

Запорізький національний технічний університет



Радіоелектроніка Інформатика Управління

2(23)'2010

Науковий журнал

Виходить двічі на рік

Видається з березня 1999 року

Зареєстрований **29 січня 2003 року**
Державним комітетом інформаційної політики,
телебачення та радіомовлення України.

Свідоцтво – серія **КВ № 6904**

Засновник і видавник – *Запорізький національний технічний університет*

Запоріжжя, ЗНТУ
2010

ISSN 1607-3274

Постановою президії ВАК України № 1-05/4 від 26.05.2010 р. журнал «Радіоелектроніка, інформатика, управління» (скорочена назва – РІУ), який видається з 1999 року, включений до переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата технічних наук та фізико-математичних наук (фізика).

Журнал є донором журналу «Telecommunications and Radio Engineering», який видається в США. Інтернет-сторінка журналу: <http://journal.zntu.edu.ua/ric/index.php?page=index>.

Статті, що публікуються в журналі, реферуються в базах даних та РЖ ВІНІТІ (Росія) і «Джерело» (Україна). Журнал РІУ міститься у міжнародній базі наукових видань Index Copernicus (<http://journals.indexcopernicus.com/index.php>), електронна копія журналу розміщена на сайті Національної бібліотеки України імені В. І. Вернадського НАН України у розділі «Наукова періодика України» за адресою: <http://nbuv.gov.ua/portal/>.

Журнал розповсюджується за Каталогом періодичних видань України (передплатний індекс – 22914).

РЕДАКЦІЙНА КОЛЕГІЯ

Головний редактор – д-р техн. наук Піза Д. М.

Заст. головного редактора – канд. техн. наук Дубровін В. І.

Члени редколегії :

д-р фіз.-мат. наук Ахметшин А. М.
д-р техн. наук Волков О. В.
д-р фіз.-мат. наук Горбань О. М.
д-р фіз.-мат. наук Горр Г. В.
д-р техн. наук Гостев В. І.
д-р фіз.-мат. наук Дробахин О. О.
д-р техн. наук Карпуков Л. М.
д-р фіз.-мат. наук Корніч Г. В.

д-р фіз.-мат. наук Матюшин В. М.
д-р фіз.-мат. наук Онуфрієнко В. М.
д-р фіз.-мат. наук Погосов В. В.
д-р техн. наук Потапенко Є. М.
д-р техн. наук Толок В. О.
д-р техн. наук Труфанов І. Д.
д-р фіз.-мат. наук Чумаченко В. П.

Рекомендовано до видання вченою радою Запорізького національного технічного університету, протокол № 9 від 26.04.2010 р.

Рукописи проходять незалежне рецензування з залученням провідних фахівців, за результатами якого редакційна колегія приймає рішення про опублікування

Журнал зверстаний редакційно-видавничим відділом
Запорізького національного технічного університету

Адреса редакції: 69063, м. Запоріжжя, вул. Жуковського, 64, ЗНТУ,
редакція журналу «РІУ»

Тел: (061) 769-82-96 – редакційно-видавничий відділ
Факс: (061) 764-21-41
E-mail: rvv@zntu.edu.ua

ЗМІСТ

РАДІОФІЗИКА 7

Vashchenko V. V., Chumachenko V. P.
SCATTERING MATRIX OF A JUMP OF MEDIUM
PARAMETERS IN A RECTANGULAR WAVEGUIDE 7

Прокопенко О. В.
ВЗАЄМНА ФАЗОВА СИНХРОНІЗАЦІЯ ДВОХ
НАНОРОЗМІРНИХ МАГНІТНИХ СТРУКТУР
З УРАХУВАННЯМ ЗАПІЗНЕННЯ СИГНАЛУ 11

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ 16

Артюшенко Б. А., Ілляшенко М. Б., Головатий А. І.
ВИЗНАЧЕННЯ ВПЛИВУ ВІДХИЛЕНЬ ПАРАМЕТРІВ
АНТЕНИ НА ХАРАКТЕРИСТИКИ ДІАГРАМИ
НАПРЯМЛЕНОСТІ МЕТОДОМ ПАРАЛЕЛЬНОГО
АНАЛІЗУ ГРАНЕЙ 16

Матюшин В. М., Жавжаров Є. Л.
ВПЛИВ АТОМАРНОГО ВОДНЮ НА СТРУКТУРИ
МЕТАЛ – НАПІВПРОВІДНИК. ВОЛЬТ-АМПЕРНІ
ХАРАКТЕРИСТИКИ 37

Гулин С. П., Гулин А. С., Дмитренко В. П., Пиза Д. М.
КОНЦЕПЦІЯ УПРАВЛЯЕМОГО ДИНАМИЧЕСКОГО
НАСЫЩЕНИЯ ПРИ МОДЕЛИРОВАНИИ
МНОГОЧАСТОТНЫХ УСТАНОВИВШИХСЯ РЕЖИМОВ
МИКРОЭЛЕКТРОННЫХ КОМПОНЕНТОВ, СТРУКТУР
И УСТРОЙСТВ 20

Шило Г. Н., Огреніч Е. В., Филобок А. В., Гапоненко Н. П.
ПРОЕКТИРОВАНИЕ РЕБРИСТЫХ ТЕПЛОТВОДОВ
МИНИМАЛЬНОЙ МАССЫ 42

Жавжаров Є. Л., Матюшин В. М.
ВПЛИВ АТОМАРНОГО ВОДНЮ НА СТРУКТУРИ
МЕТАЛ – НАПІВПРОВІДНИК. АДГЕЗІЯ ТОНКИХ
МЕТАЛЕВИХ ПЛІВОК 32

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ 48

Агеев Д. В., Копылев А. Н.
МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ ПОТОКОВ
В МУЛЬТИСЕРВИСНОЙ СЕТИ NGN ПРИ РЕШЕНИИ
ЗАДАЧ ПАРАМЕТРИЧЕСКОГО СИНТЕЗА 48

Овский А. Г., Толок В. А.
ИНСТРУМЕНТАЛЬНАЯ СИСТЕМА РЕШЕНИЯ
ДВУМЕРНЫХ И ТРЕХМЕРНЫХ ЗАДАЧ ТЕОРИИ
УПРУГОСТИ НА БАЗЕ MAPLE 74

Barkalov A. A., Zelenyova I. Y., Miroshkin A. N.
EXTENSION OF MICROINSTRUCTION FORMAT
IN COMPOSITIONAL MICROPROGRAM CONTROL
UNIT WITH ELEMENTARIZATION
OF OPERATIONAL LINEAR CHAINS 53

Олійник А. О.
МУЛЬТИАГЕНТНИЙ МЕТОД З ОПОСЕРЕДКОВАНИМ
ЗВ'ЯЗКОМ МІЖ АГЕНТАМИ 82

Бойченко О. В.
КООРДИНАЦІЯ НЕЧІТКИХ РІШЕНЬ
В БАГАТОРІВНЕВІЙ ІЄРАРХІЧНІЙ СИСТЕМІ 57

Таланін В. І., Таланін І. Е., Семікіна М. Ю.
КОМП'ЮТЕРНОЕ МОДЕЛИРОВАНИЕ ОБРАЗОВАНИЯ
ПЕРВИЧНЫХ РОСТОВЫХ МИКРОДЕФЕКТОВ
В БЕЗДИСЛОКАЦИОННЫХ МОНОКРИСТАЛЛАХ
КРЕМНИЯ 89

Дубинская Н. Г.
СТРУКТУРНАЯ ДИАГНОСТИЧЕСКАЯ МОДЕЛЬ
КОМП'ЮТЕРНОЙ СЕТИ 61

Чуб І. А., Новожилова М. В.
ОПРЕДЕЛЕНИЕ НАПРАВЛЕНИЯ СПУСКА
В ЛИНЕАРИЗОВАННОЙ ЗАДАЧЕ РАЗМЕЩЕНИЯ
НЕОРИЕНТИРОВАННЫХ ГЕОМЕТРИЧЕСКИХ
ОБЪЕКТОВ 94

Кривуля Г. Ф., Сыревич Е. Е., Карасев А. Л.
АНАЛИЗ СПИСКА СОЕДИНЕНИЙ В СИСТЕМАХ
ЛОГИЧЕСКОГО СИНТЕЗА 65

Нелася А. В., Верещак М. І.
ОЦЕНКА ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ
БИБЛИОТЕК ДЛИННОЙ АРИФМЕТИКИ
В КРИПТОГРАФИЧЕСКИХ ПРИЛОЖЕНИЯХ 67

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ 101

Авраменко В. П., Ткаченко В. Ф., Серета Л. Б.
ПРИНЯТИЕ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ
В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ И НЕЧЕТКОСТИ
ИСХОДНОЙ ИНФОРМАЦИИ 101

Асеев Г. Г.
МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОЙ ПРЕДОБРАБОТКИ
ДАННЫХ В ЭЛЕКТРОННЫХ ХРАНИЛИЩАХ 106

Бодянский Е. В., Кулишова Н. Е.
СГЛАЖИВАНИЕ ПОМЕХ НА ЦВЕТНЫХ
ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ НЕЙРО-ФАЗЗИ
ФИЛЬТРА..... 111

Зайцев С. А., Субботин С. А.
МЕТОДЫ И МОДЕЛИ АВТОМАТИЧЕСКОЙ
КЛАССИФИКАЦИИ ОБЪЕКТОВ ПО ПРИЗНАКАМ
НА ОСНОВЕ ИММУНОКОМПЬЮТИНГА..... 117

Халимов Г. З.
ОЦЕНКА ЧИСЛА РЕШЕНИЙ УРАВНЕНИЯ ГУРВИЦА
В КОНЕЧНОМ ПОЛЕ..... 124

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ..... 130

Брагина Т. И., Табуницик Г. В.
СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИТЕРАТИВНЫХ
МОДЕЛЕЙ РАЗРАБОТКИ ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ 130

Высочина О. С., Шматков С. И., Салман Амер Мухсин
АНАЛИЗ СИСТЕМ МОНИТОРИНГА
ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЕЙ..... 139

*Дубровин В. И., Афонин Ю. С., Манило Т. В.,
Харитонов В. Н.*
ВЕЙВЛЕТ-АНАЛИЗ КРАТКОВРЕМЕННЫХ
ИМПУЛЬСОВ В СИГНАЛАХ С ДАТЧИКОВ
ЧАСТОТЫ ВРАЩЕНИЯ РОТОРА ГТД 142

Корниенко С. К.
АВТОМАТИЗИРОВАННАЯ ЛОГИСТИЧЕСКАЯ
СИСТЕМА..... 148

Шкарупило В. В., Касьян К. Н., Маркин А. Г.
ИССЛЕДОВАНИЕ С ПОМОЩЬЮ СТАТИСТИЧЕСКИХ
ПРОГРАММ ЦЕЛЕСООБРАЗНОСТИ ВНЕДРЕНИЯ
СТАНДАРТА 10G-ETHERNET
В КЛАСТЕРАХ ВУЗОВ..... 153

Юр Т. В., Дубровин В. И., Харитонов В. Н.
ИССЛЕДОВАНИЕ ФИЛЬТРУЮЩИХ СВОЙСТВ
ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ..... 157

УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ..... 166

Гостев В. И.
ПРОЕКТИРОВАНИЕ НЕЧЕТКОГО РЕГУЛЯТОРА
С УНИВЕРСАЛЬНЫМИ ИДЕНТИЧНЫМИ
ВХОДНЫМИ И ВЫХОДНЫМИ ФУНКЦИЯМИ
ПРИНАДЛЕЖНОСТИ 166

CONTENTS

RADIOPHYSICS	7
<i>Vashchenko V. V., Chumachenko V. P.</i> SCATTERING MATRIX OF A JUMP OF MEDIUM PARAMETERS IN A RECTANGULAR WAVEGUIDE	7
<i>Prokopenko O. V.</i> MUTUAL PHASE-LOCKING OF TWO NANO-SIZE MAGNETIC STRUCTURES CONSIDERING AN INFLUENCE OF SIGNAL DELAY	11
RADIO ELECTRONICS AND TELECOMMUNICATIONS	16
<i>Artyushenko B. A., Il'yashenko M. B., Holovatyj A. I.</i> DETERMINATION OF ANTENNA PARAMETERS DEVIATION INFLUENCE ON ANTENNA PATTERN CHARACTERISTICS USING THE METHOD OF PARALLEL BOUNDS ANALYSIS	16
<i>Matyushin V. M., Zhavzharov E. L.</i> THE INFLUENCE OF ATOMIC HYDROGEN ON METAL-SEMICONDUCTOR STRUCTURES. VOLTAGE-CURRENT CHARACTERISTICS	37
<i>Gulin S. P., Gulin A. S., Dmitrenko V. P., Piza D. M.</i> CONCEPT OF CONTROLLED DYNAMIC SATURATION FOR MODELING MULTIFREQUENCY STEADY STATE MODES OF MICROELECTRONIC COMPONENTS, STRUCTURES AND DEVICES	20
<i>Shilo G. N., Ogrenich E. V., Filobok A. V., Gaponenko N. P.</i> DESIGNING OF RIBBED HEAT SINKS HAVING MINIMUM MASS	42
<i>Zhavzharov E. L., Matyushin V. M.</i> THE INFLUENCE OF ATOMIC HYDROGEN ON METAL – SEMICONDUCTOR STRUCTURES. ADHESION OF THIN METAL FILMS	32
MATHEMATICAL AND COMPUTER MODELLING	48
<i>Ageyev D. V., Kopyliev A. N.</i> INFORMATION STREAMS MODELING IN MULTISERVICE NETWORK NGN FOR PARAMETRICAL SYNTHESIS PROBLEM SOLVING	48
<i>Ovsky A., Tolok V.</i> DEVELOPMENT SYSTEM FOR SOLVING TWO-DIMENSIONAL AND THREE-DIMENSIONAL ELASTICITY THEORY PROBLEMS BASED ON MAPLE	74
<i>Barkalov A. A., Zelenyova I. Y., Miroshkin A. N.</i> EXTENSION OF MICROINSTRUCTION FORMAT IN COMPOSITIONAL MICROPROGRAM CONTROL UNIT WITH ELEMENTARIZATION OF OPERATIONAL LINEAR CHAINS	53
<i>Oliinyk A. O.</i> MULTIAGENT METHOD WITH INDIRECT COUPLING BETWEEN AGENTS	82
<i>Boychenko O. V.</i> CO-ORDINATION OF FUZZY DECISIONS IN MULTILEVEL HIERARCHICAL SYSTEM	57
<i>Talanin V. I., Talanin I. E., Semikina M. Yu.</i> COMPUTER MODELING OF PRIMARY GROW-IN MICRODEFECTS FORMATION IN DISLOCATION-FREE SILICON MONOCRYSTALS	89
<i>Dybinsky N. G.</i> STRUCTURAL DIAGNOSTIC MODEL OF COMPUTER NETWORK	61
<i>Chub I. A., Novozhilova M. V.</i> DETERMINATION OF DESCENT DIRECTION IN LINEARIZED PROBLEM OF NON-ORIENTED GEOMETRIC OBJECTS ARRANGEMENT	94
<i>Krivulya G. F., Syrevich Yev. Yu., Karasyov A. L.</i> NETLIST ANALYSIS IN LOGICAL SYNTHESIS SYSTEMS	65
<i>Nelasa G. V., Vereschak M. I.</i> EFFICIENCY EVALUATION OF LONG NUMBER LIBRARIES IN CRYPTOGRAPHY APPLICATIONS	67

NEUROINFORMATICS AND INTELLIGENT SYSTEMS..... 101

Avramenko V. P., Tkachenko V. P., Sereda L. B.
MANAGEMENT DECISIONS IN CONDITIONS OF
UNCERTAIN AND FUZZY SOURCE INFORMATION..... 101

Aseyev G. G.
METHODS OF INTELLIGENT INPUT DATA
PRELIMINARY PROCESSING
IN DATA WAREHOUSE 106

Bodyanskiy Ye. V., Kulishova N. Ye.
COLOR IMAGES DISTURBANCE SMOOTHING USING
NEURO-FUZZY FILTER 111

Zaitsev S. A., Subbotin S. A.
METHODS AND MODELS OF AUTOMATED OBJECTS
CLASSIFICATION BY FEATURES BASED ON
IMMUNOCOMPUTING 117

Khalimov G. Z.
ESTIMATE OF HURVITZ EQUATION SOLUTIONS
NUMBER IN FINITE FIELD 124

PROGRESSIVE INFORMATION TECHNOLOGIES 130

Bragina T. I., Tabunshchik G. V.
COMPARATIVE ANALYSIS OF ITERATIVE SOFTWARE
DEVELOPMENT MODEL 130

Vysochina O. S., Shmatkov S. I., Salman Amer M.
ANALYSIS OF TELECOMMUNICATIONS NETWORKS
MONITORING SYSTEMS 139

Dubrovin V. I., Afonin Y. S., Manilo T. V., Haritonov V. N.
WAVELET-ANALYSIS OF SHORT-TIME PULSES
IN SIGNALS OF TURBINE ENGINE ROTOR
VELOCITY SENSORS..... 142

Kornienko S. K.
AUTOMATIC LOGISTICS SYSTEM 148

Shkarupylo V. V., Kasyan K. N., Markin A. G.
THE RESEARCH OF 10G-ETHERNET STANDARD
IMPLEMENTATION IN UNIVERSITIES' CLUSTERS USING
STATISTICAL SOFTWARE..... 153

Yur T. V., Dubrovin V. I., Kharitonov V. N.
ANALYSIS OF WAVELET TRANSFORM FILTERING
PROPERTIES 157

CONTROL IN TECHNICAL SYSTEMS 166

Gostev V. I.
DESIGNING OF A FUZZY CONTROLLER WITH
UNIVERSAL IDENTICAL INPUT AND OUTPUT MEMBER-
SHIP FUNCTIONS 166

РАДІОФІЗИКА

РАДИОФИЗИКА

RADIOPHYSICS

УДК 537.87 : 621.372

Vashchenko V. V.¹, Chumachenko V. P.²¹Assistant of Zaporizhzhya National Technical University²Doctor of Physical and Mathematical Sciences, Head of Department of Zaporizhzhya National Technical University

SCATTERING MATRIX OF A JUMP OF MEDIUM PARAMETERS IN A RECTANGULAR WAVEGUIDE

For a rectangular waveguide, a generalized scattering matrix of a jump of medium parameters has been obtained under condition that the field has one variation along the direction perpendicular to E -plane. The configuration is considered as a model of a plane interface in the theory of piecewise homogeneous E -plane structures.

Key words: waveguide discontinuities, plane interface, scattering matrix.

INTRODUCTION

The theory of homogeneous E -plane waveguide junctions is similar in many respects to that of H -plane structures. Under assumption that excitation is the TE_{10} mode, a scalar formulation in terms of H_z (the component of the magnetic field perpendicular to the plane of the junction) is possible. The analysis becomes more complicated for configurations involving dielectric obstacles. Even in the case of the material parameters invariable in the z direction, the same component of the electric field arises, despite the fact that E_z is not present in the incident wave [1, 2]. The problem remains two-dimensional (the z dependence is known), but the field in the structure is expressed through two coupled scalar functions. There are only a few works (e.g. [2, 3]) giving the complete E -plane analysis of discontinuities of this type.

A particular class of E -plane junctions with dielectric inclusions is presented by configurations consisting of homogeneous parts separated by the planes parallel to the z axis or, what is same, connected by waveguides of zero length containing interfaces. The method of generalized scattering matrices [4] can give certain advantages when dealing with this category of problems. In regions between the planes, the field is representable as a

superposition of LSE and LSM modes. Since the LSE and LSM waves do not interact inside a homogeneous cell, the S -matrix of such unit is a simple union of scattering matrices obtained in solving two uncoupled scalar problems. That allows not treating a vector problem for complicated geometry. The S -matrix of the interface, situated in a connecting waveguide, can be got analytically. The objective of the present paper is to determine the said matrix. In spite of the simplicity of the problem we have not found its solution in previous publications.

PROBLEM FORMULATION

It is assumed that the medium is isotropic, boundaries of the junction are perfectly conducting and the fields have only one variation along a , the broad dimension of the waveguides. Then, omitting the time dependence $\exp(i\omega t)$, we can represent vectors \mathbf{E} and \mathbf{H} in an E -plane structure as

$$\mathbf{E} = \mathbf{E}_t(x, y) \sin \frac{\pi z}{a} + \hat{\mathbf{z}} E_z(x, y) \cos \frac{\pi z}{a}, \quad (1)$$

$$\mathbf{H} = \mathbf{H}_t(x, y) \cos \frac{\pi z}{a} + \hat{\mathbf{z}} H_z(x, y) \sin \frac{\pi z}{a}, \quad (2)$$

where $\hat{\mathbf{z}}$ is the unit vector in the z direction. With the help of Maxwell equations, the transverse with respect

to z components (denoted as t) are expressed through E_z and H_z :

$$\begin{aligned} \mathbf{E}_t &= (E_x, E_y) = \\ &= -\frac{1}{\chi^2} \left[i\omega\mu\mu_0 \left(\hat{\mathbf{x}} \frac{\partial H_z}{\partial y} - \hat{\mathbf{y}} \frac{\partial H_z}{\partial x} \right) + \frac{\pi}{a} \left(\hat{\mathbf{x}} \frac{\partial E_z}{\partial x} + \hat{\mathbf{y}} \frac{\partial E_z}{\partial y} \right) \right], \quad (3) \end{aligned}$$

$$\begin{aligned} \mathbf{H}_t &= (H_x, H_y) = \\ &= \frac{1}{\chi^2} \left[i\omega\varepsilon\varepsilon_0 \left(\hat{\mathbf{x}} \frac{\partial E_z}{\partial y} - \hat{\mathbf{y}} \frac{\partial E_z}{\partial x} \right) + \frac{\pi}{a} \left(\hat{\mathbf{x}} \frac{\partial H_z}{\partial x} + \hat{\mathbf{y}} \frac{\partial H_z}{\partial y} \right) \right], \quad (4) \end{aligned}$$

with

$$\chi^2 = k^2 - \left(\frac{\pi}{a} \right)^2, \quad (5)$$

$$k = k_0 \sqrt{\varepsilon\mu}, \quad (6)$$

$$k_0 = \omega \sqrt{\varepsilon_0\mu_0} = \frac{2\pi}{\lambda}, \quad (7)$$

ε , μ being the relative dielectric constant and the relative permeability, and λ the free space wave length. Thus, once H_z and E_z have been obtained, the rest of the electromagnetic field can be found from them.

The configuration in question is shown in Fig. 1. At the left and to the right of the interface, the field is a superposition of the *LSE* and *LSM* modes. The component H_z of the mode LSE_{1r} ($r = \overline{0, N}$), propagating along the waveguide, can be written as

$$H_z = a_r N_r^H \cos \frac{r\pi(x+b/2)}{b} e^{-\gamma_r y}, \quad (8)$$

where

$$\gamma_r = \sqrt{\left(\frac{r\pi}{b} \right)^2 - \chi^2}, \quad (9)$$

a_r is the wave amplitude, and N_r^H is a normalizing constant. The superscript H indicates, that we deal with a *LSE* wave ($H_z \neq 0$). The scattered wave amplitudes present elements of the generalized scattering matrix on assumption that the amplitude of the incident wave is unity. At that, normalizing factor is taken usually so, that for an undamped mode the quantity $\frac{|a_r|^2}{2}$ equals to the transmitted power [5]. Conventional computations result in

$$N_r^H = \sqrt{\frac{4i\chi^2}{e_r ab\omega\mu\mu_0\gamma_r}}, \quad (10)$$

where

$$e_r = \begin{cases} 2 & \text{for } r = 0, \\ 1 & \text{for } r \geq 1. \end{cases} \quad (11)$$

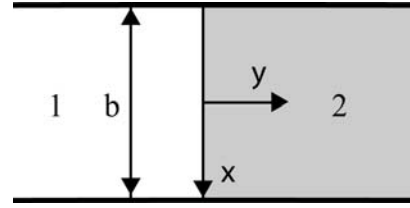


Fig. 1. Geometry of the problem

In the case of the mode LSM_{1r} ($r = \overline{1, N}$), E_z -component is of the form

$$E_z = b_r N_r^E \sin \frac{r\pi(x+b/2)}{b} e^{-\gamma_r y}, \quad (12)$$

$$N_r^E = \sqrt{\frac{4i\chi^2}{ab\omega\varepsilon\varepsilon_0\gamma_r}}. \quad (13)$$

The above representations ensure the realization of the boundary conditions at the conducting walls. We supply the used quantities with the indices 1 and 2 respectively at the left and to the right of the interface and note that modes with different values of r do not interact inside the geometry under study. After choice of r , the problem is reduced to finding sets of values of amplitudes on different sides of the plane $y = 0$, which guarantee the continuity of the tangential components of the electromagnetic field across the interface for all possible types of excitation.

SCATTERING MATRIX

In order a little bit to simplify the subsequent calculations, we will take the normalizing factors in (8) and (12) equal to one and consider their actual values in the final stage. Let us assume that the geometry is excited by the wave LSE_{1r} ($r = \overline{0, N}$) from the left. Then

$$H_z^{(1)} = \left(e^{-\gamma_r^{(1)} y} - a_r^{(1)} e^{\gamma_r^{(1)} y} \right) \cos \frac{r\pi(x+b/2)}{b}, \quad (14)$$

$$E_z^{(1)} = b_r^{(1)} e^{\gamma_r^{(1)} y} \sin \frac{r\pi(x+b/2)}{b}, \quad (15)$$

$$H_z^{(2)} = a_r^{(2)} e^{-\gamma_r^{(2)} y} \cos \frac{r\pi(x+b/2)}{b}, \quad (16)$$

$$E_z^{(2)} = b_r^{(2)} e^{-\gamma_r^{(2)} y} \sin \frac{r\pi(x+b/2)}{b}, \quad (17)$$

where $a_r^{(1)}$, $b_r^{(1)}$, $a_r^{(2)}$ and $b_r^{(2)}$ are to be found. Substituting (14) and (15) in (3) and (4), we obtain

$$\begin{aligned} H_x^{(1)} &= \left[\frac{i\omega\varepsilon_1\varepsilon_0\gamma_r^{(1)}}{\chi_1^2} b_r^{(1)} e^{\gamma_r^{(1)} y} - \frac{r\pi^2}{ab\chi_1^2} \left(e^{-\gamma_r^{(1)} y} - a_r^{(1)} e^{\gamma_r^{(1)} y} \right) \right] \times \\ &\quad \times \sin \frac{r\pi(x+b/2)}{b}, \quad (18) \end{aligned}$$

$$E_x^{(1)} = \left[\frac{i\omega\mu_1\mu_0\gamma_r^{(1)}}{\chi_1^2} \left(e^{-\gamma_r^{(1)}y} + a_r^{(1)} e^{\gamma_r^{(1)}y} \right) - \frac{r\pi^2}{ab\chi_1^2} b_r^{(1)} e^{\gamma_r^{(1)}y} \right] \times \cos \frac{r\pi(x+b/2)}{b}. \quad (19)$$

With the help of (16) and (17), components $H_x^{(2)}$ and $E_x^{(2)}$ are got analogously:

$$H_x^{(2)} = - \left[\frac{i\omega\varepsilon_2\varepsilon_0\gamma_r^{(2)}}{\chi_2^2} b_r^{(2)} e^{-\gamma_r^{(2)}y} + \frac{r\pi^2}{ab\chi_2^2} a_r^{(2)} e^{-\gamma_r^{(2)}y} \right] \times \sin \frac{r\pi(x+b/2)}{b}, \quad (20)$$

$$E_x^{(2)} = \left[\frac{i\omega\mu_2\mu_0\gamma_r^{(2)}}{\chi_2^2} a_r^{(2)} e^{-\gamma_r^{(2)}y} - \frac{r\pi^2}{ab\chi_2^2} b_r^{(2)} e^{-\gamma_r^{(2)}y} \right] \times \cos \frac{r\pi(x+b/2)}{b}. \quad (21)$$

After equating at $y = 0$ the similar tangential components of the electric and magnetic fields and some simple algebra, the amplitudes are as follows

$$(a_r^{(2)}, b_r^{(2)})^T = M^{-1}C, \quad (22)$$

$$a_r^{(1)} = 1 - a_r^{(2)}, \quad b_r^{(1)} = b_r^{(2)}. \quad (23)$$

Here, T denotes the transposition and the matrices M and C are defined by formulas

$$M_{11} = M_{22} = \frac{r\pi^2}{ab} \left(\frac{1}{\chi_1^2} - \frac{1}{\chi_2^2} \right), \quad (24)$$

$$M_{12} = -\frac{ik_0}{Z_0} \left(\frac{\varepsilon_1\gamma_r^{(1)}}{\chi_1^2} + \frac{\varepsilon_2\gamma_r^{(2)}}{\chi_2^2} \right), \quad (25)$$

$$M_{21} = ik_0Z_0 \left(\frac{\mu_1\gamma_r^{(1)}}{\chi_1^2} + \frac{\mu_2\gamma_r^{(2)}}{\chi_2^2} \right), \quad (26)$$

$$C = \left(0, 2ik_0Z_0 \frac{\mu_1\gamma_r^{(1)}}{\chi_1^2} \right)^T, \quad (27)$$

where $Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}$ is the intrinsic impedance of the free space.

The knowledge of the amplitudes allows writing down the respective elements of the normalized scattering matrix:

$$S_{rr}^{H_1H_1} = a_r^{(1)}, \quad r = \overline{0, N}, \quad (28)$$

$$S_{rr}^{E_1H_1} = \frac{N_r^{H_1}}{N_r^{E_1}} b_r^{(1)} = \frac{1}{Z_0\sqrt{\mu_1}} b_r^{(1)}, \quad r = \overline{1, N}, \quad (29)$$

$$S_{rr}^{H_2H_1} = \frac{N_r^{H_1}}{N_r^{H_2}} a_r^{(2)} = \sqrt{\frac{\mu_2\gamma_r^{(2)}}{\mu_1\gamma_r^{(1)}}} a_r^{(2)}, \quad r = \overline{0, N}, \quad (30)$$

$$S_{rr}^{E_2H_1} = \frac{N_r^{H_1}}{N_r^{E_2}} b_r^{(2)} = \frac{1}{Z_0\sqrt{\mu_1\gamma_r^{(1)}}} b_r^{(2)}, \quad r = \overline{1, N}. \quad (31)$$

In more detail,

$$S_{rr}^{E_1H_1} = \frac{2ik_0\sqrt{\varepsilon_1\mu_1}\gamma_r^{(1)}M_{11}}{\chi_1^2\det M}, \quad r = \overline{1, N}, \quad (32)$$

and

$$S_{rr}^{E_2H_1} = \frac{2ik_0\sqrt{\varepsilon_2\mu_1}\gamma_r^{(1)}\gamma_r^{(2)}M_{11}}{\chi_1\chi_2\det M}, \quad r = \overline{1, N}. \quad (33)$$

Thus, as it is seen from (32) and (33) for $r > 0$, the incoming LSE_{1r} mode excites also the nonzero LSM modes.

Now, let the LSM_{1r} ($r = \overline{1, N}$) mode be incident upon the junction from the left. Then

$$H_z^{(1)} = -a_r^{(1)} e^{\gamma_r^{(1)}y} \cos \frac{r\pi(x+b/2)}{b}, \quad (34)$$

$$E_z^{(1)} = \left(e^{-\gamma_r^{(1)}y} + b_r^{(1)} e^{\gamma_r^{(1)}y} \right) \sin \frac{r\pi(x+b/2)}{b}, \quad (35)$$

and components $H_z^{(2)}$ and $E_z^{(2)}$ keep form (16) and (17). After determining $H_x^{(1)}$, $E_x^{(1)}$, $H_x^{(2)}$, $E_x^{(2)}$ with the help of (3), (4) and taking into account the continuity of the tangential field components, we arrive again at the relation (22), where column matrix C is given by

$$C = \left(-2\frac{ik_0\varepsilon_1\gamma_r^{(1)}}{Z_0\chi_1^2}, 0 \right)^T. \quad (36)$$

Instead of (23), the following equations occur

$$a_r^{(1)} = -a_r^{(2)}, \quad b_r^{(1)} = b_r^{(2)} - 1. \quad (37)$$

The respective scattering matrix elements are of the form

$$S_{rr}^{H_1E_1} = \frac{N_r^{E_1}}{N_r^{H_1}} a_r^{(1)} = Z_0\sqrt{\frac{\mu_1}{\varepsilon_1}} a_r^{(1)}, \quad r = \overline{1, N}, \quad (38)$$

$$S_{rr}^{E_1E_1} = b_r^{(1)}, \quad r = \overline{1, N}, \quad (39)$$

$$S_{rr}^{H_2E_1} = \frac{N_r^{E_1}}{N_r^{H_2}} a_r^{(2)} = Z_0\sqrt{\frac{\mu_2\gamma_r^{(2)}}{\varepsilon_1\gamma_r^{(1)}}} a_r^{(2)}, \quad r = \overline{1, N}, \quad (40)$$

$$S_{rr}^{E_2E_1} = \frac{N_r^{E_1}}{N_r^{E_2}} b_r^{(2)} = \sqrt{\frac{\varepsilon_2\gamma_r^{(2)}}{\varepsilon_1\gamma_r^{(1)}}} b_r^{(2)}, \quad r = \overline{1, N}. \quad (41)$$

When exciting the geometry by the LSE_{1r} ($r = \overline{0, N}$) mode from the right, the $H_z^{(1)}$ is defined by (34), $E_z^{(1)}$ by (15), $E_z^{(2)}$ by (17), and

$$H_z^{(2)} = \left(-e^{\gamma_r^{(2)}y} + a_r^{(2)} e^{-\gamma_r^{(2)}y} \right) \cos \frac{r\pi(x+b/2)}{b}. \quad (42)$$

Boundary conditions at $y = 0$ yield

$$C = \left(\frac{r\pi^2}{ab} \left(\frac{1}{\chi_1^2} - \frac{1}{\chi_2^2} \right), ik_0 Z_0 \left(\frac{\mu_1 \gamma_r^{(1)}}{\chi_1^2} - \frac{\mu_2 \gamma_r^{(2)}}{\chi_2^2} \right) \right)^T, \quad (43)$$

$$a_r^{(1)} = 1 - a_r^{(2)}, \quad b_r^{(1)} = b_r^{(2)}, \quad (44)$$

$$S_{rr}^{H_1 H_2} = \sqrt{\frac{\mu_1 \gamma_r^{(1)} \chi_2^2}{\mu_2 \gamma_r^{(2)} \chi_1^2}} a_r^{(1)}, \quad r = \overline{0, N}, \quad (45)$$

$$S_{rr}^{E_1 H_2} = \frac{1}{Z_0 \sqrt{\mu_2 \gamma_r^{(2)} \chi_1^2}} \left(\epsilon_1 \gamma_r^{(1)} \chi_2^2 b_r^{(1)} \right), \quad r = \overline{1, N}, \quad (46)$$

$$S_{rr}^{H_2 H_2} = a_r^{(2)}, \quad r = \overline{0, N}, \quad (47)$$

$$S_{rr}^{E_2 H_2} = \frac{1}{Z_0 \sqrt{\mu_2}} b_r^{(2)}, \quad r = \overline{1, N}. \quad (48)$$

In the case of excitation from the right by the wave LSM_{1r} ($r = \overline{1, N}$), there hold formulas (34), (15), (16) and $E_z^{(2)}$ takes on form

$$E_z^{(2)} = \left(e^{\gamma_r^{(2)}y} + b_r^{(2)} e^{-\gamma_r^{(2)}y} \right) \sin \frac{r\pi(x+b/2)}{b}. \quad (49)$$

The enforcement of the boundary conditions leads to relations

$$C = \left(\frac{ik_0}{Z_0} \left(\frac{\epsilon_1 \gamma_r^{(1)}}{\chi_1^2} - \frac{\epsilon_2 \gamma_r^{(2)}}{\chi_2^2} \right), \frac{r\pi^2}{ab} \left(\frac{1}{\chi_2^2} - \frac{1}{\chi_1^2} \right) \right)^T, \quad (50)$$

$$a_r^{(1)} = -a_r^{(2)}, \quad b_r^{(1)} = b_r^{(2)} + 1, \quad (51)$$

$$S_{rr}^{H_1 E_2} = Z_0 \sqrt{\frac{\mu_1 \gamma_r^{(1)} \chi_2^2}{\epsilon_2 \gamma_r^{(2)} \chi_1^2}} a_r^{(1)}, \quad r = \overline{1, N}, \quad (52)$$

$$S_{rr}^{E_1 E_2} = \sqrt{\frac{\epsilon_1 \gamma_r^{(1)} \chi_2^2}{\epsilon_2 \gamma_r^{(2)} \chi_1^2}} b_r^{(1)}, \quad r = \overline{1, N}, \quad (53)$$

$$S_{rr}^{H_2 E_2} = Z_0 \sqrt{\frac{\mu_2}{\epsilon_2}} a_r^{(2)}, \quad r = \overline{1, N}, \quad (54)$$

$$S_{rr}^{E_2 E_2} = b_r^{(2)}, \quad r = \overline{1, N}. \quad (55)$$

In whole, the generalized scattering matrix S is characterized by the expression

$$S = \begin{pmatrix} S^{H_1 H_1} & S^{H_1 E_1} & S^{H_1 H_2} & S^{H_1 E_2} \\ S^{E_1 H_1} & S^{E_1 E_1} & S^{E_1 H_2} & S^{E_1 E_2} \\ S^{H_2 H_1} & S^{H_2 E_1} & S^{H_2 H_2} & S^{H_2 E_2} \\ S^{E_2 H_1} & S^{E_2 E_1} & S^{E_2 H_2} & S^{E_2 E_2} \end{pmatrix}. \quad (56)$$

Here, the submatrices are of a diagonal form with non-zero elements given by formulas (28)–(31), (38)–(41), (45)–(48) and (52)–(55). At that, $S^{HH} = (S_{mn}^{HH})$ ($m, n = \overline{0, N}$), $S^{EH} = (S_{mn}^{EH})$ ($m = \overline{1, N}, n = \overline{0, N}$), $S^{HE} = (S_{mn}^{HE})$ ($m = \overline{0, N}, n = \overline{1, N}$), $S^{EE} = (S_{mn}^{EE})$ ($m, n = \overline{1, N}$). Let us also notice, that S -matrices of the homogeneous E -plane sections mentioned in Introduction have formally form (56) too, but the blocks of the type S^{EH} or S^{HE} consist completely of zeros and the remaining blocks are dense.

CONCLUSIONS

For a rectangular waveguide, the problem of determining the scattering characteristics of the jump of medium parameters is considered under condition that the field has one variation along the direction perpendicular to E -plane. It is shown that, when exciting the structure by one of the modes of type LSE_{1m} or LSM_{1m} , the scattered field consists of the similar modes both nonzero (except for excitation by LSE_{10}). The calculation formulas for elements of the generalized scattering matrix of the unit have been derived. The examined configuration can be useful in the analysis of piecewise homogeneous E -plane structures as a model of a plane interface between sections of constant permittivity and permeability.

LIST OF REFERENCES

1. Schwinger J. Discontinuities in waveguides / J. Schwinger, D.S. Saxon. – New York : Gordon and Breach, 1968. – 163 p.
2. Kanellopoulos V.N. A complete E -plane analysis of waveguide junctions using the finite element method / V.N. Kanellopoulos, J.P. Webb // IEEE Trans. Microwave Theory Tech. – 1990. – Vol. 38. – No. 3 – P. 290–295.
3. Chumachenko V.P. Accurate model of E -plane waveguide junction with loaded rectangular coupling cavity / V.P. Chumachenko // Microwave Opt. Technol. Lett. – 2002. – Vol. 34. – No. 5 – P. 351–354.
4. Mittra R. Analytical techniques in the theory of guided waves / R. Mittra, S. Lee. – New York : Macmillan, 1971. – 324 pp.
5. Силаев М. А. Приложение матриц и графов к анализу СВЧ устройств / М.А. Силаев, С.Ф. Брянцев. – М. : Советское радио, 1970. – 248 с.

Надійшла 8.04.2010

Вашенко В. В., Чумаченко В. П.

МАТРИЦЯ РОЗСІЮВАННЯ СТРИБКА ПАРАМЕТРІВ СЕРЕДОВИЩА В ПРЯМОКУТНОМУ ХВИЛЕВОДІ

Отримана узагальнена матриця розсіювання стрибка параметрів середовища в прямокутному хвилеводі при умові, що поле має одну варіацію в напрямі перпендикулярному площині E . Конфігурація розглядається в якості

моделі плоского інтерфейсу в теорії кусково-однорідних E -площинних структур.

Ключові слова: хвилеводні неоднорідності, плоский інтерфейс, матриця розсіювання.

Ващенко В. В., Чумаченко В. П.

МАТРИЦА РАССЕЙНИЯ СКАЧКА ПАРАМЕТРОВ СРЕДЫ В ПРЯМОУГОЛЬНОМ ВОЛНОВОДЕ

Получена обобщенная матрица рассеяния скачка параметров среды в прямоугольном волноводе при условии, что поле имеет одну вариацию в направлении перпендикулярном плоскости E . Конфигурация рассматривается в качестве модели плоского интерфейса в теории кусочно-однородных E -плоскостных структур.

Ключевые слова: волноводные неоднородности, плоский интерфейс, матрица рассеяния.

УДК 537.86/87

Прокопенко О. В.

Канд. фіз.-мат. наук, доцент Київського національного університету імені Тараса Шевченка

ВЗАЄМНА ФАЗОВА СИНХРОНІЗАЦІЯ ДВОХ НАНОРОЗМІРНИХ МАГНІТНИХ СТРУКТУР З УРАХУВАННЯМ ЗАПІЗНЕННЯ СИГНАЛУ

Розглянуто можливість взаємної фазової синхронізації двох взаємодіючих нанорозмірних магнітних структур (НМС) з урахуванням часового запізнення сигналу. Показано, що залежність ширини смуги синхронізації НМС від відстані між ними має нелінійний, осцилюючий характер. Визначено оптимальне значення відстані між НМС, що може бути використано для створення масивів синхронізованих НМС.

Ключові слова: нанорозмірна магнітна структура, взаємна фазова синхронізація, запізнення сигналу, смуга синхронізації.

ВСТУП

Було теоретично передбачено Слончевським [1–2] та Берже [3], що спін-поляризований електричний струм I , який протікає крізь намагнічений феромагнітний шар, може передавати цьому шару спіновий магнітний момент, що, у свою чергу, призводить до зміни напрямку вектора намагніченості \vec{M} у феромагнітному шарі. Залежно від величини електричного струму I , більшого за деякий критичний (пороговий) струм I_{th} , цей ефект, названий spin-torque effect (дослівно – ефект спінового крутильного моменту), може призводити до появи двох явищ [4–5]: 1) «переключення» феромагнітного шару за рахунок суттєвої зміни напрямку вектора намагніченості; 2) появи у феромагнітному шарі прецесії намагніченості з частотою $\omega/2\pi$, близькою до частоти феромагнітного резонансу (ФМР). Оскільки, як правило, частота ФМР тонких магнітних плівок лежить у мікрохвильовому діапазоні, для таких систем з'являється новий фізичний механізм генерації мікрохвильових сигналів, заснований на ефекті спінового крутильного моменту. Підкреслимо, що цей механізм генерації мікрохвильових сигналів може бути використаний для створення нелінійних активних нанорозмірних магнітних структур (НМС) [1–5].

На сьогоднішній день найбільш поширеними серед НМС є структури двох типів: 1) НМС типу «наностовпчик» (nanopillar) та 2) НМС типу «наноконтакт» (nano-contact). У найпростішому випадку ці структури складаються з трьох шарів (рис. 1). Один з цих шарів є магнітожорстким шаром (шар 1 на рис. 1) – у зовнішньому магнітному полі \vec{H}_{ext} вектор намагніченості \vec{M}_{FXL} в цьому шарі має цілком визначений

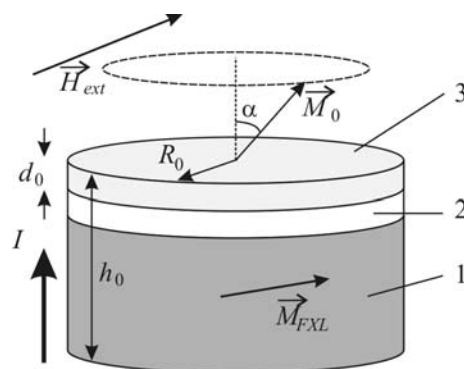


Рис. 1. Схематичне зображення тришарової нанорозмірної магнітної структури (НМС):

1 – магнітожорсткий шар, 2 – немагнітний прошарок товщиною 1×10 нм, 3 – магнітом'який шар товщиною $d_0 \approx 5 + 20$ нм; $R_0 \approx 5 + 500$ нм – радіус НМС, $h_0 \approx 20 + 50$ нм – її товщина

напрямок, а інший шар (шар 3 на рис. 1) є магнітом'яким – у полі \vec{H}_{ext} в цьому шарі існує багатодоменний стан. Середній шар (шар 2 на рис. 1) є тонким немагнітним (зазвичай Cu або MgO) прошарком (типова товщина 1×10 нм), що розділяє магнітом'який та магнітожорсткий шари. При протіканні крізь НМС спін-поляризованого струму I у шарі 3, який знаходиться у зовнішньому магнітному полі \vec{H}_{ext} , відбувається прецесія вектора намагніченості \vec{M}_0 з кутом прецесії α .

Висновки Слончевського – Берже були підтверджені експериментально, як для структур типу «наностовпчик» [6–8], так і для структур типу «нано-контакт» [9–10]. Було встановлено, що протікання крізь НМС постійного струму I , більшого за деякий критичний струм I_{th} , призводить до збудження в НМС автоколивань намагніченості. Частота цих коливань $\omega/2\pi$ залежить від величини струму накачки I та зовнішнього магнітного поля \vec{H}_{ext} і, зазвичай, становить 10×50 ГГц; теоретично вона може бути підвищена приблизно до 200 ГГц [11].

Головними перевагами наногенераторів, що досліджувались, є малі розміри (радіус $R_0 \approx 50 \div 250$ нм, висота $h_0 \approx 20 \div 50$ нм), сумісність технології їх виготовлення з сучасною мікроелектронною технологією, широкі межі перестройки частоти генерації [4–11]. Разом з тим, потужність генерації окремої НМС складає лише $P \approx 1 \div 10$ пВт [4, 6–11], що є недостатнім для переважної більшості практичних застосувань – необхідно, щоб робочий рівень потужності становив не менше ніж $P \sim 1$ мкВт [4–5]. Тому на практиці, для того щоб досягнути більших робочих потужностей, необхідно застосовувати масиви НМС. Відмітимо, що схожа проблема свого часу виникла у надпровідній електроніці при використанні джозефсонівських контактів [12] і була вирішена шляхом їх взаємної синхронізації, що згодом дозволило створити цілу низку мікрохвильових пристроїв на основі масивів синхронізованих джозефсонівських контактів [12–13].

На сьогоднішній день існує спрощена нелінійна теорія взаємної фазової синхронізації двох НМС [5, 14–15], яка не враховує скінченність швидкості поширення сигналу між НМС. Висновки цієї теорії були блискуче підтверджені цілою низкою експериментів [16–17], однак лише для систем, в яких НМС знаходились на близькій відстані одна від одної, коли зсувом фази сигналу при його поширенні в області між НМС можна знехтувати. Для створення масивів синхронізованих НМС, однак, слід враховувати

не тільки зміну амплітуди, але й зміну фази сигналу, що поширюється між НМС.

Метою даної роботи є узагальнення відомої теорії взаємної фазової синхронізації НМС [14–15] для випадку найпростішої системи, що складається з двох взаємодіючих НМС.

ПОСТАНОВКА ЗАДАЧІ

У роботах [5, 14–15] було показано, що динаміка намагніченості в НМС може бути описана аналітичною моделлю

$$\frac{dc}{dt} - i\omega(p)c + \Gamma(p)c = f(t), \quad (1)$$

де $c = c(t)$ – комплексна нормована амплітуда високочастотної складової намагніченості, що прецесує; $p = |c|^2$ – нормована потужність коливань; $\omega(p)$ – циклічна частота неізохронних (залежних від p) коливань намагніченості; $\Gamma(p)$ – загальний ефективний коефіцієнт згасання, який є сумою коефіцієнта звичайного позитивного згасання $\Gamma_+(p)$ та коефіцієнта ефективного негативного згасання $\Gamma_-(p)$ за рахунок протікання крізь НМС спін-поляризованого струму; $f(t)$ – «зовнішня сила», що описує вплив на НМС зовнішніх чинників (сигналу від іншої НМС, зовнішнього електромагнітного сигналу, теплових флуктуацій тощо); $i = \sqrt{-1}$.

Задачу про взаємну фазову синхронізацію двох НМС природно розглядати у наближенні слабого зовнішнього сигналу (сигналу від іншої НМС). Тоді, виходячи з (1) та узагальнюючи рівняння для двох слабо зв'язаних НМС [5, 14–15], динаміку такої системи автогенераторів можна описати наступною системою рівнянь:

$$\frac{dc_1}{dt} + i\omega_1(p_1)c_1 + \Gamma_1(p_1)c_1 = \Omega_{12}e^{i\beta_{12}}c_2, \quad (2)$$

$$\frac{dc_2}{dt} + i\omega_2(p_2)c_2 + \Gamma_2(p_2)c_2 = \Omega_{21}e^{i\beta_{21(12)}}c_1, \quad (3)$$

де величини, позначені індексами 1 та 2, відносяться до першої та другої НМС відповідно; Ω_{ij} – амплітуда коефіцієнта зв'язку i -ї НМС з j -ю НМС, а β_{ij} – фаза відповідного коефіцієнта зв'язку (випадок $\beta_{ij} = 0$ було розглянуто у роботах [5, 14–15]).

Рівняння (2)–(3) симетричні – при перестановці індексів 1 та 2 перше рівняння переходить у друге і навпаки. Враховуючи це, далі обмежимося розглядом лише одного з рівнянь, наприклад, рівняння (2), яке перепишемо у такому узагальненому вигляді:

$$\frac{dc_i}{dt} + i\omega_i(p_i)c_i + \Gamma_i(p_i)c_i = \Omega_{ij}e^{i\beta_{ij}}c_j, \quad (4)$$

де індексами i та j позначені величини, що відносяться до різних НМС.

У результаті розв'язання узагальненої системи рівнянь (4) у даній роботі мають бути визначені умови, за яких може відбуватись взаємна фазова синхронізація НМС та розрахована величина смуги фазової синхронізації (максимальний розкид частот $|\omega_1 - \omega_2|$, за якого може відбуватись синхронізація).

РОЗВ'ЯЗОК ЗАДАЧІ

У граничному випадку $\Omega_{ij} \rightarrow 0$ (випадок незв'язаних НМС) розв'язок рівняння (4) може бути представлений у вигляді:

$$c_i(t) = \sqrt{p_{0i}} e^{i\omega_{0i}t + i\varphi_{0i}}, \quad (5)$$

де p_{0i} – потужність коливань у стаціонарному режимі роботи НМС, який реалізується за умови $\Gamma_i(p_{0i}) = 0$; $\omega_{0i} = \omega_i(p_{0i})$ – частота коливань у стаціонарному режимі; φ_{0i} – початкова фаза коливань. Для випадку малих збурень розв'язок рівняння (4), по аналогії з (5), будемо шукати у вигляді

$$c_i(t) = \sqrt{p_i(t)} e^{i\varphi_i(t)} \approx \sqrt{p_{0i}} [1 + \rho_i(t)] e^{i\varphi_{0i}(t)}, \quad (6)$$

де $\rho_i(t) \approx [p_i(t) - p_{0i}] / 2p_{0i}$ – безрозмірна амплітуда флукутації потужності i -го автогенератора, $\varphi_i(t)$ – фаза коливань i -го автогенератора. Аналогічним чином для j -го автогенератора

$$c_j(t) = \sqrt{p_j(t)} e^{i\varphi_j(t)} \approx \sqrt{p_{0j}} [1 + \rho_j(t)] e^{i\varphi_{0j}(t)}. \quad (7)$$

Підставляючи розв'язки (6)–(7) у (4) та представляючи величини $\omega_i(p_i)$, $\Gamma_i(p_i)$ у вигляді

$$\begin{aligned} \omega_i(p_i) &\approx \omega_i(p_{0i}) + 2p_{0i} \frac{d\omega_i(p_{0i})}{dp_i} \rho_i(t), \\ \Gamma_i(p_i) &\approx \Gamma_i(p_{0i}) + 2p_{0i} \frac{d\Gamma_i(p_{0i})}{dp_i} \rho_i(t), \end{aligned} \quad (8)$$

отримуємо комплексне рівняння для величин $\rho_{i,j}(t)$, $\varphi_{i,j}(t)$, яке може бути записане як система двох рівнянь (враховуються лише величини першого порядку малізми $\sim \Omega_{ij}$ або $\rho_{i,j}(t)$):

$$\frac{d\rho_i}{dt} + 2G_i p_{0i} \rho_i = \operatorname{Re} \left\{ \Omega_{ij} \frac{\sqrt{p_{0j}}}{\sqrt{p_{0i}}} e^{-i\varphi_i} e^{i\varphi_j} e^{i\beta_{ij}} \right\}, \quad (9)$$

$$\frac{d\varphi_i}{dt} - \omega_{0i} - 2N_i p_{0i} \rho_i = \operatorname{Im} \left\{ -\Omega_{ij} \frac{\sqrt{p_{0j}}}{\sqrt{p_{0i}}} e^{-i\varphi_i} e^{i\varphi_j} e^{i\beta_{ij}} \right\}, \quad (10)$$

де $N_i = \frac{d\omega_i(p_{0i})}{dp_i}$, $G_i = \frac{d\Gamma_i(p_{0i})}{dp_i}$. Введемо коефіцієнт згасання флукутацій потужності $\Gamma_{0i} = 2G_i p_{0i}$ та безрозмірний нелінійний зсув частоти $\nu_i = N_i / G_i$, тоді система рівнянь (9)–(10) набуде вигляду

$$\frac{d\rho_i}{dt} + \Gamma_{0i} \rho_i = \Omega_{ij} \frac{\sqrt{p_{0j}}}{\sqrt{p_{0i}}} \cos(\beta_{ij} + \varphi_j - \varphi_i), \quad (11)$$

$$\frac{d\varphi_i}{dt} - \omega_{0i} - \nu_i \Gamma_{0i} \rho_i = -\Omega_{ij} \frac{\sqrt{p_{0j}}}{\sqrt{p_{0i}}} \sin(\beta_{ij} + \varphi_j - \varphi_i). \quad (12)$$

Отримана модель (11)–(12) є моделлю «потужність – фаза» для взаємодіючих автогенераторів. Коректність цієї моделі визначається умовою $|\rho_{i,j}(t)| \ll 1$.

Модель «потужність – фаза» може бути додатково спрощена, якщо ввести ефективну фазу

$$\Phi_{i,j}(t) = \varphi_{i,j}(t) + \nu_{i,j} \rho_{i,j}(t). \quad (13)$$

Виходячи з (11)–(13), ефективна фаза $\Phi_i \equiv \Phi_i(t)$ задовольняє рівнянню

$$\begin{aligned} \frac{d\Phi_i}{dt} - \omega_{0i} &= \\ &= \Omega_{ij} \frac{\sqrt{p_{0j}}}{\sqrt{p_{0i}}} [v_i \cos(\beta_{ij} + \varphi_j - \varphi_i) - \sin(\beta_{ij} + \varphi_j - \varphi_i)]. \end{aligned} \quad (14)$$

Якщо тепер знехтувати відмінністю між звичайними фазами $\varphi_{i,j}(t)$ та ефективними фазами $\Phi_{i,j}(t)$, одержимо так звану «фазову» модель:

$$\frac{d\Phi_i}{dt} - \omega_{0i} = F_{ij}(\Phi_j - \Phi_i), \quad (15)$$

де

$$\begin{aligned} F_{ij}(\Phi_j - \Phi_i) &= \\ &= \Omega_{ij} \frac{\sqrt{p_{0j}}}{\sqrt{p_{0i}}} [v_i \cos(\beta_{ij} + \Phi_j - \Phi_i) - \sin(\beta_{ij} + \Phi_j - \Phi_i)]. \end{aligned}$$

Для НМС вперше вона була представлена у роботі [18]. Ця модель є коректною за умови $|v_{i,j}| |\rho_{i,j}| \ll 1$.

Введемо нові величини $\Phi_{\pm}(t) = \Phi_i(t) \pm \Phi_j(t)$, тоді виходячи з (15), одержимо:

$$\frac{d\Phi_{\pm}}{dt} = \omega_{\pm} + F_{\pm}(\Phi_{\pm}), \quad \frac{d\Phi_{-}}{dt} = \omega_{-} + F_{-}(\Phi_{-}), \quad (16)$$

де $\omega_{\pm} = \omega_{0i} \pm \omega_{0j}$, $F_{\pm}(\Phi_{\pm}) = F_{ij}(\Phi_{\pm}) \pm F_{ji}(\Phi_{\pm})$.

Взаємна фазова синхронізація двох слабко зв'язаних НМС буде відбуватись за умови $\Phi_{-} = \text{const}$. Тоді з другого рівняння (16) відразу ж можна одержати величину смуги фазової синхронізації:

$$|\omega_i(p_i) - \omega_j(p_j)| = |F_{-}(\Phi_{-})|. \quad (17)$$

З іншого боку, перше рівняння (16) буде визначати частоту генерації НМС у режимі фазової синхронізації.

Частота генерації та ширина смуги фазової синхронізації залежать від величини фази коефіцієнта зв'язку β_{ij} . Відмітимо, що оскільки основним механізмом зв'язку НМС є зв'язок за рахунок взаємодії

намагніченості НМС зі спіновими хвилями, породженими прецесією намагніченості в іншій НМС [19], величину β_{ij} зручно представити у вигляді [18]:

$$\beta_{ij} = \omega_j(p_j)\tau_{ij} = a \frac{\omega_j(p_j)}{u_j}, \quad (18)$$

де $\omega_j(p_j)$ – циклічна частота спінової хвилі, що генерується j -ю НМС; t_{ij} – час поширення цієї спінової хвилі між i -ю та j -ю НМС; u_j – швидкість поширення спінової хвилі; a – відстань між НМС. Опис запізнення сигналу зв'язку між НМС за допомогою фазового зсуву (17) коректний за умови $\Omega_{ij}\tau_{ij} \ll 1$ (ця умова виконується у більшості практичних випадків) [18].

РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

Отримані вище результати можуть бути представлені у простому та зручному вигляді, якщо припустити, що усі параметри НМС, окрім їх власних частот ω_{0i} , ω_{0j} , є рівними. Це припущення дозволяє не тільки спростити інтерпретацію отриманих залежностей, але й природним чином відповідає випадку виготовлення масивів майже однакових НМС в умовах єдиного технологічного процесу. З урахуванням зроблених припущень з виразів (17)–(18) випливає, що величина смуги фазової синхронізації для майже ідентичних НМС визначається виразом:

$$\begin{aligned} & |\omega_i(p_i) - \omega_j(p_j)| < \Delta\omega = \\ & = 2\Omega\sqrt{1+v^2}|\cos(\beta - \arctan v)|. \end{aligned} \quad (19)$$

З (19) випливає, що існує оптимальний фазовий зсув $\beta_{\text{opt}} = \arctan v$, а отже, згідно (18), існує оп-

тимальна відстань між НМС $a_{\text{opt}} = (u_i + u_j) \times \arctan(v)/(\omega_i + \omega_j)$, при якій взаємна фазова синхронізація НМС відбувається найбільш ефективно. Оскільки для НМС нелінійний зсув частоти $v \sim 1/\alpha_G \sim 100$ (де α_G – параметр згасання Гільберта), взаємна синхронізація НМС відбувається суттєво відмінним чином, ніж взаємна фазова синхронізація традиційних автогенераторів, для яких $v \ll 1$ (рис. 2). Так, при $v > 50$, вже можна вважати, що $\beta_{\text{opt}} \approx n\pi$, де $n = 1, 2, 3, \dots$. Більш детально це питання проаналізовано в роботі [20].

Другим висновком, який можна зробити виходячи з (19), є висновок про те, що ширина смуги фазової синхронізації НМС суттєво зростає (приблизно у v разів) порівняно з аналогічною величиною для традиційних коливальних систем. Це означає, що системи з багатьма НМС можуть використовуватись при значно більшому розкиді параметрів (зокрема, власних частот), ніж традиційні автогенератори.

Результати проведеного аналізу ілюструються залежностями, показаними на рис. 2.

ВИСНОВКИ

Розроблено теоретичну модель взаємної фазової синхронізації двох слабко зв'язаних НМС і проведено її аналіз. Встановлено, що:

1) існує оптимальне значення фазового зсуву сигналу зв'язку $\beta_{\text{opt}} = \arctan v$ та оптимальне значення відстані між НМС $a_{\text{opt}} = u\beta_{\text{opt}}/\omega$, де $v \gg 1$ – нелінійний зсув частоти НМС, ω та u – циклічна частота та швидкість спінових хвиль, що генеруються НМС;

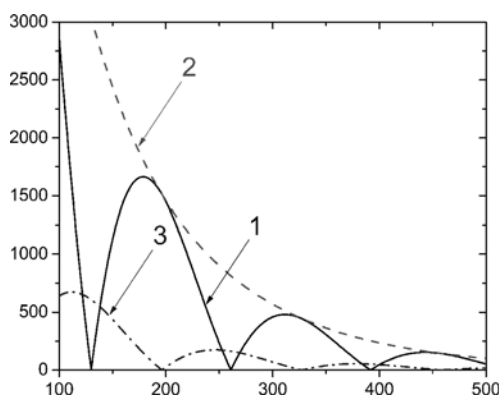


Рис. 2. Залежність частотної смуги вимушеної фазової синхронізації двох НМС від відстані a між ними (1), залежність коефіцієнта зв'язку від відстані a (2), залежність збільшеної в 10 разів частотної смуги вимушеної фазової синхронізації двох НМС за відсутності нелінійного зсуву частоти ($v = 0$) від відстані a (3)

Параметри НМС: структура нормально намагнічена ($H_{\text{ext}} = 15$ кЕ), намагніченість насичення $4\pi M_0 = 8$ кЕ, радіус НМС $R_0 = 50$ нм, коефіцієнт згасання Гільберта $\alpha_G = 0,01$, надкритичність $\zeta = I/I_{\text{th}} = 1,5$.

2) залежність смуги фазової синхронізації двох НМС суттєво залежить від нелінійного зсуву частоти і визначається множителем $\sqrt{1+v^2} \gg 1$, що призводить до значного розширення цієї смуги по відношенню до аналогічної величини для традиційних коливальних систем;

3) взаємна фазова синхронізація НМС відбувається суттєво відмінним чином, ніж взаємна фазова синхронізація традиційних автогенераторів, для яких $v \ll 1$. Так, для традиційних автогенераторів $\beta_{\text{opt}} \approx \pi/2$, у той час як для системи з двох НМС $\beta_{\text{opt}} \approx \pi$.

Автор вдячний проф. Славіну А. М. та д-ру Тиберкевичу В. С. за корисні поради і зауваження при обговоренні результатів, що увійшли в публікацію.

Робота була підтримана грантом М/175-2007 Міністерства освіти і науки України.

СПИСОК ЛІТЕРАТУРИ

1. *Slonczewski J. C.* Current-driven excitation of magnetic multilayers / J. C. Slonczewski // *Journal of Magnetism and Magnetic Materials*. – 1996. – Vol. 159. – L1–L7.
2. *Slonczewski J. C.* Excitation of spin waves by an electric current / J. C. Slonczewski // *Journal of Magnetism and Magnetic Materials*. – 1999. – Vol. 195. – L261–L268.
3. *Berger L.* Emission of spin waves by a magnetic multilayer traversed by a current / L. Berger // *Physical Review B*. – 1996. – Vol. 54. – P. 9353–9358.
4. *Ralph D. C.* Spin Transfer Torques / D. C. Ralph, M. D. Stiles // *Journal of Magnetism and Magnetic Materials*. – 2008. – Vol. 320. – P. 1190–1216.
5. *Slavin A.* Nonlinear auto-oscillator theory of microwave generation by spin-polarized current / A. Slavin, V. Tiberkevich // *IEEE Transaction on Magnetics*. – 2009. – Vol. 45, № 4. – P. 1875–1918.
6. *Kiselev S. I.* Microwave oscillations of a nanomagnet driven by a spin-polarized current / S. I. Kiselev, J. C. Sankey, I. N. Krivorotov, N. C. Emley, R. J. Schoelkopf, R. A. Buhrman, D. C. Ralph // *Nature*. – 2003. – Vol. 425. – P. 380–383.
7. *Kiselev S. I.* Current-induced nanomagnet dynamics for magnetic fields perpendicular to the sample plane / S. I. Kiselev, J. C. Sankey, I. N. Krivorotov, N. C. Emley, M. Rinkoski, C. Perez, R. A. Buhrman, D. C. Ralph // *Physical Review Letters*. – 2004. – Vol. 93, № 3. – P. 036601 (4 p.).
8. *Krivorotov I. N.* Time-Resolved Measurements of Nanomagnet Dynamics Driven by Spin-Transfer Torques / I. N. Krivorotov, N. C. Emley, J. C. Sankey, S. I. Kiselev, D. C. Ralph, R. A. Buhrman // *Science*. – 2005. – Vol. 307. – P. 228–231.
9. *Rippard W. H.* Direct current induced dynamics in $\text{Co}_{90}\text{Fe}_{10}/\text{Ni}_{80}\text{Fe}_{20}$ point contacts / W. H. Rippard, M. R. Puffall, S. Kaka, S. E. Russek, T. J. Silva // *Physical Review Letters*. – 2004. – Vol. 92, № 2. – P. 027201 (4 p.).
10. *Rippard W. H.* Current-drive microwave dynamics in magnetic point contacts as a function of applied field angle / W. H. Rippard, M. R. Puffall, S. Kaka, T. J. Silva, S. E. Russek // *Physical Review B*. – 2004. – Vol. 70. – P. 100406(R) (4 p.).
11. *Hoeffler M. A.* Theory of magnetodynamics induced by spin torque in perpendicularly magnetized thin films / M. A. Hoeffler, M. J. Ablowitz, B. Ilan, M. R. Puffall, T. J. Silva // *Physical Review Letters*. – 2005. – Vol. 95. – P. 267206 (4 p.).
12. *Лихарев К. К.* Введение в динамику джозефсоновских переходов / К. К. Лихарев. – М.: Наука, 1985. – 320 с.
13. *Ivanyuta O. M.* Microwave detection using Josephson junction arrays integrated in a resonator / O. M. Ivanyuta, Y. I. Kishenko, O. V. Prokopenko, V. M. Raksha, A. M. Klushin // *Physica Status Solidi*. – 2005. – Vol. 2, № 5. – P. 1688–1691.
14. *Slavin A. N.* Nonlinear self-phase-locking effect in an array of current-driven magnetic nanocontacts / A. N. Slavin, V. S. Tiberkevich // *Physical Review B*. – 2005. – Vol. 72. – P. 092407 (4 p.).
15. *Мелков Г.* Умови надвисокочастотної синхронізації для магнітних наноконтактів, розташованих у шарі магнітвпорядкованої речовини / Г. Мелков, О. Прокопенко // *Вісник Київського національного університету імені Тараса Шевченка. Серія: радіофізика та електроніка*. – 2007. – № 10. – С. 43–45.
16. *Kaka S.* Mutual phase-locking of microwave spin torque nano-oscillators / S. Kaka, M. R. Puffall, W. H. Rippard, T. J. Silva, S. E. Russek, J. A. Katine // *Nature*. – 2005. – Vol. 437. – P. 389–392.
17. *Mancoff F. B.* Phase-locking in double-point-contact spin-transfer devices / F. B. Mancoff, N. D. Rizzo, B. N. Engel, S. Tehrani // *Nature*. – 2005. – Vol. 437. – P. 393–395.
18. *Prokopenko O.* Mutual phase-locking of two spin-torque oscillators: Influence of time delay of a coupling signal / O. Prokopenko, V. Tyberkevych, A. Slavin // *Proceedings of the Europe International Magnetics Conference (Intermag 2008, May 4–8, 2008)*. – Madrid (Spain). – DP-06. – P. 682–683.
19. *Slavin A. N.* Theory of mutual phase-locking of spin torque nano-oscillators / A. N. Slavin, V. S. Tiberkevich // *Physical Review B*. – 2006. – Vol. 74. – P. 104401 (4 p.).
20. *Прокопенко А. В.* Влияние запаздывания сигнала на динамику намагниченности двух связанных магнитных наноконтактов / Прокопенко А. В., Верба Р. В. // *Сборник научных трудов ИРЭ НАН Украины «Радиофизика и электроника»*. – 2010. – Том 15, № 3. – С. 78–82.

Надійшла 22.03.2010

Прокопенко А. В.

ВЗАИМНАЯ ФАЗОВАЯ СИНХРОНИЗАЦИЯ ДВУХ НАНОРАЗМЕРНЫХ МАГНИТНЫХ СТРУКТУР С УЧЕТОМ ЗАПАЗДЫВАНИЯ СИГНАЛА

Рассмотрена возможность взаимной фазовой синхронизации двух взаимодействующих наноразмерных магнитных структур (НМС) с учетом временного запаздывания сигнала. Показано, что зависимость ширины полосы синхронизации НМС от расстояния между ними имеет нелинейный, осциллирующий характер. Определено оптимальное значение расстояния между НМС, что может быть использовано для создания массивов синхронизированных НМС.

Ключевые слова: наноразмерная магнитная структура, взаимная фазовая синхронизация, запаздывание сигнала, полоса синхронизации.

Prokopenko O. V.

MUTUAL PHASE-LOCKING OF TWO NANO-SIZE MAGNETIC STRUCTURES CONSIDERING AN INFLUENCE OF SIGNAL DELAY

Possibility of mutual phase-locking of two interacting nano-size magnetic structures (NMS) is considered taking into account time delay. The dependence of NMS phase-locking bandwidth on the distance between them is shown to have a nonlinear, oscillatory nature. An optimal distance between NMS is determined; the obtained results can be used for creation of phase-locked NMS arrays.

Key words: nano-size magnetic structure, mutual phase-locking, signal delay, phase-locking band.

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

РАДІОЕЛЕКТРОНИКА И ТЕЛЕКОММУНИКАЦИИ

RADIO ELECTRONICS AND TELECOMMUNICATIONS

УДК 621.396.6

Артюшенко Б. А.¹, Ільяшенко М. Б.², Головатий А. І.³¹Канд. техн. наук, старший викладач Запорізького національного технічного університету²Канд. техн. наук, доцент Запорізького національного технічного університету³Канд. техн. наук, старший викладач Тернопільського національного технічного університету ім. Івана Пулюя

ВИЗНАЧЕННЯ ВПЛИВУ ВІДХИЛЕНЬ ПАРАМЕТРІВ АНТЕНИ НА ХАРАКТЕРИСТИКИ ДІАГРАМИ НАПРЯМЛЕНОСТІ МЕТОДОМ ПАРАЛЕЛЬНОГО АНАЛІЗУ ГРАНЕЙ

Запропоновано новий метод паралельного аналізу впливу допусків для вібраторних антен – аналіз граней. Метод базується на властивостях антени: невеликій величині відхилень, їх взаємозалежності, монотонності залежності характеристик діаграми напрямленості від параметрів в межах допускової області. Розроблений метод аналізу граней базується на методі бісекцій, але обчислює значення вихідних функцій лише на вершинах допускової області. З метою зменшення часу обчислень розроблено процедуру розпаралелення.

Ключові слова: допусковий аналіз, багатопараметричний аналіз, діаграма напрямленості, паралельні обчислення, антени.

ВСТУП

Проектування випромінюючих пристроїв неможливе без використання комп'ютерних засобів моделювання. Але недоліком існуючих програмних засобів моделювання є обмеженість використання виключно номінальних значень параметрів конструкції виробу. Виготовлення (та експлуатація) антен не може не призвести до всіляких похибок та відхилень параметрів від номінальних значень. Здебільшого для аналізу впливу цих відхилень використовують методи статистичної теорії антен, метод Монте-Карло, обчислення гіршого випадку. На жаль, використання цих методів обмежене лише певними класами антенних пристроїв (статистична теорія), потребує великої кількості обчислень та людського втручання і, як наслідок, є ускладненим при використанні автоматизованих за-

собів проектування і оптимізації (і, зокрема, з використанням методів випадкового пошуку) [1]. Актуальною задачею є розробка методу аналізу впливу допусків на параметри дротових антен довільної форми та фракталоподібної антени. Цей метод має бути обчислювально легким та дозволяти отримувати дані з прийнятною точністю. Для розв'язання цієї задачі можуть бути застосовані методи випадкового пошуку, аналізу вершин, метод бісекцій, метод коефіцієнтів, інтервальні методи [2–3]. Недоліками цих методів є обчислювальна складність, або низька точність (метод коефіцієнтів, інтервальні методи).

Метою статті є розробка методу допускового аналізу фракталоподібної та дротової антени з використанням існуючих засобів комп'ютерного моделювання антен.

© Артюшенко Б. А., Ільяшенко М. Б., Головатий А. І., 2010

Для вирішення поставленої задачі необхідно:

- дослідити найбільш поширені та ймовірні відхилення параметрів антени;
- розробити засоби представлення допускових значень фракталоподібної та дротової антен;
- розробити засіб допускового аналізу антени;
- розробити засоби розпаралелення допускового аналізу.

ПРЕДСТАВЛЕННЯ ДОПУСКІВ ДРОВОВИХ ТА ФРАКТАЛОПОДІБНИХ АНТЕН

При допусковому аналізі антен приймемо такі припущення: відхилення параметрів вважатимемо взаємозалежними, відхилення не призводять до розривів, відхилення є незначними у масштабах конструкції в цілому. Для дротової та фракталоподібної антени мають місце такі відхилення: випадкові повороти провідників, їх неточна довжина, неточність виготовлення провідників (неточний радіус, нерівності поверхні, матеріал з іншими параметрами) тощо. За чисельного моделювання фракталоподібних та дротових антен зазвичай використовують метод моментів, згідно з яким антена представляється як сукупність провідників. Ці провідники розбиваються на лінійні сегменти. Тому опрацювання нерівностей поверхні є неможливим (та й непотрібним для більшості відомих авторам випадків), хоча якщо відомі місця можливих порушень, то в цьому випадку дріт можна представляти як складений з інших. Для простоти нами будуть розглянуті лише відхилення кутів нахилу дротів та їх довжини. Відомі САПР задають провідники набором координат їх кінців, для нашого ж випадку, задля спрощен-

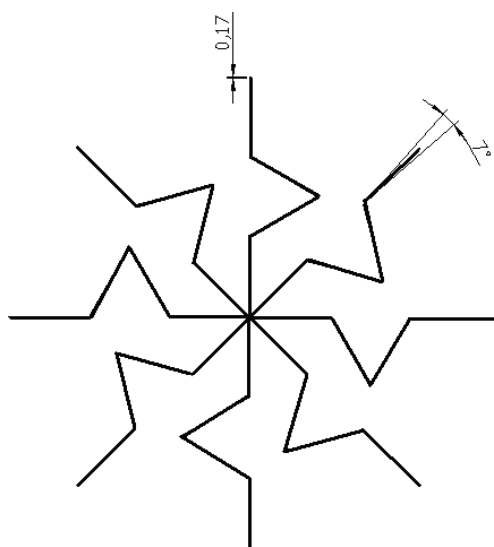


Рис. 1. Фракталоподібна антена фон Коха на першому кроці трансформації

ня було обрано таке представлення: вказівник на попередній провідник або координати початкового провідника, довжина провідника, кути нахилу в горизонтальній та вертикальній площинах. При цьому слід дотримуватись матриці зв'язності, отже, початковий провідник має задаватись користувачем, для решти, в залежності від складності задачі, зв'язок із попереднім провідником обирається користувачем або автоматично (наприклад, для фракталоподібних антен фон Коха, коли кожен «виходить» один з іншого). При флуктуації кутів нахилу одного провідника автоматично забезпечується зміна координат, з ним зв'язаних. Приклад показано на рис. 1.

МЕТОД ГРАНЕЙ

Найповніший допусковий аналіз забезпечують інтервальні методи зовнішнього оцінювання значень функції, але ці методи є досить складними в обчисленнях та їх неможливо сумістити в повному обсязі з методом моментів. Фактично ці методи базуються на методі бісекцій, тобто потребуватимуть до $2^{n \times C}$ обчислень (n – кількість параметрів, C – константа). Разом з тим, у більшості існуючих літературних джерел для задачі допускового аналізу використовується метод коефіцієнтів впливу, отже вважається достатнім перевірити значення функції при максимальних/мінімальних значеннях кожного з параметрів окремо (при номінальних значеннях решти), а потім взяти або коефіцієнти впливу, або значення в умовно найгіршій вершині. Ці методи потребують до $2n + 1$ обчислень. Тобто вважається, що вихідна функція у межах допускової області має лінійний характер. З іншого боку, при випуклій області роботопридатності достатньо перевірити вершини (аналіз вершин), що потребує до 2^n операцій. З метою зменшення обчислювальних витрат нами були поєднані метод бісекцій, метод коефіцієнтів та аналіз вершин – розроблено метод аналізу граней.

При використанні методу аналізу граней верхня та нижня границі значень вихідної функції обчислюються окремо.

На поточному кроці опрацьовуються m неопрацьованих параметрів (на першому кроці всі параметри неопрацьовані). Обчислюється значення при максимальному позитивному та негативному відхиленні кожного з неопрацьованих параметрів та центральних значеннях решти (на першому кроці центральними є номінальні значення). Обирається параметр, відхилення якого призводить до максимізації (мінімізації) вихідної функції. Якщо три точки вихідної функції на прямій зміни цього параметру є монотонно зростаючою (спадаючою) послідовністю,

то центральне значення цього параметра набуває максимального або мінімального значення (у відповідності до значень вихідної функції), цей параметр помічається як опрацьований і, за результатами перевірки критерію зупинки, робиться перехід на наступний крок. Якщо ж ці значення не сумісні з монотонністю функції, це свідчить про можливо неточні результати і слід переходити до процедури завершення.

Процедура завершення обирає найбільше (найменше) значення з отриманих раніше та виводить його користувачеві.

За критерій зупинки можливо обирати: кількість кроків (не може перевищувати кількість параметрів), час роботи, точність.

На рис. 2, а, б наведено схему роботи та розподіл часу (приблизно) запропонованого алгоритму у вигляді автомата зі станами: S0 – ініціалізація даних; S1 – вибір наступного параметра для дослідження; S2 – присвоєння обраному параметру максимального значення; S3 – обчислення діаграми напрямленості (ДН) при максимальному значенні параметра: S3.1 – підготовка даних для моделювання методу аналізу граней, S3.2 – сегментація дротів, S3.3 – обчислення матриць напруг та струмів, S3.4 – проміжні дії, S3.5 – обчислення ДН; S4 – присвоєння мінімального значення параметру та обчислення ДН; S5 – вибір найбільш впливового параметра та присвоєння найгіршого значення; S10 – завершення.

В табл. 1 представлено результати роботи запропонованого методу для аналізу впливу допусків на квадратичну характеристику діаграми напрямленості для антени, показаної на рис. 1 (допуски на всі кути – 1°, на довжини – 1 %).

Таблиця 1. Результати роботи на різних кроках

Номер кроку	Відхилення вихідної функції (%)	Номер кроку	Відхилення вихідної функції (%)
1	$1,72 \cdot 10^{-4}$	6	0,113
2	$9,26 \cdot 10^{-4}$	7	0,114
3	0,103	8	0,116
4	0,108	9	0,118
5	0,111	10	0,119

РОЗПАРАЛЕЛЕННЯ МЕТОДУ

Незважаючи на переваги методу, він потребує розпаралелення з метою зменшення часу обчислень. Можливо використовувати такі підходи: розпаралелення методу чисельного моделювання (метод моментів) та розпаралелення методу аналізу граней.

Перший підхід досить широко представлений в літературі та реалізований в існуючих САПР, але його недоліком для задачі допускового аналізу є відносно велика частина нерозпаралелених обчислень. Це видно з рис. 2, б, де наведено недетерміновані цифрові автомати різних способів [4] розпаралелення (на рис. 2 S6, S7 – відправлення та отримання даних з процесорів).

Розпаралелення аналізу граней може відбуватися таким чином: крок алгоритму складається з $ns/np/nt$ підкроків. На кожному з підкроків кожному з np процесорів відсилається по nt різних точок для аналізу, процесори відсилають результати (значення вихідної функції) і підкрок повторюється. Для паралелізації алгоритму аналізу слід визначитись з такими питаннями: кількість точок для аналізу на підкроці, чи має серверна гілка також виконувати моделювання (або виключно відсилати-приймати дані з інших, рис. 2, в та 2, г), оптимальна кількість процесорів для аналізу.

Вибір кращого засобу розпаралелення можливий лише за допомогою оцінки часових витрат на обчислення (їх наведено в табл. 2). З цих даних та з рис. 2 видно, що кращим є розпаралелення аналізу граней, при якому на підкроці мають обчислюватись всі наявні точки, а кількість процесорів має наближатись до кількості параметрів (на останньому кроці).

Таблиця 2. Розподіл часових витрат на обчислення

Стани	Час обчислень
S0 + S10	0,11
S2 + S3 + S4	$199,5 \times nt$
S6 = S7	$5,79 \cdot 10^{-5}$

Роботу було виконано у рамках науково-дослідної роботи ДБ04110 «Об'єктно-орієнтовані методи проектування радіоелектронних апаратів» (2010-2012) за програмою «Перспективні інформаційні технології, прилади комплексної автоматизації, системи зв'язку». При обчисленнях використовувався кластер НАНУ Інституту Кібернетики ім. В. Глушкова.

ВИСНОВКИ

Було розроблено метод допускового аналізу дрових антен, що дозволяє суттєво скоротити час обчислень при достатній для задачі оптимізації точності. Також досліджено можливості розпаралелення методу та розроблено засоби розпаралелення обчислень. Показано, що паралельний режим доцільно проводити за принципом зменшення частоти обміну даними між процесорами, за умови кратності кількості процесорів кількості елементів.

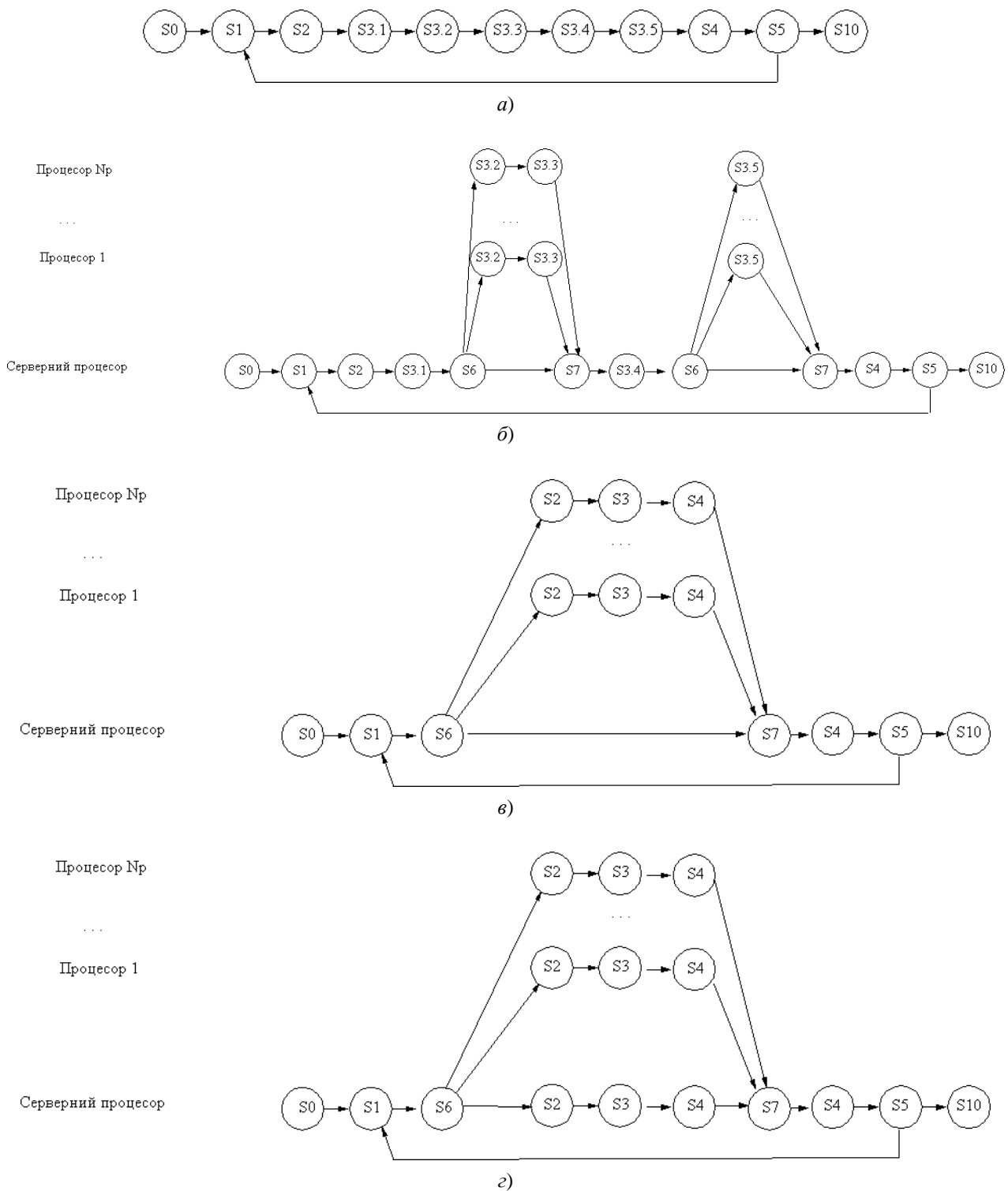


Рис. 2. Цифровий автомат роботи алгоритму аналізу граней:

a – на однопроцесорній машині; *б* – при розпаралеленні методом моментів; *в* – при розпаралеленні розподілом точок варіант 1; *г* – при розпаралеленні розподілом точок варіант 2

СПИСОК ЛІТЕРАТУРИ

1. Шифрин Я. С. Вопросы статистической теории антенн / Я. Шифрин. – М. : Сов. радио, 1970. – 384 с.
2. Вычислительные методы в электродинамике / под ред. Р. Митры – М. : Мир, 1977. – 243 с.
3. Артюшенко Б. А. Допусковый анализ фракталоподобной антенны с применением метода моментов / Артюшенко Б. А., Шило Г. Н., Кришук В. Н., Романенко С. Н. // Матеріали 17-ї Міжнародної Кримської конференції «СВЧ техніка і телекомунікаційні технології» (КрыМиКо2007). – Севастополь, 2007. – С. 402–403.
4. Антонов А. Н. Эффективная организация параллельных распределенных вычислений на основе кластерных технологий: дис. ... канд. техн. наук: 05.13.15 / Антонов А. Н. – Пенза, 2005. – 243 с.

Надійшла 20.01.2010
Після доробки 11.03.2010

Артюшенко Б. А., Ильяшенко М. Б., Головатый А. И.
ОПРЕДЕЛЕНИЕ ВЛИЯНИЯ ОТКЛОНЕНИЙ ПАРАМЕТРОВ АНТЕННЫ НА ХАРАКТЕРИСТИКИ ДИАГРАММЫ НАПРАВЛЕННОСТИ МЕТОДОМ ПАРАЛЛЕЛЬНОГО АНАЛИЗА ГРАНЕЙ

Предложен новый метод параллельного анализа влияния допусков на параметры вибраторных антенн – анализ граней. Метод базируется на свойствах антенн: небольшой величине допусков, взаимнезависимости, монотонности зависимости характеристик диаграммы направленности от

одного из параметров в пределах допускной области. Разработанный метод анализа граней базируется на методе бисекций, но вычисляет значения выходных функций лишь на вершинах допускной области. С целью снижения времени вычислений разработана процедура распараллеливания.

Ключевые слова: допусковый анализ, многопараметрический анализ, диаграмма направленности, параллельные вычисления, антенны.

Artyushenko B. A., Il'yashenko M. B., Holovatyj A. I.

DETERMINATION OF ANTENNA PARAMETERS DEVIATION INFLUENCE ON ANTENNA PATTERN CHARACTERISTICS USING THE METHOD OF PARALLEL BOUNDS ANALYSIS

A new parallel analysis method is proposed for determining the influence of tolerances on the dipole antenna parameters. The method is based on antennas characteristics: small tolerances, interdependence and monotone dependence of radiation pattern on one of the parameters within the tolerance domain. The bound analysis method is based on the bisection method, but it computes output functions only at the vertexes of the tolerance domain. To reduce the time of calculation, a parallelization procedure is developed.

Key words: tolerance analysis, polyvalent analysis, radiation pattern, parallel computation, antenna.

УДК 621.372.001.24:78

Гулин С. П.¹, Гулин А. С.², Дмитренко В. П.¹, Пиза Д. М.³

¹Канд. техн. наук, доцент Запорожского национального технического университета

²Аспирант Запорожского национального технического университета

³Д-р техн. наук, профессор Запорожского национального технического университета

КОНЦЕПЦИЯ УПРАВЛЯЕМОГО ДИНАМИЧЕСКОГО НАСЫЩЕНИЯ ПРИ МОДЕЛИРОВАНИИ МНОГОЧАСТОТНЫХ УСТАНОВИВШИХСЯ РЕЖИМОВ МИКРОЭЛЕКТРОННЫХ КОМПОНЕНТОВ, СТРУКТУР И УСТРОЙСТВ

На основе математического аппарата функциональных рядов Вольтерры – Гаусса (ФРВГ) и функций динамического насыщения (ФДН) предложена концепция управляемого динамического насыщения, реализующая обобщенный подход при исследовании и моделировании микроэлектронных нелинейных инерционных цепей и систем (НИЦС) в условиях многочастотного полиамплитудного воздействия. На ряде примеров состоятельность предложенной концепции подтверждена как с математической, так и с физической точки зрения.

Ключевые слова: нелинейные инерционные цепи и системы, концепция управляемого динамического насыщения, функциональные ряды Вольтерры – Гаусса, обобщенная мгновенная динамическая характеристика, установившийся отклик, многочастотное полиамплитудное воздействие.

ПОСТАНОВКА ЗАДАЧИ

В процессе разработки микроэлектронных (и других) систем в целом часто возникает необходимость

всесторонне исследовать поведение отдельных их компонентов, структур и устройств в условиях многочастотных полиамплитудных воздействий в режимах,

© Гулин С. П., Гулин А. С., Дмитренко В. П., Пиза Д. М., 2010

в которых наиболее полно проявляются их возможности и свойства. При этом широко используются различные методы из арсеналов экспериментальной физики и прикладной математики [1–4].

Объектами исследования могут служить как отдельные активные компоненты и структуры (диоды, транзисторы, электронные лампы, оптоэлектронные преобразователи и пр.), так и устройства и функциональные модули (ФМ): операционные, параметрические и непараметрические усилители, умножители, делители и преобразователи частоты, амплитудные ограничители, управляемые аттенюаторы, модуляторы, детекторы, компрессоры, экспандеры и конвертеры различного рода сигналов. При этом получение необходимых результатов возможно лишь на базе некоторой обобщающей концепции, которая бы позволяла всесторонне исследовать процессы, происходящие в условиях возможных вариаций параметров воздействия и окружающей среды, а также выявлять особенности и закономерности их проявлений.

Анализ многочисленных публикаций данного направления подчеркивает необходимость и значимость разработки и развития обобщенной концепции для решения сформулированных выше задач. Тем более что в процессе проведения узкоспециализированного эксперимента зачастую оказывается, что ряд факторов, характеризующих объект исследования, режим его функционирования, особенности протекания самого процесса, учтен в недостаточной степени или же не учтен вовсе.

Примером, подтверждающим необходимость разработки обобщающего подхода, является работа [5], посвященная моделированию негармонических процессов в энергетике.

Создание подобной концепции подразумевает решение четырех основных задач:

- разработки и развития методов анализа режимов нелинейных цепей и систем, обладающих достаточной степенью общности;
- разработки методов и устройств измерения многомерных параметров нелинейных цепей и систем в условиях многочастотного полиамплитудного воздействия;
- создания обобщенной математической модели для объектов исследуемых классов;
- разработки методов структурной и параметрической идентификации этой модели на основе данных эксперимента.

В данной работе предлагается один из возможных путей решения указанной проблемы – создание обобщающей концепции управляемого динамического насыщения (КУДН) для характеристики, модели-

рования и исследования режимов микроэлектронных активных компонентов, структур, устройств и систем с сосредоточенными параметрами в условиях действия многочастотных сигналов на основе математического аппарата ФРВГ и ФДН.

РЕШЕНИЕ

С целью устранения возможных просчетов и ошибок при разработке обобщающего подхода и концепции для рассматриваемой области исследований на начальном этапе предлагается физические и математические аспекты сформулированной задачи рассматривать отдельно, строго подходить к формулировкам соответствующих ограничений.

