

## СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

МІЖНАРОДНИЙ НАУКОВО-ТЕХНІЧНИЙ ЖУРНАЛ

№ 4

2015

ЗАСНОВАНО У ЛИПНІ 2001 р.

### РЕДАКЦІЙНА КОЛЕГІЯ:

#### Головний редактор

**М.З. ЗГУРОВСЬКИЙ**, акад. НАН України

#### Заступник головного редактора

**Н.Д. ПАНКРАТОВА**, проф., д.т.н., Україна

#### Члени редколегії:

**П.І. АНДОН**, акад. НАН України

**А.В. АНІСІМОВ**, чл.-кор. НАН України

**Г.-В. ВЕБЕР**, проф., Турція

**О.Д. ГВІШАНИ**, акад. РАН, Росія

**П.О. КАСЬЯНОВ**, проф., д.ф.-м.н.,  
Україна

**В.М. КУНЦЕВИЧ**, акад. НАН України

**О.А. ПАВЛОВ**, проф., д.т.н., Україна

**Т.Л. СААТІ**, проф., США

**Л. САКАЛАУСКАС**, проф., Литва

**А.М. САЛЕМ**, проф., Єгипет

**І.В. СЕРГІЄНКО**, акад. НАН України

**Х.-М. ТЕОДОРЕСКУ**, проф., Румунія

**Е.О. ФАЙНБЕРГ**, проф., США

**Я.С. ЯЦКІВ**, акад. НАН України

#### АДРЕСА РЕДАКЦІЇ:

03056, м. Київ,

пр. Перемоги, 37, корп. 35,

Навчально-науковий комплекс

«Інститут прикладного системного аналізу»

НТУУ «КПІ» МОН та НАН України

Тел.: **406-81-44**; факс: **406-81-44**

E-mail: [journal@mmsa.ntu-kpi.kiev.ua](mailto:journal@mmsa.ntu-kpi.kiev.ua)

<http://journal.iasa.kpi.ua>

### У номері:

• **Теоретичні та прикладні проблеми і методи системного аналізу**

• **Прогресивні інформаційні технології, високопродуктивні комп'ютерні системи**

• **Математичні методи, моделі, проблеми і технології дослідження складних систем**

• **Нові методи в системному аналізі, інформатиці та теорії прийняття рішень**

NATIONAL ACADEMY OF SCIENCES OF UKRAINE  
EDUCATIONAL-SCIENTIFIC COMPLEX  
«INSTITUTE FOR APPLIED SYSTEM ANALYSIS»  
AT THE NATIONAL TECHNICAL UNIVERSITY OF UKRAINE «KPI»  
OF MES AND NAS OF UKRAINE

## SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

INTERNATIONAL SCIENTIFIC AND TECHNICAL JOURNAL

№ 4

2015

IT IS FOUNDED IN JULY 2001

### EDITORIAL BOARD:

#### The editor – in – chief

**M.Z. ZGUROVSKY,** Academician of  
NASU

#### Deputy editor – in – chief

**N.D. PANKRATOVA,** Prof., Ukraine

#### Associate editors:

**F.I. ANDON,** Academician of  
NASU

**A.V. ANISIMOV,** Correspondent  
member of NASU

**E.A. FEINBERG,** Prof., USA

**A.D. GVISHIANI,** Academician of RAS,  
Russia

**P.O. KASYANOV,** Prof., Ukraine

**V.M. KUNTSEVICH,** Academician of NASU

**A.A. PAVLOV,** Prof., Ukraine

**T.L. SAATY,** Prof., USA

**L. SAKALAUSKAS,** Prof., Lithuania

**A.M. SALEM,** Prof., Egypt

**I.V. SERGIENKO,** Academician of NASU

**H.-N. TEODORESKU,** Prof., Romania

**G.-W. WEBER,** Prof., Turkey

**Ya.S. YATSKIV,** Academician of NASU

### THE EDITION ADDRESS:

03056, Kyiv,  
av. Peremogy, 37, building 35,  
Educational-Scientific Complex  
«Institute for Applied System Analysis» at the  
National Technical University of Ukraine «KPI»  
of MES and NAS of Ukraine  
Phone: 406-81-44; Fax: 406-81-44  
E-mail: journal@mmsa.ntu-kpi.kiev.ua  
<http://journal.iasa.kpi.ua>

### In the issue:

• **Theoretical and applied problems and methods of system analysis**

• **Progressive information technologies, high-efficiency computer systems**

• **Mathematical methods, models, problems and technologies for complex systems research**

• **New methods in system analysis, computer science and theory of decision making**

## Шановні читачі!

Навчально-науковий комплекс «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України видає міжнародний науково-технічний журнал

### «СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ».

Журнал публікує роботи теоретичного та прикладного характеру в широкому спектрі проблем, пов'язаних із системними дослідженнями та інформаційними технологіями.

#### Провідні тематичні розділи журналу:

Теоретичні та прикладні проблеми і методи системного аналізу; теоретичні та прикладні проблеми інформатики; автоматизовані системи управління; прогресивні інформаційні технології, високопродуктивні комп'ютерні системи; проблеми прийняття рішень і управління в економічних, технічних, екологічних і соціальних системах; теоретичні та прикладні проблеми інтелектуальних систем підтримки прийняття рішень; проблемно і функціонально орієнтовані комп'ютерні системи та мережі; методи оптимізації, оптимальне управління і теорія ігор; математичні методи, моделі, проблеми і технології дослідження складних систем; методи аналізу та управління системами в умовах ризику і невизначеності; евристичні методи та алгоритми в системному аналізі та управлінні; нові методи в системному аналізі, інформатиці та теорії прийняття рішень; науково-методичні проблеми в освіті.

**Головний редактор журналу** — ректор Національного технічного університету України «КПІ», академік НАН України Михайло Захарович Згуровський.

Журнал «Системні дослідження та інформаційні технології» включено до переліку фахових видань ВАК України.

Журнал «Системні дослідження та інформаційні технології» входить до таких наукометричних баз даних: «Україніка наукова» та УРЖ «Джерело», ВИНИТИ РАН, РИНЦ (Российский индекс научного цитирования), EBSCO, Google Scholar, IC Journals Master List, DOAJ (Directory of Open Access Journals), Index Copernicus.

Статті публікуються українською, російською та англійською мовами.

Журнал можна передплатити. **Наш індекс 23918.** Якщо ви не встигли передплатити журнал, його можна придбати безпосередньо в редакції.

Наша адреса: 03056, м. Київ, пр. Перемоги, 37, корп. 35, к. 219.

Завідуюча редакцією **С.М. Шевченко**

Редактор **А.О. Млечко**

Молодший редактор **Л.О. Тарин**

Комп'ютерна верстка та набір **А.А. Патіоха**

---

Підписано до друку 15.12.2015. Формат 70x108/16.

Папір офс. Офс. друк. Ум. друк. арк. 14. Ціна договірна.

Свідоцтво про реєстрацію КВ № 13804–2778 ПР від 17.03.2008 р.

---

Видавництво та поліграфія ВПК «Політехніка»

ВПІ НТУУ «КПІ»

03056, Київ–56, пр. Перемоги, 37, корп. 15

## **Dear Readers!**

The Educational-scientific complex «Institute for applied system analysis» at the National technical university of Ukraine «KPI» of MES and NAS of Ukraine is published of the international scientific and technical journal

### **«SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES».**

The Journal is printing works of a theoretical and applied character on a wide spectrum of problems, connected with system researches and information technologies.

#### **The main thematic sections of the Journal are the following:**

Theoretical and applied problems and methods of system analysis; theoretical and applied problems of computer science; automated control systems; progressive information technologies, high-efficiency computer systems; decision making and control in economic, technical, ecological and social systems; theoretical and applied problems of intellectual systems for decision making support; problem- and function-oriented computer systems and networks; methods of optimization, optimum control and theory of games; mathematical methods, models, problems and technologies for complex systems research; methods of system analysis and control in conditions of risk and uncertainty; heuristic methods and algorithms in system analysis and control; new methods in system analysis, computer science and theory of decision making; scientific and methodical problems in education.

**The editor-in-chief of the Journal** is Rector of the National Technical University of Ukraine «KPI», academician of the NASU Michael Zaharovich Zgurovsky.

The articles to be published in the Journal in Ukrainian, Russian and English languages are accepted. Information printed in the Journal is included in the Catalogue of periodicals of Ukraine.

# СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

4 • 2015

## ЗМІСТ

<b>ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ ПРОБЛЕМИ І МЕТОДИ СИСТЕМНОГО АНАЛІЗУ</b>	
<i>Путренко В.В.</i> Застосування трендових геопросторових статистичних моделей на основі інтелектуального аналізу даних .....	7
<b>ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, ВИСОКОПРОДУКТИВНІ КОМП'ЮТЕРНІ СИСТЕМИ</b>	
<i>Поліщук Д.О., Поліщук О.Д., Яджак М.С.</i> Комплексне детерміноване оцінювання складних ієрархічно-мережевих систем: III. Агреговане оцінювання .....	20
<i>Karayuz I.V., Bidiuk P.I.</i> Forecasting consumer price index in Ukraine with regression models and adaptive Kalman filter .....	32
<i>Прогонов Д.А., Куц С.Н.</i> Мультифрактальний флуктуационний аналіз стеганограмм .....	39
<i>Петренко О.О.</i> Порівняння типів архітектури систем сервісів .....	48
<i>Козуля Т.В., Шаронова Н.В., Білова М.О., Козуля М.М.</i> Впровадження компараторної ідентифікації для комплексної оцінки рівня безпеки об'єктів .....	63
<b>МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ, ПРОБЛЕМИ І ТЕХНОЛОГІЇ ДОСЛІДЖЕННЯ СКЛАДНИХ СИСТЕМ</b>	
<i>Марценюк В.П., Сверстюк А.С., Гвоздецька І.С.</i> Задача оптимального керування стадією елонгації полімеразно-ланцюгової реакції .....	75
<i>Yahanov P.O., Redko D.I., Redko I.V., Zakharchenko T.L.</i> Primitive programming algebra: General approach to a problem of functional completeness .....	83
<i>Пархомчук Д.М., Тимошенко Ю.О.</i> Аналітичний розв'язок некоректних задач динамічними методами .....	97
<i>Солошенко О.М.</i> Вдосконалені методи розрахунку статистики Колмогорова–Смирнова, ваги категорії змінної та значення інформації у кредитному рейтингу .....	104
<b>НОВІ МЕТОДИ В СИСТЕМНОМУ АНАЛІЗІ, ІНФОРМАТИЦІ ТА ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ</b>	
<i>Бондаренко В.Г., Трусковский К.К.</i> Хаотическая и случайная составляющие в природных временных данных .....	114
<i>Горбань И.И.</i> Почему точность измерения физических величин ограничена .....	123
Реферати .....	139
Відомості про авторів .....	146
Зміст журналу «Системні дослідження та інформаційні технології» за 2015 р. ....	148
Автори статей за 2015 р. ....	150

# SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

4 • 2015

## CONTENTS

<b>THEORETICAL AND APPLIED PROBLEMS AND METHODS OF SYSTEM ANALYSIS</b>	
<i>Putrenko V.V.</i> Using of the trend geospatial statistical models based on data mining .....	7
<b>PROGRESSIVE INFORMATION TECHNOLOGIES, HIGH-EFFICIENCY COMPUTER SYSTEMS</b>	
<i>Polishchuk D.O., Polishchuk O.D., Yadzhak M.S.</i> Complex deterministic evaluation of the complex hierarchical-network systems: III. Aggregative evaluation ...	20
<i>Karayuz I.V., Bidyuk P.I.</i> Forecasting consumer price index in Ukraine with regression models and adaptive Kalman filter .....	32
<i>Progonov D.A., Kushch S.N.</i> Multifractal detrended fluctuation analysis of steganograms .....	39
<i>Petrenko O.O.</i> A comparison of architecture types of services .....	48
<i>Kozulia T.V., Sharonova N.V., Bilova M.O., Kozulia M.M.</i> Comparative identification introduction for integrated assessment of object security level .....	63
<b>MATHEMATICAL METHODS, MODELS, PROBLEMS AND TECHNOLOGIES FOR COMPLEX SYSTEMS RESEARCH</b>	
<i>Martseniuk V.P., Sverstiuk A.S., Gvozdetska I.S.</i> Optimal control problem of the elongation stage in the polymerase chain reaction .....	75
<i>Iaganov P.O., Redko D.I., Redko I.V., Zakharchenko T.L.</i> Primitive programming algebra: General approach to a problem of functional completeness .....	83
<i>Parkhomchuk D.M., Tymoshenko Yu.O.</i> Analytic solution of ill-posed problems via dynamic methods .....	97
<i>Soloshenko O.M.</i> Improved estimation methods of the Kolmogorov-Smirnov statistic, weight of evidence and information value indicators in the credit scoring .....	104
<b>NEW METHODS IN SYSTEM ANALYSIS, COMPUTER SCIENCE AND THEORY OF DECISION MAKING</b>	
<i>Bondarenko V.G., Truskovsky K.K.</i> The chaotic and random components in time series data .....	114
<i>Gorban I.I.</i> Why the measurement precision of real physical variables is limited .....	123
References .....	139
Information about the authors .....	146
Contents of journal «System Research & Information technologies» over 2015 .....	148
Information about the authors over 2015 .....	150

**ЗАСТОСУВАННЯ ТРЕНДОВИХ ГЕОПРОСТОРОВИХ  
СТАТИСТИЧНИХ МОДЕЛЕЙ НА ОСНОВІ  
ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ**

**В.В. ПУТРЕНКО**

Досліджено застосування трендових геопросторових статистичних моделей в аналізі даних сталого розвитку України. Розглянуто математичний апарат, який використовується під час побудови трендових геопросторових статистичних моделей. На прикладі даних ієрархічної моделі сталого розвитку проведено дослідження геопросторових трендів та природи їх виникнення шляхом побудови поліноміальних регресійних моделей, що базуються на обробці геопросторово прив'язаних даних. Розроблено методику детекції геопросторових трендів в ієрархічних моделях даних на основі картографічних методів та поліномів першого і другого порядку. Запропоновано використання картометричних методів для оцінки параметрів геопросторових трендів. Обґрунтовано використання методів побудови геопросторових трендів як складової інтелектуального аналізу даних.

**ВСТУП**

Накопичення значних обсягів даних спонукало до розвитку напряму інтелектуального аналізу, який спрямовано на виявлення прихованих закономірностей та залежностей у великих сховищах даних. Одним із напрямів інтелектуального аналізу даних є інтелектуальний аналіз геопросторових даних (ІАГД), який базується на використанні інформації про геопросторове розташування об'єктів у локальних та глобальних системах координат. Серед широкого кола методів обробки та аналізу даних, які використовуються в інтелектуальному аналізі даних значне місце посідають методи трендового аналізу. Найбільший розвиток методи статистичного трендового аналізу здобули під час аналізу часових рядів та побудові прогнозних моделей динаміки явищ. Просторові тренди вже були достатньо добре вивчені в рамках статистичної обробки інформації, відображення багатомірних просторів та зменшення розмірності даних.

Проте дослідження геопросторових статистичних трендів найбільш активно почалось разом із розвитком геостатистичних методів у складі інтелектуального аналізу геопросторових даних. Основними складностями у вивченні геопросторових трендів були проблеми з накопиченням геопросторових даних, складності інтерпретації трендової моделі та значні можливості виникнення статистичного шуму, пов'язаного з особливостями

навколишнього середовища та похибок вимірів. Найбільших успіхів у дослідженні геопросторових трендів було досягнуто в метеорології та кліматології, що дозволило з'ясувати глобальні тренди у зміні температур та властивостей повітряних мас. Сьогодні моделі трендового аналізу є корисними у вивченні геологічних структур, екологічних проблем, соціально-економічних явищ тощо.

## ОГЛЯД ЛІТЕРАТУРИ

Особливістю використання трендової моделі є глобальність та абстрактність отриманих результатів оцінювання, що може ускладнювати процес їх інтерпретації для підтримки прийняття рішень. З іншого боку, результати аналізу трендової моделі можуть бути основою для виявлення глибинних зв'язків, які існують між даними.

У галузі інтелектуального аналізу даних в Україні широко відомі роботи Інституту прикладного системного аналізу НТУУ «КПІ», Світового центру даних з геоінформатики та сталого розвитку М.З. Згуровського, Н.Д. Панкратової, А.І. Петренка, А.О. Болдака [1–4]. Питання використання трендових геопросторових моделей розглядалось у роботах Ашкана Зарнані, Масуда Рахгозара, Сімона Блекмора, Шаші Шекшара, Пушенга Занга, Яна Хуанга, Ранга Раю Ватсавайя [5–8].

**Метою** роботи є виявлення та дослідження трендових геопросторових статистичних моделей в інтелектуальному аналізі складових моделі сталого розвитку України.

**Завданням** статті є визначення методологічних основ використання трендового аналізу геопросторових даних, математичного апарату побудови трендових поверхонь, аналіз за допомогою виявлення геопросторових трендів даних щодо сталого розвитку України та їх інтерпретації.

## ГЛОБАЛЬНІ МЕТОДИ ІНТЕРПОЛЯЦІЇ ТРЕНДОВИХ ПОВЕРХОНЬ

*Глобальні методи інтерполяції* одночасно використовують всі наявні дані для виконання прогнозу для всієї території, тоді як локальні методи оперують у межах невеликих зон навкруги належних інтерполяції вузлів для того, щоб забезпечити виконання оцінки тільки за даними, розміщеними в безпосередній близькості від точок прогнозу або оцінювання.

Глобальні інтерполяції, як правило, використовуються не для безпосередньої інтерполяції, а для дослідження і можливого видалення ефекту глобальних варіацій (тренду), обумовлених зовнішніми чинниками. Після того, як глобальні ефекти будуть видалені, відхилення від глобальних варіацій можуть інтерполюватися з використанням локальних методів.

Глобальні методи є простими для обчислення і часто базуються на стандартних статистичних прийомах варіаційного аналізу і регресії. До них відносять [9]:

- класифікації з використанням зовнішньої інформації;
- поліноміальну регресію з геометричними координатами;
- регресійні моделі.



Класифікаційні методи використовують досягну інформацію (таку, як ґрунтові типи або адміністративні території) для того, щоб поділити досліджувану територію на регіони, які можуть бути охарактеризовані статистичними моментами (середнім, дисперсією) атрибутів, виміряних у точках, розміщених у межах цих регіонів.

Поліноміальні інтерполятори відображають значення у точці у вигляді поліному від координат. У двовимірному випадку — для точки  $x$  із координатами  $(x, y)$   $Z^*(x, y) = P_n(x, y)$ , де  $P_n$  — полином  $n$ -го ступеня. Зазвичай на практиці для двовимірного випадку використовують один з чотирьох типів поліномів:

- площина:  $P_1(x, y) = a + bx + cy$ ;
- білінійно-сідловий:  $P_{1,5}(x, y) = a + bx + cy + dxy$ ;
- квадратичний:  $P_2(x, y) = a + bx + cy + dxy + ex^1 + fy^2$ ;
- кубічний:  $P_2(x, y) = a + bx + cy + dxy + ex^2 + fy^2 + gx^2y + ix^3 + jy^3$ .

Теоретично можна використовувати і поліноми більш високого порядку. Вони визначаються максимальним ступенем для  $x$ , максимальним ступенем для  $y$  і спільним максимальним ступенем. Всі проміжні ступені в поліномі будуть наявними.

Таким чином, горизонтальна поверхня має порядок нуль, похила плоска поверхня - перший порядок, квадратична поверхня — другий порядок, кубічна поверхня з десятьма коефіцієнтами має третій порядок. Знаходження коефіцієнтів  $a_i$  є стандартною процедурою в задачах на множинну регресію, тому обчислення легко виконуються за допомогою стандартних статистичних пакетів. Після знаходження коефіцієнтів  $a_i$  трендову поверхню може бути відображено оцінкою значень  $Z(x, y)$  у всіх точках регулярної мережі. Порядок трендової поверхні  $P$  із формальної точки зору може бути скільки завгодно великим. Збільшення порядку полінома, як правило, призводить до збільшення ступеня його відповідності реальній поверхні. Проте при цьому, з одного боку, втрачається фізичне значення побудови, з іншого — різко збільшуються вимоги до обсягу даних, необхідних для знаходження коефіцієнтів  $a_i$ .

Завдання поліноміальної інтерполяції зводиться до того, щоб визначити невідомі коефіцієнти  $a_i$  так, щоб поліноми максимально добре відповідали даним у заданих точках. Для цього знаходять мінімум за всіма коефіцієнтами ( $a, b, c, d$  тощо) функції  $\chi^2$ , яка задає інтегральну помилку інтерполяції (1) та обумовлена наступним чином:

$$\chi^2 = \sum_{i=1}^N [Z(x_i, y_i) - P_n(x_i, y_i)]^2. \quad (1)$$

Мінімізація полягає у вирішенні системи лінійних рівнянь з числом невідомих, що дорівнює кількості рівнянь. Число рівнянь (невідомих) залежить від обраного полінома.

Глобальний поліноміальний метод не є інтерполятором у строгому сенсі, швидше він відноситься до апроксиматорів. Його можна використовувати для виділення великомасштабного тренда.

Можна скористатися і локальним варіантом поліноміального методу, коли пошук коефіцієнтів проводиться тільки на основі даних, які потрапили в зону пошуку.

## МЕТОДОЛОГІЯ ОЦІНЮВАННЯ ТА АНАЛІЗУ СТАЛОГО РОЗВИТКУ

Методологія оцінювання та аналізу сталого розвитку, яку викладено в роботах [1,2] і використано в цьому дослідженні, містить модель сталого розвитку, яка є міждисциплінарним узагальненням моделей, відомих у природничих, економічних та соціальних галузях науки, та методику застосування формальних статистичних методів та методів експертного оцінювання для аналізу процесів сталого розвитку.

$$\{Q\} = j w_{sl} C_{sl} + w_{ql} \overline{C_{ql}}(I_{ec}, I_e, I_s). \quad (2)$$

Кватеріон  $\{Q\}$  містить уявну зважену скалярну частину  $j w_{sl} C_{sl}$ , яка описує безпеку життя людей і зважену дійсну векторну частину, яка описує якість життя людей у просторі трьох вимірів: економічного ( $I_e$ ), екологічного ( $I_{ec}$ ) і соціально-інституціонального ( $I_s$ ). Вагові коефіцієнти  $w_{sl}$  й  $w_{ql}$  в формулі (2) використано з метою вирівнювання масштабів компонент безпеки і якості життя.

З метою кількісного оцінювання вимірів сталого розвитку використовуємо принципи побудови ієрархічної системи показників та індексів, які визначаються як  $L_1$ -норми:

$$I_i = \sum_{j=1}^n w_j x_{i,j}, \quad i = \overline{1, m}, \quad \sum_{j=1}^n w_j = 1 \quad (3)$$

у просторі показників  $X^1 \times X^2 \times \dots \times X^m$ , що характеризують економічний, екологічний та соціально-інституціональний розвиток кожного  $i$ -ого регіону. Вагові коефіцієнти  $w_j$  в формулі (3) визначаються експертним шляхом.

Значення компоненти безпеки життя  $C_{sl}$  для  $j$ -ого регіона визначається як норма Мінковського вектора  $\vec{S}_j = (s_i^j)$ ,  $s_i^j = 1 - t_i^j$ ,  $i = \overline{1, n}$ :

$$C_{sl} = \|\vec{S}_j\| = \left( \sum_{i=1}^n (s_i^j)^p \right)^{\frac{1}{p}} \quad (4)$$

з параметром  $p = 3$ .

## АНАЛІЗ ГЕОПРОСТОРОВИХ ТРЕНДІВ ДАНИХ СТАЛОГО РОЗВИТКУ

Для проведення дослідження було використано набір даних, який описує сталий розвиток України в 2013 році Світового центру даних з геоінформатики та сталого розвитку [1]. З метою перевірки гіпотези щодо існування просторового тренду в розподілі даних про сталий розвиток було використано спочатку картографічний метод, який дозволяє оцінити загальний пат-

терн просторового розподілу значень. З цією метою було побудовано ряд тематичних карт значень індексу сталого розвитку та його складових за областями на основі поділу на п'ять груп за природними межами класифікації значень. Аналіз карт дозволяє стверджувати про наявність певної зміни у розподілі значення індексу сталого розвитку у географічному просторі. Для підтвердження цього припущення використано метод побудови поліноміальної трендової поверхні, яку було побудовано на основі точкового файлу, який містить дані про геометричні центри областей України. У результаті проведення експериментальних побудов поверхонь було з'ясовано, що практичне значення у виявленні трендів у даних мають поверхні утворені поліномами 1–3 ступенів. Якщо поліноми першого ступеня допомагають встановити загальні характеристики тренду даних, то поліноми 2–3 ступенів дозволяють підтвердити існування просторової тенденції у даних та визначити загальні риси її розподілу.

Аналіз поліноміальної поверхні 1, 2 порядків для індексу сталого розвитку дозволяє впевнено стверджувати про існування географічного тренду в розподілі індексу. Значення індексу мають загальну тенденцію до зменшення з Північного Заходу на Південний Схід. Звичайно, у цьому розподілі даних існують відхилення, які пов'язані, наприклад, з високим значенням індексу сталого розвитку для м. Києва, але вони тільки підкреслюють наявність певного географічного тренду (рис. 1).

Допоміжним інструментом аналізу тренду є графік аналізу тренду, який у площинах північ – південь, захід – схід демонструє наявність трендів у розподілі даних, які у вигляді результуючого тренду визначаються на трендовій поверхні (рис. 2). Таким чином, можна зробити припущення, що західні області України мають тенденцію до більш сталого розвитку, ніж області сходу.

Трендові поверхні можуть бути оцінені як за допомогою значень коефіцієнтів, які використовувались у поліномі, так і описом фізичних характеристик трендової поверхні. Такими характеристиками є ухил поверхні, кут між лініями скату поверхонь першого та другого поліномів, експозиція схилів. Кути нахилу прямо пропорційно та кути між лініями скату обернено пропорційно вказують на наявність вираженого тренду в даних. Кореляційний аналіз між цими показниками вказує на існування тісного зв'язку, який за індексом кореляції Пірсона дорівнює 0,816. Додатковими інструментами аналізу можуть виступати локальні методи інтерполяції даних, такі як обернено-зважені відстані та розподіл ухилів отриманих статистичних поверхонь.

Інтерпретація виявлених просторових трендів завжди пов'язана зі складнощами суб'єктивного та об'єктивного характеру, які викликані пошуком причинно-наслідкових зв'язків. У випадку аналізу просторового тренду у розподілі індексу сталого розвитку можна виходити із ієрархічної структури формування індексу, де підсумкове значення залежить від безпеки і якості життя, які, в свою чергу, складаються з екологічного, економічного та соціального вимірів та сукупності загроз, які впливають на безпеку життя. З цією метою було проведено відповідний аналіз на виявлення трендів для індексу якості та безпеки життя та трьох вимірів, які є складовими якості життя.

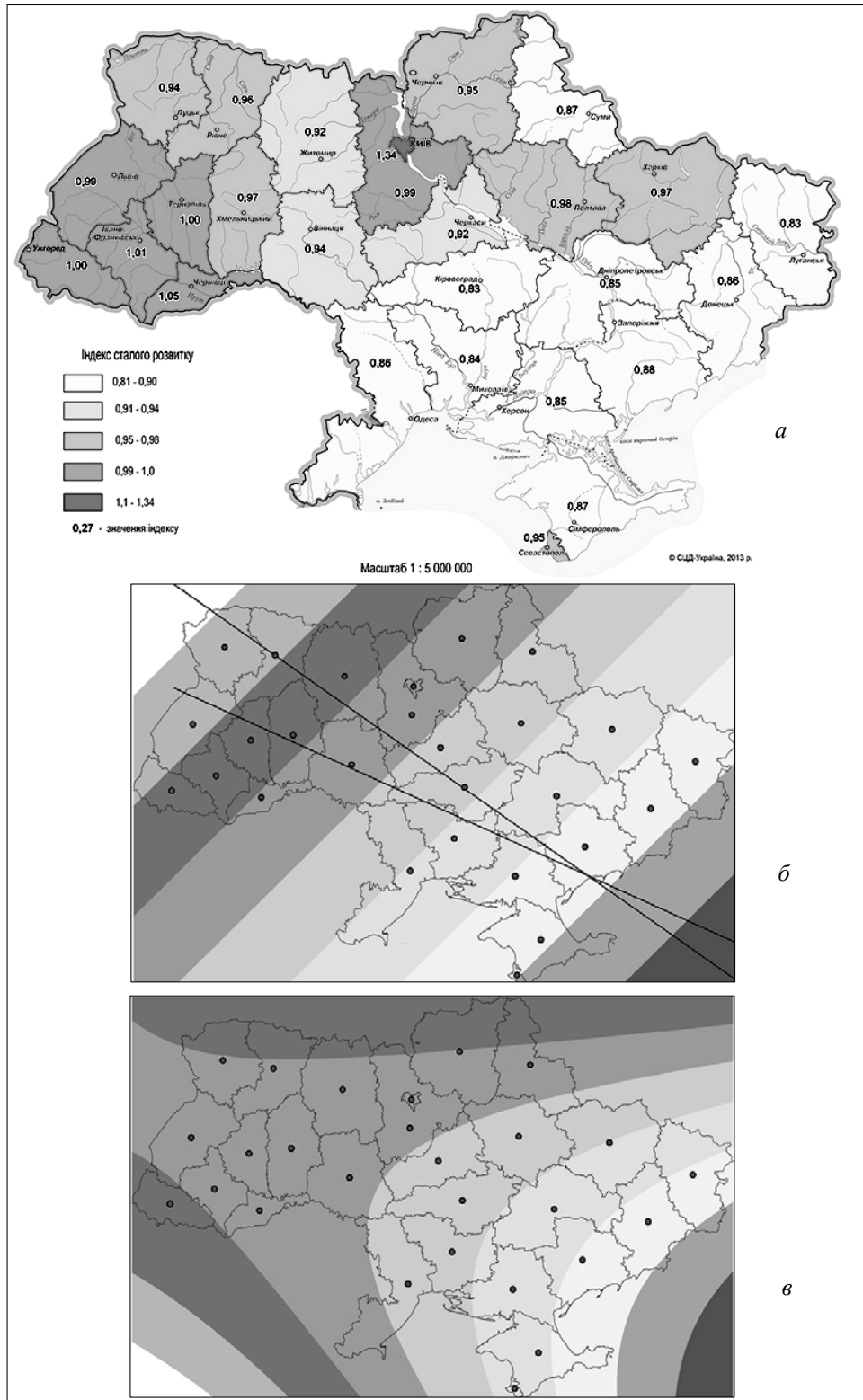


Рис. 1. Розподіл трендових моделей для індексу сталого розвитку: *а* — картограма за регіонами; *б* — поліном першого порядку; *в* — поліном другого порядку

**Таблиця 1.** Картометричні показники трендових поверхонь індексів

Індекс	Кут нахилу	Кут між лініями скату
$I_s$	0,000004	15,3142
$I_e$	0,000005	7,6633
$C_{ql}$	0,000009	23,3202
$I_{sd}$	0,000011	9,1149
$C_{sl}$	0,000014	27,1387
$I_{ec}$	0,000018	54,5763

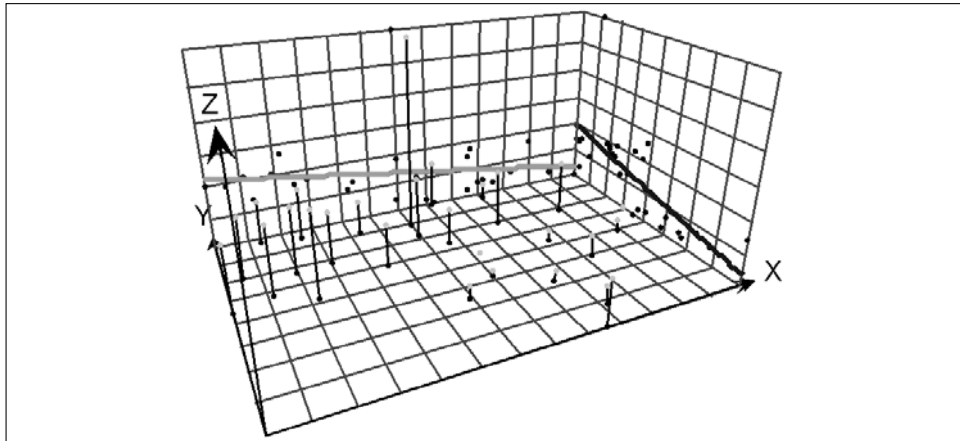


Рис. 2. Графік аналізу трендової поверхні індексу сталого розвитку

Індекс безпеки життя розраховується на підставі даних про 12 ключових загроз, що несуть суспільну небезпеку. Побудовані за його даними поліноміальні поверхні першого та другого порядку вказують на чітку наявність тренду, який дуже тісно збігається з трендом індексу сталого розвитку. Тренд характеризується чітким напрямом із заходу на схід із трохи більшим кутом між лініями скату, ніж у випадку індексу сталого розвитку (рис. 3).

Індекс якості життя, який утворено на основі економічного, соціального та екологічного вимірів характеризується трендом з півночі на південь країни для поліному 1 ступеня та трендом з північного заходу на південний схід для поліному 2 ступеня (рис. 4). При цьому трендова поверхня 2 ступеня характеризується значним плато у центральній частині України, де вплив тренду у даних нівелюється іншими факторами. У результаті аналізу двох складових сталого розвитку можна зробити висновки, що для даних індексів якості та безпеки життя тренди також простежуються. При чому індекс безпеки життя має більш виражений тренд даних, ніж індекс якості життя.

Аналіз складових якості життя дозволяє виявити ту групу показників, яка найбільше впливає на формування тренду в даних. Екологічний вимір якості життя має виражений тренд із заходу на схід країни, який візуально видно на трендових поверхнях першого та другого ступенів (рис. 5а). Це закономірно свідчить про те, що на заході країни більш сприятлива екологічна ситуація, яка пов'язана з відсутністю у більшості випадків великих джерел антропогенного забруднення та кращим забезпеченням водними ресурсами.

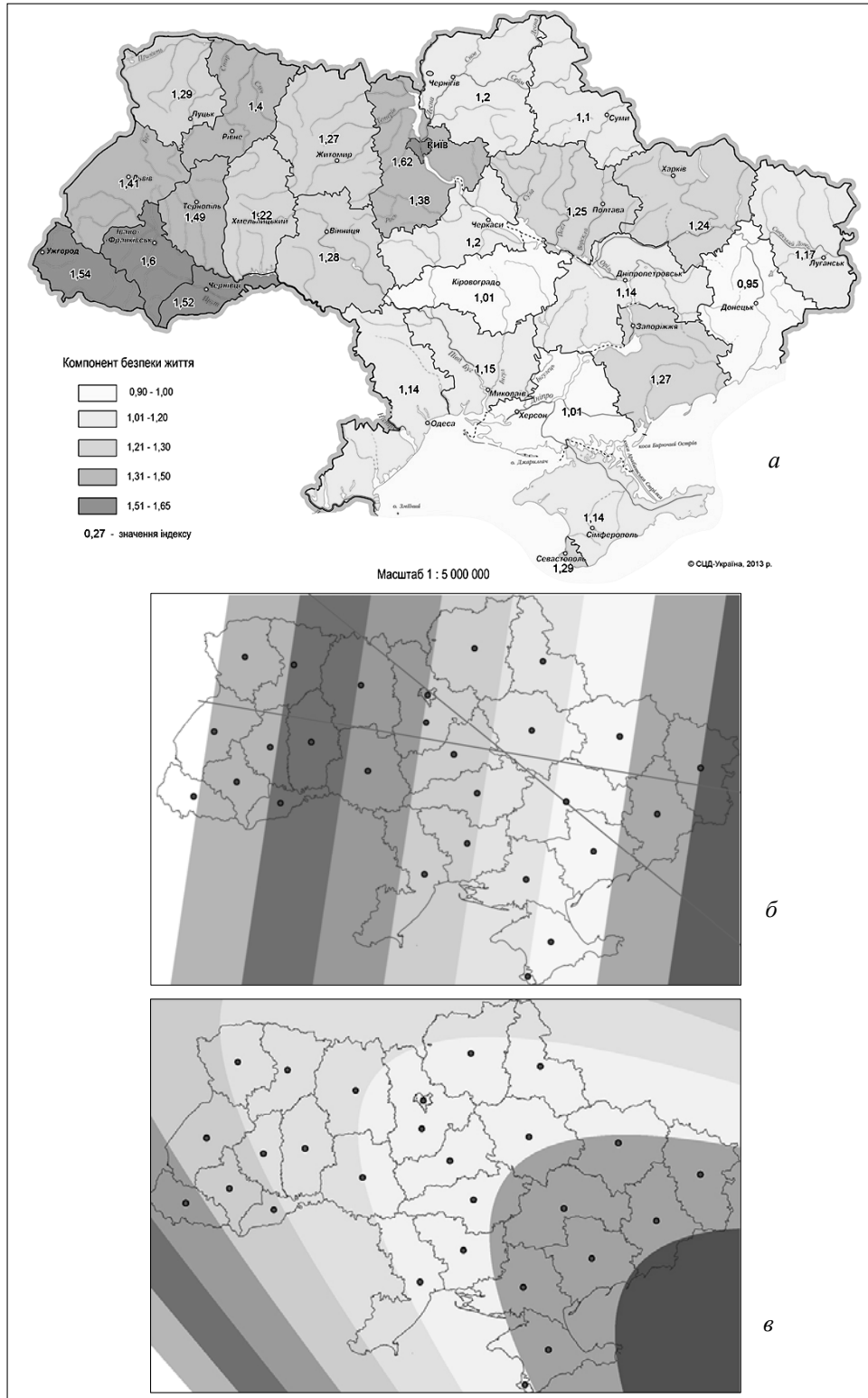


Рис. 3. Розподіл трендових моделей для індексу безпеки життя: а — картограма за регіонами; б — поліном першого порядку; в — поліном другого порядку

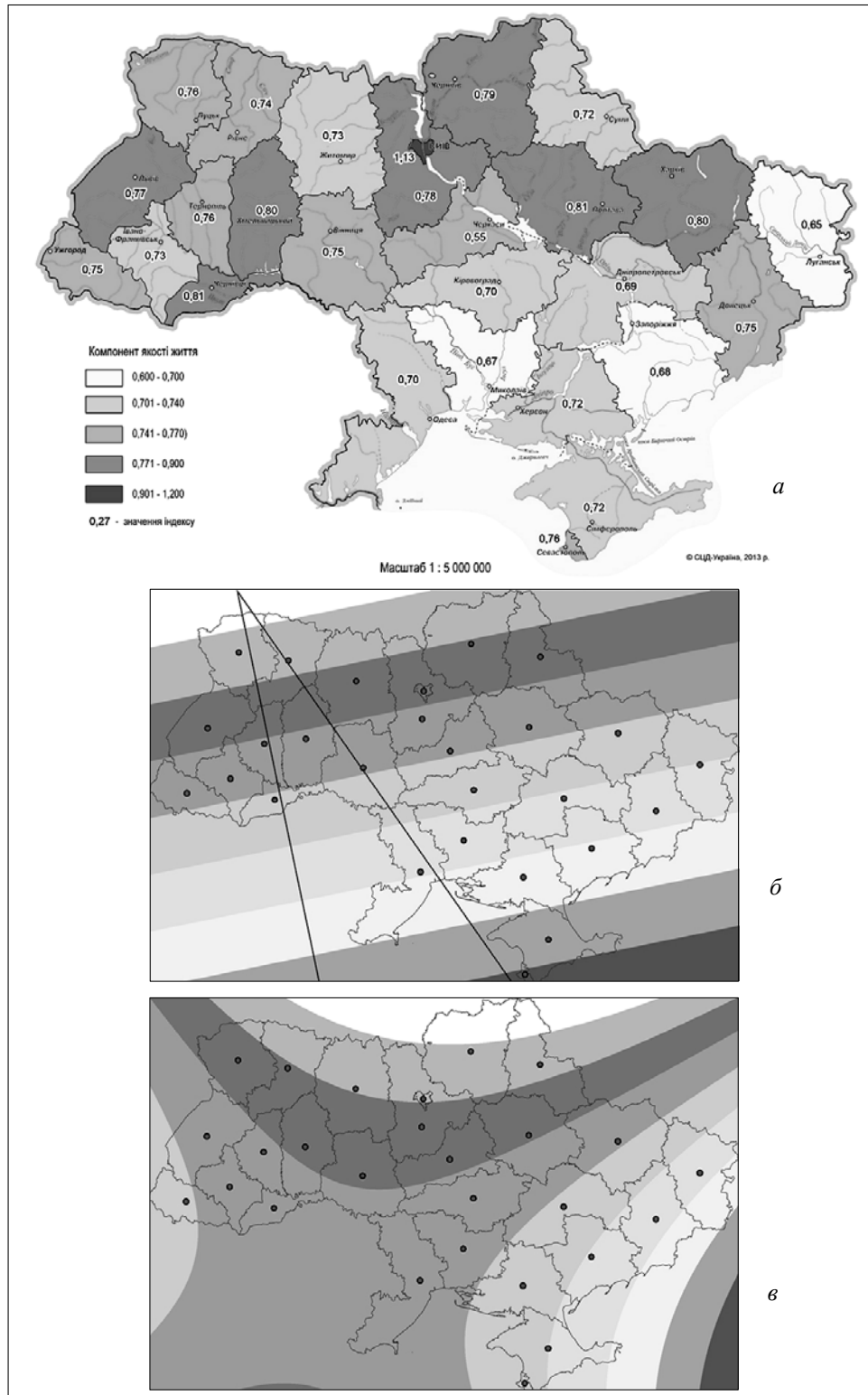


Рис. 4. Розподіл трендових моделей для індексу якості життя: а — картограма за регіонами; б — поліном першого порядку; в — поліном другого порядку

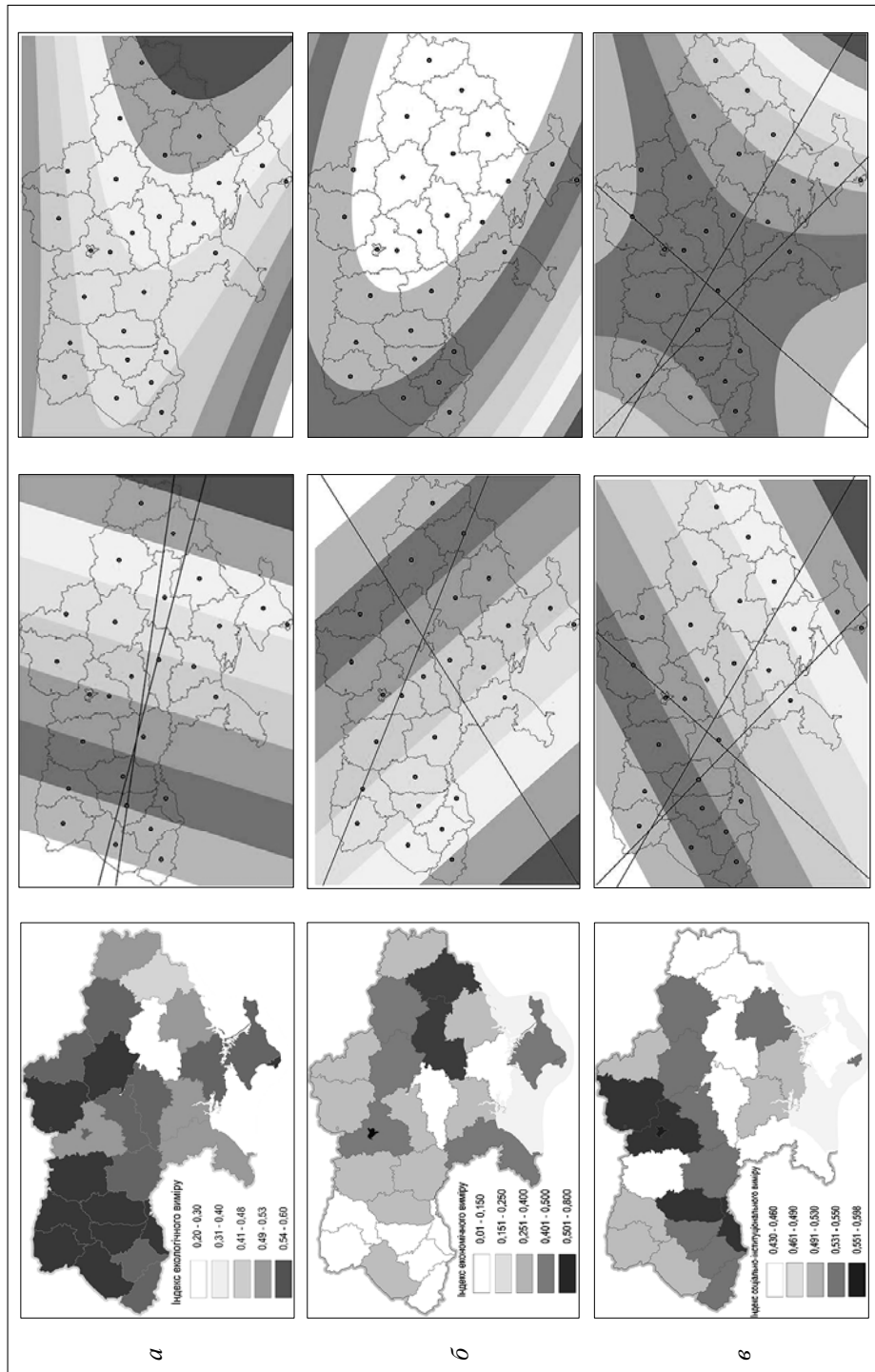


Рис. 5. Розподіл трендових моделей для: а — екологічного виміру; б — економічного виміру; в — соціально-інституціонального виміру

Трендові поверхні індексу економічного виміру демонструють зміну з південного заходу на північний схід, що пов'язано зі збільшенням економічних показників на сході країни, де сконцентровано більші виробничі потужності. Проте у випадку аналізу трендових поверхонь першого та другого ступенів існують певні розбіжності. Значною мірою це викликано впливом



столичного статусу Києва, який дає найбільший внесок в економічний розвиток (рис. 5б).

Аналіз третього соціального виміру в структурі індексу якості життя населення показав відсутність єдиного тренду в даних соціальної складової для території України. Це демонструє досить строкату структуру соціального розвитку України, яка має плато значень у центральній частині України та два тренди до зменшення показників на північному заході та південному сході країни (рис. 5в). Таким чином, гіпотеза про наявність вираженого географічного тренду в даних про соціальний розвиток не пройшла перевірку. У даних існують тільки окремі тенденції, які можуть бути визначені за допомогою локальних методів інтерполяції.

У результаті можна зазначити, що аналіз складових індексу якості життя населення показує достатньо різноманітну структуру глобальних та локальних географічних трендів в структурі даних. Найбільш виражений тренд наявний в даних екологічного виміру, що пов'язано з особливостями природокористування в різних регіонах країни. Більш складну структуру має трендова поверхня другого ступеня для економічного виміру, яку детерміновано регіональними диспропорціями в економічному розвитку та промисловому виробництві. При цьому дані екологічного та економічного вимірів мають протилежні за напрямом змін тренди. Перевірка соціального виміру на наявність глобальних трендів показала, що для соціального виміру в Україні глобальні просторові тренди відсутні і є тільки локальні регіональні тенденції, які можуть бути описані за допомогою локальних методів інтерполяції.

## МЕТОДИ ДЕТЕКЦІЇ ГЕОПРОСТОРОВИХ ТРЕНДІВ В ІЄРАРХІЧНИХ МОДЕЛЯХ ДАНИХ

На основі картографічних методів аналізу встановлено ступінь впливу складових індексу сталого розвитку на формування тренду інтегрального показника. Табл. 2 демонструє сукупний аналіз складових індексу сталого розвитку з використанням трьох методів: картографічний, трендова поверхня 1 ступеня, трендова поверхня 2 ступеня. Усю сукупність станів розподілу даних зведено до розмірності  $f(x) \in [0; 1]$ , де  $f(x)$  — функція наявності тренду даних. У цьому діапазоні  $y = 0 \Rightarrow f(x) = 0$ ,  $y = 0,5 \Rightarrow f(x) \approx 1$ ,  $y = 1 \Rightarrow f(x) = 1$ .

Це означає, що значення 0 вказує на відсутність глобального географічного тренду, 0,5 — можливість наявності анізотропії у розподілі географічних даних, 1 вказує на виражений географічний тренд у даних.

Обрахування середніх значень для складових індикаторів за формулою

$$T = \frac{\sum_{n=1}^n y}{n}, \quad (5)$$

де  $T = \{I_{sd}, C_{sl}, C_{ql}\}$ ,  $y$  — вірогідність наявності просторового тренду даних,  $n$  — кількість складових індикаторів тренду. У результаті отримано значення для індексу сталого розвитку, безпеки та якості життя, відображені в табл. 3.

**Таблиця 2.** Визначення географічного тренду у розподілі даних

Показник	Картографічний метод	Поліном 1 ступеня	Поліном 2 ступеня
Індекс сталого розвитку	1	1	1
Якість життя	0,5	1	1
Економічний вимір	0,5	1	0,5
Екологічний вимір	1	1	1
Соціальний вимір	0,5	1	0
Безпека життя	1	1	1
Злочинність	1	1	1
Зниження тривалості життя	0,5	1	0,5
Корупція	0	1	0
Соціальна нерівність	0	1	0
Рівень безробіття	1	1	1
Зношеність технологічної інфраструктури	0,5	1	1
Техногенна небезпека	0,5	1	0
Зниження добробуту населення	1	1	1
Забруднення довкілля	1	1	0,5
Дефіцит водних ресурсів	1	1	1
Смертність населення	1	1	1
Вплив на змінювання клімату	1	1	0,5

**Таблиця 3.** Порівняльний аналіз вірогідності наявності просторового тренду

Назва індексу	Середній вклад тренду при картографічному аналізі	Середній вклад тренду при поліномі 2 порядку
$I_{sd}$	0,72	0,66
$C_{sl}$	0,70	0,625
$C_{ql}$	0,66	0,50

З табл. 3 можна побачити, що індекс сталого розвитку має найбільші значення присутності тренду, а значення для індексу безпеки життя перевищують відповідні показники для індексу якості життя за обома видами аналізу, що вказує на переважання просторових тенденцій у даних індексу безпеки життя. У структурі індексу безпеки життя найбільший внесок до складових тренду роблять такі соціальні загрози як злочинність, рівень безробіття, зниження добробуту населення та смертність населення; економічний фактор зношеності технологічної інфраструктури та екологічні фактори дефіциту водних ресурсів, забруднення довкілля та впливу на зміну клімату. Наявність геопросторових трендів у цих даних пов'язано зі зміною природних умов по території України та диспропорціями регіонального розвитку в економічному та соціальному аспектах.

## ВИСНОВКИ

Трендовий аналіз є важливим інструментом інтелектуального аналізу геопросторових даних, який ґрунтується на використанні глобальних методів інтерполяції даних на основі регресійних моделей. Прикладне значення трендового аналізу полягає у можливості виявлення просторових закономірностей

у розподілі даних та встановлення загального закону їх розподілу. Експериментальне застосування моделей трендового аналізу для ієрархічної математичної моделі сталого розвитку регіонів України довело доцільність застосування геопросторових трендів під час аналізу комплексних соціальних, економічних та екологічних явищ. Було встановлено доцільність застосування для виявлення просторових тенденцій трендових поверхонь 1–3 порядків. Специфіка застосування трендових поверхонь під час аналізу геопросторових даних полягає у можливості використання додаткових методів аналізу отриманих поверхонь на основі загальноприйнятих картометричних методів визначення куту нахилу поверхонь та експозиції схилів.

Трендовий аналіз даних сталого розвитку дозволив встановити наявність загального тренду у розподілі показників індексу сталого розвитку на території України, який у подальшому може бути розкладено за складовими ієрархічної моделі. У результаті було встановлено значний вплив екологічних факторів на формування геопросторового тренду даних, а також таких соціальних загроз як злочинність, рівень безробіття, зниження добробуту та смертність населення. Використання математичного апарату для порівняння внеску складових сталого розвитку у тренд даних дозволило встановити переважання індексу безпеки життя.

Подальші дослідження можуть бути пов'язані з дослідженням інтеграції трендових моделей до системи підтримки прийняття рішень на базі ІАГД.

## ЛІТЕРАТУРА

1. Згуровський М.З., Акімова О.А., Болдак А.О., Вавулін П.Л., Джигирей І.М., Єфремов К.В., Путренко В.В. Аналіз сталого розвитку: глобальний і регіональний контексти: монографія / Міжнар. рада з науки (ICSU) та ін.; наук. кер. проекту М.З. Згуровський. — К.: НТУУ «КПІ», 2014. — Ч. 2. Україна в індикаторах сталого розвитку (2013). — 172 с.
2. Згуровський М.З., Болдак А.О., Єфремов К.В. Интеллектуальный анализ и системное согласование научных данных в междисциплинарных исследованиях // Кибернетика и системный анализ. — 2013. — № 4. — С. 62–75.
3. Згуровский М., Панкратова Н. Системный анализ: Проблемы, методология, приложения. — К.: Наук. думка, 2005. — 743 с.
4. Петренко А.І. Grid та інтелектуальна обробка даних Data Mining // Систем. дослідж. та інформ. технології. — 2008. — № 4. — С. 97–110.
5. Ashkan Zarnani, Masoud Rahgozar Mining Spatial Trends by a Colony of Cooperative Ant Agents // Spatial Data Mining: Consolidation and Renewed Bearing, April 22, 2006. — P. 64–72.
6. Blackmore B.S. The interpretation of trends from multiple yield maps // Computers and Electronics in Agriculture. — 2000. — 26(1). — P. 37–51. — [http://www.researchgate.net/publication/223399377\\_The\\_interpretation\\_of\\_trend\\_from\\_multiple\\_yield\\_maps](http://www.researchgate.net/publication/223399377_The_interpretation_of_trend_from_multiple_yield_maps).
7. Blackmore I Simon, Godwin Richard J., Fountas Spyros The Analysis of Spatial and Temporal Trends in Yield Map Data over Six Years Biosystems Engineering (2003) 84 (4). — P. 455–466.
8. Shekhar Shashi, Zhang Pusheng, Huang Yan, Vatsavai Ranga Raju. Trends in Spatial Data Mining, as a chapter to appear in Data Mining: Next Generation Challenges and Future Directions, H. Kargupta, A. Joshi, K. Sivakumar, and Y. Yesha (eds.), AAAI/MIT Press, 2003. — [http://www.spatial.cs.umn.edu/paper\\_ps/dmchap.pdf](http://www.spatial.cs.umn.edu/paper_ps/dmchap.pdf).
9. Burrough P.A., McDonnel, R.A. Principles of Geographical Information Systems for land resources assessment. — Oxford University Press, 1998. — 333 p.

*Поступила 15.05.2015*

УДК 519.6+625.1

**КОМПЛЕКСНЕ ДЕТЕРМІНОВАНЕ ОЦІНЮВАННЯ  
СКЛАДНИХ ІЄРАРХІЧНО-МЕРЕЖЕВИХ СИСТЕМ.  
ІІІ. АГРЕГОВАНЕ ОЦІНЮВАННЯ**

**Д.О. ПОЛІЩУК, О.Д. ПОЛІЩУК, М.С. ЯДЖАК**

Запропоновано методику комплексного детермінованого оцінювання складних систем з ієрархічно-мережевою структурою, складовими якої є методи локального, прогностичного, агрегованого та інтерактивного аналізу стану, якості функціонування та взаємодії об'єктів, які утворюють систему. Описано основні підходи до побудови узагальнених висновків про стан та якість функціонування системи різних рівнів ієрархії: методи «найслабшого» елемента, зваженої лінійної та нелінійної агрегації. Визначено області їхнього застосування та проаналізовано основні переваги та недоліки. Детально розглянуто різні напрямки реалізації методу зваженої лінійної агрегації оцінок елементів системи та її базових підсистем. Запропоновано принципи візуалізації результатів агрегованого оцінювання для оперативної локалізації структурних елементів системи, які отримали негативні або близькі до них висновки. Ефективність запропонованих методів проілюстровано на прикладі аналізу якості об'єктів колійного та станційного господарства залізничної транспортної системи України.

**ВСТУП**

У праці [1] було запропоновано методику комплексного детермінованого оцінювання (МКДО) стану, якості функціонування та взаємодії елементів складних ієрархічно-мережевих систем (СІМС). Ця методика є поєднанням методів локального, прогностичного, інтерактивного та агрегованого оцінювання об'єктів, які утворюють систему. За допомогою методів локального оцінювання здійснено багатокритеріальний та багатопараметричний аналіз характеристик, які описують стан та процес функціонування елементів системи. Прогностичний аналіз цих характеристик дозволяє здійснювати коротко- та довготермінове прогнозування їхньої поведінки. Методи локального та прогностичного оцінювання були описані у праці [2]. Основна проблема, яка виникає у результаті застосування цих методів, полягає у великій кількості отриманих висновків, сумарна кількість яких може сягати мільйонів одиниць. Зрозуміло, що оперативний аналіз таких обсягів інформації та прийняття на їх основі адекватних рішень — складна проблема. Одним із способів її вирішення є застосування методів агрегації локальних та прогностичних оцінок [3–6].

**Мета роботи** — побудова узагальнених висновків щодо стану та якості функціонування об'єктів системи різних рівнів ієрархії. Саме опису методів агрегованого оцінювання та принципам візуалізації отриманих за їх допомогою результатів, які дозволяють оперативно орієнтуватись у великій кількості отриманих висновків та своєчасно реагувати на виявлені недоліки, присвячено цю роботу. Проаналізовано основні підходи до побудови узагальнених висновків, визначено їх переваги та недоліки та запропоновано метод, який дозволяє формувати обгрунтовані висновки для мережевих систем із повністю впорядкованим рухом потоків.

