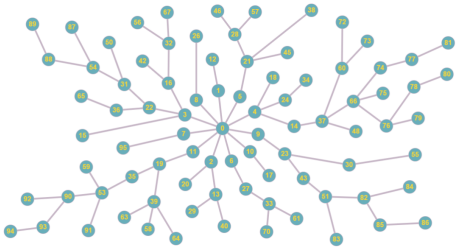


Figure 2. An tree structure in theoperative management for industry

$$\text{Maximize } r_1m_1 + r_2m_2 + \cdots + r_nm_n \to \max \tag{1}$$

Define $f_i(x_i) =$ maximum return for stages i, $i+1$, and n, given state x_i. The simplest way to determine a recursive equation is a two-step procedure [4].

Step 1.

$$f_i(x_i) = \max_{\substack{m_i=0,1,\ldots \frac{W}{w} \\ x_i=0,1,\ldots,W}} \{\, r_im_i + f_{i+1}(x_{i+1}),\} \,, i = 1, 2, \ldots, n. \tag{2}$$

Step 2.

$$f_i(x_i) = \max_{\substack{m_i=0,1,\ldots \frac{W}{w} \\ x_i=0,1,\ldots,W}} \{\, r_im_i + f_{i+1}(x_{i+1} - w_im_i),\} \,, i = 1, 2, \ldots, n. \tag{3}$$

Conclusions.. Operative management covers all functions of direct control of all process variables, monitors their values, and enables the plant engineers optimal interaction with the plant via sophisticated man-machine interfaces. That can be represented with a tree structure, and be performed efficiently by dynamic programming, a recursive equations of the two-stage procedure.

References. **1.** Richard L. Shell. and Ernest L.Hall., Handbook of industrial automation. Marcel Dekker, Inc., New York, NY 10016, 2000, p.p. 202-214 **2.** K. Mihlbachler (2015). Integrated Downstream Processing. An Enabling Manufacturing Approach [Online]. Available: http://www.lewa processtechnologies.com/sites/default/files/ assets/files/LEWA White Paper Integrated DSP v3 Combined Series Final.pdf. **3.** Manufacturing Enterprise Solutions Association – MESA Model. Http://mesa.org/en/modelstrat egicinitiatives/MESAModel.asp **4.** Taha, Hamdy A. "Operations research: an introduction —8th Univ. of Arkansas, Fayetteville,, New Jersey, Upper Saddle River NJ-07458, 2007 ed. p.p. 399-406.

Terentiev O.M., Makukha M.P., Savastynov V.V., Oparina E.L.
Institute for Applied System Analysis of NTUU "KPI", Kyiv, Ukraine

Analysis and modeling the dynamics changing of registered crimes taking into account the macroeconomic and political situation in Ukraine

Introduction. ESC IASA analytical team performs analysis of statistical information about crimes which provided by Ministry of Internal Affairs (MIA) of Ukraine. This scientific paper performs as part of the project NUKR.SFPP G4877 "Modelling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism" and aimed to present intermediate results of modeling. Based on the crimes statistical MIA data sine Dec 2002 to Dec 2014 and open data sources from Internet and State Statistical Service were performed researches about relationships between economic situation and crime changes in Kyiv.

Main analytical approaches. As analytical tools for modeling were used SAS Enterprise Miner and SAS Enterprise Guide systems. For analysis were used such economic indicators as GDP - gross domestic product (billions USD), DER - dollar exchange rate, IR - inflation rate, GR - gold reserves (billions USD). For construction the base dimensions of the space of measure in order to reduce the number of input variables used the principal component analysis. The first and second bases components calculate as:

$$PRIN_1 = 0,659 * GDP + 0,399 * DER - 0,138 * IR + 0,622 * GR, \quad (1)$$

$$PRIN_2 = 0,035 * GDP + 0,291 * DER + 0,956 * IR - 0,011 * GR. \quad (2)$$

According to inherent sense each of the main component is economical index which represent consistent pattern of related processes. As we could see the first principal component (PRIN_1) is formed mainly due to the positive influence of following parameters GDP and gold reserves. Meanwhile, the second principal component (PRIN_2) formed by using positive influence of inflation rate. On the fig. 1 shows relationships between principal components and crime changes.

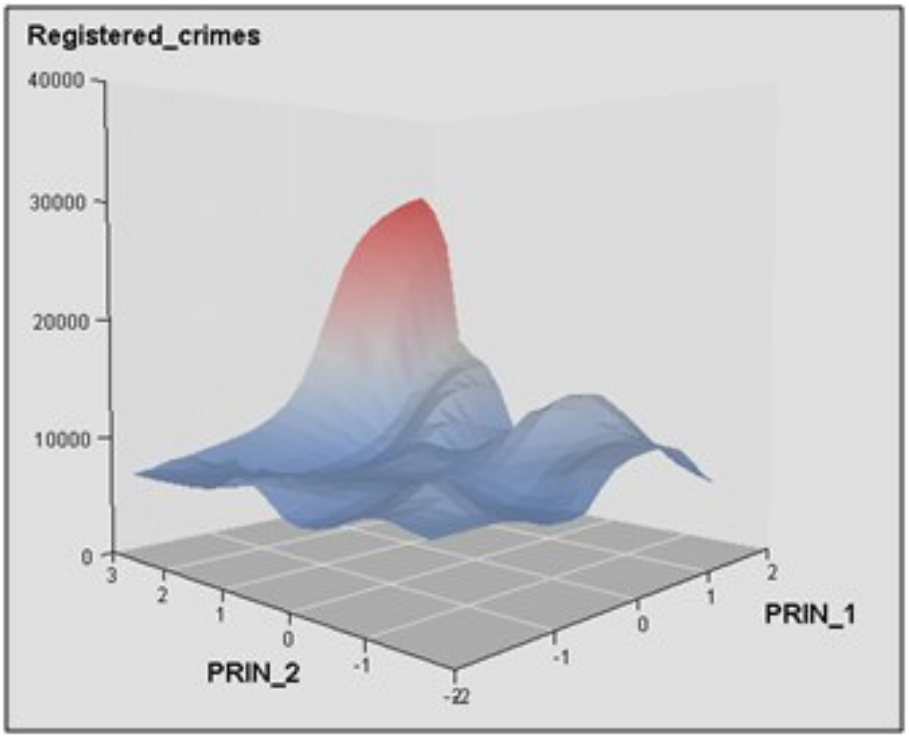


Figure 1. 3D plot in SAS Enterprise Miner.

Analysis of fig. 2 shows the presence of quarterly cycles in first principal component which caused by strong influence of GDP towards this index. Also, thorough analysis and communication with the MIA expert group revealed following interesting facts and events in the past:

- 2004 SEP - the beginning of the presidential campaign, when the main candidates for the presidency were Viktor Yushchenko and Viktor Yanukovych;
- 2012 DEC – changes in the Criminal Procedure Code of Ukraine. As a result of the changes the policemen have to register crimes that previously had not been included to the database of Ministry of Internal Affairs;
- 2013 DEC – the beginning of the political crisis 2013-2014 in Ukraine (so called dignity revolution);
- 2014 DEC – early parliamentary elections in Ukraine and the deepening of the economic crisis.

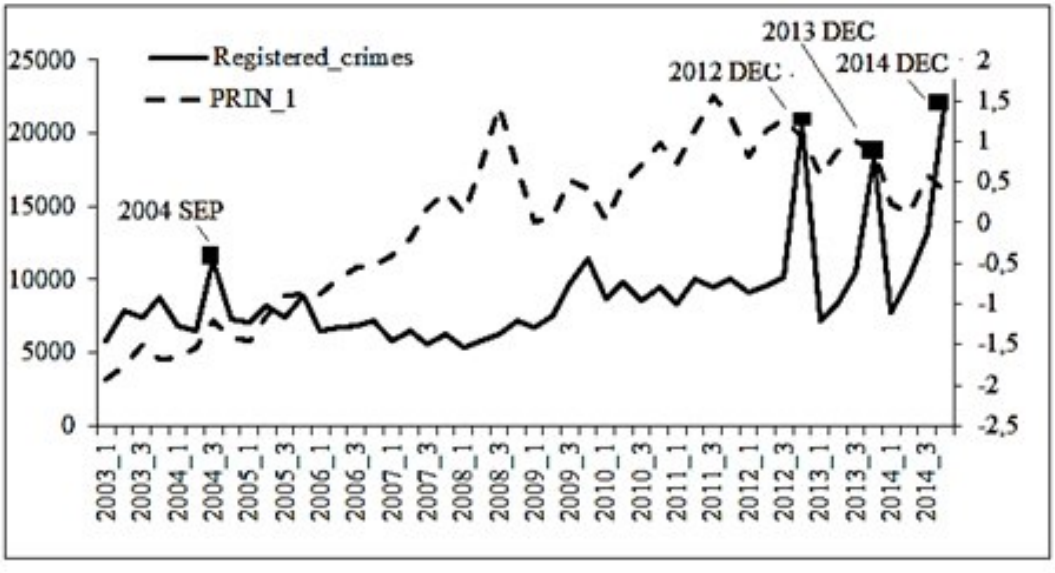


Figure 2. The horizontal axis corresponds to year-quarter, the right vertical axis corresponds to the first principal component and the left vertical axis corresponds to the total number of crimes in Kyiv.

Results. The results of modelling shows that the modern methods of data science and data mining allows find patterns among statistical data which describe crime and econometrical processes. In the next stages of research project NUKR.SFPP G4877 plans to build mathematical models use statistical except for structured and unstructured information from the Internet, and mix to models specially developed parameters and indexes based on the textual information.

References. **1.** Бідюк П.І. Прикладна статистика (навчальний посібник) / П.І. Бідюк, О.М., Терентьєв, Т.І. Просянкіна-Жарова. – Вінниця : ПП "ТД Едельвейс і К", 2013. – 288 с. –ISBN 978-966-2462-21-0. **2.** Терентьев А.Н., Домрачев В.Н., Костецкий Р.И. SAS BASE: Основы программирования (научное издание). – К: Эдельвейс, 2014. – 304 с. – ISBN 978-966-2748-49-9.

Trofimov A.S.[1,2] ***Zenkovich M.V.***[3]
[1]*SeVirt;* [2]*Moscow Power Engineering Institute (National Research University), Moscow, Russia;* [3]*National Research Nuclear University "MEPhI", Moscow, Russia*

Virtualization approaches in the area of fault-tolerant embedded systems

The paper gives introduction to existing conventional hardware (HW) based approaches of fault reservation in the area of high fault sensitive systems such as embedded equipment of automated vehicles, power plants, autonomous AI robots etc. The introduction section discuses existing limitations and disadvantages of conventional HW based approaches. Furthermore the main part of the paper describes possible non-conventional approach, which is based on the utilization of virtualization technologies. The main significance of proposed approach is considerable cost reduction in compare to multiple physical HW instances as in the conventional case. The main part of the paper also observes an example of two time reserved HW based and software (SW) based solutions. Since the complexity of nowadays embedded SW stack is very high virtualization approach gives ability to avoid fault or at least rapidly recover system state in a very short time using reserved isolated SW stack instance.

Introduction. Nowadays humanity progress more and more tightly joins with technologies created by people. Furthermore all technologies could be split into two categories. The first category could be called life-friendly and utilization of systems from this category could not be the cause of real feasible hazard to a human life. The second category could be called life-dangerous and utilization of systems from this category could potentially introduce real danger to a human life.

Fault tolerance [1] is an obligatory feature expected from life-dangerous systems. Of course the same could be applied for life-friendly systems but due to the nature of these systems we can avoid it and concentrate on life-dangerous systems. There is a set of system categories which are strictly require fault-tolerant lifecycle applied at all levels of the particular system architecture. Let's name some of them: air vehicles, cars, mechanical transport systems (such as elevators), various embedded systems, cellular, scientific and military related equipment.

Fault tolerance is a very important area of modern technologies. It relies on the top of three core characteristics of every particular fault-tolerant system [2], namely: safety, robustness and flexibility. Consequently each characteristic could be defined as a set of actions and measurements. Let's describe these characteristics. Safety could be defined as a condition of being protected against harmful conditions and events or the control of recognized hazards in order to achieve an acceptable level of risk. Robustness could be defined as the ability of the system to resist changes without adapting its initial stable configuration. Flexibility could be defined as the ability of the system to adapt when external changes occur.

This paper describes modern way of fault tolerant Hardware/Software system design through utilization of virtualization approach. It considers fault-tolerant system characteristics and finds optimal balanced solution.

Fault-tolerant systems. Fault-tolerance could be implemented in a few possible ways. The most common is redundancy scheme introduction [3]. It could be applied at the different levels of the system architecture. Basically figure 1 depicts system level redundancy and component level redundancy schemes. If the Component 1 would get in a fault state the Component 1' will replace the functionality of the Component 1. In the overall if the System 1 would get in a fault state the System 1' will replace the functionality of the System 1. As depicted at the figure 1

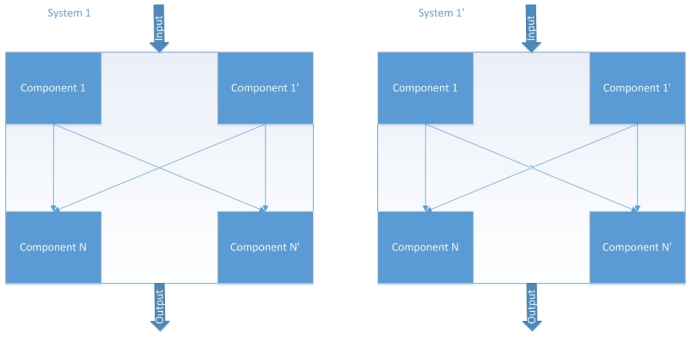


Figure 1. Redundancy design

redundancy scheme is implemented at different levels of the system, the important thing is that the redundancy scale should be various for various components. It means that components with the higher probability of fault should be covered with greater amount of redundant components and components with lesser probability of fault could be covered with only one fault recovery component.

It is need to highlight that redundant components could be in active and inactive state and this fact gives one more common way to detect and recover from the fault state. As depicted at the figure 2 systems could have the evaluator component, which is checking difference of computed result of possibly faulting component. In the worst case in the scheme introduced at figure 2 error would be detected (if there are only component and evaluator), in the best case error would be detected and recovered (if there are component, redundant component and evaluator).

Physical components redundancy is used in the most categories of fault-tolerant systems design, which introduce various disadvantages, which is needed to be taken into account during the system design. The main disadvantages of this approach are: increase complexity, increase weight, increase of power consumption etc.

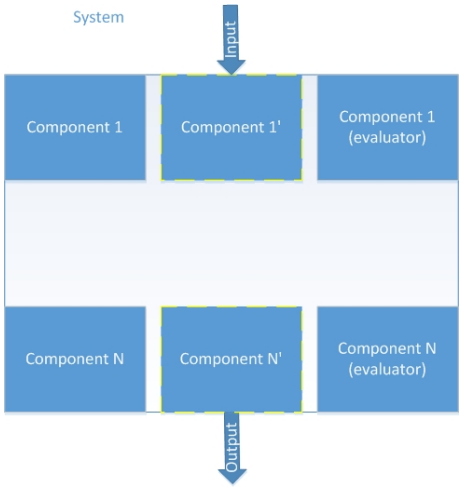


Figure 2. Error detection and recovery

Virtualization fault-tolerant systems. New approach considered in this paper proposes virtualization as a key for nowadays fault-tolerant system design. The figure 3 depicts basic concept of virtualization based fault-tolerant approach.

Hypervisor plays the role of key component in the virtualization fault-tolerant system design. It acts as a nano-kernel, strictly isolating and pinning algorithm partitions to a particular assigned CPU and system memory regions. At the same time it acts as a monitor of every partition's aliveness, which is giving one more point to the fault-tolerance.

This approach gets especially actual, when algorithm logic is mixing different processing security levels such as secure (like real-time high-priority activities) and insecure (like some entertainment possibly faulty things).

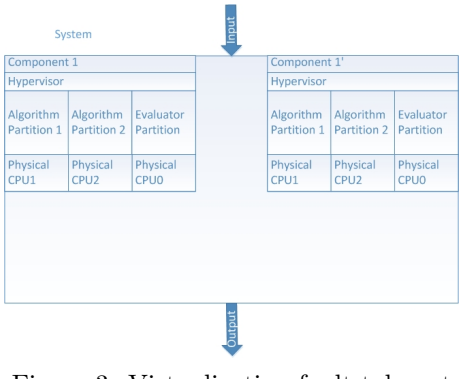


Figure 3. Virtualization fault-tolerant approach

Let's sum up the advantages and discuss possible improvements of proposed design. First of all it provides optimal combination of safety, robustness and flexibility due to the utilization of hypervisor (for instance the utilization of hypervisor provides the possibility of partial utilization of even broken HW). The second point – it reduces costs and increase integration level of the system's fault-tolerance due to the possibility of HW modules amount reduction. The third point – it could be implemented on the most system HW available on the market nowadays (the most common are Intel x86, ARM etc). The proposed approach may have one disadvantage, since the hypervisor part starts to be single point of failure it is necessary to adopt hypervisor code in order to be able to work with partially broken HW and this is the goal of the further research work.

References. **1.** Elena Dubrova.Fault-Tolerant Design, New York: Springer, 2013. **2.** Koren Israel and Krishna C. Mani. Fault-Tolerant Systems. San Francisco, CA: Morgan Kaufmann. 2007. **3.** Scott D. Sagan. Learning from Normal Accidents. Organization and Environment. March 2004.

Tsymbal M.D.
Faculty of Applied Mathematics, NTUU "KPI", Kyiv, Ukraine

Remote electronic identification methods

According to Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC [1], e-Id is one of key parts of electronic transactions. According to Final Report "eGovernment Benchmark Framework 2012-2015" [2] national citizen identity card (e-Id) can be used as a solution to implement electronic identification. To authenticate and gain access to sensitive data user should prove his identity is real and was not stolen. This task raises security problems to be solved.

Level of assurance. Usage of ISO/IEC 29115 Level of Assurance 3 (High) is reasonable to provide most secure way for remote citizen identification. It emphasizes usage of cryptically encrypted information transition and describes necessary measures for satisfying authentication.

ID card. Modern citizen ID cards come with cryptoprocessor onboard. This allows storing private key on citizen e-Id card while keeping public key in user certificate with his or her personal data and put certificate in one of card's data groups.

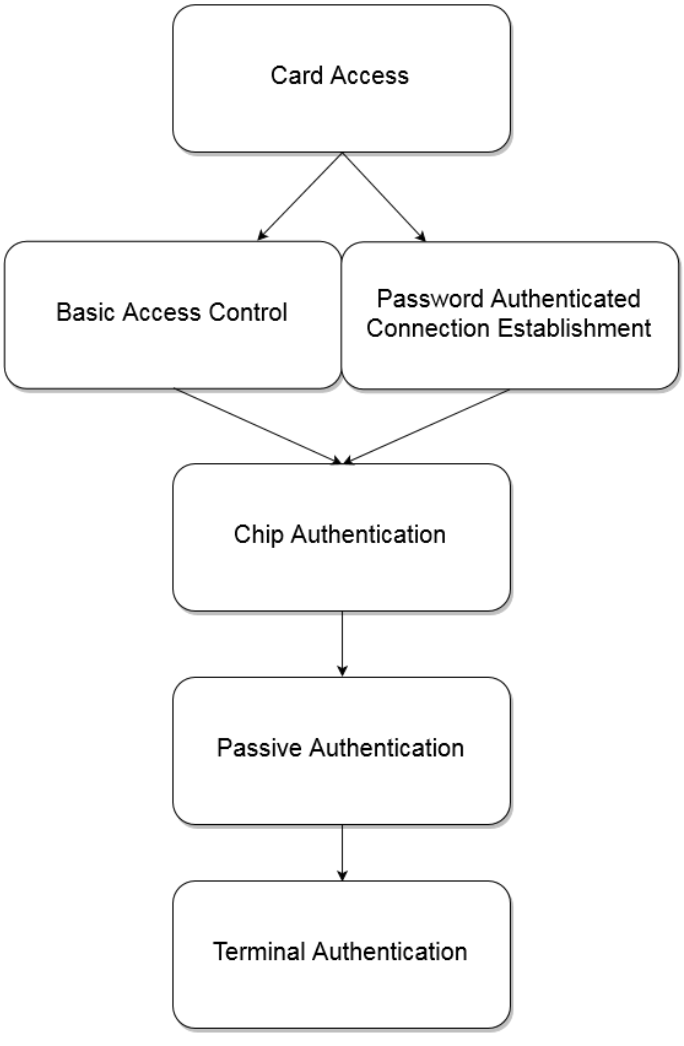


Figure 1. Card private data access algorithm

Identification process. To access card's private data user has to pass secure authentication process described in ICAO 9303 [3]. Such process is shown on figure 1. To prove identity on a web service user will go through OAuth2 with X.509 certificate [4]: user sends certificate to the server; server generates nonce and sends it to the user; user inputs PIN (every time user needs access data) and get access to own private key (private key itself does not ever leave card's chip); user signs received nonce with private key and sends signature back to the server; server verifies signature with certificate's public key; server validates certificate via CRL/OSCP server. As a result of such actions server both assures of certificate genuineness and user identity, thus completely identifying citizen. For additional security measures user may have connected phone number to which a text message would be sent to make certain users intent for authentication.

Risks. To gain access to another user's sensitive data malefactor would need to possess of that user's ID card, PIN and mobile phone. Cloning a card would not grant opportunity to pass verification since private key cannot be cloned to another card.

User-friendly process. Since describing procedure is expected to be performed by end user, identification process should not only be reliable but also user-friendly. This means server should provide one-click identification itself or use API of unified card processing software. Either way, it should be minimizing user's interference in the identification process.

Additional approach. Alternative look on matter presumes user does not need to read card every time, just provide service with card number. In such case user also needs authenticator – token, which provides one-time passwords to input with card number. In current state of computational power, such token is almost impossible to break. For additional security card number can be used as login with custom user password in pair, which is similar to regular login process, users are familiar with. However, such step may cause complications, as it requires user to remember and keep secret additional measure's value.

Hardware requirements. Using ID cards compels each user to poses a card reader to gain access to card digital information. Using SMS protection requires users to have a mobile phone. Alternative approach frees above requirements, however each user would have to be provided with authenticator instead.

ID card analogues. ID cards are commonly available. Also ID cards could be also replaced by either mobile SIM card or banking credit card, since both of above support cryptoprocessor potential.

Conclusions. Electronic identification (e-Id) is one part of electronic transactions. Secure implementation of e-Id schema must confirm Level of Assurance 3 ISO/IEC 29115 if it is used in sensitive online services. One of the common cases – usage of e-Id card with cryptoprocessor onboard, ICAO 9303 standards and common authentication protocols, such as OAuth brings Level of Assurance 3 ISO/IEC 29115 authentication schema.

References. **1.** Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures, 1999. **2.** eGovernment Benchmark Framework 2012-2015 (Method paper July 2012), European Union, 2012. **3.** Machine Readable Travel Documents, Seventh Edition, International Civil Aviation Organization specification Document 9303, 2015. **4.** OAuth 2.0 Certificate Authentication, msdn.microsoft.com/en-us/library/azure/hh127795.aspx.

***Арчвадзе Н.Н.*[1]; *Пховелишвили М.Г.*[2]; *Шецирули Л.Д.*[3]**
[1] *Тбилисский Государственный университет им.И.Джавахишвили, Тбилиси,Грузия;* [2] *Грузинский технический университет, институт вычислительной математики им. Н.Мусхелишвили;*
[3] *Батумский государственный университет им. Ш.Руставели*

Особенности параллельного программирования на языке Haskell

Haskell является языком функционального программирования, но был расширен до параллельной версии [1]. В настоящее время актуальность паралельных вычислений вызвано широко распространением многоядерных компьютеров. Под параллельными вычислениями можно понимать процессы решения задач, в которых в один и тот же момент времени могут выполняться одновременно несколько вычислительных операций.

Набор базовых операций Haskell-а изначально ориентирован на построение последовательных программ и лишь затем был расширен. Дополнительные ограничения накладывают также механизм «ленивых» вычислений и последовательная структура списков. Это приводит к изначальной ориентации на последовательные алгоритмы и не позволяет эффективно осуществить автоматическое распараллеливание программ. Несмотря на это Язык Haskell имеет средства для описания параллелизма, дополнительно введенные в качестве расширений уже существующей последовательной версии.

Использование неявного параллелизма позволяет упростить анализ механизмов управления в разрабатываемых программах[2]. Независимые функции можно вычислять в произвольном порядке или одновременно, так как данные, влияющие на окончательный результат, могут поступать только в качестве аргументов.

Рассмотрим особенности параллельного программирования на языке Haskell. Для организации параллельных вычислений в язык введены две конструкции, обеспечивающие явное управление. Комбинатор par, получает два параметра, являющиеся параллельно выполняемыми функциями. Выражение p ‘par‘ e имеет значение, определяемое e. Помимо этого, конструкция указывает на то, что p может быть выполнен новой нитью, параллельной с родительской нитью, продолжающей выполнение e. Поддержка «ленивых» вычислений накладывает свою специфику на параллелизм. Считается, что функция p может быть выполнена тогда, когда потребуется ее результат. В этом случае будет создана дополнительная нить.

Для синхронизации ранее инициированных параллельных вычислений введен оператор последовательной композиции, seq. Если e1 - не пустая конструкция, то выражение e1 ‘seq‘ e2 имеет значение e2; иначе результат – «пустое значение». Декларируется, что конструкция e1 должна быть обработана перед возвратом результата e2.

В качестве примера, использующего эти конструкции, рассмотрим следующее выражение:

```
import Control.Parallel
 main = a ‘par‘ b ‘par‘ c ‘pseq‘ print (a + b + c)
    where
        a = ack 3 10
        b = fac 42
        c = fib 34
```

В этом примере использованы известные функций: функция Аккермана, Числа Фибоначчи и факториал. Программа вычислила, что сумма равна: 1405006117752 87989854314260 624451156993638400571107б, время вычисления-95.46 secs.

В заключение следует отметить, что язык Haskell поддерживает только простейшие формы явного распараллеливания, реализованные еще в ранних императивных языках.

Литература. **1.** Simon Marlow. Parallel and Concurrent Programming in Haskell. 2013 by O'Reilly Media. **2.** Казаков Ф.А., Легалов А.И.Параллельное программирование в языках Haskell и Пифагор. **3.** N.Archvadze, M.Pkhovelishvili. Several Issues of Parallel Programs Verification for Functional Languages through Kripke Structure."Information and Computer Technologies, Modelling, Control". Proceedings. pp.545-547.

Балашевич Р.В., Гіоргізова-Гай В.Ш.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Забезпечення якості обслуговування трафіку VoIP у мережах кабельного телебачення

Сучасні комп'ютерні мережі все частіше використовуються не просто для обміну даними, а для передачі мультимедійних даних в реальному часі, що висуває нові вимоги до характеристик мереж.

Розширення кількості користувачів мережі Інтернет часто відбувається за допомогою використання вже існуючих мереж іншого призначення, наприклад, телефонних чи мереж кабельного телебачення.

Можливість забезпечення належної якості обслуговування (QoS) для VoIP та мультимедіа трафіку підвищує конкурентноздатність мереж кабельного телебачення у порівнянні з альтернативними технологіями і дозволяє їх операторам отримати додаткові прибутки від впровадження нових послуг, зокрема, IP телефонії.

Для роботи IP телефонії у відповідності із вимогами ITU-T мережеве обладнання повинне забезпечити:

- гарантії низької затримки;
- гарантії низького джиттеру;
- гарантії втрати пакетів менше 1% та відсутність втрат кількох послідовних пакетів.

Мережеве обладання в мережах з комутацією пакетів, що працює за принципом "Best effort" не може надати таких гарантій, тому необхідно застосовувати спеціальні механізми забезпечення якості обслуговування.

Для передачі даних в мережах кабельного телебачення використовується протокол канального рівня DOCSIS, який реалізує метод комутації пакетів. Але протокол DOCSIS підтримує розвинуті механізми QoS: класифікації пакетів, профілювання трафіку, підтримка різних класів сервісу (пріоритет, швидкість передачі, надійність, пульсації трафіку, тощо). Наявність різноманітних алгоритмів призначення тайм-слотів класам сервісів дозволяє зокрема для VoIP трафіку забезпечити постійну бітову швидкість передачі даних [1,2].

Стандарти PacketCable дозволяють інтегрувати механізми надання і управління QoS, що є у протоколі DOCSIS, в архітектуру VoIP мережі, визначають інтерфейси між компонентами VoIP мережі, надають можливості вирішення задач, пов'язаних з обліком і управлінням ресурсами мережі, а також з запобіганням неавторизованому використанню мережевих ресурсів [3].

Відповідно до рекомендацій стандарту PacketCable 1.5 було розроблено програмний агент для термінальної модемної станції (CMTS – Cable modem termination system), що забезпечує її взаємодію з головним компонентом VoIP мережі севером управління викликами (CMS – Call management server). Розроблене рішення орієнтовано на застосування у моделях CMTS, що використовують ОС Linux та процесори архітекур Intel x64 та ARMv8.

Тестування розробленого рішення з використанням застосування Asterisk у якості CMS та модемів Arris TM502 і Arris TM602 у якості (eMTA - Embedded multimedia terminal adapter) показало, що всі механізми QoS протоколу DOCSIS, які необхідні для якісного обслуговування VoIP сесій в мережах кабельного телебачення, були задіяні.

Література. **1.** Data Over Cable Service Interface Specifications 3.0 Physical layer, http://www.cablelabs.com/wp-content/uploads/specdocs/CM-SP-PHYv3.0-I08-090121.pdf

2. Data Over Cable Interface Secifications 3.0 Mac and Upper Layers Interface, http://www.cablelabs.com/wp-content/uploads/specdocs/CM-SP-MULPIv3.0-I01-060804.pdf

3. 3. PacketCable Dynamic Quality-of-Service Specification, http://www.cablelabs.com/wp-content/uploads/specdocs/PKT-SP-DQOS1.5-I03-070412.pdf.

Бондаренко А.С.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Использование мобильных устройств для мониторинга состояния человека с диабетом 1-го типа

Диабет 1-го типа является одним из самых распространенных хронический детских и юношеских заболеваний [2]. Причина этого типа диабета неизвестна, поэтому в настоящее время его нельзя предотвратить. При диабете типа 1 (ранее известном как инсулинозависимый), для которого характерна недостаточная выработка инсулина, необходимо ежедневное введение инсулина. [3]

Симптомы включают чрезмерное мочеотделение (полиурию), жажду (полидипсию), постоянное чувство голода, потерю веса, изменение зрения и усталость. Эти симптомы могут появиться внезапно [1].

Лечение диабета 1-го типа – это диета, уколы инсулина и физическая активность. Мобильные устройства и приложения могут помочь во всех 3 аспектах лечения.

Диета. Для контроля диеты человеку потребуется вводить данные о употребляемых продуктах питания. Обширная база блюд и продуктов поможет быстро осуществить ввод через поиск. Для каждого наименования продуктов и блюд необходимо будет указать содержание белков, жиров и углеводов, а также витаминов.

В зависимости от параметров человека врач определит требуемое количество питательных веществ. Мобильное приложение сможет уведомить пользователя о недостатке или избытке питательных веществ и витаминов.

Физическая активность. Мобильное устройство способно самостоятельно фиксировать передвижение человека, его скорость и продолжительность. Этого достаточно для вывода о расходе энергии человека в течении дня. Для более точных данных требуются специальные устройства.

Фитнес-браслеты способны предоставить информацию о частоте пульсации и температуре тела. С помощью этих данных можно сделать более точный вывод о нагрузках человека.

Уколы инсулина. Мобильное устройство может помочь контролировать уровень инсулина. Это стало возможным благодаря глюкометрам, оснащенным Bluetooth и audio-jack интерфейсами для связи с мобильными устройствами. Специальное мобильное приложение может собирать данные с глюкометров и производить анализ состояния пользователя, напоминать о необходимости укола инсулина. Также внедрение мобильных приложение для сбора информации о состоянии пациента позволяет персональному врачу удаленно контролировать показатели пользователя.

Мобильные устройства могут помочь в мониторинге состояния человека с диабетом 1-го типа. Они способны помочь в контроле над диетой, уровне физической активности и уровне инсулина. Все эти данные очень важны как для пациента, так и для его лечащего врача, для принятия правильных решений о лечении.

Носимые устройства способны обеспечить непрерывный мониторинг состояния человека. Это открывает возможности для глубокого анализа медицинских данных и выявления причин изменения состояния человека.

Литература. **1.** Диабет. Информационный бюллетень №312. Январь 2015 г. http://www.who.int/mediacentre/factsheets/fs312/ru. **2.** Improving Diabetes Care for Young People With Type 1 Diabetes Through Visual Learning on Mobile Phones: Mixed-Methods Study Dag Helge Frøisland, Eirik Årsand, Finn Skårderud J Med Internet Res 2012 (Aug 06) http://www.jmir.org/2012/4/e111/. **3.** Пхакадзе А.Г., Неборачко М.В. Регулярный контроль – основа долгой жизни без осложнений сахарного диабета. Украинский научно-практический центр эндокринной хирургии, трансплантации эндокринных органов и тканей МОЗ Украины Киев, Украина.

***Боярінова Ю.Є.*[1], *Каліновський Я.О.*[2], *Сукало А.С.*[2]**
[1]*НТУУ "КПІ", Київ, Україна;* [2]*ІПРІ НАН України, Київ, Україна*

Застосування узагальнених кватерніонів у цифровому підпису

Запрпоновано використання гіперкомплексної числової системи четвертої вимірності - узагальнених кватерніонів для побудови алгоритму цифрового підпису.

Електронний документообіг істотно скорочує об'єми паперової документації, при цьому вони можуть бути підписані електронним цифровим підписом (ЕЦП) і передані за призначенням протягом декількох секунд. Для підвищення стійкості ЕЦП можна використовувати або збільшення довжини ключа, або удосконалений спосіб представлення даних [1]. Одним з таких способів є представлення даних за допомогою гіперкомплексних числових систем (ГЧС)[2]. Так застосування ГЧС другої вимірності в алгоритмі RSA забезпечило більшу стійкість в порівнянні з алгоритмом, заснованим на операціях з дійсними числами [2].Для використання представлення даних з ГЧС в алгоритмі RSA необхідно розглянути ппредставлення функцій гіперкомплексні змінної [3]. Тобто побудувати вирази для кодування та декодування з використанням наступної властивості експоненціальної функції:

$$a^x = exp(x * lna) \tag{1}$$

Таким чином, щоб синтезувати алгоритм RSA з використанням ГЧС слід побудувати представлення експоненціальної і логарифмічної функцій в цих системах. Розглянемо клас некомутативних ГЧС четвертої вимірності, що складається з некомутативних подвоєнь ГЧС другої вимірності, одержаних за допомогою процедури подвоєння Келі-Діксона [4]. Як показали дослідження, цей клас ГЧС складається з дев'яти ГЧС:Д(C,C,4)=Н-система кватерніонів;Д(C,W,4)=АН-система антикватерніонів та ГЧС, одержаних подвоєнням систем другої вимірності у всіяких комбінаціях: Д(C,D,4),Д(D,C,4), Д(W,W,4), Д(D,D,4),Д(W,D,4), Д(D,W,4), Д(W,C,4).

Для усіх цих систем можна побудувати представлення норми, експоненти, логарифму. Представлення експоненти від гіперкомплексного числа M є частковий розв'язок гіперкомплексного лінійного рівняння з правою частиною dX/dt=MX з початковою умовою exp(0)=e1. Так, наприклад, представлення експоненти для кватерніонів має вигляд:

$$Exp(M) = e^{m_1} * (cos\bar{m} * e_1 + \frac{m_2 * e_2 + m_3 * e_3 + m_4 * e_4}{\bar{m}}) \tag{2}$$

де

$$\bar{m} = \sqrt{m_2^2 + m_3^2 + m_4^2} \tag{3}$$

Представлення логарифмічної функції можуть бути побудовані, оберненням експоненціальних функцій. Так, наприклад, показано, що головне значення логарифмічною функції від кватерніона має представлення

$$Ln(X) = ln\,|x| + \frac{1}{\bar{x}} * arccos\frac{1}{|x|} * (x_2e_2 + x_3e_3 + x_4e_4) \tag{4}$$

Такими ж методами побудовані представлення функцій і для інших ГЧС четвертої вимірності цього класу, що підтверджує можливість побудови алгоритму цифрового підпису з використанням узагальнених кватерніонів.

Література. **1.** Stinson D. R. Cryptography: Theory and Practice / D. R. Stinson. – Chapman and Hall, 2006.-593c. **2.** Y. O. Kalinovsky Applying of hypercomplex number systems to RSA-algorithms/ Y. O. Kalinovsky, Y. E. Boyarinova, I. V. Khitsko//6-th International Conference "Advanced Computer Systems and Networks: Design and Application"ACSN-2013. Львів: НВФ "Українські технології 2013 - P. 114-115 **3.** Синьков М.В. Конечномерные гиперкомплексные числовые системы. Основы теории. Применения / М.В. Синьков, Ю.Е. Бояринова, Я.А. Калиновский. – К.: Инфодрук, 2010.- 388с **4.** Kalinovsky Y. O. Computing Characteristics of One Class of Non-commutative Hypercomplex Number Systems of 4-dimension/ Y. O. Kalinovsky, D. V. Lande, Y. E. Boyarinova, A.S. Turenko. – http://arxiv.org/ftp/arxiv/papers/1409/1409.3193.pdf.

Бугаєва Л.М., Безносик Ю.О.
Хіміко-технологічний факультет НТУУ "КПІ", Київ, Україна

Програмні засоби обчислювальної гідродинаміки для розв'язання задач хімічної інженерії

Використання засобів, відомих як CFD- програми (computational fluid dynamics) в наш час все стає ширшим в різних галузях науки та техніки, що стикаються із необхідністю вирішувати задачі із розподіленими параметрами. Зазвичай, ці програми мають власні засоби формування геометрії об'єктів та здатні імпортувати образи обєктів, створених у різних CAD-системах. Важливо, що такі програми дають можливість вирішувати задачі просторового моделювання навіть інженерам без глибинних знань математичної фізики завдяки наявності бібліотек моделей процесів. В даній роботі авторами представлено приклади рішення декількох задач фізики та хімічної інженерії.

Обчислювальна гідродинаміка більш відома англійською як computational fluid dynamics (CFD) стає в наш час програмним засобом, що можуть використовувати не тільки науковці-математики, а й звичайні інженери для рішення технічних проблем, повязаних із необхідністю визначення розповсюдження деяких характеристик у просторі і часі. Найпростіший приклад, це розповсюдження забруднення від точкового джерела, що описується рівнянням дісперсії, або зміна температури в деякому об'ємі складної форми. CFD або обчислювальна гідродинаміка – це могутній наукомісткий апарат, що зв'язує в собі такі дисципліни як гідродинаміка, чисельний аналіз, теорія диференціальних рівнянь у часткових похідних, обчислювальна геометрія й комп'ютерні науки. CFD пропонує кінцево-елементну методику для чисельного рішення систем диференціальних рівнянь у часткових похідних, що звичайно описують динаміку потоку рідини або газу. В наш час CFD реалізовано у достатній кількості програмних продуктів, серед яких найвідоміші ANSYS, Solidworks, COMSOL та ін. Метою авторів роботи було на власному досвіді показати ефективність використання CFD програм, як для наукових досліджень, так і в навчальному процесі.

Разом із науковцями Інституту фізики НАНУ авторами було проведено дослідження функционування отолітового органу живих істот, який відповідає за просторову орієнтацію та поведінку, оскільки сприймає гравитацію [1]. Модель просторово неоднорідної поверхні отоліта було побудовано в програмному середовищі Solidworks. Для вирішення задачі розподілення зміщень статичних навантажень – гравітації та відцентрової сили був використаний програмний пакет COMSOL Myltiphysics 4.0. Слід відзначити, що при рішенні цієї проблеми використовувався класичний підхід до задач математичної фізики: обиралась математична модель та формувались граничні умови, після чого проводились розрахунки при різних налаштуваннях програм. Отримані результати моделювання доповнили теоретичні та експериментальні відомості про поведінку отоліту в умовах невагомості.

Також авторами було проведено моделювання та дослідження аеродинаміки процесів в топці котла із низько киплячим шаром в рамках наукового співробітництва із Інститутом теплофізики НАНУ [2–4]. Для моделювання котельної установки авторами була використана програма SolidWorks.

Рух середовища моделювався рівняннями Нав'є-Стокса, що описують в нестаціонарній постановці закони збереження маси, імпульсу і енергії цього середовища. Для моделювання турбулентних течій рівняння Нав'є-Стокса осереднюються по Рейнольдсу, тобто використовується осереднений по малому масштабом часу вплив турбулентності на параметри потоку, а великомасштабні часові зміни усереднені по малому за масштабом часу складовими газодинамічних параметрів потоку (тиску, швидкостей, температури). У результаті рівняння мають додаткові члени – напруги по Рейнольдсу, а для повноти цієї системи рівнянь використовуються рівняння переносу кінетичної енергії турбулентності та її дисипації у рамках κ-ε моделі турбулентності. Проведені розрахунки, а також чисельне моделювання аеродинаміки на

основі рівнянь Нав'є-Стокса за допомогою програми SolidWorks дозволило визначити значення основних режимних характеристик та ефективність роботи топки, отримати аеродинамічну картину поведінки газового потоку в топці котла.

Оскільки автори представляють кафедру кібернетики хіміко-технологічних процесів, то найбільший інтерес із розглянутих CFD програм для рішення дослідницьких задач та використання у навчальному процесі представляє програма COMSOL. На відміну від інших таких програмних засобів ця програма має проблемно-орієнтовані бібліотеки моделей, а саме найцікавішим є наявність бібліотеки хімічних моделей, а також підпрограми Comsol Reaction Engineering, яка дає змогу знаходити рішення кінетичних моделей хімічних реакцій.

В середовищі Comsol можливе моделювання як ізотермічних, так і інших реакторів. При розрахунку можливе врахування багатьох, необхідних для коректного рішення задачі параметрів: об'єму реактора, температури проведення реакції та ін. Реальні реактори можуть моделюватись як комбінації ідеальних реакторів. В рамках лабораторної роботи студенти навчаються використовувати Comsol на прикладі двох ідеальних реакторів безперервного перемішування з обміном, які використовуються для моделювання реального реактора з двома зонами – сильного та слабкого перемішування. Для цього необхідні два параметри, які взаємопов'язані між собою – об'єм та швидкість обміну двох областей. Ці параметри знаходяться шляхом співставлення модельних результатів з експериментальними даними. В даній лабораторній роботі студентами вирішується задача хімічної технології засобами Comsol Reaction Engineering Lab із прикладами з бібліотеки пакету, в результаті чого робиться оцінка параметрів моделі неідеального реактора. Результати, отримані за моделлю та експериментальні дані концентрації трасера у вихідному потоці реактора приблизно однакові. Наявні експериментальні дані порівнюються з даними, отриманими за моделлю в середовищі Comsol. Цей досвід дослідження у середовищі Comsol студенти в подальшому використовують при роботі над дипломом або магістерською дисертацією.

Таким чином, підсумовуючи проведену авторами роботу по впровадженю програмних засобів CFD в наукові дослідження та навчальний процес, можна зробити висновок, що найбільш привабливими для досліджень в інженерії є програми Comsol та SolidWorks.

Література. **1.** Кондрачук, О.В. Кінцево-елементне моделювання механіки отоліта засобами пакету COMSOL MULTIPHYSICS / О.В. Кондрачук, Л.М Бугаєва., І.І. Бочкор // Восточно-Европейский журнал передовых технологий – 2013. - № 3/7 (63) – с. 10-15. **2.** Бугаєва, Л.М. Дослідження котельної установки киплячого шару на базі натурних та комп'ютерних експериментів / Л.М. Бугаєва, Д.О. Серебрянський, А.А. Дейкун // Восточно-Европейский журнал передовых технологий – 2011.- 6/6(54) – с. 22-26. **3.** Безносик, Ю.А. Исследования сжигания твердого топлива в низкотемпературном кипящем слое с определением выбросов токсических веществ / Ю.А. Безносик, В.А. Логвин, К.А. Коринчук, Д.А. Киржнер // Химическая промышленность – 2014. – том 91, № 1. – с. 15-19. **4.** Безносик, Ю.А. Разработка и исследование циклофильтра для очистки промышленных газов / Ю.А. Безносик, С.В. Плашихин, Л.Н. Бугаева, А.Н. Набок // Химическая промышленность, 2014. – том 91, № 1. – с. 41-46.

Будник С.А.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Який вибір ми зробимо: HTML 5 чи Flash?

Мультимедійна платформа Flash, випущена в 1996 році компанією Macromedia і пізніше придбана Adobe, в двохтисячних стала стандартом де-факто для розробки засобів програвання відео, анімованих банерів та інтерактивних веб-сторінок. Однак з появою нового стандарту HTML5, справи Flash йдуть вже не так добре. Багато експертів впевнені в тому, що незабаром ця технологія потерпить повної поразки, і абсолютним лідером стане HTML5. Еріка Траутман, CEO компанії Traut Media, спеціалізируется на виробництві інтерактивних відео, вважає, що закінчення славної історії Flash варто очікувати вже найближчим часом [1]. Та чи так це?

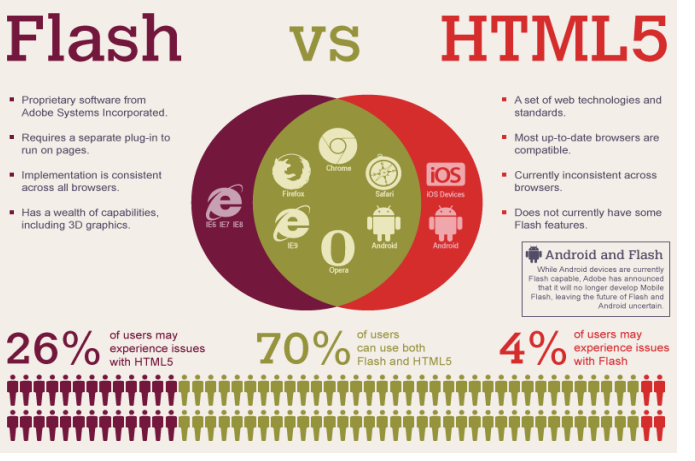


Рис. 1. Flash vs HTML5

Порівнювати ми будемо за декількома критеріями:

- Підтримка сучасними браузерами
- Файловий розподіл та структура
- Мобільні пристрої
- Ігри

Підтримка сучасними браузерами. На даний момент, 99% десктопних браузерів підтримують Flash Player, та 82% підтримують HTML 5 Canvas. Але за даними, наведеними на Рис.2 бачимо, що більшість використовує новітні версії браузерів для настільного комп'ютера, тож можемо вважати нічию у цьому пункті [2].

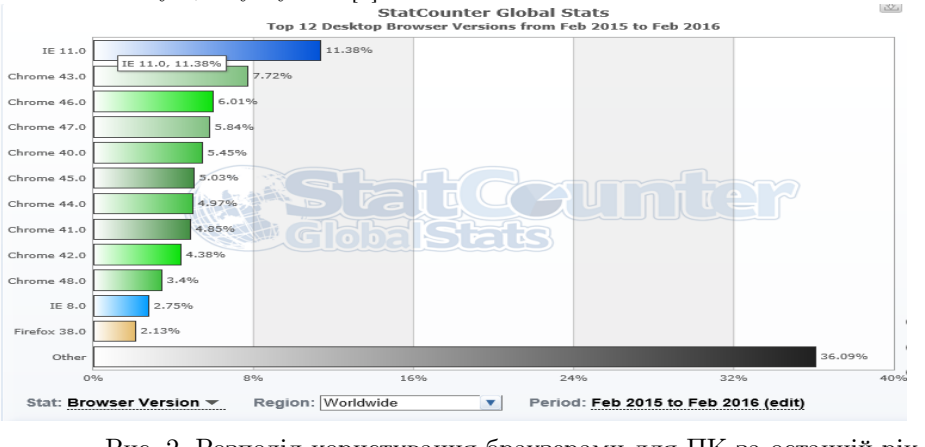


Рис. 2. Розподіл користування браузерами для ПК за останній рік [3].

Файловий розподіл та структура. Для Flash технології буде достатнім один файл з розширенням swf. Це дуже важливо для ігр і зручно для зберігання чи транспортування. Що стосується HTML5, то тут за самою природою HTML відбувається розділення коду у різні файли. Тому треба бути обачним з їхнім розташуванням та вказанням вірного шляху до них. Хоча, з іншого боку, кожен окремий файл виконує свою функцію, і при редагуванні частини коду є можливість ввести поправки, не хвилюючись за весь проект в цілому. Також сама структура нового стандарту дозволяє пошуковим системам заглядати всередину створених з його допомогою сайтів і додатків, чого не можна сказати про Flash. За допомогою HTML5 розробники можуть створювати проекти, що складаються з безлічі взаємозалежних частин, і тепер пошукові роботи вміють розуміти ці зв'язки. А проект на Flash для роботів Google представляє з себе незрозумілий для них вміст. Тому у даному пункті невелику перемогу здобув HTML5.

Мобільні пристрої. З моменту релізу HTML5 однією з головних переваг нової мови вважається її здатність працювати на мобільних пристроях. Той факт, що багато Android і iOS-пристроїв не підтримують Flash (97% мобільних браузерів підтримують HTML5 Canvas, а у Flash цей показник 0%), дозволяє використовувати його тільки на PC - цей ринок в 2013 році впав на 10 відсотків у порівнянні з 2012 [4]. За інформацією Бюро інтерактивної реклами, практично кожен другий житель США є власником мобільного телефону з доступом в інтернет, і кожен п'ятий перегляд веб-сторінок відбувається з мобільного пристрою. Цифри зростають щомісяця, і компанії, що створюють інтерактивні відео з використанням технології Flash, автоматично втрачають величезну аудиторію, яка могла б переглядати цей контент на своїх смартфонах і планшетах [5]. У даному пункті перемогу здобув HTML5.

Ігри. За інформацією блога Digital Buzz, 32% часу, витраченого користувачами мобільних пристроїв, йде на ігри. За допомогою HTML5 розробники зможуть створювати іграшки, які будуть працювати на всіх девайсах. Згідно зі звітом розробника open-source фреймворків для створення веб-додатків Sencha, понад 60% розробників повністю або частково перевели свої проекти на HTML5, і більше 70% з них заявили про те, що в цьому році користуються даною технологією більше і частіше ніж в минулому [6]. Тож в данному пункті також перемагає HTML5.

Висновок. Було досліджено два найбільш популярних стандарти та мультимедійні платформи: Flash та HTML5. Було розглянуто їхні недоліки та переваги. Виходячи з результатів аналізу HTML5 є кращим вибором, якщо дивитись наперед. Бо творці контенту, які не використовують у своїй роботі HTML5 вже відстають від тих, хто це роблять, і скоро остаточно програють. Ті, хто все ще вважають Flash відмінним вибором для онлайн-відео та ігор, в найближчому майбутньому повинні будуть змінити свою думку, і до цього їх підштовхне все триваюче проникнення мобільних пристроїв і подальше поширення HTML5.

Література. **1.** RIP Flash: Why HTML5 will finally take over video and the Web this year, [Електроний ресурс] URL: http://thenextweb.com/dd/2014/04/19/rip-flash-html5-will-take-video-web-year/ **2.** Flash vs HTML5, [Електроний ресурс] URL: http://flashvhtml.com/ **3.** Top 12 browser versions from 2015 to 2016,[Електроний ресурс] URL: http://gs.statcounter.com/#desktop-browser_version-ww-monthly-201502-201602-bar **4.** Worldwide PC Shipments Declined 6.9 Percent in Fourth Quarter of 2013, [Електроний ресурс] URL: http://www.gartner.com/newsroom/id/2647517 **5.** Every 5-th user uses mobile version , [Електроний ресурс] URL: http://www.iab.net/iabcom_launch?originalurl=/mobilecenter/makemobilework/letter **6.** Infographic: 2013 Mobile Growth Statistics , [Електроний ресурс] URL: http://www.digitalbuzzblog.com/infographic-2013-mobile-growth-statistics/

Буценко Ю.П., Савченко Ю.Г., Семенов С.В.
Факультет електроніки НТУУ "КПІ", Київ, Україна

Алгоритми формування послідовностей псевдовипадкових чисел

На сьогодні використання послідовностей псевдовипадкових чисел (ППВЧ) є доволі поширеним у багатьох процедурах криптографічного та стеганографічного захисту інформаційних ресурсів, сигнатурного діагностування цифрових пристроїв, системах зв'язку за технологією CDMA тощо. У всіх цих та інших випадках виникає актуальне питання щодо оцінки якості використовуваних ППВЧ, оскільки непередбачуваність послідовності (а це на змістовному рівні і є ключовим показником якості) визначає стійкість крипто- або стеганографічного захисту до зламу, достовірність сигнатурного методу контролю та ефективність багатьох сучасних телекомунікаційних систем.

Традиційно і до цього часу найбільш поширеним засобом формування ППВЧ залишаються регістри зсуву із зворотними зв'язками по модулю 2 або LFSR (Linear feedback shift register). Однак ППВЧ, що генеруються такими засобами, досить легко можуть бути передбачувані сучасними методами криптоаналізу (наприклад, [1]).

Пропонується для генерації ППВЧ застосувати універсальні інструменти на основі алгоритмів обчислення вихідних сигналів у скінченних цифрових автоматів. У цьому випадку функції утворення кожного двійкового числа послідовності при використанні моделі Мура можна записати у вигляді `equation`:

$$y_i(t+1) = f_i[s_1(t), s_2(t), ..., s_n(t)], \qquad i = 1, 2, ..., n, \tag{1}$$

де $y_i(t+1)$ - значення i-го біта в момент часу $(t+1)$; $s_1(t), s_2(t), ..., s_n(t)$ - двійкове число, яке відповідає внутрішньому стану автомата в попередній момент часу t; а $f_i, i = 1, 2, ..., n$ - булеві функції перходу автомата із стана в стан.

Для класичного випадку застосування LFSR функції переходу мають вигляд `equation`:

$$y_i(t+1) = \oplus \sum_{j \subset B} b_j s_j(t), y_2(t+1) = y_1(t); y_3(t+1) = y_2; ...; y_n(t+1) = y_{n-1}(t), \tag{2}$$

де коефіцієнти $b_j, j = 1, 2, .., n$ визначають ті біти, які підсумовуються в колі зворотного зв'язку по модулю 2. Конкретна множина цих бітів вибирається відповідно до степенів неприводимих (примітивних) поліномів та деяких інших обмежень, що гарантує отримання ППВЧ максимальної довжини без повторення чисел.