Исходя из теории и практики микроэлектронных компонентов и ФМ, работающих в многочастотных режимах, можно сформулировать следующие ограничения, с учетом которых должна решаться задача синтеза их моделей с физической точки зрения:

- физическую основу процесса моделирования ФМ однозначно определяет интегральная технология их производства;

- ФМ выполнены на базе схемотехники усилителей постоянного тока, что позволяет решать задачу анализа режимов и идентификации ФМ с единых методологических позиций;

- поскольку объекты исследования – реальные устройства, то условия физической реализуемости, накладываемые при синтезе их моделей, не являются определяющими;

- внутренняя структура ФМ известна лишь приближенно, что объясняется наличием неконтролируемых паразитных элементов и несовершенством технологии изготовления интегральных компонентов: p - n -переходов, конденсаторов, резисторов;

- объектом идентификации является интегральная схема, представляющая собой, в общем случае, нелинейный инерционный многополюсник (НИМ) с несколькими входами и выходами, принадлежащий к классу $ARCL$ -цепей с сосредоточенными параметрами;

- доступными для методов идентификации являются лишь внешние выводы интегральной схемы ФМ, поскольку непосредственное измерение параметров образующих их компонентов неразрушающими методами контроля крайне затруднено и/или невозможно;

- двунаправленности передачи информационного сигнала, которая, однако, для устройств низкочастотного и высокочастотного диапазонов (например, операционных усилителей (ОУ), аналоговых перемножителей) может иметь однонаправленный характер;

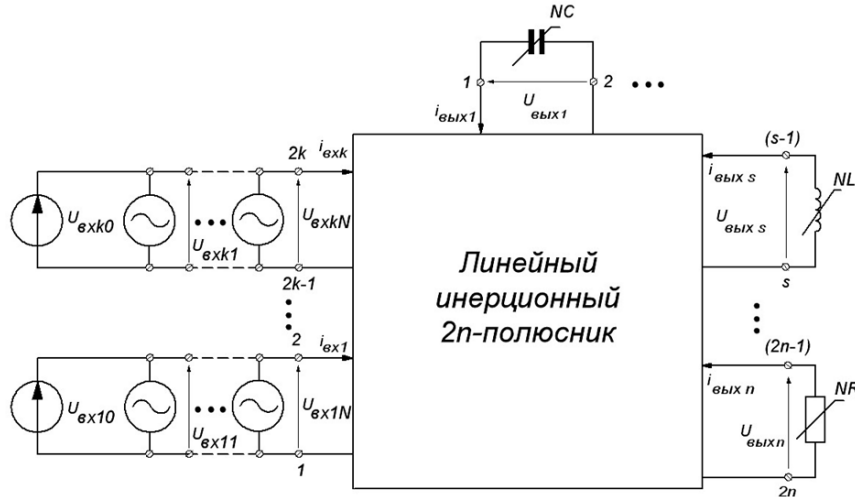


Рис. 1. Модель НИЦ:

$U_{вхl,0}, U_{вхl,i}$ – величина постоянной составляющей и амплитуда i -гармонической компоненты воздействия на l -м входе, $l = \overline{1, k}$;
 $i_{вхl}$ – мгновенное значение тока на l -м входе; $u_{выхj}, i_{выхj}$ – мгновенные значения напряжения и тока на j -выходе, $j = \overline{1, n}$

- режим объектов исследования предполагается нелинейным и установившимся;
- класс исследуемых объектов здесь ограничивается конвергентными НИЦС;
- воздействие, удовлетворяющее условию узкополосности, представляет собой многочастотный полиамплитудный тестовый сигнал (МПТС) вида

$$x(t) = X_0 + \sum_{i=1}^N X_i \cdot \cos\left(\sum_{j=1}^N n_j \cdot \omega_j \cdot t + \sum_{j=1}^N n_j \cdot \varphi_j\right), \quad (1)$$

где $X_0 = \text{const}$; X_i, ω_i, φ_i – амплитуда, круговая частота и начальная фаза i -компоненты воздействия, $i, j = \overline{1, N}$; n_j – целочисленные коэффициенты, удовлетворяющие условию $n_j \in [-N; N]$; частоты ω_i несоизмеримы между собой и образуют монотонно-возрастающий лексикографический ряд $0 < \omega_1 < \omega_2 < \dots < \omega_N$ [3, 6].

С учетом сформулированных ограничений в качестве обобщенной модели нелинейной инерционной цепи (НИЦ) рассматриваемого класса предлагается цепь, представленная на рис. 1.

НИЦ рис. 1 должна удовлетворять следующим ограничениям на вырожденные виды топологии ее эквивалентной схемы и типы управляющих переменных:

1. Граф цепи не должен содержать контуры, образованные только из независимых и (или) управляемых источников напряжения (E -контур [7]).
2. В графе цепи не должны существовать сечения, содержащие только независимые и (или) управляемые источники тока (I -сечения) [7].

3. Топология НИЦ задана графом эквивалентной схемы, который содержит s узлов и p ветвей. При этом полагаем, что цепи с взаимной индукцией отсутствуют.

4. Число нелинейных элементов (НЭ) W , из которых n_G, n_C и n_L являются, соответственно, нелинейными проводимостями, емкостями и индуктивностями, ограничено числом независимых контуров: $W = n_G + n_C + n_L \leq p - s + 1$.

5. Параметры линейных двух- и четырехполюсников известны, а их комплексные сопротивления (проводимости) могут быть вычислены для любой частоты.

6. Управляемый источник, принадлежащий CE -контур или LI -сечению, должен зависеть только от напряжения на емкости дерева или от тока индуктивности антидерева [7].

7. Соотношение «вход – выход» нелинейной емкости должно определяться функцией напряжения на ней, т. е. $q_{NC} = q_f(u)$, а нелинейные индуктивности – током, т. е. $\psi_{NL} = \psi(i)$.

Для расширения класса исследуемых НИЦ допускается применение в качестве ветви ее графа обобщенной нелинейной цепи на рис. 2.

Соотношения «вход – выход» НЭ заданы в виде

$$i_{Nw} = i_{Nw}(u_w), \quad i_{Nw} = i_{Nw}(i_w), \quad i_{Nw} = i_{Nw}(\psi_w), \\ q_{Nw} = q_{Nw}(u_w), \quad w = \overline{1, W} = (1, 2, \dots, W), \quad (2)$$

где i_{Nw}, q_{Nw} – мгновенные значения тока и заряда в НЭ; i_w, u_w, ψ_w – мгновенные значения управляющего тока, напряжения и потокосцепления НЭ.

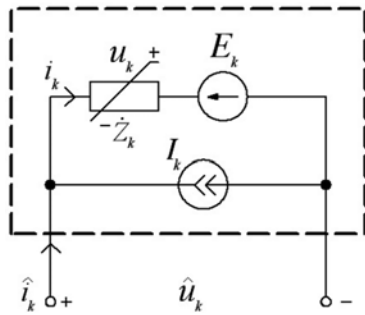


Рис. 2. Обобщенная k -ветвь графа НИЦ, содержащая независимый источник напряжения E_k , независимый источник тока I_k и нелинейное сопротивление \dot{Z}_k с двумя выводами (или управляемый напряжением источник тока)

При обобщении и унификации моделей НЭ возникла задача поиска гипотетической «универсальной» аналитической функции (или функций), которая бы описывала поведение их соотношений «вход – выход» с достаточной точностью, была бы относительно простой и, что самое главное, включала ряд параметров, вариация которых обеспечивала бы необходимую гибкость в изменении ее формы в широком диапазоне значений аргумента.

В [6] была определена система требований, которым должна удовлетворять гипотетическая универсальная функция аппроксимации. В соответствии с этими требованиями, эта функция должна:

- быть линейной для малого сигнала;
- при неограниченном возрастании ее аргумента асимптотически стремиться к некоторому регулируемому уровню насыщения;
- иметь замкнутую аналитическую форму;
- позволять варьировать крутизну наклона линейного участка функции и кривизну участка перехода от малосигнального режима к режиму среднего и большого сигнала;
- обеспечивать возможность перемещения и изменения графика функции в любом направлении оси соответствующей системы координат;
- иметь обратный оператор, выраженный в аналитическом виде;
- обеспечивать требуемую точность аппроксимации в области изменения аргумента.

С учетом сформулированных выше жестких ограничений, сформулированная задача была решена в [6, 8, 9], где для описания соотношений «вход – выход» (2) НЭ были предложены следующие функции и их модификации:

$$F1[x(t)] = F_{0,N} + A_N \cdot \left\{ 1 + \left[\frac{B_N}{x(t)} \right]^{p_N} \right\}^{-S_N^{-1}},$$

$$F2[x(t)] = F_{0,N} \cdot x(t) + A_N \cdot \left\{ 1 + \left[\frac{B_N}{x(t)} \right]^{p_N} \right\}^{-S_N^{-1}}; \quad (3)$$

$$Y1[x(t)] = Y_{0,N} - A_N \cdot \ln[1 + B_N/\exp(x(t))],$$

$$Y2[x(t)] = Y_{0,N} \cdot x(t) - A_N \cdot \ln[1 + B_N/\exp(x(t))], \quad (4)$$

где $F_{0,N}$, S_N , p_N , A_N , B_N – параметры ФДН (3), а $Y_{0,N}$, A_N , B_N – параметры ФДН (4); индекс « N » в обозначении параметров, в зависимости от типа НЭ, принимает значения G , R , C , L . В дальнейших выкладках для НЭ, управляемых напряжением, в соотношениях (3)–(4) « x » заменяется « U », а для НЭ, управляемых током, « x » заменяется « I ».

Результаты применения ФДН для решения задач аппроксимации характеристик НЭ различной физической природы показали высокую эффективность методов, основанных на принципе управляемого динамического насыщения [10–13]. Однако решения лишь задачи аппроксимации произвольной аналитической функции для становления и дальнейшего развития КУДН недостаточно.

Необходимая методологическая основа КУДН заложена в работах [6, 14–20]. Так, в [6, 14] на основе принципа управляемого динамического насыщения были разработаны методы анализа установившегося отклика безынерционной нелинейности на многочастотное воздействие, в котором описание соотношений «вход – выход» НЭ осуществлялось с помощью ФДН. В [9] обоснована возможность применения ФДН для характеристики НЭ и ФМ. В дальнейшем предложенный в [6, 14] алгоритм характеристики был распространен на резистивные и реактивные НЭ различных типов управления и нелинейные управляемые источники [17, 18]. В [19] разработана методика расчета общего числа комбинационных колебаний (КК), включая компоненты постоянной составляющей (DC -компоненты), установившегося отклика НИЦ на воздействие (1). Физически DC -компоненты могут возникать, например, в результате паразитного детектирования некоторых КК. При этом учитывается динамика рабочей точки НЭ. Недавно данный метод получил свое матричное обобщение применительно к анализу установившихся многочастотных режимов конвергентных НИЦ общего вида с НЭ, соотношения «вход – выход» которых представлены трансцендентными ФДН при любом способе задания исходных экспериментальных данных, включая табличный [20].

Полученные результаты позволяют одну из форм математической модели НИЦ (рис. 1) в однородном

У-координатном базисе представить следующей системой уравнений [20]:

$$\begin{cases} \mathbf{I}_{N, \bar{n}_{m,0}} = \mathbf{F}^{-1} \cdot \mathbf{J}_0 + \mathbf{F}^{-1} \cdot [\mathbf{A} \mathbf{Y}_{b,0} \mathbf{A}^t] \cdot [\mathbf{U}_{T, \bar{n}_{m,0}} - \mathbf{E}_{\bar{n}_{m,0}}]; & (5a) \\ \mathbf{I}_{N, \bar{n}_{l,+}} = \mathbf{F}^{-1} \cdot [\mathbf{A} \mathbf{Y}_{b,l+} \mathbf{A}^t] \cdot [\mathbf{U}_{T, \bar{n}_{l,+}} - \mathbf{E}_{\bar{n}_{l,+}}], & (5b) \end{cases}$$

где $\mathbf{I}_{N, \bar{n}_{m,0}}$ и $\mathbf{I}_{N, \bar{n}_{l,+}}$ – векторы нелинейных токов, являющиеся функциями напряжений на линейных элементах дерева графа НИЦ, параллельно которым включены инкрементальные НЭ порядка $k \geq 2$, при расчете величин DC -компонент и комплексных амплитуд КК на выходах НИЦ, соответственно; \mathbf{F} -матрица главных контуров графа эквивалентной схемы НИЦ; $\bar{n}_{m,0}$ и $\bar{n}_{l,+}$ – индексы векторов, соответствующих DC -компонентам с частотами $\omega_{\bar{n}_m} = 0$ и КК с частотами $\omega_{\bar{n}_l} > 0$ в уравнениях (5a) и (5б); \mathbf{J}_0 – вектор-столбец, учитывающий присутствие постоянных источников тока обобщенных ветвей графа НИЦ; \mathbf{A} – матрица инцидентий графа НИЦ; $\mathbf{Y}_{b,0}$, $\mathbf{Y}_{b,l+}$ – диагональная матрица проводимостей линейных двухполюсников в (5a) и (5б); $\mathbf{A} \mathbf{Y}_{b,l+} \mathbf{A}^t$ – матрица узловых проводимостей эквивалентной схемы НИЦ, рассчитываемая на частотах КК; $\mathbf{U}_{T, \bar{n}_{m,0}}$ и $\mathbf{U}_{T, \bar{n}_{l,+}}$ – векторы амплитуд напряжений на линейных элементах фундаментального дерева графа НИЦ; $\mathbf{E}_{\bar{n}_{m,0}} = [0, \dots, \dot{E}_{j, \bar{n}_{m,0}}, \dots, 0]^t$ – вектор-столбец узловых напряжений дерева графа НИЦ, ненулевой элемент которого, соответствующий j -узлу, содержит постоянную составляющую воздействия (1) и гармоническую ЭДС с набором коэффициентов $\bar{n}_{m,0}$, соответствующим m -й DC -компоненте отклика; $\mathbf{E}_{\bar{n}_{l,+}} = [0, \dots, \dot{E}_{j, \bar{n}_{l,+}}, \dots, 0]^t$ – вектор-столбец узловых напряжений дерева графа, ненулевой элемент которого, соответствующий j -узлу, содержит гармоническую ЭДС \dot{E}_{j, \bar{n}_l} , частота которой равна частоте l -го КК. В том случае, когда элемент дерева графа НИЦ не содержит источника постоянной ЭДС (источника постоянного тока), соответствующая ему координата вектора $\mathbf{E}_{k, \bar{n}_l}$ (соответственно, \mathbf{J}_0) равна 0.

Уравнение (5a), согласно [20], описывает $N_{\Sigma DC}$ DC -компонент установившегося отклика НИЦ, а уравнение (5б) характеризует $N_{(2\beta+n)}$ комбинационных и гармонических колебаний установившегося отклика.

Предложенный в [20] метод решения приведенной матричной системы уравнений (5) позволяет определить любые характеристики исследуемой НИЦ в условиях воздействия (1) с произвольным спектром, минуя расчет переходного режима и спектрального

анализа. В отличие от предложенных ранее, данный метод является, по сути дела, аналитическим и имеет перспективы своего дальнейшего развития для других классов НИЦ с учетом их режимов функционирования.

Одним из определяющих аспектов становления и дальнейшего развития КУДН является разработка методов и устройств измерения и определения параметров многомерных нелинейных передаточных функций и импульсных характеристик нелинейных устройств и систем с необходимой точностью и достоверностью. Существующие устройства, например [21, 22], не позволяют в полной мере решить данную задачу.

Для использования преимуществ предлагаемой КУДН и расширения функциональных возможностей измерителей параметров многомерных характеристик нелинейных устройств и систем, повышения точности и достоверности результатов измерений разработан новый подход и предложен ряд устройств, технические решения которых защищены патентами Украины, в частности [23, 24]. Эти устройства способны решать большинство задач, связанных с экспериментальными исследованиями установившихся многочастотных нелинейных процессов и явлений в различных областях техники.

При использовании всей совокупности составляющих предлагаемой КУДН (математические модели, методы, методики эксперимента, устройства измерения) появляется возможность определить параметры необходимых многомерных характеристик НИЦ, рассчитать частотные зависимости критериев нелинейных искажений (КНИ), комплексные амплитуды любой компоненты выходного сигнала, что позволяет с достаточной степенью корректности и строгости оценить потенциальные возможности разработанного устройства или ФМ и принять обоснованное решение по его эффективному применению.

Однако, при использовании структурных методов синтеза для разработки сложных микро- и радиоэлектронных систем существующий подход к принятию подобных решений теряет свой смысл, поскольку процесс разработки заключается в выборе совокупности ФМ, реализующих заданный алгоритм обработки входного сигнала и удовлетворяющих требуемым критериям качества из множества аналогичных устройств, предлагаемых различными фирмами-производителями. В этом случае необходимо ввести некоторые дополнительные характеристики для сравнения параметров и свойств объектов одноименного класса.

Для обоснования принятия эффективных решений в отношении качества используемых ФМ при структурном методе проектирования и разработки микро- и радиоэлектронных систем введем определенное обобщенное соотношение «вход – выход» для объектов исследования, относящихся к нелинейным неавтономным системам.

Сформулируем определение этой характеристики, которая в равной степени относится как к схемотехническим, так и к структурным конфигурациям.

Определение. Под обобщенной мгновенной динамической характеристикой (ОМДХ) по i -входу и j -выходу активного нелинейного инерционного многополюсника понимается функциональная зависимость мгновенного значения нормы отклика его j -го выхода от мгновенного значения нормы воздействия на i -м входе.

Введение ОМДХ позволяет решить несколько задач, связанных с обобщением и унификацией зависимостей подобного рода для объектов исследования самой различной физической природы. Для сравнения ОМДХ объектов исследования, принадлежащих одноименному классу, предлагается ввести понятие обобщенной нормированной ОМДХ (ОНМДХ), под которой понимается зависимость вида

$$\|W_{\text{вых}}(t)\|/\|W_{\text{вых}}(t)\|_{\text{max}} = F[\|W_{\text{вх}}(t)\|/\|W_{\lambda}\|], \quad (6)$$

где $\|W(\cdot)\|$ – норма сигнала, а $\|W_{\lambda}\|$ – норма входного сигнала, отвечающая эффекту пробоя активного компонента (например, пробоя электровакуумной лампы, лавинного пробоя p - n -перехода полупроводникового компонента).

Для практического введения ОМДХ и ОНМДХ необходимо использовать уже известные методы определения параметров постоянной составляющей, а также параметров гармонических и комбинационных компонент спектра отклика активного нелинейного безынерционного либо инерционного четырехполюсника на основе известного спектра воздействия. К ним можно отнести метод измерения параметров нелинейных искажений (например, критериев нелинейных искажений (КНИ) [3, 4]), методы измерения мощности с использованием датчиков Холла [1], а также устройства, методы и методики измерения и определения параметров многомерных характеристик НИЦ и систем, разработанные авторами [23, 24].

Следует подчеркнуть, что для корректной оценки потенциальных возможностей разработанного устройства, ФМ или системы, реализующих тот или иной алгоритм преобразования информационного сигнала, можно использовать модификации ОМДХ, учитывающие зависимость нормы выходного сигнала,

характеризующего результат полезного преобразования, от нормы воздействия, либо зависимость нормы компонент отклика, искажающих этот процесс, от нормы воздействия.

Основой для математического моделирования ОМДХ и ОНМДХ на основе КУДН и ФДН служат методы анализа, предложенные в [6, 14, 15, 20], а также методы структурной и параметрической идентификации нелинейных систем, разработанные в [25, 26].

Таким образом, задача исследования объектов рассматриваемого класса в рамках предлагаемой КУДН может быть сведена к решению задач структурной и параметрической идентификации НИЦ, эквивалентной объекту, и созданию ее обобщающей физико-математической модели.

Суммируя изложенное, сжатый смысл КУДН можно сформулировать так:

1. С точки зрения универсального информационного подхода объект исследования, принадлежащий заданному множеству, рассматривается как неавтономный нелинейный, в общем случае, инерционный $2N$ -полюсник с n входами и m выходами, а на структурном уровне – как неавтономная нелинейная, в общем случае, инерционная система с n входами и m выходами ($n, m \leq N$).

2. Внутренняя структура объекта исследования полагается либо полностью, либо частично неизвестной, т. е. объект исследования рассматривается либо как «черный», либо как «серый» ящик. В том случае, когда внутренняя структура объекта известна, процесс исследования объекта значительно упрощается, поскольку из алгоритма построения обобщающей физико-математической модели исключается ряд операций.

3. Воздействие по каждому из входов исследуемого объекта представляет собой МПТС вида (1) с конечным спектром и широким динамическим диапазоном (ДД).

4. Режим функционирования исследуемого объекта полагается установившимся.

5. Эквивалентный исследуемому объекту нелинейный инерционный $2N$ -полюсник представляется в виде двух каскадно-соединенных $2N$ -полюсников, один из которых является нелинейным безынерционным, а другой – линейным инерционным, что соответствует выводам теоремы, доказанной в [6].

6. Синтез обобщенной физико-математической модели исследуемого объекта на основе КУДН сводится к выполнению следующих операций:

а) экспериментального исследования соотношения «вход – выход» исследуемого объекта на задан-

ном множестве МПТС по каждой из пар «вход – выход»;

б) аппроксимации полученного множества характеристик «вход – выход» с помощью соответствующей ФДН;

в) полной и/или частичной идентификации внутренней структуры на основе полученного множества аппроксимированных характеристик «вход – выход»;

г) параметрической идентификации нелинейного безынерционного и линейного инерционного $2N$ -полосников с помощью множества аппроксимированных характеристик «вход – выход» и ФДН, учитывающих свойство неавтономности или автономности объекта;

д) определения комплексных амплитуд составляющих установившегося отклика на j -м выходе идентифицированного нелинейного $2N$ -полосника с n входами и m выходами ($n, m \leq N$) на МПТС его i -го входа, $i = \overline{1, n}$; $j = \overline{1, m}$;

е) расчета ОМДХ, полученной в результате структурной и параметрической идентификации нелинейного $2N$ -полосника, эквивалентного объекту исследования;

ж) сравнения расчетных и экспериментальных ОМДХ исследуемого объекта по каждой паре «вход – выход» и принятия решения по продолжению или же завершению итерационного процесса исследования по заданному критерию и заданной точности моделирования.

Таким образом, все этапы исследования, предусмотренные КУДН, достаточно наглядна, допускают достаточно простую физическую интерпретацию и при этом позволяют получить исчерпывающую информацию о характеристиках объекта, включая ОМДХ.

Возможная траектория ОМДХ, как одной из информационно емких характеристик исследуемого объекта, обусловлена прежде всего алгоритмами и принципами обработки информационного сигнала, заложенными в самом объекте исследования, а также характером физических процессов, сопровождающих их реализацию. Однако, при всем бесконечном разнообразии алгоритмов и принципов обработки информационного сигнала график функциональной зависимости «вход – выход» характеризуется неотрицательной крутизной.

Анализ результатов многочисленных работ данного направления позволяет выделить семь участков, которые охватывают весь возможный ДД по входному воздействию. График одной из возможных реализаций подобной обобщенной зависимости для объектов исследования данного класса приведен на рис. 3.

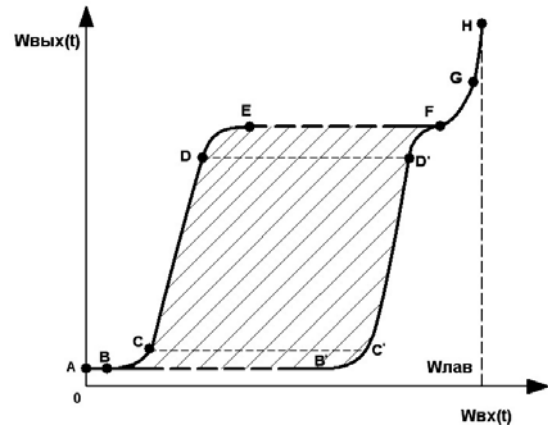


Рис. 3. ОМДХ нелинейной неавтономной системы при активном методе ее исследования

Три участка ОМДХ – АВ, CD, EF являются основными, а остальные четыре – переходными (вспомогательными). Заштрихованная область соответствует диапазону возможных изменений норм воздействия и отклика.

Каждому из семи участков ОМДХ или всей их совокупности можно поставить в частичное или полное соответствие участки статических и/или динамических характеристик микроселективных активных компонентов, структур, устройств и сред различного назначения. Используя первые пять участков ОМДХ, наиболее широко используемых на практике, можно определить значение нормы отклика, соответствующее нижнему и верхнему уровню входного воздействия, что позволяет определить соответствующие ДД по входу и выходу.

Одной из аналогий приведенной зависимости в радиотехнической литературе для переменных величин является амплитудная характеристика (АХ), а для режимов по постоянному току – семейства статических характеристик (входных, передаточных, выходных).

Как известно, экспериментальные статические характеристики полупроводниковых компонентов широко используются для характеристики их свойств в режимах постоянных токов и напряжений. Однако в технической литературе крайне редко можно найти аналитические соотношения, которые могли бы описать поведение отдельного активного прибора во всем диапазоне изменений напряжений и токов в статическом режиме. Это касается не только микроселективных полупроводниковых компонентов и структур, но и электровакуумных ламп и приборов СВЧ. Использование КУДН и ФДН позволяет решить эту задачу для активных компонентов различных видов. В [10] показана эффективность использования

ФДН для аппроксимации статических характеристик биполярных, полевых и СИТ-транзисторов.

Примером возможностей разработанной КУДН может служить решение задачи моделирования базового усилительного каскада на биполярном транзисторе, результаты которого приведены в [6]. Полученные результаты позволили решить задачу аппроксимации известной G -функции (введенной Э. Б. Грибовым в [27]) с помощью единственной трансцендентной ФДН, хотя сам автор этой функции считал эту задачу неразрешимой.

В качестве дополнительных подтверждений состоятельности КУДН рассмотрим несколько примеров из работ [11–13]. Например, в [11, 12] получено теоретическое соотношение для математического описания процессов, происходящих в фото- и терморегистрирующих средах, следующего вида

$$\alpha(t) = 1 - A \cdot \{1 + B \cdot \exp[-X(t)]\}, \quad (7)$$

где $\alpha(t)$ – степень превращения, A и B – взаимосвязанные нормировочные коэффициенты, определяемые начальными условиями процесса, которые характеризуют размерность протекания реакции в объеме реагента; $X(t) = \left(\int_0^t p(t) dt\right)$; $p(t)$ – функция оценки вероятности протекания процесса.

Более сложные виды кинетических кривых, часто наблюдаемые в ходе экспериментов и исследований топохимических превращений, описываются, согласно [12], модификацией соотношения (7) с аргументом $X(t) = \sum_{i=1}^n k_i \cdot \alpha_i \cdot t^{m_i}$, где $\sum_{i=1}^n \alpha_i = 1$; α_i – весовое содержание (доля) i -го ингредиента; n – количество

ингредиентов; m_i – параметры интенсивности процесса.

Анализ соотношения (7) и его модификаций позволяет отнести его к классу ФДН.

Таким образом, суммируя изложенное, можно утверждать, что КУДН получила подтверждение своей состоятельности не только с математической, но и с физической точки зрения, т. е. на феноменологическом уровне.

В качестве примера комплексного применения методов КУДН приведем результаты моделирования установленного режима нелинейного безынерционного устройства преобразователя частоты (УПЧ) (рис. 4) на двузатворном полевом транзисторе (ПТ) типа 2П306Б при моно- и бигармоническом воздействии. На рис. 5 и 6 приведены семейства экспериментальных (точечные кривые: I_{drain4} , I_{drain3} , I_{drain1} – по первому затвору и I_5 , I_4 , I_3 , I_1 – по второму затвору) и расчетных (непрерывные кривые: Y_{mod4} , Y_{mod3} , Y_{mod2} – по первому затвору и Y_{mod5} , Y_{mod4} , Y_{mod3} , Y_{mod2} – по второму затвору) графиков ВАХ 2П306Б.

ВАХ по первому затвору (рис. 5) аппроксимированы ФДН (3) с наборами параметров: $Y_{\text{mod4}} - B_4 = 1,8$, $S_4 = 3,2$, $A_4 = 2,48$, $p_4 = 1,3$; $Y_{\text{mod3}} - B_3 = 0,9$, $S_3 = 1,5$, $A_3 = 4,8$, $p_3 = 1,61$; $Y_{\text{mod2}} - B_2 = 1,3$, $S_2 = 1,65$, $A_2 = 7,25$, $p_2 = 1,6$. ВАХ по второму затвору (рис. 6) аппроксимированы ФДН (3) с наборами параметров, соответствующих графикам: $Y_{\text{mod5}} - Y_{50} = 0,06$, $B_5 = 0,8$, $S_5 = 3,2$, $A_5 = 2,24$, $p_5 = 3$; $Y_{\text{mod4}} - Y_{40} = 0,03$; $B_4 = 1,25$, $S_4 = 1,8$, $A_4 = 5,3$, $p_4 = 1,95$; $Y_{\text{mod3}} - Y_{30} = 0,025$, $B_3 = 1,8$, $S_3 = 2,1$,

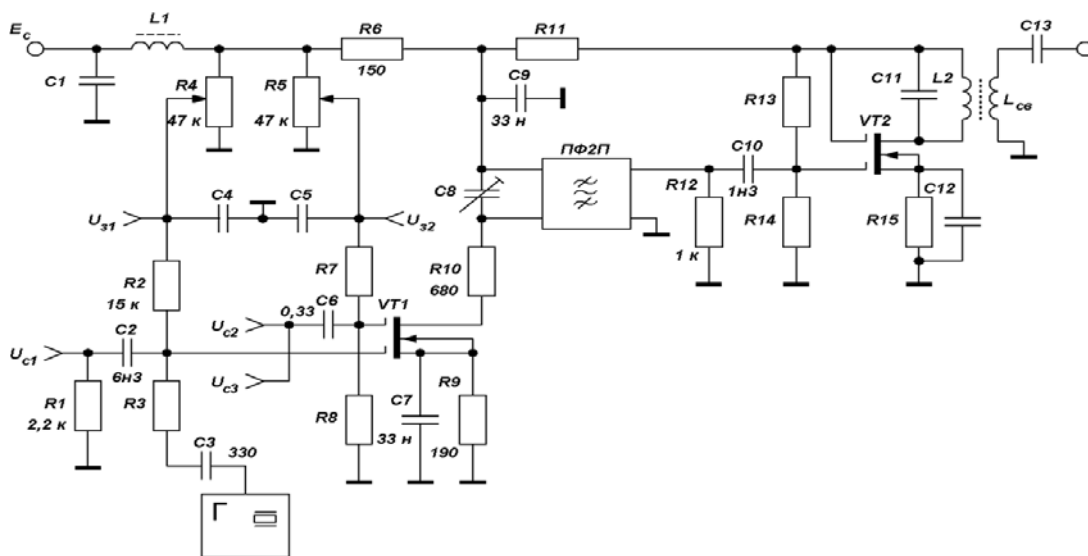


Рис. 4. Принципиальная схема УПЧ

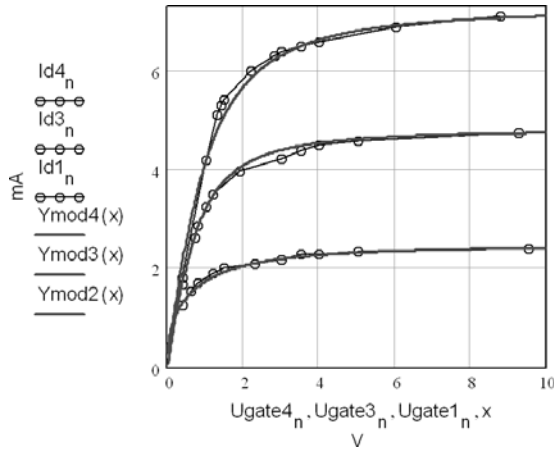


Рис. 5. ВАХ ПТ по первому затвору

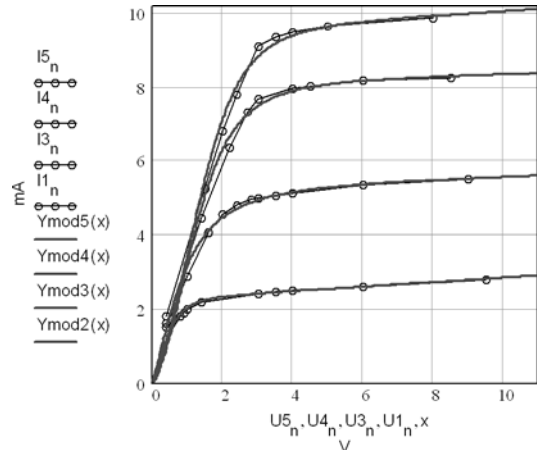


Рис. 6. ВАХ ПТ по второму затвору

$A_3 = 8,1$, $p_3 = 3,05$; $Y_{mod2} - Y_{20} = 0,055$; $B_2 = 2,2$, $S_2 = 2,9$, $A_2 = 9,5$, $p_2 = 3,7$.

Из рис. 5 и 6 следует, что свойства квадратичности ВАХ по каждому из затворов не наблюдается, как это было принято считать ранее, например в [4]. Этим и объясняется присутствие в спектре отклика УПЧ КК выше второго порядка [28].

На рис. 7 приведены графики АХ УПЧ: точечная кривая – эксперимент, а сплошная – аппроксимация; прямая линия – АХ идеального линейного ПЧ. Для ее моделирования использовалась ФДН $Y_{mod}(x)$ (3) с набором параметров: $Y_0 = 0$; $A = 2,35$; $p = 2,5$; $S = 2,35$; $B = 600$. Определение параметров ФДН осуществлялось с помощью методики [8]. Точность моделирования АХ – не ниже 0,9 %.

На рис. 8 приведены нормированные (относительно нормы воздействия) зависимости амплитуд КК первых трех порядков на выходе кварцевого фильтра: точечные кривые – результаты эксперимента, а непрерывные зависимости получены расчетным путем с применением ФДН (3), модели [16] и метода анализа спектра отклика на воздействие (1) [14].

Графики 1–3 (зависимости $dB_{1n} = F(U_{1n})$, $dB_{2n} = F(U_{2n})$, $dB_{4n} = F(U_{4n})$) соответствуют КК вида $f_1 - f_r$, графики 4–6 (зависимости $dB_{6n} = F(U_{6n})$, $dB_{12n} = F(U_{12n})$, $dB_{13n} = F(U_{13n})$) – КК вида $f_1 + f_2 \pm f_r$, а графики 7–9 (зависимости $dB_{8n} = F(U_{8n})$, $dB_{10n} = F(U_{10n})$, $dB_{11n} = F(U_{11n})$) – КК вида $f_1 - f_2$. Нумерация графиков соответствует их расположению сверху вниз. Режим ПТ по постоянному току определяется следующими величинами: $E_{31} = 0,85$ В, $E_{32} = 0,85$ В; ток стока $I_c = 1$ мА. Частота гетеродина $f_r = 1,344$ МГц.

Частоты компонент двухтонового сигнала равны: при преобразовании «вверх» – $f_{b1} = 995$ кГц, $f_{b2} =$

$= 161$ кГц, при преобразовании «вниз» – $f_{n1} = 1995$ кГц, $f_{n2} = 1849$ кГц.

Центральная частота фильтра $f_0 = 2,5$ МГц. Приведенные характеристики для КК вида $f_{пч} = f_{b1} + f_{b2} + f_r$ и $f_{пч} = f_{n1} + f_{n2} - f_r$ отличаются по амплитуде менее чем на 3,5 %, что позволяет рассматривать УПЧ в режиме безынерционного нелинейного преобразования. Граничная частота по крутизне f_s для ПТ лежит в диапазоне нескольких сот МГц.

Соответствие приведенных экспериментальных и теоретических результатов, характеризующих режим УПЧ, и отсутствие нулевых значений в графиках нормированных зависимостей КНИ, использующих КК третьего порядка в моделях НЭ на основе ФДН (3), подтверждает корректность применения предложенной математической модели, параметры которой рассчитывались на основе КУДН с помощью метода [8] и с учетом условий теоремы [9].

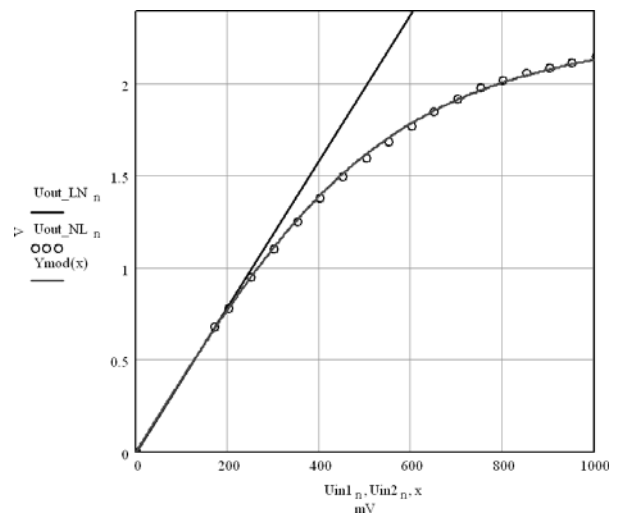


Рис. 7. Амплитудная характеристика УПЧ

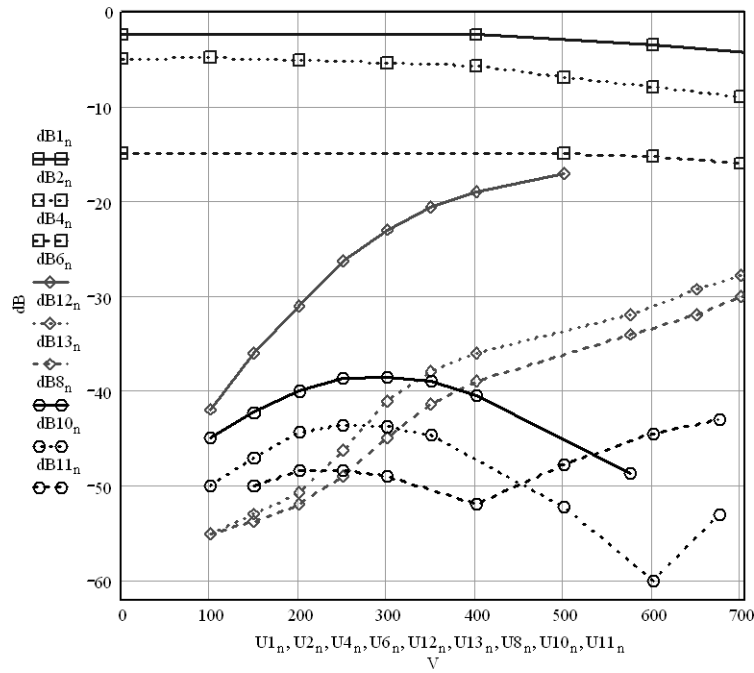


Рис. 8. Нормированные зависимости амплитуд КК 2-го и 3-го порядка

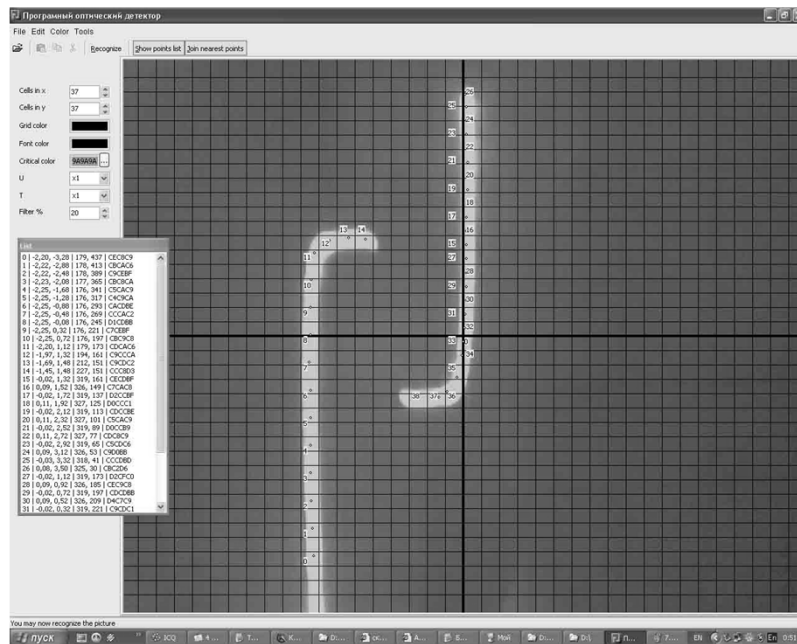


Рис. 9

В качестве заключительных примеров, подтверждающих эффективность моделирования динамических характеристик более сложных структур и устройств с помощью ФДН, приведем экспериментальные и аппроксимированные характеристики ОМДХ негаваристора [29] (рис. 9) и одного из вариантов схемы аттрактора Л. Чуа [30] (рис. 10).

На рис. 9 приведена экспериментальная ОМДХ (сплошная линия на экране осциллографа) и ее аппроксимация ФДН вида (3) (точечная кривая) для схемы негаваристора N -вида на негатроне R -типа [29].

На рис. 10 приведены изображения экспериментальной ОМДХ (сплошная линия на экране осциллографа) и ее аппроксимации трехсегментным графиком

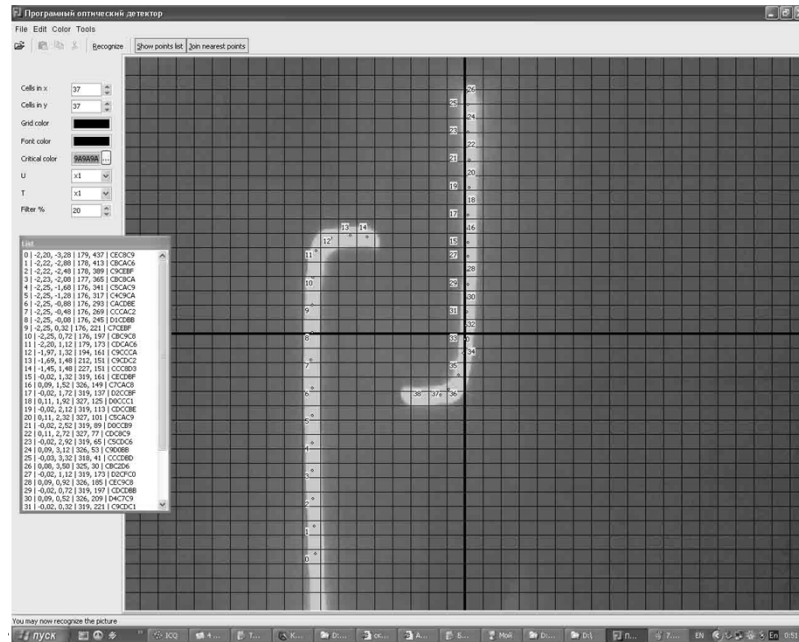


Рис. 10

ФДН вида (3) (точная кривая) для одной из схем трактора Л. Чуа [30]. Точечные кривые на рис. 9 и 10 получены с помощью специально разработанной программы цифровой обработки изображений.

Необходимо отметить, что теоретические и экспериментальные основы предложенной КУДН не претендуют на исчерпывающую полноту, хотя краеугольный камень их уже заложен. КУДН не противоречит подходу к моделированию негармонических процессов в энергетике на основе геометрического представления, изложенному в работе [5]. В то же время, даже приведенные примеры с исчерпывающей полнотой демонстрируют высокие потенциальные возможности предлагаемой КУДН для исследования и моделирования микрорелектронных НИЦ и систем в условиях многочастотного полиамплитудно-го воздействия.

ВЫВОДЫ

1. На основе математического аппарата ФРВГ и ФДН предложена обобщающая концепция управляемого динамического насыщения для отображения, моделирования и исследования режимов микрорелектронных активных компонентов, структур, устройств, принадлежащих к классам нелинейных неконвергентных цепей, на основе системного информационного подхода в условиях действия многочастотных сигналов с широким ДД. На ряде примеров состоятельность концепции подтверждена как с математической точки зрения, так и на феноменологическом уровне.

2. Отличительной особенностью предлагаемого подхода является то, что использование ФДН, параметры которых могут быть определены непосредственно из эквивалентной схемы исследуемого объекта, позволяет целенаправленно изменять его характеристики в соответствии с выбранным критерием качества.

3. Предложенная концепция имеет перспективы широкого применения и развития, поскольку позволяет с единых позиций подходить к исследованию как сложных процессов преобразования информационных сигналов нелинейными объектами различной физической природы, так и взаимодействия волновых процессов с веществом.

СПИСОК ЛИТЕРАТУРЫ

1. Макс Ж. Методы и техника обработки сигналов при физических измерениях : в 2 т. : пер. с франц. / Ж. Макс. – М. : Мир, 1983. – 2 т.
2. Голяницкий И. А. Математические модели и методы в радиосвязи / под ред. Ю. А. Громакова / И. А. Голяницкий. – М. : Эко-трендз, 2005. – 440 с. : ил.
3. Сверкунов Ю. Д. Идентификация и контроль качества нелинейных элементов радиоэлектронных систем (спектральный метод) / Ю. Д. Сверкунов. – М. : Энергия, 1975. – 96 с. : ил.
4. Богданович Б. М. Методы нелинейных функционалов в теории электрической связи / Б. М. Богданович, Л. А. Черкас, Е. В. Задедюрин, Ю. М. Вувуникян. – М. : Радио и связь, 1990. – 280 с.
5. Горбачев М. Н. Новая концепция развития теории математического моделирования негармонических электроэнергетических процессов в управляемых электрических и радиотехнических цепях и системах на основе геометрических представлений / М. Н. Горба-

- чев, А. Д. Милка // Радіоелектроніка, інформатика, управління. – 2002. – № 2. – С. 134–138.
6. Гулін С. П. Анализ спектра отклика нелинейности, представленной аналитической трансцендентной функцией, на многочастотное воздействие большой нормы / С. П. Гулін // Радіоелектроніка, інформатика, управління. – 2004. – № 1. – С. 21–28.
 7. Чуа Л. О. Машинный анализ электронных схем : алгоритмы и вычислительные методы : пер. с англ. / Л. О. Чуа, Пен-Мин Лин. – М.: Энергия, 1980. – 640 с. : ил.
 8. Гулін С. П. Определение параметров адаптивной модели нелинейных компонентов, представленной аналитической трансцендентной функцией, на основе экспериментальных характеристик / С. П. Гулін // Радіоелектроніка, інформатика, управління. – 2005. – № 2. – С. 25–32.
 9. Гулін С. П. Условия применимости модели динамического насыщения в задачах анализа спектра отклика нелинейных устройств / С. П. Гулін // Радіоелектроніка, інформатика, управління. – 2005. – № 2(14). – С. 21–28.
 10. Гулін С. П. Аппроксимация нелинейных зависимостей аналитической трансцендентной функцией высшего порядка / С. П. Гулін, А. С. Гулін. // 2-й Международный радиоэлектронный форум «Прикладная радиоэлектроника. Состояние и перспективы развития», сентябрь 19–23, МРФ–2005 : сб. научн. тр. Т. 5. СВЧ и оптоэлектроника. – Харьков : АНПРЭ, ХНУРЭ, 2005. – С. 121–124.
 11. Шустов М. А. Кинетика и формулировки законов фотохимических превращений / М. А. Шустов // Известия Томского политехнического института. – 2001. – Т. 304, вып. 1. – С. 117–122.
 12. Бойченко А. П. Основы газоразрядной фотографии / А. П. Бойченко, М. А. Шустов. – Томск : STT, 2004. – 312 с.
 13. Шустов М. А. Приборы, методы и среды регистрации и обработки информации на основе бистабильных и промежуточных состояний [Электронный ресурс] : дис. ... доктора техн. наук : 01.04.01 : защищена 14.06.07 : утв. 18.01.08 / Шустов Михаил Анатольевич. – Электрон. дан. – М. : РГБ, 2007. – 286 с. : ил. Режим доступа: <http://diss.rsi.ru/07/0806/070806026.pdf>, свободный. – Загл. с экрана.
 14. Гулін С. П. Анализ спектра отклика нелинейности, представленной функцией динамического насыщения, при многочастотном воздействии / С. П. Гулін // Радіоелектроніка, інформатика, управління. – 2008. – № 1. – С. 31–37.
 15. Гулін С. П. Расчет спектра отклика нелинейности, представленной логарифмической функцией динамического насыщения, при многочастотном воздействии // С. П. Гулін, А. С. Гулін // Радиотехника. – 2009. – № 159. – С. 373–376.
 16. Gulín S. P. Generalized model of radio engineering link on basis principle of the guided run-time saturation / S. P. Gulín, A. S. Gulín. // Наукові записки українського науково-дослідного інституту зв'язку (УНДІЗ). – 2009. – № 4(12). – С. 52–58.
 17. Гулін С. П. Характеризация резистивных нелинейных элементов на основе концепции управляемого динамического насыщения / С. П. Гулін, А. С. Гулін. – Радіоелектроніка, інформатика, управління. – 2010. – № 1. – С. 31–37.
 18. Гулін С. П. Характеризация реактивных нелинейных элементов на основе концепции управляемого динамического насыщения / С. П. Гулін // Радиотехника. – 2010. – № 161. – С. 131–142.
 19. Гулін С. П. Расчет числа компонент установившегося отклика нелинейной инерционной цепи на многочастотное воздействие, содержащее постоянную составляющую / С. П. Гулін, А. М. Гулін. – Радиотехника. – 2010. – № 162. – С. 66–77.
 20. Гулін С. П. Метод анализа установившегося многочастотного режима конвергентной нелинейной цепи на основе функционального ряда Вольтерры – Гаусса и функций динамического насыщения / С. П. Гулін. – Радиотехника. – 2010. – № 163. – (в печати).
 21. Устройство для измерения многомерных передаточных функций высокочастотных узкополосных нелинейных систем : а. с. № 1532887 СССР : МКВ 5G 01 R 27/28 / Б. И. Ивлев, С. Ю. Матвеев, В. Р. Снурницын, С. В. Трушин (СРСР). – Заявл. 12.04.88 ; опубл. 30.12.89, Бюл. № 48. – 11 с.
 22. Устройство для измерения многомерных передаточных функций нелинейных систем : а. с. № 1626199 СССР, МКВ 5G 01 R 27/28 / Б. И. Ивлев, С. Ю. Матвеев, В. Р. Снурницын, С. В. Трушин (СРСР). – Опубл. 07.02.91, Бюл. № 5. – 8 с.
 23. Пристрій для визначення N -вимірних характеристик нелінійних систем : пат. 88065 Україна : МПК₈ G01 R 27/28 // С. П. Гулін, О. С. Гулін, В. П. Дмитренко, Д. М. Піза ; заявник та патентовласник Запорізький національний технічний університет. – Заявл. 26.10.2007 ; опубл. 10.09.2009, Бюл. № 17. – 16 стр.
 24. Пристрій для визначення параметрів N -вимірних функцій високочастотних вузькосмугових нелінійних систем : пат. 47356 Україна : МПК₇ G01 R 27/28 / С. П. Гулін, О. С. Гулін, В. П. Дмитренко ; заявник та патентовласник Запорізький національний технічний університет. – Заявл. 25.08.2009 ; опубл. 25.01.2010, Бюл. № 2. – 20 стр.
 25. Богатырев Е. А. Структурное отождествление макромоделей микрорезонаторных устройств / Е. А. Богатырев, С. П. Гулін // Радиотехника. – 1985. – № 6. – С. 39–42.
 26. Гулін С. П. Параметрична ідентифікація лінійного $2N$ -полосника в частотній області із застосуванням розгалуженого дробу / С. П. Гулін, О. С. Гулін // 2-й Международный радиоэлектронный форум «Прикладная радиоэлектроника. Состояние и перспективы развития», сентябрь 19–23, МРФ–2005 : сб. научн. тр. Т. 3. Информационные технологии и сети. – Харьков : АНПРЭ, ХНУРЭ, 2005. – С. 251–254.
 27. Грибов Э. Б. Нелинейные явления в приемопередающем тракте аппаратуры связи на транзисторах / Э. Б. Грибов. – М. : Связь, 1971. – 368 с.
 28. Бобков А. М. Аппроксимация характеристик нелинейного безынерционного элемента / А. М. Бобков, Н. Н. Яковлев // Радиотехника. – 1986. – № 5. – С. 25–26.
 29. Гулін С. П. Экспериментальное исследование характеристик двухтранзисторного эквивалента R -негатрона N -типа для систем передачи информации с хаотической несущей / С. П. Гулін, А. С. Гулін // Тезиси V Международной научно-технической конференции «Сучасні інформаційно-комунікаційні технології», 5–9 жовтня 2009 р., COMINFO «2009-Livadia». – Ливадия, 2009. – С. 69–70.
 30. Chua L. O. Panoramic views of strange attractors (display instrument) / L. O. Chua, T. Sugawara // Proceedings of the IEEE. – Aug. 1987. – Vol. 75, No. 8. – Pp. 1107–1120.

Надійшла 21.01.2010
Після доробки 08.04.2010

Гулін С. П., Гулін О. С., Дмитренко В. П., Піза Д. М.
КОНЦЕПЦІЯ КЕРОВАНОВОГО ДИНАМІЧНОГО НАСИЩЕННЯ ПРИ МОДЕЛЮВАННІ БАГАТОЧАСТОТНИХ УСТАЛЕНИХ РЕЖИМІВ МІКРОЕЛЕКТРОННИХ КОМПОНЕНТІВ, СТРУКТУР І ПРИСТРОЇВ

На основі математичного апарату функціональних рядів Вольтерри – Гаусса (ФРВГ) і функцій динамічного насичення (ФДН) запропоновано концепцію керованого динамічного насичення, що реалізує узагальнений підхід до дослідження і моделювання микрорезонаторних нелінійних інерційних ланцюгів і систем (НІЛС) в умовах багаточас-

тотної поліамплітудної дії. На низці прикладів обґрунтованість запропонованої концепції підтверджена як з математичної, так і з фізичної точки зору.

Ключові слова: Нелінійні інерційні ланцюги і системи, концепція керованого динамічного насичення, функціональні ряди Вольтерри – Гаусса, узагальнена миттєва динамічна характеристика, усталений відгук, багаточастотна поліамплітудна дія.

Gulin S. P., Gulin A. S., Dmitrenko V. P., Piza D. M.

CONCEPT OF CONTROLLED DYNAMIC SATURATION FOR MODELING MULTIFREQUENCY STEADY STATE MODES OF MICROELECTRONIC COMPONENTS, STRUCTURES AND DEVICES

The authors propose the concept of controlled dynamic saturation based on the body of mathematics of Volterra – Gauss functional series and dynamic saturation functions, which realizes a generalized approach to investigation and modeling of microelectronic nonlinear inertial circuits and systems in conditions of multifrequency polyamplitude influence. A number of examples prove validity of the proposed concept from both mathematical and physical standpoint.

Key words: nonlinear inertial circuits and systems, concept of controlled dynamic saturation, Volterra – Gauss functional series, generalized instantaneous dynamic characteristic, steady-state response, multifrequency polyamplitude influence.

УДК 621.3.049

Жавжаров Є. Л.¹, Матюшин В. М.²

¹Канд. фіз.-мат. наук, доцент Запорізького національного технічного університету
²Д-р фіз.-мат. наук, професор Запорізького національного технічного університету

ВПЛИВ АТОМАРНОГО ВОДНЮ НА СТРУКТУРИ МЕТАЛ – НАПІВПРОВІДНИК. АДГЕЗІЯ ТОНКИХ МЕТАЛЕВИХ ПЛІВОК

Досліджено вплив атомарного водню на зміну властивостей тонкоплівкових структур Cu-Si. Структури на основі тонких (20–500 нм) плівок міді піддавалися дії атомарного водню при концентраціях 10^{14} – 10^{19} м⁻³. Показано, що при такій взаємодії відбувається низькотемпературне відновлення оксиду міді, підвищуються адгезія та когезія плівок міді до підкладки Si.

Ключові слова: тонкі плівки, напівпровідник, кремній, атомарний водень, рекомбінація, адгезія, дифузія.

ВСТУП

Однією з проблем, що існують в тонкоплівковій технології мікро- та нанoeлектроніки, є адгезія плівок. Найпростішим способом збільшення адгезії є очищення поверхні підкладок від забруднюючих домішок. Адгезія підсилюється, якщо при контакті утворюється хімічний зв'язок або відбувається взаємна дифузія матеріалів.

Для досягнення високої адгезії застосовують різні методи, наприклад, між підкладкою і плівкою наносять спеціальні шари, що мають високий потенціал взаємодії як з підкладкою, так і з плівкою [1]. У якості підшарів в кремнієвій технології найчастіше використовують титан, хром і ванадій. У роботі [2] для підвищення адгезії плівок нікелю до кремнію запропоновано метод акустичної стимуляції, що дозволило збільшити адгезію на третину порівняно із стандартним методом осадження з газової фази. Для підвищення адгезії плівок використовують відпал, що ініціює дифузійні процеси на межі розділу [3], а також складні багатостадійні процеси [4], необхідні для

зміни та поліпшення їх фізичних, електричних та електрофізичних параметрів. Використання відпалу, дії частинок високих енергій [5] може супроводжуватися виникненням ряду негативних явищ, таких як перерозподіл вже введеної домішки в сусідніх ділянках, дифузія фонових домішок на велику глибину, виникнення і зростання паразитних плівок, виникнення значних механічних напруг на гетеромежі. При використанні частинок високих енергій у приповерхневих шарах може виникати і залишатися значна кількість дефектів.

Тому актуальним є створення альтернативних технологій підвищення адгезії плівок широкого кола металів до напівпровідникових підкладок.

Одним з перспективних напрямів, що дозволяють проводити цілеспрямовану зміну властивостей поверхні, є дія на неї активних газових часток, зокрема атомарного водню. Поверхня, будучи центром рекомбінації атомарного в молекулярний водень, виявляється під дією енергії, що виділяється при цьому

© Жавжаров Є. Л., Матюшин В. М., 2010

(4,48 eV на 1 акт рекомбінації) [6]. У ряді робіт показано, що ця взаємодія супроводжується такими ефектами як дифузія атомів плівки в підкладку [7], розпилення атомів плівки [8], хімічна реакція [9] і таке інше. На цій підставі видається можливим збільшення адгезії тонких плівок під дією атомарного водню.

Метою роботи є створення нової технології отримання тонкопліткових структур для мікро- та наноелектроніки з підвищеною адгезією плівки до підкладки за допомогою водневої обробки.

МЕТОДИКА ПРОВЕДЕННЯ ЕКСПЕРИМЕНТУ

Для виготовлення структур Cu-Si використовувалася *n*-кремній з питомим опором $\rho = 10 \text{ Ом}\cdot\text{см}$, щільністю дислокацій 10^2 см^{-2} , орієнтацією (111). Хімічно полірована поверхня кристалів кремнію була покрита шаром природного оксиду. Кремній очищувався шляхом знежирення в CCl_4 з подальшим промиванням деіонізованою водою. Плівка міді наносилася на кристали кремнію термічним випаровуванням у вакуумі при температурі підкладки 373 К, товщина плівки складала $\sim 500 \text{ нм}$, а також при температурах підкладки, близьких до кімнатної, товщина плівок складала $\sim 20 \text{ нм}$.

Атомарний водень утворювався шляхом дисоціації молекулярного водню в плазмі, утвореній високочастотним розрядом. Молекулярний водень виділявся з водного розчину КОН (20%) шляхом електролізу при щільності струму $0,01 \text{ А/см}^2$. Далі водень пропускався через форбалон для уловлювання крапель КОН, осушений стовпчик, заповнений силікагелем, і надходив до робочої камери. Концентрація атомів водню вимірювалася калориметричним способом і складала близько 10^{19} м^{-3} при тиску в робочій камері 20 Па.

Досліджувані структури розміщувалися на відстані 25 см від області розряду, що дозволило виключити потрапляння радикалів гідроксилу та іонів водню [10]. Час впливу атомарного водню на структури варіювався від 2 до 20 хв.

Зміни властивостей плівок міді досліджувалися декількома способами. Проводився візуальний контроль зміни морфології поверхні плівки в процесі її обробки атомарним воднем через мікроскоп Биолам-М та наступним фіксуванням змін шляхом фотографування. Також проводилася оцінка зміни величини адгезії мідних плівок шляхом проведення вимірів зусилля відриву мідної плівки після обробки структури в середовищі атомарного водню. За допомогою швидко висихаючого клею на основі ціанакрилату створювалося клейове з'єднання металевого стрижня діаметром

0,5 мм з поверхнею зразка, після чого вимірювалося зусилля відриву такої системи. Місце відриву стрижня фотографувалося через мікроскоп з метою контролю адгезійних і когезійних властивостей такої системи.

РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ І ЇХ ОБГОВОРЕННЯ

Дослідження показали, що під дією атомарного водню вже після 2 хвилин обробки структури в середовищі атомарного водню інтенсивно починають протікати процеси очищення поверхні кристала і відновлення окисної плівки міді. Процес відновлення досить інтенсивно протікає вже після 5 хвилин обробки, а після 20 хвилин оксидний шар практично повністю зникає. Дослідження залежності зусилля відриву клейового з'єднання (металевого стрижня до плівки міді) від часу обробки показало, що відрив з'єднання може відбуватися різними шляхами. Процес відриву клейового з'єднання може протікати за такими напрямками: відрив плівки міді від кристала, що характеризує адгезію Cu до кристала (площа кремнію S_{Si}); відрив чи розшарування плівки міді, що характеризує когезійні властивості плівки (частина плівки міді в місці відриву S_{Cu}); відрив клею, що характеризує адгезійні властивості клею (S_{kl}) та міцність самого клею.

На рис. 1 представлено графік залежності площі відриву для кожного з компонентів (тобто S_{Si} , S_{Cu} , S_{kl}) від часу обробки.

З графіку видно, що на початковому етапі дослідження (час обробки $t < 10 \text{ хв}$) розрив з'єднання відбувається в основному за рахунок розшарування плівки міді. При подальшій обробці ($t_{\text{обробки}} > 7 \text{ хв}$) спостерігається відрив плівки від кремнію, але із збільшенням часу обробки структури площа розриву по компонентах Cu-Cu й Cu-Si(SiO_2) прагне до нуля,

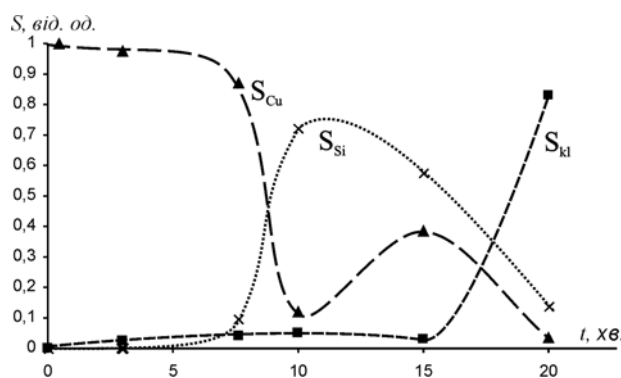


Рис. 1. Зміна характеру відриву від часу обробки атомарним воднем

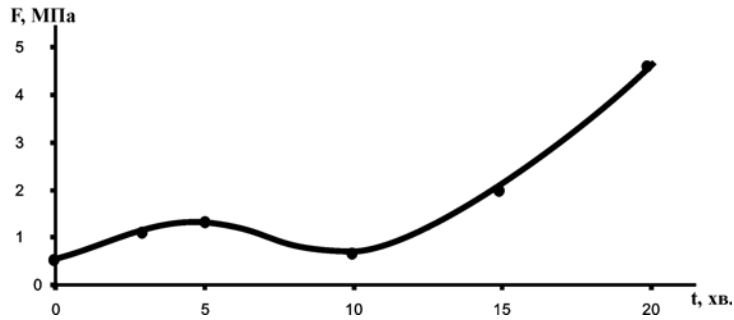


Рис. 2. Залежність зусилля відриву металевого стрижня від часу обробки структури «метал – напівпровідник»

а сам розрив відбувається переважно за рахунок клею. Тобто, найслабшим компонентом на момент обробки ~20 хвилин є клейове з'єднання.

Підрахунок площі відриву, а також експериментальні дані про масу вантажу, необхідного для розриву з'єднання, дозволило побудувати залежність зусилля відриву від часу обробки структури в середовищі атомарного водню з урахуванням площі відриву. На рис. 2 представлена усереднена для серії кристалів залежність зусилля відриву від часу обробки.

Як видно з рисунку, обробка плівки в середовищі атомарного водню приводить до збільшення зусилля відриву. При співвіднесенні графіків, представлених на рис. 1 і 2, можна відмітити, що зі збільшенням часу обробки в середовищі атомарного водню відбувається збільшення зусилля відриву. Таке явище пов'язується зі збільшенням адгезії плівки до кристала, а також збільшенням когезійних властивостей плівки, оскільки при тривалій обробці (понад 15 хв) розрив з'єднання відбувається переважно за рахунок розриву з'єднання «клей – металевий стрижень».

Зазначимо, що на залежності зусилля відриву від часу обробки (рис. 2) спостерігається ділянка спаду, яка відповідає часу обробки структури ~10 хв. Цьому часу відповідає стан, коли відрив відбувається переважно за межею поділу плівка – кремній. Враховуючи результати проведених раніше досліджень [11], відповідальним за це можна вважати процес накопичення водню на межі поділу «плівка – підкладка». Підтвердженням цього факту є графік залежності зусилля відриву плівок міді від часу обробки при менших концентраціях атомарного водню (рис. 3).

В цьому випадку, так само як і для залежності рис. 3, відбувається загальне збільшення адгезії, однак не спостерігається ділянки спаду. Обробка атомарним воднем плівок, отриманих при різних швидкостях наплення, приводить до зміни адгезії вихідної плівки до підкладки.

Згідно з [8], під дією атомарного водню можуть ініціюватися процеси дифузії атомів металу в матеріал підкладки. Отже, можна припустити, що і в даному випадку зміна властивостей контакту може бути обумовлена процесами масоперенесення атомів в приповерхневих шарах. Для дослідження цього питання було проведено визначення складу поверхні та мікропрофілювання гетероструктури методом Оже-спектроскопії. Під час профілювання поверхня травилася іонами аргону зі швидкістю ~0,5 нм/хв.

Досліджувалися декілька зразків гетероструктур Cu-Si: контрольна гетероструктура, оброблена на протязі 10 хвилин та 40 хвилин. Товщина плівки складала ~20 нм. З розподілу за глибиною складу контрольного зразку, представленого на рис. 4, можна бачити, що товщина плівки дійсно складає ~20 нм. Поверхня контрольної плівки містить забруднення у вигляді вуглецю та кисню. Оскільки концентрація кисню за глибиною спадає майже до нуля, то кисень може входити

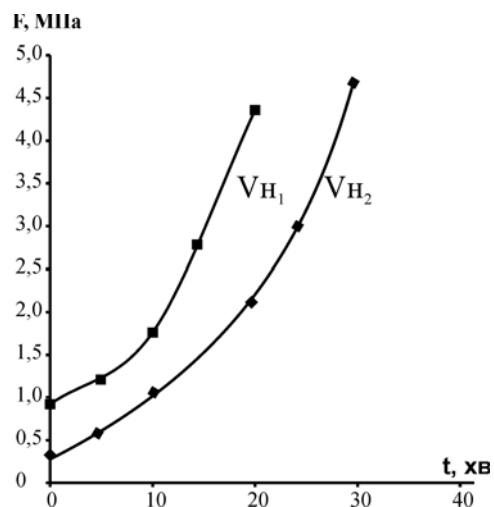


Рис. 3. Залежність зусилля відриву від часу обробки плівок Cu при $C_n \approx 10^{18} \text{ м}^{-3}$, отриманих при різних швидкостях наплення V_n ($V_{n1} > V_{n2}$)

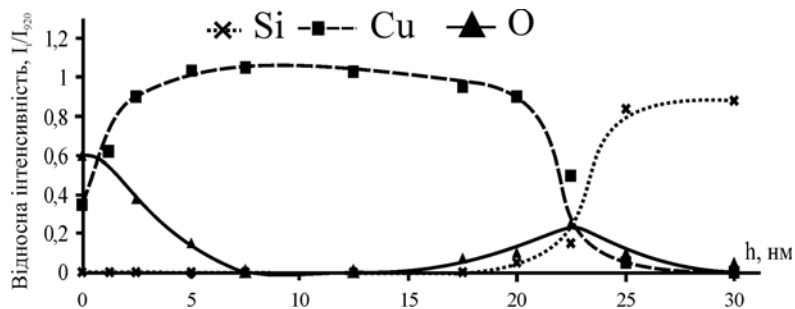


Рис. 4. Домішковий склад приповерхневих шарів контрольного зразка Cu-Si

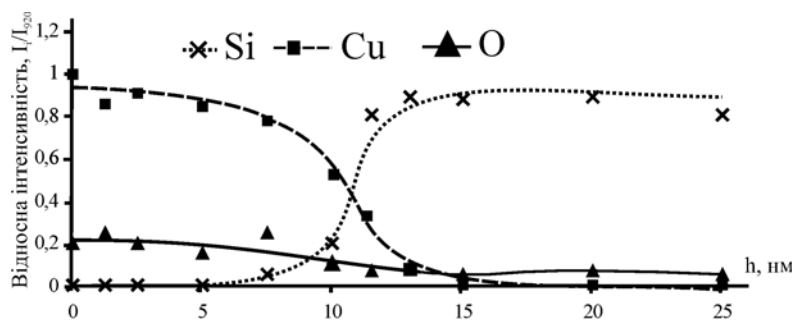


Рис. 5. Домішковий склад поверхневих шарів зразка Cu-Si, обробленого в Н на протязі 40 хв

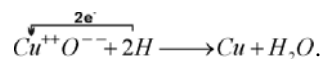
до складу окислу мідної плівки. На межі поділу металева плівка – напівпровідник спостерігається дещо збільшена концентрація кисню, що пов’язується з наявністю тонкого шару природного оксиду кремнію SiO₂, який встиг утворитися при контакті з атмосферою.

При порівнянні профілів розподілу елементів в гетероструктурах Cu-Si контрольного (рис. 4) та обробленого в Н (рис. 5) зразків можна бачити, що обробка атомарним воднем приводить до зменшення товщини плівки, оскільки границя поділу спостерігається вже на глибині ~10 нм.

Крім того, на глибині більшій ніж у контрольного зразка спостерігається наявність міді, тобто можна казати про дифузію атомів міді в приповерхневі шари Si. Це пояснює зміну адгезії тонких металевих плівок до напівпровідникових підкладок під дією атомарного водню.

Для пояснення отриманих експериментальних залежностей можна запропонувати таку якісну модель взаємодії атомарного водню із структурами на основі тонких плівок Cu. Плівка міді (сама мідна плівка та, якщо є, окисний шар), отримана без додаткового нагріву підкладки, до взаємодії з атомарним воднем досить пухка і має невелику адгезію, зусилля відриву складає менше 1 МПа. Обробка зразків у середовищі атомарного водню приводить до збільшення зусилля

відриву. При цьому збільшення зусилля на відрив на початковому етапі можна пов’язати із збільшенням когезії плівки та відновленням окису відповідно до реакції [12]:



Одночасно з відновленням, за рахунок дисипації енергії рекомбінації атомів водню на поверхні плівки, протікає процес її ущільнення, що приводить до збільшення адгезії і когезії плівки. Так, при проведенні експерименту відбувається розігрів структур до температур ~330–350 К. На початкових стадіях обробки при концентраціях атомарного водню не менше $\sim 10^{19} \text{ м}^{-3}$ протікає конкуруючий для зазначених процесів процес, що приводить до зменшення зусилля відриву. Таким процесом є процес розчинення і дифузії водню в об’єм міді до межі розділу «мідь – підкладка». Відповідно до роботи [13], водень розчиняється в більшості металів, займаючи при цьому міжвузлове положення в ґратці. Крім того, внаслідок зрощення зерен, яке відбувається в напрямку від поверхні до підкладки, на межі плівка – підкладка може концентруватися певна кількість пор з воднем. Таким чином, зменшення зусилля відриву можна пов’язати з воднем в мідній плівці та на межі поділу мідь – кремній. Із збільшенням три-

валості дії атомарного водню на структури або при концентраціях $< \sim 10^{19} \text{ м}^{-3}$ домінуючими можна вважати процеси відновлення окису міді за умови його присутності, процес ущільнення мідної плівки під дією енергії рекомбінації та масоперенесення в приповерхневих шарах підкладки, що супроводжуються в загальному випадку збільшенням когезійних і адгезійних властивостей плівки міді ($t > 15 \text{ хв}$).

ВИСНОВКИ

Дослідження структур металева плівка – підкладка показали, що дія атомарного водню приводить до модифікації не лише властивостей плівки, але і властивостей контакту металева плівка – підкладка та властивостей кремнієвих підкладок, які входять до складу гетероструктур.

На поверхні металевих плівок за наявності оксиду відбуваються хімічні реакції відновлення окислів металів. Дисипація енергії рекомбінації атомарного водню в молекулярний приводить до:

– розігріву структури (температура залежить від підкладки, на якій розміщується кристал);

– протікання реакцій відновлення окису, очищення поверхні та процесу розпилення металевої плівки, що супроводжується збільшенням когезії плівок та зменшенням їх товщини;

– ініціалізації процесу масоперенесення в приповерхневих шарах підкладки, що приводить до збільшення адгезії плівок при їх обробці понад 15 хв (збільшення зусилля відриву в ~ 6 разів).

Дисипація енергії рекомбінації стимулює розпилення плівки, перш за все на межах контакту, там, де плівка має меншу товщину. Одночасно з процесами, викликаними дисипацією енергії рекомбінації, виникає дифузія водню крізь металеву плівку та процес розчинення водню в об'ємі плівки.

Виявлені ефекти – ущільнення, підвищення адгезії – можуть бути основою для технології низькотемпературного нанесення тонких плівок з поліпшеними параметрами.

СПИСОК ЛІТЕРАТУРИ

1. *Верещака А. С.* Режущие инструменты с износостойкими покрытиями / А.С. Верещака, И. П. Третьяков // М. : Машиностроение, 1986. – С. 41.
2. *Муминов Р. А.* Акустостимулированная адгезия пленок никеля к кремнию / Р. А. Муминов, А. В. Каримов, Х. Х. Исмаилов, Б. И. Заверюхин // Поверхность. – 2000. – № 10. – С. 92.
3. *Наумов В. В.* Исследование влияния низкоэнергетической ионной стимуляции на плотность и кристаллическую структуру тонких пленок / В. В. Наумов, В. Ф. Бочкарев, О. С. Трушин, А. А. Горячев, Э. Г. Хасанов, А. А. Лебедев, А. С. Куницын // ЖТФ. – 2001. – Т. 71, вып. 8. – С. 92–97.

4. Пат. 6423201 США, МПК{7} С 23 С 28/02. Method of improving the adhesion of copper / Applied Materials, Inc., Mandrekar Tushar (US). – № 09/644968 ; заявл. 23.08.2000 ; опубл. 23.07.2002 ; НПК 205/186.
5. Пат. 9936955 США, МПК{6} H01L21/3205. Plasma annealing of substrates to improve adhesion / Chang Mei, Chen Ling, Liao Marvin, Littau Karl A, Smith David C, Chern Chyi, Tseng Jennifer, Applied Materials Inc. (US). – № 08/498990 ; заявл. 28.02.97 ; опубл. 22.02.99.
6. *Хауффе К.* Реакции в твердых телах и на их поверхности / К.Хауффе. – М. : Иностранная литература, 1962. – 320 с.
7. *Matyushin V. M.* Influence of defect generation on low-temperature diffusion on Au in Ge under the influence of atomic hydrogen / V. M. Matyushin, R. V. Martynyuk // Functional materials. – 2001. – Vol. 8, № 2. – P. 401–403
8. *Матюшин В. М.* Вплив атомарного водню на процеси масопереносу з тонких металевих плівок в системах з обмеженою розчинністю : автореф. дис. на здобуття наук. ступеня д-ра фіз.-мат. наук : спец. 01.04.13 «Фізика металів» / В. М. Матюшин. – Харків, 2004.
9. *Матюшин В. М.* Низкотемпературная диффузия индия в германии, стимулированная атомарным водородом / В. М. Матюшин // ФТП. – 2001. – Т. 35, В. 3. – С. 301–304.
10. *Волькенштейн Ф. Ф.* Радикалорекombинационная люминесценция полупроводников / Ф. Ф. Волькенштейн, А. Н. Горбань, В. А. Соколов. – М. : Наука, 1976. – 278 стр.
11. *Жавжаров Е. Л.* Модифікація параметрів тонких плівок міді під дією атомарного водню / Е. Л. Жавжаров, Н. А. Антонченко, В. М. Матюшин // Вісник Львів. ун-ту. Серія фізична. – 2008. – Вип. 42. – С. 39–46.
12. *Глинка Н. Л.* Общая химия / Н. Л. Глинка // М. : Наука, 1982. – 720 с.
13. Водород в металлах / под ред. Г. Алефельда, И. Фелькея. В 2 т. Т. 1. – М. : Мир, 1978. – 475 с.

Надійшла 22.01.2010
Після доробки 19.02.2010

Жавжаров Е. Л., Матюшин В. М.

ВОЗДЕЙСТВИЕ АТОМАРНОГО ВОДОРОДА НА СТРУКТУРЫ МЕТАЛЛ – ПОЛУПРОВОДНИК. АДГЕЗИЯ ТОНКИХ МЕТАЛЛИЧЕСКИХ ПЛЕНОК

Исследовано влияние атомарного водорода на изменение свойств тонкопленочных структур Cu-Si. Структуры на основе тонких (20–500 нм) пленок меди подвергались действию атомарного водорода с концентрацией $10^{14}–10^{19} \text{ м}^{-3}$. Показано, что при таком воздействии происходит низкотемпературное восстановление окиси меди, повышается адгезия и когезия пленок меди к подложкам Si.

Ключевые слова: тонкие пленки, полупроводник, кремний, атомарный водород, рекомбинация, адгезия, диффузия.

Zhavzharov E. L., Matyushin V. M.

THE INFLUENCE OF ATOMIC HYDROGEN ON METAL – SEMICONDUCTOR STRUCTURES. ADHESION OF THIN METAL FILMS

The influence of atomic hydrogen on the change of thin-film Cu-Si structures properties has been investigated. The structures based on thin (20–500 nm) films of copper were exposed to the action of atomic hydrogen of $10^{14}–10^{19} \text{ м}^{-3}$ concentration. It is shown that under such action low-temperature recovery of copper oxide takes place, adhesion and cohesion of copper films to silicon substrate increases.

Key words: thin films, semiconductor, silicon, atomic hydrogen, recombination, adhesion, diffusion.

ВПЛИВ АТОМАРНОГО ВОДНЮ НА СТРУКТУРИ МЕТАЛ – НАПІВПРОВІДНИК. ВОЛЬТ-АМПЕРНІ ХАРАКТЕРИСТИКИ

Представлено результати експериментального дослідження дії атомарного водню на гетероструктури на основі Si та Ge. Показано, що обробка супроводжується змінами ВАХ. Домінуючими процесами є хемосорбція водню на поверхні, модифікація гетеромежі та дифузія атомів плівки у приповерхневий шарі напівпровідника.

Ключові слова: тонкі плівки, напівпровідник, кремній, атомарний водень, рекомбінація, адгезія, дифузія.

ВСТУП

Створення необхідного вигляду вольтамперної характеристики (ВАХ) в структурах метал – напівпровідник є актуальною і практичною задачею. Але для суттєвої зміни вигляду ВАХ в таких структурах, як правило, необхідні або високі температури (більше 700 К), або дія високоенергетичних частинок, зокрема потоку електронів, іонів, лазерної обробки тощо. Тому створення необхідної ВАХ при температурах близьких до кімнатної є актуальною задачею, оскільки при цьому не змінюються основні електрофізичні параметри об'ємних частин структур. Для такої низькотемпературної обробки була використана дія атомарного водню при температурах кристалу до 320 К. Крім того, вольт-амперні характеристики структур метал – напівпровідник можуть дати цінну інформацію про гетеромежу, а також про домішковий і зарядовий стан приповерхневих шарів кристалів [1–3].

В роботі досліджено зміни ВАХ структур Ni-Si, Ni-Ge при низькотемпературній обробці їх атомарним воднем.

МЕТОДИКА ПРОВЕДЕННЯ ЕКСПЕРИМЕНТІВ

Для дослідження було обрано дві системи на основі Si та Ge. Атомарний водень утворювався шляхом дисоціації молекулярного водню в плазмі височастотного розряду. Молекулярний водень виділявся з водного розчину КОН (20 %) електролізом. Далі водень пропускався через форбалон для уловлювання крапель КОН, осушений стовпчик, заповнений силікагелем, і надходив до робочої камери. Концентрація атомів водню вимірювалася калориметрич-

ним способом і складала близько 10^{19} м^{-3} при тиску в робочій камері 20 Па.

Контакти з лицьової сторони кристалів отримувалися термічним випаровуванням у вакуумі через маску. Контакт до тильної сторони кристалу утворювався за допомогою пасту Ga, яка механічним втиранням наносилася на всю поверхню. Велика площа цього контакту забезпечила омичний характер контакту, що було перевірено експериментально.

ВАХ знімалися за допомогою автоматизованого вимірювача характеристик напівпровідникових приладів Л2-56. Вироблені зразки розміщувалися на спеціально виготовленому столику із системою рухомих притискних контактів. Система притискання складалася з двох пружних елементів з різною пружністю, що забезпечило плавне та надійне притискання контакту до металевий плівки і виключило можливість пошкодження плівки.

Величина роботи виходу електрона (РВЕ) визначалася на основі вимірювань контактної різниці потенціалів (КРП) ΔWF методом динамічного конденсатора [4]. Еталоном був зонд із золота ($WF_{\text{Au}} = 4,3 \text{ eV}$) [5]. Для виключення похибки, пов'язаної з атмосферними коливаннями, потенціал еталоного електрода перед вимірами перевірявся за еталоном зразком – Au 975 проби. Розбіжність контактної різниці потенціалів уздовж поверхні для всіх дослідних зразків не перевищувала $\pm 0,01 \text{ eV}$.

РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Дослідження потенціалу поверхні зразків показали, що обробка *n*-Si(100) та *p*-Si(111) атомарним воднем приводить до зміни поверхневого потенціалу

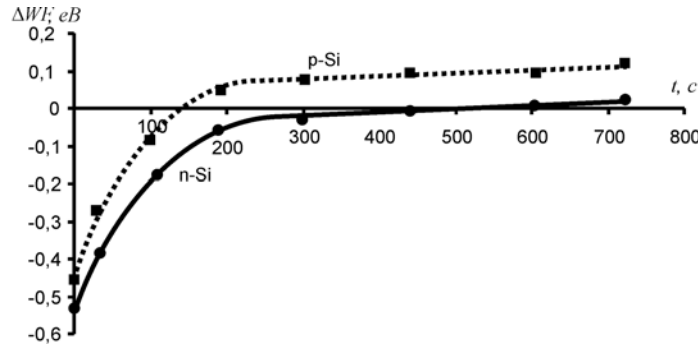


Рис. 1. Зміна потенціалу поверхні від часу обробки атомарним воднем свіжопротравлених зразків

(рис. 1). Виміри демонструють, що потенціали поверхонь реального *n*- та *p*-Si майже не відрізняються один від одного – різниця потенціалів складала $\sim 0,050 \times 0,070$ eB. Це підтверджує наявність явища «пінінгу» рівня Фермі на поверхні напівпровідника [6].

Потенціал починав змінюватися вже після 30 секунд взаємодії кристала з атомарним воднем, а після 3–5 хвилин взаємодії потенціал поверхні досягав насичення. Експериментально зафіксовано, що більшого впливу зазнає *p*-Si (111) порівняно з *n*-Si(100). Після обробки різниця потенціалів між *p*-Si (111) та *n*-Si (100) складала 0,1 eB. Така відмінність у взаємодії атомарного водню з кристалами може бути пов'язана як з різною кристалографічною орієнтацією, так і з процесом комплексоутворення водню з атомами домішки бору та фосфору [7].

В результаті досліджень впливу атомарного водню на опір підкладинок встановлено, що для всіх зразків Si (0,3 та 1 Ом·см) з початком дії атомарного водню спостерігався стрибок величини опору, що, згідно з [8], пов'язано із процесами генерації вільних носіїв заряду під дією рекомбінації атомарного в молекулярний водень та процесом адсорбції. Опір зразків змінювався під час обробки до 20 % від початкового стану, при цьому відносна зміна опору є більшою на зразках з більшим питомим опором. Після припинення дії атомарного водню опір структур з часом повертався до початкових значень. Виходячи з цього факту, за умов обробки комплексоутворення, яке вплинуло на опір зразків, не відбувається, а зміна опору Si під час обробки H обумовлена процесами адсорбції та генерації нерівноважних носіїв заряду.

Дослідження впливу атомарного водню на поверхню *n*-Si (100) та *n*-Si(111) з однаковим питомим опором показали, що зміна потенціалу для *n*-Si(111) на $\sim 0,07$ eB більша, ніж для *n*-Si(100). Згідно з [9],

щільність розірваних зв'язків на поверхні Si (111) більша, ніж на Si (100), тому більші за значенням зміни потенціалу Si (111) пов'язуються з її більшою активністю до взаємодії з воднем. Таким чином, взаємодія H з поверхнею напівпровідника за даних умов обробки не залежить від типу провідності, а визначається станом поверхні – кількістю розірваних зв'язків, адсорбованими частинками і т. ін.

Протягом деякого часу після обробки на усіх дослідних зразках спостерігається зменшення потенціалу поверхні, тобто відбувається його релаксація в бік початкового значення (рис. 2).

Після припинення дії водню потенціал поверхні повертається майже до початкового стану за 36–48 годин, якщо на зразках був зістарений шар природного окислу (витримка після травлення за кімнатних умов протягом декількох місяців). На зразках із тонким окислом, що утворився безпосередньо після обробки кристалів в HF при контакті з повітрям, потенціал поверхні значно відрізняється від початкового навіть після 240 годин витримки за кімнатних умов (різниця потенціалів обробленого та контрольного зразків складає $\sim 0,18$ eB).

Таким чином, зміна потенціалу на зразках із зістареним шаром окислу проходить переважно за рахунок фізичної адсорбції. Взаємодія H із зразками, на поверхні яких шар SiO₂ має мінімальну товщину, супроводжується ще і хімічною адсорбцією. Потенціал поверхні таких зразків за кімнатних умов змінюється незначно і залишається на рівні, що суттєво відрізняється від початкового стану до обробки.

Визначення впливу атомарного водню на електричні параметри системи «метал – напівпровідник» проводилися на двох системах: метал – немодифікований Si та метал – оброблений атомарним воднем Si. Вимірювання показали відмінності ВАХ струк-

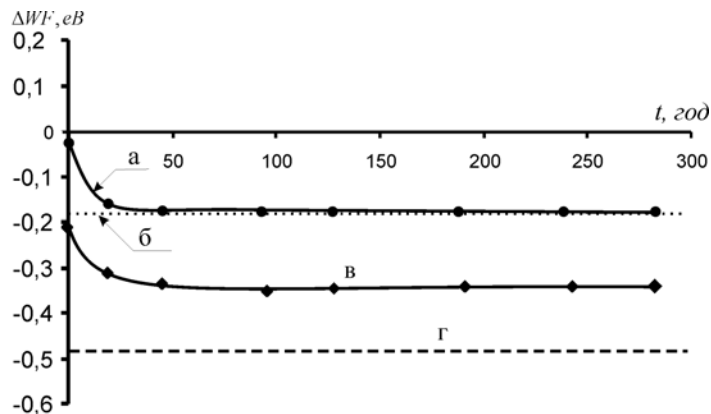


Рис. 2. Релаксація потенціалу поверхні зразків після обробки на протязі 300 с в Н:

a – зразок Si з товстим шаром природного окислу; *б* – початковий стан зразка Si з товстим шаром природного окислу; *в* – зразок Si з тонким шаром окислу; *г* – початковий стан зразка, свіжопротравленого Si

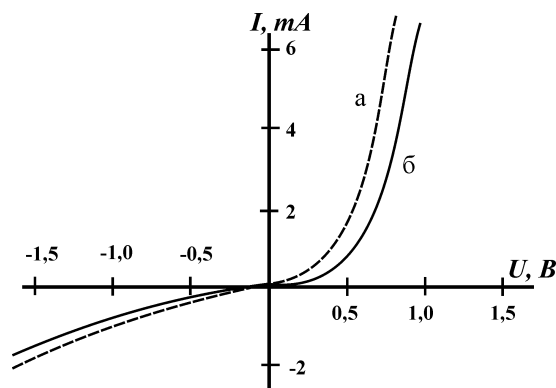


Рис. 3. Вольт-амперні характеристики структур Ni-n-Si (100) 0,3 Ом·см:

a – структури, виготовлені з контрольного Si; *б* – структури, виготовлені з обробленого в Н Si

тур, що були виготовлені з обробленого (3 хв) і необробленого атомарним воднем Si (рис. 3).

Обробка поверхні напівпровідника атомарним воднем приводить до збільшення «п’ятки» прямої гілки ВАХ. Відмінність характеристик зберігається навіть після 6 місяців витримки за кімнатних умов, що говорить про наявність незворотних змін параметрів поверхні Si.

З урахуванням явищ дифузії та дрейфу носіїв заряду через збіднений шар напівпровідника біля контакту залежність щільності струму від напруги на контакті [9]:

$$J_x = J_s [\exp(qV_a/k_bT) - 1], \quad (1)$$

де $J = (q^2 D_n N_c / k_b T) [2q(\Phi_i - V_a) N_d / \epsilon_s]^{1/2} \exp(-q\Phi_B / k_b T)$, D_n – коефіцієнт дифузії електронів; N_c – ефективна

густина станів в зоні провідності; ϵ_s – діелектрична стала напівпровідника; V_a – прикладена напруга; N_d – концентрація донорної домішки; Φ_i та Φ_B – вигин зон та поверхневий бар’єр.

В реальній структурі прикладена напруга падає не лише на переході метал – напівпровідник, але і на опорі бази та контактів. З урахуванням опорі бази R_b та площі переходу S рівняння (1) набуває вигляду

$$I = S J_s [\exp(q(V_a - IR_b) / k_b T) - 1].$$

Для врахування відмінностей реальної структури від ідеальної необхідно врахувати не тільки відмінність величини бар’єру $q\Phi_B$, яка виникає при наявності на поверхні напівпровідника поверхневих станів, але й опір базової області та опір контактів. При достатньо великих струмах через перехід залежність

струму від напруги змінюється з експоненціальної в лінійну. Це дозволяє за частиною прямої гілки ВАХ визначити еквівалентний опір структури, який включає в себе опір бази та активний опір контактів.

Для розрахунку еквівалентного опору R_b , величини бар'єру $q\Phi_B$ за експериментальними ВАХ була складена програма (розрахунок проводився ітераційним методом за критерієм найменших квадратів). Виявилось, що різниця між величинами $q\Phi_B$ зразків, оброблених в Н та ні, складає 0,130 та 0,176 eV відповідно для n -Si(100) та n -Si(111).

Отже, розрахунки параметрів контакту метал – напівпровідник корелюють з результатами, одержаними методом динамічного конденсатора і підтвер-

джують можливість довготривалої модифікації потенціалу поверхні напівпровідника під дією Н та залежність від кристалографічної орієнтації кристалів.

Дослідження ВАХ структур Ni-Si, оброблених та необроблених в Н, показали зміну вигляду ВАХ (рис. 4).

На початковому етапі обробки структура пряма гілка ВАХ майже не змінюється, а зміни зазнає зворотна гілка. Це пов'язано із наявністю процесу розпилення і очистки бокової поверхні контакту. Подальша обробка посилює тенденцію зміни ВАХ. Таке явище пов'язується з процесом дифузії атомарного водню до межі поділу метал – напівпровідник і зміною потенціалу поверхні.

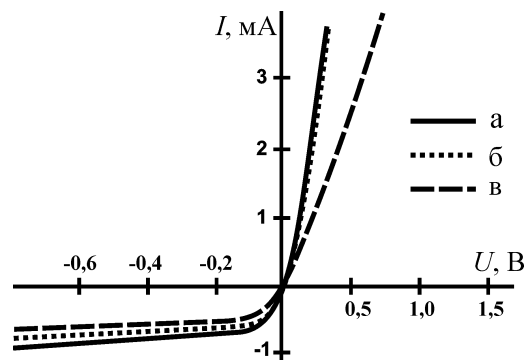


Рис. 4. Зміна ВАХ структури при обробці в Н при $C_H=10^{19} \text{ м}^{-3}$:

a – початковий стан; *b* – 5 хв обробки; *v* – 15 хв обробки

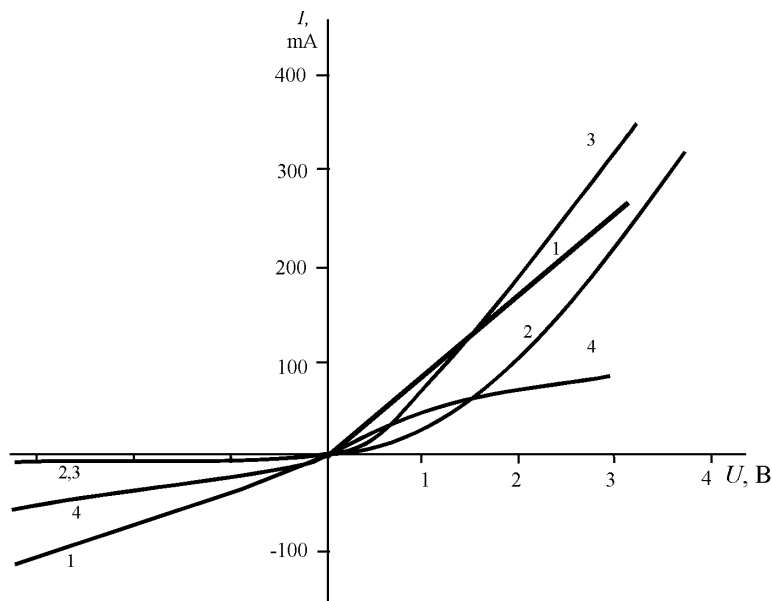


Рис. 5. ВАХ кристалів германію після обробки тонкоплівкових структур (Ni-Ge) атомарним воднем:

a – 0 хв; *b* – 15 хв., *v* – 30 хв; *z* – 60 хв

Оскільки кремній має щільну плівку природного окислу, який заважає дифузії атомів з плівки в кристал, то при взаємодії Н із системами на основі Si процес введення атомів з плівки в кристал є несуттєвим. Однак при розгляді тонкоплівкових систем на основі Ge такий процес може суттєво впливати на електрофізичні властивості структур.

Для дослідження процесу низькотемпературного легування кристала атомами плівки використовувалася система Ni-Ge [10]. Це пов'язано з тим, що Ge має шар окислу, який легко знімається атомарним воднем, тобто є так званим «відкритим» для дії атомарного водню.

Як відомо, нікель в германії є акцептором. Тому легування приповерхневих шарів *n*-германію ($p_s = 1 \text{ Ом}\cdot\text{см}$, $n = 10^{15} \text{ см}^{-3}$) нікелем з концентрацією $p > n$ може приводити до утворення *p-n*-переходу. Для виявлення таких *p-n*-переходів і їх дослідження в кристалах германію з плівками нікелю, оброблених в атомарному водні, вимірювали ВАХ (рис. 5).

Вимір еталонних зразків германію (необроблені в Н кристали, на які плівки Ni напилувалися, а потім зтравлювалися) показав, що ВАХ в цьому випадку майже лінійна (рис. 5, а).