### **АГРЕГОВАНЕ ОЦІНЮВАННЯ: ОСНОВНІ ПІДХОДИ**

Визначальна риса ієрархічно-мережевих структур полягає в тому, що кожна підсистема системи на певному рівні ієрархії поділяється на сукупність підсистем, які є підмережею мережі нижчого рівня ієрархії. Отже, на кожному ієрархічному рівні система є мережею, тобто сукупністю вузлів, поєднаних ребрами, по яких здійснюється рух потоків. Підсистеми найнижчого рівня розбиття називатимемо базовими (БПС). У якості прикладу реальної СІМС розглядаємо [7–10] залізничну транспортну систему (ЗТС) України. Згідно з територіально-ієрархічним принципом побудови системи управління, Укрзалізниця (УЗ) послідовно поділяється на 6 регіональних залізниць, 27 дирекцій залізничних перевезень, 110 дистанцій колійного господарства та більш ніж 1200 відділків, які є послідовністю станцій та міжстанційних перегонів. Така структуризація встановлює чіткий взаємозв'язок між об'єктами ЗТС та підрозділами УЗ, які відповідають за їхній стан та якість функціонування. На найнижчому рівні ієрархії ЗТС є мережею, базовими підсистемами якої є станції (вузли) та міжстанційні перегони (ребра), а потоками у ній — поїзди. На вищих рівнях ієрархії УЗ є мережею, потоками в якій є інформація, управлінські та організаційні рішення. Подібний принцип структуризації дійсний для багатьох транспортних систем інших типів, систем постачання ресурсів та товарів, економічних та фінансових систем тощо [11–14].

Кожну з БПС системи поділяємо на елементи, тобто об'єкти, які мають чітко визначене просторове розміщення, функціональне призначення та набір характеристик, які описують їхній стан і процес функціонування, з відповідними типами областей допустимих значень поведінки цих характеристик. Поведінку кожної характеристики оцінюємо за певним набором критеріїв і параметрів та називаємо локальними оцінками стану та якості функціонування елемента [2]. На основі передісторії локальних оцінок, отриманих у результаті планових досліджень системи, будуємо прогностичні оцінки поведінки характеристик на коротко- та довгострокову перспективу [2, 15]. Повертаючись до прикладу ЗТС, елементами міжстанційного перегону є ділянки, які розділено штучними спорудами (мостами, переїздами тощо), відрізняються своїми просторовими, конструктивними, геологічними та іншими особливостями і не перевищують довжину 1 км. Кількість характеристик, які описують стан елементарної ділянки колії, перевищує сорок, сумарна кількість оцінок характеристик елементів лише колійного господарства ЗТС сягає кількох мільйонів [9]. Зрозуміло, що такі обсяги інформації практично не піддаються оперативному аналізу та своєчасній реакції на

виявлені недоліки, що у випадку залізниці нерідко призводить до катастрофічних наслідків. Найбільш розповсюдженим способом, який полегшує аналіз великих масивів оціночної інформації, є формування узагальнених висновків про стан та якість функціонування об'єктів, які утворюють складну систему (СС), на різних рівнях ієрархії. Для побудови таких висновків можна використовувати три підходи. Проілюструємо їх на найпростішому прикладі системи  $S$ , яка складається з  $N$  елементів  $s_n$ ,  $n = \overline{1, N}$ . Нехай  $e(s_n)$  — оцінка якості елемента  $s_n$  у шкалі  $[e^{\min}, e^{\max}]$ .

Якість системи  $e(S)$  за першим підходом визначається якістю її «найслабшого» елемента, тобто  $e^{(1)}(S) = \min_{n=1, N} e(s_n)$ . Такий підхід використовується для оцінювання систем, збої в роботі елементів яких можуть призвести до виходу з ладу окремих підсистем або всієї СС загалом. До таких систем відноситься людський організм, незадовільне функціонування окремих органів якого може призвести до летальних наслідків, тріщина в рейці колії може стати причиною залізничної аварії та зупинити рух на лінії тощо. Можна навести ще чимало подібних прикладів.

Другим та найбільш розповсюдженим підходом [3–5] є зважена лінійна агрегація оцінок, отриманих для елементів та підсистем кожного рівня ієрархії з врахуванням їх пріоритетності для отримання узагальненого висновку наступного рівня ієрархії. Для випадку розглянутої вище системи  $S$  ця оцінка отримується за співвідношенням

$$e^{(2)}(S) = \langle \rho, \mathbf{e}(s) \rangle_{R^N} / \langle \rho, \mathbf{1} \rangle_{R^N}.$$

Тут і нижче  $\langle \cdot, \cdot \rangle_{R^N}$  — скалярний добуток в евклідовому просторі  $R^N$ ,  $\mathbf{e}(s) = \{e(s_n)\}_{n=1}^N$ ,  $\rho = \{\rho_n\}_{n=1}^N$  — вектор вагових коефіцієнтів, який визначає пріоритетність елементів системи,  $\mathbf{1} = \{1\}_{n=1}^N$ . Основним недоліком методу лінійної агрегації є нівелювання як позитивних, так і негативних оцінок. Розглянемо наступний приклад. Нехай для виготовлення певного пристрою необхідне постачання  $N$  рівноважливих комплектуючих. Навіть у випадку постачання 100 відсотків  $(N-1)$ -ї з необхідних комплектуючих та 10 відсотків  $N$ -ї комплектуючої, можна виготовити лише 10% пристроїв. У той же час, оцінка, отримана за допомогою методу лінійної агрегації при  $\rho = \{1\}_{n=1}^N$ , віднесена до стовідсоткової шкали  $[e^{\min}, e^{\max}] = [0, 100]$ , дає значення  $e^{(2)}(S) = 100 - 90/N$ . При великих значеннях  $N$  ця оцінка є близькою до 100%, що очевидно не відповідає дійсності, оскільки все ще можна виготовити лише 10% пристроїв. Описані підходи доцільно поєднувати, а саме: після усунення виявлених під час застосування першого підходу недоліків (допостачання частини комплектуючих, якої бракує), для отримання узагальненого висновку використовувати метод лінійної агрегації.

Більш адекватний дійсності результат узагальнення отримуємо за допомогою методу нелінійної агрегації [6]. Він базується на наступних міркуваннях [16]. Нехай  $\{a_n\}_{n=1}^N$  — набір дійсних чисел таких, що  $\sum_{n=1}^N a_n = A$ .

Серед усіх таких наборів величина  $\prod_{n=1}^N a_n$  досягає свого максимального значення, коли  $a_n = A/N$ ,  $n = \overline{1, N}$ . Якщо  $\{a_n\}_{n=1}^N$  є набором оцінок, це означає, що найкращий результат серед усіх наборів досягається у випадку, коли відхилення значень  $a_n$ ,  $n = \overline{1, N}$ , від серед нього значення  $a^* = A/N$  є мінімальним. У методі нелінійної агрегації для випадку розглянутої вище системи  $S$  оцінка її якості отримується [6] за співвідношенням

$$e^{(3)}(S) = \prod_{n=1}^N e(s_n) / (e^*)^{N-1},$$

де  $e^* = \sum_{n=1}^N e(s_n) / N$ . Основним недоліком, який суттєво обмежує використання третього підходу для оцінювання якості системи, є складність врахування пріоритетності об'єктів, які її утворюють. Для розглянутого вище прикладу при  $N = 3$  оцінки, отримані за допомогою розглянутих вище підходів, у відсотковій шкалі матимуть наступні значення:  $e^{(1)}(S) = 10\%$ ,  $e^{(2)}(S) = 70\%$ ,  $e^{(3)}(S) = 20,4\%$ .

#### АГРЕГОВАНЕ ОЦІНЮВАННЯ ЕЛЕМЕНТІВ СИСТЕМИ

Нехай  $SD = \{S_m\}_{m=1}^{\tilde{M}} \cup \{D_m\}_{m=1}^M$  — підсистема, яка утворюється із взаємопов'язаних БПС, тобто є сукупністю вузлів  $S_m$ ,  $m = \overline{1, \tilde{M}}$ , та ребер, які їх поєднують,  $D_m$ ,  $m = \overline{1, M}$ . У випадку ЗТС найпростішим прикладом такої підсистеми є відділок, як послідовність станцій та перегонів між ними. Поділимо ребро  $D_m$  на послідовність елементарних ділянок  $\{D_{m,n}\}_{n=1}^{N_m}$  довжиною  $X_{m,n}$ , стан кожної з яких описується набором характеристик  $\{f_{m,n,i}(x)\}_{i=1}^{I_{N_m}}$ ,  $x \in [0, X_{m,n}]$ ,  $n = \overline{1, N_m}$ . Оцінювання поведінки характеристики  $f_{m,n,i}(x)$  здійснюємо за двома параметрами [2]. За першим формуємо уточнену бальну оцінку  $e^c(f_{m,n,i})$  на основі аналізу величини максимальних збурень  $f_{m,n,i}(x)$  на елементарній ділянці  $D_{m,n}$ . За другим параметром формуємо уточнену бальну оцінку  $e^l(f_{m,n,i})$  на основі аналізу масовості збурень  $f_{m,n,i}(x)$  на елементарній ділянці.

Стан елемента за відсутності незадовільних оцінок його характеристик у рівномірній та середньоквадратичній метриках оцінюємо за співвідношеннями

$$H_{m,n}^c = \langle \mathbf{p}_m, \mathbf{e}^c(\mathbf{f}_{m,n}) \rangle_{R^{I_{N_m}}} / \langle \mathbf{p}_m, \mathbf{1} \rangle_{R^{I_{N_m}}}, \quad (1)$$

$$H_{m,n}^l = \langle \mathbf{p}_m, \mathbf{e}^l(\mathbf{f}_{m,n}) \rangle_{R^{I_{N_m}}} / \langle \mathbf{p}_m, \mathbf{1} \rangle_{R^{I_{N_m}}} \quad (2)$$

відповідно. Тут  $\mathbf{e}^c(\mathbf{f}_{m,n}) = \{e^c(f_{m,n,i})\}_{i=1}^{I_{N_m}}$ ,  $\mathbf{e}^l(\mathbf{f}_{m,n}) = \{e^l(f_{m,n,i})\}_{i=1}^{I_{N_m}}$  — вектори уточнених бальних оцінок характеристики  $f_{m,n,i}$  елемента  $n$  базової

підсистеми  $D_m$  у рівномірній та середньоквадратичній метриках,  $\mathbf{p}_m = \{\rho_{m,i}\}_{i=1}^{I_{N_m}}$  — вектор вагових коефіцієнтів, який визначає пріоритетність характеристик елементів,  $n = \overline{1, N_m}$ ,  $m = \overline{1, M}$ .

Узагальнений висновок щодо стану елементарної ділянки  $D_{m,n}$  за сукупністю параметрів оцінки визначається співвідношенням

$$H_{m,n} = (\rho^c H_{m,n}^c + \rho^l H_{m,n}^l) / (\rho^c + \rho^l), \quad (3)$$

де  $\rho^c, \rho^l$  — вагові коефіцієнти, які визначають пріоритетність параметрів оцінки,  $n = \overline{1, N_m}$ ,  $m = \overline{1, M}$ . Аналогічним чином здійснюється оцінювання характеристик елементів вузлів  $S_m$ ,  $m = \overline{1, \tilde{M}}$ , підсистеми  $SD$ .

У деяких випадках якість функціонування елемента можна визначити, використовуючи локальні та прогностичні оцінки характеристик його стану. Дійсно, основна функція залізничної колії полягає у забезпеченні плавного, безпечного руху поїздів із максимально встановленою швидкістю, реалізація якої безпосередньо залежить від стану колії. Тому можна прийняти оцінку якості функціонування залізничної колії  $e_{m,n}^f$  на ділянці  $D_{m,n}$ , що дорівнює:

- 5, якщо  $H_{m,n} = 5$ ;
- $4 + \tau / (t_c - t_n)$ , якщо  $H_{m,n} \geq 4$  і запланований згідно з нормативом ремонт зі збереженням цілочисельної частини оцінки стану можна відтермінувати на час  $\tau \leq t_c - t_n$ , де  $t_n, t_c$  — моменти наступного та того, що за ним слідує згідно з графіком ремонту;
- $3 + \tau / (t_c - t_n)$ , якщо  $H_{m,n} \geq 3$  і запланований згідно з нормативом ремонт зі збереженням цілочисельної частини оцінки стану можна відтермінувати на час  $\tau \leq t_c - t_n$ ;
- 2, якщо  $H_{m,n} = 2$ .

Узагальнений висновок щодо якості функціонування БПС  $D_m$  визначено за співвідношенням  $E_{D_m}^f = \langle \mathbf{e}_m^f, \mathbf{1} \rangle_{R^{I_{N_m}}} / N$ , де  $\mathbf{e}_m^f = \{e_{m,n}^f\}_{n=1}^N$ . Зрозуміло, що оцінки  $e_{m,n}^f$  ґрунтуються на результатах прогностичного аналізу стану колії на ділянці  $D_{m,n}$ .

Важливим показником оцінювання є рівень покриття оцінками характеристик елемента БПС із урахуванням їх пріоритетності, який визначається співвідношенням

$$C_{D_{m,n}} = 100\% \times \langle \mathbf{p}_m, \boldsymbol{\delta}_{m,n} \rangle_{R^{I_{N_m}}} / \langle \mathbf{p}_m, \mathbf{1} \rangle_{R^{I_{N_m}}}, \quad (4)$$

де  $\boldsymbol{\delta}_{m,n} = \{\delta_{m,n,i}\}_{i=1}^{I_{N_m}}$ ,

$$\delta_{m,n,i} = \begin{cases} 0, & \text{якщо оцінка } i\text{-ї характеристики не проводилася,} \\ 1, & \text{якщо оцінка } i\text{-ї характеристики проводилася.} \end{cases}$$



Узагальнення оцінок окремих характеристик у рівномірній та середньоквадратичній метриках за сукупністю елементів, які утворюють БПС, послідовно здійснюється за співвідношеннями

$$V_{m,i}^c = \langle \tilde{\mathbf{p}}_m, \mathbf{e}^c(\tilde{\mathbf{f}}_{m,i}) \rangle_{R^{N_m}} / \langle \tilde{\mathbf{p}}_m, \mathbf{1} \rangle_{R^{N_m}}, \quad (5)$$

$$V_{m,i}^l = \langle \tilde{\mathbf{p}}_m, \mathbf{e}^l(\tilde{\mathbf{f}}_{m,i}) \rangle_{R^{N_m}} / \langle \tilde{\mathbf{p}}_m, \mathbf{1} \rangle_{R^{N_m}}, \quad (6)$$

де  $\mathbf{e}^c(\tilde{\mathbf{f}}_{m,i}) = \{e^c(f_{m,n,i})\}_{n=1}^{N_m}$ ,  $\mathbf{e}^l(\tilde{\mathbf{f}}_{m,i}) = \{e^l(f_{m,n,i})\}_{n=1}^{N_m}$  — вектори уточнених бальних оцінок характеристики  $f_{m,n,i}$  для сукупності елементів БПС  $D_m$  у рівномірній та середньоквадратичній метриках,  $\tilde{\mathbf{p}}_m = \{\tilde{p}_{m,n}\}_{n=1}^{N_m}$  — вектор вагових коефіцієнтів, який визначає пріоритетність елементів у структурі БПС. Тоді узагальнений висновок щодо стану елементарної ділянки  $D_{m,n}$  для заданої характеристики визначається співвідношенням

$$V_{m,i} = (\rho^c V_{m,i}^c + \rho^l V_{m,i}^l) / (\rho^c + \rho^l), \quad i = \overline{1, I_{N_m}}. \quad (7)$$

Як видно із співвідношень (1)–(3) та (5)–(7), будуємо узагальнені оцінки у двох напрямках, які умовно називаємо «горизонтальним» ( $H$ ) й «вертикальним» ( $V$ ). Перший формує агреговані висновки про стан або якість функціонування окремого елемента БПС за сукупністю параметрів, критеріїв та характеристик, які підлягають оцінюванню. Він дозволяє ідентифікувати елементи, які отримали не задовільні або близькі до них оцінки. Другий напрямок формує агреговані висновки про стан або якість функціонування БПС системи за окремим параметром, критерієм або характеристикою елементів, які її утворюють. Він дозволяє визначити параметри, критерії або характеристики, відносно яких оцінки БПС також є незадовільними або близькими до них. Стосовно об'єктів ЗТС, «горизонтальний» напрямок оцінювання колії дає можливість виявити «критичні» ділянки окремого міжстанційного перегону, які потребують термінового ремонту. «Вертикальний» напрямок дозволяє визначити проблемні характеристики, наприклад, забрудненість земляного полотна, стан рейок або шпал на міжстанційному перегоні тощо.

Рівень покриття оцінками елементів БПС  $D_{m,n}$  з урахуванням їх пріоритетності визначимо за співвідношенням

$$C_{D_m} = \langle \tilde{\mathbf{p}}_m, \mathbf{C}_{D_m} \rangle_{R^{N_m}} / \langle \tilde{\mathbf{p}}_m, \mathbf{1} \rangle_{R^{N_m}}, \quad (8)$$

де  $\mathbf{C}_{D_m} = \{C_{D_{m,n}}\}_{n=1}^{N_m}$ .

Узагальнений висновок щодо стану БПС  $D_m$  на момент останнього планового дослідження за сукупністю елементів та їх характеристик, визначимо за співвідношенням

$$E_{D_m}^{s,P} = \langle \tilde{\mathbf{p}}_m, \mathbf{H}_m \rangle_{R^{N_m}} / \langle \tilde{\mathbf{p}}_m, \mathbf{1} \rangle_{R^{N_m}} \equiv \langle \mathbf{p}_m, \mathbf{V}_m \rangle_{R^{I_{N_m}}} / \langle \mathbf{p}_m, \mathbf{1} \rangle_{R^{I_{N_m}}}, \quad (9)$$

де  $\mathbf{H}_m = \{H_{m,n}\}_{n=1}^{N_m}$ ,  $\mathbf{V}_m = \{V_{m,i}\}_{i=1}^{I_{N_m}}$ .

За аналогічним принципом здійснюємо побудову зважених усереднених оцінок стану вузлів  $S_m$ ,  $m = \overline{1, \widetilde{M}}$ , та якості функціонування ребер і вузлів підсистеми  $SD$  на момент останнього та наступного планового дослідження.

### АГРЕГОВАНЕ ОЦІНЮВАННЯ БПС СИСТЕМИ

Наступний рівень узагальнення полягає у побудові послідовності зважених агрегованих оцінок, які формують висновок про стан та якість функціонування підсистеми  $SD$  загалом. А саме, узагальнений висновок про БПС  $D_m$  на останній та наступний моменти планових досліджень визначаємо за співвідношеннями

$$H_{D_m}^P = (\rho^{s,P} E_{D_m}^{s,P} + \rho^{f,P} E_{D_m}^{f,P})(\rho^{s,P} + \rho^{f,P}), \quad (10)$$

$$H_{D_m}^F = (\rho^{s,F} E_{D_m}^{s,F} + \rho^{f,F} E_{D_m}^{f,F})(\rho^{s,F} + \rho^{f,F}), \quad (11)$$

$$H_{D_m} = (\rho^P H_{D_m}^P + \rho^F H_{D_m}^F) / (\rho^P + \rho^F), \quad (12)$$

де  $E_{D_m}^{s,P}$ ,  $E_{D_m}^{f,P}$ ,  $E_{D_m}^{s,F}$ ,  $E_{D_m}^{f,F}$  — агреговані оцінки стану та якості функціонування БПС  $D_m$  на поточний та наступний терміни планових досліджень відповідно;  $\rho^{s,P}$ ,  $\rho^{f,P}$ ,  $\rho^{s,F}$ ,  $\rho^{f,F}$  — вагові коефіцієнти, які визначають їх пріоритетність;  $\rho^P$ ,  $\rho^F$  — вагові коефіцієнти, які визначають пріоритетність поточних та прогностичних оцінок  $H_{D_m}^P$  й  $H_{D_m}^F$  відповідно (у співвідношеннях (10)–(12) і нижче використані індекси  $s$  — state (стан),  $f$  — function (функціонування),  $P$  — Present (поточний),  $F$  — Future (наступний)). Аналогічним чином здійснюємо побудову узагальнених поточних та прогностичних висновків  $H_{S_m}^P$ ,  $H_{S_m}^F$ ,  $H_{S_m}$  щодо вузлів підсистеми  $SD$ .

Агрегований висновок щодо стану та якості функціонування всіх БПС підсистеми  $SD$  на момент останнього та наступного за планом дослідження послідовно здійснюємо наступним чином:

- стан вузлів та ребер на момент останнього планового дослідження визначаємо за співвідношенням

$$V_{SD}^{s,P} = \langle \rho_D, E_D^{s,P} \rangle_{R^M} / \langle \rho_D, \mathbf{1} \rangle_{R^M} + \langle \rho_S, E_S^{s,P} \rangle_{R^{\widetilde{M}}} / \langle \rho_S, \mathbf{1} \rangle_{R^{\widetilde{M}}}; \quad (13)$$

- стан вузлів та ребер на момент наступного планового дослідження — за співвідношенням

$$V_{SD}^{s,F} = \langle \rho_D, E_D^{s,F} \rangle_{R^M} / \langle \rho_D, \mathbf{1} \rangle_{R^M} + \langle \rho_S, E_S^{s,F} \rangle_{R^{\widetilde{M}}} / \langle \rho_S, \mathbf{1} \rangle_{R^{\widetilde{M}}}; \quad (14)$$

- якість функціонування вузлів та ребер на момент останнього планового дослідження — за співвідношенням

$$V_{SD}^{f,P} = \langle \rho_D, E_D^{f,P} \rangle_{R^M} / \langle \rho_D, \mathbf{1} \rangle_{R^M} + \langle \rho_S, E_S^{f,P} \rangle_{R^{\widetilde{M}}} / \langle \rho_S, \mathbf{1} \rangle_{R^{\widetilde{M}}}; \quad (15)$$

• якість функціонування вузлів та ребер на момент наступного планового дослідження — за співвідношенням

$$V_{SD}^{f,F} = \langle \rho_D, E_D^{f,F} \rangle_{R^M} / \langle \rho_D, \mathbf{1} \rangle_{R^M} + \langle \rho_S, E_S^{f,F} \rangle_{R^{\tilde{M}}} / \langle \rho_S, \mathbf{1} \rangle_{R^{\tilde{M}}}. \quad (16)$$

У співвідношеннях (13)–(16)  $E_D^{s,P} = \{E_{D_m}^{s,P}\}_{m=1}^M$ ,  $E_D^{s,F} = \{E_{D_m}^{s,F}\}_{m=1}^M$ ,  $E_S^{s,P} = \{E_{S_m}^{s,P}\}_{m=1}^{\tilde{M}}$ ,  $E_S^{s,F} = \{E_{S_m}^{s,F}\}_{m=1}^{\tilde{M}}$  й  $\rho_D = \{\rho_{D_m}\}_{m=1}^M$ ,  $\rho_S = \{\rho_{S_m}\}_{m=1}^{\tilde{M}}$  — вектори вагових коефіцієнтів, які визначають пріоритетність окремих ребер та вузлів підсистеми  $SD$ .

Стан та якість функціонування усіх БПС підсистеми  $SD$  з урахуванням прогнозу визначаємо за співвідношеннями

$$V_{SD}^s = (\rho^P V_{SD}^{s,P} + \rho^F V_{SD}^{s,F}) / (\rho^P + \rho^F), \quad (17)$$

$$V_{SD}^f = (\rho^P V_{SD}^{f,P} + \rho^F V_{SD}^{f,F}) / (\rho^P + \rho^F). \quad (18)$$

Узагальнений рівень покриття оцінками характеристик елементів БПС підсистеми  $SD$  визначаємо за співвідношенням

$$C_{SD} = \langle \rho_D, C_D \rangle_{R^M} / \langle \rho_D, \mathbf{1} \rangle_{R^M} + \langle \rho_S, C_S \rangle_{R^{\tilde{M}}} / \langle \rho_S, \mathbf{1} \rangle_{R^{\tilde{M}}}, \quad (19)$$

де  $C_D = \{C_{D_m}\}_{m=1}^M$ ,  $C_S = \{C_{S_m}\}_{m=1}^{\tilde{M}}$ .

Остаточний висновок щодо якості підсистеми  $SD$  отримуємо із співвідношення

$$E_{SD} = (\rho^s V_{SD}^s + \rho^f V_{SD}^f) / (\rho^s + \rho^f) \equiv \langle \rho_D, H_D \rangle_{R^M} / \langle \rho_D, \mathbf{1} \rangle_{R^M} + \langle \rho_S, H_S \rangle_{R^{\tilde{M}}} / \langle \rho_S, \mathbf{1} \rangle_{R^{\tilde{M}}}, \quad (20)$$

де  $H_D = \{H_{D_m}\}_{m=1}^M$ ,  $H_S = \{H_{S_m}\}_{m=1}^{\tilde{M}}$ ,  $\rho^s$ ,  $\rho^f$  — вагові коефіцієнти, які визначають пріоритетність оцінок стану та якості функціонування підсистеми  $SD$ . Аналогічний принцип побудови агрегованих оцінок використовуємо для формування узагальнених висновків про якість функціонування підсистем вищих рівнів ієрархії.

## ВІЗУАЛІЗАЦІЯ РЕЗУЛЬТАТІВ АГРЕГОВАНОГО ОЦІНЮВАННЯ

Принцип візуалізації результатів локального та першого рівня агрегованого оцінювання відображений у таблиці на рис. 1, у якій на перетині стовпців  $e^c(f_{m,n,i})$ ,  $e^l(f_{m,n,i})$  та рядків  $D_{m,n}$  містяться оцінки характеристик кожного елемента БПС  $D_m$ , яке здійснюється для найпростішого випадку, тобто за одним критерієм та двома параметрами. Для елементів систем інших типів таке оцінювання може здійснюватися за сукупністю більшої кількості критеріїв та параметрів [17–20]. Під час програмної реалізації рівень затемнення кожної комірки таблиці є пропорційним числовому значенню оцінки

(100% затемнення відповідає значенню «незадовільно», 0% — значенню «відмінно»). Комірка є заштрихованою, якщо оцінювання для цієї характеристики не проводилося. Підвід курсора до обраної комірки супроводжується висвітленням у ній числових значень відповідної оцінки, клікання на комірці — відображенням графіка поведінки цієї характеристики на фоні області допустимих значень та підобластей, які відповідають цілочисельним бальним оцінкам, та таблиці її числових значень [2]. Такий спосіб візуалізації отриманих результатів дозволяє оперативно локалізувати «проблемні» характеристики та частково встановити причини негативних або близьких до них висновків. У цій таблиці також відображені значення рівня покриття оцінками характеристик елемента  $C_{D_m,n}$  та рівень покриття оцінками  $C_{D_m}$  елементів БПС  $D_{m,n}$ , узагальнені оцінки у «горизонтальному» напрямку  $H_{m,n}^c$ ,  $H_{m,n}^l$ ,  $H_{m,n}$  та «вертикальному» напрямку  $V_{m,i}^c$ ,  $V_{m,i}^l$ ,  $V_{m,i}$ , а також узагальнений висновок  $E_{D_m}^{s,P}$  про стан БПС  $D_m$  загалом.

Елементарні ділянки	Оцінки									
	$e^c(f_{m,n,1})$	$e^l(f_{m,n,1})$	...	...	$e^c(f_{m,n,1N_m})$	$e^l(f_{m,n,1N_m})$	$C_{D_m,n}$	$H_{m,n}^c$	$H_{m,n}^l$	$H_{m,n}$
$D_{m,1}$							69%			
$D_{m,2}$	3,26	3,71					81%			
...							...			
$D_{m,N_m}$							24%			
$V_{m,i}^{c/l}$							$C_{D_m}$	$E_{D_m}^{s,P}$		
$V_{m,i}$										

Рис. 1. Оцінка стану елементів БПС за сукупністю характеристик

Запропонований спосіб візуалізації результатів оцінювання дає можливість уникнути недоліків лінійної агрегації, оскільки об'єкти з найнижчою оцінкою відстежуються досить легко.

Зображена на рис. 2 таблиця містить результати агрегованого оцінювання підсистеми  $SD$ , яка є лінійною послідовністю вузлів та поєднуючих їх ребер. Це є природним для багатьох СІМС (для ЗТС це станції та міжстанційні перегони, що знаходяться на лінії, яка поєднує деякі початковий та кінцевий пункти руху поїздів). На перетині стовпців  $E^{s,P}$ ,  $E^{s,F}$ ,  $E^{f,P}$ ,  $E^{f,F}$  та рядків  $D_m$ ,  $m = \overline{1, M}$ ,  $S_m$ ,  $m = \overline{1, M + 1}$ , містяться результати агрегованого оцінювання стану та якості функціонування вузлів та ребер підсистеми  $SD$  на моменти поточного та наступного планового дослідження відповідно. Як і вище, при програмній реалізації алгоритмів агрегованого оцінювання БПС рівень затемнення кожної комірки таблиці є пропорційним числовому значенню оцінки. Комірка є заштрихованою, якщо оцінювання

для цієї БПС не проводилося. Підвід курсора до обраної комірки супроводжується висвітленням у ній числових значень відповідної оцінки, клікання на комірці — відображенням таблиць, подібних зображень на рис. 1, яка містять результати локального та прогностичного оцінювання обраної БПС системи. Такий спосіб візуалізації отриманих результатів дозволяє оперативно локалізувати «проблемні» БПС. У таблиці також відображені значення рівня покриття оцінками  $C_m$  БПС підсистеми  $SD$ , узагальнені оцінки  $H_{D_m}^P$ ,  $H_{D_m}^F$ ,  $H_{D_m}$ ,  $V_{SD}^{s,P}$ ,  $V_{SD}^{s,F}$ ,  $V_{SD}^{f,P}$ ,  $V_{SD}^{f,F}$ ,  $V_{SD}^s$ ,  $V_{SD}^f$ , визначені співвідношеннями (10)–(19), відповідно, та остаточний висновок якості підсистеми  $E_{SD}$ .

$SD$	$E^{s,P}$	$E^{s,F}$	$E^{f,P}$	$E^{f,F}$	$C_{D_m/S_m}$	$H_m^P$	$H_m^F$	$H_m$
$D_1$					71%			
...					...			
$D_M$	3,54	3,22	←		74%			
$S_1$					61%			
...					...			
$S_{M+1}$					63%			
$V^{s/f,P/F}$					$C_{SD}$	$E_{SD}$		
$VH^{s/f}$								

Рис. 2. Поточна та прогностична агрегована оцінка стану та якості функціонування БПС, які утворюють окрему підсистему СІМС

Графічне відображення значень послідовностей  $H_{D_m}^P$ ,  $H_{S_m}^P$  або  $H_{D_m}^F$ ,  $H_{S_m}^F$ ,  $m = \overline{1, M}$ , з урахуванням рівня покриття оцінками об'єктів підсистеми  $SD$  дає можливість локалізувати проблемні ділянки ліній на момент останнього дослідження та дозволяє оперативно реагувати на виявлені або потенційні недоліки.

За аналогічним принципом будуються агреговані оцінки для підсистем усіх рівнів ієрархії. Пропоновані способи візуалізації результатів оцінювання дають можливість оперативно відстежувати об'єкти, для яких отримані негативні або близькі до них висновки, з метою виявлення та усунення причин їх появи. Навіть для таких систем, як ЗТС, незважаючи на величезні сумарні обсяги результатів оцінювання, кількість яких на локальному рівні сягає мільйонів одиниць, максимальна кількість кроків локалізації у пропонуваній моделі не перевищуватиме семи (УЗ → регіональна залізниця → дирекція → дистанція → відділок → БПС → елемент → характеристика).

## ВИСНОВКИ

У працях [1, 2] і цій роботі запропоновано методику комплексного детермінованого оцінювання складних систем із ієрархічно-мережевою структурою, складовими якої є методи локального, прогностичного, інтерактивного та агрегованого оцінювання її основних об'єктів. Показано, що у поєднанні та застосовуючи уточнену бальну шкалу оцінок, вони формують достатньо повне, адекватне та цілісне уявлення щодо стану та якості функціонування елементів досліджуваної системи та її підсистем на всіх рівнях структуризації. Якщо локальні та прогностичні оцінки дають можливість привертати увагу до найбільш проблемних елементів системи та своєчасно реагувати на існуючі або потенційні ризики, то результати їх агрегованого оцінювання дають обґрунтовані кількісні підстави для планування матеріальних та фінансових витрат і необхідних людських ресурсів для модернізації окремих об'єктів, підсистем чи системи загалом. Вони також є підґрунтям для формування опосередкованих, але від того не менш вагомих висновків щодо якості роботи підрозділів СУ системи. Пропоновані принципи відображення результатів оцінювання різних типів та рівнів узагальнення дають можливість оперативно орієнтуватися у великій кількості одержаних висновків та локалізувати причини виявлених недоліків. Описана методика використовується при розробці програмного забезпечення для оцінювання стану та якості функціонування колійного та станційного господарства Укрзалізниці [9, 10].

## ЛІТЕРАТУРА

1. Поліщук Д., Поліщук О., Яджак М. Комплексне оцінювання складних систем з ієрархічно-мережевою структурою: I. Опис методики // Системні дослідження та інформаційні технології. — 2015. — № 1. — С. 21–31.
2. Поліщук Д., Поліщук О., Яджак М. Комплексне оцінювання складних систем з ієрархічно-мережевою структурою: II. Локальне та прогностичне оцінювання // Системні дослідження та інформаційні технології. — 2015. — № 2. — С. 26–38.
3. Воронин А.Н. Метод многокритериальной оценки и оптимизации иерархических систем // Кибернетика и системный анализ. — 2007. — № 3. — С. 84–92.
4. Dombi J. Basic concepts for a theory of evaluation: the aggregative operator // European Journal of Operational Research. — 1982. — 10, № 3. — P. 282–293.
5. Богданчук В.З., Егоров Б.М., Катупев А.Н. Агрегирование векторных критериев. — Л.: ЛИИАН, 1990. — 214 с.
6. Polishchuk D., Polishchuk O. About evaluation of complex dynamical systems // Journal of Complex Systems. — 2013, Article ID 204304, 6 p.: Available. — <http://dx.doi.org/10.1155/2013/204304>.
7. Офіційний сайт Укрзалізниці. — <http://www.uz.gov.ua>.
8. Бакаев А.А. и др. Международные транспортные коридоры Украины: сети и моделирование. Том 1. Наземные виды транспорта. — К.: КУЭТТ, 2003. — 516 с.
9. Поліщук Д.О. Модель локального оцінювання стану елементів колійного господарства Укрзалізниці // Вісник Дніпропетровського національного університету залізничного транспорту імені акад. В. Лазаряна. — 2012. — Вип. 41. — С. 158–166.

10. *Поліщук Д.О.* Модель локального оцінювання станційного господарства Укрзалізниці // 36. наук. праць IV Конф. молодих учених із сучасних проблем механіки і математики ім. акад. Я.С. Підстригача. — 24–27 травня 2011 р. — Львів. — С. 178–181.
11. *Поліщук Д.О., Поліщук О.Д.* Моніторинг потоку транспортних мереж із частково впорядкованим рухом // 36. наук. праць XXIII наук.-техн. конф. молодих науковців Фізико-механічного інституту ім. Г.В. Карпенка НАНУ. — 23–25 жовтня, 2013. — Львів, Україна. — С.326–329.
12. *Дилигенский Н.В., Дымова Л.Г., Севастьянов П.В.* Нечеткое моделирование и многокритериальная оптимизация производственных систем в условиях неопределенности: технология, экономика, экология. — М.: Машиностроение, 2004. — 458 с.
13. *Полянский С.В., Семенов И.Б., Чижов С.А.* Комплексное оценивание в задачах управления системами социально-экономического типа. — М.: ИПУ РАН, 1996. — 48 с.
14. *Хованов Н.В.* Оценка сложных экономических объектов и процессов в условиях неопределённости // Вестник СПбГУ. Сер. 5. — 2005. — Вып. 1. — С. 139–143.
15. *Поліщук Д., Яджак М.* Використання часових рядів для прогнозування оцінки якості функціонування складних систем // 36. праць Міжн. наук. конф. «Сучасні проблеми механіки і математики». — 25–29 травня 2008. — Львів. — Т. 3. — С. 38–40.
16. *Hardy G.H., Littlewood J.E., Polia G.* Inequalities. — Cambridge University Press, 1988. — 478 p.
17. *Polishchuk O.* Optimization of estimation of man's musculoskeletal system // Computer Mathematics and Calculation Optimization. — 2001. — 2. — P. 360–367.
18. *Bar-Yam Y.* About engineering complex systems: Multiscale analysis and evolutionary engineering // Engineering Self-Organising Systems: Methodologies and Applications. — London: Springer, 2005. — P. 16–31.
19. *Bloom M., Fischer J., Orme J.* Evaluating practice. — Allyn and Bacon, 2006. — 488 p.
20. *Polishchuk D., Polishchuk O., Yadzhak M.* Solution of some problems of evaluation of the complex systems // Proc. of the 15<sup>th</sup> Int. conf. on automatic control, 23–26 September 2008. — Odesa: ONMA. — P. 968–976.

*Надійшла 28.08.2013*

## **FORECASTING CONSUMER PRICE INDEX IN UKRAINE WITH REGRESSION MODELS AND ADAPTIVE KALMAN FILTER**

**I.V. KARAYUZ, P.I. BIDYUK**

The paper considers the problem of short term forecasting of consumer price index using regression models and adaptive Kalman filter. The main purpose of the study is constructing of high quality model for forecasting of consumer price index and application of Kalman filter for computing optimal estimates of states for the process under investigation. The basic results of the study are as follows: two modifications of the Kalman filter (ordinary and adaptive), directed towards estimation of covariances for stochastic state disturbances and measurement errors. Alternative short term forecasts are generated with regression models and Kalman filters. A comparative analysis of results achieved is given. The necessary statistical data was taken from Ukrainian economy in transition.

### **INTRODUCTION**

The problem of high quality forecasting for economic and financial processes requires development and application of new modern techniques that are based on systemic approach to development of respective computational software. Most often this software is implemented in the form of modern decision support systems (DSS) that become popular as a solution instrument for many practical problems. One of the simplest definitions of DSS is as follows: DSS is computer based information processing system that provides decision making user with any help relevant to data collecting and storing, preliminary data processing, constructing necessary mathematical models, generating alternatives, selection of the best solution, convenient presentation of intermediate and final results etc [1].

To improve substantially quality of forecasts for macroeconomic processes in the frames of DSS it is useful to construct adaptive computing schemes for a model and parameter estimation [2]. According to this scheme each step of data processing is controlled by appropriate set of statistical parameters each of which characterizes specific features of data, model as a whole, model parameters and quality of forecasts estimates. A substantial help in forecasting linear and nonlinear nonstationary processes can be provided by application of adaptive Kalman filter that is useful for estimation of covariances for external stochastic disturbances and measurement noise (errors) [3]. The filter also provides a possibility for computing optimal estimates for state vector of a system under study and quality short term forecasts.

The paper considers the possibilities for modeling selected financial and economic processes with regression analysis approach and adaptive Kalman filter using statistical data from Ukrainian economy in transition.



## PROBLEM STATEMENT

The purpose of the study is as follows:

- to construct mathematical models for selected financial and economic processes providing acceptable quality of short term forecasts;
- to transform the models into state space representation and apply adaptive Kalman filter to generate optimal state vector estimates and short term forecasts;
- to perform comparison of the results achieved;
- to develop software necessary for performing computational experiments that incorporates possibilities for adaptive model structure and parameters estimation.

## PROPOSED APPROACH TO FORECASTING

We propose adaptive computing scheme that is distinguished with several possibilities for adaptation using complex quality criterion. The statistical data collected should be correctly prepared for model structure and parameter estimation. The model structure and parameter estimation is a key issue for reaching necessary quality of forecasts. It is proposed to define model structure as follows:  $S = \{r, p, m, n, d, z, l\}$ , where  $r$  is model dimensionality (number of equations);  $p$  is model order (maximum order of differential or difference equation in a model);  $m$  is a number of independent variables in the right hand side;  $n$  is a type of nonlinearity;  $d$  is a lag or output reaction delay time;  $z$  is external disturbance and its type;  $l$  are possible restrictions for variables. To perform automatic search for the “best” model it is proposed to use the following criterion:

$$V_N(\theta, D_N) = e^{|1-R^2|} + \ln\left(1 + \frac{SSE}{N}\right) + e^{|2-DW|} + \ln(1 + MSE) + \ln(MAPE) + e^U,$$

where  $\theta$  is a vector of model parameters;  $N$  is a power of time series used;  $R^2$  is a determination coefficient;  $DW$  is Durbin-Watson statistic;  $MSE$  is mean squared error;  $MAPE$  is mean absolute percentage error;  $U$  is Theil coefficient. The power of the criterion was tested experimentally with a wide set of linear and pseudolinear models with positive results.

There are several possibilities for adaptive model structure estimation: (1) automatic analysis of partial autocorrelation for determining autoregression order; (2) automatic search for the exogeneous variables lag estimate (detection of leading indicators); (3) automatic analysis of residual properties; (4) analysis of data distribution type and its use for selecting correct model parameters estimation method; (5) adaptive model parameter estimation with hiring extra data; (6) optimal selection of weighting coefficients for exponential smoothing, nearest-neighbor interpolation and some other techniques; (7) the use of adaptive approach to model type selection. The use of specific adaptation scheme depends on volume and quality of data, specific problem statement, requirements to forecast estimates, etc. In some cases it is possible to use logistic regression together with linear regression to describe the data with nonlinearities. Application of the adaptive concept described provides the following advantages: (1) automatic search

for the «best» model reduces the search time for many times; (2) it is possible to analyze much wider set of candidate models than manually; (3) the search is optimized thanks to the use of complex quality criterion; (4) in the frames of computer system developed it is possible to integrate ideologically different methods of modeling and forecasting and compute combined forecasts estimates that are distinguished with better quality.

## FILTERING ALGORITHM

Most of modern macroeconomic and financial processes on the post soviet territory are influenced by a substantial number of stochastic disturbances. Among them are the following: low qualification of managerial staff; high dependence on energy import; inconsistent and unstable laws; outdated technologies in industry and agriculture, unstable and outdated education system, high inflation (hyperinflation in 90s) etc. It was also revealed that data taken from different sources very often contain substantially different values for the same variables. It means that the measurement errors are available that also should be taken into consideration when a process model is developed. To take into consideration these disturbances and to reduce influence of measurement errors linear optimal Kalman filter is applied. The filtering procedure for scalar case is given below [3].

**Step 1.** Mathematical model of linear dynamic system:

$$x(k) = F(k)x(k-1) + C(k)u(k-1) + w(k-1),$$

where  $x(k) = y(k)$  — process state vector (selected variable);  $F(k) = \alpha_1$  — coefficient characterizing system dynamics;  $C(k)u(k-1) = \beta_1 x(k-1)$  — influence of possible control actions (regressor);  $w(k) \sim N(0, Q0, Q(k))$  — external stochastic disturbances;  $z(k) = H(k)x(k) + v(k)$  — measurement equation;  $H(k) = 1$  — measurement matrix;  $v(k) \sim N(0, R(k))$  — measurement errors; initial conditions:  $E\{x_0\} = \hat{x}_0$ ,  $E\{\hat{x}_0, \hat{x}_0^T\} = P_0 = P'_0$ ;  $P(k)$  — posterior covariance matrix for state vector estimation errors;  $P'(k)$  — prior covariance matrix for state vector estimation errors;  $K(k)$  — optimal Kalman filter coefficient; initial covariances:

$$E\{w(k), v^T(k)\} = 0, \quad E\{w(k), x_0^T(0)\} = 0, \quad E\{v(k), x_0^T\} = 0.$$

**Step 2.** State vector extrapolation:

$$\hat{x}(k) = F(k)\hat{x}(k-1) + C(k)u(k-1) = \alpha_1 y(k-1) + \beta_1 x(k-1).$$

**Step 3.** Extrapolation for covariance matrix of estimation errors:

$$P'(k) = F(k)P(k-1)F^T(k) + Q(k-1) = \alpha_1^2 P(k-1) + Q(k-1).$$

**Step 4.** Scalar coefficient of the filter is recalculated adaptively using covariance of measurement errors as root mean squared for filtering errors:  $R(k) = \sqrt{E\{\hat{x}_0, \hat{x}_0^T\}}$ . To compute the covariances an adaptive KF with moving window (AKFMW) is proposed:

$$K(k) = P'(k)H^T(k)\{H(k)P'(k)H^T(k) + R(k)\}^{-1} = \frac{P'(k)}{P'(k) + R(k)}.$$

**Step 5.** Computing of state vector estimate using the last measurement received  $z(k)$  :

$$\hat{x}(k) = \hat{x}(k) + K(k)[z(k) - H(k)\hat{x}(k)].$$

**Step 6.** Computing the posterior matrix for state vector estimation errors:

$$P(k) = [I - K(k)H(k)]P'(k).$$

**Step 7.** Go to step 2.

The statistical data for consumer price index (CPI) was taken from the state statistic tables for the period from 01.01.2000 to 01.08.2014 [4] (Fig. 1). The computations were performed using developed software [5].

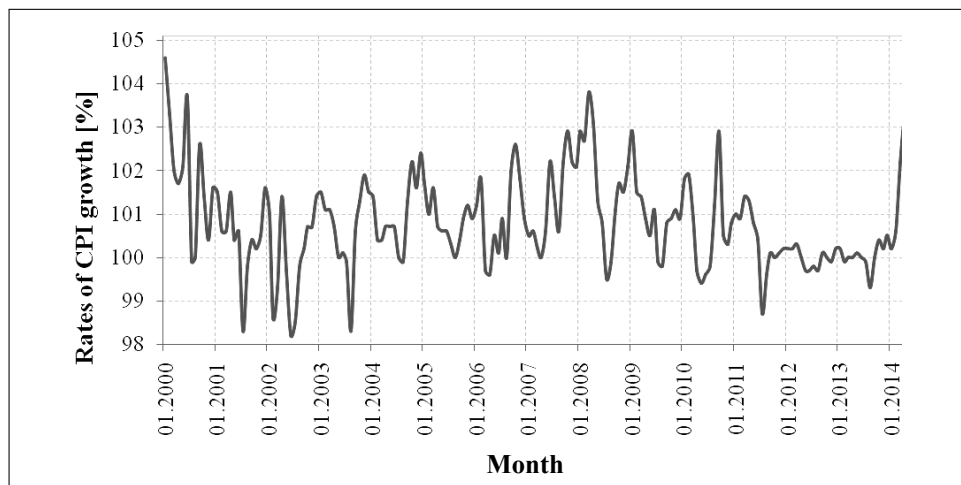


Fig. 1. Consumer price index

Constructed autoregression AR(2) model without constant is as follows:

$$y(k) = a_1y(k-1) + a_2y(k-2) + \varepsilon(k),$$

where  $y(k)$  — volume of imported natural gas;  $x(k)$  — volume of consumed gas. After estimation of the model parameters we have the equation:

$$y(k) = 0,92 \cdot y(k-1) + 0,8 \cdot y(k-2).$$

Now test the distribution law for model residuals; respective histogram is given in Fig. 2. According to statistics  $\chi^2$

$$\sum_{k=1}^r \frac{(n_k - np_k)^2}{np_k} = 3,17 < 3,84 = \chi_{\alpha,1}^2 = \chi_{0.05,1}^2$$

— the normality hypothesis is accepted.

In this expression  $n$  is a number of observations;  $r$  is a number of data intervals;  $n_k$  — is a number of measurements that are in a specific  $k$ -th interval;  $p_k$  is a probability of appearing random value in  $k$ -th interval of normal distribution;  $\alpha$  is a level of significance.

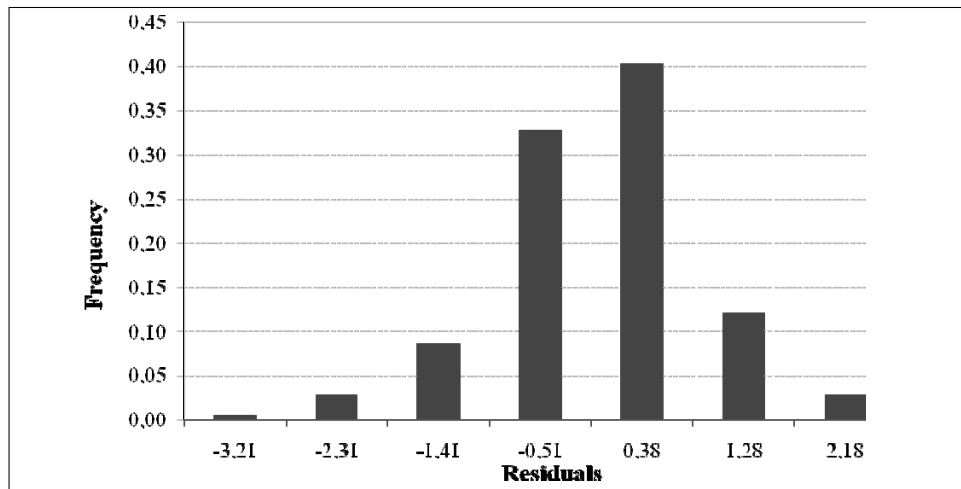


Fig. 2. Residuals histogram for CPI model AR(2)

Model quality is characterized with the following statistics:  $R^2 = 0,21$ ;  $DW = 2,07$ ;  $e^2 = 148,01$  (Table 1). Three steps forecasts quality is given in Table 2.

Table 1. Quality of the forecasts

Model type \ Statistics	MSE	MAPE	$U$	$K$
AR(2)	148,07	0,66	0,01	–
KF ( $i = 0,5$ )	36,39	0,36	0,01	0,73
<b>KF (<math>r = 1</math>)</b>	<b>22,04</b>	<b>0,26</b>	<b>0,01</b>	<b>0,61</b>
AKF	24,49	0,27	0,01	0,60
AKFMW 100	24,01	0,27	0,01	0,61
AKFMW 50	23,15	0,26	0,01	0,62
AKFMW 20	22,90	0,26	0,01	0,64
AKFMW 10	24,15	0,26	0,01	0,65
AKFMW 5	27,76	0,26	0,01	0,68
AKFMW 3	30,86	0,26	0,01	0,70

Table 2. Three-step forecast

Model type	June 2014		July 2014		August 2014	
	Value	%	Value	%	Value	%
Data	101,00		100,40		100,80	
Model	103,74	2,71	103,72	3,31	103,70	2,88
<b>KF (<math>r = 0,5</math>)</b>	<b>102,60</b>	<b>1,59</b>	<b>102,58</b>	<b>2,18</b>	<b>102,57</b>	<b>1,75</b>
KF ( $i = 1$ )	103,22	2,19	103,20	2,79	103,18	2,36
AKF	103,26	2,23	103,24	2,83	103,22	2,40
AKFMW 100	103,29	2,26	103,27	2,86	103,25	2,43
AKFMW 50	103,32	2,30	103,31	2,90	103,29	2,47
AKFMW 20	103,39	2,37	103,38	2,97	103,36	2,54

Table 2 (continued)

AKFMW 10	103,30	2,27	103,28	2,87	103,26	2,44
AKFMW 5	103,25	2,22	103,23	2,82	103,21	2,39
AKFMW 3	103,47	2,44	103,46	3,04	103,44	2,61

The Kalman filter characteristics of functioning are provided in Figs. 3–5.

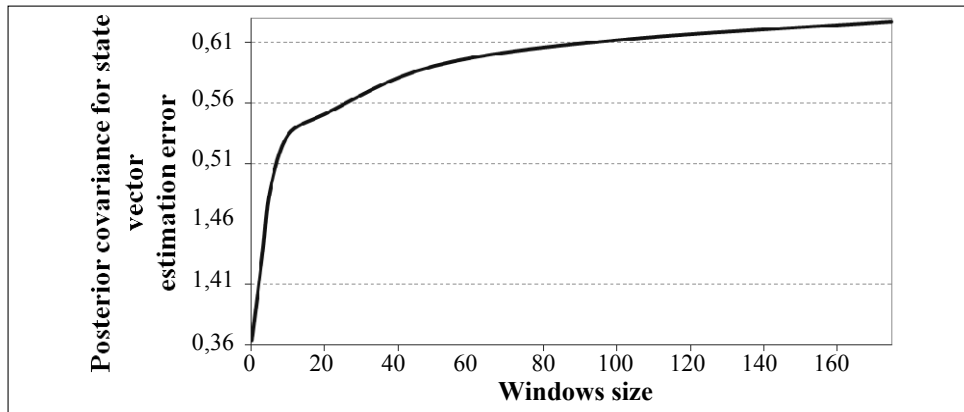


Fig. 3 Posterior covariance for state vector estimation error versus AKFMW window size

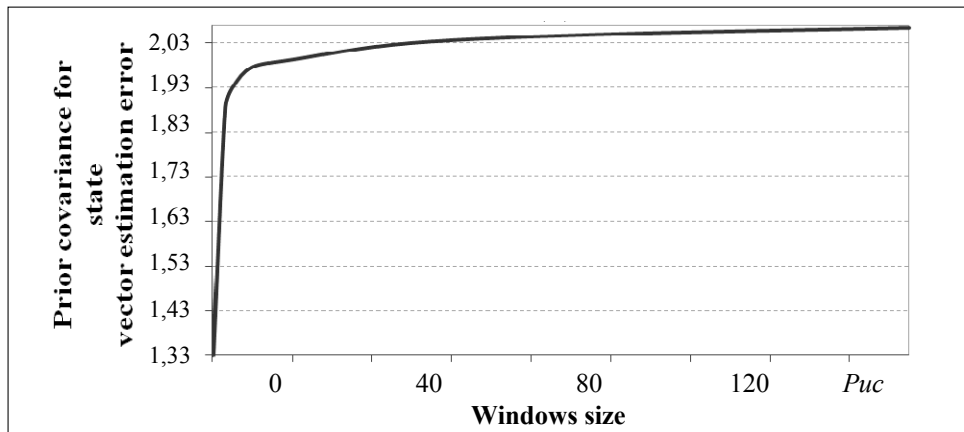


Fig. 4. Prior covariance for state vector estimation error versus AKFMW window size

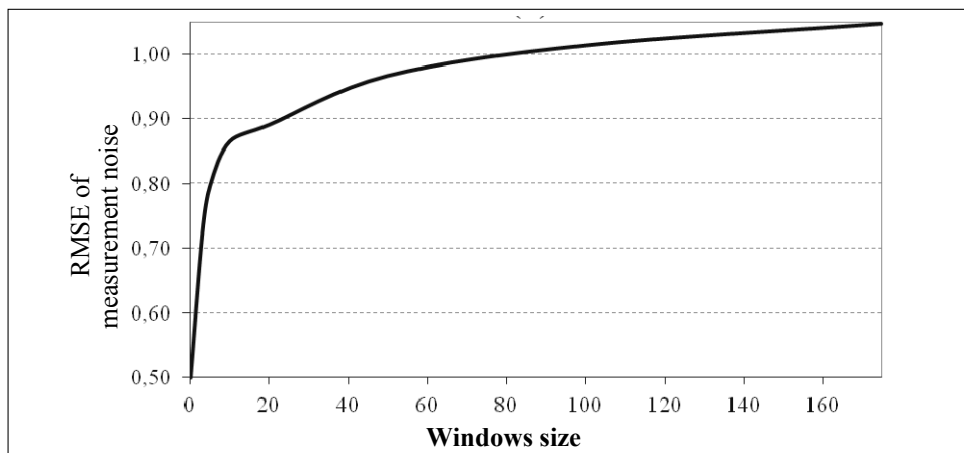


Fig. 5. RMSE of measurement noise versus AKFMW

Thus, the computational experiments performed support the idea that application of optimal filtering provides a possibility for improvement of short term forecasts using state space models. Different versions of optimal KF usually show somewhat different results though in general final result of forecasting is quite acceptable.

## CONCLUSIONS

An adaptive modeling concept is proposed based on application of simultaneous model structure and parameters estimation strategy and adaptive Kalman filter. Using statistical data for the consumer price index in Ukraine an autoregression AR(2) model was constructed for short term forecasting. To compute optimal state vector estimates two versions of Kalman filter have been applied: ordinary optimal KF and adaptive KF. The latter version was constructed with hiring moving data windows for estimation of measurement noise covariances. The moving window size has been varied in a wide range from 3 to 100. The best forecasting results were received with ordinary optimal KF with noise variance  $R=0,5$ ; and lower results were received with model AR(2) without filtering procedure.

The best one step ahead forecast was received with KF that uses  $R=1,0$ , what can be explained by higher information quality at this period of time or on purpose decreasing of CPI. Somewhat lower results of forecasting with adaptive filter could be explained by the process stationarity at the period of time examined. During such periods the adaptation is not required.

Testing of the forecasting system with wide set of macroeconomic and stock price data showed that it is possible easy to reach a value of mean absolute percentage error of about 3–4 % for short term forecasting. The use of dynamic and static estimates allows for generating necessary forecasts estimates depending on specific problem statement.

In the future studies it will be reasonable to incorporate into the software developed new forecasting techniques such as neural networks, probabilistic Bayesian networks, and immune algorithms that cover a wide class of nonlinear processes. Further steps towards automation of a model development and selection procedure are also necessary.