Зазначимо дві особливості такого алгоритму формування ППВЧ:

- кожне наступне число послідовності однозначно визначається попереднім, що суттєво зменшує складність криптоаналізу;
- повний опис алгоритму формування ППВЧ і, фактично, конкретного генератора ППВЧ (апаратного чи програмного) доволі короткий – це n бінарних коефіцієнтів b_j та n-бітове стартове слово, з якого починається генерація.

Саме друга особливість є принциповою та визначальною для оцінки якості ППВЧ. Відповідно до фундаментального постулату А.М.Колгоморова [2, 3] довжина опису алгоритму генерації ППВЧ еквівалентна її якості. Тобто отримати більш якісну ППВЧ можна лише, ускладнюючи фактично алгоритм її формування, та, як наслідок, опис такого алгоритму за довжиною, наприклад, в бінарних символах має збільшуватись за довжиною. Таке припущення виглядає досить парадоксальним, але до цього часу не було спростовано практикою реалізації генераторів ППВЧ. Якщо ж розглядати проблему якості генератора та генерованих за його допомогою послідовностей не суто формально, а на змістовному рівні, то проблема еквівалентна відповіді на досить просте питання: легко чи важко криптоаналітику знайти алгоритм, за яким утворена послідовність? Зрозуміло, чим складніший алгоритм, за яким функціонує генератор, тим важче передбачити результат його функціонування.

Першим і очевидним кроком в ускладненні алгоритму є ускладнення функцій переходів (1). Навіть не виходячи за клас регістрів зсуву (всі функції окрім $f_1(t+1)$ залишаються без змін),

можна суттєво утруднити криптоаналіз. Для цього достатньо для функції $f_1(t+1)$ вийти за межі класу лінійних булевих функцій. У випадку генерації n-розрядних двійкових чисел криптоаналітику доведеться вже перебрати не $2n$ варіантів, як це було для LFSR, а майже всі можливі булеві функції від n змінних. А загальне число таких функцій складає фантастично велике значення, а саме $M = 2^{n^2}$. Звичайно, опис конкретного генератора – це визначення лише однієї булевої функції та стартового слова.

Наступним і теж очевидним кроком є перехід від класу регістрів зсуву до довільних цифрових автоматів моделі Мура. У цьому випадку всі функції переходів стають довільними булевими функціями, а опис алгоритму формування ППВЧ подовжується в раз. Зазначимо, що коректна оцінка довжини опису повинна враховувати непридатність для використання деякої частини функцій (константи 0 та 1, вироджені функції тощо).

Таким чином, опис конкретного генератора повинен містити n описів конкретних булевих функцій від n двійкових змінних, а це у загальному випадку їх таблиці істинності. Кожна така таблиця може бути описана множиною наборів, на яких вона дорівнює 1 (або еквівалентного запису "по нулям"), а довжина опису для однієї функції складатиме максимум 2^{n-1} бітів. Для всіх функцій генератора це буде . До цієї величини необхідно додати величину вектора ініціалізації, тобто початковий внутрішній стан автомату та стартове слово, з якого починається генерація. Таким чином, повний опис алгоритму генерації конкретної послідовності при використанні універсальної автоматної моделі Мура складає: `equation`:

$$M = n + n2^{n-1}\text{біт}. \tag{3}$$

При апаратній реалізації генератора та типових значеннях $n = 12...16$ обсяг опису складатиме приблизно $10^4...10^5$ біт у порівнянні із реалізацією на LFSR, де він складає $24...32$ біти.

Слід зазначити, що використання обсягу опису алгоритму генерації послідовності ПВВЧ для визначення якості послідовності залишається дискусійним і на практиці не конкурує із використанням тестів NIST [4]. В той же час, виходячи з принципу Кірхофса, згідно з яким криптоаналітику відома загальна схема алгоритму генерації, вже на інтуїтивному рівні очевидно, що трудомісткість криптоаналізу при застосуванні LFSR та, у другому варіанті при гіпотетичному використанні скінченого цифрового автомата моделі Мура розрізняються на порядки. Так у першому випадку можуть бути застосовані досить ефективні алгебраїчні атаки, наприклад [1], а у другому – скоріш за все, лише перебір логічних функцій та їх комбінацій. Звичайно, множини функцій переходів, які доцільно використовувати в конкретних реалізаціях генераторів, потребують додаткового моделювання та перевірки на придатність, але кількість можливих варіантів навіть після такої «фільтрації» залишається песимістичною з точки зору часових витрат. Подальшим розвитком пропонованого підходу може бути перехід до моделі Мілі, що потенційно дає можливість практично без реальних обмежень збільшувати довжину генерованої послідовності.

Література. **1.** Пометун С.О., Алгебраїчні атаки на потокові шифратори як узагальнення кореляційних атак./– С.О.Пометун // Системні дослідження та інформаційні технології,– 2008, №2, , с. 29-40 **2.** Колмогоров А.Н., Три подхода к определению понятия «количество информации» / А.Н.Колмогоров //Проблемы передачи информации,– 1964 –Т. 1(1), с. 3-11 **3.** Вьюгин В.В., Колгоморовская сложность и алгоритмическая случайность / –М.: МФТИ, 2012, –131 с. **4.** Потий А., Орлова С., Статистическое тестирование генераторов случайных и псевдослучайных чисел с использованием набора статистических тестов NIST STS. / А.Потий, С.Орлова – Правове, нормативне та метрологічне забезпечення систем захисту інформації в Україны – 2001,–вип. 2, , с.206-214

Виклюк Я.І., Гусак О.М.
Приватний вищий навчальний заклад "Буковинський університет"м.Чернівці, Україна

Методи і засоби підвищення інформативності даних дистанційного зондування Землі

Актуальність. З розвитком космічних та авіаційних технологій актуальним стає вивчення та моніторинг лісових пожеж з повітря та з космосу. Одержані дані характеризуються високою оперативністю, достовірністю і можуть бути використані для підтримки прийняття рішень фахівцями під час керування процесами гасіння лісових пожеж. Однак на знімках важко безпомилково виокремити джерело запалювання, визначити масштаб пожежі, оцінити площу вигорілих ділянок, оскільки космічні та аерофотознімки містять шуми і завади, обумовлені складними умовами отримання, недосконалістю систем передачі відеоінформації тощо.

Метою даної роботи є розробка засобів та методів виокремлення епіцентрів лісових пожеж на космічних та аерофотознімках.

Космічні та аерофотознімки, відео сюжети є цифровими зображеннями і характеризуються високою розділовою здатністю. Тому для них доцільно застосовувати технології Big Data [1].

Запропонована технологія отримання зображень на яких добре помітні потенційно небезпечні ділянки в основі якої лежать наступні методи: метод вирівнювання гістограм, вейвлет – аналіз, кластерізація методами k-means, c-means, мережі Кохонена. Програмна реалізація була здійснена за допомогою платформи MatLab [2], [3]. Як видно з рисунку, розроблена програма дозволяє безпомилково виявляти джерела запалювання при різних умовах:

- на супутникових знімках з великої висоти (Рис.1.a, b);
- з квадрокоптеру в умовах відсутності лісового покрову (Рис. 1. c, d);
- на знімках лісової гущавини (Рис. 1. e, f).

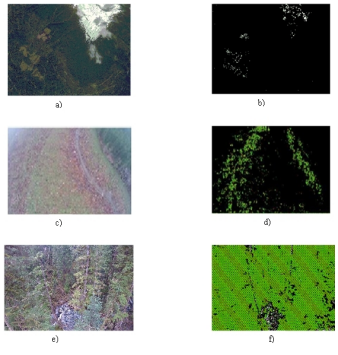


Рис. 1

Висновки.

1. Розроблена методологія кластерізації первинних зображень виявилася ефективною для розпізнавання епіцентрів лісових пожеж та виявлення вигорілих ділянок лісу.
2. Описані методи можуть знайти застосування в медицині для аналізу знімків і виявлення пухлин на ранніх стадіях, в арктичній навігації, в сфері безпеки для пошуку підозрілих предметів під час сканування багажу в аеропортах.

Література. **1.** Гонсалес, Р. Цифровая обработка изображений (Digital Image Processing) [Текст] / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2012. – 1104 **2.** Дьяконов В.П. Matlab. Обработка сигналов и изображений. Специальный справочник / Дьяконов В.П.. – СПб: Питер. – 608 с. **3.** Журавель И. М. Краткий курс теории обработки изображений [Електронний ресурс] / И. М. Журавель. – 2015. – Режим доступу до ресурсу: http://matlab.exponenta.ru/.

Галіцька В.В.
Факультет інформатики та обчислювальної техніки НТУУ "КПІ", Київ, Україна

Алгоритм штучної бджолиної колонії для задачі комівояжера

Математична постановка задачі комівояжера. Основна ідея задачі комівояжера [1] полягає у наступному: дано множину міст, а також відстань між усіма можливими парами міст. Необхідно знайти шлях, який пролягає через усі міста та повертається у початкове, окрім того сумарна довжина пройденого шляху має бути мінімальною. Тобто необхідно побудувати маршрут таким чином, щоб побувати в кожному місті по одному разу і повернутися у початкове.

Математична постановка задачі комівояжера має наступний вигляд:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min, \tag{1}$$

$$\sum_{j=1}^{n} x_{ij} = 1; (i = 1, 2, .., n), \tag{2}$$

$$\sum_{i=1}^{n} x_{ij} = 1; (j = 1, 2, .., n). \tag{3}$$

де c_{ij} — матриця відстаней між усіма містами, а n - кількість міст. (2) і (3) виступають у якості обмежень на одноразовий виїзд та в'їзд у місто відповідно.

Якщо в моделі задачі обмежитися лише умовами (2) і (3), то вона буде еквівалентною задачі про призначення, план якої не обов'язково повинен бути циклічним. Тобто, маршрут комівояжера може розпастися на декілька незв'язних між собою циклів, тоді як насправді він повинен складатися з одного циклу. Щоб забезпечити цю вимогу введемо наступне обмеження:

$$u_i - u_j + nx_{ij} \leqslant n - 1; i, j = 1, 2, .., n; i \neq j \tag{4}$$

Тобто, якщо комівояжер пересувається з i-го в j-те місто, то нерівність (4) фіксує порядкові номери цих міст. Отже математична постановка задачі комівояжера полягає у мінімізації функції (1) при обмеженнях (2), (3) і (4).

Бджолині алгоритми. Бджолині алгоритми, як і більшість комбінаторних методів, були вигадані на основі спостережень за живою природою. Як можна зрозуміти з назви, алгоритми засновуються на способі існування бджолиної колонії. У найбільш поширеному своєму варіанті бджолиний алгоритм оперує бджолами двох типів: розвідники та робочі. Розвідники шукають квіткові поля, а робочі бджоли, відповідно, збирають з цих полів нектар.

При переході від живої природи до комбінаторної оптимізації можна стверджувати, що бджоли-розвідники шукають сприятливі точки у просторі рішень поставленої задачі, а бджоли-робочі намагаються розширити ці локальні області з надією покращити поточні результати.

Найбільш розповсюдженими є дві метаевристики: оптимізація бджолиної колонії [2–5] та алгоритм штучної бджолиної колонії [6, 7].

Алгоритм штучної бджолиної колонії. ABC базується на симуляції поведінки медоносних бджіл під час пошуку їжі. Штучна бджолина колонія складається з трьох видів бджіл, що відрізняються одна від одної своєю функцією: робочі бджоли, бджоли-розвідники та незайняті бджоли-спостерігачі. Кожна робоча бджола працює виключно з одним джерелом їжі. Вона зберігає в пам'яті інформацію про це джерело, коли відлітає з вулика і ділиться цією інформацією з бджолами-спостерігачами в так званій зоні для танців. Після того як спостерігачі подивляться танці робочих бджіл, вони оберуть джерело їжі, яке в подальшому будуть намагатися покращити. Якщо робоча бджола лишає своє джерело їжі, вона перетворюється на розвідника і починає випадковим чином шукати нові джерела.

В алгоритмі ABC кожне джерело їжі відповідає можливому розв'язку задачі, а кількість нектара, що містить джерело - значенню фітнес функції відповідного розв'язку. Як відомо, на

меті оптимізації розв'яку задачі комівояжера є мінімізація довжини повного обходу всіх міст. Якість кожного розв'язку fit_i розраховується наступним чином:

$$fit_i = \frac{1}{f(x_i)} \tag{5}$$

де $f(x_i)$ - це довжина маршруту для ров'язку x_i. Максимальна кількість розвідників l обчислюється за формулою:

$$l = \frac{cs * D}{3} \tag{6}$$

де cs і D дорівнюють розміру колонії та розмірності задачі відповідно ($D = n$).

Детально комбінаторний варіант алгоритму ABC наведений нижче:

1. Ініціалізація параметрів: розмір колонії cs, максимальна кількість ітерацій $MaxNumber$.
2. Ініціалізація початкових розв'язків (позицій джерел їжі) $x_i, i = 1, 2.., cs$.
3. Розрахунок цільових функцій для наявних розв'язків.
4. Запам'ятати найкращий розв'язок.
5. $c = 0$ (лічильник кількості ітерацій)
6. REPEAT
   a) Фаза робочих бджіл: для кожної робочої бджоли:
      i. Генерація нового розв'язку v_i в околі x_i та розрахунок значення цільової функції для нього
      ii. Обрати кращий з v_i та x_i, керуючись жадібним підходом
   b) Розрахувати ймовірності P_i для знайдених попередньо розв'язків

$$P_i = \frac{0,9 * fit_i}{fit_{best}} + 0,1 \tag{7}$$

   c) Фаза спостерігачів: для кожного спостерігача:
      i. Генерація нового розв'язку v_i з x_i, в залежності від P_i та розрахунок значення цільової функції для нього
      ii. Обрати кращий з v_i та x_i, керуючись жадібним підходом
   d) Запам'ятати найкраще зі знайдених розв'язків
   e) Визначити покинуті розв'язки, і якщо вони існують замінити їх новими
7. $c = c + 1$
8. UNTIL $c = MaxNumber$.

Література. **1.** Меламед, И., Сергеев, С., & Сигал, И. (1989). Задача коммивояжера. Вопросы теории. Автомат. И Телемех., (9), 3-33. **2.** Lučić, P., & Teodorović, D. (2001). Bee system: modeling combinatorial optimization transportation engineering problems by swarm intelligence. In Preprints of the TRISTAN IV Triennial Symposium on Transportation Analysis (pp. 441–445). Sao Miguel, Azores Islands, Portugal. **3.** Lučić, P., & Teodorović, D. (2002). Transportation modeling: an artificial life approach. In Proceedings of the 14th IEEE "International Conference on Tools with Artificial Intelligence" (pp. 216-223). Washington, DC. **4.** Lučić, P., & Teodorović, D. (2003). Computing with bees: attacking complex transportation engineering problems. International Journal On Artificial Intelligence Tools, 12(3), 375-394. **5.** Lučić, P., & Teodorović, D. (2003). Vehicle routing problem with uncertain demand at nodes: the bee system and fuzzy logic approach. In Fuzzy Sets in Optimization (pp. 67-82). Springer-Verlag, Heidelberg, Berlin. **6.** Basturk, B., & Karaboga, D. (2006). An artificial bee colony (abc) algorithm for numeric function optimization. In IEEE swarm intelligence symposium 2006. Indianapolis, IN, USA. **7.** Basturk, B., & Karaboga, D. (2007). Artificial bee colony (abc) optimization algorithm for solving constrained optimization problems. In Proceedings of the 12th international fuzzy systems association world congress on foundations of fuzzy logic and soft computing (pp. 789-798). Springer, Berlin.

Галушко М.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Бібліотека наукових публікацій з використанням RDF-сховищ

При вирішенні задач, в яких дані можуть мати довільні зв'язки, виникає непередбачувана кількість зв'язків в запитах, тому для вирішення такого плану задач, зараз найбільшого розповсюдження набули RDF-сховища. Вони базуються на стандартах комітету W3C для мови опису графів (*RDF*) та для обробки графових даних (*SPARQL*). RDF значить "середовище опису ресурсів". Це модель даних, що представляє дані простими триплетами *суб'єкт – предикат – об'єкт*. Триплет також відомий як "оператор" і є базовим "явищем", або його ще можна назвати стверджуваним блоком знань в RDF. Декілька операторів комбінуються разом узгоджуючись таким чином: суб'єкт, об'єкт як вузол і предикат як ребро. Таким чином, виникає структурна мережа, також відома як RDF-граф і має вигляд вузол-ребро-вузол.

RDF може розглядатись як альтернативний до SQL підхід до організації сховищ даних у прикладних додатках, що дозволяє вирішити деякі проблеми масштабованості і паралельної обробки великих обсягів слабко структурованих даних. Нині створюються та розвиваються численні бази триплетів та споріднені їм графові бази даних, такі як Open Link Virtuoso, GraphBase, DEX, Neo4J, Infinite Graph, OrientDB, Blazegraph (RDF-сховище, раніше називалося Bigdata), FlockDB, Titan та ін. Серед них і гібридні системи, що реалізують сховищша триплетів на основі традиційних зрілих технологій реляційних СУБД (рис.1). Недоліком гібридної схеми є невисока ефективність у порівнянні з "чистими" RDF-сховищами через необхідність перетворення графової моделі даних у реляційну та навпаки.

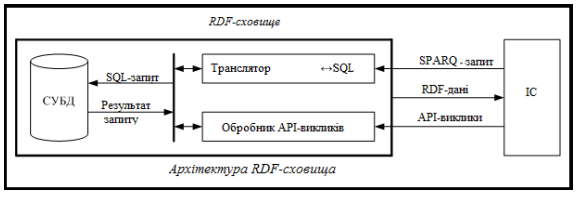


Рис. 1. Гібридна архітектура RDF-сховища

Задачею даного дослідження є виявлення ефективних способів та засобів використання RDF-сховищ у прикладних системах, таких як бібліотеки публікацій. Особливий інтерес складає можливість описувати у RDF не лише бібліографічні дані (відносно чітко структуровані), а й смислові конструкції змісту самої публікації, що є більш "тонким" рішенням, ніж просте збереження ключових слів. Це може дати якісний виграш відносно існуючих пошукових механізмів, що здійснюють пошук інформації без урахування семантики слів, які входять до запиту, а також контексту, в якому вони використовуються. Недоліком такого підходу є, звісно, необхідність вручну "семантично анотувати" наявні у бібліотеці ресурси, і дослідження засобів автоматизації такої роботи є окремою актуальною задачею.

Література. **1.** Mike Bergman. (2009). Advantages and Myths of RDF. Available: http://www.mkbergman.com/483/advantages-and-myths-of-rdf/ (дата звернення: 20.02.2016). **2.** OWL Web Ontology Language guide. W3C working draft. W3 Consortium, 2003. URL: http://www.w3.org/TR/2003/WD\T2A\textendashowl\T2A\textendashguide\T2A\textendash20030331/ (дата звернення: 20.02.2016).

***Ганиева С.А.*[1], *Танырвердиев Ч.Г.*[2], *Мутталибова Ш.Ф.*[1]**
[1] *Азербайджанский Университет Архитектуры и Строительства, Баку, Азербайджан;*
[2] *Проектно-исследовательский центр кадастра и землеустройства, Баку, Азербайджан*

Метод создания электронной кадастровой информационной базы недвижимости и цифровых кадастровых карт на примере города Сумгаит

В этой работе было показано, метод создания электронного базовой кадастровой информации и цифровых кадастровых карт объектов недвижимости в примере города Сумгаит. Работа, в основном, были выполнены в 4 этапа. В данной работе были использованы из возможностей современных технологий и многочисленных кадастровой информации, которые собрались за долгие годы

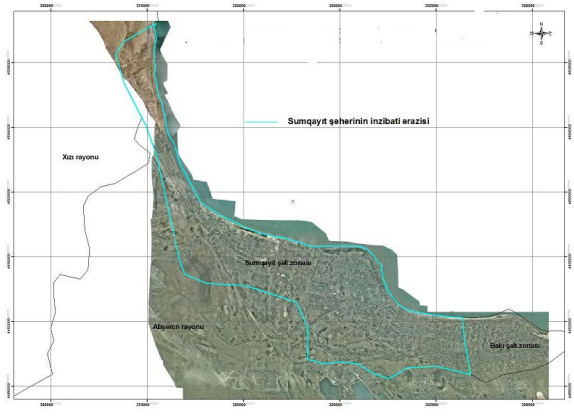
В Азербайджане собрана многочисленная кадастровая информация в различных отраслях. По причине разбросанности и разнообразия этой информации осложнилась их систематизация и управление в виде единого целого. Принимая это во внимание, создание в Азербайджане централизованно управляемой и отражающей в себе соответствие современным требованиям единой государственной кадастровой информационной базы недвижимости и составление цифровых кадастровых карт должны являться факторами, стимулирующими развитие в любой отрасли.

Объединение в единый орган всех функций по регистрации и создание единого реестра прав на недвижимость, в то же время создают предпосылки для сбора единой государственной кадастровой информации недвижимости в единый информационный банк. В Азербайджане создание данной системы в 3 этапа является одним из важнейших условий для: 1. Сбора информации, имеющейся в соответствующих отраслевых кадастровых органах, систематизации этой информации в соответствии с единой структурой с целью качественного использования и применения на производстве. 2. Внесения информаций различных организаций, занимающихся сбором кадастровой информации, в единую государственную кадастровую систему. 3. Посредством возможностей современных технологий постоянного обновления всей собираемой кадастровой информации.

В качестве примера можно указать работы по созданию электронной кадастровой информационной базы для территории города Сумгаит, выбранной для исследовательских работ с целью создания единого государственного кадастра недвижимости (рис. 1).

Город Сумгаит расположен на равнинной территории в 35 км к северо-западу от города Баку, на западном побережье Каспийского моря, в устье реки Сумгаит. В городе имеются два административно-территориальных округа, два поселка и три муниципалитета.

a)



b)

Рис. 1. a) Административная территория расположения Сумгаита; b) Территория, выбранная для исследования.

С целью создания единой государственной кадастровой информационной базы недвижимости и составления цифровых кадастровых карт, дающих возможность совершенного управления имущества на территории города Сумгаит, была собрана многочисленная информация по исследуемой территории.

Здесь к недвижимому имуществу в основном были отнесены нижеуказанные объекты: земельные участки, отдельные водные объекты, участки земельных единиц, леса, здания, квартиры, сооружения и прочие объекты, тесно связанные с землей, то есть, объекты недвижимого имущества, смена мест расположения которых невозможна без нанесения ущерба назначению объекта.

На первом этапе были выполнены работы по измерению территорий. При этом независимо от вида имущества была собрана информация об объектах недвижимости посредством определения натуральных географических позиций, расположения в системе координат WGS 84, путем использования пограничных линий всех видов объектов недвижимости, в том числе, объектов с незавершенным строительством и инженерно – коммуникационных объектов, приведенных в цифровую среду, и высококачественных электронных геодезических инструментов (электронный тахометр, GPS и т. д.). Кроме того, при этом были осуществлены следующие работы:

- были уточнены границы фактического использования на территории всех видов наземных объектов недвижимости, имеющихся на проектной территории;
- были определены виды собственности объектов недвижимости, установлены назначение земельных участков, их пользователи и собственники.

На втором этапе в камеральных условиях с использованием специальных фотограмметрических установок были определены координаты объектов недвижимости на ортофотоснимках, существующих в рамках всевозможных допускаемых ошибок. В результате границы территории исследования были уточнены посредством названий, координат и схем расположения опорных межевых знаков. На нижеуказанных рисунках один из указанных работ были приведены на примере абрис относительно точки с условным указанием показателей (рис. 2).

Информация о точках

Название города	Сумгаит	Фото	Абрис точек	Территория расположения (1:25000)
Село/ Квартал	-			
M/p №				
DOM и DG				
Точка №	1			
Координаты WGS84				
Y	123456.045			
X	7893214.123			
h	-3,928			
T. Пол	Гвоздь			
Территория расположения		На перекрестке улиц Гуртулуш и Сульх на противоположной стороне парка на асфальтовом покрытии дороги		

Рис. 2. Изображение территориально-камеральных измерений по исследуемой территории.

Картографические снимки земельных участков и объектов недвижимости приняты с нижеуказанной точностью: система координат WGS-84, ссылочная высота – уровень Балтийского моря, проекция системы координат UTM (зона 6-го уровня в пределах UTM 38–39, коэффициент масштабирования 0.9996 на Центральной меридиане).

На третьем этапе границы всех земельных участков и их частей (здания, сооружения, имущественные комплексы, административные территории, целевые эксплуатируемые земли, транспортные инфраструктуры, элементы гидрографии и т.д.), то есть, всех видов объектов недвижимости в пределах исследуемой территории были определены и сформированы в натуре в соответствии с принципами кадастрового раздела земель, а также, было предложено

формирование кадастровых номеров, присваеваемых земельным участкам, нижеуказанным образом (рис. 3).

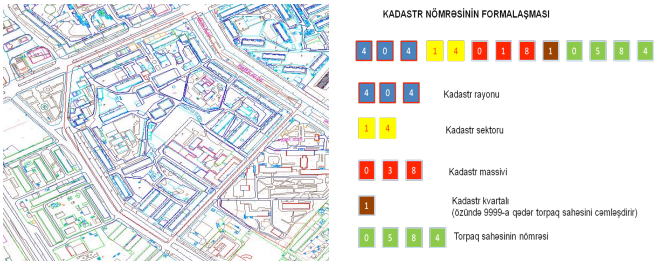


Рис. 3. Определение в натуре границ объектов недвижимости и формирование кадастровых номеров.

На 4-ом этапе в качестве заключения осуществленных территориально – геодезических работ вся информация относительно любых видов объектов недвижимости, расположенных на исследуемой территории, были систематизированы, в результате чего была создана система кадастровой информационной базы и цифровой кадастровой карты, на которой совместно можно увидеть пространственное географическое местоположение объектов недвижимости и совокупность их атрибутивных кадастровых показателей (рис. 4).

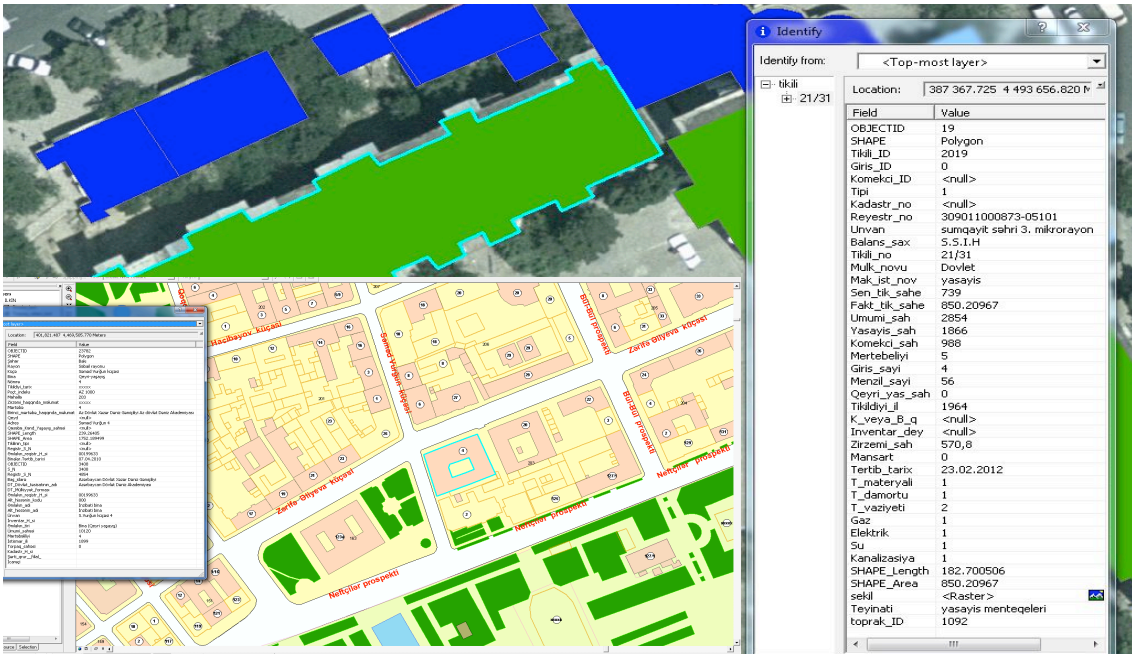


Рис. 4. Фрагмент электронной кадастровой информационной базы и цифровой кадастровой карты объектов недвижимости.

Литература. **1.** “Земельное законодательство” Издательство Закон, 2009. **2.** Сачлы Ганиева, Мубариз Исмаилов. “Земля и Многоцелевой Кадастр”. Баку, 2006. **3.** “Современный кадастр, интеграция данных” – распределённость пользователей. // Геодезия и картография, 2000. **4.** https://www.fig.net/resources/publications/figpub/cadastre2014/translation/c2014-russian.pdf

Дичко В.М.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Система співставлення супутникових зображень

Одним із найбільш поширених і зручних засобів для збору геопросторових даних на сьогодні є використання БПЛА (безпілотних літальних апаратів). У мирний час цю інформація можна використовувати для моніторингу стану доріг та трубопроводів, пошуку лісових пожеж. Проте використання БПЛА з автоматичним керуванням у військовий час може бути обмежено через можливість застосування супротивником завад сигналу GPS.

Таким чином, при вирішенні задачі керування на основі оптичної навігаційної системи, актуальною є проблема співставлення зображення, отриманого за допомогою камери БПЛА із картою, що зберігається в БПЛА. Незважаючи на відмінності зображень, деякі частини залишаються незмінними, тому для розпізнавання спільних точок доцільно застосовувати алгоритми, які знаходять ключові точки зображень та їх дескриптори [1–4], за допомогою яких ці точки можна порівнювати. Також в залежності від вимог часу роботи та якості співставлення, необхідно використовувати різні алгоритми.

У даній роботі була поставлена задача розробити систему з масштабованою архітектурою для співставлення зображень карт з можливістю вибору методів виявлення ключових точок.

Система із запропонованою архітектурою (рис. 1), завдяки використанню шаблону проектування «Стратегія», дозволяє розширення реалізацій різних кроків алгоритму, таких як методи виявлення і опису ключових точок, співставлення дескрипторів ключових точок, виявлення областей відповідностей та перетворення і накладання зображень. Таким чином, конкретні реалізації можуть бути змінені незалежно від клієнтського коду, що використовує ці методи. Дана система реалізована з використанням мови C++ та бібліотеки комп'ютерного зору OpenCV, завдяки чому забезпечується висока ефективність та можливість подальшого легкого розширення та покращення методів обробки зображень.

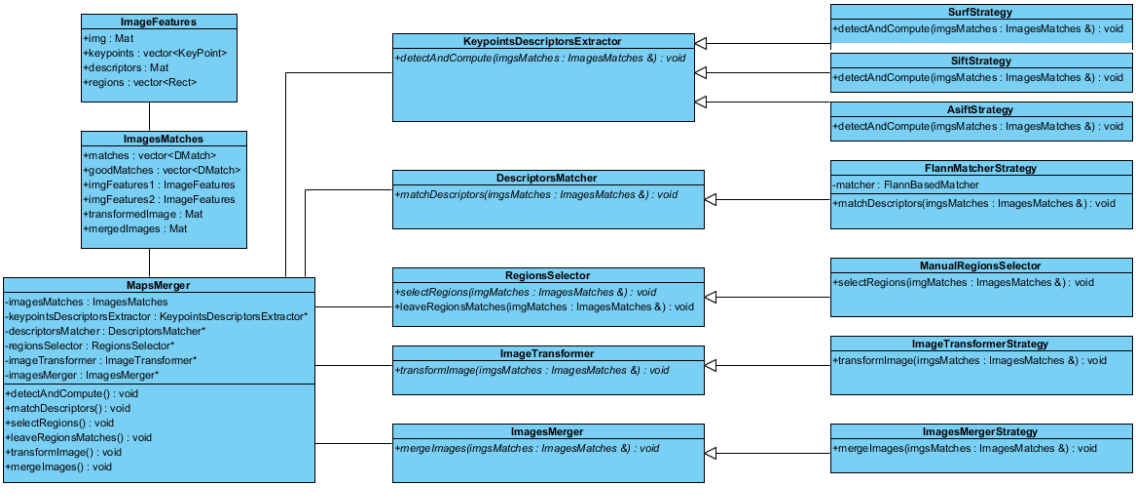


Рис. 1. Архітектура системи співставлення зображень

Література. **1.** Lowe D. G. Object Recognition from Local Scale-Invariant Features / D. G. Lowe // International conference on Computer Vision. – 1999. – P. 1150-1157. **2.** Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints / D.G. Lowe // International Journal of Computer Vision. – 2004. – Vol. 60, № 2. – P. 91-110. **3.** Bay H. Surf: Speeded up robust features / H. Bay, T. Tuytelaars, L. Van Gool // European Conference on Computer Vision. – 2006. – P. 1:404-417. **4.** Morel J. ASIFT: A New Framework for Fully Affine Invariant Image Comparison / J. Morel, G. Yu // SIAM Journal on Imaging Sciences – 2009. – Vol. 2, №2. – P. 438-469.

Діденко Д.Г., Журке Б.С.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Імітаційна модель для прогнозування фармацевтичного ринку України

Вступ. Фармацевтичний ринок – це один із потужних промислових секторів, який входить до лідерів найбільш прибуткових галузей світового господарського комплексу. Ця галузь посідає значне місце в економіці України. На даний момент часу, фармацевтичний ринок України переживає скрутні часи. Так, наприклад, обсяг проданих упаковок препаратів всіх категорій скоротився на 16,7 %, а в доларовому вираженні загальний обсяг продажів «аптечного кошика» зменшився на 40,2% [1]. З огляду на такі тенденції, необхідно створити ефективну політку управління та планування. Для розробки ефективних рішень застосовують прогнозування. Сучасні моделі для побудови прогнозу дуже трудоємкі, а також мають недоліки.

Постановка задачі. Для побудови ефективної моделі для прогнозування, часто, необхідно залучати фахівців з предметної області до якої застосовується прогноз. Сама модель має багато залежностей, які важко прослідкувати. Необхідно задавати критерії оптимальності моделі. Визначити важливі фактори, які впливають на якість прогнозу.

Реалізація. Для побудови ефективної моделі прогнозування можна застосувати імітаційне моделювання. Яке в свою чергу дозволить комбінувати декілька моделей прогнозування для отримання найкращого результату [2]. Застосування агентної моделі для імітаційного моделювання, дозволить досягти поставленої мети. Використання агентного підходу, надасть системі більшу гнучкість, зменшить навантаження на систему в якій будуть проводитись експерименти, дозволить прорахувати стан моделі через певний інтервал часу та підвищити якість прогнозу додавши складові які раніше не враховувались. Шляхом додавання нових агентів в модель. При використанні агентної моделі, не потрібно буде описувати складні формули, які змінюються в часі під впливом зовнішніх факторів, як це відбувається в інших моделях. В такій системі можна виділити два типи агентів, агент-менеджер, агент-виконавець. Агент-менеджер буде роздавати завдання іншим агентам і опрацьовувати отримані результати іншими агентами, та виводити результати. Результат представляє собою середньозважений прогноз. Агент-виконавець, представляє собою один із методів прогнозування, наприклад прогнозування за допомогою нейронних мереж або регресійних методів. Взаємодія таких агентів між собою дозволить компенсувати недоліки традиційних методів прогнозування і підвищити якість прогнозу.

Висновок. Наведена система на основі агентного підходу до імітаційного моделювання [3], дозволить отримати якісний прогноз, використовуючи всі переваги підходів до прогнозування, в будь який момент часу таку систему можна буде модифікувати додавши нового агента попередньо описавши його функціональну частину. Взаємодія користувача з такою системою являється більш спрощеною, оскільки користувач ставить завдання перед агентом, не вникаючи в суть як саме він буде її вирішувати. Також така модель достатньо проста при проектуванні, а реалізація такої системи буде займати менше часу, що дозволить отримати своєчасну інформацію, для прийняття ефективного рішення в сфері управління та планування, та зекономити фінансові ресурси.

Література. **1.** Аптечный рынок Украины по итогам I полугодия 2015 г.: Helicopter View. [Електронний ресурс].—Режим доступу: www.apteka.ua/article/339254. **2.** Alfares H.K., Nazeeruddin M. Electric load forecasting: literature survey and classification of methods // Internatioal Journak of Systems Science. 2002, Vol 33. P. 23-34. **3.** Карпов Ю. Г. Имимтационное моделирование систем. Введение в моделирование с Anylogic 5. / Карпов Ю.Г.—СПб.: БХВ-Петербург, 2005.—400с.

Діденко Д.Г., Климковська К.Д.
ННК «ІПСА», НТУУ «КПІ», Київ, Україна

Розподілене обчислення і прогнозування важливості брендів в Інтернеті

Вступ. Щодня в мережі Інтернет публікується дуже багато контенту: відгуки про компанії та думки про них, фото і відео цікавих місць, опис нових товарів і багато іншого. Для того, щоб компанія ефективно функціонувала на просторах онлайн-спільнот, необхідний регулярний моніторинг видачі інформації в мережі Інтернет і його своєчасний аналіз. Активно працюючи з моніторингом соціальних мереж, компанії мають можливість швидше за конкурентів дізнаватися про нові тренди та очікувати потенційних клієнтів [1]. Тобто в сучасному світі для вдалого розвитку компанії дуже важливо проводити моніторинг брендів та дізнаватися думку користувачів товару, наприклад через такий потужний механізм, як соціальні мережі.

Постановка задачі. Поняття *«важливість бренду»* є неоднозначним. В даному випадку воно визначається шляхом аналізу кількості позитивних та негативних відгуків про бренд. В свою чергу *«позитивність»* та *«негативність»*, тобто відтінок відгуку, теж підлягає аналізу. Важливість бренду досліджується в найбільших соціальних мережах – Facebook (https://www.facebook.com/), ВКонтакте (http://vk.com/), Twitter (https://twitter.com/) та найбільших форумах. Обчислення вавжливості бренду є розподіленим для досягнення більшої швидкості обробки даних. За результатами щоденного обчислення важливості бренду будуються прогнозні значення кількості позитивних та негативних відгуків на короткий проміжок часу (тиждень).

Реалізація алгоритму аналізу тональності. Для визначення тональності відгуків користувачів був застосований метод машинного навчання з учителем. Машинне навчання з учителем є найбільш поширеним методом, який використовується в дослідженнях. Його суть полягає в тому, щоб навчити машинний класифікатор на колекції заздалегідь розмічених текстів, а потім використовувати отриману модель для аналізу нових документів. Процес створення системи аналізу тональності складався з таких кроків:

1. Була зібрана колекція документів для навчання класифікатора (500 реальних відгуків користувачів Facebook).
2. Кожен документ з навчальної колекції був представлений у вигляді вектору ознак (комбінація уніграм та біграм слів). Уніграми і біграми дають кращі результати ніж N-грами більш високих порядків (триграми і вище), тому що вибірка навчання недостатньо велика для підрахунку N-грам вищих порядків. Тому для даного класифікатора було вибрано за ознаки уніграми, біграми та їх комбінацію.
3. Для кожного документа була вказана «правильна відповідь», тобто тип тональності (наприклад, позитивна чи негативна), за цими відповідями і був навчений класифікатор.
4. В якості алгоритму класифікації був вибраний наївний баєсівський класифікатор (NB) та метод опорних векторів (SVM).
5. Використання отриманої моделі.

Для розв'язання задачі класифікації тональності відгуків було вибрано розділення можливої тональності на 3 класи: позитивна, негативна та нейтральна. Спочатку був навчений бінарний класифікатор, який відокремлює нейтральні тексти від суб'єктивних; потім навчався класифікатор, який відокремлює позитивні думки від негативних. Так як обробка великої кількості відгуків займає багато часу і вимагає багато ресурсів, вона виконується на кластерах Hadoop за допомогою моделі розподілених обчислень MapReduce [2].

Реалізація алгоритму короткострокового прогнозування. При короткостроковому прогнозуванні, а також у випадку зміни зовнішніх умов, коли найбільш важливими є останні реалізації досліджуваного процесу, найбільш ефективними є адаптивні методи, які враховують нерівноцінність рівнів часового ряду. Адаптивні моделі прогнозування здатні швидко пристосовувати свою структуру та параметри до зміни умов. Інструментом прогнозу в адаптивних моделях є математична модель з єдиним фактором «час» [3]. Для короткострокового прогнозування

значень часового ряду була побудована нелінійна модель, яка дозволяє відобразити розвиток експоненційної тенденції, а також рядів без тенденції:

$$Y(t+k) = A_0 + A_1 exp(k) + A_2 exp(\frac{1}{k}) \quad (1)$$

t – поточний час;

k – час випередження.

Дану модель було вибрано тому, що вона є точнішою за лінійну та краще описує зміни значень часового ряду. Вона будується таким чином:

1. За декількома першими точками часового ряду методом найменших квадратів знаходяться значення параметрів A_0, A_1 та A_2.
$$Y_p(t) = A_0 + A_1 exp(t) + A_2 exp(\frac{1}{t}) \quad (2)$$
2. Використовуючи знайдені параметри, знаходиться прогнозне значення на наступному кроці:
$$Y_p(t+k) = A_0(t) + A_1(t) exp(k) + A_2(t) exp(\frac{1}{k}), k = 1 \quad (3)$$
3. Знайдемо похибку прогнозування:
$$e(t+k) = Y(t+k) - Y_p(t+k) \quad (4)$$
4. Згідно з похибкою змінюються значення параметрів моделі:
$$A_0(t+1) = A_0(t) + A_1(t) + A_2(t) + (1-\beta)^2 e(t)$$
$$A_1(t+1) = A_1(t) + A_2(t) + (1-\beta)^2 e(t)$$
$$A_2(t+1) = A_2(t) + (1-\beta)^2 e(t) \quad (5)$$
β – коефіцієнт дисконтування даних, $0<\beta<1$.
5. З моделі з виправленими параметрами знаходиться прогноз на наступний крок та відбувається перехід до п.3 якщо t<N. При t⩾N отримане значення використовується як прогнозне, не змінюючи параметри моделі.

Висновок. На сьогоднішній існує багато систем моніторингу соціальних мереж та блогів, але точність аналізу тональності користувацьких відгуків в цих системах все ще має досить низький рівень, а наведені результати аналізу не дають ніяких прогнозів на майбутнє. Тому, проаналізувавши предметну область та ознайомившись з прикладами найбільших та найпопулярніших систем моніторингу онлайн-відгуків про бренд, було прийнято рішення створити власну систему обчислення та прогнозування важливості брендів в Інтернеті, яка обчислює важливість бренду, як співвідношення кількості позитивних, негативних та нейтральних відгуків про бренд в сучасних соціальних мережах (Facebook, ВКонтакте) та найпопулярніших блогах та форумах за допомогою наївного баєсівського класифікатору та прогнозує короткострокові зміни важливості бренду на основі нелінійної моделі. Цю модель було створено но основі багатьох експериментів і дана модель вперше застосовується на практиці. Також для пришвидшення аналізу збір відгуків та обчислення їх тональності є розподіленим.

Література. **1.** Моніторинг соціальніх мереж, http://iqbuzz.ru/monitoring-soczialnyix-media-i-setej.php/. **2.** Andrew Pavlo, Erik Paulson, Alexander Rasin, Daniel J. Abadi, David J. DeWitt, Samuel Madden, Michael Stonebraker. A Comparison of Approaches to Large-Scale Data Analysis. Proceedings of the 35th SIGMOD International Conference on Management of Data, 2009, Providence, Rhode Island, USA. **3.** С.И. Татаренко. Методы и модели анализа временных рядов. -Т.:ТГТУ, 2008. -191 с.

Діденко Д.Г., Конєв Д.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Прискорення пошуку напівдублікатів текстових файлів за допомогою алгоритму MapReduce

Вступ. Задача визначення ступеня схожості текстових документів має велике прикладне значення, перш за все – для визначення ступеня оригінальності заданого тексту. З цією задачею зіштовхуються як пошукові системи в мережі Інтернет, приймаючи рішення про індексацію нових ресурсів, так і експерти наукових видань, які оцінюють оригінальність наукових робіт. Основною проблемою на шляху вирішення даної задачі є велика кількість даних, що порівнюються. Тому зараз є актуальним пошук засобів прискорення процесу пошуку напівдублікатів без зменшення точності та повноти пошуку.

Постановка задачі. Існуючі алгоритми порівняння текстових файлів працюють достатньо швидко до тих пір, поки колекція файлів достатньо мала, проте з її збільшенням швидкість суттєво падає. Одним з можливих варіантів пришвидшення є спрощення алгоритмів, але це негативним чином впливає на точність та повноту їх роботи. Отже логічним кроком буде збільшення кількості даних, що оброблюються одночасно, тобто – розподілення обчислень.

Реалізація. В якості методу порівняння документів можна взяти метод міри TF-IDF [1]. Він дає дуже високу точність [2] за рахунок того, що він враховує всю колекцію одночасно. Для збільшення швидкості обчислень використовується алгоритм MapReduce [3], який дає змогу виконувати паралельну обробку даних на кластері з декількох будь-яких комп'ютерів, що в свою чергу дозволяє значно зменшити вартість системи. Порівняння документів за допомогою міри TF-IDF виконується на основі 6-ти найбільш «важливих» слів у документі. Вага кожного слова обчислюється на основі того, наскільки часто воно зустрічається у тексті, а також кількості текстів у яких воно зустрічається. Дані про всі унікальні слова в колекції записуються у словник. У даному методі MapReduce можна застосувати одразу в декількох місцях:

- підрахунок кількості входжень кожного слова в тексті – є масштабованою задачею, тобто можна розбити текст на частини, кожну з яких відправити на обчислювальний вузол Map, та отримати загальний результат на етапі Reduce;
- оскільки при додаванні тексту розмір колекції збільшується, змінюється кількість текстів у яких зістрічається кожне слово зі словника. Це означає, що міру TF-IDF необхідно перераховувати для всіх слів в колекції. Під час виконання цієї задачі, кожен вузол Map обчислює TF-IDF лише для частини колекції. Крок Reduce в цій задачі непотрібен, тому що результати непотрібно групувати.

Висновок. Використання алгоритму MapReduce під час пошуку напівдублікатів дозволяє помітно знизити час виконання, не погіршуючи при цьому показники точності та повноти методу пошуку. Також даний алгоритм є достатньо простим в реалізації, та не потребує значних витрат. Єдиним недоліком такого підходу є більший час виконання під час роботи з малими обсягами даних, проте це не є великою проблемою враховуючи сучасні обсяги інформації.

Література. **1.** J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In Proc. CVPR, 2007 **2.** Зеленков Ю. Г., Сегалович И. В. Сравнительный анализ методов определения нечетких дубликатов для Web-документов //Труды. – 2007. – Т. 9. – С. 166-174. **3.** Andrew Pavlo, Erik Paulson, Alexander Rasin, Daniel J. Abadi, David J. DeWitt, Samuel Madden, Michael Stonebraker. A Comparison of Approaches to Large-Scale Data Analysis. Proceedings of the 35th SIGMOD International Conference on Management of Data, 2009, Providence, Rhode Island, USA.

Діденко Д.Г., Крюков О.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Модель адаптивного навчання на базі алгоритму кластеризації категорійних даних CLOPE

Дистанційна освіта є одним із основних трендів сучасного навчання і стало частиною життя більшості людей. З появою онлайн-освіти досить швидко почали розвиватися системи адаптивного навчання. Адаптивне навчання дозволяє навчати без звичних, для масової освіти, численних прогалин в індивідуальній підготовленості учнів та студентів, досягати необхідної структури та бажаного рівня знань [1].

Сучасні підходи, які застосовують Колаборативну фільтрацію, Бейесівські мережі, мережі Петрі та ланцюги Маркова, дозволяють вибрати завдання за рівнем знань учня та темами, які він знає, проте щоб досягти такої функціональності вони повинні комбінуватися, що збільшує складність системи, накладає багато обмежень та ускладнює масштабування таких платформ. Також такі системи дуже залежать від даних про інших користувачів при оцінці знань та складності матеріалів [2]. Ціль даної роботи створити досить просту модель на базі методу кластерного аналізу, яка більшою мірою залежатиме від множини знань конкретного користувача системи, та дозволить задавати навчальний «маршрут» лише з мінімальним втручанням експерта.

В якості доменної моделі навчальної системи будемо розглядати набір питань, до кожного з яких є набір деяких тегів, з кожним питання зв'язано правило, що буде з'являтися в якості підказки при неправильній відповіді. Набір тегів буде задаватися експертом при створенні питань та грати вирішальну роль при побудові семантичної мережі зв'язків між матеріалами. Всі питання будуть розподілені на кластери за тегами. В якості моделі учня в системі буде виступати окремий кластер, який буде містити пари <тег: кількість>. Що дозволить вибирати навчальний матеріал, який ближче знаходиться до знань студента.

Застосовувати для кластеризації об'єктів з категоріальним ознаками традиційні алгоритми неефективно, а часто – неможливо. Основні труднощі пов'язані з високою розмірністю і гігантським обсягом, якими часто характеризуються такі бази даних. Саме дому для подальших досліджень вибраний алгоритм CLOPE [3].

В основі алгоритму кластеризації CLOPE лежить ідея максимізації глобальної функції вартості, яка підвищує близькість транзакцій в кластерах за допомогою збільшення параметра кластерної гістограми.

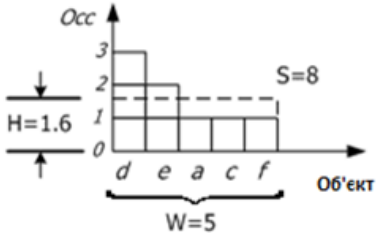


Рис. 1. Приклад гістограми кластера

Нехай є база питань D, що складається з множини питань $\{t_1, t_2, ..., t_n\}$. Кожне питання є набір тегів $\{i_1, ..., i_m\}$. Множина кластерів $\{C_1, ..., C_k\}$ є розбиття множини $\{t_1, ..., t_n\}$, таке, що $C_1...C_k = \{t_1, ..., t_n\}$ і $C_i \neq \oslash \wedge C_i \cap C_j = \oslash$, для 1 <= i , j <= k. Кожен елемент C_i називається кластером, n, m, k - кількість транзакцій, кількість об'єктів в базі транзакцій і число кластерів відповідно.

Кожен кластер C має наступні характеристики: D(C) –множина унікальних об'єктів; Occ(i,C)

– кількість входжень об'єкта i у кластер C;

Площа фігури, що уособлює кластер може бути розрахована за формулою:

$$S(C) = \sum_{i \in D(c)} Occ(i, C) = \sum_{t_i \in C} |t_i|$$

$$W(C) = |D(C)|$$

$$H(C) = \frac{S(C)}{W(C)}$$

Де H – висота кластера, а W – ширина. Для того, щоб елементи в кластері були максимально близькі, треба вибирати такі розбиття, що максимізують висоту. Проте зазвичай для оцінки використовують градієнт:

$$G(C) = \frac{H(C)}{W(C)} = \frac{S(C)}{W(C)^2}$$

Якщо розглядати модель учня як окремий кластер, тоді ми можемо вибирати наступні матеріали для користувача на основі глобальної функції оптимізації [4]:

$$Profit(C) = \frac{\sum\limits_{i=1}^{k} G(C_i) \times |C_i|}{\sum\limits_{i=1}^{k} |C_i|} = \frac{\sum\limits_{i=1}^{k} \frac{S(C_i)}{W(C_i)^r} \times |C_i|}{\sum\limits_{i=1}^{k} |C_i|}$$

Загальний алгоритм має наступний вигляд:

1. початкове розбиття питань на тематики за тегами;
2. розбиття питань вибраної користувачем теми на пули питань;
3. вибір пулу питань з найбільшим значення Profit-функції, тобто найближчий до кластеру, що уособлює знання користувача;
4. вибір питання з найбільшим значення Profit-функції. При правильній відповіді – додати відповідні теги в базу користувача, в інакшому випадку видалити відповідні теги з бази користувача (мається на увазі інкримент і дикримент значення в парі <тег: кількість>). Видалити відповідне питання з пулу питань користувача;
5. повторювати пункти 3-4;

Такий підхід дає ряд переваг перед існуючими методами:

- відсутня проблема «холодного старту». Система швидко стабілізується під кожного користувача;
- не потрібна інформація про відповіді інших користувачів;
- швидка обробка даних, так як інформація про кожний кластер зберігається і не обчислюється при кожному перезапуску;
- зручне масштабування. Додавати нові елементи в кластер можно не переобчислюючи існуючий розподіл кластерів.

Література. **1.** Сучасні тенденції в методах адаптивного навчання, http://shadowofaero.blogspot.com/2014/12/blog-post.html. **2.** Гагарін О.О.,Титенко С.В. Дослідження і аналіз методів та моделей інтелектуальних систем безперервного навчання // Наукові вісті НТУУ "КПІ". – 2007. – № 6(56). – С. 37-48. **3.** Comparison of some approaches to clusteringcategorical dataRezankova H., Husek D., Kudova P., Snasel V. **4.** CLOPE:A Fast and Effective Clustering Algorithm for Transactional Data Yiling Yang, Xudong Guan, Jinyuan You, Dept.of Computer Science & Engineering.,Shanghai JiaoTong University, Shanghai, P.R.China

Євтухов С.К.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Стеганографічний метод ін'єкціювання інформації в зображення динамічного кодування

Насьогодні актуальним питанням у сфері інформаційної безпеки (зокрема через постійне збільшення потоків інформації, розвиток мережевих і мультимедійних технологій) є захист інформації під час передачі відкритими каналами зв'язку. Сучасні клієнт-серверні системи, особливо такі, що оперують конфеденційною інформацією користувачів (онлайн-банкінги, веб-гаманці тощо) часто є вразливими через відсутність стеганографічного захисту інформації, який би дозволяв не лише приховати зміст секретного повідомлення (наприклад, хеш номеру банківської картки), а й приховати сам факт його існування. Частково це питання вирішує *цифрова стеганографія*, пропонуючи як рішення ін'єкціювання інформації у графічні контейнери.

Одним з графічних форматів є формат динамічного стиснення JPEG. Проте використання JPEG-зображень у якості контейнерів для інформації наразі не є прийнятим через складну структуру формату. А існуючі підходи не є ефективними, оскільки помітно впливають на візуальну складову JPEG-зображень, переписуючи байти пікселів. Відтак доцільним є розробка методу, алгоритму й подальша їх програмна реалізація для приховування інформації в зображення динамічного кодування формату JPEG без впливу на візуальну складову зображення.