Перед виміром оброблених в Н зразків залишки плівки нікелю зтравлювалися. Виявлено, що при обробці структури в Н протягом 15 хв. ВАХ має явно нелінійний вигляд і відповідає діодним структурам з *p-n*-переходом (рис. 5, б–г). Найбільш різка діодна залежність спостерігалася при часі обробки 30 хв. (рис. 5, в). Подальша обробка в Н приводила до «помягшення» ВАХ (крива 4, рис. 5), що обумовлено розпиленням шарів германію, легування нікелем, зниженням загальної концентрації нікелю, у тому числі і концентрації електрично активних атомів нікелю в приповерхневих шарах.

Сканування уздовж шліфа, виконаного шліф-кулею (діаметр 150 мм), показало, що глибина залягання *p-n*-переходу структури (рис. 5, в) не перевищує 0,5 мкм.

Таким чином, виявлено, що під впливом атомарного водню в системі Ni-Ge утворюється *p-n*-перехід, який має невелику глибину залягання внаслідок малого коефіцієнту дифузії нікелю в германії.

ВИСНОВКИ

Виявлено зміни параметрів кремнієвих підкладок, які потрібно враховувати при розробці технологій, пов'язаних з воднем. Навіть нетривалий контакт (3 хв) Si з Н приводить до зміни потенціалу по-

верхні (до 0,3 еВ), яка відповідає збільшенню роботи виходу. Знайдено, що взаємодія атомарного водню з поверхнею визначається товщиною, якістю окислу і орієнтацією поверхні [поверхня Si(111) зазнає більшого впливу, ніж (100)]. Зміни поверхневого потенціалу на свіжопротравлених поверхнях носять тривалий характер.

Знайдено, що під дією атомарного водню змінюється «п'ятка» прямої гілки ВАХ (у середньому на 0,2 В).

Виявлено, що процесом, який впливає на вигляд ВАХ при обробці структур метал – германій на відміну від структур на основі кремнію, є процес введення атомів металу з плівки в приповерхневі шари кристала. Така відмінність у взаємодії пов'язана з відсутністю щільного шару природного окислу на поверхні германію.

На основі виявлених процесів показана можливість керування видом ВАХ плівкових структур на основі як Si, так і Ge. Це може бути корисним при створенні низькотемпературної технології водневої обробки плівкових структур.

СПИСОК ЛІТЕРАТУРИ

1. Батавин В. В. Измерение параметров полупроводниковых материалов и структур / В. В. Батавин, Ю. А. Концевой, Ю. В. Федорович // М. : Радио и связь, 1985. – 264 с.
2. Воробьев Ю. В. Методы исследования полупроводников / Ю. В. Воробьев, В. Н. Добровольский, В. Н. Стриха // Киев : Вища школа, 1988. – 232 с.
3. Воронков И. Е. Исследование неоднородности лавинного пробоя *p-n* переходов с помощью производных их вольтамперных характеристик / И. Е. Воронков, А. Д. Ходкевич, С. П. Ходкевич // Электронная техника. – Сер. 1. – 1977. – № 8. – С. 33–44.
4. Вудраф Д. Современные методы исследования поверхности / Д. Вудраф, Т. Делчар // М. : Мир, 1989. – 564 с.
5. Таблицы физических величин / под ред. И. К. Кицина. – М. : Атомиздат, 1976. – 1006 с.
6. Проблемы физики поверхности полупроводников / под общ. ред. О. В. Снитко // К. : Наук. думка, 1981. – 332 с.
7. Sachse J.-U. Palladium-Hydrogen Related Complexes in Silicon / J.-U. Sachse, J. Weber, H. Lemke // Mater. Sci. Forum. – 1997. – Vol. 258–263. – P. 307–311.
8. Горбань А. Н. Вентильный эффект в полупроводнике с *p-n*-переходом при рекомбинации атомов водорода на его поверхности / А. Н. Горбань, Н. М. Савченко, Ю. А. Швец // Физика и техника полупроводников. – 1976. – Т. 10. – В. 12. – С. 2381–2384.
9. Малер Р. Элементы интегральных схем / Р. Малер, Т. Кейминс. – М. : Мир, 1989. – 630 с.
10. Матюшин В. М. Ускорение диффузионных процессов в твердых телах-катализаторах при протекании на их поверхности экзотермической реакции / В. М. Матюшин // Украинский химический журнал. – 1999. – Т. 65. – В. 4. – С. 361–365.

Надійшла 22.01.2010
Після доробки 19.02.2010

Матюшин В. М., Жавжаров Е. Л.

ВОЗДЕЙСТВИЕ АТОМАРНОГО ВОДОРОДА НА СТРУКТУРЫ МЕТАЛЛ – ПОЛУПРОВОДНИК. ВОЛЬТАМПЕРНЫЕ ХАРАКТЕРИСТИКИ

Представлены результаты экспериментального исследования действия атомарного водорода на гетероструктуры на основе Si и Ge. Показано, что обработка сопровождается изменениями вида ВАХ. Доминирующими процессами являются хемосорбция водорода на поверхности, модификация гетерограницы и диффузия атомов пленки в приповерхностные слои полупроводника.

Ключевые слова: тонкие пленки, полупроводник, кремний, атомарный водород, рекомбинация, адгезия, диффузия.

Matyushin V. M., Zhavzharov E. L.

THE INFLUENCE OF ATOMIC HYDROGEN ON METAL-SEMICONDUCTOR STRUCTURES. VOLTAGE-CURRENT CHARACTERISTICS

The authors present the results of experimental investigation of atomic hydrogen effect on heterostructures based on Si and Ge. It is shown that treatment is accompanied by changes of V-C characteristics. Dominant processes are hydrogen chemisorption on the surface, modification of heteroboundary and diffusion of film atoms in the semiconductor near-surface layers.

Key words: thin films, semiconductor, silicon, atomic hydrogen, recombination, adhesion, diffusion.

УДК 621.396.6.017:004.942

Шило Г. Н.¹, Огренич Е. В.², Филобок А. В.³, Гапоненко Н. П.¹

¹Канд. техн. наук, доцент Запорожского национального технического университета

²Аспирант Запорожского национального технического университета

³Студент Запорожского национального технического университета

ПРОЕКТИРОВАНИЕ РЕБРИСТЫХ ТЕПЛОТВОДОВ МИНИМАЛЬНОЙ МАССЫ

Сформированы модели элементов конструкций теплоотводов в виде тепловых линий с нагрузкой на конце. Рассмотрены режимы работы тепловых линий. Для теплоотводов минимальной массы получены соотношения, связывающие размеры элементов конструкций с параметрами границ области работоспособности. Анализируется зависимость массы теплоотводов от количества ребер.

Ключевые слова: теплоотвод, тепловая линия, минимальная масса, алгоритм оптимизации, область работоспособности, технологическое ограничение.

ВВЕДЕНИЕ

Тепловой режим теплонагруженных элементов в радиоэлектронной аппаратуре в большинстве случаев обеспечивается применением теплоотводящих устройств или вынужденным охлаждением. Оба эти пути требуют затрат материалов и влияют на массогабаритные показатели радиоаппаратуры. Одним из наиболее распространенных теплоотводящих устройств является радиатор. Существует большое количество конструкций радиаторов [1, 2], что не позволяет в общем виде решить задачу оптимизации их массогабаритных показателей. Исследованы только простые элементы конструкций [3], для которых показана возможность уменьшения массы в несколько раз.

Оптимизация теплоотводящих элементов при использовании комплексных массогабаритных показателей рассмотрена в [4], где показано, что процедура оптимизации в большинстве случаев приводила к увеличению площади, которую занимают теплоотво-

дящие элементы на печатных платах. Уменьшение этой площади осуществляется переходом к пространственным формам теплоотводящих элементов с развитыми поверхностями. Для увеличения площади теплоотвода чаще всего используются ребристые и штыревые элементы. Поскольку параметры этих элементов влияют на процессы отвода тепла конвекцией и излучением, целесообразно провести исследование влияния параметров элементов конструкций теплоотводов на их массогабаритные показатели.

Целью работы является разработка алгоритма проектирования ребристых теплоотводов минимальной массы. Для решения поставленной задачи необходимо:

- сформировать модели процессов теплопередачи в элементах конструкций теплоотводов с ребристыми поверхностями;
- разработать алгоритм оптимизации массы ребристых теплоотводов;

© Шило Г. Н., Огренич Е. В., Филобок А. В., Гапоненко Н. П., 2010

– рассмотреть особенности оптимизации массы теплоотводов при учете технологических ограничений.

1. ФОРМИРОВАНИЕ МОДЕЛЕЙ ЭЛЕМЕНТОВ КОНСТРУКЦИИ

Тепловая модель ребристого теплоотвода показана на рис. 1, где P – тепловой поток, d_r, l_r, d_p, l_p – толщина и длина ребер и участков основания. Теплоотвод имеет одинаковую ширину. Размеры остальных конструктивных элементов отличаются друг от друга.

В общем случае тепловые процессы в таких конструкциях описываются двумерными уравнениями теплопроводности. Но практика проектирования теплоотводов показывает, что обычно выполняется соотношение

$$l_i \gg d_i,$$

где l_i и d_i – длина и толщина i -го ребра и участка основания.

Это позволяет рассматривать процессы теплопередачи в элементах конструкций как одномерные и осуществлять эквивалентные преобразования тепловой модели с использованием выражения для теплового сопротивления теплонагруженного стержня. Процессы теплопередачи в такой тепловой линии будут описываться системой уравнений

$$\left. \begin{aligned} \frac{d^2\theta}{dx^2} - b^2\theta &= 0 \\ \frac{d\theta}{dx} &= -\frac{P}{\lambda S} \quad (x=0) \\ \frac{d\theta}{dx} &= -\beta\theta \quad (x=l) \end{aligned} \right\} \quad (1)$$

где θ – перепад температур относительно температуры окружающей среды; x – расстояние от начала стержня; b и β – коэффициенты тепловой нагрузки на поверхность тепловой линии и в ее конце; λ – коэффициент теплопроводности материала теплоотвода; S –

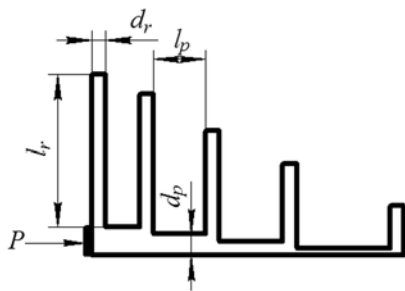


Рис. 1. Теплоотвод с ребристыми поверхностями

площадь поперечного сечения теплопередающего элемента; l – длина ребра или участка основания.

Выражение для коэффициента тепловой нагрузки на поверхность стержня имеет вид [1]

$$b^2 = \frac{\alpha U}{\lambda S},$$

где α – коэффициент теплоотдачи с поверхности теплоотвода; U – периметр поперечного сечения теплопередающего элемента.

Коэффициент тепловой нагрузки в конце линии определяется с помощью соотношения

$$\beta = \frac{1}{\lambda S R_H},$$

где R_H – нагрузочное тепловое сопротивление в конце линии.

Нагрузочное тепловое сопротивление рассчитывается как параллельное соединение входных тепловых сопротивлений ребра и следующего участка основания:

$$R_H = \frac{R_r R_p}{R_r + R_p},$$

где R_r и R_p – входное тепловое сопротивление ребра и следующего участка основания.

При расчете коэффициентов тепловой нагрузки в конце участков основания используется выражение для входного теплового сопротивления следующего участка основания, которое находится из решения системы (1):

$$\theta(x) = \frac{P}{\lambda b S} \left[\frac{b \cdot \text{ch}(bl) + \beta \cdot \text{sh}(bl)}{b \cdot \text{sh}(bl) + \beta \cdot \text{ch}(bl)} \text{ch}(bx) - \text{sh}(bx) \right]. \quad (2)$$

Из выражения (2) следует, что входное тепловое сопротивление ребер или участков основания может определяться с помощью соотношения

$$R = R_0 \frac{b \cdot \text{ch}(bl) + \beta \cdot \text{sh}(bl)}{b \cdot \text{sh}(bl) + \beta \cdot \text{ch}(bl)}, \quad (3)$$

где $R_0 = 1/\lambda b S$ – характеристическое тепловое сопротивление тепловой линии.

Графики зависимости нормированного теплового сопротивления от нормированной длины тепловой линии bl показаны на рис. 2.

В зависимости от параметров тепловой нагрузки в теплонагруженном стержне могут возникать режимы холостого хода, критический и короткого замыкания. В режиме холостого хода $\beta = 0$, а в конце стержня теплоотвод отсутствует. Выражение (3) в этом случае преобразуется к виду

$$R = R_0 \text{cth}(bl).$$

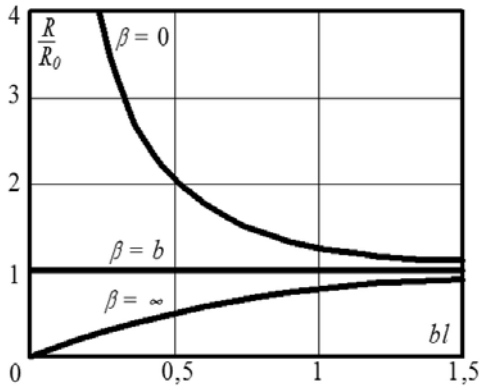


Рис. 2. Влияние тепловой нагрузки на тепловое сопротивление

Это выражение используется при расчете входных тепловых сопротивлений ребер. В режиме короткого замыкания ($\beta \rightarrow \infty$) на конце стержня поддерживается нулевая температура. Этому режиму соответствует входное тепловое сопротивление участка основания:

$$R = R_0 \text{th}(bl).$$

В критическом режиме $\beta = b$ входное тепловое сопротивление стержня не зависит от длины и равно характеристическому сопротивлению R_0 .

При использовании ребристых теплоотводов тепловой режим теплонагруженного стержня характеризуется условием

$$\beta < b, \tag{4}$$

которое необходимо выдерживать при выборе начальных приближений в алгоритме оптимизации.

2. ОПТИМИЗАЦИЯ МАССЫ ТЕПЛОТВОДОВ

Масса теплоотвода, показанного на рис. 1, определяется с помощью выражения

$$m = \rho L \sum_{i=1}^n (d_{pi} l_{pi} + d_{ri} l_{ri}), \tag{5}$$

где ρ , L и n – плотность материала, ширина и число ребер теплоотвода.

Это соотношение может использоваться как целевая функция при оптимизации массы теплоотвода. Оптимизационная задача решается при ограничении

$$R_T \leq R_{Tb}, \tag{6}$$

где R_b – входное тепловое сопротивление теплоотвода; R_{Tb} – гранично-допустимое входное тепловое сопротивление теплоотвода.

Для ограничения области работоспособности параметры тепловой линии отображаются в область обратных величин:

$$\begin{aligned} d_{pi} &= \frac{1}{x_{1i}}; & l_{pi} &= \frac{1}{x_{2i}}; \\ d_{ri} &= \frac{1}{x_{3i}}; & l_{ri} &= \frac{1}{x_{4i}}. \end{aligned} \tag{7}$$

Преобразование (7) формирует область работоспособности, ограниченную координатными плоскостями и поверхностью (6). Для упрощения процедуры оптимизации выражение (3) заменяется линейной или квадратичной моделями.

Формирование моделей области работоспособности показано на рис. 3, где Ω_w – область работоспособности, B – граничная точка области работоспособности; x_{b1} и x_{b2} – параметры тепловой линии в граничной точке; 1 – граница области работоспособности тепловой линии, 2 и 3 – границы области работоспособности при линейной и квадратичной аппроксимации входного теплового сопротивления.

Исследование показало, что лучшую сходимость итерационных процессов обеспечивает квадратичная модель выходной функции

$$R(X) = \sum_{i=1}^n \sum_{j=1}^4 c_{ij} x_{ij}^2, \tag{8}$$

где $X = \{x_{ij}\}$ – множество геометрических параметров тепловой линии; c_{ij} – коэффициенты модели.

Коэффициенты модели определяются из условия совпадения касательной к функции (3) и эллипсоида (8) [5]. Уравнение касательной к ограничивающей функции формируется из ее линейной модели:

$$R(X) = a_0 + \sum_{i=1}^n \sum_{j=1}^4 a_{ij} x_{ij}, \tag{9}$$

где $a_{ij} = \left. \frac{\partial R(X)}{\partial x_{ij}} \right|_{x_b}$ – линейные коэффициенты ограничивающей функции $R(X)$; $X_b = \{x_{bij}\}$ – множе-

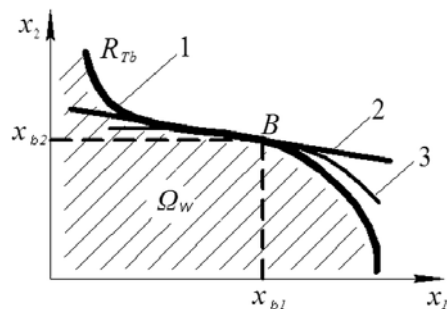


Рис. 3. Формирование области работоспособности

ство параметров элементов, соответствующих граничной точке области работоспособности; $a_0 = R_b - \sum_{i=1}^n \sum_{j=1}^4 a_{ij} x_{bij}$.

Из выражения (9) формируется уравнение касательной к функции (3) в виде

$$\frac{1}{R_b - a_0} \sum_{i=1}^n \sum_{j=1}^4 a_{ij} x_{ij} = 1. \quad (10)$$

Уравнение касательной к эллипсоиду (8) имеет вид

$$\frac{1}{R_b} \sum_{i=1}^n \sum_{j=1}^4 x_{bij} a_{ij} x_{ij} = 1. \quad (11)$$

Сравнение соотношений (10) и (11) позволяет получить выражение для коэффициентов квадратичной модели (8):

$$c_{ij} = \frac{R_b a_{ij}}{R_b - a_0 x_{bij}}. \quad (12)$$

Отображение (7) позволяет записать целевую функцию (5) в виде

$$m(x) = \rho L \sum_{i=1}^n \left(\frac{1}{x_{1i} x_{2i}} + \frac{1}{x_{3i} x_{4i}} \right) \rightarrow \max. \quad (13)$$

Оптимальное значение параметров тепловой линии определяется методом множителей Лагранжа. Для этого из соотношений (8) и (13) составляется вспомогательная функция

$$\Phi(\lambda) = \sum_{i=1}^n \left(\frac{1}{x_{1i} x_{2i}} + \frac{1}{x_{3i} x_{4i}} \right) + \lambda \sum_{i=1}^n \sum_{j=1}^4 c_{ij} x_{ij}^2, \quad (14)$$

где λ – множитель Лагранжа.

Оптимальные значения параметров теплоотводов достигаются при соотношениях

$$x_{1i} = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{\frac{c_{2i}}{c_{1i}}}, \quad x_{2i} = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{\frac{c_{1i}}{c_{2i}}}, \quad (15)$$

$$x_{3i} = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{\frac{c_{4i}}{c_{3i}}}, \quad x_{4i} = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{\frac{c_{3i}}{c_{4i}}}, \quad (16)$$

когда множитель Лагранжа принимает значение

$$\lambda = \frac{2}{R_b^2} \left[\sum_{i=1}^n \left(\sqrt[4]{c_{1i} c_{2i}} + \sqrt[4]{c_{3i} c_{4i}} \right) \right]^2.$$

Поскольку значения коэффициентов в выражениях (15) и (16) зависят от положения граничной точки, то для определения оптимальных размеров теплоотвода с количеством ребер n используется итерационный алгоритм.

Шаг 1. Определяются размеры одиночного ребра минимальной массы [3]:

$$d_m = \frac{0,632}{\alpha \lambda} \left(\frac{1}{LR_b} \right)^2;$$

$$l_m = \frac{0,7875}{\alpha} \frac{1}{LR_b}.$$

При расчете α используются выражения для вертикально ориентированной поверхности [1].

Шаг 2. Назначаются начальные значения параметров теплоотвода с учетом количества ребер:

$$d_{ri} = d_{pi} = d_m,$$

$$l_{ri} = l_{pi} = \frac{l_m}{2n-1}.$$

Шаг 3. Определяется среднеповерхностная температура ребер и участков основания путем интегрирования выражений (2) для каждого участка теплоотвода.

Шаг 4. Уточняются значения коэффициентов теплоотдачи на каждом участке теплоотвода. Используются соотношения, учитывающие влияние межреберного расстояния на коэффициенты теплоотдачи конвекцией и излучением [3].

Шаг 5. Уточняются геометрические параметры теплоотвода. Используются соотношения (15) и (16).

Шаг 6. Проверяется условие

$$\left| \frac{R(X^{(i)}) - R(X^{(i-1)})}{R(X^{(i)})} \right| \leq \varepsilon,$$

где ε – погрешность вычислений.

Если условие не выполняется, то переход на шаг 3.

Шаг 7. Проверяется условие

$$\left| \frac{R(X^{(i)}) - R_b}{R_b} \right| \leq \varepsilon,$$

где ε – погрешность вычислений.

Если условие не выполняется, то переход на шаг 3. Иначе – конец алгоритма.

Алгоритм позволяет определить оптимальные размеры теплоотвода, имеющего минимальную массу при заданном числе ребер. Для окончательного выбора конструкции теплоотвода необходимо провести исследование зависимости массы оптимальных теплоотводов от количества ребер.

3. ИССЛЕДОВАНИЕ РЕБРИСТЫХ ТЕПЛОТВОДОВ

Исследовалась зависимость массы теплоотвода от количества ребер. Исследование проводилось с целью определения оптимального числа ребер при минимизации массы теплоотвода. В исследованиях использовался теплоотвод из алюминиевого сплава Ад2.

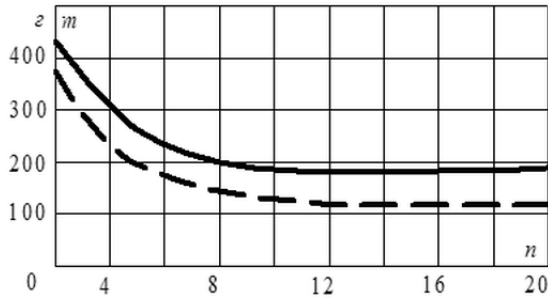


Рис. 4. Зависимость массы теплоотвода от числа ребер

Задавалась мощность $P = 10$ Вт, температура окружающей среды $t_c = 40^\circ \text{C}$, допустимое входное тепловое сопротивление $R_b = 2$ К/Вт, коэффициент черноты поверхности теплоотвода $\epsilon = 0,91$. В алгоритме учитывалась зависимость коэффициентов теплоотдачи от межреберного расстояния [3]. Результаты расчетов отображены на рис. 4 пунктирной линией.

Из рис. 4 видно, что оптимизация позволяет в несколько раз уменьшить массу теплоотвода. Изменение массы теплоотвода в 2,5 раза осуществляется при увеличении количества ребер до $n = 8$. Дальнейшее увеличение количества ребер уменьшает массу теплоотвода только на 20 %. По сравнению с пластинчатым теплоотводом [3] масса уменьшается в 5 раз.

Оптимизация приводит к конструкции теплоотвода, у которой размеры ребер и участков основания не повторяются. Толщина участков основания изменяется от 9 мм до 1 мм, а длина – от 9 мм до 14 мм. Толщина ребер изменяется от 1,2 мм до 0,4 мм, а их длина – от 136 мм до 95 мм. Длина ребер, расположенных в конце теплоотвода, уменьшается до 30 %. Необходимость обеспечения такой номенклатуры размеров снижает технологичность теплоотвода. Улучшения технологичности изделия можно достигнуть обеспечением равенства однотипных размеров. В этом случае целевая функция (5) запишется в виде

$$m = \rho L[(n - 1) \cdot d_p l_p + n \cdot d_r l_r] \rightarrow \min.$$

Вспомогательная функция (14) преобразуется к виду

$$\Phi(\lambda) = \left(\frac{n-1}{x_1 x_2} + \frac{n}{x_3 x_4} \right) + \lambda \sum_{j=1}^4 c_j x_j^2.$$

Оптимальные значения массы теплоотвода обеспечиваются при выполнении соотношений

$$x_1 = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{(n-1)^2 \cdot \frac{c_2}{c_1}},$$

$$x_2 = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{(n-1)^2 \cdot \frac{c_1}{c_2}},$$

$$x_3 = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{n^2 \cdot \frac{c_4}{c_3}},$$

$$x_4 = \frac{1}{\sqrt[4]{2\lambda}} \sqrt[8]{n^2 \cdot \frac{c_3}{c_4}},$$

$$\lambda = \frac{2}{R_b^2} \left[\sqrt[4]{(n-1)^2 c_{1i} c_{2i}} + \sqrt[4]{n^2 c_{3i} c_{4i}} \right]^2.$$

Оптимизация массы теплоотвода с равными однотипными размерами приводит к зависимости, показанной на рис. 4 сплошной линией. Из графика видно, что теплоотвод имеет минимальную массу при числе ребер $n = 12$. Дальнейшее увеличение ребер приводит к незначительному увеличению массы на 25 %. Сравнение теплоотводов с различными и равными однотипными размерами показывает, что за улучшение технологичности приходится расплачиваться увеличением массы на 25 %.

ВЫВОДЫ

Разработанный метод оптимизации массы теплоотводов позволяет проектировать теплоотводы минимальной массы путем изменения параметров элементов конструкций. Полученные оптимизационные соотношения связывают размеры участков теплоотвода с параметрами границ области работоспособности, определяющей допустимый тепловой режим теплоотводящих элементов. Улучшение сходимости алгоритма обеспечивается выбором эллипсоидной модели границ области работоспособности.

Применение разработанных оптимизационных процедур позволяет уменьшать массу теплоотводящих устройств в несколько раз по сравнению с пластинчатыми теплоотводами. Предложенный метод позволяет учитывать технологические ограничения, в частности проектировать теплоотводы с равными однотипными размерами. Но в этом случае уменьшение массы на 20–30 % меньше. Предложенный метод может использоваться и при других стратегиях проектирования теплоотводов.

СПИСОК ЛИТЕРАТУРЫ

1. Дульнев Р. Н. Тепло- и массообмен в радиоэлектронной аппаратуре / Дульнев Р. Н. – М. : Высш. шк., 1984. – 247 с.
2. Роткоп Л. Л. Обеспечение тепловых режимов при конструировании радиоэлектронной аппаратуры / Роткоп Л. Л., Спокойный Ю. Е. – М. : Советское радио, 1976. – 232 с.

3. *Ройзен Л. И.* Теловой расчет ребренных поверхностей / Ройзен Л. И., Дулькин И. Н. ; под. ред. В. Г. Фастовского. – М. : Энергия, 1977. – 256 с. с ил.
4. *Гапоненко Н. П.* Проектирование пластинчатых радиаторов с оптимальными массогабаритными параметрами / Гапоненко Н. П., Огренич Е. В. // Радиоэлектроника, информатика, управление. – 2007. – № 1. – С. 32–35.
5. *Шило Г. Н.* Геометрические методы назначения допусков / Шило Г. Н. // Проблемы управления и информатики. – 2007. – № 2. – С. 118–126.

Надійшла 12.03.2010
Після доробки 22.04.2010

Шило Г. М., Огренич Є. В., Філобок Г. В., Гапоненко М. П.

ПРОЕКТУВАННЯ РЕБРИСТИХ ТЕПЛОВІДВОДІВ МІНІМАЛЬНОЇ МАСИ

Сформовано моделі елементів конструкцій тепловідводів у вигляді теплових ліній з навантаженням на кінці. Розглянуто режими роботи теплових ліній. Для тепловідводів мінімальної маси отримано співвідношення, що

зв'язують розміри елементів конструкцій тепловідводів з параметрами границь області працездатності. Оцінено залежність маси тепловідводів від кількості ребер.

Ключові слова: тепловідвід, тепла лінія, мінімальна маса, алгоритм оптимізації, область працездатності, технологічне обмеження.

Shilo G. N., Ogrenich E. V., Filobok A. V., Gaponenko N. P.
DESIGNING OF RIBBED HEAT SINKS HAVING MINIMUM MASS

The authors present the models of heat sinks structural elements in the form of heat lines loaded at the end. Modes of heat lines operation are considered. Relationships have been obtained for heat sinks of minimum mass, relating the size of heat sinks structural elements with parameters of workability region limits. Dependence of heat sinks mass on the number of ribs has been estimated.

Key words: heat sink, heat line, minimum mass, optimization algorithm, workability region, technological limitations.

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL AND COMPUTER MODELLING

УДК 519.216

Агеев Д. В.¹, Копылев А. Н.²¹Канд. техн. наук, докторант Харьковского национального университета радиозлектроники²Ассистент Запорожского национального технического университета

МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ ПОТОКОВ В МУЛЬТИСЕРВИСНОЙ СЕТИ NGN ПРИ РЕШЕНИИ ЗАДАЧ ПАРАМЕТРИЧЕСКОГО СИНТЕЗА

В данной статье предлагаются математические модели информационных потоков и методы определения их характеристик на различных участках мультисервисной сети NGN с целью их использования при проектировании телекоммуникационных сетей.

Ключевые слова: самоподобный процесс, дисперсия, трафик, аппроксимация, агрегирование, фрактал.

ВВЕДЕНИЕ

С учетом мировых и отечественных тенденций развития телекоммуникационных систем, одной из наиболее актуальной проблем отрасли связи Украины является создание мультисервисной сети связи, которая удовлетворяла бы растущему передаваемому трафику и увеличивающемуся перечню телекоммуникационных услуг. Существующие методы, используемые при расчете параметров структурных элементов проектируемой системы, основаны на предположении об аппроксимации Марковской моделью потоков Пуассона и Эрланга, которые широко используются при проектировании сетей с коммутацией пакетов и коммутацией каналов. Перспективная сеть NGN в свою очередь обеспечивает предоставление услуг, которые предусматривают установку соединения и выделения ресурсов на все время существования соединения (что свойственно сетям с коммутацией каналов), а также без выделения ресурсов в индивидуальное пользование (что свойственно се-

тям с коммутацией пакетов). Математические модели и методы, используемые при проектировании мультисервисных телекоммуникационных систем, должны учитывать особенности обоих классов сетей.

С другой стороны, последние многочисленные исследования в области телетрафика показали, что реальный телетрафик, циркулирующий в телекоммуникационных системах, имеет более сложный характер, чем тот, что описывается классическими моделями. Основной причиной специфических свойств трафика, по мнению многих исследователей, является наличие информационных обратных связей и использование протоколов управления потоком. Исследование статистических свойств реального трафика показало наличие «тяжелых хвостов» распределения и долговременных зависимостей. Для описания подобных свойств телетрафика в работах многих исследователей предлагается использовать модели самоподобных процессов.

© Агеев Д. В., Копылев А. Н., 2010

В данной статье предлагаются математические модели информационных потоков и методы определения их характеристик на различных участках мультисервисной сети NGN.

1. ЗАДАЧИ, ВОЗНИКАЮЩИЕ ПРИ ПРОЕКТИРОВАНИИ СЕТЕЙ NGN

В рекомендациях серии Y (рек. Y.2001 и Y.2011) предложена модель инфокоммуникационной системы, которая включает четыре основных компонента:

- оборудование в помещении абонента, которое может состоять как из одного терминала, так и из нескольких;

- сеть абонентского доступа, которая обеспечивает подключение оборудования, находящегося в помещении пользователя, к транспортной сети;

- транспортную сеть, состоящую из совокупности коммутационных узлов и станций, обеспечивающих коммутацию и прозрачную передачу информации пользователя, а также обеспечивающую выход к средствам поддержки иных инфокоммуникационных услуг;

- уровень управления услугами содержит функции управления логикой услуг и приложений и представляет собой распределенную вычислительную среду, обеспечивающую предоставление инфокоммуникационных услуг, управление услугами, создание и внедрение новых услуг, взаимодействие различных услуг.

Процесс проектирования мультисервисной телекоммуникационной системы, как и любой другой сложной, пространственно разнесенной системы является сложной процедурой. Для упрощения процесса проектирования весь процесс разбивается на отдельные этапы. На каждом из этапов ставится определенная задача из общего перечня решаемых задач проектирования. Использование такого подхода позволяет упростить процесс проектирования за счет того, что задачи разных этапов можно решать максимально независимо друг от друга с использованием результатов одних из них как исходные данные для других.

Одним из этапов проектирования является выбор параметров структурных элементов мультисервисной телекоммуникационной системы. Для решения данной задачи необходимо знать характеристики информационных потоков, циркулирующих в сети. Определение параметров потока в каналах связи обычно производится в процессе проектирования в результате решения задачи распределения потоков. Для решения последней задачи необходимо знать:

- характеристики потоков, поступающих в сеть;
- методику определения характеристик группового потока;

- методику определения параметров потоков в результате агрегирования трафика в узлах сети;

- методику определения характеристик потока в результате обработки в узлах сети.

2. ОПРЕДЕЛЕНИЕ САМОПОДОБНОГО ПОТОКА

Как уже указывалось выше, современные исследования показали, что использование моделей самоподобного процесса позволяет более точно описать поведение сетевого трафика. Дадим определения строго и асимптотически самоподобных в широком смысле случайных процессов дискретного аргумента [1].

Пусть $\xi = (\xi_1, \xi_2, \dots)$ – полубесконечный отрезок стационарного в широком смысле случайного процесса дискретного аргумента t . Обозначим через $\bar{\omega} < \infty$ и $\sigma^2 < \infty$ среднее и дисперсию процесса соответственно, а через $r(k) = \frac{\Delta(\xi_{t+k} - \bar{\omega})(\xi_t - \bar{\omega})}{\sigma^2}$, $b(k) = \frac{\Delta}{\sigma^2} r(k)$, $k \in Z_+ = \{0, 1, 2, \dots\}$ автокорреляционную функцию и автоковариацию процесса ξ .

Обозначим через $\xi^{(m)} = (\xi_1^{(m)}, \xi_2^{(m)}, \dots)$ усредненный по блокам длины m процесс ξ , компоненты которого определяются равенством

$$\xi_t^{(m)} = \frac{1}{m}(\xi_{tm-m+1} + \dots + \xi_{tm}), \quad m, t \in N. \quad (1)$$

Обозначим через $r_m(k)$, $b_m(k)$ и $\sigma_m = b_m(0)$ коэффициент корреляции, автоковариацию и дисперсию процесса $\xi^{(m)}$ соответственно.

Определение [1]. Процесс ξ называется *строго самоподобным в широком смысле* (ССШС) с параметром $H = 1 - (\beta/2)$, $0 < \beta < 1$, если

$$r_m(k) = r(k), \quad k \in Z_+, \quad m \in \{2, 3, \dots\}. \quad (2)$$

То есть процесс ξ является СШС, если процесс $\xi^{(m)}$ неотличим от исходного процесса ξ , как минимум в отношении статистических характеристик второго порядка.

Определение [1]. Процесс ξ называется *асимптотически самоподобным в широком смысле* (АСШС) с параметром $H = 1 - (\beta/2)$, $0 < \beta < 1$, если

$$\lim_{m \rightarrow \infty} r_m(k) = g(k), \quad k \in N, \quad (3)$$

где $g(k) = \frac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}]$ – коэффициент корреляции СШС.

Смысл этого определения состоит в том, что ξ_t^m является АСШС процессом, если после усреднения по блокам длины m и при $m \rightarrow \infty$ он сходится к ССШС процессу.

Самоподобные процессы обладают следующими основными свойствами [2]:

- медленно затухающая дисперсия;
- наличие долговременных зависимостей;
- наличие распределения с тяжелыми хвостами.

Свойство медленно затухающей дисперсии заключается в том, что дисперсия выборочного среднего затухает медленнее, чем величина, обратная размеру выборки, т. е.

$$\sigma^2(\xi_t^{(m)}) \propto m^{(2H-2)}, \quad m \rightarrow \infty, \quad (4)$$

в то время как для традиционных случайных процессов $\sigma^2(\xi_t^{(m)}) = \sigma^2 m^{-1}$, т. е. уменьшается обратно пропорционально объему выборки.

Наличие долговременной зависимости заключается в том, что самоподобный процесс обладает гиперболически затухающей корреляционной функцией

$$r(k) \cong k^{(2H-2)} L(t), \quad k \rightarrow \infty, \quad (5)$$

где $L(t)$ – медленно меняющаяся функция на бесконечности (т. е. $\lim_{t \rightarrow \infty} \frac{L(tx)}{L(t)} = 1$ для всех $x > 0$).

Свойство наличия распределения с тяжелым хвостом заключается в том, что случайная величина ξ имеет *распределение с тяжелым хвостом*, если

$$P[\xi > x] \sim c \cdot x^{-a}, \quad x \rightarrow \infty, \quad (6)$$

где $0 < a < 2$, и называется параметром формы, c – некоторая положительная константа. В отличие от распределений с легкими хвостами, такими как экспоненциальное или гауссовское, которые имеют экспоненциальное убывание хвоста.

Коэффициент Хэрста имеет принципиальное значение в теории самоподобных процессов. Он является индикатором степени самоподобия процесса. Для процессов, обладающих эффектом самоподобия, параметр Хэрста $0,5 < H < 1$.

3. МОДЕЛИРОВАНИЕ ИСТОЧНИКОВ ИНФОРМАЦИОННЫХ ПОТОКОВ

В существующих постановках задачи параметрического синтеза телекоммуникационных систем для описания информационных потоков, поступающих в сеть, обычно используется матрица требований передачи информации между конечными узлами сети. Однако при реализации инфокоммуникационной услуги абонент сети взаимодействует с некоторым множеством узлов управления услугой на разных

уровнях семиуровневой модели ВОС, что не учитывается данным способом описания поступающего в сеть трафика.

В данной статье для задания параметров трафика, поступающего в сеть при реализации инфокоммуникационной услуги, предлагается использовать описание с использованием моделей ON-OFF источников [3]. Согласно данной модели источник информации описывается с помощью трехуровневой модели (рис. 1):

– уровень соединения. Моделирует поведение абонента при переходе в активное состояние работы с сетью (например, установка соединения с модемным пулом интернет-провайдера);

– уровень сеанса. Моделирует активность прикладной задачи в процессе работы с сетью (например, обмен транзакциями с сервером или загрузка web-страницы);

– уровень пакетов. Это нижний уровень модели, который моделирует занятость канала связи в результате передачи пакета данных.

На каждом уровне модели поведение моделируемого объекта описывается с помощью FSM (Finite state machine), которая имеет два состояния: активное и пассивное. При этом в активном состоянии происходит передача информации, в пассивном состоянии обмен информацией не происходит или обмен идет с некоторой минимальной скоростью (обмен служебной информацией для управления сетевым соединением на соответствующем уровне). Параметрически поведение модели на каждом уровне определяется двумя случайными величинами – временем между переходами в активное состояние и длительностью активного состояния.

Характеристики трафика (обычно описываемые математическим ожиданием и дисперсией скорости передачи информации) на каждом уровне определяются рекурсивно через характеристики низлежащего уровня.

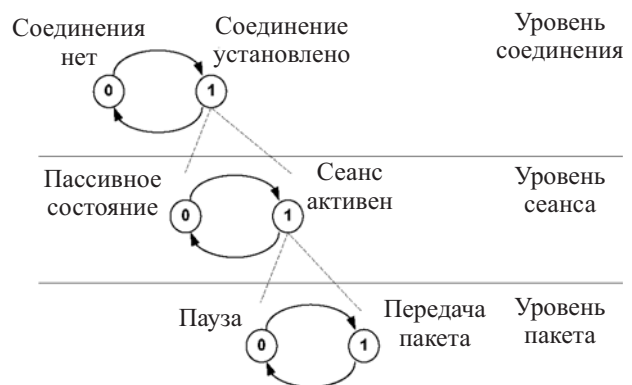


Рис. 1. Модель источника трафика

Скорость передачи информации на уровне соединения и ее стандартное отклонение σ_c определяется с учетом вероятности нахождения в пассивном P_0 и активном P_1 состоянии и скорости передачи информации в соответствующем состоянии. Расчет первых двух моментов для каждой статистической характеристики производится с учетом корректирующего коэффициента $\gamma = 2 \dots 3$. В табл. 1 приведена математическая модель, используемая для определения статистических характеристик телетрафика.

Таблица 1. Математическая модель трехуровневого источника

Уровень соединения	$\bar{V}_c = P_{0c} V_{\min} + P_{1c} V_s;$ $\sigma_c = \sqrt{m_c - \bar{V}_c^2};$ $m_c = P_{0c} (V_{\min})^2 + P_{1c} (V_s)^2;$ $V_c = \bar{V}_c + \sigma_c \gamma_c$
Уровень сеанса	$\bar{V}_s = P_{0s} V_{\min} + P_{1s} V_p;$ $\sigma_s = \sqrt{m_s - \bar{V}_s^2};$ $m_s = P_{0s} (V_{\min})^2 + P_{1s} (V_p)^2;$ $V_s = \bar{V}_s + \sigma_s \gamma_s$
Пакетный уровень	$\bar{V}_p = P_{0p} V_{\min} + P_{1p} V_{\max};$ $\sigma_p = \sqrt{m_p - \bar{V}_p^2};$ $m_p = P_{0p} (V_{\min})^2 + P_{1p} (V_{\max})^2;$ $V_p = \bar{V}_p + \sigma_p \gamma_p$

Приведенная выше модель описывает поведение одиночного источника и не отображает поведение трафика, возникающего в результате агрегирования информационных потоков от нескольких источников.

С использованием статистической аппроксимации [4], эквивалентную пропускную способность для передачи агрегированного информационного потока можно определить как

$$C = \min \left\{ \sum_{i=1}^N v_i + u \sqrt{\sum_{i=1}^N \sigma_i^2}, \sum_{i=1}^N V_i \right\} = \min \left\{ \sum_{i=1}^N v_i + u \sqrt{\sum_{i=1}^N v_i (V_i - v_i)}, \sum_{i=1}^N V_i \right\}, \quad (7)$$

где v_i – средняя скорость передачи i -го источника; σ_i – дисперсия скорости передачи i -го источника; V_i – максимальная скорость передачи i -го источника; u – нормирующий коэффициент, зависящий от допустимой ошибки ϵ :

$$u = \sqrt{-2 \ln(\epsilon) - \ln(2\pi)}. \quad (8)$$

Другой результат можно получить, используя биномиальное распределение. В данном случае каждый источник создает трафик v_0 , бит/с (в активном состоянии) и в случае постоянной скорости передачи суммарный трафик равен $N \cdot v_0$. С учетом наличия периодов неактивности эквивалентная пропускная способность $C = n \cdot v_0$, где n – эквивалентное число активных источников с постоянной скоростью передачи. Вероятность того, что k источников из N активно, можно определить, используя биномиальное распределение:

$$P(k) = C_{Np_{on}}^k (1 - p_{on})^{N-k}, \quad (9)$$

где p_{on} – вероятность того, что источник находится в активном состоянии.

Среднее количество активных источников $M[k] = N \cdot p_{on}$ и средняя скорость поступления информационного потока, создаваемого N источниками:

$$M[v] = N \cdot p_{on} \cdot v_0. \quad (10)$$

При превышении поступающим потоком величины пропускной способности узла, на котором происходит агрегирование, возникают потери. Для учета этого вводится корректирующий коэффициент $\gamma(P_B, p_{on}, N)$, зависящий от вероятности потери, вероятности активности и количества источников.

Значение пропускной способности для N агрегированных источников, бит/с:

$$C = M[v] + \gamma(P_B, p_{on}, N) \cdot \sigma(v) = \left(N \cdot p + \gamma(P_B, p_{on}, N) \cdot \sqrt{N \cdot p_{on} \cdot (1 - p_{on})} \right) \cdot v_0. \quad (11)$$

4. МОДЕЛИРОВАНИЕ ГРУППОВОГО ПОТОКА

Приведенные выше выражения позволяют найти статистические характеристики информационных потоков, поступающих на вход сети от группы абонентов. В процессе проектирования при решении задач параметрического синтеза необходимо знать характеристики групповых потоков, передаваемых по каналам связи.

В работе [5] автор приходит к выводу, что для сетей без установки соединения, где передается групповой поток от большого количества источников, в качестве модели потока можно использовать модель фрактального броуновского движения. Согласно данной модели, трафик можно описать следующим образом. Введем процесс A_t , описывающий количество информации (бит), переданное в течение временного интервала $[0, t)$:

$$A_t = vt + \sqrt{av} Z_t, \quad t \in (-\infty, \infty), \quad (12)$$

где Z_t – нормированное фрактальное Броуновское движение с параметром Хэрста H . Процесс описыва-

ется тремя параметрами: v – средняя интенсивность поступления информации, a – коэффициент разброса, H – параметр Хэрста.

Приняв $H = 0,5$, получим результаты, совпадающие с ранее известными в теории телетрафика для простейшего потока. Коэффициент разброса в этом случае равен единице ($a = 1$).

5. МЕТОДИКА ОПРЕДЕЛЕНИЯ ХАРАКТЕРИСТИК АГРЕГИРОВАННОГО ПОТОКА

В процессе передачи информационных потоков через сеть они претерпевают обработку в узлах сети. При использовании в качестве модели группового потока математической модели фрактального броуновского движения интересующими параметрами являются v , a и H .

В работе [6] было проведено исследование и установлено:

– при объединении в узле двух потоков с долговременной зависимостью с параметрами Хэрста H_1 и H_2 результирующий поток обладает долговременной зависимостью с параметром $H = \max(H_1, H_2)$;

– при объединении потока с долговременной зависимостью с параметром Хэрста H_1 и потока с кратковременной зависимостью результирующий поток обладает долговременной зависимостью с параметром $H = H_1$;

– при обработке в узле параметр Хэрста потока не изменяется.

Учитывая приведенные результаты и приняв для потока с кратковременной зависимостью значения параметров $H = 0,5$ и $a = 1$, можно задать следующие правила суммирования самоподобных потоков [7].

Таблица 2. Статистические характеристики агрегированного группового потока

Поток 1	Поток 2	Результирующий поток
H_1	H_2	$H = \max(H_1, H_2)$
$v; a; H$	$v; a; H$	$2v; a; H$
$v_1; a_1; H$	$v_2; a_2; H$	$v_1 + v_2; \frac{v_1 a_1 + v_2 a_2}{v_1 + v_2}; H$

ЗАКЛЮЧЕНИЕ

При решении задач параметрического синтеза мультисервисной телекоммуникационной сети NGN в качестве математической модели информационных потоков предлагается использовать модели самоподобных процессов.

Использование в качестве модели источника информационного потока модели трехуровневого ON-OFF источника позволяет учесть взаимодействие элементов

телекоммуникационной системы на различных уровнях модели ВОС. Приведены выражения, позволяющие определить статистические характеристики информационного потока, создаваемого пользователем на разных уровнях модели ВОС, а также характеристики потока, создаваемого группой абонентов.

В качестве модели группового потока, передаваемого в магистральных каналах телекоммуникационной системы, предлагается использовать модель фрактального Броуновского движения. Определены правила суммирования групповых потоков в узлах сети. Полученные выражения могут быть использованы при решении задачи распределения потоков с использованием моделей самоподобных процессов.

СПИСОК ЛИТЕРАТУРЫ

1. Цыбаков Б. С. Модель телетрафика на основе самоподобного случайного процесса / Б. С. Цыбаков // Радиотехника. – 1999. – № 5. – С. 24–31.
2. Шелухин О. И. Фрактальные процессы в телекоммуникациях / О. И. Шелухин, А. М. Тенякшев, А. В. Осин. – М.: Радиотехника, 2003. – 480 с.
3. Garcia A. E. Approximation to a Behavioral Model for Estimating Traffic Aggregation Scenarios // Journal of Universal Computer Science. – 2008. – Vol. 14, No. 5. – P. 731–744.
4. Wang S. Video Multiplexing with QoS Constraints / S. Wang, H. Zheng, J. A. Copeland // IEEE SPIE Conference on Internet Routing and QoS. – 1998. – P. 81–91.
5. Norros I. On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks / I. Norros // IEEE Journal on Selected Areas in Communications. – 1995. – No. 6. – P. 953–962.
6. Ng J. Statistical Delay Analysis on an ATM Switch with Self-similar Input Traffic / J. Ng, S. Song, W. Zhao // Information Processing Letters. North-Holland. – 2000. – Vol. 74. – P. 163–173.
7. Patel A. Statistical Multiplexing of Self-Similar Traffic: Theoretical and Simulation Results / A. Patel, C. Williamson // Department of Computer Science. Saskatchewan. – 1997. – P. 120–128.

Надійшла 02.12.2009
Після доробки 17.03.2010

Агеев Д. В., Копильов А. М.

МОДЕЛЮВАННЯ ІНФОРМАЦІЙНИХ ПОТОКІВ У МУЛЬТИСЕРВІСНІЙ МЕРЕЖІ NGN ПРИ ВИРІШЕННІ ЗАДАЧ ПАРАМЕТРИЧНОГО СИНТЕЗУ

У даній статті пропонуються математичні моделі інформаційних потоків і методи визначення їх характеристик на різних ділянках мультисервісної мережі NGN з метою їх використання при проектуванні телекомунікаційних систем.

Ключові слова: самоподібний процес, дисперсія, трафік, апроксимація, агрегація, фрактал.

Ageyev D. V., Kopyliev A. N.

INFORMATION STREAMS MODELING IN MULTISERVICE NETWORK NGN FOR PARAMETRICAL SYNTHESIS PROBLEM SOLVING

The authors propose mathematical models of information streams and methods of their characteristics definition at various sites of multiservice network NGN for the purpose of their usage when designing telecommunication networks.

Key words: self-similar process, dispersion, traffic, approximation, aggregation, fractal.

EXTENSION OF MICROINSTRUCTION FORMAT IN COMPOSITIONAL MICROPROGRAM CONTROL UNIT WITH ELEMENTARIZATION OF OPERATIONAL LINEAR CHAINS

The modification of synthesis method of compositional microprogram control units is directed to decrease of hardware amount in scheme of control unit in FPGA basis. Reduction of complexity of block of microinstruction addressing is reached due to of field with pseudoequivalent operational linear chain class code. Conditions of proposed method usage possibility are given. The example of method implementation is shown.

Key words: compositional microprogram control unit, operational linear chain, microinstruction addressing, class of pseudoequivalence.

INTRODUCTION

Modern industrial production requires cheap, secure and productive devices as results of design process. That's why reduces of the complexity of developing systems causes the topicality of hardware reduction problem in digital circuits [1]. It is necessary to take into account peculiarities of device structural scheme and element basis features. Some of the peculiarities, that can be used, are pseudoequivalent states and linear type of control algorithm [2].

Compositional microprogram control unit (CMCU) is reasonable to use in case of linear (percentage of operational vertices is over 75 %) algorithms realization [2]. FPGA (Field-Programmable Gate Arrays) basis is widely used nowadays for realization of control unit circuit [3, 4]. Problem of hardware amount minimization is solved by decrease in complexity of device main structural elements by means of decrease in main interconnections widths [5]. One of ways of this problem solving is proposed in the article: control unit realization as compositional microprogram one with code sharing and elementarization of operational linear chains (OLC).

The main purpose of investigation is simplification of combinational part of CMCU via implementation in the microinstruction format of additional field containing pseudoequivalent operational linear chain (POLC) class code. The main task of investigation is development of CMCU synthesis method modification that let decrease number of LUT-elements in Block of Microinstruction Addressing (BMA). Control algorithms are represented as graph-schemes (GSA).

MAIN STATEMENTS

Graph-scheme of control algorithm consists of operational and conditional vertices, making sets E_1 and E_2 accordingly, and the set of arcs E . Let us begin vertex be marked as b_0 , end – b_E . Operational vertex $b_q \in E_1$ contain set of microinstructions $Y(b_q) \subseteq Y$, where $Y = \{y_1, \dots, y_N\}$ is the set of output signals of control unit. Conditional vertex $b_g \in E_2$ contains one elements $X(b_g)$ of the logical conditions set $X = \{x_1, \dots, x_L\}$. In case of operational vertices percentage is over 75 % from total number of vertices, we talk about linear GSA.

OLC is a sequence of operational vertices of graph-scheme of algorithm. Each OLC α_g has accidental number of inputs I_g^i and only one output Q_g . Formal definitions of OLC, its input and output one can find in [4]. OLC with only one input and one output is called elementary [2].

OLC, outputs of which are connected with the input of the same vertex are called pseudoequivalent operational linear chains (POLC). Such OLCs make the class B_i . All classes are packed into the set $B = \{B_1, \dots, B_I\}$ of POLC classes.

Let GSA contains G elementary OLC α_g that form the set C .

$$R_1 = \lceil \log_2 G \rceil \quad (1)$$

bits are enough for encoding elements of the set C . Number of components in OLC α_g is marked as F_g . Maximum length $Q = \max(F_1, \dots, F_G)$ of linear chain

determines number of bits R_2 in the code for encoding OLC components, where

$$R_2 = \lceil \log_2 Q \rceil. \quad (2)$$

Elements $\tau_r \in \tau$ and $T_r \in T$ are used for encoding elementary OLC and their components accordingly. It being known that $|\tau| = R_1$ and $|T| = R_2$. Encoding of components is performed in natural order, that is

$$K(b_{gi}) = K(b_{gi-1}) + 1, \quad (3)$$

where $g = 1, \dots, G, i = 1, \dots, F_g$.

Each operational vertex $b_q \in E_1$ corresponds to microinstruction MI_q storing in control memory (CM) in the cell with address $A(b_q) = A_q$. Code sharing is obtaining of the address A_q as concatenation of OLC code and its component code.

Structure of compositional microprogram control unit with elementary OLC and code sharing can be used for interpretation of graph-scheme of control algorithm (Fig. 1). Let us call this structure U_1 .

Block of microinstruction addressing in CMCU scheme realized function of memory excitation for register RG:

$$\Psi = \Psi(X, \tau). \quad (4)$$

When signal Start is coming initial microprogram address is loaded into RG, zero value is loaded into CT, and flip-flop TF is set to "1" that allows reading microinstructions from control memory. There are two additional internal signals: y_0 and y_E . In case of $y_0 = 1$ content of CT is incremented and next vertex of current operational linear chain is addressed. If $y_0 = 0$ then OLC output is reached and BMA prepares address of next OLC using code of current POLC class. Signal y_E is used at the end of microprogram to reset flip-flop TF. The value "0" of TF output stops access to CM.

Asynchronous reset of counter must be controlled by function $\text{Start} \vee \overline{y_0}$. Signal $\overline{y_0}$ ensures loading zero value to the CT when transition to another OLC performed.

Number of terms in BMA scheme can be decreased by implementation OLC code transformer into POLC class codes [2]. But such realization demands extra FPGA recourses.

In the article complexity of code transformer is proposed to decrease by using free recourses of embedded memory.

MAIN IDEA OF PROPOSED METHOD

In initial GSA the set C_j contains OLC α_g , which are not connected to the end vertex of GSA. All operational

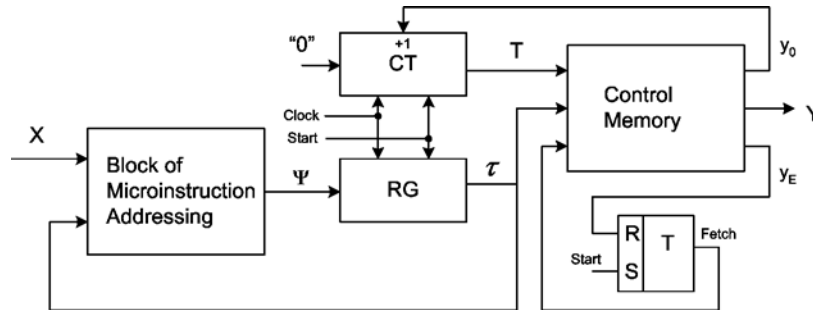


Fig. 1. Structure of compositional microprogram control unit with elementary OLC and code sharing

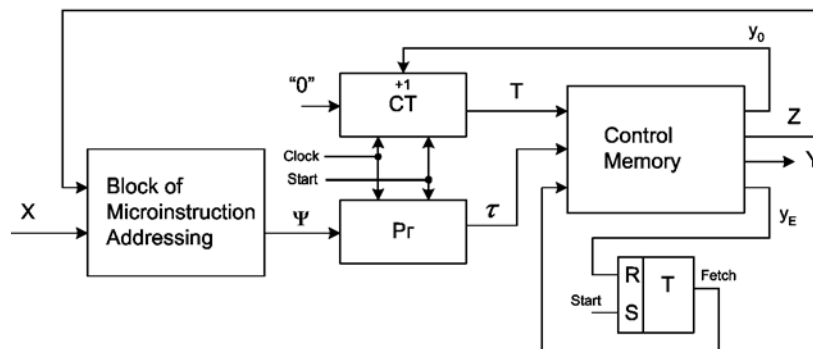


Fig. 2. Structure of compositional microprogram control unit with elementary OLC and code sharing after microinstruction format extension

linear chains are divided into classes $B_i \in \Pi_C$ of POLC. Binary code $K(B_i)$ of width R_3 is set to each class B_i , where

$$R = \lceil \log_2 I \rceil. \tag{5}$$

In (5) I is number of POLC classes. Let control memory of CMCU is realized on blocks of embedded memory with t output pins. Using unitary method of microinstructions encoding [2] we need

$$n_1 = N + 2, \tag{6}$$

bits in appropriate field, where $N = |Y|$, and constant 2 take into account internal signals y_0 and y_E . So, R_4 bits of the microinstruction may be free, where

$$R_4 = \left\lceil \frac{n_1}{t} \right\rceil * t - n_1. \tag{7}$$

If the condition

$$R_4 \geq R_3, \tag{8}$$

takes place, a field FB with the code of POLC class can be included into microinstruction format. Structure U_2 is obtained (Fig. 2).

In CMCU U_2 variables $z_r \in Z$, where $|Z| = R_3$, is bits of code $K(B_i)$. Block of microinstruction addressing performed function

$$\Psi = \Psi(Z, X). \tag{9}$$

Other blocks of CMCU U_2 perform corresponding functions to functions of CMCU U_1 blocks. Let us point out that structural elements BMA, CT, RG, TF is realized in LUT-elements, and CM is implemented in embedded memory.

The following method of CMCU U_2 synthesis is proposed in this article:

1. Construction of the sets C , C_1 , and Π_C for a GSA Γ .
2. Encoding of OLC, their components and classes $B_i \in \Pi_C$.
3. Construction of the content of control memory.
4. Construction of CMCU transition table and $\Psi = \Psi(Z, X)$ functions.
5. Synthesis of CMCU logic circuit.

EXAMPLE OF METHOD USING

Let GSA Γ_1 (Fig. 3) be characterized by next sets: $C = \{\alpha_1, \dots, \alpha_5\}$ – elementary OLC, $C_1 = C \setminus \alpha_5$ OLC without connection to the end vertex, $\Pi_C = \{B_1, \dots, B_3\}$ – classes of pseudoequivalent operational linear elementary chains, where $B_1 = \{\alpha_1\}$, $B_2 = \{\alpha_2, \alpha_3\}$, $B_3 = \{\alpha_4\}$. Number of OLC $G = 5$, $R_1 = 3$ bits from the set $\tau = \{\tau_1, \tau_2, \tau_3\}$ are enough for their encoding. Maximum length of OLC is $Q = 3$,

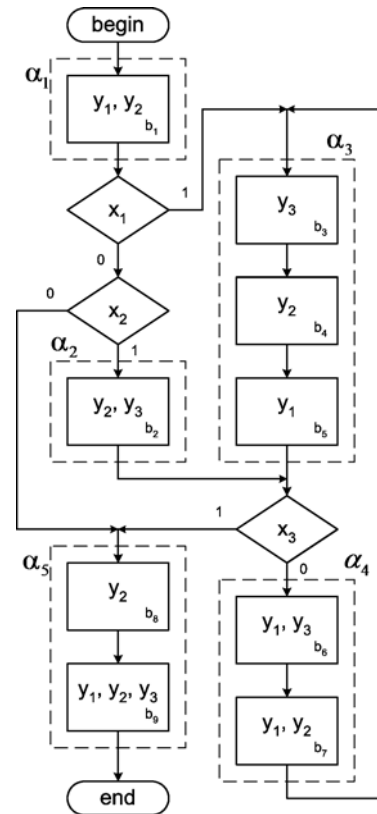


Fig. 3. Initial GSA Γ_1

let us use $R_2 = 2$ variables from the set $T = \{T_1, T_2\}$ for OLC components encoding. Total number of operational vertices is $M = 9$, this number demands $R = 4$ bit of address in CM. For encoding $I = 3$ classes $B_i \in \Pi_C$ of POLC $R_3 = 2$ bits are used.

Let us encode OLC $\alpha_g \in C$ and their components in arbitrary manner (3). Adresse $A(b_q)$ of CMCU $U_2(\Gamma_1)$ microinstruction are shown in Table 1. Here and after symbol $U_i(\Gamma_j)$ means, that CMCU U_i interprets GSA Γ_j .

Table 1. Addresses of CMCU $U_2(\Gamma_1)$ microinstruction

$T_1 T_2$ \ τ_1, τ_2, τ_3	000	001	010	011	100
00	b_1	b_2	b_3	b_6	b_8
01	–	–	b_4	b_7	b_9
10	–	–	b_5	–	–
11	–	–	–	–	–

From Table 1 one can obtain addresses, for example: $A(b_6) = 01100$, $A(b_9) = 10001$ and so on.

Codes of classes $B_i \in \Pi_C$ are set as $K(B_1) = 00, \dots, K(B_3) = 10$. Microinstruction format of CMCU U_2 includes fields y_0, y_E, FY, FB , where field FY contains

code of micro-operation set, FB – code of class $B_i \in \Pi_C$. If $y_0 = 1$, contents of FB field is ignored.

Let GSA Γ_1 includes $N = 3$ different microoperation y_n , and memory blocks with $t = 4$ output are used for realization of control memory in FPGA basis [6, 7]. In this case formula (7) gives us $R_4 = 3$ free bits. Because of $R_4 > R_3$, usage of proposed method is possible. So, in example $Z = \{z_1, z_2\}$. Contents of CMCU $U_2(\Gamma_1)$ control memory is shown in Table 2.

Table 2. Contents of CMCU $U_2(\Gamma_1)$ control memory

$A(b_q)$	y_0	FY	y_E	FB
$A(b_1)$	0	y_1, y_2	0	00
$A(b_2)$	0	y_2, y_3	0	01
$A(b_3)$	1	y_3	0	–
$A(b_4)$	1	y_2	0	–
$A(b_5)$	0	y_1	0	01
$A(b_6)$	1	y_1, y_3	0	–
$A(b_7)$	0	y_1, y_2	0	10
$A(b_8)$	1	y_2	0	–
$A(b_9)$	0	y_1, y_2, y_3	1	–

If vertex $b_q \in E_1$ is not an output of current OLC $\alpha_g \in C_1$, in memory cell with address $A(b_q)$ microoperation y_0 is written. In opposite case in FB field of this cell code $K(B_i)$ is written, where $\alpha_g \in B_i$. If vertex $b_q \in E_1$ is connected with the end of GS A than in memory cell with address $A(b_q)$ internal microoperation y_E is written.

Transitions from outputs of OLC $\alpha_g \in C_1$ are expressed by next system of formulae [2]:

$$\begin{aligned} B_1 &\rightarrow x_1 b_3 \vee \bar{x}_1 x_2 b_2 \vee \bar{x}_1 \bar{x}_2 b_8; \\ B_2 &\rightarrow x_3 b_8 \vee \bar{x}_3 b_6; \\ B_3 &\rightarrow b_3. \end{aligned} \tag{10}$$

Such system is the base for CMCU U_2 transition table formation. This table consists of next columns: $B_i, K(B_i), b_q, A(b_q), X_h, \Psi_h, h$. Their purpose became clear from Table 3.

Table 3. Fragment of CMCU $U_2(\Gamma_1)$ transition table

B_i	$K(B_i)$		b_q	$A(b_q)$			X_h	Ψ_h	h
	z_2	z_1		τ_1	τ_2	τ_3			
B_1	0	0	b_2	0	0	1	$\bar{x}_1 x_2$	D_3	1
			b_3	0	1	0	x_1	D_2	2
			b_8	1	0	0	$\bar{x}_1 \bar{x}_2$	D_1	3

Addresses of microinstruction is taken from Table 1. Let us point out, that system of memory excitation functions Ψ includes functions $\{D_1, D_2, D_3\}$. Total number of rows $H_2(\Gamma_j)$ in transition table of CMCU $U_2(\Gamma_j)$ is equal to number of terms in system transition formulae. In our example, $H_2(\Gamma_1) = 6$.

System (9) is formed according to transition table. Fragments of system Ψ can be found from Table 3:

$$\begin{aligned} D_1 &= \bar{z}_1 \bar{z}_2 \bar{x}_1 \bar{x}_2; \quad D_2 = \bar{z}_1 \bar{z}_2 x_1; \\ D_3 &= \bar{z}_1 \bar{z}_2 x_1 x_2. \end{aligned} \tag{11}$$

For minimization of terms number in (9) classes $B_i \in \Pi_C$ may be encoded with the help of EXPRESSO algorithm, for example.

Realization of logical circuit of CMCU U_2 reduces to implementation of system (9) in base of integrated circuit (FPGA) and realization of control memory on blocks of embedded or external memory. Modern CAD systems or methods [1, 2] can be used for this purpose.

CONCLUSIONS

Proposed method of microinstruction format extension for compositional microprogram control unit is oriented to LUT-elements decrease in the block of microinstruction addressing. Number of memory blocks in device and its working time are the same as for base structure CMCU U_1 with code sharing.

Disadvantage of proposed method is in its usage limitation (8).

Term number decrease in memory excitation functions can lead to decrease number of circuit levels in combinational part of device, that can increase speed of work.

Scientific novelty of proposed method modification is in usage of POLC classes and free recourses of control memory for LUT-elements number decrease in block of microinstruction addressing. Practical meaning is in chip parameters decrease. It allows realization of device with less cost.

Our future work is directed to development of CAD system for synthesis of compositional microprogram control units [4].

LIST OF REFERENCES

1. Соловьев В. В. Проектирование цифровых схем на основе программируемых логических интегральных схем. – М. : Горячая линия – ТЕЛЕКОМ, 2001. – 636 с.
2. Баркалов А. А. Синтез устройств управления на программируемых логических устройствах. – Донецк : ДНТУ, 2002. – 262 с.
3. Грушвицкий Р. И., Мурсаев А. Х., Узрюмов Е. П. Проектирование систем с использованием микросхем программируемой логики. – СПб : БХВ-Петербург, 2002. – 608 с.

4. Баркалов А. А., Тутаренко Л. А. Синтез микропрограммных автоматов на заказных и программируемых СБИС. – Донецк : УНИТЕХ, 2009. – 336 с.
5. De Micheli G. Synthesis and Optimization of Digital Circuits. – NY : McGraw-Hill, 1994. – 636 pp.
6. Virtex-6 Family Overview [Электронный ресурс] : Advance Product Specification / XILINX. – Электрон. дан. (1 файл). – [S. l.] : XILINX, 2010. – Режим доступа: http://www.xilinx.com/support/documentation/data_sheets/ds150.pdf, свободный. – Загл. с экрана. – Англ. яз.
7. Stratix III FPGA: Lowest Power, Highest Performance 65-nm FPGA [Электронный ресурс] / Altera. – Электрон. дан. – [S. l.] : Altera, 2010. – Режим доступа: <http://www.altera.com/products/devices/stratix-fpgas/stratix-iii/st3-index.jsp>, свободный. – Загл. с экрана. – Англ. яз.

Надійшла 5.04.2010

Баркалов О. О., Зеленёва І. Я., Мірошкін О. М.

РОЗШИРЕННЯ ФОРМАТУ МІКРОКОМАНД У КОМПОЗИЦІЙНОМУ МІКРОПРОГРАМНОМУ ПРИСТРОЇ КЕРУВАННЯ ІЗ ЕЛЕМЕНТАРІЗАЦІЄЮ ОПЕРАТОРНИХ ЛІНІЙНИХ ЛАНЦЮГІВ

Модифікація методу синтезу композиційного мікропрограмного пристрою керування спрямована на зменшення апаратних витрат при реалізації у FPGA базисі. Зменшення складності блоку адресації мікрокоманд досягається завдяки

полю, що містить код класу псевдоеквівалентного операторного лінійного ланцюга. Приведені умови доцільності та приклад використання модифікації методу синтезу.

Ключові слова: композиційне мікропрограмне устройство управления, операторная линейная цепь, адресация микрокоманд, класс псевдоэквивалентности.

Баркалов А. А., Зеленёва И. Я., Мирошкин А. Н.

РАСШИРЕНИЕ ФОРМАТА МИКРОКОМАНД В КОМПОЗИЦИОННОМ МИКРОПРОГРАММНОМ УСТРОЙСТВЕ УПРАВЛЕНИЯ С ЭЛЕМЕНТАРИЗАЦИЕЙ ОПЕРАТОРНЫХ ЛИНЕЙНЫХ ЦЕПЕЙ

Модификация метода синтеза композиционного микропрограммного устройства управления направлена на уменьшение аппаратных затрат при реализации в FPGA базисе. Уменьшение сложности блока адресации микрокоманд достигается за счет поля, содержащего код класса псевдоэквивалентной операторной линейной цепи. Приведены условия целесообразности и пример применения предложенной модификации метода синтеза.

Ключевые слова: композиционный микропрограммный пристрой керування, операторний лінійний ланцюг, адресация мікрокоманд, клас псевдоеквівалентності.

УДК 002.53+681.3(075.8)

Бойченко О. В.

Канд. техн. наук, професор Кримського юридичного інституту Одеського державного університету внутрішніх справ (м. Сімферополь)

КООРДИНАЦІЯ НЕЧІТКИХ РІШЕНЬ В БАГАТОРІВНЕВІЙ ІЄРАРХІЧНІЙ СИСТЕМІ

Проаналізовано методологічні засади керування складними багаторівневими інформаційними системами в умовах швидкої зміни порядку їх застосування. Запропоновано застосування комплексного системного підходу щодо координації прикладного (сеансового) та базового рівнів складних розгалужених інформаційних систем для оптимізації їхнього функціонування в умовах можливості виникнення нечітких рішень.

Ключові слова: багаторівневі ієрархічні системи, координація нечітких рішень, математичне моделювання та прогноз.

ПОСТАНОВКА ПРОБЛЕМИ

Аналіз практики застосування в діяльності організацій та установ розподілених інформаційно-телекомунікаційних систем визначає можливість виникнення обставин відсутності координації між базовим та прикладним рівнями управлінської ієрархії, що зменшує ефективність оперативного та достовірного обміну даними в діяльності закладу та потребує розробки відповідних заходів оптимізації функціонування комп'ютерної системи через застосування сучасних методів математичного моделювання та прогнозу з метою вирішення проблеми координації нечітких рішень в багаторівневій ієрархічній системі.

© Бойченко О. В., 2010

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Вирішенню проблем оптимізації складних ієрархічних систем управління стосовно координації нечітких рішень через застосування рекурентної процедури та процесу корекції в теорії нечітких множин присвячені роботи таких видатних фахівців, як Белман Р., Заде Л., Месарович М., Такахара Я., Моїсєєв Н. та інших [1–3]. Разом з тим, наявність низки проблем практики застосування управлінських ієрархічних систем визначає необхідність проведення подальших наукових досліджень з метою розробки відповідних заходів для їх вирішення. Наукова новизна запропонованої праці полягає в формуванні шляхів до

визначення якісно нового рівня комплексної системи заходів координації нечітких рішень в багаторівневій ієрархічній системі через застосування сучасних методів математичного моделювання та прогнозу.

АЛГОРИТМИ З'ЄДНАНЬ ЕЛЕМЕНТІВ ІЄРАРХІЧНИХ СИСТЕМ В УМОВАХ НЕЧІТКИХ РІШЕНЬ

Розглядаючи процедуру з'єднання елементів у багаторівневій ієрархічній системі, слід зазначити, що вона належить до класу безітеративних алгоритмів.

При цьому ухвалення оперативних рішень у складній ієрархічній системі (рис. 1) має за мету знаходження на кожному рівні вектора рішень, які забезпечують максимум системного вектора цільових функцій при координаційному завданні, отриманому від $(N + 1)$ -го рівня управління. Причому процес ухвалення рішень здійснюється дискретно в моменти часу й у загальному випадку, а крок дискретизації з управління зростає від нижчих рівнів до вищих. Частина рішень (в основному на нижчих рівнях) є управлінськими, а вищі рівні лише координують роботу підсистем нижчих рівнів.

Тому вирішення завдання оптимізації процесу управління в ієрархічній системі повинно враховувати підмножину допустимих для системи технологічно достовірних режимів $C \subset X$, тобто бути узгодженим з можливостями технології.

Спроба безпосереднього використання єдиного глобального критерію верхнього рівня з подальшою його декомпозицією для підсистем усіх рівнів робить завдання оптимізації доволі складним через низьку ефективність впливу на роботу системи нижчих

рівнів. Щодо глобальної цільової функції, то вона не залежить явним чином від рішень нижчих рівнів, які приймаються підсистемами, що ускладнює вибір режимів роботи підсистем і шляхів їх вдосконалення [2]. Тому доцільно припустити, що для кожної j -ї підсистеми i -го рівня задано свої цілі на безлічі вирішень підсистем, а системна цільова функція є векторною і залежить як від глобального критерію, так і від цільових функцій підсистем $F(\bar{x}_{ij})$.

Останнім часом прийнято досліджувати питання оптимізації для дворівневої системи та приймати це завдання як основний модуль для будь-якої N -рівневої системи [2], причому для вирішення завдання на $(i - 1)$ -му рівні вважається заданим результат оптимізації на рівні i , що можна представити таким чином:

$$\bar{x}_N = \{\bar{x}_{ij}\}, i = \overline{1, N}, j = \overline{1, M}, \quad (1)$$

де X_N – число підсистем на i -му рівні управління, за умови що всі рішення вже прийняті.

Розглядаючи різні ітеративні методи оптимізації в багаторівневих системах [3], доцільно вирізнити метод інваріантного занурення [4], у якому замість завдання з чітким рішенням x_{N+1}^0 на $(N + 1)$ -му рівні і фіксованим числом рівнів N , розглядатимемо деяку систему завдань з $N = 1, 2, 3, \dots; x_{N+1}^0 = [0, \infty]$. У такому разі максимум функцій залежатиме від N і \bar{x}_i , що є основною метою завдання. За таких умов початкове чітке завдання фактично розмивається, що приводить до нечіткого рішення для кожного рівня управління. Зазначене дозволяє застосувати для вирішення завдання апарат теорії нечітких множин.

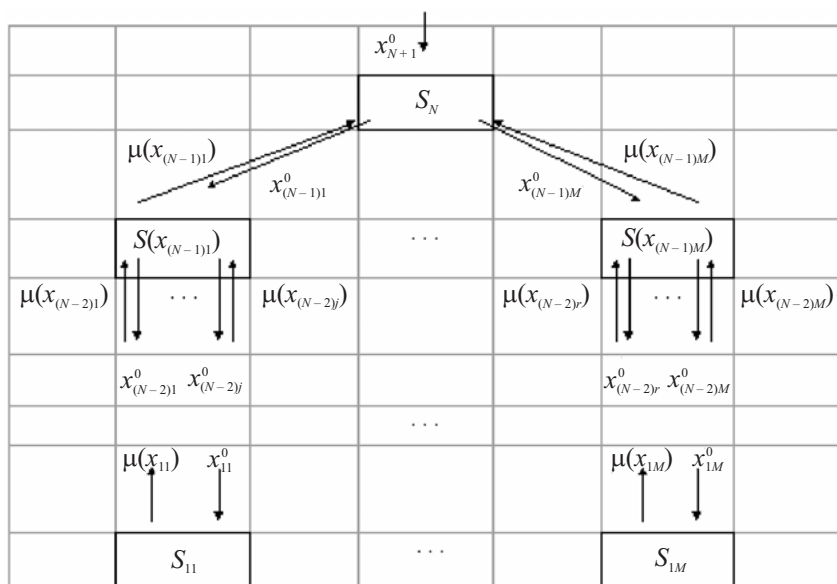


Рис. 1. Структурна схема багаторівневої ієрархічної системи

Застосування процедури навмисного внесення невизначеності або випадковості частково вирішує виключення впливу векторного критерію, дискретності та складності початкового завдання [4], але найбільш ефективним, на нашу думку, є застосування процедури раціоналізації рішення для однорівневих систем.

Для характеристики ухвалюваних рішень за окремими підсистемами j і для всієї системи в цілому використовуватимемо функцію приналежності ухвалюваних рішень до підмножини допустимих, ефективних і скоординованих рішень, причому ця функція не зводиться для чітких рішень до функції, оскільки вона містить не тільки характеристику ефективності рішення, але й характеристику його допустимості та скоординованості з цілями і обмеженнями нижчих підсистем. Функції приналежності визначаються безліччю рішень $X = \{x_{ij}\}$. Ефективність рішень задається нечіткою метою як нечітка підмножина з функцією приналежності, допустимість рішень – нечіткою підмножиною з функцією приналежності, а скоординованість – нечіткою множиною з функцією приналежності $\mu_K(\bar{x}_{ij})$.

Найчастіше для визначення функції приналежності користуються співвідношенням

$$\mu_{G \cap C \cap K}(x_{ij}) = \mu_G(x_{ij}) \wedge \mu_C(x_{ij}) \wedge \mu_K(x_{ij}) = \min\{\mu_G(x_{ij}), \mu_C(x_{ij}), \mu_K(x_{ij})\}; x_{ij} \in X. \quad (2)$$

Тоді нечітку підмножину називатимемо нечітким вирішенням підсистеми j -рівня, причому

$$\mu_D(x_{ij}) = \mu_{G \cap C \cap K}(x_{ij}). \quad (3)$$

За наявності матеріальних потоків між ієрархічними рівнями або різних рівнянь матеріального балансу, що описують групи підсистем, рішення для підсистеми r на рівні $(i + 1)$ може бути пов'язане з вирішенням підсистем i -го рівня рівнянням

$$x_{(i+1)r} = f(x_{i1}, \dots, x_{iM}). \quad (4)$$

У разі, коли безліч рішень X складається тільки з одного параметра, рівняння може бути записане у вигляді

$$x_{(i+1)r} = \sum_{j=1}^M x_{ij}. \quad (5)$$

Рішення можуть бути складнішими та характеризуватися вектором різноманітних параметрів. Тоді у загальному випадку рішення можуть ухвалюватися і для нестационарних режимів роботи системи з застосуванням диференціальних рівнянь у частинних похідних.

Тоді системне скоординоване рішення на рівні N може бути визначене таким чином:

$$D(x_N) = D(x_{11}) \cdot \dots \cdot D(x_{1M}) \cdot D(x_{21}) \cdot \dots \times \\ \times D(x_{2M}) \cdot \dots \cdot D(x_{(N-1)1}) \cdot \dots \times \\ \times D(x_{(N-1)M}) \cap C(x_N) \cap G(x_N). \quad (6)$$

Системне рішення (6) дозволяє отримати просту рекурентну процедуру для ухвалення рішень в N -рівневій ієрархічній системі.

А внаслідок інваріантності введеного опису за рівнями системи з (6) витікає, що

$$D(x_{ij}) = D(x_{11}) \cdot \dots \cdot D(x_{1M}) \cdot D(x_{21}) \cdot \dots \times \\ \times D(x_{2M}) \cdot \dots \cdot D(x_{(i-1)1}) \cdot \dots \times \\ \times D(x_{(i-1)M}) \cap C(x_{ij}) \cap G(x_{ij}), \quad (7)$$

$$D(x_{(i+1)r}) = D(x_{11}) \cdot \dots \cdot D(x_{1M}) \cdot D(x_{21}) \cdot \dots \times \\ \times D(x_{2M}) \cdot \dots \cdot D(x_{i1}) \cdot \dots \times \\ \times D(x_{iM}) \cap C(x_{(i+1)r}) \cap G(x_{(i+1)r}). \quad (8)$$

У зв'язку з тим, що в ієрархічній структурі рішення на i -му рівні управління залежить тільки від перебування на цьому рівні та рішень, що приймаються вищими $(i + 1)$ -м і нижчими $(i - 1)$ -м рівнями управління (що обумовлено самим принципом ухвалення рішень в ієрархічних системах), у рішенні враховуються всі рішення нижчих підсистем. Тоді (8) можна записати через нечіткі вирішення нижчого рівня таким чином:

$$D(x_{(i+1)r}) = [D(x_{i1}) \cdot \dots \cdot D(x_{iM})] \cap \\ \cap C(x_{(i+1)r}) \cap G(x_{(i+1)r}). \quad (9)$$

Причому для ухвалення рішення управлінська підсистема r -рівня $(i + 1)$ повинна заздалегідь визначити рішення нижчих підсистем у системне рішення $D(x_{i1}, \dots, x_{iM})$. Тоді (9) можна записати у вигляді

$$D(x_{(i+1)r}) = D(x_{i1}, \dots, x_{iM}) \cap \\ \cap C_{(i+1)r} \cap G(x_{(i+1)r}). \quad (10)$$

Наявність такої рекурентної процедури обумовлена особливістю ухвалення рішень у ієрархічній структурі та наявністю самостійних прав з обробки інформації та ухвалення рішень у кожній системі управління.

Проте на відміну від загальноприйнятого представлення рішень у чіткій формі (тобто у вигляді конкретних числових значень параметрів управління) допустимо, що вирішення всіх підсистем у процесі координації рішень є нечіткими і задаються для кожної підсистеми функціями приналежності $\mu(x)$. Тоді для функцій приналежності даної підмножини ефективних і допустимих рішень рівняння (10) набуває вигляду

$$\mu_D(x_{(i+1)r}) = \mu_D(x_{i1}, \dots, x_{iM}) \otimes \\ \otimes \mu_G(x_{(i+1)r}) \otimes \mu_C(x_{(i+1)r}), \quad (11)$$

де

$$\mu_D(x_{i1}, \dots, x_{iM}) = \mu_D(x_{i1}) \otimes \dots \otimes \mu_D(x_{iM}). \quad (12)$$

Функція приналежності характеризує міжрівневу координацію, причому для ухвалення рішення на $(i + 1)$ -му рівні необхідно спроектувати координуюче рішення. Тоді функцію можна буде представити у вигляді

$$\mu_K(x_{(i+1)r}) = \mu_D \left(\sum_{j=1}^M x_{ij} \right). \quad (13)$$

Рівняння (13) наочно показує, що рішення вищого рівня може бути виконане за допомогою різних вирішень нижчих підсистем x_{i1}, \dots, x_{iM} . Причому, аналізуючи вираз (12), можна зробити висновок про те, що за інших рівних умов робота підсистеми r рівня $(i + 1)$ буде найбільш ефективною у разі, коли кожна підсистема j -рівня і ухвалюватиме оптимальні рішення (максимізувавши свою функцію приналежності). Визначення рішень підсистем нижчого рівня є максимізуванням проекції цих рішень через

$$\mu_K \left(\sum_{j=1}^M x_{ij} \right) = \max_{\{x_{ij}\}} [\mu_D(x_{i1}) \otimes \dots \otimes \mu_D(x_{iM})]. \quad (14)$$

Це дає можливість характеризувати оптимальну стратегію подібно до принципу оптимальності Р. Беллмана [4].

Застосування принципу оптимальності дозволяє реалізувати найбільш ефективний спосіб координації рішень, що приймаються на окремих рівнях ієрархії. Цей принцип дозволяє до мінімуму скоротити обмін інформацією між рівнями й забезпечити локальну обробку інформації за підсистемами.

Але тільки використання єдиного підходу до цільових функцій і обмежень у теорії нечітких множин дозволяє отримати адекватні характеристики нижчих підсистем, що не зводяться тільки до економічних аддитивних характеристик в умовах повної визначеності.

На основі цього принципу можна побудувати рекурентну процедуру ухвалення рішень, яка складатиметься з двох етапів для кожного $(i + 1)$ -го рівня управління.

По-перше, вирішується оптимізаційне завдання на декартовому полі множин з цільовою функцією (14) для рівня $(i + 1)$ з урахуванням відповідного обмеження. Якщо підсистеми рівня i пов'язані між собою матеріальними потоками, то попередні обмеження доповнюють системою обмежень рівностей та нерівностей підсистем, що описують зв'язок, на рівні i :

$$\bar{H}(x_{i1}, \dots, x_{iM}) = 0, \quad (15)$$

$$\bar{W}(x_{i1}, \dots, x_{iM}) \geq 0. \quad (16)$$

Результатом рішення оптимізаційної задачі є функція приналежності:

$$\mu_K(x_{(i+1)r}) = \mu_D \left(\sum_{j=1}^M x_{ij} \right). \quad (17)$$

По-друге, знаходиться нечітке рішення на $(i + 1)$ -му рівні. При цьому, для ухвалення чіткого рішення на N -му рівні управління необхідно вибрати таке рішення, для якого ступінь приналежності до підмножини ефективних і допустимих рішень є максимальним, тобто

$$\mu_D(x_N^0) = \max_{\{x_{\mu}\}} [\mu_D(x_N)]. \quad (18)$$

Після ухвалення чіткого рішення на N -му рівні управління на основі рішення оптимізаційного завдання (15), (16), (17) знаходяться чіткі вирішення рівня $(N - 1)$ $x_{(N-1)1}^0, \dots, x_{(N-1)M}^0$. Далі ця процедура повторюється до тих пір, поки не будуть чітко ухвалені всі рішення в системі $\{x_{ij}\}, i = \overline{1, N}, j = \overline{1, M}$.

Підсистеми управління першого рівня можуть бути елементарними й не мати координуючих функцій, тобто

$$\mu_K(x_{1r}) = 1, \quad r = \overline{1, M}. \quad (19)$$

Активність підсистем першого рівня в цьому випадку виявляється у формуванні своїх нечітких рішень:

$$D(x_{1r}) = C(x_{1r}) \cap G(x_{1r}). \quad (20)$$

Проте за наявності взаємозв'язку підсистем першого рівня через технологічний процес виникає природна необхідність координації рішень підсистем, виходячи з цільової функції і моделей, що описують цей процес. У цьому разі координуюча функція буде отримана на основі вирішення оптимізаційних завдань для складних технологічних процесів.

Цільові функції підсистем вищих рівнів можуть бути специфічними для кожного рівня і полягати лише в забезпеченні ефективних режимів роботи нижчих підсистем:

$$\mu_G(x_{(i+1)r}) = 1. \quad (21)$$

Якщо ж вища підсистема має власну мету G , то при ухваленні рішень вона враховуватиме цілі і можливості нижчих підсистем через координуюче нечітке рішення. Ця функція приналежності гарантує призначення нижчим підсистемам режимів роботи, близьких до найбільш переважних для них режимів, тобто фактично відображає факт обліку інтересів підсистем у процесі ухвалення рішень.

У загальному випадку функція приналежності може бути умовною, тобто залежати від деякого параметра – $\mu(x_{ij}/y_{ij})$.

ВИСНОВКИ

Підсумовуючи, слід зазначити, що розглянута рекурентна процедура забезпечує вирішення внутрішніх системних конфліктів. Застосування двоетапної рекурентної процедури дозволяє вирішити у складі системи конфлікти двох видів, а саме:

– внутрішньорівневий, який виникає з приводу несутимності локальних цілей окремих елементів у межах заданої структури і обмежень;

– міжрівневий (між двома суміжним рівнями), який має місце через те, що мета (завдання) вищого елемента не збігається із забезпеченням елементам нижчого рівня оптимальних режимів.

Особливості запропонованого методу дозволяють не тільки встановлювати наявність або відсутність подібних конфліктів і їх природу на основі використання відповідних функцій приналежності, яка є характеристикою багаторівневої системи, але й приймати ці функції приналежності як орієнтири для синтезу й цілеспрямованої зміни параметрів і структури системи.

Таким чином, розглянута методика направлена на узгодження рішень за рівнями в багаторівневій ієрархічній системі в певний момент t , причому для кожного, у тому числі й нижчого рівня (основного технологічного процесу), може бути вирішено також і динамічне завдання.

СПИСОК ЛІТЕРАТУРИ

1. Беллман Р. Принятие решений в расплывчатых условиях / Р. Беллман, Л. Заде // Научно-практический сбор-

ник «Вопросы анализа и процедуры принятия решений». – М. : Мир, 1976. – С. 172–215.

2. Месарович М. Общая теория систем : монография / М. Месарович, Я. Такахара. – М. : Мир, 1978. – 345 с.
3. Моисеев Н. Н. Методы оптимизации: монография / Моисеев Н. Н. – М. : Наука, 1978. – 351 с.
4. Беллман Р. Прикладные задачи динамического программирования: монография / Р. Беллман, С. Дрейфус. – М. : Наука, 1965. – 333 с.
5. Растринин Л. А. Системы экстремального управления: монография / Растринин Л. А. – М. : Наука, 1974. – 234 с.
6. Заде Л. А. Понятие лингвистической переменной и его применение к принятию приближенных решений: монография / Заде Л. А. – М. : Мир, 1976. – 165 с.

Надійшла 15.04.2010

Бойченко О. В.

КООРДИНАЦІЯ НЕЧЕТКИХ РЕШЕНИЙ В МНОГОУРОВНЕВОЙ ИЕРАРХИЧЕСКОЙ СИСТЕМЕ

Проанализированы методологические принципы управления сложными многоуровневыми информационными системами в условиях быстрого изменения порядка их использования. Предложено применение комплексного системного подхода относительно координации прикладного (сеансового) и базового уровней сложных разветвленных информационных систем для оптимизации их функционирования в условиях возможности возникновения нечетких решений.

Ключевые слова: многоуровневые иерархические системы, координация нечетких решений, математическое моделирование и прогноз.

Boychenko O. V.

CO-ORDINATION OF FUZZY DECISIONS IN MULTI-LEVEL HIERARCHICAL SYSTEM

The author analyses methodological principles of complex multilevel information systems control in conditions of quickly changing order of their use. Application of complex system approach is proposed in relation to co-ordination of applied (session) and base levels of complex branched information systems for optimization of their functioning at probable origination of fuzzy decisions.

Key words: multilevel hierarchical systems, co-ordination of fuzzy decisions, mathematical simulation and prediction.

УДК 681. 326

Дубинская Н. Г.

Аспирант Харьковского национального университета радиозлектроники

СТРУКТУРНАЯ ДИАГНОСТИЧЕСКАЯ МОДЕЛЬ КОМПЬЮТЕРНОЙ СЕТИ

Усовершенствован структурный метод поиска явных адресуемых неисправностей компьютерной сети, который отличается от известных сочетанием структур данных условного и безусловного методов поиска дефектов.

Ключевые слова: диагностика, компьютерная сеть, поиск дефектов, модель, граф.

ВВЕДЕНИЕ

Широкое использование локальных и глобальных компьютерных сетей приводит к росту требований к надежности, отказоустойчивости и производительности

локальных вычислительных сетей (ЛВС). Высокая производительность сети обеспечивается, в первую очередь, отсутствием явных и скрытых узких мест и дефектов, приводящих как к замедлению

© Дубинская Н. Г., 2010

скорости работы в сети, так и к недостижимости отдельных элементов сети и выходу из строя коммуникационных компонентов. При этом существенным является время, затрачиваемое на восстановление работоспособности ЛВС [1]. Решение задач диагностирования ЛВС представляет собой сложную задачу. Это связано с тем, что сетевые неисправности делятся на различные типы, для поиска каждого из которых необходимо использовать различные методы и виды диагностического оборудования. Поиск и устранение неисправностей программного обеспечения (ПО), как правило, не входит в задачу диагностирования ЛВС, и отсюда возникает дополнительная проблема отделения неисправностей прикладного ПО от неисправностей сети, для определения которых разрабатывается соответствующий метод. Кроме этого, поиск неисправностей даже одного типа усложняется отсутствием единого формализованного подхода, общего алгоритма действий администратора-диагноста. В практике диагностирования ЛВС используются различные способы представления ЛВС как объекта диагностирования (ОД), каждый из которых имеет свои преимущества и недостатки, но не является единым. Отсутствие формализованного метода определения области подозреваемых неисправностей приводит к высоким временным затратам на проведение диагностического эксперимента и, следовательно, на поиск неисправности. Также в настоящий момент времени отсутствует единая формализованная методология, позволяющая локализовать любой из видов неисправностей. Это приводит к необходимости наличия у диагноста достаточно высокого уровня опыта и знаний в области сетевых технологий для обеспечения корректности постановки диагностического эксперимента. Указанные проблемы обуславливают высокие временные затраты на поиск неисправности, а также сужают диапазон субъектов, обеспечивающих корректное решение задачи поиска неисправности, что приводит к высокой трудоемкости и сложности решения задачи диагностирования ЛВС.

ПОСТАНОВКА ЗАДАЧИ

В процессе эксплуатации компьютерных сетей целесообразно иметь диагностическую систему, с использованием которой можно восстановить работоспособность ЛВС за возможно меньший промежуток времени. В этом случае стремятся уменьшить MTTR (Mean Time To Repair) – среднее время восстановления работоспособного состояния объекта после отказа T_b .

Если для конкретной ЛВС имеем определенное MTTF (Mean operating Time To Failure) – среднюю на-

работку до первого отказа T_{cp} , то коэффициент $K = T_{cp}/T_{cp} + T_b$ характеризует функционирование ЛВС при наличии отказа и необходимости восстановления работоспособного состояния компьютерной сети. Если коэффициент K приближенно равен 1, то это означает, что время T_b пренебрежимо мало по сравнению с T_{cp} .

Таким образом, актуальной задачей является разработка методов поиска неисправностей в ЛВС и их сегментах, которые обеспечат снижение трудоемкости, сокращение временных затрат на поиск неисправности в ЛВС, а также повышение эффективности диагноза за счет обеспечения требуемой глубины поиска неисправности.

ДИАГНОСТИЧЕСКАЯ МОДЕЛЬ КОМПЬЮТЕРНОЙ СЕТИ

Представим структуру ЛВС в виде ориентированного графа $G = (V, E)$, с числом узлов $0, 1, 2, \dots, V$ и с числом дуг-связей между ними $1, 2, \dots, E$ [2]. Считаем, что структура ЛВС имеет в своем составе узел – источник тестов, которым может быть заведомо исправный конечный узел (обычно сервер). Данный узел генерирует входные тестовые воздействия, принимает и анализирует выходные реакции на тесты, а также задает ориентированность графа. Обозначим узел – источник тестов в графовой модели структуры узлом с единичным номером V_1 . Назовем данный узел базовым узлом структуры. Для обеспечения процесса диагностирования базовый узел должен иметь достижимость к каждому узлу графа, что обеспечивается свойством явной и промежуточной адресуемости любого сетевого компонента и использованием адресного теста. Более того, так как базовый узел представлен заведомо исправным конечным узлом, то нет необходимости включать его в матрицу достижимостей, целью построения которой является определение одиночной неисправности (ОИ).