## REFERENCES

1. *Holsapple C.W., Winston A.B.* Decision Support Systems (a knowledge based approach). — New York: West Publishing Company, 1994. — 860 p.
2. *Бідюк П.І.* Адаптивне прогнозування фінансово-економічних процесів на основі принципів системного аналізу // Наукові Вісті НТУУ «КПІ», 2009. — № 5. — С. 54–61.
3. *Gibbs B.P.* Advanced Kalman filtering, least-squares and modeling. — New York: John Wiley & Sons, Inc., 2011. — 627 p.
4. Державна служба статистики України. — Електрон. дані. — <http://www.ukrstat.gov.ua/>.
5. *Grewal M.S., Andrews A.P.* Kalman filtering: theory and practice using Matlab. — Hoboken: John Wiley & Sons, Inc., 2008. — 575 p.

*Надійшла 15.03.2015*

From the Editorial Board: the article corresponds completely to submitted manuscript.

## МУЛЬТИФРАКТАЛЬНЫЙ ФЛУКТУАЦИОННЫЙ АНАЛИЗ СТЕГАНОГРАММ

Д.А. ПРОГОНОВ, С.Н. КУЩ

Рассмотрено использование мультифрактального флуктуационного анализа для определения типов преобразований, использованных при встраивании стегоданных в цифровые изображения. Показано, что применение как стандартных (дискретное вейвлет преобразование), так и специальных (сингулярное разложение) преобразований цифровых изображений при скрытии сообщений приводит к существенным изменениям фрактальных параметров флуктуаций яркости пикселей (ширины мультифрактального спектра). Характерные изменения результатов мультифрактального флуктуационного анализа стегограмм по сравнению с незаполненным контейнером дают возможность определять тип преобразований, использованных при встраивании стегоданных в области преобразования контейнера. Установлено, что в случае использования стандартных преобразований, применение мультифрактального флуктуационного анализа позволяет дополнительно оценить степень детализации встроженных стегоданных. Полученные результаты могут быть использованы для повышения эффективности современных методов активного стегоанализа изображений.

### ВВЕДЕНИЕ

Модернизация классических методов цифровой стеганографии, а также разработка новых методов, основанных на использовании различных преобразований, как контейнеров, так и стегоданных при формировании стегограмм, существенно усложняют проведение активного стегоанализа (АС). Определение типа преобразований, использованных для встраивания стегоданных в контейнер — цифровое изображение (ЦИ) — позволит выбрать оптимальный метод АС, обеспечивающий максимальную деструкцию стегоданных при минимальных искажениях контейнера.

**Цель работы** — определение типа преобразований, использованных для встраивания стегоданных в области преобразования контейнера, методом мультифрактального флуктуационного анализа.

### ВСТРАИВАНИЕ СТЕГОДАНЫХ В ЦИФРОВЫЕ ИЗОБРАЖЕНИЯ С ИСПОЛЬЗОВАНИЕМ РАЗЛИЧНЫХ ТИПОВ ПРЕОБРАЗОВАНИЙ

Преимуществом скрытия сообщений в области преобразования контейнера (ОПК), по сравнению с классическими методами встраивания стегоданных в пространственной области ЦИ (LSB-методами) [1–3], является их более высокая устойчивость при проведении как пассивного, так активного стегоанализа [4].

Известные методы скрытия сообщений в ОПК можно разделить на две группы [5, 6]. К первой группе относятся стегограммы, сформированные

с использованием стандартных, например, двумерного дискретного вейвлет-преобразования (ДДВП) ЦИ, а ко второй — специальных преобразований, например, сингулярного разложения (СР) ЦИ. Преимуществом использования стандартных преобразований является высокая робастность получаемых стеганограмм к существующим методам АС изображений, а применение специальных преобразований позволяет снизить эффективность известных методов пассивного стегоанализа ЦИ.

В работе рассмотрены одноэтапные методы скрытия данных в ОПК с использованием как стандартных (метод Дея), так и специальных (метод Агарваля) типов преобразований ЦИ. Исследование проводилось на тестовом пакете из 100 полноцветных ЦИ с разрешением UHD-4K (3820×2180 пикселей). В качестве стегоданных были использованы три полноцветные изображения с различной степенью детализации: портрет (565×850 пикселей), чертеж (567×463 пикселей) и карта (800×800 пикселей). Степень заполнения контейнера стегоданными варьировалась от 5% до 25% с шагом 5% и от 25% до 95% с шагом 10%.

Встраивание стегоданных в области преобразования ЦИ с использованием ДДВП проводилось по методу Дея [5]. Прямое и обратное ДДВП для отдельных каналов цвета изображений рассчитывались согласно [7]. В качестве базисных функций преобразования были выбраны вейвлет Хаара и соответствующая ему скейлинг-функция.

Согласно методу Дея, ДДВП применялось к каждому каналу цвета изображения-контейнера и стегоданных, после чего производилось суммирование полученных коэффициентов разложения. Для вариации степени «вклада» скрываемых сообщений был использован масштабирующий множитель  $\beta$  на который умножались коэффициенты ДДВП стегоданных [5]. Значения весового коэффициента  $\beta$  варьировались в диапазоне от  $\beta_{\min} = 0,02$  до  $\beta_{\max} = 0,08$  с шагом  $\Delta_{\beta} = 0,02$ , где  $\beta_{\min}$  соответствует появлению значительных искажений при экстракции стегоданных, а  $\beta_{\max}$  — возникновению видимых искажений стеганограммы.

Скрытие данных в ЦИ с использованием сингулярного разложения (СР) ЦИ, рассмотрено на примере метода Агарваля [6]. Сингулярное разложение отдельного канала цвета ЦИ  $I(x, y)$  проводилось согласно выражению [8]:

$$B = \sum_{i=1}^d \sqrt{\lambda_i} \cdot U_i V_i^T, \quad (1)$$

где  $\sqrt{\lambda_i}$  — сингулярные числа (СЧ)  $I(x, y)$ ;  $U_i, V_i$  — соответственно, левый и правый сингулярные вектора (СВ) выбранного канала цвета ЦИ.

Особенностью использования СР при встраивании стегоданных в ЦИ является независимость результатов разложения (сингулярных чисел) от вида аффинных преобразований стеганограмм (сдвига, вращения, масштабирования) [8].

Преимуществом метода Агарваля, по сравнению с другими методами, является возможность распределенной передачи СЧ и СВ, полученных согласно (1), с использованием нескольких контейнеров (мультиагентная мо-



дель). Как и для метода Дея, вариация степени искажения ЦИ при встраивании стегоданных обеспечивалась применением весового параметра  $\beta$ .

## КОРРЕЛЯЦИОННЫЕ И ФРАКТАЛЬНЫЕ ХАРАКТЕРИСТИКИ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Существующие методы пассивного стегоанализа основаны на оценке степени отличия анализируемого изображения относительно известной модели чистого контейнера [2, 3]. В этом случае широко используются статистический анализ ЦИ и специальные методы, применяемые для исследования тонкой структуры изображения, в частности флуктуаций яркости пикселей.

Одним из современных видов анализа шумовых компонент дискретных сигналов является мультифрактальный флуктуационный анализ (МФФА). Он является развитием классического флуктуационного анализа (КФА) и позволяет исследовать как корреляционные, так и фрактальные характеристики флуктуаций значений яркости пикселей исследуемого сигнала [9]. Схема алгоритма МФФА ЦИ представлена на рис. 1.

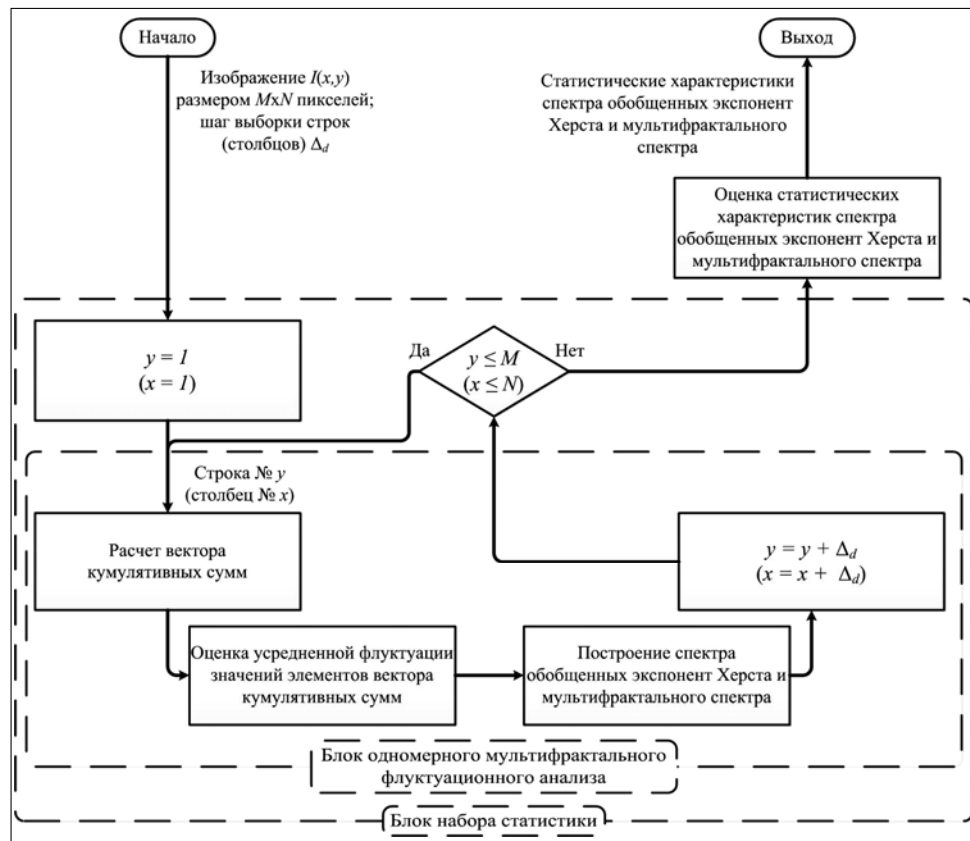


Рис. 1. Схема алгоритма мультифрактального флуктуационного анализа цифровых изображений

При проведении исследований, исходное изображение  $I(x,y)$  размером  $M \times N$  (пикселей) обрабатывалось построчно с шагом выборки  $\Delta_d$ .

На первом этапе обработки ЦИ, в блоке одномерного МФФА производился расчет вектора кумулятивных сумм (рис. 1) исследуемой строки изображения [9]:

$$C(i) = \sum_{k=1}^i (I(y, :) - \langle I(y, :) \rangle), \quad i \in \{1, 2, \dots, N\}, \quad y \in \{1, 2, \dots, M\},$$

где  $I(y, :)$  —  $y$ -я обрабатываемая строка заданного ЦИ;  $\langle \cdot \rangle$  — операция усреднения.

Полученный вектор  $C$  разбивался на  $N_s = [N/s]$  смежных блоков шириной  $s$  пикселей, где  $[\cdot]$  — операция округления к ближайшему большему целому числу.

Для оценки вариаций  $F^2(s, v)$  использовались стандартные методы:

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^s \{C[(v-1)s + i] - Q_v(i)\}^2, \quad v \in \{1, \dots, N_s\},$$

где  $Q_v$  — полином, применяемый для компенсации детерминированных компонент кумулятивных сумм в  $v$ -м блоке вектора  $C$ . Далее рассчитывалась усредненная вариация  $F_q(s)$  значений вектора кумулятивных сумм  $C$ :

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} (F^2(s, v))^{\frac{q}{2}} \right\}^{\frac{2}{q}}, \quad q \in \mathbb{R} \setminus \{0\},$$

где  $q$  — масштабирующий параметр.

Значения обобщенных экспонент Хёрста (ОЭХ)  $h_q$  рассчитывались путем линейной аппроксимации зависимости  $F_q = F(s^{h_q})$  согласно методу наименьших квадратов и использовались при построении спектра ОЭХ и мультифрактального спектра (МФС), которые связаны между собой преобразованием Лежандра:

$$f(s^{h_q}) = q(\alpha - h_q) + 1, \quad \alpha = h_q + q \frac{d(h_q)}{dq},$$

где  $\alpha$  — масштабирующая экспонента, используемая для оценки размерности Хаусдорфа  $f(\alpha)$  подмножества флуктуаций значений вектора кумулятивных сумм  $C$ , которое характеризуется вероятностью  $p$  заполнения блоков разбиения вектора  $C$  —  $p \propto s^\alpha$ .

При проведении МФФА ЦИ значения параметра  $q$  принимали дискретные значения в интервале от  $q_{\min} = (-20)$  до  $q_{\max} = 20$  с шагом  $\Delta_q = 0,25$ . Построение спектра ОЭХ и МФС проведено путем сплайн-интерполяции с использованием рассчитанных дискретных значений  $h_q$  и  $f(\alpha)$ . Расширение диапазона значений параметра  $q$  является нецелесообразным, поскольку изменения ОЭХ  $h_{q_{\min}}$  и  $h_{q_{\max}}$  не превышают  $10^{-3}$ .

Для определения оптимального шага выборки строк ЦИ  $\Delta_d^{\text{opt}}$  были исследованы изменения интегрального отклонения  $\delta$  результатов МФФА (спектра ОЭХ) при вариации шага выборки  $\Delta_d$  относительно случая обработки каждой строки ЦИ ( $\Delta_d = 1$ ):

$$\delta(\Delta_d) = \sum_{i=1}^N |S_1(i \cdot \Delta_d) - S_{\Delta_d}(i \cdot \Delta_d)|,$$

где  $S_k$  — спектр ОЭХ, полученный при  $\Delta_d = k$ ;  $N = (q_{\text{max}} - q_{\text{min}}) / \Delta_d$  — количество элементов спектра ОЭХ. Зависимость отклонения  $\delta$  и времени расчета спектра ОЭХ отдельной строки ЦИ от шага выборки  $\Delta_d$  представлена на рис. 2.

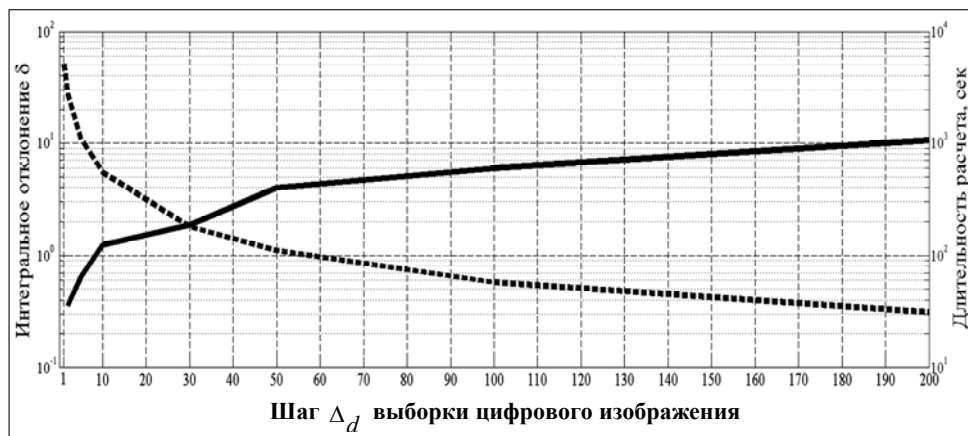


Рис. 2. Зависимость интегрального отклонения  $\delta$  и времени расчета спектра ОЭХ отдельной строки изображения от шага выборки  $\Delta_d$ : интегральное отклонение  $\delta$  — левая ось ординат, непрерывная линия; время расчета спектра ОЭХ — правая ось ординат, штриховая линия

Значение шага выборки  $\Delta_d = 1$  позволяет минимизировать интегральное отклонение  $\delta$  ( $\delta = 0|_{\Delta_d=1}$ ), но приводит к существенному возрастанию времени обработки всего изображения (рис. 2). Вследствие этого шаг  $\Delta_d^{\text{opt}}$  был выбран равным 30, что дает возможность минимизировать как интегральное отклонение  $\delta$ , так и время анализа ЦИ [10].

## РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

В данной работе, используя результаты расчетов МФФА, произведена оценка средних значений параметров спектра ОЭХ и МФС — показателя Хёрста ( $h_2 \equiv H$ ), ширины мультифрактального спектра ( $\Delta_\alpha = \alpha_{\text{max}} - \alpha_{\text{min}}$ ) и других. Дисперсия полученных результатов не превышает 15% относительно средних значений мультифрактальных характеристик шумовых компонент ЦИ.

Были исследованы спектры ОЭХ и МФС при встраивании стегоданных как в пространственной области (LSB-методы), так и в области преобразования контейнера. Примеры спектров ОЭХ и МФС для пустого контейнера и стеганограмм, сформированных согласно методам Дея и Агарваля, а также LSB-метода, при двадцатипроцентном заполнении контейнера стегоданными, представлены на рис. 3.

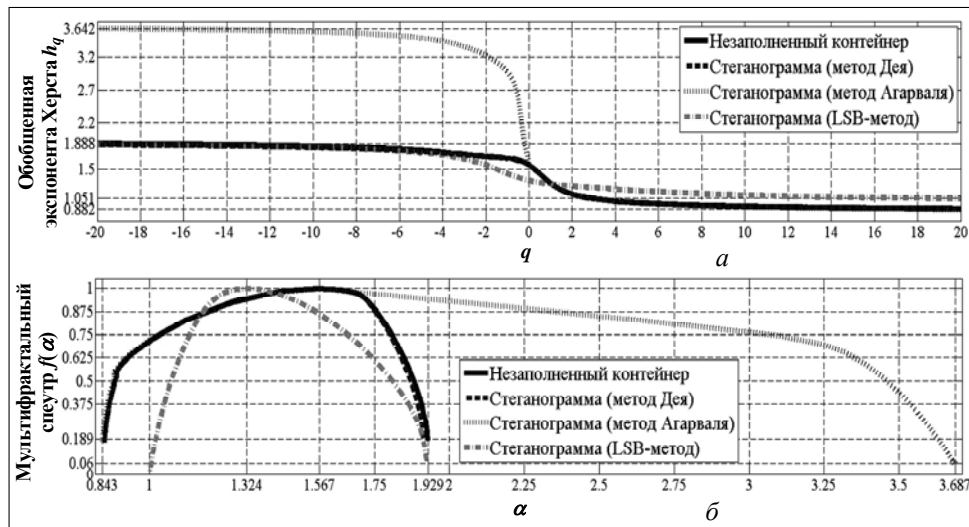


Рис. 3. Результаты мультифрактального флуктуационного анализа незаполненного контейнера и стеганограмм, сформированных согласно методам Дея и Агарваля, а также LSB-метода при 20% степени заполнения: *a* — спектр обобщенных экспонент Хёрста; *b* — мультифрактальный спектр

Встраивание данных в ОПК ЦИ приводит к существенным изменениям корреляционных и фрактальных характеристик флуктуаций яркости пикселей ЦИ с малой амплитудой (рис. 3). При использовании метода Дея и Агарваля происходит, соответственно, сужение и расширение спектра МФС (рис. 3б).

На рис. 3 также приведены результаты МФФА для случая встраивания стегоданных в пространственной области ЦИ согласно LSB-методу. Характер изменений параметров шумов изображений, при использовании LSB-метода, существенно отличается от случая применения методов Дея и Агарваля. Для LSB-метода характерно сглаживание спектра ОЭХ (рис. 3а), смещение и сужение МФС (рис. 4б), что дает возможность определить область скрытия сообщений.

В работе определены зависимости значений показателя Хёрста (ПХ) от степени заполнения контейнера для трех ЦИ стегоданных (рис. 4).

Как видно из рис. 4, даже в случае сильного заполнения контейнера (более 75%), изменения значений показателя Хёрста (ПХ)  $\Delta_H$  являются незначительными как при использовании стандартных ( $\Delta_H = 0,03$ ), так и специальных ( $\Delta_H = 0,05$ ) типов преобразований ЦИ. Вследствие этого значительно снижается эффективность использования данного критерия для распознавания области встраивания стегоданных.

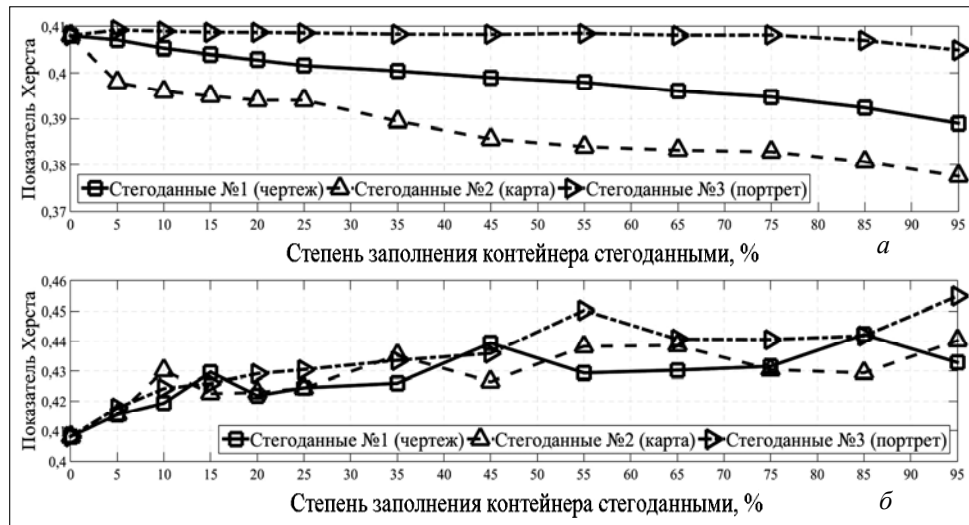


Рис. 4. Зависимости значений показателя Хёрста от степени заполнения контейнера стегоданными для случая скрытия сообщений с использованием метода Дея (а) и метода Агарвала (б)

В связи с низкой эффективностью применения ПХ для выявления стеганограмм и распознавания типа использованных преобразований при скрытии сообщений, были исследованы фрактальные свойства шумов ЦИ (ширина  $\Delta_\alpha$  МФС). Зависимость ширины МФС от степени заполнения контейнера стегоданными представлена на рис. 5.

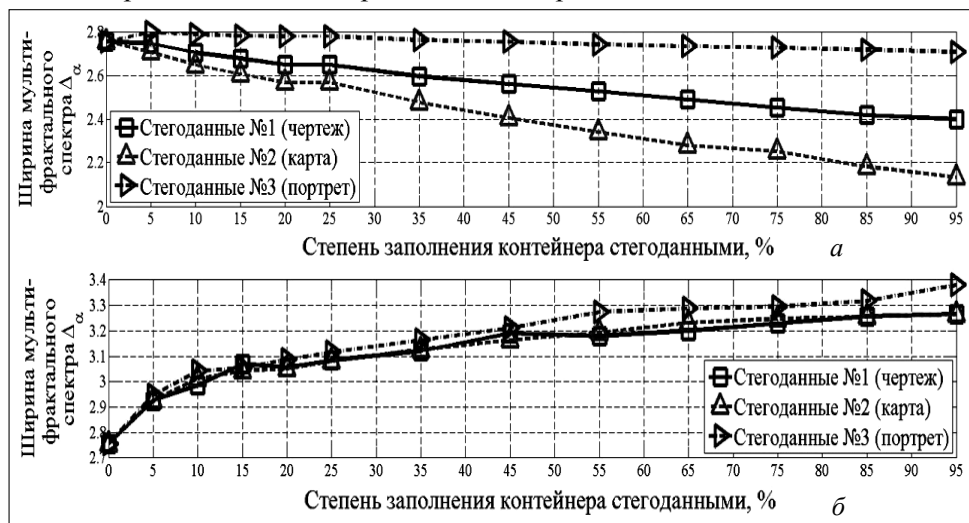


Рис. 5. Зависимость ширины мультифрактального спектра флуктуаций яркости пикселей стеганограмм от степени заполнения контейнера стегоданными и скрытия сообщений с использованием метода Дея (а) и метода Агарвала (б)

Встраивание данных в ОПК приводит к существенному изменению ширины МФС  $\Delta_\alpha$  (рис. 5). В случае использования метода Дея (рис. 5а) происходит уменьшение значений  $\Delta_\alpha$  (на 0,4 в случае 95%), а в случае метода Агарвала (рис. 5б) — увеличение ширины МФС (на 0,6 в случае 95%) относительно случая незаполненного контейнера. Различие в тенденциях

изменений параметра  $\Delta_\alpha$  при использовании стандартных и специальных типов преобразований ЦИ дает возможность распознавать область встраивания стегоданных в изображения.

По результатам анализа полученных данных выявлена зависимость значения параметра  $\Delta_\alpha$  при встраивании стегоданных согласно методу Дея от степени детализации стегоданных: при возрастании количества мелких деталей в стегоданных происходит уменьшение отклонения ширины МФС относительно случая незаполненного контейнера (рис. 5а). Данная особенность обусловлена тем, что при возрастании количества небольших деталей на встраиваемых изображениях происходит перераспределение энергии пикселей в области высоких частот.

## ВЫВОДЫ

В результате исследований возможности применения мультифрактального флуктуационного анализа для определения типа преобразований ЦИ, использованных при встраивании стегоданных в контейнеры на основе одноэтапных методов скрытия в ОПК, установлено, что:

1. Использование классического флуктуационного анализа для определения типа преобразования, использованного при встраивании стегоданных в контейнер, из-за слабых изменений показателя Хёрста является неэффективным.
2. На основе анализа результатов МФФА стеганограмм, возможно определение области преобразования и типа преобразования, использованного для встраивания стегоданных в ОПК.
3. В результате анализа ширины мультифрактального спектра возможно проведение оценки степени детализации стегоданных, встроенных в ОПК, при использовании стандартных преобразований ЦИ.

Полученные результаты могут быть использованы для повышения эффективности современных методов активного стегоанализа цифровых изображений.

## ЛИТЕРАТУРА

1. *Конахович Г., Пузыренко А.* Компьютерная стеганография. Теория и практика. — К.: МК-Пресс, 2006. — 288 с.
2. *Ramkumar M.* Data Hiding in Multimedia. Theory and Applications. — Doctoral Dissertation, Department of ECE, New Jersey Institute of Technology. — University Heights, 1999. — 70 p.
3. *Katzenbeisser S., Petitcolas P.* Information Hiding Techniques for Steganography and Digital Watermarking. — Boston: Artech House, 2000. — 237 p.
4. *Черепанова К.В., Прогонов Д.О., Куц С.М.* Активний стегоаналіз повідомлень, прихованих в цифровому зображенні // Матеріали XII Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених «Теторетичні та прикладні проблеми фізики, математики та інформатики». — К.: ВПК «Політехніка», 2014. — С. 219–221.

5. Dey N.A., Roy A.B., Dey S. novel Approach of Color Image Hiding using RGB Color Planes and DWT. // International Journal of Computer Applications. — 2011. — **36**, № 5. — P. 19–24.
6. Agarwal R., Santhanam M.S. Digital watermarking in the singular vector domain // International Journal of Image and Graphics. — 2008. — **8**. — P. 351–362.
7. Гонсалес Р., Вудс Р. Цифровая обработка изображений / Изд. 3-е, исправленное и дополненное. — М.: Техносфера, 2012. — 1104 с.
8. Голяндина Н.Э. Метод «Гусеница»-SSA: Анализ временных рядов. — Учеб. пособие. — СПб., 2004. — 76 с.
9. Kantelhardt J., Zschiegner S., Koscielny-Bunde E., Bunde A., Havlin S., Stanley E. Multifractal Detrended Fluctuation Analysis of Nonstationary Time Series // Physica A. — 2002. — **316**, Issue 1–4. — P. 87–114.
10. Прогонов Д.О., Куц С.М. Виявлення стеганограм з даними, прихованими в області перетворення цифрових зображень // Вісник Національного технічного університету України «Київський політехнічний інститут». Серія — Радіотехніка. Радіоапаратобудування. — 2014. — № 57. — С. 128–142.

*Поступила 01.06.2014*

## ПОРІВНЯННЯ ТИПІВ АРХІТЕКТУРИ СИСТЕМ СЕРВІСІВ

О.О. ПЕТРЕНКО

Розглянуто сучасні архітектури систем сервісів SOA (service-oriented architecture — сервісно-орієнтована архітектура) й EDA (event-driven architecture - подійно-орієнтована архітектура), їх переваги і вади, можливість і доцільність побудови об'єднаної сервісно-орієнтованої архітектури EDSOA (event-driven service-oriented architecture — подійно-керована сервісно-орієнтована архітектура). Показано, що події з'єднують сервіси за допомогою передачі стану бізнес-процесу від одного сервісу, який визначає і публікує події, до інших сервісів, які запускаються конкретними подіями. В свою чергу, обґрунтовано, що сервіси об'єднують події за допомогою передачі даних про перехід одного стану процесу в інший. Особливу увагу в роботі приділено питанням ефективної реалізації запропонованого гібридного рішення архітектури EDSOA до моделювання бізнес-процесів як сервісів.

### ВСТУП

Нещодавно з ініціативи фірми IBM було сформовано науку про сервіси, менеджмент і SSME (service science management and engineering — наука про сервіси, менеджмент і інжиніринг), що покликана дослідити основні принципи функціонування складних систем сервісів, шляхи створення, масштабування і вдосконалення таких систем. Варто особливо підкреслити, що розвиток науки про сервіси спирається на два таких відомих технічних нововведення в інформаційних технологіях, як технологія SaaS, коли програмне забезпечення використовується і орендується через інтернет як хмарний сервіс, і SOA (service-oriented architecture — сервісно-орієнтована архітектура) як базовий стиль архітектури у процесі проектування складних розподілених систем.

Масштаб, складність і взаємозалежність сучасних систем сервісів у зв'язку з глобалізацією, демографічними змінами та технологічними розробками досягли безпрецедентного рівня. У найбільш розвинених країнах більше 70% ВВП формується індустрією сервісів, у якій на сьогодні зайнято (за інформацією Міжнародної організація праці) більше половини людства [1, 2]. Сьогоднішня глобальна економіка вимагає від компаній розширення своїх бізнес-процесів за межі організацій та інтеграції бізнесу з партнерами. Необхідність інтеграції та взаємодії додатків у рамках сукупності великої кількості інформаційних систем самого підприємства або кількох підприємств, об'єднаних у цілий партнерський ланцюжок, також справляють істотний вплив на доцільний вибір архітектури систем сервісів.

Поняття «сервіс» і «процес» є взаємозалежними і їх можна використовувати на різних рівнях узагальнення. Наприклад, невеликий процес може бути організований у вигляді окремого сервісу у тому випадку, якщо його можна *типізувати*. У той же час процес може бути розбито на окремі сервіси, що взаємодіють між собою в рамках процесу. Якщо сервісів буде дуже



багато, то не буде досягнуто необхідної типізації, але якщо сервіси будуть занадто високорівневими, то їх застосування буде ускладнено через специфічні відмінності у бізнес-процесах. Відображаючи логіку бізнесу у формі веб-сервісів, можна формувати потоки завдань (workflows), спеціально налаштованих з урахуванням потреб підприємства.

Термін «сервіс» має два основних загально прийнятих значення: економічне, бізнес-орієнтоване значення, наприклад, як у виразі сервісний сектор економіки, IT-орієнтоване значення, веб-сервіс або SOA. У першому випадку акцент робиться на взаємодію у процесі обміну і на нематеріальний характер сервісу, а у другому — на технології програмного забезпечення, яке дозволяє підтримувати сумісність різних програмних модулів або агентів. У сучасних систем сервісів для економіки два значення зливаються, оскільки впровадження сучасної системи сервісів передбачає також інтеграцію інформаційних систем як підсистем організаційної системи систем. У той же час зростає сервісний домен з орієнтацією на сучасні виробничі процеси, наприклад, через «сервітизацію» виробництва [3]. Ці сервісо-орієнтовані процеси також сприяють інтеграції виробничих процесів з інформаційними системами і конкретними IT-технологіями.

Більшість сучасних знань було розроблено в період виробничої економіки, зокрема, в галузі систем управління технічними об'єктами за допомогою IT-технологій. Індустрія послуг потребує створення своєї наукової бази, розроблення методик й інструментарію для розроблення систем сервісів, розгортання підготовки відповідних кадрів, спроможних забезпечити подальше поширення і зміцнення індустрії послуг.

## **SOA**

SOA замість монолітної пропонує блокову систему із взаємодіючих компонентів, у якій різні функціональні модулі додатків (сервіси), призначено для *управління бізнес-процесами*, що пов'язані між собою за допомогою чітко визначених інтерфейсів. Інтерфейси самі по собі незалежні від оточення і платформи, і тому така модель отримала назву моделі «слабкого зв'язку». Фактично суть концепції SOA полягає в уніфікації та автоматизації бізнес-процесів за допомогою типових компонентів — сервісів (наприклад, веб-сервіси використовують один стандарт — розширювану мову розмітки XML). При цьому створення, впровадження або зміна бізнес-процесу є компоновкою (оркестровкою) раніше розроблених сервісів, призначених для автоматизації бізнес-функцій.

Для бізнесу SOA означає збільшення задоволеності клієнтів, реальну гнучкість бізнесу, швидкий час виходу на ринок, простоту співробітництва і низьку вартість бізнесу. Для IT-організацій SOA означає більшу продуктивність, зниження витрат на IT за рахунок прискорення розробки додатків, більш м'якого повторного використання сервісів, більш високої якості додатків, і в цілому більш швидкого реагування на запити бізнес-клієнтів для поліпшення і модифікації системи. На додаток до цього існує можливість використання сервісів незалежних постачальників, що забезпечує ще більшу цінність SOA. Слово «гнучкість» часто згадується у ході обговорення переваг SOA і може бути інтерпретовано, з одного боку, як можливість

змінювати бізнес-процеси відповідно до змін вимог ринку і вимог клієнтів і, з другого, як здатність виконувати бізнес-процеси швидше або запускати швидше нові процеси, продукти і сервіси. Гнучкість і швидкість — реальні й відчутні переваги переходу на SOA і багаторазове використання сервісів.

SOA є архітектурним підходом, який дозволяє розкласти функціональність додатків на множину сервісів, розміщених у мережі. Веб-сервіси є технологію реалізації концепції дизайну SOA. Існує велика кількість досліджень, присвячених вимогам надійності в SOA і, більш конкретно, веб-сервісів. Веб-спільнота розробила ряд специфікацій, які підтримують надійний обмін повідомленнями, управлінням транзакціями, реплікаціями і безпекою. Сервіси взаємодіють один з одним за допомогою шаблону синхронного запиту і відповіді, що забезпечує щільне зчеплення між компонентами системи (рис. 1).

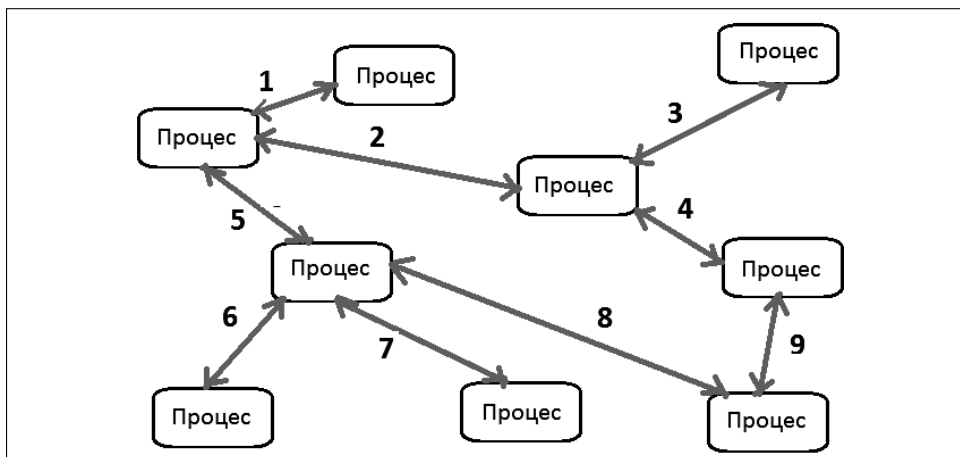


Рис. 1. Архітектура SOA, керована запитом, де 1–9 — послідовність запитів

До базових принципів SOA треба віднести [4]:

- **Слабке зв'язування сервісів** — одна з найбільших переваг SOA. Слабко зв'язані модулі мають кілька добре відомих залежностей, а тісно пов'язані модулі мають багато невідомих залежностей. У суперечність зі слабким зв'язком між сервісами і можливістю їх повторного використання застосування сервісу може бути ґрунтовним і щільним.

- **Стандартизація** — стандарти SOA відкриті в тому сенсі, що будь-який виробник програмного забезпечення має право використовувати ці стандарти у процесі розробки архітектури SOA. Крім того, процес створення та перегляду стандартів є більш-менш демократичним, де будь-яка зацікавлена сторона має право брати участь у всіх засіданнях, які призводять до рішень про стандарт.

- **Модульність** — SOA реалізує сервіси, що підтримують чітко визначені модульні бізнес-функції та інформаційні процеси. Ці модулі можуть надалі бути використані повторно і бути частиною бізнес-процесу. Модульність сервісів призводять до поділення додатків на безліч дрібних модулів. Кожен модуль відповідає за одну окрему функцію у додатку.

- **Скомпонованість** — здатність ефективно складати сервіси є найважливішою вимогою для досягнення деяких із найбільш фундаментальних ці-

лей сервіс-орієнтованих застосувань. Сервіси, які складаються, здатні брати участь у якості ефективних модулів. Модульна сервісна структура дозволяє створювати нові інформаційні системи, про які розробники сервісів можуть не мати жодного уявлення під час проектування сервісів.

- **Повторюваність** — повторне використання сервісів як основна частина аналізу обслуговування та процесів проектування систем сервісів. Згідно з цим принципом сервіс може бути використаний більш ніж за одним сценарієм у різних бізнес-процесах або інформаційних системах.

- **Відкритість** — сервіси мають бути легко ідентифіковані й зрозумілі, щоб забезпечити можливість їх повторного використання. Тому у процесі проектування сервісів необхідно враховувати «якість доступу» до сервісів та їх індивідуальні властивості незалежно від того, чи вони зареєстровані в депозитарії, чи ні.

- **Абстракція** — цей принцип підкреслює необхідність приховувати так багато з основних деталей сервісів, як це можливо. Це забезпечує безпосередньо описане раніше слабке зв'язування сервісів.

- **Зв'язок «один-на-один»** — один із конкретних сервісів викликається у часі тільки одним із споживачів, але зв'язок є двонаправленим.

- **Синхронність** — відповіді на запити відправляються назад до споживача в синхронному режимі.

- **Запуск процесу** — потік управління зніціюється клієнтом (споживачем послуг).

- **Зернистість сервісів** — рівень деталізації обслуговування. Сервіси в SOA є модулями бізнес-логіки досить високого рівня, завдяки чому взаємодія між ними зводиться до обмеженого числа повідомлень за змістом бізнес-логіки замість безлічі низькорівневих викликів, що враховують деталі реалізації сервісів. Такий підхід знижує навантаження на мережу і сприяє більш високій продуктивності системи.

- **Відсутність стану** — сервіси проектуються так, щоб залишитися зі станом тільки за необхідності. Сервісам не варто покладатися на тривалий зв'язок між споживачем і постачальником, вони не мають також покладатися на попередні виклики.

Орієнтація на сервіси та SOA може краще використовувати тоді, коли процеси або їх частини є стандартними, коли вони часто повторюються без змін, або коли декільком користувачам потрібно той же компонент процесу для виконання своїх завдань. Виклик (споживання) сервісів у SOA реалізується віддалено за допомогою віддаленого виклику процедури RPC (remote procedure call — віддалений виклик процедур) на вимогу споживача сервісів. SOA добре зарекомендувала себе для побудови великих корпоративних програмних додатків. Ціла низка розробників та інтеграторів пропонують інструменти і рішення на основі SOA (наприклад, платформи Intel SOA Expressway, JBoss SOA Platform, IBM WebSphere, Software AG webMethods, Oracle, BEA Aqualogic, Microsoft Windows Communication Foundation, SAP NetWeaver, TIBCO).

Однак практичне застосування всього потенціалу SOA ускладнюється через часту необхідність використовувати окремі сервіси з несумісними інтерфейсами. Крім того, у бажанні створення підприємств реального часу, які

постійно підключені й завжди доступні в інтернеті, організації стикаються з більш різноманітними бізнес-сценаріями і виявляють потреби в альтернативних шаблонах проектування на доповнення до синхронної керованої запитом SOA.

## EDA

EDA (event-driven architecture — подійно-орієнтована архітектура) містить в собі три базові компоненти: генератор події (датчик), оброблювач події і менеджер подій (відповідач). Менеджер подій реєструє всі події, які виникають в системі (зовнішні і внутрішні), відповідним чином ідентифікує і передає оброблювачу подій. Якщо потрібний оброблювач за яких-небудь причин не доступний, подія в залежності від конкретної реалізації або ставиться в чергу, або передається іншому оброблювачу. Завдяки такій схемі система стає максимально гнучкою і чутливою до змін інформаційного середовища — у разі виникнення принципово нової події досить підключити новий оброблювач, не зачіпаючи ядра системи. Подібні системи зазвичай будуються на тригерах, що реагують на певні події та ініціюють відповідні веб-сервіси, рис. 2 [5].

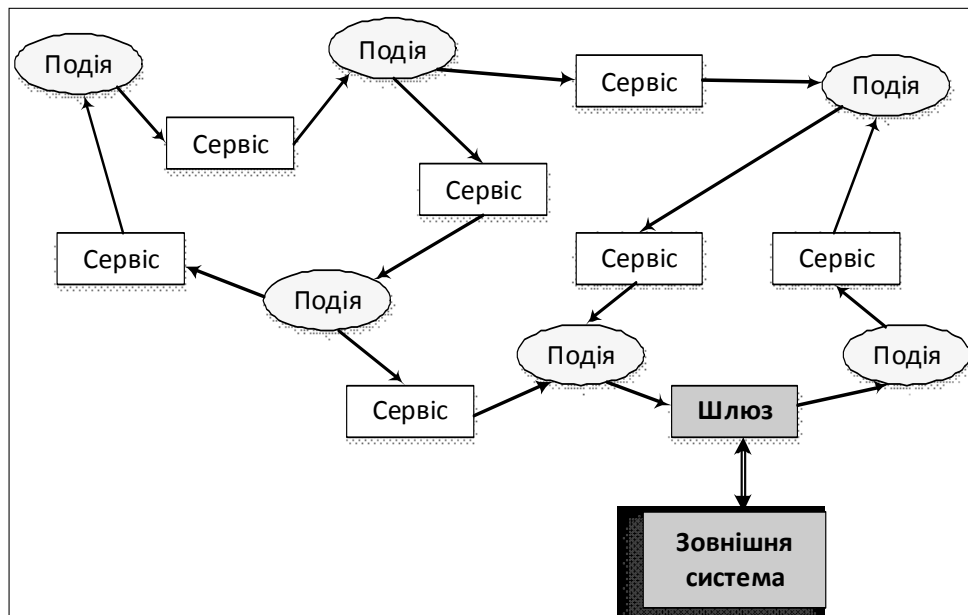


Рис. 2. EDA архітектура з подіями

Обидві нові архітектури SOA й EDA відрізняються від традиційної архітектури тим, що бізнес-процес розбивається на малі й повторно використовувані процедури. Ці процедури формалізуються і реалізуються у вигляді ІТ-компонентів, які слабо пов'язані і можуть бути інтегровані в динамічний і гнучкий засіб. SOA і EDA намагаються замінити і розбити старі і великі успадковані системи й архітектури на більш гнучкі і багаторазово використовувані бізнес- та ІТ-компоненти. Співвідношення між SOA й EDA є актуальною темою для обговорення. Споживачі і дослідники дотримуються протилежних думок з приводу того, як сервіси та події мають взаємодіяти на

різних рівнях, включаючи бізнес, інформацію та інформаційні системи. До базових принципів EDA треба віднести:

- **Роз'єднаність** — джерело події знає тільки створену подію і не має жодного уявлення щодо її подальшого оброблення заходу або про зацікавлені сторони. Цей принцип використовується у ході інтеграції роз'єднаних додатків і бізнес-процесів. Процес складається з декількох етапів. У EDA етапи не залежать фізично або логічно один від одного, так що кожен може бути змінено, не викликаючи побічних ефектів, на інші, поки повідомлення щодо подій не змінюються.

- **Повторюваність** — вирішальним фактором у проектуванні подій є те, що вони мають бути спроектованими таким чином, щоб відповідати не тільки нинішнім вимогам, а й майбутнім потребам і поза сферою використання. Події мають бути розроблені в загальному вигляді, який забезпечує їх повторне використання.

- **Повідомлення у режимі реального часу** — технічна інфраструктура має підтримувати створення в режимі реального часу подій і їх доставку. Персонал інфраструктури забезпечує в режимі реального часу перетворення інформації в знання, а потім в інтелектуальні наступні дії.

- **Свобода діяти** — повідомлення не вказує, які дії виконуватиме споживач події. Це звіт, а не вимога. Споживачу притаманна логіка, яка визначає, як він буде реагувати. Величезну кількість подій може бути вироблено, але тільки деякі з них будуть споживаними.

- **Зв'язок «один-до-багатьох»** — на одну конкретну подію може відгукнутися багато передплатників.

- **Асинхронність** — підтримка асинхронних операцій через повідомлення про події.

- **Запуск процесу** — потік управління, який визначається одержувачем, базується на розміщених подіях.

- **Відсутність стану компонентів обробки подій** — компоненти оброблення подій не мають стану, але сама подія буде мати стан, визначений у ній.

Для оброблення подій та ініціалізації веб-сервісів можуть бути використані три стратегії: *проста*, *потоківна* і *комплексна*. Просте оброблення полягає в ініціалізації веб-сервісів по мірі реєстрації відповідних подій. Основна перевага такої системи — це робота в режимі реального часу. Мінус очевидний — не враховується фактор пріоритету. Потоківне оброблення враховує існуючі залежності веб-сервісів і об'єднує кілька сервісів в один загальний потік. Цей підхід виправдовує себе в системах з великими накладними витратами на пошук і отримання інформації з бази даних. І, нарешті, комплексне оброблення — це так зване управління за відхиленнями. Будь-яка подія розцінюється як вихід системи зі стану рівноваги. Ініціалізація веб-сервісів переслідує єдину мету — повернути систему в стан рівноваги (між справою задовольняючи потреби клієнтів системи).

Новітній хмарний механізм оброблення подій CEP (complex event processing — комплексне оброблення подій) знаходиться в експлуатації з 2007 року [6]. У ньому реалізовано такі методи, як виявлення складних візерунків багатьох подій, кореляції подій і абстракції, ієрархії подій і виявлення від-

носин між ними, таких як причинність, підпорядкованість, синхронізація зв'язку з процесами, пов'язаними з подіями. У той час як SOA і оркестрування автоматизують процеси, СЕР автоматизує інтелектуальну кореляцію і співвідношення між рішеннями, прийнятими персоналом. Система СЕР дозволяє користувачам оброблювати випадкові, не пов'язані події, визначати тенденції і передбачати результати. Заходи можуть бути вжито для запобігання негативного наслідку від виникнення події. Шаблони реакцій також можуть бути проаналізовано, щоб поліпшити основні процеси та інформаційні системи з метою запобігання майбутніх помилок.

Бізнес-аналітики та керівники середньої ланки добре розуміють процеси, керовані подіями, оскільки вони пов'язані з стандартними бізнес-процедурами, такими як відносини замовника і постачальника, операційні процедури, процеси і потоки завдань. Переваги ведення бізнесу в режимі реального часу з EDA зрозуміло:

- Рішення приймаються, а потреби користувачів відразу задовольняються.
- Бізнес-процеси адаптується швидко, щоб задовольнити попит або вирішити проблему.
- CRM (customer relationship management — управління відносинами з клієнтами), підтримка, маркетинг і навіть ланцюжки постачань, можуть бути миттєво змінено.
- Весь бізнес можна розглядати в режимі реального часу для підвищення ефективності підприємства.

## ПОРІВНЯННЯ АРХІТЕКТУР EDA Й SOA

EDA й SOA можуть бути реалізовані і функціонувати паралельно без будь-яких суперечностей. Існує три основні передумови для цієї природної співпраці. Перша — це спільні та визначені бізнес-цілі, друга — використання роз'єднаних компонентів і спільної моделі даних, третя — використання спільної інфраструктури і технологій. EDA й SOA є взаємодоповнюючими у багатьох аспектах. Додавши EDA на вершину SOA архітектури, можна отримати нові можливості. Обидві архітектури SOA й EDA є досить новими концепціями і знаходяться на ранніх стадіях становлення. Багато компаній сьогодні здійснюють впровадження SOA та реалізацію очікуваних користі й цінностей [7]. EDA протягом багатьох років використовувалася як технічний шаблон проектування, який забезпечує масштабованість і високу продуктивність транзакцій. EDA як бізнес-концепція та архітектурний стиль все ще є новою і не підтримується прийнятими платформами і методами проектування. Ще однією серйозною проблемою є управління подіями.

Проведення порівняння EDA й SOA можна сьогодні лише на основі літературних джерел, оскільки бракує інформації щодо практичного досвіду організацій, де EDA експлуатуються. Обидві SOA й EDA націлені на узгодження бізнесу та IT-технології, але за допомогою різних засобів. У той час, як SOA використовує бізнес-сервіси, EDA використовує бізнес-події для подолання розриву між бізнесом та IT. Результати порівняння зведено в таблиці, де виділено три основні напрямки: відображення бізнес-процесів

і бізнес-функцій, використання даних та інфраструктури, здатність до інтеграції.

**Таблиця.** Порівняльний аналіз EDA і SOA

SOA	EDA
<b>Відображення бізнес-процесів і бізнес-функцій</b>	
Є архітектурним стилем, який визнає сервіси (функціональність), що відображають етапи бізнес-процесу і які створюють більш високий рівень гнучкості бізнесу та ІТ-середовища	Є архітектурним стилем, що визнає події (повідомлення), що відображають зміни стану процесу, на які бізнес реагує згідно з планом і які забезпечують більш високий рівень гнучкості бізнесу та ІТ-середовища
Фокусування на декомпозицію бізнес-процесу на сервіси, які слабо пов'язані між собою	Фокусування на виявленні бізнес-подій, які визначаються змінами в стані бізнес-процесу підприємства
Заохочення до повторного використання сервісів (реактивність)	Заохочення до обміну інформацією та використання бізнес-аналітики (активність)
<b>Використання даних</b>	
Дані запитується тими, хто зацікавлений у цій інформації	Дані поширюється в режимі реального часу серед тих, хто зацікавлений у цій інформації і підписався на неї
Сприяє повторному використанню і спільному вживанню даних у різних додатках і для декількох каналів доступу кінцевих користувачів	Сприяє миттєвій ідентифікації та реагуванню на події / інформацію, що управляють бізнес-процесом
Дані як сервіс є архітектурним підходом, що послаблює тісні зв'язки між даними додатками, так що дані можуть контролюватися і поділятися по всьому підприємству	Надмірність даних необхідна для збільшення рівня декомпозиції бізнес-процесу
<b>Інфраструктура і здатність до інтеграції</b>	
Застосовується поетапне розроблення і підтримка великих розподілених додатків	Застосовується поетапне розроблення і підтримка великих розподілених додатків
Застосовується шаблон «Запит / Відповідь»	Застосовується шаблон «Публікація / Підписка»
Сервіси викликаються незалежно від технології їх виготовлення та місця знаходження	Події формують повідомлення, що відправляються незалежним програмним модулям, які абсолютно нічого не знають один про одного
Одночасно лише один із конкретних сервісів викликається одним споживачем (зв'язок «один-до-одного»)	На одну конкретну подію може відгукнутися багато передплатників (зв'язок «один-до-багатьох»)
Сприяє інтеграції шляхом впровадження стандартів і ведення кроків обгортки, на шарування і композиції	Сприяє інтеграції, оскільки на повідомлення про події можуть підписатися всі зацікавлені й потім обирати необхідні з них
Виведення може використовувати синхронні та асинхронні моделі	Використовуються лише асинхронні моделі

Таким чином обидві архітектури SOA й EDA забезпечують перехід організації зі старою архітектурою бізнес-процесів, заснованою на підтримці незалежними монолітними прикладними додатками, до нового типу архітек-

тури, яка базується на незалежних сервісах і подіях, яка є більш динамічною і допускає багаторазове використання компонентів. Обидві парадигми ведуть до довгострокових інвестицій у різних галузях, де вони можуть принести користь, наприклад, шляхом:

- доповнення існуючих інформаційних систем і представлення їх як сервісів, тим самим збільшити рентабельність і продовжити термін використання спадщини, а також знизити витрати на розробку нової функціональності;
- використання існуючих сервісів і подій для підтримки нових бізнес-процесів, які можуть бути розгорнуті на ринку більш швидко і з більш низькою вартістю;
- повторного використання інваріантних міжгалузевих сервісів і подій для декількох бізнес-процесів, підрозділів або підприємств, що може привести до збільшення добутку, тобто кращого повернення від інвестиції.

Подійно-керований шаблон EDA сприяє декомпозиції пов'язаної системи (або бізнес-процесу) у порівнянні з вимогою слабо-пов'язаної системи для SOA. Тобто функціональні процеси в системі відправника, де відбулася бізнес-подія, не залежать від наявності та завершення віддалених процесів в розподілених низькорівневих системах. У той час, як в архітектурі, керованій запитом, система відправника повинна точно знати, які розподілені сервіси відгукуються на виклик, тобто залежить від наявності та завершення цих віддалених сервісів.

Архітектура EDA, керована подіями, має значення тільки тоді, коли події опубліковані, споживається і поширюється *в режимі реального часу*. Якщо цей підхід застосовується в пасивних сценаріях оброблення подій, наприклад, орієнтованих на фіксовану графіку, то його ефективність не дуже відрізняється від традиційних методів інтеграції даних, що використовуються сьогодні.

EDA притаманна тісна інтеграція з бізнес-процесами, у той час як для SOA завжди слабким місцем були точки входу кінцевих користувачів в силу їх неоднорідності за безліччю параметрів. На думку деяких експертів, як тільки SOA стає тісно зав'язана на бізнес-процесах організації, вона перероджується в подійно-керовану архітектуру і набуває не властиву статичним системам гнучкість й адаптивність. Головна складність у ході розробки подієво-керованих додатків — це гарантування, що всі оброблювачі подій будуть завершуватися досить швидко, щоб не блокувалися системні виклики. Тобто не для всіх додатків ця архітектура підходить, але вона успішно може працювати, якщо в якості оброблювачів подій використовуються веб-сервіси.

## ОБ'ЄДНАНА АРХІТЕКТУРА СИСТЕМИ СЕРВІСІВ

З порівняльного розгляду SOA й EDA очевидний тісний зв'язок між двома їх основними компонентами (подіями та сервісами). Дві основні відносини між EDA й SOA теж очевидні. У першому відношенні через події підключаються сервіси за допомогою передачі стану процесу і даних від одного сервісу, який виявляє і публікує події, до інших сервісів, які запускаються



у процесі появи конкретних подій. У другому відношенні сервіси генерують події, переводячи бізнес-процес з одного стану в інший. Іншими словами, **подія фіксує стан, а сервіс змінює стан**. Взаємодія між SOA й EDA при цьому здебільшого відбувається на двох різних рівнях (рис. 3).

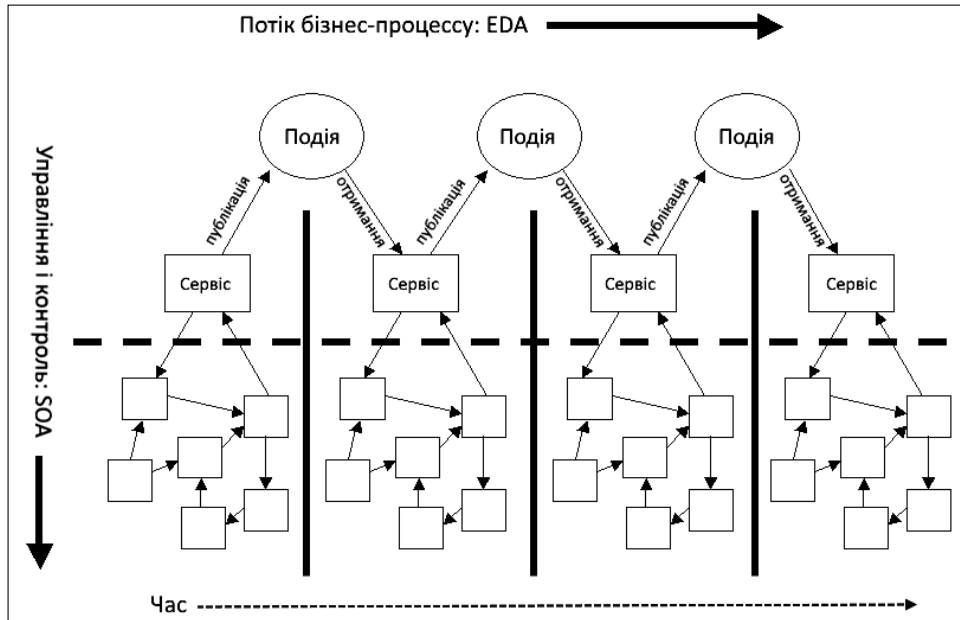


Рис. 3. Взаємодія SOA й EDA

- На першому рівні події з'єднують сервіси за допомогою передачі стану процесу від одного сервісу, який визначає і публікує події, до інших сервісів, які запускаються конкретними подіями.
- На другому рівні сервіси об'єднують події за допомогою передачі даних про перехід одного стану процесу в інший (споживач отримує подію і вживає необхідні заходи, які можуть призвести до виклику нових сервісів і формування нових подій).

Важливо також відзначити, що взаємодія між SOA й EDA також може відбуватися на рівні інформаційної системи, а не тільки на рівні бізнес-процесів. В об'єднаних SOA й EDA сервіси більше не є тільки споживачами подій. Зовнішні / внутрішні інформаційні системи можуть також вживатися і виробляти події.