Поняття стеганосистеми. Стеганосистемою називається п'ятірка $\{C,M,S,E,D\}$ що процес стеганографічних перетворень кодування (E) і декодування (D) описується наступними правилами:

$$E\text{: } C \times M \to S, \tag{1}$$

$$D\text{: } S \to M. \tag{2}$$

де $M = \{m_1 ... m_n\}$ - множина всіх повідомлень. $C = \{c_1 ... c_q\}$ - множина усіх контейнерів. $S = \{(c_1, m_1), (c_1, m_2), ...(c_1, m_n)...(c_q, m_n)\}$ - множина контейнерів-результатів. Множину контейнерів C обирають таким чином, щоб після перетворень контейнер-оригінал і контейнер-результат були максимально схожими. Схожість оцінюється за допомогою функції схожості наступним чином: множина контейнерів C вважається непорожньою множиною, на якій визначена функція $sim(c) \to (-\infty; 1]$, яка і є функцією схожості на множині C. При цьому вважається, що

$$\exists\ x,\ y \in C\text{:} \begin{cases} sim(x,\ y) = 1, & x = y, \\ sim(x,\ y) < 1, & x \neq y. \end{cases} \tag{3}$$

Стеганосистема може вважатися надійною, якщо $sim[c,\ E(c,\ m)] \approx 1$ для всіх повідомлень і контейнерів. Функції кодування і декодування можуть набувати довільних значень (але мають відповідати одне одній), проте на практиці для підвищення надійності на них накладають певні обмеження виду:

$$E(c,\ m) \approx E(c + \delta,\ m), \tag{4}$$

$$D[E(c,\ m)] \approx D[E(c + \delta,\ m)] \approx m. \tag{5}$$

Тобто, модифікація на незначну величину δ контейнера с не повинна впливати на зміст прихованого повідомлення m [1].

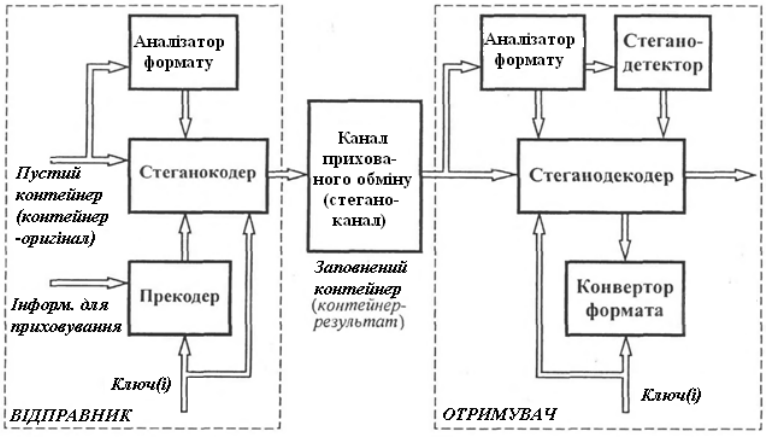
Структура контейнеру. Файл JPEG містить послідовність маркерів, кожен з яких починається з байта 0xFF, що свідчить про початок маркера, і байта-ідентифікатора. Деякі маркери складаються тільки з цієї пари байтів, інші ж містять додаткові дані, які складаються з двухбайтового поля з довжиною інформаційної частини маркера (включаючи довжину цього поля, але не враховуючи двох байтів початку маркера тобто 0xFF і ідентифікатора) і власне даних. Така структура файлу дозволяє швидко відшукати маркер з необхідними даними

(наприклад, з довжиною рядка, числом рядків і числом колірних компонентів стисненого зображення) [2]. В ході експериментальних досліджень було встановлено, що внесення даних між маркерами не призводить до порушення цілісності формату файлу.

Пропонується. метод стеганографічного приховування інформації в графічних контейнерах формату JPEG, принципова відмінність якого від популярних використовуваних підходів полягає у використанні недекларованих зон графічного формату (міжмаркерного простору) для збереження закодованої інформації та дозволяє підвищити ефективність збереження даних при впливі на графічну складову контейнеру. Головною умовою використання міжмаркерного простору є встановлення власного маркеру, що починається з байта 0xFF [3]. Під час приховування повідомлення інформаційні блоки зберігаються за байтами маркерів так, що модифікація структури файлу не призводить до зміни кольорової палітри. Спектральний аналіз зображень показав абсолютну відповідність контейнера-оригінала, підготовленого контейнера та контейнера-результата. Статистичні показники - математичне сподівання, дисперсія і стандартне відхилення - співпадають для трьох контейнерів.

Для підвищення надійності системи повідомлення слід шифрувати одним з криптографічних алгоритмів. Перед розбиттям на окремі інформаційні блоки для подальшого внесення в міжмаркерний простір для шифрованого повідомлення розраховується контрольна сумма, яка також вноситься у контейнер окремим блоком після спеціально створеного маркера.

Передача даних відбувається по відкритому каналу зв'язку з вільною передачею ключа кодування/ декодування даних та контейнеру з закодованими даними.



Висновок. Розглянуто метод стеганографічного приховування інформації в графічних контейнерах формату JPEG, принципова відмінність якого у порівнянні з популярними підходами полягає у використанні недекларованих зон графічного формату для збереження закодованої інформації та дозволяє підвищити ефективність збереження даних при впливі на графічну складову контейнеру. При апробації програмної реалізації даного підходу спектральний аналіз зображень показав абсолютну відповідність контейнера-оригінала, підготовленого контейнера та контейнера-результата. Отримані статистичні показники RGB-складових для трьох зображень однакові. Відтак запропонований метод унеможливлює викриття факту передачі прихованої інформації відкритими каналами зв'язку шляхом аналізу графічної складової контейнеру.

Література. 1. Бельська В.Ю. Метод надлишкового представлення та збереження даних у графічних контейнерах / Бельська В.Ю, Костенко П.П., Сухарев О.Є, Шевченко В.О. // Електромеханічні і енергозберігаючі системи. Щоквартальний науково виробничий журнал. – Кременчук: КНУ– Вип. 4 (12). – 132с. С.113 – 117 2. Структура JPEG, wikipedia.org/wiki/JPEG 3. Декодування формату JPEG, habrahabr.ru/post/102521

Журавлёв П.В.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Выбор архитектуры корпоративных серверных приложений

Вступление. Одним из важных решений при проектировании корпоративных приложений является выбор архитектуры развертывания проекта. Неправильный выбор может ухудшить его структуру и привести к потере времени и увеличению стоимости разработки, затруднить его дальнейшее развитие. Большие программные системы могут разрабатываться как монолитное приложение либо могут быть разделены на отдельно развертываемые компоненты, взаимодействующие друг с другом.

Монолитная архитектура. Корпоративные приложения часто имеют трехслойную архитектуру: клиентский слой с пользовательским интерфейсом, слой данных, серверное приложение, реализующее бизнес логику. Ядро приложения - бизнес логика, которая реализована модулями, определяющими сервисы, объекты предметной области и события. Ядро окружено адаптерами, которые формируют интерфейсы для взаимодействия с другими слоями. Несмотря на модульную архитектуру, приложение упаковано и развернуто как монолит – единый логический исполняемый модуль, компоненты которого разделяют общую память и ресурсы. Такая архитектура имеет следующие особенности, которые обеспечивают ее популярность на начальном этапе развития проекта:

- Простота развертывания – требуется всего лишь скопировать упакованное приложение на сервер и запустить
- Простота тестирования – приложение можно запустить на одном комьпютере и произвести неразрывное тестирование
- Имеется возможность горизонтального масштабирования приложения путем выполнения копий приложения во разных процессов с использованием балансировщика нагрузки

По мере того, как монолитные программные системы разрастаются с добавлением новой функциональности, становится трудным сохранять оптимальную модульную структуру. Увеличивается зацепление компонентов системы что приводит к таким недостаткам:

- Высокая сложность, большая трата времени на исправление ошибок и внедрение новой функциональности
- Высокое время запуска системы и необходимость развертывания всего проекта после внесения небольших изменений вносит дополнительные временные затраты
- Сложность масштабирования
- Низкая надежность, ошибка в одном модуле может влиять на целую систему
- Сложность внедрения нового инструментария и модулей

Микросервисы. Микросервисы представляют собой такой подход к проектированию распределенных систем, при котором их компонентами является множество взаимодействующих друг с другом сервисов с высокой степенью детализации и с собственными жизненными циклами. Такой декомпозиции приложения можно добиться, применяя принципы предметно ориентированного проектирования. Из бизнес модели выделяется множество ограниченных контекстов, каждый их которых описывает свою предметную область. Каждый модуль проектируется в пределах своего ограниченного контекста тем самым реализуя принцип единственной ответственности, что избавляет систему от недостатков монолитной архитектуры. Микросервисы разворачиваются на разных серверах и обмениваются сообщениями через API. Основные преимущества:

- Простая для понимания структура компонентов системы, каждый модуль имеет четко очерченные границы
- Сервисы могут разрабатываться разными командами разработчиков, которые могут ипользовать разнообразные стеки технологий, облегчается процесс внедрения и адаптации новых технологий

- Упрощается процесс масштабирования отдельных модулей, замена модулей становится менее затратной процедурой

Присутствуют также и недостатки:

- Увеличивается стоимость поддержки инфраструктуры развертывания, усложняется управление системой
- Управление данных становится децентрализованным, увеличивается количество проблем, связаных с согласованностью данных
- Поскольку обмен сообщениями между серверами является более затратной операцией, чем взаимодействие компонент программы в пределах одного процесса, возможно падение производительности при интенсивном обмене данными.

Микросервисы работают корректно в том случае, если границы между сервисами являются четкими и стабильными. Поскольку рефакторинг функциональности между серверами сложнее аналогичного в монолитной архитектуре, попытка начать разрабатывать систему с использованием микросервисов в пределах новой или плохо изученной предметной области может привести к неправильному определению ограниченных контекстов.

Применение. Если проект только начинает разрабатываться, более оптимальной является монолитная архитектура, поскольку ее недостатки слабо проявляются в небольших системах, а преимущества в виде простого развертывания и тестирования хорошо подходят для первых этапов разработки.

Микросервисная архитектура подходит в том случае, если предметная область проекта хорошо изучена и система достаточно разрослась. Программные модули можно развернуть в качестве микросервиса в таких ситуациях:

- Необходимо применить новую технологию
- Определенный модуль потребляет большое количество ресурсов, что приводит к снижению производительности всей системы
- Модули потребляют разные типы ресурсов, например такты CPU, оперативную память, сетевой трафик и поддержка сервера, который предоставляет одновременно все требующиеся ресурсы является слишком затратной
- Предприятие организовано в виде множества небольших команд разработчиков

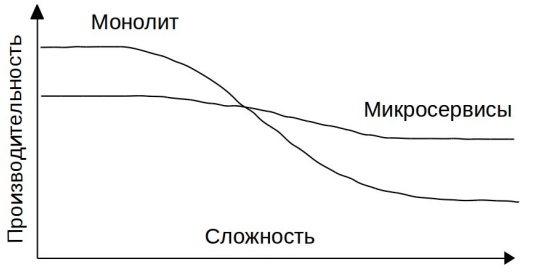


Рис. 1. С повышением сложности системы производительность микросервисов падает медленнее, чем производительность монолитной системы

Вывод. Таким образом риски и затраты сводятся к минимуму при такой стратегии: проект начинает разрабатываться как монолит, далее при необходимости модули постепенно выделяются в микросервисы. При этом монолитное ядро системы может постепенно уменьшиться до размеров микросервиса.

Литература. **1.** Newman, Sam. Building Microservices: Designing Fine-grained Systems. O'Reilly Media, 2015. **2.** Fowler, Martin. Microservices http://martinfowler.com/articles/microservices.html

Забелин С.И.
ННК ИПСА, Киев, Украина

Контрмеры против Arp-Spoofing в SDN

SDN - сети будущего. Программно-конфигурируемые сети (ПКС, англ. Software Defined Networks - SDN) – это актуальный тренд в сетевых технологиях, в которой уровень управления сетью отделён от устройств передачи данных и реализуется программно. Непосредственно пересылку данных выполняют специальные коммутаторы SDN на основе универсальных таблиц потоков, которые формируются и поддерживаются внешним управляющим компонентом - контроллером сети. Ключевая особенность управления SDN в том, что как таблицы потоков в коммутаторах, так и правила функционирования контроллеров доступны для управления внешними программными средствами через специализированные API. [1] Из-за централизованной структуры SDN может стать потенциальной целью злоумышленников. [2]

SDN и сетевые атаки. ARP-spoofing (ARP — poisoning) — разновидность сетевой атаки типа MITM (англ. Man in the middle), применяемая в сетях с использованием протокола ARP. В основном применяется в сетях Ethernet. Атака основана на недостатках протокола ARP. [3]

Для эксперимента на стенде использовалась Arp атака для активного контроля сетевого трафика дезинформированного узла, воздействуя на него по схеме «ложный объект РВС».

Методика исследования. Для исследования был выбран контроллер SDN – POX. Контроллер имеет поддержку высокоуровневых API SDN, такие как виртуализация и обнаружение топологий. POX реализован на языке Python. [4]

Экспериментальные исследования были организованы на основе эмулятора сетей Mininet и контроллера POX. Организация Arp-Spoofing была смоделирована на разных топологиях построения сети и разном количестве сетевого оборудования.

Модифицированный контроллер AntiArpPox. Стандартный контроллер не имеет защит от Arp-Spoofing и с помощью API был модифицирован и назван AntiArpPox. Результатом исследования стал модифицированный контроллер AntiArpPox, в котором устранены уязвимость стандартной версии POX. В AntiArpPox реализованы функции анализа сетевого трафика внутри сети и на основе полученных данных обнаружения Arp-Spoofing, и впоследствии AntiArpPox создаёт фильтры которые, предотвращают возможные Arp-Spoofing.

Исходя из результатов экспериментальных исследований AntiArpPox на незащищенном и защищенном контроллерах SDN в среде Mininet был сделан вывод о его полной работоспособности.

Выводы. Используя преимущества SDN, а именно централизацию управления, гибкость, программное конфигурирование позволяют создавать высокозащищенные контроллеры, что является важным шагом в повышении безопасности сетей. Была разработана методика защиты от Arp-Spoofing и проведены экспериментальные исследования, которые были направлены на подтверждение применения разработанных методов и инструментов в практической работе.

Литература. **1.** T. Nadeau and K. Gray, *SDN: Software Defined Networks*, Washington, O'Reilly Media, 2013. **2.** T. Limoncelli. (2012, June 20). *Openflow: a radical new idea in networking* [Online]. Available: http://queue.acm.org/detail.cfm?id=2305856 **3.** Wikipedia contributors. (2012, April 12). *ARP-spoofing* [Online]. Available: https://ru.wikipedia.org/wiki/ARP-spoofing **4.** Ali Al-Shabibi. (2014, January 25). *POX Wiki* [Online]. Available: https://openflow.stanford.edu/display/ONL/POX+Wiki

Загороднюк А.А.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Использование ORM при разработке корпоративных приложений

Введение. ORM (англ. Object-Relational Mapping, рус. объектно-реляционное отображение) — технология программирования, которая связывает базы данных с концепциями объектно-ориентированных языков программирования, создавая «виртуальную объектную базу данных». Существуют как проприетарные, так и свободные реализации этой технологии. [4]

Вся прелесть и сила ORM в том, что он предоставляет разработчику возможность манипулировать данными на уровне объектов. Можно сказать, что выстраивается некий "мост" между реляционными данными и ООП. Этот мост позволяет абстрагироваться от реализации. Вы можете писать красивый, удобный и качественный объектно ориентированный код, который Вам привычен. Это сильно упрощает тестирование проекта. У ORM 3 основных преимущества перед обычными(чистыми) запросами к базе данных.

Принцип работы ORM. ORM имееют ключевую особенность - отображение, которое используется для привязки объекта к его данным в БД. ORM как бы создает и использует «виртуальную» схему базы данных в памяти, с помощью которой манипулирует данными уже на уровне объектов. Отображение показывает как объект и его свойства связанны с одной или несколькими таблицами и их полями в базе данных. ORM использует информацию этого отображения для управления процессом преобразования данных между базой и формами объектов, а также для создания SQL-запросов для вставки, обновления и удаления данных в ответ на изменения, которые приложение вносит в эти объекты. Мы так же можем на уровне сущностей создавать методы, которые являются аналогами триггеров в СУБД(beforeSelect, afterInsert и т.д.). Говоря конкретнее, использование ORM решает проблему так называемой парадигмы «несоответствия», которая гласит о том, что объектные и реляционные модели не очень хорошо работают вместе. Реляционные базы представляют данные в табличном формате, в то время как объектно-ориентированные языки представляют их как связанный граф объектов. Основные проблемы и несоответствия возникают во время сохранения этого графа объектов в реляционную базу или его загрузки. ORM же это с легкостью решает.

Преимущества ORM.

- *Краткость.* Первый плюс - это краткость записи. Потому что можно сделать вот так: Cars.selectAsObject(); Вот так мы выберем все автомобили из нашей таблицы. Просто и удобно. Аналог на SQL: SELECT * FROM Cars. После того, как мы сформировали запрос, его потом ещё и исполнить надо, и результат получить. То есть строк кода будет 3-4, вместо одной.
- *Универсальность.* Заключается в том, что мы можем взять и сменить PostgreSQL на Oracle (если у нас есть драйвер для данной СУБД. Если нет - то пишется он относительно просто и быстро) и наше приложение готово к работе. Это очень удобно использовать на платформах, используя СУБД заказчика. У разных СУБД немного разный синтаксис, и поэтому без ORM надо будет править часть запросов. А это значит, что придется всё приложение перелопачивать, а если нет тестов, то можно только надеяться, что мы всё нашли и все будет работать как надо. И всё это вместо того, чтобы поменять одну настройку в конфиге.
- *Объекты.* Ну и последний плюс - в том, что на выходе мы получаем не просто массив с данными (что ORM тоже позволяет делать), а полноценные объекты, внутри которых есть и методы и свойства, которые были определены у данного класса. Это очень сильно ускоряет процесс разработки, что является критичных в корпоративных приложениях.

Литература. **1.** Dataobjects, dataobjects.net. **2.** Internetika, internetka.in.ua/orm-intro. **3.** Madgeek, madgeek.com/Articles/ORMapping/EN/mapping.htm. **4.** Wikipedia, en.wikipedia.org/wiki/Object-relational_mapping.

Ивченко Д.А.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Использование мобильных устройств для мониторинга состояния человека с астмой

Вместе с диабетом и хроническими сердечно-сосудистыми заболеваниями хронические легочные заболевания в виде бронхиальной астмы являются одними из самых распространенных на планете. По имеющимся оценкам 300 миллионов человек во всем мире страдают от астмы [1]. Около 10% от общего количества – это дети. Только в США их насчитывается 25 млн., а в Соединенном Королевстве -5,4 млн.

К 2025 году ожидается увеличение числа пациентов с астмой более чем на 400 млн. Каждый год регистрируется всё больше больных с тяжёлым течением бронхиальной астмы, в том числе – среди детей. С ростом распространенности астмы растут расходы системы здравоохранения на стационарное лечение, возрастают нагрузки на медицинский персонал [2].

Астма определется как хроническая гиперчувствительность дыхательных путей, которая проявляется в виде повторяющиися периодов обострений(приступов). Симптомы приступов включают одышку, затрудненное дыхание и кашель. Они возникают как следствие сокращения гладкой трахеобронхиальной мускулатуры, отека и гиперсекреции слизи, приводящих к сужению дыхательных путей. Больше половины всех астматиков в мире страдают от аллергии, индуцированной аллергенами из окружающей среды: пыльцой с растений, дымом, бытовой пылью, клещами, перьями, мехом и др.

Астма по частоте приступов подразделяется на четыре стадии в соответствии с руководящими принципами GINA (Global Initiative Asthma) [3].

1. Начальная: прерывистый легкий кашель и хрипы менее чем два раза в неделю, ночные приступы менее чем два раза в месяц;
2. Мягкая: приступы случаются, по крайней мере, один раз в неделю;
3. Средняя: сужаются проходы в легких и трахеи, активизируются кашель и свистящее дыхание;
4. Тяжелая: приступы постоянно случаются как днем, так и ночью и могут длиться более нескольких дней, постоянный кашель и хрипы.

Цели лечения астматиков с помощью мобильных приложений на дому.

- Предупреждение пациентов о потенциально возможных приступах, исходя из анализа данных о состоянии пациента, наличии аллергенов и загрязнителей.
- Предоставление пользователю информации о том, как лечится астма, где и что вызывает приступы астмы в целях более эффективного предупреждения и лечения осложнений астмы.
- Мониторинг применения пациентом ингаляторов с встроенными системами GPS. Поддержка ежедневного здоровья и активности с учетом индивидуальных особенностей пациента и состояния окружающей среды (уровня вредных выбросов в атмосфере, наличия цветочной пыльцы, параметров температуры и влажности воздуха и т.д.).
- Предотвращение или устранение последующих осложнений от болезни.

Виды помощи астматикам, оказываемой мобильными приложениями.

- Измерение пикового объема выдыхаемого воздуха из легких пациента.
- Разработка плана действий при астме, выполняемая лечащим врачом пациента.

Таким образом, план действий при астме может включать.

1. Список возбудителей астмы, которые провоцируют проявление симптомов астмы, и как их избежать.
2. Список пиковых показаний пикфлоуметра и его зон, учитывающий личностные физиологические особенности пациента.

3. Список рутинных симптомов астмы, таких как кашель, свистящее дыхание, стеснение в груди, одышка и избыток слизи, а также то, что должен делать пациент, если эти симптомы возникают.
4. Наименования и ежедневную дозу лекарства против астмы, которое должно быть принято, даже если у пациента нет симптомов астмы.
5. Наименования и дозы быстродействующего или спасательных лекарств, которые должны быть приняты при появлении симптомов астмы.
6. Наименования и дозы препарата неотложной помощи, который должен быть принятым, когда пациент испытывает приступ астмы.
7. Телефоны экстренной связи и места оказания неотложной помощи.
8. Ведение дневника при астме.
9. Удаленный мониторинг состояния пациента лечащим врачом.
10. Интегрирование в социальные сети, что дает возможность делится своим прогресов в борьбе с болезнью.
11. Синхронизацию записей дневника при астме с персональной электронной медицинской картой ПЭМК.

Примеры уже существующих решений.

Приложение	Описание
BreatheEasy	BreatheEasy представляет собой приложение, которое разработано для лечения астмы врачами в Новой Зеландии. Это приложение позволяет контролировать пиковые значения потока воздуха и напоминать больным, чтобы они использовали их ингаляторы. Приложение имеет возможности, позволяющие создавать для астматиков дневник заболевания. Такой дневник симптомов полезен для медицинских работников, чтобы дистационно контролировать прогресс лечения.
My Asthma Log	Это приложение позволяет записывать историю всех приступов астмы у пациента и выполнение назначений доктора по приему препаратов. Приложение также предоставляет пациентам информацию и фотографии противоастматических препаратов. Есть интеграция с социальными сетями и сайтом астмы Великобритании.
Asthmapolis	Asthmapolis– это бесплатное приложение, которое можно скачать для вашего смартфона. Трекер крепится к ингалятору и подключается к смартфону через Bluetooth. Это позволит отслеживать местонахождение вашего ингалятора. Собранные данные могут также информировать, насколько долго вы пользовались ингалятором. Тенденции изменения показателей могут помочь врачам определить степень симптомов, чтобы вы могли лучше подготовиться к последующим приступам.

Литература. **1.** J. Finkelstein, M. R. Cabrera, and G. Hripcsak, “Internet-based home asthma telemonitoring: can patients handle the technology?” Chest, vol. 117, no. 1, pp. 148–155, Jan 2000. **2.** Национальная программа «Бронхиальная астма у детей. Стратегия лечения и профилактика» (3-е изд., испр. и перераб.). М.: 2008; 108. **3.** GINA Classification of Asthma Severity http://goo.gl/7l5xFX

Калінін В.В., Боярінова Ю.Є.
НТУУ "КПІ", Київ, Україна

Огляд існуючих нейропакетів та перспективи їх застосування

Розглянуто моделі нейропакетів та визначені переваги нейропакету NeuroSolution.

Для вирішення різноманітних задач необхідні моделі нейроних мереж. Модель нейроної мережі визначається моделями нейронів та структурою зв'язків мережі. Для моделювання роботи та навчання нейроної мережі використовуються нейропакети. Нейропакетом(НП) називається програмна система, що емулює середу нейрокомп'ютера(НК) на звичайному комп'ютері. Відповідно до неї НП діляться на сім основних класів.

1. Інструментарій побудови НП.
2. Універсальні НП. Під універсальністю розуміється можливість моделювання штучної нейронної мережі (ШНМ) різної структури і з різними алгоритмами навчання.
3. Спеціалізовані НП, що використовують нейрони складної функціональності і включають спеціалізовані засоби для: обробки зображень; розпізнавання образів; розпізнавання рукописних і друкованих символів; розпізнавання мови; управління динамічними системами; фінансового аналізу та ін.
4. Нейронні зв'язки.
5. Пакети генетичного навчання ШНМ.
6. Пакети нечіткої логіки, що використовують ШНМ.
7. Інтегровані пакети.

Найбільший інтерес викликає останній клас НП тому, що він призначений для створення програмних модулів, що реалізують модель ШНМ з ядерної організацією. Процес розробки складається з чотирьох етапів: 1) візуальне проектування структури і топології ШНМ; 2) визначення синаптичної карти і функцій активації нейронів; 3) навчання побудованої ШНМ; 4) тестування навченої ШНМ (в тому числі оцінювання швидкості роботи ШНМ).

Приклади нейропакетів цього класу: Neuro Office дозволяє сформувати навчену ШНМ з програмним інтерфейсом на базі технології COM (Component Object Model — об'єктна модель компонентів). Пакет включає три основні компоненти:

NeuroView + - редактор структури і топології ШНМ;

NeuroEmulator - засіб для визначення синаптичної карти і функцій активації нейронів, навчання і тестування ШНМ;

ActiveX-елемент «нейронна мережа» - компонент, що вбудовується в додатки і реалізує програмний інтерфейс з ШНМ.

У NeuroView + можна одночасно редагувати кілька ШНМ. Розміри формованих ШНМ програмно не обмежені. У файл з описом ШНМ також заносяться дані при графічному поданні мережі на екрані дисплея і коментарі розробника. Прикладом використання НП може бути банківська діяльність, в якій використовують нейронні мережі для контролю кредитних карток. Гарним прикладом може бути нейропакет NeuroSolution фірми NeuroDemension Inc. Пакет NeuroSolution призначений для роботи в операційних системах Windows. НП забезпечується генератором вихідного коду і засобами. що дозволяють використовувати зовнішні модулі при проектуванні і навчанні нейроної мережі, що робить його дуже зручним в використанні. Пакет підтримуєпрограми, створені за допомогою C++. NeuroSolution є гнучкою відкритою системою, яку можна доповнювати і модифікувати. Все вище сказанедозволяє зробити висновок в доцільності використання НП NeuroSolution в банківський діяльності.

Література. 1. Мілль Д. С. Автоматическое порождение гипотез в интеллектуальных системах / Д. С. Мілль, Ч. С. Пірс, К. Р. Поппер. – Москва: Книжний дім "Ліброком 2009. – 528 с. **2.** Башмаков І. О. Интеллектуальные информационные технологии / І. О. Башмаков, О. І. Башмаков. – Москва: МГТУ имені Н.Е. Баумана, 2005. – 304 с. **3.** Заєнцев І. В. Нейронные сети. Основные модели. / І. В. Заєнцев. – Воронеж, 1999. – 76 с.

Калінін В.В., Черненко П.Р.
НТУУ "КПІ", Київ, Україна

Метадані у алгоритмах пошуку

Розглянуто властивості метаданих, проблематика та їх використання для пошуку.

У сучасному суспільстві обмін і отримання інформації відіграє важливу роль, тому знаходження необхідної інформації за найкоротші терміни - це завдання, яке може бути вирішене різними способами. При цьому, методи пошуку можуть значно розрізнятися, як по ефективності і якості пошуку, так і по типу інформації, яка буде отримана. Тому для більш вдосконаленого пошуку слід використовувати нові методи.

Розглянемо основні методи пошуку інформації в Інтернеті. Насамперед, це використання пошукових машин. В даний час цей метод - один з основних і фактично єдиний при проведенні попереднього пошуку. В результаті пошуку з'явиться список ресурсів Інтернету, який потрібно детально розглянути.

Застосування пошукових машин засноване на використанні ключових слів, що передаються пошуковим серверам в якості аргументів пошуку. В якості методу може розглядатися і безпосередній пошук з використанням гіпертекстових посилань, метаданих.

Метадані (metadata) - це інформація про документ, що розуміється ПК та володіє властивістю внутрішньої інтерпретації. В загальному випадку метадані являють собою характеристики, що описують дані для їх ідентифікації, пошуку, оцінки, управління.

Протягом усього процесу пошуку слід поєднувати різні види і методи пошуку в двох основних напрямках: з одного боку, від загальних класифікаційних пошукових систем до предметно - орієнтованих вказівників ресурсів, далі до конкретних документів; з іншого - від унікальних слів і фраз запиту до предметно - орієнтованих вказівників ресурсів та / або до конкретних документів.

Одним з найбільш природних шляхів вирішення проблеми пошуку вважаємо концентрацію не на наборах даних, в яких слід вести пошук, а на асоційовані з ними набори метаданих, що містять широкий спектр зовнішніх характеристик, за якими споживач може досить просто сформувати свого роду запит для пошуку необхідних документів.

Набори метаданих можуть, зрозуміло, створюватися спеціальними програмними комплексами, що входять до складу пошукових систем, шляхом автоматичної обробки просканованої інформації. Однак, незрівнянно більшої ефективності можна було б досягти, організувавши зберігання метаданих безпосередньо в інформаційних документах.

Для цих цілей як не можна краще підходять XML(Extensible Markup Language)-технології, які отримали останнім часом широке поширення. Так, наприклад, до складу веб-сайту міг би входити XML-документ, що містить якійсь уніфікований (або стандартизований) опис структури і основних характеристик цього сайту, призначених для побудови різного роду класифікаторів.

На сьогоднішній день основними такими мовами є Розширювана Мова Розмітки XML і Засоби Опису Ресурсів RDF (Resource Description Framework).

Пошукові запити, що використовують метадані можуть врятувати користувача від зайвої ручної роботи з фільтрації. Інформуючи комп'ютер про те, які елементи даних зв'язані і як ці зв'язки враховувати, стає можливим здійснювати достатньо складні операції по фільтрації та пошуку.

Література. **1.** Григорьев А.Н., Ландэ Д.В. Адаптивный интерфейс уточнения запросов к системе контент-мониторинга InfoStream//Труды Международного семинара "Диалог'2005". - 2005. - С. 109-111 **2.** Gianna M. Del Corso, Antonio Gull?, Francesco Romani. Ranking a Stream of News. In Processing of the 14th International World Wide Web Conference, 2005 (www2005.org/cdrom/docs/p97.pdf)

Капустін В.В., Маслянко П.П.
Національний технічний університет України "КПІ", факультет прикладної математики, Київ, Україна

Система мультипоточного тестування веб застосунків на основі алгоритмів паралельних обчислень

ВСТУП. Програмне забезпечення (англ.software) – сукупність програм системи обробки інформації і програмних документів, необхідних для експлуатації цих програм [1].

Розрізняють 2 види тестування: ручне та автоматизоване.

Автоматизоване тестування з кожним днем стає все більше і більше поширеним, зокрема у великих проектах. Для автоматизованого тестуванная постає проблема, коли кількість автоматизованих тестів стає достатньо великою, що призводить і до збільшення часу, який необхідний для отримання результату по всім тестам [2].

В цій ситуації одним з рішенням є: запуск тестів в декілька потоків. Тобто, вся група тестів розбивається на відповідну кількість потоків, яку задає сам користувач. Далі вся група тестів розбивається на відповідну кількість підгруп, у зв'язку з чим час проходження тестів зменшується.

ПОСТАНОВКА ЗАДАЧІ. Об'єкт дослідження: алгоритми паралельних обчислень, види програмного забезпечення, типи програмного забезпечення, методи тестування, види тестування, процеси паралельного тестування.

Предмет дослідження: розробка системи мультипоточного тестування веб застосунків на основі алгоритмів паралельних обчислень з метою зменишити час, відведеного на тестування програмного продукту.

Метою роботи є розробка системи мультипоточного тестування веб застосунків на основі алгоритмів паралельних обчислень з метою зменшення часу необхідного на тестування програмних продуктів. Забезпечення можливості запуску тестів в декілька потоків, можливість перезапуску тестів, що не спрацювали, генерація HTML та XML звітів.

Результати роботи: система мультипоточного тестування веб застосунків з можливостями перезапуску тестів, що не спрацювали, генерації HTML та XML звітів, потужною підсистемою фільтрації категорій тестів.

На рисунку 1 представлено структуру системи паралельного тестування.

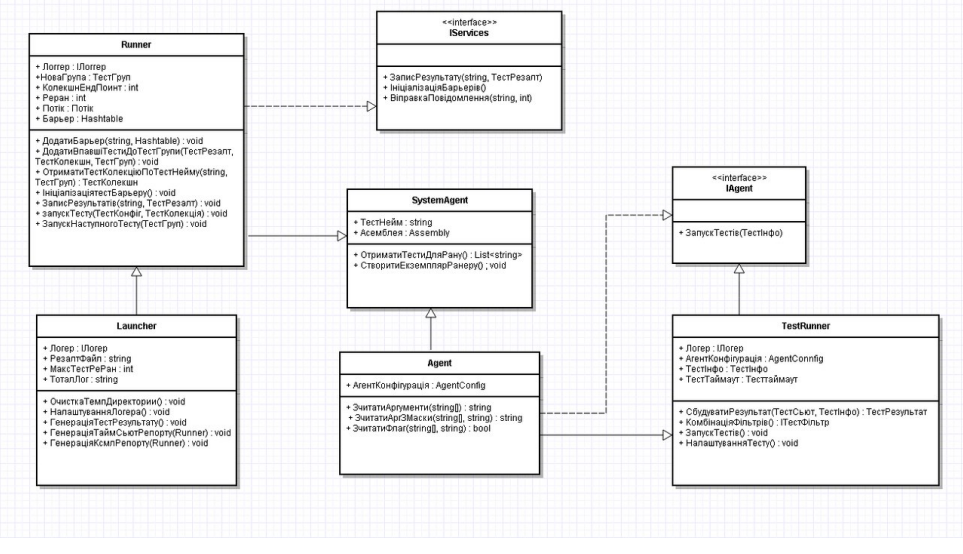


Рис. 1. Стркутурне представлення системи

- Launcher – відповідальний за створення Runner. Також сценарії з тестами зчитуються за допомгою Launcher;
- Runner – запускає тести та забезпечує механізм синхронізації та збору результатів;
- Agent – має можливість запускати тести віддалено, створюючи екземпляри SystemAgent;
- System agent – отримує тести, які необхідно запустити та створює екземпляр Test runner;
- Test runner – запускає тести та відправляє їх до Runner;

На рисунку 2 представлено логіку роботи системи. Блок 'Початок запуску тестів' відповідає за збір всіх тестів, які необхідно запустити. Далі весь набір тестів розбиваться на 'Початок запуску тест групи', кількість таких груп залежить від того в скільки потоків користувач хоче запускати тести. Потім кожна група запускає тести, які потрапили до відповідної групи.

Далі ми отримаємо результати запуску тестів по кожній з підгруп. Після цього, відбувається об'єднання результатів по кожній з груп та генерується звіт по результатам тестування.

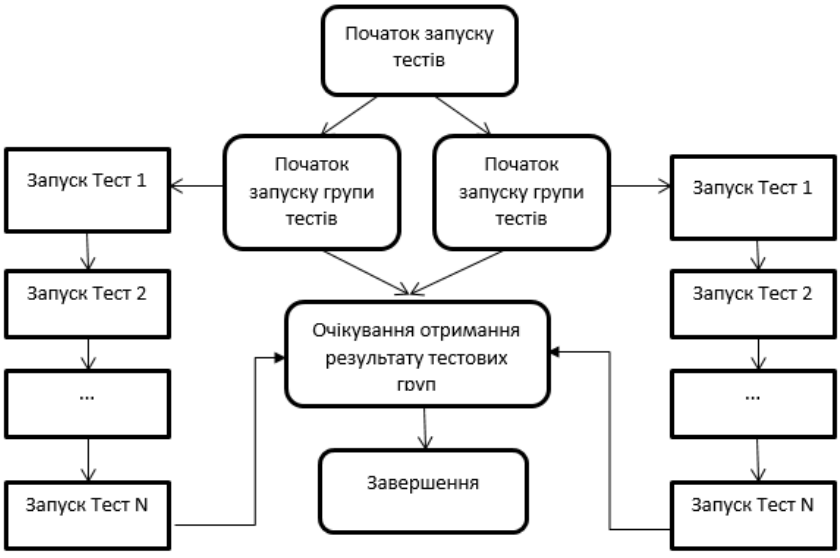


Рис. 2. Логіка роботи системи

ВИСНОВКИ. Було проаналізовано необхідність тестування програмного продукту, були виділені переваги мультипоточного тестування. В результаті аналізу моделей, методів, засобів тестування було встановлено, що на сьогоднішній день нема системи, яка б могла повністю вирішити поставлену задачу.

Розроблена система відрізняється від існуючих рішень наступними показниками: система безкоштовна, генерація звіту у форматі XML та HTML, можливість перезапуску тестів, що не спрацювали, потужна підсистема фільтрація категорії тестів, фіксація часу, який був затрачений на запуск тестів по кожному з потоків.

Отже, система дозволяє вирішити раніше некоректно вирішені задачі, такі як:

- Можливість генерації звіту у форматі XML та HTML;
- Можливість перезапуску тестів, що не спрацювали;

ЛІТЕРАТУРА.

Література. **1.** Dorothy Graham, Erik Van Veenendaal, Isabel Evans, Rex Black. Foundations of software testing. - London, 2007 – 7p. **2.** Selenium Grid. [Електронний ресурс]. – Режим доступу: http://www.seleniumhq.org/projects/grid/

Карпенко О.В., Голубовський О.І.
Факультет прикладної математики НТУУ "КПІ", Київ, Україна

Методи забезпечення якості обслуговування в мережах IP

На сьогоднішній день разом зі збільшенням швидкостей передачі даних в комп'ютерних мережах зростає частка інтерактивного трафіку, чутливого до параметрів середовища транспортування. Тому завдання забезпечення якості обслуговування QoS (Quality of Service) стає все більш актуальним.

Під QoS розуміють ймовірність передачі до кінцевої станції потоків або окремих пакетів з виконанням заданих вимог або ймовірність того, що мережа відповідає цим вимогам (гарантує їх виконання).

Комп'ютерні прикладні сервіси, які генерують мережевий трафік, мають різні вимоги до обслуговування створеного трафіку і до робочих характеристик мережі, в якій він транспортується.

Здатність мережі забезпечувати різні рівні обслуговування, що вимагаються тими чи іншими мережевими програмами, поряд з проведенням контролю за різними характеристиками передачі (пропускною спроможністю, затримкою, джитером, втратою пакетів тощо) може бути класифікована за трьома перерахованим нижче категоріями [1].

1. Негарантована доставка даних (best-effort service), яка забезпечує зв'язність вузлів мережі без гарантії часу і самого факту доставки пакета в точку призначення. При цьому відкидання пакетів може статися тільки в разі переповнення буфера вхідної або вихідної черги в маршрутизаторах.

2. Диференційоване обслуговування (differentiated service), що передбачає розділення трафіку на класи на основі вимог до якості обслуговування. Кожен клас трафіку диференціюється і обробляється мережею відповідно до заданих для цього класу механізмів QoS.

3. Інтегрований сервіс (integrated service), який забезпечує наскрізну якість обслуговування і передбачає резервування мережевих ресурсів з метою забезпечення специфічних вимог до обслуговування трафіку. Відповідно до інтегрованого сервісу виконується попереднє резервування мережевих ресурсів по всьому маршруту передачі.

В роботі проведені дослідження і порівняльний аналіз основних методів забезпечення якості обслуговування в мережах IP, а саме: диференційованого обслуговування та інтегрованого сервісу.

У ході даного дослідження, а також спираючись на числові експериментальні дані, отримані шляхом математичного моделювання, в роботах [2, 3], нами було встановлено наступні закономірності:

- модель Diffserv є більш гнучкою порівняно з Intserv, оскільки при великих значеннях навантаження забезпечує менший час затримки передачі пакетів;

- модель Intserv має низьку ефективність масштабування мережі та має високі вимоги до продуктивності маршрутизаторів, порівняно з Diffserv;

- модель Intserv забезпечує кращі показники продуктивності (затримки передачі та втрати пакетів) при невеликих кількостях потоків трафіку;

- модель Diffserv забезпечує кращі показники затримки обробки трафіку в чергах маршрутизаторів при низьких числових значеннях пропускної спроможності каналу.

Література. **1.** Шринивас Вегешна. Качество обслуживания в сетях IP (Cisco). – 2003. – 368 С. **2.** Анализ показателей качества обслуживания услуг Triple Play infocom.uz/2012/02/27/analiz-pokazatelej-kachestva-obsluzhivaniya-uslug-triple-play/ **3.** Герцій О. А. Математичне моделювання та оцінка параметрів якості зв'язку IP-мереж./ О. А. Герцій //Збірник наукових праць Державного економіко-технологічного університету транспорту. Сер.: Транспортні системи і технології. - 2013. - Вип. 22. - С. 144-151.

Кириленко Е.Г., Стрельчук Е.А.
Национальный аэрокосмический университет им. Н.Е. Жуковского ХАИ, Харьков, Украина

Методический инструментарий для оценки удобства использования пользовательского интерфейса информационного веб-сайта

Несмотря на обилие дизайнерских рекомендаций и руководств для создания работоспособного веб-сайта, оценка удобства использования пользовательского интерфейса (ПИ) информационных веб-сайтов по-прежнему является актуальной проблемой. По мнению экспертов в области проектирования сайтов, около девяноста процентов сайтов являются недостаточно удобными для использования из-за устойчивого роста новых сайтов и острой нехватки специалистов, которые обеспечивают удобство использования ПИ веб-сайта. В нашем исследовании удобством использования ПИ будем считать свойство системы, продукции или услуги, при наличии которого пользователь может применить продукцию в определенных условиях использования для достижения установленных целей с необходимой результативностью, эффективностью и удовлетворенностью [1].

В зависимости от цели оценки исследователи выделяют три подхода оценки удобства использования ПИ веб-сайта. Первый подход включает методы оценки удобства использования с участием пользователя. Основной целью методов этого подхода является оценка и выявление проблем удобства использования ПИ веб-сайта. К преимуществам подхода относятся возможность применения на ранних стадиях разработки, а также возможность получения количественных показателей оценки удобства использования. Недостатком метода считают сложность определения количества пользователей, необходимых для выявления потенциальных проблем удобства использования.

Второй подход включает методы оценки удобства использования с участием экспертов. Спецификой этого подхода является то, что эксперты анализируют интерфейс и оценивают удобство использования ПИ веб-сайта, используя руководящие принципы разработки ПИ, стандарты проектирования ПИ, задачи пользователей или свой опыт. К преимуществам подхода можно отнести возможность использования методов на ранних стадиях процесса разработки веб-сайта, быструю обратную связь с дизайнерами. К недостаткам – необходимость наличия опытных экспертов и неоднозначность трактовки принципов успешного дизайна, эвристик разными экспертами.

Третий подход включает методы, в которых в качестве инструмента оценки удобства использования ПИ веб-сайта используют программное обеспечение (ПО). Эти методы для оценки удобства использования ПИ как отдельных веб-страниц, так и веб-сайта в целом используют большой набор измеряемых объективных количественных параметров, связанных с навигацией, информационным наполнением, дизайном веб-сайта. Эти методы дизайнеры могут использовать как на ранних стадиях проектирования интерфейса веб-сайта, так и в процессе уточнения дизайна.

В рамках проводимого исследования была разработана модель оценки удобства использования ПИ информационного веб-сайта, которая включает пять параметров оценки удобства использования ПИ веб-сайта, двадцать пять метрик их измерения и позволяет оценить три компоненты ПИ веб-сайта: информационную, навигационную, графическую. Разработано специализированное ПО для автоматического расчета метрик оценки удобства использования ПИ информационного веб-сайта (ИВС). Предложенный методический инструментарий позволил снизить уровень субъективизма оценки за счет автоматизации измерения метрик удобства использования ПИ ИВС. На примере ИВС отечественных университетов была проведена оценка удобства использования ПИ веб-сайтов, выявлены и проанализированы проблемы удобства использования ИВС, а также разработаны рекомендации по решению проблем.

Литература. 1. ISO 9241-210. Ergonomics of humansystem interaction. Part 210: Human-centred design for interactive systems [Internet source] – http://www.iso.org/iso/catalogue_detail.htm.

Киричек Г.Г., Котов Р.О.
Запорізький національний технічний університет, Запоріжжя, Україна

Автоматичне розгортання мережних сервісів

Приводяться рішення, які націлені на скорочення часу та підвищення якості розгортання мережних сервісів, з використанням автоматичного тестування вихідного коду.

Вступ. Організація процесу розробки програмного забезпечення включає декілька етапів [1]: передпроектна підготовка; специфікація вимог до системи; проектування, розробка, тестування та впровадження системи. Складність програм порівняна зі складністю сучасних машин, тому вимагає від розробника постійного контролю якості коду сервісу, методом тестування. Виходячи з аналізу основних завдань розробки [1-3], можна зробити висновок, що участь людини повинна бути найменшою в процесі виконання завдань. Базуючись на дослідженнях прийнято рішення реалізувати систему автоматичного тестування та розгортання мережних сервісів з тестуванням конфігураційних параметрів, використовуючи її запуск.

Постановка завдання. Проаналізувавши середовища розробки програмних продуктів маємо сервіси тестування та автоматичного розгортання проектів [2, 4-5] - сервіси безперервної інтеграції. У кожного з них є алгоритм роботи, який дозволяє автоматизувати вище названі процеси. Відсутність поєднань готових рішень для розгортання великих систем, приводить до написання подібних скриптів багатьма розробниками [1-4]. Тому прийнято рішення поєднати системи безперервної інтеграції з сервісами, призначеними для спрощення розгортання, а саме систем автоматичного конфігурування. При розробці програм важливим є використання систем контролю версій. В якості такої системи обрано Git, як провідної серед розробників [6]. Проаналізувавши тенденції розгортання додатків зроблено висновок, що найчастіше для цього використовуються хмарні технології [7-8], які зменшують витрати на конфігурацію та підтримку працездатності серверів і мережного обладнання, збільшуючи доступність додатків за рахунок готових рішень провайдерів [9]. Актуальність обумовлюється і тим, що в Україні вже існують проекти переходу на хмарні технології IT-інфраструктур державних органів.

Реалізація системи. Для отримання обчислювальної потужності при розгортанні проекту обираємо провайдера послуг хмарного хостингу або віртуального виділеного серверу. Підприємства активно впроваджують SaaS, використовуючи хостинги та послуги при зберіганні інформації. Доставка SaaS-додатків через хмару надає їм можливість звільнитись від проблем конфігурування [9]. Після аналізу хмарних технологій [8] та сервісів, обрано Amazon Web Services. Перехід до безперервної інтеграції знижує трудомісткість інтеграції і робить її передбачуваною при усуненні помилок [4]. Для розгортання та тестування працездатності проекту обрано CircleCI, який інтегрується до системи контролю версій, підтримує налагодження неправильного розгортання. Конфігурація розгортання і тестування знаходиться в проекті.

В ролі системи керування конфігураціями обрано Ansible, яка переконфігуровує систему додаванням лише декількох нових рядків у сценарій [8]. Оскільки при тестуванні та налагодженні скриптів систем керування конфігураціями використання хмари є дорогим, а пересилка даних займає багато часу, вирішено, для оцінювання правильності конфігурування, ці процеси проводити локально. Для цього використовуємо Vagrant, який на базі провайдерів віртуальних машин створює кластери на локальній машині і дозволяє конфігурувати кластер за допомогою вбудованого модулю provisioner з підтримкою Ansible [10]. Спираючись на розроблену модель системи (рис.1), наводимо етапи її реалізації. Перед початком роботи створюємо обліковий запис AWS. Після реєстрації, для конфігурування сервісів за допомогою веб-інтерфейсу, маємо доступ до консолі менеджменту. Оскільки для розгортання додатку потрібні комп'ютери, зконфігуруємо доступ до них. Для реалізації безпеки створюємо ключі та виконуємо налаштування ролей доступу до портів і

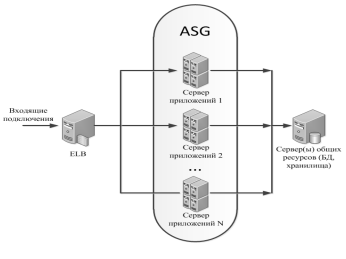


Рис. 1. Модель системи

конфігуруємо спільні ресурси. Так як розгортання проекту виконується автоматично, задаємо ролі доступу до сервісів Amazon - Identity and Access Management і надаємо права новому акаунту. При цьому генеруються повноваження автентифікації, які складаються з ідентифікатору доступу і паролю. Далі використовуємо сервіси надання віртуальних станцій Elastic Compute Cloud (EC2) для створення кластеру, автоматичного балансування (Elastic Load Balancing) при розподіленні навантаження між машинами в кластері і для аналізу метрик віртуальних машин (CloudWatch), а також автоматичного горизонтального масштабування машин, яке полягає у зміні кількості доступних серверів Auto Scaling. Окремо використовуємо сервіс створення та збереження знімків машин Amazon Machine Image (AMI). Конфігурація системи безперервного розгортання знаходиться у кореневому каталозі репозитарію у файлі circle.yml, а скрипт тестування у performtest.sh. Оскільки розгортання проекту виконується за допомогою скриптів з іншого репозиторію, то його додано як підмодуль програми і він клонується у директорію deployment проекту. Для зв'язку з проектом проводимо конфігурацію з веб-консолі CircleCI. При підготовці серверу встановлюємо програмне забезпечення для роботи з AWS, розширюємо набір функцій Ansible для розгортання кластеру, модифікуємо права доступу до AWS і ключі машин, які створені раніш. Конфігураційні скрипти Ansible розміщуються у файлах, написані на мові розмітки YAML і мають назву playbooks. Після кластеру створюється сервер з параметрами, які відповідають завданню з ключем та отримується адреса сервера для подальшого використання. При розгортанні можна використовувати сервіси [9], це дозволяє уникнути конфігурування Ansible. Далі створимо образ машини для запуску автоматичного масштабування (Auto Scaling). Після цього модифікуємо балансувальник навантаження, який розподіляє запити між серверами і слідкує за доступними на даний час машинами. Потім замінемо кластер або створимо новий з конфігурацією автоматичного масштабування. Далі очистимо образи і конфігурацій запуску і зупинимо сервер. При тестуванні скриптів на локальній машині використовуємо Vagrant, який створює задану кількість віртуальних серверів. Конфігурація знаходиться у файлі Vagrantfile і створює віртуальну машину. Для виконання запуску віртуального серверу використовуються команда vagrant up, яка виконує запуск і скрипти Ansible, а при ввімкненому сервері для конфігурації використовуємо команду vagrant provision. Автоматичне конфігурування і розгортання кластеру виконує скрипт deploy.sh.

Висновки. В ході роботи розроблено систему автоматизації процесу тестування і розгортання додатків будь-якої складності в середовищі віртуальної приватної хмари. Наведено метод тестування, з використанням розгортання автоматично масштабованого кластеру, який доводить працездатність системи і може використовуватись при розробці у подібних проектах.

Література. **1.** Build-Deploy-Test. Непрерывная интеграция – [Електронний ресурс] Режим доступу: habrahabr.ru/company/icl_services/blog/262173/. **2.** Непрерывная интеграция в процессе гибкой разработки – [Електронний ресурс] Режим доступу: www.ibm.com/developerworks/ru/library/r-continuous-integration-agile-development/. **3.** Киричек Г.Г. Модель оцінки плагіату програмного коду на основі системи контролю версій / Г.Г.Киричек,О.О.Киричек // Східно-європейський журнал передових технологій.– №2/2(56).–2012. –С.25-28. **4.** TeamCity – [Електронний ресурс] Режим доступу: jetbrains.ru/products/teamcity/. **5.** Система управления конфигурацией Ansible – [Електронний ресурс] Режим доступу: blog.selectel.ru/sistema-upravleniya-konfiguraciej-ansible. **6.** The Simplest Way to Build and Deploy Web Applications to the Cloud? – [Електронний ресурс] Режим доступу: www.ctl.io/blog/post/the-simplest-way-to-build-and-deploy-web-applications-to-the-cloud. **7.** Five requirements for deploying an application in a public cloud – [Електронний ресурс] Режим доступу: searchcloudcomputing.techtarget.com/tip/Five-requirements-for-deploying-an-application-in-a-public-cloud. **8.** Amazon Web Services – [Електронний ресурс] Режим доступу: aws.amazon.com/ru/. **9.** Облачные решения (SaaS) – [Електронний ресурс] Режим доступу: store.softline.ru/saas/. **10.** Хамбл Д. Непрерывное развертывание ПО / Д.Хамбл, Д.Фарли.– М.: «Вильямс», 2011.– С.432.

Кириченко В.Е., Терентьєв О.М., Связінська Н.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Застосування наївного та дереводоповненого байєсівський класифікаторів для прогнозування кредитоспроможності фізичних осіб

Вступ. На сьогоднішній день задача інтелектуального аналізу даних є однією з найбільш актуальних, оскільки у сучасному світі все частіше виникає необхідність роботи з великими об'ємами даних. У даній роботі розглядається один із методів вирішення даної задачі – побудова мережі Байєса, а саме наївного та дереводоповненого класифікаторів.

Наївний (Naive) та дереводоповнений (TAN) байєсівські класифікатори – це ймовірнісні графічні моделі, що використовуються для моделювання великих масивів даних, які мають багато невизначеностей серед своїх взаємозалежних наборів характеристик. Дані моделі широко використовуються для сегментації зображень, медичної діагностики і інших задач кластеризації та класифікації даних [1].

Практичне застосування наївного та дереводоповненого класифікаторів для прогнозування кредитоспроможності фізичних осіб. Мережа Байєса являє собою зручний інструмент для опису досить складних процесів і подій з невизначеностями. Одним із лідерів прогнозної аналітики в світі є компанія SAS, що розробила спеціалізований інструмент SAS Enterprise Miner для вирішення задач інтелектуального аналізу даних [2]. У 2015 році компанія SAS в рамках продукту SAS Enterprise Miner запропонувала новий компонент HP Bayesian Network Classifier. За допомогою даного компонента можна будувати мережі Байєса різних типів, у тому числі й наївну та дереводоповнену мережі [3].

Дереводоповнений класифікатор – це покращений наївний класифікатор Байєса, що враховує ще один рівень взаємодії між параметрами системи, тобто кожна змінна може залежати від однієї іншої змінної, при цьому залежність між характеристиками даної моделі є більш реалістичною, аніж у наївному класифікаторі [1]. TAN має ненабагато більшу обчислювальну складність, ніж наївний байєсівський класифікатор, при цьому показує кращу точність.