Сигналы – реакции на адресные тесты снимаются с выходных контрольных точек, которые будем называть пассивными. При этом должна обеспечиваться контрдостижимость к любому узлу графа от пассивных контрольных точек для обеспечения получения реакции на тест базовым узлом. Выполнение данного условия для модели ЛВС как ОД обеспечивается отсутствием в сети нескольких маршрутов от определенного узла-источника к узлу-приемнику, исходя из определения ЛВС.

Понятия достижимости и контрдостижимости обычно иллюстрируются с использованием матриц или соответствующих им множеств [3]. Матрица дос-

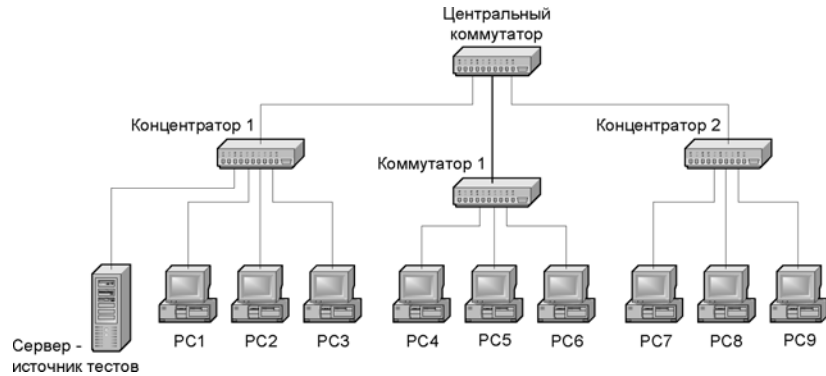


Рис. 1. Фрагмент ЛВС с сервером – источником тестов

тижимостей $M = [m_{ij}]$ описывает возможные пути от узла V_i к узлу V_j . При этом элемент матрицы равен:

$$m_{ij} = \begin{cases} 1 \rightarrow \exists V_i V_j; \\ 0 \rightarrow \bar{\exists} V_i V_j, \end{cases} \quad (1)$$

где $V_i V_j$ – путь от узла-источника тестов V_i к узлу-приемнику V_j .

Множество узлов $B(V_i)$ графа G , достижимых из заданного узла V_i , состоит из таких элементов V_j , для которых L_{ij} -й элемент в матрице достижимостей равен 1. Очевидно, что все диагональные элементы в матрице B равны 1, так как каждый узел достижим из себя самого с помощью пути длиной 0. Пусть $T_1(V_i)$ является множеством узлов V_j , которые достижимы из V_i путями длины 1; $T_2(V_i)$ – множество узлов, достижимых из V_i с использованием путей длины 2. Аналогично $T^k(V_i)$ является множеством узлов, достижимых из V_i путями длины K . Все множество узлов, достижимых из V_i , можно представить в следующем виде:

$$B(V_i) = \{V_i\} \cup \{T^1(V_i)\} \cup \{T^2(V_i)\} \cup \dots \cup \{T^k(V_i)\}. \quad (2)$$

Аналогично определим матрицу контрдостижимостей $Q(V_i) = [q_{ij}]$, элемент которой равен:

$$q_{ij} = \begin{cases} 1 \rightarrow \exists V_i V_j; \\ 0 \rightarrow \bar{\exists} V_i V_j. \end{cases} \quad (3)$$

Контрдостижимым множеством $Q(V_i)$ графа G является такое множество узлов, что из его любого узла может быть достигнут узел V_i .

Построение матрицы достижимостей предваряется выполнением процедуры индексирования ориентированного графа, представляющего здесь структурную модель ЛВС как ОД.

1. Присвоение номера 1 источнику тестов.

2. Индексация промежуточных систем (Intermediate Systems).

3. Индексация конечных систем (End Systems).

Ниже представлен фрагмент ЛВС (рис. 1), для которой будет построена модель в виде индексированного ориентированного графа (рис. 2). Применение структурного метода к предложенной модели ЛВС обуславливается выполнением следующих условий [1]:

- ориентированность графа модели сети, обеспечивающаяся, как было сказано ранее, наличием адреса-источника и адреса-приемника у любого тестового воздействия;

- однонаправленность графа модели сети, обеспечивающаяся за счет наложения ограничения на двунаправленный диагностический сетевой трафик (от центрального диагностического узла поступает запрос на остальные узлы сети об их состоянии, на который каждый из узлов шлет требуемую информацию). Согласно указанному ограничению, будет считаться, что реакция на тест будет получена источником тестов по умолчанию (недоступность диагностируемого узла ввиду его неисправности также считается реакцией на запрос, так как источник тестов получит соответствующее уведомление) и будет доставлена к источнику тестов по маршруту тестового воздействия, что обеспечивается определением ЛВС.

Таким образом, каждому направленному маршруту «источник тестов – приемник тестов» ставится

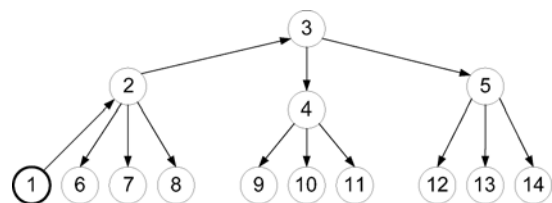


Рис. 2. Графовая структурная модель фрагмента ЛВС

в соответствии маршрут «приемник тестов – источник тестов», не обозначенный на графовой модели ЛВС, но подразумевающийся как имеющий место по умолчанию (рис. 2).

Матрица достижимостей для структурного метода поиска неисправностей строится по полученной графовой модели ЛВС с учетом всех наложенных на нее ограничений, а также в соответствии с определением матрицы достижимостей, приведенным выше.

Ниже представлена матрица достижимостей M_j , построенная для графовой модели фрагмента ЛВС (рис. 2), и соответствующий вектор элементарных проверок R_i , состоящий из вершин – конечных узлов рассматриваемой сети:

$$M_j = \begin{array}{c|cccccccccccccc} & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ \hline 2 & 1 & . & . & . & . & . & . & . & . & . & . & . & . \\ 3 & 1 & 1 & . & . & . & . & . & . & . & . & . & . & . \\ 4 & 1 & 1 & 1 & . & . & . & . & . & . & . & . & . & . \\ 5 & 1 & 1 & . & 1 & . & . & . & . & . & . & . & . & . \\ 6 & 1 & . & . & . & 1 & . & . & . & . & . & . & . & . \\ 7 & 1 & . & . & . & . & 1 & . & . & . & . & . & . & . \\ 8 & 1 & . & . & . & . & . & 1 & . & . & . & . & . & . \\ 9 & 1 & 1 & 1 & . & . & . & . & 1 & . & . & . & . & . \\ 10 & 1 & 1 & 1 & . & . & . & . & . & 1 & . & . & . & . \\ 11 & 1 & 1 & 1 & . & . & . & . & . & . & 1 & . & . & . \\ 12 & 1 & 1 & . & 1 & . & . & . & . & . & . & 1 & . & . \\ 13 & 1 & 1 & . & 1 & . & . & . & . & . & . & . & 1 & . \\ 14 & 1 & 1 & . & 1 & . & . & . & . & . & . & . & . & 1 \end{array}$$

$$R_i = \begin{array}{c} R_6 \\ R_7 \\ R_8 \\ R_9 \\ R_{10} \\ R_{11} \\ R_{12} \\ R_{13} \\ R_{14} \end{array}$$

Исходя из определения ЛВС как ОД, графовая модель сети может быть представлена в расширенном виде, где каждая вершина, представляющая конечный узел сети, рассматривается как совокупность компонентов конечного узла, каждый из которых может быть носителем явной адресуемой неисправности. Таким образом, рабочая станция сегмента ЛВС 10BaseT может быть представлена расширенной вершиной $PC_i = (ПК_i, СА_i, К_i, П_i)$, где $ПК_i$ – рабочая станция, $СА_i$ – сетевой адаптер, $К_i$ – кабельный сег-

мент от рабочей станции к коммутирующему устройству, $П_i$ – порт концентратора/коммутатора.

Каждый из указанных компонентов будет рассматриваться при построении матрицы достижимостей и определении области подозреваемых неисправностей.

ДИАГНОСТИРОВАНИЕ НЕИСПРАВНОСТЕЙ В ЛВС

В случае предположения о наличии в сети единичной неисправности для ее нахождения применяется методика на основе формулы (4), которая обычно используется для анализа многовыходовых схем при структурном методе поиска дефектов [1]:

$$D = \bigcap M_j^{R_i=1} \setminus \bigcup M_j^{R_i=0}, \tag{4}$$

где M_j – строка матрицы достижимости.

Затем для полученной области применяются традиционные деревья с равной ценой элементарной проверки. Если $OH D = \emptyset$, то в модели предполагается наличие кратной неисправности. В случае возникновения кратной неисправности в сети (например, данная ситуация естественна в шинной топологии), применяется формула (5):

$$D = \bigcup M_j^{R_i=1} \setminus \bigcup M_j^{R_i=0}. \tag{5}$$

ВЫВОДЫ

Структурная диагностическая модель компьютерной сети, предложенная в статье, позволила реализовать структурный метод поиска явных адресуемых неисправностей в сегменте ЛВС, что является новым подходом к решению задачи определения ОН явно адресуемого типа. Данный метод представляет собой сочетание структур данных условного и безусловного методов поиска дефектов, опирается на свойство явной и промежуточной адресуемости сетевого компонента и использует адресный тест для определения состояния сетевого компонента как вид теста. При этом обеспечивается элементарная проверка с наименьшими аппаратными затратами за заданное время.

СПИСОК ЛИТЕРАТУРЫ

1. *Бондаренко М. Ф.* Проектирование и диагностика компьютерных систем и сетей / М. Ф. Бондаренко, Г. Ф. Кривуля, В. Г. Рябцев, С. А. Фрадков, В. И. Хаханов. – Харьков : ХНУРЭ, 2000. – 306 с.
2. *Кривуля Г. Ф.* Компьютерная диагностика сложных систем / Г. Ф. Кривуля // Информационно-управляющие системы на железнодорожном транспорте. – 1996. – № 3, 4. – С. 24–29.
3. *Кристофидес Н.* Теория графов / Н. Кристофидес. – М. : Мир, 1978. – 432 с.

Надійшла 12.03.2010

Дубінська Н. Г.

СТРУКТУРНА ДІАГНОСТИЧНА МОДЕЛЬ КОМП'ЮТЕРНОЇ МЕРЕЖІ

Удосконалено структурний метод пошуку явних несправностей комп'ютерної мережі, що адресуються, який відрізняється від відомих поєднанням структур даних умовного і безумовного методів пошуку дефектів.

Ключові слова: діагностика, комп'ютерна мережа, пошук дефектів, модель, граф.

Dybinsky N. G.

STRUCTURAL DIAGNOSTIC MODEL OF COMPUTER NETWORK

The paper is dedicated to solving the tasks of time decreasing, cost reducing and efficiency increasing for the procedure of network fault location and diagnosis formulation due to providing appropriate depth of fault location. The structural method of explicit addressable faults location has been proposed.

Key words: diagnosis, computer network, search for defects, model, graph.

УДК 004.657

Кривуля Г. Ф.¹, Сыревич Е. Е.², Карасев А. Л.³

¹Д-р техн. наук, заведуючий кафедрой Харьковского национального университета радиозлектроники

²Канд. техн. наук, доцент Харьковского национального университета радиозлектроники

³Ассистент Харьковского национального университета радиозлектроники

АНАЛИЗ СПИСКА СОЕДИНЕНИЙ В СИСТЕМАХ ЛОГИЧЕСКОГО СИНТЕЗА

Рассмотрена проблема анализа результатов логического синтеза для их последующего анализа на соответствие правилам синтеза. Предложена табличная модель, состоящая из получаемого после фазы распространения словаря атрибутов. Определены этапы поиска по заданным параметрам при использовании разработанной модели.

Ключевые слова: логический синтез, HDL, атрибут линии, макрос, поиск по параметрам.

1. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ И ПОСТАНОВКА ЗАДАЧИ

Маршрут проектирования современных систем на кристалле с использованием языков описания аппаратуры (HDL) позволяет значительно сократить время проектирования заказных или полузаказных СБИС за счет ускорения процесса создания прототипа. При этом при проектировании прототипа используют микросхемы перепрограммируемой логики, а именно FPGA. Процесс логического синтеза на самом раннем этапе позволяет оценить качество модели на языке описания аппаратуры с точки зрения ее «пригодности» к реализации в аппаратуре. В процессе синтеза используется три типа информации: HDL-модель схемы, множество ограничений, накладываемых на физическую схему, и множество примитивов, которые могут использоваться для построения схемы. Благодаря тому, что одна и та же схема может быть описана множеством способов, перед системой синтеза, которая выводит (синтезирует) конкретную аппаратную реализацию из HDL-описания, возникает многокритериальная задача [1].

За процессом логического вывода следует процесс оптимизации для того, чтобы уменьшить размер полученной схемы или увеличить ее быстродействие.

Учитывая тот факт, что результаты этапа синтеза являются базой для дальнейшей реализации проекта, его эффективность во многом определяет быстродействие и ресурсоемкость проекта. Результатом этапа синтеза является список соединений. Шаблон – это последовательность операторов языка, которая распознается системой синтеза и однозначно отображается в соответствующую аппаратную структуру. То есть, шаблон – это некоторый библиотечный элемент, существующий во взаимосвязи с другими.

2. ПОСТАНОВКА ЗАДАЧИ

Задача поиска по заданным параметрам (поиск шаблона) в списке соединений является основной при удовлетворении требованиям к синтезу, представлении результатов синтеза на заданном уровне абстракции и т. д. в ходе логического синтеза [2, 3]. Недостаток методов, используемых в существующих системах, состоит в том, что поиск происходит в последовательном режиме, что сильно замедляет реализацию алгоритмов поиска. Таким образом, поиск эффективных методов решения данной задачи является актуальным.

Цель работы – сокращение времени анализа результатов логического синтеза HDL-моделей на соответствие правилам синтеза.

Объект исследований – результаты логического синтеза цифрового устройства на основе его HDL-модели.

Предмет исследований – методы предварительного анализа списка соединений после этапа логического синтеза.

3. ТАБЛИЧНАЯ МОДЕЛЬ СПИСКА СОЕДИНЕНИЙ

HDL-код, проходя через этап синтеза, преобразуется в схемную реализацию, графическое представление которой формирует RTL схему (схему на уровне регистровых передач), а внутреннее представление, которое используется для передачи информации о модели, формирует список соединений.

В исследовании предлагается разработать специализированный метод представления информации об иерархии схемных соединений. Основными ключевыми понятиями, на которых базируется данный метод, являются понятия макроса, инстанса и атрибута. Макрос – это библиотечный элемент системы синтеза, т. е., другими словами, шаблон, который соответствует фрагменту кода (исходя из общих правил синтеза и ограничений, накладываемых системой синтеза). Макрос задается:

$$M = F(W, P), \quad (1)$$

где M – макрос; F – функция вычисления макроса; W – тип макроса, $W \in \{\text{flip-flop, mux, tristate и т. п.}\}$; P – типы портов, $P \in \{\text{in, set, reset, clk, и т. д.}\}$.

Инстанс (от англ. *instance* – реализация) – это описание связей определенного макроса в конкретном списке соединений, то есть полный путь до конкретной схемной реализации фрагмента кода внутри всей модели. Инстанс может быть представлен в следующем виде:

$$I = F(M, l), \quad (2)$$

где I – инстанс; F – функция получения инстанса; M – макро, которые задает инстанс; l – линии.

Атрибут – способ представления дополнительной информации на линии схемы, то есть это признак, описывающий особенности этой линии: порт, выход триггера, выход комбинационного элемента. В общем случае атрибут – это некоторая дополнительная информация, которая может быть приписана к типам, полям, методам, свойствам и некоторым другим конструкциям схемы. Данная информация никаким образом не сможет повлиять на работу схемы, программы и т. д. Атрибут вносит ясность при анализе и поиске отдельных элементов. В данной работе атрибут используется как дополнение (пояснение) к свойствам линии схемы. Атрибут представлен как

$$A = F(I, L), \quad (3)$$

где A – атрибут; L – линия связи между элементами; I – инстанс; F – функция вычисления атрибута.

Предлагаемая табличная модель представления иерархии списка соединений содержит пять связанных ключами таблиц: таблица производителей, таблица примитивов, таблица инстансов, таблица атрибутов, таблица линий.

Такая модель позволяет организовать систему ключей, облегчающую поиск по заданному параметру (с точки зрения анализа на соответствие правилам синтеза это поиск инстансов с линиями, имеющими определенный атрибут).

4. ЭТАПЫ АНАЛИЗА СПИСКА СОЕДИНЕНИЙ

Как было отмечено выше, на основании списка соединений формируется ряд таблиц и связи между ними. Затем выполняется анализ табличной модели путем формирования системы запросов. Можно обобщить этапы предлагаемого метода анализа:

1. Информация о схеме представляется в виде таблиц и связей между ними.

2. Выполняется фаза распространения атрибутов, которая приводит к формированию системы таблиц (см. рис. 1). Инстанс и множество входных/выходных линий данного инстанса с атрибутами на них образуют целевой объект и условия его существования.

3. Параметры поиска задаются в терминах: тип целевого объекта и условия, в которых он должен находиться. То есть, сначала выполняется поиск макроса (таким образом, сужается область поиска), а затем среди отобранных макросов выполняется поиск инстанса с определенными атрибутами на линиях.

4. Поиск кандидатов, удовлетворяющих параметрам, происходит путем выбора из заранее сформированных подмножеств определенного типа целевого элемента и анализа условий, в которых он существует.

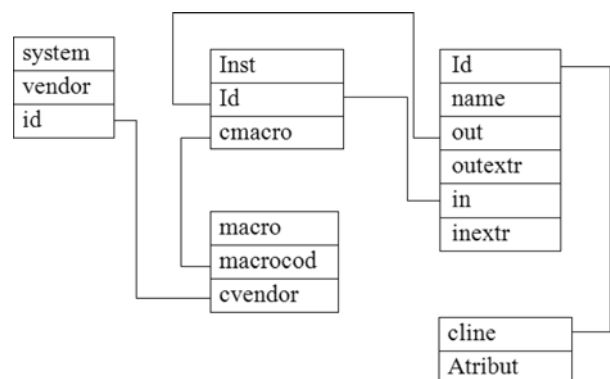


Рис. 1. Структура взаимодействия таблиц

Необходимо подчеркнуть, что фаза распространения атрибутов выполняется один раз. Метод позволяет на схеме, представленной в виде базы данных после фазы распространения атрибутов, находить всех кандидатов для любого числа запросов без повторного распространения атрибутов. Такой подход гарантирует существенное уменьшение суммарного времени на поиск. С помощью данного метода решаются следующие задачи: определение глобальных сигналов синхронизации, асинхронных управляющих сигналов (сброс/установка), определение и анализ доменов синхронизации, проверка правил проектирования, проверка правил тестопригодного проектирования, поиск циклов с обратными связями.

ВЫВОДЫ

Задача поиска по заданным параметрам в некоторой внутренней модели решается на разных этапах проектирования аппаратуры. Особенно заметна роль таких задач на этапе логического синтеза. С помощью поиска по шаблону можно находить некорректно спроектированные цепи сброса или синхронизации, «случайные» триггеры-защелки, использование комбинационной логики для управления синхронизацией триггеров и т. д. При древовидной либо графовидной структуре поиск выполняется последовательно, начиная от корневой вершины [4]. Применение табличного представления позволяет свести поиск к реализации запросов по системе таблиц.

УДК 003.26+004.272.4+004.415.2+004.051

СПИСОК ЛИТЕРАТУРЫ

1. *Yadav M.* Hardware Architecture of a Parallel Pattern Matching Engine / Yadav M., Venkatachaliah A., Franzon P. // Proceedings of the ISCAS 2007. – Pp. 1369–1372.
2. *Tarau P.* Exact Combinational Logic Synthesis and Non-Standard Circuit Design / Tarau P., Luderman B. // Proceedings of the CF'08, May 5–7, 2008, Ischia, Italy. – Pp.15–24.
3. *Tarau P.* Revisiting Exact Combinational Circuit Synthesis / Tarau P., Luderman B. // Proceedings of the SAC'08, Fortaleza, Cear'a, Brazil. – Pp. 1620–1621.
4. *Clifford R.* From coding theory to efficient pattern matching / Clifford R., Efremenko K., E. Porat, Rothschild A. // Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, 2009, New York, USA. – Pp. 778–784.

Надійшла 15.03.2010

Кривуля Г. Ф., Сиревич Є. Ю., Карасьов А. Л.

АНАЛІЗ СПИСКУ З'ЄДНАНЬ У СИСТЕМАХ ЛОГІЧНОГО СИНТЕЗУ

Розглянуто проблему аналізу результатів логічного синтезу для їх подальшого аналізу на відповідність правилам синтезу. Запропоновано табличну модель, що складається із словника атрибутів, отриманого після фази поширення. Визначено етапи пошуку за заданими параметрами при використанні розробленої моделі.

Ключові слова: логічний синтез, HDL, атрибут лінії, макрос, пошук за параметрами.

Krivulya G. F., Syrevich Yev. Yu., Karasyov A. L.

NETLIST ANALYSIS IN LOGICAL SYNTHESIS SYSTEMS

The problem of logical synthesis results analysis for their further analysis for accordance to the rules of synthesis is considered. A tabular model is proposed, consisting of an attributes dictionary obtained after the distribution phase. Stages of search by the preset parameters are determined when using the developed model.

Key words: logical synthesis, HDL, line attribute, macro, pattern matching.

Неласая А. В.¹, Верещак М. И.²

¹Старший преподаватель Запорожского национального технического университета

²Студент Запорожского национального технического университета

ОЦЕНКА ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ БИБЛИОТЕК ДЛИННОЙ АРИФМЕТИКИ В КРИПТОГРАФИЧЕСКИХ ПРИЛОЖЕНИЯХ

В статье проводится анализ эффективности целочисленных операций в современных библиотеках длинной арифметики. Выбрана программно-аппаратная платформа, допускающая выполнение параллельных алгоритмов, для построения эффективных процедур определения параметров криптографических систем.

Ключевые слова: криптографическая система, асимметричная криптография, эллиптическая кривая, порядок группы, библиотека длинной арифметики, анализ производительности, параллельные вычисления, графический процессор.

ВВЕДЕНИЕ

Потребность широкого применения электронных способов передачи и обработки информации выво-

дит на качественно новый уровень проблемы защиты электронной информации. Большинство угроз целостности и конфиденциальности информации, цирку-

лирующей в компьютерных системах, может быть перекрыто на основе использования механизмов безопасности, реализуемых с помощью криптографических методов защиты.

Традиционно криптографические преобразования делятся на две большие ветви – симметричные и асимметричные. Симметричные криптопреобразования являются более быстрыми и используются, собственно, для шифрования файлов. Появление асимметричной криптографии позволило решить задачу аутентификации электронной информации в условиях, когда обменивающиеся информацией стороны не доверяют друг другу. Эту задачу невозможно было бы решить с использованием только лишь симметричных алгоритмов.

Асимметричные криптографические преобразования используются в системах направленного шифрования и цифровой подписи для обеспечения целостности, неопровержимости и аутентичности электронных ресурсов. В частности, задача аутентификации документов решается с помощью использования механизма электронной цифровой подписи, который реализуется методами криптографии с открытым ключом.

Современные стандарты цифровой подписи (в том числе Украинский ДСТУ 4145-2002) и направленного шифрования основаны на использовании операций в группах точек эллиптических кривых. Для обеспечения достаточной стойкости длина секретного ключа должна превышать 160 бит. Размер элементов основного поля, над которым определяется кривая, имеет такой же порядок. Отсюда ясно, что для программной реализации алгоритмов криптографических преобразований на эллиптических кривых не обойтись без длинной арифметики. Естественным обобщением эллиптических кривых являются гиперэллиптические кривые – кривые более высокого рода [1, 2], которые позволяют решать аналогичные задачи.

Стойкость алгоритмов асимметричной криптографии обеспечивается использованием алгебраических структур больших порядков. Наиболее важным параметром, с точки зрения обеспечения стойкости, является порядок группы точек эллиптической кривой и, соответственно, порядок якобиана гиперэллиптической кривой. Построенная криптосистема будет достаточно надежной при условии соответствия этого параметра ряду ограничений [3]. Однако, задача определения порядка в общем случае не является тривиальной и требует значительных затрат временных и технических ресурсов.

Цель данной работы состоит в проведении сравнительного анализа эффективности использования

библиотек длинной арифметики для целочисленных вычислений и выборе программно-аппаратного решения для повышения производительности процедуры определения порядка кривой.

1. КРАТКАЯ ХАРАКТЕРИСТИКА СОВРЕМЕННЫХ БИБЛИОТЕК ДЛИННОЙ АРИФМЕТИКИ

В вычислительной технике под длинной арифметикой понимается выполнение операций над числами, разрядность которых превышает длину машинного слова данной вычислительной машины. Частный случай – арифметика произвольной точности – относится к арифметике, в которой длина чисел ограничена только объемом доступной памяти. Именно арифметика произвольной точности находит широкое применение в расчетах, связанных с криптографией. В языках программирования, как правило, нет встроенных средств для выполнения расчетов с произвольной точностью. Для этого необходимо использовать сторонние библиотеки либо создавать собственные решения. На сегодняшний день существует множество различных готовых библиотек в разных языках программирования для выполнения таких расчетов. При этом возникает проблема выбора инструментария, наиболее подходящего для решения конкретных прикладных задач в каждом случае. В данной статье проводится анализ производительности с точки зрения удобства использования в алгоритмах асимметричной криптографии.

Прежде чем остановить свой выбор на какой-либо из библиотек, либо прийти к выводу о необходимости создания собственного программного продукта, необходимо проанализировать и сравнить существующие решения.

Все множество предлагаемых библиотек можно классифицировать по нескольким параметрам. Первый – лицензия, под которой распространяется программный продукт. Ее необходимо учитывать при публикации своих исследований, проведенных с использованием той или иной библиотеки. Различают коммерческие (IMSL), бесплатные для некоммерческого использования (LiDIA, MIRACL) и бесплатные (GMP, NTL, CLN, MPI, Imath) программные продукты. Следует отметить, что распространяемых под различными разновидностями свободных лицензий программных продуктов сейчас значительно больше, нежели коммерческих. Причиной тому, на наш взгляд, является то, что большая часть таких библиотек рождалась в стенах университетов в приложении к различным научным проектам. Впоследствии эти наработки публиковались, и за их доработку

и совершенствование бралось множество программистов по всему миру.

Вторым важным критерием, который нужно учесть при выборе библиотеки, является перечень поддерживаемых типов и структур данных, от которых напрямую зависит эффективность производимых расчетов. Это:

- целые числа произвольной длины (знаковые и беззнаковые);
- рациональные дроби;
- числа с плавающей точкой произвольной точности;
- комплексные числа;
- векторы;
- матрицы;
- полиномы.

В задачах, с которыми сталкивается современная криптография, наиболее востребованными являются расчеты целочисленных значений произвольной длины (присутствуют во всех библиотеках) и полиномиальная арифметика (реализация которой встречается значительно реже). Поэтому результаты сравнения библиотек, приведенные в данной статье, могут оказаться не вполне подходящими для иных предметных областей. В табл. 1 приведены основные характеристики современных библиотек длинной арифметики [4].

GMP – это свободная библиотека длинной арифметики, поддерживающая работу со знаковыми целыми, рациональными числами и числами с плавающей запятой [5]. Преимущества GMP:

- практически не имеет ограничений точности вычислений, кроме ограниченного объема доступной памяти;
- имеет богатый набор функций и удобный интерфейс;
- поддерживает большинство современных платформ: Unix-подобные операционные системы, такие как GNU/Linux, Solaris, HP-UX, Mac OS X/Darwin, BSD, AIX; Windows. Существуют 32-разрядная и 64-разрядная версии GMP;
- допускает использование наиболее эффективных алгоритмов и оптимизированного под различные современные процессорные системы ассемблерного кода во всех внутренних циклах.

GMP является одной из самых быстрых на сегодняшний день библиотек длинной арифметики. Основные области применения библиотеки – это криптографические системы и исследования, обеспечение безопасности межсетевых взаимодействий, алгебраические пакеты. Библиотека GMP является частью проекта GNU.

CLN – это библиотека для расчетов с использованием всех существующих числовых типов [6]. В этой библиотеке реализованы классы для таких типов данных, как: целые числа, рациональные дроби, числа с плавающей точкой, комплексные числа, унивариантные полиномы, вычисления по модулю. CLN использует GMP в качестве вычислительного ядра и незначительно уступает ей в скорости расчетов. Богатый набор поддерживаемых типов, удобный программный интерфейс, прозрачные механизмы взаимодействия и преобразования друг в друга различных структур данных вместе с достаточно высокой скоростью вычислений обеспечивают широкую распространенность данной библиотеки. CLN лежит в основе многих программных продуктов, связанных с научными исследованиями и математическими расчетами: Scilab, Octave (свободные аналоги MatLAB), maxima (пакет для алгебраических вычислений, свободный аналог MAPLE) и др.

NTL – это высокопроизводительная библиотека C++, предоставляющая структуры данных и алгоритмы для работы с целыми произвольной длины, векторами, матрицами, полиномами, числами с плавающей точкой произвольной точности [7]. Полиномиальная арифметика в NTL – одна из самых быстрых на сегодняшний день. Она не раз использовалась при становлении «мировых рекордов» в скорости алгоритмов факторизации и определении порядка эллиптических кривых. Все алгоритмы NTL реализованы на C++, что обеспечивает высокую мобильность данной библиотеки. Несмотря на наличие собственного вычислительного ядра, NTL позволяет использовать вместо него библиотеку GMP.

MIRACL – единственная из перечисленных библиотек, распространяемая под коммерческой лицензией [8]. Однако авторами разрешается использование MIRACL для некоммерческих нужд. Основные преимущества данной библиотеки:

- наличие реализованных C/C++ интерфейсов «из коробки»;
- богатая библиотека специализированных функций для вычислений в области эллиптической криптографии;
- наличие различных алгоритмов решения одних и тех же задач с возможностью выбора оптимального варианта для текущих нужд.

2. АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ

Анализ производительности исследуемых библиотек проводился с применением простых, аналогичных по структуре программ, выполняющих различные арифметические операции. Всего было проведено

Таблица 1. Основные характеристики библиотек длинной арифметики

Пакет / название библиотеки	Поддерживаемые типы данных	Языки программирования	Лицензия
apfloat	Decimal floats, integers, rationals, and complex	Java and C++	LGPL and Freeware
ARPREC and MPFUN	Integers, binary floats, complex binary floats	C++ with C++ and Fortran bindings	BSD
Base One Number Class	Decimal floats	C++	Proprietary
bbnum library	Integers and floats	Assembler and C++	New BSD
phpseclib	Decimal floats	PHP	LGPL
BigDigits	Naturals	C	Freeware [1]
BigFloat	Binary Floats	C++	GPL
BigNum	Binary Integers, Floats (with math functions)	C# / .NET	Freeware
C++ Big Integer Library	Integers	C++	Public domain
CLN, a Class Library for Numbers	Integers, rationals, and complex	C and C++	GPL
Computable Real Numbers	Reals	Common Lisp	
IMSL		C	Commercial
decNumber	Decimals	C	ICU licence (MIT licence) [2]
FMLIB	Floats	Fortran	
GNU Multi-Precision Library (and MPFR)	Integers, rationals and floats	C and C++ with bindings (GMPY,...)	LGPL
GNU Multi-Precision Library for .NET	Integers	C# / .NET	LGPL
Eiffel Arbitrary Precision Mathematics Library	Integers	Eiffel	LGPL
HugeCalc	Integers	C++ and Assembler	Proprietary
IMath	Integers and rationals	C	Freeware
IntX	Integers	C# / .NET	New BSD
JScience LargeInteger	Integers	Java	
libmpdec	Decimals	C and C++	Simplified BSD
LibTomMath	Integers	C and C++	Public domain
LiDIA	Integers, floats, complex floats and rationals	C and C++	Free for non-commercial use
MAPM	Integers and decimal floats	C (bindings for C++ and Lua)	Freeware
MIRACL	Integers and rationals	C and C++	Free for non-commercial use
MPI	Integers	C	LGPL
MPArith	Integers, floats, and rationals	Pascal / Delphi	zlib
mpmath	Floats, complex floats	Python	New BSD
NTL	Integers, floats	C and C++	GPL
bigInteger (and bigRational)	Integers and rationals	C and Seed7	LGPL
TTMath library	Integers and binary floats	Assembler and C++	New BSD
vecLib. framework	Integers	C	Proprietary
W3b. Sine	Decimal floats	C# / .NET	New BSD

4 теста на компьютере следующей конфигурации:
CPU Intel Celeron M 1,73 GHz, RAM DDR2 2 Gb.

Тест 1. Работа с целыми числами.

Тесты 2–3. Полиномиальная арифметика.

Тест 4. Вычисление точек эллиптической кривой.

Целочисленные вычисления. Этот тест (табл. 2) представляет собой вычисление факториала числа без применения каких-либо оптимизирующих алгоритмов или функций. При этом в GMP использован C-интерфейс, а во всех остальных библиотеках – C++-интерфейс.

Таблица 2. Время вычисления факториала N!

N	GMP, с	NTL, с	CLN, с	MIRACL, с
1000	<0,01	<0,01	<0,01	0,01
5000	0,01	0,01	0,02	0,16
10000	0,05	0,05	0,08	0,65
20000	0,24	0,23	0,34	2,71
50000	1,67	1,66	2,46	! Overflow!
100000	7,22	7,19	12,06	! Overflow!

Полиномиальные расчеты. Второй тест (табл. 3) представляет собой последовательное умножение полинома первой степени на самого себя. Результирующее значение на каждой итерации – полином, степень которого линейно зависит от номера текущей итерации. Этот эксперимент по своей сути соответствует основной операции метода [9] определения порядка кривой.

Таблица 3. Время последовательного умножения полинома первой степени на самого себя

Количество итераций	NTL, с		CLN, с	
	Длина коэффициентов полинома			
	16 бит	32 бит	16 бит	32 бит
50	0,03	0,04	0,04	0,05
100	0,15	0,15	0,17	0,18
200	0,6	0,59	0,67	0,73
500	3,78	3,78	4,33	4,66
1000	15,84	15,16	17,24	18,29

Третий тест (табл. 4) представляет собой последовательное возведение полинома в квадрат. Длина ко-

Таблица 4. Время возведения полинома в квадрат

Количество итераций	Степень полинома	NTL, с	CLN, с
4	8	<0,01	<0,01
6	32	<0,01	0,01
8	128	0,01	0,2
12	2048	0,17	50,16
16	32768	3,86	inf
18	131072	17,08	inf
20	524288	! overflow!	inf

эффициентов – 32 бит. При этом вычислительная сложность и объем используемой программой памяти растет очень быстро. Такая операция чаще всего встречается в полиномиальной арифметике, применяемой в криптографии.

Скалярное умножение на эллиптической кривой. Последний тест (табл. 5) был проведен для оценки производительности встроенных функций библиотеки MIRACL для операций на эллиптических кривых. По результатам предыдущих тестов ясно, что CLN сильно отстает в скорости вычислений по сравнению с NTL, поэтому в таблице приведены данные сравнительного анализа производительности только NTL и MIRACL. В данном тесте определялось время умножения точки кривой, определенной над простым полем с размером модуля 192 бита, на длинное число размером 256 бит. Эта операция является основой для процедур вычисления открытого ключа, формирования и проверки цифровых подписей. Для чистоты эксперимента за один запуск производится 1000 операций умножения. Тест был повторен 3 раза.

Таблица 5. Скалярное умножение на эллиптической кривой

Номер запуска	NTL		MIRACL	
	Общее время, мс	Одна операция, мс	Общее время, мс	Одна операция, мс
1	6820	6,82	2845	2,84
2	5678	5,68	3048	3,05
3	7532	7,53	3209	3,21

Этот эксперимент демонстрирует, что, не смотря на то, что MIRACL проигрывает в производительности целочисленных вычислений библиотеке NTL, производительность специализированных функций превосходит аналогичные по реализации в NTL более чем в 2 раза.

По результатам анализа можно сделать некоторые выводы относительно исследованных библиотек и дать им краткие характеристики:

– Как было сказано выше, библиотека GMP является частью проекта GNU, что в значительной мере определяют идеологию ее разработки: каждый отдельно взятый продукт данного проекта выполняет определенную возложенную на него задачу и только ее. В случае с GMP этот принцип выливается в ограниченный набор поддерживаемых типов данных – в GMP реализованы лишь базовые операции над скалярными величинами. Упор в этой библиотеке делается на оптимизацию вычислительного ядра под каждую из популярных моделей процессоров и своевременные обновления алгоритмов с появлением новых платформ, и

с этой задачей разработчики справляются в полной мере. Таким образом, у GMP нет определенной направленности, что делает возможным и желательным использование ее в качестве базы для разработки более узконаправленных продуктов.

– Библиотека CLN, пожалуй, имеет наиболее широкий спектр применений в области математических исследований. Однако универсальность этой библиотеки имеет и свою обратную сторону – как показывают тесты, ее скорость в сравнении с NTL и GMP падает в 1,5–2 раза. Эта библиотека прекрасно подходит для программных продуктов, где существует необходимость в поддержке значительного количества различных типов данных и скорость вычислений не является критичной, например математических пакетов общего назначения (наподобие MAPLE, MathCAD и пр.).

– Библиотека NTL имеет более узкую направленность. Здесь основной упор сделан на целочисленную арифметику и полиномиальные расчеты. Применяемые в ней алгоритмы полиномиальных вычислений на сегодняшний день обладают рекордными показателями скорости. Следует также заметить, что в полной мере потенциал NTL раскрывается в сочетании с GMP. Таким образом, NTL доказывает целесообразность использования GMP в качестве основы для создания более узконаправленных библиотек.

– Библиотека MIRACL может наиболее эффективно использоваться для решения специализированных задач области эллиптической криптографии.

3. ИСПОЛЬЗОВАНИЕ СОВРЕМЕННЫХ АППАРАТНЫХ РЕШЕНИЙ

Традиционно для решения объемных вычислительных задач, к которым относится и задача определения порядка кривой, используются параллельные методы вычислений, требующие наличия специализированной высокопроизводительной техники. Параллельные вычисления в крупных масштабах (десятки и сотни вычислительных узлов) требуют очень серьезных капиталовложений на приобретение и обеспечение аппаратной базы. Первый вариант – это кластеры, состоящие из нескольких десятков вычислительных узлов, управляемых одной операционной системой. При этом каждый вычислительный узел способен функционировать автономно, как самостоятельная вычислительная система. Подобные системы очень громоздки, имеют большой расход электроэнергии и требуют применения сложных и высоконадежных систем охлаждения наряду с системой обеспечения безопасности. Другой вариант – это технология облачных вычислений (cloud computing) и

GRID-системы. Их преимуществом перед кластерами является то, что подобные системы достаточно легко могут быть развернуты практически на любой имеющейся в наличии аппаратной базе. При этом они легко масштабируемы и практически не имеют ограничений по наращиванию мощностей. Недостатком облачных вычислений является то, что в качестве вычислительных узлов выступают отдельные рабочие станции, объединенные локальной сетью, вследствие чего они опять-таки требовательны к ресурсам. К тому же передача данных по локальной сети происходит сравнительно медленно, поэтому для получения высокой эффективности в такой системе каждую задачу необходимо разбивать на крупные блоки, каждый из которых будет выполняться на своем узле, что значительно затрудняет процесс разработки и подходит не для каждой задачи.

Альтернативное решение пришло с довольно неожиданной стороны. По мере развития игровой компьютерной индустрии в течение многих лет велись разработки по наращиванию мощности графических адаптеров, так как на них возлагалась задача быстрого пересчета в режиме реального времени массивов данных, связанных с моделированием трехмерного виртуального пространства. Для решения этой задачи в графические адаптеры было внедрено множество отдельных процессоров с усеченным по сравнению с CPU набором команд, которые способны, работая параллельно, выполнять различные арифметические операции. Когда количество и мощность этих процессоров достигли определенного уровня, оказалось, что они могут быть эффективно использованы не только для расчетов координат трехмерного пространства, но и для любых других математических расчетов. Так родилась идея использования возможностей графических адаптеров для организации параллельных вычислений, превратившаяся впоследствии в отдельный проект.

На сегодняшний день компанией NVIDIA разработан и запущен проект под названием CUDA [10, 11]. Это программно-аппаратная архитектура, позволяющая производить вычисления с использованием графических процессоров NVIDIA, поддерживающих технологию GPGPU. CUDA SDK позволяет программистам реализовывать на специальном упрощенном диалекте языка программирования Си алгоритмы, выполнимые на графических процессорах NVIDIA, и включать специальные функции в текст программы на Си. CUDA дает разработчику возможность по своему усмотрению организовывать доступ к набору инструкций графического ускорителя и управлять его памятью, организовывать на нем сложные параллель-

ные вычисления. Все современные графические адаптеры NVIDIA поддерживают технологию CUDA. Кроме этого, NVIDIA выпустила семейство вычислительных систем Tesla на основе графических процессоров с архитектурой CUDA, которые могут быть использованы для научных и технических вычислений общего назначения. Технические решения на базе данной технологии значительно дешевле, чем аналогичные им по мощности кластерные системы, и обладают значительно меньшими показателями энергопотребления и тепловыделения в процессе работы, за счет чего снижаются затраты на их обеспечение.

Технология CUDA является достаточно перспективной для решения поставленных в данной работе задач. На данный момент на CUDA разработана библиотека вычислений произвольной точности с действительными числами для применения в задачах моделирования сложных процессов, описываемых дифференциальными уравнениями [12]. Однако в криптографических приложениях основная масса вычислений выполняется над длинными целыми числами. Программного инструментария, реализующего подобную функциональность для GPU, пока не существует. Исходя из этого, для обеспечения возможности эффективной реализации алгоритмов определения порядков кривых на графических ускорителях первоочередной задачей является разработка библиотеки длинной целочисленной (в том числе модульной) арифметики на CUDA. В этом направлении авторами планируется проводить дальнейшие исследования.

ЗАКЛЮЧЕНИЕ

Проведенный анализ современных программных и аппаратных средств позволил определить особенности использования и эффективность для конкретных операций распространенных библиотек длинной арифметики GMP, CLN, NTL и MIRACL. Эту информацию могут использовать разработчики программного обеспечения при выборе библиотеки для своих нужд.

Анализ аппаратных решений позволил выбрать достаточно мощную и недорогую технологию параллельных вычислений, которую планируется использовать для повышения производительности процедур определения параметров криптосистем на эллиптических и гиперэллиптических кривых.

СПИСОК ЛИТЕРАТУРЫ

1. *Koblitz N.* Hyperelliptic cryptosystem / N. Koblitz // *Journal of Crypto.* – 1989. – № 1. – P. 139–150.
2. *Menezes A.* An Elementary Introduction to Hyperelliptic Curves [Электронный ресурс] : published as Technical Report CORR 96-19 Department of C&O University of Waterloo/ Menezes A., Wu Y., Zuccherato R. – Электрон.

дан. – Ontario, Canada, 1996. – P. 1–35. – Режим доступа: www.cacr.math.uwaterloo.ca/techreports/1997/corr96-19.ps, свободный. – Загл. с экрана.

3. *Неласа Г. В.* Удосконалення методів перетворень в якобіанах гіпереліптичних кривих для криптографічних додатків : автореф. дис. на здобуття наук. ступеня канд. техн. наук : спец. 05.13.21 «Системи захисту інформації» / Г. В. Неласа. – Харків, 2010. – 22 с.
4. Arbitrary-precision arithmetic [Электронный ресурс]. – Электрон. дан. – Режим доступа: http://en.wikipedia.org/wiki/Arbitrary-precision_arithmetic, свободный. – Название с титул. экрана.
5. GMP «Arithmetic without limitation» The GNU Multiple Precision Arithmetic Library [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://gmplib.org/>, свободный. – Название с титул. экрана.
6. CLN-Class Library for Numbers [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://www.ginac.de/CLN/>, свободный. – Название с титул. экрана.
7. NTL: A Library for doing Number Theory [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://www.shoup.net/ntl/>, свободный. – Название с титул. экрана.
8. Multiprecision Integer and Rational Arithmetic C/C++ Library [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://www.shamus.ie/>, свободный. – Название с титул. экрана.
9. *Bostan A.* Linear recurrences with polynomial coefficients and application to integer factorization and Cartier-Manin operator / Bostan A., Gaudry P., Schost É. // *Proceedings of Fq7, Lecture Notes in Comput. Sci.* Vol. 2948. – Berlin : Springer-Verlag. 2004. – P. 40–58.
10. CUDA Toolkit 3.0 (March 2010) : Download Quick Links [Электронный ресурс] / NVIDIA. – Электрон. дан. – Режим доступа: http://developer.nvidia.com/object/cuda_3_0_downloads.html, свободный. – Загл. с экрана.
11. *Борсеков А. В.* Основы работы с технологией CUDA / А. В. Борсеков, А. А. Харламов. – М. : ДМК Пресс, 2010. – 232 с.
12. Научно-образовательный центр «Параллельные вычисления» [Электронный ресурс]. – Электрон. дан. – Режим доступа: www.parallel-compute.com, свободный. – Название с титул. экрана.

Надійшла 15.02.2010

Неласа Г. В., Верещак М. І.

ОЦІНКА ЕФЕКТИВНОСТІ ВИКОРИСТАННЯ БІБЛІОТЕК ДОВГОЇ АРИФМЕТИКИ В КРИПТОГРАФІЧНИХ ДОДАТКАХ

В статті проводиться аналіз ефективності операцій з цілими числами в сучасних бібліотеках довгої арифметики. Вибрано програмно-апаратну платформу, що допускає виконання паралельних алгоритмів, для побудови ефективних процедур визначення параметрів криптографічних систем.

Ключові слова: криптографічна система, асиметрична криптографія, еліптична крива, порядок групи, бібліотека довгої арифметики, аналіз ефективності, паралельні обчислення, графічний процесор.

Nelasa G. V., Vereschak M. I.

EFFICIENCY EVALUATION OF LONG NUMBER LIBRARIES IN CRYPTOGRAPHY APPLICATIONS

Research of integer arithmetic efficiency in modern long number libraries is carried out. The soft hardware with parallel commands for construction of efficient procedures of cryptosystem parameters determination is chosen.

Key words: cryptosystem, open key cryptography, elliptic curve, order of group, long number library, efficiency evaluation, parallel computing, graphic processor unit.

ИНСТРУМЕНТАЛЬНАЯ СИСТЕМА РЕШЕНИЯ ДВУМЕРНЫХ И ТРЕХМЕРНЫХ ЗАДАЧ ТЕОРИИ УПРУГОСТИ НА БАЗЕ MAPLE

В статье приведено описание инструментальной системы решения двумерных и трехмерных задач теории упругости. Вычислительное ядро системы составляет Maple. Произведена формализация основных алгоритмов с целью нахождения операций программирования указанных задач. В качестве примеров работы с системой приведены решения конкретных задач.

Ключевые слова: препроцессор, процессор, постпроцессор, метод начальных функций.

ВВЕДЕНИЕ

В настоящее время не теряют своей актуальности аналитические методы, которые позволяют получать точные и обобщенные результаты. Благодаря развитию систем компьютерной математики (СКМ) стала возможной реализация операций аналитического дифференцирования и интегрирования. Это открывает большие возможности в применении СКМ для математического моделирования. Однако системы не обладают достаточно развитыми способами решения технических задач. Большинство из них ориентировано на применение в алгебре и математическом анализе [1]. Поэтому в процессе математического моделирования их необходимо дорабатывать, чтобы иметь возможность применить для решения той или иной сложной технической задачи. Создав препроцессор в СКМ и воспользовавшись вычислительными функциями СКМ, можно построить средство для решения конкретных задач.

В предлагаемой работе дана характеристика одного из средств для решения статических задач линейной теории упругости – инструментальной системы, основанной на базе ядра СКМ Maple. Система Maple была выбрана авторами как система, наиболее подходящая по входному языку. С помощью программирования в ее среде можно создать эффективное вычислительное средство для решения задач теории упругости. В разработанной инструментальной системе решаются задачи теории плит, пластин, полуплоскости и полупространства. Производятся сложные аналитические операции над формульным заданием выражений теории упругости. Вопросами разработки инструментальных систем для решения задач теории упругости с внедрением аналитических методов занимались Кабулов В. К., Толлок В. О. [2]; построением инструментальной системы на ос-

нове метода конечных элементов занимался Гоменюк С. И. [3]. Их работы способствовали разработке новой системы на основе Maple.

ОПИСАНИЕ СИСТЕМЫ

Входной язык – расширенный язык Maple. Алфавит языка Maple включает 26 букв латинского алфавита, строчных (a–z) и прописных (A–Z), десять цифр (0–9) и 32 специальных символа. Подробное описание языка приведено в работе [4]. Расширения языка процедуры и функции: `sdiff` – символическое дифференцирование; `sint` – символическое интегрирование; `operators` – процедура вывода операторов решения плиты, пластины; `mch` – процедура разложения гиперболического косинуса в ряд по Маклорену; `msh` – разложение гиперболического синуса; `p` – перевод символических рядов в числовые для двумерной задачи; `nm` – перевод в числовые ряды для трехмерного случая; `integral` – процедура интегрирования по параметру; `galerkin2D` – процедура Галеркина для двумерной задачи; `galerkin3D` – трехмерной; `maklogen` – превращение тригонометрических функций в разложения Маклорена; `symbolism` – проведение символических операций дифференцирования и интегрирования; `semi` – процедура решения для полуплоскости и полупространства; `seminitial` – процедура начальной инициализации, задание фундаментального решения.

Предлагаемая система состоит из трех основных блоков: препроцессора, процессора и постпроцессора. Препроцессор – набор процедур, предназначенный для автоматизации подготовки исходных данных. В системе он занимается выводом решения в операторно-символической форме и заданием начальных функций и граничных условий задачи.

Препроцессор отвечает за:

- а) выбор размерности задачи (двумерная или трехмерная);
- б) выбор системы координат (декартова, полярная, цилиндрическая);
- в) реализация символьных операций дифференцирования и интегрирования;
- г) вывод решения в символьной форме;
- д) вывод решения в виде операторов [5].

Препроцессор занимается обработкой вводимых пользователем данных. Перед решением задачи на языке инструментальной системы необходимо описать входные и выходные параметры, далее с помощью подключения библиотеки осуществить процесс решения поставленной задачи.

Процессор – важная часть работы системы. Процесс его работы можно условно разбить на следующие этапы:

- а) перевод решения из символьной формы в интегрально-дифференциальную;
- б) проведение аналитического дифференцирования, интегрирования [6];
- б) вывод аналитического решения задачи.

В роли постпроцессора выступают функции и процедуры, осуществляющие числовой счет по выведенной аналитической формуле. Далее графические процедуры системы визуализируют результат.

Автоматизация аналитического вывода решения и последующий его перевод с символьного представления в формульный вид представляет собой отдельную сложную задачу, которая была осуществлена путем создания библиотеки подпрограмм и функций для Maple. В библиотеке содержатся команды препроцессора и процессора.

Строение препроцессора

Алгоритм 1. В препроцессор системы вошел алгоритм построения аналитического решения для балок, плит и пластин в двумерной и трехмерной постановке. Суть его в следующем: решается система из четырех уравнений теории упругости для двумерной постановки и из шести для трехмерной [7], [8].

Используется решение бигармонического уравнения $\nabla\nabla U = 0$ в форме операторно-символьных Φ -функций. Процесс вывода детально описывается в работе [5]:

$$f(x, y, z) = f_0(x, y)\cos(\gamma z) + f_1(x, y)\sin(\gamma z) + f_2(x, y)z\cos(\gamma z) + f_3(x, y)z\sin(\gamma z), \quad (1)$$

где $\gamma = \sqrt{\frac{d^2}{dx^2} + \frac{d^2}{dy^2}} = \sqrt{\alpha^2 + \beta^2}$ – сумма производных $\alpha = \frac{d}{dx}$, $\beta = \frac{d}{dy}$; \sin , \cos – символические разложения Маклорена.

В случае двумерной задачи Φ -функция записывается в виде

$$f(x, y) = f_0(x)\cos(\alpha y) + f_1(x)\sin(\alpha y) + f_2(x)y\cos(\alpha x) + f_3(x)y\sin(\alpha y),$$

где $\alpha = \frac{d}{dx}$ – символическое обозначение производной.

Далее программируется второй метод из работы [5]. Однако, чтобы проводить аналитические операции над символьными разложениями Маклорена \sin и \cos , как над обычными функциями, требовалось доказать, что эти разложения обладают свойствами тригонометрических функций. Именно этому вопросу и была посвящена работа [9]. Благодаря доказательству закона ортогональности для общей трехмерной задачи все символьные преобразования над разложениями Маклорена \sin и \cos в системе считаются допустимыми и дают правильные результаты. По методу Φ -функций решаются уравнения теории упругости [5], [7], [8].

Этот процесс реализован в процедуре operators.

```
> S1:= eval(value(S), z = 0);
S2:= eval(diff(value(S), z), z = 0);
> S3:= eval(diff(value(S), `$(z, 2)), z = 0);
[S1[], S2[], S3[]];
solve(% , [u1, u2, u3, v1, v2, v3, w1, w2, w3, xx1, xx2, xx3,
yy1, yy2, yy3, zz1, zz2, zz3]); assign(%); factor(%);
> R:= collect(simplify(algsubs(alpha^2+beta^2 = gamma^2,
%)), gamma, factor)[];
subs(R, U); simplify(%); collect(% , u0); collect(% , v0);
collect(% , w0); collect(% , zz0); collect(% , yy0);
U:= sort(collect(% , xx0), [u0, v0, w0, zz0, yy0, xx0]);
subs(R, V); simplify(%); collect(% , u0); collect(% , v0);
collect(% , w0); collect(% , zz0); collect(% , yy0);
V:= sort(collect(% , xx0), [u0, v0, w0, zz0, yy0, xx0]);
subs(R, W); simplify(%); collect(% , u0); collect(% , v0);
collect(% , w0); collect(% , zz0); collect(% , yy0);
W:= sort(collect(% , xx0), [u0, v0, w0, zz0, yy0, xx0]);
subs(R, Z); simplify(%); collect(% , u0); collect(% , v0);
collect(% , w0); collect(% , zz0); collect(% , yy0);
Z:= sort(collect(% , xx0), [u0, v0, w0, zz0, yy0, xx0]);
subs(R, Y); simplify(%); collect(% , u0); collect(% , v0);
collect(% , w0); collect(% , zz0); collect(% , yy0);
Y:= sort(collect(% , xx0), [u0, v0, w0, zz0, yy0, xx0]);
subs(R, X); simplify(%); collect(% , u0); collect(% , v0);
collect(% , w0); collect(% , zz0); collect(% , yy0);
X:= sort(collect(% , xx0), [u0, v0, w0, zz0, yy0, xx0]);
```

В процедуре используются следующие операторы и функции языка Maple. Функция eval(<expr>, <пере-

менная = значение>) – вычисление выражения при заданном значении входящей в него переменной; <expr> – выражение, которое необходимо вычислить, <переменная> – имя входящей в <expr> переменной. Функция value(<expr>) – вычисление выражения, в которое входят инертные переменные. Оператор diff(<expr>, <переменная> \$ <порядок производной>) – взятие производной в аналитической форме. Функция solve(<уравнение|[система уравнений]>{система уравнений}>, <решение|[решения]>{решения}>) – аналитическое решение уравнения или системы уравнений. При этом система уравнений может задаваться в виде списка перечисляемых уравнений, обладающего упорядоченностью, или в виде множества уравнений. Порядок следования уравнений во множестве не важен. Список уравнений заключается в прямые скобки [], множество – в фигурные {}. После вывода решений переменным необходимо присвоить их значения, для этого используется функция assign(<expr>). Присвоенные значения переменных преобразовываются с помощью функции factor(%) – разложение полинома на множители. Функция algsubs(<переменная1 = значение1>, <переменная2 = значение2>, ..., <переменнаяN = значениеN>, <expr>) – алгебраическая замена переменных их значениями и алгебраическое преобразование выражения <expr>. Функция subs(<переменная1 = значение1>, <переменная2 = значение2>, ..., <переменнаяN = значениеN>, <expr>) – замена переменных их значениями без алгебраического преобразование выражения. Функция simplify(<expr>|<методы упрощения>|<assume = свойство>) – упрощение математического выражения <expr> по заданным методам <методы упрощения>. В качестве методов могут быть выбраны опции exp, ln, sqrt, trig, radical, power, свойство assume – указание диапазона значений, принимаемых переменными в упрощении, может принимать значения complex – комплексные значения, real – действительные числа, positive – положительные числа, integer – целые числа, RealRange(a, b) – интервал действительных чисел (a, b). Функция collect(<expr>, <набор переменных>|<методы преобразования>) – функция приведения подобных членов выражения <expr> по выбираемым именам переменных <набор переменных>; <методы преобразования> – перечисление методов преобразований выражений за скобками приведенных членов. Может принимать значения combine – представление в форме произведения скобок, factor – разложение на множители, normal – сокращение дробей, simplify – полное упрощение, recursive – приведение подобных членов при степенях первой переменной в списке, далее по второй и т. д., distributed – при-

ведение коэффициентов при членах, содержащих всевозможные произведения степеней переменных в списке или множестве, причем суммарная степень всех переменных возрастает от наименьшей к наибольшей. По умолчанию в collect работает метод recursive. Функция sort(<expr>) – сортировка полинома по степеням от наибольшей к наименьшей.

Получается общее решение, из которого выделяются операторы [9]. Этот процесс осуществляет процедура operat.

```
> operat:=proc(res, per, vec)
local oper, dn, fr;
if has(oper, sin) then
oper:=simplify(eval(subs(vec, res)/per,
gamma=sqrt(alpha^2+beta^2)));
oper:=collect(simplify(subs(alpha^2+beta^2=gamma^2,
oper)),[gamma, cos, sin, nu]);
dn:=simplify(numer(oper)); fr:=denom(oper);
dn:=collect(dn,[gamma, cos, sin, nu, z]);
fr:=collect(fr, gamma);
dn:=subs(alpha^2+beta^2=gamma^2, dn);
dn/fr;
else
oper:=simplify(eval(subs(vec, res)/per,
gamma=sqrt(alpha^2+beta^2)), power, radical, sqrt);
oper:=combine(simplify(subs(alpha^2+beta^2=gamma^2,
oper), power, radical, sqrt));
dn:=numer(oper); fr:=denom(oper);
dn:=collect(dn,[per, nu, gamma, z, exp]);
fr:=collect(fr, gamma);
dn:=subs(alpha^2+beta^2=gamma^2, dn);
dn/fr; fi;
end proc;
```

В указанных выше процедурах добавлены процедуры numer(<expr>) – выделение числителя выражения <expr> и denom(<expr>) – знаменателя, соответственно, и оператор has(<expr>, <подвыражение>), который возвращает логическое значение true, если <подвыражение> содержится в <expr>.

Алгоритм 2. В препроцессор был введен алгоритм, позволяющий получать решения задач теории упругости для полуплоскости. Основная идея алгоритма принадлежит Толоку В. О., аспирантом Овским А. Г. он был реализован в Maple.

Используются функции библиотеки Vlasov. m. Основная идея алгоритма следующая.

Решение бигармонического уравнения $\nabla\nabla W = 0$ строится в виде $w(x, y) = w_0(x, y) + u\psi(x, y)$, где $w_0(x, y)$ и $\psi(x, y)$ – гармонические функции-решения уравнения $\Delta_2 w = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}$. Общее решение краевой задачи в случае полуплоскости для уравнения Лапласа $\Delta_2 w = 0$ имеет вид $w_0(x, y) =$

$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y w_0(\xi) d\xi}{(x-\xi)^2 + y^2}$. Для начальной прямой решение

бигармонического уравнения будет иметь вид

$$\frac{\partial w}{\partial y} \Big|_{y=0} = \frac{\partial w_0}{\partial y} + \psi + y \frac{\partial \psi}{\partial y} \Big|_{y=0};$$

$$w_1(x, y) = \frac{\partial w_0}{\partial y} + \psi, \quad (2)$$

где $w_1 = w_1(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y w_1(\xi) d\xi}{(x-\xi)^2 + y^2}$, $w_1(\xi)$ – находится из уравнений теории упругости в процессе решения задач [10]. Выразим гармоническую составляющую ψ . Решив уравнение, получим:

$$w(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2y^3 w_0(\xi) d\xi}{((x-\xi)^2 + y^2)^2} + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y w_1(\xi) d\xi}{(x-\xi)^2 + y^2}. \quad (3)$$

Запишем выражение (3) для $\sigma_y(x, y)$, имеем

$$\sigma_y(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2y^3 w_y^0(\xi)}{((x-\xi)^2 + y^2)^2} d\xi + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y^2 w_y^1(\xi)}{((x-\xi)^2 + y^2)} d\xi, \quad (4)$$

где $\sigma_y^1 = \frac{\partial \sigma_y^0}{\partial y} = \frac{\partial \tau_0}{\partial x}$ – выражение для производной напряжения σ_y на прямой $y = 0$, согласно Власову [7].

Произведя все подстановки и учитывая свойство подынтегральных выражений, получим:

$$\sigma_y(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2y^3}{r^4} \sigma_y^0(\xi) d\xi + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2(x-\xi)y^2}{r^4} \tau_{xy}^0(\xi) d\xi. \quad (5)$$

Здесь $r^4 = ((x-\xi)^2 + y^2)^2$ – сокращенное обозначение. Далее подставляя в уравнения теории упругости для двумерного случая [8] и произведя интегрирования, получим выражения для перемещений и напряжений в случае плоской деформации:

$$\tau_{xy}(x, y) = \int_0^{x-\xi} \frac{\partial \sigma_y}{\partial y} dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2y^2(x-\xi)}{r^4} \sigma_y^0(\xi) d\xi + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2y(x-\xi)^2}{r^4} \tau_{xy}^0(\xi) d\xi;$$

$$\sigma_x = - \int_0^{x-\xi} \frac{\partial \sigma_{xy}}{\partial y} dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2y(x-\xi)^2}{r^4} \sigma_y^0(\xi) d\xi + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2(x-\xi)^3}{r^4} \tau_{xy}^0(\xi) d\xi;$$

$$u = \int_0^{x-\xi} \frac{1}{2G} [(1-\nu)\sigma_x - \nu\sigma_y] dx = \frac{1}{2G\pi} \int_{-\infty}^{\infty} \left[(1-2\nu) \arctg \frac{x-\xi}{y} - \frac{(x-\xi)y}{r^2} \right] \sigma_y^0(\xi) d\xi + \frac{1}{2G\pi} \int_{-\infty}^{\infty} \left[(1-\nu)(\ln r^2 + C) - \frac{y^2}{r^2} \right] \tau_{xy}^0(\xi) d\xi;$$

$$v = \int_0^{x-\xi} \left(\frac{\tau_{xy}}{G} - \frac{\partial u}{\partial y} \right) dx = \frac{1}{2G\pi} \int_{-\infty}^{\infty} \left[(1-\nu)(\ln r^2 + C) - \frac{y^2}{r^2} \right] \sigma_y^0(\xi) d\xi - \frac{1}{2G\pi} \int_{-\infty}^{\infty} \left[(1-2\nu) \arctg \left(\frac{x-\xi}{y} \right) - \frac{(x-\xi)y}{r^2} \right] \tau_{xy}^0(\xi) d\xi. \quad (6)$$

Преобразования (5), (6) в разработанной системе реализуются с помощью следующих алгебраических операций:

$$\sigma_y := \phi \sigma_0 + \psi \tau_0;$$

$$\tau := - \frac{\beta \phi \sigma_0}{\alpha} - \frac{\beta \psi \tau_0}{\alpha};$$

$$\sigma_x := \frac{\beta^2 \phi \sigma_0}{\alpha^2} + \frac{\beta^2 \psi \tau_0}{\alpha^2};$$

$$\sigma_y := - \alpha \phi u_0 - \alpha \psi u_1 - \frac{\alpha^2 \phi v_0}{\beta} - \frac{\alpha^2 \psi v_1}{\beta};$$

$$u := \frac{1}{2} \frac{\beta^2 \phi \sigma_0}{\alpha^3 G} + \frac{1}{2} \frac{\beta^2 \psi \tau_0}{\alpha^3 G} - \frac{1}{2} \frac{\beta^2 \nu \phi \sigma_0}{\alpha^3 G} - \frac{1}{2} \frac{\beta^2 \nu \psi \tau_0}{\alpha^3 G} - \frac{1}{2} \frac{\nu \psi \tau_0}{\alpha G} - \frac{1}{2} \frac{\nu \phi \sigma_0}{\alpha G};$$

$$v := \frac{1}{2} \frac{\beta \nu \phi \sigma_0}{\alpha^2 G} + \frac{1}{2} \frac{\beta \nu \psi \tau_0}{\alpha^2 G} - \frac{1}{2} \frac{\nu \psi \tau_0}{\beta G} - \frac{1}{2} \frac{\nu \psi \tau_0}{\beta G} - \frac{1}{2} \frac{\nu \phi \sigma_0}{\beta G} + \frac{1}{2} \frac{\phi \sigma_0}{\beta G}.$$

Здесь $\alpha = \frac{d}{dx}$, $\beta = \frac{d}{dy}$, $\frac{1}{\alpha} f = \int_0^{x-\xi} f dx$, $\frac{1}{\beta} f = \int_0^y f dy$ – замена операций интегрирования и дифференцирования на умножение и деление; $\phi = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2y^3}{((x-\xi)^2 + y^2)^2} d\xi$, $\psi = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y^2}{((x-\xi)^2 + y^2)} d\xi$ – сокращенные обозначения; G – модуль сдвига.

Преобразования в препроцессоре осуществляются в строгом порядке без нарушения преобразований.

Все операции дифференцирования и интегрирования сведены к умножению и делению.

Строение процессора

После предварительной подготовки и указания граничных условий начинает работать процессор. Он переводит операторно-символическую форму решения в формульную. Осуществляется для каждого из полученных операторов с помощью процедуры `nm` для трехмерного и `n` для двумерного случая. Учитывается свойство тригонометрических функций в рядах (2) и (3) менять знак при дифференцировании. После взятия всех производных разложения Маклорена в символическом решении результат преобразуется в числовой [8].

Текст на языке Maple процедур `nm` и `n`:

```
> vlasov[nm]:=proc(res, per)
local i:: integer, rs, op1, op2, op3, po:: anything, str:: string;
global z, alpha, beta, gamma;
op1:=op(per);
op2:=op(op1[1]);
op3:=op(op2[1]);
rs:= nops(op2[1]); po:=1;
for i from 1 to rs do
po:=po*op3[i]; od;
rs:=res;
if po<>0 then
rs:=subs(sin=sinh, cos=cosh, rs);
rs:=subs(alpha^3=diff(po, x$3)/po, beta^3=diff(po, y$3)/
po, rs);
rs:=subs(alpha^2=diff(po, x$2)/po, beta^2=diff(po, y$2)/
po, rs);
rs:=subs(alpha=diff(po, x)/po, beta=diff(po, y)/po, rs); fi;
rs*po;
end proc;
> vlasov[n]:=proc(res, per)
local i:: integer, rs, po:: anything, str:: string;
global y, alpha, gamma;
po:=op(per)[1];
rs:=res;
if po<>0 then
rs:=subs(sin(alpha*y)=sinh(gamma*y),
cos(alpha*y)=cosh(gamma*y), rs);
rs:=expand(rs);
rs:=subs(alpha^3=diff(po, x$3)/po, rs);
rs:=subs(alpha^2=diff(po, x$2)/po, rs);
rs:=subs(alpha=diff(po, x)/po, rs); fi;
rs*po;
end proc;
```

После этого применяется процедура Галеркина и получается результат в числовой форме:

```
> vlasov[galerkin2D]:= proc (inp, r:: anything)
local st; global gamma, y, S;
integral(inp*r, x = 0.. S)
end proc;
> vlasov[galerkin3D]:= proc (inp, r1, r2:: anything)
local str, t;
```

```
global gamma, z, Q, S, R;
str:= inp;
t:= remove(has, str, r1);
str:= select(has, str, r2);
str:= simplify(select(has, str, r1)/(r1*r2));
(1/4)*str*t*S*R end proc;
```

В процедурах используются функции работы со структурой выражений `op(<expr>)` – подъем по бинарному дереву структуры выражения вверх, в результате получается последовательность выражений. Функция `nops(<expr>)` – подсчет количества выражений в последовательности выражений `op`. Функция `expand(<expr>)` раскрывает скобки в выражении `<expr>`. Используются также процедуры выделения подвыражения из заданного выражения `select(has, <expr>, <выражение>)`, где `has` указывает на выделение подвыражения, где встречается выражение имени, или оператора. Функция `remove(has, <expr>, <выражение>)`, наоборот, удаляет выделяемое подвыражение.

В библиотеку вошли созданные автором процедура интегрирования и процедура взятия модуля числа. Стандартные процедуры системы Maple не рассчитаны на обработку сложных аналитических выражений. Этот недостаток Maple вызывал зависание вычислительного ядра инструментальной системы при расчете задач методом начальных функций [6]. Он был программным способом устранен автором. Ниже приведен текст на языке Maple процедур интегрирования и взятия модуля числа.

```
> vlasov[fabs]:=proc(inp)
local per;
per:=evalf(inp);
if per<0 then per:=-per; fi;
per;
end proc;
> vlasov[integral]:=proc(input, perem)
local arg, gran1, gran2, result;
arg:=op(perem)[1]; op(rhs(perem));
gran1:=%[1]; gran2:=%[2];
result:=int(input, arg);
simplify(eval(result, arg=gran2)-eval(result, arg=gran1));
end proc;
```

Для того чтобы решить задачи в системе, необходимо запустить среду Maple и подключить откомпилированную библиотеку к создаваемому документу. Для этого в режиме `worksheet` Maple набирается следующая последовательность команд [5]:

```
> restart;
> read(`c:/Vlasov. m`);
> with(vlasov);
```

Решение задач теории плит, пластин

Для решения задач теории упругости для плит, пластин необходимо вызвать процедуру `operators`

со следующим списком параметров в скобках: *vaxes* – обязательный параметр, задает систему координат (*cartesian* – обычная декартова система, *polar* – полярная); *vdimension* – обязательный параметр, задающий размерность задачи (2 – плоская задача, 3 – трехмерная); *vchange* – необязательный параметр, он меняет ось, относительно которой строится решение в плоской задаче (принимает значение *yes*). По умолчанию строится решение относительно оси *y*, на практике нужно получать решение и для оси *x*; *vtype* – тип решения, *bigarmonic* – бигармоническое (*exp* – экспоненциальное). Для трехмерной задачи вместо необязательного параметра *vchange* используется обязательный *vaxe* (*x*, *y*, *z* – оси вывода). После вызова процедуры *operators* и задания начальной инициализации автоматически получаем решение общей задачи теории упругости. Пример вызова процедуры указан ниже:

```
> operators(vaxes=cartesian, vdimension=3, vaxe=z,
vtype=bigarmonic):
> operators(vaxes=cartesian, vdimension=2,
vtype=bigarmonic):
```

Общее решение представляется в операторно-символической форме и хранится в переменных *U*, *V*, *X*, *Y*, *sigma_x*, *sigma_y*, *tau* в случае двумерной задачи и в *U*, *V*, *W*, *X*, *Y*, *Z*, *sigma_x*, *sigma_y*, *tau* – в трехмерном случае. Операторно-символическая форма, содержащая дифференциалы высших порядков, слишком сложна для проведения вычислительных операций. Поэтому, после подстановки начальных функций в общее дифференциальное решение двумерной и трехмерной задачи и проведения сложных математических операций, на выходе получается решение, содержащее числовые ряды [6]. Для двумерного случая эти операции осуществляются путем вызова процедур библиотеки *n(expr)*, где *expr* – решение в операторной форме (а именно, через операторы *Luu*, *Luv* и т. д.) и *nim(expr)* для трехмерной задачи. После перевода в числовую форму над решением производится процедура Галеркина. Для двумерной задачи это процедура *galerkin2D(expr, sin(a[n]*x))*, трехмерной – *galerkin3D(expr, sin(a[n]*x)*cos(b[m]*y))*. После этого строятся уравнения, с помощью которых решается задача и находятся коэффициенты *An*, *Bn* для двумерной задачи и *Anm*, *Bnm*, *Cnm* – для трехмерной.

Решение задач для полуплоскости, полупространства

Эти задачи в системе решаются следующим образом. Вызываются процедуры:

```
> semi(saxes = cartesian, sdimension = 2, stype = napr);
```

```
> phi := \frac{2 \cdot y^3}{((x - xi)^2 + y^2)^2}; psi := \frac{y^2}{(x - xi)^2 + y^2};
R := sqrt((x - xi)^2 + y^2);
> seminital(phi, psi, R):
```

Для решения задачи используются две процедуры. Первая процедура *semi* имеет следующий список параметров в скобках: *saxes* – обязательный параметр, задает систему координат (*cartesian* – обычная декартова система, *polar* – полярная); *sdimension* – обязательный параметр, задающий размерность задачи (2 – плоская задача, 3 – трехмерная); *stype* – тип решения; *napr* – напряжения (*perem* – перемещения). Вызвав процедуру *semi*, выбираем тип задачи. Передав два параметра процедуре *seminital*, задаем, над какими функциями производятся операции. Третий параметр – переменная, относительно которой производятся упрощения. В поставленных задачах – это переменная *R*.

ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ ТЕОРИИ УПРУГОСТИ В ИНСТРУМЕНТАЛЬНОЙ СИСТЕМЕ

При решении множества задач теории упругости для плит, оболочек, пластин их сводят с помощью различных инженерных гипотез и допущений к общим модельным задачам. Математическая модельная задача значительно проще при реализации расчета. Математическая модель вместе с математическим выражением инженерных гипотез составляет расчетную модель конструкции.

Расчетная модель дает математическую зависимость между основными компонентами теории упругости и новыми величинами, относительно которых будут построены все соотношения. Чаще всего в моделях используются перемещения, напряжения и величины, их заменяющие.

Рассмотрим детально процесс построения модели решения двумерных и трехмерных задач для плиты и массива.

Задача 1. Рассмотрим свободно опертую по краям плиту длины *S* и толщины *h*, изгибаемую нагрузкой постоянной интенсивности *P* [10].

Выбираются начальные функции в виде

$$U_0(x, y) = \sum_{n=1}^{\infty} \sin A_n \frac{n\pi}{S} x,$$

$$V_0(x, y) = \sum_{n=1}^{\infty} B_n \cos \frac{n\pi}{S} x.$$

На языке системы они записываются в виде:

```
> Uo:= Sum(Anm*sin(a[n]*x), n=1.. p);
> Vo:=Sum (Bnm*cos(a[n]*x), n=1.. p);
```

В результате коэффициенты A_n и B_n на Maple:

$$\begin{aligned}
 A_n = & - \left(4h \sin(n\pi)^3 \left(e^{\frac{n\pi h}{S}} \right)^2 - 1 \right) P S^2 e^{\frac{n\pi h}{S}} (-1 + \nu) / \left(-2n^2 \pi^2 h^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^2 + n^2 \pi^2 h^2 \sin(n\pi)^4 + \right. \\
 & + n^2 \pi^2 h^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^4 + S^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^4 + S^2 \sin(n\pi)^4 - h^2 n^2 \pi^2 \cos(n\pi)^2 \sin(n\pi)^2 \left(e^{\frac{n\pi h}{S}} \right)^4 - \\
 & - 2h^2 n^2 \pi^2 \cos(n\pi)^2 \sin(n\pi)^2 \left(e^{\frac{n\pi h}{S}} \right)^2 + 2hn^2 \pi^2 \sin(n\pi) - 2hn^2 \pi^2 \left(e^{\frac{n\pi h}{S}} \right)^2 S \sin(n\pi)^2 - \\
 & \left. - h^2 n^2 \pi^2 \cos(n\pi)^2 \sin(n\pi)^2 - 2S^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^2 + h^2 n^4 \pi^4 + 2h^2 n^4 \pi^4 \left(e^{\frac{n\pi h}{S}} \right)^2 + h^2 n^4 \pi^4 \left(e^{\frac{n\pi h}{S}} \right)^2 \right); \\
 B_n = & 4 \sin n\pi \left(hn\pi \cos(n\pi) \sin(n\pi) \left(e^{\frac{n\pi h}{S}} \right)^2 + hn\pi \cos(n\pi) \sin(n\pi) - hn^2 \pi^2 \left(e^{\frac{n\pi h}{S}} \right)^2 - \right. \\
 & - hn^2 \pi^2 + S \sin(n\pi)^2 \left(e^{\frac{n\pi h}{S}} \right)^2 - S \sin(n\pi)^2 \left. \right) P S^2 e^{\frac{n\pi h}{S}} (-1 + \nu) / \left(n\pi \left(-2n^2 \pi^2 h^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^2 + \right. \right. \\
 & + n^2 \pi^2 h^2 \sin(n\pi)^4 + n^2 \pi^2 h^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^4 + S^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^4 + S^2 \sin(n\pi)^4 - \\
 & - h^2 n^2 \pi^2 \cos(n\pi)^2 \sin(n\pi)^2 \left(e^{\frac{n\pi h}{S}} \right)^4 - 2h^2 n^2 \pi^2 \cos(n\pi)^2 \left(e^{\frac{n\pi h}{S}} \right)^2 + 2hn^2 \pi^2 S \sin(n\pi)^2 - \\
 & - 2hn^2 \pi^2 \left(e^{\frac{n\pi h}{S}} \right)^4 S \sin(n\pi)^2 - h^2 n^2 \pi^2 \cos(n\pi)^2 \sin(n\pi)^2 - 2S^2 \sin(n\pi)^4 \left(e^{\frac{n\pi h}{S}} \right)^2 + \\
 & \left. \left. + h^2 n^4 \pi^4 + 2h^2 n^4 \pi^4 \left(e^{\frac{n\pi h}{S}} \right)^2 + h^2 n^4 \pi^4 \left(e^{\frac{n\pi h}{S}} \right)^4 \right) \right).
 \end{aligned}$$

Задача 2. Рассмотрим длинный в направлении x трехмерный массив, нагруженный постоянной нагрузкой Q в направлении этой оси, а также оси y . Размеры массива: длина (ось x) – S , ширина (по оси y) – R , толщина (ось z) – H .

Выбираются начальные функции в виде:

$$\begin{aligned}
 U_0(x, y) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{nm} \sin \frac{n\pi}{S} x \cos \frac{m\pi}{R} y, \\
 V_0(x, y) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} B_{nm} \cos \frac{n\pi}{S} x \sin \frac{m\pi}{R} y, \\
 W_0(x, y) &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} C_{nm} \cos \frac{n\pi}{S} x \sin \frac{m\pi}{R} y.
 \end{aligned}$$

На языке системы они записываются в следующем виде:

> Uo:=Sum(Sum(Anm*sin(a[n]*x)*cos(b[m]*y), m=1.. p), n=1.. p);
 > Vo:=Sum(Sum(Bnm*cos(a[n]*x)*sin(b[m]*y), m=1.. p), n=1.. p);
 > Wo:=Sum(Sum(Cnm*sin(a[n]*x)*sin(b[m]*y), m=1.. p), n=1.. p);

Результаты для коэффициентов A_{nm} , B_{nm} , C_{nm} , выведенные системой в трехмерном случае:

$$\begin{aligned}
 A_{nm} &= 0; \\
 B_{nm} &= 0;
 \end{aligned}$$

$$\begin{aligned}
 C_{nm} = & (4QS^2R^2\gamma(-1 + \nu + \cos(m\pi) - \cos(m\pi)\nu + \\
 & - \cos(n\pi)\nu + \cos(n\pi)\nu - \cos(n\pi)\cos(m\pi) + \\
 & + \cos(n\pi)\cos(m\pi)\nu) / (\pi^4 nm(\gamma h \cosh(\gamma h)n^2 R^2 + \\
 & + h \cosh(\gamma h)\gamma m^2 S^2 - \sinh(\gamma h)n^2 R^2 - \sinh(\gamma h)m^2 S^2)).
 \end{aligned}$$

Далее в общее численное решение подставляются эти коэффициенты и строятся циклы, в которых производится суммирование. ЭВМ оперирует с дискретным набором данных, в суммах ограничиваем число членов до двадцати. Результаты аналитического и числового решения представлены в работах [6], [8].

Для полупространства решались задачи Фламана и Черутти.

Задача 3. Пусть изотропная среда занимает бесконечную область, ограниченную прямой. Такое полубесконечное тело называется полуплоскостью с границей. Рассмотрим задачу о действии нагрузки на границу полуплоскости $y \geq 0$:

$$\tau_{xy}|_{y=0} = P(x, \xi),$$

$$\sigma_y|_{y=0} = R(x, \xi).$$

При заданных условиях компоненты напряженно-деформированного состояния не зависят от координаты w по z , а перемещение $w \equiv 0$. Имеем плоское напряженно-деформированное состояние и плоскую деформацию. Краевая задача замыкается условием обращения в ноль компонентов напряженно-деформированного состояния на бесконечности.

Для задачи Черутти $P(x, \xi) = \delta(x - \xi)P$ – сосредоточенная сила, $R(x, \xi) = \delta(x - \xi)R$, $R = 0$.

Результат решения задачи Черутти для напряжений на Maple:

$$\text{sigmay} := \frac{y^2(-2x + 2\xi)}{\pi((x - \xi)^2 + y^2)^2};$$

$$\text{sigmax} := -\frac{2(x^3 - 3x^2\xi + 3x\xi^2 - \xi^3)}{\pi(x^2 - 2x\xi + \xi^2 + y^2)^2};$$

$$\text{tauxy} := \frac{-\frac{2y}{x^2 - 2x\xi + \xi^2 + y^2} + \frac{2y^3}{(x^2 - 2x\xi + \xi^2 + y^2)^2}}{\pi}.$$

В случае задачи Фламана $R(x, \xi) = \delta(x - \xi)R$, $P(x, \xi) = \delta(x - \xi)P$, $P = 0$.

Решение на Maple для напряжений:

$$\text{sigmay} := \frac{2y^3}{\pi((x - \xi)^2 + y^2)^2};$$

$$\text{sigmax} := \frac{2y(x^2 - 2x\xi + \xi^2)}{\pi(x^2 - 2x\xi + \xi^2 + y^2)^2};$$

$$\text{tauxy} := \frac{\frac{2y^2x}{(x^2 - 2x\xi + \xi^2 + y^2)^2} - \frac{2y^2\xi}{(x^2 - 2x\xi + \xi^2 + y^2)^2}}{\pi}.$$

ВЫВОДЫ

Таким образом, в статье была охарактеризована инструментальная система, осуществляющая процесс моделирования решения задач теории упругости для плиты, полуплоскости в двумерной и массива в трехмерной постановке. Результатами работы системы являются решения для прогибов двумерного и трехмерного тел. При программировании аналитических операций возникли сложности технического характера, связанные с недоработкой вычислительных средств Maple. Внедрением новых программных продуктов в ядро системы эти сложности были устранены.

СПИСОК ЛИТЕРАТУРЫ

1. Аладьев В. З. Системы компьютерной алгебры: Maple: Искусство программирования / Аладьев В. З. – М. : Лаборатория базовых знаний, 2006. – 792 с.
2. Кабулов В. К. Алгоритмическая система для решения задач теории упругости и пластичности, Сб. / И. К. Кабулов, В. А. Толоч // Вопросы вычислительной и прикладной математики. – 1970. – № 4. – С. 37–45.
3. Гоменюк С. И. Объектно-ориентированные модели и методы анализа механических процессов / С. И. Гоменюк. – Никополь : Никопольская коммунальная типография, 2004. – 316 с.
4. Матросов А. В. Mapleб. Решение задач высшей математики и механики / Матросов А. В. – СПб. : БХВ-Петербург, 2002. – 528 с.
5. Овский А. Г. Использование системы Maple при реализации метода начальных функций Власова / Е. Е. Галан, Овский А. Г., В. А. Толоч // Вісник Запорізького національного університету : Збірник наукових статей. Фізико-математичні науки. – 2008. – № 1. – С. 16–26.
6. Овский А. Г. Моделирование схемы решения трехмерной задачи теории упругости в системе Maple / А. Г. Овский, В. О. Толоч // Гідроакустичний журнал. – 2008. – № 3. – С. 88–97.
7. Власов В. З. Балки плиты и оболочки на упругом основании. / В. Власов, Н. Леонтьев – Москва : ФИЗМАТГИЗ, 1960. – 491 с.
8. Толоч В. А. Операторно-символьные ряды Власова В. З. в решении задач теории упругости в системе Maple / В. В. Шапар, В. А. Толоч // Гідроакустичний журнал. – 2006. – № 3. – С. 66–74.
9. Овский А. Г. Использование системы компьютерной математики Maple для доказательства закона ортогональности матриц прямого и обратного преобразования, составленных Власовым В. З. / А. Г. Овский, В. А. Толоч // Радіоелектроніка, інформатика, управління. – 2008. – № 1. – С. 78–85.
10. Горшков А. Г. Теория упругости и пластичности : учеб. для вузов / Горшков А. Г., Старовойтов Э. И., Талаковский Д. В. – М. : ФИЗМАТЛИТ, 2002. – 416 с.

Надійшла 1.07.2009
Після доробки 12.05.2010

Овський О. Г., Толоч В. А.

ІНСТРУМЕНТАЛЬНА СИСТЕМА РОЗВ'ЯЗАННЯ ДВОВИМІРНИХ І ТРИВИМІРНИХ ЗАДАЧ ТЕОРІЇ ПРУЖНОСТІ НА БАЗІ MAPLE

У статті наведено опис інструментальної системи розв'язання двовимірних і тривимірних задач теорії пружності. Обчислювальне ядро системи складає Maple. Здійснено формалізацію основних алгоритмів з метою знаходження операцій програмування вказаних задач.

Як приклади роботи з системою наведено розв'язки конкретних задач.

Ключові слова: препроцесор, процесор, постпроцесор, метод початкових функцій.

Ovsky A., Tolok V.

DEVELOPMENT SYSTEM FOR SOLVING TWO-DIMENSIONAL AND THREE-DIMENSIONAL ELASTICITY THEORY PROBLEMS BASED ON MAPLE

The authors describe the development system for solving two-dimensional and three-dimensional problems of the theory of elasticity. Maple is a computing kernel of the system. The basic algorithms are formalized for the purpose of finding programming operations of the specified problems. Solutions of specific problems are given as examples of system operation.

Key words: preprocessor, processor, postprocessor, method of initial functions.

УДК 004.93

Олійник А. О.

Канд. техн. наук, старший викладач Запорізького національного технічного університету

МУЛЬТИАГЕНТНИЙ МЕТОД З ОПОСЕРЕДКОВАНИМ ЗВ'ЯЗКОМ МІЖ АГЕНТАМИ

Досліджено метод оптимізації на основі мультиагентного підходу. Розроблено мультиагентний метод з опосередкованим зв'язком між агентами, що передбачає виконання додаткових етапів відбору, схрещування та мутації. Проведено експерименти по розв'язанню задач оптимізації багатовимірних нелінійних функцій, що дозволило дослідити ефективність запропонованого методу.

Ключові слова: агент, генетичні оператори, мультиагентний метод, оптимізація, цільова функція.

ВСТУП

Останнім часом широкого застосування набули напрямки штучного інтелекту, основні ідеї яких узяті з природи. До таких напрямків можна віднести: штучні нейронні мережі [1, 2], методи еволюційного пошуку [3, 4], методи оптимізації, засновані на моделюванні колективного інтелекту соціальних тварин [5, 6], штучні імунні системи [7] та ін. [8–13]. Суттєвими перевагами таких напрямків штучного інтелекту є можливість їхнього застосування до розв'язання слабоформалізованих задач та здатність до видобування знань з даних.

Зокрема, мультиагентні методи інтелектуальної оптимізації не висувають додаткових вимог до виду цільової функції, що дозволяє їх застосовувати при розв'язанні оптимізаційних задач, які виникають при побудові моделей складних багатовимірних об'єктів, процесів та систем (на етапах відбору інформативних ознак, структурно-параметричного синтезу моделей, спрощення побудованих моделей). Крім того, на відміну від градієнтних методів багатовимірної нелінійної оптимізації, що традиційно застосовуються при побудові моделей, мультиагентні методи є методами глобальної оптимізації та не схильні до заціклення у локальних оптимумах.

© Олійник А. О., 2010

Для розв'язання завдань дискретної та безперервної оптимізації ефективно застосовується метод оптимізації на основі моделювання переміщення бактерій (Bacteria Foraging Optimization, BFO) [11]. Однак моделі поведінки живих істот, що лежать в основі роботи даного методу, є неповними та не зовсім відповідають своїм природним аналогам. Зокрема, в методі BFO не досить ефективним є механізм утворення нової множини рішень, що призводить до тривалих розрахунків при виконанні оптимізаційного процесу [11].

Тому актуальною є розробка доповнень існуючих моделей, що використовуються для моделювання поведінки живих істот, які забезпечують більш повну відповідність природним аналогам і кращу ефективність роботи зазначених методів оптимізації.

ПОСТАНОВА ЗАВДАННЯ

Метод BFO ефективно застосовується для розв'язання завдань оптимізації [11], однак математичні моделі, розроблені для даного методу, навіть на думку їх розробників [11] є неповними та не досить ефективними. Тому необхідно доробити відомі математичні моделі BFO для більш адекватного моделювання еволюційних процесів життєдіяльності бактерій.

Метою даної роботи є удосконалення мультиагентного методу оптимізації на основі моделювання переміщення бактерій шляхом розширення його математичної моделі.

Основними завданнями роботи є:

- аналіз методу оптимізації на основі моделювання переміщення бактерій;
- включення операторів еволюційного пошуку у вже існуючу математичну модель, яка застосовується в методі оптимізації на основі моделювання переміщення бактерій.
- розробка програмного забезпечення, що реалізує запропоновану математичну модель, і його тестування.

МЕТОД ОПТИМІЗАЦІЇ НА ОСНОВІ МОДЕЛЮВАННЯ ПЕРЕМІЩЕННЯ БАКТЕРІЙ

Даний метод призначений для знаходження мінімуму функції $J(X)$, $X \in R^p$, при невідомому градієнті $\nabla J(X)$, де X – позиція агента (бактерії) в просторі пошуку R^p , а за допомогою $J(\theta)$ моделюються корисні та шкідливі властивості середовища, тобто $J(X)$ характеризує, де перебувають аттрактанти та ренти. Таким чином, $J < 0$, $J = 0$, $J > 0$ означає, що бактерія перебуває в корисному, нейтральному та шкідливому середовищі, відповідно.

В методі BFO важливим етапом є моделювання хемотаксису. Хемотаксис – це рухова реакція бактерії у відповідь на появу в середовищі аттрактанта (аттрактант – речовина, що притягує бактерії) або репелента (репелент – речовина, що відштовхує бактерії) [3]. У природніх умовах аттрактантами є речовини, корисні для бактерій, а репелентами – ті, які бактеріям шкідливі.

При наявності просторових змін концентрації аттрактантів або репелентів частота перекидань, а, отже, і довжина вільного пробігу бактерії змінюються. Довжина вільного переміщення бактерії, що рухається в сторону зростаючої концентрації аттрактанта, збільшується, а при русі в сторону зростаючої концентрації репелента зменшується.

Таким чином, можна виділити наступні хемотаксичні дії бактерії, які застосовуються в методі BFO [13]:

– якщо бактерія перебуває в нейтральному середовищі, то чергуються перекидання та пересування, за рахунок чого виконується пошук;

– якщо бактерія переміщається за градієнтом аттрактанта, то переміщення триває в цьому ж напрямку. Таким чином, забезпечується пошук більш сприятливого навколишнього середовища (глобальний пошук);

– якщо відбувається переміщення в напрямку, протилежному градієнту репелента, то забезпечується вихід з несприятливого навколишнього середовища.

Таким чином, бактерія може переміщатися в областях з корисними речовинами й у той же час уникати небезпечних речовин.

Враховуючи аналогії з природніми процесами, метод BFO передбачає виконання таких етапів:

- 1) ініціалізація;
- 2) моделювання хемотаксису;
- 3) відтворення;
- 4) виключення й розсіювання;
- 5) перевірка критеріїв зупинення;
- 6) перезапуск агентів.

Таким чином, метод BFO містить етапи, що властиві відомим методам оптимізації (ініціалізація та перевірка критеріїв зупинення), мультиагентним методам (відтворення та перезапуск агентів) та природнім процесам переміщення бактерій (моделювання хемотаксису та виключення й розсіювання).

Будемо використовувати $X(j, k, l) = \{X^i(j, k, l), i = 1, 2, \dots, S\}$ для опису позиції кожного члена популяції S агентів (бактерій) на j -му хемотаксичному кроці, k -му кроці відтворення та на l -й події виключення-розсіювання.

На етапі *ініціалізації* в методі BFO задаються параметри, що впливають на роботу методу: S – кількість бактерій, N_{re} – кількість кроків відтворення, N_s – кількість кроків-повторень на одному хемотаксичному кроці, N_c – кількість хемотаксичних кроків, N_{ed} – кількість подій виключення-розсіювання; P_{ed} – імовірність розсіювання.

На цьому етапі також випадковим чином розподіляються початкові значення X^i , $i = 1, 2, \dots, S$ по простору пошуку та розраховуються початкові значення цільової функції для кожного агента J^i .

При *моделюванні хемотаксису* над кожним агентом виконуються оператори кувиркання, переміщення та ковзання.

Алгоритм виконання етапу моделювання хемотаксису методу BFO може бути поданий такою послідовністю кроків.

Крок 1. Встановити: $i = i + 1$.

Крок 2. Виконати оператор кувиркання для i -го агента. Моделювання кувиркання досягається за рахунок генерації вектора випадкових чисел $\phi(j) \in R^p$:

$$\phi = \frac{\Delta}{\sqrt{\Delta^T \Delta}},$$

де Δ – вектор випадкових чисел в інтервалі $[-1; 1]$. Вектор ϕ являє собою множину довжин для відповідних вимірів.

Крок 3. Переміщення. Розраховується нове положення i -ї бактерії за формулою:

$$X^i(j+1, k, l) = X^i(j, k, l) + C(i)\varphi(j),$$

де $C(i) > 0$ – розмір кроку в певному напрямку, що дозволяє моделювати процес кувиркання.

Для нової позиції $X^i(j+1, k, l)$ розраховується відповідне значення цільової функції $J(i, j+1, k, l)$.