Об'єднання концепцій SOA й EDA, пропонується називати EDSOA (event-driven service-oriented architecture — подійно-керована сервісно-орієнтована архітектура). Тільки формування бізнес-сервісів і бізнес-подій, а потім їх цільова ув'язка для рішень завдань бізнесу дозволяє домогтися стратегічних переваг для підприємства. При цьому EDA використовується для того, щоб «накинути» на всю програмну інфраструктуру організації «мережу» програмних датчиків і програмних агентів, а також апаратних датчиків, які стежать за подіями у всіх апаратних і програмних компонентах, відстежують значущі для бізнесу події і передають їх у центр прийняття рішень із сигналами, що асоційовані з цими подіями. Це дозволяє технологічно управляти бізнесом не наосліп, а маючи чітку картину всього, що відбувається в конкретний момент на підприємстві.

Як було зазначено раніше, важливо аналізуючи відносини між EDA й SOA враховувати рівень деталізації (зернистості) полій і сервісів. У добре визначеній архітектурі рівень деталізації бізнес-подій збігається з зернистістю бізнес-сервісів. Хоча рівень деталізації обох сервісів та подій має велике значення, поки не зовсім зрозуміло, як обрано потрібний рівень деталізації на етапі визначення сервісів і подій. І навіть менш очевидно поки, як переконатися, що сервіси і події вибрані на одному рівні.

Якщо говорити про SOA й EDA конвергенцію, можна вважати найбільш зручним рішенням цієї задачі використання сервісної шини підприємства ESB (enterprise service bus — сервісна шина підприємства), яка діє в якості проміжного шару (або посередника) для забезпечення комунікації та інтеграції великомасштабних гетерогенних прикладних процесів. ESB підтримує всі можливості: як SOA, так й EDA — оскільки підтримує синхронні, а також асинхронні взаємодії між однією або багатьох зацікавлених сторін. Це не стандарт, ні специфікація, а, швидше, велика кількість відкритих та комерційних ESB, які розроблялися багатьма постачальниками і налаштовані відповідно до їх потреб. До найбільш відомих реалізацій ESB можна віднести Open ESB [8], Apache ServiceMix [9], JBoss ESB [10], IBM WebSphere ESB [11], Celtix IONA й ObjectWeb [12]. Варто також звернути увагу на вільно поширювану шину *Mule* [13].

Як показано на рис. 4, ESB містить бізнес-процеси, сервіси та події, які споживаються або виробляються. Використання ESB в якості проміжного ПЗ забезпечує наступні групи сервісів [14].

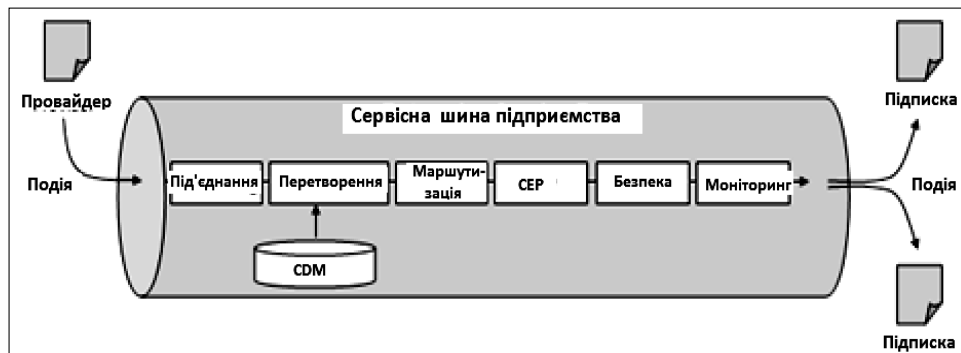


Рис. 4. Среднемесячные числа Вольфа

- **Транспортування:** транспортний сервіс гарантує доставку повідомлень та їх обмін між різними прикладними процесами. В ESB можна застосувати динамічну маршрутизацію (також на основі змісту) і відправку запитів за декількома напрямками. Це дозволяє ESB балансувати навантаження або механізми відмовостійкості.

- **Оброблення подій:** сервіс подій дозволяє ESB обробляти події, можливо, аналізувати та контролювати складні серії взаємопов'язаних подій. Для цього, більшість продуктів ESB забезпечує реалізацію специфікацій, присвячених організації подій, таких як WS-Notification й WS-Eventing [15].

- **Оркестрування:** цей сервіс засновано на використанні BPEL технологій [15] і дозволяє ESB організовувати виконання ряду взаємозв'язаних сервісів. Під час роботи з подіями на кожному етапі бізнес-процесу сервіси вищого рівня організують роботу сервісів нижчого рівня.

- **Пошуку:** цей сервіс, включений до ESB, сприяє тому, що прикладні процеси виявляють відповідні сервіси, з якими вони взаємодіють [16].

- **Посередництва:** цей сервіс посередництва є фундаментальним для напрямку бізнес-інтеграції. Він відповідає за: перетворення одного протоколу зв'язку на інший для того, щоб зробити можливим спілкування між гетерогенними середовищами; перетворення змісту будь-якого повідомлення, а також збагачення повідомлення додатковою інформацією для того, щоб будь-які дані, що передаються, були зрозумілими для будь-якого прикладного додатку. Крім перетворень, сервіс є відповідальним за безпеку, яка має вирішальне значення в міжорганізаційних взаємодіях, за вимоги якості QoS (наприклад, вимірювань, кешування, виявлення відмови і наступне відновлення).

Основні обмеження ESB пов'язані з їх застосуванням тільки в суворо контрольованих умовах, де адміністратори можуть правильно конфігурувати, розгортати і, нарешті, тонко налаштувати проміжне ПЗ, щоб отримати максимальну продуктивність і максимальний рівень надійності. Як тільки взаємодії перетинають кордони контрольованого середовища (яке, можливо, охоплює кілька незалежних адміністративних доменів), можливості цих платформ можуть непередбачувано деградувати — і загальний рівень обслуговування не буде більше легко керованим.

ESB є посередником, який пов'язує сервіси разом. ESB може бути реалізована різними способами: за допомогою класичних повідомлень; EAI (enterprise application integration — інтеграція корпоративних додатків); брокерських технологій; шляхом використання конкретної платформи компонентів, наприклад, інтеграції шин у сервері додатків J2EE (java 2 enterprise edition — обчислювальна корпоративна платформа Java). ESB також може бути поєднанням обох EAI й технології серверів додатків, але реалізація не має впливати на загальну архітектуру.

ESB виступає як інтелектуальний шар для підключення інформаційних систем, різних даних та інших сервісів, які зазвичай розподілено по всьому IT-середовищу підприємства. Вона поєднує в собі можливості синхронного та асинхронного обміну повідомленнями з можливостями інтелектуальних перетворень і маршрутизації. Це гарантує, що повідомлення передаються надійно. ESB дозволяє інтелектуальне оброблення запитів на обслуговування і відповідей, але також подій і повідомлень.

ESB не є новим продуктом, а новою концепцією інтеграції інформаційних систем, координації ресурсів та управління інформацією. На відміну від багатьох попередніх підходів до підключення інформаційних систем, ESB забезпечує підключення програмного забезпечення, яке працює на різних платформах, написаних на різних мовах програмування, і використовує різні технології. ESB є архітектурним шаблоном, який полегшує і спрощує бізнес-інтеграцію через посередницькі й транспортні сервіси. Вона виступає посередником для всіх комунікацій та взаємодій між різнорідними вузлами, як у сервіс-орієнтованій архітектурі (синхронний підхід «один-до-одного»), так у подійно-керованій архітектурі (асинхронний підхід «багато-до-багатьох»). На сьогодні ESB є найбільш ефективним засобом вирішення складних проблем інтеграції та є технічним рішенням, яке забезпечує найбільшу гнучкість бізнесу й ефективне з'єднання між різнорідними додатками.

Варто визнати, що EDSOA дозволяє домогтися поставлених цілей тільки в тому випадку, якщо:

- в інформаційній системі компанії буде структуровано й оформлено програмні компоненти, що реалізують сервіси, важливі для бізнесу (SOA складова);
- інформаційну систему буде насичено програмними датчиками, що формують значущі для бізнесу події (EDA складова);
- оперативне прийняття управлінських рішень буде автоматизовано за схемою, у якій реакція на бізнес-подію викликає старт бізнес-процесу, який ініціює запит на бізнес-сервіси (концепція EDSOA).

На завершення варто навести слова відомого в світі фахівця Юхима Натиса з компанії Gartner Research, який радить «розглядати SOA як довгострокову програмну архітектуру та інженерну практику, що вимагає наявності відповідних кваліфікованих кадрів, але не як спосіб вирішення поточних проблем департаментів інформаційних технологій. Великі вигоди вимагають великих інвестицій. Ухвалення бізнес-семантики для сервісів і додавання інтегрованої підтримки бізнес-подій вимагає великих систематичних зусиль, і ці кроки необхідні для реалізації потенційних можливостей архітектури бізнес-компонентів у всій її повноті».

## ВИСНОВКИ

SOA й EDA є новими поняттями для багатьох організацій і вимагають зміни менталітету і підходів як всередині бізнесу, так і в ІТ-середовищі. Наприклад, австралійська компанія Springboard Research, опитавши майже 3 тис. ІТ-керівників в Австралії, Індії, Китаї та Сінгапурі (аж ніяк не найвідсталіший регіон), встановила, що тільки 21% з них розуміють суть концепції SOA [17]. За традицією ІТ-додатки і процеси орієнтовано на використання передбачуваних і повторюваних подій. Коли ж відбуваються помітні події, які відрізняються винятковістю, то це створює складнощі для значної частини бізнесу. Це робить доцільним впровадження об'єднаної EDSOA архітектури як основної для побудови складних сучасних систем сервісів. Але недосвідченість розробників прикладних додатків у розробці систем, керованих подіями, і труднощі в декомпозиції бізнес-процесів на окремі компоненти стримують у цей час широке застосування парадигми EDSOA. Тим не менш, є також обнадійливі ознаки зацікавленості великих ІТ-виробників у новій технології (IBM, Microsoft, Google, Oracle тощо), які розпочали дослідження підходів з усунення вказаних труднощів [18–24].

Дуже важливо правильно виділити частину бізнесу організації, якій буде користь від EDSOA. Цей крок вимагає формування архітектури підприємства, де буде визначено сферу використання задіяних сервісів і подій. Бізнес-сервіси і події мають бути об'єднані таким чином, щоб забезпечити гнучку реконфігурацію процедур оброблення, можливо, шляхом використання декількох реакцій на подію. Усі ці заходи вимагають нової компетенції. Визначені сервіси, події й архітектура інтегруються в складний бізнес-процес.

Конвергенція SOA й EDA потребує розробки нових надійних інфраструктур проміжного ПЗ для реалізації складних сценаріїв прикладних до-

датків (наприклад, ланцюгів постачань, фінансових операцій). Ці підходи, будучи об'єднаними разом, можуть забезпечити використання їх взаємно додаткових характеристик. У цьому сенсі технології ESB можуть бути використані для реалізації конвергенції SOA й EDA. Однак відсутність стандарту й існування безлічі конкретних постачальників ESB реалізацій (чи комерційних, чи відкритих) ставить підприємства перед важким вибором ESB, яка найкраще підходить до конкретних конфігурацій розгортання бізнес-процесу. Це може негативно позначитися на гнучкості, яка забезпечується цими рішеннями і контентом вимог бізнесу.

Крім того, існуючі ESB є складними продуктами, які вимагають великих зусиль для конфігурації, розгортання і налаштування. Це пояснюється, очевидно, великими масштабами впровадження систем сервісів, неоднорідністю і сильною динамікою, яка характеризує намічені сценарії, де проміжне ПЗ має охоплювати кілька незалежно керованих адміністративних доменів. З цього погляду теперішні ESB представляють собою досить статичні рішення, позбавлені гнучкості, яка може бути забезпечена за допомогою автоматичної конфігурації, самооптимізації, самостійної адаптації та самозахисту функціональності. Таким чином, бажано, щоб у майбутньому було здійснено заходи з удосконалення ESBS з метою досягнення нових рівнів гнучкості в складних умовах застосування в індустрії сервісів.

## ЛІТЕРАТУРА

1. *Maglio P.* Service systems, service scientists, SSME, and innovation // *Communications of ACM*. — 2006. — 49, issue 7. — P. 81–85.
2. *Петренко О.О.* Объекты и методы науки о сервисах // *Системні дослідження та інформаційні технології*. — 2015. — № 2. — С. 75–82.
3. *Succeeding through service innovation: A service perspective for education, research, business and government*. — [http://www.ifm.eng.cam.ac.uk/uploads/Resources/Reports/080428cambridge\\_ssme\\_whitepaper.pdf](http://www.ifm.eng.cam.ac.uk/uploads/Resources/Reports/080428cambridge_ssme_whitepaper.pdf).
4. *Service Oriented Architecture: SOA Features and Benefits*. — [https://www.open-group.org/soa/source-book/soa/soa\\_features.htm](https://www.open-group.org/soa/source-book/soa/soa_features.htm).
5. *A model-driven ontology approach for manufacturing system interoperability and knowledge sharing*. — [http://www.researchgate.net/profile/Keith\\_Case/publication/257001871\\_A\\_model-driven\\_ontology\\_approach\\_for\\_manufacturing\\_system\\_interoperability\\_and\\_knowledge\\_sharing/links/00b49530dc9487310e000000.pdf](http://www.researchgate.net/profile/Keith_Case/publication/257001871_A_model-driven_ontology_approach_for_manufacturing_system_interoperability_and_knowledge_sharing/links/00b49530dc9487310e000000.pdf).
6. *Esper: Event Processing for Java*. — <http://www.espertech.com/products/esper.php>.
7. *Industrial SOA*. — <http://www.oracle.com/technetwork/articles/soa/ind-soa-toc-1934143.html>.
8. *Enterprise Integration EAI vs. SOA vs. ESB*. — [http://ggatz.com/images/Enterprise\\_20Integration\\_20\\_20SOA\\_20vs\\_20EAI\\_20vs\\_20ESB.pdf](http://ggatz.com/images/Enterprise_20Integration_20_20SOA_20vs_20EAI_20vs_20ESB.pdf).
9. *Apache Software Foundation*. — <http://servicemix.apache.org/home.html>.
10. *BossESB*. — <http://www.jboss.org/community/docs/DOC-10326>.
11. *ESB: The SOA communication center*. — <http://www-01.ibm.com/software/integration/wsesb/>.
12. *Celtix: The Open Source Java Enterprise Service Bus*. — <http://celtix.objectweb.org>.
13. *Event-driven services in SOA: Design an event-driven and service-oriented platform with Mule*. — <http://www.javaworld.com/article/2072262/soa/event-driven-services-in-soa.html>.

14. *Combining* Service-Oriented Architecture and Event-Driven Architecture using an Enterprise Service Bus. — <http://www.ibm.com/developerworks/library/ws-soa-eda-esb/>.
15. *OASIS* Web Services Notification. — [https://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=wsn](https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsn).
16. *OASIS* Web Services Business Process Execution Language. — [https://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=wsbpel](https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsbpel).
17. *EDA* как очередная инкарнация SOA. — <http://ecm-journal.ru/post/EDA-kak-ocherednaja-inkarnacija-SOA.aspx>.
18. *Advancing* soa with an event-driven architecture. — [http://www.intersystems.com/assets/Ensemble\\_SOA-EDA-6a940c1315b29eec545fa18db122dd47.pdf](http://www.intersystems.com/assets/Ensemble_SOA-EDA-6a940c1315b29eec545fa18db122dd47.pdf).
19. *New* to SOA and web services. — <http://www.ibm.com/developerworks/webser-vices/newto/>.
20. *Mixing* Event Driven Computing, SOA, BPM and BI for Instant Responsiveness. — [http://www.ebizq.net/blogs/firstlook/2007/10/post\\_17.php](http://www.ebizq.net/blogs/firstlook/2007/10/post_17.php).
21. *Event-Driven* SOA: Events Meet Services. — <http://www.oracle.com/technetwork/articles/soa/schmutz-soa-eda-405955.html>.
22. *SOA* and BPM Are Better Together. — <ftp://public.dhe.ibm.com/software/eg/soa/garbetter.pdf>.
23. *Implementing* an Event-driven Service-oriented Architecture in TIP. — <http://www.cs.waikato.ac.nz/pubs/wp/2010/uow-cs-wp-2010-02.pdf>.
24. *Event-Driven* Architecture and SOA — Allies or Enemies? — <http://wis.vsb.cz/ekf/php/tsw/getfile.php?prispevekid=935>.

Надійшла 31.08.2015

## ВПРОВАДЖЕННЯ КОМПАРАТОРНОЇ ІДЕНТИФІКАЦІЇ ДЛЯ КОМПЛЕКСНОЇ ОЦІНКИ РІВНЯ БЕЗПЕКИ ОБ'ЄКТІВ

Т.В. КОЗУЛЯ, Н.В. ШАРОНОВА, М.О. БІЛОВА, М.М. КОЗУЛЯ

Запропоновано новий підхід до оцінки стану природно-техногенних об'єктів з позицій сталого розвитку, що ґрунтується на методі компараторної ідентифікації. Розроблено алгоритмічне забезпечення реалізації методології компараторної екологічної ідентифікації з встановлення рівня екологічності еколого-соціально-економічних систем, розроблено математичні моделі стану системних об'єктів. На основі поданого методу отримано рейтинг еколого-економічних систем з безпечності екологічного стану та комплексну оцінку екологічності території забруднення важкими металами з встановленням механізмів саморегулювання якості систем. Розрахунок оцінки екологічності досліджених об'єктів за наданою методикою проведено відповідно програмного забезпечення, яке дозволяє автоматизувати аналіз стану складних систем.

### ВСТУП

Доцільність досліджень з ідентифікації несприятливих техногенних впливів на природне середовище (ПС), пов'язаних з динамічністю, інтенсивністю і небезпечністю наслідків для стану і функціонування соціально-екологічних систем, обумовлена недостатньою розробленістю теоретичних положень щодо узгодженості різнорідних оцінок стану складових еколого-соціально-економічних систем. Необхідність таких робіт визначена виконанням державних екологічних проектів з реалізації Концепції екологічної політики України, спрямованої на втілення загальноприйнятих принципів сталого розвитку (СР), які полягають у поєднанні та узгодженості в екологічній оцінці економічного, екологічного та людського розвитку за умови відсутності зниження якості й безпеки життя людей, збереження природного стану НС й соціального прогресу, відповідно до потреб людини [1–3].

Аналіз ситуації щодо розробок методичного забезпечення оцінки різних факторів впливу на об'єкти навколишнього природного середовища (НПС), показав, що у ході розв'язання задачі уніфікації та узгодження задач структурної і параметричної ідентифікації доцільно звернутися до інтелектуальних моделей, які б поєднували функції узагальненої корисності з оцінкою альтернатив розв'язання. Для задач синтезу й ідентифікації моделей багатofакторного оцінювання й оптимізації найбільш прийнятним є метод компараторної ідентифікації, запропонований Ю.П. Шабановим-Кушнаренком [4], який було розвинено у роботах М.Ф. Бондаренка [5], Е.Г. Петрова [6] та їх наукових шкіл.

**Метою роботи** є запровадження положень теорії компараторної ідентифікації для формування методичного забезпечення з оцінювання екологічного стану системних об'єктів, пов'язаних із багатofакторним аналізом.

У ході дослідних робіт поставлено та вирішено такі задачі:

- обґрунтовано необхідність і доцільність оцінки екологічної якості складних систем за концепцією сталого розвитку у ході використання методу компараторної ідентифікації;
- розроблено алгоритмічне забезпечення реалізації запропонованої методики з встановлення рівня екологічності природно-техногенних об'єктів;
- встановлено математичні моделі стану еколого-економічних об'єктів із застосуванням елементів синергетики — ентропійну та інформаційну функції як індикатори екологічності систем;
- проаналізовано практичні результати екологічних розрахунків відповідно до встановлених оцінок якості за існуючими методиками, визначено ефективність запровадження компараторної ідентифікації для системи оцінювання стану складних екологічних систем.

## ОСНОВНІ НАУКОВІ РЕЗУЛЬТАТИ І ЇХ ПРАКТИЧНА ОЦІНКА

Моделі індексного оцінювання стану системних об'єктів природно-техногенного походження орієнтовані на встановлення явних (прямих) властивостей системи, рівень яких безпосередньо визначається властивостями елементів, що утворюють систему. За концепцією СР екологічність як рівень якості НПС досягається за умови цілісності еколого-соціально-економічної системи, що передбачає визначення екологічної безпеки не за усередненими проміжними результатами, а з врахуванням усього спектру різномірних даних моніторингу. Саме багатофакторність аналізу вихідної інформації та її різноманітність за природою та динамікою складає проблематику об'єктивного оцінювання системних об'єктів.

Таким чином, необхідною є розробка більш складної моделі розрахунку індексу сталого розвитку, яка враховувала б не тільки параметричну, а й структурну ідентифікацію моделі та емерджентні властивості системи. Для вирішення цього питання запропоновано звернутися до методу компараторної ідентифікації, який дозволяє відмовитися від проміжних оцінок, що дають накопичення похибок, замінивши їх аналізом (порівнянням) ситуації, як цілого.

Для реалізації компараторної ідентифікації екологічної відповідності прийнято деяку сукупність об'єктів  $X_0, X_1, X_2, \dots, X_n$  загальною кількістю  $n$ , кожен з яких характеризується  $m$  певних параметрів початкового стану  $S$ :

$$\begin{aligned} X_0(x_{00}, x_{01}, x_{02}, \dots, x_{0m}), S_0; \\ X_1(x_{10}, x_{11}, x_{12}, \dots, x_{1m}), S_1; \\ \dots \\ X_n(x_{n0}, x_{n1}, x_{n2}, \dots, x_{nm}), S_n; \\ i = \overline{1; n}, j = \overline{1; m}. \end{aligned} \quad (1)$$

Задача екологічного оцінювання полягає у встановленні структурної здатності системного утворення призводить до виникнення процесу, який



спричиняє перехід систем у новий екологічно відповідний стан  $S_1$  або збереження стану  $S_0'$  під час реалізації небезпеки і виникненні суттєвих дестабілізуючих порушень у них (рис. 1).

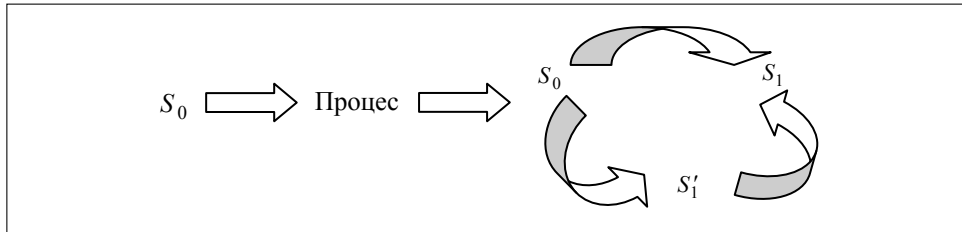


Рис. 1. Схема екологічної відповідності розвитку систем

Відповідно до методу компараторної ідентифікації запропоновано розробити шкалу відповідності екологічній якості під час оцінювання стану природно-техногенних об'єктів дослідження. Для фіксованого  $j$  обирається одне еталонне значення параметру  $y_j$ , яке буде мінімальним або максимальним значенням параметру  $x_{ij}$  серед усіх  $X_n$  в залежності від фізичного змісту кожного з параметрів. Загальна кількість еталонних значень буде складати  $m$ .

У ситуації оцінки за різнорідними щодо фізичного змісту параметрами необхідним є їх нормування, що полягає у порівнянні встановленої властивості системи  $x_{ij}$  з її еталонним значенням  $y_j$  й знаходженні частки

$$n_{ij} = \frac{x_{ij}}{y_j}. \quad (2)$$

Для характеристики стану системи запропоновано три варіанти відношення:  $n_{ij} = 1$  — оптимальний,  $n_{ij} \gg 1$  та  $n_{ij} \ll 1$  — незадовільний, який свідчить про високий рівень відхилення від відповідності вимогам екологічного розвитку.

З метою встановлення міри відповідності запропоновано визначення відхилення від еталонного значення, прийнятого на одиницю, у вигляді:

$$\Delta_{ij} = \frac{|n_{ij} - 1|}{\max n_j - \min n_j}. \quad (3)$$

Оптимальним варто вважати значення  $\Delta_{ij} = 0$ , допустимим під час оцінки якості для змінених систем прийнято відхилення 20% [7], що відповідає припустимим або малим ризикам, а звідси значення параметра для відношення відповідності прийняти  $\Delta_{ij} \in [0; 0,2]$ .

Для вибору відповідного стану екологічності досліджених об'єктів усі значення  $\Delta_{ij}$  проходять через компаратор. Якщо значення  $\Delta_{ij}$ , що надходить на вхід компаратора, знаходиться в межах  $[0; 0,2]$ , то відповідь на виході компаратору буде ТАК, тобто 1. В іншому випадку на виході компаратора буде НІ, тобто параметр оцінювання дорівнює 0.

Таким чином, кожен з об'єктів  $X_i$  буде характеризуватись множеною  $k_{im}$ , де кожне  $k_{ij}$  дорівнює 0 або 1. Надалі використовується компаратор, що має  $m$  входів та один вихід. Такий компаратор застосовується для кожного  $X_i$ : на вхід подаються відповідні значення  $k_{ij}$  і, якщо всі вони дорівнюють 1, — на виході буде значення ТАК. Такий об'єкт є «відповідним» — він не потребує заходів управління або регулювання його якості. Однак, якщо хоч один з  $k_{ij}$  дорівнює 0, — об'єкт потребує регулюючих змін для повернення початкових параметрів або управляючих дій з досягнення відповідності вимогам екологічної безпеки.

Запропоновану методику екологічної оцінки відповідності запропоновано для визначення рейтингу об'єктів за їх станом. На перше місце рейтингу пропонується розмістити об'єкт, для якого найбільша кількість характеристик  $k_{ij}$  приймає значення 1. Сформована таким чином нова вибірка систем оцінювання більше не буде містити в собі вищезазначений об'єкт. Розрахунки починаються з пошуку еталонних значень параметрів  $y_j$ . За тим же принципом обирається об'єкт, який посідає друге місце рейтингу. Формується оновлена вибірка. Таким чином, розрахунки проводяться до моменту повного розміщення досліджених об'єктів у певній послідовності за значеннями екологічних відхилень (рис. 2).

Запропоновану методику компараторної ідентифікації екологічності реалізовано для визначення стану регіонів з високо розвинутою промислово-соціальною інфраструктурою ( $X_i$ ) за такими показниками:  $x_{i1}$  — приріст населення за останній рік;  $x_{i2}$  — народжуваність;  $x_{i3}$  — кількість померлих віком до одного року;  $x_{i4}$  — середня заробітна плата;  $x_{i5}$  — кількість безробітних працездатного віку;  $x_{i6}$  — реалізована промислова продукція, млн. грн;  $x_{i7}$  — викиди забруднюючих речовин, тис. т;  $x_{i8}$  — кількість утворених відходів, тис. т;  $x_{i9}$  — утилізовано відходів, тис. тонн [8].

За формулами (2, 3) та алгоритмом (рис. 2) наведено рейтингові розрахунки, у результаті яких отримано таку послідовність стану областей за рівнем екологічності: 1. Київська; 2. Волинська; 3. Автономна Республіка Крим; 4. Івано-Франківська; 5. Закарпатська; 6. Рівненська; 7. Чернівецька; 8. Запорізька; 9. Полтавська; 10. Сумська; 11. Миколаївська; 12. Кіровоградська; 13. Житомирська; 14. Тернопільська; 15. Донецька; 16. Дніпропетровська; 17. Харківська; 18. Херсонська; 19. Луганська; 20. Львівська; 21. Чернігівська; 22. Хмельницька; 23. Черкаська; 24. Одеська; 25. Вінницька.

Результати рейтингу збігаються з оцінкою стану екологічної безпеки областей України за трьома класами небезпеки, встановленими за інтегральним показником ризику (перші 8 місць рейтингу більше, ніж на 50% мають низьку оцінку ризику, а з 9-ї позиції переважає оцінка високого і середнього ризику) [9].

Відповідно до запропонованої методики оцінки екологічної відповідності встановлено рейтинг більш небезпечних промислових об'єктів Харківської області за рівнем небезпеки їх стану:  $X_1$  — ПАТ «Євроцемент-Україна»,

$X_2$  — ПАТ «Харківська ТЕЦ-5»,  $X_3$  — Філія «Теплоелектроцентрально» ТОВ «ДВ нафтогазовидобування»,  $X_4$  — Зміївська ТЕС ПАТ «Центренерго». Аналіз виконано у рамках восьми параметрів, кожен з яких позначає кількість викидів у тоннах на рік певної речовини:  $x_{i1}$  — метали та їх сполуки;  $x_{i2}$  — речовини у вигляді зважених твердих частинок;  $x_{i3}$  — сполуки азоту;  $x_{i4}$  — діоксид та інші сполуки сірки;  $x_{i5}$  — оксид вуглецю;  $x_{i6}$  — неметанові леткі органічні сполуки;  $x_{i7}$  — метан;  $x_{i8}$  — діоксид вуглецю [10].

Рейтинг з небезпечності екологічного стану, отриманий після проведення розрахунків: 1 — Зміївська ТЕС ПАТ «Центренерго»; 2 — Філія «Теплоелектроцентрально» ТОВ «ДВ нафтогазовидобування»; 3 — ПАТ «Харківська ТЕЦ-5»; 4 — ПАТ «Євроцемент-Україна», що відповідає офіційним даним, представленим у «Доповіді про стан навколишнього природного середовища в Харківській області у 2012 році» [10].

Для встановлення об'єктивної комплексної оцінки екологічності систем (КЕС), що є об'єктом управління якістю НПС, запропоновано запровадити структурну і параметричну ідентифікацію рівноваги систем і необоротних процесів, визначених самоорганізацією об'єкта [11]. Складний природно-техногенний об'єкт дослідження визначається у вигляді системної моделі — виділення економічного, екологічного і соціального аспектів аналізу, для якої передбачається перехід від системного аналізу стану мікрооб'єкту до визначення процесів стабілізації відповідно до положень теорії синергетики та негентропії.

На основі ймовірно-ентропійних показників стану систем і процесів ( $P$ ,  $S$  й параметр стану  $x$ ) запропоновано перейти від результатів аналізу статистичних спостережень до характеристики термодинамічних потоків (процесів), які дозволяють утримати систему в стані рівноваги ( $\Delta S \rightarrow 0$ ) чи зменшити негативні впливи між системами за рахунок трансформаційних перетворень ( $S \rightarrow \max$ ), або фіксувати дестабілізацію в об'єкті  $\Delta S \rightarrow \max$ .

Загальний підхід у визначенні екологічності чи рівня екологічної безпеки системного об'єкта дослідження передбачає таку послідовність розв'язку задачі екологічної якості:

$$S_0 \xrightarrow{S_w} < S_1 \xrightarrow{q(x,t)} S_1^1 \xrightarrow{A_w} < S_2 \xrightarrow{D_w \rightarrow H'_w} S_{\max} (\Delta S \rightarrow 0), \quad (4)$$

де  $S_w$  — стохастичний оператор під час дії природних впливів, зовнішнього регулювання в межах природоохоронних заходів;  $q(x,t)$  — зв'язок між системами, речовинно-енергетичний потік;  $A_w$  — оператор переходу стохастичних функцій у детерміновані, який визначає перехід системи у новий стаціонарний стан у результаті самоорганізації;  $D_w$  — детермінований оператор, який відповідає за функцію виходу, що приводить об'єкт у рівноважний екологічний стан за умови стабілізації зовнішнього і внутрішнього гомеостазу — оператор  $H_w$  схеми «вхід-вихід», який для системного утворення визначається досягненням максимального ентропійного стану  $S_{\max}$  й відсутністю дестабілізуючих явищ ( $\Delta S \rightarrow 0$ ).

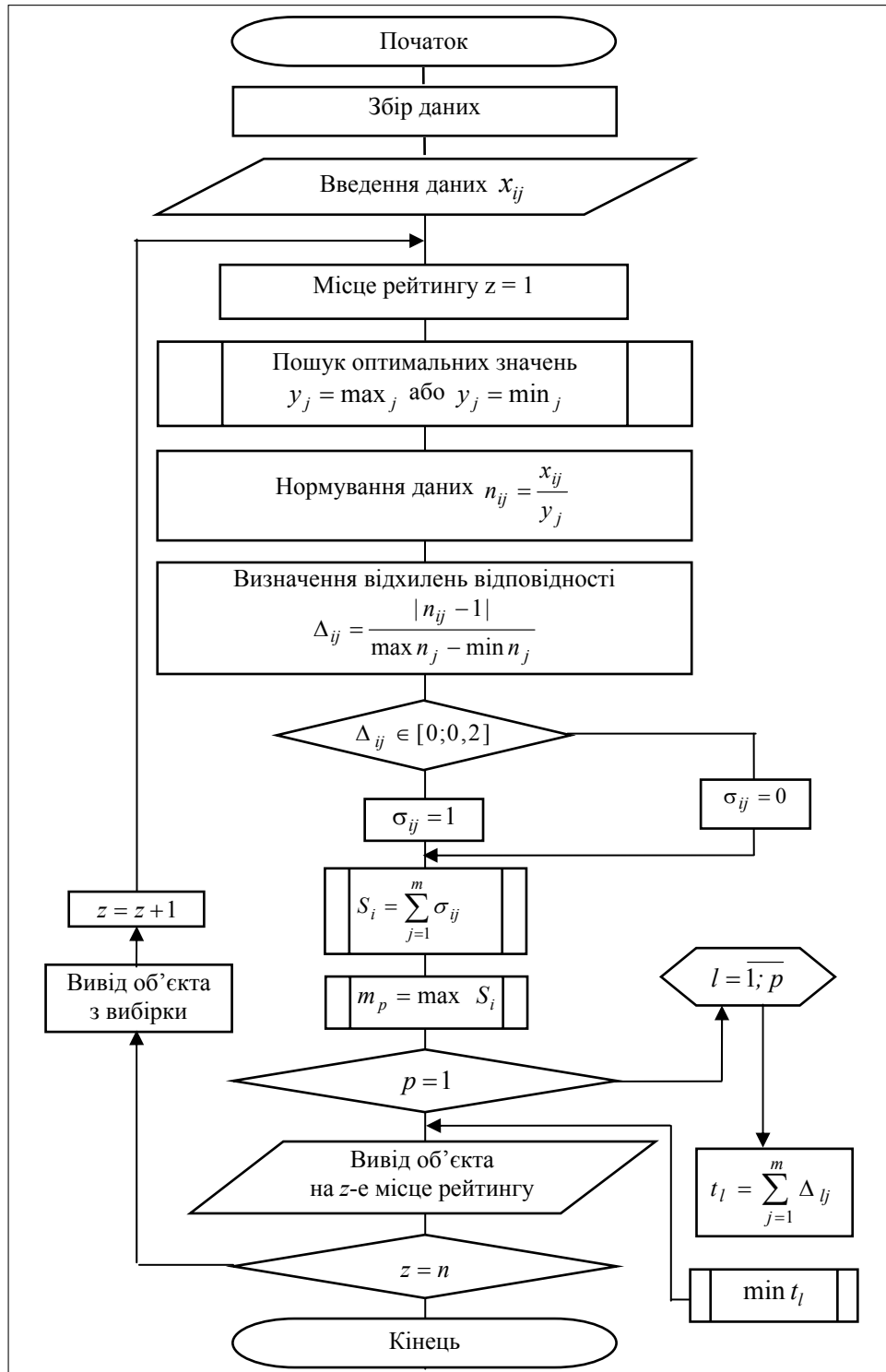


Рис. 2. Схема алгоритму оцінки якості за екологічним компаратором

Для комплексної оцінки екологічності територіально-об'єктових систем компараторна ідентифікація надає кількісне значення у двох парамет-

рах — 0 й 1, що дозволяє поєднати за параметром  $\Delta S$  зміни у стані систем і об'єкта й імовірності  $P$  порушення зв'язків у середовищі [5, 11–13].

За теорією компараторної ідентифікації [5]: компаратор — вимірювач відповідності структури, функціональності систем внутрішньому гомеостазу системного об'єкта, поданої кортежем  $X^s$  й  $X^t$  вхідного впливу  $X$  виду

$$F(A, X^{(s)}) = F(A, X^{(t)}), F(A, X^{(s)}) < (\leq) F(A, X^{(t)}), s, t = \overline{1, k}, s \neq t. \quad (5)$$

Зазначені кортежі є складовими термодинамічного потоку, визначеного як деякий зв'язок — відношення  $R$  між складовими системи і об'єкта відповідно до параметра  $A$  моделі і оператора  $F$  (структура) (2), що реалізує внутрішній гомеостаз складових систем. За однозначності характеристики стану систем через імовірнісно-ентропійні параметри оцінку екологічності надано видом оператора  $F$  й параметрів моделі  $A$ , для яких невідповідність вимогам безпеки (ситуація нееквівалентності) призводить до нерівності (5).

Під час вивчення територіально-об'єктових екосистем поряд з варіантами поведінки  $X = (x_1, x_2, \dots, x_n)$  для  $x_i \in X, i = \overline{1, n}$  розглядаються поодинокі кількісні вимірювання для окремих систем (наприклад, під час дослідження стану екосистеми «грунт» — вимірювання фізичних параметрів і хімічного складу як індикаторів їх екологічності і відповідності гомеостазу):

$$K(x_i) = \langle k_1(x_i), k_2(x_i), \dots, k_m(x_i) \rangle.$$

На основі аналізу окремих характеристик систем за певних умов стану системного об'єкта визначають імовірнісну оцінку реалізованої структури і її відповідність вимогам екологічної безпеки, поданих у вигляді  $P: X \rightarrow V$ , де  $P$  — оператор моделі оцінювання, а  $V = P(X)$  — модель багатofакторних оцінок альтернатив рішення ( $v_i = P(x_i), i = \overline{1, n}$ ).

Оцінка екологічності систем і рівня безпеки передбачає застосування *вимірювальної процедури екологічного ризику*, яка і реалізує предикат виду

$$D_1(v_q, v_n) = \begin{cases} 1 & \text{при } v_q = v_n, \\ 0 & \text{при } v_q \neq v_n, \end{cases}$$

$$E_1(x_q, x_n) = D_1[P(x_q), P(x_n)], \forall x_q, x_n \in X, \quad (6)$$

де  $v_q, v_n$  — оцінка корисності дослідженого і природного стану, як  $v_q = P(x_q), v_n = P(x_n)$  з врахуванням витрат на підтримку екологічності систем.

Аналіз потоків між системами та імовірності виникнення негативного фактору впливу  $v_q, v_r \in V$  виявляють зміни у результаті трансформаційних процесів щодо складу потоку, прояву ефектів у НПС, які приводять до збільшення  $S$  й  $\Delta S$  внутрішнього простору об'єкта. Результати аналізу подають за компаратором і предикатом у вигляді (позитивні результати — 1):

$$D_2(v_q, v_r) = \begin{cases} 1 & \text{при } v_q \geq v_r, \\ 0 & \text{при } v_q < v_r, \end{cases}$$

$$E_2(x_q, x_r) = D_2[P(x_q), P(x_r)], \forall x_q, x_r \in X. \quad (7)$$

Екологічність цього стану системи визначається незмінністю ентропійної функції оцінки і  $\Delta S \rightarrow 0$ , негативною оцінкою є збільшення імовірності перебігу процесів із дестабілізації систем  $S_q > S_r$ ,  $\Delta S > 0$  і появою ініціюючого впливу на них у внутрішньому середовищі об'єкта  $P(x_q) > P(x_r)$ .

Для моделі виду (4) стан систем визначається досягненням максимального значення ентропійної функції  $S_{\max}$  і відсутністю дестабілізуючих явищ ( $\Delta S \rightarrow 0$ ), надають оператор з деталізацією, зважаючи на пріоритетність збереження стійкості природних екосистем ( $x_1$ ):

$$P(x_s) > (\geq) P(x_1), \quad x_s, x_i \in X, \quad s = \overline{2, n}, \quad s \neq 1, \\ P(x_2) < (\leq) P(x_1), \quad P(x_3) < (\leq) P(x_1). \quad (8)$$

За узгодженням імовірнісно-ентропійних показників екологічності і компараторної ідентифікації (4–7) надано загальну модель ( $M$ ) оцінювання:

$$V_M(x_i) = P_M(A_M, K(x_i)), \quad i = \overline{1, n}, \quad (9)$$

де  $V_M(x_i)$  — узагальнена оцінка корисності альтернатив з оцінки екологічної якості систем, вибору їх структури, прийнятого рішення;  $P_M$  — оператор моделі оцінювання — структурна ідентифікація як реалізація економічних, соціальних і екологічних складових і їх зв'язку;  $K(x_i)$  —  $m$ -мірна кількісно-вимірjana вхідна дія (характеристика стану);  $A_M$  —  $r$ -мірний вектор характеристик моделі об'єкта — параметрична ідентифікація (стан систем, потоків, як зв'язок між системами).

За даними моніторингу для системного об'єкта визначають функцію

$$Y(x_i) = F(K(x_i)),$$

де  $Y(x_i)$  — скалярна багатофакторна оцінка станів або змін у складових і об'єкті (за рівнем задачі дослідження),  $x_i \in X$ ;  $K(x_i) = \{k_1(x_i), k_2(x_i), \dots, k_m(x_i)\}$ ,  $i = \overline{1, n}$  — фактори оцінювання  $x_i$ , для яких вводиться коефіцієнт ізоморфізму з метою досягнення однорідності  $K(x_i) — A = \langle a_1, a_2, \dots, a_r \rangle$ .

Таким чином, загальна оцінка якості забезпечується завдяки досягненню однорідності подання досліджених факторів стану систем і характеристик процесів у них у вигляді  $S, \Delta S, P, Risk$ .

Реалізацію методики комплексної оцінки якості НПС розглянуто на прикладі дослідження екологічного стану ландшафтно-геохімічних комплексів, які знаходяться під впливом потоку забруднення важкими металами (ВМ).

За компараторною ідентифікацією (відношення (5–9)) визначають геохімічний підпростір  $A_1, A_2, \dots, A_m$  ( $m$  — розмірність), для якого  $(x_1, x_2, \dots, x_m)$  складені з факторів впливу у вигляді  $x_1 \in A_1, x_2 \in A_2$ , що становлять предметний простір з катіонних і аніонних форм ВМ  $A_1 = A_2 = \dots = A_m = U$ :

$$U = \left\{ \underbrace{\text{Zn, Co, Ni, Pb, Sr, Cu}}_{x_1}, \underbrace{\text{Mo, Cr, V}}_{x_2} \right\}.$$

Для оцінки безпеки надходження ВМ в об'єкти НПС розглядають відношення  $\{x_1, x_2\}$   $m = 2$ ,  $A_1 = \{\text{Zn, Co, Ni, Pb, Sr, Cu}\}$ .  $A_2 = \{\text{Mo, Cr, V}\}$ , тоді  $S = A_1 \times A_2$  є множина пар виду  $(x_1, x_2)$ , для яких відношення формуються за значенням ентропійного стану, тобто аналізу процесів змін і самоорганізаційних здатностей системи ґрунт. Відношення, що є частинами одного простору, однотипні, реалізовані операціями: об'єднання — диз'юнкції  $\vee \cup$  — або; переріз — кон'юнкція  $\wedge \cap$  — і. Для аналізу стану важких елементів у ґрунті прийнято для значень імовірності відхилення кількості їх від нормативно встановленого обмеження у межах малого ризику — 20%:

$$P(x_1 - x_n) = \begin{cases} 1, & \text{якщо } x \leq 0,2, \\ 0, & \text{якщо } x \geq 0,2. \end{cases}$$

Ентропія у ході самовільних процесів збільшується і відповідно цього ідентифікується стан безпеки знаходження елементу у ґрунті (таблиця).

**Таблиця.** Результати ентропійної і компараторної ідентифікації стану елементів у ґрунтах

Елемент	Ентропійний стан	Компаратор	Cr	V
Zn	-2,265	1	1	1
Co	-1,619	0	1	1
Ni	-1,616	0	1	1
Pb	-2,042	1	1	1
Sr	-1,687	1	1	1
Cu	-1,844	1	1	1
Mo	-1,462	0	1	1
Cr	-2,639	1	1	1
V	-2,434	1	1	1

Відношення з відображення «перевищення» природного рівня завдяки техногенному надходженню і взаємодії у ґрунті зі складовими ґрунтів (сорбція) і між собою з урахуванням об'єднання цих процесів для кожного елемента (кон'юнкція  $P' \cap Q$ ) має такий остаточний вид  $P' = x^0 y^0 \vee x^0 y^1 \vee x^0 y^2 \vee x^0 y^3 \vee x^0 y^4$  згідно з сформованою базою оцінки їх стану:

Важкі метали	Zn	Ni	Pb	Sr	Cu
Zn	1	1	1	1	1
Ni	0	0	0	0	0
Pb	1	1	1	1	1
Sr	1	1	1	1	1
Cu	0	0	0	0	0

Відповідно до отриманої досконалої диз'юнктивної нормальної форми (ДДНФ):  $P' = x^0 y^0 \vee x^0 y^1 \vee x^0 y^2 \vee x^0 y^3 \vee x^0 y^4$  отримують такі дані:

Важкі метали	Zn	Pb	Sr	Cr	V
Zn	0	0	0	1	1
Pb	0	0	0	1	1
Sr	0	0	0	1	1
Cr	1	1	1	0	0
V	1	1	1	0	0

Урахування перерізу станів Zn відповідно до стану присутніх інших важких металів (ВМ) визначено ДКНФ виду

$$x^0 y^3 \wedge x^0 y^4 \wedge x^1 y^3 \wedge x^1 y^4 \wedge x^2 y^3 \wedge x^2 y^4 \wedge x^3 y^0 \wedge x^3 y^1 \wedge x^3 y^2 \wedge x^4 y^0 \wedge x^4 y^1 \wedge x^4 y^2. \quad (10)$$

Зафіксовані процеси самовільних процесів накопичення ВМ можливі при малому значенні  $P$  відхилення від мінімального їх вмісту, стабільне забезпечення «позитивного стану» визначено значним відхиленням від допустимого обмеження наявності елементу у ґрунті та перебігу при цьому зазначених процесів:  $x^i \leq -2,3$  при  $P = 0,1$  і  $y^j \leq -2,99$  при  $\alpha = 0,05, P = 0,95$ .

З урахуванням даних таблиці і виразу (10) отримують таку оцінку стану елементів за компараторною ідентифікацією:

Zn	$x^0 = -2,78$	$1 \cdot 1 \wedge 1 \cdot 1$	<b>1</b>
Pb	$x^1 = -2,04$	$0 \cdot 1 \wedge 0 \cdot 1$	<b>0</b>
Sr	$x^2 = -2,02$	$0 \cdot 1 \wedge 0 \cdot 1$	<b>0</b>
Cr	$x^3 = -2,7$	$1 \cdot 1 \wedge 1 \cdot 1 \wedge 1 \cdot 1$	<b>1</b>
V	$x^4 = -2,46$	$1 \cdot 1 \wedge 1 \cdot 1 \wedge 1 \cdot 1$	<b>1</b>

Таким чином, феноменологічні знання про поведінку ВМ в об'єктах НПС обґрунтовуються ентропійною оцінкою їх стану та компараторною ідентифікацією з урахуванням «стан – процес» завдяки застосуванню ДДНФ і ДКНФ. Це дозволяє встановити фактор небезпеки, імовірність нівелювання його негативного впливу за рахунок процесів трансформації у міграційних потоках у наявності інших негативних складових фактору.

Аналогічно проводиться оцінка стану здоров'я населення, що є відображенням рівня екологічності дослідженої території.

Таким чином, сформована екологічна оцінка є комплексним критерієм щодо визначення стану еколого-соціально-економічної системи і суттєвих змін, пов'язаних з підтримкою гомеостазу в природно-техногенних об'єктах. Розрахунок оцінки екологічності досліджених об'єктів за наданою методикою (4–10) проводиться відповідно до програмного забезпечення, що дозволяє автоматизувати аналіз стану складних систем — Microsoft Visual Studio 2005 (рис. 3).



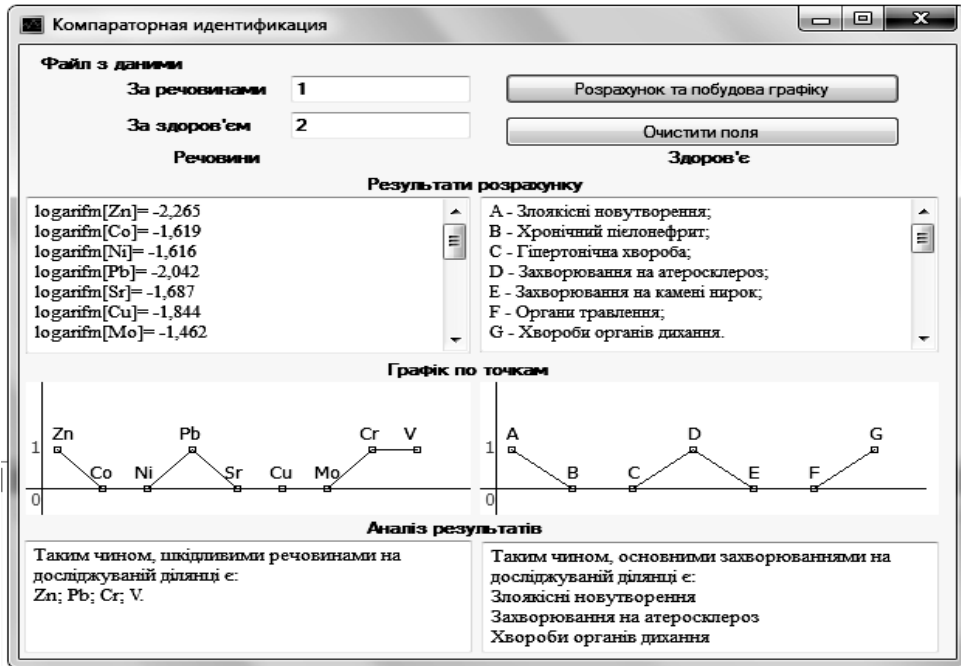


Рис. 3. Приклад оцінки рівня екологічності території

## ВИСНОВКИ

Отже, у роботі запропоновано новий підхід до оцінки стану об'єктів навколишнього середовища з позицій сталого розвитку, що ґрунтується на методі компараторної ідентифікації. Запропонована методика оцінки якості системних природно-техногенних об'єктів (рис. 1–3) дозволила отримати такі науково-практичні результати:

1) відповідно до удосконалення системного аналізу для розв'язання задач сталого розвитку запропоновано методика оцінки екологічної якості системних об'єктів із застосуванням теорії компараторної ідентифікації;

2) розроблено математичне (1–9) та алгоритмічне (рис. 2) забезпечення оцінки відповідності стану системного рівня моделей природно-техногенних комплексів вимогам екологічності;

3) використано надану методика для ранжування за екологічністю областей і підприємств за соціально-екологічними показниками, показано відповідність отриманих результатів встановленим даним [9, 10];

4) обґрунтовано ефективність запровадження компараторної ідентифікації для оцінки якості природно-техногенних об'єктів за ентропійною характеристикою стану і процесів, визначення факторів стабілізації екологічності.

Таким чином, у методиці оцінки стану об'єктів запроваджено компараторну ідентифікацію відповідності вимогам екологічності з урахуванням зв'язків між системами різної природи та визначенням стану об'єкту не за усередненими проміжними результатами, а за всіма даними інформаційного простору функціональності системного об'єкта дослідження.

## ЛІТЕРАТУРА

1. Рамперсанд Х. Универсальная система показателей: Как достигать результатов, сохраняя целостность. — М.: Альпина Бизнес Букс, 2006. — 352 с.
2. Згуровский М.З., Гвишиани А.Д. Глобальное моделирование процессов устойчивого развития в контексте качества и безопасности жизни людей. — К.: Політехніка, 2008. — 331 с.
3. Згуровський М.З., Статюха Г.О., Войтко С.В., Мельниченко А.А., Болдак А.О., Джигирей І.М. Аналіз сталого розвитку — глобальний та регіональні контексти. Україна в індикаторах сталого розвитку — К.: Політехніка, 2010. — 359 с.
4. Бондаренко М.Ф., Шабанов-Кушнарченко Ю.П. Мозгоподобные структуры: Справочное пособие. Том первый. Под ред. акад. НАН Украины И.В. Сергиенко. — К.: Наукова думка, 2011. — 460 с.
5. Бондаренко М.Ф., Шабанов-Кушнарченко С.Ю., Шабанов-Кушнарченко Ю.П. Про загальну теорію компараторної ідентифікації // Біоніка інтелекту: наук.-техн. журнал. — 2008. — № 2 (69). — С. 13–22.
6. Петров К.Э. Компараторная идентификация модели формирования индекса устойчивого развития // Системні дослідження та інформаційні технології. — 2009. — № 1. — С. 36–46.
7. Машина Н.І. Економічний ризик і методи його вимірювання. — К.: Центр навчальної літератури, 2003. — 188 с.
8. Державна служба статистики України. — <http://www.ukrstat.gov.ua/>.
9. Качинський А.Б. Екологічна безпека України: системний аналіз перспектив покращення. — К.: Екологічна безпека, 2001. — 251 с.
10. Доповідь про стан навколишнього природного середовища у Харківській області у 2012 році. — [http://www.menr.gov.ua/docs/activity-dopovidi/regionalni/regionalni-dopovidi-u-2012-rotsi/kharkivska\\_2012.pdf](http://www.menr.gov.ua/docs/activity-dopovidi/regionalni/regionalni-dopovidi-u-2012-rotsi/kharkivska_2012.pdf).
11. Sharonova N.V., Kozulia T.V. Entropy as Substratum of identifying the Corporative Ecological system (CES) condition // Вестник Херсонского национального технического университета. — 2008. — № 2 (31). — С. 518–527.
12. Козуля Т.В. Процеси екологічного регулювання. Концепція корпоративної екологічної системи: монографія. — Харків: НТУ «ХПИ», 2010. — 588 с.
13. Козуля Т.В., Білова М.О. Оцінка якості системних об'єктів навколишнього середовища на основі метода компараторної ідентифікації // Проблеми інформаційних технологій. — 2013. — № 13. — С. 78–84.

Надійшла 13.05.2015

УДК 510.87:544.431.7:577.21

## **ЗАДАЧА ОПТИМАЛЬНОГО КЕРУВАННЯ СТАДІЄЮ ЕЛОНГАЦІЇ ПОЛІМЕРАЗНО-ЛАНЦЮГОВОЇ РЕАКЦІЇ**

**В.П. МАРЦЕНЮК, А.С. СВЕРСТЮК, І.С. ГВОЗДЕЦЬКА**

Застосовано загальну методологію керування для отримання розв'язку задачі оптимального перебігу стадії елонгації у полімеразно-ланцюговій реакції. У досліджуваній моделі стадії елонгації використано рівняння Арреніуса, яке враховує залежність швидкості реакції від абсолютної температури. Цю залежність може бути використано для дослідження стадії елонгації полімеразно-ланцюгової реакції, оскільки керуючим впливом при цьому є саме температура. Розглянуто схему температурного керування. При реалізації полімеразно-ланцюгової реакції застосовано принцип максимуму Понтрягіна до задачі оптимального керування та сформульовано необхідну умову оптимальності. Отримані результати необхідні для чисельного обчислення оптимального керування досліджуваної стадії та допоможуть мінімізувати необхідний час для реалізації стадії елонгації, що в майбутньому дозволить мінімізувати час проведення полімеразно-ланцюгової реакції.

### **ВСТУП**

Полімеразна ланцюгова реакція (ПЛР) є експериментальним методом молекулярної біології, під час якого значно збільшуються малі концентрації бажаних фрагментів дезоксирибонуклеїнової кислоти (ДНК) у біологічному матеріалі (пробі). Крім простого збільшення числа копій ДНК (цей процес називається ампліфікацією), ПЛР дозволяє проводити безліч інших маніпуляцій з генетичним матеріалом, і широко використовується в біологічній та медичній практиках, наприклад для клонування генів, створення й визначення генетично модифікованих організмів, діагностики захворювань (спадкових, інфекційних), ідентифікації малих кількостей ДНК, встановлення батьківства [1].

ПЛР базується на багаточисленному копіюванні (селективній ампліфікації) досліджуваної ДНК ферментом ДНК-полімеразаю. Утворені копії ДНК ідентифікують за допомогою методу електрофорезу.

Для проведення ПЛР виконується 20–35 циклів [2], кожен з яких складається з трьох стадій (рисунок).

1. Дволанцюгову ДНК-матрицю нагрівають до 367–369 К (або до 371 К, якщо використовується особливо термостабільна полімераза) на 0,5–10 хв., щоб ланцюги ДНК розділилися. Ця стадія називається денатурацією — руйнуються водневі зв'язки між двома ланцюгами. Іноді перед першим циклом проводять попереднє прогрівання реакційної суміші протягом 2–5 хв. для повної денатурації матриці й праймерів.

2. Коли ланцюги розійшлися, температуру знижують, щоб праймери могли зв'язатися з одноланцюговою матрицею. Ця стадія називається відпадом, її температура залежить від праймерів і зазвичай вибирається на 4–5 К нижче за їх температуру плавлення. Час стадії — 0,5–2 хв.

3. ДНК-полімераза реплікує матричний ланцюжок, використовуючи праймер як затравку. Це так звана стадія елонгації. Температура елонгації залежить від полімерази. Полімерази *Taq* й *Pfu*, що найчастіше використовуються, найактивніші за 345 К. Час елонгації залежить як від типу ДНК-полімерази, так і від довжини фрагмента, який ампліфікують. Середня швидкість елонгації — 1000 пар основ за 1 хв. Після закінчення всіх циклів часто проводять додаткову стадію фінальної елонгації, щоб добудувати всі одноланцюжкові фрагменти. Ця стадія триває 10–15 хв.