Для порівняння класифікаторів використовується стандартний загальновідомий серед фахівців набір даних (під назвою German Bank Accepts), що описує процес кредитування фізичних осіб на суму до 120 тис. доларів, що містить 20 змінні (19 вхідну змінну і одну цільову змінну) та 5837 спостережень.

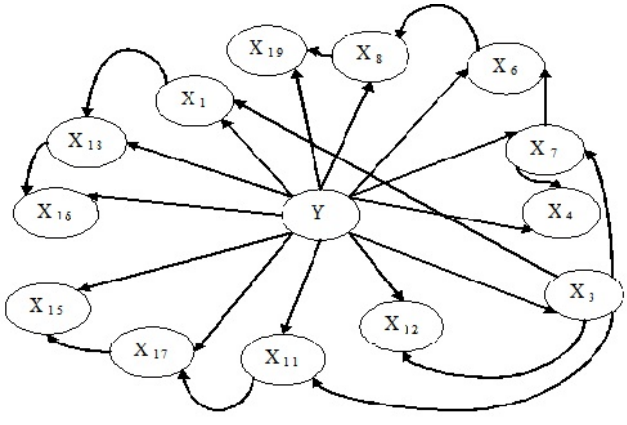


Рис. 1. Дереводоповнена байєсівська мережа для набору даних German Bank Accepts

Найбільшу кількість інформації для прийняття рішення для видачі кредиту містять характеристики: скоринговий бал бюро кредитних історій, загальна сума погашених кредитів по закритим лініям, термін з моменту відкриття першої кредитної лінії.

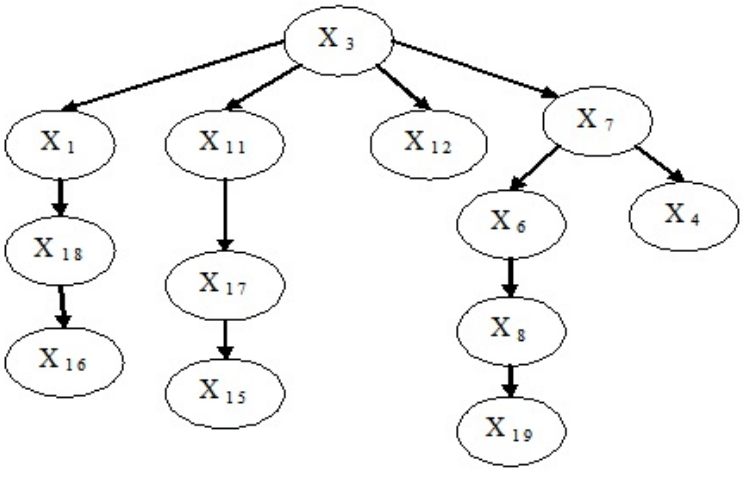


Рис. 2. Дерево вхідних змінних для набору даних German Bank Accepts

Табл. 1. Значення статистичних характеристик тестової вибірки наївної та дереводоповненої моделі, застосованих до даних German Bank Accepts

Назва моделі	Відсоток неправильно класифікованих спостережень	ROC індекс	Коефіцієнт GINI
Наївна	34,93	0,702	0,405
TAN	32,53	0,726	0,452

Висновок. За усіма представленими характеристиками дереводоповнена мережа має кращі показники, ніж наївна. ROC індекс і коефіцієнт GINI для обох класифікаторів знаходяться у межах допустимих значень, що свідчить про середню якість моделей. Результати моделювання показали, що TAN класифікатор має більшу складність і потребує більших витрат часу. Вона дає кращу точність результатів у випадку якщо існують достатньо високі кореляції між характеристиками позичальників кредитів.

Література. **1.** Padmanaban Harini. Comparative Analysis of Naive Bayes and Tree Augmented Naive Bayes Models // Master's Theses and Graduate Research, San José State University. – 2014. – 65 p. **2.** Терентьев А.Н. SAS BASE: Основы программирования (научное издание) / Терентьев А.Н., Домрачев В.Н., Костецкий Р.И. – К: Эдельвейс, 2014. – 304 с. – ISBN 978-966-2748-49-9 **3.** SAS Enterprise Miner 13.2: Reference Help. SAS Documentation. – SAS Institute Inc., Cary, 2015. – 320 p. **4.** German Credit Accept https://sites.google.com/site/data4mining/datasets/german_credit_accept

Кислий Р.В., Письменний І.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Мікросервісна архітектура в побудові подійно-орієнтованої системи моніторингу та аналізу стану здоров'я пацієнта

Своєчасне звернення пацієнта та повнота даних, якими володіє медична установа, є ключовими факторами якості наданих медичних послуг. Для того, щоб досягти більшої ефективності лікування хворого, необхідно брати до уваги не тільки традиційні фактори, що використовуються в медицині, а й враховувати фізичні навантаження, дієту, чинники зовнішнього середовища, тощо. [1, 2]

Дані питання вирішуються побудовою системи моніторингу фізичного стану пацієнта та подальшим аналізом отриманих даних, як в реальному часі, так і передбачаючи можливі загострення хвороби.

Дана система повинна відповідати наступним вимогам:

1. Бути стійкою до падінь окремих її вузлів: система повинна працювати при відключенні окремих її частин.
2. Легко розширюватися, додавання нових модулів та рівнів обробки не повинно впливати на вже існуючі.
3. Не залежати від джерел даних:
   a) Дані мають надаватися пацієнтом (ведення щоденника, натільні датчики).
   b) Дані мають збиратися з сторонніх джерел (метеодані).
4. Бути масштабованою.

Виходячи з наведених вище вимог для проектування системи було обрано мікросервісну архітектуру. [3]

Означення (Мікросервісна архітектура)**.** Це підхід до проектування, за якого єдина система будується з набору невеликих сервісів, кожен з яких працює в окремому процесі та комунікує з іншими через стандартні протоколи, наприклад HTTP. Кожен сервіс відповідає за окремий етап бізнес процесу та автоматично розгортається. Таким чином, мікросервіси можуть функціонувати в гетерогенному серидовищі. [4]

Однією з найпоширеніших платформ для побудови мікросервісної архітектури є Spring Cloud Netflix (Рис. 1).

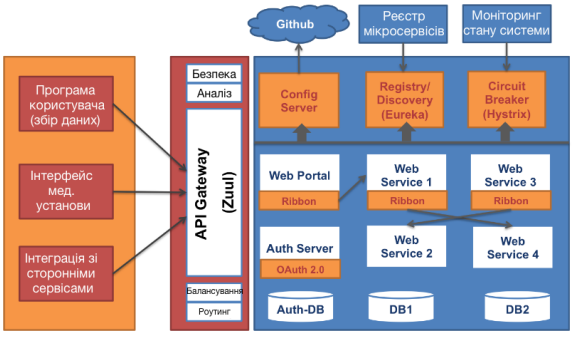


Рис. 1. Архітектура моніторингової системи на базі Spring Cloud Netflix

Література. **1.** The Diabetes Risk Score a practical tool to predict type 2 diabetes risk [Електронний ресурс] / Jaana Lindstrom, Jaako Tuomilehto - Режим доступу:http://nute.ufsc.br/bibliotecas/upload/artig **2.** Anatoly I. Petrenko Mobile health applications to support diabetic patient and doctor // Design Test Symposium (EWDTS). -2014. -pp. 1-6 **3.** Anatoly I. Petrenko Service-Oriented Computing in a Cloud Computing Environment // Comput. Sci. Appl. Volume 1, Number 6. -2014. -pp. 349-358. **4.** Microservices [Електронний ресурс] / James Lewis, Martin Fowler - Режим доступу: http://martinfowler.com/articles/microservices.html

Козленко А.С., Подладчиков В.Н.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Минимизация смещения последовательных изображений на основе метода наименьших квадратов

С развитием информационных технологий изображения, полученные с помощью технических средств, занимают все большую долю информационных носителей. Если же одно изображение несет в себе визуальную информацию, то последовательность изображений одного и того же объекта может дать более полные данные о свойствах фиксируемого объекта или объектов. Этот принцип в современном мире широко используется в робототехнике в качестве оптического зрения, в повседневной жизни последовательность изображений с малым временным интервалом записи называют видеосъёмкой и широко используется в быту, в медицине, например, при использовании компьютерной томографии, когда с помощью сканирующей аппаратуры делают снимки срезов внутренних органом для дальнейшей диагностики, в геологии последовательность изображений создается при использовании электронных микроскопов для последующего анализа породы и поиска полезных ископаемых, в биологии при наблюдении за животными и т. д. При дальнейшем развитии информационных технологий область применения последовательных изображений существенно расширится.

Как и при любых других измерениях, полученная с помощью сканирующей аппаратуры, последовательность изображений $image1, image2, image3, image4, \ldots$ имеет ошибку в виде смещения, когда в процессе создания снимков или кадров объект перемещался или же сама сканирующая аппаратура двигалась относительно фиксированного объекта. Решение этой проблемы чрезвычайно актуально т. к. из-за смещения повышается погрешность при анализе полученной последовательности изображений.

Для решения этой проблемы предложен метод поиска смещения объекта на изображении на основе минимизации суммы квадратов отклонений интенсивностей последовательных изображений $image1, image2$. Визуально алгоритм совмещения показан на Рис. 1.

Процесс минимизации смещения происходит итерационно по всей последовательности. На начальной итерации первое изображение $image1$ считается основным на основе которого калибруется следующее $image2$. Из $image1$ считываем центральный прямоугольник меньше размера $image1$ и циклично сравниваем его с частью $image2$ такого же размера. Разность интенсивностей в каждом пикселе по всем трем или одному цветовому каналу за записываем в матрицу смещения I_{ij}. Возводим элементы матрицы смещения I_{ij} в квадрат, а затем поэлементную суму полученной матрицы записываем в общую матрицу смещения I. После окончания цикла сравнивания выполняется поиск наименьшего элемента общей матрицы смещения. Индексы искомого элемента укажут на сколько $image2$ смещен относительно $image1$ по осям декартовой системы координат. После этого происходит перезапись $image2$ с найденными параметрами. На следующей итерации процесс минимизации смещения происходит аналогично с $image2$ в качестве основного изображения и $image3$ в качестве калибруемого.

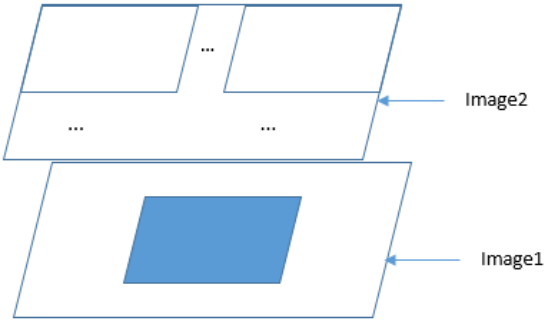


Рис. 1. Визуализация алгоритма совмещения

Литература. **1.** У. Прэтт. Цифровая обработка изображений, volume 2. М. Мир, 1982. **2.** А. Ушольд. Системы стабилизации изображения. Foto and Video, 7:76–78, 2007. **3.** А.А. Бузников and А.В. Купянский. Динамическое совмещение полутоновых аэрокосмических и графических изображений. Изв. Вузов, сер. "Геодезия и аэрофотосъемка"., 3:102–107, 1993.

Копичко С.М., Бобрик І.М.
Факультет прикладної математики НТУУ "КПІ", Київ, Україна

Аналіз моделей управління запасами для СППР

Вступ. Оптимізація управління товарними запасами - це одне за найбільш потенціальних напрямків підвищення результативності компаній, що працюють у роздрібній торгівлі, дистрибуції, виробництві. Система підримки прийняття рішень (СППР) в управлінні запасами покликана допомогти менеджменту компанії забезпечити постійну наявність потрібного товару при мінімальних запасах у системі розподілу. Результатом роботи такої системи повинно бути зниження надлишки, скорочення втрачених продажів, прискорення оборотності.

Огляд основних моделей управління запасами. Управління запасами полягає у вирішенні двох основних завдань [1]:

1. Визначення розміру необхідного запасу, тобто норми запасу, і частоти його поповнення;
2. Створення системи контролю за фактичним розміром запасу і своєчасним його поповненням у відповідності зі встановленою нормою.

Другу задачу можна вирішити за допомогою автоматизації управління. Для розв'язання першої необхідно обрати метод, що буде відповідати таким критеріям: ефективність (забезпечення наявності потрібного товару у потрібному місці у потрібний час при одночасній мінімізаціїї запасів), простота у реалізаціїї, можливість швидкого впровадження.

Зазвичай, для визначення норм запасів використовують три групи методів:

- евристичні методи,
- методи техніко-економічних розрахунків,
- економіко-математичні методи.

Евристичні методи передбачають використання досвіду спеціалістів та є здебільшого ручним способом. Метод техніко-економічних розрахунків дозволяє досить точно визначати необхідний розмір запасів, але відрізняється великою трудомісткістю. Економіко-математичні методи ґрунтуються на прогнозах та є найбільш поширеними в автоматизованих системах. [1] Проте, звачаючи на те, що попит у роздрібній торгівлі ймовірнісний, то його прогнозування потребує поєднання декількох різних методів, що значно ускладнює як реалізацію, так і подальше впровадження. Впродовж останніх десятиліть все більшої популярності набирають новаторські методи, такі як *MRP, JIT, TOC*. Перші два краще підходять для виробництва [2], а *теорію обмежень* (англ. Theory of Constraints, TOC) можно застосовувати у різноманітних галузях, в т.ч. і у роздрібній торгівлі. Для управління запасами у ритейлі теорія обмежень пропонує інструмент *динамічного управління буфером* (ДУБ, англ. DBM), тобто нормою запасів, яку необхідно зберігати у кожній ланці системи розподілу. Спираючись на рівень залишку, буфер динамічно реагує на зміни у попиті та поповненні ланки, не використовуючи при цьому складних математичних розрахунків [3]. Це робить даний метод виграшним по всім встановленим критеріям.

Висновки. Зважаючи на широкий ассортимент мереж роздрібної торгівлі, складність управління кожною номенклатурною позицією та прагнення до скорочення товарних запасів, зростає необхідність автоматизації управління запасами, а ефективне управління залишається важливою та актуальною проблемою. У роботі були розглянуті моделі та методи управління запасами, проаналізована їх ефективність, складність реалізації та впровадження. Була виказана пропозиція внесення модифікацій у алгоритм динамічного управління буфером для низькообертових товарів, виконана програмна реалізація цього методу та проведено тестування програмного продукту.

Література. **1.** Алесинская Т.В. Основы логистики. Функциональные области логистического управления / Т.В. Алесинская – Таганрог: Изд-во ТТИ ЮФУ, 2009. – 79 с. **2.** Капінос Г.І. Операційний менеджмент / Г.І. Капінос, І.В. Бабій – К.: «Центр учбової літератури», 2013. – 352 с. **3.** Eliyahu M. Goldratt Isn't It Obvious! / E. M. Goldratt – North River Press 2009.

Корепанова Н.Л., Лебедева М.А.
Севастопольский государственный университет, Севастополь, Крым

Помехоустойчивая облачная платформа на базе ОС LINUX

Одной из задач построения современных облачных платформ является разработка программного обеспечения для безопасного хранения файлов пользователей в распределенной облачной системе [1]. Основным недостатком многих современных облачных хранилищ является сложность разрабатываемых программных модулей, вызванная необходимостью поддержки различных файловых систем и СУБД, кросс-платформенной разработкой и поддержкой хранилищ различного типа [2]. Исходя из этой задачи, была разработана облачная система, предоставляющей возможность надежного хранения данных пользователя в качестве платформы как сервиса. Целевой операционной системой для разработки и эксплуатации платформы явилась операционная система Linux. В ходе разработки были выделены следующие компоненты системы: клиент, управляющий сервер и хранилище, каждый из них содержит надежные сервисы, обеспечивающие необходимый функционал и взаимодействие компонентов системы, представленные на рисунке 1.

Рис. 1. Базовая архитектура облачной платформы

Данные сервисы обеспечивают возможность хранения данных, распределение данных по серверам, механизм резервного копирования и предоставление пользователю виртуальной среды, также предоставляется возможность сбора статистики по пользователям, хранилищам и файлам пользователя. В качестве служебной информации выступают таблицы статистики, информации о физических устройствах, копиях и типов запросов. Разработанные сущности и связи полностью соответствуют их программной реализации для возможности использования

функций облачной среды. В реализованной платформе присутствуют несколько базовых уровней предоставления ресурсов, каждый из которых представляет определенный набор возможностей в пределах используемого облака: SaaS, PaaS, IaaS. В качестве языка программирования для разработки облачной платформы был выбран язык C++, преимуществом которого явилось большое количество разработанных библиотек, позволяющих лаконично реализовать сложный функционал и взаимодействовать с сторонними системами с помощью специального интерфейса. Кроме того C++ отличается высокой производительностью разработанных приложений. В соответствии с поставленной задачей разработка надежной платформы была произведена с использованием операционной системы Linux. Использование данной системы позволило использовать системные вызовы для работы с файловой системой Linux для возможности гибкого доступа к файлам и каталогам. Служебными сервисами данной облачной системы являются сервис балансировки нагрузки, сервис обхода хранилищ и сервис статистики. Их особенностью является отсутствие вызовов пользователя, которые обязательно приведут к выполнению функций этих сервисов. Для обеспечения сетевых коммуникаций разрабатываемого приложения используются TCP сокеты. Задачей TCP сокетов является организация передача данных на транспортном уровне модели TCP/IP. Механизм TCP обеспечил надежную передачу потока данных, предполагающую предварительное соединение. Надежность достигается путем дублирования утерянных в процессе передачи фрагментов данных и отправки подтверждения о получении переданных пакетов. Сокет является конечной точкой сетевой коммуникации. При отправке данных приложение просто записывает данные в сокет. Дальнейшие операции по буферизации и транспортировке ложатся на аппаратуру и протоколы более низкого уровня. Аналогичным образом происходит чтение из сокета. Для определения адреса конкретного сокета используется комбинация IP-адреса и 16-битного порта. IP-адрес определяет хост в сети, а порт – конкретный сокет на этом хосте. Для хранения данных облачной платформы на управляющем сервере использовалась СУБД MySQL. Преимуществом использования данной системы является высокая скорость обработки данных на малом объеме, бесплатная лицензия а также возможность быстрого развертывания СУБД на различных платформах. MySQL обладает возможностями для написания запросов различной сложности благодаря языку запросов SQL. Для удобства использования системы разработчиками предлагается набор API для различных языков, предлагающий удобный интерфейс для использования функционала СУБД. Основной библиотекой, которая использовалась для написания облачного приложения, явилась библиотека libsocket. Данная библиотека разработана для упрощения использования интернет и UNIX сокетов. Libsocket состоит из двух частей, написанных на языке C и C++. Объектно-ориентированная часть библиотеки представляет собой набор классов, представляющих различные типы сокетов, таких как TCP сокеты, UNIX дейтаграммы и другие. Реализация облачной платформы была проверена на примере перегруженных операторов потокового ввода и вывода для работы с сокетами. Набор функций принимающих объекты типа std::string и поддержка STL для функционала библиотеки обеспечили корректную и удобную интеграцию с другими приложениями и библиотеками. Возможности библиотеки позволяют использовать технологии IPv4, IPv6, TCP, UDP, UNIX Domain Sockets и другие. Проведенные тесты данной системы показали высокую производительность при обработке данных пользователя. Разработанные механизмы распределения файлов по хранилищам позволили эффективно использовать доступное пространство для размещения данных пользователя и обеспечить эффективность их хранения.

Литература. **1.** Клементьев И.П., Устинов В.А. Введение в Облачные вычисления. Издательство: УГУ, 2009.– 223 с. **2.** Егоров Г.А., Шяудкулис В.И., Финотти М., Беляков М.И. Принципы практической реализации современных архивных хранилищ данных. // Информационные технологии и вычислительные системы. 2013. – №1. – С. 21–32. [PDF]

Кошкин Е.Г.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Гибкая методология управления проектами как инструмент для качественного контроля разработки программного обеспечения

В современном мире высоких информационных технологий невозможно обойтись без действительно оптимального контроля над созданием программного обеспечения. При разработке ПО разработчикам приходится ежедневно принимать решения различной степени значимости, причем число уровней абстракции, по которым распределяются эти решения, может доходить до девяти [3]. Но не контролируемое принятие решений и не следование определенному плану влекут огромные потери во времени и в бюджете, а также в качестве программного обеспечения. Поэтому было введено понятие модели процесса разработки: некое следование определенному плану действий или так называемых процедур. При этом часть таких процедур направлена непосредственно на конструирование проекта (разработка архитектуры, детальное проектирование, написание кода), а прочие - на управление им [1].

Существуют много классических устоявшихся моделей процесса разработки. Ярким представителем таких моделей является «Каскадная модель», в которой сам процесс разработки напоминает поток: последовательное прохождение таких этапов, как анализ требований, проектирование, реализация, тестирование, интеграция и поддержка. Но в последнее время, когда приходится вносить быстрые критические изменения на этапах разработки, «Каскадная модель» начинает уступать гибким методологиям. Это объясняется тем, что она не имеет достаточной гибкости.Гибкая методология разработки - серия подходов к разработке программного обеспечения, ориентированных на использование интерактивной разработки, динамическое формирование требований и обеспечение их реализации в результате постоянного взаимодействия внутри самоорганизующихся рабочих групп, состоящих из специалистов различного профиля [2]. Наиболее популярные представители: экстремальное программирование, DSDM, Scrum, FDD. Практически все гибкие методологии созданы для минимизации рисков путём сведения разработки к серии коротких итераций, которые обычно длятся не длительное время (1-2 недели). Результатом каждой итерации является дополнение функционала главного проекта в общем. После окончания итерации команда расставляет новые приоритеты в дальнейшем процессе разработки. Рассмотрим подробнее поведения итераций и их контроль на популярной технологии Scrum. Scrum представляет набор принципов, на которых строится процесс разработки, позволяющий в жёстко фиксированные и небольшие по времени итерации, называемые внутри этой технологии спринтами (sprints), предоставлять конечному пользователю работающее ПО с новыми возможностями, для которых определён наибольший приоритет. Реализуемые возможности ПО в очередном спринте определяются в его начале, и они не могут быть изменены на протяжении текущей итерации. Небольшая длительность спринта позволяет добиться гибкости, а также приносит предсказуемость в процесс разработки. Обычно длительность одного спринта от 2 до 4 недель. Чем короче спринт, тем чаще выходят версии ПО. Это позволяет собирать отзывы от потребителя в конце каждой итерации, что в свою очередь уменьшает вероятность работы в неправильном направлении.

В итоге, гибкая методология позволяет своевременно отслеживать часто меняющиеся требования к ПО, при этом сохраняя бюджет проекта. Внесение изменений требований к ПО от заказчика происходит на максимально эффективном уровне. Процесс внедрения работающего функционала и оптимизации сокращается до реально коротких сроков. А минимизация рисков обеспечивается путем контроля над разработкой и внедрением со стороны заказчика.

Литература. **1.** Ramsin, R. Process-Centered Review of Object-Oriented Software Development Metodologies / R. Ramsin, R. Paige // ACM Computer Surveys. - February 2008. - V.40, №1. **2.** What is Agile Software Development?. Agile Alliance, www.agilealliance.org/agile101/what-is-agile/ **3.** Макконнелл, С. Совершенный код. Мастер-класс / С. Макконнелл. - М.: Рус. Редакция; СПб.: Питер, 2007. - 896 стр.: ил.

Кравчук Є.С.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

TensorFlow, як нове слово в області систем машинного навчання

Вступ. Нещодавно, 9 листопада 2015 року, компанія Google пренесла свою бібліотеку для машинного навчання у вільний доступ. Компанія пояснює цей вчинок тим, що бібліотека недостатньо "розумна" і, на думку розробників TensorFlow, завдяки широкому розповсюдженню штучного інтелекту навчання буде відбуватися швидше.

TensorFlow - це біліотека програмного забезпечення з відкритим вихідним кодом для числових розрахунків з використанням графів потоку данних. Гнучка архітектура дозволяє розгортати обчислення на більше ніж одному центральному чи графічному процесорі з єдиним API. Спочатку TensorFlow розроблявся дослідникамі та інженерами, що працюють в Google Brain Team з метою машинного навчання і дослідження нейронних мереж, але система стала використовуватися і для значно ширшого кола завдань.

Граф потоку даних. Графи потоків даних описують математичні обчислення за допомогою направленого графа вузлів та ребер. Вершини типово реалізують математичні операції, але можуть також представляти точки для введення даних в систему, виводу результатів, чи зчитування/запису персистентних змінних. Ребра описують відношення вводу/виводу між вузлами. Ребра даних працюють з багатовимірними масивами даних динамічного розміру, або тензорами. Потік тензорів через граф це саме те, через що TensorFlow отримав своє ім'я. Вузли, що прив'язані до обчислювальних пристроїв, асинхронно і паралельно виконують обчислення як тільки всі тензори на їх вхідних ребрах стають доступними.

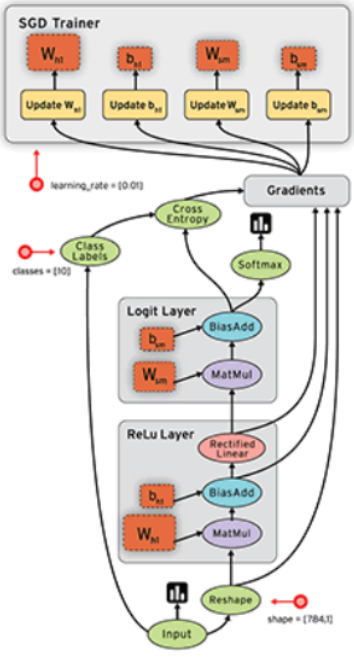


Рис. 1. Граф потоку даних

Ключові особливості.

1. Гнучкість: розробник конструює граф і програмує внутрішній цикл, який керує обчисленнями.
2. Портативність: бібліотека може працювати як на GPU так і на CPU. Більше того, підтримуються серверні, хмарні, десктопні та мобільні платформи. Можна переносити

натреновані моделі з однієї платформи на інші без внесення змін до коду.

3. Автоматичне диференціювання:користувач лише будує архітектуру, об'єднує її зі своєю цільовою функцією, вводить дані, а TensorFlow порахує похідні.
4. Підтримка різних мов: є підтримка Python та C++, ведеться розробка інтерфейсів для Go, Java, JavaScript, Lua та R.
5. Максимальна продуктивність: низькорівнева підтримка потоків, черг та асинхронного програмування дає можливість призначити виконання різних частин графу окремим пристроям, доступних користувачу.

Приклад. Розглянемо як виглядає код TensorFlow з використанням Python API. Програма має деякі дані в двох вимірах і знаходить лінію (y = x*W + b), що розділяє ці дані на два класи. Перша частина цього коду будуєте граф потоку даних. TensorFlow не починає обчислення, доки не створена сесія і не викликано метод run().

```
import tensorflow as tf
import numpy as np

# Створюємо 100 (x, y) точок даних у NumPy, y = x * 0.1 + 0.3
x_data = np.random.rand(100).astype(np.float32)
y_data = x_data * 0.1 + 0.3

# Намагалося знайти значення для W та b такі,  що y_data = W * x_data + b
# (Користувачу відомо,  що  W має бути  0.1 і b 0.3, але Tensorflow визначить це сам)
W = tf.Variable(tf.random_uniform([1], -1.0, 1.0))
b = tf.Variable(tf.zeros([1]))
y = W * x_data + b

# Мінімізуємо середньоквадратичні  помилки.
loss = tf.reduce_mean(tf.square(y - y_data))
optimizer = tf.train.GradientDescentOptimizer(0.5)
train = optimizer.minimize(loss)

#Спочатку проводиться ініціалізація змінних
init = tf.initialize_all_variables()

# Запуск сесії графа
sess = tf.Session()
sess.run(init)

# Пошук лінії
for step in xrange(201):
    sess.run(train)
    if step % 20 == 0:
        print(step, sess.run(W), sess.run(b))

# Після навчання, TensorFlow навчився,що найкраща лінія
# має параметри W: [0.1], b: [0.3]
```

Рис. 2. Код програми

Висновок. Завдяки створенню біліотеки TensorFlow було зроблено величезний крок у розвитку систем машинного навчання. А той факт, що бібліотека є у вільному доступі та завдяки мультиплатформенності та гнучкості може працювати як на смартфоні, так і на тисячі вузлів у центрах обробки даних, дозволяє всім бажаючим застосовувати її для широкого спектру завдань..

Література. **1.** Інтернет-ресурс. - Режим доступу:https://www.tensorflow.org/. **2.** Інтернет-ресурс. - Режим доступу:https://habrahabr.ru/post/270543/. **3.** Інтернет-ресурс. - Режим доступу:https://databricks.com/blog/2016/01/25/deep-learning-with-spark-and-tensorflow.html. **4.** Інтернет-ресурс. - Режим доступу:http://www.kdnuggets.com/2015/11/google-tensorflow-deep-learning-disappoints.html.

Крижанівський В.В.[1,2]
[1]*Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна;* [2]*Høgskolen i Sørøst-Norge, Porsgrunn, Norway*

Система оптичного 3D сканування

Робота присвячена проблемі реконструкції 3Д моделі об'єкту в режимі реального часу за допомогою відеокамери. Система дає можливість отримати макети невеликих об'єктів у звичайних умовах без використання спеціального обладнання.

Широке застосування об'ємні вимірювальні системи знайшли насамперед в інженерії. Основні напрямки - проектування, контроль та інспектування об'єктів. 3D Технології сканування також застосовуються в будівництві, архітектурі, медицині, кіноіндустрії, музейну справу, промисловому дизайні і в індустрії розваг. За допомогою тривимірного сканування можна оцифровувати культурну спадщину, археологічні об'єкти, предмети мистецтва. [1]

Існуючі підходи. Можна виділити два основні методи, що використовуються в системах об'ємного сканування - контактний і безконтактний. На даний момент найбільш широко застосовуються вимірювальні системи з безконтактними лазерними датчиками, що замінили контактні сенсори. Це дозволяє більш точно сканувати об'єкти і отримувати текстурну модель. Лазерні тривимірні сканери працюють як з дрібними, так і з дуже великими об'єктами, що також розширює сферу їх застосування в самих різних галузях.

Та коли нема можливості використовувати дороге обладнання, можна скористатись оптичним методом реконструкції 3D моделі. В даній роботі описане сканування об'єкта з відео в режимі реального часу, що дозволяє коригувати процес сканування для отримання бажаного результату.

Модель оптичного 3D сканування.

1. *Калібрування камери:* Реальні камери мають недоліки, такі як спотворення в об'єктиві, неквадратні пікселями і змінна фокусна відстань. Камери мають бути відкалібровані, що зазвичай робиться перед процесом реконструкції.
2. *Визначення точок інтересу:* Виводом програми є «хмара» точок в 3D. Ці точки знаходяться на 2D-зображеннях до визначення їх положення в просторі. Цей крок полягає у визначенні точок, що легко знайти на зображенні. Ці точки, як правило, на кутах на границях об'єкту. [2]
3. *Встановлення відповідності точок інтересу:* Після того, як точки інтересу були знайдені на 2D зображеннях, вони мають бути співставлені. Це означає, що потрібно визначити відповідність однієї і тієї ж точки на кожному зображенні. У випадку з відео робиться припущення, що зміна положення камери на наступному кадрі є незначною, так що точка буде лежати близько до її відповідності на наступному кадрі.
4. *Обчислити координати камери:* На основі переміщення точок від одного зображення до іншого, можна розрахувати відносне положення і обертання камери маючи при цьому достатню кількість точок відповідності. Так може бути знайдена фундаментальна матриця F, а в разі повністю відкаліброваної камери і домінантна матриця E. [2]
5. *Тріангуляція точок:* Після того, як було отримане відносне положення камер на сусідніх зображеннях, стає можливою тріангуляція точки в просторі.
6. *Блочне вирівнювання:* Використання двох сусідніх зображень є одним із способів отримати інформацію про положення точки у просторі, але в цій роботі використовується відео. Використання послідовності кадрів для отримання точного положення точки називається *блочним вирівнюванням.*

Зміст доповіді. В запропонованій роботі будуть результати практичних досліджень запропонованої моделі, порівняння з існуючими підходами, а також аналіз якості роботи програми.

Література. 1. Лысыч М. Н., Шабанов М. Л., Жадобкина В. В. Современные системы 3D сканирования // Молодой ученый. — 2014. — №20. — С. 167-171 **2.** 1. Richard Szeliski «Computer Vision: Algorithms and Applications» — 2010 — Ch 4,6,7,11,12

Круковский М.Ю.
Институт Проблем Математических Машин и Систем НАНУ

Дискретное моделирование рабочих групп

1. Введение

Интенсификация вычислительных и информационных технологий идет практически во всех отраслях деятельности человеческой деятельности [1]. Рабочие группы, а особенно малые рабочие группы стали основной техникой в производстве программного обеспечения. Применение аппарата дискретной математике для моделирования рабочих групп позволяет использовать простой эффективный инструмент для построения рабочих моделей.

2. Основные понятия и определения

Применение графов и других инструментов дискретной математики при моделировании рабочих групп рассмотрено в [2]. Применение неориентированных графов позволяет устанавливать отношения связности между узлами. Применение ориентированных графов позволяет устанавливать направленность отношений. Это используется для моделирования рабочих групп. В качестве частной задачи рассматриваются прикладные системы информационных процессов в системах электронного документооборота [3].

3. Построение модели Применения ориентированных графов позволяет установить направленность информационных процессов рабочей группы. В динамической модели, направления движения информации может меняться, маршруты могут адаптироваться в зависимости от внешних и внутренних условий рабочей группы. Для создания динамических систем рабочих групп в модель направленных графов добавляется свойство, которое позволяет менять характеристики связи, включая направление. Кроме этого, используется свойство, определяющее вес связи. Это надо для расчёта маршрута движения документов по информационным процессам рабочей группы. Свойство ориентирования вводится как функция, что позволяет реализовать ситуативную природу динамичности информационных процессов в рабочих группах. Граф задается с помощью матрицы инцидентности, где каждый элемент матрицы определяет связность между соответствующими узлами, которые представляют члена рабочей группы или подгруппу. При этом, значение поля означает не только наличие связи, но и характеристики этой связи. На значение также может накладываться характеристика вероятности, которая используется при расчёте маршрутов. Маршруты определяются как общая вероятность одновременного ориентирования ребер на всем пути.

4. Выводы Применение аппарата дискретной математики при моделировании рабочих групп позволяет строить простые модели. Модели имеют управляющие связи, реализованные с использованием инструментария ориентирующихся графов.

Литература. **1.** Теслер Г.С. Новая кибернетика.- Киев: Логос, 2004. – 401с. **2.** Круковский М.Ю. Система управления процессами Workflow. – Киев: Логос, 2006. – 191 с. **3.** Круковский М.Ю. Графовая модель композитного документооборота // Математичні машини і системи. – 2005. – № 1. – С. 120–136.

Круш І.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Особливості побудови веб систем на базі мікросервісів

Мікросервісна архітектура або, мікросервіси - це один із підходів розробки програмного забезпечення(ПЗ), який набув широкого розповсюдження за останні роки. Мікросервісна архітектура розглядається як набір незалежно розміщених невеликих сервісів, кожен з яких має свій процес і які спілкуються через чітко визначений легкий протокол для обслуговування бізнес-потреб [1].

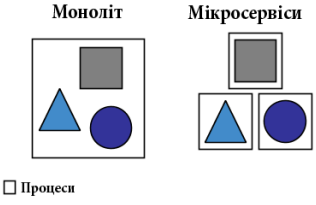


Рис. 1. Різниця між монолітною та мікросервісною архітектурою

Порівнючи цей підхід з іншими рішеннями, варто розглянути його протилежність - монолітний архітектурний стиль (Рис. 1). На відміну від мікросервісів, монолітне програмне забезпечення завжди будується як автономна одиниця. У клієнт-серверній моделі серверний додаток - моноліт, який обробляє HTTP запити, виконує бізнес-логіку та отримує/обновляє дані в БД. Основною проблемою такої архітектури є те, що всі цикли обробки інформації дуже тісно зв'язані один з одним. Зміни, внесені в одну невелику частину програми, можуть вимагати повторної збірки та встановлення на серверах нової версії ПЗ. За умов масштабування певної частини додатку, доведеться використовувати збільшення кількості процесів цілої програми замість окремих компонент. У таких випадках вирішенням усіх проблем може стати використання мікросервісної архітектури [1].

Мікросервісна архітектура допомагає у випадку еволюційних систем, коли важко передбачити, скільки пристроїв і які саме будуть намагатись одночасно отримати доступ до вашого ПЗ. Прикладом сценарію, за якого масштабування системи провели виділенням на мікросервіси, є нова версію веб системи The Guardian's. Ядро додатку спочатку було монолітним, але після того, як програма зіткнулась з непередбаченими проблемами з продуктивністю, замість переписування ПЗ з нуля була обрана стратегія використання мікросервісів. Обробка даних була організована шляхом взаємодії кількох HTTP API, що дозволило суттєво знизити навантаження на ядро і паралельно опрацьовувати додаткову інформацію [2].

Також за останні кілька років широкого поширення набула ідея використання незалежних ізольованих контейнерів. Спеціальна контейнеризація одного сервісу чи процесу дозволяє легко керувати чи оновлювати мікросервіс. Найбільш поширеним вирішенням проблеми створення кластеру є ПЗ для створення контейнерів Docker з використанням таких фреймворків як Rancher для їх менеджменту. Такий стек дозволяє організувати кластер з багатьох мікросервісів на одному або на кількох серверах, координувати взаємодію між ними, швидко оновлювати та встановлювати додатки [3].

Резюмуючи, побудова системи на базі мікросервісів дає нам переваги у легкості підтримки та пришвидшення роботи системи, а також дозволяє оптимально використовувати обмежені ресурси. Найближчим часом ця практика стане загальноприйнятою, а монолітна архітектура буде використовуватись лише для слабо навантажених систем.

Література. **1.** Chris Richardson. (2012, Feb. 2). *Pattern: Microservices Architecture* [Online]. Available: http://microservices.io/patterns/microservices.html **2.** Tom Huston. *What is Microservices Architecture?* [Online]. Available: https://smartbear.com/learn/api-design/what-are-microservices/ **3.** Nate Slater. *Using Containers to Build a Microservices Architecture* [Online]. Available: http://bit.ly/1P0v7RS

Кузнецова Ю.А.
Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", Харьков, Украина

Решение оптимизационной задачи визуализации управляющих алгоритмов методами программирования в ограничениях

Трудоёмкость разработки, тестирования и верификации управляющего алгоритма (УА) составляет, по экспертным оценкам, около 57 % от общей трудоёмкости проекта. Учитывая сложность управляющего алгоритма, а также большое количество разнообразных параметров, которыми оперирует алгоритмы данного типа, задача построения изображения графа [1] УА усложняется из-за невозможности отображения на экране монитора всех вершин и рёбер графа (режимов и переходов управляющего алгоритма).

Поэтому существует необходимость в разработке методов визуализации графов для представления УА, позволяющих отображать информацию (представлять граф управляющего алгоритма) таким образом, чтобы пользователь мог сфокусировать свое внимание на существенных (интересующих его) объектах (вершинах), игнорируя второстепенные.

Управляющий алгоритм (CA – Controlling Algorithm) может быть визуально представлен одним из трёх способов:

1. В общем случае визуализации двух типов вершин «Data» и «Task» и связей между ними:

$$CA_{DT,TD} = \langle Data_{CA}, Task_{CA}, Data_{CA} \times Task_{CA}, Task_{CA} \times Data_{CA} \rangle. \quad (1)$$

2. В частном случае визуализации только вершин типа «Data» и связей между ними:

$$CA_{DD} = \langle Data_{CA}, Data_{CA} \times Data_{CA} \rangle. \quad (2)$$

3. В частном случае визуализации только вершин типа «Task» и связей между ними:

$$CA_{TT} = \langle Task_{CA}, Task_{CA} \times Task_{CA} \rangle. \quad (3)$$

Частные случаи реализуются удалением промежуточных вершин из графа общего случая. Согласно классификации [2] визуализация графов (1) – (3) относится к классу планарных, с ортогональной сеткой, в узлах которой размещаются вершины «Data» и «Task».

В основе метода визуализации УА лежит **оптимизационная постановка задачи визуализации управляющих алгоритмов**, которая сводится к нахождению экстремального значения целевой функции на основе полученных исходных данных, удовлетворяющих заданным ограничениям, изменяя при этом значения переменных.

Исходные данные для решения оптимизационной задачи: длина $ScreenWidth$ и высота $ScreenHeigth$ рабочей области экрана в пикселях; множество объектов данных $Data$; множество элементарных задач-преобразователей информации $Task$; матрица, характеризующая правила инициирования задач при изменении (обновлении) данных DT; матрица, характеризующая получение (порождение) новых данных в результате работы задач-преобразователей информации TD.

Ограничения: только положительные координаты: $\forall\ i$, $x_i{>}0$, $y_j{>}0$; ограничения на размер вершины $Data_i$; ограничения на размер вершины $Task_j$; ограничения на толщину линий TD и DT, $Data_i$ и $Task_j$; ограничения на длину направленных рёбер; ограничения на размер текста в комментариях.

Переменные: для всех вершин $Data_i$ их длина $Size_{hor_i}$ и высота $Size_{vert_i}$; для всех вершин $Task_j$ их радиус R_j; толщина $LineThick$ и длина $LineLength$ линий; размер текста в комментариях к вершинам $Data_i$, $Task_j$.

Целевая функция показывает эффективность применения нелинейных преобразований к изображению графа. Решение задачи оптимизации сводится к определению такой величины *коэффициента искажения*, которая обеспечивает максимум целевой функции, характеризующей эффективность применения нелинейных преобразований к изображению графа.

Для решения оптимизационной задачи визуального представления управляющего алгоритма применены методы и алгоритмы недоопределенного программирования (Н-модели) – новой

теории и технологии эффективного решения широкого спектра проблем: от прикладных расчетов до обработки знаний и задач искусственного интеллекта.

Н-модели активно развивались академиком Нариньяни в рамках более общего направления "constraint programming", которое качественно расширяет возможности работы с информацией и вычислительными моделями повышенной сложности.

В терминах Н-моделей геометрические свойства визуального представления каждой вершины и каждого из направленных рёбер графа являются структурным расширением интервального н-объекта.

Способ представления недоопределенного значения влияет как на качество полученных результатов, так и на вид ограничений, связывающих это значение.

В зависимости от характера представляемой информации недоопределенные значения могут быть представлены в виде целочисленных и вещественных интервалов, множеств, перечислений и других, более специальных, конструкций [3].

Отношения непосредственного предшествования между вершинами $Data_i$ и $Task_j$, характеризующие частичную упорядоченность этих вершин, задаются в виде неравенств и входят в общую систему C – множества ограничений на н-объектах из V *(множество н-объектов v из заданной предметной области)*.

В начале вычислений границы вершин графа совпадают с границами области для изображения, а в результате вычислений происходит уточнение координат вершин графа.

Координаты двух типов вершин $Data_i$ и $Task_j$, их размеры (высота $Size_{vert_i}$ и длина $Size_{hor_i}$, толщина линий квадрата $LineThick$ вершины $Data_i$; радиус R и толщина окружности $LineThick$ вершины $Task_j$), а также длина $LineLength$ и толщина $LineThick$ направленных рёбер DT и TD определены как множество вещественных переменных. А их минимальные и максимальные значения представлены целыми константами. Совместное использование нескольких эстетических критериев определено во множестве констант логического типа.

Таким образом, исходная постановка оптимизационной задачи при использовании н-моделей заменена на поиск и определение такого множества значений параметров графа, при котором все показатели эффективности изображения графа удовлетворяют соответствующим ограничениям.

В рамках недоопределенного программирования известны десятки алгоритмов распространения ограничений, которые обеспечивают удаление избыточных ограничений, сужение области поиска, проверку совместности и поиск решения [4].

Изложенные теоретические положения позволяют строить интерактивные изображения графов, в отличие от известных методов, ориентированных на рисование статических графов.

Следовательно, ключевые свойства программного обеспечения, реализующего управляющий алгоритм, могут быть достигнуты при решении оптимизационной задачи визуализации управляющих алгоритмов методами программирования в ограничениях, что в итоге способствует повышению наглядности программного обеспечения, снижению трудоёмкости процессов тестирования и отладки управляющих алгоритмов реального времени, а также возможности их дальнейшей модификации и редактирования без привлечения высокооплачиваемых квалифицированных программистов.

Литература. **1.** Касьянов, В. Н. Графы в программировании: обработка, визуализация и применение [Текст] / В. Н. Касьянов, В. А. Евстигнеев. – СПб. : БХВ-Петербург, 2003. – 1104 с. **2.** Di Battista, G. Algorithms for Drawing Graphs: an Annotated Bibliography [Text] / G. Di Battista, P. Eades, R. Tamassia, I. G. Tollis // Computational Geometry, Theory and Applications. – 1994. – Vol. 4. – P. 235 – 282. **3.** Телерман, В. В. Удовлетворение ограничений в задачах математического программирования [Текст] / В. В. Телерман, Д. М. Ушаков // Вычислительные технологии. – 1998. – Т. 3, № 2. – С. 45 – 54. **4.** Tsang, E. Foundations of constraint satisfaction [Text] / E. Tsang // University of Essex. Colchester. Essex. – UK : Academic Press Limited, 1996. – 421 p.

***Лавренюк А.М.*[1], *Лавренюк С.І.*[2]**
[1] *Фізико-технічний інститут НТУУ "КПІ", Київ, Україна;* [2] *Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна*

Аспекти використання C++ AMP для моделювання та обробки даних

Задачі моделювання для планування сейсмічних спостережень, обробки та інтерпретації сейсмічних даних, наприклад, 3D міграція дуплексних хвиль потребують високопотужних обчислювальних ресурсів [1], [2]. Вже є розроблені та апробовані алгоритми моделювання та обробки сейсмічних даних великих об'ємів з використанням грід-мереж та обчислювальних кластерів з CPU [3]. На даному етапі для прискорення обчислень активно використовуються графічні процесори (GPU), що вже зарекомендували себе, як обчислювальні ресурси, у різних галузях науки, серед них і моделювання та обробка сейсмічних даних. Графічні процесори ефективні для прискорення обчислень, завдячуючи їх будові, що дозволяє виконувати велику кількість обчислень за одиницю часу на одному GPU [4]. При цьому обчислення на GPU більш енергоефективні, ніж на CPU. Для обчислень GPU вже доступні не тільки на потужних кластерах, але і на робочих станціях чи ноутбуках. Для ефективного і повноцінного використання обчислювальних потужностей GPU необхідні ефективні програмні засоби для розробки нових чи модернізації існуючих програм під GPU. C++ AMP дає зручний інструмент використання GPU для прискорення обчислень при моделюванні та обробці даних.

Вступ. Для використання GPU компанією Nvidea розроблена програмно-апаратна архітектура паралельних обчислень CUDA. Значними недоліками якої є закритість стандарту - розвивається тільки однією компанією і працює лише з обладнанням Nvidea. Окрім Nvidea існують і інші виробники GPU, такі як Intel, Amd, у яких є своя велика частка ринку. Так, наприклад, GPU Intel також використовуються і в інтегрованому відеоадаптері в материнську плату. Для роботи в гетерогенному середовищі, з обладнанням різних виробників, спільнотою розробників була створена платформа OpenCL (Open Computing Language) - відкритий стандарт для універсального паралельного програмування різних типів процесорів. До недоліків OpenCL можна віднести те, що OpenCL базується на мові C, що унеможливлює використання певних конструкцій мови C++ і потребує значних часових затрат на програмування і стиковку двох частин програми - тієї, що запускається на хості (CPU) та взаємодіє з GPU і програм-ядер, які виконуються уже на самому GPU. При використанні OpenCL потребується значних часових затрат на написання нових програм, або адаптацію існуючих програм, що були розроблені для роботи на CPU.

Платформу C++ AMP (Accelerated Massive Parallelism) розроблено компанією Microsoft. C++ AMP є відкритим стандартом і являє собою бібліотеку та певне функціональне розширення мови C++, яке в сукупності дозволяє проводити гетерогенні обчислення в єдиній програмі на C++. В Visual Studio версії 12 і вище включені нові інструменти та механізми для підтримки відлагодження та профілювання програм C++ AMP, зокрема відлагодження коду, що виконується на GPU, та візуалізація GPU-паралелізму. На перший погляд платформа C++ AMP включає сильні сторони CUDA та OpenCL і простіша у застосуванні.

Використання C++ AMP. Завдячуючи C++ AMP, розробники програм на C++ отримали в своє розпорядження знайомі інструменти для створення програм, що можуть досягати відчутного прискорення при виконанні алгоритмів, котрі допускають розпаралелювання за даними. C++ AMP робить ресурси GPU більш доступними розробникам програм, позаяк вже багато настільних комп'ютерів чи ноутбуків містять один або декілька обчислювачів GPU. Це дає змогу розв'язувати деякі задачі сейсмічного моделювання чи обробки сейсмічних даних вже на персональних комп'ютерах. Так як C++ AMP взаємодіє з GPU через бібліотеку DirectX версії вище 11, що встановлюється на комп'ютер разом з сучасною операційною системою, розробнику програми на C++ AMP чи користувачу такої програми не потрібно встановлювати додаткові бібліотеки чи драйвери.

Алгоритм сейсмічного моделювання чи обробки сейсмічних даних з використанням GPU

схематично можна представити так: завантаження програми-ядра до GPU, копіювання даних з пам'яті CPU в пам'ять GPU, обробка даних, зміна вхідних параметрів, обробка даних, копіювання даних з пам'яті GPU в пам'ять CPU. Процес завантаження програми-ядра, вхідних даних і вихідних відбувається один раз або декілька разів. А от програма-ядро при моделюванні чи обробці даних виконується дуже багато разів. В загальному випадку сейсмічного моделювання чи обробки сейсмічних даних час роботи програми можна представити формулою:

$$T_{all} = T_p + T_{in} + T_{out} + T_{wk_{1..n}}, \quad (1)$$

де T_{all} - загальний час обробки даних чи моделювання, T_p - час завантаження програми-ядра, T_{in} - час завантаження вхідних даних, T_{out} - час завантаження результату обчислень, час виконання програми-ядра T_{wk}, а n-кількість ітерацій або запусків програми-ядра до завершення моделювання чи обробки даних.

При аналізі швидкодії алгоритмів з обробки великих об'ємів сейсмічних даних, розроблених на C++ AMP та OpenCL, доцільно знехтувати часом одиничних процесів і аналізувати час процесів, які повторюються багато разів.

Серед особливостей обробки великих об'ємів сейсмічних даних високопродуктивними кластерами з GPU слід виділити те, що багато високопродуктивних кластерів працюють не на операційних системах Microsoft. Як CUDA так і OpenCL стабільно працюють на різних програмних платформах. C++ AMP розроблена Microsoft під операційні системи Microsoft. Поки, що відомі компілятори C++ для xUnix не підтримують платформу C++ AMP. Але завдячуючи відкритості стандарту C++ AMP, вже є спроби реалізації C++ AMP компіляторів для систем на xUnix. Однак іще немає стабільних версій таких компіляторів.

Проведені експерименти у середовищі Windows 7.0 на GPU Amd та NVidea показують, що C++ AMP дає приріст продуктивності у порівнянні з OpenCL при використанні класів array і майже не дає приросту продуктивності при використанні класів array_view C++ AMP.

Висновки. Платформа C++ AMP:

- дає можливість простіше використовувати GPU, насамперед, що містяться в ноутбуках або настільних комп'ютерах для розв'язування задач з обробки великих об'ємів даних та моделювання;
- з меншими часовими і ресурсними затратами адаптувати існуючі програми для обчислень на CPU до використання GPU у порівнянні з OpenCL або CUDA.

Література. **1.** Marmalevskyi N., Kostyukevych A., Dubrova G. Duplex Wave Migration and Corner Reflector Approximation // 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013 London, UK – 2013 – Vol 6. – P. 5253. **2.** Мармалевський Н.Я. Міграція дуплексних хвиль - новий засіб формування сейсмічних зображень субвертикальних границь / Н.Я.Мармалевський, Ю.В.Роганов, З.В.Горняк та ін. // Зб. наук. праць. УкрДГРІ.-2007.-**№2**. **3.** Lavreniuk S., Roganov Y., Tulchinsky V., Kolomiets O. Synergy of 2.5D approach and grid technology for synthesis of realistic 3D/3C seismograms in anisotropic media // 73rd EAGE Conference & Exhibition, Extended Abstracts – 2011. – P. 289. **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**– P. 69-75.

***Ланде Д.В.*[1], *Сулема О.К.*[2]**
[1] *Інститут проблем реєстрації інформації НАН України, Київ, Україна;* [2] *Факультет прикладної математики НТУУ "КПІ", Київ, Україна*

Ієрархізація моделі предметної галузі шляхом визначення центральних вузлів

Вступ. Метою даної роботи є дослідження моделі предметної галузі. Така модель представляє собою мережу, утворену з вузлів (понять) та зв'язків між ними. У загальному випадку ця мережа являє собою онтологію.

Онтологія – це формальне представлення знань. Онтології визначають концепти (тобто поняття), які належать до певної предметної галузі, та задають відношення між цими концептами [1]. Часто вони містять десятки та сотні тисяч понять, тому для практичних задач доцільним є певне спрощення онтології, виділення в ній основних зв'язків та визначення можливого підпорядкування понять головним концептам. Одним з шляхів досягнення цього спрощення є так звана «ієрархізація» графа мережі онтології, тобто побудова відповідного дерева, вершини якого – це поняття, пов'язані між собою найбільш вагомими зв'язками. Такий підхід надасть можливість автоматично формувати пошукові образи запитів та, як наслідок, забезпечить швидку навігацію за визначеною тематикою.

Формування моделі предметної галузі. Термінологічні онтології можуть бути отримані різними методами. Зокрема, у даній роботі використовується методологія автоматизованого формування моделі предметної галузі шляхом зондування великої інформаційної мережі [2]. Під зондуванням тут розуміється отримання з великих інформаційних мереж, для яких складно або неефективно проводити повне сканування, невеликої вибірки найголовніших понять. Для цього авторами був застосований алгоритм сканування реальної мережі тегів сервісу Google Scholar Citations з тим, щоб отримати набір концептів (тегів, понять) як основи майбутньої онтології.

Побудова реальної мережі. В рамках дослідження, викладеного у даній роботі, за алгоритмом зондування [2] була побудована модель предметної галузі за тематикою Information Retrieval. На графі цієї мережі (рис. 1) чітко прослідковуються три головні концепти – information_retrieval (тобто власне тематика предметної галузі), artificial_intelligence та data_mining. Треба зазначити, що на етапі визначення переліку базових концептів було сформовано словник, який включає в себе наступні поняття: _mining, _information, _search та _intelligence. Відповідно, концепти, які містять будь-які інші поняття, вважатимуться суттєвим відходженням від головної тематики.