Крок 4. Ковзання. Якщо в позиції $X^i(j+1, k, l)$ значення $J(i, j+1, k, l)$ краще, ніж у позиції $X^i(j, k, l)$, тобто виконується умова: $J(i, j+1, k, l) < J(i, j, k, l)$, тоді проводиться наступний хемотаксичний крок з тим же вектором φ і у тому ж напрямку (перехід до кроку 3), і таке повторення може виконуватися N_s раз. Якщо умова не виконується, то відбувається перехід до кроку 5.

Крок 5. Якщо $i < S$, то виконується перехід до кроку 1, в іншому випадку – перехід до наступного етапу.

Моделювання процесу хемотаксису відбувається N_c разів.

На етапі *відтворення* менш здоровіші бактерії вмирають (агенти з гіршими значеннями цільової функції видаляються), а інші, більш здоровіші, розділяються на дві бактерії, при цьому нові бактерії розміщуються в ту ж саму точку простору пошуку. За рахунок такого підходу забезпечується незмінність загальної кількості бактерій. Для цього все бактерії сортується відповідно до отриманих значень цільової функції, після чого гірша половина (менш здоровіші бактерії) відкидається, а краща (більш здоровіші бактерії) – дублюється.

Етап *виключення й розсіювання* призначений для зменшення ймовірності зациклення в локальних оптимумах. Життя популяції бактерій у навколишньому середовищі може змінюватися поступово (наприклад, шляхом споживання корисних речовин) або раптово у зв'язку з деяким іншим впливом. Може відбутися так, що всі бактерії в області загинуть, або колонія бактерій буде розсіяна в іншу частину навколишнього середовища. Даний ефект може перешкодити можливому хемотаксичному прогресу, але в той же час цей ефект і допомагає, оскільки, у випадку розсіювання, бактерії можуть розміститися близько гарних джерел з корисними речовинами. Таким чином, виключення й розсіювання допомагають знизити ймовірність стагнації, тобто зациклення в локальному оптимумі, що часто спостерігається в традиційних градієнтних методах оптимізації.

Відповідно до даного підходу кожна бактерія з ймовірністю P_{ed} розміщується у випадково обраній точці простору пошуку. Таким чином, перевіряється

умова: $U^i < P_{ed}$, де U^i – випадкове число в інтервалі $[0; 1]$ для i -ї бактерії. Якщо дана умова виконується, то бактерія переміщується в позицію $X^i(j, k, l)$, отриману випадковим чином.

Після виконання N_{ed} разів етапів моделювання хемотаксису, відтворення та виключення й розсіювання виконується збереження кращого розв'язку J_{best} та відповідної позиції X_{best} у якій досягається кращий розв'язок J_{best} .

Пошук оптимуму припиняється, якщо досягнуто *критеріїв зупинення* методу, зокрема, якщо знайдено прийнятний розв'язок, або у випадку, якщо були виконані всі цикли для всіх бактерій.

Якщо критерії зупинення пошуку не досягнуто, то виконується етап *перезапуску агентів*: вибираються нові випадкові позиції для кожного агента X^i , $i = 1, 2, \dots, S$ та розраховуються відповідні значення цільової функції J^i , $i = 1, 2, \dots, S$, після чого знову виконується етап моделювання хемотаксису.

Метод оптимізації на основі моделювання переміщення бактерій застосовувався для розв'язку різних завдань оптимізації: знаходження оптимального потоку енергії [14], динамічного розподілу ресурсів [15] та ін. [16, 17]. Результати проведених експериментів з вирішення тестових завдань [9, 14–17] показали, що застосування методу BFO є ефективним і перспективним, оскільки результати, отримані за допомогою даного методу, характеризувалися більш високою або достатньою точністю в порівнянні з відомими методами оптимізації.

Можна виділити наступні переваги методу BFO у порівнянні з традиційними методами багатовимірної оптимізації:

- метод не висуває ніяких вимог до цільової функції, тобто функція може бути мультимодальною, негладкою, мати розриви;

- для знаходження оптимуму цільової функції не потрібно розраховувати градієнт цільової функції, тобто вона може бути недиференційованою;

- метод є методом глобального пошуку, у той час як більшість класичних методів оптимізації є методами локального пошуку;

- метод має достатню кількість параметрів, що настроюються, що дозволяє забезпечити його кращу адаптацію для розв'язання того або іншого практичного завдання оптимізації;

- метод є мультиагентним методом, за рахунок чого досягається зниження часу виконання пошуку, оскільки пошук виконується деякою кількістю агентів паралельно;

– метод є добре розширюваним, тобто в його роботу можна легко додавати нові оператори й тим самим створювати гібридні системи.

До недоліків методу варто віднести такі:

- випадковий характер швидкості збіжності;
- сильна залежність роботи методу від параметрів, що настроюються;
- неповнота використовуваної моделі поведінки бактерій, зокрема не досить ефективний механізм відтворення агентів.

МУЛЬТИАГЕНТНИЙ МЕТОД З ОПОСЕРЕДКОВАНИМ ЗВ'ЯЗКОМ МІЖ АГЕНТАМИ

Як відзначалося вище, в методі BFO не досить ефективним є механізм відтворення, який призначений для утворення нової множини рішень, що призводить до тривалих розрахунків при виконанні оптимізаційного процесу.

Для усунення зазначеного недоліку в розробленому мультиагентному методі з опосередкованим зв'язком між агентами пропонується розширити модель методу BFO шляхом впровадження генетичних операторів відбору, схрещування та мутації.

Пропонується використовувати генетичні оператори:

- для відбору – пропорційний відбір;
- для схрещування – випадковий вибір батьківської пари з арифметичним схрещуванням;
- для мутації – проста мутація.

У запропонованому методі взаємозв'язок між агентами відбувається за рахунок моделювання опосередкованого зв'язку між ними, оскільки агенти моделюють виділення речовини в залежності від якості середовища, після чого наступні агенти при своєму переміщенні враховують властивості середовища, які у свою чергу залежать як від об'єктивних факторів (значення цільової функції), так і від виділених іншими агентами речовин. Таким чином зв'язок між агентами відбувається не за рахунок прямого (безпосереднього) обміну інформацією між ними, а за допомогою інформації, отримуваної ними через навколишнє середовище.

Запропонований мультиагентний метод з опосередкованим зв'язком між агентами можна представити у вигляді послідовності таких етапів.

Етап 1. *Ініціалізація*. Задаються параметри, що впливають на роботу методу: S , N_{re} , N_s , N_c , N_{ed} , P_{ed} . Випадковим чином розподілити початкові значення X^i , $i = 1, 2, \dots, S$ у просторі пошуку. Розраховуються початкові значення цільової функції для кожної бактерії J^i .

Для кожної бактерії виконується розрахунок величини J , що оцінює ефективність застосування агента для оптимізації цільової функції J^* :

$$J^i = J^{*i} + J_{ar}(X^i),$$

де величина $J_{ar}(X^i)$ розраховується за формулою

$$J_{ar}(X) = \exp(M - J(X))J_{cc}(X),$$

де M – параметр, що настроюється; $J_{cc}(X)$ розраховується за формулою

$$J_{cc}(X) = \sum_{i=1}^S J_{cc}^i = \\ = \sum_{i=1}^S \left[-d_{attract} \exp\left(-w_{attract} \sum_{j=1}^p (X_j - X_j^i)^2\right) \right] + \\ + \sum_{i=1}^S \left[h_{repellent} \exp\left(-w_{repellent} \sum_{j=1}^p (X_j - X_j^i)^2\right) \right],$$

де $d_{attract}$ – глибина атрактанту (корисних речовин) клітини; $w_{attract}$ – міра ширини атрактанту $h_{repellent} = d_{attract}$ – висота шкідливої речовини (репелента); $w_{repellent}$ – міра ширини репелента.

Таким чином, клітки виключають одна одну за допомогою локального споживання, а також за рахунок того, що клітини не є їжею одна для одної.

Етап 2. *Моделювання хемотаксису*. Цей етап доцільно виконувати у такій послідовності кроків.

Крок 1. Установити: $i = i + 1$.

Крок 2. Кувіркання й переміщення. Розраховується нове положення i -го агента за формулою

$$X^i(j+1, k, l) = X^i(j, k, l) + C(i)\varphi(j).$$

Після цього розраховується значення цільової функції $J(i, j+1, k, l)$, встановлюється: $J(i, j+1, k, l) = J(i, j+1, k, l) + J_{ar}(X^i(j+1, k, l))$.

Якщо $J(i, j+1, k, l) < J_{best}$, то встановити: $J_{best} = J(i, j+1, k, l)$, $X_{best} = X^i(j+1, k, l)$.

Крок 3. Ковзання. Поки виконується умова $J(i, j+1, k, l) < J(i, j, k, l)$, повторюється крок 2. Таке повторення може повторюватися N_s разів. Якщо умова не виконується, то виконується перехід до кроку 4.

Крок 4. Якщо $i < S$, то здійснюється перехід до кроку 1, в іншому випадку перехід до наступного етапу.

Етап моделювання хемотаксису виконується N_c разів.

На третьому етапі виконується *відтворення* агентів, для чого застосовуються еволюційні оператори відбору, схрещування та мутації.

Для відбору агентів з метою генерації нових агентів пропонується використовувати пропорційний відбір, алгоритм виконання якого подано нижче.

Крок 1. Знайти середнє значення цільової функції J_{cp} популяції як середнє арифметичне значень цільових функцій усіх особин:

$$J_{cp} = \frac{1}{S} \sum_{i=1}^S J^i.$$

Крок 2. Для кожного агента обчислити відношення, що визначає ймовірність бути відібраним до схрещування:

$$P_s(i) = \frac{J^i}{J_{cp}}.$$

Крок 3. В залежності від величини $P_s(i)$ сформува-ти масив агентів, допущених до схрещування. Формування масиву допущених до схрещування агентів здійснюється шляхом стохастичного залишкового відбору: якщо $P_s(i) < 1$, то агент вважається добре пристосованим й допускається до схрещування.

Для генерації нових агентів на основі агентів, відібраних за допомогою оператора відбору, виконується випадковий вибір батьківських пар.

Крок 1. Для вибору пари батьків задати ймовірність схрещування P_c . Величина P_c залежить від розв'язуваного завдання, і в загальному випадку таку ймовірність доцільно обирати [3] з інтервалу $[0,6; 0,99]$.

Крок 2. Довільним чином пронумерувати всіх агентів, відібраних для схрещування.

Крок 3. Вибрати першого батька. Для цього, починаючи з першого, перебрати всіх агентів популяції доти, поки випадково обране число з інтервалу $[0; 1]$ не буде меншим за P_c . Коли ця умова виконається для одного з агентів, цей елемент стане першим батьком.

Крок 4. Продовжити перегляд популяції, починаючи з наступного після першого батька розв'язку. Перегляд продовжувати доти, поки знову випадково обране число не буде меншим за P_c . Агент, для якого виконається така умова, буде другим батьком.

Описаним способом складаються пари доти, поки не вибереться потрібна кількість пар батьків.

Далі виконується арифметичне схрещування.

При використанні арифметичного схрещування на основі двох батьків X^1 та X^2 створюються два нащадки X^{1n} і X^{2n} , значення i -х координат x^{1n}_i та x^{2n}_i яких розраховуються за формулами:

$$x^{1n}_i = k \cdot x^1_i + (1 - k)x^2_i;$$

$$x^{2n}_i = k \cdot x^2_i + (1 - k)x^1_i,$$

де $k \in [0; 1]$ – деякий дійсний коефіцієнт, який вибирається випадковим чином на кожній ітерації.

З метою організації пошуку в різних областях простору пошуку та виходу з локальних оптимумів на етапі відтворення агентів виконується оператор простої мутації.

Крок 1. Вибрати випадковим образом координату, що буде змінюватися.

Крок 2. У заданому інтервалі припустимих значень координати вибрати нове значення, що не дорівнює поточному.

Етап 4. *Виключення й розсіювання.* Перевіряється умова $U^i < P_{ed}$. Якщо ця умова виконується, то агент переміщається в позицію $X^i(j, k, l)$, отриману випадковим чином. Після перевірки цієї умови для усіх агентів виконується пошук та збереження кращого розв'язку J_{best} та відповідної йому позиції X_{best} .

Етап 5. *Перевірка критеріїв зупинення пошуку.* Якщо знайдено прийнятний розв'язок або були виконані всі цикли для всіх агентів, то виконується зупинення пошуку, в іншому випадку виконується перезапуск агентів.

Етап 6. *Перезапуск агентів:* вибираються нові випадкові позиції для кожного агента X^i , $i = 1, 2, \dots, S$, розраховуються відповідні значення цільової функції J^i , $i = 1, 2, \dots, S$. Лічильники циклів скидаються в 0: $j = 0$; $k = 0$; $l = 0$, після чого знову виконується етап моделювання хемотаксису.

ЕКСПЕРИМЕНТИ ТА РЕЗУЛЬТАТИ

Для перевірки ефективності запропонованого мультиагентного методу з опосередкованим зв'язком між агентами було розроблено програмне забезпечення, що реалізує запропонований та відомі мультиагентні методи. Програмне забезпечення розроблялося за допомогою середовища розробки пакету Matlab 7.0.

За допомогою розробленого програмного забезпечення вирішувалося завдання знаходження оптимуму таких тестових функцій:

$$\begin{aligned} f_1 = & 5 \cdot e^{-0,1 \cdot ((x_1 - 15)^2 + (x_2 - 20))^2} - \\ & - 2 \cdot e^{-0,08 \cdot ((x_1 - 20)^2 + (x_2 - 15))^2} + \\ & + 3 \cdot e^{-0,08 \cdot ((x_1 - 25)^2 + (x_2 - 10))^2} + \\ & + 2 \cdot e^{-0,1 \cdot ((x_1 - 15)^2 + (x_2 - 20))^2} - \\ & - 2 \cdot e^{-0,5 \cdot ((x_1 - 5)^2 + (x_2 - 10))^2} - \\ & - 4 \cdot e^{-0,1 \cdot ((x_1 - 15)^2 + (x_2 - 5))^2} - \\ & - 2 \cdot e^{-0,5 \cdot ((x_1 - 21)^2 + (x_2 - 25))^2} + \\ & + 2 \cdot e^{-0,5 \cdot ((x_1 - 25)^2 + (x_2 - 16))^2} + \\ & + 2 \cdot e^{-0,5 \cdot ((x_1 - 5)^2 + (x_2 - 14))^2} \end{aligned}$$

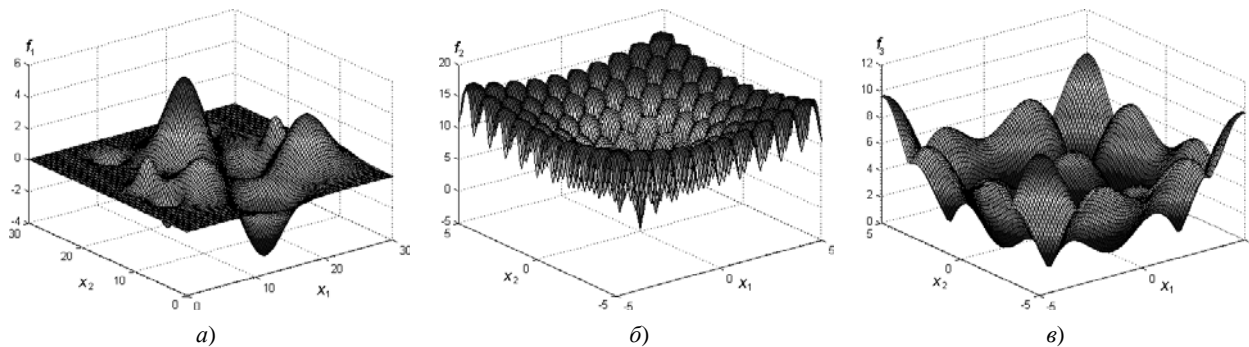


Рис. 1. Графіки тестових функцій f_1 (а), f_2 (б) та f_3 (в)

при умовах $0 \leq x_1 \leq 30; 0 \leq x_2 \leq 30;$

$$f_2 = 20 + e - 20 \exp\left(-0, 2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right), n = 2$$

при $-5 \leq x_1 \leq 5; -5 \leq x_2 \leq 5;$

$$f_3 = \sum_{i=1}^n |x_i \cdot \sin(x_i) + 0, 1 \cdot x_i|, n = 2$$

при $-5 \leq x_1 \leq 5; -5 \leq x_2 \leq 5.$

Графіки тестових функцій наведені на рис. 1.

Кожний метод запускався по 10 разів для знаходження оптимуму кожної функції. Використовувалися такі значення параметрів, що настроюються: $S = 20; N_{re} = 6; N_s = 4; N_c = 20; N_{ed} = 2; P_{ed} = 0,25; d_{attract} = 0,1; w_{attract} = 0,2; h_{repellant} = d_{attract}; w_{repellant} = 10.$

У процесі проведення експериментів розраховувалися такі показники: f' – кращий отриманий оптимум; Δf – відхилення отриманого оптимуму від реального; c_f – кількість обчислень цільової функції, i^* – номер хемотаксичного кроку, на якому було уперше знайдено результуючий кращий розв'язок; \bar{f} – середнє значення отриманих оптимумів функції для всіх запусків методу; $D(f)$ – дисперсія отриманих значень оптимумів.

Отримані в ході виконання експериментів значення показників для відомого методу оптимізації на основі моделювання переміщення бактерій наведені в табл 1.

Значення показників, отриманих у ході експериментів, для запропонованого мультиагентного методу з опосередкованим зв'язком між агентами наведені в табл. 2.

Графіки роботи методів у процесі знаходження оптимальних значень функцій представлені на рис. 2.

На підставі отриманих результатів експериментів можна зробити висновок, що запропонований метод характеризується більшою точністю отриманих результатів, кращою стабільністю одержуваних розв'язків, про що свідчать значення середніх і дисперсій, отриманих внаслідок декількох запусків кожного з методів. При цьому можна відзначити, що оптимальні результати запропонований метод знаходить на більш пізніх етапах своєї роботи, що у свою чергу сприяє одержанню більш оптимальних розв'язків. Також перевагою запропонованого методу є те, що його застосування потребує меншої кількості обчислень значень функції, що оптимізується.

Таблиця 1. Результати експериментів для методу BFO

	f^*	f'	Δf	c_f	i^*	\bar{f}	$D(f)$
f_1	-3,9865	-3,9863	0,0002	9252	48	-3,7868	0,6317
f_2	0	0,0065	0,0065	8626	207	0,0657	0,0434
f_3	0	0,0004	0,0004	9721	76	0,0037	0,0045

Таблиця 2. Результати експериментів для запропонованого мультиагентного методу

	f^*	f'	Δf	c_f	i^*	\bar{f}	$D(f)$
f_1	-3,9865	-3,9864	0,0001	5982	88	-3,9821	0,0052
f_2	0	0,0018	0,0018	6527	228	0,0278	0,0198
f_3	0	0,0002	0,0002	5790	62	0,0025	0,0040

ВИСНОВКИ

У роботі одержав подальший розвиток метод оптимізації на основі моделювання переміщення бактерій.

Наукова новизна роботи полягає в тому, що розроблено мультиагентний метод з опосередкованим зв'язком між агентами, в якому на етапі відтворення агентів застосовуються генетичні оператори відбору, схрещування й мутації, що дозволяє більш ефективно

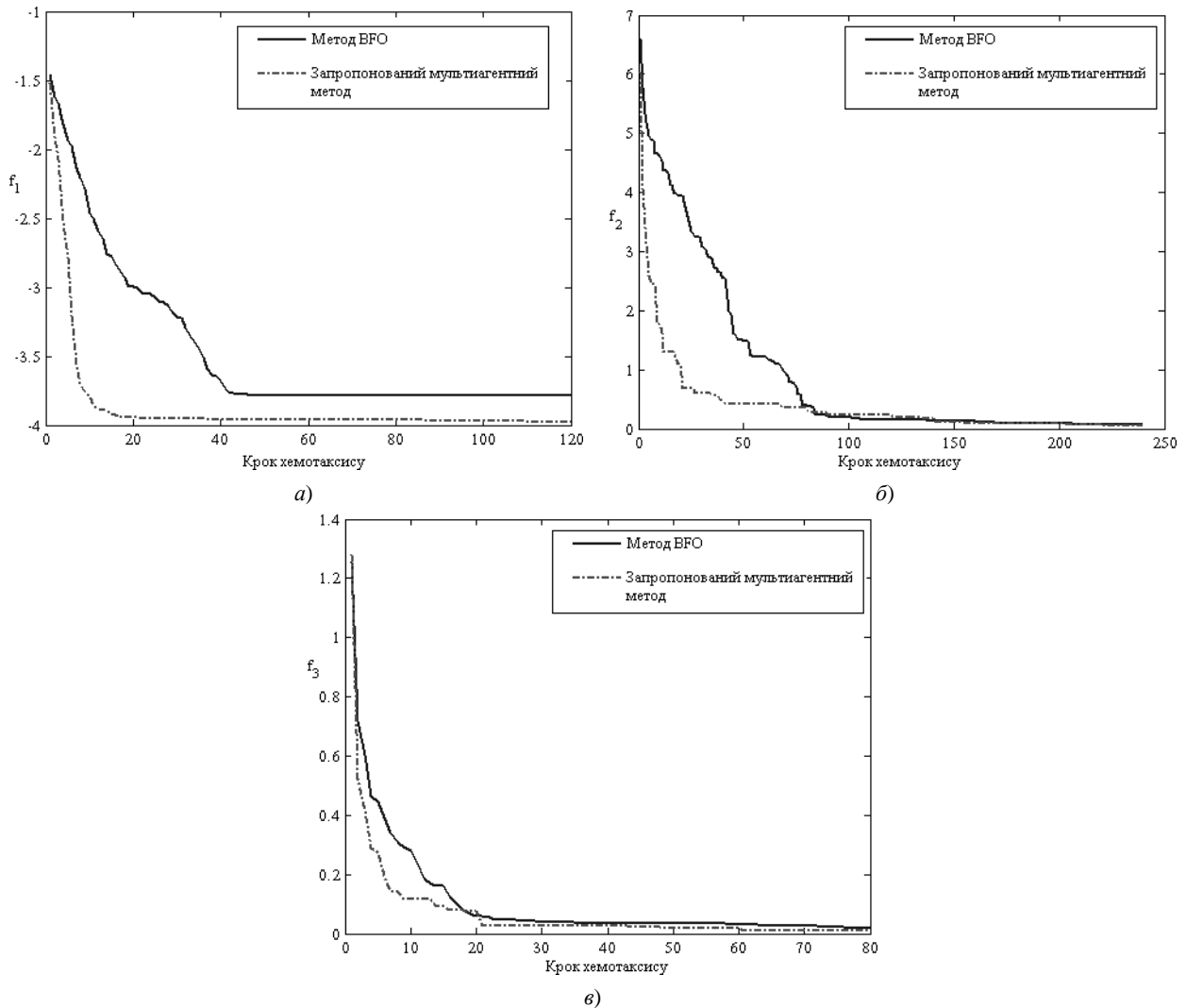


Рис. 2. Процес оптимізації тестових функцій f_1 (а), f_2 (б) та f_3 (в)

но в порівнянні з базовим BFO-методом досліджувати простір пошуку, що приводить до зменшення часу розрахунків при виконанні оптимізаційного процесу.

Практична цінність роботи полягає у тому, що розроблено програмне забезпечення, яке реалізує запропонований метод та дозволяє виконувати оптимізацію багатовимірних нелінійних цільових функцій без висування додаткових вимог до них.

На основі результатів проведених експериментів можна зробити висновок, що запропоновані модифікації методу BFO підвищують ефективність оптимізаційного процесу, що проявляється в: одержанні більш прийнятних значень цільової функції, більшій стабільності результатів, меншій кількості обчислень цільової функції. Таким чином, запропонований метод може бути рекомендований до застосування при вирішенні практичних завдань оптимізації.

Роботу виконано у рамках науково-дослідної роботи кафедри програмних засобів ЗНТУ «Інформаційні технології автоматизації розпізнавання образів і прийняття рішень для діагностики в умовах невизначеності на основі гібридних нечіткологічних, нейромережових і мультиагентних методів обчислювального інтелекту» (номер державної реєстрації 0109U007673).

СПИСОК ЛІТЕРАТУРИ

1. Хайкин С. Нейронные сети: полный курс / С. Хайкин. – СПб. : Издательский дом «Вильямс», 2005. – 1104 с.
2. Руденко О. Г. Штучні нейронні мережі / О. Г. Руденко, С. В. Бодянский. – Харків : Компанія СМІТ, 2006. – 404 с.
3. Субботін С. О. Неітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережових моделей: монографія / С. О. Субботін, А. О. Олійник, О. О. Олійник ; під заг. ред. С. О. Субботіна. – Запоріжжя : ЗНТУ, 2009. – 375 с.

4. The Practical Handbook of Genetic Algorithms. Applications / Ed. L. D. Chambers. – Florida : CRC Press, 2000. – Vol. I. – 520 p.
5. Beni G. Swarm Intelligence / G. Beni, J. Wang // Annual Meeting of the Robotics Society: Proceedings of Seventh International Conference. – Tokyo : RSJ Press, 1989. – P. 425–428.
6. Bonabeau E. Swarm Intelligence: From Natural to Artificial Systems / E. Bonabeau, M. Dorigo, G. Theraulaz. – New York : Oxford University Press, 1999. – 320 p.
7. De Castro L. N. Artificial Immune Systems. Part I. Basic Theory And Applications / L. N. De Castro, F. J. Von Zuben. – Technical Report No. Rt Dca 01/99. – Brazil : Feec/Unicamp, 1999. – 95 p.
8. Dorigo M. The Ant System: Optimization by a colony of cooperating agents / M. Dorigo, V. Maniezzo, A. Colomi // IEEE Transactions on Systems, Man, and Cybernetics. – 1996. – Part B, № 26(1). – P. 29–41.
9. Прогрессивные технологии моделирования, оптимизации и интеллектуальной автоматизации этапов жизненного цикла авиадвигателей: Монография / А. В. Богуслаев, Ал. А. Олейник, Ан. А. Олейник и др.; под ред. Д. В. Павленко, С. А. Субботина. – Запорожье : Мотор Сич, 2009. – 468 с.
10. Sumpter D. J. Formalising the Link between Worker and Society in Honey Bee Colonies / D. J. Sumpter, D. S. Broomhead // Lecture Notes In Computer Science: Proceedings of the First International Workshop on Multi-Agent Systems and Agent-Based Simulation. – MABS'98, LNAI, 1998. – P. 95–110.
11. Passino K. M. Biomimicry of bacterial foraging for distributed optimization and control / K. M. Passino // IEEE Control System Magazine. – 2002. – № 3 (22). – P. 52–67.
12. Madigan M. T. Biology of Microorganisms / M. T. Madigan, J. M. Martinko, J. Parker. – New Jersey : Prentice Hall, 1997. – 1036 p.
13. Berg H. C. Random Walks in Biology / H. C. Berg. – Princeton : Princeton University Press, 1993. – 164 p.
14. Tripathy M. Transmission Loss Reduction Based on FACTS and Bacteria Foraging Algorithm / M. Tripathy, S. Mishra, L. L. Lai, Q. P. Zhang // Proceedings of 9th International Conference of Parallel Problem Solving from Nature. – Heidelberg : Springer, 2006. – P. 222–231.
15. Analysis And Design Of Intelligent Systems Using Soft Computing Techniques / Ed. Melin P., Castillo O. R., Ramirez E. G., Kacprzyk J. – Heidelberg : Springer, 2007. – 855 p.
16. Mishra S. A hybrid least square-fuzzy bacterial foraging strategy for harmonic estimation / S. Mishra // IEEE Transactions on Evolutionary Computation. – 2005. – № 1(9). – P. 61–73.
17. Kim D. H. Bacterial Foraging Based Neural Network Fuzzy Learning / D. H. Kim, C. H. Cho // Proceedings of the 2nd Indian International Conference on Artificial Intelligence (ICAI-2005). – Pune : ICAI, 2005. – P. 2030–2036.

Надійшла 18.02.2010

Олейник А. А.

МУЛЬТИАГЕНТНЫЙ МЕТОД С КОСВЕННОЙ СВЯЗЬЮ МЕЖДУ АГЕНТАМИ

Исследован метод оптимизации на основе мультиагентного подхода. Разработан мультиагентный метод с косвенной связью между агентами, предусматривающий выполнение дополнительных этапов отбора, скрещивания и мутации. Проведены эксперименты по решению задач оптимизации многомерных нелинейных функций, что позволило исследовать эффективность предложенного метода.

Ключевые слова: агент, генетические операторы, мультиагентный метод, оптимизация, целевая функция.

Oliiynk A. O.

MULTIAGENT METHOD WITH INDIRECT COUPLING BETWEEN AGENTS

The optimization method based on multiagent approach has been investigated. A multiagent method with indirect coupling between agents is developed, providing execution of additional stages of selection, crossover and mutation. Experiments have been made aimed at solving the problems of multidimensional nonlinear functions optimization, which made it possible to evaluate efficiency of the proposed method.

Key words: agent, genetic operators, multiagent method, optimization, efficiency function.

УДК 621.315.592

Таланин В. И.¹, Таланин И. Е.², Семикина М. Ю.³

¹Д-р физ.-мат. наук, заведующий кафедрой Классического частного университета (г. Запорожье)

²Канд. физ.-мат. наук, профессор Классического частного университета (г. Запорожье)

³Старший преподаватель Классического частного университета (г. Запорожье)

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ОБРАЗОВАНИЯ ПЕРВИЧНЫХ РОСТОВЫХ МИКРОДЕФЕКТОВ В БЕЗДИСЛОКАЦИОННЫХ МОНОКРИСТАЛЛАХ КРЕМНИЯ

Представлена математическая модель и компьютерная программа расчета для процесса образования первичных ростовых микродефектов в бездислокационном монокристалле кремния на основе диссоциативного процесса диффузии.

Ключевые слова: ростовые микродефекты, бездислокационный монокристалл кремния, первичные ростовые дефекты, гетерогенный механизм дефектообразования.

ВВЕДЕНИЕ

Образующиеся в процессе роста бездислокационного монокристалла кремния ростовые микродефек-

ты определяют не только исходную дефектную структуру монокристаллов, но и последующие процессы трансформации этой структуры в результате

технологических воздействий. Поэтому исследование механизма образования таких дефектов помогает решить проблему управления дефектной структурой кристалла и проблему фундаментальных взаимодействий точечных дефектов.

На сегодняшний день подходы к решению проблемы дефектообразования в монокристалле кремния можно разделить на две группы: теоретические и экспериментальные. К первой группе относится широко распространенная модель динамики точечных дефектов [1, 2], основой которой является модель В. Воронкова [3, 4]. Основа модели Воронкова – признание критичной роли процесса «быстрой рекомбинации» собственных точечных дефектов около фронта кристаллизации. Модель динамики точечных дефектов предполагает образование в процессе охлаждения кристалла в узком температурном интервале ниже 1200 °С междузельных дислокационных петель и вакансионных микропор в различных областях кристалла в зависимости от температурных условий роста. Кроме того, предполагается, что образование всех видов микродефектов имеет гомогенный характер и игнорируется процесс взаимодействия собственных точечных дефектов с атомами примеси [2]. В модели динамики точечных дефектов неявно предполагается, что в зависимости от соотношения V/G (V – скорость роста кристалла, G – осевой температурный градиент) образующиеся в температурном интервале 1420...1200 °С дефекты представляют собой либо малые вакансионные микропоры, либо малые междузельные дислокационные петли [3].

Ко второй группе относятся подходы, связанные с экспериментальными исследованиями физической природы и характеристик ростовых микродефектов [5]. На основании экспериментальных исследований был построен качественный гетерогенный механизм образования и трансформации ростовых микродефектов, который базируется на следующих положениях [6]:

- рекомбинация собственных точечных дефектов вблизи температуры плавления ничтожна из-за наличия рекомбинационного барьера;

- распад пересыщенного твердого раствора собственных точечных дефектов во время охлаждения кристалла проходит по двум механизмам: вакансионному и междузельному;

- фоновые примеси кислорода и углерода принимают непосредственное участие в образовании и трансформации ростовых микродефектов.

Основные положения гетерогенного механизма находятся в противоречии с теоретическими положениями

модели Воронкова. Недостатком гетерогенного механизма является отсутствие теоретической (математической) модели образования дефектов как в температурном интервале 1420...1200 °С, так и в температурном интервале 1200...900 °С.

Целью настоящей работы является обоснование теоретической модели гетерогенного механизма образования ростовых микродефектов как результата взаимодействия «примесь – собственный точечный дефект».

Различия в физической сущности образования первичных и вторичных ростовых микродефектов дают возможность рассматривать дефектную структуру бездислокационных монокристаллов кремния в рамках гетерогенного механизма образования и трансформации ростовых микродефектов как структуру, состоящую из двух подсистем: подсистемы первичных ростовых микродефектов (включает в себя взаимодействие «примесь – собственный точечный дефект») и подсистемы вторичных ростовых микродефектов (включает в себя взаимодействие «собственный точечный дефект – собственный точечный дефект») [7].

Исходя из анализа кинетических процессов взаимодействий примесей, в качестве математической модели образования подвижных комплексов можно выбрать приближение сильного комплексообразования модели последовательной диффузии, которая допускает простую физическую интерпретацию и наиболее адекватна рассматриваемой физической модели. В условиях гетерогенного механизма образования и трансформации ростовых микродефектов речь может идти о модели последовательной диффузии, когда собственные точечные дефекты можно рассматривать как компонент, находящийся в начальный момент времени в объеме образца, а фоновые примеси – как компонент на границе образца и вводимый в образец. При этом не делается различия между диффузией примеси вглубь с поверхности образца и выделением из пересыщенного твердого раствора [8].

1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Математическая модель процесса образования первичных ростовых микродефектов была сформулирована на основе диссоциативного процесса диффузии примесей. Полное решение задачи о последовательной диффузии компонента A в образец, однородно легированный компонентом B с учетом комплексообразования $A + B \leftrightarrow AB$, было приведено в работе [9].

Искомое уравнение диффузии рассматривалось в виде

$$\left. \begin{aligned} \frac{1}{2} \left[(N_A - N_B - k + \sqrt{k^2 + 2k(N_A + N_B) + (N_A - N_B)^2}) \right]' + 2\lambda N_A' &= 0 \\ \frac{1}{2} \left[(N_B - N_A - k + \sqrt{k^2 + 2k(N_A + N_B) + (N_A - N_B)^2}) \right]' + 2\lambda d^2 N_B' &= 0 \end{aligned} \right\} (1)$$

где $\lambda = \frac{x}{2\sqrt{D_A \cdot t}}$ – подстановка Больцмана, штрих обозначает дифференцирование по λ , $d^2 = D_A/D_B$,

D_A, D_B – коэффициенты диффузии компонентов; N_A, N_B – полные концентрации компонентов; $k(T)$ – константа реакции комплексообразования. Начальные и граничные условия:

$$\begin{aligned} N_A(x, 0) &= 0; \\ N_B(x, 0) &= N_B(\infty); \\ N_A(0, t) - Q(0, t) &= H_A(0); \\ \frac{\partial}{\partial x} [N_B(x, t) - Q(x, t)]_{x=0} &= 0. \end{aligned}$$

Решение задачи диффузии компонента A в полубесконечный образец, однородно легированный компонентом B , при условии отсутствия испарения компонента B из образца и наличия свободного компонента A на границе образца языком полных компонентов имеет вид:

$$H_A = N_A - Q = \begin{cases} (N_A(0) - N_{B1}) \left[1 - \frac{\operatorname{erfc}(\lambda/d_A)}{\operatorname{erfc}(\lambda_0/d_A)} \right], & \lambda \leq \lambda_0; \\ 0, & \lambda > \lambda_0, \end{cases} (2)$$

$$H_B = N_B - Q = \begin{cases} 0, & \lambda \leq \lambda_0; \\ N_B(\infty) \left[1 - \frac{\operatorname{erfc}(\lambda/d_B)}{\operatorname{erfc}(\lambda_0/d_B)} \right], & \lambda > \lambda_0, \end{cases} (3)$$

$$Q = \begin{cases} N_{B1}, & \lambda \geq \lambda_0; \\ N_{B1} \frac{\operatorname{erfc}(\lambda/d_Q)}{\operatorname{erfc}(\lambda_0/d_Q)}, & \lambda < \lambda_0, \end{cases} (4)$$

$$\begin{aligned} N_{B1} T(\lambda_0/d_Q) &= N_B(\infty) T(\lambda_0/d_B); \\ N_{B1} T(\lambda_0/d_Q) &= \{N_A(0) - N_{B1}\} S(\lambda_0/d_A), \end{aligned} (5)$$

где $T(x) = \frac{\exp(-x^2)}{\sqrt{\pi x} \operatorname{erfc}(x)}$; $S(x) = \frac{\exp(-x^2)}{\sqrt{\pi x} \operatorname{erf}(x)}$; $N_A(0) - N_{B1} = H_A(0)$.

В условиях физической модели (гетерогенного механизма образования ростовых микродефектов) было принято, что компонент A является фоновой примесью (кислорода или углерода), а компонент B – собственные точечные дефекты (вакансии или собственные междуузельные атомы). Для вакансионного механиз-

ма рассматривалось взаимодействие «кислород + вакансия» ($O + V$), для междуузельного механизма – «углерод + собственный междуузельный атом» ($C + I$) для нелегированных бездислокационных монокристаллов FZ-Si и CZ-Si. В расчетах были использованы следующие величины: $H_A(0) \text{ (см}^{-3}\text{)} = H_O(0) = 4 \cdot 10^{15}$ (FZ-Si); $H_A(0) \text{ (см}^{-3}\text{)} = H_O(0) = 8 \cdot 10^{16}$ (CZ-Si); $C_B \text{ (см}^{-3}\text{)} = C_V = 8,84 \cdot 10^{14}$; $D_A \text{ (см}^2\text{/с)} = D_O = 0,17 \exp(-2,54/kT)$; $D_B \text{ (см}^2\text{/с)} = D_V = 4 \cdot 10^{-5}$; $H_A(0) \text{ (см}^{-3}\text{)} = H_C(T_m) = 4 \cdot 10^{15}$ (FZ-Si); $H_A(0) \text{ (см}^{-3}\text{)} = H_C(0) = 1 \cdot 10^{16}$ (CZ-Si); $C_B \text{ (см}^{-3}\text{)} = C_I = 6,31 \cdot 10^{14}$; $D_A \text{ (см}^2\text{/с)} = D_C = 1,9 \exp(-3,1/kT)$; $D_B \text{ (см}^2\text{/с)} = D_I = 4,75 \cdot 10^{-4}$.

Приближенные аналитические выражения согласуются с гетерогенным механизмом образования ростовых микродефектов, и используемый математический аппарат позволяет учитывать и анализировать взаимодействие собственных точечных дефектов не только с фоновыми примесями кислорода и углерода, но и с другими примесями (переходных металлов, азота, водорода и др.), а также взаимодействия типа «примесь – примесь» [10].

2. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Первым этапом эксперимента является процесс формирования исходных данных для его проведения. Исходные данные, которые при необходимости могут изменяться, хранятся во внешнем файле. Константы, количество которых невелико, прописаны в самой программе. Затем поток данных направляется в блок расчета, где, в соответствии с вышеизложенными правилами обработки исходных значений, формируется результат расчета. Завершающим этапом работы программы является визуальное представление полученного результата и его дальнейший вывод. Вывод организован таким образом, что в одно окно формы выводятся все полученные данные в текстовом виде, а в другое – те же данные, только в графическом виде.

Минимальные системные требования для работы с программой:

- а) процессор – 700 МГц;
- б) оперативная память – 128 Мб;
- в) платформа – Microsoft .NET Framework 2.0;
- г) операционная система – Microsoft Windows XP/Vista.

В качестве среды разработки была использована среда Dev-C++ версии 4.9.8.0. Такой выбор обусловлен простотой самой среды, а также наличием в математической модели функций erf и erfc (2)–(5). Именно в среде Dev-C++ сложность реализации (синтаксиса) этих функций сведена к минимуму.

Функциональная часть главной формы сочетает в себе модуль ввода и проверки. При запуске программы появляется форма Form1, на которой располагаются меню, поля ввода исходных данных и кнопка «Расчитать». Главное меню содержит в себе такие подменю, как «Выход», «Помощь» и «О программе». Меню «Выход» завершает работу программы. При нажатии на кнопку меню «Помощь» открывается форма Help, а при нажатии кнопки меню «О программе» – форма About. Форма Help представляет собой инструкцию пользователя, вторая же содержит информацию о программе и ее разработчиках.

Выбор метода выращивания для кристалла кремния осуществляется за счет изменения содержимого области «Метод выращивания». При этом меняются и значения исходных данных в соответствии с выбранным методом. Эти значения устанавливаются программой по умолчанию и при желании их можно изменить.

На рис. 1 представлен интерфейс программы, а также исходные данные программы. Результат работы программы представлен на рис. 2.

При нажатии на кнопку «Расчитать» исходные данные проходят проверку на правильность ввода. Проверяется наличие данных в полях формы и соответствие их ограничениям (например, скорость роста при методе выращивания – БЗП должна быть в диапазоне от 1 до 9 мм/мин). При успешном прохождении проверки значения параметров передаются в форму Result для дальнейшего расчета. Если хотя бы одно значение введено неверно, то прохождение про-

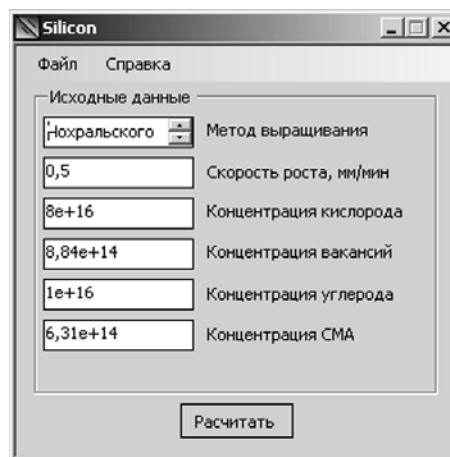


Рис. 1. Интерфейс программы Silicon

верки будет проведено повторно, при этом сопровождаясь сообщением об ошибке соответствующего параметра.

На первом этапе расчетной части программы объявляются все необходимые переменные, а также производится расчет параметров, которые носят статический характер. Алгоритм расчета построен в соответствии с гетерогенным механизмом образования микродефектов в кремнии и разделен на две ветви: вакансионную и межузельную. Первым производится расчет для вакансионной ветви. Параметр Choice определяет, в соответствии с каким методом выращивания монокристаллов кремния будет произведен расчет. Если значение этого параметра равно единице, то расчет производится для метода бестигельной

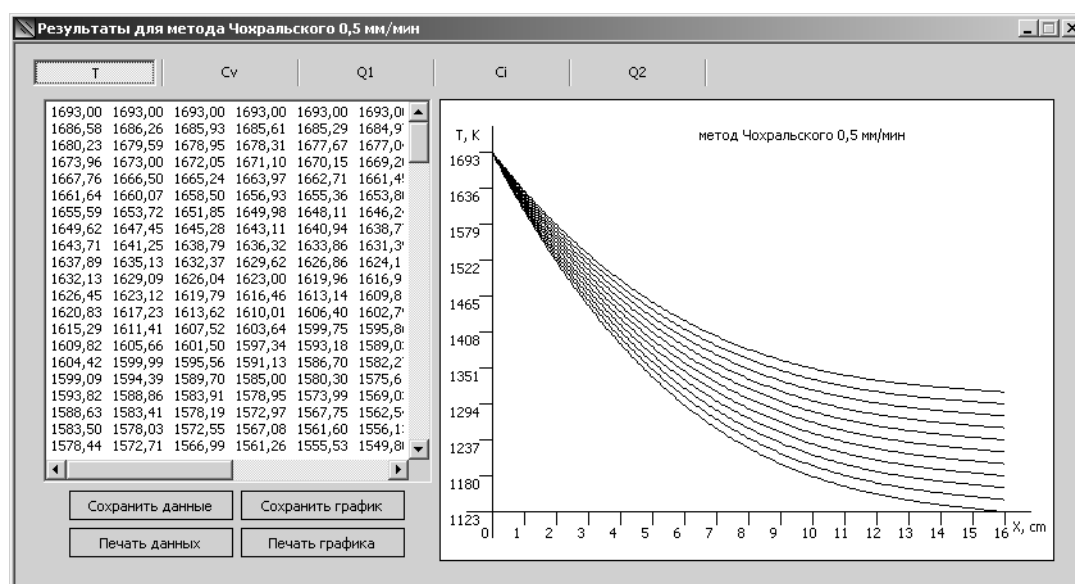


Рис. 2. Вывод результатов в программе

зонной плавки, в противном случае – для метода Чохральского. Необходимости в условии равенства параметра Choice двойке или любому другому значению, отличному от единицы, нет, так как еще на форме Form1 этот параметр может принять одно из двух значений, иначе не пройдет этап проверки.

Следующим шагом является главный цикл, в котором рассчитываются динамические величины, зависящие от длины кристалла и радиуса. Также в теле цикла производится вывод рассчитанных данных на вкладке формы Result. Для того, чтобы при построении графических зависимостей не возвращаться к расчетам, определены динамические массивы.

Далее следует расчетная часть для междоузельной ветви, во многом идентичная вакансионной ветви. Сначала определяются и рассчитываются параметры статического характера, соответствующие данной ветви гетерогенного механизма, затем запускается главный цикл. Главный цикл также содержит механизмы для вывода расчетных характеристик на вкладки формы Result.

За расчетами следует этап построения графических зависимостей. Сначала строятся кривые температурных характеристик $T(x)$, затем концентраций вакансий $C_V/C_V(0)$, комплексов $(O + V) C_{O1}/C_{O1}(0)$, концентраций собственных междоузельных атомов $C_I/C_I(0)$ и, наконец, комплексов $(C + I) C_{C2}/C_{C2}(0)$.

ВЫВОДЫ

Сформулирована математическая модель процесса образования первичных ростовых микродефектов на основе диссоциативного процесса диффузии. Рассмотрены случаи взаимодействия «кислород – вакансия» $(O + V)$ и «углерод – собственный междоузельный атом кремния» $(C + I)$ вблизи фронта кристаллизации для бездислокационных монокристаллов кремния, полученных методами бестигельной зонной плавки и Чохральского. Полученные приближенные аналитические выражения согласуются с гетерогенным механизмом образования ростовых микродефектов.

Разработано приложение для программной реализации математической модели образования первичных ростовых микродефектов в бездислокационных монокристаллах кремния. Проведен численный расчет модели, подтвердивший основные положения гетерогенного механизма образования микродефектов.

Разработанный программный продукт предполагает высокую точность и скорость расчетов, а также представление данных в удобочитаемом виде. Основным преимуществом является совместимость формата импортируемого файла с программным комплексом Microsoft Office Excel. Это дает возмож-

ность накопления данных и представления их в графическом виде. Как следствие, такая база данных может послужить хорошей платформой для дальнейшего анализа.

СПИСОК ЛИТЕРАТУРЫ

1. *Kulkarni M. S.* Quantification of defect dynamics in unsteady-state and steady-state Czochralski growth of monocrystalline silicon / Kulkarni M.S., Voronkov V., Falster R. // J. Electrochem. Soc. – 2004. – Vol. 151. – No. 5. – P. G663–G669.
2. *Sinno T.* Modeling microdefect formation in Czochralski silicon / Sinno T., Brown R.A., Von Ammon W., Dornberger E. // J. Electrochem. Soc. – 1998. – Vol. 145, № 1. – P. 302–318.
3. *Voronkov V. V.* Mechanism of swirl defects formation in silicon / Voronkov V.V. // J. Cryst. Growth. – 1982. – Vol. 59, № 3. – P. 625–642.
4. *Voronkov V. V.* Nucleation of oxide precipitates in vacancy-containing silicon / Voronkov V.V., Falster R. // J. Electrochem. Soc. – 2002. – Vol. 149. – P. G167.
5. *Talanin V. I.* Nucleation, growth and transformation of microdefects in FZ-Si / Talanin V. I., Talanin I. E. // New Research on Semiconductors / Ed. T. B.Elliot. – New York : Nova Sci. Publ., 2006. – P. 35–59.
6. *Talanin V. I.* Classification of microdefects in semiconducting silicon / Talanin V. I., Talanin I. E. // Defect & Diffusion Forum. – 2004. – Vol. 230–232, № 1. – P. 177–198.
7. *Talanin V. I.* Моделирование и свойства дефектной структуры бездислокационных монокристаллов кремния / Таланин В. И. // Дні науки. Збірник тез доповідей. – Запоріжжя: ЗИГМУ, 2007. – 275 с.
8. *Talanin V. I.* Взаимодействие точечных дефектов в процессе роста бездислокационных монокристаллов кремния / Таланин В. И. // Известия вузов. Материалы электронной техники. – 2007. – № 4. – С. 27–40.
9. *Булярский С. В.* Термодинамика и кинетика взаимодействующих дефектов в полупроводниках / Булярский С. В., Фистуль В. И. – М. : Наука, 1997. – 352 с.
10. *Talanin V. I.* About formation of grown-in microdefects in dislocation-free silicon single crystals / Talanin V. I., Talanin I. E., Voronin A. A. // Can. J. Phys. – Vol. 85, № 12. – P. 1459–1471.

Надійшла 22.02.2010
Після доробки 23.03.2010

Таланін В. І., Таланін І. Є., Семікіна М. Ю.

КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ УТВОРЕННЯ ПЕРВИННИХ РОСТОВИХ МІКРОДЕФЕКТІВ У БЕЗДИСЛОКАЦІЙНОМУ МОНОКРИСТАЛІ КРЕМНІЮ

Представлено математичну модель і комп'ютерну програму розрахунку для процесу утворення первинних ростових мікродфектів у бездислокаційному монокристалі кремнію на основі дисоціативного процесу дифузії.

Ключові слова: ростові мікродфекти, бездислокаційний монокристал кремнію, первинні ростові мікродфекти, гетерогенний механізм дефектоутворення.

Talanin V. I., Talanin I. E., Semikina M. Yu.

COMPUTER MODELING OF PRIMARY GROW-IN MICRODEFECTS FORMATION IN DISLOCATION-FREE SILICON MONOCRYSTALS

The mathematical model and computer program for primary growing microdefects formation process in a dislocation-free silicon monocrystal based on a dissociation diffusion process are represented in this article.

Key words: grow-in microdefects, dislocation-free silicon monocrystal, primary grow-in microdefects, heterogeneous defect formation mechanism.

¹Канд. техн. наук, докторант Университета гражданской защиты Украины МЧС Украины (г. Харьков)²Д-р физ.-мат. наук, заведующая кафедрой Харьковского государственного технического университета строительства и архитектуры МОН Украины

ОПРЕДЕЛЕНИЕ НАПРАВЛЕНИЯ СПУСКА В ЛИНЕАРИЗОВАННОЙ ЗАДАЧЕ РАЗМЕЩЕНИЯ НЕОРИЕНТИРОВАННЫХ ГЕОМЕТРИЧЕСКИХ ОБЪЕКТОВ

Проведено исследование оптимизационной задачи размещения многоугольных неориентированных объектов в полосе, рассматриваемой в линейризованной постановке. Выделены дополнительные свойства области допустимых решений, предложена схема метода поиска локального минимума функции цели. Рассмотрена схема определения направления движения по границе компоненты связности линейризованной области допустимых решений.

Ключевые слова: размещение неориентированных многоугольников, линейная аппроксимация, область допустимых решений, локальный минимум.

ВВЕДЕНИЕ

Рассмотрим задачу оптимального размещения конечного набора неориентированных многоугольных геометрических объектов в заданной многоугольной области. Данная задача принадлежит к классу задач оптимизационного геометрического проектирования [1–3] и представляет интерес как с теоретической, так и с практической точек зрения. К задачам такого рода относятся, например, задачи раскроя изотропного материала (металлопрокат, ткань, стекло, пластмасса и т. д.) на многоугольные заготовки. При этом, если над геометрическим объектом допустимо аффинное преобразование поворота, то такой геометрический объект называют неориентированным.

1. ПОСТАНОВКА ЗАДАЧИ

Имеется набор $R = \{R_i\}$, $i = \overline{1, N}$ выпуклых неориентированных многоугольных объектов, заданных в арифметическом евклидовом пространстве E^2 , и область размещения $R_0 = \{(x, y) \in E^2 | x \in [0, z], y \in [0, W], W = \text{const}, z = \text{var}\}$ – полубесконечная полоса. Объект R_i задается упорядоченным набором $\{(x_i^n, y_i^n)\}$, $n = \overline{1, n_i}$ координат его вершин в собственной системе координат $X_i O_i Y_i$, $i = \overline{1, N}$ (рис. 1).

Положение объекта R_i в общей системе координат XOY , связанной с областью R_0 , определяется вектором параметров размещения $u_i = (v_i, \varphi_i) =$

$= (x_i, y_i, \varphi_i)$, задающим начало O_i системы координат $X_i O_i Y_i$, при этом $O_i \in \text{int } R_i$.

Компонента $v_i = (x_i, y_i)$ задает трансляцию O_i , а угол φ_j – поворот системы координат $X_i O_i Y_i$ относительно O_i .

Необходимо разместить набор R_i , $i = \overline{1, N}$ объектов без взаимных наложений в области R_0 так, чтобы длина занятой части полосы z была минимальной.

Математическая модель оптимизационной задачи размещения имеет вид:

$$\text{найти: } u^* = \arg \min_{u \in D \subset E^{3N+1}} cu, \quad (1)$$

где $u = (u_1, u_2, \dots, u_N, z) = (x_1, y_1, \varphi_1, x_2, y_2, \varphi_2, \dots, x_N, y_N, \varphi_N, z)$; $c = (0, 0, \dots, 0, 1)$; $D = D_1 \cap D_2$ – область допустимых решений задачи; подобласть $D_1 \subset E^{3N+1}$ определяется ограничениями на размещение объектов R_i в полосе R_0 :

$$R_i \subset R_0, \quad i = \overline{1, N}; \quad (2)$$

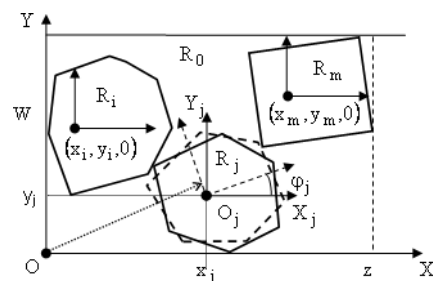


Рис. 1. Область размещения R_0 с размещаемыми объектами R_i, R_j, R_m

$D_2 \subset E^{3N+1}$ задається умовами взаємного попарного перетинання об'єктів (R_i, R_j) :

$$\text{int } R_i \cap \text{int } R_j = \emptyset, \quad i, j = \overline{1, N}, \quad i \neq j. \quad (3)$$

Аналитическое описание условия (2) задается системой $F_l(u) \geq 0 := \{f_i^{hn}(u_i) \geq 0, n = \overline{1, n_i}; h = \overline{1, 4}; i = \overline{1, N}$ нелинейных неравенств, причем функции $f_i^{hn}(u_i)$ имеют вид:

$$f_i^{hn}(u_i) \in \{x_i - \xi_i^n; y_i - \psi_i^n; -y_i + W - \psi_i^n; z - x_i - \xi_i^n\}, \quad (4)$$

где $\xi_i^n = |x_i^n \cos \varphi_i + y_i^n \sin \varphi_i|$; $\psi_i^n = |y_i^n \cos \varphi_i - x_i^n \sin \varphi_i|$.

Условие (3) задается системой $F_2(u) \geq 0$ наборов $F_{ij}^h(u_i, u_j) \geq 0$ нелинейных неравенств вида:

$$F_{ij}^h(u_i, u_j) \geq 0; \quad f_{ij}^{hl}(u_i, u_j) \geq 0, \\ h = \overline{1, n_i}, \quad l = \overline{1, n_j}, \quad i, j = \overline{1, N}, \quad i \neq j.$$

При этом функции f_{ij}^{hl} в зависимости от типа касания [3] пары (R_i, R_j) имеют вид:

$$(A \cos \varphi_i + B \sin \varphi_i)(x_j - x_i) + (B \cos \varphi_i - A \sin \varphi_i)(y_j - y_i) + \\ + (A x_j^k + B y_j^k) \cos(\varphi_i - \varphi_j) - (B x_j^k - A y_j^k) \sin(\varphi_i - \varphi_j) - \\ - (B y_i^l + A x_i^l) \quad (5)$$

или

$$(A \cos \varphi_j + B \sin \varphi_j)(x_j - x_i) + (B \cos \varphi_j - A \sin \varphi_j)(y_j - y_i) - \\ - (A x_i^l + B y_i^l) \cos(\varphi_j - \varphi_i) + (B x_i^l - A y_i^l) \sin(\varphi_j - \varphi_i) + \\ + (B y_j^k + A x_j^k), \quad (6)$$

где $x_i^l, y_i^l, x_j^k, y_j^k$ – координаты вершин объектов R_i, R_j ; A, B – функции координат их вершин.

Данная задача является многомерной многоэкстремальной задачей нелинейного невыпуклого программирования со специфичной областью допустимых решений, что затрудняет или делает невозможным применение классических методов условной оптимизации [4]. Эффективные точные методы решения задач практической размерности отсутствуют.

Актуальность рассматриваемого класса задач обуславливает необходимость создания программного обеспечения процесса решения, основанного на современном математическом инструментарии.

В статье [5] предложена методика линейной аппроксимации области D и построено преобразование $\mathfrak{S}(D, \varepsilon): D \xrightarrow{\mathfrak{S}} D^L$, такое, что его применение к нелинейным функциям ограничений задачи (1) с точностью $\varepsilon > 0$ продуцирует аппроксимационное

множество D^L (соответственно $D_1 \xrightarrow{\mathfrak{S}} D^{L_1}, D_2 \xrightarrow{\mathfrak{S}} D^{L_2}$) с кусочно-линейной границей [5].

Целью настоящей статьи является выделение дополнительных свойств линеаризованной области D^L допустимых решений и построение метода поиска локального минимума функции цели задачи, основанного на использовании идеи активного набора [5].

2. ОБЩАЯ СХЕМА АКТИВНОГО НАБОРА

Приведем интерпретацию общей схемы активного набора в применении к задаче линейного программирования (LP) вида:

$$\text{найти } \min_{u \in D \subset E^n} c^T u, \quad (7)$$

$$A u \geq b,$$

где A есть $m \times n$ – матрица коэффициентов, $m > n$, $A = (a_j), j = \overline{1, m}$.

В схеме активного набора $(h + 1)$ -я итерация имеет вид [5]

$$u^{h+1} = u^h + \Delta u^h \cdot p^h, \quad (8)$$

где u^h – значение на предыдущей итерации; Δu^h – шаг, p^h – направление спуска.

Активным набором называют те ограничения задачи LP (7), которые обращаются в текущей точке u^h в равенства: $a_j^T u^h = b_j$.

Рабочим списком называют те ограничения активного набора, которые учитываются при расчете направления поиска (не все обращающиеся в равенства в точке u^h ограничения задачи (7) попадут в рабочий список). Обозначим I – множество индексов ограничений рабочего списка, $|I| = n$.

Для задачи LP (7) можно организовать движение к оптимуму u^* так, что число ограничений $a_j^T u^h = b_j$ в рабочем списке на каждой итерации будет совпадать с числом переменных: $\bar{A} u^h = \bar{b}, \bar{A} = [\bar{a}_{ij}]_{n \times n}$, причем матрица \bar{A} ограничений невырожденная, то есть u^h – вершина D [5]. В точке u^h по схеме активного набора определяются знаки компонент вектора множителей Лагранжа λ . Последний представляет собой решение невырожденной системы линейных уравнений

$$\bar{A}^T \lambda = c. \quad (9)$$

Если все компоненты $\lambda \geq 0$, то u^h есть решение LP. Если найдется отрицательная (например, $\lambda_s < 0$), то целевую функцию можно уменьшить, двигаясь по направлению, вдоль которого левая часть s -го огра-

ничення возрастает (исключается из рабочего списка), а остальные ограничения рабочего списка остаются равенствами (левые части таких ограничений не изменяются: $a_j^{-T} p = 0 \quad j \neq s$).

Направление p определяется как решение линейной системы вида

$$\bar{A}p = e_s, \quad (10)$$

где через e_s обозначен s -й столбец единичной матрицы, т. е. p – s -й столбец матрицы \bar{A}^{-1} .

3. ИССЛЕДОВАНИЕ СВОЙСТВ ОБЛАСТИ D^L

Аппроксимационное многогранное множество D^{L_1} в диапазоне изменения углового параметра $\varphi_i \in [-0,36; 0,36]$ с заданной погрешностью ϵ_{\cos} [5] задается системой $F_1^L(u) \geq 0 = \{\mathfrak{S}_i^{hn}(u_i) \geq 0, n = \overline{1, n_i}; h = \overline{1, 4}; i = \overline{1, N}$ наборов $\mathfrak{S}_i^{hn}(u_i) \geq 0$ линейных неравенств. При этом функции набора $\mathfrak{S}_i^{lj}(u_i)$ имеют вид:

$$\begin{aligned} \mathfrak{S}_i^{ln} &\approx x_i - x_i^n g^L(\varphi_i) - y_i^n \varphi_i = \\ &= \langle -x_i + x_i^n g_k^L(\varphi_i) + y_i^n \varphi_i, \end{aligned} \quad (11)$$

где $g_k^L(\varphi_i) = a_k^i \varphi_i + b_k^i$ (при $k = \overline{1, 3}$ коэффициенты a_k^i, b_k^i приведены в табл. 1).

Аналогично записываются аппроксимации $\mathfrak{S}_i^{hn}(u_i)$ функций f_i^{hn} , $h = \overline{2, 4}$.

Многогранное множество D^{L_2} задается системой $F_2^L(u) \geq 0$ наборов $F_{ij}^L(u_i, u_j) \geq 0$ линейных неравенств вида:

$$\begin{aligned} F_{ij}^L(u_i, u_j) \geq 0 &:= \langle \mathfrak{S}_{ij}^{hl}(u_i, u_j) \geq 0, \\ h = \overline{1, n_i}, l = \overline{1, n_j}, i, j = \overline{1, N}, i \neq j. \end{aligned} \quad (12)$$

При этом набор $\mathfrak{S}_{ij}^{hl}(u_i, u_j) \geq 0 = \langle \xi_{ij}^{hl}(u_i, u_j) \geq 0$ является линейной аппроксимацией функций $f_{ij}^{hl_1}$ вида (4), (5). Последние с учетом применения функции $g_k^L(\varphi_i) = a_k^i \varphi_i + b_k^i$ могут быть выписаны в виде

$$\begin{aligned} C_{1k}^i \varphi_i x_j - C_{1k}^i \varphi_i x_i + C_{2k}^i \varphi_i y_j - C_{2k}^i \varphi_i y_i + C_1^j \varphi_i \varphi_j + \\ + b_k^i (Ax_j - Ax_i + By_j - By_i) + \varphi_i (C_1^j a_k^i + C_2^j) + \\ + \varphi_j (C_1^j a_h^j - C_2^j) + C_1^j (b_k^i + b_h^j - 1) - C_i^k, \end{aligned} \quad (13)$$

где $C_{1k}^i = B + Aa_k^i$; $C_{2k}^i = Ba_k^i - A$; $C_1^j = Ax_j^1 + By_j^1$; $C_2^j = (Bx_j^1 - Ay_j^1)$; $C_i^k = (By_i^k + Ax_i^k)$.

Кусочно-линейная аппроксимация Γ_v^L функций $\Gamma_v = \varphi_i \omega_v$, $\omega_v \in \{x_j, x_i, y_j, y_i, \varphi_j\}$, $v = \overline{1, 5}$ в заданных диапазоне изменения параметров осуществляется с помощью набора T линейных функций вида:

$$A_{vt} \omega_v + B_{vt} \varphi_i + D_{vt}, \quad t = \overline{1, T}.$$

Тогда функция $f_{ij}^{hl_1}$ имеет вид:

$$\begin{aligned} (C_{1k}^i A_{1t} + b_k^i A) x_j - (C_{1k}^i A_{2t} + b_k^i A) x_i + \\ + (C_{2k}^i A_{3t} + b_k^i B) y_j - (C_{2k}^i A_{4t} + b_k^i B) y_i + \\ + (C_1^j B_{5t} + C_1^j a_h^j - C_2^j) \varphi_j + (C_{1k}^i B_{1t} - C_{1k}^i B_{2t} + \\ + C_{2k}^i B_{3t} - C_{2k}^i B_{4t} + C_1^j A_{5t} + C_1^j a_k^i + C_2^j) \varphi_i + D, \end{aligned}$$

где $D = C_{1k}^i D_{1t} - C_{1k}^i D_{2t} + C_{2k}^i D_{3t} - C_{2k}^i D_{4t} + C_1^j D_{5t} + C_1^j (b_k^i + b_h^j - 1) - C_i^k$.

Отметим, что одноименные (относительно t) коэффициенты A_{vt}, B_{vt}, D_{vt} , $v = \overline{1, 4}$ равны для всех t .

Таким образом, задача (1) приобретает вид:

$$\text{найти } u^* = \arg \min_{u \in D^L \subset E^{3N+1}} z. \quad (14)$$

Приведем основные свойства области D^L , необходимые для дальнейших построений.

Свойство 1. Область D^L – несвязное ограниченное точечное множество, имеющее кусочно-линейную границу $\Psi^L = Fr D^L$, $\Psi \subset E^{3N+1}$:

$$D^L = \bigcup_{q=1}^Q D_q^L,$$

где D_q^L – компонента связности области D^L .

Свойство 2. Функции ограничений (3), (4) являются невыпуклыми (обратно выпуклыми), следовательно, наборы вида (11), (12) описывают невыпуклые линейные многообразия.

Таблица 1

φ_i	b_k^i	a_k^i	φ_i	b_k^i	a_k^i
$[\varphi^1; \varphi^2] = [0; 0,12]$	1	-0,6	$[-\varphi^2; -\varphi^1] = [-0,12; 0]$	1	0,6
$[\varphi^2; \varphi^3] = [0,12; 0,24]$	1,0144	-0,18	$[-\varphi^3; -\varphi^2] = [-0,24; -0,12]$	1,0144	0,18
$[\varphi^3; \varphi^4] = [0,24; 0,36]$	1,0432	-0,3	$[-\varphi^4; -\varphi^3] = [-0,36; -0,24]$	1,0432	0,3

Свойство 3. Каждое неравенство вида (4), (5), становясь активным ограничением, определяет условие касания объекта границы области размещения или условие касания пары размещаемых объектов.

Свойство 4. Число ограничений S , описывающих область D^L задачи (1), квадратично зависит от числа размещаемых объектов и равно

$$S = S_1 + S_2 = O\left(4NK \max_{i=1, N} n_i\right) + O\left(KTN^2 \max_{i=1, N} n_i\right), \tag{15}$$

где S_1, S_2 – число ограничений, описывающих подобласти D^{L_1}, D^{L_2} соответственно, K, T – параметры, определяемые точностью аппроксимации.

Свойство 5. Каждая компонента связности D_q^L области D^L может быть представлена в виде объединения конечного числа G_q выпуклых многогранных подобластей $D_{q_g}^L$ вида

$$D_q^L = \bigcup_{g=1}^{G_q} D_{q_g}^L.$$

При этом границу области $D_{q_g}^L$ задает набор уравнений, содержащий, по крайней мере, по одному из ограничений каждого набора вида (11), (12).

Свойство 6. В силу линейности функций ограничений области D^L и функции цели (1) оптимальная точка u^* задачи (как и любое текущее приближение u^h) описывается системой $F^*u = 0$ $I = 3N + 1$ активных в точке u ограничений из систем $\{F_1^L(u) \geq 0, F_2^L(u) \geq 0\}$ наборов линейных неравенств. При этом система $F^h u = 0$ (особенно на начальных этапах алгоритма оптимизации) может содержать также уравнения вида $\varphi_i = (\pm)\varphi^k$ (табл. 1).

Свойство 7. Для точки u^h , включая u^* , справедливо соотношение

$$u^h \in \bigcap_{r>1} D_{q_r}^L,$$

где $D_{q_r}^L$ – выпуклая подобласть области D^L .

4. МЕТОД ПОИСКА НАПРАВЛЕНИЯ ДВИЖЕНИЯ (МПНД)

Общая схема метода такова. Пусть имеется допустимая начальная точка u^0 ($h = 0$), определяемая системой уравнений $F^h u = 0$.

Шаг 1. Проверка условий останова. Если выполнено условие $\lambda_i \geq 0, i = \overline{1, I}$, где λ_i – множители Лагранжа, определяемые по формуле

$$(F^h)^T \lambda = e_{3N+1},$$

то u^h – искомое решение.

Шаг 2. Расчет допустимого направления поиска p^h . Если при некотором $s \in I, \lambda_s < 0$, направление p^h есть s -й столбец матрицы $(F^h)^{-1}$.

Реализация данной схемы МПНД обладает рядом особенностей, учитывающих характер формирования области допустимых решений задачи (14).

4.1. Особенности определения направления поиска p^h

Пример 1. Пусть имеется объект R_1 с вершинами $\{(-5,3),(-3,-3),(2,1)\}$ и полоса $R_0, W = 10$. Начальная точка $u^0 = (x_1^0, y_1^0, \varphi_1^0, z_1^0) = (5, 3, 0, 7)$ (объект R_1^1 и его полюс Π_1 на рис. 2). Положим $u^0 \in D_{q_1}^L$. Система $F^0 u = b$, задающая точку u^0 (рабочий список), и система $F_{q_1} u \geq b_{q_1}$, описывающая выпуклое многогранное множество $D_{q_1}^L$ (рассматривается ослабленная задача без части ограничений, не нарушаемых в точке u^0), имеют соответственно вид:

$$F_{q_1} u^0 = b_{q_1} \Leftrightarrow \begin{cases} x_1 + 3, 3\varphi_1 = 5, \\ y_1 + 3, 18\varphi_1 = 3, \\ \varphi_1 = 0, \\ -x_1 - 0, 88\varphi_1 + z_1 = 2, \end{cases}$$

$$F_{q_1} u = b_{q_1} \Leftrightarrow \begin{cases} x_1 + 3, 3\varphi_1 \geq 5, \\ y_1 + 3, 18\varphi_1 \geq 3, \\ \varphi_1 \geq 0, \\ -x_1 - 0, 88\varphi_1 + z_1 \geq 2, \\ -\varphi_1 \geq 0, 12. \end{cases} \tag{16}$$

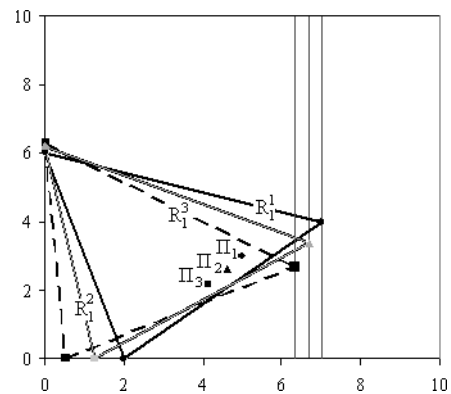


Рис. 2. Иллюстрация примера 1

Вектор множителей Лагранжа, вычисляемый как решение системы $(\bar{F}^0)^T \lambda = c$ (формула 9): $\lambda = (1; 0; -2,42; 1)$. Таким образом, из системы $F^0 u = b$ исключается третье уравнение. Вектор направления движения p для системы $F^0 u = b$, вычисленный по формуле (10), имеет вид: $p = (-3,3; -3,18; 1; -2,42)$. На рис. 2 приведены результаты двух следующих итераций: $R_1^2: u^1 = (4,604; 2,6184; 0,12; 6,7096)$ и $R_1^3: u^2 = (4,136; 2,1936; 0,24; 6,3184)$.

Замечание 1. Согласно Свойствам 2, 4 точка u^1 является единственной смежной вершиной областей $D_{q_1}^L$ и $D_{q_2}^L$, а вершина $u^2 \in D_{q_2}^L \cap D_{q_3}^L$ (при этом угол $\varphi_1 > 0$). Согласно (11) системе неравенств $F_{q_i} u \geq b_{q_i}$, задающие области $D_{q_i}^L$, $i = 2,3$, имеют вид:

$$F_{q_2} u \geq b_{q_2} \Leftrightarrow \begin{cases} x_1 + 3,9\varphi_1 \geq 5,07, \\ y_1 + 3,54\varphi_1 \geq 3,04, \\ \varphi_1 \geq 0,12, \\ -x_1 - 0,64\varphi_1 + z_1 \geq 2,029, \end{cases}$$

$$F_{q_3} u \geq b_{q_3} \Leftrightarrow \begin{cases} x_1 + 4,5\varphi_1 \geq 5,22, \\ y_1 + 3,9\varphi_1 \geq 3,13, \\ \varphi_1 \geq 0,24, \\ -x_1 - 0,4\varphi_1 + z_1 \geq 2,09, \\ -\varphi_1 \geq 0,3. \end{cases} \quad (17)$$

Замечание 2. Согласно Свойствам 2, 4 точка u^0 является единственной смежной вершиной областей $D_{q_1}^L$ и $D_{q_4}^L$, где коэффициенты системы ограничений $F_{q_4} u \geq b_{q_4}$ множества $D_{q_4}^L$ есть функции отрицательного угла поворота φ_1 (табл. 1):

$$F_{q_4} u = b_{q_4} \Leftrightarrow \begin{cases} x_1 + 2,7\varphi_1 = 5, \\ y_1 + 2,82\varphi_1 = 3, \\ -\varphi_1 = 0, \\ -x_1 - 1,12\varphi_1 + z_1 = 2, \end{cases}$$

$$F_{q_4} u \geq b_{q_4} \Leftrightarrow \begin{cases} x_1 + 2,7\varphi_1 \geq 5, \\ y_1 + 2,82\varphi_1 \geq 3, \\ -\varphi_1 \geq 0, \\ -x_1 - 1,12\varphi_1 + z_1 \geq 2. \end{cases}$$

Очевидно, в общем случае заранее неизвестно, какая из областей $-D_{q_1}^L$ или $D_{q_4}^L$ является перспективной для дальнейшего спуска. Поэтому задача определения направления движения p предполагает проверку знаков множителей Лагранжа как для ограничений рабочего списка $F_{q_4} u = b_{q_4}$ (для данной системы $\lambda = (1; 0; 1,58; 1)$), так и для ограничений рабочего списка $F_{q_1} u^v = b_{q_1}$. Другими словами, задача предполагает наличие перебора вариантов, что нарушает концепцию направленного спуска. На рис. 3 представлены проекции ограничений областей $D_{q_1}^L$ или $D_{q_4}^L$ на плоскость $X(Z)O\varphi$. Ряды 1x (1z) и 2x (2z) соответствуют 1-му (4-му) ограничениям области $D_{q_1}^L$ ($D_{q_4}^L$). Ряд 5 соответствует границе области $D_{q_1}^L$ (область $D_{q_1}^L$ отмечена штриховкой вида \diagup).

Утверждение 1. Для определения направления движения в невыпуклой области $D_{q_1}^L$ необходимо рассмотреть комплиментарную область $D_{q_{-compl}}^L$ (на рис. 3. область $D_{q_1}^L$ отмечена штриховкой вида \diagdown), которую формируют ограничения областей $D_{q_1}^L$ и $D_{q_4}^L$, взятые с обратным знаком.

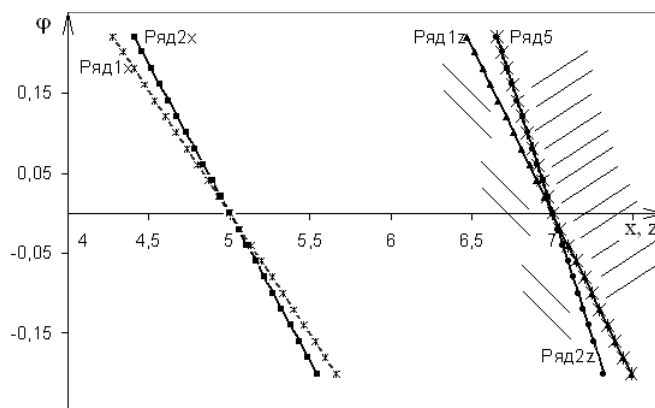


Рис. 3. Иллюстрация замечания 2

Доказательство Утверждения 1 непосредственно следует из факта осуществления спуска по границе области D_q^L , при этом $FrD_q^L = D_{q_compl}^L$ и свойства выпуклости области $D_{q_compl}^L$.

Замечание 3. Во всех рассмотренных случаях Примера 1 аналитические описания соответствуют одним и тем же геометрическим ограничениям (условиям касания).

4.2. Особенности проверки условий останова

Алгоритм проверки свойства оптимальности текущего приближения u^h в силу Свойств 6, 7 и Замечания 2 включает следующие шаги.

Шаг 1. Выделение набора выпуклых множеств $D_{q_r}^L, r = 1, 2, \dots, r_q$, (Свойство 7), вершиной которых является точка u^h и организация их перебора, i -я итерация которого состоит из двух составляющих:

Шаг i1. Построение на каждом из множеств $D_{q_r}^L$ двух систем уравнений, задающих данную точку u^h , при этом экзогенные коэффициенты систем уравнений вычисляются в зависимости от знака приращения углового параметра.

Шаг i2. Построение и проверка векторов множителей Лагранжа для двух систем уравнений, построенных на предыдущем шаге.

Шаг 2. Если для всех построенных систем наборы множителей Лагранжа содержат неотрицательные компоненты, то u^h – оптимальная точка.

Пример 2. На рис. 4 вектор $u_3 = (4,136; 2,1936; 0,24; 6,3184)$, определяющий положение объекта R_2^3 , задается следующими системами уравнений:

$$Fu^{3+} = b^3 \Leftrightarrow \begin{cases} x_1 + 1,72\varphi_1 = 4,06, \\ y_1 + 3,82\varphi_1 = 3,04, \\ -y_1 + 3,36\varphi_1 = -4,97, \\ -x_1 - 1,82\varphi_1 + z_1 = 3,04, \end{cases}$$

$$Fu^{3-} = b^3 \Leftrightarrow \begin{cases} x_1 + 0,28\varphi_1 = 4,06, \\ y_1 + 3,36\varphi_1 = 3,04, \\ -y_1 + 2,64\varphi_1 = -4,97, \\ -x_1 - 2,18\varphi_1 + z_1 = 3,04. \end{cases}$$

Векторы множителей Лагранжа для систем $Fu^{3+} = b^3$ и $Fu^{3-} = b^3$ соответственно имеют вид: $\lambda^+ = (1; 0,014; 0,014; 1)$, $\lambda^- = (1; 0,31; 0,31; 1)$. Следовательно, по критерию останова данная точка u_3 является оптимальной.

Пример 3. Рассмотрим точку u^0 для задачи размещения двух многоугольников (на рис. 5 соответствующее

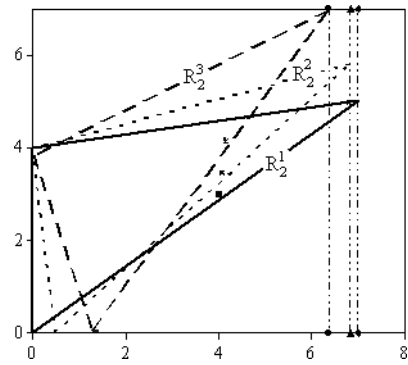


Рис. 4. Иллюстрация критерия останова

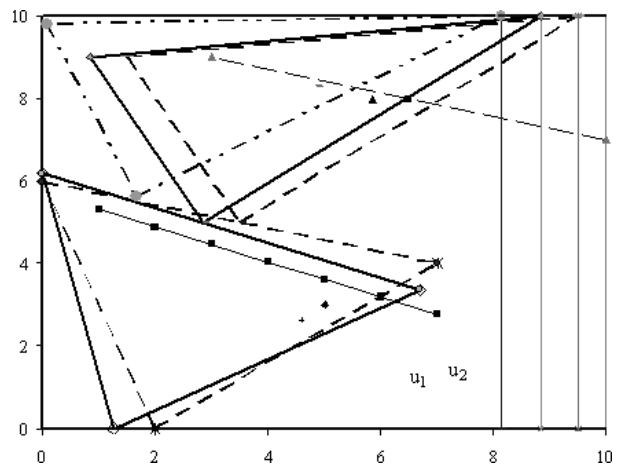


Рис. 5. Иллюстрация примера 3

положение многоугольников показано штриховой линией). Точка $u^0 = (5; 3; 0; 6,5; 8; 0; 9,5)$ определяется системой вида:

$$\begin{cases} -2,07x_1 - 6,98y_1 - 13,63\varphi_1 + 2,07x_2 + 6,98y_2 + \\ + 16,37\varphi_2 = 38, \\ x_1 + 3,3\varphi_1 = 5, \\ y_1 + 3,18\varphi_1 = 3, \\ \varphi_1 = 0, \\ -x_2 + 1,82\varphi_2 + z = 3, \\ -y_2 + 3,12\varphi_2 = -8, \\ \varphi_2 = 0, \end{cases}$$

где коэффициенты первого ограничения рассчитаны по формуле (13).

Вектор множителей Лагранжа $\lambda = (0,48; 1; 3,37; -7,44; 1; 3,37; -20,26)$ содержит два отрицательных множителя: λ_4 и λ_7 , следовательно, из системы можно исключить четвертое и седьмое уравнения.

Положим, исключается четвертое уравнение. Вектор p направления спуска имеет вид: $p = (-3,3; -3,18;$

1; -7,43626; 0; 0; -7,43626). Простейший шаг в направлении p приводит в точку $u^1 = (4,6; 2,62; 0,12; 5,85; 8; 0; 8,85)$ (рис. 5, сплошная линия), определяемую системой

$$\begin{cases} -2,83x_1 - 6,71y_1 - 10,39\varphi_1 + 2,83x_2 + 6,71y_2 + \\ + 11,76\varphi_2 = 38,39, \\ x_1 + 3,9\varphi_1 = 5,07, \\ y_1 + 3,54\varphi_1 = 3,03, \\ \varphi_1 = 0,12, \\ -x_2 + 1,82\varphi_2 + z = 3, \\ -y_2 + 3,12\varphi_2 = -8, \\ \varphi_2 = 0. \end{cases}$$

На рис. 5 штрихпунктирной линией показано положение 2-го многоугольника при выборе в качестве удерживающего 7-го уравнения и соответствующего направления спуска $p = (0; 0; 0; -18,43; 3,12; 1; -20,25)$.

ВЫВОДЫ

Таким образом, в работе проведено исследование оптимизационной задачи размещения многоугольных неориентированных объектов в полосе, рассматриваемой в линеаризованной постановке, и выделены дополнительные свойства линеаризованной области допустимых решений задачи. На этой базе разработан метод поиска локального минимума функции цели, основанный на использовании идеи активного набора, рассмотрена схема определения направления движения по границе компоненты связности линеаризованной области допустимых решений задачи. Приведены примеры реализации шагов алгоритма спуска. Изложенный материал является теоретическим базисом при разработке программного обеспечения процесса решения практических задач оптимизации размещения неориентированных многоугольных объектов, основанного на современном математическом инструментарии.

СПИСОК ЛИТЕРАТУРЫ

1. Стоян Ю. Г. Математические модели и оптимизационные методы геометрического проектирования / Стоян Ю. Г., Яковлев С. В. – К. : Наук. думка, 1986. – 268 с.
2. Гиренко К. А. Математична модель та метод розв'язання задачі розміщення неорієнтованих складених геометричних об'єктів: автореф. дис. на здобуття наук. ступеня канд. техн. наук: спец. 01.05.02 «Математичне моделювання і обчислювальні методи» / К. А. Гиренко. – Харків, 2009. – 18 с.
3. Новожилова М. В. Методологія розв'язання оптимізаційних нелінійних задач геометричного проектування / Новожилова М. В. // Вісник Запорізького державного університету. – 1999. – № 1. – С. 56–60.
4. Гилл Ф. Практическая оптимизация / Гилл Ф., Мюррей У., Райт М. – М. : Мир, 1985. – 509 с.
5. Чуб И. А., Новожилова М. В. // Системи обробки інформації. – 2009. – Вып. 4 (78). – С. 160–163.

Надійшла 15.12.2009

Чуб І. А., Новожилова М. В.

ВИЗНАЧЕННЯ НАПРЯМКУ СПУСКУ В ЛІНЕАРИЗОВАНІЙ ЗАДАЧІ РОЗМІЩЕННЯ НЕОРІЄНТОВАНИХ ГЕОМЕТРИЧНИХ ОБ'ЄКТІВ

Проведено дослідження оптимізаційної задачі розміщення багатокутних неорієнтованих об'єктів у смугі, розглянутої в линеаризованій постановці. Виділено додаткові властивості області припустимих рішень, запропоновано схему методу пошуку локального мінімуму функції мети. Розглянуто схему визначення напрямку руху по межі компоненти зв'язності линеаризованої області припустимих рішень.

Ключові слова: розміщення неорієнтованих багатокутників, лінійна апроксимація, область припустимих рішень, локальний мінімум.

Chub I. A., Novozhilova M. V.

DETERMINATION OF DESCENT DIRECTION IN LINEARIZED PROBLEM OF NON-ORIENTED GEOMETRIC OBJECTS ARRANGEMENT

The problem of non-oriented polygonal objects arrangement optimization is considered in the linearized formulation. Additional properties of the constraint region have been studied. The method of searching an objective local minimum is proposed. The scheme for determining a descent direction along the boundary of connectivity component of the linearized constraint region is described.

Key words: non-oriented polygons arrangement, linear approximation, constraint region, local minimum.

НЕЙРОИНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

НЕЙРОИНФОРМАТИКА И ИНТЕЛЕКТУАЛЬНЫЕ СИСТЕМЫ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

УДК 681.518.2:658.5

Авраменко В. П.¹, Ткаченко В. Ф.², Серета Л. Б.³

¹Д-р техн. наук, профессор Харьковского национального университета радиозлектроники

²Профессор, заведующий кафедрой Харьковского национального университета радиозлектроники

³Старший преподаватель Харьковского национального университета радиозлектроники

ПРИНЯТИЕ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ И НЕЧЕТКОСТИ ИСХОДНОЙ ИНФОРМАЦИИ

Исследованы интеллектуальные вычислительные процедуры принятия управленческих решений в условиях неопределенности и нечетко заданной исходной информации. Разработаны гибкие математические модели и регуляризованные процедуры принятия многокритериальных эффективных управленческих решений.

Ключевые слова: принятие решений, интеллектуальные процедуры, условия неопределенности и нечеткости

ВВЕДЕНИЕ

Важная роль в повышении эффективности производства отводится методам и средствам искусственного интеллекта, которые позволяют повысить качество принимаемых решений в условиях неопределенности цели, функционирования и проявления внешней среды. Источниками неопределенности может выступать несогласованность или противоречивость требований плана и выделенных ресурсов; отклонение исходных предпосылок от фактических условий; наличие дрейфа характеристик и зашумленность измерений, слабая структурируемость и плохая обусловленность системы; варьирование исходных данных и т. п. [1, 2].

Если до недавнего времени считалось, что существует два основных подхода к принятию управленческих решений в условиях неопределенности – детерминированный и вероятностный, то в последнее десятилетие с полным основанием можно говорить о появлении третьего, принципиально ново-

го подхода, основанного на нечетко множественном их описании [3, 4].

Решения, принимаемые в условиях неопределенности исходных данных и проявления внешней среды, всегда приводят к худшим результатам, чем при полной определенности. В этом случае отыскивается квазиоптимальное решение, которое является не безусловно лучшим, а лучшим в некотором смысле, например, при многократном повторении или в смысле близости к некоторому предпочтительному решению с точки зрения лица, принимающего решения (ЛПР).

При формализации задачи принимаются попытки снизить меру неопределенности путем привлечения дополнительной информации. Фактически осуществляется переход от полной неопределенности к стохастической неопределенности или нечетким множествам, что приводит к множеству допустимых

квазиоптимальных решений. Выбор лучшего из квазиоптимальных решений производится на основании различных методов интеллектуализации процедур принятия управленческих решений применительно к конкретному виду неопределенности [1–4].

Таким образом, *актуальной задачей* является исследование характеристик неопределенностей и выявление их влияния на структуру и параметры математических моделей принятия управленческих решений с учетом этих неопределенностей и нечеткости множеств принимаемых решений.

Целью исследования является разработка, совершенствование и реализация гибких математических моделей принятия управленческих решений в условиях различного рода детерминированных и вероятностных неопределенностей, а также нечеткости исходных данных и принимаемых решений.

ФАКТОРЫ НЕОПРЕДЕЛЕННОСТИ И НЕЧЕТКОСТИ ОПИСАНИЯ ЗАДАЧ

Большинство задач принятия управленческих решений приходится решать в условиях неопределенности – неполноты, недостоверности и несвоевременности поступления информации. Существует большое количество источников и форм проявления неопределенности [2]:

- 1) неопределенность, вызванная недостатком информации и ее достоверности;
- 2) неопределенность, связанная с ограничениями времени принятия решения, заданной точности и объема оперативной памяти ЭВМ;
- 3) неопределенность, обусловленная высокой платой за получение определенности;
- 4) неопределенность, возникающая из-за неадекватности модели по причинам неопределенности цели и критериям функционирования, неопределенности проявления внешней среды по отношению к рассматриваемой системе, неопределенности при выборе рациональной сложности модели;
- 5) неопределенность, порождаемая действиями людей в процессе производства и принятия решений;
- 6) преднамеренно организованная неопределенность для сокрытия поведения системы и ее ресурсов.

В настоящее время личностный фактор принятия управленческих решений становится одним из тех источников, которые вносят существенную неопределенность, неточность и нечеткость в описание многих важных технических задач. Для преодоления сложившегося положения Л. А. Заде разработал аппарат нечетких множеств, который позволяет решать задачи классической математики с существенной неопределенностью и нечеткой логикой [3, 4].

Принципиальное отличие неопределенности и нечеткости аппарата принятия решений заключается в том, что в отличие от классической вероятностной неопределенности, пересечение взаимно исключающих событий в нечеткой логике не является пустым множеством, а сумма вероятностей взаимноисключающих событий может быть больше единицы.

Основным инструментом нечеткого моделирования является функция принадлежности. Эта функция предназначена для преобразования лингвистических переменных, называемых терм-множествами (переменных, которым сложно придать точную количественную оценку, такую, например как «малый», «средний», «большой» и т. д.). Терм представляет собой множество значений, которые может принимать нечеткая переменная.

Функцией принадлежности $\mu_A(X)$ является математическая функция, задающая уверенность, с которой элементы некоторого множества X принадлежат заданному нечеткому множеству A . Чем больше аргумент соответствует нечеткому множеству A , тем больше значение $\mu_A(X)$, т. е. тем ближе значение функции к 1. Основанием для построения функции принадлежности могут служить экспертные оценки, справочные или нормативные данные.

Построение функций принадлежности можно осуществлять с использованием прямых и косвенных методов [5].

Прямые методы характеризуются тем, что эксперт непосредственно задает правила определения значений функции принадлежности $\mu_A(X)$, характеризующей элемент x . Примерами прямых методов являются непосредственное задание функции принадлежности таблицей, графиком или формулой. Недостатком этой группы методов является большая доля субъективизма.

В косвенных методах значения функции принадлежности выбираются таким образом, чтобы соответствовать заранее сформулированным условиям. Экспертная информация является только исходной информацией для дальнейшей обработки. К группе данных методов можно отнести методики построения функций принадлежности на основе парных сравнений, с использованием статистических данных, на основе ранговых оценок и т. д.

Основными этапами нечеткого вывода являются: 1) формирование базы правил системы нечеткого вывода; 2) фазсификация входных параметров; 3) агрегирование; 4) активизация подусловий в нечетких правилах продукций; 5) аккумуляция заключений; 6) дефазсификация. Данная схема относится к алгоритму нечеткого вывода Мамдани, который нашел широкое

применение в различных нечетких экспертных системах [5].

Оценка достоверности принимаемых решений осуществляется на основе вероятностных или нечетких моделей, разработка которых в общем случае представляет собой довольно сложную задачу, требующую введения различных ограничений в условиях неопределенности, нечеткой логики или зашумленных данных.

МОДЕЛИ ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

Принятие управленческих решений в общем случае сводится к постановке и решению задач стохастического программирования [1]

$$x^* = \arg \min_{x \in X} \left\{ \frac{E_{\omega} Q[c(\omega), x]}{E_{\omega} G[a(\omega), x]} \geq E_{\omega} b(\omega), x \geq 0 \right\}, \quad (1)$$

где x – вектор искоемых переменных; X – допустимое множество (детерминированное или вероятностное); $E_{\omega}\{\cdot\}$ – оператор усреднения в условиях действия случайного вектора ω ; $Q[c(\omega), x]$, $G[a(\omega), x]$, $c(\omega)$, $a(\omega)$, $b(\omega)$ – соответственно целевая функция, функция ограничений, коэффициенты целевой функции, функции ограничений и условий функционирования, зависящие от случайного фактора ω .

Вероятностная неопределенность позволяет построить стохастические модели различной структуры. Существуют задачи, в которых нарушение ограничений при какой-либо реализации случайных условий приводит к весьма нежелательным последствиям. В таких случаях под планом понимаются лишь векторы x , которые удовлетворяют всем функциональным $a(\omega)$ и ресурсным ограничениям $b(\omega)$, появляющимся с ненулевой вероятностью.

Задачи принятия управленческих решений часто сводятся к задачам терминального управления [2]

$$Q^*(x) = \text{extr}_{x \in X} Q(c, x, \gamma, \eta), \quad (2)$$

$$G_j(a, b, x, \gamma, \eta) \geq R_j, \quad j = \overline{1, m_1}, \quad (3)$$

$$G_j(a, b, x, \gamma, \eta) = R_j, \quad j = \overline{m_1, m}, \quad (4)$$

где x – искомая переменная; a, b, c – параметры модели; $Q(\cdot)$, $G(\cdot)$ – критерий эффективности и ограничения задачи; γ – фактор неопределенности, η – регуляризирующий параметр, зависящий от способа внесения определенности.