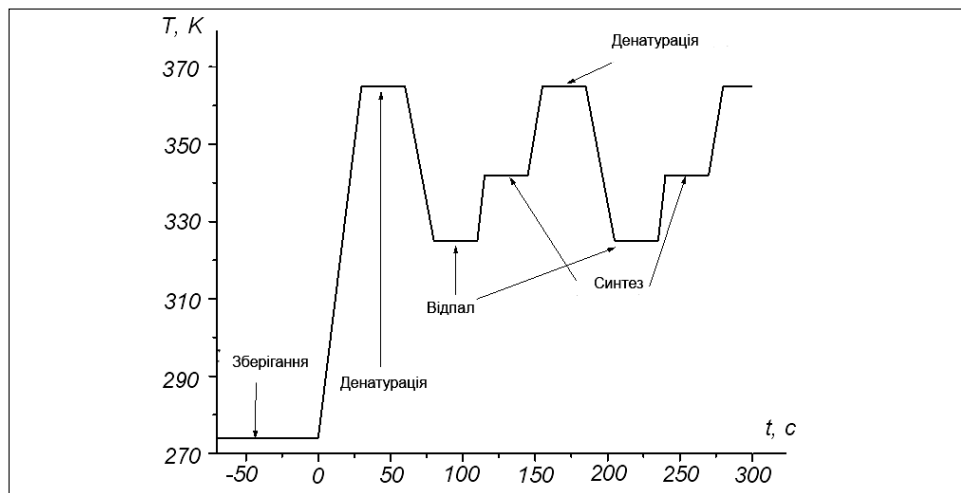


Рисунок. Схема температурного керування для реалізації ПЛР [3]

Для ефективного проведення ПЛР необхідно забезпечити багатостадійний циклічний режим зміни температури. Кожна стадія циклу (денатурація, відпал, елонгація) має відбуватися за певних температур та протягом відповідного часу. В іншому випадку необхідних перетворень молекул ДНК може не відбуватися. Таким чином задача математичного моделювання і відповідних розрахунків полягає в оцінюванні мінімально необхідного часу реалізації кожної стадії циклу, що в загальному випадку забезпечить досягнення мінімального часу проведення ПЛР.

У багатьох роботах [4–6] наведено різні моделі ПЛР, однак для ефективного використання методів ПЛР, доцільно провести дослідження керованості цих моделей. Таким чином задача математичного моделювання і відповідних розрахунків полягає в оцінюванні мінімально необхідного часу реалізації кожної стадії циклу, що в загальному випадку забезпечить досягнення мінімального часу та ефективності проведення ПЛР.

**Мета роботи** — дослідження керованості стадією елонгації ПЛР, яку запропоновано в роботі [6].

## МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ

**Теоретичні основи оптимального керування в ПЛР.** У задачах оптимального керування ПЛР за допомогою температури варто розглядати таку множину керувань:

$$U = \{u(t) : a \leq u(t) \leq b, \quad t_1 \leq t \leq t_2, \quad u(t) \text{ — вимірна}\}.$$

Тут  $a, b, t_1, t_2 > 0$ .

Припускається, що стан системи  $x(t) \in R^n$  у випадку заданого керування  $u \in U$  визначається системою звичайних диференціальних рівнянь:

$$\begin{aligned} \frac{dx(t)}{dt} &= f(t, x, u), \\ x(t_0) &= x_0, \end{aligned} \quad (1)$$

де  $f : R \times R^n \times R \rightarrow R^n$  є неперервною і має неперервні перші частинні похідні відносно  $x$  та  $u$ . Оскільки припускається, що  $u(t)$  є вимірною та обмеженою, то права частина системи (1) є неперервною відносно  $x$  й лише вимірною відносно  $t$  для фіксованого  $x$ . Отже, розв'язки (1) є абсолютно неперервними функціями, які задовольняють (1) майже скрізь. За таких умов існування розв'язку (1)  $x(t, u)$  доведено в роботах [7, 8].

Задача оптимального керування містить критерій якості  $J[u]$  вигляду:

$$J[u] = \int_{t_1}^{t_2} L(t, x, u) dt + \phi(x(t_2)),$$

де  $L$  — задана дійснозначна функція;  $\phi$  — неперервнодиференційовна дійснозначна функція. Метою є знаходження керування  $u^* \in U$  такого, що

$$J[u^*] = \inf_{u \in U} J[u]. \quad (2)$$

Після того, як описано модель та визначено критерій якості, в теорії оптимального керування ставлять ряд задач [9]:

- доведення існування оптимального керування;
- опис побудови оптимального керування;
- доведення єдиності оптимального керування;
- чисельне обчислення оптимального керування;
- дослідження залежності оптимального керування від параметрів моделі.

Достатні умови існування оптимального керування для задачі (1)–(2) без термінальної складової в критерії якості наведено в роботах [9, 10].

**Теорема 1.** Розглядається задача оптимального керування (1)–(2) на фіксованому інтервалі  $[t_1, t_2]$ . Припустимо, що:

- існує стала  $M > 0$  така, що  $\|x(t, u)\| \leq M$  для всіх  $u \in U$  та  $t_1 \leq t \leq t_2$ ;
- $L$  є напівнеперервною знизу;
- множина  $D^+ = \{(y^0, y) : \exists v \in U, y = f(t, x, v), y^0 \geq L(t, x, v)\}$  є опуклою для  $(t, x) \in [t_1, t_2] \times \{x : |x| \leq M\}$ .

Тоді існує оптимальне керування  $u^* \in U$ .

Опис побудови оптимального керування для задачі (1)–(2) дає принцип максимуму Понтрягіна з термінальною складовою [11, 12].

**Теорема 2.** Нехай  $u^* \in U$  — оптимальне керування в задачі (1)–(2). Тоді існує спряжена функція  $\lambda : R \rightarrow R^n$  така, що  $x(t, u^*)$ ,  $u^*$ ,  $\lambda$  задовольняють систему:

$$\begin{aligned} \frac{dx(t)}{dt} &= f(t, x, u^*), \\ x(t_1) &= x_1 \end{aligned} \quad (3)$$

та спряжену систему:

$$\begin{aligned} \frac{d\lambda(t)}{dt} &= -\frac{\partial H}{\partial x} = -L_x(t, x, u^*) - \lambda^T f_x(t, x, u^*), \\ \lambda(t_2) &= \phi'(x(t_2)) \quad \text{— (умова трансверальності),} \end{aligned} \quad (4)$$

де функція Гамільтона–Понтрягіна  $H$  задається як

$$H(t, x, u) = L(t, x, u) + \lambda^T f(t, x, u). \quad (5)$$

### ЗАДАЧА ОПТИМАЛЬНОГО КЕРУВАННЯ СТАДІЇ ЕЛОНГАЦІЇ ПЛР

Розглядається модель стадії елонгації ПЛР, яку запропоновано в роботі [6]. Для стадії елонгації

$$\frac{ds'}{dt} = -k_2 s' q + k_{-2} c, \quad (6)$$

$$\frac{dq}{dt} = -k_2 s' q + k_{-2} c + k_3 c n, \quad (7)$$

$$\frac{dc}{dt} = k_2 s' q - k_{-2} c - k_3 c n, \quad (8)$$

$$\frac{dn}{dt} = -k_3 c n, \quad (9)$$

$$\frac{dd}{dt} = k_3 c n, \quad (10)$$

де  $k_{-2}$ ,  $k_2$  — пряма і зворотня швидкості реакції для утворення комплексу,  $q$  —  $Taq$  молекула,  $c$  — число копій,  $n$  — нуклеотидна послідовність для елонгації,  $k_{-3}$ ,  $k_3$  — пряма і зворотня швидкості реакції для елонгації,  $d$  — дволанцюгова ДНК.

У той же час за проведення фази елонгації ПЛР керуючим впливом є температура [7–8]. Залежність швидкості реакції  $k$  від абсолютної температури  $T$  описується рівнянням Арреніуса [8]:

$$k = A e^{-E_a / RT}, \quad (11)$$

де  $A$  характеризує частоту зіткнень молекул,  $R$  — універсальну газову сталу,  $E_a$  — енергію активації.

Виходячи з рівняння Арреніуса (11) система диференціальних рівнянь для стадії елонгації (6)–(10) може бути уточнена таким чином:

$$\frac{ds'}{dt} = -k_2 e^{-\frac{r}{T}} s' q + k_{-2} e^{-\frac{r}{T}} c, \quad (12)$$

$$\frac{dq}{dt} = -k_2 e^{-\frac{r}{T}} s' q + k_{-2} e^{-\frac{r}{T}} c + k_3 e^{-\frac{r}{T}} cn, \quad (13)$$

$$\frac{dc}{dt} = k_2 e^{-\frac{r}{T}} s' q - k_{-2} e^{-\frac{r}{T}} c - k_3 e^{-\frac{r}{T}} cn, \quad (14)$$

$$\frac{dn}{dt} = -k_3 e^{-\frac{r}{T}} cn, \quad (15)$$

$$\frac{dd}{dt} = k_3 e^{-\frac{r}{T}} cn, \quad (16)$$

з відповідними початковими умовами:

$$s'(t_1) = s_0, q(t_1) = q_0, c(t_1) = c_0, n(t_1) = n_0, d(t_1) = d_0.$$

У системі диференціальних рівнянь (12-16)  $r = \frac{E_a}{R}$  — стала. Вважаємо,

що  $T = T(t)$  — функція керування. Припускаємо, що  $T(t) \in [T_e^{\min}, T_e^{\max}]$ .

Для стадії елонгації метою є отримати якнайбільше двохланцюгових ДНК  $d$ , при цьому витративши якнайменше  $Taq$  молекул, тобто:

$$J(s', q, c, n, d, T) = \int_{t_1}^{t_2} (d^2(t) - Wq^2(t)) dt \rightarrow \inf_{T \in U},$$

де  $W > 0$  — ваговий коефіцієнт,  $U$  — множина кусково-неперервних функцій  $T(t) \in [T_e^{\min}, T_e^{\max}]$ .

Біологічно значимою областю є

$$\Omega_1 = (s', q, c, n, d, T) \in R_+^5, \quad (17)$$

що накладає фазові обмеження:

$$s' \geq 0, q \geq 0, c \geq 0, n \geq 0, d \geq 0. \quad (18)$$

Отже, метою є визначення оптимального керування  $T^* \in U$ , що задовольняє:

$$J[T^*] = \inf_{T \in U} J[T]. \quad (19)$$

На основі теореми 1 бачимо, що оптимальне керування в задачі (12)–(19) існує, оскільки підінтегральний вираз у критерії якості є опуклою функцією, а траєкторія системи належить простору  $L^\infty$ .

Застосуємо теорему 2 для отримання необхідних умов оптимальності. Функція Гамільтона–Понтрягіна має вигляд:

$$H = d^2 - Wq^2 + \lambda_1 (-k_2 e^{-\frac{r}{T}} s' q + k_{-2} e^{-\frac{r}{T}} c) +$$

$$\begin{aligned}
 & + \lambda_2(-k_2 e^{-\frac{r}{T}} s'q + k_{-2} e^{-\frac{r}{T}} c + k_3 e^{-\frac{r}{T}} cn) + \\
 & + \lambda_3(k_2 e^{-\frac{r}{T}} s'q - k_{-2} e^{-\frac{r}{T}} c - k_3 e^{-\frac{r}{T}} cn) + \lambda_4(-k_3 e^{-\frac{r}{T}} cn) + \lambda_5(k_3 e^{-\frac{r}{T}} cn). \quad (20)
 \end{aligned}$$

Отже, з теореми 2 маємо спряжену систему:

$$\begin{aligned}
 \frac{d\lambda_1}{dt} &= -\frac{\partial H}{\partial s'} = k_2 e^{-\frac{r}{T}} q(\lambda_1 + \lambda_2 + \lambda_3), \\
 \frac{d\lambda_2}{dt} &= -\frac{\partial H}{\partial q} = 2Wq + k_2 e^{-\frac{r}{T}} s'(\lambda_1 + \lambda_2 - \lambda_3), \\
 \frac{d\lambda_3}{dt} &= -\frac{\partial H}{\partial c} = k_{-2} e^{-\frac{r}{T}} (\lambda_1 - \lambda_2 + \lambda_3) - k_3 e^{-\frac{r}{T}} n(\lambda_2 - \lambda_3 - \lambda_4 + \lambda_5), \quad (21) \\
 \frac{d\lambda_4}{dt} &= -\frac{\partial H}{\partial n} = k_3 e^{-\frac{r}{T}} c(-\lambda_2 + \lambda_3 + \lambda_4 - \lambda_5), \\
 \frac{d\lambda_5}{dt} &= -\frac{\partial H}{\partial d} = -2d.
 \end{aligned}$$

Позначимо:

$$\begin{aligned}
 \Phi(t) &= \lambda_1(-k_2 e^{-\frac{r}{T}} s'q + k_{-2} e^{-\frac{r}{T}} c) + \lambda_2(-k_2 e^{-\frac{r}{T}} s'q + k_{-2} e^{-\frac{r}{T}} c + k_3 e^{-\frac{r}{T}} cn) + \\
 & + \lambda_3(k_2 e^{-\frac{r}{T}} s'q - k_{-2} e^{-\frac{r}{T}} c - k_3 e^{-\frac{r}{T}} cn) + \lambda_4(-k_3 e^{-\frac{r}{T}} cn) + \lambda_5(k_3 e^{-\frac{r}{T}} cn). \quad (22)
 \end{aligned}$$

Враховуючи (22), функцію Гамільтона–Понтрягіна (20) запишемо у вигляді:

$$H = d^2 - Wq^2 + e^{-\frac{r}{T}} [\Phi(t)]. \quad (23)$$

Звідси бачимо, що максимальні значення  $H$  будуть при  $T = T^*(t)$ , де:

$$T^*(t) = \begin{cases} T_e^{\min}, & \text{якщо } \Phi(t) > 0, \\ T_e^{\max}, & \text{якщо } \Phi(t) < 0, \\ \text{будь-яке з } [T_e^{\min}, T_e^{\max}], & \text{якщо } \Phi(t) = 0. \end{cases} \quad (24)$$

Отже, оптимальні траєкторія  $(s^*, q^*, c^*, n^*, d^*)$  на керування  $T^*$  можуть бути побудовані в результаті розв'язку крайової задачі:

$$\begin{aligned}
 \frac{ds'}{dt} &= -k_2 e^{-\frac{r}{T}} s'q + k_{-2} e^{-\frac{r}{T}} c, \\
 \frac{dq}{dt} &= -k_2 e^{-\frac{r}{T}} s'q + k_{-2} e^{-\frac{r}{T}} c + k_3 e^{-\frac{r}{T}} cn,
 \end{aligned}$$



$$\begin{aligned}
 \frac{dc}{dt} &= k_2 e^{-\frac{r}{T}} s' q - k_{-2} e^{-\frac{r}{T}} c - k_3 e^{-\frac{r}{T}} c n, \\
 \frac{dn}{dt} &= -k_3 e^{-\frac{r}{T}} c n, \\
 \frac{dd}{dt} &= k_3 e^{-\frac{r}{T}} c n, \\
 \frac{d\lambda_1}{dt} &= -\frac{\partial H}{\partial s'} = k_2 e^{-\frac{r}{T}} q (\lambda_1 + \lambda_2 + \lambda_3), \\
 \frac{d\lambda_2}{dt} &= -\frac{\partial H}{\partial q} = 2Wq + k_2 e^{-\frac{r}{T}} s' (\lambda_1 + \lambda_2 - \lambda_3), \\
 \frac{d\lambda_3}{dt} &= -\frac{\partial H}{\partial c} = k_{-2} e^{-\frac{r}{T}} (\lambda_1 - \lambda_2 + \lambda_3) - k_3 e^{-\frac{r}{T}} n (\lambda_2 - \lambda_3 - \lambda_4 + \lambda_5), \\
 \frac{d\lambda_4}{dt} &= -\frac{\partial H}{\partial n} = k_3 e^{-\frac{r}{T}} c (-\lambda_2 + \lambda_3 + \lambda_4 - \lambda_5), \\
 \frac{d\lambda_5}{dt} &= -\frac{\partial H}{\partial d} = -2d.
 \end{aligned} \tag{25}$$

З крайовими умовами:

$$s'(t_1) = s_0, \quad q(t_1) = q_0, \quad c(t_1) = c_0, \quad n(t_1) = n_0, \quad d(t_1) = d_0;$$

$$\lambda_1(t_2) = 0, \quad \lambda_2(t_2) = 0, \quad \lambda_3(t_2) = 0, \quad \lambda_4(t_2) = 0, \quad \lambda_5(t_2) = 0.$$

**Теорема 3.** Для досить малого значення  $t_2$  розв'язок системи (25) є єдиним.

**Доведення.** Припустимо навпаки, що існують два розв'язки (25), а саме:

$$T^* = (s'^*, q^*, c^*, n^*, d^*, \lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*, \lambda_5^*),$$

$$T^{**} = (s'^{**}, q^{**}, c^{**}, n^{**}, d^{**}, \lambda_1^{**}, \lambda_2^{**}, \lambda_3^{**}, \lambda_4^{**}, \lambda_5^{**}).$$

Праві частини системи (25) є Лівшицевими функціями аргументів  $s', q, c, n, d, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ . Звідси існує стала  $C > 0$  така, що:

$$\|T^*(t) - T^{**}(t)\| \leq \int_{t_1}^{t_2} C \left( \|T^*(s) - T^{**}(s)\| \right) ds. \tag{26}$$

Застосовуючи до (26) теорему про середнє значення, маємо, що існує момент часу  $\xi: t_1 \leq \xi \leq t_2$  такий, що:

$$\|T^*(t) - T^{**}(t)\| \leq t_2 C \left( \|T^*(\xi) - T^{**}(\xi)\| \right)$$

при всіх  $t \in [t_1, t_2]$ . Якщо ж виберемо  $t_2$  таким, що  $t_2 < \frac{1}{C}$ , то отримуємо суперечність.

## ВИСНОВКИ

Сформульовано задачу оптимального керування протікання стадії елонгації у ПЛР. Застосовано принцип максимуму Понтрягіна до задачі оптимального керування та сформульовано необхідну умову оптимальності. Отримані результати необхідні для чисельного обчислення оптимального керування досліджуваної стадії та допоможуть мінімізувати необхідний час для реалізації стадії елонгації, що в майбутньому дозволить мінімізувати час проведення полімеразно-ланцюгової реакції.

У подальших дослідженнях потрібно дослідити керованість інших стадій ПЛР та провести відповідне чисельне обчислення оптимального керування.

## ЛІТЕРАТУРА

1. *Сверстюк А.С., Бігуняк Т.В., Перевізник Б.О.* Огляд методів та моделей полімеразно-ланцюгової реакції // Медична інформатика та інженерія. — № 3. — С. 97–100.
2. *Путинцева Г.Й.* Медична генетика: підручник. — К.: Медицина, 2008. — С. 81–87.
3. *Буляница А.Л.* Математическое моделирование циклических режимов управления температурой при реализации полимеразно-цепной реакции (ПЦР) для метода молекулярных колоний (ММК) на микрочипе // Научное приборостроение. — 2011. — **21**, № 1. — С. 87–96.
4. *Arnheim N.* Polymerase chain reaction strategy // Annual review of biochemistry. — 1992. — **61**, XIV+1359P. — P. 131–156.
5. *Xiangchun X., Sinton D., Dongqing L.* Thermal end effects on electroosmotic flow in capillary // Int. J. of Heat and Mass transfer. — 2004. — **47**. — P. 3145–3157.
6. *Stone E., Goldes J., Garlick M.* A multi-stage model for quantitative PCR // Mathematical biosciences and engineering. — [http://hs.umt.edu/math/research/technical-reports/documents/2009/10\\_pcrpaper4.pdf](http://hs.umt.edu/math/research/technical-reports/documents/2009/10_pcrpaper4.pdf)
7. *Lukes D.L.* Differential Equations: Classical to Controlled // Academic Press, New York. — 1982. — **162**. — P. 322.
8. *Piccinini L. C., Stampacchia G., Vidossich G.* Ordinary Differential Equations in Rn. Problems and Methods Ordinary // Berlin-Heidelberg-New York-Tokyo, Springer-Verlag. — 1984. — **XII**. — P. 385.
9. *Macki J., Strauss A.* Introduction to Optimal Control Theory, Springer-Verlag // New York. — 1982. — **XIV**. — P. 168.
10. *Fleming W.H., Rishel R.W.* Deterministic and Stochastic Optimal Control, Springer Verlag // New York. — 1975. — **XIII**. — P. 222.
11. *Kamien M.I., Schwartz N.L.* Dynamic Optimization // North-Holland, Amsterdam. — 1991. — **3**. — P. 272.
12. *Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе П.В., Мищенко Е.Ф.* Математическая теория оптимальных процессов. — М.: Наука, 1983. — 393 с.
13. *Kelly K., Kostin M.* Non-Arrhenius rate constants involving diffusion and reaction // Journal of Chemical Physics. — 1986. — **85**, issue 12. — P. 7318.

Надійшла 28.02.2014

## **PRIMITIVE PROGRAMING ALGEBRA: GENERAL APPROFCH TO A PROBLEM OF FUNCTIONAL COMPLETENESS**

**P.O. YAHANOV, D.I. REDKO, I.V. REDKO, T.L. ZAKHARCHENKO**

The goal of the research is development of scientific foundations of programming problems solutions genesis. Investigations carried out are based on algebraic research methods of programs and compositional programming methods. Basis of the last ones consists of program algebras with special classes of functions as carriers, and compositions that represent abstractions from program synthesis tools as operations. Problems of completeness in classes of computable functions that took one of the most important places in programming problems are well defined and solved in the context of program algebras. Universal method for the problem of completeness solution in primitive program algebras (PPA) on different classes of computable functions proposed in the article. Results achieved are presented as series of original statements, lemmas and theorems. The results can be applied in algebraic characteristics research of different computable functions classes in problems of programming language semantics formalization

### **INTRODUCTION**

Today's posture in IT field and, particularly, in programming, considerably defined by process of more and more vast it's penetration in all aspects of human's life. Naturally, with every step taken in that direction, requirements made to quality of product produced and effectiveness of its production are constantly increasing. Despite of impressive and speaking for themselves results achieved with programming activity (PA) today, it becomes more obvious that the results in majority are extensive, so sustainment of this tendency becomes more problematic and impossible in foreseeable future. The reason is typical for nowadays understanding of PA, particularly, programming, its excessive simplicity, which is not corresponding level of complexity of problematic indicated.

As for programming, simplicity of its understanding led to the fact, that, mainly, attention paid to results of programming without consideration of processes, which made that results possible. It makes process of programming problems solution too subjective, regarding intuitive principles of paramount importance. These facts are not allowing us seriously discuss problems of software quality management, effectiveness of its production and preservation of investments. An avalanche-like increase of number of such facts stimulated discussions about development of crisis in programming, depression in IT industry etc [1–3]. Now, not a crisis of the field should be discussed itself, but crisis of its ways of development! Statements, made above, one more time demonstrate that contemporary programming, and the overall field can not effectively develop now exclusively on objectively-intuitive basis, which is the source of different concepts of PA. Long ago, problems of the field became so significant and so complex, that intuitive considerations must be objectified adequately and supplemented with precise researches and developments as far as possible. The matter is in the main

intuition carrier of PA — programming as process of software creation. Questions connected with revelation of programming languages semantic play here vital role. That is why research of following problematic is **objective** of the paper. Paramount role here plays compositional paradigm [4–8], as methodological consideration basis of whole diversity of general as well as particular software creation methods. Namely that methods, to be more precise, their explications in the form of different classes of compositions build up the **object** of the research. The **subject** of the research is problem of computable functions characteristics classes creation. The functions are on different carriers in primitive programming algebras (PPAs) [9–11]. The main attention is paid to search of such algebra's generative sets and bases. In process of research, along with general results, description of computable functions class on records also received. It supplements analogical results for natural numbers tuples, graph transformers and multi-sets [10–15].

All undefined generic mathematical conceptions and designations are interpreted in sense [16], and concepts of numerations theory and theory of algorithms interpreted as in [17, 18].

## GENERAL STATEMENTS

The carrier of PPA are  $n$ -ary functions and  $n$ -ary predicates (or simply functions and predicates) ( $n=1,2,\dots$ ). The signature of PPA (denoted as  $\Omega$ ) consists of superposition, branch, loop operations, which are represent adequate specifications of the main methods of software or computational hardware design, which are peculiar to majority of high-level programming languages [1, 4–7, 9]. Let us make formal definitions of the operations. Termal, rather than operator notation of functions will be preferred for convenience and compactness [18]. Usage of specific notation form in every individual case conditioned by the fact that different notations present fundamentally different viewpoints on the entity they describe. In other words, operator notation used in cases when it is important to reveal genesis of entity described, termal notation is important for description of result genesis. Although, those forms are interchangeable like texts in certain senses, those keypoints arrangement is important because it presents completely natural dominant of genesis relatively to its result.

Let  $m$  functions  $f_1, \dots, f_m$  of the same arity (for example,  $k$ ) of type  $A^k \rightarrow B$  be defined on certain set  $A$  with values from set  $B$  (it is no need to preserve  $A \cap B = \emptyset$ , moreover  $A \cap B \neq \emptyset$  is acceptable case too). Also, let  $m$ -ary function  $f$  with values in certain set  $C$  be defined on set  $B$ . Consider  $k$ -ary function  $g: A^k \rightarrow C$  with value  $g(\langle a_1, \dots, a_k \rangle) \cong f(f_1(\langle a_1, \dots, a_k \rangle), \dots, f_m(\langle a_1, \dots, a_k \rangle))$  on argument  $\langle a_1, \dots, a_k \rangle$ . In this case function  $g$  is the result of a  $(m+1)$ -ary *superposition* application, denoted as  $S^{m+1}$ , to  $m$ -tuple of functions  $\langle f, f_1, \dots, f_m \rangle$ , i.e.  $g = S^{m+1}(\langle f, f_1, \dots, f_m \rangle)$ . Hereinafter in this document the designation “ $\cong$ ” means the generalized equality [19].

Now, additionally let function  $h$  of type  $A^k \rightarrow B$  and  $m$ -valued function  $\beta: B \rightarrow \{1, 2, \dots, m\}$  be defined.  $k$ -ary function  $g: A^k \rightarrow B$  is built from func-

tions  $h, f_1, \dots, f_m$  by  $(m+1)$ -ary parametric *branch* operation  $\diamond_{\beta}^{m+1}$  if for any argument  $\langle a_1, \dots, a_k \rangle \in A^k$  the value of function  $g(\langle a_1, \dots, a_k \rangle)$  defined as:

$$g(\langle a_1, \dots, a_k \rangle) \cong f_r(\langle a_1, \dots, a_k \rangle), \text{ if } \beta(h(\langle a_1, \dots, a_k \rangle)) = r, (1 \leq r \leq m).$$

Note, that described parametric branch operation represents adequate specification of well-known method of software design — *case\_of*. Ternary *branch* operation  $\diamond$ , which puts in correspondence to two functions  $f_i: A^k \rightarrow B, i=1,2$  and one predicate  $p: A^k \rightarrow \{T, F\}$  a  $k$ -ary function  $g \equiv \diamond(\langle p, f_1, f_2 \rangle)$  with values defined on any argument  $\langle a_1, \dots, a_k \rangle \in A^k$  as:

$$g(\langle a_1, \dots, a_k \rangle) \cong \begin{cases} f_1(\langle a_1, \dots, a_k \rangle), p(\langle a_1, \dots, a_k \rangle) \cong T \\ f_2(\langle a_1, \dots, a_k \rangle), p(\langle a_1, \dots, a_k \rangle) \cong F \end{cases}$$

may be useful partial case.

Finally, let us complete our list of definitions with  $k$ -ary predicate  $p: A^k \rightarrow \{T, F\}$ . Consider  $k$ -ary function  $g: A^k \rightarrow B$  with value  $g(\langle a_1, \dots, a_k \rangle)$  on arbitrary argument  $\langle a_1, \dots, a_k \rangle \in A^k$  equal to the first component of the first tuple from sequence of tuples  $[\langle a_1^i, \dots, a_k^i \rangle]_{i=0,1,2,\dots}$ , where  $a_j^0 = a_j, j=1,2,\dots,k$  and  $a_j^{i+1} = f_j(\langle a_1^i, \dots, a_k^i \rangle), j=1,2,\dots,k$ , for which (denote it as  $\langle a_1^s, \dots, a_k^s \rangle$ , for example)  $p(\langle a_1^s, \dots, a_k^s \rangle) = F$  in case if for all  $r < s$ , if such argument exists, value of  $p(\langle a_1^r, \dots, a_k^r \rangle) \cong T$ . Function  $g$  built by application of  $(m+1)$ -loop operation to functions of  $(m+1)$ -tuple  $\langle p, f_1, \dots, f_m \rangle$ , i.e.  $g = *^{m+1}(\langle p, f_1, \dots, f_m \rangle)$ . Thus, according to statements made before,  $g(\langle a_1, \dots, a_k \rangle) = a_1^s$ .

Note, previously, in order to denote introduced operations, we used exclusively operator notation. By using termal notation of operations from PPA signature, we will evidently denote only variables with considerably used values. For instance, for loop operation notation like  $p(x_1, \dots, x_m) *_{y_1 \dots y_k} \langle f_1(z_1^1, \dots, z_{m_1}^1), \dots, f_m(z_1^k, \dots, z_{m_k}^k) \rangle$  will be used with those variables denoted, on which functions and predicate considerably depend. At the same time, function  $f_j(z_1^j, \dots, z_{m_j}^j)$  changes variable  $y_j, j=1,2,\dots,k$  and variable  $y_1$  considered as an “output”. Operator notation can be easily reconstructed from this notation.

Let us declare certain countable set  $D$  and for any natural number  $k > 0$  consider classes  $\Phi^k$  of partial  $k$ -ary functions and predicates of types:  $D^k \rightarrow D$  and  $D^k \rightarrow \{T, F\}$  accordingly, and  $\Phi \equiv \bigcup_k \Phi^k, k=1,2,3,\dots$  of partial multiplace functions and predicates on set  $D$ . Further, functions and predicates on  $D$  will be denoted as  $D$ -functions ( $D$ -predicates) and will belong to set  $\Phi$ . Computability

on  $D$  is defined as numeric computability [17, 20]. PPA with carrier formed from partial recursive functions (computable functions, pr-functions) and partial recursive predicates (computable predicates, pr-predicates on  $D$  will be denoted as  $A_D^{pr}$ . Generative set of  $A_D^{pr}$  will be defined as its *complete system* (CS).

Complete system of PPA will be its  $I_m^n$ -basis, if any system produced by exclusion of any elements from the CS, except selective function, will not be complete.

Some terms, designations, properties and associated results, showed below, may be useful during study of PPA complete systems [10].

**Property 1.**  $n$ -ary function  $f$  preserves set  $L \subseteq D$ ,  $L \neq \emptyset$ , if  $f(\underbrace{L \times \dots \times L}_n) \subseteq L$ . Here  $f(\underbrace{L \times \dots \times L}_n) \equiv \{f(\langle a_1, \dots, a_n \rangle) \mid \langle a_1, \dots, a_n \rangle \in \underbrace{L \times \dots \times L}_n\}$ .

Now, let  $D$  be set of composite objects (composed from certain components). Assume that universal set of such components is countable. Denote it as  $B$ . Lets declare surjective map  $\beta: D \rightarrow 2^B$ , where  $2^B$  is set of all finite subsets of set  $B$ . Hence, for any  $d \in D$   $\beta(d) \in 2^B$  is set of elements from  $B$ , which  $d$  composed of. Those elements are called denotates of  $d$ .

**Property 2.**  $n$ -ary function  $f$   $\beta$ -preserves denotates, if finite set  $B_f \subset B$  exists and for any  $d \equiv \langle d_1, \dots, d_n \rangle \in \text{dom } f$  expression  $\beta(f(\langle d_1, \dots, d_n \rangle)) \subseteq \bigcup_{i=1}^n \beta(d_i) \cup B_f$  is correct.

Note that described properties of  $D$ -functions are preserved in signature  $\Omega$  [10]. This allows to form several simple and essential conditions of completeness of  $A_D^{pr}$  CS [10].

**Statement 1.** Any complete system of  $A_D^{pr}$  contains at least one  $D$ -function that does not preserve set  $L$  for any non-empty set  $L$  ( $L \subset D, L \neq \emptyset$ ).

**Statement 2.** Any complete system of  $A_D^{pr}$  contains at least one  $D$ -function, which does not  $\beta$ -preserve denotates.

## THE CONCEPT OF COMPLETENESS IN CLASSES OF PR-FUNCTIONS AND PR-PREDICATES

Consider general method for PPA of  $D$ -functions and  $D$ -predicates complete systems finding. It will be represented by series of interconnected results, introduced as proved lemmas and theorems. First, let us define some notions, useful conventions and denotations.

Let two countable sets  $D_1$  and  $D_2$  be defined. Assume that for every of those sets exists effective numeration  $\alpha_1: N \rightarrow D_1$  and  $\alpha_2: N \rightarrow D_2$ . Also, PPAs

$A_{D_1}^{\text{pr}}$  and  $A_{D_2}^{\text{pr}}$  are defined. Elements of sets  $D_1$  and  $D_2$  designated with lowercase letters:  $a^1, b^1, \dots$  and  $a^2, b^2, \dots$ , may be with subscript. Let complete system  $\sigma_{D_1}$  of PPA  $A_{D_1}^{\text{pr}}$  is defined and injective constructive mappings  $\varphi: D_2 \rightarrow D_1$  and  $\Phi: D_1 \rightarrow D_2$  are given. Sets  $\varphi(D_2) \equiv \{\varphi(d) \mid d \in D_2\}$  and  $\Phi(D_1) \equiv \{\Phi(d) \mid d \in D_1\}$  are recursive [18]. Consider approach to solution of completeness problem for algebraic structure  $A_{D_2}^{\text{pr}}$ .

To designate  $D_1$ - and  $D_2$ -functions, lowercase ( $f, g, \dots$ ) and uppercase ( $F, G, \dots$ ) letters accordingly will be used. Letters  $p, r, \dots$  and  $P, R, \dots$  are used for designation of  $D_1$ - and  $D_2$ -predicates accordingly. When using termal notation, variables for  $D_1$ -functions and  $D_1$ -predicates are designated with lowercase roman letters  $x, y, z, \dots$ , and  $D_2$ -function and  $D_2$ -predicates — with lowercase Greek letters  $\tau, \xi, \pi, \dots$ , subscripts and superscripts may be used in both cases.

**Definition 1.**  $\varphi(D_2)$ -function  $f(x_1, \dots, x_n)$  is  $D_1$ -image of  $D_2$ -function  $F(\tau_1, \dots, \tau_n)$ , if for any  $a_1^2, \dots, a_n^2 \in D_2$  expression  $f(\varphi(a_1^2), \dots, \varphi(a_n^2)) \equiv \varphi(F(a_1^2, \dots, a_n^2))$  is true.

**Definition 2.**  $\varphi(D_2)$ -predicate  $p(x_1, \dots, x_n)$  is  $D_1$ -image of  $D_2$ -predicate  $P(\tau_1, \dots, \tau_n)$ , if for any  $a_1^2, \dots, a_n^2 \in D_2$  expression  $p(\varphi(a_1^2), \dots, \varphi(a_n^2)) \equiv P(a_1^2, \dots, a_n^2)$  is true.

Lets show that relations «to be an image of function» and «to be an image of predicate», declared with definitions listed, preserve property of partial recursiveness. In other words, listed theorem below is true.

**Theorem 1.**  $D_1$ -image of  $D_2$ -pr-function ( $D_2$ -pr-predicate) is  $D_1$ -pr-function ( $D_1$ -pr-predicate).

Indeed, it is easy to check that  $\varphi$ , as mapping of numerated set  $\langle D_2, \alpha_2 \rangle$  to numerated set  $\langle \varphi(D_2), \alpha_2 \cdot \varphi \rangle, \alpha_2 \cdot \varphi: N \rightarrow \varphi(D_2): \forall k = 1, 2, \dots, \alpha_2 \cdot \varphi(k) \equiv \varphi(\alpha_2(k))$  is pr-equivalence ([21], p. 150–160), because of constructiveness of mapping  $\varphi$  and effectiveness of numerations  $\alpha_1$  and  $\alpha_2$ . Hereinafter in this document the designation  $\alpha_2 \cdot \varphi$  means standard multiplication of functions  $\alpha_2$  and  $\varphi$ :  $\text{dom}(\alpha_2 \cdot \varphi) \subseteq \text{dom}(\alpha_2)$ ,  $\text{ran}(\alpha_2 \cdot \varphi) \subseteq \text{ran}(\varphi)$  and for any  $d \in \text{dom}(\alpha_2 \cdot \varphi)$  value of this function  $\alpha_2 \cdot \varphi(d) \equiv \varphi(\alpha_2(d))$ .

After application of theorem 2.1.5 [17], lemma 1 will be true.

**Lemma 1.**  $D_1$ -image of  $D_2$ -pr-function ( $D_2$ -pr-predicate) is  $\varphi(D_2)$ -pr-function ( $\varphi(D_2)$ -pr-predicate).

Hence, recursiveness of set  $\varphi(D_2)$  results.

**Lemma 2.** Any  $\varphi(D_2)$ -pr-function is  $D_1$ -pr-function. The same for  $\varphi(D_2)$ -pr-predicates.

This lemma results truth of theorem 1.

**Definition 3.**  $D_2$ -function  $F(\tau_1, \dots, \tau_n)$  is  $D_2$ -model of  $D_1$ -function  $f(x_1, \dots, x_n)$ , if expression  $F(\Phi(a_1^1), \dots, \Phi(a_n^1)) \cong \Phi(f(a_1^1, \dots, a_n^1))$  holds for any  $a_1^1, \dots, a_n^1 \in D_1$ .  $D_2$ -model of  $D_1$ -predicate defined in the same way.

Let  $\psi = \varphi \cdot \Phi$  ( $f \cdot g$  is standard function multiplication, i.e. such function for which  $\text{dom}(f \cdot g) \subseteq \text{dom}(f)$ ,  $\text{ran}(f \cdot g) \subseteq \text{ran}(g)$  and which any  $d \in \text{dom}(f \cdot g)$  maps to value  $f \cdot g(d) \equiv g(f(d))$ , if  $f(d) \in \text{dom}(g)$ ). Obviously, that  $\psi : D_2 \rightarrow \Phi(\varphi(D_2))$ ,  $\Phi(\varphi(D_2)) \subseteq D_2$  — bijection. Thus, it is possible to assume that mapping, which inverse mapping to  $\psi$ , exists. Some mapping extension  $\psi^{-1} : \Phi(\varphi(D_2)) \rightarrow D_2$  will be designated as  $\chi : D_2 \rightarrow D_2$ . In other words,  $D_2$ -functions  $\psi$  and  $\chi$  are playing roles of coding and decoding functions accordingly. Let  $\sigma_{D_2}$  is designator for set of  $D_2$ -functions and  $D_2$ -predicates for which, firstly,  $D_2$ -model of  $D_1$ -function ( $D_1$ -predicate) from CS  $\sigma_{D_1}$  may be built from  $D_2$ -functions and  $D_2$ -predicates of set  $\sigma_{D_2}$  by finite number of application of operations from signature  $\Omega$  and, secondly,  $D_2$ -functions  $\psi$  and  $\chi$  may be built from  $D_2$ -functions and  $D_2$ -predicates of  $\sigma_{D_2}$  in analogical manner.

**Definition 4.** Sextuple  $\Sigma \equiv \langle D_1, D_2, \sigma_{D_1}, \sigma_{D_2}, \psi, \chi \rangle$  is called allowable system (AS) and tuple  $\langle D_1, \sigma_{D_1} \rangle$  is its basis.

Obviously, that in context of coding and decoding function lemma 3 true.

**Lemma 3.** Let  $F(\tau_1, \dots, \tau_n)$  is  $D_2$ -pr-function, and  $H(\pi_1, \dots, \pi_n)$  is  $D_2$ -model of  $D_1$ -image of function  $F(\tau_1, \dots, \tau_n)$ , then  $F(a_1^2, \dots, a_n^2) \cong \chi(H(\psi(a_1^2), \dots, \psi(a_n^2)))$  is true for any  $a_1^2, \dots, a_n^2 \in D_2$ .

**Lemma 4.** Let  $P(\tau_1, \dots, \tau_n)$  be  $D_2$ -pr-predicate, and  $R(\pi_1, \dots, \pi_n)$  be  $D_2$ -model of  $D_1$ -image of predicate  $P(\tau_1, \dots, \tau_n)$ , then  $P(a_1^2, \dots, a_n^2) \cong R(\psi(a_1^2), \dots, \psi(a_n^2))$  is true for any  $a_1^2, \dots, a_n^2 \in D_2$ .

Hence, theorem 2 is true.

**Theorem 2.**  $\sigma_{D_2}$  is CS of PPA  $A_{D_2}^{\text{pr}}$ .

Considered, that there are few as general as possible requirements to sets  $D_1, D_2$  and to its elements the nature of our constructions are maximally general. This allows to formulate simple, but effective condition of completeness of functions system in PPA.

So, if  $D_1, D_2, \sigma_{D_1}, \sigma_{D_2}, \psi, \chi$  are objects, mentioned above, then theorem 3 is true.

**Theorem 3.** If  $D_1, D_2, \sigma_{D_1}, \sigma_{D_2}, \psi, \chi$  is AS, then  $\sigma_{D_2}$  is CS of PPA  $A_{D_2}^{\text{pr}}$ .

Results gained are giving complete enough idea about building method of complete systems for PPA of partially recursive functions and predicates on countable sets. This method will be applied below in order to solve problem of PPA completeness in class of pr-functions and pr-predicates on pragmatically significant in programming data type — set of records.



## PPA OF PR-FUNCTIONS AND PR-PREDICATES ON SET OF RECORDS

Number of different intuitive interpretations of term «record» exists in information technologies and programming. Despite the fact that some interpretations of «record» significantly differ one from another, all of them tend to use adequately a concept of *named structures* to describe complex aggregated entities. Often, those interpretations burdened with minor partial details, blurring significance of naming mechanisms. However, as experience shows, named structures form «common denominator», through which all other aspects of problems solved should be considered. Namely, this tendency is basis for all following constructions.

Let  $V$  and  $W$  are non-empty countable sets of elements, interpreted as sets of names and values (denotates) accordingly. In general case it is allowed, that some names may play role of values and vice versa, i.e. it is possible that  $V \cap W \neq \emptyset$ .

We need to define some denotations, introduce main and auxiliary definitions in order to go further. Some of definitions will be given now, others — later, as may be necessary. All undefined terms and designations are given in [7].

One of the main concepts of this section is record. Set of all records on sets of names  $V$  and values  $W$  designated as  $Z^{(V,W)}$ . Now, introduce definition of record.

**Definition 5.** Record on sets of names  $V$  and values  $W$  (or simply record, it is clear from context what do  $V$  and  $W$  mean) is finite functional binary relation between set of names  $V$  and set of values  $W$ .

To designate record uppercase letters  $I, J, K, \dots$  will be used. Lowercase letters  $u, v, w, \dots$  are used to designate names of record elements, letters  $a, b, c, d, \dots$  are their values and letters  $\lambda, \mu, \eta, \dots$  are elements of records. In all cases subscripts may be used. Left subscripts and (or) superscripts may be used to designate names and (or) values of elements of record may be used. For example, let  $\lambda = (v, a)$ . Then such designations of this element of record as  ${}^v\lambda$ ,  ${}_a\lambda$  and  ${}_a^v\lambda$  may be used,

Hereinafter in the article so called «schemes», which represent name templates of correspondent records, may be used along with records.

**Definition 6.** The scheme of record  $K$  is finite set of names  $\{v_1, \dots, v_n\}$ , which represent projection of the record by the first component, i.e.  $\{v_1, \dots, v_n\} = \text{pr}_1(K)$ , where  $\text{pr}_i$  is function of projection by  $i$ -th component of  $m$ -ary relation ( $1 \leq i \leq m$ ) [7].

Scheme of the record  $I$  is designated as  $sh(I) = \{v_1, \dots, v_n\}$ , and record itself named for compactness as  $sh(I)$ -record or record of  $sh(I)$  type. In case when type of record  $I$  must be defined explicitly, designation  $I^{sh(I)}$  will be used. Set of all records of  $\{v_1, \dots, v_n\}$  type designated as  $Z[\{v_1, \dots, v_n\}]$ . Couple particular cases of those notations take place:  $I^\emptyset = \emptyset$  and  $Z[\emptyset] = \{I^\emptyset\}$ . As follows from above, it is obvious that  $Z^{(V,W)} = \bigcup_{V' \in 2^V} Z[V']$ .

In unfolded notation record is designates as  $I^{\{v_1, \dots, v_n\}} \equiv \{(v_1, d_1), \dots, (v_n, d_n)\}$ .

For correct usage of numeric computability on set of records, it is required to prove existence of effective numeration of set  $Z^{(V,W)}$ . Given the countability of sets  $V, W$ , as well as the fact that in this case is not so much important form of presentation names and values, how important their fundamentally different role, without limitation of subsequent constructions generality, we can assume that  $V = W = N$ . It can be deduced from a context which of the roles meant in every single case. Therefore, any further formal constructions will be carried out on set  $Z^{(N,N)}$ , which is special case of  $Z^{(V,W)}$ .

Few steps need to be done to construct the numeration. Firstly, we need to take in account that for any non-empty record  $I = \{(v_1, d_1), \dots, (v_m, d_m)\}$  its number concur with number of finite set  $M_I = \{n_1, \dots, n_m\}$ , where  $n_i$  is number of named element of record  $(v_i, d_i)$  (effective numeration of set  $N^2$  defined in [12]). Number (unique identifier) for set  $M_I$  itself defined, for example, as  $\alpha'(M_I) \equiv 2^{n_{j_1} + 1} \cdot 3^{n_{j_2} + 1} \cdot \dots \cdot p_{m-1}^{n_{j_m} + 1}$ , where  $n_{j_1} < n_{j_2} < \dots < n_{j_m}$ ,  $n_{j_s} \in M$ ,  $s = 1, \dots, m$ , and  $p_i$ ,  $i = 0, 1, 2, \dots$  is  $i$ -th prime number ( $p_0 = 2$ ,  $p_1 = 3$ ,  $p_2 = 5$ , ...). Than numeration  $\alpha_{Z^{(N,N)}}$  of set  $Z^{(N,N)}$  defined through piecewise scheme

$$\alpha'_{Z^{(N,N)}}(K) = \begin{cases} 1, & \text{if } I = \emptyset, \\ \alpha'(M_K), & \text{else,} \end{cases}$$

where  $K$  is certain record. Namely,  $\alpha_{Z^{(N,N)}} \equiv (\alpha'_{Z^{(N,N)}})^{-1}$ .

Now, consider to find of complete system of PPA  $A_{Z^{(N,N)}}^{\text{Pr}}$  itself. From the results gained above, conclusion followed that the solution of this problem reduces to the corresponding AS construction. Refer to concept of multi-set, mentioned in [14, 15]. Let  $U$  be some finite, may be empty, set.

**Definition 7.** Multi-set  $\alpha$  with  $U$  basis is finitely defined function of  $\alpha : U \rightarrow N^+$  type, where  $N^+ \equiv N \setminus \{0\} = \{1, 2, 3, \dots\}$ . If designation of  $\alpha$  basis is necessary, notation  $\alpha^U$  will be used.

It would not be a great loss of generality, if we would assume that  $U \subseteq N$ . Collection of all multi-sets with basis  $U$  designated as  $M_U$ . Then, obvious, that  $M \equiv \bigcup_{U \in 2^N} M_U$  is set of all multi-sets (on  $N$ ).

Elements of set  $M$  are designated with lowercase Greek letters  $\alpha, \beta, \delta, \dots$ , may be with subscripts and superscripts. Element of multi-set will be designated as tuple  $\langle a, d \rangle$ , every component of which may be with subscript and superscript. Here  $a$  as the first component of tuple, its argument, the second is  $d$  — value (denotate, multiplicity). Two terms are related to multi-sets for convenient: characteristics  $\chi_\alpha$  and full image  $f[\alpha]$ . The first one is parametric function  $\chi_\alpha : D \rightarrow N$  with values defined with piecewise scheme:

$$\chi_\alpha(a) = \begin{cases} \alpha(a), & \text{if } a \in \text{dom } \alpha, \\ 0, & \text{else,} \end{cases} \text{ for all } a \in N.$$

The second one is creates multi-set  $f[\alpha]^{f(U)}$  from multi-set  $\alpha^U$  and given function  $f: N \rightarrow N$ , where  $f(U)$  is full image of set  $U$  relatively to function  $f$ , and characteristics of arbitrary argument  $a$  of this multi-set defined as:  $\chi_{f[\alpha]^{f(U)}}(a) = \sum_{a' \in f^{-1}(a)} \chi_\alpha(a')$ . Here  $f^{-1}(a)$  is full image of element  $a$  relatively to function  $f$ . In case of empty set of summarands, sum assumed to be 0.

Now consider PPA  $A_M^{pr}$  of  $M$ -pr-functions and  $M$ -pr-predicates. Following collection  $\sigma_M$  of  $M$ -pr-functions and  $M$ -pr-predicates is of interest for us. It includes *predicate of equality*  $\alpha = \beta: \alpha = \beta \Leftrightarrow \forall a(a \in N \Rightarrow \chi_\alpha(a) = \chi_\beta(a))$ ; *function of unification*  $\cup_{All}$ , which from two multi-sets  $\alpha$  and  $\beta$  produces such multi-set  $\alpha \cup_{All} \beta$ , that for any argument  $a$  its characteristics equals to  $\max(\chi_\alpha(a), \chi_\beta(a))$ , i.e.  $\chi_{\alpha \cup_{All} \beta}(a) = \max(\chi_\alpha(a), \chi_\beta(a))$ ; *function of direct junction*  $\otimes$ , which from two arbitrary multi-sets  $\alpha^{U_\alpha}$  and  $\beta^{U_\beta}$  produces new multi-set  $(\alpha^{U_\alpha} \otimes \beta^{U_\beta})^{U_\alpha \times U_\beta}$ , characteristics of arguments  $\langle a_1, a_2 \rangle$  defined as:  $\chi(\langle a_1, a_2 \rangle) = \chi_{\alpha^{U_\alpha}}(a_1) \cdot \chi_{\beta^{U_\beta}}(a_2), \forall a_1, a_2 \in N$ ; *functions of addition*  $\oplus$  and *subtraction*  $\div$ , defined with expressions  $\alpha \oplus \beta = +[\alpha \otimes \beta]$  and  $\alpha \div \beta = -[\alpha \otimes \beta]$  accordingly (here « $\cdot$ » — truncated distinction [18, 19]); *constant functions*  $\{1^1\}(\alpha)$  and  $\emptyset_m(\alpha)$ , which produce multi-sets  $\langle 1, 1 \rangle$  and  $\emptyset_m \equiv \alpha^\emptyset$  accordingly; *function of multiplicity*  $\varphi$  which produces from two multi-sets  $\langle n, 1 \rangle$  and  $\langle r, 1 \rangle$  multi-set  $\langle n, r \rangle$ ; and selection functions  $I_m^n$ . The significance of collection  $\sigma_M$  described above is in the truth of.

**Theorem (about multi-set PPA completeness).** Collection  $\sigma_M \equiv \{=, \cup_{All}, \oplus, \div, \{1^1\}, \emptyset_m, \varphi, I_m^n\}_{m=1, \dots, n}^{n=1, 2, 3, \dots}$  is complete system of PPA  $A_M^{up}$  [14, 15].

The choice of multi-sets caused by relative simplicity of injective mappings  $\varphi: Z^{(N, N)} \rightarrow M$  and  $\Phi: M \rightarrow Z^{(N, N)}$  nature and obvious recursiveness of sets  $\varphi(Z^{(N, N)})$  and  $\Phi(M)$ . These facts are creating reliable basis for solution completeness problem of PPA  $A_{Z^{(N, N)}}^{up}$ . Injective mapping of set of records to multi-sets  $\varphi: Z^{(V, W)} \rightarrow M$  is defined as

$$\varphi(K) = \begin{cases} \langle \langle v_1, d_1 + 1 \rangle, \langle v_2, d_2 + 1 \rangle, \dots, \langle v_m, d_m + 1 \rangle \rangle, & \text{if } K = \{(v_1, d_1), \dots, (v_m, d_m)\} \in Z^{(N, N)}, \\ \alpha^\emptyset, & \text{if } K = \emptyset. \end{cases}$$

Inverse mapping  $\Phi: M \rightarrow Z^{(N, N)}$  is defined analogically

$$\Phi(\delta) = \begin{cases} \{(a_1, d_1 - 1), (a_2, d_2 - 1), \dots, (a_m, d_m - 1)\}, & \text{if } \delta = \{ \langle a_1, d_1 \rangle, \dots, \langle a_m, d_m \rangle \} \in M, \\ \emptyset, & \text{if } \delta = \emptyset. \end{cases}$$

**Lemma 5.**  $M$ -image of  $Z^{(N,N)}$ -pr-function ( $Z^{(N,N)}$ -pr-predicate) is  $\varphi(Z^{(N,N)})$ -pr-function ( $\varphi(Z^{(N,N)})$ -pr-predicate).

From the lemma 5 it can be concluded that any  $\varphi(Z^{(N,N)})$ -pr-function is  $M$ -pr-function. Analogical conclusion may be done for  $\varphi(Z^{(N,N)})$ -pr-predicates. Thus consequence 1 is true.

**Consequence 1.**  $M$ -image of  $Z^{(N,N)}$ -pr-function ( $Z^{(N,N)}$ -pr-predicate) is  $M$ -function ( $M$ -predicate).

Consider following  $Z^{(N,N)}$ -pr-functions and  $Z^{(N,N)}$ -pr-predicates with simple, but representative examples for some of them. Beforehand let us define auxiliary parametric operation of record projection  $pr_{\{v_{i_1}, \dots, v_{i_k}\}}$ , which maps any record  $I \in Z^{(N,N)}$  to new record  $pr_{\{v_{i_1}, \dots, v_{i_k}\}}(I) \equiv I \cap (\{v_{i_1}, \dots, v_{i_k}\} \times N)$ . So, *predicate of equality*  $=_Z$  is analogical to predicate of equality for multi-sets; *deletion by example*  $\div^Z: I \div^Z J \equiv pr_{pr_1(I) \setminus pr_1(J)}(I)$ , particularly if  $pr_1(I) \cap pr_1(J) = \emptyset$ , then  $I \div^Z J = I$ , for example:  $\{(1,3), (2,10), (5,7)\} \div^Z \{(1,1), (2,5), (3,7)\} = \{(5,7)\}$ ;  $\{(5,7)\} \div^Z \{(6,5), (3,7)\} = \{(5,7)\}$  and  $\{(6,5), (3,7)\} \div^Z \emptyset = \{(6,5), (3,7)\}$ ; *records overlapping*  $\nabla$ : for any  $I, J \in Z^{(N,N)}$   $I \nabla J \equiv J \cup pr_{pr_1(I) \setminus pr_1(J)}(I)$ , particularly,  $I \nabla I^\emptyset = I$ ;  $I^\emptyset \nabla J = I^\emptyset$ , and in case if  $pr_1(I) \cap pr_1(J) = \emptyset$   $I \nabla J = J \nabla I = J \cup I$ , for example,  $\{(1,3), (5,7)\} \nabla \{(1,1), (2,5), (3,7)\} = \{(1,1), (2,5), (3,7), (5,7)\}$ ; *append to record*

$$U^+ : U^+(I) \equiv \begin{cases} I \cup \{(\max(pr_1(I)) + 1, 0)\}, & \text{if } I \neq I^\emptyset, \\ \{(0,0)\}, & \text{if } I = I^\emptyset, \end{cases} \text{ for any } I \in Z^{(N,N)}.$$

For example, for  $I = \{(1,3), (2,10)\}$  and  $J = I^\emptyset$ , function will result  $U^+(I) = \{(1,3), (2,10), (3,0)\}$  and  $U^+(J) = \{(0,0)\}$  accordingly; *selection by maximal name*  $\max$ :  $\max(I) \equiv pr_{\max(pr_1(I))}(I)$ ,  $\max(I^\emptyset) = I^\emptyset$ ; *zeroing of values*  $\{0\} : \{0\}(I) \equiv J$ , where  $pr_1(J) = pr_1(I) \& pr_2(J) = \{0\}$ . For example,  $\{0\}(\{(1,3), (2,10)\}) = \{(1,0), (2,0)\}$ ; *increment*  $\uparrow$ : maps any non-empty record  $I \in Z^{(N,N)}$  to record  $\uparrow(I)$ ,  $\uparrow(I) \equiv \{(v, a + 1) \mid \forall (v, a) \in I\}$ . *decrement*  $\downarrow$ : maps any non-empty record  $I \in Z^{(N,N)}$  to record  $\downarrow(I)$ , which  $\downarrow(I) = \left\{ (v, b) \mid v : \exists (v, a) \in I \& b : b = \begin{cases} a - 1, & a > 0, \\ 0, & a = 0 \end{cases} \right\}$ . In case if  $I = I^\emptyset$ ,  $\uparrow(I^\emptyset) = \downarrow(I^\emptyset) = I^\emptyset$ .