Рис. 1. Граф онтології за тематикою Information Retrieval

Визначення центрального вузла. З метою визначення центрального вузла отримана мережа концептів досліджується за критеріями центральності. Для цього авторами був використаний запропонований раніше метод багатокритеріальної оцінки центральності вузлів [3]. В результаті його застосування авторами було визначено, що за критерієм мінімальної кількості рівнів, як і за критерієм максимального степеня вершини, найоптимальнішим вузлом є поняття data_mining, в той час як критерій мінімального середнього шляху показав, що оптимальним вузлом є концепт artificial_intelligence. Слід також зазначити, що критерій максимального степеня вершини може давати різний результат залежно від коефіцієнтів при *LevelsNumber* та *AveragePath* функції пристосованості *Fitness* [3]:

$$Fitness(i) = \alpha \cdot LevelsNumber(i) + \beta \cdot AveragePath(i) + \gamma \cdot VertexDegree(i),$$

де *LevelsNumber*, *AveragePath*, *VertexDegree* – відповідні критерії, що оцінюються. В резуль-

таті, якщо більш пріоритетним є критерій мінімальної кількості рівнів, то найоптимальнішим вузлом вважатиметься artificial_intelligence; якщо ж пріоритетом є критерій мінімального середнього шляху, то «найкращим» вузлом стає поняття data_mining.

Таким чином, для центральних вузлів онтології спостерігається деякий відхід від головної тематики information_retrieval, хоча й в межах визначеного на першому етапі словника споріднених понять.

Побудова ієрархій понять. Базуючись на отриманих результатах, тепер можна побудувати з вихідної квазіієрархічної онтологічної мережі концептів відповідні ієрархії. Для цього авторами був використаний алгоритм порівневої побудови за всіма суміжними вершинами [4]. В результаті, було отримано дві можливі ієрархії – з коренем data_mining, обраним за критерієм мінімальної кількості рівнів (рис. 2), та artificial_intelligence, обраним за критерієм мінімального середнього шляху (рис. 3).

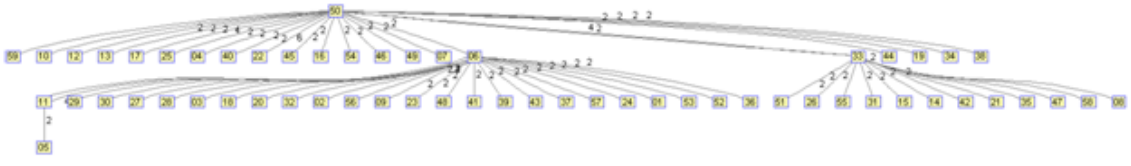


Рис. 2. Ієрархія концепту data_mining

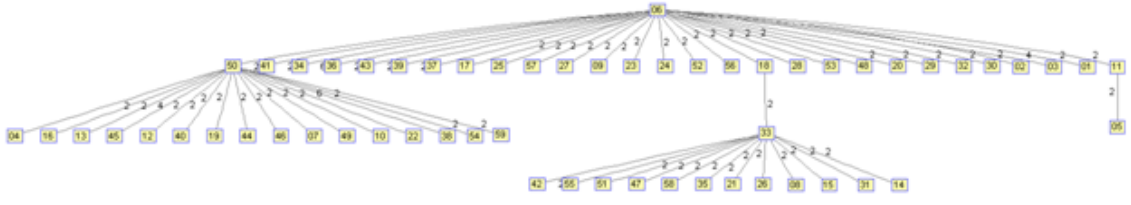


Рис. 3. Ієрархія концепту artificial_intelligence

Як бачимо, в обох ієрархіях збереглось підпорядкування понять головним концептам – data_mining (вузол 50), artificial_intelligence (вузол 6), information_retrieval (вузол 33). Таким чином, побудова ієрархії із видаленням «зайвих» зв'язків [4] не вплинула на основні зв'язки між поняттями.

Висновки. В результаті проведеного дослідження авторами був запропонований метод побудови ієрархії понять, який в подальшому може бути застосований у класифікаторах. З метою перевірки його адекватності була розглянута реальна практична задача ієрархізації онтології за тематикою Information Retrieval. Авторами було виявлено, що в результаті використання методології багатокритеріальної оцінки центральності вузлів спостерігається деяке відходження від головної тематики, і ця проблема потребує більш детального опрацювання. Зокрема, мають бути вдосконалені методи зондування сервісів Google для досягнення більшої релевантності при побудові ієрархій концептів.

Література. **1.** Лапшин В. А. Онтологии в компьютерных системах [Текст] / В. А. Лапшин // М. : Научный мир, 2010. - 224 с. **2.** Ландэ Д. В. Построение модели предметной области путем зондирования сервиса Google Scholar Citations [Текст] / Д. В. Ланде // Онтология проектирования, 2015. – N 3(17). – С. 328-335. **3.** Ланде Д. В. Порівняльна оцінка критеріїв центральності в ієрархічних мережах [Текст] / Д. В. Ланде, О. К. Сулема // Information Technology and Security. July-December 2015. Vol. 3. Iss. 2. – К., 2016. – С. 80-87. **4.** Сулема О. К. Знаходження оптимальної ієрархії у квазіієрархічному графі за критеріями центральності [Текст] / О. К. Сулема, Д. В. Ланде // Реєстрація, зберігання і обробка даних, 2015. – Т. 17. – N 4. – С. 3-10.

Литвинов В.А., Майстренко С.Я., Хурцилава К.В.
Институт проблем математических машин и систем НАНУ, Киев, Украина

Классификация основных правил и методов проверки целостности информационных ресурсов (ИР) ГИС

Под ИР ГИС понимается комплексная база данных (КБД), включающая пространственные и тематические (атрибутивные), компоненты. Целостность (consistency, integrity) КБД определяется соответствием компонентов правилам непротиворечивости (НП-правилам), устанавливающим необходимые (но, как правило, недостаточные) условия истинности данных [1]. Процесс обнаружения ошибок в КБД , нарушающих её целостность, заключается в проверке соблюдения синтаксических и семантических НП-правил для проверяемых данных. Под синтаксическими понимаются ошибки, нарушающие самые общие НП-правила, относящиеся к формам представления данных и их сочетаний, а под семантическими – ошибки, нарушающие НП-правила, связанные со смысловым содержанием данных (топологическими свойствами объектов и/или значениями атрибутов).

Существуют различные критерии, на основе которых возможно построение классификации ошибок в ГИС (например, [2]). В принятой нами классификации, ориентированной на заключительные этапы подготовки ИР ГИС, иерархическая структуризация типовых ошибок основана на критерии разделения пространственных и тематических компонентов данных по видам НП-правил - с учетом имеющегося в настоящее время готового инструментария проверки корректности топологии геоданных ГИС (в частности, ArcGis 10 [3]).

Классификация НП-правил и методов их проверки. 1. Для пространственных ошибок выделяются следующие классы и подклассы.

1.1. Синтаксические ошибки;

1.1.1. Ошибки в представлении полигональных объектов;

1.1.2. Ошибки в представлении линейных объектов;

1.1.3. Ошибки с точечными объектами.

Для проверки НП-правил класса 1.1 и исправления соответствующих ошибок современные ГИС общего назначения (в частности ArcGis) представляют готовый программный инструмент топологических проверок и исправления ошибок (11 конкретных правил для полигональных объектов, 16 правил для линейных объектов и 6 правил для точечных объектов).

Для диагностики специфических ошибок, не покрывающихся ограниченными штатными возможностями используемых ГИС, разрабатываются дополнительные предметно-ориентированные модули (например, для Mapinfo Professional [4]).

1.2 Семантические ошибки топологии

Здесь можно выделить два подкласса ошибок, наполнение которых существенно зависит от проблемной ориентации системы:

1.2.1. Ошибки позиционной точности;

1.3.1. Ошибки в представлении линейных объектов.

Кардинальным способом обнаружения и исправления и тех и других ошибок является "ручная"сводка между собой границ топологических объектов с выполнением геодезических работ. Для предварительного обнаружения сомнительных объектов представляется перспективным ретроспективный анализ пространственно-временных зависимостей [5], ([6]) между границами объектов, установленных в разные моменты времени.

2. Атрибутивные ошибки.

2.1. Синтаксические атрибутивные ошибки;

2.1.1. Нарушения алфавита представления атрибута;

2.1.2. Нарушение принятых форматов представления атрибутов и записей;

2.1.3. Прочие отклонения от установленных синтаксических правил.

Ошибки класса 2.1 практически полностью могут быть обнаружены на этапе формирования

таблиц БД, так что их появление в ИР ГИС маловероятно (хотя и возможно, особенно для подкласса 2.1.3).

2.2 Семантические атрибутивные ошибки

Систематизация семантических ошибок в табличных данных, организованных на основе ERM-моделей реляционных СУБД, основана на учете: 1) уровней организации контролируемых элементов; 2) характера ошибки.

С точки зрения первого фактора выделяются:

2.2.1. Ошибки в одном атрибуте записи таблицы БД; 2.2.2. Ошибки, охватывающие несколько атрибутов одной записи; 2.2.3. Ошибки, охватывающие атрибуты разных записей; 2.2.4. Ошибки, охватывающие атрибуты разных таблиц.

С точки зрения второго фактора конкретные НП-правила каждого уровня организации могут быть представлены следующим обобщенным правилом:

2.2.k.j Значение аргументов НП-правил должны удовлетворять заданному логико-арифметическому соотношению $ЛАС_{2.2.k.j}$ [7].

Решение задачи контроля ошибок 2.2 включает:

1) анализ смысла атрибутов и установление ЛАС, определяющих НП-правила;

2) организация базы знаний (базы НП-правил);

3) выполнение контрольных проверок.

Важное значение имеет ранжирование ЛАС, с учетом ожидаемой результативности.

3. Пространственно-атрибутивные ошибки.

Ошибки данного класса заключаются в нарушении соответствия между тематической и пространственной базами данных (компонентами базы данных).

Типовые правила, определяющие проверки соответствия, охватывают:

3.1. Отсутствие разногласий в идентификационных данных; 3.2. Соответствие площадей полигональных объектов; 3.3. Соответствие длин линейных объектов.

В принципе ошибки класса 3 можно считать разновидностью атрибутивных ошибок, НП-правила которых охватывают значения данных, находящихся в базе геоданных.

В целом, приведенная классификация носит общий, "скелетный"характер и для конкретных приложений нуждается в развитии и конкретизации. В частности, классификация положена в основу структуры модуля диагностики целостности ИР "ГИС-Леспроект" [8].

Литература. **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.** Овчинникова А. Классификация кадастровых ошибок [Электронный ресурс]. - Режим доступа: www.zemvest.ru/jurnal/archiv-jurnala/9-2013/07/. **3.** ArcGis 10. Geodatabase topology rules and topology error fixes [Электронный ресурс]. - режим доступа: help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Geodatabase_topology_rules_and_topology_error_fixes/001t000000sp000000/. **4.** Федорова Н. Модуль проверки пространственных данных на наличие топологических ошибок [Электронный ресурс]. - режим доступа: http://www.itpgrad.ru/node/1575. **5.** 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. **6.** 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. **7.** Кузьменко Г.Е., Литвинов В.А., Пилипенко Ю.Г., Ходак В.И. Задача повышения качества информационной базы агентно-ориентированных многоуровневых интеллектуализованных СППР // Математичні машини і системи.-2002.-№3.-С.53-61. **8.** Майстренко С. Я. Система «ГІС-ЛІСПРОЕКТ» як прототип геоінформаційної складової кадастрової системи // Математичні машини і системи.-2015 .-№3. С.93-99..

Лютенко И.В., Лиликович С.А.
НТУ "ХПИ", Харьков, Украина

Измерение субъектной диверсности в условиях многоверсионной разработки программного обеспечения

Избыточность и диверсность – фундаментальные стратегии повышения надёжности систем. Диверсность обеспечивается разнообразием платформ, разработчиков. Последнее принято называть субъектной диверсностью [1].

Использование метрик программного кода, таких как метрики Холстеда, позволяет численно оценить разнообразие программных реализаций [2], а в случае использования ПЛИС подобная методика может быть распространена на код, задающий принципиальную электрическую схему аппаратной платформы [3].

В то же время измерению уровня разнообразия разработчиков уделяется меньше внимания. Разработка подобного аппарата может позволить оценить уровень разнообразия разработчиков, влияющего на уровень диверсификации проекта в целом, еще до начала разработки.

В ходе работы авторами выдвинута гипотеза о том, что разнообразие разработчиков можно оценить через:

1. Формальные метрики исходного кода (S). На этапе подготовки команды кандидаты получают одинаковые тестовые задания, которые необходимо реализовать с одинаковым набором инструментов. Если задание достаточно тривиально, то можно считать что любые различия в реализации достигаются за счёт различий между исполнителями.
2. Профессиональные оценки (P). Разнообразие разработчиков можно оценить с помощью опросника, который поможет оценить подход к работе.
3. Тип личности(C). Особенности личности могут влиять на разработку неопределённым образом. Принимая во внимание ряд показателей, указывающих на тот или определённый тип личности, можно попытаться учесть это влияние.

При использовании опросников с вариантами ответа «да» и «нет» оценки P легко могут быть представлены в виде бинарного множества. Различия между двумя множествами P_A и P_B, которые образованы кандидатами A и B соответственно, могут быть оценены с помощью оператора d:

$$d_w(P_A, P_B) = |\{x | x = k_i \forall P_{Ai} \neq P_{Bi}\}|.$$

Здесь k_i – это весовой коэффициент, задающий важность i-го различия в ответах кандидатов.

Оценки S и C зависят от используемой методики, но как правило носят абсолютный характер и могут быть просуммированы с оценками P для нахождения обобщённого уровня разнообразия разработчиков C_l и C_m – значения $V(C_l, C_m)$:

$$V(C_l, C_m) = K_1 \sum_{i=0}^{j} \left[w_i \left| s_i^{C_l} - s_i^{C_m} \right| \right] + K_2 d_w(C^{C_l}, C^{C_m}) + K_3 d_w(P^{C_l}, P^{C_m}). \quad (1)$$

Здесь w_i представляет весовой коэффициент i-й формальной метрики программного кода.

Очевидно, что предложенный метод оценки нуждается в разработке или адаптации существующих опросников, таких как MBTI. Также требуется настройка коэффициентов w, k, K.

Литература. **1.** Ian Sommerville. Software Engineering. — Pearson Education, 2007. — ISBN: 9780321313799. **2.** Использование метрик Холстеда при оценке безопасности критического программного обеспечения / В.С. Харченко, В.В. Скляр, А.А. Гордеев и др. — 2003. **3.** Скляр В.В., Головир В.А. Метрики оценки сложности проектов ПЛИС, реализующих алгоритмы управления технологическим оборудованием АЭС // Радіоелектронні і комп'ютерні системи. — 2006. — No 7 (19). — С. 82–86.

Магас В.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Інтелектуальний аналіз просторових даних

Вступ. Складність обробки даних і виявлення просторових взаємозв'язків обмежує ефективність застосування традиційних методів інтелектуального аналізу даних для отримання просторових моделей. Тож, у зв'язку зі зростаючою необхідністю ефективних інструментів для добування інформації з геоданих та подальшим прогнозуванням на їх основі певних процесів виник та нині активно розвивається ітелектуальний аналіз просторових даних.

Інтелектуальний аналіз просторових даних - це процес виявлення раніше невідомих, але потенційно корисних та цікавих закономірностей з великих наборів просторових даних. На даний час надзвичайно мало готових інструментальних засобів для розробки програмних рішень саме для інтелектуального аналізу просторових даних.Як наслідок, дослідження можливості використання наявних, традиційних засобів інтелектуального аналізу даних та існуючих СУБД для розробки додатку для інтелектуального аналізу просторових даних є дуже актуальним завданням.

Методи аналізу. Ґрунтуючись на загальних принципах аналізу даних, методи інтелектуального аналізу просторових даних поділяються на чотири головні напрямки: кластеризація і метод визначення відхилення, метод асоціації та колокації, класифікація, аналіз тенденцій.

1. Кластеризація та метод визначення відхилення. Просторова кластеризація - це процес групування набору просторових об'єктів в групи, що звуться кластерами. Об'єкти в кластерах відзначаються високим ступенем схожості. Кластеризація часто використовується при роботі з великими наборами географічних даних.

2. Метод асоціації та колокації. У той час, коли кластеризація виконується на даних, які можна чітко розподілити за певною характеристикою, метод асоціації та колокації знаходить своє застосування у випадках, коли нам потрібно знайти закономірності асоціювання одного чи кількох просторових об'єктів з іншими.

3. Класифікація. Як відомо, кожен об'єкт, що знаходиться в базі даних характеризується певними ознаками. Класифікація ж, у свою чергу, є методом пошуку правил, що описують поділ бази даних на чітко заданий набір класів. Часто її вважають прогностичним методом аналізу даних.

4. Аналіз тенденцій. Просторовою тенденцією прийнято вважати зміну однієї чи більше непросторової характеристики, коли вона просторово віддаляється від початкового об'єкта. Тому, метод визначення просторової тенденції використовується для знаходження моделі зміни характеристики з огляду на сусідній вплив певного просторового об'єкта.

Структура програмного додатку. Пропонується розробити прототип веб-додатку на основі добревідомих бібліотек інтелектуального аналізу даних, таких як Weka, поєднаних з базою просторових даних на основі реляційних або графових баз даних, дослідивши переваги та недоліки такого "симбіозу". Умовно структуру такого додатку можна поділити на три рівні, як показано на Рисунку 1.

Перший рівень, тобто, інтерфейс користувача використовується для завдань вводу і виводу. Дані беруться зі сховища з використанням інтерфейсу бази даних, що дозволяє оптимізувати запити.

Другий рівень переважно використовується для управління даними, вибору алгоритмів обробки та зберігання отриманих знань. Для прикладу, на даному рівні вирішується чи певні атрибути відповідають поставленим завданням, або вибираються об'єкти, використання яких забезпечує відповідні результати.

Третій і останній рівень включає базу просторових даних і базу даних предметної області. Вхідні дані інтелектуального аналізу просторових даних значно складніші ніж класичного, адже вони включають розширені об'єкти, такі як лінії, багатокутники та інші.

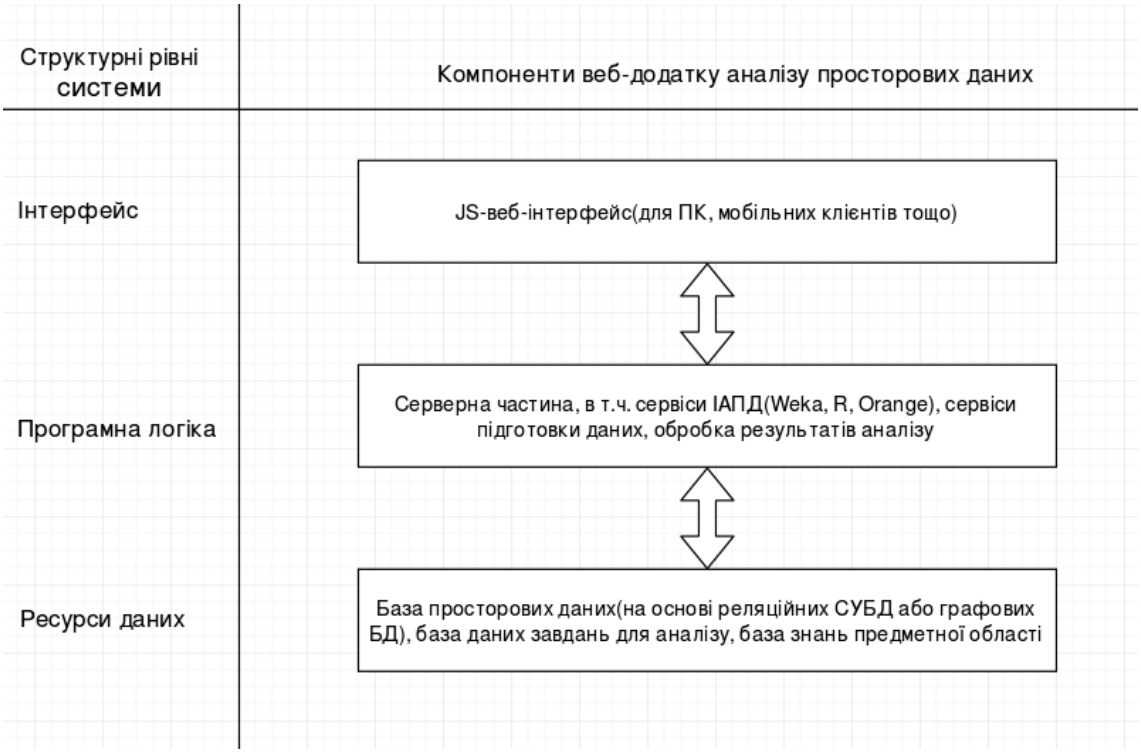


Рис. 1. Структура веб-додатку аналізу просторових даних

Приклади застосування. Інтелектуальний аналіз просторових даних пропонує значні вигоди для прикладного прийняття рішень на основі геоінформаційної системи. Останнім часом, завдання інтеграції цих двох технологій стало вирішальним фактором, особливо в різних організаціях державного і приватного секторів, що володіють величезними базами географічно пов'язаних даних, адже вони усвідомили величезний потенціал, що міститься у цій інформації. До них можна віднести:

- організації, що потребують аналізу чи поширення статистичних геоданих;
- організації, які займаються сферою охорони здоров'я і шукають закономірності поширення певної хвороби;
- природоохоронні установи, які оцінюють вплив зміни характеру землекористування в області зміни клімату;
- геомаркетингові компанії, що займаються сегментацією клієнтів, грунтуючись на їхньому просторовому розташуванні.

Висновок. Завдяки інтелектуальному аналізу просторових даних виникла можливість пошуку взаємозв'язків та особливостей у великих наборах просторових даних і моделювання поведінки відповідних об'єктів дослідження. Варто зазначити, що як алгоритми, так і програмні засоби аналізу просторових даних представляють широке поле для подальших досліджень та вдосконалень.

Література. **1.** Deren Li, Shuliang Wang, Deyi Li. Spatial Data Mining: Theory and Application, 2015 **2.** Інтернет-ресурс. - Режим доступу: http://www.spatial.cs.umn.edu/Book/sdb-chap7.pdf **3.** Інтернет-ресурс. - Режим доступу: http://www.ismll.uni-hildesheim.de/lehre/spatial-10w/script/ch7_partA.pdf **4.** Інтернет-ресурс. - Режим доступу: http://edugi.uji.es/Bacao/Geospatial%20Data%20Mining.pdf **5.** Інтернет-ресурс. - Режим доступу: http://www.spatialdatamining.org/sites/spatialdatamining.net/files/publications/intro403.pdf

Марковський О.П., Гура Е.С.
НТУУ "КПІ Київ, Україна

Метод побудови ортогональних систем нелінійних булевих SAC–функцій

Ортогональні системи булевих функцій широко використовуються в якості незворотних перетворень для широкого класу алгоритмів захисту інформації. Для того, щоб успішно протистояти диференційному та нелінійному криптоаналізу, функції, що складають систему мають відповідати критерію строго лавинного ефекту (Strict Avalanche Criterion) та мати високу не лінійність. Актуальною є проблема ефективного синтезу таких перетворень.

Більшість існуючих методів трансформують ортогональне лінійне перетворення в таке, що відповідає SAC. Проте, нелінійність отриманих перетворень не висока. Ціллю досліджень є розробка методу побудови перетворень, що мають високу нелінійність. Для досягнення поставленої цілі пропонується метод побудови системи функцій f_1,f_2,...,f_n від n булевих змінних x_1,x_2,...,x_n що зводиться до наступної послідовності дій:

1) Розділити множину ω=$\{x_1,x_2,\ldots,x_n\}$ на дві підмножини: μ, що містить парну кількість m змінних та β: ω=$\mu\cup\beta$.

2) Сформувати n ортогональну систему функцій ϕ_i, i=1,..,n, у вигляді $\phi_i = \lambda_i(\mu) \oplus \delta_i(\beta) \oplus \varphi(\beta)$, де $\lambda_i(\mu)$ - сума непарної кількості змінних множини μ, $\delta_i(\beta)$ та $\varphi(\beta)$ - відповідно лінійна та нелінійна функції змінних множини β.

3) На множині наборів з парною кількістю одиниць вибрати 2^{n-3} пар наборів таких, що змінні множини μ на них приймають протилежні значення. Значення функцій ϕ_1, $\phi_{2,..}$,ϕ_n інвертуються на цих наборах в результаті чого отримується система функцій f_1,f_2,...,f_n що відповідає SAC.

Запропонований метод ілюструється прикладом синтезу ортогональної системи нелінійних SAC-функцій від 4-х змінних x_1, x_2, x_3, x_4. Множина цих змінних розділяється на дві підмножини: μ=$\{x_1,x_2\}$ та β=$\{x_3,x_4\}$. Лінійні функції вибираються у вигляді: λ_1=x_1, λ_2=x_1, λ_3=x_2, $\lambda_4 = x_2$ та $\delta_1 = x_3$, $\delta_2 = 1$, $\delta_3 = 0$, δ_4= x_4. Нелінійна функція має вигляд: $\varphi = x_3{\cdot}x_4$. Після цього, згідно з п.2 формуються функції ϕ_1, ϕ_2, ϕ_3, ϕ_4 у вигляді: $\phi_1 = x_1 \oplus x_3 \oplus x_3{\cdot}x_4$, $\phi_2 = x_1 \oplus x_3{\cdot}x_4 \oplus 1$, $\phi_3 = x_2 \oplus x_3{\cdot}x_4$ та $\phi_4 = x_2 \oplus x_4 \oplus x_3{\cdot}x_4$. Згідно п.3 вибирається 2 пари наборів змінних на яких змінні множини μ приймають протилежні значення: <{0,0,1,1},{1,1,1,1}>, <{0,1,0,1},{1,0,0,1}>. Для одержання функцій f_1, f_2, f_3, f_4 значення ϕ_1, ϕ_2, ϕ_3, ϕ_4 інвертуються на цих 4-х наборах. Значення функцій ϕ_1, ϕ_2, ϕ_3, ϕ_4 та f_1, f_2, f_3, f_4 наведені в таблиці.

Допоміжні та отримані функції f_1,f_2, f_3,f_4 Таблиця

x_1x_2 x_3 x_4	ϕ_1 ϕ_2 $\phi_3\phi_4$	f_1 f_2 f_3 f_4	x_1x_2 x_3 x_4	ϕ_1 ϕ_2 $\phi_3\phi_4$	f_1 f_2 f_3 f_4
0 0 0 0	0 1 0 0	0 1 0 0	1 0 0 0	1 0 0 0	1 0 0 0
0 0 0 1	0 1 0 1	0 1 0 1	**1 0 0 1**	**1 0 0 1**	**0 1 1 0**
0 0 1 0	1 1 0 0	1 1 0 0	1 0 1 0	0 0 0 0	0 0 0 0
0 0 1 1	**0 0 1 0**	**1 1 0 1**	1 0 1 1	1 1 1 0	1 1 1 0
0 1 0 0	0 1 1 1	0 1 1 1	1 1 0 0	1 0 1 1	1 0 1 1
0 1 0 1	**0 1 1 0**	**1 0 0 1**	1 1 0 1	1 0 1 0	1 0 1 0
0 1 1 0	1 1 1 1	1 1 1 1	1 1 1 0	0 0 1 1	0 0 1 1
0 1 1 1	0 0 0 1	0 0 0 1	**1 1 1 1**	**1 1 0 1**	**0 0 1 0**

Отримані функції f_1, f_2, f_3, f_4 задовольняють критерію лавинного ефекту та мають максимальну нелінійність 6.

Таким чином, в результаті проведених досліджень запропоновано ефективний метод побудови булевих перетворень, що мають високі криптографічні властивості і, відповідно, забезпечують захисті від злому методами диференційного та нелінійного криптоаналізу.

Марченко О.І., Погорелов В.В.
Факультет прикладної математики, НТУУ "КПІ", Київ, Україна

Використання розширеної форми статичного однократного присвоєння для трансляції програм

У галузі проектування та розробки сучасних компіляторів можна виокремити наступні тенденції розвитку внутрішніх форм представлення програми (ВФПП):

1. Інформація про оператори присвоєння набуває другорядної ролі.
2. Потік управління відображається непрямим способом.
3. Залежності між даними стають чітко вираженими.

У той же час, однією з поширених ВФПП є граф потоку управління (ГПУ), який не відповідає тенденціям розвитку ВФПП у сучасних компіляторах із-за таких особливостей [1]:

1. Вершини, що відповідають операторам присвоєння, є єдиним елементом ГПУ, за допомогою якого можна простежити передачу даних від оператора до оператора.
2. Потік управління представлений явно за допомогою дуг, що інцидентні вершинам сусідніх операторів. Це у свою чергу призводить до надлишкового впорядкування операторів — впорядкованими є як залежні один від одного, так і незалежні оператори.
3. Залежності між місцями використання змінних та їх значеннями визначені непрямим чином. Для їх виявлення необхідно виконати аналіз досяжності визначення змінних.

Недоліки ГПУ можна частково усунути, перетворивши цей граф у форму статичного однократного присвоєння (ФСОП) [2]. Для приведення ГПУ до ФСОП потрібно виконати наступні кроки:

1. Для кожної змінної програми x, для якої зустрічається більше одного ОП (позначимо їх a_j) та x використовується у лівій частині a_j, визначаються нові змінні x_j. Ліва частина усіх a_j замінюється відповідними змінними x_j.
2. Вводяться спеціальні функції ϕ_k, що використовуються у точках конвергенції x_j. Ці функції потрібні для зв'язування різних визначень змінної x, що досяжні у точках конвергенції.

У ФСОП відсутні визначення функцій ϕ_k. Тому на практиці доцільно використовувати розширену форму статичного однократного присвоєння (РФСОП). РФСОП — це ФСОП, що доповнена визначеннями функцій ϕ_k. Повна інформація про функції ϕ_k спрощує процедуру генерацію вихідного коду. У РФСОП вводяться кілька типів ϕ-функцій:

1. γ-функція – функція, що відповідає заголовку умовного оператора. Функція має вигляд $x_3 = \gamma(p, x_1, x_2)$, де p – умова переходу конструкції $if - then - else$, x_1 – змінна, що відповідає умові p, а x_2 – змінна, що відповідає інверсній умові $\bar{p}$.
2. μ-функція – функція, що відповідає заголовку циклу. Функція має вигляд $x_2 = \mu(x_1, x_3)$, де x_1 – змінна, що є початковою установкою циклу, x_3 – змінна, що використовується в тілі циклу для обчислення результату.
3. ν-функція визначає значення змінної на виході циклу. Функція має вигляд $x_3 = \nu(p, x_{final})$, де p – умова завершення циклу, x_{final} – значення змінної, що доступне за межами циклу.

У порівнянні з ГПУ РФСОП надає наступні переваги:

1. Усувається частина побічних ефектів.
2. Залежності між місцями використання змінних та значеннями змінних в цих місцях виражені однозначно.

Література. **1.** Lawrence A. C., "Optimizing compilation with the Value State Dependence Graph", 2007, 183 p. **2.** Ballance R. A., Maccabe A. B., Ottenstein K. J., "The program dependence Web: A representation supporting control-, data-, and demand-driven interpretation of imperative languages", In Proceedings of the ACM SIGPLAN Conference on PLDI, 1990, Pp. 257–271.

Медведський А.М.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Застосування WhitestormJS для оптимізації 3Д графіки в онлайн іграх

Останні дослідження показують збільшення використання різного роду пристроїв порівняно з класичними комп'ютерами. Кожна з індустрій намагається охопити своїми продуктами ті пристрої, котрі завжди з нами. Не виключенням є і сфера ігор. Головною проблемою для них завжди була кросплатформенність, і саме онлайн-ігри постали як її рішення. Перехід у формат кліент-серверних реалізацій дав змогу використовувати браузери у ролі кліента та забезпечив можливість кроссплатформенності ігрових додатків. Доки все обмежувалось 2d графікою і залученням Flash, особливих проблем не виникало. Однак з появою HTML5 та canvas стало можливим створення 3d графіки без використання додаткових плагінів [4].

Однією з можливостей реалізувати та працювати з 3d є використання бібліотеки Three.js. Вона вирізняється своєю потужністю і дозволяє створювати прискорену 3d графіку як частину сайту без підключення додатків для браузера. Це стало можливим завдяки WebGL, як контексту елементу canvas [1].

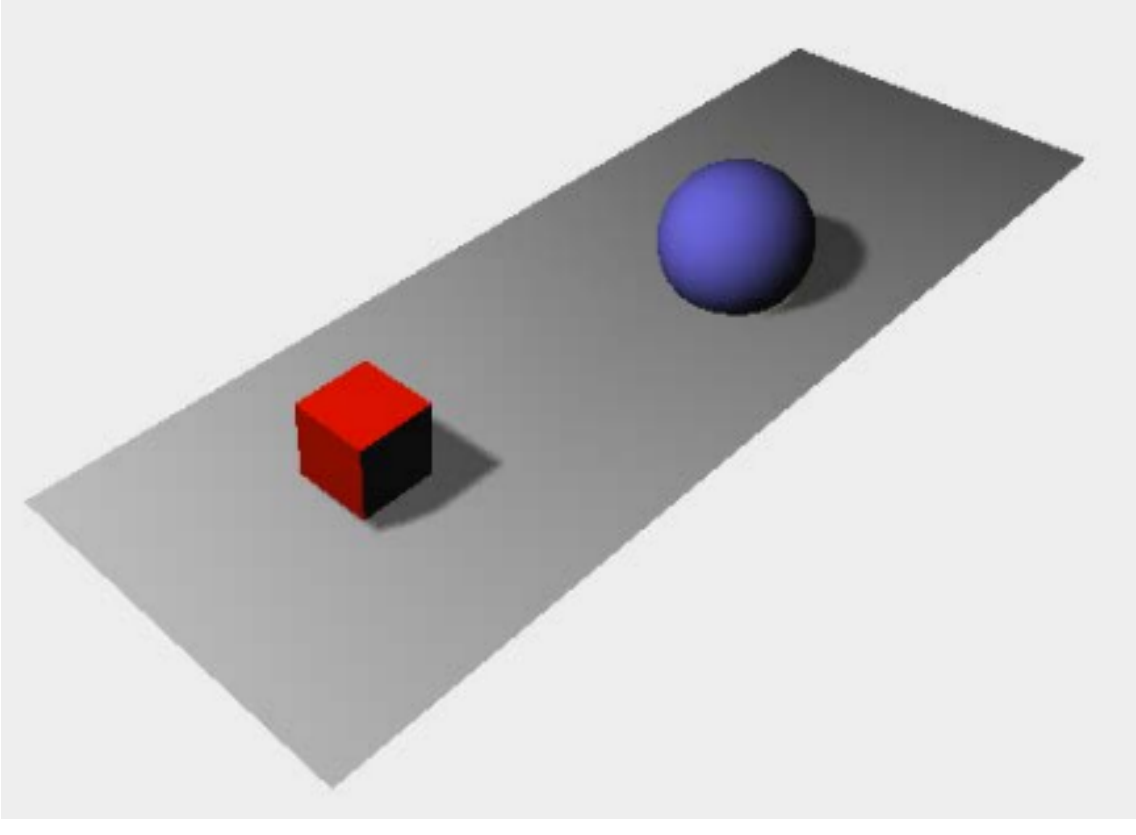


Рис. 1. Приклад використання ThreeJS

Застосування цієї бібліотеки розширило можливості веб-ігор, але не вирішило проблеми фізики руху об'ємних тіл. Як вирішення, з'явилась бібліотека WhitestormJS. Це 3d JavaScript движок, що базується саме на Three.JS. Він використовує фізику і ефекти бібліотеки для визначення WhitestormJS API, котрий містить корисні сценарії для генерації місцевості, скайбокси, анімації та симуляції пост-ефектів [5].

Whitestorm спрощує алгоритм крафтингу Three.JS до Javascript методів з параметрами. Загалом це дуже потужна і в той же час легка у використанні бібліотека, схожа на jQuery.

```
var express = require('express');
var app = express();

var UAParser = require('ua-parser');

var routes = require('./routes/index');
var staticRoute = require('./routes/test');

var UACheck = function (req, res) {
    var ua = req.headers['user-agent'];
    ua = UAParser.parseUA(ua)['family'];
    if (ua == 'FacebookBot' || ua == 'TwitterBot' || ua == 'Googlebot' || ua == 'Other'
    ) {
        staticRoute(req, res);
    } else {
        routes(req, res);
    }
}

app.use('/', UACheck);
```

Рис. 2. Приклад ініціалізації

Однією з переваг цієї бібліотеки є компактність коду без втрати функціональності. Whitestorm грунтується на тій же ідеї. Варто лише ввести параметри, які зроблять 3D об'кти відмінними від інших. Багато інших опцій генеруються за замовчуванням. Також, для прикладу, кожний проект ThreeJS має свою власну функцію перебудови в залежності від ширини. За умови існування таких 10 проектів, потрібно буде написати цю просту функцію 10 разів. Для Whitestorm достатньо лише autoresize - true параметр до функції ініціалізції [2].

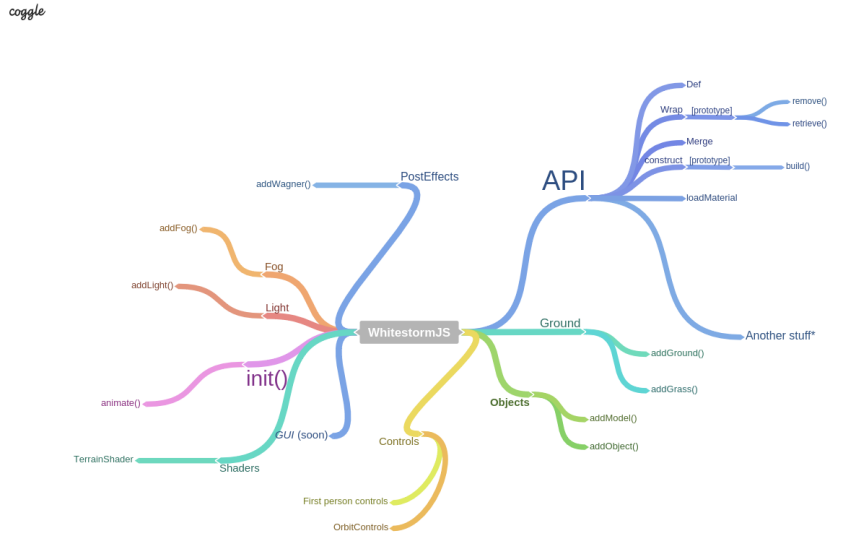


Рис. 3. Структура WhitestormJS

Загалом, використання WhitestormJS оптимізує роботу ThreeJS, пришвидшує та стандартизує розробку 3d графіки для веб-ігор [3].

Література. **1.** WhitestormJS http://whitestormjs.xyz. **2.** SurviveJS http://survivejs.com/blog/whitestormjs-interview. **3.** Github https://github.com/WhitestormJS/whitestorm.js. **4.** Wikipedia https://en.wikipedia.org/wiki/WhitestormJS. **5.** Gitter https://gitter.im/WhitestormJS/whitestorm.js.

Михалько В.Г.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Використання асинхронної моделі програмування для розробки високонавантажених веб-серверів

При розробці серверного програмного забезпечення, яке повинно паралельно обробляти велику кількість клієнтських запитів, класичним підходом є створення окремого потоку (thread) для кожного нового клієнта. Якщо при обробці запиту треба зробити звернення до файлової системи, бази даних чи якогось веб-сервісу, використовуються блокуючі операції вводу / виводу. Тобто потік робить виклик відповідної блокуючої функції, і чекає на результат. Так як такі операції займають відносно багато часу, то потік фактично знаходиться в стані простою, але при цьому несе додаткове навантаження на операційну систему.

Типовий алгоритм для створення веб-серверу, використовуючи цей підхід, виглядає таким чином:

```
ServerSocket server = new ServerSocket(PORT);
while (true) {
    Socket clientSocket = server.accept();  // очікування наступного запиту

    Input input = clientSocket.input();     // вхідний буфер
    Output output = clientSocket.output();  // вихідний буфер

    Task task = createTask(input, output) /*   завдання по обробці запиту,
                                               включаючи виконання блокуючих
                                               операцій вводу виводу
                                          */

    runTaskInNewThread(task);               // виконання завдання в новому потоці
}
```

Рис. 1. Типова реалізація блокуючого веб-серверу

З точки зору ефективності, така реалізація є прийнятною, якщо кількість потоків порівняна з кількістю фізичних ядер у процесорі. Але у реальних високонавантажених веб-сервісах кількість одночасних запитів значно перевищує кількість фізичних ядер, тому подібний підхід проявляє ряд недоліків, серед яких основними є:

- Зменшення продуктивності, внаслідок чого клієнтам потрібно довго чекати на отримання відповіді
- Складність написання коректного коду для міжпотокової взаємодії

Якщо на один фізичний процесор припадає багато потоків, то операційна система має їх періодично переключати. У кожного потоку є свій стан, який зберігається в регістрах і в кеші. Коли ОС надає потоку можливість продовжити виконання, то вона також оновлює дані у регістрах і кеші, щоб відповідний код міг коректно працювати. При невеликій кількості потоків, час, який необхідний для оновлення регістрів і кеша, є незначним у порівнянні з часом виконання основної логіки програми, але при значній їх кількості, дії операційної системи для переключення потоків можуть створити значний спад продуктивності системи. Це ілюструє так звана "проблема 10000 з'єднань" [1]. Її суть полягає в тому, що мережеві інтерфейси сучасних комп'ютерів можуть обробляти 10000 запитів і більше, але внаслідок неефективних алгоритмів, тобто використання великої кількості потоків і блокуючого вводу / виводу, система в цілому не справляється з цим навантаженням.

Одним із варіантів вирішення проблеми з продуктивністю - є використання пулів потоків (thread pool). Пул містить в собі визначену кількість потоків, і при необхідності може надавати їх для використання. Це підхід частково вирішує проблеми з ефективністю, але все одно має обмеження. Якщо кількість одночасних запитів запитів значно перевищує кількість потоків в пулі, то клієнту все одно доводиться дуже довго чекати на відповідь, що часто не є припустимо.

Окрім того, взаємодія між потоками є доволі складною, і це вимагає додаткових навичок від розробників програмного забезпечення. Дуже часто виникають помилки, які спричинені непра-

вильним кодом міжпотокової взаємодії. Внаслідок цього може бути порушена конзистетність системи, і вона працюватиме некоректно.

Цих всіх недоліків не містить асинхронна модель програмування. В її основі лежить цикл подій (eventloop). Цикл подій потребує тільки один потік для роботи. Основна його функція - приймати завдання (task), і виконувати їх у встановленому порядку. Зазвичай у нього є черга з завданнями, з якої вони і витягуються у визначеному порядку. Єдина блокуєча операція при такому підході - це отримання наступного завдання з черги, якщо вона пуста.

Розглянемо спрощену схему по обробці клієнтського запиту, використовуючи асинхронний підхід:

Рис. 2. Схема роботи циклу подій

Як видно зі схеми, при надходженні нового запиту, він спочатку реєструється в черзі. Через деякий час завдання по обробці цього запиту вигягується з черги і починає виконуватись. Одна з основних вимог до завдання, яке буде виконуватись в циклі подій - це відсутність блокуючих операцій, бо у разі їх наявності, всі клієнти, які звертаються до до цієї системи, незалежно від складності обробки їхнього запиту, будуть змушені чекати тривалий проміжок часу, і може здаватися що система “зависла”.

Класичні операції по вводу / виводу - блокуючі, через це їх не можна використовувати при асинхронному підході. Якщо завданню треба зробити довготривалий виклик (наприклад звернення до файлової системи, передача даних по мережі, і т.д.), то воно реєструє цей виклик, використовуючи спеціальні зовнішні сервіси, які надаються сучасними операційні системи [2] та бібліотеками [3], і прикріплює до нього колбек(callback). Після цього керування знову передається в цикл подій. Коли зовнішній сервіс виконав виклик, то він сповіщає про це, реєструючи в черзі завдання для виконання прикріпленого раніше колбека, надавши йому відповідні дані.

Використання циклу подій разом з неблокуючими операціями вводу / виводу, дозволяє значно покращити продуктивність, через те що немає додаткового навантаження, яке створюється потоками.

В цілому, правильне використання асинхронної моделі програмування збільшує продуктивність системи. Також, написання коду стає значно легшим, бо в такому підході відсутні багато класичних проблем, пов’язаних з міжпотоковою взаємодією і синхронізацією. Зважаючи на ці всі обставини, така модель є дуже вдалим вибором при розробці високонавантажених веб-серверів і систем.

Література. **1.** Проблема 10000 з’єднань, https://goo.gl/hLdCzk. **2.** Asynchronous IO in modern OS, https://goo.gl/ELlR87. **3.** DataKernel Framework, http://datakernel.io/.

Михняк О.О.
Національний технічний університет України "КПІ", факультет електроніки, Київ, Україна

Комп'ютерна система обробки зображень на сітківці ока

Актуальність теми. На сьогоднішній день, в умовах стрімкого науково-технічного прогресу збільшується час використання різноманітних гаджетів (смартфони, планшети, ноутбуки та інше), що негативно впливає на зір людини. Вже не говорячи, що під час бойових дій найчастіше страждає сітківка ока. Стан офтольмологія на Україні викликає серйозне занепокоєння.

Якісні офтольмологічні клініки розташовані лише в великих містах України та зазвичай належать приватним особам. Це пов'язано з великою вартістю офтольмологічного обладнання та з не зацікавленістю населення регулярного відвідування лікаря-офтольмолога. По статистиці найчастіше відбувається візит до лікаря-офтольмолога через погіршення зору. Тобто після наслідків хвороби, а не для її профілактики.

Мета роботи. Метою даної роботи є розроблення програмного забеспечення, сервісу (бази-даних) та в подальношу реалізації апаратної частини, для здешевлення діагностики очних хвороб.

Основні переваги. Даний продукт, має свою унікальність. Оскільки більшість програмного забеспечення, яке використовується в офтольмології розраховане на отримання якісного зображення та в подальшому його аналізі.

Продуктивним є метод, згідно якого зображення очного дна, яке не відповідає найвищим стандартам якості, може бути оброблене за допомогою спеціальних комп'ютерних програм з метою отримання інформації про стан відповідних структур ока. Подібні методи використовують, зокрема, у рентгендіагностиці, знижуючи таким чином дозу опромінення пацієнта.

Тобто прямими "конкурентами"є програми по обробці зображень. Але вони не пропонують сервісу (бази-даних) взаємодії лікаря та пацієнта. Також мають "зайвий а іноді й не достатній функціонал. Наприклад в офтольмології часто потрібно проводити дослідження аберацій сітківки ока за допомогою інфрачервона трансскліральной офтальмоскопії. [1].

Отже, основними перевагами є:

- Створено декілька стандартних наборів фільтрів (функцій), що обхвачує різноманітні випадки;
- Основні функції створені на базі плагінів, тобто дозволяють легко переключатись між необхідними пакетами фільтрів;
- Підтримується кросплатформеність, відсутні незручності при використанні різних операційних систем користувачем (лікарем) або пацієнтом;
- Розробляється база-даних (сервіс), для взаємодії лікаря та пацієнта. Тим самим не привязуючи пацієнта до конкретних офтольмологічних клінік;
- Не потребує спеціальних навичок, для пацієнтів щоб користуватьсь сервісом;
- Надається можливість обробляти не тільки зображення, але й відео файли. Тобто розкладати відео потік по кадрам та обробляти найбільш вдалі;
- Розробляється додаткове апаратне забеспечення, яке доповнює можливості програми. В апаратній частині використовуються спеціальні методи підсвітки очного дна, що значно збільшує якість зображень та їх інформаційну цінність. [2].

Тобто розробляється цілий сервіс, за допомогою якого можна суттєво здешевити діагностику очних хвороб залишаючись на високому рівні якості.

Практичне значення. На разі, з метою здешевлення діагностики очних хвороб розробники медичного обладнання широко використовують сучасні IT-технології. Повідомляється про

використання можливостей iPhone зі спеціальним програмним забезпеченням «Filmic Pro» для отримання якісного зображення очного дна для подальшого аналізу. [3].

Запропонований сервіс можна використовувати:

- Самостійно пацієнтом, з використанням підручних засобів (камери смартфона, фотоапарату та інше) для отриманя попереднього аналізу стану сітківки;
- "Польовим"лікарем, для зберанян попередньої іформації про стан зору пацієнту. Також передбачає використання підручних засобів, але знаючи специфіку ока людини дає змогу отримання більш якісних зображень;
- Лікарем-офтольмологом з використанням не дорогої апаратури.

Рузультат роботи. Данна система реалізована за допомогою інтегрованого середовища розробки програмного забезпечення Microsoft Visual Studio на мові програмування C#, тим самим підтримуючи кросплатформеність. Тобто відсутні незручності при використанні різноманітних операційних систем користувачем.

Програма отримала позитивні відгуки від експертів, і наразі проходить тестові випробування.

Програма має тестову назву "Ukrainian Eyes Book"або ж "U Eyes Book".

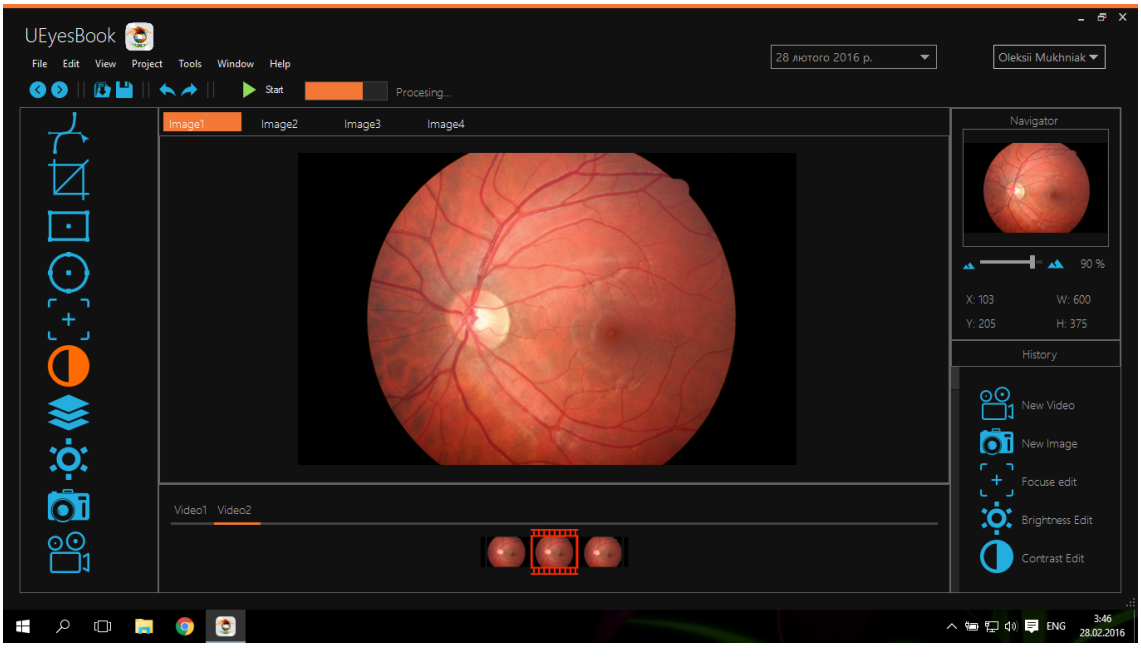


Рис. 1. Інтерфейс користувача (лікаря-офтольмолога). Зовнішній вигляд програми

Література. **1.** Плюто И.В. Инфракрасная транссклеральная офтальмоскопия: физические и технологические аспекты метода / И. В. Плюто, А. П. Шпак. — К.:ИМФ НАНУ, 2005. — 44 с.

2. Плюто І.В., Шпак А.П., Запорожець А.О., Замурняк О.Г., Тяжка Н.П., Панасенко Г.М., Тіницька Я.Ю. / Удосконалення технології транссклерального просвічування для аналізу зображення очного дна. Наукові технології, 2009, № 3 (7 – 8). С. 86 - 90.

3. Використання можливостей iPhone для отримання зображення очного дна для подальшого аналізу, http://hronika.info/tehnologii/8120-amerikanskie-uchenye-prevratili-iphone-v-oftalmoskop-foto.html.

Морозов К.В., Потапова Е.Р.
Факультет прикладной математики, НТУУ "КПИ", Киев, Украина

О самотестировании многопроцессорных систем

В современном мире все большую популярность получают отказоустойчивые многопроцессорные системы (ОМС), которые используются как для организации различного рода вычислений, так и для управления сложными объектами. Одним из факторов функционирования ОМС является ее реконфигурирование, для чего необходимо самотестирование системы. Процесс самотестирования может быть основан на выполнении серии проверок одних процессорных элементов (ПЭ) другими, выполняемых совместно с основными (рабочими) функциями системы. Результатом этого процесса является информация о том, какие ПЭ в данный момент являются исправными, а какие вышли из строя. Эта информация может быть представлена в виде так называемого вектора состояния системы, каждый элемент которого соответствует одному из процессоров системы и принимает значение 1, если процессор исправен и 0, если он неисправен.

Отметим, что если проверяющий процессор – неисправен, результат проверки может оказаться не соответствовующим действительности. Эти особенности отражены в модели Препарата-Метса-Чена (ПМЧ-модель), в соответствии с которой результатом проверки одного ПЭ другим является 1 (если обнаружена неисправность), либо 0 (если неисправность не обнаружена). Иными словами, результатам проверок исправным процессором можно доверять. Если же тестирующим является неисправный процессор, то эти результаты могут быть искаженными (грубо говоря, эти результаты не зависят от состояния тестируемого процессора).

Учитывая, что проверки выполняются совместно с основными функциями системы и требуют для себя часть процессорного времени, желательной является минимизация их количества. Отметим также, что в реальных системах ПЭ обычно имеют непосредственные связи только с некоторыми ближайшими соседями, таким образом, выбор пар "проверяющий-проверяемый"весьма ограничен.

Построим выражение $f(\mathbf{v})$ на основе результатов проведенных проверок ПЭ:

$$f(\mathbf{v}) = f_1(\mathbf{v}) \wedge f_2(\mathbf{v}) \wedge ... \wedge f_s(\mathbf{v}),$$

где $\mathbf{v}$ — вектор состояния системы, s — количество проверок, а $f_k(\mathbf{v})$ — булево выражение, соответвующее результату k-й проверки. Если эта проверка заключалась в тестировании j-го процессора j-м процессором и ее результат (в соответствии с моделью ПМЧ) имеет значение t_i, то

$$f_k(\mathbf{v}) = v_i v_j \vee \bar{v_i}, \text{ если } t_i = 0;$$

либо

$$f_k(\mathbf{v}) = v_i \bar{v_j} \vee \bar{v_i}, \text{ если } t_i = 1;$$

где v_i и v_j — компоненты вектора состояния системы, соответствующие j-му и j-му процессорам.