Детерминированный эквивалент задачи (1) для нормального распределения случайных величин c_i , a_{ij} , b_j можно представить в виде [1, 2]

$$x^* = \text{Arg min}_{x \in X^n} \left\{ \frac{\sum_{i=1}^n \bar{c}_i x_i}{\sum_{i=1}^n \bar{a}_{ij} x_i} \geq \bar{b}_j + \xi_j, \quad j = \overline{1, m} \right\}, \quad (5)$$

$$\bar{x}_i \leq x_i \leq x_i^+, \quad i = \overline{1, n}, \quad (6)$$

$$\xi_j = t_{\alpha_i} \sqrt{\sum_{i=1}^n \sigma_{ij}^2 x_i + Q_j^2}, \quad j = \overline{1, m}, \quad (7)$$

где ξ_j – слагаемое, учитывающее неопределенность задания коэффициентов функциональных ограничений через вероятностные характеристики – закон распределения, доверительную вероятность и дисперсии σ_{ij}^2 и θ_j^2 случайных величин a_{ij} и b_j ; \bar{c}_i , \bar{a}_{ij} , \bar{b}_j – математическое ожидание случайных величин c_i , a_{ij} , b_j ; t_{α_i} – значение t , соответствующее уровню доверительной вероятности соблюдения ограничений α_i .

Многокритериальную задачу принятия решения сформулируем следующим образом. Пусть на некотором линейном пространстве x^n определены m функций $y_i = \phi_i(x)$, которые можно рассматривать как частные критерии эффективности. Требуется определить такое значение x^* , которое доставляет экстремум (для определенности максимум) нескольким частным критериям y_i , $i = \overline{1, k}$, удерживая остальные y_i , $i = \overline{k+1, m}$ на некотором уровне b_i , $i = \overline{k+1, m}$, т. е.

$$Q(x) = \sum_{i=1}^k \alpha_i \phi_i(x) \rightarrow \max_{x \in X^n}, \quad (8)$$

$$R_i(x) = \phi_i(x) \geq b_i, \quad i = \overline{k+1, m}, \quad (9)$$

$$x_i^- \leq x \leq x_i^+, \quad x_i \geq 0, \quad i = \overline{1, k}. \quad (10)$$

Система функциональных ограничений (9) содержит аналитические зависимости типа балансовых уравнений и статистические – в виде производственных функций. Параметры последних оцениваются по результатам наблюдений, а поэтому содержат вероятностную неопределенность. Применение интегральных оценок наталкивается на неопределенность исходных данных (среднее \pm разброс) и приводит к нескольким квазиоптимальным решениям. Выбор лучшего из них производится ЛПР.

Неопределенность принимаемых решений при многокритериальной оптимизации резко усиливается, поскольку она включает в себя неопределенность исходных данных локальных задач оптимизации, неопределенность вычислительных процедур и неопределенность свертки локальных критериев.

Наименьшей неопределенностью обладает вероятностный метод свертки, когда выбирается главный критерий, рассчитывается статистическая связь между ним и частными критериями, определяются коэффициенты при частных критериях, которые служат весовыми коэффициентами при свертке.

Стоимостной метод устанавливает веса, пропорциональные стоимости улучшения показателя, например, на определенный процент. При экспертном методе веса устанавливаются по важности критериев. ЛПР может устанавливать важность частных критериев по своему усмотрению, поэтому этот метод свертки обладает наибольшей неопределенностью.

Компенсация неопределенности осуществляется различными приемами, учитывающими дополнительную информацию о свойствах объекта и проявлении внешней среды. Каждый способ повышения определенности ориентирован на соответствующий фактор неопределенности. Важным средством повышения определенности является обратная связь, позволяющая реагировать на результаты принятия решения и осуществлять адаптацию характеристик системы, проявления внешней среды и свойств алгоритмов принятия решения.

**МЕТОДЫ РЕШЕНИЯ
МНОГОКРИТЕРИАЛЬНЫХ ЗАДАЧ
В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ**

В функциональных ограничениях $\varphi(x)$ задачи принятия управленческих решений (2)–(4) можно выделить фактор неопределенности в виде отдельного параметра γ :

$$\gamma = \varphi(x, a, \gamma). \tag{11}$$

Основными причинами появления неопределенности γ являются случайные отклонения параметров технологических процессов, непредсказуемые волевые решения организационных вопросов, изменяющиеся неслучайным образом условия работы и проявления внешней среды, старение и износ оборудования. В задачах оптимизации неопределенность проявляется в недостоверности и неполноте исходных данных, неадекватности используемых соотношений и противоречивости системы ограничений. Неопределенность многокритериальных задач включает в себя неопределенность, вносимую исходными данными, регуляризацией при оценивании параметров функциональных ограничений, аппроксимацией некорректно поставленной задачи и сверткой частных критериев.

Учет неопределенности осуществляется с целью определить степень влияния γ на качество получа-

емых решений и по возможности принять меры, ослабляющие это влияние. Существует несколько путей «избавиться» от неопределенности. Наиболее простой из них – заменить в модели (11) случайные изменяющиеся компоненты γ усредненными их характеристиками и перейти к модели

$$y = \varphi(x, a; \bar{\gamma}), \tag{12}$$

где $\bar{\gamma} = E(\gamma)$ – математическое ожидание случайного вектора γ .

Формально избавиться от неопределенности можно путем усреднения не только самого фактора, но и реакции системы. В последнем случае модель (12) заменяется выражением

$$y = \overline{\varphi(x, a; \gamma)} = E\{\varphi(x, a; \gamma)\}. \tag{13}$$

Несмотря на то, что дальнейшие исследования моделей (12) и (13) можно производить детерминированными методами, полученные результаты должны интерпретироваться с учетом вероятностной природы вектора γ . Общая задача (2)–(4) принятия управленческих решений с учетом компонент фактора неопределенности может быть представлена моделью

$$Q(x, \gamma_1) = \{(\varphi_1(x), \dots, \varphi_k(x); \gamma_1)\} \rightarrow \max_{x \in D}, \tag{14}$$

$$D(x; \gamma_2, \gamma_3) = \{x \in R^n \mid R_i = \varphi_i(x, a; \gamma_2) \geq B_i, \\ a_j(x^l, \gamma_3) \leq x_j \leq b(x^l, \gamma_3), x_j \geq 0, j = \overline{1, n}\}, \tag{15}$$

где γ_1 – компонента, отражающая неопределенность процедуры свертки критериев; γ_2 – случайная компонента, отражающая дрейф характеристик; γ_3 – детерминированная компонента, изменения начальных условий (исходных данных).

Таким образом, желание избавиться от вероятностной неопределенности приводит к постановкам задач в классе стохастического программирования, решение которых сопряжено со значительными вычислительными трудностями и внесением новых неопределенностей. Поэтому целесообразно вначале решить задачу детерминированной оптимизации при фиксированных значениях γ^l , а затем исследовать устойчивость и чувствительность полученного решения к проявлениям внешней среды.

**ПРИНЯТИЕ РЕШЕНИЙ ПРИ НЕЧЕТКОЙ
ИСХОДНОЙ ИНФОРМАЦИИ**

При выработке управленческих решений часто встречаются задачи, в которых исходные данные нечетко сформулированы и плохо определены. Такие задачи содержат большое число неопределенностей типа *много, мало, плохо, хорошо*, которые не имеют

аналогов в языке традиционной математики. Поэтому подобные описания средствами традиционной математики сильно огрубляют математическую модель принятия решений.

Для решения такого класса задач потребовалось создание нового математического аппарата для формализации нечетких понятий и суждений, которыми оперирует человек при описании своих представлений о системе, своих желаний и целей. Как уже упоминалось, такой математический аппарат получил название теории нечетких или размытых множеств.

Аппарат нечетких множеств позволяет широко использовать надежные и проверенные математические подходы при решении задач, которые раньше с трудом подлежали математическому описанию или вообще не подлежали формализации. При этом возникла возможность соединения строгости и точности классической математики с имеющейся неопределенностью, нечеткостью и неоднозначностью многих практических ситуаций.

Задача принятия управленческих решений при нечеткой исходной информации может быть сформулирована следующим образом. Пусть имеется некоторое множество $X = \{x_1, x_2, \dots, x_n\}$ альтернатив принимаемых решений. Реализация каждой альтернативы приводит к наступлению некоторого исхода, совокупность которых представляет собой множество $Y = \{y_1, y_2, \dots, y_m\}$. В результате решения задачи требуется построить такую модель выбора альтернатив, которая бы обеспечила лучший в некотором смысле результат этого выбора.

При постановке, анализе и решении задач математического программирования при нечетко заданной исходной информации принято различать следующие подходы [3–5]:

1) задачи достижения поставленной цели для случая пересечения нечеткого множества целей $G(X)$ и четкого множества допустимых альтернатив $C(X)$;

2) задачи достижения поставленной цели для случая пересечения нечеткого множества целей $G(X)$ и нечеткого множества допустимых альтернатив $C(X)$;

3) задачи достижения поставленной цели для случая непересечения нечеткого множества целей $G(X)$ и нечеткого множества допустимых альтернатив $C(X)$ методом взаимной «подтяжки» друг к другу области целей и ограничений;

4) задачи достижения четко поставленной цели $G(X)$ на заданном нечетком множестве допустимых альтернатив $C(X)$;

5) нечеткий вариант задач математического программирования, которые решаются по принципу многоальтернативной оптимизации.

РЕГУЛЯРИЗУЮЩИЕ ПРОЦЕДУРЫ ПРИНЯТИЯ МНОГОКРИТЕРИАЛЬНЫХ РЕШЕНИЙ

Принятие управленческих решений в условиях неопределенности целей, критериев и ограничений часто осуществляется по совокупности локальных критериев $Q_1(x), \dots, Q_i(x)$, которые рассматриваются как локальные критерии качества. Задача многокритериальной оптимизации состоит в том, чтобы на множестве допустимых решений

$$x = \{x: x \in E_+, g(x) \geq 0, h(x) = 0\} \quad (16)$$

определить точку x^* , в которой векторный критерий эффективности достигает максимально-предпочтительного значения

$$Q^*(x) = \langle Q_1^*(x), \dots, Q_i^*(x) \rangle \rightarrow \max_{x \in X} Q(x). \quad (17)$$

Поскольку получить решение x^* , доставляющее максимальные значения одновременно всем локальным критериям эффективности, не представляется возможным, то важным этапом решения многокритериальной задачи является установление компромисса о том, какое решение считать наиболее предпочтительным по совокупности локальных критериев. Наиболее разработанными в настоящее время считаются процедуры выделения эффективных решений.

Подмножество Ω_3 множества допустимых решений x является областью эффективных решений (областью компромиссов или областью решений, оптимальных по Парето)

$$\Omega_3 = \{x: x \in X, \langle x': Q(x') \geq Q(x) \rangle \cap X \neq \emptyset\}, \quad (18)$$

если все принадлежащие ему решения не могут быть улучшены по всем локальным критериям без снижения уровня хотя бы одной из компонент векторного критерия $Q = \langle Q_1(x), \dots, Q_i(x) \rangle$. Совокупность эффективных значений векторного критерия образует паретову границу подмножества $E(\Omega_3)$. Допустимый план, который не лежит на паретовой границе, не является эффективным, поскольку его можно улучшить с точки зрения некоторого критерия, а, следовательно, и глобальной цели, с которой этот локальный критерий согласован.

Множественность эффективных решений является скорее достоинством, а не недостатком, поскольку «жесткие» схемы получения единственного решения неадекватны сущности векторной оптимизации, а имеющаяся «свобода» выбора предпочтительного решения из множества эффективных позволяет учесть неопределенность целей и критериев.

ВЫВОДЫ

Поскольку мир неопределенностей и нечеткостей исходной информации чрезвычайно разнообразен, то при исследовании конкретной предметной области целесообразно выделять наиболее характерные для нее неопределенности и нечеткости с точки зрения решаемых задач.

В задачах принятия управленческих решений широко распространена неопределенность обстановки, причинами которой являются случайные отклонения параметров процессов, непредсказуемые волевые решения организационных вопросов, изменяющиеся неслучайным образом условия работы и проявления внешней среды, старение и износ оборудования.

В задачах оптимизации неопределенность проявляется в недостоверности и неполноте исходных данных, неадекватности используемых соотношений и противоречивости системы ограничений. Неопределенность многокритериальных задач включает в себя неопределенность, вносимую исходными данными, регуляризацией при оценивании параметров функциональных ограничений, аппроксимацией некорректно поставленных задач и сверткой частных критериев.

Решение конкретных задач с учетом нечеткости характеристик объекта, проявления внешней среды и нарушения исходных предпосылок требует разработки интеллектуальных (регуляризованных, эвристических) процедур принятия управленческих решений с ориентацией на существующий вид неопределенности и нечеткость исходной информации.

УДК 004.032.26

СПИСОК ЛИТЕРАТУРЫ

1. Авраменко В. П. Управление производством в условиях неопределенности / Авраменко В. П. – Киев : УМК ВО, 1992. – 48 с.
2. Авраменко В. П. Концепции интеллектуализации процедур принятия управленческих решений в условиях неопределенности / Авраменко В. П. // АСУ и приборы автоматизации. – 1998. – Вып. 108. – С. 42–58.
3. Пономарев А. С. Нечеткие множества в задачах автоматизированного управления и принятия решений / Пономарев А. С. – Харьков : НТУ «ХПИ», 2005. – 258 с.
4. Раскин Л. Г. Нечеткая математика. Основы теории. Приложения / Раскин Л. Г., Серая О. В. – Харьков : Парус, 2008. – 352 с.
5. Леоненков А. Нечеткое моделирование в среде Matlab и fuzzyTECH / Леоненков А. – СПб. : БХВ-Петербург, 2005. – 716 с.

Надійшла 12.04.2010

Авраменко В. П., Ткаченко В. П., Серета Л. Б.
ПРИЙНЯТТЯ УПРАВЛІНСЬКИХ РІШЕНЬ В УМОВАХ НЕВИЗНАЧЕНОСТІ ТА НЕЧІТКОСТІ ВИХОДНОЇ ІНФОРМАЦІЇ

Досліджено інтелектуальні обчислювальні процедури прийняття управлінських рішень в умовах невизначеності та нечітко заданої початкової інформації. Розроблено гнучкі математичні моделі та регуляризовані процедури прийняття багатокритеріальних управлінських рішень.

Ключові слова: прийняття рішень, інтелектуальні процедури, умови невизначеності та нечіткості.

Avramenko V. P., Tkachenko V. P., Sereta L. B.
MANAGEMENT DECISIONS IN CONDITIONS OF UNCERTAIN AND FUZZY SOURCE INFORMATION

Intelligent computer procedures of making management decisions in conditions of uncertain and fuzzy source information were studied. Flexible mathematic models and regularized procedures of making multicriterion management decisions were developed.

Key words: decision-making, intellectual procedures, conditions of uncertainty and fuzziness.

Асеев Г. Г.

Докт. техн. наук, заведующий кафедрой Харьковской государственной академии культуры

МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОЙ ПРЕДОБРАБОТКИ ДАННЫХ В ЭЛЕКТРОННЫХ ХРАНИЛИЩАХ

Представлен один из возможных формализмов задачи предобработки входных сигналов для сигмоидных обучающихся нейронных сетей на основании минимизации критерия Липшица.

Ключевые слова: предобработка, нейронная сеть, электронное хранилище, константа Липшица, синапс, сумматор.

ВВЕДЕНИЕ

Возрастание объемов информации, необходимой для принятия решений, приводит к резкому увеличе-

нию количества документов, в связи с чем традиционные методы работы с последними становятся малоэффективными. На помощь приходят корпоратив-

© Асеев Г. Г., 2010

ные информационные системы (КИС), которые позволяют создавать и обрабатывать документы электронными средствами. КИС обеспечивают процесс создания, управления доступом и распространения больших объемов информации в компьютерных сетях, а также контроль над ее потоками в организации. Информация рассматривается не только как некая принадлежащая предприятию ценность, но и как исходная точка для построения информационной системы, обслуживающей предприятие. Следует отметить, что правильное понимание информационного аспекта служит необходимой предпосылкой для построения высококачественных и целостных информационных систем. По этой причине ведущим направлением при создании КИС является переход от процедурно-ориентированных методов разработки [1, 2] к информационно-ориентированным [3].

Обычно информация в виде документов хранится в специальных хранилищах данных, где за время своего существования большинством организаций накоплены огромные объемы данных, которые могли бы быть использованы в практической деятельности: дать ответ о том, что нужно клиентам, наиболее предпочтительным для организации, как наиболее эффективно разместить ресурсы, каким образом минимизировать потери и т. д.

ПОСТАНОВКА ПРОБЛЕМЫ

В настоящее время потоки текстовой и числовой информации ежедневно порождаются и оседают в хранилищах данных. Насколько полно на практике используются все те закономерности, которые кроются в этих данных и, возможно, представляют большую ценность? Можно предположить, что процент переработки «сырых» данных в практически значимые знания пока что весьма скромнен.

Всякое познание может быть представлено в виде модели. Модель – это искусственно создаваемая система, в которой отражено сходство структуры и функции с системой-оригиналом. Существуют два вида моделей: предсказательные и описательные. В первых используется один набор данных с известными результатами для построения моделей, которые явно предсказывают результаты для других наборов данных [3, 4], а во вторых описываются зависимости в существующих данных [5]. Созданная модель не сможет претендовать на абсолютное знание, но даст аналитику определенное преимущество уже самим фактом обнаружения альтернативной статистически значимой модели.

В последнее время активно развиваются новые описательные методы анализа данных и извлечения знаний, базирующиеся на иных, нежели традиционная интегро-дифференциальная парадигма, подходах. Имеются в виду методы эволюционного моделирования. Под этим термином подразумеваются искусственные нейронные сети, генетические алгоритмы, RBF сети и пр.

Рассмотрим искусственные нейронные сети, а именно, компоненту предобработки входных данных. Существует множество различных видов нейронных сетей [6–9]. Однако, для большинства нейронных сетей характерно наличие такого интервала входных сигналов, в пределах которого сигналы различимы. Для различных нейронных сетей эти интервалы различны. До настоящего времени результаты попыток формализовать или унифицировать подходы к предобработке входных сигналов были незначительны. В настоящей статье дадим один из возможных формализмов этой задачи.

ОПИСАНИЕ ПРЕДОБРАБОТЧИКА

В качестве примера рассмотрим сети с сигмоидными нелинейными преобразователями. Описываемые ниже способы предобработки могут быть применены для сетей с произвольными нелинейными преобразователями.

Нейрон. Нейроны, используемые в большинстве нейронных сетей, имеют структуру, приведенную на рис. 1, где применены следующие обозначения: x – вектор входных сигналов нейрона; α – вектор синаптических весов нейрона; Σ – входной сумматор нейрона; $p = p(\alpha, x)$ – выходной сигнал входного сумматора; σ – функциональный преобразователь; y – выходной сигнал нейрона.

Обычно нейронным сетям присваивают названия по виду функции $\sigma(p)$. Хорошо известны и наиболее часто используются следующие два вида сигмоидных сетей:

$$S_1: \quad \sigma(p) = 1/(1 + \exp(-cp)),$$

$$S_2: \quad \sigma(p) = p/(c + |p|),$$

где c – параметр, называемый «характеристикой нейрона». Обе функции имеют идентичные графики.

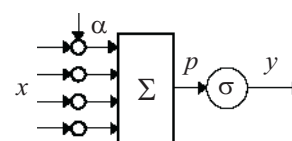


Рис. 1. Пример нейрона

Кждому типу нейрона соответствует свой интервал приемлемых входных данных. Как правило, этот диапазон либо совпадает с диапазоном выдаваемых выходных сигналов (например, для сигмоидных нейронов с функцией S_1), либо является объединением диапазона выдаваемых выходных сигналов и отрезка, симметричного ему относительно нуля (например, для сигмоидных нейронов с функцией S_2). Этот диапазон будем обозначать как $[a, b]$.

Простейшая предобработка числовых признаков. Числовые сигналы рекомендуется масштабировать и сдвигать так, чтобы весь диапазон значений попадал в диапазон приемлемых входных сигналов. Эта предобработка проста и задается следующей формулой:

$$c' = \frac{(c - c_{\min})(b - a)}{c_{\max} - c_{\min}} + a, \quad (1)$$

где $[a, b]$ – диапазон приемлемых входных сигналов; $[c_{\min}, c_{\max}]$ – диапазон значений признака c , c' – предобработанный сигнал, который будет подан на вход сети. Предобработку входного сигнала по формуле (1) будем называть простейшей предобработкой.

Оценка способности сети решить задачу. В данном подразделе рассматриваются только сети, все элементы которых непрерывно зависят от своих аргументов. Предполагается, что все входные данные предобработаны так, что все входные сигналы сети лежат в диапазоне приемлемых входных сигналов $[a, b]$. Будем обозначать векторы входных сигналов через x^i , а требуемые ответы сети через f^i . Компоненты векторов будем обозначать нижним индексом, например, компоненты входного вектора через x_j^i . Будем полагать, что в каждом примере ответ является вектором чисел из диапазона приемлемых сигналов $[a, b]$. В случае обучения сети задаче классификации требуемый ответ зависит от вида используемого интерпретатора ответа.

Нейронная сеть вычисляет некоторую вектор-функцию F от входных сигналов. Эта функция зависит от параметров сети. Обучение сети состоит в подборе такого набора параметров сети, чтобы величина $\sum_{i,j} [F_j(x^i) - f_j^i]^2$ была минимальной (в идеале равна нулю). Для того, чтобы нейронная сеть могла хорошо приблизить таблично заданную функцию f , необходимо, чтобы реализуемая сетью функция F при изменении входных сигналов с x^i на x^j могла изменить значение с f^i на f^j . Очевидно, что наиболее трудным для сети должно быть приближение функции в точках, в которых при малом изменении входных

сигналов происходит большое изменение значения функции. Таким образом, наибольшую сложность будет представлять приближение функции f в точках, в которых достигает максимума выражение $\|f^i - f^j\| / \|x^i - x^j\|$. Для аналитически заданных функций величина $\sup_{x,y} \|f(x) - f(y)\| / \|x - y\|$ называется константой Липшица. Исходя из этих соображений, можно дать следующее определение сложности задачи.

Сложность аппроксимации таблично заданной функции f , которая в точках x^i принимает значения f^i , задается выборочной оценкой константы Липшица, вычисляемой по следующей формуле:

$$\Lambda_i = \max_{i \neq j} \|f^i - f^j\| / \|x^i - x^j\|. \quad (2)$$

Оценка (2) является оценкой константы Липшица аппроксимируемой функции снизу.

Для того, чтобы оценить способность сети заданной конфигурации решить задачу, необходимо оценить константу Липшица сети и сравнить ее с выборочной оценкой (2). Константа Липшица сети вычисляется по следующей формуле:

$$\Lambda_n = \sup_{x,y} \|F(x) - F(y)\| / \|x - y\|. \quad (3)$$

В формулах (2) и (3) можно использовать произвольные нормы. Однако для нейронных сетей наиболее удобной является евклидова норма. Далее везде используется евклидова норма.

Ниже опишем способ вычисления оценки константы Липшица сети (3) сверху. Очевидно, что в случае $\Lambda_n < \Lambda_i$ сеть принципиально не способна решить задачу аппроксимации функции f .

Оценка константы Липшица сети. Оценку константы Липшица сети будем строить в соответствии с принципом иерархического устройства сети. При этом потребуются следующие правила:

1. Для композиции функций $f \circ g = f(g(x))$ константа Липшица оценивается как произведение констант Липшица:

$$\Lambda_{f \circ g} \leq \Lambda_f \Lambda_g. \quad (4)$$

2. Для вектор-функции $f = (f_1, f_2, \dots, f_n)$ константа Липшица равна:

$$\Lambda_f = \sqrt{\sum_{i=1}^n \Lambda_{f_i}^2}. \quad (5)$$

Для непрерывных функций константа Липшица является максимумом производной в направлении $r = (r_1, r_2, \dots, r_n)$ по всем точкам и всем направлениям. При этом вектор направления имеет единичную

длину $\sum_{i=1}^n r_i^2 = 1$. Напомним формулу производной функции $f(x_1, x_2, \dots, x_n)$ в направлении r :

$$\frac{\partial f}{\partial r} = \sum_{i=1}^n r_i \frac{\partial f}{\partial x_i}. \quad (6)$$

Синапс. Обозначим входной сигнал синапса через x , а синаптический вес через α . Тогда выходной сигнал синапса равен αx . Поскольку синапс является функцией одной переменной, то константа Липшица равна максимуму модуля производной – модулю синаптического веса:

$$\Lambda_s = |\alpha|. \quad (7)$$

Умножитель. Обозначим входные сигналы умножителя через x_1, x_2 . Тогда выходной сигнал умножителя равен $f^* = x_1 x_2$. Используя (6), получаем $\Lambda_{f^*} = \sup_{x, r} |r_1 x_2 + r_2 x_1|$. Выражение $r_1 x_2 + r_2 x_1$ является скалярным произведением векторов (r_1, r_2) и, учитывая единичную длину вектора r , достигает максимума, когда эти векторы сонаправлены. То есть, при векторе

$$r = \left(\frac{x_2}{\|x\|}, \frac{x_1}{\|x\|} \right), \quad \|x\| = \sqrt{x_1^2 + x_2^2}.$$

Используя это выражение, можно записать константу Липшица для умножителя:

$$\Lambda_{f^*} = \sup_{x, r} |r_1 x_2 + r_2 x_1| = \sup_x \frac{|x_1^2 + x_2^2|}{\|x\|} = \sup_x \|x\|. \quad (8)$$

Если входные сигналы умножителя принадлежат интервалу $[a, b]$, то константа Липшица для умножителя может быть записана в следующем виде:

$$\Lambda_{f^*} = \sqrt{2} \max\{|a|, |b|\}. \quad (9)$$

Точка ветвления. Поскольку в точке ветвления не происходит преобразование сигнала, то константа Липшица для нее равна единице.

Сумматор. Производная суммы по любому из слагаемых равна единице. В соответствии с (6) получаем:

$$\Lambda_\Sigma = \sup_r \left| \sum_{i=1}^n r_i \right| = \sqrt{n}, \quad (10)$$

поскольку максимум суммы при ограничении на сумму квадратов достигается при одинаковых слагаемых.

Нелинейный Паде преобразователь. Нелинейный Паде преобразователь или Паде элемент имеет два входных сигнала и один выходной. Обозначим входные сигналы через x_1, x_2 . Используя (6), можно записать константу Липшица в следующем виде:

$$\Lambda_\pi = \sup_{r, x} \left| \frac{r_1 x_2}{x_2^2} - \frac{r_2 x_1}{x_2^2} \right| = \sup_{r, x} \left| \frac{r_1 x_2 - r_2 x_1}{x_2^2} \right|.$$

Знаменатель выражения под знаком модуля не зависит от направления, а числитель можно преобразовать так же, как и для множителя. После преобразования получаем:

$$\Lambda_\pi = \sup_x \frac{\|x\|}{x_2^2}. \quad (11)$$

Нелинейный сигмоидный преобразователь. Нелинейный сигмоидный преобразователь, как и любой другой нелинейный преобразователь, имеющий один входной сигнал x , имеет константу Липшица, равную максимуму модуля производной:

$$\Lambda_\phi = \sup_x \frac{\|x\|}{x_2^2}. \quad (12)$$

Адаптивный сумматор. Для адаптивного сумматора на n входов оценка константы Липшица, получаемая через представление его в виде суперпозиции слоя синапсов и простого сумматора, вычисляется следующим образом. Используя формулу (7) для синапсов и правило (5) для вектор-функции, получаем следующую оценку константы Липшица слоя синапсов:

$$\Lambda_L = \sqrt{\sum_{i=1}^n \Lambda_{S_i}^2} = \sqrt{\sum_{i=1}^n |\alpha_i|^2} = \|\alpha\|. \quad (13)$$

Используя правило (4) для суперпозиции функций и оценку константы Липшица для простого сумматора (10), получаем:

$$\Lambda_A = \sup_{x, r} \left| \sum_{i=1}^n r_i \alpha_i \right| = \left| \sum_{i=1}^n \alpha_i \frac{\alpha_i}{\|\alpha\|} \right| = \|\alpha\|. \quad (14)$$

Очевидно, что оценка (14) точнее, чем оценка (13).

Константа Липшица слоистой сигмоидной сети. Рассмотрим слоистую сигмоидную сеть со следующими свойствами: а) число входных сигналов – n_0 ; б) число нейронов в i -м слое – n_i ; в) каждый нейрон первого слоя получает все входные сигналы, а каждый нейрон любого другого слоя получает сигналы всех нейронов предыдущего слоя; г) все нейроны всех слоев имеют вид, приведенный на рис. 1 и имеют одинаковую характеристику; д) все синаптические веса ограничены по модулю единицей; е) в сети m слоев.

В этом случае, учитывая формулы (4), (5), (12) и (14), константу Липшица i -го слоя можно оценить следующей величиной:

$$\Lambda_i \leq \sqrt{\sum_{j=1}^{n_i} (\Lambda_\phi \Lambda_{A_j})^2} = \Lambda_\phi \sqrt{\sum_{j=1}^{n_i} \|\alpha^j\|^2} \leq \Lambda_\phi \sqrt{n_{i-1} n_i}.$$

Используя формулу (4), получаем оценку константы Липшица всей сети:

$$\Lambda_n \leq \prod_{i=1}^m \Lambda_i \leq \Lambda_\phi^m \sqrt{n_0 n_m} \prod_{i=1}^{m-1} n_i.$$

Если используются нейроны типа S_1 , то $\Lambda_\phi = c$ и оценка константы Липшица сети равна:

$$\Lambda_{S_1} \leq c^m \sqrt{n_0 n_m} \prod_{i=1}^{m-1} n_i.$$

Для нейронов типа S_2 $\Lambda_\phi = 1/c$ и оценка константы Липшица сети равна:

$$\Lambda_{S_2} \leq c^{-m} \sqrt{n_0 n_m} \prod_{i=1}^{m-1} n_i.$$

Обе последние формулы подтверждают экспериментально установленный факт, что чем круче характеристическая функция нейрона, тем более сложные функции (функции с большей константой Липшица) может аппроксимировать сеть с такими нейронами.

ПРЕДОБРАБОТКА, ОБЛЕГЧАЮЩАЯ ОБУЧЕНИЕ

При обучении нейронных сетей иногда возникают ситуации, когда дальнейшее обучение нейронной сети невозможно. В этом случае необходимо выявить и проанализировать причины. Возможны несколько видов анализа. Одной из возможных причин является высокая сложность задачи, определяемая как выборочная оценка константы Липшица.

Для упрощения задачи необходимо уменьшить выборочную оценку константы Липшица. Наиболее простой способ добиться этого – увеличить расстояние между входными сигналами. Рассмотрим пару примеров – x^i, x^j , таких, в которых $\Lambda_i = \|f^i - f_j\| / \|x^i - x^j\|$. Определим среди координат векторов x^i и x^j координату, в которой достигается минимума величина $|x_l^i - x_l^j|$, исключив из рассмотрения совпадающие координаты. Очевидно, что эта координата является «узким местом», определяющим сложность задачи. Следовательно, для уменьшения сложности задачи требуется увеличить расстояние между векторами x^i и x^j , а наиболее перспективной координатой для этого является l -я. Однако увеличение расстояния между x_l^i и x_l^j не всегда имеет смысл. Дело в том, что все параметры, как правило, измеряются с конечной точностью. Поэтому, если величина $|x_l^i - x_l^j|$ меньше точности измерения l -го параметра, значения x_l^i и x_l^j можно считать совпадающими. Таким образом, для изменения масштаба надо выби-

рать тот из входных параметров, для которого значение $|x_l^i - x_l^j|$ минимально, но превышает точность измерения этого параметра.

Предположим, что все входные параметры преобразованы в соответствии с формулой (1). Перенумеруем примеры обучающего множества так, чтобы были верны следующие неравенства: $x_1^1 < x_1^2 < \dots < x_1^N$, где N – число примеров в обучающем множестве. При этом придется исключить ряд пар «параметр – ответ» с совпадающими значениями параметра. Если в какой-либо из таких пар значения ответов различаются, это снижает возможную полезность данной процедуры.

Наиболее простой путь – разбить диапазон l -го параметра на два. Зададимся точкой x . Будем кодировать l -й параметр двумя входными сигналами в соответствии с табл. 1.

Таблица 1. Кодирование параметра после разбиения на два сигнала

Значение	Первый сигнал	Второй сигнал
$x_l^i < x$	$\frac{(x_l^i - a)(b - a)}{(b - a)} + a$	a
$x_l^i > x$	b	$\frac{(x_l^i - a)(b - a)}{(b - x)} + a$

При таком кодировании критерий Липшица, очевидно, уменьшится. Вопрос о выборе точки x может решаться по-разному. Простейший путь – задать $x = (a + b)/2$. Более сложный, но часто более эффективный – подбор x , исходя из требования минимальности критерия Липшица.

ВЫВОДЫ

При больших объемах хранилищ данных практически невозможно напрямую применить методы Data Mining. Если не будут произведены формализация и сбор данных и их представления, то тогда плохое «качество» исходных данных явится одной из самых серьезных проблем. В связи с тем, что в большинстве случаев источником информации для аналитических систем является корпоративное хранилище данных, в котором аккумулируются сведения из множества разнородных источников, острота проблемы существенно возрастает.

Необходимость предварительной обработки данных при их анализе возникает независимо от того, какие технологии и алгоритмы используются. Более того, эта задача может представлять самостоятельную ценность в областях, не имеющих непосредственного отношения к анализу данных. При исполь-

зованим же механізмом аналізу, в основі яких лежать самообучаючіся алгоритми, такі як нейронні мережі, дерева рішень і проче, хороше якість даних є ключовим вимога. В даній роботі був дан один з можливих формалізмів цієї задачі к передобротці входних сигналів на основі мінімізації критерію Липшица для сигмоїдних навчаючихся нейронних мереж.

СПИСОК ЛІТЕРАТУРИ

1. *Асеев Г. Г.* Логико-алгебраический подход к описанию информационных процессов электронного документооборота / Г. Г. Асеев, Т. Г. Белова // АСУ и приборы автоматики. – 2004. – Вып. 126. – С. 161–167.
2. *Асеев Г. Г.* Концепция математической модели информационных потоков электронного взаимодействия в документообороте / Г. Г. Асеев, Т. Г. Белова // Вестник НТУ «ХПИ». – 2003. – № 6. – С. 161–168.
3. *Асеев Г. Г.* Электронный документооборот : учеб. / Г. Г. Асеев. – К. : Кондор, 2007. – 500 с.
4. *Асеев Г. Г.* Методы интеллектуального анализа данных в электронных хранилищах / Г. Г. Асеев // Бионика интеллекта : науч.-техн. журнал. – 2008. – № 1(70). – С. 28–33.
5. *Асеев Г. Г.* Проблема обнаружения нового знания в хранилищах данных методами Knowledge Discovery in Databases / Г. Г. Асеев // Вестник НТУ «ХПИ». – № 19. – С. 62–70.
6. *Галушкин А. И.* Нейрокомпьютеры и их применение. Кн. 1: Теория нейронных сетей / А. И. Галушкин. –

М. : Изд. предприятие редакции журнала «Радиотехника», 2000. – 354 с.

7. *Кохонен Т.* Анализ финансовых данных с помощью самоорганизующихся карт / Тейво Кохонен, Гвидо Дебук. – М. : Альпина, 2001. – 380 с.
8. *Корнеев В. В.* Базы данных. Интеллектуальная обработка информации / В. В. Корнеев, А. Ф. Гарев, С. В. Васютин, В. В. Райх. – М. : Нолидж, 2000. – 352 с.
9. *Царегородцев В. Г.* Предобработка обучающей выборки, выборочная константа Липшица и свойства обученных нейронных сетей // Материалы X Всеросс. семинара «Нейроинформатика и ее приложения» / В. Г. Царегородцев. – Красноярск, 2002. – С. 146–150.

Надійшла 22.03.2010

Асеев Г. Г.

МЕТОДИ ІНТЕЛЕКТУАЛЬНОЇ ПЕРЕДОБРОБКИ ДАНИХ В ЕЛЕКТРОННИХ СХОВИЩАХ

Подано один з можливих формалізмів задачі передобротці входних сигналів для сигмоїдних нейронних мереж, що навчаються, на підставі мінімізації критерію Липшица.

Ключові слова: передобротка, нейронна мережа, електронне сховище, константа Липшица, синапс, суматор.

Aseyev G. G.

METHODS OF INTELLIGENT INPUT DATA PRELIMINARY PROCESSING IN DATA WAREHOUSE

One of the possible formalisms is presented of the problem of input signals preliminary processing for sigmoidal learnable neural networks on the basis of Lipschitz criterion minimization.

Key words: pre-treatment, neural network, electronic data warehouse, Lipschitz constant, synapse, summator.

УДК 519.7: 655.3.022.16

Бодянский Е. В.¹, Кулишова Н. Е.²

¹Д-р техн. наук, профессор Харьковского национального университета радиозлектроники

²Канд. техн. наук, доцент Харьковского национального университета радиозлектроники

СГЛАЖИВАНИЕ ПОМЕХ НА ЦВЕТНЫХ ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ НЕЙРО-ФАЗЗИ ФИЛЬТРА

Рассмотрена проблема подавления помех на цветных цифровых изображениях. Для решения предложен нейро-фаззи фильтр, который обеспечивает высокое качество фильтрации с сохранением контуров, текстур и плавных тоновых переходов.

Ключевые слова: цветные изображения, вычислительный интеллект, нейро-фаззи фильтрация, помеха.

ВВЕДЕНИЕ

В структуре полиграфических изданий зачастую значительную долю составляют иллюстрации, которые могут быть получены либо фотографированием, либо создаются с помощью специальных графических пакетов. В настоящее время цифровые или оцифрованные фотографии используются для иллюстраций гораздо чаще, в связи с чем возникает множе-

ство проблем по подготовке таких изображений к печати [1, 2].

На оцифрованных изображениях, как правило, присутствует шум в виде пятен различного рода [3–5], порождаемых различного рода возмущениями:

– мелкие черные пятна, обусловленные наличием пыли на поверхности. Они встречаются на многих фотографиях. Размеры пятен очень малы, в связи с чем их можно считать точечными объектами;

- пятна от грязи или клякс при попадании на фотografiю чернил или другой жидкости;

- пятна, обусловленные химическими реакциями (гидролиз, выделение химических соединений);

- другие виды пятен (отпечатки пальцев, плесень и так далее);

- пятна, возникающие из-за зернистости фотопленки или фотобумаги.

Для сглаживания резких перепадов яркости, которыми и сопровождается присутствие пятен и дефектов на изображении, достаточно часто используют линейные алгоритмы фильтрации [6, 7], которые являются оптимальными при гауссовом распределении сигналов и помех. Реальные изображения, строго говоря, не подчиняются нормальному распределению вероятностей. Поддаваясь локальному гауссову описанию в пределах ограниченных участков, многие реальные изображения в этой связи плохо представляются как глобально гауссовы объекты. Именно это и служит причиной плохой передачи границ при линейной фильтрации.

Эффективным методом для подавления шумов является медианная фильтрация [8], в процессе которой сначала производится сортировка по величине пикселей, попавших в окно фильтра, а затем замена значения опорной точки величиной яркости пикселя, расположенного на медиане этого ряда. В связи с этим медианный фильтр не искажает ступенчатые и линейно изменяющиеся сигналы, сохраняя контуры изображений.

Сглаживающие фильтры, как линейные, так и нелинейные, имеют общий недостаток: подавляя верхние частотные компоненты изображения, они не только уменьшают влияние шума, но и приводят к потере мелких объектов изображения и множества тоновых переходов, которые не являются ступенчатыми. Кроме того, процедура упорядочивания пикселей в пределах окна требует значительных вычислительных затрат даже для небольших по размеру изображений.

Неравномерность изображений приводит к существованию более чем одного типа признаков изображений, которые доносят независимые формы информации зрителю. Хотя вариаций среди разных изображений может быть множество, большое количество изображений можно характеризовать малым числом типов признаков. Они сводятся к гладким областям, текстурам и границам [9]. Гладкие области обычно соответствуют преобладающим площадям изображений, потому что поверхности искусственных или естественных объектов при их регистрации с расстояния всегда рассматриваются как гладкие. В проти-

воположность гладким областям, границы составляют очень малую долю площадей изображений. Тем не менее, большая часть информации об изображении доносится через эти границы. Возникновение текстур всегда свойственно присутствию натуральных объектов в изображении. Учитывая разные свойства упомянутых типов признаков, обычно необходимо придавать пространственную адаптивность системам обработки изображений для получения оптимальных результатов.

1. ПОСТАНОВКА ЗАДАЧИ

Стандартные фильтры, работающие с окном, ориентированы на неизменные характеристики опорного сигнала, они не могут учитывать изменяющиеся локальные характеристики изображений. Их действие не зависит от характера изображения: текстура это или гладкая цветовая область. Для решения подобной проблемы используется фаззи подход непосредственно к обработке пикселей окна.

Пусть, $K_1 \times K_2$ – многоканальное изображение $x:Z_1 \rightarrow Z_m$, соответствующее двумерной матрице из m компонентов например, $x_i = [x_{i1}, x_{i2}, \dots, x_{im}]$. Для стандартных цветных RGB изображений обычно $m = 3$. Компоненты x_{ik} , для $k = 1, 2, \dots, m$ и $i = 1, 2, \dots, K_1 K_2$, изображают k -е элементы векторного входного сигнала x_i . Поэтому каждый многоканальный образец x_i может считаться m -мерным вектором. Под действием некоторого скользящего окна $W = \{x \in Z; i = 1, 2, \dots, N_w\}$, ограниченного размером N_w , центр окна $x_{(N_w+1)/2}$ заменяется некоторой функцией местной окрестности. Этот оператор окна скользит над изображением, чтобы воздействовать индивидуально на все пиксели изображения.

Общая форма сглаживающего фильтра, зависящего от изображения, может быть определена как нечеткое взвешенное усреднение входящих векторов в пределах окна [10]:

$$\tilde{x}_{(N+1)/2} = f\left(\sum_{i=1}^{N_w} w_i^C x_i\right) = f\left(\frac{\sum_{i=1}^{N_w} w_i^C x_i}{\sum_{i=1}^{N_w} w_i^C}\right), \quad (1)$$

где $f(\bullet)$ – нелинейная функция, w_i^C – веса фильтра, эквивалентные уровням нечеткой функции принадлежности, ассоциированной с входным цветовым вектором x_i . Можно использовать разные функции для задания меры расстояния в цветовом пространстве между векторами x_i . Это позволяет формулировать разные критерии обучения для синтеза алгоритмов настройки подобных адаптивных фильтров [11, 12].

Целью данной работы является синтез архитектуры и алгоритма обучения нейро-фаззи фильтра, предназначенного для сглаживания помех различного вида на цветных изображениях с максимальным сохранением контуров, текстур и плавных тоновых переходов.

2. АРХИТЕКТУРА НЕЙРО-ФАЗЗИ ФИЛЬТРА

На рис. 1 приведена архитектура нейро-фаззи фильтра для обработки цветных изображений, являющегося по сути модификацией адаптивной нейро-сетевой системы нечеткого вывода ANFIS [13, 14].

Данная архитектура имеет шесть слоев обработки информации и рецепторный (нулевой) слой, на который поступает входной сигнал в виде матрицы, столбцы которой соответствуют разным цветовым каналам (R, G, B) обрабатываемого изображения. В результате векторизации этот сигнал преобразуется в вектор с элементами x_i^C , где $i = 1, 2, \dots, n; n + 1, \dots, 2n; 2n + 1, \dots, 3n$, C принимает значения R, G или B, соответствующие различным каналам, а n определяет число одновременно обрабатываемых пикселей (или, иначе, размер окна).

В первом скрытом слое осуществляется фаззификация каждой компоненты входного сигнала x_i^C с помощью набора функций принадлежности $\mu_{ji}^C(x_i^C)$, $j = 1, 2, \dots, m$, в качестве которых могут быть использованы любые колоколообразные конструкции (треугольные, кубические сплайны, обобщенные функции Гаусса), в простейшем случае – гауссовы функции вида

$$\mu_{ji}^C(x_i^C) = \exp\left(-\frac{1}{2}\left(\frac{x_i^C - c_j^C}{\sigma_j^C}\right)^2\right), \quad (2)$$

где c_j^C определяет положение центра j -й функции принадлежности C -го цвета, а σ_j^C – ее ширину.

Во втором скрытом слое производится раздельное агрегирование уровней принадлежности входных сигналов по каждому цветовому каналу на основе вычисления нечеткой T -нормы вида

$$w_j^C = \prod_{l=1}^n \mu_{jl}^C(x_l^C), \quad (3)$$

а в третьем реализуется операция нормализации

$$\bar{w}_j^C = \frac{w_j^C}{\sum_{j=1}^m \prod_{l=1}^n \mu_{jl}^C(x_l^C)}, \quad (4)$$

обеспечивающая выполнение стандартного условия

$$\sum_{j=1}^m \bar{w}_j^C = 1. \quad (5)$$

В четвертом скрытом слое рассчитывается $3m$ функций консеквента, как правило, линейных (система Такаги – Сугено – Канга первого порядка [15, 16])

$$f_j^C(x^C) = p_{j0}^C + \sum_{l=1}^n p_{jl}^C x_l^C \quad (6)$$

и производится их умножение на нормализованные уровни принадлежности, вычисленные в третьем слое, в виде $\bar{w}_j^C f_j^C(x)$.

В пятом скрытом слое, образованном тремя сумматорами, для каждого цветового канала вычисляются отфильтрованные оценки

$$\tilde{x}^C = \sum_{j=1}^m \bar{w}_j^C f_j^C(x^C) = \sum_{j=1}^m \frac{\prod_{l=1}^n \mu_{jl}^C(x_l^C)}{\sum_{j=1}^m \prod_{l=1}^n \mu_{jl}^C(x_l^C)} f_j^C(x^C), \quad (7)$$

которые в результате операции конкатенации S в выходном (шестом) слое формируют выход фильтра \tilde{x}^C для каждого пикселя обрабатываемого изображения.

Таким образом, рассматриваемая нейро-фаззи система производит фильтрацию цветных изображений, искаженных разного рода возмущениями и помехами.

3. ОБУЧЕНИЕ НЕЙРО-ФАЗЗИ ФИЛЬТРА

Обучение рассматриваемого нейро-фаззи фильтра состоит в нахождении $3(n+1)m$ параметров p_{jl}^C ($j = 1, 2, \dots, m; l = 0, 1, \dots, n; C = R, G, B$) моделей консеквента (6), обеспечивающих наилучшее в смысле принятого критерия обучения качество обработки изображения.

Переписывая выражение (7) в виде

$$\begin{aligned} \tilde{x}^C &= \sum_{j=1}^m \bar{w}_j^C f_j^C(x_i^C) = \sum_{j=1}^m \bar{w}_j^C \left(p_{j0}^C + \sum_{l=1}^n p_{jl}^C x_l^C \right) = \\ &= \bar{w}_1^C (p_{10}^C + p_{11}^C x_1^C + \dots + p_{1n}^C x_n^C) + \\ &+ \bar{w}_2^C (p_{20}^C + p_{21}^C x_1^C + \dots + p_{2n}^C x_n^C) + \dots + \\ &+ \bar{w}_m^C (p_{m0}^C + p_{m1}^C x_1^C + \dots + p_{mn}^C x_n^C), \end{aligned} \quad (8)$$

вводя векторы преобразованных входов моделей $f^C(x^C) = (\bar{w}_1^C, \bar{w}_1^C x_1^C, \dots, \bar{w}_1^C x_n^C, \bar{w}_2^C, \bar{w}_2^C x_1^C, \dots, \bar{w}_2^C x_n^C, \dots, \bar{w}_m^C, \bar{w}_m^C p_{m1}^C, \dots, \bar{w}_m^C p_{mn}^C)^T$ и оцениваемых параметров $p^C = (p_{10}^C, p_{11}^C, \dots, p_{1n}^C, p_{20}^C, p_{21}^C, \dots, p_{2n}^C, p_{m0}^C, p_{m1}^C, \dots, p_{mn}^C)^T$ размерности $(n+1)m$, можно записать обобщенную модель консеквента C -го цвета в виде

$$\tilde{x}^C = p^{CT} f^C(x^C), \quad (9)$$

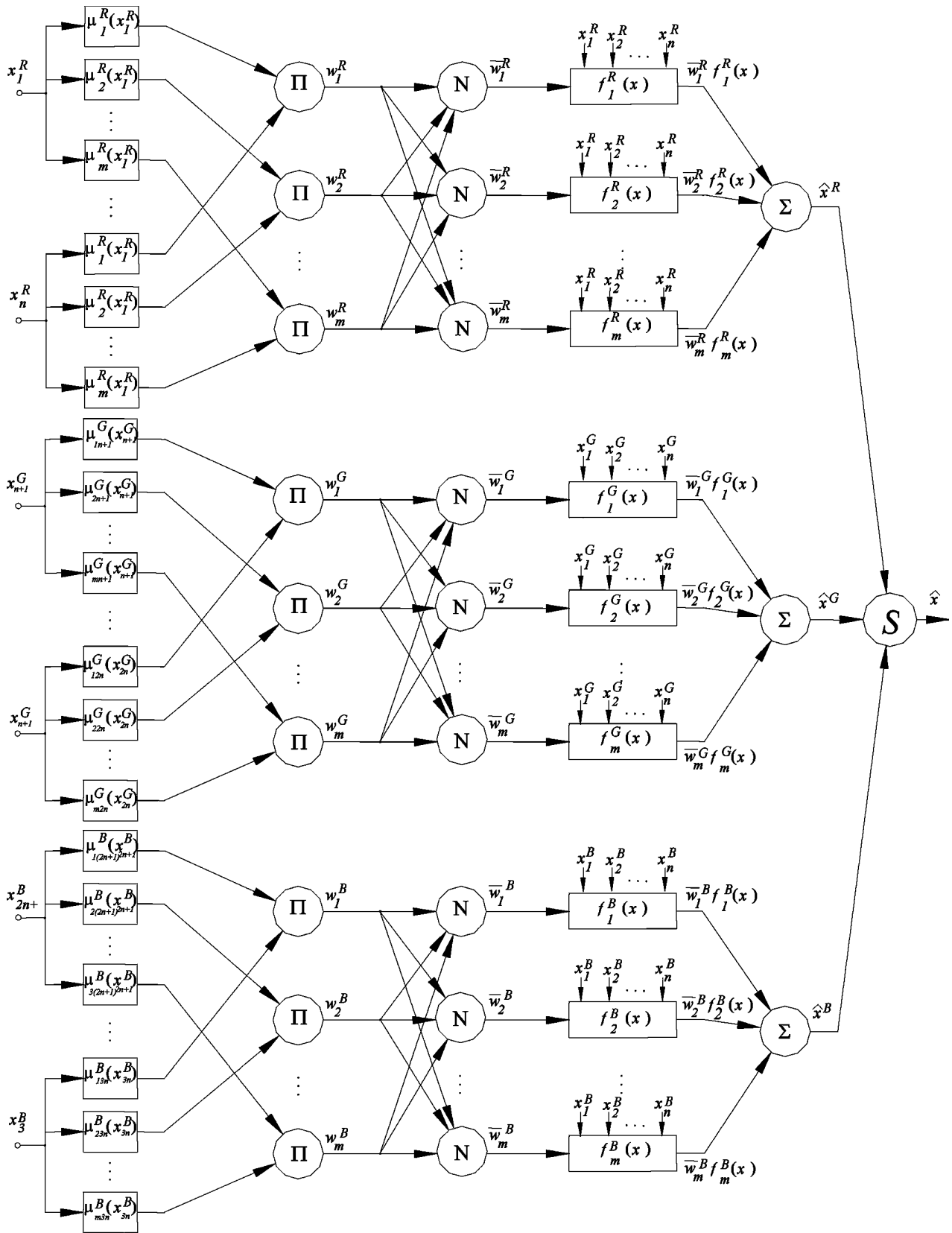


Рис. 1. Архитектура нейро-фаззи фільтра

которая для k -го пикселя обрабатываемого изображения принимает вид

$$\tilde{x}^c(k) = p^{CT} f^c(x^c(k)), \quad k = 1, 2, \dots, N. \quad (10)$$

Минимизация стандартного квадратичного критерия обучения

$$E^C = \sum_{k=1}^N (\tilde{x}^c(k) - p^{CT} f^c(x^c(k)))^2 \quad (11)$$

ведет к обычной оценке наименьших квадратов:

$$p^C = \left(\sum_{k=1}^N f^c(x^c(k)) f^{CT}(x^c(k)) \right)^+ \left(\sum_{k=1}^N f^c(x^c(k)) x^c(k) \right) \quad (12)$$

(здесь $(\bullet)^+$ – символ псевдообращения по Муру – Пенроузу), использованной, в частности, в нейро-фаззи системе обработки изображений [16].

Необходимость обработки нестационарных сигналов в реальном времени не позволяет применять пакетный алгоритм (12), в котором все данные учитываются с одинаковыми весами. В этом случае более предпочтительными являются адаптивные алгоритмы с конечной памятью, наиболее популярным из которых является экспоненциально взвешенный рекуррентный метод наименьших квадратов в форме

$$\begin{cases} p^c(k+1) = p^c(k) + \\ + \frac{D(k)(x^c(k+1) - p^{CT}(k)f^c(x^c(k+1)))f^c(x^c(k+1))}{\beta + f^{CT}(x^c(k+1))D(k)f^c(x^c(k+1))}, \\ D(k+1) = \frac{1}{\beta} \left(D(k) - \frac{D(k)f^c(x^c(k+1))f^{CT}(x^c(k+1))D(k)}{\beta + f^{CT}(x^c(k+1))D(k)f^c(x^c(k+1))} \right), \end{cases} \quad (13)$$

где $0 < \beta \leq 1$ – параметр «забывания» устаревшей информации.

Громоздкость этого алгоритма, а также возможная численная неустойчивость («взрыв параметров») при малых значениях β делают более предпочтительным использование градиентных процедур, в частности, адаптивного алгоритма обучения нейро-фаззи сетей [17–19]

$$\begin{cases} p^c(k+1) = p^c(k) + \beta^{-1}(k+1) \times \\ \times (x^c(k+1) - p^{CT}(k)f^c(x^c(k+1)))f^c(x^c(k+1)), \\ \beta(k+1) = \beta\beta(k) + \|f^c(x^c(k+1))\|^2, \quad 0 \leq \beta \leq 1, \end{cases} \quad (14)$$

устойчивого при любых значениях параметра β и обеспечивающего компромисс между фильтрующими и следящими свойствами. Заметим также, что при $\beta = 0$ (14) превращается в оптимальный по быстродействию алгоритм обучения Уидроу – Хоффа, а при $\beta = 1$ – в процедуру стохастической аппроксимации Гудвина – Рэмеджа – Кэйнеса.

4. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ

Эффективность предложенного нейро-фаззи фильтра была исследована с помощью 262×394 цветного изображения (рис. 2, а). Изображение было искажено мультипликативным шумом (рис. 2, б), для его подавления применены линейный сглаживающий, медианный и нейро-фаззи фильтры. Для всех фильтров выбрано окно размерами 5×5 пикселей. Результаты фильтрации показаны на рис. 2, в, г, д.

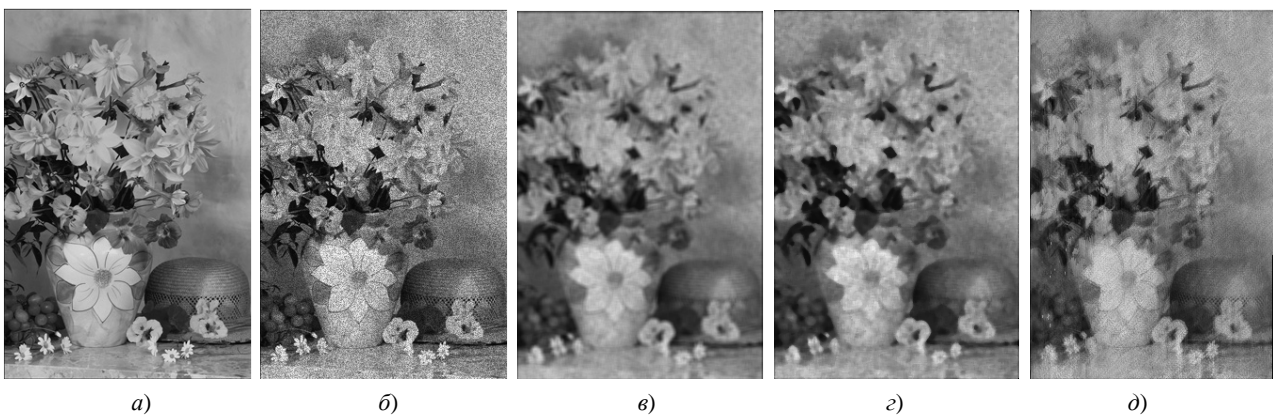


Рис. 2. Оригинальное изображение (а), изображение с помехами (б) и результаты фильтрации: линейное сглаживание (в); медианная фильтрация (г); ANFIS-фильтрация (д)

Поскольку в данной работе рассматривается обработка изображений на допечатной стадии полиграфического процесса, можно провести визуальную оценку результатов фильтрации. На рис. 2 можно наблюдать, что линейный сглаживающий фильтр сильно искажает изображение, размывая контуры, а при наблюдении всего изображения в целом можно отметить наличие размытых пятен помехи. Медианная фильтрация подавляет помеху, уничтожая при этом и малые детали, и мало заметные плавные тоновые переходы, делая изображение слишком искусственным. Результат применения нейро-фаззи фильтра демонстрирует остаточные проявления помехи при увеличении изображения, однако в целом на изображении эти следы не так заметны по сравнению с линейным сглаживанием. В то же время после ANFIS-фильтрации сохраняются мелкие детали в виде тонких штрихов, которые были потеряны при медианной фильтрации.

ВЫВОДЫ

Предложенный нейро-фаззи фильтр обеспечивает достаточно высокое качество подавления помех на цветных цифровых изображениях, характеризуется высокой скоростью обучения и может найти применение в решении широкого круга практических задач. В частности, фильтр можно использовать при подготовке к полиграфической печати фотографий, сделанных цифровыми фотоаппаратами.

СПИСОК ЛИТЕРАТУРЫ

1. Ализар А. А. Фильтры и эффекты для цифровой фотографии / А. А. Ализар, К. Г. Афанасьев. – СПб. : Питер, 2007. – 144 с.
2. Фрейзер Б. Усиление резкости фотографий в Adobe Photoshop / Б. Фрейзер. – М. : Вильямс, 2007. – 304 с., ил.
3. Stanco F. Towards the automated restoration of old photographic prints: A survey/ A. de Polo, G. Ramponi, F. Stanco // IEEE Int. Conf. EUROCON 2003, Ljubljana, Slovenia, Sept. 22–24. – Ljubljana, 2003. – P. 213–217.
4. Гладких В. Сканирование фотоленки и динамический диапазон / В. Гладких // DigitalWare. – 2003. – № 1. – P. 67–71.
5. Варламов А. Д. Исследование и классификация дефектов архивных документов/ С. С. Садыков, А. Д. Варламов // Технологии информационного общества – Интернет и современное общество : труды VII Всеросс. Конференции. – С. Пб. : Изд-во Филологического ф-та СПбГУ, 2004. – С. 119–121.
6. Buades A. A review of image denoising algorithms, with a new one / A. Buades, B. Coll, J. M. Morel // SIAM Multiscale Modeling and Simulation. – 2005. – V. 4. – P. 490–530.
7. Gonzalez R. C. Digital Image Processing, 2nd ed. / R. C. Gonzalez, R. E. Woods. – Prentice Hall, NJ : Upper Side River, 2002. – 616 p.
8. Pitas I. Order statistics in digital image processing / I. Pitas, A. N. Venetsanopoulos // Proc. IEEE. – 1992. – 80. – № 12. – Pp. 1892–1919.
9. Perry S. W. Adaptive Image Processing: A Computational Intelligence Perspective / S. W. Perry, H.-S. Wong, L. Guan. – Boca Raton ; London ; New York ; Washington, DC : CRC Press LLC, 2002. – 258 p.
10. Lukac R. Vector filtering for color imaging / R. Lukac, K. N. Plataniotis, A. N. Venetsanopoulos, B. Smolka, K. Martin // IEEE Signal Processing Mag. – 2005. – 22. – № 1. – P. 74–86.
11. Chatterjee P. Clustering-based denoising with locally learned dictionaries / P. Chatterjee, P. Milanfar // IEEE Trans. on Image Processing. – 2009. – 18. – № 7. – P. 1438–1451.
12. Morillas S. Fuzzy peer groups for reducing mixed gaussian impulse noise from color images/ S. Morillas, V. Gregori, A. Hervas // IEEE Trans. on Image Processing. – 2009. – 18. – № 7. – P. 1452–1466.
13. Jang J.-S. R. Neuro-Fuzzy Modeling: Architectures, Analyses and Applications / J.-S. R. Jang. – Berkeley, CA : University of California, 1992. – 155 p.
14. Jang J.-S. R. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence / J.-S. R. Jang, C.-T. Sun, E. Mizutani. – Upper Saddle River, NJ : Prentice Hall, 1997. – 614 p.
15. Takagi T. Fuzzy identification of systems and its application to modeling and control / T. Takagi, M. Sugeno // IEEE Trans. on Systems, Man and Cybernetics. – 1985. – 15. – P. 116–132.
16. Thanh N. M. Image denoising using adaptive neuro-fuzzy system / N. M. Thanh, M.-S. Chen. // IAENG Int. J. of Applied Mathematics. – 2007. – 36. – P. 1–7.
17. Bodyanskiy Ye. An adaptive learning algorithm for a neuro-fuzzy network / Ye. Bodyanskiy, V. Kolodyazhniy, A. Stephan // Computational Intelligence. Theory and Applications / [Ed. By B. Reusch]. – Berlin ; Heidelberg ; New York : Springer, 2001. – P. 68–75.
18. Otto P. A new learning algorithm for a forecasting neuro-fuzzy network / P. Otto, Ye. Bodyanskiy, V. Kolodyazhniy / Integrated Computer-Aided Engineering. – 2003. – 10. – № 4. – P. 399–409.
19. Digital Color Imaging Handbook / [ed. By G. Sharma]. – Boca Raton ; London ; New York ; Washington DC, 2003. – 764 p.

Надійшла 15.04.2010

Бодяньський Є. В., Кулішова Н. Є.

ЗГЛАДЖУВАННЯ ЗАВАД НА КОЛЬОРОВИХ ЗОБРАЖЕННЯХ ЗА ДОПОМОГОЮ НЕЙРО-ФАЗЗИ ФІЛЬТРА

Розглянуто проблему заглушення завад на кольорових цифрових зображеннях. Для вирішення проблеми запропоновано нейро-фаззі фільтр, який забезпечує високу якість фільтрації із зберіганням контурів, текстур та плавних тонових переходів.

Ключові слова: кольорові зображення, обчислювальний інтелект, нейро-фаззі фільтрація, завада.

Bodyanskiy Ye. V., Kulishova N. Ye.

COLOR IMAGES DISTURBANCE SMOOTHING USING NEURO-FUZZY FILTER

The problem of color images disturbance smoothing is considered. A neuro-fuzzy filter is proposed for solving the problem. This filter provides high quality of filtration, saving contours, textures and smooth tone gradations.

Key words: color images, computational intellect, neuro-fuzzy filtering, disturbance.

МЕТОДЫ И МОДЕЛИ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ОБЪЕКТОВ ПО ПРИЗНАКАМ НА ОСНОВЕ ИММУНОКОМПЬЮТИНГА

Проведен анализ основных моделей искусственных иммунных систем. Экспериментально оценена работа различных видов детекторов применительно к решению задачи автоматической классификации сельскохозяйственных растений по результатам дистанционного зондирования. Предложен метод классификации растений на основе моделей иммунокомпьютинга, позволяющий проводить обучение на экземплярах только одного класса и обеспечивающий необходимый уровень точности распознавания.

Ключевые слова: искусственная иммунная система, отрицательный отбор, Hypercell.

ВВЕДЕНИЕ

Проблема автоматизации разделения объектов на классы по признакам актуальна для разнообразных приложений в технической и биомедицинской диагностике, а также находит свое применение в растениеводстве при решении проблемы борьбы с сорными растениями [1].

При традиционном подходе обработка пропашных культур ядохимикатами для уничтожения сорных растений осуществляется путем поверхностного внесения гербицидов. При этом большое количество ядохимикатов попадает в окружающую среду, загрязняя почву, поверхностные и подземные воды, а также попадая через культурные растения и влияя негативным образом на человека и животных. При этом эффективность традиционного подхода является весьма невысокой вследствие дороговизны ядохимикатов, потерь из-за загрязнения окружающей среды, а также способности ядохимикатов бороться преимущественно только с растениями одного класса (например, в случае использования гербицидов селективного действия необходимо классифицировать однодольные и двудольные сорняки).

В [1] предлагается способ обработки пропашных культур, направленный на индивидуальную обработку сорных растений и основанный на использовании автоматической классификации растений по их биологическим особенностям, проявляющимся при дистанционном зондировании в спектрах отражения. В качестве признака для процедуры распознавания принимается измеряемая величина, соответствующая коэффициенту спектральной яркости (КСЯ).

Ранее для решения задачи разделения растений на культурные и сорные применялись многослойные [2] и радиально-базисные [3] нейронные сети. Недостаток полученных моделей заключается в том, что методы их обучения требуют наличия экземпляров каждого класса, в т. ч. и сорных растений. Однако получение информации обо всех возможных видах сорных растений в разных фазах роста для построения распознающих моделей является затруднительным.

В качестве системы классификации растений перспективным является использование моделей искусственных иммунных систем (ИИС), поскольку одна из основных задач биологической иммунной системы – разделение объектов на классы, а именно на «свои» (self) и «чужие» (non-self) клетки. При этом существует возможность обучения с помощью экземпляров только одного класса. Используя иммунную память, иммунная система поддерживает идеальный баланс между экономией ресурсов и выполнением своих задач за счет сохранения минимально необходимых знаний, получаемых в процессе обучения. Поскольку формирование иммунного ответа всегда происходит случайным образом, детектор может определять инородные тела даже в тех случаях, когда не существует явных представлений об их структуре.

Исходя из этого, для решения задачи классификации растительных объектов по результатам дистанционного зондирования авторами предлагается использование ИИС, которые обладают свойствами обучаемости распознаванию образов даже в случае сложно делимых классов [4].

Цель работы заключается в построении такой модели ИИС для классификации растений, которая позволяла бы выполнять наиболее точное распознавание

с минимальными временными затратами, основываясь на экземплярах растений только одного класса.

1. ПОСТАНОВКА ЗАДАЧИ

Пусть задана выборка $X \subset B$, каждый экземпляр которой описывается n признаками $x^s = \{x_j^s\}$, где s – номер экземпляра, j – номер признака, $x^s \in X$. Для экземпляров x^s обучающей выборки известны номера классов экземпляров $y = \{y^s\}$. Тогда задача автоматической классификации может быть представлена как задача построения модели зависимости номера класса y от значений признаков экземпляра $x \in B$ по обучающей выборке данных $\langle x, y \rangle$.

Применительно к задаче классификации растений на основе методов иммунокомпьютинга необходимо синтезировать соответствующую модель ИИС, которую в общем виде можно описать кортежем $\langle B, M, A \rangle$, где B – область допустимых значений входных данных, M – множество детекторов, представляющих иммунную память, A – правило, позволяющее распознавать экземпляры, полагаясь на знания, хранящиеся в иммунной памяти.

В рассматриваемой задаче множество B определяет возможные значения спектров растений, а классификацию необходимо провести для выделения культурных и сорных (однодольных и двудольных) растений.

2. СИНТЕЗ ИММУННОЙ МОДЕЛИ

Построение иммунного классификатора на основе ИИС, как правило, происходит в три этапа [5].

Вначале следует определить способ описания структуры исследуемого объекта. Их существует два – с помощью вектора вещественных чисел и с помощью битовых строк.

Следующим этапом будет определение меры для оценки подобия (аффинности) двух объектов. В качестве данной меры используются различные метрики. В случае вектора вещественных чисел аффинность элементов можно измерить, используя Евклидово расстояние или расстояние Манхэттена, для битовых строк мерой аффинности выступает расстояние Хемминга, tsb -метрика, метрика g -chunks.

Завершающим этапом является выбор модели ИИС, которая и будет проводить процесс классификации. Среди основных моделей ИИС выделяют модель клональной селекции, модель отрицательного отбора и искусственные иммунные сети.

Поскольку описание признаков в рассматриваемой задаче определяется значениями КСЯ, при проведении синтеза иммунной модели для классифика-

ции растений в качестве представления детектора предлагается использовать вектор вещественных чисел. В качестве меры аффинности в таком случае используется Евклидово расстояние, т. к. эта метрика отличается простотой и наглядностью. «Своими» экземплярами будем считать культурные растения, а «чужими» – сорные.

Исходные данные, использованные для оценки работоспособности методов, с разрешения автора взяты из [6], а сравнимые показатели качества распознавания – из работы [7]. Особенностью решаемой задачи является то, что исходное пространство признаков характеризуется высокой размерностью – значения КСЯ измерялись в 55 диапазонах. Кроме того, в случае исходной выборки ограниченного объема (248 экземпляров), сложно выполнить эффективное обучение, поскольку при малом объеме обучающей выборки, которая является частью исходной, снижается вероятность построения адекватной модели для распознавания.

Для решения задачи использовались две основные иммунные модели – модель отрицательного отбора и модель клональной селекции. Рассмотрим особенности каждой из них.

Модель отрицательного отбора была предложена С. Форрест [8]. Основным компонентом данной модели является специфическая область иммунной памяти, хранящая информацию только о «своих» клетках. В биологической иммунной системе ей соответствует тимус. Работа модели отрицательного отбора происходит в два этапа и включает в себя обучение и классификацию. В процессе обучения создаются случайные детекторы, и выполняется сопоставление каждого сгенерированного детектора набору «своих» клеток. В случае совпадения новый детектор удаляется из популяции. Детекторы, не совпадающие ни с одной из «своих» клеток, пополняют иммунную память и становятся частью ИИС, определяющей область «чужих» клеток. Создание случайных детекторов и их последующий отбор проводится до тех пор, пока не будет получено требуемое число детекторов. Для обеспечения необходимого уровня точности распознавания производится вероятностная оценка количества детекторов.

Процесс классификации устроен таким образом, что поступающая клетка считается «чужой», если она попадает в область детекторов иммунной памяти, и «своей» – в противном случае.

Модель отрицательного отбора широко используется в задачах обнаружения аномалий, например, сетевых атак, при двухклассовом распознавании образов и во многих других приложениях [9]. Это объясняется

простотой, гибкостью и относительно высоким быстродействием данной модели.

Модели, основанные на теории клональной селекции, способны улучшать свои характеристики во время обучения за счет этапа «усиления аффинности» [10]. Модели клональной селекции, как правило, реализуют парадигму положительного отбора, т. е. детекторы, хранящиеся в иммунной памяти, описывают пространство «своих» клеток. Метод обучения модели включает в себя следующие этапы.

1. Задается обучающая выборка Ag , состоящая из антигенов. Антигенами будем считать такие экземпляры из обучающей выборки, которые принадлежат к классу «своих». Создается временная популяция из n случайных клеток $D = \{d_1, d_2, \dots, d_n\}$, $d_i = \text{rand}()$, где $\text{rand}()$ – функция, возвращающая вектор $x \in B$, состоящий из случайных элементов, $i = 1, \dots, n$.

2. Рассчитывается матрица взаимной аффинности элементов обучающей выборки и детекторов в популяции $Aff = D \times Ag$.

3. Детекторы популяции подвергаются клонированию и мутации, при этом для детектора создается тем больше клонов, чем выше его аффинность антигенам (что приводит к повышению вероятности покрытия антигена), а степень мутации созданных клонов обратно пропорциональна аффинности (что снижает область поиска для более оптимальных детекторов). Процесс клонирования приводит к созданию некоторого числа детекторов, идентичных родительскому. Процесс мутации случайным образом изменяет эти детекторы, как правило, с помощью случайной или неравномерной (non-uniform) мутации.

4. Из полученных клонов отбираются элементы с максимальным уровнем аффинности, и они заменяют родительский детектор в популяции.

Элементы из популяции с наиболее высоким уровнем аффинности переносятся в иммунную память, после чего этапы 1–4 повторяются, пока не будет достигнут критерий останова. В качестве критерия останова, как правило, применяют ограничение на время обучения, на максимальное количество итераций, на размер популяции или на предельное количество неэффективных итераций (т. е. итераций, которые не приводят к добавлению новых детекторов в память).

Благодаря высокой скорости сходимости метода обучения, модель клональной селекции находит широкое применение в задачах оптимизации, поскольку она нечувствительна к локальным оптимумам. Ее использование в качестве классификатора также оправдано

высоким быстродействием и способностью обучаться на экземплярах только одного класса («своего»).

Применительно к задаче распознавания растений обучающую выборку получаем из исходной путем случайного выбора из нее около трети экземпляров (от 81 до 87). Однако построение модели зависимости номера класса от значений всех 55 признаков экземпляра выборки не дало удовлетворительных результатов. Даже при различных параметрах (критериях останова, максимальном размере популяции, пороге минимальной аффинности, различных схемах мутации) показатели точности классификации оставались в диапазоне 27–33 %.

Причина такого снижения эффективности распознавания описана в [11]. Как известно, объем гиперсферы радиусом r в n -мерном пространстве можно вычислить как

$$V(n, r) = \frac{r^n \pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)},$$

где $\Gamma(n+1) = n!$, $n \in N$ и $\Gamma\left(n + \frac{1}{2}\right) = 1, 3, 5, \dots, \frac{2n-1}{2^n \sqrt{n}}$, для вещественных аргументов.

Из данной формулы видно, что объем n -мерной гиперсферы стремится к нулю, если размерность стремится к бесконечности. Также установлено, что при больших размерностях точки, равномерно распределенные в гиперсфере, преимущественно сосредоточены в тонком слое у поверхности гиперсферы. Это объясняет тот факт, что, например, гиперсфера единичного радиуса обладает сравнительно большим объемом по отношению к радиусу лишь при размерности пространства менее 15.

В работе [12] предлагалось использовать в качестве меры аффинности нормы большего порядка, чем у Евклидовой метрики. С другой стороны, такой подход усилит влияние шумов при обучении. Незначительные отклонения будут многократно усилены за счет вознесения в большую степень.

Таким образом, для построения модели зависимости классов растений от значений КСЯ является актуальным снижение размерности задачи за счет отбора признаков. Поскольку оценка информативности признаков в случае использования ИИС тесно связана с характеристиками иммунной модели, которая будет применяться во время классификации, то для отбора признаков был использован подход «настройки» (wrapper). В частности, применялся метод QuickReduct [13], поскольку он прост в реализации и приводит к хорошим результатам при небольшом количестве

Таблиця 1. Результаты одномерной классификации растений с помощью моделей ИИС

Модель ИИС	I		II		III	
	№ признака	E, %	№ признака	E, %	№ признака	E, %
Отрицательный отбор	4	27,82	3	32,26	1	33,47
Клональная селекция	3	32,26	1	33,47	5	41,94

Таблиця 2. Результаты классификации растений с помощью моделей ИИС по двум признакам

Модель ИИС	I		II		III	
	№ признака	E, %	№ признака	E, %	№ признака	E, %
Отрицательный отбор	3, 8	41,94	2, 3	42,34	2, 9	43,55
Клональная селекция	1, 6	11,69	1, 9	12,90	3, 7	16,13

признаков в результирующем множестве. При этом данный метод в процессе работы не генерирует избыточное количество больших наборов признаков, а, начиная с пустого множества, «жадным» способом добавляет по одному признаку в набор до тех пор, пока добавление очередного признака не перестанет улучшать точность классификации.

В табл. 1 приведены результаты одномерной классификации растений с помощью моделей ИИС. В столбцах, описывающих точность классификации, указаны значения для трех наиболее информативных признаков с точки зрения моделей: указывается номер признака, по которому проводилась классификация, и процент ошибочных решений E.

Анализируя полученные данные, следует отметить, что поскольку обе модели ИИС в равной степени хорошо справляются с разделением нелинейно разделимых классов, их использование приводит к близким результатам. В частности, одномерная классификация позволила установить, что наилучшее разделение по одному признаку происходит при анализе спектров низких частот. Однако количество ошибочных решений в этом случае достаточно высоко, следовательно, требуется выполнить отбор еще одного признака для классификации.

Проведенные эксперименты показали, что дальнейший перебор комбинаций признаков является вычислительно сложным. Поскольку все признаки однотипные, то снижение размерности задачи возможно за счет выполнения линейной свертки групп по пяти признакам.

В табл. 2 приведены результаты классификации растений по двум признакам. Обучающая выборка отбиралась аналогично предыдущему случаю.

Очевидно, что модель отрицательного отбора показывает гораздо менее точные результаты по сравнению с моделью клональной селекции. Этот факт

объясняется тем, что размер допустимого пространства признаков значительно превышает область, покрываемую «своими» детекторами.

Так, если радиус детектора r , то область, покрываемая им в двухмерном пространстве имеет площадь $\pi \cdot r^2$. В то время как допустимая область значений признаков нормирована и её площадь равна единице. В рассматриваемой выборке насчитывается в среднем около 30 уникальных (отдаленных более чем на r) «своих» экземпляров. Соответственно, при пороге аффинности, равном радиусу детектора, их площадь составляет $30 \cdot \pi \cdot r^2$. Иными словами, для полного покрытия «своей» области детекторами радиуса r необходимо около 30 детекторов, а для покрытия чужой области такими же детекторами их количество должно быть около 100, если принять радиус детектора равным 0,05.

Увеличивая радиус детектора, можно восстановить баланс между «своими» и «чужими» областями, однако в этом случае ошибка классификации все равно останется высокой, поскольку увеличение радиуса детектора неизбежно приводит к снижению точности распознавания.

3. ОПТИМИЗАЦИЯ МОДЕЛИ ИИС ДЛЯ РЕШАЕМОЙ ЗАДАЧИ

Для использования парадигмы отрицательного отбора в дальнейших исследованиях необходимо использовать улучшенные модели, которые позволили бы сократить число детекторов без существенного снижения точности распознавания. В следующих тестах будут применяться модели V-Detector [14] и Hypercell.

Основная идея первой модели заключается в том, что детекторы могут обладать варьирующимися ха-

рактическими, в частности, для вещественных детекторов такой характеристикой является их радиус.

Метод обучения такой модели может выглядеть следующим образом.

1. Установить иммунную память $M = \emptyset$.
2. Сгенерировать случайный детектор d радиусом r_d .
3. Определить расстояние до ближайшего антигена r_{ag} .
4. Уменьшить радиус детектора таким образом, чтобы он не перекрывал ни один из антигенов $r_d = r_d - r_{ag}$.
5. Если $r_d > 0$, т. е. если детектор не перекрывает ни один антиген, то добавить его в иммунную память $M = M \cup \{d\}$.

Изложенный метод предполагает, что большие участки, не принадлежащие к «своей» области, будут покрыты более крупными детекторами, в то время как более детальные контуры области будут «уточняться» детекторами меньшего радиуса. Данный метод отличается высокой скоростью сходимости и при этом обеспечивает высокую точность распознавания.

Альтернативой для такой модификации модели отрицательного отбора может служить предложенная авторами модель Hypercell, которая представляет собой объект, объединяющий свойства детектора и модели и способный автономно принимать решение о том, какие свойства в данный момент лучше применять. Когда Hypercell ведет себя как детектор, она может вычислять аффинность другим детекторам и антигенам. В случае поведения модели, Hypercell может создавать внутри себя другие Hypercell и контролировать их популяцию за счет метадинамики – генерации новых детекторов или удаления неэффективных. Таким образом, модель Hypercell представляет собой древовидную иерархию, в которой листья имитируют детекторы, а остальные узлы – модели. Древовидная структура определяет сильные стороны модели Hypercell,

такие как высокая скорость работы, снижение негативного эффекта оверфиттинга (запоминания) за счет отсека избыточных узлов дерева. Двойственное поведение элементов дерева обеспечивает адаптивность модели и её способность обучаться. В табл. 3 и 4 приведены результаты классификации растений с помощью описанных выше моделей по одному и двум признакам, соответственно.

Исходя из табл. 3 и 4, рассмотренные выше модели отличаются высокой скоростью и эффективностью по сравнению с моделью отрицательного отбора. В то же время они так же обучаются с использованием экземпляров одного класса и толерантны к «своим» клеткам.

Для подтверждения качества обучения моделей был использован метод случайных подвыборок [15]. Идея этого метода заключается в том, что для обучения модели формируется подвыборка путем случайного независимого отбора экземпляров исходной выборки, при этом вероятность отбора для каждого экземпляра одинакова. Поскольку используются различные части выборки, то случайные закономерности в различных подвыборках отбрасываются, в то время как действительно устойчивые закономерности на различных подвыборках только «подтверждаются».

Такое преобразование исходной выборки позволило оценить способность моделей аппроксимировать входные данные. Выяснилось, что в моделях клональной селекции и V-Detector наблюдалось явление оверфиттинга. Рассмотрим причины этого явления.

Основная идея модели клональной селекции заключается в генерации случайных детекторов таким образом, чтобы на каждой итерации расстояние между детекторами в иммунной памяти и экземплярами обучающей выборки сокращалось. При большом количестве итераций такой подход приводит к тому, что детекторы памяти фактически «запоминают» обучающую выборку, что ухудшает способность модели искать зако-

Таблица 3. Результаты одномерной классификации растений с помощью моделей V-Detector и Hypercell

Модель ИИС	I		II		III	
	№ признака	$E, \%$	№ признака	$E, \%$	№ признака	$E, \%$
V-Detector	3	32,26	2	33,06	1	33,47
Hypercell	2	32,26	3	33,87	4	37,90

Таблица 4. Результаты классификации растений с помощью моделей V-Detector и Hypercell по двум признакам

Модель ИИС	I		II		III	
	№ признака	$E, \%$	№ признака	$E, \%$	№ признака	$E, \%$
V-Detector	1, 3	5,26	1, 4	10,89	3, 7	14,11
Hypercell	1, 3	8,06	1, 2	10,48	3, 8	16,94

номерности в выборке. Снизить влияние этого эффекта можно, уменьшив число итераций.

Модель V-Detector склонна к переобучению, поскольку плотность покрытия пространства признаков детекторами постоянна и не зависит от концентрации антигенов. Иными словами, любой детектор покрывает область, вплотную прилегающую к его зоне. В процессе обучения покрытие детекторов описывает все пространство за исключением элементов обучающей выборки, т. е. происходит явление оверфиттинга, при котором запоминаются области, комплементарные обучающей выборке (дополняющие её до области допустимых значений *B*). Возможным решением этой проблемы может стать случайная вариация радиуса детектора – увеличивать радиус не до момента обнаружения антигена, а на некоторую случайную величину, близкую к данной.

Данные обучения с учетом предложенных модификаций приведены в табл. 5. Отбрасывая заведомо непригодные признаки, в тесте выполнялся отбор их комбинаций в группах 0, 1 и 4. Как видно из табл. 5, исследуемые модели ИИС показывают хорошие результаты классификации для решаемой задачи, что позволяет рекомендовать их для применения на практике. Следует отметить, что ошибка классификации в большей степени зависит от правильности распознавания экземпляров «чужого» класса. Объясняется это тем, что обучение производится с помощью экземпляров культурных растений, т. е. «своего» класса, а значит и ошибка распознавания таких растений минимальна (как правило, 1–2 экземпляра). Наиболее информативные признаки, согласно результатам, соответствуют длинам волн в 310–320, 360–370, 410 нм.

Поскольку иммунная система по своей природе предназначена для проведения двухклассовой классификации, наиболее очевидное решение проблемы разделения сорных растений на однодольные и двудольные заключается в использовании метода дихотомии. Поэтому вначале необходимо выделить только сорные растения. Эти данные являются исходными для построения еще одной модели ИИС. «Своими» экземплярами можно считать однодольные растения, а «чужими» – двудольные. Несмотря на свою простоту, такой подход существенно сокращает репрезентативность обучающей выборки, формируемой из исходной выборки малого объема, что в условиях решаемой задачи недопустимо.

Вместо этого предлагается построение нескольких моделей. Первая производит распознавание однодольных сорных растений («своих») и остальных («чужих») растений. Вторая модель отличает только двудольные растения от прочих, а третья определяет культурные растения. Каждый экземпляр тестовой выборки проходит классификацию с помощью всех трех моделей. Таким образом, получаем многослойную иммунную систему, в которой каждый слой принимает решение об отнесении экземпляра к определенному классу. Решение является однозначным, если только одна модель отнесла экземпляр к «своему» классу, в противном случае система не может классифицировать объект, и, соответственно, определить класс можно только с определенной вероятностью пропорционально значениям меры аффинности различных моделей.

В табл. 6 приведены результаты классификации сорных растений на однодольные и двудольные с применением описанной выше схемы.

Таблица 5. Результаты классификации растений с помощью модифицированных моделей по двум признакам

Модель ИИС	I		II		III	
	№ признака	<i>E</i> , %	№ признака	<i>E</i> , %	№ признака	<i>E</i> , %
V-Detector	1, 7	8,87	1, 2	9,68	1,6	10,08
Hypercell	1, 6	8,06	1, 10	10,08	1, 2	10,48
Клональная селекция	1,7	8,06	1, 6	10,08	1, 23	10,48

Таблица 6. Результаты классификации сорных растений на однодольные и двудольные с помощью многослойной модели ИИС

Модель ИИС	I		II		III	
	№ признака	<i>E</i> , %	№ признака	<i>E</i> , %	№ признака	<i>E</i> , %
V-Detector	1, 6	10,48	1, 2	10,89	1, 7	11,69
Hypercell	1, 7	10,48	1, 6	11,29	1, 23	12,10
Клональная селекция	1, 7	11,29	1, 2	11,69	7, 23	12,90

Полученные результаты позволяют рекомендовать модели V-Detector, Hypercell и клональной селекции для классификации растений не только на культурные и сорные, но и для дальнейшего разделения сорных растений на однодольные и двудольные.

Вышеизложенные методы были программно реализованы в виде библиотеки для синтеза моделей иммунокомпьютинга JAIMS (Java Artificial Immune Systems). Библиотека написана на языке программирования Java и позволяет выполнять построение различных моделей иммунокомпьютинга, а также проводить их обучение и использование на практике. Данная библиотека использовалась для проведения экспериментов, результаты которых приведены в табл. 3 и 4.

Проведенные эксперименты подтвердили работоспособность разработанного математического обеспечения и позволяют рекомендовать его для использования на практике, в частности, при решении задачи автоматизации классификации растений.

ЗАКЛЮЧЕНИЕ

С целью решения актуальной проблемы автоматизации процесса классификации растительных объектов по признакам разработано математическое обеспечение, позволяющее синтезировать распознающие модели на основе методов иммунокомпьютинга.

Научная новизна работы заключается в том, что впервые предложен метод классификации растительных объектов на основе модели Hypercell. Получили дальнейшее развитие модели V-Detector и клональной селекции, использующие для обучения экземпляры только одного класса растений (культурные на первом уровне дихотомии, сорные однодольные – на втором), что позволяет помимо разделения на сорные и культурные классифицировать также однодольные и двудольные сорные растения. Предложенный метод дает высокую точность распознавания растений.

Практическая ценность работы состоит в том, что разработана программа для классификации объектов на основе иммунокомпьютинга, а также получены модели принятия решений в задачах дистанционного зондирования сельскохозяйственных растений с помощью отрицательного отбора и клональной селекции, позволяющие определять зависимость номера класса растений от спектральных признаков.

Разработанное математическое обеспечение успешно прошло стадию тестирования, обладает высокими показателями точности и может быть рекомендовано для использования на практике. Использование способа [1] в совокупности с предложенным

математическим обеспечением позволит сократить расход ядохимикатов для уничтожения сорных растений, снизить их влияние на окружающую среду, повысить урожайность культурных растений.

Дальнейшие исследования могут быть сосредоточены на исследовании влияния различных мер аффинности на точность классификации растений с помощью моделей иммунокомпьютинга, а также на повышении скорости построения модели Hypercell за счет нахождения оптимального соотношения размеров детекторов на разных уровнях иерархии модели.

Работа выполнена в рамках госбюджетной научно-исследовательской работы кафедры программных средств Запорожского национального технического университета «Информационные технологии автоматизации распознавания образов и принятия решений для диагностики в условиях неопределенности на основе гибридных нечеткологических, нейросетевых и мультиагентных методов вычислительного интеллекта» (номер государственной регистрации 0109U007673).

СПИСОК ЛИТЕРАТУРЫ

1. Патент 27425 Україна МПК²⁰⁰⁶ А 01 В 39/00. Спосіб прополки просапних культур і пристрій для його здійснення / Д. М. Піза, О. Д. Піза, І. М. Татарчук. – Запор. держ. техн. ун-т. Заявл. 17.12.1992. Заявка № 94010291. Опубл. 15.09.2000. Бюл. № 4, 2000. – 6 с.
2. Дубровин В. И. Распознавание растений по результатам дистанционного зондирования на основе многослойных нейронных сетей / В. И. Дубровин, С. В. Морщавка, Д. М. Пиза, С. А. Субботин // Математические машины и системы. – 2000. – № 2, 3. – С. 113–119.
3. Дубровин В. И. Применение радиально-базисных нейронных сетей для обработки данных дистанционного зондирования растений / В. И. Дубровин, С. В. Морщавка, Д. М. Пиза, С. А. Субботин // 3-я Международная конференция и выставка «Цифровая обработка сигналов и ее применение»: сборник трудов. Москва, 29 ноября – 1 декабря 2000 г. – М.: РНТОРЭС им. А. С. Попова, 2000. – Т. 3. – С. 48–53.
4. Искусственные иммунные системы и их применение / под ред. Д. Дасгупты; пер. с англ. А. А. Романюхи. – М.: Физматлит, 2006. – 344 с.
5. Tarakanov A. O. Immunocomputing: Principles and Applications / A. O. Tarakanov, V. A. Skormin, S. S. Sokolova. – New York: Springer-Verlag, 2003. – P. 195.
6. Морщавка С. В. Радиоэлектронная система распознавания растительных объектов: дис. ... канд. техн. наук: 05.12.17 / Морщавка Сергей Владимирович. – Запорожье, 2004. – 162 с.
7. Піза Д. М. Вибір ефективного методу розпізнавання рослин за коефіцієнтами спектральної яскравості / Д. М. Піза, С. В. Морщавка, Ю. В. Скоробогатов // Радіоелектроніка. Інформатика. Керування. – 1999. – Vol. 1. – С. 83–87.
8. Forrest S. Self-nonsel self discrimination in a computer / S. Forrest, A. S. Perelson, L. Allen, R. Cherukuri // IEEE Computer Society Symposium «Research in Security and Privacy»: Proceedings, Oakland, May, 16–18, 1994. – Los Alamitos: IEEE Computer Society Press, 1994. – P. 202–212.
9. Stibor T. A Comparative Study of Real-Valued Negative Selection to Statistical Anomaly Detection / T. Stibor,

- J. Timmis, C. Eckert // 4th International Conference on Artificial Immune Systems (ICARIS): Proceedings / Eds.: C. Jacob et al., Banff, Alberta, Canada, August, 14–17, 2005. – Berlin : Springer-Verlag, 2005. – P. 262–275.
10. *De Castro L. N.* The clonal selection algorithm with engineering applications / L. N. de Castro, F. J. von Zuben // Genetic and Evolutionary Computation Conference (GECCO'00) : proceedings, Las Vegas, July 2000. – California : Morgan Kaufmann Publishers, 2000. – P. 36–37.
 11. *Stibor T.* On the Use of Hyperspheres in Artificial Immune Systems as Antibody Recognition Regions / T. Stibor, J. Timmis, C. Eckert // 5th International Conference on Artificial Immune Systems (ICARIS): Proceedings / Eds.: H. Bersini, J. Carneiro, Oeiras, September 4–6, 2006. – New York : Springer-Verlag, 2006. – P. 215–228.
 12. *Verleysen M.* Learning high-dimensional data / M. Verleysen // Limitations and future trends in neural computation. – 2003. – Vol. 186. – P. 141–162.
 13. *Jensen R.* Computational Intelligence and Feature Selection: Rough and Fuzzy Approaches / R. Jensen, Q. Shen. – UK : IEEE Press and Wiley & Sons, 2008. – P. 340.
 14. *Ji Z.* V-Detector: An Efficient Negative Selection Algorithm with «Probably Adequate» Detector Coverage / Z. Ji, D. Dasgupta // Information Sciences. – 2009. – Vol. 179. – P. 1390–1406.
 15. *Buhlmann P.* Analysing Bagging / P. Buhlmann, B. Yu // The Annals of Statistics. – 2002. – Vol. 30. – P. 927–961.

Надійшла 29.03.2010

УДК 681.3.06

Халимов Г. З.

Канд. техн. наук, доцент Харківського національного університету радіоелектроніки

ОЦЕНКА ЧИСЛА РЕШЕНИЙ УРАВНЕНИЯ ГУРВИЦА В КОНЕЧНОМ ПОЛЕ

Представлены результаты по оценкам числа решений уравнения Гурвица в конечном поле и практический алгоритм нахождения решений.

Ключевые слова: кривая Гурвица.

ВВЕДЕНИЕ

Алгеброгеометрический подход при решении ряда задач в кодировании и криптографии связывается с полем рациональных функций на многообразиях, которые во многих случаях определяются точками алгебраических кривых. Можно отметить, например, традиционные исследования по эллиптическим кривым в криптографии цифровой подписи или по кривым Эрмита, Сузуки для задач универсального хеширования и помехоустойчивого кодирования. Кривые Гурвица относятся к классу замечательных кривых с большим числом точек.

Основные результаты по кривым Гурвица представлены в работах F. Torres [1, 2], P. Carbonne,

Зайцев С. О., Субботин С. О.

МЕТОДИ ТА МОДЕЛІ АВТОМАТИЧНОЇ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ ЗА ОЗНАКАМИ НА ОСНОВІ ІМУНОКОМП'ЮТИНГУ

Проведено аналіз основних моделей штучної імунної системи. Експериментально оцінено роботу різних видів детекторів у вирішенні задачі автоматичної класифікації сільськогосподарських рослин за результатами дистанційного зондування. Запропоновано метод класифікації рослин на основі моделей імунокомп'ютингу, що дозволяє проводити навчання на екземплярах лише одного класу та забезпечує необхідний рівень точності розпізнавання.

Ключові слова: штучна імунна система, негативна селекція, Hypercell.

Zaitsev S. A., Subbotin S. A.