Designate set of  $Z^{(N,N)}$ -pr-functions and  $Z^{(N,N)}$ -pr-predicates described as

$$\sigma_{Z^{(N,N)}} = \left\{ =_Z, \nabla, \div^Z, U^+, \max, \{0\}, \uparrow, \downarrow, I_m^n \right\}, \quad n=1,2,\dots, m=1,\dots,n.$$

Analogically to previous section, consider  $Z^{(N,N)}$ -functions  $\psi$  and  $\chi$ — coding and decoding functions, accordingly.  $\psi \equiv \varphi \cdot \Phi$  and  $\chi$  is certain extension of mapping  $\psi^{-1}$ .

Therefore lemma 6 takes place.

**Lemma 6.**  $Z^{(N,N)}$ -model of  $M$ -function ( $M$ -predicate) from set  $\sigma_M$  may be created from functions of  $\sigma_{Z^{(N,N)}}$  set with PPA operations.

It is easy to build  $Z^{(N,N)}$ -models for  $M$ -pr-predicate of equality and  $M$ -pr-functions  $\cup_{All}, \div, \{1^1\}, \emptyset_m, \varphi$ . Let us build model of function for multi-sets addition  $\oplus$ . Now, introduce few auxiliary  $Z^{(N,N)}$ -functions and  $Z^{(N,N)}$ -predicates. Namely *identically false* and *identically true* predicates *Fal* and *Tru*:

$Fal \equiv S^3(=, I_1, S^2(U^+, I_1))$  and  $Tru \equiv S^3(=, I_1, I_1)$ ; *predicate of inequality*  $Neq$ :  $Neq \equiv \diamond(S^3(=, I_1, I_2), Fal, Tru)$ ; *constant empty record*  $\emptyset^Z$ :  $\emptyset^Z \equiv S(\div^Z, I_1, I_1)$ ; *selection by pattern*  $Sel$ :  $Sel(I_1, I_2) \equiv S^3(\div^Z, I_1, S^3(\div^Z, I_1, I_2))$ ; for example,  $Sel(\{(1,1), (2,2)\}, \{(2,3), (4,5)\}) = \{(2,2)\}$  — function «selects» from record  $I_1$  those components, names of which are in record  $I_2$ , i.e.  $I_2$  is a kind of pattern for selection from  $I_1$ ; *maximal addition*  $+^{\max}$  of pair of records with same schemes:  $+^{\max} \equiv *^3 S^3(Neq(I_2^2, S^2(\{0\}, I_2^2)), S^2(\uparrow, I_1^2), S^2(\downarrow, I_2^2))$ . For example, for records  $I_1 = \{(1,1), (2,2)\}$  and  $I_2 = \{(1,3), (2,5)\}$  we will get  $+^{\max}(I_1, I_2) = \{(1,1+5), (2,2+5)\} = \{(1,6), (2,7)\}$ . Note, that operation of maximal addition in general case is non-commutative, i.e.  $+^{\max}(I_1, I_2) \neq +^{\max}(I_2, I_1)$ . Commutative property preserved only for records of special type, for example, for same-scheme single-element records. For instance, if  $I_1 = \{(2,2)\}$  and  $I_2 = \{(2,5)\}$ , then  $+^{\max}(I_1, I_2) = +^{\max}(I_2, I_1) = \{(2,7)\}$ .

Now we can get down directly to building of  $Z^{(N,N)}$ -pr-function  $\oplus^Z$  — model of  $M$ -function  $\oplus$ . Assume that records  $I_1 = \{(v_1^1, d_1^1), \dots, (v_{n_1}^1, d_{n_1}^1)\}$  and  $I_2 = \{(v_1^2, d_1^2), \dots, (v_{n_2}^2, d_{n_2}^2)\}$  are given and  $sh(I_1) \cap sh(I_2) = \{v_{r_1}, \dots, v_{r_s}\}$ .  $\oplus^Z$  operation «breaks» records  $I_1$  and  $I_2$  to «segments», designated as  $I_{1_1}, I_{1_2}, I_{2_1}$  and  $I_{2_2}$  with schemes  $sh(I_{1_1}) = sh(I_1) \setminus \{v_{r_1}, \dots, v_{r_s}\}$ ,  $sh(I_{1_2}) = \{v_{r_1}, \dots, v_{r_s}\}$ ,  $sh(I_{2_1}) = sh(I_2) \setminus \{v_{r_1}, \dots, v_{r_s}\}$  and  $sh(I_{2_2}) = \{v_{r_1}, \dots, v_{r_s}\}$ . Thus, resulting record may be represented as:  $I_3 = I_{1_1} \cup I_{2_1} \cup I_{1_2} \oplus^Z I_{2_2}$ . Note,  $sh(I_{1_2}) = sh(I_{2_2})$ . As for first two items  $I_{1_1}$  and  $I_{2_1}$ , they are easily created with earlier defined function  $\div^Z$ . Namely,  $I_{1_1} = I_1 \div I_2$  and  $I_{2_1} = I_2 \div I_1$ . As for  $I_{1_2} \oplus^Z I_{2_2}$ ,  $I_{1_2}$  and

$I_{2_2}$  are easily defined with usage of function  $Sel: I_{1_2} = Sel(I_1, I_2)$  and  $I_{2_2} = Sel(I_2, I_1)$ . Now model  $Z^{(N,N)}$ -function  $\oplus^Z$  for same-scheme records  $I_{1_2}$  and  $I_{2_2}$ . It is easy to convince that  $\oplus^Z$  may be represented as:

$$\oplus^Z = *^4 \underbrace{(Neq(I_1^2, I_2^2))}_p, \underbrace{I_1^2 \nabla (\max(I_2^2) + \max(Sel(I_1^2, \max(I_2^2))))}_{f_1}, \underbrace{I_2^2 \div^Z \max(I_2^2)}_{f_2}.$$

Obviously, in case when  $sh(I_1) \cap sh(I_2) = \emptyset$ , record  $I_{1_2} \oplus^Z I_{2_2}$  is empty too, i.e.  $I_{1_2} \oplus^Z I_{2_2} = \emptyset$  and, consequently, the result is  $I_3 = I_1 \oplus^Z I_2 = I_{1_1} \cup I_{2_1}$ . Taking in account that  $sh(I_{1_1}) \cap sh(I_{2_1}) = \emptyset$ , the result may be expressed as  $I_3 = I_1 \oplus^Z I_2 = I_1 \nabla I_2 = I_{1_1} \nabla I_{2_1}$ .

**Lemma 7.** Functions  $\psi$  and  $\chi$  may be built from functions of set  $\sigma_{Z^{(N,N)}}$  by finite number of applications of PPA.

Correctness of the result is obvious, because of noted similarity of records and multi-sets, simplicity of coding and decoding mappings ( $\varphi$  and  $\Phi$ ), and adduced earlier statements. Hence, lemma 8 is correct.

**Lemma 8.**  $\langle M, Z^{(N,N)}, \sigma_M, \sigma_{Z^{(N,N)}}, \psi, \chi \rangle$  — AS with basis  $\sigma_M$ .

So, theorem 4 is true.

**Theorem 4.**  $\sigma_{Z^{(N,N)}} = \left\{ =_Z, \nabla, \div^Z, U^+, \max, \{0\}, \uparrow, \downarrow, I_m^n \right\}_{m=1, \dots, n}^{n=1, 2, \dots}$  — genera-

tive system of PPA  $A_{Z^{(N,N)}}^{up}$ .

Considering given above statements 1, 2 certain conclusions respectively to possible reducibility of  $\sigma_{Z^{(N,N)}}$  may be made. Equality predicate cannot be ex-

cluded from  $\sigma_{Z^{(N,N)}}$  because it is sole predicate in CS.  $\div^Z (U(I), \nabla, \max, \{0\})$  — the only function in CS that does not preserve set  $Z^{(N,N)} \setminus \{I^\emptyset\}$  ( $Z^{(N,N)} \setminus \{(0,0)\}$ ),  $\bigcup_{i \in N} Z[\{i\}]$ ,  $\bigcup_{i \in N} Z[\{i, i+1\}]$ ,  $\bigcup_{i \in N} Z[\{i\}] \setminus \{(i,0)\}$ ). Moreover,

$U(I)$  does not  $\beta$ -preserve denotations with given such estimation  $\beta: Z^{(N,N)} \rightarrow 2^N$  that  $\beta(I^\emptyset) = \emptyset$  and  $\beta(\{(v_1, d_1), \dots, (v_n, d_n)\}) = \{v_1, \dots, v_n, d_1, \dots, d_n\}$ ,  $n=1, 2, 3, \dots$ . As for increment  $\uparrow$  and decrement  $\downarrow$  functions, they are in  $\sigma_{Z^{(N,N)}}$  simultaneously for convenience and symmetry, however they are not independent. For example, decrement may be easily produced by PPA operations from rest of the functions and predicate of  $\sigma_{Z^{(N,N)}}$  ( $S^4(*^4(S^3(Neq, I_2^3, I_3^3), S(\uparrow, I_1^3), (\uparrow, I_2^3), I_3^3), \{0\}, S(\uparrow, I_1^1), I_1^1))$ ). The fact that  $\uparrow$  as well as  $\downarrow$  does not preserve denotations with given estimation  $\beta$ : for which  $\beta(I^\emptyset) = \emptyset$  and  $\beta(\{(v_1, d_1), \dots, (v_n, d_n)\}) = \{d_1, \dots, d_n\}$ ,  $n=1, 2, 3, \dots$  directly results truthfulness of theorem 5.

**Theorem 5.**  $\sigma_{Z^{(N,N)}}^{I_m^n} \equiv \left\{ =_Z, \nabla, \dot{\div}^Z, U^+, \max, \{0\}, \uparrow, I_m^n \right\}_{m=1,2,\dots,n}^{n=1,2,\dots} — I_m^n$ -basis of

PPA  $A_{Z^{(N,N)}}^{Pr}$ .

## CONCLUSIONS

Modern IT problematic needs direct consideration of not only programming problems solutions, but processes, which lead to them. That is why researches of such processes organization structures are of paramount importance today. A special place in those researches takes problematic, connected with building of algebraic characteristics of pragmatically conditioned function classes, particularly, with solutions of completeness problems in corresponding algebras. In the paper these questions discussed on basis of primitive program algebras. Method of generative sets finding in PPA presented here, and applied to research of class of partially-recursive functions on records, which is of theoretical and applied importance. Using concepts of complete and allowable systems, and results received, especially criteria of completeness, universality of proposed method in classes of computable functions on different carriers proved.

Results received form foundations for development of adaptive programming environments. Next steps in this direction will be related with investigation of general concept of composition and development of functions exploring reduction methods connected with it as means of pragmatically driven decomposition of programming problems.

## REFERENCES

1. *Dijkstra E.* A Discipline of Programming. — Prentice Hall, Inc., 1978. — 275 p.
2. *Brooks F.P.* The Design of Design: Essays from a Computer Scientist. — Addison-Wesley, 2010. — 448 p.
3. *Brooks F.P.* The Mythical Man-Month: Essays on Software Engineering. — Addison-Wesley, 1995. — 304 c.
4. *Redko V.N.* Fundamentals of compositional programming // Cybernetics and System Analysis. — 1979. — № 3. — P. 3–13.
5. *Redko V.N.* Semantical structures of software // Cybernetics and System Analysis. — 1981. — № 1. — P. 3–19.
6. *Redko V.N.* Universal program logics and their application // Proc. of 4<sup>th</sup> soviet-wide symp. — Kishenev, 1983. — P. 310–326.
7. *Basarab I.A., Nikitchenko N.S., Redko V.N.* Compositional databases. — K.: Lybid, 1992. — 92 p.
8. *Redko I.V., Redko V.N.* Existential basis of compositional paradigm // Programming and Computer Software. — 2008. — № 2. — P. 3–12.
9. *Bui D.B., Redko V.N.* Primitive program algebras I, II // Cybernetics. — 1985. — № 1. — P. 28–33.
10. *Bui D.B., Redko I.V.* Primitive program algebras of computable functions // Cybernetics. — 1987. — № 3. — P. 68–74.
11. *Bui D.B., Redko I.V.* Primitive program algebras of functions, which preserve denotates // Report of Ukrainian AS. — 1988. — № 9. — P. 66–68.
12. *Yershov U.L.* Theory of numerations. — M.: Nauka, 1977. — 416 p.

13. *Redko I.V., Snigur N.M.* Primitive program algebra of computable functions on graph // *Naukovi Visti NTUU "KPI"*. — 2011. — № 4. — P. 75–80.
14. *Bogatyryova Y.O.* Computability on finite sets and multi-sets // *Taras Shevchenko Kiev Nation University bulletin. Ser.: phys.-math. sciences.* — 2010. — № 4. — P. 88–96.
15. *Bogatyryova Y.O.* Concept of multi-set. Structure of multi-sets family // *Academitian M. Kravtchuk 13<sup>th</sup> international scientific conference, proceedings of (Kyiv, may 13–15, 2010).* — K.: NTUU "KPI", 2009. — 60 p.
16. *Maltsev A.I.* *Algorhythmic systems.* — M.: Nauka. — 1970. — 392 p.
17. *Maltsev A.I.* *Constructive algebras. 1* // *Uspekhi matematicheskikh nauk.* — 1961. — **6.** — № 3. — P. 3–60.
18. *Maltsev A.I.* *Algorithms and recursive functions* // *Groningen: Wolters-Noordhoff, 1970.* — 391 p.
19. *Cutland N.* *Computability. An introduction in recursive function.* — M.: Mir, 1983. — 256 p.
20. *Yershov A.P.* *Computability in arbitrary fields and bases* // *Semiotics and Informatics.* — 1982. — № 19. — P. 3–58.
21. *Maltsev A.I.* *Selected works* // *Mathematical logic and general theory of algebraic systems.* — 1976. — **2.** — M.: Nauka, 1976. — 388 p.

*Received 13.05.2015*

---

From the Editorial Board: the article corresponds completely to submitted manuscript.



## АНАЛІТИЧНИЙ РОЗВ'ЯЗОК НЕКОРЕКТНИХ ЗАДАЧ ДИНАМІЧНИМИ МЕТОДАМИ

Д.М. ПАРХОМЧУК, Ю.О. ТИМОШЕНКО

Апроксимація неперервних диференціальних та інтегральних рівнянь скінченними дискретними алгебраїчними системами, локальна лінеаризація систем нелінійних рівнянь за заданою інформацією у разі вирішення обернених задач зводиться до задач розв'язку систем лінійних алгебраїчних рівнянь. Матриці таких систем зазвичай є погано обумовленими, тому задачі розв'язання таких систем є некоректними, оскільки порушується третя умова коректності за Адамаром. Для розв'язання некоректних задач запропоновано динамічний метод регуляризації [1]. З метою зменшення часу роботи алгоритму, що пропонується динамічним методом запропоновано модифікований метод — динамічний метод другого порядку. Розроблено математичний апарат та на його основі запропоновано алгоритм для модифікованого методу, а також показано його ефективність на практичному прикладі.

### ВСТУП

Ж. Адамар на початку ХХ століття ввів наступне поняття коректності задачі. Задача називається *коректною*, якщо виконано наступні вимоги:

- задача має містити розв'язок;
- вирішення задачі є єдиним;
- розв'язок задачі неперервно залежить від вхідних даних.

Якщо порушено хоча б одну з наведених вище вимог, то така задача називається некоректною.

Адамар навів наступний приклад некоректно поставленої задачі. Як відомо, диференціальне рівняння Лапласа описує фізичний процес — стаціонарний розподіл температури у просторовому тілі:

$$\begin{cases} u_{yy}(x, y) = -u_{xx}(x, y), & y > 0, \\ u|_{y=0} = 0, \quad u_y|_{y=0} = \frac{1}{k} \sin kx. \end{cases} \quad (1)$$

Нескладно показати, що розв'язком такого рівняння буде функція

$$u_k(x, y) = \frac{\operatorname{sh} ky}{k^2} \sin kx.$$

При  $k \rightarrow \infty$  видно, що  $\frac{1}{k} \sin kx \rightarrow 0$  по  $x$  (тут збіжність розуміється у рівномірному сенсі); тоді розв'язок також має прямувати до нуля. Однак, у загальному випадку, коли  $x \neq \pi n$ ,  $n = 0, \pm 1, \dots$ ,  $u_k(x, t) \rightarrow 0$ ,  $k \rightarrow \infty$ . Тому неперервної залежності від початкових умов немає, і відповідно, задачу (1) поставлено некоректно [1].

**Мета роботи** — модифікація динамічного методу регуляризації некоректних задач для поліпшення певних характеристик методу.

## ПОСТАНОВКА ЗАДАЧІ

Цілі цього дослідження можна умовно поділити на 2 частини: модифікація існуючого динамічного методу регуляризації та розробка відповідного чисельного методу, а також демонстрація роботи розробленого методу на практичному прикладі.

Надалі будемо розглядати систему лінійних рівнянь виду

$$Ax = b, A \in \mathbf{R}^{M \times N}, b \in \mathbf{R}^N, x \in \mathbf{R}^M, \quad (2)$$

де матриця  $A$  погано обумовлена, що унеможливлює застосування прямих методів [2].

Серед сучасних методів розв'язання задачі (1) виділяють методи *регуляризації* — додавання певної інформації в умову задачі так, щоб отримана задача стала коректною. Найвідомішими методами регуляризації є: регуляризація Тихонова та регуляризація Лаврентьєва.

В роботі [3] розглядається динамічний метод регуляризації некоректних задач, відповідно до якого за наближений розв'язок системи рівнянь (2) приймається розв'язок наступної задачі Коші:

$$\begin{cases} x' + Ax = b, \\ x(0) = x_0. \end{cases}$$

Загальну схему застосування цього методу можна описати таким чином: нехай відомо, що в задачі (2) матриця  $A$  відома з деякою точністю  $\alpha$ , а права частина — з точністю  $\beta$ . Тоді, виходячи з умови задачі, обирається деякий функціонал  $H(\alpha, \beta, x(\theta^*))$  і деякий рівень зупинки  $e_d$ .  $\theta^*$  обирається як перший момент часу, коли виконано умову  $H(\alpha, \beta, x(\theta^*)) \leq e_d$ , а у якості наближеного розв'язку (2) обирають  $x^* = x(\theta^*)$ .

У роботі досліджується наступна система диференціальних рівнянь другого порядку:

$$\begin{cases} \xi_1 x'' + \xi_2 x' + Ax = b, \\ x(0) = x_1, x'(0) = x_2, \end{cases} \quad (3)$$

де  $\xi_1$  та  $\xi_2$  — параметри.

## РОЗВ'ЯЗОК СИСТЕМИ

Наведемо основні кроки розв'язання системи (3).

Спочатку введемо наступні позначення.

Нехай  $a_i$  —  $i$ -й стовпчик матриці  $A^T A$ ,  $\delta(a_i, a_2, \dots, a_n)$  — визначник матриці, у якій  $(a_i, a_2, \dots, a_n)$  є стовпчиками,  $\lambda_i$  —  $i$ -е власне число матриці  $A^T A$ .

Застосуємо до (3) перетворення Лапласу й отримаємо:

$$\xi_1(p^2 X(p) - px_1 - x_2) + \xi_2(pX(p) - x_1) + AX(p) = b/p. \quad (4)$$

Вище отримано систему лінійних алгебраїчних рівнянь, розв'язком якої буде образ розв'язку системи (3). Систему (4) було розв'язано за методом Крамера.

Таким чином,

$$X_i(p) = \frac{\Delta_i(p)}{\Delta(p)}. \quad (5)$$

Користуючись означенням характеристичного поліному, можна стверджувати, що визначник системи (4) дорівнює

$$\begin{aligned} \Delta(p) = & \delta((\xi_1 p^2 + \xi_2 p)e_1 + a_1, (\xi_1 p^2 + \xi_2 p)e_2 + \\ & + a_2, \dots, (\xi_1 p^2 + \xi_2 p)e_n + a_n). \end{aligned} \quad (6)$$

Далі розпишемо  $\Delta_{\text{sys}i}(p)$ :

$$\begin{aligned} \Delta_i(p) = & \delta((\xi_1 p^2 + \xi_2 p)e_1 + a_1, (\xi_1 p^2 + \xi_2 p)e_2 + a_2, \dots \\ & \dots, (\xi_1 p^2 + \xi_2 p)e_{i-1} + a_{i-1}, b1/p + \xi_1 x_2 + \xi_2 x_1 + \\ & + \xi_1 p x_1, (\xi_1 p^2 + \xi_2 p)e_{i+1} + a_{i+1}, \dots, (\xi_1 p^2 + \xi_2 p)e_n + a_n). \end{aligned} \quad (7)$$

Введемо наступні позначення, щоб переписати вираз (7) у більш компактній формі:

$$\begin{aligned} D_1(p) = & \delta((\xi_1 p^2 + \xi_2 p)e_1 + a_1, (\xi_1 p^2 + \xi_2 p)e_2 + a_2, \dots, (\xi_1 p^2 + \xi_2 p)e_{i-1} + \\ & + a_{i-1}, b1/p, (\xi_1 p^2 + \xi_2 p)e_{i+1} + a_{i+1}, \dots, (\xi_1 p^2 + \xi_2 p)e_n + a_n); \end{aligned}$$

$$\begin{aligned} D_2(p) = & \delta((\xi_1 p^2 + \xi_2 p)e_1 + a_1, (\xi_1 p^2 + \xi_2 p)e_2 + \\ & + a_2, \dots, (\xi_1 p^2 + \xi_2 p)e_{i-1} + a_{i-1}, \xi_1 x_2 + \xi_2 x_1, (\xi_1 p^2 + \xi_2 p)e_{i+1} + \\ & + a_{i+1}, \dots, (\xi_1 p^2 + \xi_2 p)e_n + a_n); \end{aligned}$$

$$\begin{aligned} D_3(p) = & \delta((\xi_1 p^2 + \xi_2 p)e_1 + a_1, (\xi_1 p^2 + \xi_2 p)e_2 + \\ & + a_2, \dots, (\xi_1 p^2 + \xi_2 p)e_{i-1} + a_{i-1}, \xi_1 p x_1, (\xi_1 p^2 + \xi_2 p)e_{i+1} + \\ & + a_{i+1}, \dots, (\xi_1 p^2 + \xi_2 p)e_n + a_n). \end{aligned}$$

Тоді

$$\Delta_i(p) = D_1(p) + D_2(p) + D_3(p). \quad (8)$$

Підставляючи (8) у (5), отримаємо

$$x_i(p) = \frac{D_1(p)}{\Delta(p)} + \frac{D_2(p)}{\Delta(p)} + \frac{D_3(p)}{\Delta(p)}. \quad (9)$$

Для явного наведення розв'язку системи (3) залишилося знайти обернене перетворення Лапласу від кожного доданку (9) та просумувати. Наведемо остаточний результат таких перетворень. Для цього введемо наступні позначення:

$$\begin{aligned}
 c_i &= \frac{-\xi_2 + \sqrt{\xi_2^2 - 4\lambda_i \xi_1}}{2\xi_1}, \quad d_i = \frac{-\xi_2 - \sqrt{\xi_2^2 - 4\lambda_i \xi_1}}{2\xi_1}, \\
 \Delta_i^j &= \delta(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_{j-1}, e_j, a_{j+1}, \dots, a_n), \\
 \nabla_i^j &= \delta(a_1, \dots, a_{i-1}, \xi_1 x_2 + \xi_2 x_1, a_{i+1}, \dots, a_{j-1}, e_j, a_{j+1}, \dots, a_n), \\
 \aleph_i^j &= \delta(a_1, \dots, a_{i-1}, \xi_1 x_1, a_{i+1}, \dots, a_{j-1}, e_j, a_{j+1}, \dots, a_n), \\
 F_1(p) &= (\Delta_i + (\xi_1 p^2 + \xi_2 p) \sum_{j=1, j \neq i}^n \Delta_i^j + \\
 &+ (\xi_1 p^2 + \xi_2 p)^2 \sum_{1 < j_1 < j_2 < n}^{j_1, j_2 \neq i} \Delta_i^{j_1, j_2} + \dots + p(\xi_1 p^2 + \xi_2 p)^{n-1} b_i), \\
 F_2(p) &= p \prod_{i=1}^n (p - c_i)(p - d_i), \\
 G_1(p) &= (\nabla_i + (\xi_1 p^2 + \xi_2 p) \sum_{j=1, j \neq i}^n \nabla_i^j + (\xi_1 p^2 + \xi_2 p)^2 \sum_{1 < j_1 < j_2 < n}^{j_1, j_2 \neq i} \nabla_i^{j_1, j_2} + \dots \\
 &+ (\xi_1 p^2 + \xi_2 p)^{n-1} (\xi_1 x_2 + \xi_2 x_1)_i), \quad G_2(p) = \prod_{i=1}^n (p - c_i) * (p - d_i), \\
 H_1(p) &= (\aleph_i + (\xi_1 p^2 + \xi_2 p) \sum_{j=1, j \neq i}^n \aleph_i^j + \\
 &+ (\xi_1 p^2 + \xi_2 p)^2 \sum_{1 < j_1 < j_2 < n}^{j_1, j_2 \neq i} \aleph_i^{j_1, j_2} + \dots + (\xi_1 p^2 + \xi_2 p)^{n-1} (\xi_1 x_1)_i).
 \end{aligned}$$

Користуючись наведеними вище позначеннями, розв'язок системи (3) можна записати як

$$\begin{aligned}
 x_i(t) &= \frac{\Delta_i}{\Delta} + \sum_{j=1}^n \left( e^{c_j t} \left( \frac{F_1(c_j)}{F_2'(c_j)} + \frac{G_1(c_j) + H_1(c_j)}{G_2'(c_j)} \right) + \right. \\
 &\left. + e^{d_j t} \left( \frac{F_1(d_j)}{F_2'(d_j)} + \frac{G_1(d_j) + H_1(d_j) \frac{d_j}{c_j}}{G_2'(d_j)} \right) \right). \quad (10)
 \end{aligned}$$

### АНАЛІЗ ОТРИМАНОГО РОЗВ'ЯЗКУ

**Наслідок 1.** Для стійкості розв'язку (10) достатньо, щоб було виконано умову ( $\forall i \in [1, n]: c_i < 0, d_i < 0$ ), тобто

$$\forall i \in [1, n]: \frac{-\xi_2 + \sqrt{\xi_2^2 - 4\lambda_i \xi_1}}{2\xi_1} < 0, \quad \frac{-\xi_2 - \sqrt{\xi_2^2 - 4\lambda_i \xi_1}}{2\xi_1} < 0. \quad (11)$$

Визначимо, які умови необхідно накласти на  $\xi_1$  та  $\xi_2$ , щоб наведена вище умова виконувалась.

Розглянемо два випадки.

**Перший випадок.** Нехай  $\xi_1 > 0$ , тоді (11) еквівалентно

$$\forall i \in [1, n]: \sqrt{\xi_2^2 - 4\lambda_i \xi_1} < 2\xi_1 \xi_2,$$

звідки

$$\forall i \in [1, n]: -4\lambda_i \xi_1 < 0, \quad 0 < \xi_2.$$

Враховуючи той факт, що  $0 < \lambda_i$ , остаточно отримуємо, що  $\xi_1 > 0$ ,  $\xi_2 > 0$  і цієї умови достатньо, щоб рішення (10) було стійким.

**Другий випадок.** Нехай  $\xi_1 < 0$ , тоді (11) еквівалентно

$$-\xi_2 - \sqrt{\xi_2^2 - 4\lambda_i \xi_1} > 0,$$

звідки  $\forall i \in [1, n]: -4\lambda_i \xi_1 < 0, \quad 0 < \xi_2$ . Як і у попередньому випадку, отримуємо  $\xi_1 > 0$ , але це суперечить початковому припущенню  $\xi_1 < 0$ . Таким чином достатньою умовою для стійкості рішення (10) є  $\xi_1 > 0, \quad \xi_2 > 0$ .

**Наслідок 2.** Дослідимо асимптотичні властивості розв'язку (10), якщо виконано умови, які наведено у попередньому наслідку. Справді, при  $t \rightarrow \infty$  виконано:  $e^{c_j t} \rightarrow 0, \quad e^{d_j t} \rightarrow 0$ , тому з виразу (8) отримуємо:

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{\Delta_i}{\Delta}.$$

Таким чином розв'язок системи (3) збігається до точного розв'язку системи (2).

**Наслідок 3.** Вище було отримано, що умови ( $\xi_1 > 0, \quad \xi_2 > 0$ ) достатньо для того, щоб розв'язок (3) був стійким. З метою підвищення швидкодії методу отримуємо умови, які необхідно накласти на  $\xi_1$  та  $\xi_2$ . Для цього необхідно, щоб коефіцієнти у показниках експонент рішення (10) відповідали умовам:

$$\frac{\xi_2 + \sqrt{\xi_2^2 - 4\lambda_i \xi_1}}{2\xi_1} > \lambda_i, \quad \frac{\xi_2 - \sqrt{\xi_2^2 - 4\lambda_i \xi_1}}{2\xi_1} > \lambda_i.$$

Дослідимо це. Для цього достатньо перевірити тільки другу умову. Отже:

$$\xi_2 - \sqrt{\xi_2^2 - 4\lambda_i \xi_1} > 2\xi_1 \lambda_i \quad \text{або} \quad \xi_2 - 2\xi_1 \lambda_i > \sqrt{\xi_2^2 - 4\lambda_i \xi_1}.$$

Звідки  $\xi_2^2 - 4\xi_1 \lambda_i + 4\xi_1^2 \lambda_i^2 > \xi_2^2 - 4\lambda_i \xi_1, \quad \xi_2 - 2\lambda_i \lambda_i > 0$ . Таким чином, остаточно отримуємо

$$\xi_2 > 2\lambda_i \xi_1. \quad (12)$$

Для її виконання необхідно знайти модуль найбільшого власного числа матриці А та прийняти, наприклад, що  $\xi_1 = 1$ , а  $\xi_2 = 2\lambda_{\max} + 1$ .

### АЛГОРИТМ РОЗВ'ЯЗАННЯ НЕКОРЕКТНИХ ЗАДАЧ ДИНАМІЧНИМ МЕТОДОМ ДРУГОГО ПОРЯДКУ

Спираючись на отриманий вище результат та використовуючи правило зупинки, яке було запропоновано у динамічному методі [3]. В роботі розгля-

дається наступний алгоритм динамічного методу другого порядку для розв'язання некоректних задач:

1. Звести обернену задачу до задачі вигляду (2).
2. Визначити  $\lambda_{\max}$  — найбільше власне число матриці  $A^T A$ .
3. Обрати значення  $\xi_1, \xi_2$  виходячи з умови  $\xi_2 > 2\xi_1 \lambda_{\max}$ .
4. Визначити вигляд функціоналу  $H(\alpha, \beta, x(\theta^*))$ .
5. Визначити рівень зупинки  $e_d$ .
6. Чисельно розв'язувати систему (3) до виконання умови  $H(\alpha, \beta, x(\theta^*)) \leq e_d$ .
7. За наближений розв'язок задачі (3) приймається  $x^* = x(\theta^*)$ .

### ПРАКТИЧНИЙ РЕЗУЛЬТАТ

Задля демонстрації роботи запропонованого методу було виконано наступний чисельний експеримент.

Розглянемо наступне рівняння:

$$\int_0^1 K(t, \tau) y(\tau) d\tau = f(t), \quad (13)$$

де  $K(t, \tau) = \frac{1}{\pi(1 + (t - \tau)^2)}$  (розподіл Коші з математичним сподіванням

у точці  $t$ ). Ядра такого типу зустрічаються у задачах гравіметрії [4]. Нехай  $y(\tau) = (1 - \tau^2)^2$ . Вважаємо, що праву частину рівняння (13) можна замірити лише з певною точністю  $\beta$ . Задача полягає у визначенні значень  $y$  на деякій дискретній сітці. Розв'яжемо її відповідно наведеного вище алгоритму.

Застосуємо наступну схему для лінеаризації задачі (13):

$$\sum_{j=1}^n K(x_i, x_j) y_j h = f(x_i).$$

Для остаточного формування системи рівнянь (2) введемо наступні позначення:

$$A = (a_{ij}), \quad a_{ij} = K(x_i, x_j) h, \quad b = (b_i), \quad b_i = f(x_i), \quad x = (x_i).$$

У якості кількості замірів правої частини та  $n$  візьмемо 50. Число обумовленості отриманої матриці  $\text{cond}(A) = 6,7665 \cdot 10^{18}$ , а отже задача є некоректною. Найбільше власне число матриці  $A^T A$   $\lambda_{\max} = 0,078402$ . Згідно з (12) припускаємо, що  $\xi_1 = 1$  та  $\xi_2 = 1$ . Враховуючи той факт, що матриця  $A$  в цій задачі задана точно, вважаймо, що  $e_d = 1$ . Тоді

$$H(\alpha, \beta, x(\theta^*)) = \frac{\|Ax(\theta^*) - b\|}{\beta}. \quad (14)$$

Таким чином, чисельно розв'язуючи систему (3), знаходиться значення  $\theta^*$  згідно з (14).

Для оцінки результатів роботи алгоритму було підраховано похибку знайденого результату роботи алгоритму:

$$\|y^* - y\| = \sqrt{\sum_{i=1}^{50} (y_i^* - y_i)^2},$$

а також похибку відносно правої частини  $f^* = Ay^*$ :

$$\|f^* - f\| = \sqrt{\sum_{i=1}^{50} (f_i^* - f_i)^2}.$$

Нижче наведено порівняльну таблицю результатів розв'язання задачі за регуляризацією за Тихоновим та динамічним методом першого (ДМ1) та другого (ДМ2) порядків.

**Таблиця.** Порівняння результатів

Похибка $y$			Похибка $f$		
Тихонов	ДМ1	ДМ2	Тихонов	ДМ1	ДМ2
0,0771	0,0764	0,0764	0,00052	0,00059	0,00059

Як бачимо, ДМ1 та ДМ2 надали кращий результат за шуканою функцією і дещо гірший результат для правої частини рівняння (13).

## ВИСНОВКИ

У роботі розглянуто модифікацію динамічного методу регуляризації шляхом введення в рівняння другої похідної. Знайдено аналітичний розв'язок такої системи диференціальних рівнянь другого порядку. Проведено аналіз на стійкість отриманих рішень. Розроблено правила з вибору параметрів динамічного методу другого порядку. Знайдено певні залежності між параметрами системи, які дозволяють збільшити швидкість методу. Запропоновано чисельний алгоритм розв'язку динамічного методу другого порядку. Наведено практичне застосування розробленого методу для вирішення інтегрального рівняння Фредгольму першого порядку, яке часто зустрічається у задачах гравіметрії. Виконано порівняння наближених рішень, що знайдені динамічними методами першого та другого порядку та регуляризацією за Тихоновим.

Практичне значення роботи полягає у можливості створення замкнених систем автоматичного керування для обернених задач об'єктів, які описуються неперервними диференціальними та інтегральними рівняннями.

## ЛІТЕРАТУРА

1. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. — М.: Наука, 1979. — 284 с.
2. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений: Пер. с англ. Икрамова Х.Д. — М.: Мир, 1980. — 277 с.
3. Гутенмахер Л.И., Тимошенко Ю.А., Тихончук С.Т. О динамическом методе решения некорректных задач // Докл. АН СССР. — 1977. — 237. — № 4. — С. 776–778.
4. Васильева А.Б., Тихонов А.Н. Интегральные уравнения. — М.: Физматлит, 2002. — 158 с.

Надійшла 08.07.2013

## ВДОСКОНАЛЕНІ МЕТОДИ РОЗРАХУНКУ СТАТИСТИКИ КОЛМОГОРОВА-СМИРНОВА, ВАГИ КАТЕГОРІЇ ЗМІННОЇ ТА ЗНАЧЕННЯ ІНФОРМАЦІЇ У КРЕДИТНОМУ РЕЙТИНГУ

О.М. СОЛОШЕНКО

Розроблено вдосконалені методи обчислення показників статистики Колмогорова-Смирнова, ваги категорії змінної та значення інформації без явного розбиття оригінальної вибірки на дві підмножини, з виведенням відповідних формул для аналізу предикативної (прогностичної) сили категоріальних змінних у задачах кредитного рейтингу та інших областях практичного застосування методів бінарної класифікації. Здійснено узагальнення класичних формул статистики Колмогорова-Смирнова, ваги категорії змінної та показника значення інформації шляхом перетворення агрегатних виразів для дискретних розподілів та кумулятивних функцій розподілу з застосуванням скалярного добутку векторів та операторів проектування, а також оператора умовної перестановки. Запропоновано вдосконалені формули обчислення статистики Колмогорова-Смирнова, ваги категорії змінної та індексу значення інформації, що узагальнено описуються в термінах дискретного безумовного розподілу вхідної змінної та умовного розподілу бінарної цільової змінної.

### ВСТУП

Практично у всіх системах та підходах побудови кредитних рейтингових моделей — скорингових моделей (скорингових карт) у задачах ризик-менеджменту щодо моделювання кредитних ризиків, для внутрішньої задачі аналізу прогностичної (предикативної) сили вхідних характеристик з метою оцінювання доцільності їх включення у модель логістичної регресії, використовуються значення  $WoE$  (Weight of Evidence — вага категорії змінної) для категорій категоріальної або дискретизованої змінної та показник  $IV$  (Information Value — значення інформації) на основі відстані Кульбака-Лейблера [1]. Одним з індикаторів оцінки якості ймовірнісних (або рейтингових) прогнозів на тестовій (валідаційній) вибірці з бінарною цільовою змінною є статистика  $KS$  (Kolmogorov-Smirnov statistic — статистика Колмогорова-Смирнова) [2], яка дозволяє оцінювати нерівність функцій розподілу для двох взаємовиключних класів. Статистику  $KS$  також застосовно на етапі аналізу характеристик на навчальній вибірці як альтернативу показнику  $IV$ . У такому разі, ця статистика буде в точності відповідати показнику якості прогнозів однофакторної моделі у термінах класичних показників якості прогнозів бінарного класифікатора. Вона відображає роздільну здатність класифікатора відносно двох підмножин, що відповідають двом значенням цільової змінної, тобто якість ранжування елементів всієї множини відносно цільової змінної [2]. Цей факт пояснюється тим, що у випадку використання класичних моделей зважування факторів типу логістичної регресії, що зберігають монотонність виходу моделі відносно єдиного входу, або у випадку використання дерев рішень, що збігаються з вхідною категоріальною змінною при використанні єдиного вхідного категоріального параметру, зберіга-



ється ранжування елементів вибірки [2]. Статистику KS застосовують як для дискретних (категоріальних), так і для неперервних розподілів.

Класичні формули обчислення *WoE* та *IV* оперують розбиттям на два окремі умовні розподіли категорій певної змінної на власне виділених окремо класах одиничних та нульових значень бінарної цільової змінної [1], а обчислення статистики Колмогорова-Смирнова передбачає побудову емпіричних функцій розподілу безпосередньо розглядаючи всю відому неагреговану множину елементів [2].

Актуальність дослідження полягає у практичній цінності наведення відповідних формул у термінах та поняттях безумовного дискретного розподілу (total distribution) змінної, що аналізується, та у термінах умовних ймовірностей нульових значень цільової змінної (bad rate) за кожною з категорій змінної, що аналізується, оскільки два наведені розподіли найбільш ілюстративні для відображення таблиць та графіків аналізу характеристик (зокрема, групи ризику фінального рейтингового балу) [2]. Також актуальність розроблення формул розрахунку ключових показників предикативності категоріальних змінних у кредитному скорингу (рейтингу) саме за допомогою різноманітних агрегатних показників обумовлена новітніми технологіями розробки баз даних. Вони не відповідають реляційній моделі та призначені для роботи з надзвичайно великими масивами даних [3]. Ще одним аспектом актуальності пропонованих методів є забезпечення можливості точної кількісної оцінки ключових індикаторів, використовуючи лише класичні таблицю та графік аналізу характеристик [2]. Також з використанням альтернативних формул можлива організація додаткової перевірки коректності розрахунку даних статистичних показників, відновлення індикаторів властивостей оригінальної вибірки. Альтернативні формули відображають важливу інтерпретацію числових значень ваги категорії змінної.

## **ПОСТАНОВКА ЗАДАЧІ**

Об'єктами дослідження є класичні формули KS, *WoE* та *IV*.

Предметом дослідження є методи перетворення агрегатних виразів для дискретних розподілів та кумулятивних функцій з застосуванням скалярного добутку та операторів проектування, а також оператора умовної перестановки.

**Мета роботи** — наведення вдосконалених методів обчислення показників статистики Колмогорова-Смирнова, ваги категорії змінної та значення інформації без явного розбиття оригінальної вибірки на дві підмножини, з виведенням відповідних формул для аналізу предикативної (прогностичної) сили категоріальних змінних у задачах кредитного скорингу та інших областях практичного застосування методів бінарної класифікації. Тобто необхідно розробити вдосконалені методи розрахунку ключових показників предикативної сили довільної категоріальної змінної у кредитному скорингу за відомих вхідних ймовірностях безумовного дискретного розподілу категоріальної змінної та умовних ймовірностях частоти нульових значень цільової змінної, тобто за узгодження вхідних векторів ймовірностей розподілу вхідної змінної та умовних ймовірностей цільової змінної.

### КЛАСИЧНІ МЕТОДИ ОЦІНЮВАННЯ KS, WOE ТА IV

Класичні формули для обчислення показників IV та  $WoE$  за відомих категорій та відомих значеннях цільової змінної кожного елемента множини вибірки для певної дискретної або дискретизованої вхідної змінної (категоріальної), щоб оцінити предикативну (прогностичну) силу вхідної характеристики, мають такий вигляд [1, 4]:

$$WoE_i = \ln\left(\frac{g_i}{b_i}\right),$$

$$IV = \sum_{i=1}^c (g_i - b_i) \ln\left(\frac{g_i}{b_i}\right) = \sum_{i=1}^c (g_i - b_i) WoE_i.$$

Категоріальний показник  $g_i$  — це відносна кількість елементів з одиничним («good») бінарним цільовим результатом у сегменті категорії до загальної кількості елементів з одиничним цільовим результатом всіх категорій:

$$g_i = \frac{G_i}{\sum_{i=1}^c G_i} = \frac{G_i}{G}.$$

Тобто оперуємо розподілом елементів з одиничним цільовим результатом за дискретними або дискретизованими значеннями змінної (категоріями), тому має місце рівність:

$$\sum_{i=1}^c g_i = 1.$$

Аналогічно, категоріальний показник  $b_i$  — це відносна кількість елементів з нульовим («bad») цільовим результатом у сегменті категорії до загальної кількості елементів з нульовим цільовим результатом всіх категорій:

$$b_i = \frac{B_i}{\sum_{i=1}^c B_i} = \frac{B_i}{B}.$$

Тобто також оперуємо розподілом елементів з нульовим цільовим результатами за дискретними або дискретизованими значеннями змінної (категоріями), тому має місце рівність:

$$\sum_{i=1}^c b_i = 1.$$

Взаємозв'язок значення інформації з відстанню Кульбака-Лейблера у теорії інформації [5] описується рівністю значення інформації сумі двох несиметричних відстаней Кульбака-Лейблера відносно кожного з розподілів [1, 5]:

$$IV = \sum_{i=1}^c g_i \ln\left(\frac{g_i}{b_i}\right) + \sum_{i=1}^c b_i \ln\left(\frac{b_i}{g_i}\right) = D_{KL}(\vec{g}, \vec{b}) + D_{KL}(\vec{b}, \vec{g}).$$

Класична формула обчислення статистики Колмогорова-Смирнова має такий вигляд [2]:

$$KS = \max_{x \in X} |F_B(x) - F_G(x)|.$$

Основною модифікацією статистики Колмогорова-Смирнова, що використовується на практиці, є показник рівня статистичної значимості (*p-value*) для розподілу Колмогорова, що пов'язаний з поняттям броунівського мосту [6]. Значення рівня статистичної значимості (*p-value*) записується з використанням функції розподілу таким чином:

$$\begin{aligned} PV &= 1 - F(KS) = 1 - \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2 KS^2} = \\ &= 1 - \left( 1 + 2 \sum_{j=1}^{\infty} (-1)^j e^{-2j^2 KS^2} \right), \\ PV &= -2 \sum_{j=1}^{\infty} (-1)^j e^{-2j^2 KS^2}. \end{aligned}$$

Перевагами використання статистики Колмогорова-Смирнова та відповідного значення статистичної значимості є двостороння обмеженість (на відміну від значення інформації), наочність (оскільки статистика Колмогорова-Смирнова — це максимальна абсолютна різниця функцій розподілу на спільній області визначення), зв'язок з поняттям броунівського мосту [6]. Основним недоліком статистики Колмогорова-Смирнова є відображення різниці між розподілами за допомогою максимуму, а не інтегрального показника, прикладом якого може слугувати індекс Джині [4].

#### ВДОСКОНАЛЕНІ МЕТОДИ ОЦІНЮВАННЯ KS, WOE ТА IV

У наведеній задачі мають місце вхідні вже агреговані дані без наведення оригінальної множини вибірки — матриця  $M = (\bar{t} \quad \bar{p})$  розмірності  $c \times 2$ , перший стовпець якої відповідає безумовному розподілу категорій вхідної змінної (total distribution), а другий — умовним ймовірностям частот елементів з нульовими значеннями бінарної цільової змінної (bad rate).

Має місце така рівність:

$$\sum_{i=1}^c t_i = 1.$$

Відповідну ймовірнісну вхідну матрицю зручно представити у вигляді графіку аналізу вхідної характеристики відносно бінарної цільової змінної, де гістограмі відповідає безумовний розподіл категорій вхідної характеристики, а ламаній лінії — відсоток елементів з нульовим цільовим результатом (умовна ймовірність). Наведемо приклад розподілу клієнтів банку за інтервалами віку клієнта та відсоток випадків некредитоспроможності для кожної вікової категорії (рисунки).

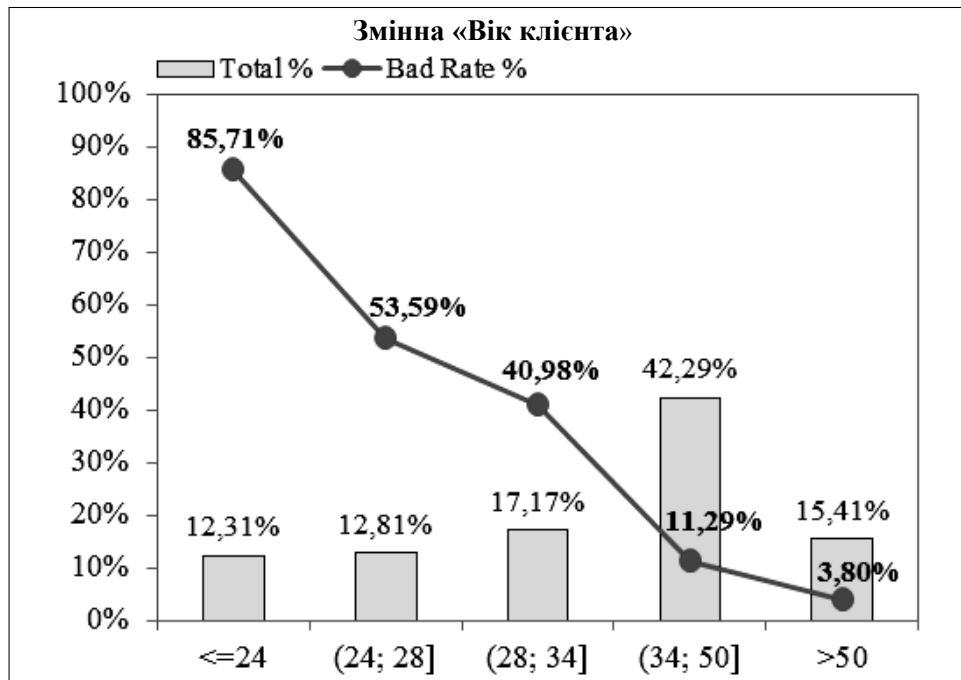


Рисунок. Графік аналізу характеристики «вік клієнта» з наведенням розподілу вибірки за категоріями та графік відсотку некредитоспроможних випадків для категорій

Щодо аналізу характеристик насамперед перепишемо формулу  $WoE_i$ , використовуючи «bad rate» на рівні категорії та середньозважений «bad rate»:

$$WoE_i = \ln\left(\frac{g_i}{b_i}\right) = \ln\left(\frac{\frac{G_i}{G}}{\frac{B_i}{B}}\right) = \ln\left(\frac{B}{G}\right) - \ln\left(\frac{B_i}{G_i}\right),$$

$$WoE_i = \ln\left(\frac{\frac{B}{B+G}}{1 - \frac{B}{B+G}}\right) - \ln\left(\frac{\frac{B_i}{B_i+G_i}}{1 - \frac{B_i}{B_i+G_i}}\right), \quad WoE_i = \ln\left(\frac{p_{w.avg}}{1 - p_{w.avg}}\right) - \ln\left(\frac{p_i}{1 - p_i}\right),$$

$$WoE_i = \ln\left(\frac{\frac{\sum_{j=1}^c p_j t_j}{\sum_{k=1}^c t_k}}{1 - \frac{\sum_{j=1}^c p_j t_j}{\sum_{k=1}^c t_k}}\right) - \ln\left(\frac{p_i}{1 - p_i}\right),$$

$$WoE_i = \ln\left(\frac{(\bar{p}, \bar{t})}{1 - (\bar{p}, \bar{t})}\right) - \ln\left(\frac{p_i}{1 - p_i}\right). \quad (1)$$

Враховано рівність:

$$\sum_{i=1}^c t_i = 1.$$

Середньозважений за допомогою скалярного добутку «bad rate» відповідає загальному «bad rate» на всій множині вибірки  $\frac{B}{B + G}$ .

Наведемо вдосконалену формулу IV:

$$IV = \sum_{i=1}^c (g_i - b_i) WoE_i = \sum_{i=1}^c \left( \frac{(1 - p_i)t_i}{\sum_{l=1}^c (1 - p_l)t_l} - \frac{p_i t_i}{\sum_{r=1}^c p_r t_r} \right) WoE_i.$$

Ще раз врахуємо рівність:

$$\sum_{i=1}^c t_i = 1.$$

Остаточна формула IV:

$$IV = \sum_{i=1}^c \left( \frac{(1 - p_i)t_i}{1 - (\bar{p}, \bar{t})} - \frac{p_i t_i}{(\bar{p}, \bar{t})} \right) \left( \ln\left(\frac{(\bar{p}, \bar{t})}{1 - (\bar{p}, \bar{t})}\right) - \ln\left(\frac{p_i}{1 - p_i}\right) \right). \quad (2).$$

Для прикладу з рисунку округлене значення IV буде дорівнювати 1,97.

Остаточні суть методів обчислення Weight of Evidence (1) та Information Value (2) полягає у використанні скалярного добутку вектору розподілу категорій змінної та відповідного вектору умовних ймовірностей, що відображає ймовірності набуття нульового значення для цільової змінної, а також у використанні інших перетворень від відповідних векторів агрегатних даних. Скалярний добуток відповідає середньому значенню ймовірності набуття нульового значення цільової бінарної змінної на всій вибірці.

Позначимо псевдопроектор з простору  $R^n$  на підпростір меншої розмірності  $R^m$ , що відповідає  $m$  першим координатам, як  $P_{n,m}$ . Суть псевдопроектора відображається такою формулою:

$$P_{n,m} : \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \\ \dots \\ x_n \end{pmatrix} \rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix}.$$

Цей псевдооператор проектування можна зобразити у вигляді матричного оператора (прямокутної матриці):

$$P_{n,m} : \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{pmatrix} \in \text{Mat}(m \times n).$$

Властивість елементів матриці:

$$p_{i,j} = \begin{cases} 1, & (i = j) \wedge (i \leq m); \\ 0, & \text{else.} \end{cases}$$

Введемо позначення оператора проектування перших  $m$  координат на підпростір розмірності  $m$  з довільного простору не меншої розмірності, ніж його образ, тобто довільної (без фіксації) розмірності області визначення  $n$ , де  $n \geq m$ , як  $P_m$ .

Основна властивість будь-якого проектора, що збігається з його означенням описується такою формулою (запишемо для  $P_m$ ) [7]:

$$P_m^2 = P_m.$$

Детальніше основну властивість проектора запропонованого типу можна описати таким чином:

$$\forall n \geq m \quad \forall \vec{x} \in R^n : P_m(P_m(\vec{x})) = P_{m,m}(P_{n,m}(\vec{x})) = P_{n,m}(\vec{x}) = P_m(\vec{x}) \in R^m.$$

Суть означення оператора проектування полягає у властивості ідемпотентності — точній рівності «проекції від проекції» власне значенню проекції [7, 8].

Також лінійний оператор проектування  $P$  можна означити як такий, що задається квадратною матрицею  $n \times n$ , тобто, коли розмірність образу збігається з розмірністю області визначення, а порядок (набір) координат, що проектуються, може бути довільним, при цьому проектування може відбуватись за допомогою лінійних комбінацій координат [7, 8]. Тоді основна властивість ( $P^2 = P$ ), що збігається з означенням оператора проектування, можлива, наприклад, завдяки таким умовам на елементи матриці такого оператора [7, 8]:

$$\begin{cases} i = j : p_{i,j} \in \{1; 0\}, \\ i \neq j : p_{i,j} = 0. \end{cases}$$

Також для означення лінійного оператора проектування за допомогою квадратної матриці можливе використання довільної ідемпотентної матриці [8].

Надалі будемо використовувати лише  $P_m$  — вищезначений оператор проектування перших  $m$  координат на підпростір розмірності  $m$  з простору довільної нефіксованої розмірності  $n \geq m$ . Також формули нижче будуть справедливими у випадку використання замість оператора  $P_m$  також звичайних квадратних діагональних матриць  $P_c^m$  розмірністю  $c \times c$ , що мають діагональ з першими  $m$  елементами рівними одиниці, а іншими діагональ-

ними елементами рівними нулю, при цьому виконується властивість (означення) проєктора:  $(P_c^m)^2 = P_c^m$ .

Введемо оператор ранжування (перестановки) одного вектора як перестановку його координат, що відповідає сортуванню другого вектора по спаданню координат  $R(\vec{x}, \vec{y}) : R^n \times R^n \rightarrow R^n$ .

Суть оператора сортування першого вектора відносно другого вектора по спаданню координат можна представити через функцію рангу  $r(i, \vec{x})$ , яку визначено на натуральних числах (але не більше розмірності власне вектора), що повертає початковий номер позиції координати ще не відсортованого вектора для заданої як аргумент координати вже відсортованого по спаданню вектора:

$$R(\vec{x}, \vec{y}) : \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \rightarrow \begin{pmatrix} x_{r(1, \vec{y})} \\ x_{r(2, \vec{y})} \\ \dots \\ x_{r(n, \vec{y})} \end{pmatrix}.$$

Суть запропонованого методу обчислення статистики Колмогорова-Смирнова полягає у використанні скалярного добутку, спеціального оператора сортування та операторів проєкування й описується таким чином:

$$KS = \max_{i=1, \dots, c} \left| \frac{(P_i(R(\vec{t}, \vec{p})), P_i(R(\vec{p}, \vec{p})))}{(\vec{t}, \vec{p})} - \frac{(P_i(R(\vec{t}, \vec{p})), P_i(R(\vec{e} - \vec{p}, \vec{p})))}{(\vec{t}, \vec{e} - \vec{p})} \right|. \quad (3).$$

Тут  $\vec{e}$  є одиничним вектором розмірності  $c$ .

Введемо позначення для композиції операторів перестановки та проєкування:

$$P_{\vec{p}, i}(\bullet) = P_i(R(\bullet, \vec{p})). \quad (4).$$

Враховуючи позначення (4), формулу (3) можна переписати таким чином:

$$KS = \max_{i=1, \dots, c} \left| \frac{(P_{\vec{p}, i}(\vec{t}), P_{\vec{p}, i}(\vec{p}))}{(\vec{t}, \vec{p})} - \frac{(P_{\vec{p}, i}(\vec{t}), P_{\vec{p}, i}(\vec{e} - \vec{p}))}{(\vec{t}, \vec{e} - \vec{p})} \right|. \quad (5).$$

Для прикладу з рисунку округлене значення  $KS$  для дискретизованої змінної «вік клієнта» буде дорівнювати 56,60%.

## ВИСНОВКИ

Запропоновано альтернативні методи обчислення та формули розрахунку статистики  $KS$ ,  $WoE$  та  $IV$  виходячи з відомого розподілу категорій та відомих умовних ймовірностей нульових значень цільової змінної, що дозволяють аналізувати характеристики та предикативну силу на навчальній та довільній вибірках, маючи лише відповідний графік агрегованих відносних значень. Предикативна сила змінної згідно з класичними методами скорингу [1, 2, 4] дорівнює якості прогнозів однофакторної моделі для категоріальної змінної, оскільки використання монотонних функцій типу логістичного пе-

ретворення або оптимального на навчальній вибірці дерева рішень, що точно відповідатиме власне категоріальній змінній, не змінює порядок категорій відносно умовного розподілу цільової змінної. При цьому присвоєний скоринговий бал відносно довільної шкали у випадку використання логістичної регресії буде монотонною функцією від умовного розподілу цільової змінної — долі нульових значень у категорії.

Основною відмінністю та практичною цінністю запропонованих формул відносно класичних є оперування лише агрегатними величинами без використання розбиття на дві окремі підмножини.

Ключовими особливостями запропонованих методів є використання скалярного добутку векторів з метою зважування величин, умовних перестановок та операторів проектування.