В докладе предложен подход для нахождения вектора состояния системы, основанный на решении булевого уравнения вида: $f(\mathbf{v}) = 1$.

Следует сказать, что такое уравнение может иметь больше одного решения, часть из которых можно отсеять за счет введения ограничения M на максимальное (допустимое) количество неисправных процессоров, т.е. на количество нулей в векторе $\mathbf{v}$. Отметим, что для модели ПМЧ в случае полного диагностического графа $M \leqslant \frac{n-1}{2}$, где n — число процессоров. Если с учетом введенного ограничения получено единственное решение, то оно соответствует истинному состоянию системы. В случае, когда решений несколько, необходимо провести дополнительные проверки (если это возможно), благодаря которым будет дополнено выражение $f(\mathbf{v})$ и могут быть отсеяны лишние варианты.

Предложенный метод может быть использован и для тех случаев, когда диагностический граф не является полным.

Осадчий Д.Ю.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Auto-Layout как способ построения графических и мультимедийных интерфейсов для приложений операционной системы iOS

Введение. В современном мире мобильных устройств мы имеем большое количество разных вариаций мобильных устройств. Каждое устройство обладает определенными характеристиками, которые определяются производителем. Одним из ключевых компонентов является экран. С появлением различных размеров экранов мы столкнулись с проблемой адаптивного дизайна для каждого устройства. Чтобы устранить данную проблему в операционной системе *iOS6*, компания *Apple* добавила *Auto-Layout*. *Auto-Layout* это инструмент для разработки графических интерфейсов, который позволяет разместить графические элементы на экране, согласно определенным правилам.

Проблема. Задача построения пользовательского интерфейса лишена сложности, если размер окна остается фиксированным, но если размеры окна меняются, то позиция элементов должна адаптироваться к новому разрешению. До операционной системы *iOS6*, дизайн нужно было описывать программно для поддержки адаптивной разметки. В случае добавления дополнительно языка в интерфейс, разработчик был обязан описывать новый интерфейс для каждого языка.

Решение. *Auto-Layout* динамически вычисляет размер и позицию всех *View* в нашей *View* иерархии, базируясь на ограничениях короче наложены на *View*. Для примера вы можете наложить ограничение на кнопку, такое, что ее расположение будет горизонтально в центре, относительно *View* картинки, а также верхний край кнопки будет на 8 точек ниже *View* картинки. В случае изменения позиции *View* картинки, позиция кнопки будет автоматически изменена, удовлетворяя условиям ограничений.

Подход построения интерфейса на основе ограничений, позволяет строить динамические пользовательские интерфейсы которые реагируют на внутренние и внешние изменения.

Внутренние изменения происходят когда размер или форма родительского View изменилась. С каждым изменением вы должны обновлять разметку вашей View иерархии для использования свободного пространства. Источники которые могут быть причиной внутренних изменений:

- Пользователь изменил размер окна.
- Происходит переворот экрана.
- Вы хотите поддерживать различные размеры View.
- Активная запись или разговорная панель инструментов на верху экрана.

Внутренние изменения происходят когда размеры View изменяются. Источники внутренних изменений:

- Содержание View поменялось.
- Пользователь изменил язык.
- Пользователь изменил размер шрифта в локальных настройках операционной системы.

Вывод. Для создания динамического и адаптивного пользовательского интерфейса приложения в операционной системе *iOS* понадобится инструмент *Auto-Layout*, который позволяет описывать поведение View, вызванных внутренними или внешними изменениями. С помощью определенных правил и ограничений, разработчики получили возможность создавать единный интерфейс приложения для различных размеров экранов.

Литература. **1.** Apple Inc, Understanding Auto Layout, https://developer.apple.com/library/ios/documentation/UserExperience/Conceptual/AutolayoutPG/. **2.** Matthijs Hollemans, Auto-Layout in iOS7 https://www.raywenderlich.com/50317/beginning-auto-layout-tutorial-in-ios-7-part-1. **3.** Erica Sadun, Mobile Programming Series: iOS Auto Layout Demystified, Second Edition, p.41-44, 2013

Пашаев Н.М., Рагимов Р.М., Пашаева М.М., Мутталибова Ш.Ф.
Научно-Исследовательский Институт Аэрокосмической Информатики НАКА, Баку, Азербайджан

Аэрокосмический мониторинг динамики развития элементов наземной инфраструктуры Апшеронского региона на основе космических изображений и технологии ГИС

В последнее время в Азербайджане было развернуто интенсивное строительство объектов промышленного и гражданского назначения. В результате этого процесса выросло число поселков и деревень расположенных в Баку и пригородах, что приведет к увеличению и масштабу влияния антропогенных процессов на окружающую среду. Вследствие роста численности населения увеличивается также масштаб строительных работ, уменьшаются зеленые насаждения, и наносится серьезный урон окружающей среде и экологии, одним словом, в ландшафте происходят серьезные изменения. В результате всего этого процесса экологические проблемы в городе и пригородах существенно обострились [1].

При оценке и исследовании геоландшафтных изменений, происходящих в Апшеронском экономическом районе, были использованы современные геоинформационные технологии и космические снимки высокого пространственного разрешения. При этом решалось задача оценки динамики изменения геопространственных элементов Апшеронского экономического района, была дана оценка влияния роста населения, на процессы, происходящие в ландшафте, а полученные результаты предоставлены в виде многослойных цифровых электронных карт.

Для определения изменений геоландшафте Апшеронского района, произошедшых в результате прироста населения, были использованы топографические карты масштаба 1:5000, а также космические снимки исследуемого региона со спутника Landsat TM. Для определения геоландшафтных изменений на исследуемой территории, в среде программы Arc ГИС были созданы электронные карты различных тематических слоев [2].

При оценке динамики изменений геопространственных элементов исследуемой части Апшеронского экономического района были использованы 17 топографических карт масштаба 1:5000 и космические снимки, сделанные в разное время. После сканирования эти карты были привязаны к координатной системе Пулкова – 42, переведены в цифровой формат и после георефернции были включены в географическую информационную систему (рис. 1).

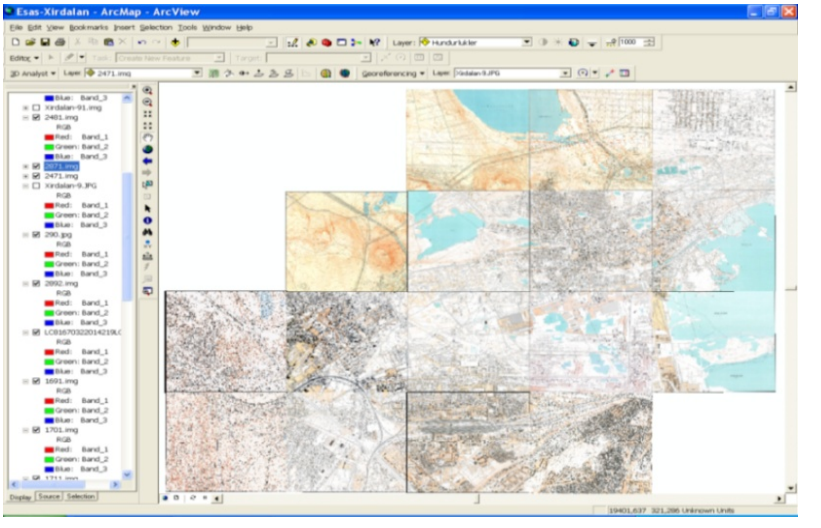


Рис. 1. Топографическая карта масштаба 1:5000 исследуемой части Апшеронского экономического района после георефернции

С целью определения связи геоландшафтных изменений, произошедших в регионе в 2000–2015 годы с увеличением численности населения, полученные сведения о темпах прироста населения на исследуемой территории, приведены в виде графиков (рис.2). Во время исследования было выявлено, что этот показатель по сравнению с общим темпом прироста поселения по стране является очень высоким. Можно прийти к выводу, что прирост населения на исследуемой территории происходит в основном за счет внутренней миграции.

Цифровые электронные карты Апшеронского экономического района были составлены с помощью модуля Arc Map программы Arc ГИС и обновлены на основе космических снимков. Данные, накопленные о территории, были включены в базу данных (БД) ГИС и некоторые тематические слои, созданные на основе топографических карт. Данные слои были раздельно сгруппированы на основе показанной в базе данных ГИС атрибутивной информации и размещены в виде векторных моделей на карте под определенными названиями [3]. Объекты с общей символизацией (населенные пункты, дороги, зеленые насаждения, рельеф, нефтегазовые месторождения) были обобщены в виде векторных слоев (рис. 3, рис. 4, рис. 5, рис. 6).

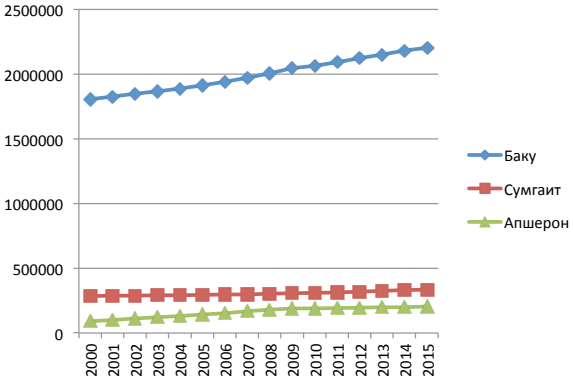
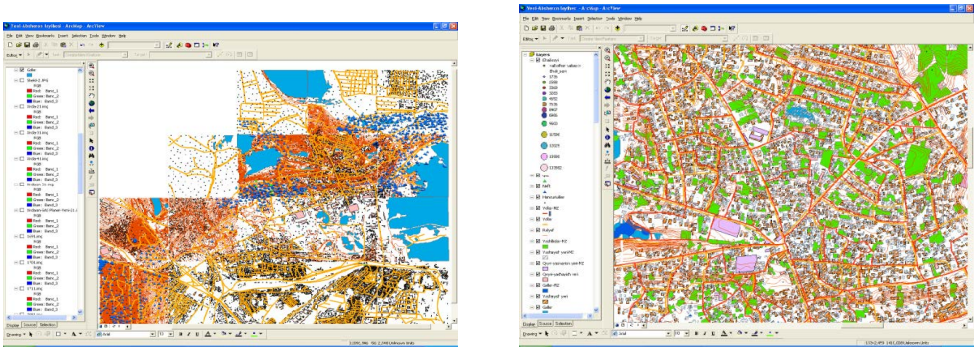
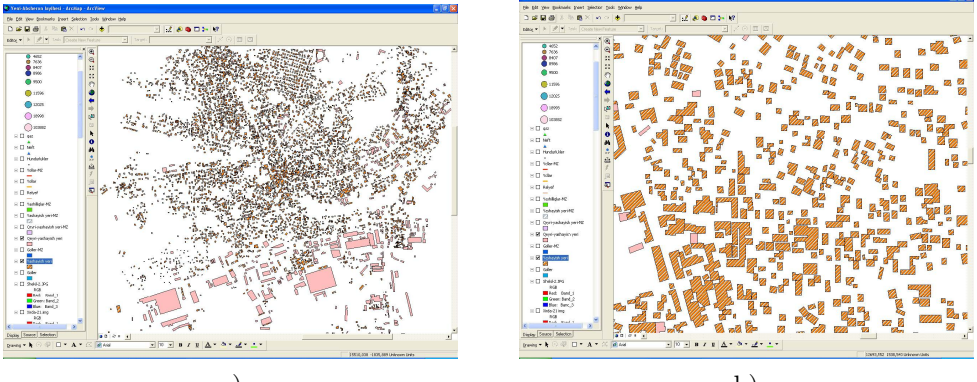


Рис. 2. Прирост поселения за 2000–2015 г.



a) b)

Рис. 3. a) Электронные карты отражающие общий вид исследуемой территории, b) электронные цифровые карты созданные в среде Arc Map и включающие все тематические слои



a) b)

Рис. 4. a) Электронные карты включающие только населенные пункты исследуемой территории, b) Цифровые электронные карты населенных пунктов созданные при помощи Arc Map

Для определения геоландшафтных изменений произошедших в Апшеронском экономиче-

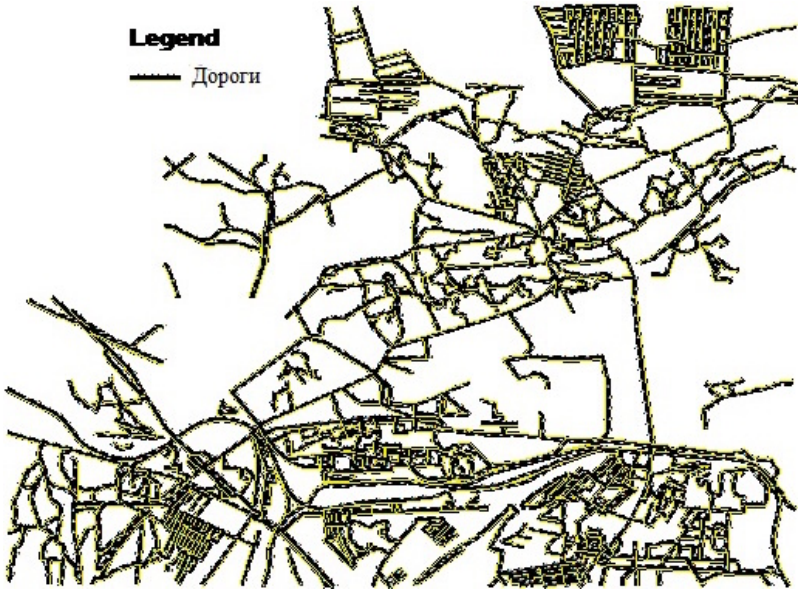
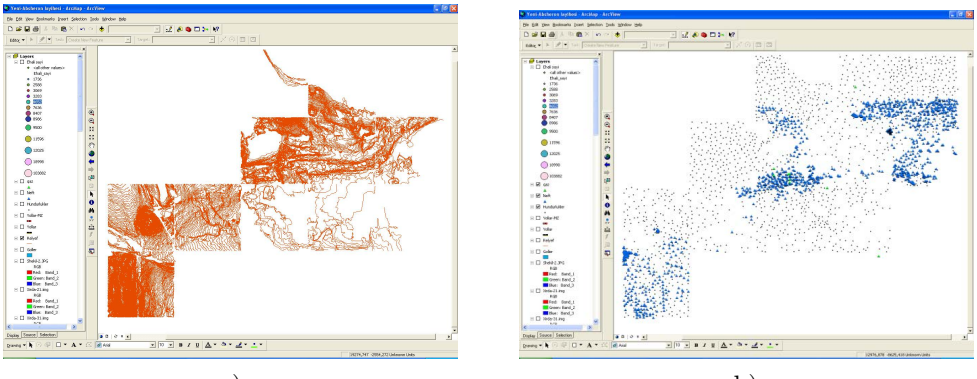


Рис. 5. Цифровые электронные карты показывающие дороги на исследуемой территории

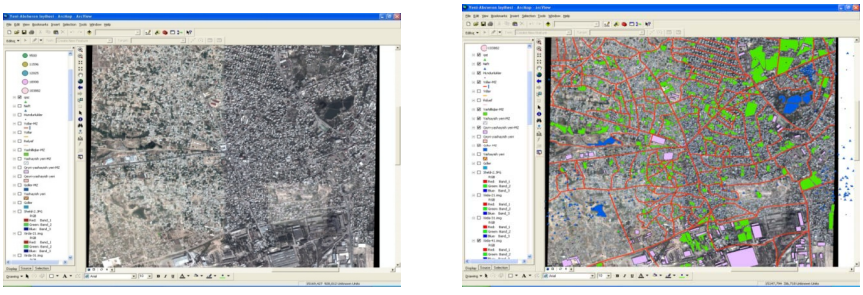


a) b)

Рис. 6. a) Цифровые электронные карты, показывающие рельеф исследуемой территории, b) Цифровые электронные карты, показывающие высоту распределение рельефа и нефтегазовые месторождения

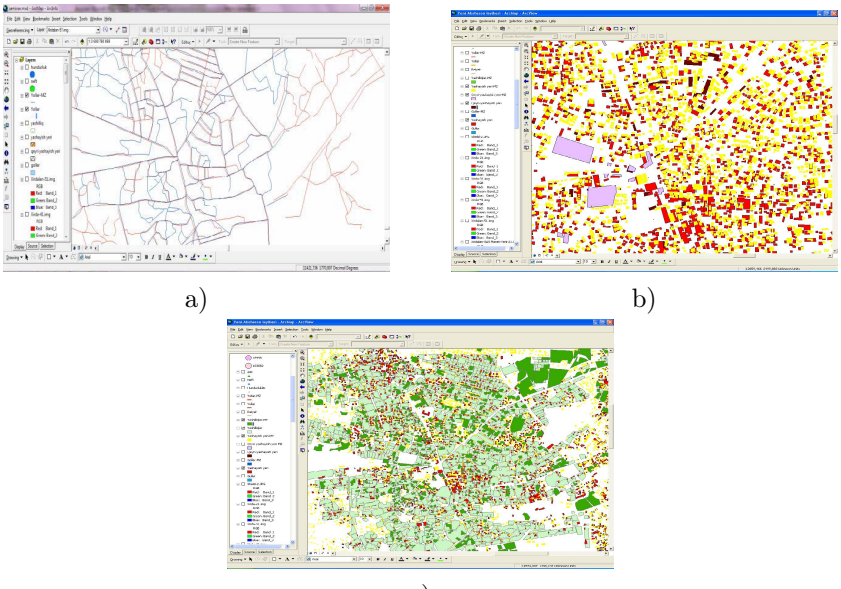
ском районе в результате прироста населения, были использованы космические снимки Landsat TM 2015г. После совмещения снимка Landsat TM с цифровые электронными картами были выявлены изменения и сделаны соответствующие обновления. На (рис. 7 и рис. 8) показаны топографические карта 2005 года и космический снимок 2015 года соответственно [4].

В результате сравнения цифровых электронных карт Апшеронского экономического района созданных на основе космических снимков и топографических карт исследуемой территории, было выявлено, что ускоренный прирост населения привел к увеличению количество и объема строительных работ, увеличению числа жилых зданий, нежилых объектов и дорог, построенных в эти годы. Увеличение объема строительных работ привело к уменьшению территорий занятых растительностью, а это привело к ухудшению экономической ситуации в районе.



a) b)

Рис. 7. a) Космический снимок, исследуемой территории, b) различные слои указанные на космическом снимке – населенные пункты, незаселенные места, дороги, гидрографические элементы, озеленения



a) b)

c)

Рис. 8. Обновления на картах, показывающих исследуемую территорию: a) красным цветом показаны дороги на топографической карте, синим цветом показаны дороги на космическом снимке, b) красным цветом показаны поселенные пункты на топографической карте, желтым цветом – на космическом снимке поселенные пункты, c) светло зеленым цветом показаны зеленые насаждения на топографической карте, темно зеленым – то же на космическом снимке

Литература. **1.** P.Y. Nağiyev, X.ə.əsədov, Ş.Q.Mahmudova. Kosmik təsvirlərin rəqəmli təhlili nəticəsində Abşeron yarimadasinda torpaq sahələrinin ekoloji vəziyyətinin öyrənilməsi. Azərbaycan Elmi, 2006, № 4–5, səh.30–37. **2.** T.İ. Süleymanov, Ş.F.Müttəlibova, E.R. ələsgərov. Coğrafi məlumat sisteminin proqram mühitində rəqəmli xəritələrin yaradilmasi. Труды Международной конференции "Проблемы кибернетики и информатики", Баку, 24–26 октябрь, 2006 г., С.96–98. **3.** Абламейко С.В., Апарин Г.П., Крючков А.Н. Географические информационные системы. Создание цифровых карт. – Мн.: Ин-т техн. кибернетики НАН Беларуси, 2000, 276 с. **4.** Журкин И.Г., Цветков В.Я. Геоинформационное моделирование в ГИС при обработке данных дистанционного зондирования // Исследование Земли из космоса, 1998, №6, С.66–72.

Поворознюк Н.І., Гридчук Д.Т., Шибін О.В.
ФЕА, НТУУ "КПІ", Київ, Україна

Частотно-часовий аналіз сигналів при аускультації

Вступ. Серцево-судинні захворювання складають значну частку захворювань у цілому і продовжують зростати по мірі старіння населення. За даними Американської асоціації серця (AHA — American Heart Association) втрати бюджету країни, обумовлені серцево-судинними захворюваннями, складали $ 448 млрд в 2008 році. Раннє виявлення і профілактика серцево-судинних захворювань дає змогу не тільки значно скоротити ці витрати, а й продовжити середню тривалість життя. Одним з недорогих і ефективних засобів ранньої діагностики серцево-судинних захворювань є аускультація — прослуховування звуків серця [1–3].

Рішення задачі. Під час роботи серця виникають звукові нестаціонарні коливання, обумовлення відкриванням і закриванням клапанів серця, а також хаотичні коливання (шуми серця) обумовлені турбулентним режимом руху крові по судинах. Розрізняють чотири види звукових коливань (heart sound) — S1, S2, S3, S4, причому S3, S4 виникають не завжди і мають значно меншу інтенсив-ність. Шуми можуть виникати як у період систоли, так і у період діастоли. Відхилення від нормальної роботи серця, обумовлені його захворюваннями, спричинюють зміну характеристик як звуків серця, так і шумів. Отже, за допомогою аускультації, тобто прослуховуванням звуків серця, можна діагностувати його захворювання на ранніх стадіях [4, 5]. Безпосереднє прослуховування звуків серця обмежено можливостями людського слуху, оскільки більша частина частотного діапазону і діапазону інтенсивності звуків серця знаходиться поза межами людського слуху. Для розширення можливості прослуховування звуків серця французький лікар Лане (René Théophile Hyacinthe Laennec) у 1819 році винайшов і виготовив стетоскоп — пристрій для прослуховування звуків серця. Цей пристрій значно збільшив можливості аускультації. У той же час, діагностика відхилень у роботі серця за допомогою стетоскопа вимагає від медичного персоналу високої кваліфікації, яка досягається тривалим інтенсивним навчанням і багатолітньою практикою. На точність діагностики впливає багато суб'єктивних чинників. Бурхливий розвиток електроніки, мікропроцесорної техніки дає змогу створити електронний стетоскоп, який значно розширює частотний діапазон і підвищує чутливість аускультації. Звукові сигнали діяльності серця, сприйняті стетоскопом, називаються фонокардіограмами. Сучасні методи цифрової обробки сигналів, зокрема методи штучного інтелекту, дають змогу вилучати із фонокардіограм цінну інформацію, що значно підвищує точність діагностики. Звукові сигнали діяльності серця — це нестаціонарні сигнали, частотні компоненти яких змінюються з плином часу. Інформація у таких сигналах міс-титься як у частотній області, так і в області часу. Для аналізу нестаціонарних сигналів розроблені складні методи аналізу у частотно-часовій області. Одним із таких методів є вейвлет-аналіз. У середовищі MATLAB/Simulink було здійснено вейвлет-аналіз фонокардіограм з різними відхиленнями у роботі серця.

Висновок. Результати моделювання показали високу частотно-часову роздільну здатність вейвлет-аналізу, що дає змогу отримати широкий спектр діаг-ностичних ознак і тим самим істотно підвищити точність діагностики захворювань серця.

Література. **1.** M.E. Tavel, Cardiac auscultation: a glorious past—and it does have a future! Circulation 113 (9) (2006) 1255–1259. **2.** Rangayyan,R.M and Lehner ,R.J (1988). Phonocardiogram signal analysis: a review. CRC Critical Reviews in Biomedical Engineering 15 (3), 211-236. **3.** R. L. Watrous, "Computer-Aided Auscultation of the Heart: From Anatomy and Physiology to Diagnostic Decision Support," IEEE EMBC, 2006, pp. 140-143. **4.** A. K. Bhoil, K. S. Sherpa, B.Khandelwal Multidimensional Analytical Study of Heart Sounds: A Review Int. J. Bioautomation, 2015, 19(3), pp. 351-376. **5.** S.K. Randhawa, M. Singh Classification of Heard Sound Signal Using Multi-modal Feature Procedia Computer Science, 2015, 58 pp. 165-171.

Прогонов Д.О.
Фізико-технічний інститут, НТУУ "КПІ", Київ, Україна

Стегодетектор на основі мультифрактального аналізу цифрових зображень

Для виявлення та протидії роботі прихованих каналів несанкціонованої передачі конфіденційних даних в інформаційно-комунікаційних системах широко використовуються методи пасивного та активного стегоаналізу. Одним з найбільш розповсюджених підходів до виявлення повідомлень (стегоданих), прихованих у файли-контейнери, зокрема цифрові зображення (ЦЗ), є використання стегодетекторів, створених на основі методів статистичного стегоаналізу [1]. Дані стегодетектори дозволяють з високою ймовірністю виявляти стегодані, вбудовані у просторовій області зображення-контейнеру (ЗК) згідно LSB-методів, а також в області перетворення контейнеру (ОПК) з використанням двовимірного дискретного косинусного або вейвлет перетворень (ДДКП, ДДВП). Поява новітніх методів вбудовування стегоданих в ОПК, заснованих на використанні як спектральних (ДДВП, ДДКП), так і спеціальних (наприклад, сингулярний розклад) перетворень матриць яскравостей пікселів ЦЗ, призвела до суттєвого зниження ефективності відомих статистичних стегодетекторів [2]. Тому актуальною і важливою задачею є пошук нових методів виявлення заповнених ЗК (стеганограм) незалежно від області прихування повідомлень.

В роботах [3, 4] було показано, що застосування методів структурного аналізу ЦЗ, а саме мультифрактального аналізу (МФА), дозволяє з високою точністю виявляти факт прихування повідомлень в ОПК навіть у випадку слабкого заповнення ЗК стегоданими ($\Delta_C < 10\%$) та мінімальної енергії прихованих повідомлень. Метою даної роботи є розробка стегодетектору на основі МФА цифрових зображень для виявлення стеганограм з даними, вбудованими в ОПК з використанням спектральних та спеціальних перетворень зображення-контейнеру.

В роботі був розроблений стегодетектор SD_{MFA} на основі ансамблю лінійних дискримінантів Фішера. В якості елементів простору ознак SD_{MFA} були використані сформовані у роботах [3,4] кластери демаскуючих ознак стеганограм з даними, вбудованими в ОПК. Загальна розмірність простору ознак стегодетектору склала 14 параметрів (характеристик спектру узагальнених фрактальних розмірностей та мультифрактального спектру ЦЗ).

Дослідження точності виявлення стеганограм з даними, вбудованими в ОПК, при використанні розробленого стегодетектору SD_{MFA} проводилося на стандартному тестовому пакеті цифрових зображень MIRFlickr-25k. В якості стегоданих були використанні кольорові зображення з різним ступенем деталізації: креслення, карта та портрет. Формування стеганограм проводилося згідно одноетапним методам Дея та Агарваля, багатоетапним методам Джозефа та Хана, а також комплексним методам Елайона та Гунджаля [3, 4]. Для оцінки точності виявлення стеганограм при використанні стегодетектору SD_{MFA} були використані стандартні метрики ROC-аналізу – Area under ROC curve (AUC), Matthews Correlation Coefficient (MCC).

За результатами проведених досліджень були побудовані криві залежності значень метрик AUC та MCC від ступеня заповнення ЗК стегоданими при варіації типу та енергії прихованих повідомлень. Встановлено, що SD_{MFA} дозволяє з високою точністю (AUC>0.95) виявляти факт прихування повідомлень в ОПК навіть у випадку малої енергії прихованих повідомлень.

Література. **1.** Fridrich J. Steganography in Digital Media: Principles, Algorithms, and Applications / Fridrich J. – Cambridge University Press, 2009 – p. 437. **2.** Progonov D. Passive Steganalysis of Multidomain Embedding Methods / Progonov D., Kushch S. // International Journal "Information Theories & Applications". – 2015. – Vol. 22, Iss. 1. – pp. 86–99. **3.** Прогонов Д. Спектральный анализ стеганограмм / Прогонов Д., Кущ С. // Научный журнал "Радиоэлектроника, информатика, управление". – 2015. – № 2 (33). – с. 71-81. **4.** Progonov D. On the Multifractal Analysis of Steganograms Progonov D., Kushch S. – Proceeding of the 4th International Scientific Conference of Students and Young Scientist TAAC-2014. – Kyiv: Bukrek, 2014. – pp. 32-38.

Роганов Е.В.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Сравнительный анализ эффективности алгоритмов трекинга

Разработка трекеров остается сложной проблемой в области компьютерного зрения на протяжении десятилетий. Возможная область применения трекинга очень широка, от наблюдения до мониторинга для умных комнат [1].

Отслеживание объектов зависит от нескольких факторов, таких как наличие априорного знания об объекте, а также от количества и типа параметров трекинга (местоположение, масштаб и контур вокруг объекта). Для применения оптимального алгоритма в конкретной ситуации, необходимо знать сильные и слабые стороны различных алгоритмов.

Постановка задачи. В данной работе проведен сравнительный анализ шести алгоритмов в условиях трекинга небольших объектов на видео с низким разрешениям. Трекинг в таких условиях остается сложной задачей по разным причинам: изменения врешнего вида объекта, изменение освещенности, зашумленный фон, различные степени движения объекта. В частности, анализ проведен в следующих сценариях: обработка ситуаций заграждения объекта, пренебрежимо малое движение, низкая контрастность, зашумленный фон.

Алгоритмы. Для анализа было выбрано шесть современных алгоритмов трекинга: инкрементальный трекер (IVT) [2], алгоритм многоэкземплярного обучения (MIL) [3], L_1 трекер [4], трекер наименьших частичных квадратов (PLS) [5], разреженный онлайн трекер (SOT), пространственно-временной трекер (STC).

Результаты. Общая эффективность алгоритмов трекинга оценивалась по сравнению с размеченным вручную эталоном. Точность высчитывалась с помощью двух метрик: ошибки локализации (евклидово расстояние между центрами) и точности перектытия (взаимное перекрытие ограничивающих прямоугольников используя коэффициент Дайса).

Таблица 1. Метрики эффективности

Алгоритм	Локализация	Перекрытие (%)
PLS	1.8	70.96
SOT	4.7	63.61
IVT	4.7	63.61
L_1	7.8	66.61
STC	10.0	61.65
MIT	70.7	34.54

Алгоритм PLS показал наибольшую эффективность в целом (с ошибкой локализации 1.8 пикселей). Алгоритм MIL показал плохие результаты в сценариях зашумленного фона, низкой контрастности и незначительного движения, однако работает исключительно хорошо в случае заграждения объекта благодаря самообучаемости и возможности начать повторно отслеживать объект после временной потери. Алгоритм L_1 теряет цель в случае заграждения объекта и не может повторно захватить ее. Трекеры SOT и IVT сравнимы в эффективности и оба работают плохо в случае заграждений. STC трекер отлично отслеживает загражденные объекты, однако качество трекинга значительно уменьшается на менее контрастных видео из-за малого пространственно-временного контекста. Таким образом, только алгоритм PLS работает одинаково хорошо во всех случаях, в то время как остальные рассмотренные алгоритмы не справляются с задачей хотя бы в одном сценарии.

Литература. **1.** Maggio E., Cavallaro A. Video Tracking: Theory and Practice. Somerset, San-Francisco: Wiley&Sons, 2011. **2.** D. Ross, J. Lim, R.-S. Lin, M.-H Yang, "Incremental learning for robust visual tracking", International Journal of Computer Vision, 2007. **3.** B. Babenko, M.-H. Yang, S. Belongie, "Visual tracking with online multiple instance learning IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2009. **4.** Bao, C., Wu, Y., Ling, H., "Real time robust L1 tracker using accelerated proximal gradient approach", IEEE Conf on Computer Vision and Pattern Recognition (CVPR), pp. 1830–1837, 2012. **5.** Q. Wang, F. Chen, W. Xu, M.-H. Yang, "Object tracking via partial least squares analysis", IEEE Trans. on Image Processing, pp. 4454–4465, 2012.

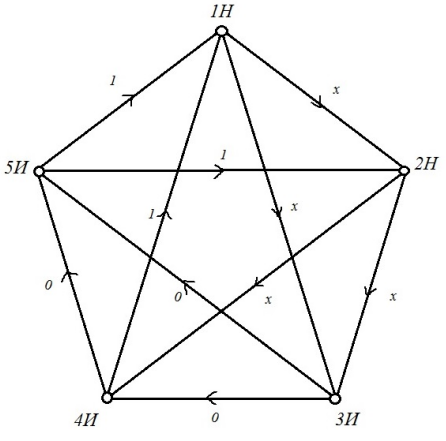
Романкевич В.А., Дейнеко А.Л., Олейник В.В.
Факультет прикладной математики НТУУ "КПИ", Киев, Украина

О диагностировании многопроцессорных систем с ограничением на число допустимых неисправностей

Многопроцессорные системы с каждым годом находят все более широкое применение, в частности, в системах управления сложными объектами. Такие системы обладают широкими возможностями и высокими показателями по быстродействию и надежности. Последнее имеет место, в первую очередь, в реконфигурируемых многопроцессорных системах, где функции, которые должен выполнять отказавший процессор, могут быть подхвачены теми, которые остались работоспособными. Сказанное справедливо для так называемых отказоустойчивых многопроцессорных систем (ОМС), для которых характерно определенное ограничение по допустимому количеству T неисправных процессоров, после появления которых система может продолжать функционировать в полном объеме. Неисправные процессоры в ОМС определяются путем взаимного тестирования [1]. Отметим, что надежность, безопасность, диагностируемость, отказоустойчивость, - основные компоненты теории гарантоспособности.

В общем случае число каналов связи между процессорами серьезно зависит от величины T, и способы определения исправности тестируемого процессора бывают разными. Например, при T=1состояние процессора (исправен - неисправен) может быть определено по большинству результатов тестирования, и тогда тестирующих должно быть не менее трех.

В докладе рассматривается случай T=2 и предлагается использовать направленный диагностический граф, названный конструктивно регулярным, который имеет всего две входящие и две выходящие дуги для каждой вершины. Одно из преимуществ такого графа – число процессоров может быть любым (что не свойственно, например, матричной системе коммутации или гиперкубу). Пример конструктивно регулярного графа для случая 5-и процессоров приведен на рисунке.



Вершины пронумерованы (как процессоры в ОМС), буквы Н и И означают, соответственно неисправное и исправное состояние. В качестве модели неисправностей выбрана широко известная модель Препарата – Метца – Чена, и стрелки обозначены знаками, соответсвующими результатам тестирования. Напомним, что эта модель предполагает результат v_{ij} тестирования i-м процессором j-го следующим образом: если i,j – исправны, v_{ij}=0; если i - исправен, j – неисправен, v_{ij}=1; если i – неисправен, v_{ij}=x. Здесь x означает неопределенный результат: x принимает значение либо 1, либо 0 независимо от состояния тестируемого процессора. Суть предлагаемого метода – выполнение всех взаимных проверок и последующий анализ, который включает в себя поиск 0-цепочек (множество вершин, объединенных дугами со знаками 0). Последний элемент цепочки имеет вес, равный количеству вершин цепочки. Если этот вес превышает величину T, то соответствующий процессор исправен. В нашем примере 5-й элемент имеет вес 3, что превышает T=2, следовательно, он исправен. В докладе доказывается, что имеет место полное диагностирование, то есть определение состояния всех процессоров при любом взаимном расположении исправных и неисправных и при любом (не менее 5-и) числе процессоров системы.

Литература. 1. Рабочее диагностирование безопасных информационно-управляющих систем / А.В.Дрозд, В.С.Харченко, С.Г.Антощук, Ю.В.Дрозд, М.А.Дрозд, Ю.Ю.Сулима / Под ред. А.В.Дрозда и В.С.Харченко.- Харьков, 2012.- 614с.

Романкевич В.А., Малышева М.О., Примак И.К.
Факультет прикладной математики НТУУ "КПИ", Киев, Украина

Об одном способе формирования реберных функций GL-модели

При проектировании реконфигурируемых отказоустойчивых многопроцессорных систем (ОМС) управления сложными объектами важным этапом является расчет показателей их надежности. Для осуществления подобных расчетов используются различные модели, позволяющие получить отражение реакции ОМС на появление отказов. Ниже остановимся на одной из них – графо-логической модели (GL-модели), которая использует элементы теории графов и булевой алгебры для отображения состояния системы. Расчет осуществляется статистическим методом.

GL-модель представляет собой неориентированный циклический связный граф, ребрам которого ставятся в соответствие реберные функции, зависящие от состояния модулей в системе (если элемент системы отказал, то описывающая его булева переменная обнуляется; если элемент работоспособен, переменная равна 1). Ребра, булевы функции которых при конкретном состоянии системы равны 0, считаются отсутствующими, что влияет на связность графа. Работоспособность системы описывается связностью графа, при отказе системы связность нарушается.

ОМС, состоящая из n модулей сохраняющая работоспособность при отказе любых m из них, называется базовой и обозначается K(m, n) [1, 2].

Иногда с целью повышения надежности системы разработчику приходится блокировать отказ системы, при появлении отказов компонентов повышенной кратности. Известны два метода модификации GL-моделей для отображения подобного блокирования:

- проведение внутренних ребер;
- модификация реберных функций.

В докладе выбран второй метод и предлагается алгоритм построения модели с целью уменьшения сложности реберных функций.

Рассмотрим K(3,7). По алгоритму, предложенному в работе [3], элементы системы разбиваются на две группы, составляется таблица, на основании которой формируются реберные функции:

$$f_1 = x_1 \vee x_2 \vee x_3$$

$$f_2 = x_4 \vee x_5 \vee x_6 x_7$$

$$f_3 = x_4 x_5 \vee x_6 \vee x_7$$

$$f_4 = x_4 x_5 x_6 x_7 \vee (x_1 \vee x_2)(x_1 x_2 \vee x_3)$$

$$f_5 = x_1 x_2 x_3 \vee (x_4 \vee x_5)(x_6 \vee x_7)(x_4 x_5 \vee x_6 x_7)$$

Пусть разработчик трансформировал систему таким образом, что она сохраняет работоспособность при появлении отказа элементов $\{x_1, x_2, x_4, x_6\}$. Для того, чтобы модель сохраняла адекватность поведению системы в потоке отказов, её необходимо модифицировать.

При появлении соответствующего вектора состояния ОМС (т.е. с 0 значением состояния модулей x_1, x_2, x_4, x_6 функции f_4 и f_5 становятся равными нулю. Решением может быть объединение одной из этих функций с соответствующей конституентой «1» для того, чтобы граф не потерял связность в данном конкретном случае. Тогда данные функции будут иметь следующий вид:

$$f_4 = x_4 x_5 x_6 x_7 \vee (x_1 \vee x_2)(x_1 x_2 \vee x_3\ ) \vee \overline{x_1 x_2} x_3 \overline{x_4} x_5 \overline{x_6} x_7$$

$$f_5 = x_1x_2x_3 \vee (x_4 \vee x_5)(x_6 \vee x_7)(x_4x_5 \vee x_6x_7) \vee \overline{x_1x_2}x_3\overline{x_4}x_5\overline{x_6}x_7$$

Выполнив минимизацию данных функций, получим:

$$f_4 = x_1x_2 \vee x_1x_3 \vee x_2x_3 \vee x_3\overline{x_4}x_5\overline{x_6}x_7 \vee x_4x_5x_6x_7$$

$$f_5 = x_1x_2x_3 \vee \overline{x_1x_2}x_3x_5x_7 \vee x_4x_5x_6 \vee x_4x_5x_7 \vee x_4x_6x_7 \vee x_5x_6x_7$$

Полученные выше результаты предполагают разбиение множеств элементов на подмножества одинаковой мощности, не конкретизируя, в какое из подмножеств какие элементы входят.

В докладе предлагается разбивать элементы по множествам в зависимости от частоты вхождения их в векторы, которые необходимо блокировать. Рассматривая предложенный ранее пример модели, в которой должен блокироваться один вектор с нулевыми позициями $\{x_1, x_2, x_4, x_6\}$, элементы системы можно разделить на те, которые входят в блокируемый вектор, и те, которые не входят. Соответственно, множество переменных системы будет разбито на два подмножества $\{x_3, x_5, x_7\}$ и $\{x_1, x_2, x_4, x_6\}$.

Используя метод, аналогичный указанному выше, получаем следующие реберные функции:

$$f_1 = x_3 \vee x_5 \vee x_7$$

$$f_2 = x_1 \vee x_2 \vee x_4x_6$$

$$f_3 = x_1x_2 \vee x_4 \vee x_6$$

$$f_4 = x_1x_2x_4x_6 \vee (x_3 \vee x_5)(x_3x_5 \vee x_7)$$

$$f_5 = x_3x_5x_7 \vee (x_1 \vee x_2)(x_4 \vee x_6)(x_1x_2 \vee x_4x_6)$$

Теперь при появлении вектора с 0 значениями на позициях $\{x_1, x_2, x_4, x_6\}$ обнуляются функции f_2 и f_3. Тогда, достаточным условием корректного функционирования системы при указанных отказах является добавление к обнуляющимся функциям конъюнкции элементов, которые не входят в блокируемый вектор. После модификации функции f_2 и f_3 приобретут следующий вид:

$$f_2 = x_1 \vee x_2 \vee x_4x_6 \vee x_3x_5x_7$$

$$f_3 = x_1x_2 \vee x_4 \vee x_6 \vee x_3x_5x_7$$

Понятно, что предложенный способ приводит к более простым реберным функциям. Кроме того, упрощается процесс преобразования функций, т.к. предложенный способ не требует дизъюнкции функций с конституэнтой «1» и последующей их минимизации.

Литература. **1.** Романкевич А.М., Орловский А.С., Романкевич В.А. Алгоритм синтеза базовой GL-модели // Вісник НТУУ "КПІ".- Інформатика, управління та ОТ.-2001.- №32.- С.132-135. **2.** Романкевич А.М., Иванов В.В., Романкевич В.А. Анализ отказоустойчивых многомодульных систем со сложным распределением отказов на основе циклических GL-моделей // Электронное моделирование.-№5, т.26, 2004, с.67-81 **3.** Романкевич А.М., Карачун Л.Ф., Романкевич В.А. Графо-логические модели для анализа сложных отказоустойчивых вычислительных систем // Электронное моделирование.- 2001.-т.23, №1.- С.102-111.

Романовський О.В., Сало А.М.
Кафедра електронних обчислювальних машин, Національний університет «Львівська політехніка», Львів, Україна

Технології захисту платіжних транзакцій у кіберфізичних системах

Платіжна транзакція (ПТ) – це дія або серія дій, яка виконується користувачем, в результаті якої здійснюються операції переказу, вводу або виводу грошей на/із рахунку. З точки зору платіжної системи (ПС), транзакція – це взаємодія користувача із прикладним програмним забезпеченням (ПЗ), яке здійснює зміни або надає доступ до бази даних (БД). Роглядаючи ПС, як складову кіберфізичних систем (КФС), варто зазначити, що захист ПТ є однією із пріорітених задач, оскільки виявлення і усунення недоліків значно підвищить безпеку ПТ та КФС вцілому. На даний момент у КФС розрізняють онлайн ПТ (76,9%), які здійснюються у режимі реального часу, і оффлайн транзакції (23,1%) (рис.1).

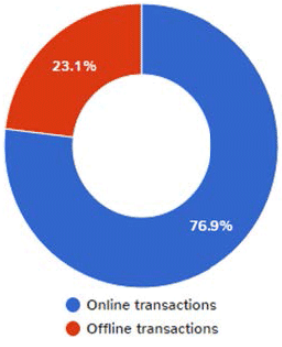


Рис. 1. Порівняльна діаграма онлайн/оффлайн транзакцій у кіберфізичних системах

ПТ досить динамічно розвиваються у високорозвинених країнах, таких як США, Китай та країнах ЄС. Про це свідчать такі статистичні дані: поточне значення ПТ роздрібної торгівлі у США сягає 294 млн.$, а до 2018 року цей показник досягне рівня 414 млн.$; 93 % (у США) і 81% (у країнах ЄС) підприємств роздрібної торгівлі бажають встановити собі POS-термінал; у 2016 році Китай очікує від ПТ прибуток у розмірі 562,7 млрд.$; «мобільні» ПТ також не відстають - 174 млн. споживачів США (72 %) в даний час здійснюють ПТ зі смартфонів і 93 млн. (38 %) – з планшетів.

Процес захисту ПТ у КФС – це сукупність дій, що пов'язані із захистом транзакції від моменту її формування клієнтом, до передачі інформації банку еквайеру (БЕ), який відправляє запит на авторизацію у процесінговий центр (ПЦ), а ПЦ в свою чергу має підтвердити інформацію про дійсність платежу, або відхилити транзакцію (рис.2.). Отже, можна зробити висновок, що захист ПТ у КФС складається з 2-х основних компонент: «клієнтської частини» та «сукупності банківських серверів».

Рис. 2. Схема захисту платіжних транзакцій у кіберфізичних системах

Розглянувши рис. 2, захист «клієнтської частини» можна поділити на 3 підкатегорії: захист мобільних пристроїв, захист веб-сайтів (веб-додатків) та захист терміналів оплати. Захист мобільних пристроїв в свою чергу можна поділити ще на 2 підкатегорії:

- програмний захист: наприклад, встановлення антивірусного ПЗ, відмова від «root» (для Android-користувачів) та jailbreak (для Apple-користувачів), що надасть безпеку конфіденційної інформації та унеможливить використання вашого пристрою для DDoS-атак;
- апаратний захист: захист NFC-чіпу від атак з використанням експлойта, «прослуховування», модифікування даних та використання ретрансляції, а також, апаратний захист від «root» на основі eFuse-чіпа (наразі, єдиний представник – Motorola Droid X), який був розроблений для запобігання модифікації ПЗ телефону та викрадення інформації та коштів з Android Pay. Захист веб-сайтів (веб-додатків) здійснюється програмними методами, такими як протокол 3D Secure, який заснований на принципі 3-х доменів (звідси і назва): домен еквайера, домен емітента та домен сумісності; технологія OTP (one time password) – пароль, який використовується для одного сеансу автентифікації (наприклад, паролі на скретч-картах або SMS-паролі). Захист терміналів оплати є ва-

жливою складовою, оскільки POS-термінали (Point-of-Sale) забезпечують розрахунок продавця і покупця методом перевірки дебетових та кредитних карток безпосередньо в місцях продажу товарів та послуг (наприклад супермаркет, кінотеатр, АЗС тощо). На даний момент, аналітики виділили 3 найпопулярніші сервіси захисту POS-терміналів:

- токенізація - платіжні дані заміняються на їх на шифр при проведенні карти через POS-термінал, а номер платіжного інструменту замінюється на набір випадкових символів і втрачає сенс для злочинців, але даний сервіс не може забезпечити безпеку даних в зараженому платіжному терміналі. Зафіксовані випадки, коли хакери встановлювали шкідливе ПЗ на термінал і програмно перехоплювали платіжні дані до перетворення в шифр;
- EMV - вразливість карт із магнітною смугою змусила міжнародні платіжні системи створити новий стандарт платіжних інструментів - карти на основі мікрочіпа, які дозволяють генерувати динамічний код для кожної транзакції і проводити платежі на безконтактній основі, але ця технологія відносно малорозвинена (на сьогоднішній день стандарту EMV відповідає тільки 32% світових платежів); - сервіс шифрування P2PE дозволяє закодувати платіжну інформацію всередині терміналу і далі передавати її по мережі в захищеному вигляді, але дана система коштує доволі дорого (пристрій, який керує закодованою інформацією може коштувати до 100000 $);

Усі ці методи відіграють певну роль у вдосконаленні захисту транзакцій, однак тільки в комплексі три технології захисту зможуть ефективно боротися з платіжним шахрайством.. Захист «сукупності банківських серверів» реалізується згідно стандарту безпеки даних PCI DSS (Payment Card Industry Data Security Standard). Даний стандарт детально описує основні 6 областей контролю безпеки даних тримачів дебетових та кредитних карток, що оброблюються, зберігаються та передаються та 12 основних вимог по їх безпеці, таких як: побудова і підтримка захищених мереж (встановлення та коректна робота брандмауерів (firewall) для більш надійного інформаційного захисту даних та зміна передбачених виробниками паролів), захист персональних даних власників карт (надійний захист даних користувачів та їх шифрування в разі передачі через мережу Internet), контроль програм управління «вразливих місць» системи (постійне оновлення антивірусу, забезпечення його постійного функціонування та розробка безпечних додатків та систем вцілому), забезпечення строгого контролю доступу (можливі обмеження доступу зі сторони обслуговуючого персоналу, наявність власних ID у кожного працівника, що працює із системою та максимальне унеможливлення фізичного доступу до персональних даних власників), тестування та моніторинг мережі («логування» усіх сеансів доступу до мережевих ресурсів і даних тримачів карт та постійна перевірка системи), політика «конфіденційності» (розробка політики конфіденційності, її виконання та постійна підтримка даної програми).

Проаналізувавши статистику минулих років можна із впевненістю сказати, що кількість ПТ зростає і буде зростати, а ПС, як складові КФС потребують всебічного захисту: як зі сторони клієнтів, так і зі сторони банків, процесінгових центрів тощо.

Література. **1.** Карнаухова Е. Ю. «Мошенничество в сфере безналичных расчетов с использованием банковских платежных карт» [Текст] // Юридические науки: проблемы и перспективы: материалы II междунар. науч. конф. (г. Пермь, январь 2014 г.). — Пермь: Меркурий, 2014. — С. 42-44. **2.** Сайт «MagazinePaySpace» psm7.com/novyj-opasnyj-virus-voruet.... **3.** NFS, ru.wikipedia.org/wiki/Near_Field_Communication. **4.** graphs.net, graphs.net/online-vs-offline-shopping-stats.html. **5.** nchannel nchannel.com/blog/retail-data-ecommerce-statistics/. **6.** П.Ю. Грицюк, Ю.І. Грицюк «ОСОБЛИВОСТІ ЗАХИСТУ ЕЛЕКТРОННИХ ПЛАТІЖНИХ СИСТЕМ У МЕРЕЖІ ІНТЕРНЕТ», Національний лісотехнічний університет України, Науковий вісник НЛТУ України. – 2013. – Вип. 23.10, с. 314 – 331. **7.** Сайт «Информзащита», itsecurity.ru/study-center/blog/zashchita-bank. **8.** 3-D Secure, ru.wikipedia.org/wiki/3-D_Secure. **9.** PCI DSS, ru.wikipedia.org/wiki/PCI_DSS. **10.** Система LiqPay, uk.wikipedia.org/wiki/LiqPay

Рубець А.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Моделювання систем високої доступності

На відміну від розробника, робота системного адміністратора включає забезпечення безперервної роботи сервісів, оновлення систем та програмного забезпечення, резервування, впровадження нових сервісів, підтримку мережі та багато інших. Тому однією з проблем для системного адміністратора, окрім впровадження та підтримки серверної інфраструктури, є не повна або відсутня система для її попереднього тестування. Окрім цього, всі задачі повинні бути протестовані для реальних систем для досягнення мінімального часу простою у надзвичайних ситуаціях і без системи моделювання ця задача стає досить трудомісткою.

Для забезпечення коректного процесу тестування та моделювання необхідно виконати наступні кроки: автоматизація процесу конфігурації, створення моделі реальної серверної інфраструктури та створення безперервної системи тестування.

Автоматизація конфігурування. Перший етап включає перехід від ручного налаштування серверів на систему автоматичних конфігурації, що дозволяє на основі описаних завдань здійснювати інсталяцію систем та сервісів в автоматичному режимі для заданого середовища без участі людини. Крім цього, увесь опис для цієї системи необхідно перенести у систему контролю версій, для сумісної роботи команди підтримки серверної інфраструктури. У цьому випадку зупинимось на системі Ansible[1], як основа для опису серверів та їх базової конфігурації. У системі Ansible присутня можливість використання мови програмування Python, що дозволяє досить легко інтегрувати будь-які додаткові можливості поверх вже існуючих. Для контролю версій кодової бази використовується досить популярний сервіс git[3], що має багато додаткових інструментів, включаючи інструменти, що спрощують процес тестування.

Створення моделі серверної інфраструктури. Для реалізації практичної моделі можна використовувати повномасштабні моделі, але це передбачає збільшення витрат та збільшення часу на інсталяцію серверів та сервісів з нуля. Другий варіант, що є менш витратним і більш гнучким, включає створення зменшеної моделі середовища за допомогою систем віртуалізації. Для Windows систем систем найкраще підходить технологія HyperV від Microsoft. У випадку Linux зручним засобом буде використання контейнерної технології Docker[2], що дозволяє досить просто та швидко створювати необхідні конфігурації. Хоча у разі використання системи Docker виникає ряд проблем, що пов'язаний з відсутністю системи контролю запущених сервісів, наприклад systemD, оскільки, системи автоматичних конфігурацій покладаються саме на них. Для вирішення проблеми було розроблено шаблон контейнера, що включає емуляції даного сервісу.

Створення системи тестування. Для забезпечення системи тестування використовується система безперервної інтеграції, що забезпечує автоматичний запуск тестування за допомогою системи git-hook у сервісі git. Також реалізований додаток, що виконує функції супервізора, який пов'язує усі вищеописані системи та надає інструментарій для моніторінгу процесу моделювання та генерації звітів з подальшої відправкою їх на пошту.

Висновки. Питання автоматизації стоїть досить гостро, оскільки, системи високої доступності накладають досить суворі умови щодо надійності систем та сервісів, які лежать в їх основі. Даний сервіс забезпечує спеціаліста з інформаційних технологій досить потужним інструментом для тестування та моделювання серверної інфраструктури, для створення единої бази тестів та опису моделей надзвичайних ситуацій.

Література. **1.** Ansible, http://docs.ansible.com/ **2.** Docker, https://docs.docker.com/ **3.** Git, https://git-scm.com/documentation

Савчук В.В., Лозицький О.А., Пасічник В.В.
Інститут комп'ютерних наук та інформаційних технологій НУ "Львівська політехніка", Львів, Україна

Динамічне формування персоніфікованого мультимедійного контенту туристичного путівника міських екскурсійних маршрутів

Одним із актуальних завдань, що постали перед розробниками інтелектуальної системи «Мобільний інформаційний асистент туриста» [1] є автоматизоване динамічне формування індивідуального мультимедійного контенту призначеного для мобільного інформаційного супроводу туриста під час здійснення ним екскурсії.

Передумовою його вирішення є створення інформаційно-технологічного інструментарію для динамічного формування та озвучення особистісно-орієнтованого екскурсійного контенту, що подається у відповідному програмно-алгоритмічному застосунку користувацького мобільного пристрою (планшета чи смартфона) [2].