METHODS AND MODELS OF AUTOMATED OBJECTS CLASSIFICATION BY FEATURES BASED ON IMMUNOCOMPUTING

Basic artificial immunity models have been analyzed. Experiments were carried out and various detectors were estimated by solving the problem of plants recognition on remote sensing results. The method based on the immunocomputing principles was proposed. It allows to perform learning procedure using single class exemplars and ensures necessary recognition accuracy.

Key words: artificial immune system, negative selection, Hypercell.

T. Nenocq в [3], R. Pellikan, P. Beelen [4, 5] и автора статьи [6, 7]. В работе [1] введено определение обобщенных кривых Гурвица и установлен морфизм между обобщенными кривыми Гурвица и Ферма. Здесь же определены условия максимальности для обычных кривых Гурвица и обобщенных кривых при ограничении на выбор показателей степени кривой. Связь между кривыми Гурвица и Ферма представлена P. Carbonne, T. Nenocq в [3]. В работе [4] предложена техника построения кривых на основе формального полинома и определен класс кривых Гурвица, как обобщение кватрики Клейна. В работе [5] представлены соотношения для рода кривой. Некоторые оценки числа решений кривой Гурвица для про-

© Халимов Г. З., 2010

извольного конечного поля представлены в [6]. Теорема о существовании нетривиальных кривых Гурвица и правила построения нетривиальных кривых также получены в [6]. Впервые получена максимальная кривая Гурвица с третьим значением рода для максимальных кривых в [7].

Целью статьи является нахождение оценок для числа решений уравнения Гурвица в конечном поле. В разделе 1 приводятся определения и основные результаты по кривым Гурвица. В разделе 2 получены оценки числа точек для кривых Гурвица в конечном поле. В разделе 3 представлен практический алгоритм вычисления числа точек кривой Гурвица.

1. ОПРЕДЕЛЕНИЕ И ОСНОВНЫЕ РЕЗУЛЬТАТЫ ПО КРИВЫМ ГУРВИЦА В КОНЕЧНОМ ПОЛЕ

Кривые Гурвица H_n определяются выражением

$$X^n Y + Y^n Z + XZ^n = 0 \tag{1}$$

и имеют частные производные вида

$$\begin{aligned} F_X &= nX^{n-1}Y + Z^n, & F_Y &= nY^{n-1}Z + Z^n, \\ F_Z &= nZ^{n-1}X + Y^n. \end{aligned} \tag{2}$$

Несингулярность кривых Гурвица над F_q определяется условиями вида:

- 1) n и q должны быть взаимно простыми;
- 2) $\gcd(n^2 - n + 1, q) = 1$.

Для вычисления рода кривой Гурвица H_n используем выражение рода для кривой вида

$$X^a + X^b Y^c + Y^d = 0. \tag{3}$$

В работе [4] (замечание 43) доказывается, что для несингулярной модели кривой (2) справедливо

$$g \leq 1 + \frac{1}{2} \{ |ac + bd - ad| - \gcd(a - b, c) - \gcd(b, c - d) - \gcd(a, d) \}. \tag{4}$$

Если характеристика поля F_q не делит $\gcd(a - b, c)$, $\gcd(b, c - d)$, $\gcd(a, d)$ и $ac + bd - ad$, тогда справедливо равенство. В силу соотношения (4) для кривой Гурвица H_n выражение для рода имеет следующий вид

$$g = \frac{n^2 - n}{2}. \tag{5}$$

Существует обобщение кривых Гурвица $H_{n,l}$, которое имеет вид

$$X^n Y^l + Y^n Z^l + X^l Z^n = 0, \tag{6}$$

где $n \geq l \geq 2$, $\Delta(n, l) = n^2 - nl + l^2 \geq 2$ и частные производные

$$\begin{aligned} F_X &= nX^{n-1}Y^l + lX^{l-1}Z^n; & F_Y &= nY^{n-1}Z^l + lX^n Y^{l-1}; \\ F_Z &= nZ^{n-1}X^l + lY^n Z^{l-1}. \end{aligned} \tag{7}$$

Несингулярность кривых Гурвица $H_{n,l}$ над F_q определяется условием [1]:

$$(\Delta(n, l), \text{char}(F_q)) = 1. \tag{8}$$

Род кривой $H_{n,l}$, как следует из (4) и отмечается в [5], равен

$$g = \frac{n^2 - nl + l^2 + 2 - 3\gcd(n, l)}{2}. \tag{9}$$

Справедливо следующее утверждение.

Утверждение 1. В простом поле F_q все кривые Гурвица H_n являются несингулярными.

Доказательство следует из того, что всегда выполняются условия: $\gcd(n, q) = 1$ и $\gcd(n^2 - n + 1, q) = 1$.

2. ОЦЕНКИ ЧИСЛА ТОЧЕК ДЛЯ КРИВЫХ ГУРВИЦА В КОНЕЧНОМ ПОЛЕ

Ниже рассмотрим основные свойства кривых Гурвица для случая произвольного конечного поля F_q , прежде всего оценки для числа точек кривых.

Следующая теорема определяет число решений кривой Гурвица для произвольного конечного поля.

Теорема 1. Пусть кривая $X^n Y^l + Y^n Z^l + X^l Z^n = 0$ определена над конечным полем F_q и является несингулярной. Пусть $\gcd(n, l, q - 1) = c > 1$ и $\gcd(n^2 - nl + l^2, c(q - 1)) = c^2 d$, тогда имеем следующую оценку для числа точек кривой Гурвица:

$$N = tc^2 d + 3, \tag{10}$$

где $0 \leq t \leq q - 1$.

Доказательство. Решения для кривой $X^n Y^l + Y^n Z^l + X^l Z^n = 0$ есть точки проективного пространства P^2 : $(\gamma:\zeta:1)$, $\gamma, \zeta \in F_q$ и $\gamma \neq 0, \zeta \neq 0$, удовлетворяющие уравнению

$$\gamma^n \zeta^l + \zeta^n + \gamma^l = 0. \tag{11}$$

Выражение (11) сократим на $\gamma^l \neq 0$

$$\gamma^{n-l} \zeta^l + \zeta^n \gamma^{-l} + 1 = 0. \tag{12}$$

Пусть $\gamma = \beta^i, \zeta = \beta^j$, где β – образующий элемент поля F_q , тогда

$$\beta^{i(n-l)+jl} + \beta^{-il+jn} + 1 = 0. \tag{13}$$

Равенство (13) относительно неизвестных i и j имеет решения, если

$$\begin{cases} i(n-l) + jl \equiv a \pmod{(q-1)}; \\ -il + jn \equiv b \pmod{(q-1)}; \\ \beta^a + \beta^b + 1 = 0. \end{cases} \quad (14)$$

Система тождественных уравнений (14) первой степени относительно неизвестных i и j для значений a и b , удовлетворяющих условию $\beta^a + \beta^b + 1 = 0$, имеет следующие решения.

Пусть $\gcd(n, l, q-1) = c > 1$, тогда левые и правые части уравнений системы (14) можно сократить на значение c :

$$\begin{cases} \frac{i(n-l)}{c} + j\frac{l}{c} \equiv \frac{a}{c} \pmod{\frac{(q-1)}{c}}; \\ -i\frac{l}{c} + j\frac{n}{c} \equiv \frac{b}{c} \pmod{\frac{(q-1)}{c}}. \end{cases} \quad (15)$$

Выразим неизвестные i и j значения в уравнениях системы (15). После элементарных преобразований система тождественных уравнений преобразуется к системе линейных уравнений

$$\begin{cases} i\frac{(n^2-nl+l^2)}{c^2} = \frac{an-bl}{c^2} + \frac{k_1n-k_2l(q-1)}{c}; \\ j\frac{(n^2-nl+l^2)}{c^2} = \frac{al+b(n-l)}{c^2} + \frac{k_1l+k_2(n-l)(q-1)}{c}. \end{cases} \quad (16)$$

Множество коэффициентов k_1 и k_2 определяют все решения для i и j при зафиксированных значениях a и b . Так как c делит n и l , систему линейных уравнений (16) преобразуем к системе тождественных уравнений по модулю $(q-1)c$:

$$\begin{cases} i(n^2-nl+l^2) \equiv an-bl \pmod{c(q-1)}; \\ j(n^2-nl+l^2) \equiv al+b(n-l) \pmod{c(q-1)}. \end{cases} \quad (17)$$

Пусть $\gcd(n^2-nl+l^2, c(q-1)) = c^2d$, тогда для зафиксированных значений a и b , удовлетворяющих уравнению $\beta^a + \beta^b + 1 = 0$ и условию делимости $an-bl$ и $al+b(n-l)$ на c^2d , получим c^2d решений. Пусть число пар степеней a и b , удовлетворяющих условию делимости, равно t . С учетом особых точек $(1:0:0)$, $(0:1:0)$, $(0:0:1)$, число точек кривой $H_{n,l}$ будет равно $N = tc^2d + 3$.

Частные результаты по кривым Гурвица представлены следствиями 1–4.

Следствие 1. Если $\gcd(n^2-nl+l^2, q-1) = 1$ и $\gcd(n, l) = 1$, тогда число точек несингулярной кривой Гурвица $H_{n,l}$ будет равно

$$N = q + 2. \quad (18)$$

Действительно, если $\gcd(n^2-nl+l^2, q-1) = 1$, а степени n и l не имеют общих сомножителей, тогда при зафиксированных значениях a и b уравнения (7) имеют единственное решение. Общее число пар степеней a и b , удовлетворяющих условию $\beta^a + \beta^b + 1 = 0$, равно $q-1$. С учетом особых точек $(1:0:0)$, $(0:1:0)$, $(0:0:1)$, число точек кривой $H_{n,l}$ будет равно $N = q + 2$.

Данный случай описывает кривые Гурвица $H_{n,l}$ с тривиальным числом точек, равным $q + 2$, и родом

$$g = \frac{n^2-nl+l^2-1}{2}.$$

Следствие 2. Если $\gcd(n^2-nl+l^2, q-1) = 1$ и $\gcd(n, l) = c$, тогда число точек несингулярной кривой Гурвица $H_{n,l}$ будет равно

$$N = tc^2 + 3. \quad (19)$$

Так как $\gcd(n, l) = c$ и $\gcd(n^2-nl+l^2, q-1) = 1$, тогда $\gcd(n^2-nl+l^2, c(q-1)) = c^2$.

Для зафиксированных значений a и b уравнения (15) имеют c^2 решений, если a и b делятся на c . Пусть число пар степеней a и b , удовлетворяющих условию делимости, равно t . Тогда число точек кривой $H_{n,l}$ будет равно $N = tc^2 + 3$.

В данном случае род кривых Гурвица $H_{n,l}$ равен $g = \frac{n^2-nl+l^2+2-3c}{2} = \frac{c(c\delta-3)}{2} + 1$, где $\delta = \frac{n^2-nl+l^2}{c^2} > 1$.

Замечание 1. Если $\gcd(n^2-nl+l^2, q-1) = 1$ и $\gcd(n, l) = c$, тогда кривая Гурвица $H_{n,l}$ покрывается кривой Ферма Fr_c $X^c + Y^c + Z^c = 0$, рода $g = \frac{(c-1)(c-2)}{2}$. Следует это из оценок для числа точек кривых. Род кривых Ферма Fr_c меньше рода кривых Гурвица $H_{n,l}$ данного типа.

Следствие 3. Пусть $\gcd(n^2-nl+l^2, q-1) = d$ и $\gcd(n, l) = 1$, тогда число точек несингулярной кривой Гурвица $H_{n,l}$ будет равно

$$N = td + 3. \quad (20)$$

Пусть $\gcd(n^2-nl+l^2, q-1) = d$ и $\gcd(n, l) = 1$. Для фиксированных значений a и b уравнения (5) имеют d решений, если $an-bl$ и $al+b(n-l)$ делятся на d . Пусть число пар степеней a и b , удовлетворяющих условию делимости, равно t . Тогда число точек кривой $H_{n,l}$ будет равно $N = td + 3$.

Данный случай кривых является наиболее интересным. Докажем следующие полезные леммы.

Лемма 1. Пусть $n > 0$ есть целое число и $\Delta(n, l) = n^2 - n + 1$ имеет простые делители $d > 3$. Тогда $d \equiv 1 \pmod 6$.

Пусть d – простое число есть делитель $\Delta(n, l)$. Тогда

$$n^2 - n + 1 \equiv 0 \pmod d. \quad (21)$$

Применим модульную операцию к членам суммы $\Delta(n, l)$. С учетом подстановки $n \equiv n_d \pmod d$ получим

$$n_d^2 - n_d + 1 \equiv 0 \pmod d. \quad (22)$$

Вычисления по модулю простого числа можно свести к вычислениям в конечном поле F_{d_i} . Пусть α – образующий элемент поля и пусть $n_d = \alpha^s$, тогда имеем уравнение

$$\alpha^{2s} - \alpha^s + 1 = 0. \quad (23)$$

Так как $\alpha^{\frac{d-1}{2}} = -1$, получим

$$\alpha^{2s} + \alpha^{\frac{d-1}{2}} \alpha^s + 1 = 0. \quad (24)$$

Трехчлен (24) с единицей равен нулю, если его элементы образуют мультипликативную подгруппу третьего порядка $1, \beta, \beta^2$. Здесь $\beta = \alpha^{2s}, \beta^2 = \alpha^{\frac{d-1}{2} + s}$. Получим решение относительно параметра $s = \frac{d-1}{6}$. Таким образом, каждый делитель $\Delta(n, l)$, если он больше 3, имеет свойство $d_i \equiv 1 \pmod 6$.

Лемма 2. Пусть $n > 1$ есть целое число и $\Delta(n, l) = n^2 - n + 1$. Тогда одним из делителей $\Delta(n, l)$ может быть простой делитель, равный 3.

Действительно, пусть $d = 3$ есть делитель $\Delta(n, l)$. Тогда с учетом подстановки $n \equiv n_d \pmod 3$ получим

$$n_d^2 - n_d + 1 \equiv 0 \pmod 3. \quad (25)$$

Решение уравнения (25) возможно только в случае, если $n_d = 1$ и $n_d = 2$. Из уравнения (24) следует, что $s = 1$ и $\alpha^2 + \alpha^1 + 1 = 0$ справедливо для подгруппы второго порядка.

Лемма 3. Пусть $n > 0$ есть целое число и $\Delta(n, l) = n^2 - n + 1$ имеет степенные делители $d^e, d > 3$ и $d \equiv 1 \pmod 6$.

Пусть d^e – делитель $\Delta(n, l)$. Тогда

$$n^2 - n + 1 \equiv 0 \pmod{d^e}. \quad (26)$$

Применим модульную операцию к членам суммы $\Delta(n, l)$. С учетом подстановки $n \equiv n_d \pmod{d^e}$ получим

$$n_d^2 - n_d + 1 \equiv 0 \pmod{d^e}. \quad (27)$$

Вычисления по модулю d^e можно свести к вычислениям в кольце целых чисел F_{d^e} , порядка $p = d^e - d^{e-1}$. Пусть α – образующий элемент кольца и пусть $n_d = \alpha^s$, тогда имеем уравнение

$$\alpha^{2s} - \alpha^s + 1 = 0. \quad (28)$$

Так как $\alpha^{\frac{p}{2}} = -1$, получим

$$\alpha^{2s} + \alpha^{\frac{p}{2}} \alpha^s + 1 = 0. \quad (29)$$

Трехчлен (29) с единицей равен нулю, если его элементы образуют мультипликативную подгруппу третьего порядка $1, \beta, \beta^2$. Здесь $\beta = \alpha^{2s}, \beta^2 = \alpha^{\frac{p}{2} + s}$. Получим решение относительно параметра $s = \frac{p}{6}$. По условию $p \equiv 0 \pmod 6$, в силу того, что $d \equiv 1 \pmod 6$, в разложении $\Delta(n, l)$ делитель d^e может иметь место. \diamond

Обобщение полученных результатов на случай, когда $l \neq 1$ в параметре $\Delta(n, l) = n^2 - nl + l^2$, представлено в лемме 4.

Лемма 4. Пусть $n, l > 0$ есть целые взаимно простые числа, $\gcd(n, l) = 1$ и $\Delta(n, l) = n^2 - nl + l^2$. Тогда одним из делителей $\Delta(n, l)$ может быть простой делитель, равный 3, простые делители $d > 3$ со свойством $d \equiv 1 \pmod 6$, а также степенные делители $d^e, d > 3$ и $d \equiv 1 \pmod 6$.

Пусть d – есть делитель $\Delta(n, l)$. Рассмотрим выражение

$$n^2 - nl + l^2 \equiv 0 \pmod d. \quad (30)$$

Применим модульную операцию к членам суммы $\Delta(n, l)$. С учетом подстановки $n \equiv n_d \pmod d, l \equiv l_d \pmod d$ получим

$$n_d^2 - n_d l_d + l_d^2 \equiv 0 \pmod d. \quad (31)$$

Вычисления по модулю числа d можно свести к вычислениям в конечном поле F_{d_i} . Пусть α – образующий элемент поля и пусть $n_d = \alpha^s, l_d = \alpha^t$, тогда имеем уравнение

$$\alpha^{2s} - \alpha^s \alpha^t + \alpha^{2t} = 0. \quad (32)$$

Сделаем замену переменных $I = s - t$. После преобразований с учетом $\alpha^{\frac{d-1}{2}} = -1$ получим

$$\alpha^{2t} (\alpha^{2i} + \alpha^{d-1-i} \alpha^i + 1) = 0. \quad (33)$$

Таблица 1. Уравнения кривых Гурвица и оценки для числа точек и рода в конечном поле F_q

Уравнение кривой Гурвица	Свойства кривой	Число точек кривой N	Род кривой g
$X^n Y + Y^n Z + XZ^n$	$\gcd(n^2 - n + 1, q - 1) = 1$	$q + 2$	$n(n - 1)/2$
$X^n Y^l + Y^n Z^l + X^l Z^n$	$\gcd(n^2 - nl + l^2, q - 1) = 1; \gcd(n, l) = 1$	$q + 2$	$(n^2 - nl + l^2 - 1)/2$
$X^n Y^l + Y^n Z^l + X^l Z^n$	$\gcd(n^2 - nl + l^2, q - 1) = 1; \gcd(n, l) = c$	$tc^2 + 3$	$(n^2 - nl + l^2 + 2 - 3c)/2$
$X^n Y + Y^n Z + XZ^n$	$\gcd(n^2 - n + 1, q - 1) = d$	$td + 3$	$n(n - 1)/2$
$X^n Y^l + Y^n Z^l + X^l Z^n$	$\gcd(n^2 - nl + l^2, q - 1) = d; \gcd(n, l) = 1$	$td + 3$	$(n^2 - nl + l^2 - 1)/2$
$X^n Y^l + Y^n Z^l + X^l Z^n$	$\gcd(n^2 - nl + l^2, c(q - 1)) = c^2 d;$ $\gcd(n, l, (q - 1)) = c > 1$	$tc^2 d + 3$	$(n^2 - nl + l^2 + 2 - 3c)/2$

Трехчлен в выражении (33) с единицей равен нулю, если его элементы образуют мультипликативную подгруппу третьего порядка $1, \beta, \beta^2$. Выполнение условий существования мультипликативной подгруппы третьего порядка определяются в леммах 1 ÷ 3, и свойства делителей $\Delta(n, l)$ переносятся на данную лемму.

Параметр $\Delta(n, l) = n^2 - nl + l^2$ имеет следующие свойства:

1. $\Delta(n, l) = \Delta(l, n)$. (34)

2. $\Delta(n, l) = \Delta(n, n - l)$. (35)

3. Если $\gcd(n, l) = 1$ и $p > n > 1$ является делителем $\Delta(n, l)$, тогда

$$\Delta(n, l) = \Delta(p - n, p - l). \quad (36)$$

4. Если $\gcd(n, l) = 1$ и p является делителем $\Delta(n, l)$, тогда $\Delta(n', l')$ будет иметь делители не выше $p > n > l$, где $n \equiv n' \pmod{p}, l \equiv l' \pmod{p}$.

5. $\Delta(cn, cl) = c^2 \Delta(n, l)$. (37)

Рассмотрим доказательство свойства 4. Так как $n', l' < p$, возьмем верхнюю оценку $n' = l' = p$. Получим $\Delta(n', l') = p^2 - pp + p^2 = p^2$. Отсюда следует, что один делитель p , другие не больше чем p .

Доказательства свойств 1–3, 5 простые.

Основные результаты по оценкам для числа точек и рода для кривых Гурвица представлены в табл. 1.

В табл. 1 представлено все многообразие кривых Гурвица, которое вытекает из результатов теоремы 1.

3. ПРАКТИЧЕСКИЙ АЛГОРИТМ ВЫЧИСЛЕНИЯ ЧИСЛА ТОЧЕК КРИВОЙ ГУРВИЦА

В соответствии с теоремой 1, практический алгоритм вычисления числа точек кривой Гурвица включает следующие шаги:

1. Для заданной кривой $X^n Y^l + Y^n Z^l + X^l Z^n = 0$ в конечном поле F_q вычисляем параметр $\Delta(n, l) = n^2 - nl + l^2$ и определяем по результатам вычислений $\gcd(n, l, q - 1)$ и $\gcd(n^2 - nl + l^2, q - 1)$, к какому множеству кривых Гурвица относится кривая по классификации, представленной в табл. 1, и по какой оценке определяется число точек.

2. Для конечного поля F_q с элементами β^i подсчитываем число пар степеней a и b образующего элемента β мультипликативной группы порядка $q - 1$, которые удовлетворяют уравнению $\beta^a + \beta^b + 1 = 0$ и условию делимости $an - bl$ и $al + b(n - l)$ на $\Delta(n, l)$. В результате получим t пар. С учетом особых точек $(1:0:0), (0:1:0), (0:0:1)$ вычисляем значение для числа точек кривой Гурвица $N = t\Delta(n, l) + 3$.

ВЫВОДЫ

1. Оценки для числа точек кривых Гурвица, представленные теоремой 1, являются новыми, и при доказательстве теоремы 1 используется конструктивный алгоритм вычисления числа точек кривой Гурвица в конечном поле.

2. Сложность вычислений практического алгоритма определяется размерностью конечного поля, что существенно меньше перебора всех решений по точкам проективного пространства.

СПИСОК ЛИТЕРАТУРЫ

1. Torres F. Plan maximal curves / Torres F // Acta Arithmetica. – 2001. – Vol. 98, No. 2. – P. 165–179.
2. Cossidente A. Curves of large genus covered by the Hermitian curve / Cossidente A., Korchm'aros G. and Torres F. // Commutative Algebra. – 2000. – Vol. 28, No. 10. – P. 4707–4728.
3. Carbonne P. Decomposition de la Jacobienne sur les corps finis / Carbonne P., Henocq T. // Bulletin Polish Academy of Sciences Mathematics. – 1994. – Vol. 42, No. 3. – P. 207–215.
4. Pellikan R. The Klein quartic, the Fano plan and curves representing design / Pellikan R. // In Codes, Curves and

- Signals: Common Threads in Communications. – Dordrecht : Kluwer Academy Publication, 1998. – P. 9–20.
5. *Beelen P.* The Newton polygon of plane curves with many rational points / Beelen P. and Pellikan R. // Designs, Codes and Cryptography. – 2000. – No. 21. – P. 41–67.
 6. *Халимов Г. З.* Оценка параметров кривых Гурвица для целей универсального хеширования / Халимов Г. З. // Сборник трудов I Международной научно-технической конференции «Компьютерные науки и технологии» (Белгород, Россия, 8–10 октября 2009). – 2009. – Ч. 2. – С. 118–121.
 7. *Халимов Г. З.* Максимальные кривые Гурвица для целей универсального хеширования/ Халимов Г. З. // Материалы XI Международной научно-практической конференции «Информационная безопасность» (Таганрог, Россия, 23–25 июня 2010), ТТИ ЮФУ. – 2010. – Ч. 3. – С. 144–146.

Надійшла 12.04.2010

Халимов Г. З.
ОЦІНКА ЧИСЛА РІШЕНЬ РІВНЯННЯ ГУРВИЦА
В КІНЦЕВОМУ ПОЛІ

Представлено результати по оцінках числа рішень рівняння Гурвица в кінцевому полі та практичний алгоритм знаходження рішень.

Ключові слова: крива Гурвица.

Khalimov G. Z.
ESTIMATE OF HURVITZ EQUATION SOLUTIONS
NUMBER IN FINITE FIELD

The author presents the results of estimation of Hurvitz equation solutions number in finite field and practical algorithm of solutions finding.

Key words: Hurvitz curve.

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

PROGRESSIVE INFORMATION TECHNOLOGIES

УДК 004.03:004.051

Брагина Т. И.¹, Табунщик Г. В.²¹Студент Запорожского национального технического университета²Канд. техн. наук, доцент Запорожского национального технического университета

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИТЕРАТИВНЫХ МОДЕЛЕЙ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

В статье выполнен обзор моделей разработки программных проектов, их жизненного цикла, достоинств и недостатков. Проведен анализ наиболее распространенных моделей: Rational Unified Process, Microsoft Solutions Framework, Personal Software Process / Team Software Process и Agile (eXtreme Programming, Crystal, Feature Driven Development). Сделаны выводы относительно факторов, влияющих на выбор модели, и предложены рекомендации по выбору соответствующей модели программного обеспечения в зависимости от характеристик проекта.

Ключевые слова: программный проект, стандарты качества, факторы выбора модели.

Методология управления проектами включает совокупность моделей, методов и программных продуктов, применяемых при разработке и реализации проектов различных классов и типов. Специфика проектов по разработке программных продуктов заключается в том, что результат разработки нематериален – это коллективные ментальные модели, записанные на языке программирования. Это приводит к тому, что большая часть проектов разработки программного обеспечения (ПО) завершается со срывами сроков, перерасходом бюджета, а часть проектов не завершается в принципе. Вышесказанное подтверждают исследования Standish Group [1], которые показали, что только 35 % проектов завершились в срок.

Одной из главных задач, напрямую влияющих на эффективность разработки ПО, является выбор модели процесса разработки. Проблемой является то, что не существует единого оптимального решения. Модель должна определяться для каждого конкретного

проекта и может меняться в широком диапазоне, в зависимости от масштаба, новизны и критичности проекта, распределения участников, требований заказчика. На выбор модели также влияет возможность дальнейшей сертификации, что позволяет компании-разработчику получить преимущество перед конкурентами, привлечь новых заказчиков, повысить имидж организации.

Поэтому выбор модели процесса разработки ПО является очень актуальной проблемой при управлении проектом и коллективом разработчиков.

СРЕДСТВА ОЦЕНКИ КАЧЕСТВА МОДЕЛЕЙ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Существуют десятки различных подходов к обеспечению качества ПО, некоторые из которых (ГОСТ 28195-89, ГОСТ 19.102, DO-178B, ISO 9000 и др.)

© Брагина Т. И., Табунщик Г. В., 2010

постепенно устарели и содержат различные недостатки, например:

- недостаточная подробность стандарта, возможность самых различных его толкований в зависимости от представлений аудитора;

- неточность оценки качества процессов, задействованных при создании и внедрении программного обеспечения;

- отсутствие в стандарте механизмов, способствующих улучшению существующих процессов.

В начале 90-х годов прошлого столетия эксперты разработали более совершенные решения в области обеспечения качества, что привело к созданию целого ряда новых стандартов и методологий. Далее опишем три широко используемых в США и Европе стандарта – ISO 9001:2008, Capability Maturity Model (CMM) и ISO/IEC 15504 (SPICE), и государственные стандарты, используемые на территории стран СНГ.

Стандарт ISO 9001:2008. Стандарт ISO 9001:2008 был принят Международной организацией по стандартизации 14 ноября 2008 года. Он направлен на применение «процессного подхода» при разработке, внедрении и улучшении результативности системы менеджмента качества с целью повышения удовлетворенности потребителей путем выполнения их требований [2]. Для успешного функционирования организация должна определить и осуществлять менеджмент многочисленных взаимосвязанных видов деятельности. Деятельность, использующая ресурсы и управляемая с целью преобразования входов в выходы, может рассматриваться как процесс. Преимущество процессного подхода состоит в непрерывности управления, которое он обеспечивает на стыке отдельных процессов в рамках их системы, а также при их комбинации и взаимодействии.

В отличие от других, стандарт ISO 9001 имеет самую большую общность и зону применения. Однако ISO9001 затрагивает узкую область, оставляя за рамками своего внимания другие проблемы (не описываются связь между организационной поддержкой и реализацией проекта, роль измерения в системе управления качеством, необходимость постоянной оптимизации управления) процесса производства ПО и менеджмента этого процесса.

При применении в системе менеджмента качества процессный подход подчеркивает важность понимания и выполнения требований; необходимости рассмотрения процессов с точки зрения добавляемой ими ценности; достижения запланированных результатов выполнения процессов и обеспечения их результативности; постоянного улучшения процессов, основанного на объективном измерении.

С точки зрения многих западных и японских компаний, соответствие требованиям ISO 9001 – это крайне низкий уровень гарантий качества, однако это тот минимальный уровень, который дает возможность вхождения в рынок.

В Украине приказом Госпотребстандарта Украины от 22.06.2009 № 225 утвержден национальный стандарт ДСТУ ISO 9001:2009 «Системы управления качеством. Требования» (ISO 9001:2008, IDT) с введением в действие 01.09.2009.

Модель зрелости Capability Maturity Model for Software (SW-CMM). В середине 80-х годов XX столетия по заказу Министерства обороны США Институт программной инженерии (Software Engineering Institute – SEI), входящий в состав Университета Карнеги – Меллона, разработал стандарт SW-CMM [3] в качестве эталонной модели организации разработки ПО при реализации крупномасштабных программных проектов.

Данная модель определяет пять уровней зрелости процесса разработки ПО, каждый следующий уровень включает в себя все ключевые характеристики предыдущих: Начальный, Повторяемый, Определенный, Управляемый, Оптимизируемый.

В результате аттестации компании присваивается определенный уровень, который в дальнейшем может повышаться или (теоретически) понижаться. Данный стандарт предполагает возможность сертификации только одного процесса или подразделения организации.

Использование CMM затрудняют следующие проблемы: стандарт CMM является собственностью SEI и не является общедоступным; оценка качества процессов организаций может проводиться только специалистами, прошедшими специальное обучение и аккредитованными SEI; стандарт ориентирован на применение в относительно крупных компаниях; стандарт преувеличивает роль формальных описаний процессов; в стандарте не рассмотрены вопросы об отборе, повышении квалификации и сохранении компетентных сотрудников.

Модель зрелости Capability Maturity Model Integration (CMMI). После появления CMM стали разрабатываться специализированные модели зрелости для создания информационных систем, для процесса выбора поставщиков и некоторые другие. На их основе была разработана интегрированная модель CMMI.

CMMI уточняет и совершенствует предшествовавшие модели CMM, а также учитывает основные требования существующих международных стандартов в области менеджмента программных средств. С 2006 года действует версия CMMI – 1.2, которая

является модернизацией версии СММІ – 1.1. Предполагается выпустить в 2010 году новую, существенно улучшенную модель СММІ – 1.3.

Значительное внимание в СММІ уделяется процессам разработки и учету итераций при изменении требований заказчиков, их прослеживанию к функциям, компонентам, тестам и документам проекта. Существует два варианта модели СММІ – 1.1 для обеспечения непрерывного оценивания комплекса процессов в определенной области создания программных средств или для поэтапного оценивания и совершенствования зрелости предприятия, а также для организации жизненного цикла (ЖЦ) комплексов программ в целом. Стандарт регламентирует построение всех моделей по единой схеме [4].

Недостаток заключается в следующем. Поскольку модель СММІ предполагает строгое и формальное планирование, отслеживание проекта и полное документирование (иными словами – большие накладные расходы), эти модели вряд ли применимы к разработкам небольших систем в рамках маленьких проектов с небольшим количеством сотрудников, а также проектов с быстро меняющимися требованиями или с короткими циклами разработки.

Модели СММ/СММІ подходят для длинных проектов (программ) и больших организаций, которые создают сложные многокомпонентные системы с поставками новых версий каждые 9–12 месяцев и относительно большим числом сотрудников.

Стандарт SPICE. В 1991 году Международная организация по стандартизации инициировала работу по созданию единого стандарта оценки программных процессов. Стандарт получил имя SPICE (сокращение от **S**oftware **P**rocess **I**mprovement and **C**apability **d**etermination – определение возможностей и улучшение процесса создания программного обеспечения). Официально стандарт называется «ISO/IEC 15504: Information Technology – Process Assessment» [5]. Международный стандарт версии ISO/IEC 15504 в настоящее время включает 6 частей, а 7-я часть в настоящее время находится в окончательной форме и начата работа по 8-й части.

ISO/IEC 15504 является эталонной моделью, которая определяет «Измерение процесса» и «Измерение зрелости (возможностей)» качества процесса.

«Измерение процесса» разделяет процессы на пять категорий: клиент – поставщик; инженерные; поддержки; управление; организация.

«Измерение зрелости (возможностей)» определяется с помощью девяти атрибутов процесса, сгруппированных на 5 уровнях зрелости:

1.1. Представление процесса (Process Performance);

2.1. Управление процессом (Performance Management);

2.2. Работа по управлению продуктами (Work Product Management);

3.1. Определение процесса (Process Definition);

3.2. Развертывание процесса (Process Deployment);

4.1. Измерение процесса (Process Measurement);

4.2. Контроль процесса (Process Control);

5.1. Инновация процесса (Process Innovation);

5.2. Оптимизация процесса (Process Optimization).

Каждый атрибут процесса оценивается по шкале: не достигнуто (0–15 %); частично достигнуто (15–50 %); в основном достигнуто (50–85 %); полностью достигнуто (85–100 %).

ISO/IEC 15504 может использоваться в двух контекстах: для совершенствования качества процессов разработки и для определения качества.

Достоинства ISO / IEC 15504:

– ISO/IEC 15504 является общедоступным национальным стандартом;

– данный стандарт пользуется поддержкой международного сообщества;

– SPICE возможно использовать и в небольших компаниях – об этом свидетельствуют результаты работы проекта SPIRE [6].

Недостатки SPICE [4]:

– ISO/IEC 15504 не является столь успешным, как СММ, т. к. СММ активнее спонсируется, был создан раньше и впоследствии был заменен СММІ, которая включает в себя множество идей ISO / IEC 15504, а также сохраняет преимущества СММ;

– по ISO/IEC 15504 не выдаются сертификаты соответствия. Он не является нормативным, хотя соответствует современным международным требованиям к качеству.

Система государственных стандартов. В настоящее время на территории России и других стран СНГ действуют старые ГОСТы 19-й и 34-й серий и более новый ГОСТ Р ИСО МЭК 122207 (в России).

ГОСТ 19 «Единая система программной документации» и ГОСТ 34 «Стандарты на разработку и сопровождение автоматизированных систем» ориентированы на последовательный подход к разработке ПО. Разработка в соответствии с этими стандартами проводится по этапам, каждый из которых предполагает выполнение строго определенных работ, и завершается выпуском достаточно большого числа весьма формализованных и обширных документов. Таким образом, строгое следование этим ГОСТам не только приводит к водопадному подходу, но и требует очень высокой степени формализованности разработки.

На основе этих стандартов разрабатываются программные системы по госзаказам в странах СНГ.

ГОСТ 12207 (перевод на русский язык ISO/IEC 12207), в отличие от стандартов 19-й и 34-й серий, описывает разработку ПО как набор основных и вспомогательных процессов, которые могут действовать от начала и до завершения проекта. Модель ЖЦ может выбираться исходя из особенностей проекта. Таким образом, этот ГОСТ явно не запрещает применение итеративного подхода, но и явно не рекомендует его использование. ГОСТ 12207 также более гибок в части требований к формальности процесса разработки. В нем содержатся только указания на необходимость документирования основных результатов процессов, но нет перечней требуемых документов и указаний относительно их содержания. Таким образом, ГОСТ 12207 допускает итеративную и менее формализованную разработку ПО.

ХАРАКТЕРИСТИКА МОДЕЛЕЙ С ИТЕРАТИВНЫМ ПРОЦЕССОМ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Под моделью ПО обычно понимается структура, определяющая последовательность выполнения и взаимосвязи процессов, действий и задач на протяжении ЖЦ. Далее рассмотрены наиболее распространенные модели, ориентированные на итеративный процесс разработки: Rational Unified Process, Microsoft Solutions Framework, Personal Software Process / Team Software Process и Agile(eXtreme Programming, Crystal, Feature Driven Development).

Методология Rational Unified Process (RUP). Унифицированный процесс RUP [7] был разработан Филиппом Крачтенем (Philippe Kruchten), Иваром Якобсоном (Ivar Jacobson) и другими сотрудниками компании «Rational Software» в качестве дополнения к языку моделирования Unified Modeling Language (UML). Особенностью RUP является то, что в результате работы над проектом создаются и совершенствуются модели ПО. Вместо создания громадного количества бумажных документов RUP опирается на разработку и развитие семантически обогащенных моделей, всесторонне представляющих разрабатываемую систему.

RUP, в отличие от большинства других методологий, позволяет в широком диапазоне выбирать степень формализации и итеративности процесса разработки в зависимости от особенностей проектов и разрабатывающей организации. RUP наиболее часто используют компании с большим количеством разработчиков (свыше 40–50 человек).

ЖЦ проекта RUP состоит из фаз и дисциплин. Фазы RUP: начало (Inception), проектирование (Elaboration), разработка (Construction) и внедрение (Transition). Последовательность этих фаз фиксирована, но число итераций, необходимых для завершения каждой фазы, определяется индивидуально для каждого конкретного проекта [8]. Каждая фаза состоит из дисциплин: «Бизнес-моделирование» (Business modeling); «Управление требованиями» (Requirements); «Анализ и Проектирование» (Analysis and Design); «Реализация» (Implementation); «Тестирование» (Test); «Развертывание» (Deployment), «Управление проектом» (Project management); «Управление изменениями» (Change management); «Среда» (Environment).

В RUP все дисциплины выполняются практически во всех фазах ЖЦ программных средств (ПС). Однако, в зависимости от фазы, меняются текущие цели проекта и соотношение между объемами работ, соответствующих различным дисциплинам.

Методология RUP использует программные решения Rational Application Developer, Rational Rose, Rational Asset Manager, Rational Software Architect, Rational Software Modeler, Rational Suite, Rational Requisitepro, Rational Clearquest, Rational Clearcase и другие [9].

Достоинства RUP:

- очень широкий диапазон решений в части формализации процесса разработки;
- можно использовать как основу для водопадного стиля разработки, так и в качестве гибкого процесса;
- снижение основных рисков заказчика и разработчика;
- экономия ресурсов за счет автоматизации регрессионного тестирования;
- улучшение качества ПО за счет многократных проверок изменений;
- улучшение качества тестирования за счет использования современных технологий.

Использование RUP гарантирует получение 3-го уровня CMM.

Недостатком является недостаточный уровень формализма, который может приводить к несогласованности решений, принимаемых участниками проекта, к непродуктивным затратам ресурсов на переработку кода и на повторное решение типовых проблем.

Модель Microsoft Solutions Framework (MSF). В 1994 году Microsoft выпустила MSF – набор концепций и рекомендуемых моделей, которые позволяют разрабатывать и внедрять распределенные информационные системы масштаба предприятия на основе технологий и инструментальных средств фирмы Microsoft. MSF базируется на практических результатах организации распределенных вычислений и применения

клиент-серверных технологий, полученных как фирмой Microsoft, так и ее партнерами и заказчиками [10].

Процесс разработки в соответствии с MSF является итеративным и включает в себя следующие основные фазы: разработка концепции (единого видения), планирование, разработка, стабилизация (тестирование), внедрение. Каждая фаза цикла заканчивается главной вехой (контрольной точкой). Соответственно главные вехи будут иметь названия: концепция продукта утверждена, планы продукта утверждены, разработка завершена, готовность решения утверждена, внедрение завершено. Веха является точкой синхронизации достигнутых результатов и ожиданий заказчика, а также анализа проектной среды. В решении о закрытии очередной фазы должны принимать участие ответственные представители всех ролевых кластеров (разработка, тестирование, внедрение, управление проектом и пр.) [10].

MSF предлагает достаточно нестандартные подходы к организационной структуре, распределению ответственности и принципам взаимодействия внутри команды. MSF предлагает разбиение больших команд (более 10 человек) на малые многопрофильные группы направлений (feature teams). Эти малые коллективы работают параллельно, регулярно синхронизируя свои усилия [11].

Microsoft предлагает инструменты для реализации MSF: MSF for Agile Software Development, MSF for CMMI Process Improvement, Visual Studio Team System (Team Suite; Team Edition for Database Professionals; Team Suite for Software Architects; Team Suite for Software Developers; Team Suite for Software Testers; Team Foundation Server; Team Test Load Agent).

Применение MSF позволит сертифицировать организацию на 2–3-й уровень CMM.

Достоинствами MSF являются:

- систематизация и структуризация информации в форме базы знаний;
- нестандартные подходы к организационной структуре, распределению ответственности и принципам взаимодействия внутри команды;
- MSF легко масштабируется – минимальный размер проектной группы в MSF-проекте – 3 человека, но применять данную методологию можно для коллективов и в десятки, сотни, тысячи человек;
- MSF не навязывает использование каких-либо конкретных инструментов и программных средств.

Недостатки MSF: не описывает детально важнейшие роли заказчика и пользователя; не рассматриваются методы управления группой проектов.

Personal Software Process / Team Software Process (PSP/TSP). В конце 90-х годов SEI разработал модель PSP/TSP, основанную на модели зрелости

CMM [12]. PSP определяет требования к компетенциям разработчика. Согласно этой модели каждый программист должен уметь: учитывать время, затраченное на работу над проектом; учитывать найденные дефекты; классифицировать типы дефектов; оценивать размер задачи; осуществлять систематический подход к описанию результатов тестирования; планировать программные задачи; распределять их по времени и составлять график работы; выполнять индивидуальную проверку проекта и архитектуры; осуществлять индивидуальную проверку кода; выполнять регрессионное тестирование.

PSP включает 4 процесса, которые последовательно добавляются разработчиком: PSP0 – базовый собственный процесс (The Baseline Personal Process); PSP1 – планирование (Personal Planning Process); PSP2 – управление качеством (Personal Quality Management); PSP3 – циклический процесс (Cyclic Personal Process).

TSP делает ставку на самоуправляемые команды численностью 3–20 разработчиков. Команды должны установить собственные цели, составить свой процесс и планы, отслеживать работу, поддерживать мотивацию и максимальную производительность.

Методология является эффективной, может обеспечить существенные преимущества и стать моделью для любой организации, заинтересованной в совершенствовании своих процессов разработки. Совместное использование организациями технологий PSP и TSP способно сформировать процесс совершенствования в рамках модели зрелости процессов разработки ПО и сделать нормой в организации пятый уровень CMM.

PSP/TSP – это чистая методология по обеспечению качества проектов, рассчитанная на использование UML при проектировании, и любые CASE-средства, в частности они ориентированы на использование продуктов компании Microsoft.

Достоинства: лучшее планирование времени и бюджета; управление качеством продукта; уменьшение времени разработки; аккуратный сбор метрик; оценка времени работы через оценку объема артефактов; использование исторических данных в планировании; раннее обнаружение дефектов.

Недостатки: сложность в учете рисков, учете недооценки сложности.

Гибкая методология разработки ПС (Agile). В феврале 2001 года 17 разработчиков различных методологий собрались для того, чтобы попытаться обнаружить что-нибудь общее в своих подходах. Результатом стал Манифест гибкой разработки [13]. Тогда же появился термин Agile («гибкий, шустрый»), объединяющий все методологии.

Agile – гибкая методология разработки, это концептуальный каркас, в рамках которого выполняется разработка программного обеспечения. Основная идея всех гибких моделей заключается в том, что применяемый в разработке ПО процесс должен быть адаптивным. Они декларируют своей высшей ценностью ориентированность на людей и их взаимодействие, а не на процессы и средства [7].

CASE-средства, которые могут помочь внедрить и улучшить процесс создания ПО на базе Agile: XPArchitect; Borland ALM(в частности Silk).

Рассмотрим более детально несколько гибких методологий: eXtreme Programming, Crystal и Feature Driven Development.

1. Экстремальное программирование (eXtreme Programming, или XP) – методология, разработанная Кентом Бекем (Kent Beck) [12], Уордом Каннингемом (Ward Cunningham) и Роном Джеффрисом (Ron Jeffries), является сегодня наиболее известной из гибких методологий. XP проповедует коммуникабельность, простоту, обратную связь. Она описывается как набор практик: игра в планирование, короткие релизы, метафоры, простой дизайн, переработки кода (refactoring), разработка «тестами вперед», парное программирование, коллективное владение кодом, 40-часовая рабочая неделя, постоянное присутствие заказчика и стандарты кода. При использовании XP тщательное предварительное проектирование ПО заменяется, с одной стороны, постоянным присутствием в команде заказчика, а с другой – регулярными переработками кода (так называемый рефакторинг). Основой проектной документации считается тщательно прокомментированный код. Очень большое внимание в методологии уделяется тестированию.

ЖЦ проекта в XP [8] состоит из последовательности релизов. Каждый релиз – это полноценная версия продукта, которую может использовать заказчик, и содержащая дополнительную функциональность по сравнению с предыдущим релизом. Релиз появляется в результате одной или нескольких итераций, длящихся от одной до четырех недель.

В XP не рекомендуется тратить много времени на планирование; сам процесс планирования называется игрой (planning game). Подробный план составляется только на очередную итерацию и ближайшие один-два релиза.

Недостатки:

- XP могут эффективно использоваться только в команде опытных разработчиков;

- команду нельзя разбивать на несколько частей, возможный размер команды ограничен числом в 10–15 человек, т. к. большую роль в XP играет прямое общение;

- не всегда можно обеспечить постоянное присутствие представителя заказчика в проектной команде;
- нежелательно использовать XP в проекте с фиксированной ценой.

Несмотря на все перечисленные ограничения, XP может показать исключительную эффективность благодаря крайне низким накладным расходам.

2. Crystal – семейство методологий, определяющих необходимую степень формализации процесса разработки в зависимости от количества участников и критичности задач. Алистэр Коуберн (Alistair Cockburn) в начале 90-х по заказу от компании IBM написал работу на тему методологии разработки ПО [14]. При этом его подход существенно отличался от подхода большинства других методологов. Его теории основаны не только на личном опыте, но и на постоянных исследованиях других проектов и процессов.

Степень важности нарастает по вертикальной оси. Величина команды нарастает по горизонтальной оси. В результате получается семейство методологий. Чем ниже критичность и чем меньше команда, тем более «легкую» методологию нужно использовать. Самой легкой из всего семейства является методология Crystal Clear.

Главные принципы данной методологии:

- вся команда разработчиков (до 6 человек) находится в одном помещении;

- частые поставки продукта позволяют выработать «ритм» проекта;

- обмен информацией с реальными пользователями позволяет быстро получать обратную связь и ликвидировать недостатки на ранних стадиях;

- средства контроля версий кода обеспечивают коллективное владение кодом.

Семейство методологий Crystal построено на итеративной разработке. Продолжительность итерации может изменяться в пределах от 1 до 4 месяцев. Чем меньше итерация, тем лучше.

Важной особенностью Crystal является непрерывная настройка методологии. Это позволяет постепенно адаптировать ее для конкретной команды и конкретного проекта. Ни в одной другой методологии настройке не уделяется такого внимания.

Достоинством Crystal Clear является то, что она максимально проста в использовании, требует минимальных усилий для внедрения, ориентирована на человеческие привычки, описывает естественный порядок разработки ПО.

Недостатки:

- уступает XP по производительности;

- очень сложно заранее предсказать, какие промежуточные продукты необходимы.

3. Feature Driven Development(FDD). Эта методология была разработана Джеффом Де Люка (Jeff De Luca) и Питером Коадом (Peter Coad) в 1997 году [15]. Это функционально-ориентированная разработка, которая оперирует понятием функции или свойства (feature) системы, достаточно близким к понятию сценария использования, применяемому в RUP. Отличие – это дополнительное ограничение: «каждая функция должна допускать реализацию не более чем за две недели».

FDD включает пять процессов, причем последние два повторяются для каждой функции [15]: разработка общей модели; составление списка необходимых функций системы; планирование работы над каждой функцией; проектирование функции; конструирование функции. Работа над проектом предполагает частые сборки и делится на итерации, каждая из которых реализуется с помощью определенного набора функций.

Достоинства FDD: делает процесс легче; дает возможность планирования и предварительного проектирования, создания детального дизайна, есть управление приоритетами и возможна трассировка требований; рассчитан на работу в больших командах.

Но в тоже время, недостаток документации может значительно увеличивать стоимость последующего сопровождения продукта, поскольку внесение каких-либо изменений в него потребует очень больших усилий.

СРАВНЕНИЕ МОДЕЛЕЙ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Поскольку каждая из рассмотренных моделей обладает своими достоинствами и недостатками, не существует единой оптимальной модели ПО. Она должна определяться для каждого конкретного проекта и может меняться в очень широком диапазоне. Основными факторами, влияющими на выбор модели и степень формализации процесса, являются [16]:

– масштаб проекта – чем больше людей участвует в проекте, тем более формально он должен вестись;

– распределение участников – чем компактнее расположены участники, чем легче им общаться между собой, тем менее формализованным может быть проект;

– критичность проекта – проекты, в которых ошибки могут привести к тяжелым последствиям, вплоть до гибели людей, должны проводиться гораздо более формально, чем те проекты, в которых ошибки приводят только к временным неудобствам;

– новизна проекта – если разработчики используют новые технологии и мало знакомы с предметной областью, то следует более тщательно, более формально прорабатывать проектные решения. Однако если над проектом работает небольшая группа высококвалифицированных программистов, более предпочтительным будет тщательное оформление уже отработанных решений перед их тиражированием;

– требования заказчика – существенно различаются в зависимости от отрасли и статуса организации-заказчика;

– ожидаемая долговечность проекта – если разрабатываемое для конкретного заказчика ПО планируется через пару лет заменить новым, то вряд ли имеет смысл тратить много сил на документацию. Если же срок жизни проекта предполагается достаточно длинным, то без хорошей документации не обойтись.

Опираясь на данные факторы, будем классифицировать модели итеративной разработки ПО по «весу» – количеству формализованных процессов (большинство процессов или только основные) и детальности их регламентации. Чем больше процессов документировано, чем более детально они описаны, тем больше «вес» модели. Достоинства и недостатки модели в зависимости от ее веса рассмотрены в табл. 1.

Таблица 1. Характеристика моделей процессов разработки ПО

Вес	Модели	Достоинства	Недостатки
Тяжелый	RUP, MSF (для CMMI)	Процессы рассчитаны на среднюю квалификацию исполнителей. Невысокие требования к стабильности команды. Нет ограничений на распределение участников. Отсутствуют ограничения по объему и сложности выполняемых проектов. Используются для проектов с высокими рисками. Легче получить высокий уровень сертификации стандартов качества	Требуют существенной управленческой надстройки. Более длительные стадии анализа и проектирования. Более формализованные коммуникации. Большие затраты на документацию длительного процесса сопровождения ПО
Легкий	MSF (для Agile), Agile, PSP/TSP	Меньше непроизводительных расходов, связанных с управлением проектом, рисками, изменениями, конфигурациями, сопровождением ПО. Упрощенные стадии анализа и проектирования, основной упор на разработку функциональности, совмещение ролей. Неформальные коммуникации	Используется для проектов с невысокими рисками и небольшого масштаба и сложности. Участники проекта должны располагаться компактно. Плохо подходят для новых проектов. Требуют более квалифицированной и стабильной команды. Сложнее получить высокий уровень сертификации стандартов качества

Таблиця 2. Сравнение моделей разработки программных проектов

Параметры Содержание	Модели		PSP/TSP	MSF	RUP	Agile	FDD
	RUP	XP					
Команда	Абстрактный общий процесс, на основе которого организация должна создать конкретный специализированный процесс, ориентированный на ее потребности	Согласованный набор концепций, моделей и правил, описывающих управление людьми и процессами в процессе разработки решения; опирается на практический опыт Microsoft	Набор требований к компетенциям разработчика, делает ставку на самоуправляемые команды	Группы по 3–10 человек Может быть и тяжелым, и легким	Итеративная инкрементная разработка; автоматическое регрессионное тестирование; состав документов заранее определяется участниками проекта	Функционально-ориентированная разработка, которая оперирует понятием функции или свойства системы	20–30 человек
Вес	Свыше 40–50 человек	Тяжелый	3–20 разработчиков	Легкий	Легкий	До 6 человек	
Используемые стандарты	CMML, ISO, SPICE	CMML, ISO, SPICE	CMML, SPICE	CMML, ISO, SPICE	CMML, SPICE	CMML, SPICE	CMML, SPICE
Достоинства	<ol style="list-style-type: none"> Снижение основных рисков заказчика и разработчика. Экономия ресурсов за счет автоматизации регрессионного тестирования. Улучшение качества ПО за счет многократных проверок. Улучшение качества тестирования за счет использования современных технологий. 	<ol style="list-style-type: none"> Систематизация и структуризация информации в форме базы знаний. Нестандартные подходы к организации структуры, распределению ответственности и принципам взаимодействия внутри команды 	<ol style="list-style-type: none"> Лучшее планирование времени и бюджета. Управление качеством продукта. Уменьшение времени разработки. Аккуратный сбор метрик. Оценка времени работы через оценку объема артефактов. Раннее обнаружение дефектов 	<ol style="list-style-type: none"> Максимально проста в использовании. Требует минимальных усилий для внедрения; Ориентирована на человеческие прививычки; Описывает естественный порядок разработки ПО 	<ol style="list-style-type: none"> Дает возможность планирования и предварительного проектирования, создания детального дизайна. Есть упорядоченные приоритетные и возможные трассировка требований 	<ol style="list-style-type: none"> Дает возможность планирования и предварительного проектирования, создания детального дизайна. Есть упорядоченные приоритетные и возможные трассировка требований 	
Недостатки	Несогласованность решений, непродуктивные затраты ресурсов на переработку кода и на повторное решение типовых проблем в случае недостаточного уровня формализма	Не описывает детально важнейшие роли заказчика и пользователя; не рассматриваются методы управления группой проектов	Проблемы при учете рисков, сложности	Не описывает детально важнейшие роли заказчика и пользователя; не рассматриваются методы управления группой проектов	Рассчитан только на команду опытных разработчиков, не разбитую на несколько частей; проблемы разрешения нетехнологических рисков	Уступает XP по производительности, очень сложно за ранее предсказать, какие промежуточные продукты необходимы	Недостаток документации; значительно увеличивается стоимость последующего сопровождения продукта
CASE-средства	IBM Rational	Visual Studio Team System	Любые CASE-средства, в частности продукты компании Microsoft	Visual Studio Team System	Любые CASE-средства, в частности продукты компании Microsoft	Уступает XP по производительности, очень сложно за ранее предсказать, какие промежуточные продукты необходимы	Уступает XP по производительности, очень сложно за ранее предсказать, какие промежуточные продукты необходимы

В табл. 2 приводится сравнение моделей, поддерживающих итеративную разработку, по содержанию, весу, ориентации на количество разработчиков, используемым стандартам для обеспечения качества, достоинствам и недостаткам.

Если организация достаточно велика (больше десятка разработчиков) и выполняет проекты для различных отраслей или для систем разного уровня критичности – желательно выбирать для каждого проекта свой оптимальный уровень формализации и модель ПО.

Организациям с большим количеством людей или организациям-компаньонам, проводящим разработку совместно, но разделенным географически, следует выбирать RUP. В случае управления проектом с большим количеством данных целесообразно использовать MSF, т.к. систематизация и структуризация информации в форме базы данных позволит упростить работу с ними.

Если команда разработчиков состоит из 1–20 человек, которые расположены компактно, следует выбрать любую из гибких методологий или PSP/TSP. Последнюю желательно использовать, если ошибки в проекте не являются критичными, а команда организации состоит из небольшого количества разработчиков и является самоуправляемой.

Для критических и малознакомых проектов необходимо выбирать MSF либо RUP, для некритических подойдут PSP/TSP и Agile. Гибкие методологии также можно успешно применить при разработке новых технологий.

Если заказчик настаивает на личном участии в процессе разработки, следует выбирать гибкие методологии, в частности XP. Если же для контроля используется строгая отчетность, выбор следует остановить на RUP, MSF.

В случае повторного использования ПО целесообразно использовать RUP.

ВЫВОДЫ

Учитывая вышеизложенное, можно сделать следующие выводы:

1. Каждая из рассмотренных моделей предназначена для разработки определенного класса ПО с разными требованиями к командам разработчиков. Модели ориентированы на соответствующее ПО, однако эффективны и при использовании в качестве независимых методологий.

2. Для обеспечения качества ПО широко используются стандарты ISO9001, CMM, ISO/IEC 15504. На использование стандартов CMM/CMMI и ISO/IEC не

влияет используемая модель разработки ПО. Однако стандартизация процессов, ориентированных на Agile методологию, с использованием стандарта ISO9001 практически невозможна.

3. Для выбора модели итеративного процесса разработки ПО одним из наиболее эффективных является показатель «вес модели», который учитывает масштаб, новизну и критичность проекта, распределение участников и требования заказчика.

СПИСОК ЛИТЕРАТУРЫ

1. *Rubinstein D.* Standish Group Report: There's Less Development Chaos Today [Электрон. ресурс]. – Режим доступа: <http://www.sdtimes.com/content/article.aspx?ArticleID=30247>.
2. ISO 9001:2008 [Электрон. ресурс]. – Режим доступа: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=46486.
3. Capability Maturity Model for Software, Version 1.1 (CMU/SEI-93-TR-24) [Текст] / Paulk, Mark C., and others, Software Engineering Institute, Carnegie Mellon University. – Pittsburgh, Pa. : SEI, CMU, 1993. – 88 с.
4. *Лунаев В. В.* Модели зрелости программной инженерии – CMMI. Содержание и применение [Текст] / В. В. Лунаев // Информационный бюллетень. – 2006. – № 6(157). – С. 1–15.
5. ISO/IEC 15504 [Электрон. ресурс]. – Режим доступа: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=50519.
6. Software Process Improvement in Regions of Europe [Электрон. ресурс]. – Режим доступа: <http://www.cse.dcu.ie/spire>.
7. *Bragina T., Tabunshchik G.* Comparative Analysis of Software Development Models for Electro-technical Systems // Proc. of Int. Conf. on Modern Problem of Radio Engineering, Telecommunications and Computer Science TC-SET'2010, February 19–23, 2010, Lviv–Slavsko, Ukraine. – P. 347.
8. *Рахимбердиев А.* Современные процессы разработки программного обеспечения [Текст] / А. Рахимбердиев // RSDN Magazine. – 2006 – № 4. – С. 3–10.
9. IBM [Электрон. ресурс]. – Режим доступа: <http://www-142.ibm.com/software/products/ru/ru/category/rational/SW700>.
10. *Keeton M.* Microsoft Solutions Framework (MSF) : A Pocket Guide (Paperback) [Текст] / Keeton M. – Wan Haren Publishing, 2005. – 112 p.
11. *Turner M. S. V.* Microsoft Solutions Framework Essentials: Building Successful Technology Solutions [Текст] / Turner M. S. V. – Microsoft Press 1 ed., – 2006. – 336 p.
12. *Брагина Т. И., Табуницкий Г. В.* Классификация моделей итеративной разработки программного обеспечения // Материалы всеукраинской научно-практической конференции «Системный анализ. Информатика. Управление», САИУ-2010, Март 04–05, 2010, Запорожье, Украина. – С. 23–25.
13. Agile Manifesto [Электрон. ресурс]. – Режим доступа: <http://agilemanifesto.org>.
14. *Cockburn A.* Surviving Object-Oriented Projects [Текст] / A. Cockburn. – Addison Wesley, 1997. – 272 p.
15. *Палмер С. Р.* Практическое руководство по функционально-ориентированной разработке ПО [Текст] / Стивен Р. Палмер, Джон М. Фелсинг. – Вильямс, 2002. – 304 с.
16. *Архипенков С.* Руководство командой разработчиков программного обеспечения. Прикладные мысли [Текст] / С. Архипенков. – М. : Москва, 2008. – 80 с.

Надійшла 24.02.2010
Після доробки 22.03.2010

Брагіна Т. І., Табунщик Г. В.

ПОРІВНЯЛЬНИЙ АНАЛІЗ ІТЕРАТИВНИХ МОДЕЛЕЙ РОЗРОБКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

У статті виконано огляд моделей розробки програмних проєктів, їх життєвого циклу, переваг і недоліків. Проведено аналіз найбільш поширених моделей: Rational Unified Process, Microsoft Solutions Framework, Personal Software Process / Team Software Process і Agile (eXtreme Programming, Crystal, Feature Driven Development). Зроблено висновки щодо факторів, що впливають на вибір моделі, та запропоновано рекомендації з вибору відповідної моделі програмного забезпечення в залежності від характеристик проєкту.

Ключові слова: програмний проєкт, стандарти якості, фактори вибору моделі.

Bragina T. I., Tabunshchik G. V.

COMPARATIVE ANALYSIS OF ITERATIVE SOFTWARE DEVELOPMENT MODEL

The authors have made a review of software projects development models, their life cycle, their advantages and disadvantages. The most common models were analyzed, such as Rational Unified Process, Microsoft Solutions Framework, Personal Software Process / Team Software Process and Agile (eXtreme Programming, Crystal, Feature Driven Development). Conclusions were drawn on the factors which influence the model choice, and recommendations were proposed for appropriate software model selection depending on project characteristics.

Ключові слова: software project, quality standards, factors for model selection.

УДК 621.391

Высочина О. С.¹, Шматков С. И.², Салман Амер Мухсин³

¹Аспирант Харьковского национального университета им. В.Н. Каразина

²Канд. техн. наук, заведующий кафедрой Харьковского национального университета им. В.Н. Каразина

³Аспирант Харьковского национального университета радиозлектроники

АНАЛИЗ СИСТЕМ МОНИТОРИНГА ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЕЙ

Рассматриваются наиболее известные системы мониторинга телекоммуникационной сети, представлен их сравнительный анализ, сформулированы общие требования и синтезирована общая архитектура подобных систем. Предлагаемые решения не способны прогнозировать состояние телекоммуникационной сети. Поэтому в такие системы необходимо включать дополнительные модули обработки статистической информации.

Ключевые слова: диагностика состояния телекоммуникационных сетей, нейронные сети, мониторинг состояния, адаптивное управление.

1. ПОСТАНОВКА ЗАДАЧИ

Телекоммуникационная отрасль переживает сегодня значительные преобразования, переходу от традиционных сетей с коммутацией каналов к пакетной передаче данных сопутствует лавинообразный рост предоставляемых абонентам услуг. При этом в условиях постоянного повышения сложности информационных и телекоммуникационных систем надежность телекоммуникационной сети и качество предоставляемых сервисов приобретают особую важность. Современная телекоммуникационная инфраструктура представляет собой сложную гетерогенную сеть, включающую телекоммуникационное, серверное и программное обеспечение различных производителей, работающее в различных стандартах и под управлением различного программного обеспечения. Сложность и масштабность сетевой инфраструктуры предопределяют высокий уровень автоматизиро-

ванных средств мониторинга и управления, которые должны использоваться для обеспечения надежной работы сети.

Целью создания системы мониторинга телекоммуникационной инфраструктуры является [1–2]:

- обеспечение высокой скорости обработки запросов пользователей на предоставление требуемых информационных ресурсов и сервисов;
- предоставление программно-аппаратных средств по управлению информационными и телекоммуникационными ресурсами;
- создание эффективной службы диагностики и своевременного оповещения для предупреждения аварийных ситуаций и повышения отказоустойчивости телекоммуникационных систем;
- выполнение сбора, обработки, хранения и отображения полной информации о состоянии всех компонентов телекоммуникационной и информационной

инфраструктуры сети в реальном времени независимо от архитектуры сети, типа коммутатора и поставщика.

При этом должны обеспечиваться:

– невмешательство в работу сетевого оборудования благодаря оверлейной архитектуре сбора данных;

– повышенная безопасность, поскольку зонды не зависят от коммутаторов и потери информации в результате отказов или перегрузок в сети отсутствуют;

– постоянный сбор статистической информации, который позволяет создавать крупномасштабные базы данных, необходимые для проведения псевдооперативного и статистического анализа сети, трафика и технических параметров обслуживания;

– управление сетью и обеспечение обслуживания в режиме реального времени – управление параметрами качества обслуживания;

– интеграция с усовершенствованными инструментальными средствами анализа данных;

– создание единого информационного центра обработки данных о состоянии систем и сети.

На практике к платформам системы мониторинга телекоммуникационной инфраструктуры предъявляются следующие основные требования [3–4]:

– масштабируемость;

– поддержка распределенной архитектуры клиент – сервер;

– открытость, позволяющая управлять оборудованием различных производителей;

– разграничение функции персонала поддержки.

Целью настоящей работы является анализ систем мониторинга телекоммуникационной сети, формирование общих требований и обобщение архитектуры подобных систем.

2. СРАВНИТЕЛЬНЫЙ АНАЛИЗ СИСТЕМ МОНИТОРИНГА

Системы мониторинга существуют на рынке телекоммуникаций много лет и стремительно развиваются с развитием отрасли в целом. Предлагаемые на мировом рынке системы мониторинга схожи по выполняемым функциям, все они предоставляют почти одинаковый минимальный набор возможностей. Наиболее интересными являются следующие системы мониторинга, сравнение которых представлено в табл. 1:

- Argus;
- Intellipool Network Monitor;
- AdRem NetCrunch;
- IPHost Network Monitor;
- NetMRI;
- NetQoS Performance Center;
- OPNET ACE Live;
- Opsview;

– Performance Co-Pilot;

– Scrutinizer;

– Orion;

– Zenoss.

Таблица 1. Сравнительный анализ систем мониторинга

Система мониторинга	Параметры								
	1	2	3	4	5	6	7	8	9
Argus	+	-	-	-	+	+	+	+	+
Intellipool Network Monitor	+	-	-	+	-	+	+	-	+
AdRem NetCrunch	-	+	-	+	-	+	+	+	-
IPHost Network Monitor	+	+	-	+	-	+	-	+	-
NetMRI	-	+	-	+	-	+	+	+	+
NetQoS Performance Center	+	+	+	+	-	+	+	+	+
OPNET ACE Live	+	+	+	+	-	+	-	+	+
Opsview	+	+	-	+	+	+	+	+	+
Performance Co-Pilot	-	+	-	-	+	-	-	+	+
Scrutinizer	+	+	-	-	-	+	+	+	+
Orion	+	+	+	+	+	+	+	+	+
Zenoss	+	+	+	+	-	+	+	+	+

Сравнение систем мониторинга проводилось по следующим параметрам:

1. Формирование отчетов SLA (Service Level Agreement) согласно требованиям.

Контроль гарантированных параметров качества обслуживания SLA, определяющих межоператорские взаимоотношения.

2. Формирование трендов.

Выявление основных тенденций динамики показателей качества работы телекоммуникационной сети.

3. Прогнозирование трендов.

Прогнозирование изменения динамики показателей качества работы телекоммуникационной сети.

4. Анализ топологии сети.

Сбор информации об элементах сети.

5. Использование агентной модели мониторинга.

Наличие устройств, осуществляющих сбор и передачу информации о работе сети.

6. Поддержка SNMP.

Использование протокола SMNP для обмена информацией о состояниях объектов наблюдения в режиме реального времени.

7. Протоколирование событий.

Формирование подробных записей о состоянии элементов сети.

8. Датчики внештатных ситуаций.

Наличие устройств для оповещения о возникновении критических ситуаций, негативной тенденции изменения показателей качества работы телекоммуникационной сети.

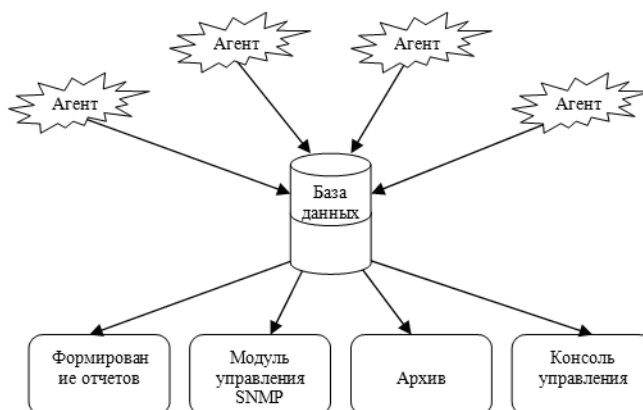


Рис. 1. Общая архитектура системы мониторинга

9. Распределенный мониторинг.

Мониторинг сигнального обмена на предмет соответствия работы оборудования определенным спецификациям протоколов.

Проведенный анализ показал, что предлагаемые на мировом рынке системы мониторинга схожи по выполняемым функциям. Все они предоставляют почти одинаковый минимальный набор возможностей, однако каждая из них обладает определенными недостатками: в большинстве систем вообще не реализованы возможности прогнозирования трендов, а в системах, где они реализованы, построение происходит на основе устаревшей статистической информации. Подобное прогнозирование не учитывает фрактальность трафика, нелинейность характеристик и нестационарность процессов.

Обобщив предлагаемые решения, можно синтезировать общую архитектуру системы мониторинга и управления. Все рассмотренные системы мониторинга основаны на использовании агентного подхода. Агенты собирают статистическую информацию о работе элементов сети и передают ее в центральную базу данных, после чего она обрабатывается управляющими модулями. В состав системы мониторинга должны входить следующие компоненты: формирование отчетов, модуль управления SNMP, архив и консоль управления. Модуль формирования отчетов позволяет формировать из имеющихся данных информацию для принятия управленческих решений. Модуль управления SNMP отвечает за сбор информации с агентов мониторинга и взаимодействие с системами управления. Архив позволяет упорядочить хранение статистической информации и организовать последующую работу с ней. Консоль управления реализует функции конфигурирования и управления системой.

В общем виде система мониторинга, удовлетворяющая современным требованиям, представлена на рис. 1.

Внедрение систем мониторинга позволяет решить множество задач, в числе которых:

- сокращение сроков и затрат на выполнение текущих задач, включая активацию услуг;
- повышение отдачи от существующих ресурсов сети и улучшение качества планирования их будущего развития;
- снижение потребности в персонале и, как следствие, сокращение текущих расходов;
- более полная реализация потенциала современного сетевого оборудования за счет разработки и реализации новых услуг;
- сведение к минимуму рисков потерь доходов;
- сокращение сроков реагирования на происходящие в сети события;
- привлечение высокодоходных клиентов за счет предоставления дополнительных услуг на основе гарантированного качества;
- сокращение сроков ввода в эксплуатацию новых услуг;
- повышение качества и оперативности обслуживания пользователей сети за счет четкой координации и информационной поддержки работ;
- обеспечение координации взаимодействия многочисленного персонала удаленных подразделений в режиме реального времени.

ВЫВОДЫ

Анализ задач, стоящих перед телекоммуникационными сетями в настоящее время, позволил сформировать набор требований, предъявляемых к современным системам мониторинга. Рассмотрены наиболее популярные системы мониторинга, представленные на мировом рынке. Проведенный анализ

показал, что все рассмотренные системы мониторинга или не способны прогнозировать состояние сети, или выполняют прогнозирование достаточно приближенно. Для более эффективного управления качеством обслуживания предоставляемых услуг (сервисов) в телекоммуникационной сети подобные системы должны поддерживать функции не только формирования трендов на основе полученной статистики, но и выполнять прогнозирование состояния сети. Под прогнозированием в данном случае понимается предсказание характера изменений основных показателей качества сети (среднепутевая задержка, джиттер, количество потерянных пакетов, среднее время простоя в очереди и т. д.). На основании такой информации алгоритмы динамической маршрутизации смогут более эффективно перераспределять информационные потоки. В основе модулей прогнозирования могут лежать различные методы классификации, но проведенный анализ показал, что в данном случае наиболее эффективным средством прогнозирования являются нейронные сети. Анализ функциональности систем мониторинга и предъявляемых к ним требований позволил синтезировать общую архитектуру подобных систем.

СПИСОК ЛИТЕРАТУРЫ

1. *Сторожук Д. О.* Методы и алгоритмы для систем мониторинга локальных сетей / Д. О. Сторожук. – М., 2008. – 121 с.

УДК 004.67:004.9

Дубровин В. И.¹, Афонин Ю. С.², Манило Т. В.³, Харитонов В. Н.⁴

¹Канд. техн. наук, профессор Запорожского национального технического университета

²Ассистент Запорожского национального технического университета

³Аспирант Запорожского национального технического университета

⁴Инженер-конструктор ГП «Ивченко-Прогресс» (г. Запорожье)

ВЕЙВЛЕТ-АНАЛИЗ КРАТКОВРЕМЕННЫХ ИМПУЛЬСОВ В СИГНАЛАХ С ДАТЧИКОВ ЧАСТОТЫ ВРАЩЕНИЯ РОТОРА ГТД

В статье рассмотрена проблема обнаружения кратковременных импульсов в сигналах с датчиков частоты вращения ротора ГТД, которые возникают в момент аварийного рассоединения трансмиссии. Приведена классификация данных импульсов. Анализ каждого вида импульса выполнен с использованием вейвлет-преобразования, как наиболее подходящего для выявления локальных особенностей характеристик сигнала. Рассмотрены перспективные задачи в данном направлении.

Ключевые слова: сигнал ДЧВ, вейвлет-анализ, кратковременный импульс, коэффициенты вейвлет-разложения, вейвлеты Добеши.

ВВЕДЕНИЕ

Сигналы, получаемые с датчиков частоты вращения (ДЧВ) ротора ГТД, являются нестационарными,

2. *Новиков Ю. В.* Локальные сети. Архитектура. Алгоритмы. Проектирование. Мониторинг и анализ сетей / Новиков Ю. В., Кондратенко С. В., Уилсон Э. – М. : ЭКОМ, 2000. – 308 с.
3. *Уилсон Э.* Мониторинг и анализ сетей / Уилсон Э. – М. : ЛОРИ, 2002. – 350 с.
4. *Бараш Л.* Мониторинг трафика в сетях с коммутацией пакетов / Бараш Л. // Компьютерное обозрение. – 2009. – № 37 (654). – С. 20–25.

Надійшла 17.03.2010

АНАЛІЗ СИСТЕМ МОНІТОРИНГУ ТЕЛЕКОМУНІКАЦІЙНИХ МЕРЕЖ

Розглядаються найбільш відомі системи моніторингу телекомунікаційної мережі, представлено їх порівняльний аналіз, сформульовано загальні вимоги та синтезовано загальну архітектуру подібних систем. Запропоновані рішення не здатні прогнозувати стан телекомунікаційної мережі. Тому в такі системи необхідно включати додаткові модулі обробки статистичної інформації.

Ключові слова: діагностика стану телекомунікаційних мереж, нейронні мережі, моніторинг стану, адаптивне управління.

Vysochina O. S., Shmatkov S. I., Salman Amer M.

ANALYSIS OF TELECOMMUNICATIONS NETWORKS MONITORING SYSTEMS

The most known systems of telecommunications network monitoring are examined, their comparative analysis is presented, general requirements are formulated and general architecture of the similar systems is synthesized. The proposed solutions are not able to predict the telecommunications network state. Therefore in such systems it is necessary to include additional modules of statistical data processing.

Key words: telecommunication networks state diagnostics, neural networks, state monitoring, adaptive control.

имеют сложные частотно-временные характеристики.

В задаче обнаружения аварийного рассоединения трансмиссии анализ данных сигналов должен прово-

© Дубровин В. И., Афонин Ю. С., Манило Т. В., Харитонов В. Н., 2010

диться в режимі реального часу. Основним інтересуючим в даній задачі параметром являється поведінка несущої частоти сигналу ДЧВ [1]. Крім того, момент аварійного роз'єднання трансмісії ГТД може характеризуватися наявністю короткочасних імпульсів або скачків параметрів сигналів ДЧВ. В даній роботі розглянуті способи виявлення даних імпульсів різних видів в сигналах з ДЧВ. Важним критерієм для вибору математичного апарату обробки сигналу на предмет наявності таких імпульсів являється критерій швидкості та точності їх виявлення. Опереваючись на цьому, в якості основного математичного апарату обрано вейвлет-перетворення, так як вейвлет-фільтри дозволяють з високою точністю виявити локальні зміни характеристик сигналу при малих часових витратах.

КЛАСИФІКАЦІЯ ІМПУЛЬСІВ

Для аналізу скачкообразних змін сигналів необхідно знати можливі види цих змін. На рис. 1 наведено основні можливі види імпульсів в момент аварійного роз'єднання трансмісії ГТД.

К першому класу віднесені сильні імпульси з амплітудою більшою, ніж амплітуда основного сигналу, і часом існування від

5×10^{-5} до 9×10^{-5} секунд, т. е. часом, рівним 3–5 дискретам сигналу при частоті дискретизації 64 кГц.

К другому класу – імпульси середньої потужності з амплітудою в діапазоні від 100 % до 10 % від амплітуди основного сигналу і часом існування від 5×10^{-5} до 9×10^{-5} секунд, т. е. часом, рівним 3–5 дискретам сигналу.

К третьому – слабкі імпульси з амплітудою, рівною не більше 5–10 % від амплітуди основного сигналу, і часом існування від 5×10^{-5} до 9×10^{-5} секунд, т. е. часом, рівним 3–5 дискретам сигналу.

К четвертому – сильні імпульси з амплітудою, рівною амплітуді основного сигналу або більшою, з часом існування від $1,5 \times 10^{-4}$ до $7,8 \times 10^{-4}$ секунд, т. е. часом, рівним 10–50 дискретам сигналу.

АНАЛІЗ ІМПУЛЬСІВ «ПЕРВОГО КЛАСА»

Розглянемо сигнал з імпульсом «первого класу» (рис. 1, а) і його дискретне вейвлет-перетворення до другого рівня розкладання з використанням вейвлет-функції Добеші десятого порядку (db10), яке представлено на рис. 2.

Деталізуючі коефіцієнти швидкого вейвлет-перетворення (БВП) [2, 3] на продовженні

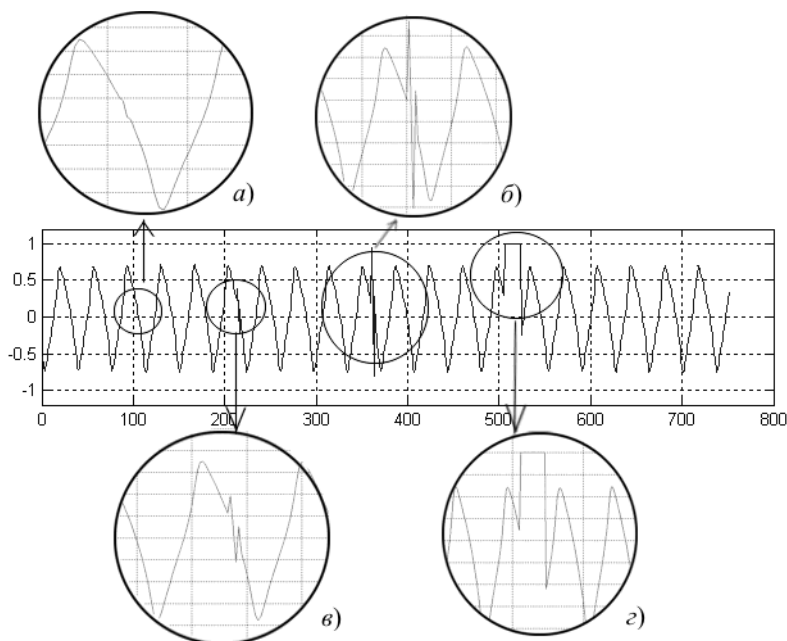


Рис. 1. Основні види імпульсів в момент аварійного роз'єднання трансмісії ГТД:

а – імпульс «третього класу»; б – імпульс «первого класу»; в – імпульс «второго класу»; з – імпульс «четвертого класу»

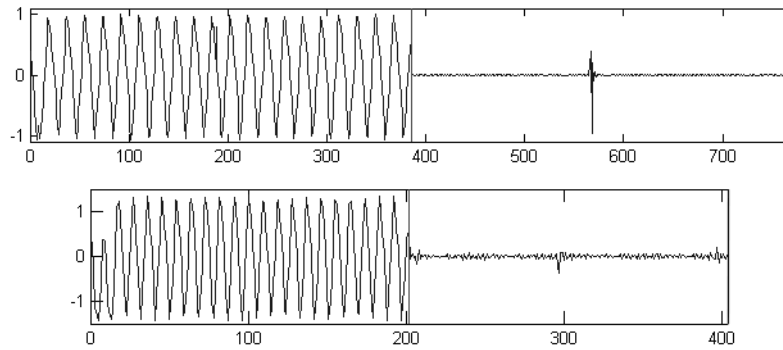


Рис. 2. Вейвлет-преобразование сигнала с импульсом «первого класса»

сигнала, за исключением времени возникновения импульса, представляют собой последовательность, дисперсия которой $\delta_{normal} = 3,6 \times 10^{-3}$, а максимум и минимум соответственно: $\max(coefs_{normal}) = 9 \times 10^{-3}$ и $\min(coefs_{normal}) = -(10,7) \times 10^{-3}$. При этом в момент возникновения импульса детализирующие коэффициенты достигают больших величин относительно остального сигнала. Максимальное значение коэффициента в момент импульса: $\max(coef_{peak}) = 572 \times 10^{-3}$, а минимальное $\min(coef_{peak}) = -351 \times 10^{-3}$.

Рассчитав по формуле (1) относительную мощность (в дБ) коэффициентов вейвлет-преобразования основного сигнала и части сигнала, которая содержит импульс, получим 16,71 дБ.

$$W = 10 \log_{10} \frac{P_{normal}}{P_{peak}}, \quad (1)$$

где W – относительная мощность, дБ; P_{normal} и P_{peak} – соответственно, размах вейвлет-коэффициентов на участке без импульса и на участке с импульсом, рассчитываются по формуле (2):

$$P_k = |\max(coefs_k) - \min(coefs_k)|, \quad (2)$$

где $coefs_k$ – последовательность коэффициентов вейвлет-преобразования.

Такая относительная мощность говорит о том, что, анализируя коэффициенты вейвлет-преобразования, а именно детализирующие коэффициенты, можно с высокой вероятностью выявить появление импульсов «первого класса».

ИССЛЕДОВАНИЕ ИМПУЛЬСОВ «ВТОРОГО КЛАССА» И «ТРЕТЬЕГО КЛАССА». ВЫБОР ОПТИМАЛЬНЫХ ПАРАМЕТРОВ ВЕЙВЛЕТ-РАЗЛОЖЕНИЯ

Следующие два вида сигнала с импульсами «второго класса» и «третьего класса» похожи на сигналы с импульсами «первого класса» по возможности при-

менения метода обработки. Но при обработке сигналов такого рода необходимо учесть, что мощность импульса значительно меньше, чем в сигнале «первого класса», что также скажется и на получаемых в результате вейвлет-преобразования детализирующих коэффициентах. Для обработки сигналов с импульсами «второго класса» и «третьего класса» предлагается использовать пакетное вейвлет-преобразование (ПВП) [3–5]. Для анализа сигнала с помощью ПВП в окнах, соответствующих детализирующим коэффициентам БВП, при выборе метода вейвлет-преобразования следует руководствоваться обеспечением максимальной скорости получения результатов вейвлет-разложения [6].

На рис. 3 и 4 приведены результаты ПВП второго уровня сигналов с импульсами «второго класса» и «третьего класса» соответственно, где в качестве вейвлет-функции использован вейвлет Добеши десятого порядка (db10).

Из рисунков видно, что на первом уровне ПВП, так же как и для сигналов с «первым классом» импульсов, детализирующие коэффициенты возрастают в местах появления импульсов. Но при этом отчетливо видно, что величина коэффициентов в местах импульсов значительно уменьшается при уменьшении мощности самих импульсов и становится сравнимой с детализирующими коэффициентами остальной части сигнала.

В табл. 1 и 2, соответственно, представлены сводные данные по различным окнам ПВП сигнала с импульсом «второго класса» и «третьего класса».

Оптимальным уровнем и окном являются те, при которых W принимает максимальное значение:

$$Level:Window_{optimal} = Level>window, \quad (3)$$

при которых

$$W_{Level:Window} = \max(W), \quad \forall Level = 1, 2, \forall Window = 1, 2^{Level}. \quad (4)$$

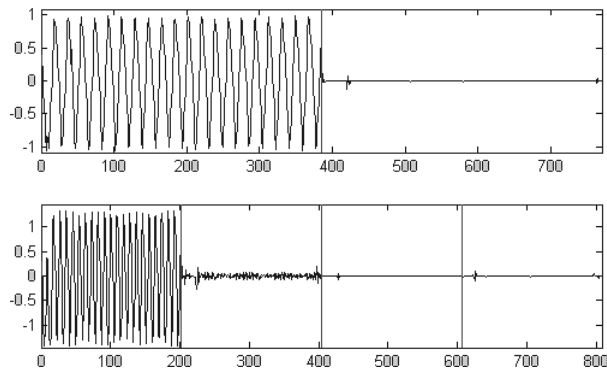


Рис. 3. ПВП сигнала с импульсом «второго класса»

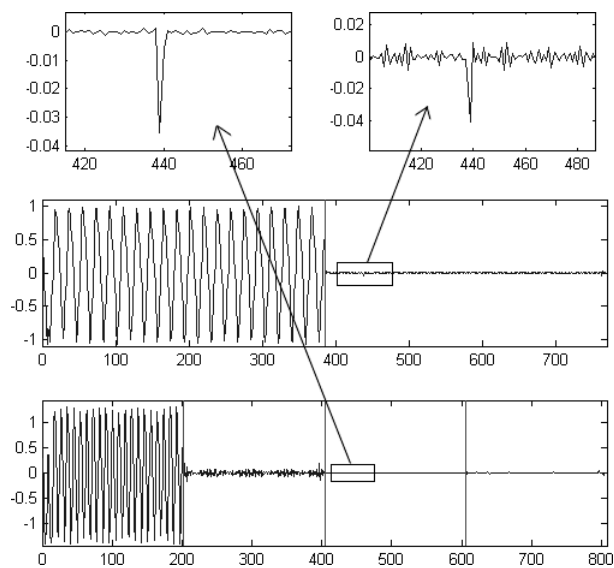


Рис. 4. ПВП сигнала с импульсом «третьего класса»

Таблица 1. Сравнительная характеристика окон ПВП сигнала с импульсом «второго класса»

Уровень : № окна Характеристика	1:1	1:2	2:1	2:2	2:3	2:4
δ_{normal}	0,6283	0,0036	0,8864	0,0386	0,00054	0,005
$\min(coefs_{normal})$	-1,0631	-0,0107	-1,4156	-0,0768	-0,0012	-0,0129
$\max(coefs_{normal})$	0,9921	0,0090	1,3128	0,0717	0,0017	0,0147
$\min(coefs_{peak})$	-1,0631	-0,1509	-1,4156	-0,2764	-0,0822	-0,1139
$\max(coefs_{peak})$	0,9921	0,0757	1,3195	0,1646	0,0490	0,1108
W	0	10,61	0	4,73	16,56	9,11

Таблица 2. Сравнительная характеристика окон ПВП сигнала с импульсом «третьего класса»

Level : Window Характеристика	1:1	1:2	2:1	2:2	2:3	2:4
δ_{normal}	0,6283	0,0036	0,8864	0,0386	0,00054	0,005
$\min(coefs_{normal})$	-1,0631	-0,0107	-1,4156	-0,0768	-0,0012	-0,0129
$\max(coefs_{normal})$	0,9921	0,0090	1,3128	0,0717	0,0017	0,0147
$\min(coefs_{peak})$	-1,0631	-0,0110	-1,4156	-0,0768	-0,0077	-0,0129
$\max(coefs_{peak})$	0,9921	0,0119	1,3128	0,0717	0,0042	0,0147
W	0	0,654	0	0	6,13	0

Сравнивая полученные коэффициенты в разных окнах, полученных после ПВП сигнала с импульсом, можно заметить, что третье окно второго уровня преобразования, как для сигналов с импульсом «второго класса», так и с импульсом «третьего класса», содержит коэффициенты, которые позволяют с наибольшей точностью определить наличие импульса и время его возникновения в сигнале. ПВП на дальнейших уровнях является менее эффективным, и импульс становится менее выраженным в получаемых коэффициентах.

Исходя из выше сказанного, получаем, что для выявления импульсов «второго класса» и «третьего класса» в сигнале достаточно производить вычисления, необходимые только для получения коэффициентов в третьем окне второго уровня ПВП. Для того чтобы иметь возможность производить вычисления отдельно взятых окон ПВП [7], необходимо рассчитать вейвлет-пакет для конкретного окна. Расчет конкретного вейвлет-пакета выполняется на основе выражений

$$\Psi_{2j}(t) = \sqrt{2} \sum_n h_n \psi_j(2t - n), \quad (5)$$

$$\Psi_{2j+1}(t) = \sqrt{2} \sum_n g_n \psi_j(2t - n), \quad (6)$$

где h_n, g_n – вейвлет-последовательность и масштабирующая последовательность, n – количество коэффициентов последовательности h и g , ψ – вейвлет-пакет, ψ_0 – исходная вейвлет-функция.

В данном случае зависимость между сквозной нумерацией окон и уровнем и номером окна на этом уровне представлена в выражении

$$\begin{aligned} \text{>window} &= (2^{\text{Level}} - 2) + \text{Window} - 1 = \\ &= 2^{\text{Level}} + \text{Window} - 3. \end{aligned} \quad (7)$$

Для получения коэффициентов на заданном уровне и заданном окне данного уровня необходимо выполнить свертку сигнала с соответствующим вейвлет-пакетом:

$$\begin{aligned} \text{coefs}_{\text{Level:Window}} &= \Psi_{\text{>window}} \times s = \\ &= \Psi_{2^{\text{Level}} + \text{Window} - 3} \times s. \end{aligned} \quad (8)$$

В рассмотренном решении по обнаружению в сигнале импульсов с «первого класса» по «третий класс» в качестве вейвлет-функции использовалась функция Добеши десятого порядка (db10). Вейвлеты Добеши увеличивают свою чувствительность к выделению импульсов, особенно слабых, от низких к высшим порядкам. При этом бесконтрольное увеличение порядка вейвлета Добеши бессмысленно, нужная точность для обнаружения «третьего класса» импульсов достигается на десятом уровне и дальнейшее увеличение порядка вейвлета приведет лишь к увеличению затрат времени на производимые вычисления.

ИССЛЕДОВАНИЕ ИМПУЛЬСОВ «ЧЕТВЕРТОГО КЛАССА»

Для определения в сигнале импульса «четвертого класса» недостаточно простого выявления увеличения значений коэффициентов относительно среднего значения в сигнале без импульса. Это обусловлено тем, что импульс сильно растянут во времени и вносит значительно больший вклад в коэффициенты ПВП, нежели импульсы «первого» – «третьего класса», и, как следствие, может быть принят за отказ работы датчика. Пример ПВП до второго уровня приведен на рис. 5.

Из рисунков видно, что на первом уровне ПВП, так же как и для сигналов с «первым классом» импульсов, детализирующие коэффициенты возрастают в местах появления импульсов. Но при этом отчетливо видно, что величина коэффициентов в местах импульсов значительно уменьшается при уменьшении мощности самих импульсов и становится сравнимой с детализирующими коэффициентами остальной части сигнала.

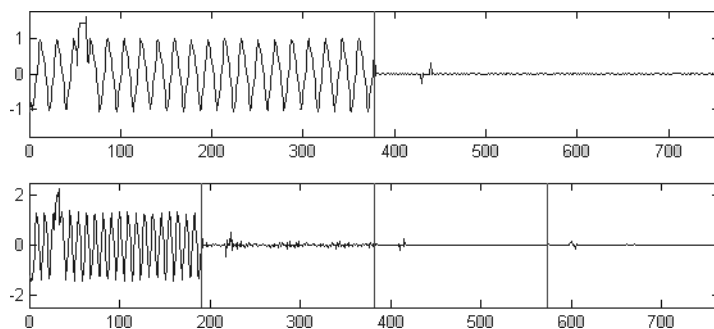


Рис. 5. ПВП сигнала с импульсом «четвертого класса»

В качестве базового вейвлета был выбран вейвлет Добеши третьего порядка (db3). Такое снижение порядка базового вейвлета скажется на его длине и длине его вейвлет-пакетов, т. е. длина вейвлет-функции будет меньше за счет того, что ее можно описать меньшим количеством точек. При длине вейвлет-функции меньшей отрезка сигнала, который содержит исследуемый импульс «четвертого класса», часть коэффициентов преобразования будет иметь одинаковые значения, поскольку значение функции (пика импульса) не будет изменяться во времени.

Кроме того, что после вейвлет-преобразования в детализирующих коэффициентах возможно обнаружение последовательности коэффициентов, принимающих одинаковые значения, необходимы дополнительные условия, которые смогли бы охарактеризовать появление в сигнале именно импульса, а не просто определенной последовательности значений сигнала, которые могут породить такие же последовательности коэффициентов, как и импульс. В качестве таких дополнительных условий можно рассмотреть восходящий и нисходящий фронты импульса «четвертого класса». Как видно из рис. 5, во втором окне первого уровня и во втором – четвертом окнах второго уровня разложения имеются выразительные всплески или отклонения коэффициентов от среднего значения всего остального сигнала. Эти всплески являются достаточно выраженными за счет резко возрастающего и убывающего фронтов импульса. Важным является тот факт, что знаки всплесков (максимального отклонения от средней величины остального сигнала), соответствующие возрастающему и ниспадающему фронту импульса, являются противоположными.

Учитывая все вышеизложенное, для определения наличия в сигнале импульса «четвертого класса» необходимо выполнить свертку сигнала (8) с вейвлет-пакетом соответствующего уровня и окна, определить в полученных вейвлет-коэффициентах наличие последовательности из нескольких одинаковых по значению величин, а также установить наличие «всплесков» на краях найденной последовательности с противоположными знаками и заданной относительной мощностью. Важным вопросом при выявлении импульсов «четвертого класса» остается выбор оптимального уровня и окна для анализа вейвлет-коэффициентов.

ВЫВОДЫ

Выявление импульсов в сигналах с датчиков частоты вращения является одной из проблем в задаче

выявления аварийного рассоединения трансмиссии, так как момент обрыва вала характеризуется скачком параметров сигнала, получаемых с ДЧВ.

Для выявления импульсов различного рода выбран математический аппарат вейвлет-анализа. Выбор обусловлен хорошей частотно-временной локализацией вейвлет-преобразования и высокой скоростью обработки.