Перевагами запропонованих методів розрахунку статистичних показників прогностичної сили категоріальних змінних є:

- відсутність необхідності розбиття початкової вибірки на дві підмножини, що відповідають двом значенням бінарної змінної;
- використання безумовного дискретного розподілу категоріальної змінної та умовного розподілу бінарної цільової змінної, що відповідає класичним таблиці та графіку аналізу характеристик у кредитному скорингу [2];
- можливість швидкого розрахунку статистичних показників лише на основі наявних агрегованих даних класичного аналізу категоріальної змінної та можливість організації додаткової перевірки розрахунків згідно з класичними формулами;
- математична наочність запропонованих формул в поняттях скалярного добутку, операторів проектування та умовних перестановок;
- можливість точного відновлення значень, що описують детальні властивості оригінальної вибірки (наприклад,  $WoE$ ), за агрегатними ймовірнісними показниками аналізу характеристик, які явно не використовуються в класичних формулах розрахунку значень  $KS$ ,  $WoE$  та  $IV$ , але мають місце в запропонованих альтернативних формулах;
- зручність використання запропонованих методів у термінах агрегатів нереляційних систем керування базами даних [3], що дозволяє проводити аналіз характеристик та підрахунок показників, які розглядаються, одночасно, а не послідовно, із забезпеченням високої швидкодії на надзвичайно великих масивах даних (Big Data) [3].

Ще одним важливим висновком для ризик-менеджменту [1, 2, 4] є більш наочна інтерпретація показника  $WoE$  за допомогою запропонованої формули (1) як ступеню відхилення долі нульових значень цільової змінної по окремій категорії вхідної змінної відносно загальної (середньозваженої) долі нульових значень цільової змінної на всій вибірці. Згідно з інтерпретацією, негативне значення  $WoE$  означає перевищення відносно середнього значення на всій вибірці, а позитивне значення  $WoE$  означає, що значення долі нульових значень цільової змінної по даній категорії (bad rate) нижче, ніж на всій вибірці, нульове — точна рівність долі по категорії долі на всій вибірці.



Перспективи подальших досліджень містять вдосконалення методів обчислення інших показників предикативності (прогностичної сили) змінних, використовуючи лише агреговані показники значень ймовірностей (умовних та безумовних), а також застосування математичної методології кредитного скорингу поза межами управління ризиками.

#### **ЛІТЕРАТУРА**

1. *Siddiqi Naeem*. Credit risk scorecards: developing and implementing intelligent credit scoring. — Hoboken: John Wiley & Sons, Inc., 2006. — 196 p.
2. *Мэйз Элизабет*. Руководство по кредитному скорингу. — Минск: Гревцов Паблішер, 2008. — 464 с.
3. *Фаулер Мартин, Садаладж Дж. Прамодкумар*. NoSQL: новая методология разработки нереляционных баз данных. — Минск: ООО «И.Д. Вильямс», 2013. — 192 с.
4. *Thomas C. Lyn, Edelman B. David, Crook N. Jonathan*. Credit Scoring and its Applications. — Philadelphia: Society for Industrial and Applied Mathematics, 2002. — 248 p.
5. *Kullback Solomon*. Information Theory and Statistics. — Hoboken, NJ: John Wiley & Sons, 1959. — 395 p.
6. *Булінский А.В., Ширяев А.Н.* Теория случайных процессов. — М.: Физматлит, 2005. — 408 с.
7. *Треногин В.А.* Функциональный анализ. — М.: Наука, 1980. — 495 с.
8. *Мальцев А.И.* Основы линейной алгебры. — М.: Наука, 1975. — 400 с.

*Надійшла 22.09.2014*

УДК 519.946; 519.254

## ХАОТИЧЕСКАЯ И СЛУЧАЙНАЯ СОСТАВЛЯЮЩИЕ В ПРИРОДНЫХ ВРЕМЕННЫХ ДАННЫХ

В.Г. БОНДАРЕНКО, К.К. ТРУСКОВСКИЙ

Для временных данных предложен метод определения соотношения детерминированной и стохастической составляющих. Для решения данной задачи выполнен ряд вычислительных экспериментов, использующих имитационное моделирование логистической хаотической последовательности и значений фрактального броуновского движения с различными показателями Харста —  $H$ . В полученной аддитивной смеси задается соотношение энергий детерминированной и случайной составляющих. Для больших значений показателя Харста хаотическое слагаемое оказывается более агрессивным: контрольные статистики смеси значимо отличаются от эталонных значений, соответствующих фрактальному броуновскому движению. Для малых значений  $H$  (антиперсистентный случай) наблюдается обратный результат. Рассмотренные примеры реальных временных данных описываются антиперсистентной моделью.

### ВВЕДЕНИЕ

Анализ временных рядов — последовательных измерений  $x_1, \dots, x_n$  некоторой величины произвольной природы — состоит, прежде всего, в построении адекватной математической модели. Если эти данные априорно являются детерминированными, то есть члены последовательности  $\{x_k\}$  представляются значениями некоторой динамической системы, то модель определяется уравнением

$$x_k = f(x_{k-1}, \dots, x_{k-p}),$$

описывающим эту систему. Если же временной ряд имеет стохастическую природу, то его модель определяется равенством

$$x_k = X\left(\frac{k}{n}\right),$$

где  $X(t)$  — реализация некоторого случайного процесса  $\xi(t)$ ,  $t \in [0:1]$ .

Реальные временные данные, описывающие природные явления, как правило, содержат и детерминированную (хаотическую), и случайную составляющую. Хаотичность означает крайне нерегулярное поведение последовательности. В первом приближении хаотичность можно определить

свойством системы разводить первично близкие траектории в ограниченной области фазового пространства. В данной работе предложен алгоритм оценивания доли каждой из этих составляющих в наблюдаемых временных рядах.

## ПОСТАНОВКА ЗАДАЧИ

Предполагается, что наблюдаемые значения являются аддитивной смесью детерминированной хаотичной и случайной последовательностей:

$$x_k = \alpha u_k + \beta v_k,$$

где  $u_k$  — значения некоторой динамической системы,  $v_k$  — значения случайного процесса. Последовательности  $\{u_k\}$ ,  $\{v_k\}$  нормированы по энергии, то есть

$$\frac{1}{n} \sum u_k^2 = \frac{1}{n} \sum v_k^2 = 1.$$

Тогда величина  $a = \frac{\beta}{\alpha}$  определяет долю стохастичности в наблюдаемых данных.

В качестве модели для хаотической составляющей выберем логистическую динамическую систему:

$$u_k = 4u_{k-1}(1 - u_{k-1}), \quad (1)$$

решения которой образуют хаотическую последовательность.

Для модели стохастической составляющей положим

$$v_k = \sigma B_H\left(\frac{k}{n}\right), \quad (2)$$

где  $B_H(t)$  — фрактальное броуновское движение — гауссов случайный процесс, определенный соотношениями:

$$E B_H(t) = 0, \quad B_H(0) = 0, \quad E[B_H(t) B_H(s)] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}).$$

Свойства моделей, определенных равенствами (1) и (2), подробно изучены в работах [1] и [2]. Выбор именно этих моделей объясняется их универсальностью. Так, fBm (fractional Brownian motion — фрактальное броуновское движение) применимо к описанию широкого класса природных явлений [3].

**Цель работы** — определение соотношения  $a$  стохастической и детерминированной составляющих в реальных наблюдаемых данных  $x_1, \dots, x_n$  в предположении, что они образуют аддитивную смесь

$$x_k = u_k + a v_k, \quad (3)$$

где последовательности  $u_k$  и  $v_k$  определены соотношениями (1) и (2).

**ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ**

Приведем некоторые результаты, изложенные в работе [4] о статистической обработке временного ряда  $v_1, \dots, v_n$ ,  $\bar{v} = \frac{1}{n} \sum v_k = 0$ , описываемого моделью (2). Параметры модели  $\sigma$  (волатильность) и  $H$  (параметр Харста) предполагаются неизвестными. Алгоритм их оценивания состоит в следующем. Рассмотрим статистику

$$Q(H) = \frac{0,8}{R_1} \sqrt{\frac{(S_H^{-1}y, y)}{n}},$$

где вектор  $y = (y_1, \dots, y_n)$ ,  $y_k = v_k - v_{k-1}$ ,  $k = 1, \dots, n$ ,  $R_1 = \overline{|y|} = \frac{1}{n} \sum |y_k|$ ,  $S$  — корреляционная матрица вектора  $y$  с элементами

$$\rho(y_j, y_k) = \frac{1}{2} (|k - j + 1|^{2H} + |k - j - 1|^{2H} - 2|k - j|^{2H}).$$

Значения  $Q$  вычисляются перебором  $H$  (в матрице  $S_H^{-1}$ ) с некоторым шагом: то значение  $H$ , при котором  $|Q - 1| \rightarrow \min$  объявляется оценкой  $\hat{H}$  параметра Харста. Оценка волатильности определяется формулой

$$\hat{\sigma} = 1,25 R_1 n^{\hat{H}}.$$

Эффективность алгоритма оценивания параметра Харста проверена в [4]; значения fBm получены методом имитационного моделирования.

Для произвольного априорно стохастического временного ряда  $x_1, \dots, x_n$ , ( $\bar{x} = 0$ ) в работе [5] предложен алгоритм его аппроксимации fBm. Идея алгоритма — представление наблюдаемой траектории  $x(t)$  в виде

$$x(t) = \Phi(\sigma B_H(\cdot))(t), \quad x_k = x\left(\frac{k}{n}\right),$$

где  $\Phi$  — некоторое преобразование в пространстве траекторий. Реализация алгоритма базируется на значении статистики

$$d(y) = \frac{R_1^2}{R_2},$$

где  $R_2 = \overline{y^2} = \frac{1}{n} \sum y_k^2$ .

Если  $d(y)$  значительно отличается от  $\frac{2}{\pi} = 0,637$  (случай гауссовых приращений), перейдем к новой последовательности

$$z_k = \text{sgn } y_k |y_k|^{\frac{1}{\lambda}}, \quad \lambda > 0, \tag{4}$$

где  $\lambda$  определяется из уравнения  $d = \frac{1}{\sqrt{\pi}} \frac{\Gamma^2\left(\frac{\lambda+1}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right)}$ , тогда  $d(z) = \frac{2}{\pi}$ .

Таким образом, исходные данные приведены к виду

$$x_k = \sum_{j=1}^k \operatorname{sgn} y_j |z_j|^\lambda, \quad (5)$$

где  $z_1, \dots, z_n$  предполагаются приращениями fBm.

Последнее предположение представляет собой статистическую гипотезу  $T$  и нуждается в проверке. Если гипотеза  $T$  верна, то формула (4) есть представление исходного временного ряда функционалом от fBm (аппроксимация фрактальным броуновским движением). Алгоритм проверки гипотезы (то есть качества аппроксимации) базируется на предельных теоремах, доказанных в [6] для статистик

$$A_n = \frac{1}{n} \sum \xi_k z_k^3, \quad B_n = \frac{1}{n^{1+H}} \sum \xi_k^2 z_k^3, \quad D_n = \frac{1}{n^{2H}} \sum \xi_k z_k^3,$$

где  $\xi_k = \sum_{j=1}^{k-1} z_j$ .

Если гипотеза  $T$  —  $z_k$  являются приращениями fBm — верна, то имеет место сходимость ( $n \rightarrow \infty$ ):

$$A_n \rightarrow -1,5R_2^2 \equiv A, \quad B_n \rightarrow 3R_2^{5/2}\eta, \quad \text{если } H \in \left(0; \frac{1}{2}\right),$$

$$D_n \rightarrow 1,5R_2^2 B^2(1), \quad \text{если } H \in \left(\frac{1}{2}; 1\right),$$

где случайная величина  $\eta \sim N\left(0; \frac{1}{2H+2}\right)$ .

При проверке гипотезы  $T$  для статистик  $B_n$  и  $D_n$  задается уровень значимости  $\alpha = 0,1$  гипотеза принимается, если имеют место неравенства

$$\left|A_n + 1,5R_2^2\right| < 0,3; \quad |B_n| < \beta_1; \quad 0 < D_n < \beta_2, \quad (6)$$

где  $\beta_k$  — квантили предельных распределений, соответствующие

$$\alpha = 0,1, \quad \beta_1 = \frac{4,95R_2^{2,5}}{\sqrt{2H+2}}, \quad \beta_2 = 4,08R_2^2.$$

В работе [7] гипотеза  $T$  проверена для реальных временных данных стохастической природы.

В табл. 1 представлены значения оценки  $\hat{H}$  статистик  $A_n, B_n, D_n$  для генерируемого fBm, предельные значения  $A_n$  и значения квантилей.

**Таблица 1.** Сравнительные значения статистик для генерируемого fBm  $10 B_H(t)$ ,  $n = 2000$

H	0,1	0,2	0,3	0,4	0,6	0,7	0,8	0,9
$\hat{H}$	0,1	0,2	0,3	0,4	0,6	0,7	0,8	0,9
$A_n$	-794	-37,3	-1,18	-0,051	-	-	-	-
$A$	-789	-35	-1,52	-0,068	-	-	-	-
$B_n$	-1100	-94	0,33	-0,02	-	-	-	-
$\beta_1$	8407	160	3,46	0,622				
$D_n$	-	-	-	-	0,00034	0	0	0
$\beta_2$	-	-	-	-	0,044	0,010	0,0022	0,0004

Табл. 1 иллюстрирует «идеальный» случай, критерии согласия для гипотезы  $T$  выполнены для всех контрольных статистик. Данные таблицы используются для сравнения с аналогичными статистиками при анализе временных рядов, содержащих хаотическую составляющую.

**ОПРЕДЕЛЕНИЕ КАЧЕСТВА АППРОКСИМАЦИИ СМЕСИ**

Аппроксимация временных данных  $x_1, \dots, x_n$  по формулам (4), (5) признается удовлетворительной, если контрольные статистики удовлетворяют неравенствам (6). Применение алгоритма аппроксимации к логистической последовательности (1) дает следующие результаты (табл. 2).

**Таблица 2.** Результаты аппроксимации для логистической последовательности ( $n = 2000$ )

$x_0$	0,1	0,2	0,3	0,4	0,6	0,7	0,8	0,9
$\hat{H}$	0,25	0,15	0,25	0,15	0,15	0,4	0,15	0,25
$A_n$	1,20	0,62	1,38	0,60	0,62	-0,043	0,56	1,26
$B_n$	-3,65	-1,90	-4,51	-1,81	-1,90	-0,03	-1,63	-3,93
$A = \lim A_n$	-0,068	-0,077	-0,071	-0,080	-0,080	-0,085	-0,079	-0,068
$\beta_1$	0,065	0,080	0,070	0,082	0,082	0,082	0,082	0,065

Из табл. 2 следует, что оценка  $\hat{H}$  зависит от начального условия  $x_0$ , во всех случаях аппроксимация соответствует антиперсистентному процессу. Значения контрольных статистик далеки от предельных теоретических значений ( $A_n > 0$ ,  $|B_n| > \beta_1$ ), то есть аппроксимация неудовлетворительна. Исследуем качество аппроксимации для аддитивной смеси (3), увеличивая долю  $a$  стохастической составляющей.

Данные табл. 3 свидетельствуют об «агрессивности» хаотической составляющей по отношению к стохастической для  $H_{fBm} \geq 0,2$  (для  $a = 1; a = 2$ ) и для  $H_{fBm} \geq 0,3$  если  $a = 10$ . Для указанных значений  $H$  ге-

нерируемого  $f_{Bm}$  характер смеси определяет логистическая последовательность, то есть значения контрольных характеристик далеки от предельных (неравенства (6) не выполнены), что указывает на невозможность аппроксимации. Для  $H_{f_{Bm}} = 0,1$  наблюдается обратная картина — отклонение статистик от предельных значений такое же, как для «чистого»  $f_{Bm}$  (табл. 1), то есть смесь допускает качественную аппроксимацию при сопоставимых энергиях ее составляющих. Для  $H_{f_{Bm}} = 0,2$  удовлетворительная аппроксимация возможна лишь для  $a = 10$ .

Таким образом, персистентность ( $\hat{H} > 0,5$ ) исследуемого ряда при выполнении неравенства  $D_n < \beta_2$  означает его стохастическую природу: антиперсистентность ( $\hat{H} = 0,1 - 0,2$ ;  $A_n \approx A$ ,  $|B_n| < \beta_1$ ) допускает наличие хаотической компоненты.

**Таблица 3.** Результаты аппроксимации смеси ( $a = 1$ ;  $a = 2$ ;  $a = 10$ ;  $n = 2000$ )

$H_{f_{bm}}$		$\hat{H}$	$A_n$	$B_n$	$D_n$	$A = \lim A_n$	$\beta_1$	$\beta_2$
0,1	$a=1$	0,1	-2,78	-1,14	-1215	-2,95	7,77	5,72
	$a=2$	0,1	-19,4	-11,6	-8497	-19,4	82,0	14,7
	$a=10$	0,1	-8289	-21038	-3623570	-8239	157803	302
0,2	$a=1$	0,15	10,0	-83	2046	-0,84	1,57	3,05
	$a=2$	0,15	6,83	-39,3	1397	-1,26	2,6	3,7
	$a=10$	0,2	-48,8	-182	-4663	-45,5	227	22,5
0,3	$a=1$	0,15	11,4	-104	2330	-0,8	1,49	2,98
	$a=2$	0,15	10,7	-88,2	2194	-1,0	2,0	3,35
	$a=10$	0,25	-8,23	8,53	-368	-14,4	52,8	12,6
0,4	$a=1$	0,15	11,8	-113	2413	-0,75	1,37	2,88
	$a=2$	0,15	11,9	-111	2430	-0,81	1,52	3,0
	$a=10$	0,3	1,37	-4,11	28,6	-3,14	7,73	5,90
0,5	$a=1$	0,15	11,85	-118	2424	-0,72	1,31	2,84
	$a=2$	0,15	12,1	-123	2480	-0,74	1,31	2,90
	$a=10$	0,3	12,1	-39,0	253	-1,34	2,66	3,85
0,6	$a=1$	0,15	11,0	-101	2265	-0,72	1,31	2,83
	$a=2$	0,15	10,7	-93	2180	-0,72	1,31	2,83
	$a=10$	0,15	7,27	-38,0	1487	-0,78	1,43	2,93
0,7	$a=1$	0,15	11,9	-115	2432	-0,72	1,31	2,83
	$a=2$	0,15	12,3	-122	2517	-0,72	1,31	2,83
	$a=10$	0,2	15,6	-120	1487	-0,75	1,35	2,89
0,8	$a=1$	0,15	10,95	-101	2239	-0,72	1,31	2,83
	$a=2$	0,15	10,4	-92,5	2132	-0,72	1,31	2,83
	$a=10$	0,15	6,32	-38,0	1293	-0,73	1,32	2,84
0,9	$a=1$	0,15	11,0	-101	2253	-0,72	1,31	2,83
	$a=2$	0,15	10,6	-93,2	2158			
	$a=10$	0,15	6,82	-40,0	1395	-0,74	1,35	2,87

## ПРИМЕРЫ АППРОКСИМАЦИИ РЕАЛЬНЫХ ВРЕМЕННЫХ ДАННЫХ

В качестве примеров рассмотрим следующие реальные данные.

1. Концентрация планктона (мг/л) и хлорофилла (мкг/л) в озере Нарочь — всего 132 наблюдений, полученных усреднением многолетних измерений [8].

2. Скорость ветра в штате Аризона за 01.01.2001–01.12.2013 всего 156 наблюдений [9].

3. Осцилляция волн в Северной Атлантике — 10.1980–10.2014 всего 409 наблюдений [10].

Графическое изображение данных приведено на рис. 1, 2, 3, 4.

Применим к этим данным метод аппроксимации согласно описанному выше алгоритму. Значения оценок  $d$ ,  $\lambda$ ,  $\hat{H}$ , контрольных статистик  $A_n$ ,  $B_n$  и параметров  $A$  и  $\beta_2$  представлены в табл. 4.

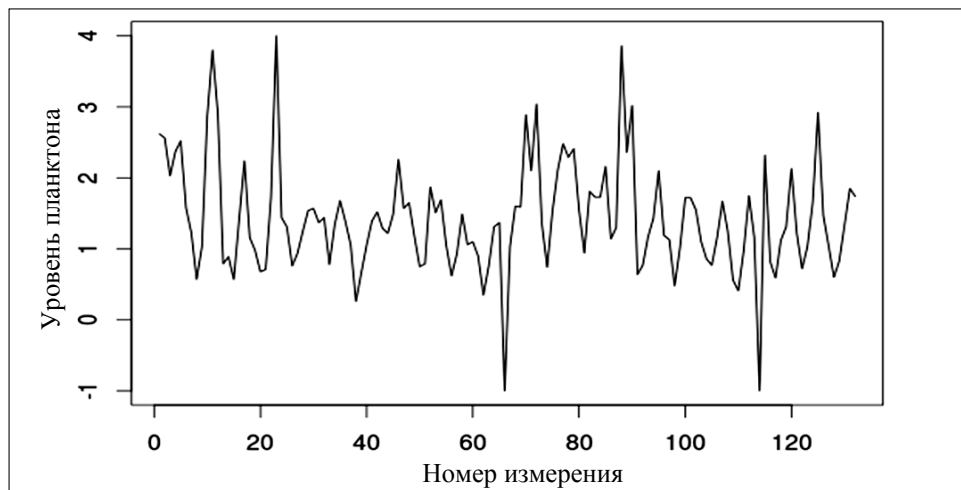


Рис. 1. Концентрация планктона

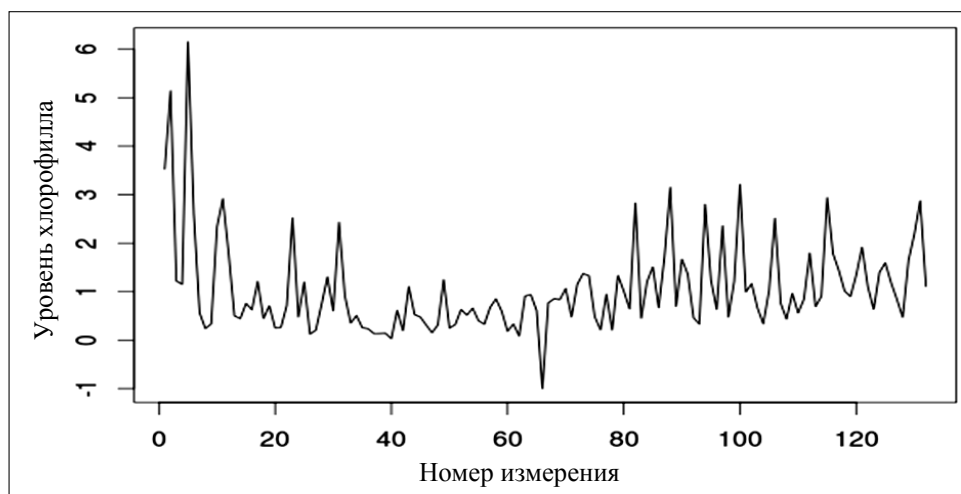


Рис. 2. Концентрация хлорофилла



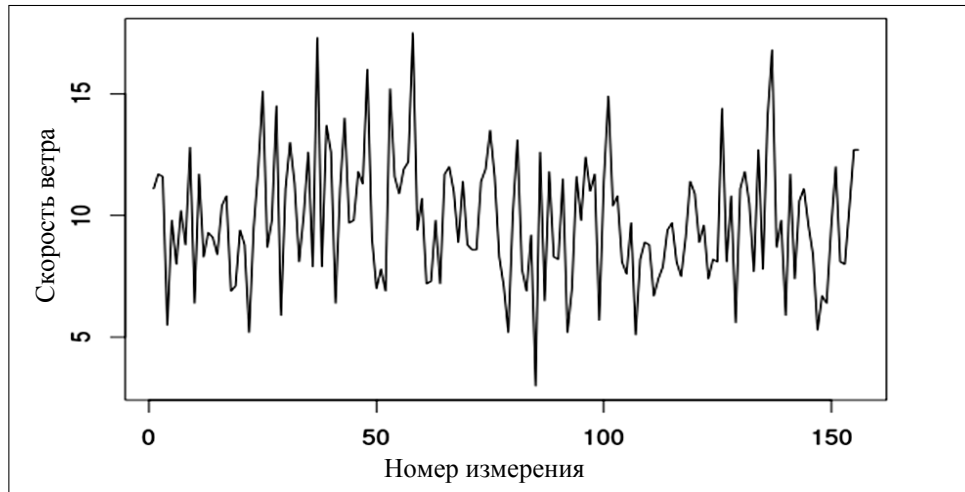


Рис. 3. Скорость ветра

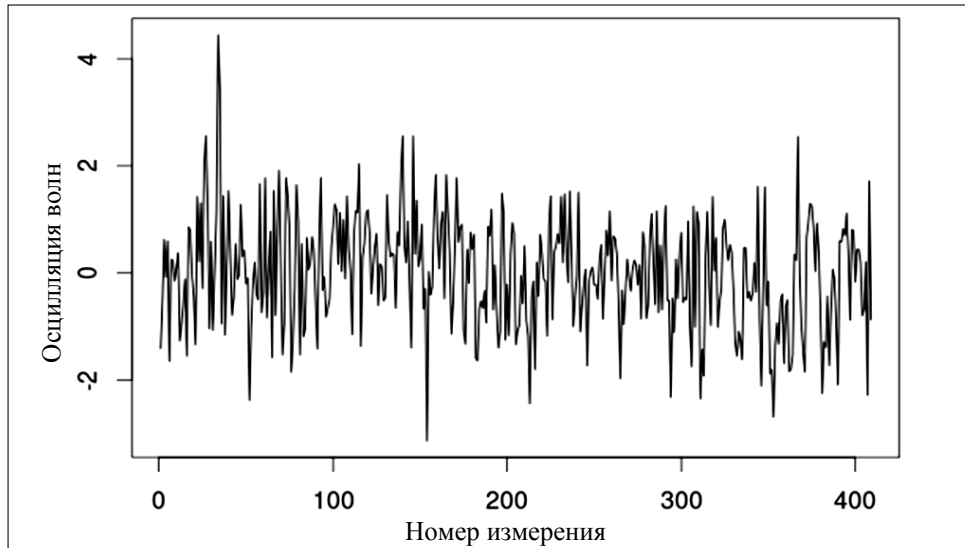


Рис. 4. Осциляция волн

**Таблица 4.** Параметры аппроксимации реальных данных

Характеристики	Планктон	Хлорофилл	Скорость ветра	Осциляция волн в Атлантике
$D$	0,55	0,48	0,63	0,61
$\lambda$	1,2	1,44	1	1,07
$\hat{H}$	0,3	0,1	0,4	0,1
$A_n$	-0,41	-1,26	-274	-3,44
$A = \lim A_n$	-0,48	-1,17	-267	-3,57
$B_n$	0,19	1,2	-123	-6,38
$\beta_B$	0,74	2,46	1921	9,88

Из табл. 4 следует антиперсистентность всех исследуемых данных. Контрольные статистики всех примеров удовлетворяют неравенствам (6), то есть аппроксимация приведенных временных рядов удовлетворительна. Возможно наличие детерминированной составляющей для концентрации хлорофилла и осцилляции океанских волн.

## ВЫВОДЫ

Рассмотрена следующая задача анализа временных рядов: определить соотношение детерминизма (как правило, хаотической природы) и стохастичности наблюдаемой временной последовательности. Известные методы решения такой задачи основаны на вычислении доминантного показателя Ляпунова и требуют большого объема выборки. В работе предложен новый метод вычисления искомого соотношения, использующий идею аппроксимации временного ряда фрактальным броуновским движением. Соотношение хаотичность/стохастичность в исследуемом временном ряде определяется значениями статистик, соответствующих этой аппроксимации.

## ЛИТЕРАТУРА

1. Кузнецов С.П. Динамический хаос. — М.: Физматлит, 2006. — 296 с.
2. Mishura Y. Stochastic Calculus for Fractional Brownian Motion and Related Processes. Lecture Notes in Mathematics. V. 1929 // Springer-Verlag, 2008. — 392 p.
3. Beran J. Statistics for Long-Memory Processes / Beran J. — Chapman and Hall. — 1995. — 315 p.
4. Бондаренко В.В. Итерационный алгоритм оценивания параметров фрактального броуновского движения // Проблемы управления и информатики. — 2012. — № 4. — С. 28–33.
5. Бондаренко В.В. Аппроксимация временного ряда степенной функцией фрактального броуновского движения // Проблемы управления и информатики. — 2013. — № 3. — С. 113–116.
6. Nourdin I., Nualart D., Tudor C. Central and non-central limit theorems for weighted power variations of fractional Brownian motion // Ann. Institute H. Poincaré Probab Statist. — 2010 — **46**, № 4. — P. 1055–1079.
7. Бондаренко В.В. Проверка качества аппроксимации временного ряда фрактальным броуновским движением // Проблемы управления и информатики. — 2014. — № 3. — С. 102–108.
8. Данные Научно-практического центра АН Республики Беларусь. — <http://biobel.bas-net.by/zoo>.
9. University of Arizona Tree Ring Data: <http://rda.ucar.edu/#!lfd?nb=y&b=topic&v=Paleoclimate>.
10. Climate Prediction Center(USA). — <http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/pna.shtml>.

Поступила 21.12.2014

## ПОЧЕМУ ТОЧНОСТЬ ИЗМЕРЕНИЯ ФИЗИЧЕСКИХ ВЕЛИЧИН ОГРАНИЧЕНА

И.И. ГОРБАНЬ

Приведены результаты исследований, проясняющие механизм ограничения точности измерений реальных физических величин. Показано, что ключевую роль в ограничении точности играет неидеальный характер феномена статистической устойчивости, проявляющийся в отсутствии сходимости (несостоятельности) статистических оценок. Выявлены особенности действия закона больших чисел и центральной предельной теоремы теории вероятностей в условиях нарушения статистической устойчивости. Результаты теоретических исследований подтверждены экспериментально. Установлено, что реальные оценки носят гиперслучайный характер. Гиперслучайная природа реальных оценок проявляется при больших объемах выборки.

### ВВЕДЕНИЕ

Научно-технический прогресс неразрывно связан с повышением точности измерений. Несмотря на колоссальные усилия, предпринимаемые на протяжении столетий, каждый шаг по повышению точности измерений дается с большим трудом. Естественно возникает вопрос: существует ли предел точности измерений реальных физических величин?

Вопрос не нов. Благодаря классическим работам Фишера, Крамера, Рао, Ван Триса, Тихонова и др. в XX веке в теории вероятностей сформировалось широко распространенное мнение, что точность, определяемая объемом выборки и свойствами помехи, при неограниченном увеличении объема выборки теоретически может возрасти до бесконечности [1–3].

Этот оптимистический вывод, к сожалению, не находит подтверждения на практике. Инженеры и физики знают, что путем увеличения объема выборки во многих случаях можно повысить точность измерений, но не беспредельно — рано или поздно начинают сказываться те или иные факторы, ограничивающие возможность повышения точности.

Исследование причин противоречия между теорией и практикой привело к пониманию, что известный издавна феномен статистической устойчивости массовых явлений, являющийся физической основой теории вероятностей, носит неидеальный (ограниченный) характер.

Теория вероятностей [4] изучает законы массовых явлений, описывая их с помощью случайных (стохастических) математических моделей. В основе их построения лежит физическая гипотеза абсолютной (идеальной) статистической устойчивости частоты реальных событий, из которой следует абсолютная статистическая устойчивость (статистическая прогнозируемость) параметров и характеристик любых физических явлений — реальных событий, величин, процессов и полей.

На протяжении столетий считалось, что гипотеза идеальной статистической устойчивости адекватно отражает реалии физического мира. Однако экспериментальные исследования различных физических величин и процессов на больших интервалах наблюдения показали [5–17], что эта гипотеза экспериментально не подтверждается.

При небольших временных, пространственных или пространственно-временных интервалах наблюдения увеличение объема статистических данных приводит к уменьшению уровня флуктуаций статистических оценок, что создает иллюзию идеальной статистической устойчивости. Но, начиная с некоторого критического объема данных, уменьшение уровня флуктуаций прекращается и они приобретают незатухающий характер. Дальнейшее увеличение числа данных либо практически не сказывается на уровне флуктуаций оценок, либо даже приводит к их росту.

Этот эффект можно объяснить изменчивостью характеристик и параметров реальных объектов и изменчивостью условий их наблюдения. Изменения происходят на всех уровнях, в том числе статистическом.

Происходящие изменения отражаются на физических закономерностях (в частности описываемых законом больших чисел и центральной предельной теоремой теории вероятностей) и приводят к нарушению статистической устойчивости, что влечет за собой ограничение потенциальной точности измерений.

Исследования нарушений статистической устойчивости явлений и разработка эффективных средств адекватного описания мира с учетом таких нарушений привели к построению новой физико-математической теории — теории гиперслучайных явлений [5, 6, 9].

Под гиперслучайным явлением (гиперслучайным событием, величиной или функцией) понимается множество  $G$  случайных явлений (случайных событий, величин или функций), ассоциируемых с различными условиями наблюдения  $g \in G$ . Для каждого  $g$ -го случайного элемента этого множества определена вероятностная мера, однако для самих условий вероятностная мера не определена.

Вопрос о потенциальной точности измерений с учетом нарушений статистической устойчивости исследовался в ряде работ, в частности [5, 6, 17–19].

**Цель работы** — обобщение, уточнение и детализация полученных результатов на основе недавно разработанных принципов описания расходящихся и многозначных функций (отображений) [20–22].

## ОБОБЩЕННЫЙ ПРЕДЕЛ ЧИСЛОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

В статьях [20–21] обобщены понятия предела и сходимости числовой последовательности  $\{x_n\} = \{x_n\}_{n=1, \infty}$ . Поясним суть этих обобщений.

Согласно классическим представлениям [23–24] последовательность  $\{x_n\}$  считается сходящейся при  $n \rightarrow \infty$ , если существует необходимо единственный предел  $a = \lim_{n \rightarrow \infty} x_n$  (в частности, бесконечно большой по модулю).

Последовательность, не имеющая единственного предела, считается расходящейся.

Расходящиеся последовательности можно описать различными способами. Один из них состоит в следующем.

Из любой последовательности можно сформировать множество частичных последовательностей (подпоследовательностей), получаемых из исходной последовательности вычеркиванием некоторых ее членов.

Известно, что если последовательность сходится, то сходятся все ее частичные последовательности. Если же последовательность расходится, то не обязательно расходятся все ее частичные последовательности. Некоторые из них могут сходиться к определенным пределам, называемые предельными точками или частичными пределами.

В работах [20–21] введено понятие обобщенного предела  $\text{LIM}_{n \rightarrow \infty} x_n$  последовательности  $\{x_n\}$ , под которым понимается множество всех ее предельных точек с мерами. Такое множество точек с мерами названо спектром предельных точек  $\tilde{S}$ :

$$\tilde{S} = \text{LIM}_{n \rightarrow \infty} x_n. \quad (1)$$

Здесь и далее факт многозначности подчеркнут знаком тильда над буквой, обозначающей многозначную величину или функцию.

Спектр  $\tilde{S}$  рассматривается как обобщение понятия предела на случай любой последовательности, в том числе расходящейся. Если последовательность  $\{x_n\}$  сходится, то спектр предельных точек состоит из одного элемента (числа), если же расходится, то спектр состоит из множества чисел. В последнем случае можно говорить о сходимости последовательности к множеству чисел (в частном случае — к интервалу).

Спектр может быть охарактеризован различными способами. Простейшими параметрами, характеризующими разброс предельных точек и их местоположение, являются граничные точки: минимальная  $a_i$  и максимальная  $a_s$ . Более полное описание спектра дает функция распределения предельных точек

$$\tilde{F}(x) = \text{LIM}_{N \rightarrow \infty} \frac{n(x)}{N}, \quad (2)$$

где  $n(x)$  — количество членов начального отрезка  $\{x_n\}_{n=1, \overline{N}}$  последовательности  $\{x_n\}$ , меньших  $x$ .

Если последовательность  $\{x_n\}$  сходится к числу  $a$ , то функция распределения предельных точек представляет собой однозначную функцию  $F(x)$  в виде функции единичного скачка в точке  $a$  (рис. 1 а) (тогда мера равна единице в точке  $a$  и нулю во всех остальных точках). Если же эта последовательность расходится (сходится к множеству чисел), то функция распределения предельных точек представляет собой или однозначную убывающую функцию  $F(x)$ , отличную от функции единичного скачка, (рис. 1 б), или многозначную функцию  $\tilde{F}(x)$  (рис. 1 в).

На спектр предельных точек можно взглянуть под другим углом зрения. В первом случае (рис. 1 а) обобщенный предел (1) числовой последова-

тельности  $\{x_n\}$  представляет собой детерминированную однозначную величину  $a$  (число без меры или множество чисел с единичной мерой в точке  $a$  и нулю во всех остальных точках), во втором случае (рис. 1 б) — случайную величину, а в третьем (рис. 1 в) — гиперслучайную величину.

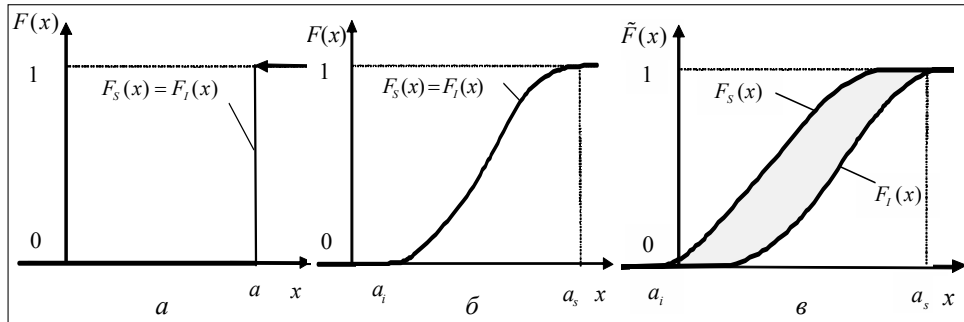


Рис. 1. Функция распределения  $\tilde{F}(x)$  предельных точек и границы  $F_I(x)$ ,  $F_S(x)$  функции распределения предельных точек для последовательности  $\{x_N\}$ , сходящейся к числу  $a$  (а) и к множеству чисел  $x \in [a_i, a_s]$ , описываемому однозначной (б) и многозначной (в) функциями распределения (соответственно  $F(x)$  и  $\tilde{F}(x)$ )

Строго говоря, число, рассматриваемое как одноточечное множество без меры, и как множество чисел с мерой, равной единицы в одной точке и нулю — в остальных точках — разные понятия. Однако, поскольку для дальнейшего изложения это обстоятельство не является существенным, различие между ними не делается.

Для описания гиперслучайной величины можно использовать [5, 6, 9] различные однозначные характеристики и параметры, в частности, нижнюю  $F_I(x)$  и верхнюю  $F_S(x)$  границы функции распределения  $\tilde{F}(x)$  (рис. 1), а также (при определенных условиях) — плотности распределения этих границ

$$f_I(x) = \frac{dF_I(x)}{dx}, \quad f_S(x) = \frac{dF_S(x)}{dx},$$

моменты границ (математические ожидания границ  $m_I$ ,  $m_S$ , дисперсии границ  $D_I$ ,  $D_S$ ), границы моментов (границы математического ожидания  $m_i$ ,  $m_s$ , границы дисперсии  $D_i$ ,  $D_s$ ) и другие характеристики и параметры теории гиперслучайных явлений [6, 9].

## ЗАКОН БОЛЬШИХ ЧИСЕЛ ПРИ НАРУШЕНИИ СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ

Закон больших чисел был открыт Я. Бернулли [25] и опубликован после его смерти триста лет назад (в 1713 г.). Известно несколько вариантов этого закона, один из которых был сформулирован Чебышевым в виде следующей теоремы [6, 26].

**Теорема.** Пусть  $X_1, X_2, \dots, X_N$  — начальный отрезок последовательности попарно независимых случайных величин с математическими ожи-

данными  $m_1, m_2, \dots, m_N$  и ограниченными дисперсиями. Тогда при устремлении объема выборки  $N$  к бесконечности среднее  $m_{xN}^* = \frac{1}{N} \sum_{n=1}^N X_n$  случайных величин  $X_1, X_2, \dots, X_N$  (выборочное среднее) стремится по вероятности к среднему  $m_{xN} = \frac{1}{N} \sum_{n=1}^N m_n$  математических ожиданий  $m_1, m_2, \dots, m_N$ :

$$\lim_{N \rightarrow \infty} P\{|m_{xN}^* - m_{xN}| > \varepsilon\} = 0 \quad (\varepsilon > 0). \quad (3)$$

При конечном объеме выборки  $N$  величина  $m_{xN}^*$  — случайная величина, а величина  $m_{xN}$  — детерминированная величина (число). Величина  $m_{xN}^*$  описывается функцией распределения  $F_{m_{xN}^*}(x)$ , а  $m_{xN}$  — функцией распределения  $F_{m_{xN}}(x)$  в виде функции единичного скачка в точке  $m_{xN}$ .

Обратим внимание на одно чрезвычайно важное обстоятельство, долгое время остававшееся незамеченным: закон больших чисел не гарантирует, что при  $N \rightarrow \infty$  выборочное среднее  $m_{xN}^*$  и среднее математических ожиданий  $m_{xN}$  имеют обычные пределы. Этот закон утверждает лишь сходимость по вероятности выборочного среднего к среднему математических ожиданий, не требуя при этом их сходимости к определенным числам.

Пределы последовательностей  $\{m_{xN}^*\}$  и  $\{m_{xN}\}$  могут быть многозначными величинами — случайными или гиперслучайными. В дальнейшем, следуя указанной выше договоренности, касающейся обозначений однозначных и многозначных величин и функций, однозначные пределы последовательностей  $\{m_{xN}^*\}$  и  $\{m_{xN}\}$  будем обозначать  $m_x^*$  и  $m_x$ , а многозначные —  $\tilde{m}_x^*$  и  $\tilde{m}_x$ .

Заметим, что различные варианты сходимости последовательности гиперслучайных величин (в частном случае рассматриваемой здесь числовой последовательности) к гиперслучайной величине и сходимости последовательности гиперслучайных функций (в частном случае рассматриваемой ниже последовательности детерминированных функций) к гиперслучайной функции рассмотрены в монографии [6].

Вне зависимости от того, являются ли рассматриваемые обобщенные пределы однозначными или многозначными, согласно выражению (3) по мере увеличения объема выборки  $N$  выборочное среднее  $m_{xN}^*$  приближается к среднему математических ожиданий  $m_{xN}$ .

При  $N \rightarrow \infty$  возможны два случая:

1. Величина  $m_{xN}^*$  сходится к однозначному выборочному среднему математических ожиданий  $m_x$  (числу).
2. Величина  $m_{xN}^*$ , становясь в пределе многозначной величиной  $\tilde{m}_x^*$ , сходится к многозначной величине  $\tilde{m}_x$ .

Случай 1 — идеализированный, рассматриваемый в теории вероятностей [1–3]. В этом случае предел  $m_x$  среднего математических ожиданий описывается функцией распределения  $F_{m_x}(x)$  в виде функции единичного скачка в точке  $m_x$ . К ней стремится функция распределения  $F_{m_{xN}}^*(x)$  выборочного среднего  $m_{xN}^*$  при  $N \rightarrow \infty$  (рис. 2 а).

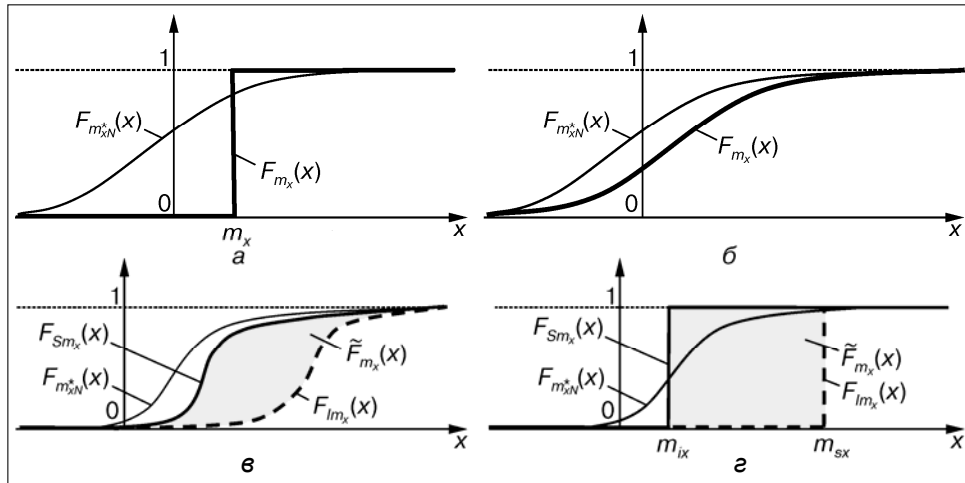


Рис. 2. Формирование предельной функции распределения  $F_{m_x}^*(x)$  выборочного среднего случайной величины при  $N \rightarrow \infty$ , когда предельные выборочное среднее и среднее математических ожиданий — числа (а), когда они — случайные величины (б) и когда они — гиперслучайные величины (в, з) (в — общий случай, з — частный случай)

Случай 2 более реалистичен. Здесь предельное выборочное среднее  $\tilde{m}_x^*$  и предельное среднее математических ожиданий  $\tilde{m}_x$  описываются многозначными спектрами  $\tilde{S}_{m_x}^*$  и  $\tilde{S}_{m_x}$ . При этом возможны два варианта.

2.1. Предельное выборочное среднее  $\tilde{m}_x^*$  и предельное среднее математических ожиданий  $\tilde{m}_x$  являются случайными величинами. При этом спектры  $\tilde{S}_{m_x}^*$  и  $\tilde{S}_{m_x}$  характеризуются однозначными функциями распределения  $F_{m_x}^*(x)$  и  $F_{m_x}(x)$  (рис. 2 б).

2.2. Величины  $\tilde{m}_x^*$ ,  $\tilde{m}_x$  являются гиперслучайными величинами. При этом спектры  $\tilde{S}_{m_x}^*$  и  $\tilde{S}_{m_x}$  характеризуются многозначными функциями распределения  $\tilde{F}_{m_x}^*(x)$  и  $\tilde{F}_{m_x}(x)$  (рис. 2 в).

Т.к. сходимость по распределению последовательности случайных величин менее сильная, чем сходимость по вероятности [3, 26], в варианте 2.1 предельная функция распределения  $F_{m_x}^*(x)$  совпадает с предельной функцией распределения  $F_{m_x}(x)$ .



Сходимость по распределению последовательности гиперслучайных величин также менее сильная, чем сходимость по вероятности [6]. Поэтому в варианте 2.2 предельная функция распределения  $\tilde{F}_{m_x}^*(x)$  совпадает с предельной функцией распределения  $\tilde{F}_{m_x}(x)$ . При этом нижняя граница  $F_{Im_x}^*(x)$  предельной функции распределения  $\tilde{F}_{m_x}^*(x)$  совпадает с нижней границей  $F_{Im_x}(x)$  предельной функции распределения  $\tilde{F}_{m_x}(x)$ , а верхняя граница  $F_{Sm_x}^*(x)$  предельной функции распределения  $\tilde{F}_{m_x}^*(x)$  — с верхней границей  $F_{Sm_x}(x)$  предельной функции распределения  $\tilde{F}_{m_x}(x)$ . На рис. 2 в расположенная между указанными границами область неопределенности затемнена.

В статье [21] доказана теорема, из которой следует, что, если функция распределения, описывающая спектр последовательности средних детерминированных величин — многозначная функция, то соответствующая область неопределенности непрерывная. На основании этой теоремы область неопределенности функции распределения  $\tilde{F}_{m_x}(x)$  — непрерывная.

Интервал, в котором флуктуирует выборочное среднее  $m_{xN}^*$  при  $N \rightarrow \infty$ , характеризуется нижней границей  $m_{ix}^*$  достижения функцией  $F_{Sm_x}^*(x)$  минимального (нулевого) значения и верхней границей  $m_{sx}^*$  достижения функцией  $F_{Im_x}^*(x)$  максимального (единичного) значения. Естественно эти границы совпадают с соответствующими границами  $m_{ix}, m_{sx}$  функций  $F_{Sm_x}(x), F_{Im_x}(x)$ :  $m_{ix}^* = m_{ix}, m_{sx}^* = m_{sx}$ . Указанные границы могут быть как конечными, так и бесконечными.

Заметим, что в работах [6, 18] исследован частный случай, когда верхняя  $F_{Sm_x}^*(x)$  и нижняя  $F_{Im_x}^*(x)$  границы функции распределения  $\tilde{F}_{m_x}^*(x)$  описываются функциями единичного скачка соответственно в точках  $m_{ix}^*, m_{sx}^*$ , а верхняя  $F_{Sm_x}(x)$  и нижняя  $F_{Im_x}(x)$  границы функции распределения  $\tilde{F}_{m_x}(x)$  — функциями единичного скачка соответственно в точках  $m_{ix}, m_{sx}$  (рис. 2 з).

Таким образом, теоретически среднее случайной выборки может сходиться к определенному числу, стремиться к плюс или минус бесконечности или флуктуировать в пределах интервала  $[m_{ix}, m_{sx}]$ . В последнем случае имеет место сходимость выборочного среднего к интервалу.

Особый интерес представляют первый и третий случаи. В первом случае при увеличении объема выборки флуктуации среднего затухают и погрешность измерения стремится к нулю, а в третьем случае — эти флуктуации не затухают и погрешность ограничена снизу длиной интервала  $[m_{ix}, m_{sx}]$ .

Для выяснения характера сходимости реальных физических процессов была проанализирована динамика поведения выборочных средних множества процессов различной физической природы на больших интервалах наблюдения.

Исследовались [6–17] колебания напряжения городской электросети на 60-ти часовых интервалах наблюдения; колебания магнитного поля Земли за 13 лет; колебания высоты морских волн и периода их следования за 15 месяцев; колебания температуры воздуха и количества осадков в различных регионах за многие десятки лет, колебания скорости ветра в районе Чернобыля за 11 лет; колебания температуры воды в Тихом океане за 8 месяцев; колебания интенсивности излучения разных астрофизических объектов за 16 лет и др.

Выяснилось, что интервал статистической устойчивости всех исследованных реальных физических процессов ограничен, в результате чего все процессы проявляют выраженную тенденцию сходимости не к конкретным числам, а к интервалам.

Незатухающими флуктуациями при больших объемах выборки  $N$  выборочного среднего  $m_{xN}^*$  в пределах этих интервалов можно объяснить ограниченную точность реальных измерений.

## ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА ПРИ НАРУШЕНИИ СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ

Более глубокое понимание описанного эффекта дает центральная предельная теорема теории вероятностей для случайных величин.

Рассмотрим центральную предельную теорему в варианте, изложенном в работе [26].

**Теорема (Линдберга–Феллера).** Пусть  $X_1, X_2, \dots, X_N$  — в общем случае неоднородная случайная выборка, элементы которой взаимно независимы и описываются функциями распределения  $F_{x_n}(x)$  с математическими ожиданиями  $m_{x_n}$  и дисперсиями  $D_{x_n}$  ( $n = \overline{1, N}$ ). Выполняется условие Линдберга: при любом  $\varepsilon > 0$

$$\lim_{N \rightarrow \infty} \frac{1}{B_N^2} \sum_{n=1}^N \int_{|x-m_{x_n}| > \varepsilon B_N} (x-m_{x_n})^2 dF_{x_n}(x) = 0 \quad (4)$$

где  $B_N^2 = \sum_{n=1}^N D_{x_n}$  — сумма дисперсий  $D_{x_n}$  случайных величин  $X_n$  ( $n = \overline{1, N}$ ).

Тогда функция распределения  $F_{m_{xN}^*}(x)$  выборочного среднего  $m_{xN}^*$  сходится равномерно к гауссовской функции распределения

$$F(x/m_{xN}, D_{xN}) = \Phi\left(\frac{x-m_{xN}}{\sqrt{D_{xN}}}\right) \quad (5)$$

с математическим ожиданием  $m_{xN}$  и дисперсией  $D_{xN} = B_N^2 / N^2$ :

$$\lim_{N \rightarrow \infty} F_{m_{xN}}^*(x) = \lim_{N \rightarrow \infty} F(x/m_{xN}, D_{xN}), \quad (6)$$

где  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-z^2/2) dz$ .

Условие (4) является необходимым и достаточным условием сходимости выборочного среднего к гауссовскому распределению.

Используя описанную в [26] схему вывода равенства (6), можно доказать более общее утверждение, а именно: разность между функцией распределения  $F_{m_{xN}}^*(x)$  выборочного среднего  $m_{xN}^*$  и гауссовской функцией распределения  $F(x/m_{xN}, D_{xN})$ , описываемое выражением (5), сходится равномерно к нулю:

$$\lim_{N \rightarrow \infty} [F_{m_{xN}}^*(x) - F(x/m_{xN}, D_{xN})] = 0. \quad (7)$$

Между формулами (6) и (7) имеется существенное различие. Формула (6) предполагает наличие у выборочного среднего  $m_{xN}^*$  однозначной предельной функции распределения  $F_{m_x}^*(x)$ , к которой стремится функция  $F_{m_{xN}}^*(x)$  при  $N \rightarrow \infty$ , и наличие однозначной предельной функции распределения  $F_{m_x}(x) = F(x/m_x, D_x)$ , к которой стремится функция  $F(x/m_{xN}, D_{xN})$  при том же условии, где  $m_x$  и  $D_x$  — соответственно математическое ожидание и дисперсия предельной гауссовской функции распределения.

Формула же (7) допускает, что рассматриваемые предельные функции распределения — многозначные функции.

Многозначность предельной функции распределения, к которой стремится функция  $F(x/m_{xN}, D_{xN})$ , обусловлена многозначностью математического ожидания и (или) дисперсии. Поэтому в выражении  $\tilde{F}_{m_x}(x) = \tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$ , представляющем предельную функцию распределения, фигурируют многозначные параметры  $\tilde{m}_x$  и  $\tilde{D}_x$ . Поскольку эти параметры в общем случае гиперслучайные, гиперслучайной оказывается и функция  $\tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$ . Ее можно трактовать как множество однозначных гауссовских функций распределения с различными однозначными математическими ожиданиями  $m_x \in \tilde{m}_x$  и дисперсиями  $D_x \in \tilde{D}_x$ .

Выражение (7) означает, что имеет место сходимость по распределению последовательности детерминированных функций  $F_{m_{xN}}^*(x)$  к гиперслучайной функции  $\tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$ . Иными словами многозначные предельные функции распределения  $\tilde{F}_{m_x}^*(x)$ ,  $\tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$  описываются одинаковыми множествами однозначных функций распределения.

Когда параметры  $m_x$  и  $D_x$  — числа, причем  $D_x = 0$ , предельная гауссовская функции распределения  $F_{m_x}(x) = F(x/m_x, D_x)$  имеет вид функции единичного скачка, изображенной на рис. 2 а полужирной линией. Когда  $m_x$  и  $D_x$  — однозначные величины, но  $D_x \neq 0$ , то эта функция распределения имеет вид кривой, изображено на рис. 2 б полужирной линией.

Когда предельное математическое ожидание  $\tilde{m}_x$  — многозначная величина или когда и предельное математическое ожидание  $\tilde{m}_x$  и предельная дисперсия  $\tilde{D}_x$  — многозначные величины, то предельная функция распределения  $\tilde{F}_{m_x}(x) = \tilde{F}(x/\tilde{m}_x, \tilde{D}_x)$  — многозначная функция. На рис. 2 в, г она изображена в виде затемненной области.

Когда предельное математическое ожидание  $m_x$  — число, а предельная дисперсия  $\tilde{D}_x$  — многозначная величина, предельная функция распределения  $\tilde{F}_{m_x}(x) = \tilde{F}(x/m_x, \tilde{D}_x)$  — тоже многозначная функция. При этом она имеет определенную специфику: состоит из двух областей, соприкасающихся в одной точке (затемненные области на рис. 3). Абсцисса точки соприкосновения этих областей равна математическому ожиданию  $m_x$ .

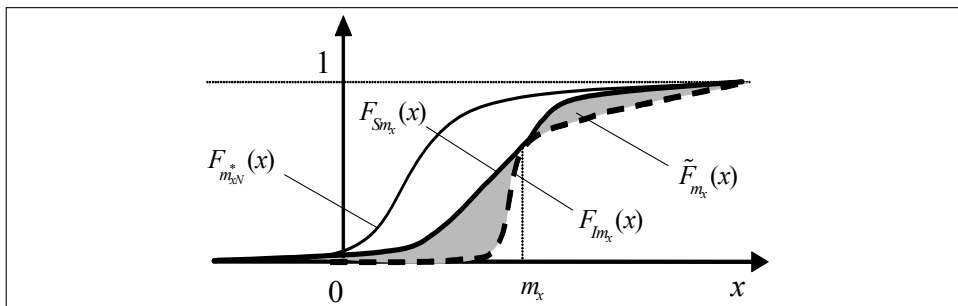


Рис. 3. Формирование предельной функции распределения  $\tilde{F}_{m_x^*}(x)$  выборочного среднего случайной величины при  $N \rightarrow \infty$ , когда математическое ожидание  $m_x$  — число, а дисперсия  $\tilde{D}_x$  — многозначная величина

## ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Для иллюстрации сходимости выборочных средних реальных процессов к интервалам ниже приведено описание и результаты двух из упомянутых выше экспериментов, касающихся исследований статистической устойчивости напряжения городской электросети и колебаний интенсивности излучения пульсара PSRJ 1012+5307 в рентгеновском диапазоне частот. Выбор именно этих примеров связан с тем, что колебание напряжения сети — одно из наиболее неустойчивых колебаний, а колебание интенсивности излучения пульсара — одно из наиболее устойчивых.

При изучении колебаний напряжения электросети напряжение подавалось на вход 16 разрядного аналого-цифрового преобразователя компьютера

через понижающий трансформатор и согласующее устройство (делитель напряжения).