Одним з ключових елементів інтелектуальної системи «MIAT» при забезпеченні виконання функцій інформаційного супроводу користувача під час його перебування на певному туристичному маршруті є синхронізація різнотипового мультимедійного контенту, його впорядкування та інтеграція у єдиній базі даних/знань. Саме такі завдання виконує мова, створена консорціумом W3C на базі засобів XML, і призначена, зокрема, для синхронізації компонентів мультимедійної інформації Synchronized Multimedia Integration Language (SMIL) [3].

Мова SMIL дозволяє описувати динаміку мультимедійних об'єктів і синхронізувати відтворення окремих їх компонентів. SMIL є мовою розмітки, яка дозволяє одним тегом описати певний складний сценарій для окремих фігур (наприклад, переміщення фігури, її трансформацію чи візуалізацію тощо).

Книга формату DAISY (Digital Accessible Information System – доступна цифрова інформаційна система) являє собою набір файлів поданих здебільшого у форматах XML та SMIL, з яких можна виділити зв'язки та посилання на окремі частини книги (речення, абзаци, підрозділи, розділи і т. ін.) і шляхи доступу до аудіо, відео та графічних файлів. Засоби мови SMIL дають змогу отримувати інформацію про тривалість аудіо та відео файлів, їх накладання, а також відповідність текстовим фрагментам і рисункам, які зустрічаються в книзі [4].

Для всіх елементів книги задається початкова відмітка програвання, а також, при відомій тривалості медіа файлу визначається кінцева відмітка програвання, або знаючи кінцеву відмітку – можна визначити тривалість відповідного фрагменту.

Можна отримати або змінити контрольні точки відтворення синхронізованих елементів книги, доступатись до відповідного контенту відео/аудіо/тексту, а саме до початкових та кінцевих відміток програвання і тривалості звучання (перегляду) конкретного мультимедійного елементу.

Виклад основного матеріалу. «Кванти знань» являють собою певні абзаци DAISY-книги, які містять вичерпну інформацію про туристичні об'єкти, такі як архітектурні пам'ятки, музеї, галереї, замки, палаци, пам'ятники, тощо (див. рис. 1). Інформація, що подана в DAISY-путівнику (туристичний путівник-книга у форматі DAISY), супроводжується додатковими відомостями, що асоціюються з місцем розташування тих чи інших туристичних об'єктів.

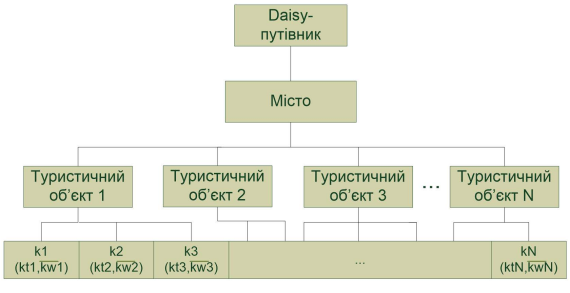


Рис. 1. Структура DAISY-путівника

Кожному туристичному об'єкту відповідають декілька «квантів знань», що відрізняються

за тривалістю відтворення, типом подання відомостей відповідно до можливих варіантів специфікацій профілів користувачів.

Екскурсійний маршрут задається у формі напрямленого графа, вершинами якого є цільові точки (див. рис. 2), ребрами є переходи між ними, що навантажуються часовими значеннями їх тривалості, петлі в графі відображають можливість зупинок в тому чи іншому туристичному місці, а відповідні петлі навантажуються часовою тривалістю такої зупинки (див. рис. 3).

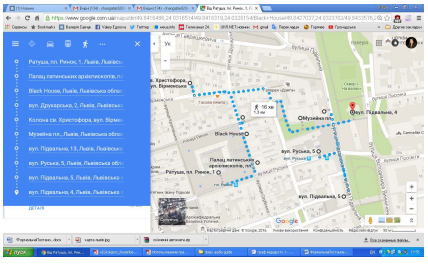


Рис. 2. Картографічне подання екскурсійного маршруту в центральній частині м. Львів

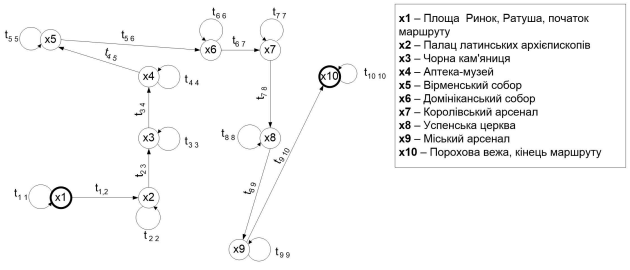


Рис. 3. Графове подання екскурсійного маршруту прокладеного в центральні частині м. Львів

Користувач мобільного програмно-алгоритмічного застосунку повинен мати можливість самостійно обирати ключові точки майбутнього екскурсійного маршруту та задавати послідовність їх відвідування. За допомогою інструменту *Google Maps* турист вибирає цільові точки маршруту та порядок їх відвідування, а також уточнює якою саме траєкторією буде відбуватись його екскурсійний тур. Маршрут руху туриста під час спланованої ним індивідуальної екскурсії задається переліком цільових точок:

$$\overline{x} = [x_1,\ \ldots,\ x_n] \tag{1}$$

Для прикладу на рисунку 3 вершини графа x_1 та x_{10} є відповідно початковою та кінцевою точками маршруту, а $(x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9)$ – проміжними.

Тривалість перебування користувача на екскурсійному маршруті залежить від тривалостей переходів між туристичними об'єктами (точками інтересів), та тривалостей зупинок відведених на огляд туристичних об'єктів. Часова матриця переміщень подається тривалостями переміщень між відповідними цільовими точками на маршруті, а її діагональні елементи подають значення тривалості зупинок біля точок інтересів на екскурсійному маршруті.

Мультимедійний інформаційний контент формується для кожної цільової точки маршруту в порядку їх відвідування та залежить від цільового об'єкту, його порядкового номеру на маршруті, тривалостей зупинки на зазначеному об'єкті і переходу до наступного об'єкту, наявних у базі даних «квантів знань», тривалості їх відтворення та відповідності профілю користувача, а також інформації відібраної для супроводу туриста на попередніх об'єктах туристичного маршруту. Контент інформаційного супроводу екскурсії формується як результат об'єднання відібраних "квантів знань що співставлені з відповідними туристичними об'єктами маршруту.

Література. **1.** Savchuk V.V. Dynamic characteristics of perspective touristic information technologies / Artemenko O.I., Kunanec N.E., Pasichnyk V.V., Savchuk V.V. // ECONTECHMOD: international quarterly journal.journal – 2015, – Vol. 04, No. 4. – pp.107-118 **2.** Savchuk V.V. Mobile information technologies for tourism domain (Мобільні інформаційні технології для туристичної галузі) / Pasichnyk V.V., Savchuk V.V. // ECONTECHMOD: international quarterly journal. – 2015, - Vol. 04, No. 2. - 25–32. **3.** Synchronized Multimedia Integration Language (SMIL 3.0), W3C Recommendation 01 December 2008, https://www.w3.org/TR/SMIL3. **4.** XHTML+SMIL Profile, W3C Note 31 January 2002, http://www.w3.org/TR/XHTMLplusSMIL.

Сергеєв Д.С.
Факультет інформатики та обчислювальної техніки НТУУ "КПІ", Київ, Україна

Методика оцінки якості роботи природно-мовних пошукових систем

Задача зберігання природно-мовної інформації (ПМІ) у структурованому вигляді має досить довгу історію. Перші спроби створити онтологію (модель світу) на основі текстової інформації мали місце ще у 1974 (мова *KL-ONE*) [1], незабаром після створення перших текстових моніторів — терміналів. Основною перешкодою на шляху популяризації цих онтологій була їх повнота, або здатність до автоматичного наповнення: формування навіть простої моделі світу вимагало значних витрат як робочого часу оператора, так і машинного часу. Ця проблема була вирішена з появою і розвитком мережі Інтернет: тепер тисячі людей добровільно вносили окремі частки інформації про світ у єдину мережу. Але слідом за зростанням загальної кількості документів виникла нова проблема — категоризації, а незабаром й індексації вмісту мережі — тобто, потреба у пошукових системах (ПС). Першою ластівкою тут стала *JumpStation* [2], яка реалізовувала ідею повноцінної пошукової системи —- такої, що здатна самостійно рекурсивно обходити посилання, індексувати отримані документи та надавати можливість пошуку в отриманій базі. Її наступники, такі як *Yahoo*, *Google* та *AltaVista* — всесвітньо відомі й сьогодні.

Таким чином, на початку 2000-них склалася ситуація, неймовірно сприятлива для виникнення нових природно-мовних пошукових систем (ПМПС): мережа Інтернет — величезна за обсягом динамічна база природно-мовної інформації та вибухове зростання можливостей комп'ютерів, від дискового простору до швидкості зв'язку. Одночасно виникла й проблема категоризації, порівняння та оцінки ПМПС. З одного боку, майже усі наявні системи добре виконували прості задачі, такі як пошук за ключовим словом. З іншого боку, для роботи за складнішими сценаріями, як-то використання ПМПС у якості складових частин інших програмних засобів, вони виявилися пристосованими погано. Звісно, окремі системи добре виконували поставлені задачі, але про їх порівняння чи взаємодію годі було й говорити. На жаль, це твердження актуальне і сьогодні.

Аналіз проблеми. Очевидно, що отримання адекватної оцінки роботи ПМПС можливе лише за умови наявності об'єктивної методики — послідовності чітко визначених кроків, в результаті виконання яких можна отримати формальне представлення якості пошуку. Спробуємо визначити, яким чином наразі виконується оцінка та порівняння ПМПС і чи є об'єктивні причини, які стояли б на заваді появі загальної чіткої методики що їх об'єднує.

Навіть при поверхневому пошуку можемо побачити досить багато різноманітних критеріїв оцінки роботи ПМПС [3,4], найбільш популярними з яких є повнота й релевантність. Але їх опис в більшості випадків обмежується загальною словесною характеристикою, без формального їх визначення та обґрунтування. Пояснення, чому цей процес не зайшов далі, можемо отримати, якщо трохи ширше поглянемо на наявні ПМПС та спробуємо їх класифікувати.

Досить чітко виділяються основні групи. Перша це "спеціалізовані" проекти, створені для перевірки наукових гіпотез або виконання специфічних задач — сюди входять і такі потужні системи як *WordNet*, і експертні системи, і деякі спеціалізовані ПС (такі як системи пошуку наукових статей), і багато інших. Такі проекти (навіть великі за масштабом та аудиторією) не націлені на конкуренцію з ПС загального призначення, кількість їх користувачів порівняно невелика, а критерії оцінки ефективності їх роботи можуть бути дуже специфічним. Системи "загального призначення", натомість, характеризуються великою кількістю користувачів і, що важливо, активною конкуренцією між собою (наприклад, *Google* та *Yahoo*). У ПС загального призначення якість пошуку є основним ресурсом, оскільки від неї залежить популярність ПС, а від популярності, в свою чергу — кількість користувачів. Таким чином, будь-які можливості для покращення, в тому числі й ефективні системи порівняння з конкурентами, розглядаються як закрита інформація й не призначені для розповсюдження.

На основі цієї класифікації досить чітко виділяються й відповідні групи користувачів. Основ-

на аудиторія спеціалізованих ПМПС це науковці або спеціалісти, які, як правило, визначають критерії якості пошуку власноруч згідно вимог конкретного проекту. Користувачі ПМПС загального призначення, з іншого боку — це не-спеціалісти, які, в залежності від складності задачі, або не потребують детальної оцінки якості роботи ПМПС, або використовують закриті алгоритми оцінки. Отже, хоча загальна методика оцінки була б корисна для усіх користувачів (і як спільний базис порівняння, і як основа для створення власної методики, і як допомога при виборі ПМПС згідно відомих вимог), у більшості випадків користувачі просто не готові витрачати власні ресурси задля її створення.

Постановка задачі. Для отримання об'єктивної оцінки якості пошуку необхідно виконати два основні кроки. В першу чергу потрібно формально, в максимально точних термінах, визначити основні критерії якості пошуку і представити їх у вигляді математичної моделі. Після цього, на основі обраних критеріїв, необхідно підібрати еталонні запити, створити відповідні вибірки набори результатів пошуку та порівняти їх з фактичними результатами.

Як було сказано вище, окремі текстові описи критеріїв якості пошуку уже існують і використовуються. Тобто, мова йде про формалізацію й уніфікацію усього процесу оцінки. При цьому створювана методика повинна бути, з одного боку, досить загальною (однаково підходити для усіх систем), а з іншого — універсальною (дозволяти вносити суттєві модифікації у процес оцінки). Відповідно, представимо лише основні кроки і їх загальні характеристики та особливості.

Визначення критеріїв. Будь-який метод оцінки ПМПС вимагає, в першу чергу, чіткого опису критеріїв якості пошуку, оскільки саме на них ґрунтуються усі подальші кроки. Навіть якщо ці критерії точно не відомі, їх завжди можна виділити з загальних вимог до ПМПС. Як приклад таких характеристик можемо взяти повноту (усі документи з бази, що відповідають запиту, повинні бути представлені у результатах) і релевантність (у результатах пошуку не повинно бути документів, що не відповідають запиту). Вже на такому рівні деталізації їх можливо представити у вигляді залежностей, функцій тощо.

Зауважимо, що критерії оцінки можуть значно відрізнятися в залежності від поставленої задачі. Це можуть бути і як суто штучні вузькоспеціальні вимоги (скажімо, пошук лише за одним із декількох можливих значень слова), так і розширення набору загальних критеріїв (наприклад, врахування швидкості обробки запиту). Більш того, навіть базові критерії можуть бути доповнені або модифіковані — наприклад, при аналізі релевантності може враховуватися не лише релевантність, а й розподіл результатів за релевантністю в межах вибірки.

Математична модель. Отримані на попередньому кроці критерії тепер необхідно формалізувати. Оскільки методика розрахована на роботу з довільними ПМПС, при побудові моделі можемо спиратися лише на те, що залишається незмінним для усіх методів — на структуру самого процесу пошуку. Опишемо його як процес вибірки з загальної бази обмеженої кількості документів, що відповідають запиту, у набір результатів. На основі цього представимо якість пошуку за певним критерієм x як функцію

$$Qx = f(s, e, q, r) \tag{1}$$

, де

Qx — функція довільного вигляду, що показує значення якості за даним показником x; s — база документів, по якій проводиться пошук; q — набір еталонних запитів; e та r — еталонні результати пошуку та реальні результати пошуку за запитами q відповідно.

Слід зазначити, що документи у базі, еталонній вибірці та реальній виборці можуть мати різний вигляд для різних методів оцінки, але в рамках одного метода формат документів має бути однаковим. Наприклад, для пошуку за релевантністю достатньо текстових об'єктів; якщо у пошуку необхідно враховувати час створення документів, кожний документ повинен містити відповідне поле. Отже, ще при підборі критеріїв необхідно враховувати, чи в змозі обрані ПМПС надати усі необхідні дані.

Крім того, під час пошуку саме джерело даних повинно залишатись незмінним. Тобто, якщо

база змінюється в реальному часі (наприклад, якщо воно залежить від мережі Інтернет), то для точної оцінки необхідно або зберігати певний знімок його стану, або додавати до запиту допоміжні поля, які забезпечать такий самий стан бази (наприклад, не враховувати документи, створені після певної відмітки часу).

Побудова вибірок. Очевидно, що еталонні запити та вибірки повинні бути науково обґрунтованими — однаково оброблятись системою при повторному використанні, максимально охоплювати діапазон можливих значень критеріїв, забезпечувати статистично значимі результати і т.д.. Але, крім того, складність ПМ як об'єкту дослідження додає певні специфічні вимоги: дані повинні бути однозначними (містити настільки істинну й чітку інформацію, наскільки це можливо); природними (а не штучно синтезованими текстами подібними до "*lorem ipsum*"); якісними (містити мінімум помилок на всіх рівнях від граматики до семантики);

Досить якісною еталонною базою видається спеціалізована та навчальна література (книги, підручники, посібники тощо). На відміну від художньої літератури ці джерела, як правило, прості для ручного аналізу і містять мінімум художніх засобів; на відміну від синтетично створених текстів, вони мають цілісну семантику; на відміну від нерецензованих джерел (як-то записи користувачів на власних сторінках або у системах обміну повідомленнями), вони граматично правильні й структурно повні.

Висновки. Зазначимо ще кілька загальних факторів, які можуть внести похибки у процес оцінки якості пошуку на усіх його етапах.

По-перше, в більшості випадків необхідно враховувати розподіл можливостей ПС між власне системою пошуку та її оболонкою (*middleware*). У багатьох ПС до початку власне пошуку відбувається підготовка запиту — обробка синонімів, корекція граматичних помилок тощо. Запити необхідно формувати таким чином, щоб ця обробка була мінімальною — адже результатом методу повинна стати оцінка якості саме пошуку, а не, скажімо, алгоритму перевірки правопису або якості словнику синонімів.

По-друге, досить часто можливо модифікувати запит для отримання кращих результатів пошуку у тій чи іншій ПС, зазвичай шляхом надання запиту певної структури і, тим самим, віддаленням його від ПМ. Отже, для збереження якості пошуку запити потрібно створювати максимально близько до ПМ.

По-третє, в більшості випадків не є доцільним оцінка усіх можливих комбінацій значень критеріїв. Оскільки критерії якості пошуку, як правило, досить загальні, вони оцінюють різні частини функціональності ПС, і, відповідно, мало залежать один від одного. Отже, їх комбінування лише збільшує вимоги метода до ресурсів, не покращуючи його якості.

Зауважимо, що запропонована методика не є "срібною кулею", оскільки для створення конкретного методу під конкретні вимоги все одно необхідно фактично заново виконати усі описані кроки. Значною частиною роботи залишається створення еталонних наборів даних (запитів і результатів пошуку), що надихає на думку що є сенс розглянути автоматичне формування таких вибірок. Втім, сама наявність такої методики дозволяє уникнути деяких поширених помилок в процесі оцінки різнопланових ПС і взагалі спростити цей процес. Автор сподівається що може стати в нагоді — як спеціалістам, так і практикам — для вибору ПМПС для виконання відомих задач або, навпаки, визначення сильних і слабких сторін даної ПМПС.

Література. **1.** Brachman R. An Overview of the KL-ONE Knowledge Representation System / R. Brachman, J. Schmolze. // Cognitive Science. – 1985. – №9. – С. 171–216. **2.** The History of Search Engines - An Infographic [Електронний ресурс] – http://www.wordstream.com/articles/internet-search-engines-history **3.** Thelwall M. Quantitative comparisons of search engine results / M. Thelwall. // Journal of the American Society for Information Science and Technology. – 2008. – №59. – С. 1702–1710. **4.** Кураленок И. Е. Оценка систем текстового поиска : дис. канд. техн. наук : 05.13.01 / Кураленок Игорь Евгеньевич – Санкт-Петербург, 2004. – 105 с.

Сергєєв А.В.
Київський національний університет імені Тараса Шевченка, Київ, Україна

Роль децентралізованих мереж у побудові некомерційних систем доставки контенту

У випадку значної географічної віддаленості від сервера, у клієнта глобальної географічноо розосередженої мережі можуть проявлятися проблеми, пов'язані із затримкою доступу до потрібної йому інформації.

Для їх подолання у комерційних мережах прийнято використовувати системи доставки контенту (CDN - Content Distribution Network) - структуру, що оптимізує доставку контенту до кінцевих користувачів. Робота CDN організована на основі взаємодії між локальними серверами, що зберігають копії даних основного серверу та намагаються підримтувати їхню актуальну версію. Для забезпечення відмовостійкості та комфортної роботи такої системи необхідно підтримувати і постійно оновлювати велику кількість локальних серверів, що потребує значних фінансових ресурсів. Некомерційним глобальним мережам потрібно шукати підхід при якому навантаження на систему було розподілено рівномірно між усіма учасниками мережі, щоб уникнути перенавантаження та виходу з ладу цілих сегментів мережі.

У побудові некомерційних систем доставки контенту доцільно використовувати децентралізовані (однорангові - p2p) мережі, в яких кожен елемент одночасно є клієнтом та сервером. У порівнянні з клієнт-серверною архітектурою, однією з основних переваг однорангових мереж є надійність - при виході з ладу підмножини елементів, система збереже хоча б часткову працездатність, в той час як при відмові центрального серверу класичної централізованої системи з ладу виходить вся мережа.

Децентралізована мережа потребує ефективних алгоритмів пошуку та розподілу інформації між учасниками. Для реалізації таких алгоритмів прийнято використовувати розподілені хеш-таблиці (DHT - Distributed Hash Table) [1] - оверлейну інфраструктуру, що допомагає ефективно знаходити значення за певним ключем, тобто працює за принципом асоціативного масиву. У DHT кожному елементу мережі призначається ідентифікатор (ID), а ключем файлу є результат хеш-функції де параметром є його назва, причому ID елементів та ключі файлів належать до одного ключового простору. Розподіл посилань на файли у DHT відбувається на основі розрахунку певної метрики відстані, яка зазвичай індивідуальна для кожної системи. Посилання на файл буде зберігатися у того вузла, у якого відстань між ID до ключа файлу є найменшою. Серед сервісів побудованих на основі DHT найпопулярнішим є BitTorrent. До цього часу залишається відкритим питання вибору найоптимальнішої реалізації розподіленої хеш-таблиці для побудови децентралізованих мереж.

Незважаючи на значну кількість реалізацій однорангових мереж на основі DHT [2], питання побудови CDN на базі децентралізованої інфраструктури залишається відкритим. Наразі існує лише декілька спроб реалізації такої мережі [3, 4]. Серед основних недоліків даних систем можна виділити недостатню увагу приділену їх безпеці.

Реалізація CDN на основі децентралізованої мережі з ефективними алгоритмами пошуку та розподілу інформації допоможе у створенні некомерційних географічно розосереджених мереж, використання яких буде комфортним для кінцевих користувачів.

Література. **1.** M. Harren, J.M. Hellerstein, R. Huebsch, B.T. Loo, S. Shenker, I. Stoica, Complex queries in dhtbased peer-to-peer networks, in: Proceedings of the 1st International Workshop on Peer-to-peer Systems, March 2002. **2.** A. Rowstron, and P. Druschel, "Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems." In Proceedings of the 18th IFIP/ACM Int'l Conf. on Distributed Systems Platforms (Nov. 2001). **3.** Manal El Dick, Esther Pacitti, Bettina Kemme, FlowerCDN: a hybrid P2P overlay for efficient query processing in CDN, ACM EDBT 2009, pp. 427-438. **4.** M. J. Freedman, E. Freudenthal, and D. Mazieres. Democratizing content ' publication with Coral. In NSDI, Mar. 2004.

Скукіс О.Є.
Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна

Підхід до побудови програмних систем для розв'язання задач транспортної логістики

Відомо, що 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].

Для автоматизації розв'язання типових задач транспортної логістики запропоновано структуру та склад інформаційної системи (ІС), розроблено програмне забезпечення її підсистем.

Структура ІС та її функціональні можливості реалізовано наступними підсистемами.

Формування замовлення на виконання перевезень. Функціонування підсистеми забезпечує виконання наступних процесів: підготовка заявки Замовника; вибір Виконавця та підготовка заявки Виконавцю; формування Замовлення; формування оптимального маршруту; розрахунок вартості перевезень.

Контроль виконання замовлень. Підсистема надає можливість відслідковувати етапи виконання замовлень Виконавцями, проводить моніторинг використання транспортних засобів та розхідних матеріалів.

Звітно-аналітична підсистема. Забезпечує формування супровідних документів процесу перевезень, формування звітів планово-економічного характеру, розрахунок рентабельності використання вибраних транспортних засобів.

Підсистема класифікаторів та довідників. Забезпечує уніфікацію оброблювальної інформації. Підсистема дозволяє систематизувати бази даних Замовників, Виконавців, будувати різноманітні вибірки звітно-аналітичною підсистемою, будувати діаграми та графіки динаміки використання ресурсів Виконавця.

Адміністрування ресурсів ІС. Підсистема забезпечує ведення довідника користувачів системи, контролює розподіл прав доступу до елементів програмної системи відповідно до визначених ролей, веде хронологію та історію виконаних операцій користувачем.

Розробка ІС проводилася із застосуванням об'єктно-орієнтованої технології, яка надає можливість використовувати її властивості спадкоємства та поліморфізму. Застосування поняття онтології дозволило абстрагувати опис класів об'єктів, зокрема опис їх властивостей, методів обробки і аналізу подій, дозволило розробити структуроване програмне забезпечення як в частині предметного наповнення, так і в частині побудови інтерфейсу користувача [6]. Інструментом реалізації ІС були засоби опису типів даних (*.tlb), механізми підвантажуваних бібліотек (*.dll), мова програмування C++ [7].

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.** Gruber T.R. A translation approach to portable ontologies. // Knowledge Acquistion. 1993. N 5 (2). **7.** Гради Буч. Объектно-ориентированный анализ и проектирование с примерами приложений на C++. – Бином, 1998. – 560 с.

Слободянюк Б.Р., Павловський В.І.
Національний технічний університет України "КПІ", ФПМ, Київ, Україна

Засоби інтеграції успадкованих систем з новітніми системами предачі даних

Постановка задачі. Сучасний стан розвитку суспільства характеризується безперервно зростаючими вимогами до телекомунікаційних систем і сервісів, що призводить до їх постійних змін і динамічного розвитку. Вони еволюціонують від простих, автономних до складних, інтегрованих систем, заснованих на вимогах взаємозв'язку компонентів та технологіях, що постійно змінюються і ускладнюються.

З іншого боку за останні 15 років було створено та широко впроваджено багато дорогих і "важких"– монолітних систем, що використовуються у банківських структурах, на кшталт JPMorgan Chase та телекомунікаційних гігантах типу Blackberry. Практично будь-яка з цих систем після створення і впровадження швидко застаріває відносно поточного стану IT-технологій, "опирається"змінам і має тенденцію поступового перетворення в негативний баласт компанії, але проста заміна таких систем заздалегідь неприйнятна в зв'язку з їх вартістю та масштабами використання.

Таким чином успадковані системи (Legacy Systems), що побудовані на застаріваючих технологіях, архітектурах і платформах, а також їх програмне і інформаційне забезпечення, при проектуванні яких не були передбачені заходи для їх поступового переростання в системи з новою якістю, потребують перебудови (Legacy Transformation) в відповідності з новими вимогами бізнес-процесів і технологій. В процесі такої трансформації необхідно, щоб нові модулі системи і решта компонентів успадкованих систем зберігали існуючу функціональність та мали змогу до взаємодії з новими система передачі даних (СПД).

З метою суттєвого подовження терміну експлуатації успадкованих систем актуальною є задача створення засобів, які б дозволяли модернізувати такі системи, не зачіпаючи їх основу, з тим, щоб в умовах швидкого технологічного розвитку забезпечувати необхідну тривалість їх життєвого циклу, включаючи постійні покращення властивостей і обов'язкове підтримання існуючої функціональності [1].

Аналіз існуючих рішень. На сьогодні відомі інструментальні засоби, що дозволяють провести інтеграцію успадкованих СПД з сучасними глобальними телекомунікаційними системами. Ці засоби побудовані в рамках архітектурного шару ESB (Enterprise Service Bus), що реалізується, як правило, на основі технологій, що надають базові сервіси на основі поширених стандартів на підставі подій або керованих повідомлень. У реальності системи ESB надають деякий рівень абстракції над технологіями і реалізаціями передачі повідомлень, маршрутизації, зберігання, оркестровки і т.п. Часто це дозволяє архітекторам і розробникам будувати досить складні інтеграційні структури не вдаючись до низькорівневого програмування і обходячись тільки конфігуруванням. Конфігурації можуть задаватися у вигляді окремих XML-файлів або за допомогою спеціальних графічних інструментів [2].

Як приклад можна навести ESB-систему Oracle Enterprise Service Bus - основний компонент сервісно-орієнтованої архітектури Oracle, яка забезпечує слабозв'язаний фреймворк взаємного обміну повідомлення. Служба ESB розробляється і конфігурується за допомогою середовища Oracle JDeveloper та інтерфейсу Oracle ESB Control. Після цього вона реєструється на ESB Server. ESB Server підтримує багатозв'язність (Multibinding) протоколів обміну повідомленнями, включаючи протоколи HTTP/SOAP, JMS, JCA, WSIF та моделі взаємодії синхронний/асинхронний, запит/відповідь або публікація/підписка [3].

Але використання існуючих засобів побудови ESB-систем вимагає високої кваліфікації виконавців, багато часу та значного обсягу ручної праці, що призводить до великої кількості помилок, що вимагають подальшого виправлення.

Виходячи з важливості, масштабності та складності задач трансформації успадкованих систем актуальною є задача створення засобів ESB-систем які б забезпечували простоту

розширення існуючих ESB-систем і мали простий рівень взаємодії з користувачем, надавали можливості по включенню додаткової бізнес-логіки і рознесенню навантаження між вузлами СПД та дозволяли вбудовувати власну обробку виключень, що необхідно для стабільної роботи систем, що працюють з даною СПД. Також необхідно враховувати ситуації, коли окремі вузли СПД можуть виходити з ладу. Для швидкого відновленні повноцінної працездатності СПД необхідно мати зручний сервіс, який зберігає та відновлює усі необхідні налаштування даної системи, щоб в разі виходу з ладу якогось вузла, ці налаштування можна було перенести на інший вузол, тим самим скоротивши час, коли усі вузли не можуть взаємодіяти спільно.

Запропоновані засоби. На основі визначених вимог пропонуються засоби інтеграції різнорідних СПД в глобальну телекомунікаційну систему, які, на відміну від існуючих, дозволяють забезпечити функціональну розширюваність системи за рахунок можливості впровадження необхідної функціональності в існуючу систему. До особливостей запропонованих засобів слід віднести їх спрямованість на роботу з успадкованими системами в умовах постійної динаміки змін IT-технологій.

В основу реалізації запропонованих засобів покладена сервісно-орієнтована архітектура SOA побудови програмних систем, яка дозволяє організувати взаємозв'язок між СПД за допомогою сервісів. Ці сервіси об'єднуються в певний контейнер, який є проміжним вузлом між ESB та успадкованими системами.

Контейнером для даних засобів виступає Application Server Apache Tomcat EE. Даний контейнер обраний згідно вимог до реалізації стандартів та технологій SOA, таких як JAX-WS (Web Services) або JAX-RS (RESTfull Services). Також він дозволяє реалізувати додаткову функціональність з допомогою технологій Java EE.

До запропонованих засобів інтеграції різнорідних СПД в глобальну телекомунікаційну систему входять:

- Сервіс генерації звітів з роботи системи та передачі інформації по каналам між вузлами системи;
- Сервіс планового збереження (з певним інтервалом в часі) інформації про налаштування системи та основних модулів;
- Сервіси роботи з успадкованими БД - читання і запис сутностей;
- Сервіси роботи з успадкованими СПД - відправка і отримання повідомлень;
- Обробник виключень з можливістю додавання методів обробки специфічних виключень, що необхідно для стабільної роботи систем, що працюють з даною СПД;
- Інтерфейс користувача, через який відбувається конфігурація усіх засобів, підключення систем, налаштування зв'язків успадкованої системи з ESB системою.

Усі основні класи і методи запропонованих засобів є абстрактними, що дозволяє розробникам, в разі необхідності, вносити свої зміни відносно тих особливостей, які представлені в конкретній системі.

Запропоновані засоби інтеграції різнорідних СПД мають такі особливості:

1. Телекомунікаційна система як єдина точка доступу, що підтримує мобільність та заміну вузлів СПД;
2. Можливість додавання функціональності, необхідної користувачам системи, зокрема, обробка специфічних виключень;
3. Трансформація повідомлень системи-постачальника в повідомлення, зрозумілі системі-споживачеві і навпаки.

Література. **1.** Технология открытых систем. Под ред. А. Я. Олейникова.,Учебное пособие / М., ИРЭ РАН. 1999 - 97 с. - [Електронний ресурс] - Режим доступ:http://ict.informika.ru/ft/005023/posobie.pdf. **2.** Архитектурный принцип Шины в ESB [Електронний ресурс] - Режим доступу: http://goo.gl/J1JY8x. **3.** Oracle Enterprise Service Bus [Електронний ресурс] - Режим доступу: https://ru.wikipedia.org/wiki/Oracle_Enterprise_Service_Bus.

Сутула О.В.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Порівняння методів розробки для мобільної платформи Android

Зі збільшенням числа мобільних додатків і їхнього стрімкого розвитку, постало питання про використання сучасних мов та засобів розробки програмного забезпечення для мобільних платформ. [1] Останнім часом існує три основних напрямки в розробці Android додатків, а саме: класична розробка за допомогою мови Java, використання мов, що запускаються на Java Virtual Machine (JVM), а також фреймворки для розробки мобільних додатків за допомогою Javascript, HTML, CSS.

- класична розробка за допомогою мови Java,
- мови, що використовують Java Virtual Machine (JVM) як оточення для запуску,
- фреймворки для розробки мобільних додатків за допомогою Javascript, HTML, CSS.

За умови проведення розробки на мові Java з використанням XML розмітки досягається найбільша продуктивність мобільного додатку, але при цьому збільшується час розробки. [2]

Android, для вирішення проблеми з обмеженнями мобільних платформ, використовує Dalvik virtual machine, що може виконувати JVM подібний код після трансформації в Dalvik код. Це дає можливість використовувати мови написані на основі JVM.[3]

В останні роки з'явилось декілька нових мов які використовують Java Virtual Machine як оточення для запуску, такі як Scala, Clojure, GO, Kotlin. Ці мови підтримують нові парадигми програмування, що зменшують надлишковість коду, додають нові абстракції та можливості. Через підключення додаткових бібліотек, що необхідні для підтримки нових мов, збільшується розмір додатку. А під час роботи додатків витрачається більше заряд. Але використання оперативної пам'яті зменшується, адже функціональні мови, а саме Scala, Clojure, GO, Kotlin, використовують менше пам'яті для своїх об'єктів і зберігають їх менше часу, що забезпечує кращу роботу Garbage collector.

Фреймворки, наприклад PhoneGap, Apache Cordova, дозволяють створювати додатки для всіх мобільних платформ (iOS, Android, і т.д.) за допомогою Javascript, HTML5, CSS3. Цей спосіб має найбільший рівень абстракції, відповідно й найнижчу продуктивність, найбільшу витрату оперативної пам'яті і найгірші показники енергоспоживання, але вирішує проблему дублювання опису однотипного фунціоналу для різних платформ. Також, у зв'язку з широким спектром підтримуваних платформ, деякі низькорівневі функції платформи можуть бути недоступні.[4]

Висновки:

Враховуючи особливості різних підходів, при розробці мобільного програмного забезпечення, вибір технології слід засновувати на таких факторах як: складність проекту, час на розробку, оптимальний об'єм споживаних ресурсів, вимоги до продуктивності роботи ПЗ.

Доцільно використовувати Java для задач пов'язаних з швидкодією додатку та мінімізацією використання ресурсів. В проектах з часовими обмеженнями альтернативою є мови, що породжують абстракції. Якщо проект не має суттєвих вимог по швидкодії, не використовує низькорівневих функцій платформи і аналогічний додаток потрібен на всі платформи, доцільно застосувати фреймворки типу PhoneGap та Apache Cordova.

Література. **1.** V. Pankratius, F. Schmidt, and G. Garreton. Combin- ing functional and imperative programming for multi- core software: An empirical study evaluating scala and java. In Software Engineering (ICSE), 2012 34th Inter- national Conference on, pages 123 –133, june 2012. **2.** A. Sewe, M. Mezini, A. Sarimbekov, D. Ansaloni, W. Binder, N. Ricci, and S. Z. Guyer. new scala() in- stanceof java: A comparison of the memory behaviour of java and scala programs. In Proceedings of the International Symposium on Memory Management, ISMM '12, pages 97–108, New York, NY, USA, 2012. ACM. **3.** PhoneGap - Overview http://www.tutorialspoint.com/phonegap/phonegap_overview.htm. **4.** Four Ways To Build A Mobile Application, Part 3: PhoneGap https://www.smashingmagazine.com/2014/02/four-ways-to-build-a-mobile-app-part3-phonegap/.

Сухомлин І.Б.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна; Національний технічний університет України "Київський політехнічний інститут", Київ, Україна

Технологія тестування мобільних додатків для контролю здоров'я, що підключаються до зовнішніх приладів

Введення. На відміну від комп'ютерів, мобільні телефони можуть вважатися найбільш «персональними» приладами, використовуваними щодня: рідко мобільний телефон знаходиться далі 1 м від користувача, зате набагато частіше - в руці або кишені.Технічні можливості сучасних сматрфонів і їх близькість до людини дозволяють отримувати інформацію про стан здоров'я людини в реальному часі. За допомогою вбудованих сенсорів, таких як GPS-приймач, Bluetooth-канал, акселератор, доступ до календаря, контактів та іншої персональної інформації, мобільні додатки можуть контролювати, де перебуває користувач і що він робить а також підключатись до зовнішніх пристроїв (тонометр, апарат ЕКГ і електронні ваги) і відображати отримані виміри. Ця інформація дозволяє в режимі реального часу забезпечувати підтримку пацієнта, гарантуючи допомогу в той момент, коли вона необхідна.

Тестування передачі і обробки даних. Розглянемо електронні ваги, що підключаються до смартфона користувача про каналу Bluethooth.

Після підключення зовнішнього приладу в процесі передачі даних з'являється ще одна ланка, по якій вимірювання передаються з приладу на смартфон через Bloethooth.

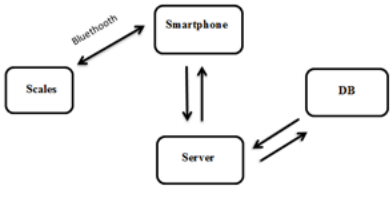


Рис. 1. Схема передачі даних

Отже, можна виділити такі обов'язкові етапи тестування передачі і обробки вимірювань користувача:

1. Передача вимірювань з підключеного пристрою на смартфон
2. Передача вимірювань в JSON форматі з телефона на сервер для обчислення додаткових показників
3. Формування SQL-запиту до бази даних для додання нових даних в журнал вимірювань, для оновлення даних.
4. Передача вибірки даних з бази даних на сервер.
5. Обробка даних на сервері і відправка на мобільний пристрій.

Висновки. Бездротове підключення мобільного телефону до приладів, що фіксують різні фізіологічні показники, дозволяє зберігати біометричні вимірювання в зручній формі, переправляти їх в будь-яке місце (лікар, клініка, аптека) для роботи з ними. Запропоновані етапи тестування такого роду додатків охоплюють всі ланки передачі і обробки даних, що забезпечує коректну роботу мобільного додатку.

Література.

Література. **1.** УДК 007.3 Абдуллаев В. Г., Чуба И. В., Аскеров Т. К. Мобильные приложения для здоровья **2.** Інтернет ресурс - Хабрахабр https://habrahabr.ru/post/269135/ **3.** УДК 621.397 С.П. Ботуз. Модели трафика мобильных приложений **4.** Інтернет ресурс- Хабрахабр https://habrahabr.ru/post/246877/

***Сухоручкина О.Н.*[1], *Прогонный Н.В.*[1], *Воронов М.А.*[2]**
[1] *Международный научно-учебный центр информационных технологий и систем НАН Украины и МОН Украины, Киев, Украина;* [2] *Национальный технический университет «КПИ», Киев, Украина*

Технология использования измерений сенсоров-дальномеров в задачах управления автономным мобильным роботом

В задачах навигации автономных мобильных роботов (МР) важную роль играет формирование в рабочей памяти системы управления (СУ) робота адекватного представления о расположении и форме окружающих объектов. Известны различные технические средства получения данных о внешнем мире МР и различные программные решения интерпретации этих данных для формирования 2D или 3D моделей этого мира в памяти робота. Использование на борту робота современных специализированных устройств локации и автоматического построения трехмерных моделей внешнего мира, таких как лазерные сканеры, система MS Kinect и другие комплексные аппаратно-программные ToF-системы (Time of Flight Systems), существенно повышает его стоимость и оправдано при разработке роботов, предназначенных выполнять задания особой сложности. Для интенсивно развивающегося рынка сервисных роботов, ориентированных на непрофессионального пользователя, предпочтительны более простые и дешевые источники локационных данных, такие как сенсоры-дальномеры, работающие в инфракрасном или ультразвуковом диапазонах. В данной работе источником измерительной информации являются сенсоры-дальномеры Sharp GP2Y0A02YK [1].

Организация сенсорного восприятия внешнего мира как информационного процесса, отвечающего за получение текущих измерений дальномеров, их анализ, интерпретацию, представление в определенных структурах памяти и методы использования, зависит от избранного принципа реализации СУ автономного робота и возложенных на нее задач. Как правило, измерения сенсоров-дальномеров используются в задачах:

– построения и текущей коррекции карты окружающего робот пространства;

– реактивного управления для избежания столкновений МР с препятствиями;

– определения целевого положения исполнительного органа робота для успешного взятия различных объектов.

Реализованный нами принцип интеллектуального управления МР порождает еще одну задачу — наблюдение за соответствием фактических измерений сенсоров-дальномеров ожидаемым согласно процессам 3D моделирования текущих состояний робота в контексте исполняемого им задания. Наш МР способен автономно выполнять в помещениях различные задания пользователя, связанные с исследованием объектов окружающего пространства, движением робота в нестационарной среде, поиском и перемещением определенных объектов.

Характерная особенность данного подхода состоит в формировании на верхнем уровне управления в режиме реального времени информационного эквивалента результата текущего действия робота, как объекта управления, в виде ожидаемых параметров его состояния. Одними из наблюдаемых параметров состояния МР являются текущие измерения всех установленных на нем сенсоров-дальномеров. Специальный программный механизм компарирования способен фиксировать несоответствие фактического состояния ожидаемому. Это позволяет активирующей подсистеме СУ своевременно принимать решения об актуальных для текущей ситуации действиях робота [2]. Например, расхождение значения усредненных по некоторой выборке фактических дальнометрических измерений с ожидаемыми по модельным оценкам, свидетельствует о потере адекватности текущей модели взаимного положения робота и окружающих его объектов. Управление движением МР по спланированной ранее траектории теряет актуальность. Первостепенной становится задача локальной коррекции модели окружающего робот мира.

Информационные процессы, связанные с дальнометрическими измерениями, потребовали формирования различных структур представления этой информации в рабочей и долговременной памяти СУ робота и моделей его пространственного восприятия, а также программной

реализации различных методов обработки и анализа сенсорной информации. Геометрические характеристики зон чувствительности сенсоров-дальномеров, модели интерпретации их измерений и причины, порождающие неустранимые неопределенности в оценке точности этих измерений, показаны в [1].

Формирование пространственного восприятия роботом окружающих объектов состоит из ряда этапов:

– фильтрация текущих измерений сенсора-дальномера;

– преобразование измерений сенсора-дальномера в пространственные координаты точек поверхностей объектов;

– разбиение множества точек по принадлежности различным объектам; – построение контуров или элементов контуров объектов; – минимизация данных для представления моделей объектов в долговременной памяти МР.

При этом помимо вычислительной эффективности, необходимой для систем реального времени, используемые алгоритмы должны быть одинаково применимы при неопределенном характере входных данных, поскольку измерения сенсоров-дальномеров происходят при движении МР в среде с произвольными физическими характеристиками окружающих объектов. Форма, размеры и число объектов заранее неизвестны. Также неизвестны цветовые характеристики поверхностей объектов, что сказывается на степени неопределенности погрешности сенсорных измерений.

Программное обеспечение СУ МР состоит из нескольких модулей, обеспечивающих использование дальнометрической информации и обслуживаемых различными вычислительными устройствами. За нижний уровень работы с сигналами сенсоров Sharp GP2Y0A02YK отвечает специализированный узел сенсорной сети [1]. Ресурсами микропроцессора узла выполняется преобразование сигнала с частотой 40 мс из измеряемого выходного напряжения в оценку дальности до объекта относительно корпуса сенсора и первичная фильтрация. Модуль СУ бортового компьютера МР отвечает за формирование выборок сенсорных измерений с предобработкой их согласно контексту исполняемого роботом задания. Пакеты сенсорных измерений в структурах данных, заказанных верхним уровнем управления, передаются по Wi-Fi сети в компьютер супервизора, где используются во всех упомянутых выше задачах управления автономными действиями робота.

В репертуаре СУ нашего МР присутствует ряд заданий, алгоритмы исполнения которых опираются на интерпретацию текущей информации сенсоров-дальномеров. Успешное автономное выполнение роботом заданий «Исследуй объект», «Уточни обстановку», «Возьми объект» и «Подойди к объекту» в различных внешних условиях, в том числе среди движущихся случайных образом объектов-препятствий и целевых объектов, свидетельствует об удачной практической реализации всех составляющих информационного процесса, связанного с дальнометрическими измерениями.

Литература. **1.** Сухоручкина О.Н., Прогонный Н.В., Ланбин В.С., Людовик Т.В. Об обработке дальнометрической сенсорной информации // Управляющие системы и машины. — 2006. — № 5. — С. 77–84. **2.** 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.

Томашевський В.М., Дифучин А.Ю.
Факультет інформатики та обчислювальної техніки НТУУ "КПІ", Київ, Україна

Візуальне середовище моделювання дискретно-подійних систем

Програмне забезпечення з імітаційного моделювання систем охоплює широке різноманіття підходів: імітація з використанням колекцій блоків, налаштованих на виконання тих чи інших функцій – очікування, обробка, комплектація, транспортування та інші (Arena, ExtendSim, PlantSim) [1]; математичне моделювання з використанням динамічних блоків (VisSim, Simulink) [2]; гібридне моделювання з використанням блоків обох попередніх підходів (AnyLogic) [3]; відкриті бібліотеки програм для моделювання дискретно-подійних систем (DESMO-J) [4]; імітація на основі опису процесу функціонування системи у вигляді стохастичної мережі Петрі (CPN Tools) [5]. Проте пошуки нових ефективних підходів до моделювання складних систем з великою кількістю взаємопов'язаних елементів та різнорідністю процесів функціонування елементів і підсистем постійно продовжуються.

У роботі [6] викладені теоретичні засади технології Петрі-об'єктного моделювання, основними перевагами якої є уніфікований опис динаміки як елементів управління, так і об'єктів управління, можливість створювати класи типових елементів зі схожою динамікою, конструювання динаміки системи на основі динаміки її елементів, наявність математичного опису динаміки моделі та поліноміальна складність обчислювального алгоритму. Динаміка Петрі-об'єктної моделі описується стохастичною мережею Петрі, математичні рівняння якої представлені в роботі [7].

Метою даного дослідження є створення веб-орієнтованого програмного забезпечення з компонентами візуальної розробки динаміки Петрі-об'єктної моделі, які збільшують швидкість конструювання моделі та її відлагодження.

Процес розробки моделі відбувається у такий спосіб. Виділяються елементи моделі та розробляються структурні зв'язки між ними. Проектування динаміки елементів моделі здійснюється на основі стохастичної мережі Петрі з багатоканальними та конфліктними переходами. На першому кроці процес функціонування об'єкта поділяється на події та визначаються умови, при яких вони відбуваються. Події представляються переходами, а умови – позиціями мережі Петрі. Маркери у позиціях характеризують виконання умови. Наявність кількох маркерів у позиціях означає багатократне виконання умови. Для кожного переходу задається часова затримка, з якою виконується подія. Часова затримка може бути задана детермінованою чи стохастичною величиною. Для миттєвої події задається детерміноване нульове значення. Інформаційні зв'язки надають можливість перевірити наявність умови без змінювання цієї умови при виконанні події, що розширює можливості мережі Петрі для представлення систем з елементами управління. Зауважимо, що при розробці динаміки об'єкта відтворюється функціонування об'єкта з урахуванням інформації, яка надходить від інших об'єктів.

Технологія розробки динаміки об'єкта передбачає такі кроки:

1. Створення мережі Петрі з використанням модуля візуальної розробки та збереження її у форматі JSON
2. Тестування та корегування розробленої мережі Петрі з використанням модуля імітаційного моделювання
3. Збереження мережі Петрі у вигляді метода для створення об'єктів із заданими параметрами
4. Запуск моделі на заданому часовому проміжку

Програмна реалізація моделі складається з таких кроків:

1. Розробка класів Петрі-об'єктів на основі створених мереж Петрі
2. Створення множин Петрі-об'єктів, з яких складається модель
3. Розробка зв'язків Петрі-об'єктів
4. Запуск моделі на заданому часовому проміжку

За результатами імітаційного моделювання визначаються вихідні характеристики моделі

на основі статистичних оцінок стану позицій та переходів мережі Петрі, а також кінцевого маркірування мережі Петрі.

Програмне забезпечення реалізоване з використанням Ruby on Rails та JavaScript. Динаміка та структура моделі створюються користувачем шляхом маніпуляцій з відповідними графічними об'єктами. Клієнтська частина додатку містить інтерфейс користувача та реалізований за допомогою JavaScript-фреймворку AngularJS та HTML5 Canvas. Алгоритми імітації реалізовані у вигляді Ruby gem'а(бібліотеки класів). Серверна частина являє собою RESTful JSON API, що дозволяє використовувати всі можливості розробленого програмного продукту іншим розробникам програмного забезпечення. Архітектура клієнт-серверного додатку представлена на рисунку 1.

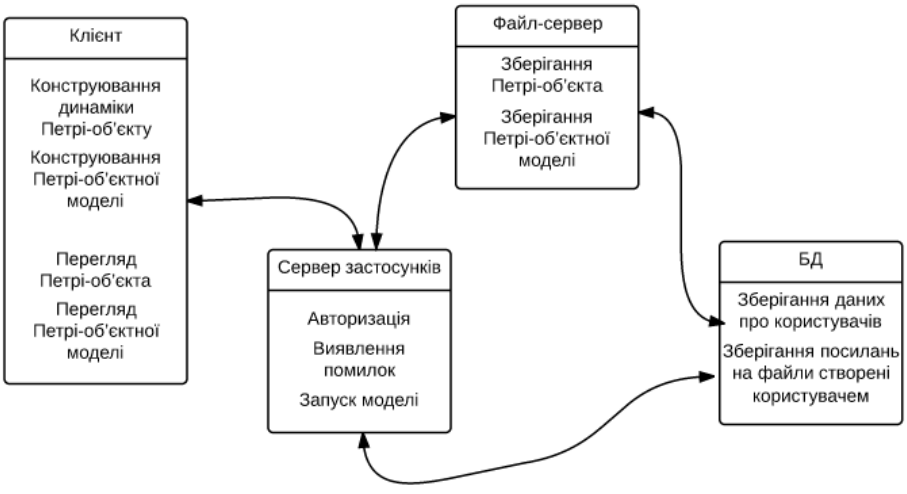


Рис. 1. Архітектура клієнт-серверного додатку

Таким чином, формалізм Петрі-об'єктного моделювання в поєднанні з відповідним програмним забезпеченням надає можливість візуального програмування складної динаміки моделі. Реалізація технології моделювання як клієнт-серверного додатку дозволить співпрацювати дистанційно віддаленим розробникам імітаційних моделей, накопичувати на одному веб-ресурсі та ефективно використовувати універсальні бібліотеки об'єктів моделювання, налаштованих на певну предметну галузь.

1. Law, A. M. Simulation Modeling and Analysis [Text] / A. M. Law // New York: McGraw-Hill International, 2014. - 804 p.

2. Томашевський, В. М. Моделювання систем [Текст] // К: Видавнича група ВНУ, 2005. - 352 с.

3. Карпов, Ю. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5 [Текст] / Ю. Карпов // СПб.: «БХВ-Петербург», 2005. – 400 с.

4. University of Hamburg, Department of Computer Science: DESMO-J - A Framework for descret4e-Event Modeling and Simulation J web-site [Electronic resource] - Mode of access: http://desmoj.sourceforge.net/ (viewed on Mart 9, 2016) - Title from the screen.

5. Bause, F. Stochastic Petri Nets: An Introduction to the Theory [Text] / F. Bause, Pieter S. Kritzinger // Cape Town: Vieweg+Teubner, 2002. – 218 p.

6. Стеценко, И. В. Теоретические основы Петри-объектного моделирования систем [Текст] / И. В. Стеценко // Математичні машини і системи, 2011. - №4. – С. 136-148.

7. Stetsenko, I. V. State equations of stochastic timed petri nets with Informational relations [Text] / I. V. Stetsenko // Cybernetics and Systems Analysis, 2012. – Vol.48, No. 5. - P.784-797.

***Ухина А.В.*[1], *Ситников В.С.*[1], *Ситникова В.А.*[2]**
[1] *Одесский национальный политехнический университет, Одесса, Украина;* [2] *Одесский национальный медицинский университет, Одесса, Украина*

Управление процессом предварительной обработки информации в экспериментальных медико-биологических исследованиях при использовании специализированных компьютерных систем

Современные медико-биологические исследования строятся на концепции доказательной медицины. В настоящее время технология рандомизированных контролируемых испытаний считается стандартом качества научных исследований эффективности лечения [1]. Для этого необходимо учитывать условия проведения эксперимента или исследования (клинические или лабораторные), снизить субъективный фактор исследователя и пациента, а также уменьшить систематическую ошибку, повысить объективность данных и т.п. Один из путей решения этой задачи фильтрация экспериментальных данных для устранения артефактов и шумов измерения. Однако применение стандартных и жестких алгоритмов фильтрации усложняет работу и приводит к многократной обработки исходных данных в специализированных компьютерных системах [2–4].

Неопределенность условий фильтрации в таких системах приводит к задаче адаптации или перестройки работы тракта предварительной обработки экспериментальных данных на основе заданных критериев. При этом используемые адаптивные алгоритмы в основном акцентируют внимание на подавлении сигнала без перестройки частотного диапазона.

Построение таких трактов базируется на перестраиваемых алгоритмах второго порядка общего вида

$$y[n] = a_0 x[n] + a_1 x[n-1] + a_2 x[n-2] - b_1 y[n-1] - b_2 x[n-2],$$

где a_0, a_1, a_2, b_1, b_2 – действительные коэффициенты, $x[n]$ и $y[n]$ – соответственно входной и выходной набор данных.

Например, с технической точки зрения для фильтрации экспериментальных данных в основном используются полосовые частотно-зависимые алгоритмы. Особенность такого алгоритма состоит в необходимости перестройки его относительной центральной частоты фильтрации $\overline{\omega_0}$, а возможно и полосы пропускания $\Delta\overline{\omega}$, а не только коэффициента усиления [5, 6].