Снятие данных осуществлялось с частотой дискретизации 5 кГц. По каждым 1024 отсчетам вычислялись действующие (эффективные) значения напряжения, записываемые в память компьютера. Запись велась сеансами на протяжении двух месяцев с перерывами в несколько дней. Продолжительность каждого сеанса составляла около 60 часов. За время сеанса записывалось  $N = 2^{20} \approx 1$  млн отсчетов напряжения.

Обработка полученных записей сводилась к вычислению и анализу выборочных средних  $m_{xN}^*$ , оценок параметров статистической неустойчивости (в частности описанных в статьях [16, 17] параметров  $\gamma_N^*$ ,  $\mu_N^*$ ,  $h_N^*$  и  $\Gamma_N^*$ ,  $M_N^*$ ,  $H_N^*$ , позволяющих оценить интервал статистической неустойчивости), оценок функций распределения колебания напряжения  $F_g^*(x)$  на прилегающих друг к другу интервалах наблюдения длительностью около часа ( $g = \overline{1,64}$ ) и оценок функций распределения выборочного среднего  $F_{m_{xN}^*}^*(x)$ .

Результаты обработки одной из записей приведены на рис. 4.

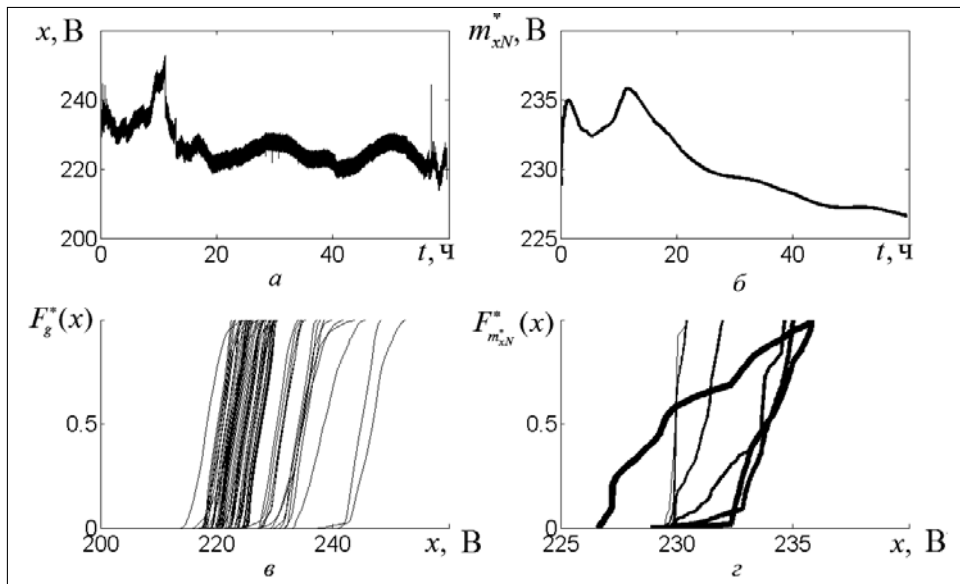


Рис. 4. Колебание напряжения  $x$  городской электросети (а) и его выборочного среднего  $m_{xN}^*$  (б), оценки функции распределения напряжения  $F_g^*(x)$  на 64 прилегающих друг к другу интервалах наблюдения (в) и оценки функции распределения выборочного среднего напряжения  $F_{m_{xN}^*}^*(x)$  при различном объеме выборки  $N = 2^r$  ( $r = 8, 10, 12, 14, 16, 18, 20$ ) (г) (толщина линий возрастает с увеличением значения параметра  $r$ )

Как видно из рис. 4 а, качество электросети низкое. В данном случае напряжение колеблется между  $x_i = 215$  В и  $x_s = 255$  В.

На рассматриваемом временном интервале выборочное среднее не проявляет тенденции к стабилизации (рис. 4 б), что свидетельствует о явном нарушении статистической устойчивости процесса.

Кривые функции распределения  $F_g^*(x)$ , соответствующие разным значениям параметра  $g$ , сильно отличаются друг от друга (в первую очередь по своему местоположению) (рис. 4 в), что свидетельствует о выраженной нестационарности исследуемого колебания.

Кривые функции распределения выборочного среднего  $F_{m_{xN}}^*(x)$ , полученные для нарастающих по экспоненциальному закону объемов выборки (рис. 4 з), демонстрируют полное отсутствие тенденции стремления этой функции распределения к какой-то определенной предельной функции распределения  $F_{m_x}(x)$  и, тем более, стремления выборочного среднего  $m_{xN}^*$  к какому-то определенному предельному значению  $m_x$ .

При небольших значениях параметра  $r$  (8 и 10) по виду кривых функции распределения выборочного среднего  $F_{m_{xN}}^*(x)$  (рис. 4 з) можно предположить наличие стремления ее к гауссовскому распределению (причем, как и предсказывает теория вероятностей, с уменьшающейся дисперсией при росте объема выборки). Однако при больших значениях  $r$  (начиная с 10 до 20), как видно из рисунка, предполагаемая тенденция не наблюдается — закон распределения выборочного среднего  $F_{m_{xN}}^*(x)$  становится явно негауссовским.

Дисперсия выборочного среднего  $m_{xN}^*$  с увеличением объема выборки то нарастает (для значений  $r$  от 8 до 14 и от 18 до 20), то спадает (для значений  $r$  от 14 до 18). В целом при переходе от малых к большим объемам выборки дисперсия не только не проявляет тенденции стремления к нулю, предсказываемой теорией вероятностей, а возрастает, причем во много раз (размах выборочного среднего увеличивается примерно от 1 В до 8 В).

Судя по полученным результатам, функция распределения выборочного среднего стремится к многозначной функции  $\tilde{F}_{m_x}(x)$  общего вида (типа рис. 2 в). К такому же выводу приводит анализ результатов обработки данных, полученных в других сеансах записи колебаний напряжения электро-сети.

Результаты аналогичных расчетов для колебаний интенсивности излучения пульсара PSRJ 1012+5307 приведены на рис. 5.

Исходные данные взяты с [27]. Измерения интенсивности  $x$  проводились в период с 1 января 1996 г. по 31 декабря 2011 г. Средняя периодичность измерений составляла 2,8 часов.

Колебание интенсивности пульсара, как следует из рис. 5а, напоминает белый шум. Показание прибора, регистрирующего интенсивность излучения, меняется в диапазоне от  $x_i = -15$  В до  $x_s = 15$  В.

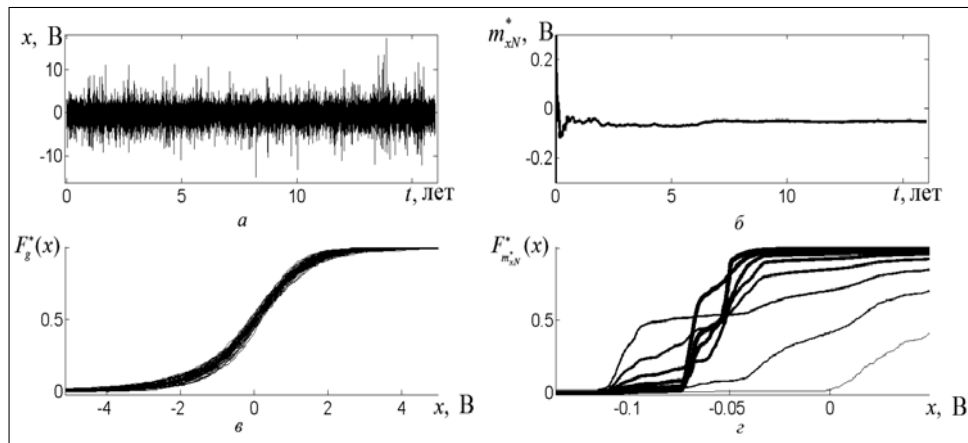


Рис. 5. Колебание интенсивности излучения пульсара PSRJ 1012+5307 (а) и его выборочного среднего  $m_{xN}^*$  (б), оценки функции распределения  $F_g^*(x)$  на 64 прилегающих друг к другу интервалах наблюдения (в) и оценки функции распределения выборочного среднего  $F_{m_{xN}^*}^*(x)$  при различном объеме выборки  $N = 2^r$  ( $r = \overline{8,15}$ ) (г) (толщина линий возрастает с увеличением значения параметра  $r$ )

На рис. 5 б кривая выборочного среднего на первый взгляд проявляет тенденцию к стабилизации, что говорит об отсутствии явно выраженных нарушений статистической устойчивости процесса. На отсутствие явных нарушений статистической устойчивости указывает и близость кривых оценок функций распределения напряжения  $F_g^*(x)$ , рассчитанных для прилегающих друг к другу интервалов наблюдения длительностью три месяца ( $g = \overline{1,64}$ ) (рис. 5 в).

Однако кривые оценок функции распределения выборочного среднего  $F_{m_{xN}^*}^*(x)$ , полученные для нарастающих по экспоненциальному закону объемов выборки (рис. 5 г), демонстрируют отсутствие стремления выборочного среднего  $m_{xN}^*$  к какому-то определенному предельному значению  $m_x$  и даже отсутствие тенденции стремления функции распределения  $F_{m_{xN}^*}^*(x)$  к какой-то определенной предельной функции распределения  $F_{m_x}(x)$ .

По динамике изменения функции распределения выборочного среднего  $F_{m_{xN}^*}^*(x)$  при небольших значениях параметра  $r$  (от 8 до 13) можно предположить наличие стремления этой функции распределения к гауссовскому распределению (причем, как и предсказывает теория вероятностей с уменьшающейся дисперсией при росте объема выборки). Но при больших значениях  $r$  (начиная с 13 до 15) предполагаемая тенденция не прослеживается — закон распределения выборочного среднего  $F_{m_{xN}^*}^*(x)$  становится явно негауссовским.

При переходе от малых к большим объемам выборки дисперсия выборочного среднего  $m_{xN}^*$  вначале проявляет тенденцию к уменьшению (при изменении  $r$  от 8 до 13), но затем перестает уменьшаться. Размах выборочного среднего остается примерно на одном и том же уровне (примерно 0,04 В).

Судя по виду кривых функции распределения выборочного среднего, можно предположить стремление функции  $F_{m_{xN}}^*(x)$  к многозначной функции  $\tilde{F}_{m_x}(x)$  типа, изображенной на рис. 3.

Подобным образом были исследованы выборочные средние и других перечисленных выше процессов. В этих исследованиях ни разу не была зафиксирована тенденция стремления оценки функции распределения выборочного среднего  $F_{m_{xN}}^*(x)$  к какому-то определенному закону распределения, а тем более к гауссовскому закону с дисперсией, стремящейся к нулю.

Полученные результаты экспериментальных исследований показывают, что, хотя на небольших интервалах усреднения выборочные средние реальных физических величин хорошо описываются стохастическими моделями, в действительности эти средние носят гиперслучайный характер и на больших интервалах наблюдения, превосходящих интервал статистической устойчивости, представляются стохастическими моделями неадекватно.

## ЗАКЛЮЧЕНИЕ

Проведенные теоретические и экспериментальные исследования проясняют механизм ограничения потенциальной точности измерения реальных физических величин и подтверждают результаты предыдущих исследований, касающихся действия закона больших чисел и центральной предельной теоремы в условиях нарушения статистической устойчивости [5, 6, 18, 19].

Ограничение точности можно объяснить неидеальным характером феномена статистической устойчивости, проявляющимся в отсутствии сходимости (несостоятельности) статистических оценок. При возрастании объема статистических данных вначале при небольших интервалах усреднения уровень флуктуаций статистических оценок уменьшается. Но из-за нарушения статистической устойчивости, начиная с некоторого критического объема данных (интервала статистической устойчивости), эти флуктуации приобретают незатухающий характер, что приводит к ограничению точности реальных измерений.

Потенциальная точность измерений определяется диапазоном, в котором флуктуирует оценка. При большом объеме выборки описать реальные оценки случайными моделями, характеризуемыми определенными однозначными функциями распределения, невозможно. Адекватное описание обеспечивают гиперслучайные модели, характеризуемые многозначными функциями распределения.

Экспериментальные исследования подтверждают, что при небольших интервалах усреднения гиперслучайная природа реальных выборочных средних практически не проявляется и поэтому эти выборочные средние



хорошо описываются стохастическими моделями. Однако при больших интервалах усреднения проявляется гиперслучайная их природа и тогда для их описания приходится использовать более сложные гиперслучайные модели.

## ЛИТЕРАТУРА

1. Тихонов В.И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем. — М.: Радио и связь, 1991. — 608 с.
2. Ван Трис Г. Теория обнаружения, оценок и модуляции. — М.: Советское радио, 1972, Т. 1. — 743 с.
3. Горбань І.І. Теорія ймовірностей і математична статистика для наукових працівників та інженерів. — К.: ПІММС НАН України, 2003. — 245 с.
4. Колмогоров А.Н. Основные понятия теории вероятностей. — М.: ОНТИ, 1974. — 120 с.
5. Горбань І.І. Теория гиперслучайных явлений. — К.: НАН Украины, 2007. — 181 с.
6. Горбань І.І. Теория гиперслучайных явлений: физические и математические основы. — К.: Наукова думка, 2011. — 320 с.
7. Gorban I.I. Disturbance of statistical stability / In the book “Information Models of Knowledge”. Kiev — Sofia: ITHEA, 2010. — P. 410.
8. Gorban I.I. Disturbance of statistical stability (part II) // Information Theories and Applications. — 2011. — **18**, № 4. — P. 321–334.
9. Gorban I.I. Hyper-random Phenomena: Definition and Description // Information Theories and Applications. — 2008. — **15**, № 3. — P. 203–211.
10. Горбань І.І. Нарушение статистической устойчивости физических процессов // Математические машины и системы. — 2010. — № 1. — С. 171–184.
11. Горбань І.І. Эффект статистической неустойчивости в гидрофизике // Труды десятой Всероссийской конференции «Прикладные технологии гидроакустики и гидрофизики». — СПб: Наука, 2010. — С. 199–201.
12. Горбань І.І. Статистическая устойчивость колебаний температуры воздуха и осадков в районе Москвы // Математические машины и системы. — 2011. — № 3. — С. 97–104.
13. Горбань І.І., Горбань Н.І., Новотрясов В.В., Яроцук І.О. Исследование статистической устойчивости колебаний температуры шельфовой зоны окраинных морей // Материалы докладов седьмого всероссийского симпозиума «Физика геосфер». — Владивосток: Дальнаука, 2011. — С. 542–547.
14. Горбань І.І. Статистическая устойчивость излучения астрофизических объектов // Математические машины и системы. — 2012. — № 2. — С. 155–160.
15. Горбань І.І. Статистически неустойчивые процессы: связь с фликкер, неравновесными, фрактальными и цветными шумами // Известия вузов. Радиоэлектроника. — 2012. — **55**, № 3. — С. 3–18.
16. Горбань І.І. Критерии и параметры статистической неустойчивости // Математические машины и системы. — 2012. — № 4. — С. 106–114.
17. Горбань І.І. Феномен статистической устойчивости // Журнал технической физики. — 2014. — № 3 — С. 22–30.
18. Горбань І.І. Особенности закона больших чисел при нарушениях статистической устойчивости // Известия вузов. Радиоэлектроника. — 2011. — № 7. — С. 31–42.
19. Горбань І.І. Закон больших чисел при нарушениях статистической устойчивости // Математические машины и системы. — 2011. — № 4. — С. 107–115.

20. *Gorban I.I.* Divergent and multiple-valued sequences and functions // Problems of Computer Intellectualization. Book 28. — Kyiv–Sofia: ITHEA, 2012. — P. 395.
21. *Горбань И.И.* Расходящиеся последовательности и функции // Математические машины и системы. — 2012. — № 1. — С. 106–118.
22. *Горбань И.И.* Многозначные величины, последовательности и функции // Математические машины и системы. — 2012. — № 3. — С. 147–161.
23. *Фихтенгольц Г.М.* Курс дифференциального и интегрального исчисления. — М.-Л.: Гос. изд-во физ.-мат. литературы, 1958. — Т. 1. — 607 с.
24. *Ильин В.А., Садовничий В.А., Сендов Бл.Х.* Математический анализ. — М.: Изд-во московского ун-та, 1985. — Т. 1. — 660 с.
25. *Бернулли Я.* О законе больших чисел. — М.: Наука, 1986. — 176 с.
26. *Гнеденко Б.В.* Курс теории вероятностей. — М.: Изд-во физ.-мат. литературы, 1988. — 448 с.
27. *All-Sky Monitor (ASM) team at the Kavli Institute for Astrophysics and Space Research at the Massachusetts Institute of Technology.* — [http://xte.mit.edu/ASM\\_lc.html](http://xte.mit.edu/ASM_lc.html).

*Поступила 16.04.2014*

ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ ПРОБЛЕМИ  
І МЕТОДИ СИСТЕМНОГО АНАЛІЗУ

THEORETICAL AND APPLIED PROBLEMS AND  
METHODS OF SYSTEM ANALYSIS

УДК 681.58

**Применение трендовых геопространственных статистических моделей на основе интеллектуального анализа данных / Путренко В.В.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 7–19.

Исследовано применение трендовых геопространственных статистических моделей в анализе данных устойчивого развития Украины. Рассмотрен математический аппарат, используемый при построении трендовых геопространственных статистических моделей. На примере данных иерархической модели устойчивого развития проведено исследование геопространственных трендов и природы их возникновения путем построения полиномиальных регрессионных моделей, основанных на обработке координированных данных. Разработана методика детекции геопространственных трендов в иерархических моделях данных на основе картографического метода и полиномов первого, второго порядков. Предложено использование картометрических методов для оценки параметров геопространственных трендов. Обосновано использование методов построения геопространственных трендов как составляющей интеллектуального анализа данных. Рис.: 5. Табл.: 3. Библиогр.: 9 назв

UDC 681.58

**Using of the trend geospatial statistical models based on data mining / Putrenko V.V.** // System Research and Information Technologies. — 2015. — № 4. — P. 7–19.

An application of geospatial trending statistical models for analysis the data on the sustainable development of Ukraine was studied. The mathematical apparatus was used in the construction of trend geospatial statistical models. Using the hierarchical model data of sustainable development as an example, a study on geospatial trends and the nature of their origin by constructing polynomial regression models based on the processing of geospatial data was conducted. The method for detecting geospatial trends in hierarchical data models was developed based on cartographic methods and the first and second order polynomials. The use of cartometric methods for estimating the parameters of geospatial trends was proposed. The application of methods for constructing geospatial trends as part of data mining was justified. Fig.: 5. Tabl.: 3. Refs.: 9 titles.

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ,  
ВИСОКОПРОДУКТИВНІ КОМП'ЮТЕРНІ  
СИСТЕМИ

PROGRESSIVE INFORMATION TECHNOLOGIES,  
HIGH-EFFICIENCY COMPUTER SYSTEMS

УДК 519.6+625.1

**Комплексное детерминированное оценивание сложных иерархически-сетевых систем: III. Агрегированное оценивание / Полищук Д.А., Полищук А.Д., Яджак М.С.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 20–31.

Предложена методика комплексного детерминированного оценивания сложных систем с иерархически-сетевой структурой, составляющими которой являются методы локального, прогностического, агрегированного и интерактивного анализа состояния, качества функционирования и взаимодействия объектов, образующих систему. Описаны основные подходы к построению обобщенных выводов о состоянии и качестве функционирования системы разных уровней иерархии: методы «слабейшего» элемента, взвешенной линейной и нелинейной агрегации. Определены области их использования и проанализированы основные преимущества и недостатки. Детально рассмотрены различные направления реализации метода взвешенной линейной агрегации оценок элементов системы и ее базовых подсистем. Предложены принципы визуализации результатов агрегированного оценивания для оперативной локализации структурных элементов системы, получивших отрицательные или близкие к ним выводы. Эффективность предлагаемых методов проиллюстрирована на примере анализа качества объектов путевого и станционного хозяйства железнодорожной транспортной системы Украины. Рис.: 2. Библиогр.: 20 назв.

UDC 519.6+625.1

**Complex deterministic evaluation of the complex hierarchical-network systems: III. Aggregative evaluation / Polishchuk D.O., Polishchuk O.D., Yadzhak M.S.** // *System Research and Information Technologies*. — 2015. — № 4. — P. 20–31.

Methods of complex deterministic evaluation of the complex systems with a hierarchical-network structure, components of which are methods of local, forecasting, aggregative, and interactive estimation of the state, function quality, and interaction of objects comprising the system are proposed. The main approaches to the construction of generalized conclusions about the state and quality of functioning of the system at different levels of the hierarchy (methods of the “weakest” element, weighted linear and nonlinear aggregation) are described. Areas of their use are defined and the main advantages and drawbacks are analyzed. The various directions of implementation of the weighted linear aggregation method for evaluations of elements of the system and its basic subsystems are considered in detail. Principles for visualizing the results of the aggregate evaluation for timely localization of the structural elements of the system which received negative or close to them conclusions are proposed. Effectiveness of the proposed methods is illustrated by the analysis of quality of the objects of the track and station facilities of the rail transport system of Ukraine. Figs.: 2. Refs.: 20 titles.

---

УДК 519.766.4

**Прогнозування індексу споживчих цін в Україні з використанням регресійних моделей і фільтра Калмана / Караюз І.В., Бідюк П.І.** // *Системні дослідження та інформаційні технології*. — 2015. — № 4. — С. 32–38.

Роботу присвячено розв'язанню задачі короткострокового прогнозування індексу споживчих цін в Україні на основі регресійних моделей і адаптивного фільтра Калмана. Побудовано адекватну модель для прогнозування індексу споживчих цін і застосовано адаптивний фільтр Калмана для отримання оптимальних оцінок стану досліджуваного процесу і обчислення короткострокового прогнозу. Основні результати роботи: реалізація і застосування двох модифікацій фільтра Калмана (звичайний та адаптивний), орієнтовані на оцінювання коваріації випадкових збурень стану та похибок вимірів. Альтернативні регресійні моделі та оцінки короткострокових прогнозів, отримані на основі фільтра. Надано порівняльний аналіз отриманих результатів. Для аналізу використано статистичну інформацію перехідної економіки України. Рис.: 5. Табл.: 2. Бібліогр.: 5 назв.

УДК 519.766.4

**Прогнозирование индекса потребительских цен в Украине с использованием регрессионных моделей и фильтра Калмана / Караюз И.В., Бидюк П.И.** // *Системні дослідження та інформаційні технології*. — 2015. — № 4. — С. 32–38.

Работа посвящена решению задачи краткосрочного прогнозирования индекса потребительских цен в Украине на основе регрессионных моделей и адаптивного фильтра Калмана. Построена адекватная модель для прогнозирования индекса потребительских цен и использован адаптивный фильтр Калмана для получения оптимальных оценок состояний исследуемого процесса и краткосрочных прогнозов. Основные результаты работы: две модификации фильтра Калмана (обычный и адаптивный), ориентированные на оценивание ковариации случайных возмущений состояния и погрешностей измерений. Альтернативные регрессионные модели и оценки краткосрочных прогнозов получены при помощи фильтра. Дан сравнительный анализ достигнутых результатов. Для анализа использована статистическая информация переходной экономики Украины. Рис.: 5. Табл.: 2. Библиогр.: 5 назв.

---

УДК 004.932.2

**Мультифрактальный флуктуационный анализ стеганограм / Прогонов Д.О., Куш С.М.** // *Системні дослідження та інформаційні технології*. — 2015. — № 4. — С. 39–47.

Розглянуто використання мультифрактального флуктуационного аналізу для визначення типів перетворень, які були використано у вбудовуванні стегоданих у цифрові зображення. Показано, що застосування як стандартних (дискретне вейвлет перетворення), так і спеціальних (сингулярний розклад) перетворень цифрових зображень у прихованні повідомлень призводить до значних змін фрактальних параметрів флуктуацій яскравості пікселів (ширини мультифрактального спектру). Характерні зміни результатів мультифрактального флуктуационного аналізу стеганограм дають можливість визначати тип перетворень, використаних у ході вбудовування стегоданих в області перетворення контейнеру. Встанов-

лено, що у випадку використання стандартних перетворень, застосування мультифрактального флуктуаційного аналізу дозволяє додатково оцінити ступінь деталізації вбудованих стегоданих. Отримані результати можуть бути використаними для підвищення ефективності сучасних методів активного стегоаналізу зображень. Рис.: 5. Бібліогр.: 10 назв.

UDC 004.932.2

**Multifractal detrended fluctuation analysis of steganograms / Progonov D.A., Kushch S.N.** // System Research and Information Technologies. — 2015. — № 4. — P. 39–47.

The paper investigates the multifractal detrended fluctuation analysis usage for identification the transformations, which have been employed for embedding the stego data in digital images. It is shown that applying both the standard (the discrete wavelet transform) and special (singular value decomposition) transforms for message hiding leads to significant changes of fractal parameters of pixel brightness of images (the width of the multifractal spectrum). Distinctive changes of the multifractal detrended analysis results for steganograms in comparison with unfilled cover images allow to determine the type of a transformation, which has been applied for message hiding in the transform domain of a container. It has been determined that applying the multifractal detrended analysis makes it possible to additionally estimate the level of detail of embedded stegodata in case of using the standard transformations. Obtained results can be used for increasing the efficiency of modern methods of the active steganalysis of digital images. Fig.: 5. Refs.: 10 titles.

---

УДК 004.046: 004.896: 004.942

**Сравнение типов архитектуры систем сервисов / Петренко А.А.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 48–62.

Рассмотрены современные архитектуры систем сервисов SOA (service-oriented architecture — сервисно-ориентированная архитектура) и EDA (event-driven architecture — событийно-ориентированная архитектура), их преимущества и недостатки, возможности и целесообразность построения объединенной сервисно-ориентированной архитектуры EDSOA (event-driven service-oriented architecture — событийно-управляемая сервисно-ориентированная архитектура). Показано, что события соединяют сервисы посредством передачи состояния бизнес-процесса от одного сервиса, который определяет и публикует события, к другим сервисам, которые запускаются конкретными событиями. В свою очередь, обосновано, что сервисы объединяют события посредством передачи данных о переходе одного состояния процесса в другой. Особое внимание в работе уделяется вопросам эффективной реализации предложенного гибридного решения архитектуры EDSOA и его применению для моделирования бизнес-процессов как сервисов. Рис.: 4. Табл.: 1. Библиогр.: 24 назв.

UDC 004.046: 004.896: 004.942

**A comparison of architecture types of services / Petrenko O.O.** // System Research and Information Technologies. — 2015. — № 4. — P. 48–62.

This paper studies modern architectures of services systems SOA (service-oriented architecture) and EDA (event-driven architecture), their advantages and drawbacks, capabilities, and feasibility of constructing the unified service-oriented architecture EDSOA (event-driven service-oriented architecture). It is shown that events connect services by transferring the state of a business process from one service, which define and publish events, to other services, which are started by actual events. In its turn, it is justified, that services unify events by transferring the data about moving one state of a process into another. In this work, a special attention is paid to questions related to an effective implementation of a proposed hybrid solution of architecture EDSOA and its application for modeling business processes as services. Fig.: 4. Tabl.: 1. Refs.: 24 titles.

---

УДК 519.713: 504.064

**Внедрение компараторной идентификации для комплексной оценки уровня безопасности объектов / Козуля Т.В., Шаронова Н.В., Белова М.А., Козуля М.М.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 63–74.

Предложен новый подход к оценке состояния природно-техногенных объектов с позиции постоянного развития, которое базируется на методе компараторной идентификации. Представлено алгоритмическое обеспечение реализации методологии компараторной экологической идентификации по установлению уровня экологичности эколого-социально-экономических систем, разработаны математические модели состояния системных объектов. На основе предложенного метода получен рейтинг эколого-экономических систем по безопасности экологического состояния и комплексную оценку экологичности территории за-

грязнения тяжелыми металлами с установлением механизмов саморегулирования качества систем. Расчеты оценки экологичности исследованных объектов по представленной методике проведены на основе программного обеспечения, которое позволяет автоматизировать анализ состояния сложных систем. Рис.: 3. Табл.: 1. Библиогр.: 13 назв.

UDC 519.713: 504.064

**Comparative identification introduction for integrated assessment of object security level / Kozulia T.V., Sharonova N.V., Bilova M.O., Kozulia M.M.** // System Research and Information Technologies. — 2015. — № 4. — P. 63–74.

A new approach for the natural and anthropogenic objects state-assessment from the continuous development standpoint based on the comparative identification method is proposed in this work. The algorithms for the comparative ecological identification methodology implementation were developed on ecological compatibility leveling of ecological-social-economic systems and mathematical models of the system objects' state. The rating of the ecological state safety for ecology-economical systems and a comprehensive assessment of ecological compatibility for a territory polluted by heavy metals with establishing the self-regulated system quality mechanism were developed on the basis of the presented method. The calculations of ecological compatibility assessment of objects under the study based on presented methods were done using the software that automated the complex systems state analysis. Fig.: 3. Tabl.: 1. Refs.: 13 titles

---

МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ, ПРОБЛЕМИ І  
ТЕХНОЛОГІЇ ДОСЛІДЖЕННЯ СКЛАДНИХ  
СИСТЕМ

MATHEMATICAL METHODS, MODELS,  
PROBLEMS AND TECHNOLOGIES FOR  
COMPLEX SYSTEMS RESEARCH

УДК 510.87:544.431.7:577.21

**Задача оптимального управления стадией элонгации полимеразной цепной реакции / Марценюк В.П., Сверстюк А.С., Гвоздецкая И.С.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 75–82.

Применена общая методология управления для получения решения задачи оптимального течения стадии элонгации в полимеразной цепной реакции. В исследуемой модели стадии элонгации использовано уравнение Аррениуса, которое учитывает зависимость скорости реакции от абсолютной температуры. Данная зависимость может быть использована при исследовании стадии элонгации полимеразной цепной реакции, поскольку управляющим воздействием при этом является именно температура. Рассмотрена схема температурного управления при реализации полимеразной цепной реакции. Применен принцип максимума Понтрягина к задаче оптимального управления и сформулировано необходимое условие оптимальности. Полученные результаты необходимы для численного вычисления оптимального управления исследуемой стадией и помогут минимизировать необходимое время для реализации стадии элонгации, что в общем позволит минимизировать время проведения полимеразной цепной реакции. Рис.1. Библиогр.: 13 назв.

UDC 510.87:544.431.7:577.21

**Optimal control problem of the elongation stage in the polymerase chain reaction / Martseniuk V.P., Sverstiuk A.S., Gvozdetska I.S.** // System Research and Information Technologies. — 2015. — № 4. — P. 75–82.

The general methodology of optimal control for obtaining the solution of the optimal flow problem in the elongation stage of a polymerase chain reaction is applied. The examined model of the elongation stage takes into account the dependence of reaction's rate on the absolute temperature, which is described by the Arrhenius equation. This equation can be used in the investigation of the elongation stage of a polymerase chain reaction, since the temperature controls the process. A scheme of the temperature control in the process of a polymerase chain reaction is examined. Pontryagin's maximum principle for the optimal control problem is used and the necessary condition for optimality is formulated. The obtained results are required for the numerical calculations of optimal control of the examined stage and help to minimize the required duration of the elongation stage, which will allow to minimize the duration of a polymerase chain reaction in general. Fig.: 1. Refs.: 13 titles.

УДК 519.683.8

**Примітивна програмна алгебра: загальний підхід до вирішення проблеми функціональної повноти / Яганов П.О., Редько Д.І., Редько І.В., Захарченко Т.Л.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 83–96.

Основним напрямком дослідження є розробка наукових засад генезису рішень програмістських задач. Проведено побудови, що базуються на алгебраїчних методах дослідження програм та методах композиційного програмування. В основі останніх лежать програмні алгебри з функціями спеціального класу у якості носія, і композиціями, які представляють абстракції інструментів програмного синтезу, у якості операцій. У рамках так званих програмних алгебр строго поставлено та вирішено проблеми отримання характеристик репрезентативних класів обчислюваних функцій, проблеми знаходження породжуючих сукупностей та базисів, що займають одне з чільних місць у програмістській проблематиці. Запропоновано загальний метод вирішення згаданих проблем у примітивних програмних алгебрах (ППА) над різними класами обчислюваних функцій. Отримані результати викладено у вигляді низки оригінальних тверджень, лем та теорем. Вони можуть бути використані у ході дослідження алгебраїчних характеристик різних класів обчислюваних функцій в задачах формалізації семантик мов програмування. Бібліогр.: 21 назва.

УДК 519.683.8

**Примитивная программная алгебра: общий подход к решению проблемы функциональной полноты / Яганов П.А., Редько Д.И., Редько И.В., Захарченко Т.Л.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 83–96.

Основным направлением исследования есть разработка научных основ генезиса решенных программистских задач. Проведены построения, которые базируются на алгебраических методах исследования программ и методах композиционного программирования. В основе последних лежат программные алгебры с функциями специального класса в качестве носителя, и композициями, которые представляют абстракции инструментов программного синтеза, в качестве операций. В рамках так называемых программных алгебр строго поставлены и решены проблемы получения характеристик репрезентативных классов вычислимых функций, проблемы нахождения порождающих совокупностей и базисов, которые занимают одно из главных мест в программистской проблематике. Предложен общий метод решения упомянутых проблем в примитивных программных алгебрах (ППА) над разными классами вычислимых функций. Полученные результаты изложены в виде ряда оригинальных утверждений, лемм и теорем. Они могут быть использованы при исследовании алгебраических характеристик разных классов вычислимых функций в задачах формализации семантик языков программирования. Библиогр.: 21 назв.

---

УДК 519.6

**Аналитическое решение некорректных задач динамическими методами / Пархомчук Д.М., Тимошенко Ю.А.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 97–103.

Аппроксимация непрерывных дифференциальных и интегральных уравнений конечными дискретными алгебраическими системами, локальная линеаризация нелинейных уравнений по заданой информации при решении обратных задач сводятся к задачам решения систем линейных алгебраических уравнений. Матрицы таких систем обычно плохо обусловлены, поэтому задачи их решения некорректны, поскольку нарушается третье условие корректности по Адамару. Для решений некорректных систем предложен динамический метод регуляризации некорректных задач [1]. С целью уменьшения времени работы алгоритма, который предлагается динамическим методом, предложен модифицированный метод — динамический метод второго порядка. Разработан математический аппарат и на его основании предложен алгоритм для модифицированного метода, а также показана его эффективность на практическом примере. Табл.: 1. Библиогр.: 4 назв.

UDC 519.6

**Analytic solution of ill-posed problems via dynamic methods / Parkhomchuk D.M., Tymoshenko Yu.O.** // System Research and Information Technologies. — 2015. — № 4. — P. 97–103.

The inverse problems of continuous differential and integral equations approximation with finite discrete algebraic systems and the problems of local linearization of nonlinear equations by the provided information are reduced to solving the linear algebraic systems. Matrices of such systems are usually ill-conditioned due to ill-posed problems according to Hadamard correctness. As a solu-

tion to these problems a dynamical method for regularization was proposed [1]. In order to reduce the computation time of the algorithm, a second order modification of the dynamical method is proposed. This paper provides mathematical tools based on this method. A practical example shows its effectiveness. Tabl.: 1. Refs.: 4 titles

УДК 303.732.4:519.254

**Усовершенствованные методы расчета статистики Колмогорова–Смирнова, веса категории переменной и значения информации в кредитном рейтинге / Солошенко А.Н.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 104–113.

Разработаны усовершенствованные методы вычисления показателей статистики Колмогорова–Смирнова, веса категории переменной и значения информации без явного разбиения оригинальной выборки на два подмножества, с приведением соответствующих формул для анализа предикативной (прогностической) силы категориальных переменных в задачах кредитного рейтинга и других областях практического применения методов бинарной классификации. Произведено обобщение классических формул статистики Колмогорова–Смирнова, веса категории переменной и показателя значения информации путем преобразования агрегатных выражений для дискретных распределений и кумулятивных функций распределения с применением скалярного произведения векторов и операторов проектирования, а также оператора условной перестановки. Предложены усовершенствованные формулы вычисления статистики Колмогорова–Смирнова, веса категории переменной и индекса значения информации, которые обобщенно описаны в терминах дискретного безусловного распределения входящей переменной и условного распределения бинарной целевой переменной. Рис.: 1. Библиогр.: 8 назв.

UDC 303.732.4:519.254

**Improved estimation methods of the Kolmogorov-Smirnov statistic, weight of evidence and information value indicators in the credit scoring / Soloshenko O.M.** // System Research and Information Technologies. — 2015. — № 4. — P. 104–113.

The improved evaluation methods of the Kolmogorov-Smirnov statistic, Weight of Evidence and Information Value indicators are developed without explicit splitting of the original sample into two subsets with developing corresponding formulas for the predictive (forecasting) power analysis of categorical variables in the credit scoring tasks and other fields of practical application of binary classification methods. The generalization of the classical formulas for the Kolmogorov-Smirnov statistic, Weight of Evidence and Information Value indicators have been performed by means of the aggregate expressions transformation for discrete distributions and cumulative distribution functions applying the inner product of two vectors, projection operators, and also a conditional substitution operator. The improved estimation formulas for the Kolmogorov-Smirnov statistic, Weight of Evidence and Information Value indices are proposed and generally described in terms of the discrete unconditional distribution of the input variable and the conditional distribution of the binary target variable. Fig.: 1. Refs.: 8 titles.

**НОВІ МЕТОДИ В СИСТЕМНОМУ АНАЛІЗІ,  
ІНФОРМАТИЦІ ТА ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ**

**NEW METHODS IN SYSTEM ANALYSIS, COMPUTER  
SCIENCE AND THEORY OF DECISION MAKING**

УДК 519.946; 519.254

**Хаотична та випадкова складові у природних часових даних / Бондаренко В.Г., Трусковський К.К.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 114–122.

Для часових рядів запропоновано метод визначення співвідношення детермінованої та стохастичної складових. Для розв'язку цієї задачі виконано ряд обчислювальних експериментів з використанням імітаційного моделювання логістичної послідовності та значень фрактального броунівського руху із різними показниками Харста —  $H$ . В отриманій адитивній суміші задається співвідношення енергій детермінованої та випадкової складових. Для великих значень показника Харста хаотичний доданок виявляється більш агресивним: контрольні статистики суміші суттєво відрізняються від еталонних значень, що відповідають фрактальному броунівському руху. Для малих значень  $H$  (антиперсистентний випадок) має місце обернений результат. Розглянуто приклади реальних часових даних, що відповідають антиперсистентній моделі. Рис.: 3. Табл.: 4. Бібліогр.: 10.



UDC 519.946; 519.254

**The chaotic and random components in time series data / Bondarenko V.G., Truskovsky K.K.** // System Research and Information Technologies. — 2015. — № 4. — P. 114–122.

We proposed a method for determining the ratio of deterministic and stochastic components for observed real data. We illustrated a number of numerical experiments which used simulation modelling of the logistic chaotic sequence and the values of fractional Brownian motion with different values of Hurst exponent  $H$ . In the additive mixture, the ratio of the energies of deterministic and random components are defined. The chaotic term turns out to be more aggressive for large values of Hurst exponent: the control statistics of the mixture are different from the reference values corresponding to the fractional Brownian motion. Another situation takes place for small values of  $H$  (antipersistent case). The considered examples of time series data are described by an antipersistent model. Figs.: 3. Tabl.: 4. Refs.: 10 titles.

---

УДК 53.01:53.05 + 519.2

**Чому точність вимірювання фізичних величин обмежена / Горбань І.І.** // Системні дослідження та інформаційні технології. — 2015. — № 4. — С. 123–138.

Наведено результати досліджень, що прояснюють механізм обмеження точності вимірювання реальних фізичних величин. Показано, що ключову роль в обмеженні точності відіграє неідеальний характер феномену статистичної сталості, що проявляється у відсутності збіжності (неспроможності) статистичних оцінок. З'ясовано особливості дії закону великих чисел і центральної граничної теореми теорії ймовірностей в умовах порушення статистичної сталості. Результати теоретичних досліджень підтверджено експериментально. Звернено увагу, що реальні оцінки носять гіпервипадковий характер. Гіпервипадкова природа реальних оцінок проявляється при великих об'ємах вибірки. Рис.: 5. Бібліогр.: 27 назв.

UDC 53.01:53.05 + 519.2

**Why the measurement precision of real physical variables is limited / Gorban I.I.** // System Research and Information Technologies. — 2015. — № 4. — P. 123–138.

The research results explaining the restriction causes of measurement precision of real physical variables are presented. It is shown that the imperfect character of the phenomenon of statistical stability that manifests itself in misconvergence (inconsistency) of statistical estimates is a key element in the precision limitation. The peculiarities of the law of average and the central limit theorem of probability theory under conditions of violation of statistical stability are determined. The results of theoretical research are confirmed by experiments. It is drawn to attention that real estimates have hyper-random character. Their hyper-random nature becomes evident when the sample size is large. Figs.: 5. Refs.: 27 titles.

---

## ВІДОМОСТІ ПРО АВТОРІВ

**Бідюк Петро Іванович,**

професор, доктор технічних наук, професор кафедри математичних методів системного аналізу Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ

**Білова Марія Олександрівна,**

магістр кафедри комп'ютерного моніторингу і логістики Національного технічного університету України «Харківський політехнічний інститут», Україна, Харків

**Бондаренко Віктор Григорович,**

професор, доктор фізико-математичних наук, професор кафедри математичних методів системного аналізу Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ

**Гвоздецька Ірина Степанівна,**

кандидат технічних наук, доцент кафедри медичної фізики діагностичного та лікувального обладнання Тернопільського державного медичного університету ім. І.Я. Горбачевського, Україна, Тернопіль

**Горбань Ігор Ілліч,**

професор, доктор технічних наук, головний науковий співробітник Інституту проблем математичних машин і систем НАН України, Київ

**Захарченко Тарас Леонідович,**

аспірант кафедри конструювання електронно-обчислювальної апаратури Національного технічного університету України «КПІ», Київ

**Караюз Ірина Валентинівна,**

аспірантка Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ

**Козуля Марія Михайлівна,**

магістр кафедри комп'ютерного моніторингу і логістики Національного технічного університету України «Харківський політехнічний інститут», Україна, Харків

**Козуля Тетяна Володимирівна,**

доктор технічних наук, професор кафедри комп'ютерного моніторингу і логістики Національного технічного університету України «Харківський політехнічний інститут», Україна, Харків

**Куш Сергій Миколайович,**

доцент, кандидат технічних наук, доцент Фізико-технічного інституту Національного технічного університету України «КПІ», Україна, Київ

**Марценюк Василь Петрович,**

професор, доктор технічних наук, професор Університету Бельсько-Бялої, Польща, професор кафедри медичної інформатики Тернопільського державного медичного університету ім. І. Я. Горбачевського, Україна, Тернопіль

**Пархомчук Дмитро Михайлович,**

аспірант Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, старший аналітик департаменту ринкових, роздрібних та операційних ризиків АТ «Райффайзен банк Аваль», Україна, Київ

**Петренко Олексій Олексійович,**

аспірант Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ

**Поліщук Дмитро Олександрович,**

інженер з програмного забезпечення комп'ютерів відокремленого підрозділу «Інформаційно-обчислювального центру» Державного територіально-галузевого об'єднання «Львівська залізниця», Україна, Львів

**Поліщук Олександр Дмитрович,**

старший науковий співробітник, кандидат фізико-математичних наук, старший науковий співробітник Інституту прикладних проблем механіки і математики ім. Я.С. Підстригача НАН України, Львів

**Прогонов Дмитро Олександрович,**

аспірант Фізико-технічного інституту Національного технічного університету України «КПІ», Україна, Київ

**Путренко Віктор Валентинович,**

кандидат географічних наук, старший науковий співробітник, докторант Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ

**Редько Дмитро Ігорович,**

студент кафедри теорії та технології програмування Київського національного університету імені Тараса Шевченка, Україна, Київ

**Редько Ігор Володимирович,**

професор, доктор фізико-математичних наук, професор кафедри конструювання електронно-обчислювальної апаратури Національного технічного університету України «КПІ», Україна, Київ

**Сверстюк Андрій Степанович,**

доцент, кандидат технічних наук, доцент кафедри медичної інформатики Тернопільського державного медичного університету ім. І.Я. Горбачевського, Україна, Тернопіль

**Солошенко Олександр Миколайович,**

аспірант Національного технічного університету України «КПІ», Україна, Київ

**Тимошенко Юрій Олександрович,**

кандидат технічних наук, доцент кафедри математичних методів системного аналізу Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ

**Трусовський Кирил Константинович,**

студент Навчально-наукового комплексу «Інститут прикладного системного аналізу» НТУУ «КПІ» МОН та НАН України, Київ

**Шаронова Наталія Валеріївна,**

професор, доктор технічних наук, завідувач кафедри інтелектуальних комп'ютерних систем Національного технічного університету України «Харківський політехнічний інститут», Україна, Харків

**Яганов Петро Олексійович,**

доцент, кандидат технічних наук, доцент кафедри конструювання електронно-обчислювальної апаратури Національного технічного університету України «КПІ», Україна, Київ

**Яджак Михайло Степанович,**

старший науковий співробітник, доктор фізико-математичних наук, провідний науковий співробітник Інституту прикладних проблем механіки і математики ім. Я.С. Підстригача НАН України, Львів

Зміст журналу  
«Системні дослідження та інформаційні технології»  
за 2015 р.

**ЗМІСТ № 1**

<i>Згуровский М.З., Петренко А.И.</i> Цифровая наука в программе «Горизонт 2020» .....	7
<i>Полищук Д.О., Полищук О.Д., Яджак М.С.</i> Комплексне детерміноване оцінювання складних ієрархічно-мережових систем: Частина I. Опис методики .....	21
<i>Зімчук І.В., Іщенко В.І., Канкін І.О.</i> Синтез алгоритмів цифрового управління для автоматичних слідкувальних систем .....	32
<i>Смирнов С.А., Терещенко І.М.</i> Процедура обчислення внутрішньої валюти в рефлексивних іграх .....	39
<i>Махорт А.П.</i> Про рівновагу відкритої економічної системи за наявності невикористаного капіталу та заданих рівнів споживання .....	45
<i>Коваль А.В., Бойко Ю.Д., Волкова Е.А.</i> Особенности сценарно-целевого подхода к анализу объектов действенной аналитики .....	57
<i>Зайченко Е.Ю.</i> Оптимальная ремаршрутизация потоков в компьютерных сетях с технологией MPLS .....	68
<i>Олішук А.А., Subbotin S.A.</i> Neural network synthesis based on evolutionary optimization .....	77
<i>Щербовських С.В.</i> Математична модель надійності для аналізу причин непрацездатності системи із роздільним навантажувальним резервуванням .....	87
<i>Гарт Л.Л., Манойло М.В.</i> О некоторых алгоритмах регуляризации для решения интегральных уравнений .....	99
<i>Зак Ю.А.</i> Альтернативные алгоритмы дефазификации .....	111
<i>Романенко В.Д., Милявский Ю.Л.</i> Управление соотношениями координат когнитивной модели сложной системы при неустойчивом импульсном процессе .....	121
<i>Бомба А.Я., Сінчук А.М., Яроцак С.В.</i> Методи комплексного аналізу при моделюванні процесу витіснення нафти теплоносієм із врахуванням ефекту гідророзриву .....	130
Реферати .....	141
Відомості про авторів .....	149

**ЗМІСТ № 2**

<i>Згуровский М.З., Павлов А.А., Халус Е.А.</i> Задача построения допустимого расписания с максимально поздним моментом запуска и минимальным суммарным опережением .....	7
<i>Глибовець М.М., Зінчук С.О.</i> Використання моделі акторів для реалізації розподілених генетичних алгоритмів .....	16
<i>Полищук Д.О., Полищук О.Д., Яджак М.С.</i> Комплексне детерміноване оцінювання складних ієрархічно-мережових систем. Частина II. Локальне та прогностичне оцінювання .....	26
<i>Лавданский А.А., Фауре Э.В.</i> Оценка статистических свойств последовательностей на выходе комбинационного генератора с помощью графических тестов .....	39
<i>Іваненко В.І., Пасічніченко І.О.</i> Очікувана корисність у ситуаціях прийняття рішень з випадковими у широкому сенсі наслідками .....	51
<i>Ови Нафас Агаи аг Гамши, Зайченко Ю.П., Войтенко О.С.</i> Анализ финансового состояния и прогнозирование риска банкротства банков .....	59
<i>Петренко А.А.</i> Цели и объекты науки о сервисах .....	75
<i>Гоголадзе Н.Г.</i> Моделирование и прогнозирование процесса инфляции в Грузии .....	83
<i>Панкратова Н.Д., Бузань И.В., Дашук В.А.</i> Восстановления функциональных закономерностей на основе многочленов Гегенбауэра .....	88
<i>Ротштейн О.П., Ракитянська Г.Б.</i> Діагностика на основі багатовимірних нечітких відношень .....	97
<i>Podolsky S.V., Zorin Yu.M.</i> O(1) delta part computation technique for the quadratic assignment problem .....	112
<i>Савченко І.А.</i> Эволюция объекта исследования с привлечением модифицированного метода морфологического анализа .....	122
<i>Левченко Є.Г., Рабчун Д.І.</i> Кореляція витрат в багаторубіжних системах захисту інформації .....	131
Реферати .....	141
Відомості про авторів .....	149

### ЗМІСТ № 3

<i>Панкратова Н.Д., Бідюк П.І., Рубець М.Г.</i> Інформаційна система для моделювання та оцінювання фінансових операційних ризиків за допомогою байєсівської мережі .....	7
<i>Путренко В.В.</i> Системні основи інтелектуального аналізу геопросторових даних .....	20
<i>Гаращенко Ф.Г., Кулян В.Р., Юнькова О.О.</i> Ідентифікація параметрів моделей динаміки активів .....	34
<i>Романенков Ю.О., Зейнієв Т.Г.</i> Завдання контуру стратегічного управління ефективністю бізнес-процесів в організації .....	43
<i>Трофимчук А.Н., Васянин В.А.</i> Время работы алгоритма Краскала с древовидной и списочной структурой данных .....	48
<i>Радовановіч М., Виклюк Я.І., Міленковіч М., Йовановіч А., Вуковіч Д., Стеванчевіч М., Мацюк Н.О., Леко Т.Б.</i> Концептуальне моделювання спалахів лісових пожеж на основі онтологічного підходу DataMining. Частина 2 .....	62
<i>Качинский А.Б., Азаркова Н.В.</i> Характер связанности элементов системы обеспечения безопасности гидротехнических сооружений .....	72
<i>Matsuki Y., Bidiuk P.I., Danylov V.J., Yevtushenko K.I.</i> Expected effect of environmental investment on market share of oil company .....	84
<i>Зражевська Н.Г.</i> Метод згладженої автокореляційної функції для прогнозування варіації гетероскедастичних часових рядів .....	97
<i>Романенко В.Д., Милявський Ю.Л.</i> Адаптивное координирующее управление соотношениями координат вершин взаимодействующих когнитивных карт в режиме импульсных процессов .....	109
<i>Недашкова Н.И.</i> Построение доверительных интервалов для весов альтернатив решений на основе экспертных оценок парных сравнений .....	121
<i>Братусь О.В., Подладчиков В.М.</i> Ідентифікація змінних параметрів моделі для побудови алгоритму прогнозування .....	131
Реферати .....	142
Відомості про авторів .....	149

### ЗМІСТ № 4

<i>Путренко В.В.</i> Застосування трендових геопросторових статистичних моделей на основі інтелектуального аналізу даних .....	7
<i>Поліщук Д.О., Поліщук О.Д., Яджак М.С.</i> Комплексне детерміноване оцінювання складних ієрархічно-мережевих систем: III. Агреговане оцінювання .....	20
<i>Karayuz I.V., Bidiuk P.I.</i> Forecasting consumer price index in Ukraine with regression models and adaptive Kalman filter .....	32
<i>Прогонов Д.А., Куц С.Н.</i> Мультифрактальный флуктуационный анализ стеганограмм ...	39
<i>Петренко О.О.</i> Порівняння типів архітектури систем сервісів .....	48
<i>Козуля Т.В., Шаронова Н.В., Білова М.О., Козуля М.М.</i> Впровадження компараторної ідентифікації для комплексної оцінки рівня безпеки об'єктів .....	63
<i>Марценюк В.П., Сверстюк А.С., Гвоздецька І.С.</i> Задача оптимального керування стадією елонгації полімеразно-ланцюгової реакції .....	75
<i>Yahanov P.O., Redko D.I., Redko I.V., Zakharchenko T.L.</i> Primitive programing algebra: General approach to a problem of functional completeness .....	83
<i>Пархомчук Д.М., Тимошенко Ю.О.</i> Аналітичний розв'язок некоректних задач динамічними методами .....	97
<i>Солошенко О.М.</i> Вдосконалені методи розрахунку статистики Колмогорова–Смирнова, ваги категорії змінної та значення інформації у кредитному рейтингу ....	104
<i>Бондаренко В.Г., Трусковский К.К.</i> Хаотическая и случайная составляющие в природных временных данных .....	114
<i>Горбань И.И.</i> Почему точность измерения физических величин ограничена .....	123
Реферати .....	139
Відомості про авторів .....	146
Зміст журналу «Системні дослідження та інформаційні технології» за 2015 р. ....	148
Автори статей за 2015 р. ....	150

## АВТОРИ СТАТЕЙ ЗА 2015 р.

Агаркова Наталія Володимирівна, № 3  
Бідюк Петро Іванович, № 3, 4  
Білова Марія Олександрівна, № 4  
Бойко Юрій Дмитрович, № 1  
Бомба Андрій Ярославович, № 1  
Бондаренко Віктор Григорович, № 4  
Братусь Олена Вікторівна, № 3  
Бузань Ігор Валентинович, № 2  
Васянін Володимир Олександрович, № 3  
Виклюк Ярослав Ігорович, № 3  
Войтенко Ольга Сергіївна, № 2  
Волкова Катерина Алімівна, № 1  
Вукович Дарко, № 3  
Гарашенко Федір Георгійович, № 3  
Гарт Людмила Лаврентіївна, № 1  
Гвоздецька Ірина Степанівна, № 4  
Глибовець Микола Миколайович, № 2  
Гоголадзе Ніно Гурамівна, № 2  
Горбань Ігор Ілліч, № 4  
Данилов Валерій Якович, № 3  
Дашук Валерій Олександрович, № 2  
Євтушенко Катерина Ігорівна, № 3  
Зайченко Олена Юріївна, № 1  
Зайченко Юрій Петрович, № 2  
Зак Юрій Олександрович, № 1  
Захарченко Тарас Леонідович, № 4  
Згуровський Михайло Захарович, № 1, 2  
Зейнієв Теймур Гидаятович, № 3  
Зімчук Ігор Валерійович, № 1  
Зінчук Сергій Олександрович, № 2  
Зорін Юрій Михайлович, № 2  
Зражевська Наталія Григорівна, № 3  
Іваненко Віктор Іванович, № 2  
Іщенко Володимир Іванович, № 1  
Йованович Ана, № 3  
Канкін Іван Олегович, № 1  
Караюз Ірина Валентинівна, № 4  
Качинський Анатолій Броніславович, № 3  
Коваль Олександр Васильович, № 1  
Козуля Марія Михайлівна, № 4  
Козуля Тетяна Володимирівна, № 4  
Кулян Віктор Романович, № 3  
Куц Сергій Миколайович, № 4  
Лавданський Артем Олександрович, № 2  
Левченко Євген Григорович, № 2  
Леко Тарас Богданович, № 3  
Манойло Марія Володимирівна, № 1  
Марценюк Василь Петрович, № 4  
Махорт Андрій Пилипович, № 1  
Мацукі Йошіо, № 3  
Мацюк Наталія Олександрівна, № 3  
Міленкович Мілан, № 3  
Мілявський Юрій Леонідович, № 1, 3  
Недашківська Надія Іванівна, № 3  
Ові Нафас Агаі аг Гаміш, № 2  
Олійник Андрій Олександрович, № 1  
Павлов Олександр Анатолійович, № 2  
Панкратова Наталія Дмитріївна, № 2, 3  
Пархомчук Дмитро Михайлович, № 4  
Пасічніченко Ілля Олександрович, № 2  
Петренко Анатолій Іванович, № 1  
Петренко Олексій Олексійович, № 2, 4  
Подладчиков Володимир Миколайович, № 3  
Подольський Сергій Валентинович, № 2  
Поліщук Дмитро Олександрович, № 1, 2, 4  
Поліщук Олександр Дмитрович, № 1, 2, 4  
Прогинов Дмитро Олександрович, № 4  
Путренко Віктор Валентинович, № 3, 4  
Рабчун Дмитро Ігорович, № 2  
Радованович Мілан, № 3  
Ракитянська Ганна Борисівна, № 2  
Редько Дмитро Ігорович, № 4  
Редько Ігор Володимирович, № 4  
Романенко Віктор Демидович, № 1, 3  
Романенков Юрій Олександрович, № 3  
Ротштейн Олександр Петрович, № 2  
Рубець Микола Григорович, № 3  
Савченко Ілля Олександрович, № 2  
Сверстюк Андрій Степанович, № 4  
Сінчук Алесь Михайлівна, № 1  
Смирнов Сергій Анатолійович, № 1  
Солошенко Олександр Миколайович, № 4  
Стеванчевіч Мілан, № 3  
Субботін Сергій Олександрович, № 1  
Терещенко Іван Миколайович, № 1  
Тимошенко Юрій Олександрович, № 4  
Трофімчук Олександр Миколайович, № 3  
Трусовський Кирил Константинович, № 4  
Фауре Еміль Віталійович, № 2  
Халус Олена Андріївна, № 2  
Шаронова Наталія Валеріївна, № 4  
Щербовських Сергій Володимирович, № 1  
Юнькова Олена Олександрівна, № 3  
Яганов Петро Олексійович, № 4  
Яджак Михайло Степанович, № 1, 2, 4  
Ярошак Сергій Вікторович, № 1