Анализ полосовых алгоритмов такого обобщенного вида показал, что действительные коэффициенты при входных данных имеют следующие соотношения

$$a_0 > 0, a_1 = 0, a_2 < 0, a_0 = -a_2.$$

Однако, частотные свойства алгоритмов зависят от коэффициентов b_1, b_2. Исследования показали, что при управлении обработки в таких алгоритмах коэффициенты b_1, b_2 и a_0 зависят от параметров фильтрации: относительной центральной частоты $\overline{\omega_0}$ и полосы пропускания $\Delta\overline{\omega}$. Частотный анализ таких алгоритмов дает возможность получить соотношения, которые позволяют описать эти зависимости и выполнить управление свойствами алгоритма

$$\begin{cases} b_2 = \frac{\cos(\Delta\overline{\omega}) - \sin(\Delta\overline{\omega})}{\cos(\Delta\overline{\omega}) + \sin(\Delta\overline{\omega})}; \\ b_1 = -2\frac{\cos(\overline{\omega})}{\cos(\Delta\overline{\omega}) - \sin(\Delta\overline{\omega})}; \\ a_0 = \frac{\sin(\Delta\overline{\omega})}{\cos(\Delta\overline{\omega}) + \sin(\Delta\overline{\omega})} \end{cases}$$

Очень важным моментом при управлении свойствами алгоритма является вопрос обеспечения устойчивости алгоритма фильтрации. Устойчивость алгоритма второго порядка описывается треугольником устойчивости в координатах коэффициентов b_1 и b_2. Для полосовых частотно-зависимых алгоритмов треугольник устойчивости приведен на рис.1.

Кружочками показаны соответствующие нормированные центральные частоты, которые увеличиваются с низу в верх, изменяясь от 0 до 1.

Анализ устойчивости полосовых алгоритмов второго порядка показал, что:

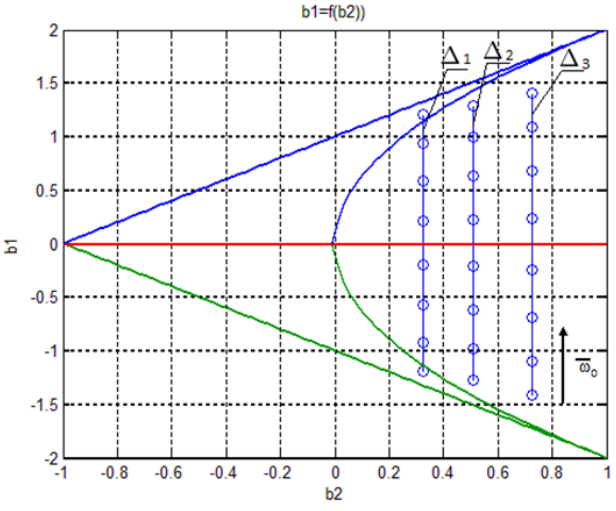


Рис. 1. Треугольник устойчивости для цифровых полосовых алгоритмов второго порядка, где полоса пропускания $\Delta_1 > \Delta_2 > \Delta_3$

1. При уменьшении полосы пропускания Δ коэффициент b_2 смещается в сторону увеличения до 1. При увеличении полосы пропускания Δ коэффициент b_2 стремится к 0. Поэтому при изменении полосы пропускания Δ возможен выход на границу устойчивости или выход за область устойчивости компоненты, что приводит к «разносу» выходных данных.
2. При изменении разрядности представления коэффициентов алгоритма в разрядной сетке специализированной компьютерной системы при исследовании устойчивости показали, что коэффициент b_1 имеет больший разброс, чем коэффициент b_2. Чем меньше величина коэффициента b_1 тем больше разброс, который может достигать 12 % относительной погрешности.

Таким образом, исследования управления свойствами алгоритма фильтрации второго порядка позволило получить простые соотношения, которые можно реализовать на микропроцессорном блоке предварительной обработки в составе специализированной компьютерной системы, а анализ устойчивости выявить критические области выхода за пределы устойчивости.

Литература. **1.** Клюшин Д.А., Петунин Ю.И. Доказательная медицина. Применение статистических методов. — М.: ООО "И.Д.Вильямс 2008. — 320 с. **2.** Моисеевич Г.И., Розенфельд Б.И., Шифрин М.А. Очерки о совместной работе математиков и врачей / Под ред. С.Г. Гиндикина. — М.: Едиториал УРСС, 2005. — 320 с. **3.** Котов Ю.Б. Новые математические подходы к задачам медицинской диагностики. М.: Едиториал УРСС, 2004. — 328 с. **4.** Лапач С.Н., Чубенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Exel. — К.: МИРИОН, 2001. — 408 с. **5.** Гайдаржи Г.М., Кравченко С.О., Ситников В.С. Специализированная компьютерная система сбора и обработки биомедицинских сигналов головного мозга // Строительство, материаловедение, машиностроение: Сб. научн. тр. — Вып. 78. — Серия: Компьютерные системы и информационные технологии в образовании, науке и управлении. — Приднепровская гос. акад. строительства и архитектуры. — Днепропетровск, 2014. — С. 57-61. **6.** Ухина А.В., Ситников В.С. Управление перестройкой полосовой частотнозависимой компоненты второго порядка для специализированной компьютерной системы // Матеріали XXII міжнар. конф. з автомат. упр-ня "Автоматика 2015"— (Одеса, Україна, 10 - 11 вересня, 2015) — Одеса, 2015. — С. 186-188.

Шевчик М.О.
Факультет прикладної математики НТУУ "КПІ", Київ, Україна

Особливості діагностики нерозпізнанного інтернет трафіку

Приводиться опис алгоритмів аналізу інтернет потоків, що можуть знаходити заголовки протоколів пакетної комутації, користуючись складними сигнатурами. Це дає змогу відновлювати потоки даних, та проводити дігностику роботи мережі.

Завдання аналізу мережевого трафіку набуває все більшої актуальності у зв'язку з розвитком і впровадженням нових мережевих технологій (і, як наслідок, збільшенням обсягу даних, що передаються по мережі), а також появою великої кількості нових мережевих протоколів прикладного рівня [1].

Розвиток інтернет-технологій призводить і до збільшення обсягу випуску різного інтернет-обладнання. На жаль, через це часто виникають проблеми з налаштуванням такого обладнання, що створює потребу у спеціальному програмному забезпеченнні, яке займається аналізом інтернет трафіку для визначення причин несправності мережі.

Такі програми керуються загальноприйнятими правилами побудови пакетів даних для того, щоб розібрати кожний з таких пакетів на складові частини, і надати системному адміністратору зручний інструмент для пошуку помилок в налаштуваннях мережного обладнання. Але такі програми безсильні в умовах, коли структура інтернет трафіку деформована, тобто не відповідає вищезгаданим загальним правилам [2].

В цьому випадку можна використовувати допоміжні алгоритми аналізу, що засновані на принципі розумного перебору даних. Ці алгоритми унікальні для кожного протоколу передачі даних, і виконують операції пошуку певних сигнатур серед байтового потоку інтернет трафіку.

В якості прикладу можна привести алгоритм пошуку сигнатури протоколу IP версії 4.

<table>
<tr><td>1 4</td><td>5 8</td><td>9 16</td><td>17 19</td><td>20 32</td></tr>
<tr><td>1.Vers</td><td>2. HLEN</td><td>3. Type of Service</td><td colspan="3">4. Total Length</td></tr>
<tr><td colspan="3">5. Identification</td><td>6. Flags</td><td>7. Fragment Offset</td></tr>
<tr><td colspan="2">8. Time to Live</td><td>9. Protocol</td><td colspan="3">10. Header Checksum</td></tr>
<tr><td colspan="5">11. Source IP address</td></tr>
<tr><td colspan="5">12. Destination IP address</td></tr>
<tr><td colspan="5">13. IP option</td></tr>
<tr><td colspan="5">14. Data</td></tr>
</table>

Рис. 1. Структура пакету IPv4

Для протоколу IPv4 перше поле (версія) завжди буде містити число 4. Отже, почати перебір потрібно з пошуку байтів, які мають вигляд 0004****, де * - будь яка цифра. Після цього можна вважати знайдений байт першим байтом заголовку. Базуючись на цьому, потріно вирахувати контрольну суму заголовку пакету (checksum), та порівняти її з полем Header Checksum. Якщо ці дві контрольні суми співпадуть, то можна з високою імовірністю (більше 0.9999) сказати, що знайдена послідовність байт є заголовком пакету IPv4.

Такі алгоритми можна застосувати майже для будь якого протоколу пакетної комутації даних. Загальною рисою цих алгоритмів є виділення певних ознак протоколу, завдяки яким можна з найбільшою точністю та найменшими затратами комп'ютерних ресурсів визначити положення першого байту заданого протоколу.

Після отримання цієї інформації можна продовжувати розбір байтового потоку, користуючись отриманими даними. Це дає системному адміністратору чи аналітику змогу отримати повну картину несправностей мережі, і зрозуміти що саме стало їх причиною та як їх виправити. Завдяки чому значно полегшується задача діагностики роботи мережі.

Література. **1.** Партика Т.Л. Информационная безопасность (2-е издание)// М.: «Academia», 2007. – 456 с. **2.** Уилсон Ед. Мониторинг и анализ сетей. Методы выявления неисправностей // Москва: Издательство «Лори», 2002. – 561 с.

Шеренковський А.О.
Інститут прикладного системного аналізу НТУУ "КПІ", Київ, Україна

Програмне забезпечення медіа-серверу Інтернет радіо

Незважаючи на велику кількість медіа контену в мережі Інтернет, радіомовлення через всесвітню павутину у сьогоденні не лише не втрачяає своїх позицій, а й навпаки знаходить все більше й більше прихильників. За данними статистики, кількість інтернет радіослухачів у США станом на 2014 рік була близькою до 160 млн. осіб, що складає близько 50% всього населення Сполученних Штатів, та близько 67% американців, що користуються інтеренетом.Разом із тим зростає час, який користувачі проводять слухаючи онлайн-радіо [1].

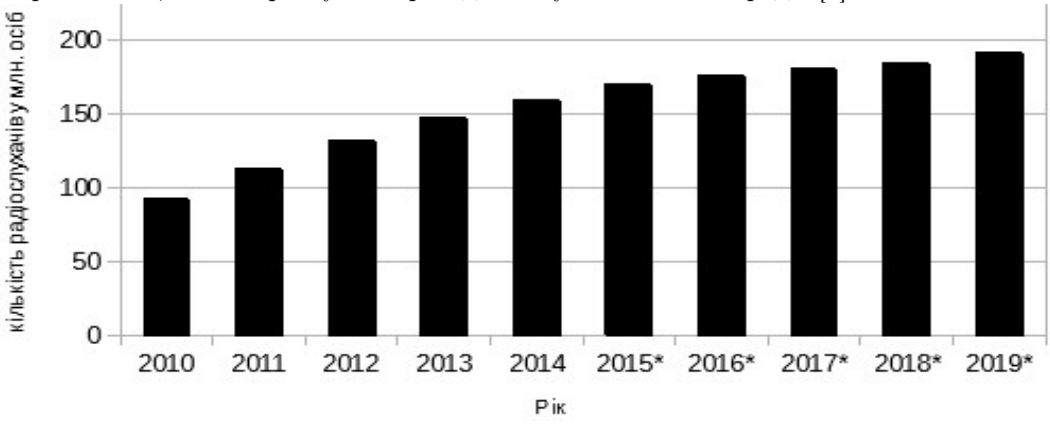


Рис. 1. Графік збільшення кількості інтернет-радіослухачів(* - прогнозоване збільшення) [2].

В данній роботі будуть розглядатися декілька альтернатив медіа-серверів для Інтернет радіомовлення. Для вирішення завдання слід відповісти на наступні питання:

- Які є готові альтернативи реалізацій медіа-серверів для Інтернет радіомовлення?
- Які є преваги та недоліки, знайдених реалізацій?

Структура. У Інтернет радіомовленні є 3 основних компоненти:

- Станція - генерує потік аудіоданих, та направляє його на сервер. Потребує малої пропускної здатності мережі, бо створює лише 1 поток(на сервер).
- Сервер - приймає потік аудіоданих від станції, та направляє його всім підключеним клієнтам. Потребує набільшої пропускної здатності мережі, бо генерує окремий потік до кожного клієнту.
- Клієнт - приймає поток аудіоданих від сервера та перетворює його на аудіосигнал, який сприймається кінцевим користувачем. Потребує малої пропускної здатності як і станція, бо лише приймає один потік від серверу.

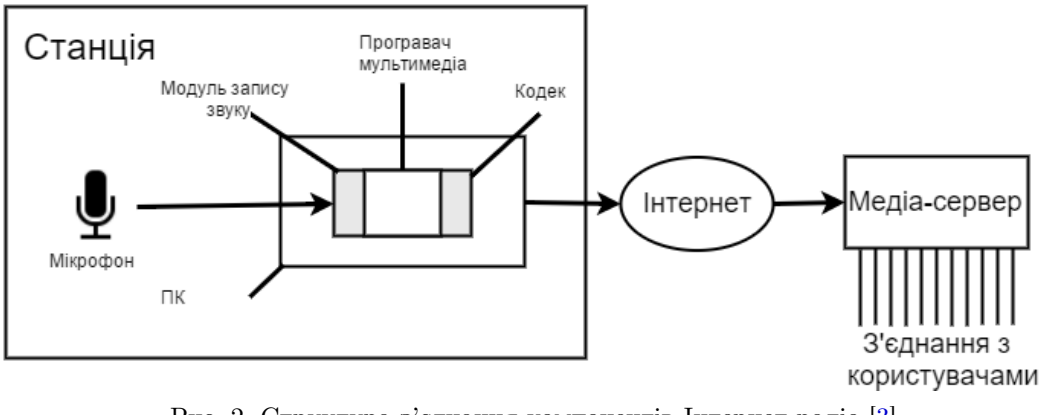


Рис. 2. Структура з'єднання компонентів Інтернет радіо [3].

Медіа-сервер. Найважливішою частиною Інтернет радіомовлення є медіа-сервер - програмне забезпечення, що до якого під'єднуються усі потоки аудіоданих, як вхідні, так і вихідні. Існують двай найпопулярніших медіа-сервера для радіомовлення: SHOUTcast та IceCast.

- SHOUTcast - це кроссплатформне програмне забезпечення для потокової передачі данних через мережу Інтернет, що поширюється у вільному доступі, проте з закритим кодом. Плюсами даного продукту є його легкість у налаштуванні та простота використання. Для цього ПЗ наявна велика кількість додаткових скриптів та інших додатків. Дає змогу потокового мовлення аудіо або відео потоків у форматах MP3, AAC, AACplus, NSV. SHOUTcast сервери і клієнти доступні для FreeBSD, Linux, Mac OS X, Microsoft Windows і Solaris. Також існуть клієнтські версії для таких операційних систем, як: Android, iOS, Windows Mobile, Blackberry OS та ін. Основним мінусом данного продукту є велика завантаженність пам'ятіб CPU та використання великої кількості трафіку. Дуже мало вбудованних фунцій та функцій адміністративного центру, а сам сервер продовжує працювати навіть при відсутності користувачів. Недоліком можно вважати відсутність можливості створення більше одного аудіопотоку на одному сервер. Це говорить про необхідність встановлення декількох медіа-серверів для паралельного мовлння декількох різних аудіопотоків.
- IceCast - безкоштовне програмне запезпечення від Xiph.Org Foundation. IceCast також є кроссплатформеним, проте на відміну від SHOUTcast має відкритий, що дає змогу самостійно змінювати, додавати або видаляти певний функціонал. У наявності користувача є велика кількість вбудованих функцій, а також великий набір зручних функцій адміністративної панелі. Підримує велику кількість аудіо та відео форматів, серед яких, як розповсюдженні MP3, AAC, так і менш відомі Ogg Vorbis, Opus, Theora та NSV. На відміну від SHOUTcast, менша навантаженність CPU(в деяких ситуаціях до 10 разів), оперативної пам'яті та мережевого трафіку, а мовлення відбувається лише при наявності слухачів. Значною перевагою, порівняно з аналогом, є можливість створення декількох паралельних аудіопотоків для радіомовлення різних станцій. Мінусами данного програмного забезпечення є більша складність в установці та налаштуванні, а встановлення можливе лише на UNIX-подібні операційні системи, або на операційну систему Windows. Наявно мало додатків та скриптів.

Висновки. Було досліджено два набільш популярних варіанти програмного забезпечення для потокової передачі даних медіа-серверів: IceCast та SHOUTcast. Було розглянуто іхні недоліки та переваги. Виходячи з результатів аналізу IceCast є кращим вибором за багатьма показниками, головнии серед яких є навантаженість на апаратну частину серверу. Надати перевагу SHOUTcast можна лише при неохідності швидкого налаштування серверу, а при потребі у існуванні клієнтських программ для багатьох операційних систем, включаючи мобільні операційні системи.

Література. **1.** Internet Radio Trends Report 2015, [Електроний ресурс] URL: http://xappmedia.com/wp-content/uploads/2015/01/Internet-Radio-Trends-Report-2015_january.pdf **2.** Number of monthly digital radio listeners in the United States from 2010 to 2019, [Електроний ресурс] URL: http://www.statista.com/statistics/191657/us-weekly-internet-radio-listeners-since-2009/ **3.** Таненбаум Э. Компьютерные сети. – М.: «Питер», 2012. – 767 с

Шишков О.Н.
Институт прикладного системного анализа НТУУ "КПИ", Киев, Украина

Распознавание рукописного ввода с помощью рекуррентный нейронных сетей

Вступление. С развитием мобильных телефонов и планшетов появилась возможность вводить всё больше информации на устройство. Примером тому может быть серия смартфонов и планшетов Samsung Note с стилусом S Pen, который позволяет вводить информацию неотличимо от бумажных носителей. Для упрощения взаимодействия с устройством, система должна не только фиксировать ввод, но и иметь возможность анализировать и распознавать закономерности. Наиболее распространённым сценарием взаимодействия с телефоном или планшетом является ввод текстовой информации.

Постановка задачи. Задача распознавания рукописного ввода заключается в автоматическом преобразовании пространственной информации о тексте в последовательность символов из заранее заданного алфавита. Данные ввода поступают в виде траектории, которая задается последовательностью точек.

Процесс распознавания рукописного ввода состоит из следующих этапов:

1. Предобработка. Прореживание, фильтрация, выравнивание и выделение признаков.
2. Классификация. Для каждой точки вычисляется вероятность ее принадлежности к каждому символу из алфавита.
3. Декодирование. Из вероятностей, полученных на предыдущем шаге, и из дополнительной информации о языке (словарь, языковая модель), находится наиболее вероятное слово или предложение.

В данной работе предлагается метод, который позволяет улучшить скорость и качество классификации точек в существующих систем онлайн распознавания.

Данные. Для обучения и тестирования использовались открытый наборы онлайн рукописных данных Unipen [3] и IAM-OnDB [4]. Пример данных Unipen представлен на рис. 1.

Каждый образец данных состоит из набора точек, которые описывают траекторию движения стилуса при написании, и текст, который писал пользователь. Информация о выравнивании между точками и текстом отсутствует, неизвестно какой букве принадлежит каждая точка.

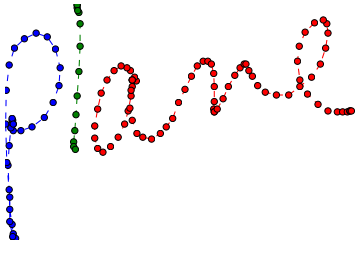


Рис. 1. Пример рукописного ввода. Слово "plane"

Метод решения задачи. Основным методами для распознавания рукописного ввода являются Скрытые Марковские Модели (англ. Hidden Markov Model, HMM). Сложностью использования данного подхода является большое количество гиперпараметров, которые необходимо настраивать вручную. Для каждого символа из алфавита необходимо подобрать количество состояний Скрытой Марковской Модели, которая его описывает, в зависимости от сложности и вариативности написания символа.

Вместо Скрытых Марковских Моделей предлагается использовать рекуррентные нейронные сети. Обучение нейронной сети проще и занимает меньше времени. Необходимо настраивать всего несколько гиперпараметров. Наиболее подходящими для данной задачи являются рекуррентные сети с долгой кратковременной памятью (англ. Long Short-Term Memory, LSTM) [1].

Алгоритм Connectionist Temporal Classification (CTC) [1] объединяет в себе некоторые идеи из HMM, динамического программирования и алгоритма прямого и обратного прохода. При использовании CTC нет необходимости находиться выравнивание. Результатом работы CTC

являются значения ошибки нейронной сети и градиентов, необходимых для метода обратного распространения ошибки во времени (англ. Backpropagation through time, BPTT).

CTC может быть легко встроен в существующие системы обучения нейронных сетей как новый слой без модификации алгоритма обучения или других слоев. Данный подход показал хорошие результаты при распознавании речи [2], и является наиболее перспективным для распознавания рукописного ввода, так как эти задачи имеют схожую структуру и отличаются только предобработкой и способом выделения признаков.

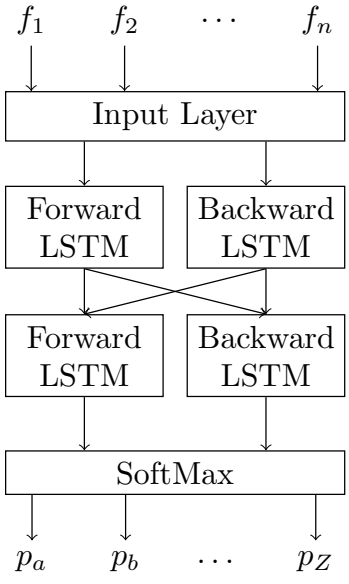


Рис. 2. Архитектура сети

Архитектура сети. Сеть состоит из 4 слоев (рис. 2.). Первый слой не производит вычислений и используется для загрузки данных в сеть. По размеру совпадает с количеством используемых признаков. Далее идут два скрытых BLSTM слоя. Выбор размера BLSTM слоя зависит от поставленной задачи, сложности данных и необходимого времени распознавания. С увеличением BLSTM слоя повышается точность, но увеличивается время распознавания и возрастает риск переобучения. Последний слой - Softmax. Каждый выход последнего слоя соответствует символу из алфавита или фиктивной метке.

Результаты. На рис. 3. изображен результат распознавания слова "plane". Как видно из графика, после обучения нейронной сети с CTC выходы состоят из пиков в 1-2 точках в случае наличия символа. Сплошные линии соответствуют символам из алфавита, пунктирной линией обозначена фиктивная метка. Предложенный метод работает и показывает хорошие результаты разпознавания на тестовых данных.

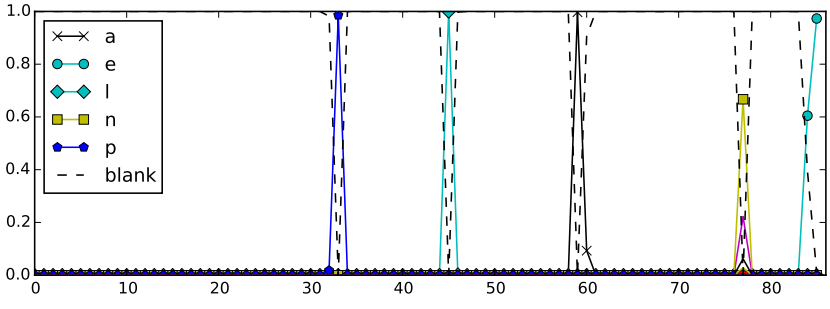


Рис. 3. Выходы нейронной сети после распознавания слова "plane"

Литература. **1.** Supervised Sequence Labelling with Recurrent Neural Networks / Alex Graves., 2012. – 146 p. **2.** Fast and Accurate Recurrent Neural Network Acoustic Models for Speech Recognition / H.Sak, A. Senior, K. Rao, F. Beaufay. // CoRR. – 2015. **3.** UNIPEN Project of On-Line Data Exchange and Recognizer Benchmarks / I. Guyon, L. Schomaker, R. Plamondon, M. Liberman, S. Janet. // Proc. of the 14th International Conference on Pattern Recognition, Jerusalem (Israël). – 1994. – pp. 29-33. **4.** IAM-OnDB - an on-line English sentence database acquired from handwritten text on a whiteboard / M. Liwicki, H. Bunke. // Proc. 8th Int. Conf. on Document Analysis and Recognition, volume 2. – 2005. – pp. 956-961.

Шліхтенко Н.М., Ярмак Д.О.
Факультет інформатики та обчислювальної техніки НТУУ "КПІ", Київ, Україна

Паралельні алгоритми розв'язування систем лінійних алгебраїчних рівнянь та алгебраїчної проблеми власних значень

Сучасні потреби у вирішенні якісно нових науково-технічних завдань призводять до необхідності обробки великих об'ємів інформації за прийнятний час, зокрема – розв'язання систем лінійних алгебраїчних рівнянь (СЛАР) з великими матрицями коефіцієнтів. СЛАР виникають при вирішенні низки прикладних задач, що описуються системами диференціальних, інтегральних або нелінійних рівнянь, у задачах математичного програмування, апроксимації функцій, статистичної обробки даних, при дискретизації задач на власні значення проекційно-різницевим методом та в багатьох інших випадках. Порядок матриць, що виникають в таких задачах, є дуже високим і сягає десятків мільйонів[1]. Важливою особливістю є те, що ці матриці є зазвичай розрідженими (кількість елементів, що не дорівнюють нулю, становить k_n, де n - це порядок матриці, $k << n$), симетричними і додатньовизначеними (або напіввизначеними).

Оскільки порядки матриць надзвичайно високі, то розв'язування таких задач вимагає значних обчислювальних ресурсів, незважаючи на розріджену структуру матриць. Наразі ефективне обчислення можна забезпечити за допомогою багатопроцесорних комп'ютерів, розпаралеливши обчислення. Тому актуальним є створення ефективних паралельних алгоритмів для розв'язування задач із розрідженими матрицями. Ефективний алгоритм - це такий алгоритм, що дозволяє отримати достовірний розв'язок задачі з мінімальним використанням ресурсів комп'ютера – процесорів, пам'яті, часу. Ефективність паралельних алгоритмів зазвичай оцінюється за допомогою коефіцієнтів прискорення і ефективності.

В роботі наведені паралельні алгоритми для двох видів задач із розрідженими матрицями:

- СЛАР з розрідженою симетричною додатно визначеною матрицею:
$$Ax = b; \tag{1}$$
- алгебраїчна проблема власних значень (АПВЗ) з розрідженими симетричними матрицями, одна з яких додатно визначена, а інша - додатно напіввизначена:
$$Ax = \lambda Bx. \tag{2}$$

Для розв'язування СЛАР із симетричною додатно визначеною матрицею застосовується метод Холецького або метод Гауса без вибору головного елемента[2]. При цьому варто виділити три основні підзадачі:

1. Розвинення матриці СЛАР у добуток трьох матриць
$$A = LDL^T \text{ або } A = U^T DU, \tag{3}$$
де L - нижня трикутна матриця, U - верхня трикутна матриця, D - діагональна матриця.
2. Розв'язування СЛАР з нижньою трикутною матрицею:
$$Ly = b \text{ або } U^T y = b. \tag{4}$$
3. Розв'язування СЛАР з верхньою трикутною матрицею:
$$L^T x = D^{-1} y. \tag{5}$$

Для розв'язування АПВЗ з розрідженими матрицями застосовують метод ітерацій на підпросторі. Він полягає у побудові послідовності підпросторів, яка збігається до одного підпростору, що містить у собі шукані власні вектори, що відповідають мінімальним власним значенням. Важливо відмітити, що при застосуванні цього методу виникає необхідність розв'язання СЛАР, тобто використовуються вищезгадані методи для розв'язування СЛАР.

Література. **1.** Джордж А., Лю Дж. Численное решение больших разреженных систем уравнений. – М.: Мир, 1984. – 333 с. **2.** А.Н. Химич, И.Н. Молчанов, А.В. Попов. Параллельные алгоритмы решения задач вычислительной математики – Киев: Наук. думка, 2008. – 248 с.

***Шпортько О.В.*[1], *Шпортько Л.В.*[2]**
[1]*Рівненський державний гуманітарний університет, Рівне, Україна;* [2]*ДВНЗ "Рівненський коледж економіки та бізнесу", Рівне, Україна*

Особливості реалізації розбиття даних зображень на однорідні блоки перед застосуванням арифметичного кодування

Обгрунтована необхідність розбиття стиснутих даних перед застосуванням арифметичного кодування на блоки, які мають між собою суттєво різні відносні частоти окремих елементів. Проаналізовані особливості застосування контекстно-незалежного кодування в процесі прогресуючого ієрархічного стиснення зображень без втрат.

На сьогодні одним з основних методів стиснення зображень є контекстно-незалежне кодування, яке ставить у відповідність елементам, які зустрічаються частіше, коди не більшої довжини, ніж для елементів, котрі трапляються рідше. На практиці найчастіше використовуються два альтернативні контекстно-незалежні методи: кодування Хафмана (HUFF) [1] та арифметичне кодування (ARIC) [2], які зменшують розмір стиснутих зображень приблизно на третину. Кодування Хафмана ставить у відповідність кожному елементу префіксний код з цілої кількості бітів. Арифметичне ж стиснення, на відміну від кодування HUFF, для кожного чергового елемента визначає інтервал з довжиною, пропорційною його ймовірності (частоті) [2], і тому середня довжина арифметичного коду наближається до ентропії. Враховуючи те, що термін дії основних патентів на ARIC вже минув і таке стиснення точніше кодує окремі елементи від кодування HUFF, а довжина арифметичного коду найімовірнішого елемента може бути меншою одного біта, дослідимо особливості застосування арифметичного кодування в процесі стиснення зображень без втрат.

Опрацювання яскравостей пікселів зображень у популярних графічних форматах, які виконують стиснення без втрат, на сьогоднішній день найчастіше здійснюється послідовно по рядках зверху вниз, а у кожному рядку – поспіль зліва направо [3]. Як наслідок, вивести стиснуте зображення у цих форматах можливо лише після декодування всіх пікселів, а декомпресія фотокарток чи малюнків з мільйонами пікселів при такому способі обходу може тривати декілька секунд незалежно від розміру області та роздільної здатності пристрою виводу. Поряд з цим, для прискорення виводу великих зображень у форматах компресії з втратами найчастіше застосовують прогресуюче (поступальне) ієрархічне опрацювання пікселів [2]. В процесі застосування цього способу обходу піксели зображення обробляють пошарово, збільшуючи щоразу роздільну здатність, причому в процесі послідовного опрацювання даних чергового шару використовують дані попередніх шарів. Зображення з пікселів чергового шару фактично є зменшеною у декілька разів (найчастіше – у чотири) копією зображення з пікселів наступного шару, а останній шар співпадає з вхідним зображенням. Тому під час прогресуючого ієрархічного декодування деталі зображення проявляються поступово. Зупинити таке декодування можливо вже після декомпресії шару з кількістю пікселів, не меншою від області виводу по кожній з осей, не очікуючи відтворення всіх пікселів зображення. Отже, розробка контекстно-незалежних способів кодування для графічного формату компресії зображень без втрат з використанням принципів прогресуючого ієрархічного опрацювання є на сьогодні актуальним завданням.

З опису принципів ARIC слідує, що для виконання такого стиснення кодеру і, тим більше, декодеру мають бути відомі ймовірності (частоти) окремих елементів. Сьогодні широко використовуються дві стратегії формування цих частот: статична та адаптивна. У випадку статичної стратегії частоти окремих елементів передаються декодеру у стиснутих даних явно. При використанні ж адаптивної стратегії інтервали елементів формуються синхронно кодером та декодером в процесі опрацювання даних. Адаптивна стратегія забезпечує, як правило, краще стиснення (оскільки не потребує зберігання у стиснутих даних частот окремих елементів та враховує нерівномірності розподілу поточної послідовності елементів, а не потоку загалом [4]), але суттєво сповільнює роботу декомпресора, бо вимагає перерахунку ймовірностей елементів у процесі декодування. Формати графічних файлів мають забезпечувати, насамперед,

швидке декодування, тому для реалізації прогресуючого стиснення зображень без втрат ми використали статичну стратегію формування інтервалів елементів.

Для однозначного декодування блоку статичних кодів ARIC на його початку має міститися заголовок, який дає змогу кожному елементу поставити в однозначну відповідність інтервал з довжиною, пропорційною його ймовірності, оскільки ймовірності елементів в різних блоках можуть відрізнятися. На перший погляд здається, що стиснення даних зображення в єдиний блок покращує коефіцієнт стиснення (КС), адже тоді буде використано лише один заголовок блоку замість багатьох. З іншого боку, розглянемо, наприклад, послідовність елементів "1, 2, 1, 1, 2, 1, 2, 2, 2, 1, 3, 4, 3, 4, 4, 3, 3, 3, 4, 4", яка містить чотири різні елементи "1", "2", "3", "4" з однаковою частотою 5 та ймовірністю 0.25. Якщо згенерувати арифметичні коди для всієї послідовності, як єдиного блоку, то будь-якому елементу буде поставлено у відповідність код довжиною 2 біти і вся послідовність буде закодована в 40 бітів. Якщо ж розбити цю послідовність на два рівні блоки, то в кожному з них зустрінуться лише по два різні елементи з ймовірностями 0.5 і кожному з цих елементів буде поставлено у відповідність код довжиною 1 біт, а дані всієї послідовності закодуються в 20 бітів. Тобто розбиття даних на блоки не збільшує (найчастіше – зменшує) загальну довжину арифметичного коду (без врахування довжин їх заголовків). Приріст довжини ентропійного коду (до якого наближається довжина коду ARIC) внаслідок поєднання блоків зростає зі збільшенням різниці між відносними частотами елемента у вхідних послідовностях.

Отже, створення малих блоків ARIC призводять до зайвих витрат пам'яті для зберігання заголовків, а великих – до неврахування змін характеристик даних. Тому **стиснуті дані перед застосуванням ARIC доцільно розбивати на блоки, які мають між собою суттєво різні відносні частоти (ймовірності) окремих елементів**.

Як відомо, підвищити ефективність контекстно-незалежного кодування в процесі стиснення зображень без втрат намагаються за допомогою предикторів, які під час обходу прогнозують значення яскравості кожної компоненти чергового піксела (наприклад, для найпоширеніших 24-бітних зображень – це яскравості червоної, зеленої та синьої компоненти, записані цілими числами в окремих байтах), використовуючи значення яскравостей тих самих компонентів опрацьованих раніше суміжних пікселів, адже дані яскравості мають між собою найбільшу степінь кореляції [5]. Після застосування предикторів розподіли частот значень компонентів окремих блоків зображення близькі до розподілу Лапласа [2] і тому найбільший вплив на довжину ентропійного коду стиснутих даних мають елементи, близькі до нуля. Ентропія даних зображень в процесі ієрархічного обходу пікселів значно змінюється (зменшується) при переході між шарами та проходами внаслідок збільшення нерівномірності розподілу, адже під час таких переходів відносні частоти елементів, близьких до нуля, змінюються суттєво. Тому **дані різних шарів і проходів мають стискуватися у різні блоки ARIC**.

Розбиття даних зображень на блоки в наших програмних реалізаціях ієрархічного стиснення без втрат виконується автоматично при переходах між шарами та проходами і в середині проходів, оскільки максимальний розмір блоку даних обмежується 64 Кб. Створення таких додаткових блоків зменшує розмір арифметичних кодів, хоча й генерує їх додаткові заголовки, тому й значно не впливає на середній КС, але прискорює кодування на 2.27 %, що й свідчить про доцільність такого розбиття.

Література. **1.** Huffman D. A Method for the Construction of Minimum Redundancy Codes / D. Huffman // Proceedings of the IRE. – Sept. 1952. – Vol. 40(9). – P. 1098–1101. **2.** Сэломон Д. Сжатие данных, изображений и звука / Д. Сэломон. – М. : Техносфера, 2006. – 368 с. **3.** Миано Дж. Форматы и алгоритмы сжатия изображений в действии: учеб. пособ. / Дж. Миано. – М. : Триумф, 2003. – С. 249-318. – (Практика программирования). **4.** Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М. : ДИАЛОГ-МИФИ, 2003. – 384 с. **5.** Прэтт Э. Цифровая обработка изображений: Пер. с англ. / Э. Прэтт. – М.: Мир, 1982. – Кн. 2, 480 с.

Authors · Авторы · Автори

Ademola Samuel Adegoke, 296
Antonyuk, Sofia A., 266
Arkhipov, Dmytro V., 186
Artyukhov, Vitaliy G., 266
Azzuz, Iskandar D., 187
Bakun, Sabina A., 299
Basyuk, Taras M., 301
Bidyuk, Petro I., 42, 188
Bilushchak, Yurii I., 24
Bolea, Cecilia, 44
Chenarani, Ali, 26
Chernukha, Olha Y., 24
Cojocaru, Svetlana, 28
Coman, Lilia, 28
Davydok, Anastasiia Y., 24
Didenko, Dmytro G., 303
Dozirtsiv, Volodymyr O., 38
Druzhinin, Evgeny A., 26
Dzhygyrey, Iryna M., 42
Fomin, Oleksandr O., 30
Goncharuk, Volodymyr Y., 24
Gorban, Nataliia V., 32
Guskova, Vera G., 196
Herasymiv, Ihor S., 33
Horos, Grigore, 28
Ivanova, Yuliia V., 304
Kavrelishvili Avtandil G., 306
Kharchenko, Kostiantyn V., 268
Korshevnyuk, Lev O., 304
Kovalenko, Anatolii E., 33
Kovalenko, Anatoly E., 269
Kyselova, Anna G., 195
Levin, Dmytro K., 190
Livinskiy, Olexander V., 194
Lulu, Karim M., 34
Lytvyniuk, Anton A., 299
Makukha, Mykhailo P., 318
Malynovskyi, Artyom Dmytryevych, 307
Markov, Dmitriy K., 310
Markova, Valentyna M., 191
Mieshkov, Oleksandr Yu., 35
Nikolaiev, Sergii Sergiiovych, 307
Novikov, Oleksandr O., 35
Oparina, Elena L., 318
Ovi Nafas Aghaei Agh Ghamish, 198
Paliichuk, Liliia S., 32
Pankratova, Nataliya D., 38
Parkhomchuk, Dmytro M., 39
Pavlenko, Vitaliy D., 30
Petic, Mircea, 28
Povisenko, Sergey V., 312
Prosyankina-Zharova, Tatiana I., 304
Radchenko, Olexandra Yu., 194
Radovanovic, Milan, 314
Reznik, Ilia M., 192
Romanko, Oleksii R., 194
Savastiyanov, Volodymyr V., 40
Savastynov, Volodymyr V., 318
Shaptala, Roman V., 195
Shatalov, Vladimir M., 192
Shcherbachenko, Larysa V., 41
Sidletskiy, Viktor M., 316
Sobko Valentyna Oleksandrivna, 42
Strbac, Dragoljub, 314
Teodorescu, Horia Nicolai, 44
Terentiev, Oleksandr M., 42
Terentiev, Oleksandr N., 304
Terentiev, Olexandr M., 299, 318
Titchiev, Inga I., 46
Trofimov, Alexander S., 320
Tsymbal, Maksym D., 322
Tymoshenko, Yuriy O., 39
Vertiletskij, Vladimir G., 188
Zaychenko, Yurii P., 196, 198
Zenkovich, Mikhail V., 320
Zorin, Yurii M., 48
Алєксєєв Микола Олександрович, 211
Алєксєєв Миколай Олександрович, 199

Алфимцев Александр Николаевич, 254
Андреев Николай Варфоломеевич, 50
Андреєва Марина Олександрівна, 51, 108
Андрущенко Валентина Борисівна, 52
Аникин Вадим Юрьевич, 201
Арчвадзе Натела Нодариевна, 324
Балашевич Роман Валерыйович, 325
Барановский Алексей Николаевич, 53
Бардадим Іван Володимирович, 54
Безлюднов Данїїл Євгенович, 199
Безносик Олександр Юрійович, 285
Безносик Юрій Олександрович, 328
Бідюк Петро Іванович, 145, 202, 203, 222, 262
Білова Марія Олексіївна, 96
Бобрик Ірина Максиміліанівна, 368
Богун Ксенія Володимирівна, 203
Бойко Оксана Михайлівна, 205
Бондаренко Алексей Сергеевич, 326
Боярінова Юлія Євгенівна, 356
Боярінова Юлія євгенівна, 327
Бритов Олексій Анатолійович, 271
Бугаєва Людмила Миколаївна, 328
Будник Сергій Анатолійович, 330
Булах Богдан Вікторович, 273
Буценко Юрій Павлович, 56, 332
Вартанян Василий Михайлович, 249
Васильев Владимир Иванович, 57
Вергун Костянтин Валерійович, 152
Вероцький Денис Сергійович, 59
Виклюк Ярослав Ігорович, 334
Вишталь Дмитрий Михайлович, 57
Волкова Виолетта Николаевна, 60
Воронов Михаил Александрович, 422
Гайвась Богдана Іванівна, 63
Галіцька Вікторія Володимирівна, 335
Галушко Марія Олегівна, 337
Гамоцька Сніжана Леонідівна, 65
Ганиева Сачлы Абдулхак кызы, 338
Ганчукова Дар'я Володимирівна, 207
Гарт Людмила Лаврентьевна, 92
Гіоргізова-Гай Вікторія Шалвівна, 325
Глоба Лариса Сергіївна, 209
Годлевский Михаил Дмитриевич, 149
Головань Михайло Володимирович, 131
Головатюк Володимир Олегович, 67
Голоскокова Анна Александровна, 149
Голубовський Олег Ігорович, 360
Гончар Вячеслав Олександрович, 69
Горбань Ірина Миколаївна, 70
Горелова Галина Викторовна, 72
Городецкий Виктор Георгиевич, 74
Гридчук Денис Тарасович, 402
Губарев Вячеслав Федорович, 76
Гура Едуард Сергійович, 388
Гусак Олександр Вадимович, 211
Гусак Олена Михайлівна, 334
Даник Юрий Григорьевич, 77
Даник Юрій Григорович, 79
Дейнеко Андрей Леонидович, 405
Декрет Маргарита Володимирівна, 213
Денисенко Олександр Іванович, 81
Денисов Сергій Віталійович, 274
Дидковская Марина Витальевна, 240
Дифучин Антон Юрійович, 424
Дичко Віктор Миколайович, 341
Діденко Дмитро Георгієвич, 342
Діденко Дмитро Георгійович, 343, 345, 346
Дмитрієнко Валерій Дмитрович, 82
Дмитрук Анатолій Андрійович, 63
Дмитрук Вероніка Анатоліївна, 63
Довніч Олександр Сергійович, 101
Донець Надія Петрівна, 83
Донченко Володимир Степанович, 84
Дорофеев Юрий Иванович, 14
Дуда Олексій Михайлович, 214, 215
Дюков Игорь Николаевич, 77
Дюков Ігор Миколайович, 79
Дяденко Ольга Николаевна, 76
Дяченко Олексій Миколайович, 85
Євтухов Сергій Костянтинович, 348
Ємельянова Дар'я Ігорівна, 98
Железко Борис Александрович, 236
Журавель Сергей Вадимович, 94
Журавлёв Павел Владимирович, 350
Журке Богда Сергійович, 342
Забелин Станислав Игоревич, 352
Загороднюк Андрей Александрович, 353
Зайченко Юрий Петрович, 217
Заковоротний Олександр Юрійович, 82
Згуровський Михайло Захарович, 16
Зінько Тарас Петрович, 84
Зражевська Наталія Григорівна, 86
Зубарева Поліна Костянтинівна, 89
Ивченко Дмитрий Анатолиевич, 354
Іваненко Євгеній Вікторович, 151
Іванущак Наталія Михайлівна, 218

Каковська Тетяна Ігорівна, 220
Калінін Владислав Валерійович, 356, 357
Каліновський Яків олександрович, 327
Калюжний Максим Сергійович, 88
Каніовська Ірина Юріївна, 89, 102
Капустін Володимир Віталійович, 358
Капшук Олег Олексійович , 276
Караюз Ірина Валентинівна, 222
Карпенко Олег Вікторович, 360
Качаловський Андрій Сергійович, 91
Квитко Марк Владимирович, 223
Кириленко Елена Георгиевна, 361
Киричек Галина Григорівна, 362
Кириченко Вікторія Едуардівна, 364
Кириченко Людмила Олеговна, 53
Кирюша Богдан Анатоллійович, 278
Киселева Елена Михайловна, 92, 94
Кисельов Геннадій Дмитрович, 250
Кислий Роман Володимирович, 366
Климковська Катерина Дмитрівна, 343
Кобицкая Юлия Александровна, 53
Коваль Олександр Васильович, 209
Козленко Анастасия Сергеевна, 367
Козловська Вікторія Олександрівна, 227
Козуля Марія Михайлівна, 178
Козуля Тетяна Володимирівна, 96, 98
Колбасин Вячеслав Александрович, 100
Комлєв Олександр Олександрович, 101
Кондратенко Наталія Романівна, 224
Кондратова Людмила Павловна, 132, 137
Кондратьєв Ігор Іванович, 226
Конєв Дмитро Володимирович, 345
Копичко Сергій Миколайович, 368
Кораб Катерина Олександрівна, 102
Корбутяк Василь Михайлович, 173
Корепанова Наталия Леонидовна, 369
Корнага Ярослав Ігоревич, 182
Корначевский Ярослав Ильич, 287
Котов Роман Олександрович, 362
Кошкин Егор Геннадиевич, 371
Кравець Ірина Олександрівна, 227
Кравченко Анатолій Михайлович, 104
Кравчук Євген Сергійович, 372
Крамар Олександр Володимирович, 273
Краснюк Роман Петрович, 174
Крижанівський Володимир Віталійович, 374
Круковский Максим Юрьевич, 375
Круш Ігор Володимирович, 376
Крюков Олександр Олександрович, 346
Кузенков Олександр Олександрович, 105
Кузнецова Юлия Анатольевна, 377
Кузнєцова Наталія Володимирівна, 106
Кунанець Наталія Едуардівна, 215
Кунанець Наталя Едуардівна, 214
Курдуп Іван Олександрович, 182
Курінний Ігор Валерійович, 228
Лабжинський Володимир Анатолійович, 56
Лавренюк Алла Миколаївна, 379
Лавренюк Сергій Іванович, 379
Ланде Дмитро Володимирович, 52, 381
Лебедева Марина Анатольевна, 369
Левко Лілія Русланівна, 51, 108
Левкович Мар'яна Володимирівна, 158
Ликтей Юрій Тарасович, 158
Лиликович Сергей Александрович, 385
Липак Галина Ігорівна, 214
Липчак Ілля Романович, 202
Литвиненко Артур Миколайович, 279
Литвинов Валерий Андроникович, 383
Лісун Андрій Григорович, 169
Ліщук Катерина Ігорівна, 110
Лозицький Олександр Анатолійович, 411
Лопатин Алнксей Константинович, 111
Любашенко Наталья Дмитриевна, 57
Любчик Леонид михайлович, 14
Лютенко Ирина Викторовна, 385
Магас Валентин Васильович, 386
Мазурік Олексій Юрійович, 113
Майстренко Светлана Яковлевна, 383
Малышева Милена Олеговна, 406
Марковський Олександр Петрович, 388
Марченко Олександр Іванович, 114, 389
Масліков Віталій, 280
Маслянко Павло Павлович, 117, 358
Матяш Владислав Вадимович, 92
Мацюк Олександр Васильович, 214, 215
Медведський Андрій Миколайович, 390
Мезенцев Микола Вікторович, 82
Мейтус Владимир Юльевич, 229, 231
Мигаль Валерий Павлович, 119
Мигаль Галина Валерьевна, 119
Миселюк Артур Юрійович, 233
Михалько Віталій Геннадійович, 392
Михняк Олексій Олександрович, 394
Мілявський Юрій Леонідович, 20
Молчанов Олексій Андрійович, 114

Морозов Константин Вячеславович, 396
Мосійчук Марія Дмитрівна, 121
Мурга Николай Алексеевич, 234
Мутталибова Шафаг Фируддин кызы, 338, 398
Мухін Вадим Євгенович, 182
Надточій Тарас Вікторович, 282
Небесний Руслан Михайлович, 214
Недашковская Надежда Ивановна, 123
Новицкая Екатерина Георгиевна, 236
Новогрудська Ріна Леонідівна, 209
Огурцов Максим Игоревич, 238
Олейник Владимир Витальевич, 405
Олефір Олександр Степанович, 121, 220, 260
Омельчук Антон Анатольевич, 124
Орєхов Олексндр Арсенійович, 126
Орєхова Наталія Артурівна, 126
Осадчий Дмитрий Юрьевич, 397
Осадчук Николай Павлович, 74
Павловський Володимир Ілліч, 418
Павлюк Юлія Русланівна, 127
Падалко Вадим Геннадійович, 105
Панкратов Владимир Андреевич, 129
Панкратова Наталия Дмитриевна, 136
Панкратова Наталія Дмитрівна, 131, 134, 135
Панкратова Наталья Дмитриевна, 132
Пасічник Володимир Володимирович, 214, 215, 411
Пашаев Намик Мухтар оглы, 398
Пашаева Малахат Мухтар кызы, 398
Перепеличный Ярослав Игоревич, 240
Петренко Анатолій Іванович, 17
Петренко Олексій Олексійович, 283
Печурин Николай Капитонович, 137
Печурин Сергей Николаевич, 137
Пилипенко Інна Віталіївна, 241
Писарчук Алексей Александрович, 77
Писарчук Олексій Олександрович, 79
Письменний Ігор Олександрович, 366
Пікож Андрій Віталійович, 138
Піпіч Артем Андрійович, 243
Поворознюк Анатолий Иванович, 244
Поворознюк Назар Іванович, 402
Поворознюк Оксана Анатольевна, 244
Погорелов Володимир Володимирович, 389
Погорілий Сергій Дем'янович, 142
Погребной Евгений Александрович, 140
Подладчиков Владимир Николаевич, 141, 367
Подладчіков Володимир Миколайович, 127, 207
Поливода Оксана Валериевна, 124
Попова Юлия Александровна, 141
Попович Євген Сергійович, 285
Поршнєва Катерина Олегівна, 246
Потапова Екатерина Романовна, 396
Потебня Артем Віталійович, 142
Примак Иван Константинович, 406
Притоманова Ольга Михайловна, 94
Прогонный Николай Витальевич, 422
Прогонов Дмитро Олександрович, 403
Прудський Юрій Юрійович, 117
Пудло Ілля Віталійович, 145
Путренко Віктор Валентинович, 147
Пховелишвили Мераб Гайозович, 324
Рагимов Рауф Махмуд оглы, 398
Ржепішевський Андрій Леонідович, 247
Роганов Егор Владимирович, 404
Романенко Віктор Демидович, 20
Романенков Юрий Александрович, 249
Романкевич Виталий Алексеевич, 405, 406
Романов Валерий Владимирович, 287
Романов Олексій Миколайович, 153
Романовський Олександр Віталійович, 408
Рубець Антон Вікторович, 410
Рубин Эдуард Ефимович, 149
Рудакова Анна Владимировна, 124
Савченко Ілля Олександрович, 151
Савченко Юлій Григорович, 332
Савчук Валерія Володимирівна, 411
Сало Андрій Миколайович, 408
Сальніков Микола Миколайович, 152
Саржинська Марина Сергіївна, 250
Сахаров Сергій Юрійович, 176
Сащук Ігор Миколайович, 153
Свістунов Сергій Якович, 290
Связінська Наталія Олександрівна, 364
Селищев Владислав Леонидович, 92
Семенов Сергій Васильович, 332
Сенченко В'ячеслав Родіонович, 209
Сергєєв Данило Сергійович, 413
Сергєєв Антон Валерійович, 416
Сирик Сергей Валентинович, 155
Ситников Валерий Степанович, 426
Ситникова Варвара Александровна, 426
Скірко Ілля Олегович, 252
Скотаренко Федір Миколайович, 84

Скукіс Олексій Євгенійович, 417
Слободюк Алексей Юрьевич, 253
Слободянюк Богдан Русланович, 418
Слухай Яна Олександрівна, 157
Сльота Максим Русланович, 134
Снігур Ольга Олексіївна, 224
Соколовський Ярослав Іванович, 158
Соловйов Володимир Миколайович, 162
Стасюк Андрій Вікторович, 164
Статкевич Виталий Михайлович, 50
Стельмах Роман Петрович, 158
Степаненко Альона Ігорівна, 165
Стефанишин Дмитро Володимирович, 167, 173
Стоянова Анастасия, 168
Стрельчук Елена Анатольевна, 361
Сукало Аліна Сергіївна, 327
Сулема Ольга Костянтинівна, 381
Супруненко Оксана Олександрівна, 169
Сутула Олександр Віталійович, 420
Сухомлин Ірина Борисівна, 421
Сухоручкина Ольга Николаевна, 422
Танырвердиев Чингиз Гариб оглы, 338
Таранян Арам Робертович, 254
Тарасенко Оксана Геннадіївна, 271
Тарасюк Тарас Сергійович, 135
Терентьєв Олександр Миколайович, 364
Тесленко Олександр Кирилович, 292
Тертишний Андрій Андрійович, 256
Тимчук Сергей Витальевич, 77
Тимчук Сергій Віталійович, 79
Томашевський Валентин Миколайович, 424
Торський Адріан Романович, 63
Тумко Володимир Степанович, 227
Ухина Анна Владимировна, 426
Филатова Анна Евгеньевна, 258
Фіногенов Олексій Дмитрович, 285
Харчун Володимир Михайлович, 171
Ходзинский Александр Николаевич, 238
Ходневич Ярослав Васильович, 173
Хоменко Ольга Володимирівна, 70
Хурцилава Константин Викторович, 383
Цверчкова Анна Александровна, 155
Цегелик Григорій Григорович, 174
Цепкова Дарья Павловна, 231
Чабаненко Дмитро Миколайович, 162
Черненко Ольга Борисовна, 111
Черненко Павло Романович, 357
Черный Юрий Юрьевич, 60
Чертов Олег Романович, 176
Четырбок Пётр Васильевич, 217
Чкалов Олексій Валерійович, 165
Шаповал Іван Іванович, 292
Шаронова Наталья Валеріївна, 178
Шевченко Дар'я Сергіївна, 180
Шевчик Михайло Олександрович, 428
Шелест Антон Викторович, 136
Шеренковський Артем Олегович, 429
Шецирули Лия Дмитриевна, 324
Шибін Олександр Володимирович, 402
Шишков Олег Николаевич, 431
Шліхтенко Неоніла Миколаївна, 433
Шпортько Леся Василывна, 434
Шпортько Олександр Володимирович, 434
Штельмах Сергій Олександрович, 260
Щодро Олексій Євгенович, 173
Юрченко Володимир Вікторович, 290
Юрченко Дмитро Володимирович, 176
Яковенко Мар'яна Миколаївна, 261
Яковлева Алла Петрівна, 182
Яковлєва Олександра Євгенівна, 262
Янковский Иван Павлович, 240
Яркун Володимир Ігорович, 158
Ярмак Дмитро Олегович, 433
Ясинський Володимир Кирилович, 83