Для выявления импульсов предложено рассчитывать относительную мощность последовательных участков коэффициентов вейвлет-разложения, сделан вывод об информативности данной характеристики. Сравнение с порогом относительной мощности позволит судить о наличии импульса в сигнале. Важным является подбор такого порога, который обеспечит достоверность принятия решения о наличии импульса. Не менее важным является выбор параметров вейвлет-преобразования сигнала (порядка вейвлета, уровня разложения, при ПВП – уровня и окна данного уровня).

Перспективным в данном направлении является: анализ вейвлетов других семейств, поиск дополнительных информативных характеристик, которые бы позволили ускорить процесс принятия решений; исследование поведения датчиков при их выходе из строя с целью последующего учета такого поведения при анализе сигналов ДЧВ, что позволит повысить достоверность принятия правильного решения.

СПИСОК ЛИТЕРАТУРЫ

1. *Афонин Ю. С.* Методы обработки сигналов в задаче обнаружения аварийного рассоединения трансмиссии газотурбинного двигателя / Ю. С. Афонин, В. И. Дубровин, В. Н. Харитонов // Вестник двигателестроения : научно-технический журнал. – 2009. – № 3. – С. 15–18: рис., граф.
2. *Дремин И. М.* Вейвлеты и их использование / И. М. Дремин, О. В. Иванов, В. А. Нечитайло // Успехи физических наук. – 2000. – Т. 171, № 5. – С. 465–501.
3. *Дьяконов В. П.* Вейвлеты. От теории к практике / Дьяконов В. П. – Изд. 2-е, перераб. и доп. – М. : Солон-Пресс, 2004. – 400 с.
4. *Чуи К.* Введение в вейвлеты / К. Чуи. – М. : Мир, 2001. – 412 с.
5. *Jaideva C.* Fundamentals of wavelets. Theory, algorithms and applications / Goswami, Andrew K. Chan. – Wiley Interscience, 1999. – 324 p.
6. *Афонин Ю. С.* Быстрое и пакетное вейвлет-преобразование в задачах моделирования объекта / Ю. С. Афонин, В. И. Дубровин, Т. В. Манило // 5-я международная научно-практическая конференция «Математическое и программное обеспечение интеллектуальных систем» : сб. докладов. – Днепропетровск, 2007. – С. 12–13.
7. *Манило Т. В.* Оптимизация схемы пакетного вейвлет-разложения для сокращения времени обработки одномерных сигналов / Т. В. Манило, Ю. С. Афонин // 12-й международный молодежный форум «Радиоэлектроника и молодежь в XXI в.» : сб. материалов форума. – Харьков : ХНУРЭ, 2008. – Ч. 2. – С. 130.

Надійшла 15.02.2010

Дубровін В. І., Афонін Ю. С., Маніло Т. В., Харитонов В. М.

ВЕЙВЛЕТ-АНАЛІЗ КОРОТКОЧАСНИХ ІМПУЛЬСІВ В СИГНАЛАХ З ДАТЧИКІВ ЧАСТОТИ ОБЕРТАННЯ РОТОРА ГТД

В статті розглянута проблема виявлення короткочасних імпульсів в сигналах з датчиків частоти обертання ротора ГТД, які виникають в момент аварійного роз'єднання трансмісії. Подана класифікація даних імпульсів. Аналіз кожного виду імпульсу виконаний з використанням вейвлет-перетворення, як найбільш доцільного для виявлення локальних особливостей характеристик сигналу. Розглядаються перспективні задачі в даному напрямку.

Ключові слова: сигнал ДЧО, вейвлет-аналіз, короткочасний імпульс, коефіцієнти вейвлет-перетворення, вейвлети Добеши.

УДК 004.042

Корниенко С. К.

Канд. техн. наук, доцент Запорозького національного технічного університета

АВТОМАТИЗИРОВАННАЯ ЛОГИСТИЧЕСКАЯ СИСТЕМА

Рассмотрены основные понятия информационной логистики. Проведен анализ информационных потоков торгового предприятия и основных требований, предъявляемых к логистическим системам. Предложена архитектура и описаны функции разработанной логистической информационной системы.

Ключевые слова: информационные потоки, логистика, логистическая информационная система, торговая логистика.

ВВЕДЕНИЕ

В настоящее время коммерческие отношения формируются в условиях высокой конкуренции, неопределенности и неустойчивости рыночной среды. Для того чтобы добиться успеха в предпринимательской деятельности, уже недостаточно использовать лишь маркетинговые подходы, требуется применение современных высокоэффективных способов и методов управления потоковыми процессами.

Результатом управления товарным и информационными потоками в ручном режиме является недостаточная оперативность, полнота и достоверность получения данных. Как следствие – принятие неверных или запоздавших управленческих решений, невозможность предвидеть и моделировать ситуации, низкая эффективность менеджмента в целом.

Спецификой современной торговой компании является широкая номенклатура товаров, которые поступают от разных поставщиков. При этом учет каждого вида товаров имеет свои нюансы: нормы естественного убытка, контроль сроков годности, нормативный запас и т. д. В этих условиях нужны не

только аккуратность и точность, но и высокая скорость работы с информационным потоком и быстрое принятие решений, что невозможно обеспечить только за счет человеческого ресурса [1, 2].

Обеспечить это можно только современными автоматизированными методиками, базирующимися на основе компьютерных технологий. Крайне важно, чтобы собранная информация была структурирована с учетом потребностей потенциальных пользователей и хранилась в форме, позволяющей использовать современные технологии доступа и обработки [3, 4].

Наиболее прогрессивным научно-прикладным направлением в данной области является информационная логистика.

АНАЛИЗ ИНФОРМАЦИОННЫХ ПОТОКОВ ТОРГОВОГО ПРЕДПРИЯТИЯ

Логистика – часть экономической науки и область деятельности, предмет которых заключается в организации и регулировании процессов продвижения товаров от производителей к потребителям, функционирования сферы обращения продукции, това-

© Корниенко С. К., 2010

ров, услуг, управления товарными запасами, создания инфраструктуры товародвижения.

Информационная логистика является неотъемлемой частью всей логистической системы. Объектом изучения информационной логистики являются информационные потоки, отражающие движение материальных, финансовых и других потоков, влияющих на производственный процесс. Основная цель – обеспечение логистических систем информацией в нужные сроки, в нужном объеме и в нужном месте [5].

Эффективное использование информационной логистики заключается в рациональном управлении информационными потоками по всей логистической сети на всех иерархических уровнях.

Информационный поток – это поток сообщений в речевой, документной (бумажной и электронной) и других формах, сопутствующий материальному или сервисному потоку в рассматриваемой логистической системе и предназначенный в основном для реализации управляющих воздействий [6, 7].

Разнообразные информационные потоки, циркулирующие внутри между элементами логистической системы, логистической системой и внешней средой, образуют своеобразную логистическую информационную систему, которая может быть определена как интерактивная структура, состоящая из персонала, оборудования и процедур (технологий), объединенных связанной информацией, используемой логистическим менеджментом для планирования, регулирования, контроля и анализа функционирования логистической системы.

Необходимым условием для успешного функционирования любой сложной системы является нормальное функционирование следующих процессов:

- целенаправленного сбора, первичной обработки и предоставления доступа к информации;
- каналов организации доступа пользователей к собранной информации;
- своевременного получения информации и ее использования для принятия решений.

Логистика оперирует многочисленными показателями и характеристиками информационных потоков:

- номенклатурой передаваемых сообщений;
- типами данных;
- документами;
- массивами данных;
- интенсивностью и скоростью передачи данных;
- специальными характеристиками (пропускной способностью информационных каналов, защитой от несанкционированного доступа, помехозащищенностью и др.).

Между информационным и материальным потоками отсутствует изоморфность (т. е. однозначное соответствие, синхронность во времени возникновения). Как правило, информационный поток либо опережает материальный, либо отстает от него. В частности, само зарождение материального потока обычно является следствием информационных потоков в ходе, например, переговоров по сделкам купли-продажи товаров, составления контрактов и т. д. Типичным является наличие нескольких информационных потоков, сопровождающих материальный поток (рис. 1).

Информационные потоки в торговой логистике содержат следующую информацию [8–10]:

- накладные, поступающие вместе с товаром;
- информацию о поступлении и размещении грузов на складах;
- данные о транспортных тарифах и о возможных маршрутах и типах транспорта;
- изменения в динамических моделях состояния запасов;
- различную нормативно-справочную производственную информацию;
- текущие сведения о производственных мощностях;
- текущие сведения о поставщиках и производителях;
- текущие сведения о продажах;
- изменения в динамических моделях портфеля заказов;
- текущие данные о складах;
- данные о финансовых потоках.

Разнообразные информационные потоки являются теми связями, которые объединяют в единое целое различные функциональные подсистемы логистической информационной системы. В каждой из этих функциональных подсистем реализуются материальные потоки, соответствующие целям, обеспечиваемым этими подсистемами.

Информационные потоки объединяют эти подсистемы в единое целое, так что отдельные цели каждой подсистемы подчиняются общей цели всего производственно-сбытового комплекса. Именно это является основной концепцией логистики.

Анализ структуры информационных потоков позволяет сформулировать перечень основных функций проектируемой логистической информационной системы:

- учет складских запасов;
- учет поставок товаров;
- учет продаж;
- ведение статистики коммерческой деятельности;
- генерация отчетной документации.

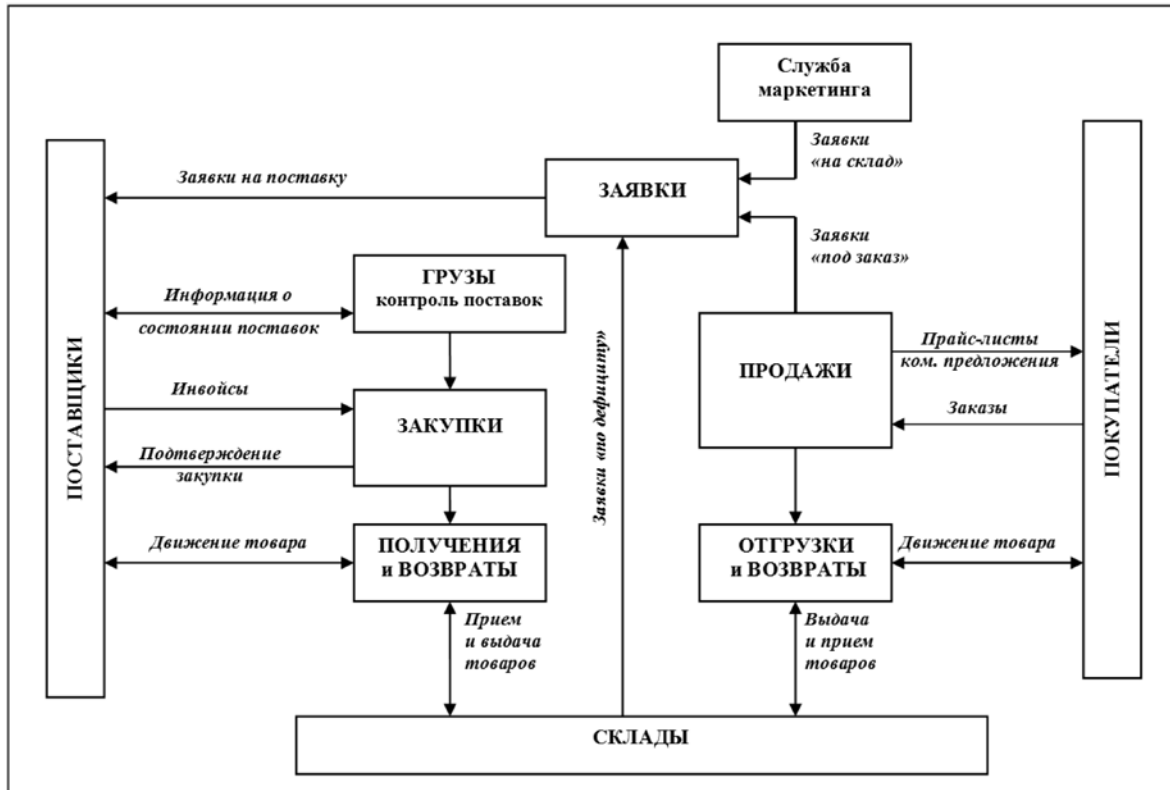


Рис. 1. Схема інформаційних потоків торгового центра

АРХИТЕКТУРА ЛОГИСТИЧЕСКОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

К логистической информационной системе (ЛИС) предъявляется ряд общих технологических требований [2, 4]:

- модульный принцип построения системы, что обеспечит экономию ресурсов при развитии ее функциональных возможностей;

- возможность подключения различных источников данных, интеграция с бухгалтерскими и учетными системами, а также с глобальными информационными системами и специализированными продуктами третьих фирм;

- поддержка единой базы данных клиентов, поставщиков и партнеров и обеспечение для них доступа к системе с разграничением прав доступа.

В ЛИС должны быть реализованы следующие принципы:

- полнота и удобство использования информации для пользователя (предоставлять информацию в том объеме, в то время и в том месте, которые наилучшим образом обеспечивают выполнение соответствующих логистических функций и операций);

- точность и своевременность информации;

- поддержка единой технологии обработки информации;

- работа в едином или в нескольких согласованных информационных пространствах;

- ориентированность на улучшение качества обслуживания потребителей при снижении логистических издержек;

- гибкость (максимальная приспособленность к требованиям всех участников логистического процесса);

- контролируемый доступ к информационным ресурсам;

- настраиваемость системы;

- электронный обмен данными между участниками движения товаров.

Особое значение приобретает организация оперативного обмена информацией как по горизонтали (между структурными подразделениями), так и по вертикали – между руководителями и подчиненными.

Информационная система должна предоставлять информацию в нужное время, требуемого качества и только «нужным потребителям» (защита информации) для принятия ими эффективных решений.

Система движения и распределения информации призвана обеспечивать организацию совместной работы всех подразделений фирмы. Руководители должны иметь информацию в концентрированном виде (только необходимые данные без второстепенных де-

талей, обобщенную, наглядно представленную), но не искаженную и соответствующую тем задачам, которые решаются на данном уровне управления.

В настоящее время на украинском рынке представлено ограниченное число программных продуктов для автоматизации логистической деятельности.

Общим недостатком этих систем является их «закрытость», как с точки зрения расширения и изменения их функциональности, так и при интеграции их с другими программными продуктами, используемыми в компании [11–14].

Кроме того, их характеризует низкая масштабируемость: модернизация одного из компонентов может потребовать внесения изменений в других, в том числе собственно бухгалтерских, модулях.

На кафедре программных средств Запорожского национального технического университета разработана логистическая информационная система, предназначенная для управления информационными потоками торгового центра (ТЦ).

Цель разработанной системы своевременно обеспечить менеджера всей информацией, необходимой для принятия решения. Руководство центра может получать ежедневную, еженедельную, ежеквартальную или годовую информацию о том, какие виды товаров продаются, по какой цене, в каких количествах, что нужно заказать повторно, каковы текущие затраты и доходы, какими они могут быть в следующем году и т. д.

Система обладает открытой модульной структурой (рис. 2). Это позволяет максимально эффективно решать весь спектр задач, стоящих перед организацией, обеспечивает наращивание функциональных возможностей системы и ее гибкую настройку на возможные изменения в структуре информационных потоков. Кроме того, оптимизируются затраты на внедрение за счет поэтапной реализации проекта [15].

К основным функциям системы относятся:

- ведение клиентской базы;
- учет складских запасов;
- учет поставок товаров;
- учет продаж;
- ведение статистики производственной деятельности;
- ведение справочной системы;
- генерация отчетной документации;
- экспорт и импорт данных из других систем (1С, Excel и др.)

Логика работы ЛИС построена на принципе единой базы данных и разных сценариев работы для различных бизнес-процессов. Это позволяет непрерывно работать с информацией о каждом клиенте, заказе или товаре, исключает незапланированное дублирование функций и оптимизирует трудозатраты сотрудников торгового центра.

Система предполагает различные сценарии работы с системой нескольких категорий пользователей:

- руководство ТЦ;
- сотрудники коммерческого отдела;

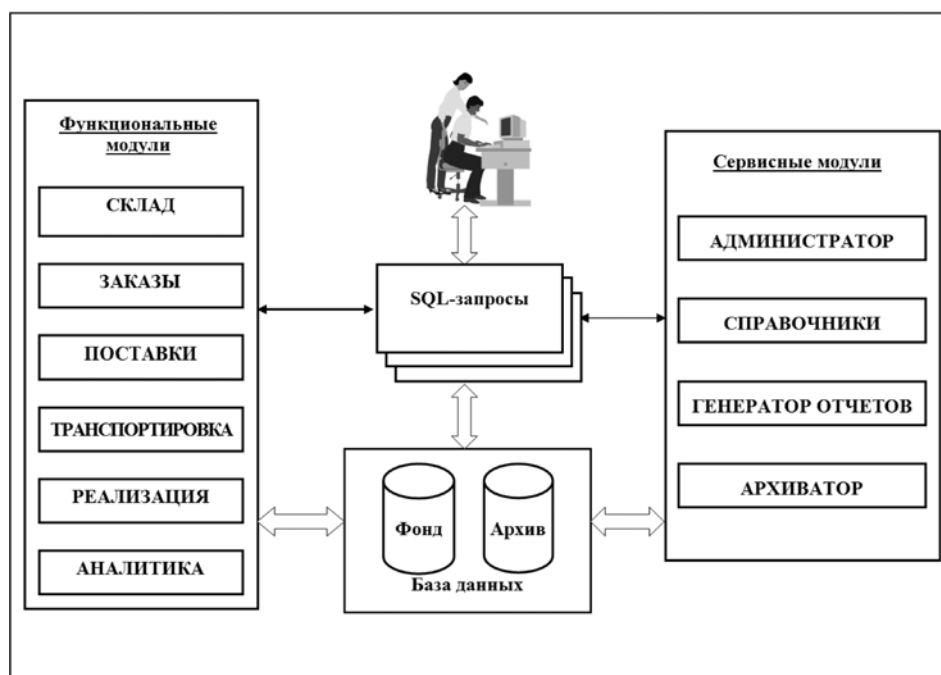


Рис. 2. Архитектура системы

- сотрудники консолідаційного відділу;
- сотрудники транспортної відділу;
- сотрудники складського відділу.

Настройка системи і права доступу базуються на особливостях ролевих функцій підрозділів ТЦ. Наприклад, з клієнтської базою працює тільки комерційний відділ. Для решти відділів клієнтська інформація доступна в формі оброблюваної заявки. Справочник по консолідаційним складам призначений тільки для відділу консолідації, справочник по перевозчикам – для транспортної відділу.

Доступ в режимі адміністрування дозволяє не тільки коректувати базу даних, але і змінювати різні налаштування, параметри доступу і т. п. Їм володіє тільки керівництво і адміністрація системи.

Модулі системи діляться на дві групи:

- функціональні;
- сервісні.

К функціональним модулям відносяться:

- «Склад»;
- «Закази»;
- «Поставки»;
- «Транспортировка»;
- «Реализация»;
- «Аналитика».

Модуль «Склад» призначений для учета і обробки інформації, пов'язаної з товарами, зберігаються на складах (повна характеристика всіх товарів: товарна група, найменування товару, виробник, поставщик, ціна закупочна, ціна отпускна, кількість на складі), а також рухом товарів на складах ТЦ.

Модуль «Закази» служить для учета і обробки інформації, пов'язаної з заказами на поставку товару для ТЦ:

- обробка заявок секцій і мелкооптових магазинів;
- формування і контроль виконання заказів.

Модуль «Поставки» призначений для учета і обробки інформації, пов'язаної з поставкою товару в ТЦ, до якої відносяться дані:

- поставщиках товару;
- складі поставок;
- доставці товару.

Модуль «Транспортировка» служить для учета і обробки обобщенних даних по співпраці з транспортними підприємствами:

- свідень о транспортних агентствах;
- договорів з транспортними агентствами;
- транспортних витрат.

Модуль «Реализация» призначений для учета і обробки інформації, пов'язаної з реалізацією товару, до якої відносяться дані:

- споживачах (найменування і адреса всіх організацій – споживачів продукції);
- складі покупок.

Модуль «Аналитика» служить для поточного і перспективного аналізу ефективності комерційної діяльності торговельної організації. До його функцій відносяться, наприклад, наступні:

- виявлення тенденцій в динаміці продажів;
- аналіз рентабельності товарів;
- аналіз даних за попередні/наступні періоди;
- аналіз абсолютних і відсоткових змін результатів роботи;
- оцінка результатів по календарним і фінансовим рокам.

До сервісним модулям відносяться:

- «Адміністратор» (розподіл прав доступу до інформації);
- «Справочники» (редагування нормативно-справочної інформації);
- «Генератор звітів» (генерація звітної документації);
- «Архіватор» (перенесення оброблених даних з бази даних системи в архів з метою зменшення обсягу робочої області бази даних і підвищення швидкості обробки інформації).

ВИВОДИ

Внедрення розробленої автоматизованої логістичної системи на торговельному підприємстві забезпечує:

- визначення оптимального рівня товарних запасів, усунення заговарювання і дефіциту товарів;
- підвищення точності прогнозів, якісніше покращення процесів підготовки і прийняття рішень;
- зменшення трудомісткості процесів обробки і використання даних;
- зменшення витрат на неликвідні товари, зростання щоденної виручки в довгостроковій перспективі;
- оцінку рентабельності товарів, що дозволяє краще формувати асортимент, підвищувати частку ходових товарів, а в кінцевому підсумку ефективніше використовувати оборотні засоби, торговельні і складські приміщення;
- скорочення транспортних витрат, чого вдасться досягти за рахунок високої узгодженості учасників в питаннях використання транспорту;

– повышение оперативности работы;
 – возможность консолидировать данные о продажах в едином центре;
 – отсутствие ошибок учета, сокращение затрат времени на подготовку отчетности, повышение эффективности работы персонала.

Совокупный эффект от использования логистики, как правило, превышает сумму эффектов от улучшения перечисленных показателей. Это объясняется возникновением у логистически организованных систем ценной для рынка способности обеспечивать поставку нужного груза, необходимого качества, в нужном количестве, в нужное время, в нужное место с минимальными затратами.

СПИСОК ЛИТЕРАТУРЫ

1. *Ферн Дж.* Логистика и управление розничными продажами / Дж. Ферн. – Новосибирск : Сибирское университетское изд-во, 2007. – 263 с.
2. *Taylor G. Don.* Introduction to Logistics Engineering / G. Don Taylor. – CRC Press, 2009. – 320 p.
3. *Naasis H.-D.* Dynamics in Logistics / H.-D. Naasis, H.-J. Kreowski, B. Scholz-Reiter. – Springer, 2008. – 465 p.
4. *Сергеев В. И.* Логистика: информационные системы и технологии / В. И. Сергеев, М. Н. Григорьев, С. А. Уваров. – М. : Альфа-Пресс, 2008. – 435 с.
5. *Джонсон Д.* Современная логистика / Д. Джонсон, Д. Вуд, Д. Вордлоу. – М. : Вильямс, 2009. – 546 с.
6. *Mangan J.* Global Logistics and Supply Chain Management / John Mangan, Chandra Lalwarni, Tim Butcher. – John Wiley & Sons, 2008. – 372 p.
7. *Сток Д. Р.* Стратегическое управление логистикой / Д. Р. Сток. – М. : Инфра-М, 2005. – 390 с.
8. *Неруш Ю. М.* Коммерческая логистика / Ю. М. Неруш. – М. : Банки и биржи, 2004. – 347 с.

9. *Джестон Н.* Управление бизнес-процессами. Практическое руководство по успешной реализации проектов / Н. Джестон. – М. : Альпина, Бизнес Букс, 2008. – 439 с.
10. *Harrison Alan.* Logistics management and Strategy: Competing through the supply chain / Alan Harrison, Remko I. van Hoek. – FT Prantice Hall, 2008. – 316 p.
11. *Аникин Б. А.* Логистика / Б. А. Аникин, Т. А. Родкина. – М. : Проспект, 2010. – 406 с.
12. *Круминьш Н.* Логистика в Восточной Европе / Н. Круминьш, К. Витолиньш. – М. : SIA, 2007. – 305 с.
13. *Джонсон Д.* Современная логистика / Д. Джонсон, Д. Вуд, Д. Вордлоу. – М. : Вильямс, 2009. – 420 с.
14. S2B – Solutions to business [Электронный ресурс]. – Электрон. дан. – К., 2010. – Режим доступа: <http://www.s2b.com.ua/about.html>, свободный. – Загл. с экрана.
15. *Корнієнко С. К.* Системи баз даних: організація та проектування : навч. посібник / С. К. Корнієнко. – Запоріжжя : ЗНТУ, 2006. – 252 с.

Надійшла 14.01.2010
 Після доробки 29.03.2010

Корнієнко С. К.

АВТОМАТИЗОВАНА ЛОГІСТИЧНА СИСТЕМА

Розглянуто основні поняття інформаційної логістики. Проведено аналіз інформаційних потоків торговельного підприємства та основних вимог, які висуваються до логістичних систем. Запропоновано архітектуру та описано функції розробленої логістичної інформаційної системи.

Ключові слова: інформаційні потоки, логістика, логістична інформаційна система, торгова логістика.

Kornienko S. K.

AUTOMATIC LOGISTICS SYSTEM

The main concepts of information logistics are considered. Information flows of mercantile business and main requirements to the logistic system are analyzed. Architecture of the designed logistic information system is proposed and its functions are described.

Key words: information flows, logistics, logistics information system, trade logistics.

УДК 004.75

Шкарупило В. В.¹, Касьян К. Н.², Маркин А. Г.³

¹Студент Запорозького національного технічного університета

²Канд. техн. наук, доцент Запорозького національного технічного університета

³Старший преподаватель Запорозького національного технічного університета

ИССЛЕДОВАНИЕ С ПОМОЩЬЮ СТАТИСТИЧЕСКИХ ПРОГРАММ ЦЕЛЕСООБРАЗНОСТИ ВНЕДРЕНИЯ СТАНДАРТА 10G-ETHERNET В КЛАСТЕРАХ ВУЗОВ

Предложена математическая модель комплексной оценки получаемого эффекта от перевода кластеров ВУЗов на стандарт 10G-Ethernet, основанная на данных квартальных отчетов, предоставляемых аналитической компанией IDG. Приведены результаты проверки модели, которые свидетельствуют об обоснованности перехода на новый стандарт в 2010 году.

Ключевые слова: ВУЗ, Gigabit Ethernet, 10G-Ethernet, кластер.

ВВЕДЕНИЕ

Эффективность работы кластерных систем во многом зависит от пропускной способности канала,

посредством которого отдельные его узлы обмениваются между собой промежуточными данными. Необходимость значительного повышения числа переда-

© Шкарупило В. В., Касьян К. Н., Маркин А. Г., 2010

ваемых бит за единицу времени диктуется высокой сложностью современных исследовательских задач, что, в свою очередь, предъявляет высокие требования к пропускной способности тракта. Широкое же распространение кластерных систем объясняется их гибкостью (отсутствие совместно используемой оперативной памяти и отдельной операционной системы для каждого узла) [1]. Наиболее эффективной областью применения кластеров являются научные приложения, характеризующиеся высокой степенью параллелизма. Существует множество методов оценки экономической эффективности внедрения новых сетевых технологий. Особенностью же данной модели является то, что она основывается на учете стоимости внедрения. Следует отметить также, что вопрос снижения финансовых расходов приобретает особую актуальность, в частности, в ВУЗах, что существенно повышает важность приоритета минимизации стоимости.

Предполагается, что на текущий момент кластеры базируются на физическом уровне протокола TCP/IP – Gigabit Ethernet, и мы рассматриваем вопрос повышения пропускной способности канала в 10 раз. Статистические модели получены с использованием программного пакета TableCurve v5.01 [2].

ПОСТАНОВКА ЗАДАЧИ

В данной работе предлагается подход, основанный на статистических данных, полученных компанией IDG за указанный временной отрезок. Подход также предполагает введение двух критериев: первый характеризует степень присутствия стандарта на рынке локальных сетей, второй – стоимость внедрения нового стандарта.

В работе предлагается производить оценку эффективности перехода со стандарта Gigabit Ethernet на стандарт 10G-Ethernet, т. е. предполагается, что первый уже эксплуатируется учебным учреждением. Возможность реализации транков на Fast Ethernet не рассматривается по причине несоответствия современным требованиям. Следует также отметить, что имеется и более производительное решение, действующие образцы которого уже существуют, при этом обеспечивается пропускная способность в 100 Гб/с, однако эти образцы далеки от широкомасштабного рыночного внедрения, ввиду чрезвычайно высокой стоимости реализации.

Необходимо определить, является ли экономически обоснованным перевод кластера на стандарт 10G-Ethernet. Обоснованность данного перехода будем строить на основании статистических данных квартальных отчетов аналитической компании IDG за пе-

риод с 2004 по 2009 год включительно. Эти данные будут затрагивать две области: первая – доли рынка локальных сетей стандартов Gigabit Ethernet и 10G-Ethernet; вторая – полная относительная стоимость внедрения каждого из вышеназванных стандартов.

МОДЕЛЬ ОЦЕНКИ ЭФФЕКТИВНОСТИ

Пусть существует некоторый набор компьютеров, рассматриваемый операционной системой, системным программным обеспечением, приложениями и пользователями как единая система.

Введем следующее обозначение: n – размер выборки (количество лет, которые охватывают полученные нами статистические данные).

Предположим, что будем оперировать квартальными данными, охватывающими некоторое число лет, тогда имеет смысл представить этот диапазон в виде отдельных рядов:

$$s_1, \dots, s_n, \quad (1)$$

где s_1 – доля рынка локальных сетей, занимаемая стандартом Gigabit Ethernet за 1-й год выборки; s_n – за n -й год. Годовые значения будем получать путем линейного интерполирования квартальных. Соответственно, для стандарта 10G-Ethernet введем ряд

$$ts_1, \dots, ts_n. \quad (2)$$

Относительная стоимость внедрения Gigabit Ethernet за каждый год, входящий в заданный диапазон, представляется рядом вида

$$c_1, \dots, c_n. \quad (3)$$

Для стандарта 10G-Ethernet относительная стоимость внедрения будет представлена в виде следующего ряда:

$$tc_1, \dots, tc_n. \quad (4)$$

Анализируя долю присутствия на рынке локальных сетей стандарта 10G-Ethernet, на основании данных IDG, начиная с 2004 года, и, используя пакет TableCurve v5.01, мы сделали вывод, что лучше всего подходит следующая статистическая модель:

$$ts_i = \sqrt{a + b \cdot i}, \quad (5)$$

где a, b – коэффициенты, причем $a \ll b$; i – порядковый номер года в ряде, начиная с 1-го; ts_i – доля рынка локальных сетей, занимаемая стандартом 10G-Ethernet за i -й год выборки.

В случае с Gigabit Ethernet, статистическая модель имеет вид линейной зависимости

$$s_i = a + b \cdot i, \quad (6)$$

где a, b – коэффициенты, причем $a \ll b$; s_i – доля рынка локальных сетей, занимаемая стандартом Gigabit Ethernet за i -й год выборки.

Касательно стоимости внедрения, статистические модели для двух стандартов будут иметь одинаковый вид (разница будет заключаться в коэффициентах a, b):

$$tc_i = 1/(a + b \cdot i), \quad (7)$$

$$c_i = 1/(a + b \cdot i), \quad (8)$$

причем $a > b$; tc_i – относительная стоимость внедрения 10G-Ethernet в i -м году; c_i – относительная стоимость внедрения Gigabit Ethernet в i -м году.

Теперь введем следующие критерии: K_i^{ts} – критерий степени присутствия на рынке в i -м году; K_i^{tc} – критерий стоимости внедрения в i -м году.

K_i^{ts} определяется как отношение (5) к (6):

$$K_i^{ts} = ts_i/s_i. \quad (9)$$

K_i^{tc} , в свою очередь, определяется отношением (7) к (8):

$$K_i^{tc} = tc_i/c_i. \quad (10)$$

Будем рассчитывать значения K_i^{ts} и K_i^{tc} за n -й и за $(n+1)$ -й годы. Если $K_{n+1}^{ts} > K_n^{ts}$ и $K_{n+1}^{tc} > K_n^{tc}$, то положительная тенденция снижения стоимости внедрения 10G-Ethernet, с ростом доли рынка, сохраняется. Однако это замечание не учитывает резкость возрастания критериев: в идеале, более быстрыми темпами должен уменьшаться именно критерий стоимости внедрения K_i^{tc} , т.к. увеличивается не только доля рынка локальных сетей, но и уменьшается себестоимость отдельных компонент, т.е. предлагается оценить относительные значения того, как сильно изменились величины критериев, по сравнению с предыдущим годом.

Выразим через P_{ts} относительный прирост K_{n+1}^{ts} , в сравнении с K_n^{ts} :

$$P_{ts} = (K_{n+1}^{ts} - K_n^{ts})/K_n^{ts}. \quad (11)$$

Аналогично, выразим через P_{tc} относительное уменьшение K_{n+1}^{tc} , в сравнении с K_n^{tc} , поскольку рассматриваем снижение полной стоимости внедрения:

$$P_{tc} = (K_{n+1}^{tc} - K_n^{tc})/K_n^{tc}. \quad (12)$$

Переход на стандарт 10G-Ethernet можно считать обоснованным, если выполняется следующее условие:

$$P_{tc} > P_{ts}. \quad (13)$$

Чем больше P_{tc} , в сравнении с P_{ts} , тем выше получаемый эффект от перехода на новый стандарт.

РЕЗУЛЬТАТЫ ПРОВЕРКИ МОДЕЛИ

Приведем график (рис. 1), полученный на основании статистических данных компании IDG [3], характеризующий долю рынка, занимаемую стандартами Gigabit Ethernet и 10G-Ethernet в период с 2004 по 2009 год включительно.

Для ts_i формулы (5) значения коэффициентов a, b имеют следующий вид:

$$a = 0,018, \quad b = 0,007.$$

ts_6 , согласно рис. 1, получаем как среднеарифметическое по четырем квартальным значениям за 2009 год:

$$ts_6 = (0,232 + 0,236 + 0,24 + 0,243)/4 = 0,238.$$

Для получения прогноза на 2010 год подставим в (5) вместо i – значение 7, т.к. исходные статистические данные охватывают временной промежуток в 6 лет.

Получим: $ts_7 = 0,26$.

Для s_i формулы (6) значения коэффициентов a, b имеют следующий вид:

$$a = 0,39, \quad b = 0,038.$$

s_6 получаем как среднеарифметическое по четырем квартальным значениям за 2009 год: $s_6 = (0,61 + 0,62 + 0,633 + 0,642)/4 = 0,627$.

Для получения прогноза на 2010 год, подставим в (6) вместо i – значение 7.

Получим: $ts_7 = 0,656$.

Тогда, согласно (9), найдем K_6^{ts} и K_7^{ts} (за 2009 и 2010 годы соответственно):

$$K_6^{ts} = 0,238/0,627 = 0,38;$$

$$K_7^{ts} = 0,26/0,656 = 0,4.$$

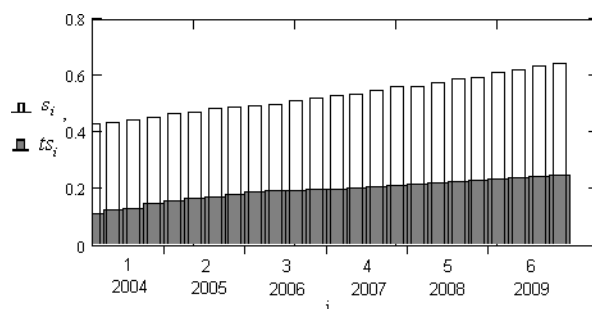


Рис. 1. Рынок локальных сетей

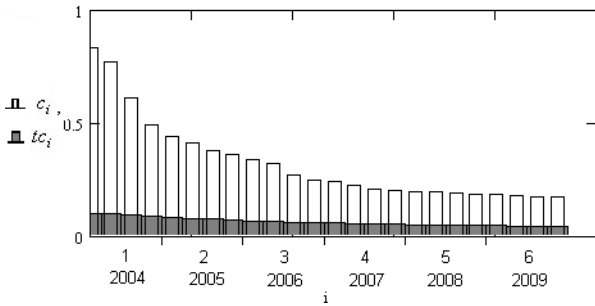


Рис. 2. Относительная стоимость внедрения стандартов Gigabit Ethernet и 10G-Ethernet

Согласно (11), найдем P_{ts} :

$$P_{ts} = (0,4 - 0,38) / 0,38 = 0,05.$$

Из полученного значения P_{ts} видно, что доля рынка стандарта 10G-Ethernet растет на 5 % быстрее, нежели рыночная доля Gigabit Ethernet, что дает основания делать предположения о более стремительном снижении цен на первый из вышеназванных стандартов, в сравнении с последним.

Далее делается попытка подтвердить данное предположение. Приведем график (рис. 2), полученный на основании статистических данных компании IDG, характеризующий относительную стоимость внедрения стандартов Gigabit Ethernet и 10G-Ethernet в период с 2004 по 2009 год включительно.

Для tc_i формулы (7) значения коэффициентов a, b имеют следующий вид:

$$a = 4,58, b = 1,33;$$

$$tc_6 = (0,084 + 0,082 + 0,08 + 0,078) / 4 = 0,081;$$

$$tc_7 = 0,072.$$

Для c_i формулы (8) значения коэффициентов a, b имеют вид:

$$a = 0,58, b = 0,92;$$

$$c_6 = (0,19 + 0,18 + 0,176 + 0,17) / 0,18;$$

$$c_7 = 0,14;$$

$$K_6^{tc} = 0,081 / 0,18 = 0,45;$$

$$K_7^{tc} = 0,072 / 0,14 = 0,51.$$

Согласно (13), найдем P_{tc} :

$$P_{tc} = (0,51 - 0,45) / 0,45 = 0,13.$$

Из значения P_{tc} видим, что стоимость внедрения 10G-Ethernet снижается на 13 % быстрее, нежели таковая для Gigabit Ethernet, что дает основания говорить о тенденции увеличения привлекательности стандарта 10G-Ethernet на рынке локальных сетей.

Видим, что $P_{tc} > P_{ts}$, а это говорит о том, что наши предположения верны (стоимость нового стандарта снижается быстрее, в сравнении с Gigabit Ethernet, чем растет доля рынка локальных сетей).

Эти данные демонстрируют позитивную тенденцию дальнейшего проникновения стандарта 10G-Ethernet на рынок локальных сетей и дают все основания рассматривать переход на него со стороны ВУЗов в ближайшем будущем (2010 год).

ВЫВОДЫ

Разработана математическая модель комплексной оценки эффекта, получаемого от перевода кластеров ВУЗов на стандарт 10G-Ethernet. Результаты экспериментальных исследований свидетельствуют о большем темпе снижения стоимости внедрения 10G-Ethernet, в сравнении с Gigabit Ethernet, что говорит о повышении экономической привлекательности первого, для внедрения в ВУЗах. В дальнейшем планируется расширить данную модель и применительно к 100G-Ethernet.

СПИСОК ЛИТЕРАТУРЫ

1. Топорков В. В. Модели распределенных вычислений / В. В. Топорков – М. : ФИЗМАТЛИТ, 2004. – 240 с.
2. Тюрин Ю. Н. Анализ данных на компьютере / Ю. Н. Тюрин, А. А. Макаров ; под. ред. В. Э. Фигурнова. – 3-е изд., перераб. и доп. – М. : ИНФРА-М, 2003. – 544 с.
3. Jack W. Plunkett. Plunkett's Entertainment & Media Industry Almanac / Jack W. Plunkett. – Houston, Texas 77254-1737 USA : Plunkett Research, 2009. – P. 590.

Надійшла 25.11.2009
Після доробки 02.04.2010

Шкарупило В. В., Касьян К. М., Маркін О. Г.

ДОСЛІДЖЕННЯ ЗА ДОПОМОГОЮ СТАТИСТИЧНИХ ПРОГРАМ ДОЦІЛЬНОСТІ ВПРОВАДЖЕННЯ СТАНДАРТУ 10G-ETHERNET У КЛАСТЕРАХ ВУЗІВ

Запропонована математична модель комплексної оцінки ефекту від переведення кластерів ВНЗів на стандарт 10G-Ethernet, що базується на даних квартальних звітів, наданих компанією IDG. Наведені результати перевірки моделі, що свідчать про обгрунтованість переходу на новий стандарт в 2010 році.

Ключові слова: ВНЗ, Gigabit Ethernet, 10G-Ethernet, кластер.

Shkarupilo V. V., Kasyan K. N., Markin A. G.

THE RESEARCH OF 10G-ETHERNET STANDARD IMPLEMENTATION IN UNIVERSITIES' CLUSTERS USING STATISTICAL SOFTWARE

The mathematical model of final effect complex estimation, as a result of universities' clusters shifting from Gigabit Ethernet standard to 10G-Ethernet, based on IDG company's quarter data, is proposed. Model testing results represent the fact, that shifting to a new standard in 2010 is well-grounded.

Key words: University, Gigabit Ethernet, 10G-Ethernet, cluster.

ИССЛЕДОВАНИЕ ФИЛЬТРУЮЩИХ СВОЙСТВ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

В работе исследованы фильтрующие свойства непрерывного вейвлет-преобразования. Приведен метод уменьшения влияния краевых эффектов на результаты вейвлет-анализа. Выполнено обоснование выбора типа вейвлета для анализа нестационарных сигналов. Предложен метод построения банка вейвлетных фильтров с необходимыми анализирующими свойствами. Разработанный метод применен к задаче диагностирования подшипников качения по вибрационному сигналу.

Ключевые слова: нестационарный сигнал, вейвлет-фильтрация, банк фильтров, вейвлет Морле.

ВВЕДЕНИЕ

Наиболее широко используемым методом обработки цифровых сигналов является преобразование Фурье. Однако оно имеет ряд недостатков, которые привели к появлению и разработке новых усовершенствованных методов цифрового анализа нестационарных сигналов. Среди недостатков преобразования Фурье можно выделить усреднение характерных особенностей по всей длительности сигнала, что делает нецелесообразным применение данного метода при необходимости анализа изменений сигнала во времени.

На сегодняшний день все большее распространение получает вейвлет-преобразование [1–4]. Этот инструмент позволяет получить частотно-временное представление анализируемого сигнала.

Однако зачастую на практике вейвлет-анализ сигналов выполняется без учета особенностей настройки вейвлетов, которые существенно влияют на результаты преобразования и возможность их правильной интерпретации. Применение вейвлет-преобразования без учета частотных характеристик вейвлета может давать искаженную информацию об особенностях анализируемого сигнала. Такой подход может привести к неуниверсальности разрабатываемых методов в силу того, что при изменении характеристик исследуемого сигнала применение одних и тех же настроек метода может давать совершенно непредсказуемые результаты.

ПОСТАНОВКА ЗАДАЧИ

Для успешного использования вейвлет-преобразования необходим тщательный анализ параметров мате-

ринского вейвлета. Это позволит гибко управлять настройкой вейвлета для получения необходимых анализирующих свойств вейвлет-преобразования и правильно интерпретировать получаемые результаты.

Целью данной работы является исследование фильтрующих возможностей непрерывного вейвлет-преобразования, обоснование выбора материнского вейвлета и разработка метода построения анализирующего банка вейвлетных фильтров.

ПРЕДСТАВЛЕНИЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ В ВИДЕ ОПЕРАЦИИ ФИЛЬТРАЦИИ

Непрерывное вейвлет-преобразование представляет собой разложение анализируемого сигнала в базисе анализирующей вейвлетной функции.

Базис вейвлет-преобразования строится путем масштабных преобразований и переносов материнского вейвлета $\psi(t)$ с непрерывными значениями базисных параметров – масштабного коэффициента a и параметра сдвига b :

$$\Psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right), \quad a, b \in R, \quad (1)$$

где $\Psi_{a,b}(t)$ – дочерний вейвлет базиса.

В интегральной форме вейвлет-преобразование непрерывного сигнала $s(t)$ может быть выражено:

$$W(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} s(t) \Psi_{a,b}^*(t) dt, \quad (2)$$

где $*$ – оператор комплексного сопряжения.

Используя равенство Парсеваля, которое описывает связь между функциями и их образами Фурье, выражение (2) может быть записано в альтернативной форме:

$$W(a, b) = \sqrt{a}F^{-1}\{S(f)\Psi^*(af)\}, \quad (3)$$

где F^{-1} – оператор обратного преобразования Фурье; $\Psi(af)$ – преобразование Фурье дочернего вейвлета; $S(f)$ – преобразование Фурье анализируемого сигнала; f – частота; a – масштабный коэффициент.

Такое представление вейлет-преобразования позволяет значительно ускорить вычисление вейлет-коэффициентов за счет использования быстрого преобразования Фурье (БПФ). К тому же, преобразование Фурье материнского вейвлета обычно известно в аналитическом виде и, следовательно, количество вычислений сокращается.

Выражение (3) также показывает, что вейлет-преобразование может быть рассмотрено как специальная операция фильтрации с вейвлетом в качестве ядра фильтра.

ЧАСТОТНО-ВРЕМЕННЫЕ ХАРАКТЕРИСТИКИ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Рассмотрим частотно-временные характеристики непрерывного вейлет-преобразования [5, 6].

Выражение (2) представляет вейлет-преобразование как средство анализа во временно-масштабной области. В таком виде интерпретация масштабов и их соответствие частотным диапазонам могут быть затруднены. Если же рассматривать вейлет-преобразование в частотной области, частотный образ материнского вейвлета соответствует полосовому фильтру, характеристики которого определяются типом и настройками вейвлета.

Пусть вейвлет $\psi(t)$ и его образ Фурье $\Psi(f)$ являются функциями-окнами с центрами t_0 и f_0 , радиусами δ_f и δ_t соответственно.

На рис. 1 изображено частотно-временное разрешение непрерывного вейлет-преобразования. Оно иллюстрирует анализирующие возможности вейвлетов на различных масштабах.

Благодаря свойствам смещения и масштабирования материнского вейвлета каждый дочерний вейвлет на масштабе a будет иметь ширину окна, равную $2\delta_f/a$, и центральную частоту, равную f_0/a . Их отношение не зависит от масштаба преобразования и является постоянной величиной $2\delta_f/f_0$, значение которой задается параметрами материнского вейвлета.

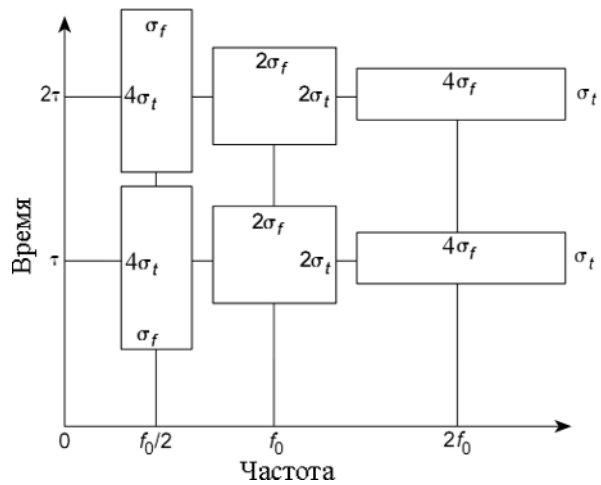


Рис. 1. Частотно-временное разрешение вейлет-преобразования

Таким образом, вейлет-преобразование также может быть рассмотрено как операция фильтрации с постоянной добротностью (с постоянной Q), при которой отношение среднеквадратической ширины полосы пропускания к центральной частоте фильтра является постоянной величиной.

Если рассмотреть частотно-временное окно, характеризующее разрешение вейлет-анализа, можно отметить, что его площадь постоянна и равна $4\delta_f\delta_t$. При масштабировании материнского вейвлета радиус окна дочернего вейвлета во временной области увеличивается прямо пропорционально масштабу ($\delta_t a$), а в частотной радиус окна уменьшается обратно пропорционально масштабу (δ_f/a). Таким образом, вейлет-преобразование предоставляет переменное разрешение в частотно-временной плоскости, которое дает ему большое преимущество при анализе сигналов, так как быстрые вариации сигналов (высокочастотные характеристики) хорошо локализованы, а для выявления медленно меняющихся характеристик необходимо большое временное окно.

Согласно принципу неопределенности Гейзенберга между временной и частотной локализациями площадь частотно-временного окна вейлет-преобразования ограничена $\delta_f\delta_t \geq (4\pi)^{-1}$, что определяет ограничение разрешающей способности вейлет-анализа [1].

ВЫБОР МАТЕРИНСКОГО ВЕЙВЛЕТА

Среди существующих вейлет-функций, используемых в непрерывном вейлет-анализе, вейвлет Морле и вейвлет Мексиканская шляпа более всего по форме подобны импульсным составляющим нестационарных сигналов.

ционарных сигналов, что делает их наиболее подходящим средством анализа такого вида сигналов.

Оба эти вейвлета являются экспоненциально убывающими как во временной, так и в частотной области. Принципиальное различие между ними заключается в следующем: вейвлет Морле имеет узкий спектральный диапазон и продолжителен во временной области, тогда как вейвлет Мексиканская шляпа более локализован во временной области за счет более широкого частотного представления [3]. В силу принципа неопределенности Гейзенберга невозможно получить идеальную локализацию одновременно в частотной и временной областях.

При выборе анализирующего вейвлета для фильтрации сигналов предпочтение было отдано вейвлету Морле, который имеет узкий образ Фурье. Присутствие доминирующей частоты позволяет варьировать избирательность вейвлета Морле в частотной области.

Кроме того, в качестве анализирующего был выбран комплексный тип вейвлета Морле. Его преобразование Фурье равно нулю для отрицательных частот, что позволяет разделить фазовые и амплитудные компоненты сигнала при выполнении вейвлет-преобразования.

Во временной области комплексный вейвлет Морле представляет собой комплексную экспоненту, модулируемую функцией Гаусса:

$$\psi(t) = \frac{\sigma}{\sqrt{\pi}} e^{-\sigma^2 t^2} e^{i2\pi f_0 t}, \quad (4)$$

где t – время; σ и f_0 – параметры вейвлета, задающие его форму.

В частотной области вейвлет Морле имеет форму Гауссова окна с центральной частотой f_0 и шириной σ :

$$\Psi(f) = \Psi^*(f) = e^{-\frac{\pi^2}{\sigma^2}(f-f_0)^2}, \quad (5)$$

где $\Psi(f)$ – преобразование Фурье вейвлета Морле; f_0 – центральная частота анализирующего вейвлета; σ – ширина анализируемого диапазона частот; * означает комплексное сопряжение.

Так как $\Psi(f)$ является действительным числом, то $\Psi(f) = \Psi^*(f)$.

Таким образом, частотный диапазон, покрываемый окном вейвлета Морле, практически ограничен интервалом $[f_0 - \sigma/2, f_0 + \sigma/2]$.

На рис. 2 приведен пример представления вейвлета Морле во временной и частотной областях.

Строго говоря, вейвлет Морле не удовлетворяет условию допустимости для вейвлетов:

$$C_\psi = \int_0^\infty \frac{|\Psi(f)|^2}{f} df < \infty. \quad (6)$$

Это условие также подразумевает, что вейвлет не должен иметь нулевой частотной компоненты $\Psi(0) = 0$ или же, другими словами, вейвлет должен иметь нулевое среднее.

Это означает, что вейвлет Морле не обладает компактным носителем. Однако наибольшая часть его энергии сосредоточена на интервале, ширина которого определяется его полосой пропускания, и на практике при $f_0 \gg 0$ вейвлет Морле может быть использован с минимальной погрешностью [3].

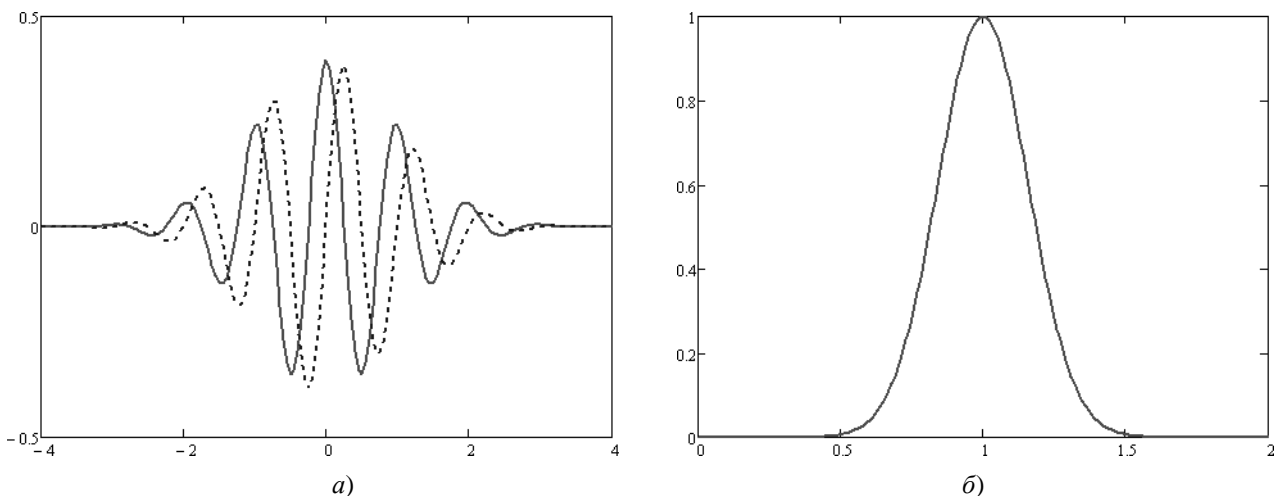


Рис. 2. Представление вейвлета Морле при $\sigma = 0,7$ и $f_0 = 1$:

a – во временной области: реальная часть (–) и мнимая (–); b – в частотной области

КРАЕВЫЕ ЭФФЕКТЫ И ОБЛАСТЬ ДОСТОВЕРНОСТИ ДЛЯ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Рассмотрим механизм возникновения граничного эффекта при выполнении вейвлет-преобразования сигнала [7, 8].

Значение анализируемой функции в точке t_0 влияет на значения коэффициентов преобразования в растущих с масштабом временных диапазонах, в так называемом угле влияния (рис. 3).

Если анализирующий вейвлет $\psi(t)$ хорошо локализован в интервале $2\delta_t$ для $a = 1$, то коэффициенты вейвлетного спектра, соответствующие моменту времени t_0 будут содержаться в конусе влияния, ограниченном прямыми $a = t_0 - t\delta_t$ и $a = t_0 + t\delta_t$ (рис. 3, а). Коэффициент $W(a_0, b_0)$ в свою очередь зависит от значений ряда из определяемого тем же углом влияния временного диапазона около значения b_0 (рис. 3, б).

Диапазон этот тем больше, чем больше анализируемый масштаб a . Таким образом, высокочастотная

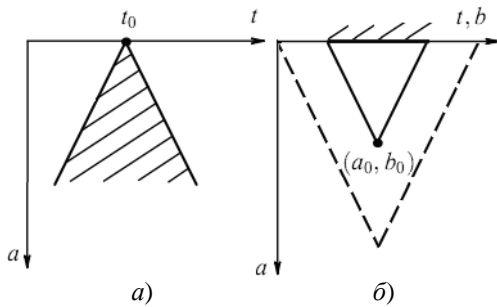


Рис. 3. Угол влияния вейвлет-преобразования: а – влияние точки анализируемой функции на значения вейвлет-коэффициентов; б – зависимость коэффициента $W(a_0, b_0)$ от значений анализируемой функции (пунктиром обозначена область достоверности)

(мелкомасштабная) информация вычисляется на основе малых по длительности отрезков сигнала, а низкочастотная – на основе больших отрезков сигнала.

Максимальный угол влияния отмечает область достоверности – область, вне которой коэффициенты $W(a, b)$ вычисляются уже по отрезкам, выходящим за границы ряда, и поэтому амплитуды вейвлет-коэффициентов носят искаженный характер. Заметим, что при нелинейном изменении масштабов на графиках вейвлет-преобразования угол влияния примет нелинейный вид.

Таким образом возникает граничный эффект – область, в которой коэффициенты вейвлет-преобразования искажаются разрывной природой границы сигнала. Пример граничного эффекта, возникающего при вейвлет-преобразовании, приведен на рис. 4.

Как видно из рис. 4, амплитуды вейвлет-преобразования в области краевых эффектов сравнимы с амплитудами коэффициентов, определяющих особую точку сигнала.

При выполнении вейвлет-преобразования конечной временной реализации длины N в частотной области ошибки в расчете коэффициентов $W(a, b)$ обусловлены природой преобразования Фурье, которое предполагает, что сигнал $s(t)$ периодичен и имеет период длиной N .

Для уменьшения ошибки граничных эффектов ряд дополняется с учетом его поведения (средним значением, известным временным ходом и т. д.) [3].

В данной работе для анализа вибрационных сигналов предлагается использовать дополнение временного ряда нулями до длины $2N$. Такой подход приводит к появлению сильной неоднородности на границах исходного временного ряда. Однако возмущения, вносимые этой неоднородностью, лежат в области очень больших временных масштабов, в то время как спектр исходной неоднородности (которая фор-

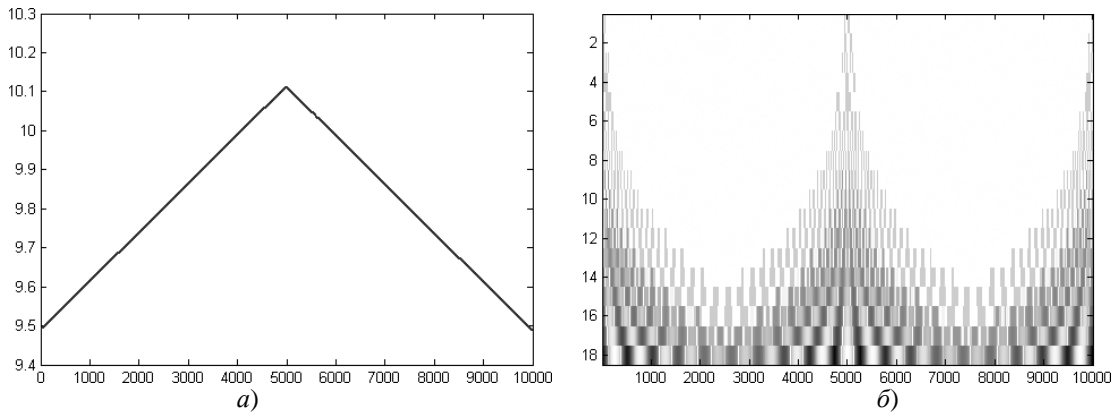


Рис. 4. Иллюстрация граничного эффекта, возникающего при вейвлет-анализе: а – пример анализируемого сигнала; б – вейвлет-преобразование сигнала

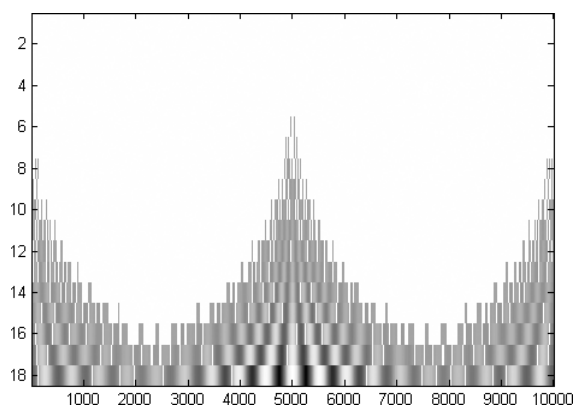


Рис. 5. Вейвлет-преобразование сигнала 4, a после дополнения нулями

мально существенно меньше внесенной) лежал бы в области масштабов анализируемого сигнала. Так что внесение такой неоднородности приводит к уменьшению амплитуды $W(a, b)$ около границ исходного ряда. Очевидно, чем больше будет длина дополнения, заполненного нулями, тем меньше влияние краевых эффектов. Использование для анализа ряда, в котором длина последовательности нулей равна длине исходного сигнала, является оптимальной в смысле баланса скорости вычислений, затрат машинной памяти и точности осуществления вейвлет-преобразования вблизи границ исходного временного ряда [8] (рис. 5).

На рис. 5 показано вейвлет-преобразование сигнала, приведенного на рис. 4, a , с использованием дополнения ряда нулями. Как видно, если сравнить рис. 4, b и рис. 5, влияние краевых эффектов значительно уменьшилось. Амплитуды в области краевых эффектов значительно ниже амплитуд коэффициентов, определяющих особую точку.

РАСЧЕТ ОБЛАСТИ ДОСТОВЕРНОСТИ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ, ВЫПОЛНЯЕМОГО ПРИ ПОМОЩИ ВЕЙВЛЕТА МОРЛЕ

Пусть вейвлет Морле $\psi(t)$ представляет собой функцию-окно с радиусом δ_t . Тогда каждый дочерний вейвлет на масштабе a будет иметь радиус окна, равный $\delta_t a$. Как видно из рис. 3, b , если точка $W(a, b)$ расположена от границы ряда на расстоянии меньшем, чем $\delta_t a$, то для ее расчета необходимы значения, выходящие за пределы ряда. Таким образом, количество точек от краев сигнала, которые носят искаженный характер, может быть задано следующим образом:

$$n = \delta_t a * f_d, \quad (7)$$

где f_d – частота дискретизации сигнала.

Так как вейвлет Морле является комплексной экспонентой, модулируемой функцией Гаусса (см. выр. 4), его ширина во временной области может быть определена при помощи правила трех сигм как:

$$\delta_t = \frac{3}{\sigma\sqrt{2}}. \quad (8)$$

Используя выражения (7)–(8), получим 2 области: область достоверности вейвлет-преобразования и область, в которой нельзя пренебречь наличием краевых эффектов. Предлагается отбрасывать данную область краевых эффектов при дальнейшей манипуляции с рассчитанными коэффициентами вейвлет-преобразования.

На рис. 6 приведен пример вейвлет-преобразования 4, b и 5 с рассчитанным максимальным углом влияния.

Заметим, что с уменьшением количества точек после отсекания области, содержащей краевые

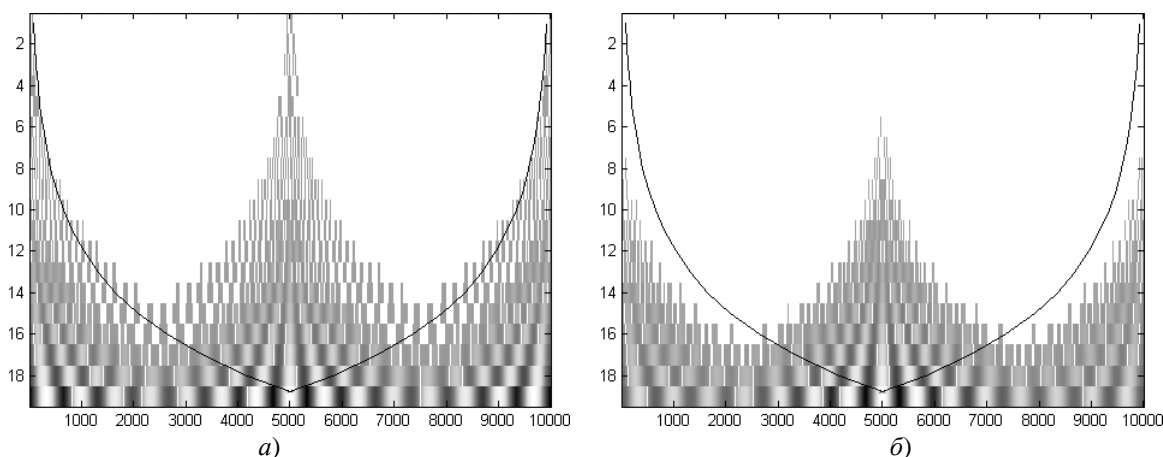


Рис. 6. Вейвлет-преобразование анализируемого сигнала (рис. 4, a) с обозначенной областью достоверности:

a – без дополнения сигнала нулями; b – с дополнением сигнала нулями

эффекты, падает разрешение по частоте для спектра вейвлет-преобразования:

$$\Delta f = \frac{1}{(N - 2^n)f_d}, \quad (9)$$

где N – количество отсчетов сигнала; f_d – частота дискретизации сигнала; n – длина области краевых эффектов.

МЕТОД ПОСТРОЕНИЯ АНАЛИЗИРУЮЩЕГО БАНКА ФИЛЬТРОВ

Для разделения различных частотных компонент сигналов используются банки фильтров. Как было отмечено, вейвлет-преобразование сигналов можно рассматривать как фильтрацию с постоянной относительной полосой частот. Поэтому по аналогии с обычными банками (гребенками) из $1/n$ -октавных полосовых фильтров может быть построен вейвлетный банк фильтров.

Для построения анализирующего банка фильтров на основе комплексного вейвлета Морле предлагается использовать следующую стратегию.

Добротность вейвлета Морле в качестве фильтра определяется как

$$Q_w = \frac{\omega}{f_0}. \quad (10)$$

При использовании набора масштабов a_i набор соответствующих центральных частот f_i и параметров ширины диапазона частот σ_i комплексного вейвлета Морле могут быть получены с помощью константы Q_w :

$$f_i = \frac{f_0}{a_i} \text{ и } \sigma_i = f_i Q_w. \quad (11)$$

Задавая исходные параметры вейвлета f_0 и σ , а также набор масштабов для расчета непрерывного вейвлет-преобразования, можно построить банк фильтров с необходимыми анализирующими параметрами.

Для f_0 максимальным приемлем значение $f_0 = 0,8 * f_{Nyq}$, где f_{Nyq} – частота Найквиста сигнала.

Такое значение не позволит первому фильтру в банке выйти за границу f_{Nyq} .

В качестве основания для расчета масштабов вейвлет-преобразования выберем $a_i = (2^{1/k})^i$, что позволит строить банк фильтров с k фильтрами на октаву.

Параметр добротности выберем таким образом, чтобы ширина диапазона каждого фильтра равнялась расстоянию между фильтрами

$$Q_w = 2^{1/k} - 1. \quad (12)$$

При помощи данного параметра можно варьировать ширину полосы пропускания фильтров и степень перекрытия фильтров в банке.

После определения обозначенных параметров при помощи выражения (5) может быть построен банк фильтров.

На рис. 7, а показан банк фильтров, полученный при помощи данной процедуры, а на рис. 7, б – банк фильтров с добротностью, увеличенной в два раза.

Используя выражение (7), производится вейвлет-фильтрация анализируемого сигнала.

Так как используемый вейвлет Морле является комплексным, то получаемые вейвлет-коэффициенты для каждого масштаба/фильтра также являются комплексными значениями. Реальная часть вейвлет-коэффициентов $W_{re}(a, b)$ представляет собой отфильтрованный сигнал, мнимая часть $W_{im}(a, b)$ является фазой отфильтрованного сигнала.

Для анализа особенностей сигнала также предлагается использовать спектр огибающей вейвлет-преобразования, который позволяет демодулировать сигнал в необходимом частотном диапазоне:

$$EW(a, b) = [W_{re}(a, b)^2 + W_{im}(a, b)^2]^{1/2}. \quad (13)$$

Правильная настройка банка фильтров и учет области достоверности вейвлет-преобразования позволяет проанализировать особенности сигнала наилучшим образом.

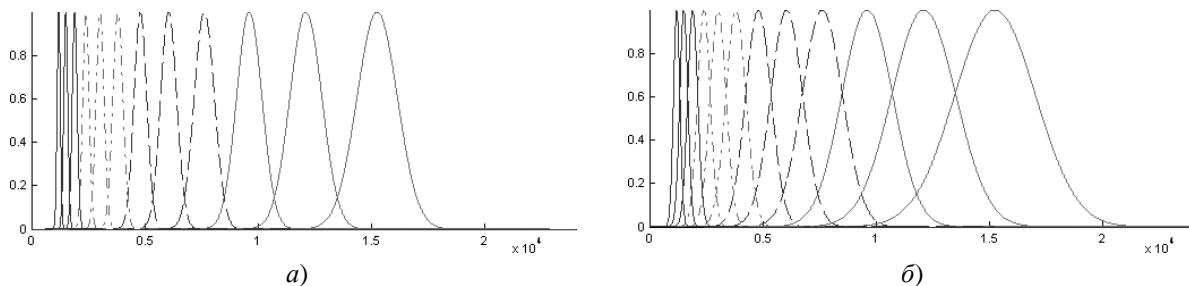


Рис. 7. Банк вейвлет-фильтров с тремя фильтрами на октаву, покрывающий 4 октавы: а – при добротности $Q_w = 2^{1/k} - 1$; б – при добротности, увеличенной в два раза: $Q_w = 2 \cdot (2^{1/k} - 1)$

В качестве признаков для выбора оптимального фильтра из построенного банка предлагается использовать различные статистические параметры. Например, эксцесс коэффициентов вейвлет-преобразования может быть использован для получения отфильтрованного сигнала с наибольшей импульсностью [9].

При настройке вейвлетного банка фильтров следует также учитывать требования к необходимой точности обработки сигнала. Достижение высокой точности и детальности описания характеристик сигнала приводит к увеличению длительности обработки сигнала. Поэтому одной из главных задач является рациональная настройка метода обработки сигнала, чтобы с приемлемой скоростью получить требуемое описание сигнала.

Для примера в табл. 1 приведено время обработки сигналов при помощи банков вейвлетных фильтров в зависимости от длины сигнала и характеристик банка фильтров. Приведенные значения получены для конкретного программно-аппаратного комплекса.

Таблица 1. Время обработки сигнала при помощи банка вейвлетных фильтров в зависимости от длины сигнала и характеристик банка фильтров

Длина сигнала, отсчеты	Время обработки, с		
	6 фильтров в банке	12 фильтров в банке	24 фильтра в банке
1024	0,0156	0,0313	0,0469
2048	0,0313	0,0469	0,0781
4096	0,0313	0,0625	0,1406
8192	0,0496	0,0938	0,2188
16384	0,0781	0,1719	0,4219
32768	0,1875	0,4219	1,0469

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ПРЕДЛОЖЕННОГО МЕТОДА К ЗАДАЧЕ ДИАГНОСТИКИ ПОДШИПНИКОВ КАЧЕНИЯ

В качестве примера для демонстрации фильтрующих возможностей вейвлет-преобразования были использованы вибрационные сигналы тестовой установки с дефектными подшипниками качения [10]. На установку помещали подшипник с точечным дефектом на внешней дорожке качения и подшипник с точечным дефектом на внутренней дорожке. Дефекты размером 0,007 дюйма нанесены при помощи эрозионного станка.

Скорость вращения вала установки примерно равна $f = 29$ Гц. Частота дискретизации сигнала 48 кГц. Подшипник имеет следующие параметры: количество шариков $z = 9$, угол контакта шарика с дорожка-

ми $\varphi = 0^\circ$, диаметр шарика $d = 7,95$ мм, диаметр сепаратора $D = 38,5$ мм.

Таким образом, может быть рассчитана характеристическая частота перекачивания тел качения по внешней дорожке:

$$F_{\text{ВРФО}} = 0,5zf\left(1 - \frac{d}{D}\cos\varphi\right) = 104 \text{ Гц}$$

и частота перекачивания тел качения по внешней дорожке:

$$F_{\text{ВРФИ}} = 0,5zf\left(1 + \frac{d}{D}\cos\varphi\right) = 158 \text{ Гц.}$$

На рис. 8 приведены результаты обработки сигнала с подшипником с дефектом на внешней дорожке качения, на рис. 9 – с дефектом на внутренней дорожке качения.

Масштабу 7 соответствует центральная частота фильтра 3809,8 Гц с шириной полосы пропускания 990,23 Гц, а масштабу 8 – 3023,8 Гц и 785,95 Гц соответственно.

Рис. 8 и 9 иллюстрируют эффективность фильтрации нестационарных сигналов при помощи комплексного вейвлета Морле.

ВЫВОДЫ

В работе решена актуальная задача исследования фильтрующих свойств непрерывного вейвлет-преобразования. Приведен метод уменьшения влияния краевых эффектов на результаты вейвлет анализа. Выполнено обоснование выбора типа вейвлета для анализа нестационарных сигналов. Предложен метод настройки материнского вейвлета Морле и построения банка фильтров с необходимыми анализирующими свойствами. Предложенный метод эффективно применен к задаче диагностирования подшипников качения по вибрационному сигналу.

В качестве направления дальнейшего исследования может быть выделена задача построения вектора характеристических признаков анализируемого сигнала из получаемых вейвлет-коэффициентов с целью постановки диагноза и принятия решения о состоянии исследуемого объекта.

СПИСОК ЛИТЕРАТУРЫ

1. Чуи К. Введение в вейвлеты / К. Чуи. – М. : Мир, 2001. – 412 с.
2. Малла С. Вэйвлеты в обработке сигналов / С. Малла. – М. : Мир, 2005. – 671 с.
3. Addison P. S. Illustrated wavelet transform handbook. Introductory Theory and Applications in Science, Engineering, Medicine and Finance / Paul S. Addison. – Bristol : Institute of Physics Publishing, 2002. – 400 p.
4. Дьяконов В. П. Вейвлеты. От теории к практике / В. П. Дьяконов. – Изд. 2-е, перераб. и доп. – М. : Солон-Пресс, 2004. – 400 с.

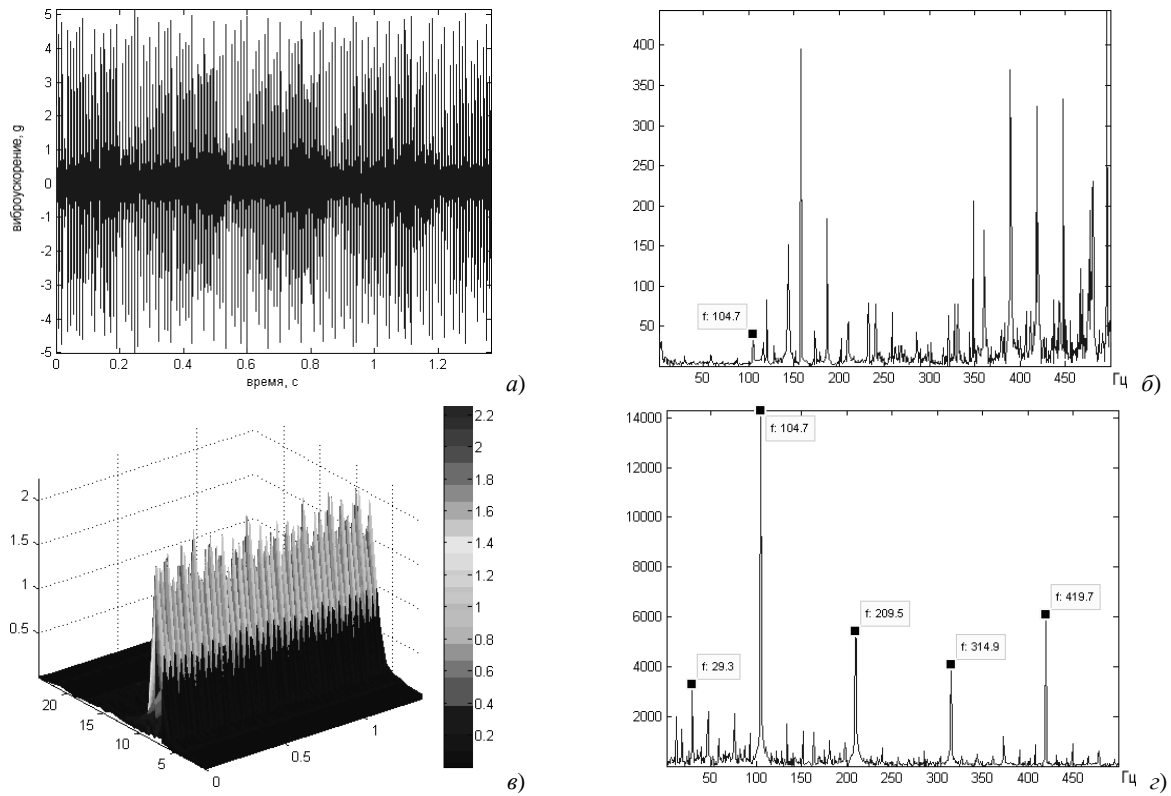


Рис. 8. Обработка сигнала с дефектом на внешней дорожке качения:
a – исходный сигнал; *b* – спектр необработанного сигнала; *в* – вейвлет-преобразование сигнала;
г – спектр огибающей вейвлет-коэффициентов на масштабе номер 8

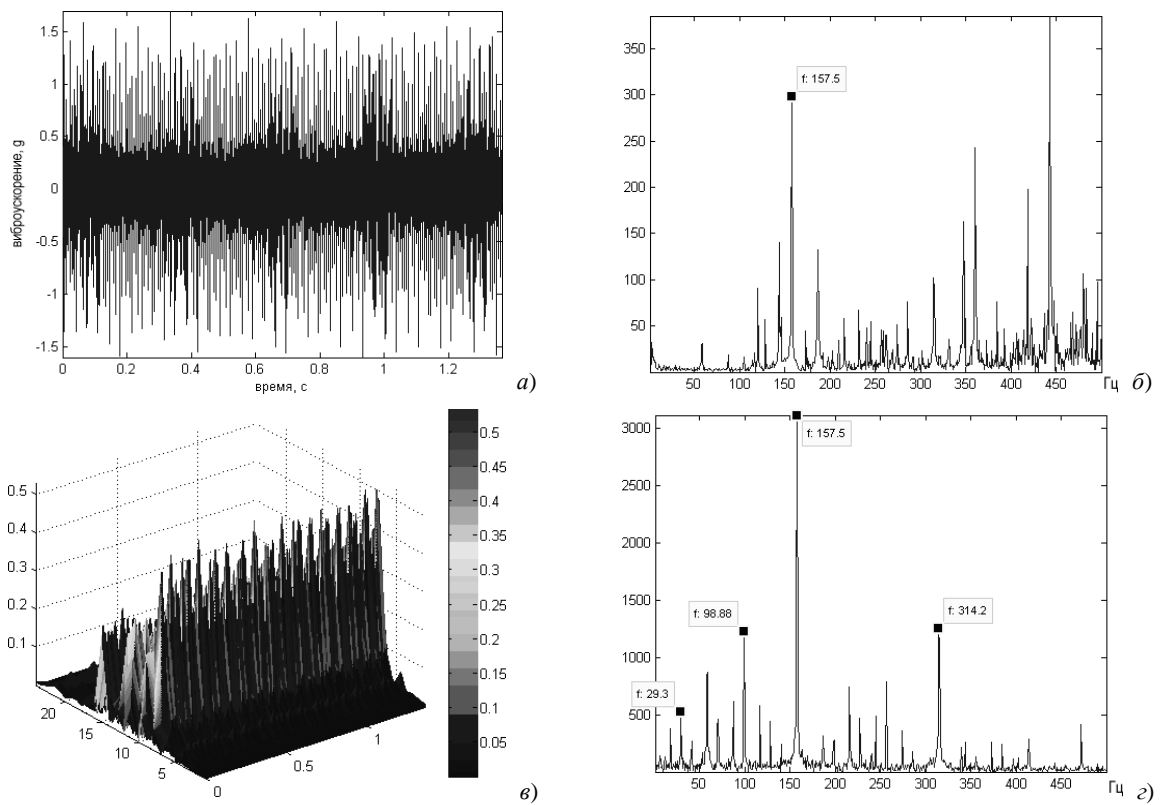


Рис. 9. Обработка сигнала с дефектом на внутренней дорожке качения:
a – исходный сигнал; *b* – спектр необработанного сигнала; *в* – вейвлет-преобразование сигнала;
г – спектр огибающей вейвлет-коэффициентов на масштабе номер 8

5. *Torrence C.* A Practical Guide to Wavelet Analysis / C. Torrence, G. P. Compo // Bulletin of the American Meteorological Society. – 1998. – Vol. 79. – Pp. 61–78.
6. *Najmi A.* The Continuous Wavelet Transform and Variable Resolution Time–Frequency Analysis / A. Najmi, J. Sadowsky. // JOHNS HOPKINS APL TECHNICAL DIGEST. – 1997. – Vol. 18, No. 1. – Pp.134–140.
7. *Астафьева Н.* Вейвлет анализ: основы теории и примеры применения / Н. Астафьева // Успехи физических наук. – 1996. – Т. 166, № 11. – С. 1145–1170.
8. *Короновский А. А.* Непрерывный вейвлетный анализ и его приложения / А. А. Короновский, А. Е. Храмов. – М. : ФИЗМАТЛИТ, 2003. – 176 с.
9. *Shir L.* Rolling element bearing fault detection using improved envelope analysis / L. Shir, B. Randall, J. Antoni // IMECHE conference transactions. – 2004. – Vol 2. – P. 301–312.
10. Case Western Reserve University Bearing Data Center. Seeded Fault Test Data [Электронный ресурс]. – Режим доступа: http://www.eecs.case.edu/laboratory/bearing/welcome_overview.htm.

Надійшла 19.04.2010

Юр Т. В., Дубровін В. І., Харитонов В. Н.

ДОСЛІДЖЕННЯ ФІЛЬТРУЮЧИХ ВЛАСТИВОСТЕЙ ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ

В роботі досліджено фільтруючі властивості безперервного вейвлет-перетворення. Наведено метод зменшення

впливу крайових ефектів на результати вейвлет-аналізу. Виконано обґрунтування вибору типу вейвлету для аналізу нестационарних сигналів. Запропоновано метод побудови банку вейвлетних фільтрів з необхідними аналізуючими властивостями. Розроблений метод застосовано до задачі діагностування підшипників кочення за вібраційним сигналом.

Ключові слова: нестационарний сигнал, вейвлет-фільтрація, банк фільтрів, вейвлет Морле.

Yur T. V., Dubrovin V. I., Kharitonov V. N.

ANALYSIS OF WAVELET TRANSFORM FILTERING PROPERTIES

Filtering properties of continuous wavelet transform are investigated in the paper. The method that decreases the influence of edge effects on wavelet analysis results is described. The choice of wavelet for non-stationary signals analysis is grounded. The method of construction of a filter bank with required analyzing properties is proposed. The given method is applied to the problem of rolling bearing diagnostics by vibration signals.

Key words: non-stationary signal, wavelet filtering, filter bank, Morlet wavelet.

УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

CONTROL IN TECHNICAL SYSTEMS

УДК 62-55:681.515

Гостев В. И.

Д-р техн. наук, заведующий кафедрой Государственного университета информационно-коммуникационных технологий (г. Киев)

ПРОЕКТИРОВАНИЕ НЕЧЕТКОГО РЕГУЛЯТОРА С УНИВЕРСАЛЬНЫМИ ИДЕНТИЧНЫМИ ВХОДНЫМИ И ВЫХОДНЫМИ ФУНКЦИЯМИ ПРИНАДЛЕЖНОСТИ

Получены аналитические выражения для управляющих воздействий на выходе нечеткого регулятора при универсальных функциях принадлежности и изложены вопросы проектирования нечеткого регулятора.

Ключевые слова: автоматическое управление, нечеткий регулятор, проектирование, нечеткая логика, функции принадлежности, Matlab.

ВВЕДЕНИЕ

В работе [1] изложен новый метод проектирования нечетких регуляторов. Актуальной является задача проектирования нечетких регуляторов с различными типовыми функциями принадлежности на основе изложенного метода. В данной работе проектируется нечеткий регулятор с универсальными идентичными входными и выходными функциями принадлежности, структурная схема которого представлена на рис. 1 и подробно описана в работах [1–3]. Нечеткий регулятор состоит из трех блоков: блока *формирователя величин* $A(t)$ и $B(t)$ (блок 1), блока *сравнения величин* $A(t)$ и $B(t)$ и *расчета* u_c (блок 2) и блока *нормировки выходной переменной* (блок 3).

РЕШЕНИЕ ЗАДАЧИ

Рассмотрим на универсальном множестве $U = [0, 1]$ два нечетких подмножества с функциями принадлежности вида (см. рис. 2)

$$\mu_1(u) = \begin{cases} \frac{(1-a-u)^c}{(1-a)^c}, & 0 \leq u \leq 1-a; \\ 0, & 1-a \leq u \leq 1, \end{cases}$$

$$\mu_2(u) = \begin{cases} 0, & 0 \leq u \leq a; \\ \frac{(u-a)^c}{(1-a)^c}, & a \leq u \leq 1. \end{cases} \quad (1)$$

Функции принадлежности назовем универсальными, если из них можно получить не менее трех элементарных функций. Записанные выше функции

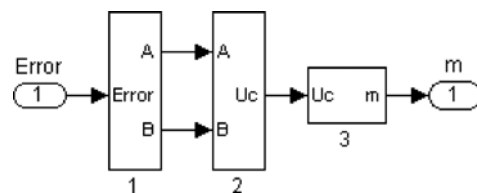


Рис. 1.

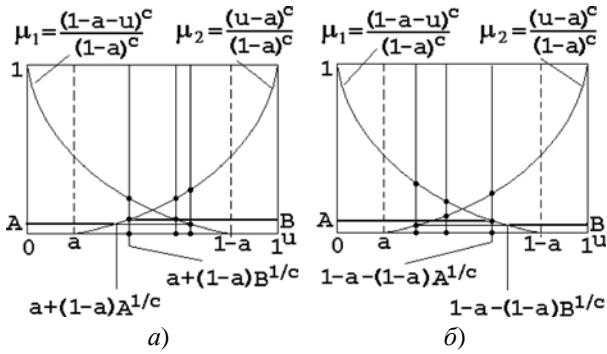


Рис. 2.

принадлежности отвечают такому условию. Действительно, из формул (1) при $a = 0$ получаем формулы для возведенных в степень треугольных функций принадлежности (рис. 3.41 [2]), при $a = 0, c = 1$ получаем формулы для треугольных функций принадлежности (рис. 3.42 [2]), при $c = 1$ получаем формулы для треугольных функций принадлежности с увеличенным наклоном (рис. 3.47 [2]).

На входе нечеткого регулятора в блоке *формирователя величин A(t) и B(t)* переменные ошибка системы θ^* , первая производная ошибки $\dot{\theta}^*$ и вторая производная ошибки $\ddot{\theta}^*$, поступающие в регулятор с шагом квантования h , пересчитываются в переменные u_1^*, u_2^*, u_3^* по формулам

$$\left. \begin{aligned} u_1^* &= (\theta^* + A_m) / (2A_m); \\ u_2^* &= (\dot{\theta}^* + B_m) / (2B_m); \\ u_3^* &= (\ddot{\theta}^* + C_m) / (2C_m) \end{aligned} \right\} \quad (2)$$

(это формулы для пересчета значений сигналов в значения элементов единого универсального множе-

ства $U = [0, 1]$ при симметричных диапазонах изменения переменных $A_m = \theta_{\max} = -\theta_{\min}, B_m = \dot{\theta}_{\max} = -\dot{\theta}_{\min}, C_m = \ddot{\theta}_{\max} = -\ddot{\theta}_{\min}$) и рассчитываются значения функций принадлежности для переменных u_1^*, u_2^*, u_3^* . Точками на универсальном множестве отмечены возможные для какого-то момента времени значения переменных u_1^*, u_2^*, u_3^* .

По алгоритму Мамдани определяются величины A и B по формулам

$$\left. \begin{aligned} A &= \min[\mu_1(u_1^*), \mu_1(u_2^*), \mu_1(u_3^*)]; \\ B &= \min[\mu_2(u_1^*), \mu_2(u_2^*), \mu_2(u_3^*)]. \end{aligned} \right\} \quad (3)$$

На основании формул (1)–(3) построен блок *формирователя величин A(t) и B(t)* (блок 1 на рис. 1). Этот блок показан на рис. 3.

В схеме *формирователя величин A(t) и B(t)* при настройке нечеткого регулятора перестраиваются граничные значения диапазонов $A_m = \theta_{\max} = -\theta_{\min}, B_m = \dot{\theta}_{\max} = -\dot{\theta}_{\min}, C_m = \ddot{\theta}_{\max} = -\ddot{\theta}_{\min}$. Коэффициенты a и c при настройке диапазонов обычно задаются постоянными. При изменении значений коэффициентов граничные значения диапазонов требуется подстраивать.

Отметим весьма существенный факт. Какие бы значения ни принимали переменные u_1^*, u_2^*, u_3^* на универсальном множестве $U = [0, 1]$, в зависимости от соотношений величин A и B «результующая фигура» может принимать только две конфигурации: при $A \leq B$ первая конфигурация показана на рис. 2, *a*; при $A \geq B$ вторая конфигурация показана на рис. 2, *б*.

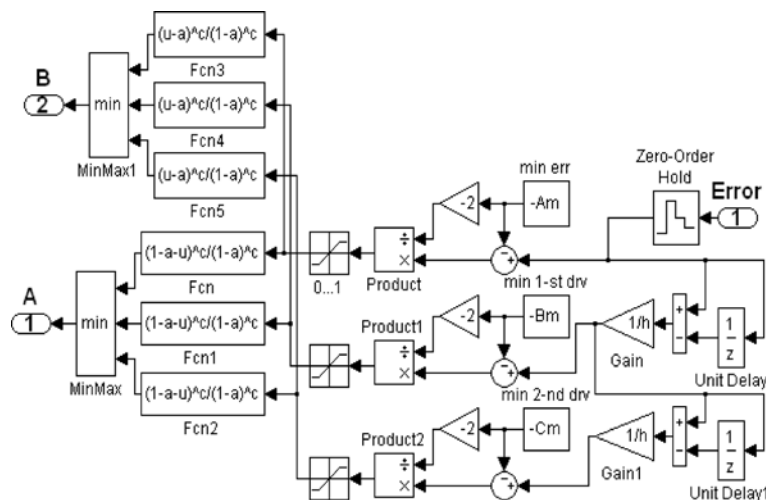


Рис. 3.

Абсцисса «центра тяжести результирующей фигуры» при $A \leq B$ определяется по формуле

$$u_c = \frac{A \int_0^{a+(1-a)A^{1/c}} u du + \frac{1}{(1-a)^c} \int_{a+(1-a)A^{1/c}}^{a+(1-a)B^{1/c}} (u-a)^c u du + B \int_{a+(1-a)B^{1/c}}^1 u du}{A \int_0^{a+(1-a)A^{1/c}} du + \frac{1}{(1-a)^c} \int_{a+(1-a)A^{1/c}}^{a+(1-a)B^{1/c}} (u-a)^c du + B \int_{a+(1-a)B^{1/c}}^1 du} \quad (4)$$

После несложных вычислений находим

$$u_c = \frac{\frac{B}{2} + \frac{a^2}{2}(A-B) + \frac{a(1-a)c}{c+1} \left(A^{\frac{c+1}{c}} - B^{\frac{c+1}{c}} \right) + \frac{(1-a)^2 c}{2(c+2)} \left(A^{\frac{c+2}{c}} - B^{\frac{c+2}{c}} \right)}{B + a(A-B) + \frac{(1-a)c}{c+1} \left(A^{\frac{c+1}{c}} - B^{\frac{c+1}{c}} \right)} \quad (5)$$

Абсцисса «центра тяжести результирующей фигуры» при $A \geq B$ определяется по формуле

$$u_c = \frac{A \int_0^{1-a-(1-a)A^{1/c}} u du + \frac{1}{(1-a)^c} \int_{1-a-(1-a)A^{1/c}}^{1-a-(1-a)B^{1/c}} (1-a-u)^c u du + B \int_{1-a-(1-a)B^{1/c}}^1 u du}{A \int_0^{1-a-(1-a)A^{1/c}} du + \frac{1}{(1-a)^c} \int_{1-a-(1-a)A^{1/c}}^{1-a-(1-a)B^{1/c}} (1-a-u)^c du + B \int_{1-a-(1-a)B^{1/c}}^1 du} \quad (6)$$

После несложных вычислений находим

$$u_c = \frac{\frac{B}{2} + \frac{(1-a)^2}{2}(A-B) + \frac{(1-a)c}{c+1} \left(A^{\frac{c+1}{c}} - B^{\frac{c+1}{c}} \right) + \frac{(1-a)^2 c}{2(c+2)} \left(A^{\frac{c+2}{c}} - B^{\frac{c+2}{c}} \right)}{B + (1-a)(A-B) + \frac{(1-a)c}{c+1} \left(A^{\frac{c+1}{c}} - B^{\frac{c+1}{c}} \right)} \quad (7)$$

На основании формул (5) и (7) построен блок сравнения величин $A(t)$ и $B(t)$ и расчета u_c (блок 2 на рис. 1). Этот блок показан на рис. 4.

В качестве примера приведем следующие результаты расчетов, которые легко проверяются путем моделирования блока сравнения величин $A(t)$ и $B(t)$ и расчета u_c (см. рис. 4):

– идентичные возведенные в степень треугольные функции принадлежности (рис. 3.41 [2])

$$\begin{cases} a = 0, c = 3, \\ A = 0,01; B = 0,3 \Rightarrow u_c = 0,7272; \\ A = 0,3; B = 0,01 \Rightarrow u_c = 0,2728; \end{cases}$$

– идентичные треугольные функции принадлежности (рис. 3.42 [2])

$$\begin{cases} a = 0, c = 1, \\ A = 0,2; B = 0,4 \Rightarrow u_c = 0,5608; \\ A = 0,4; B = 0,2 \Rightarrow u_c = 0,4392; \end{cases}$$

– идентичные треугольные функции принадлежности с увеличенным наклоном (рис. 3.47 [2])

$$a = 0,25, c = 1,$$

$$\begin{cases} A = 0,1; B = 0,3 \Rightarrow u_c = 0,6082; \\ A = 0,3; B = 0,1 \Rightarrow u_c = 0,3918; \end{cases}$$

– идентичные универсальные функции принадлежности (рис. 2)

$$\begin{cases} a = 0,2, c = 3, \\ A = 0,01; B = 0,3 \Rightarrow u_c = 0,7707; \\ A = 0,3; B = 0,01 \Rightarrow u_c = 0,2293. \end{cases}$$

В блоке сравнения величин $A(t)$ и $B(t)$ и расчета u_c производится расчет ненормированного выхода регулятора. Далее полученное значение u_c в блоке нормировки выходной переменной пересчитывается в выходное напряжение регулятора по формуле

$$m^* = m_{\min}(1 - 2u_c). \quad (8)$$

Блок нормировки выходной переменной, построенный по формуле (8), показан на рис. 5.

В блоке нормировки выходной переменной при настройке нечеткого регулятора перестраиваются граничные значения диапазона $D_m = m_{\max} = -m_{\min}$.

В динамике при изменении переменных $u_1(t)$, $u_2(t)$, $u_3(t)$ текущие величины $A(t)$ и $B(t)$ поступают

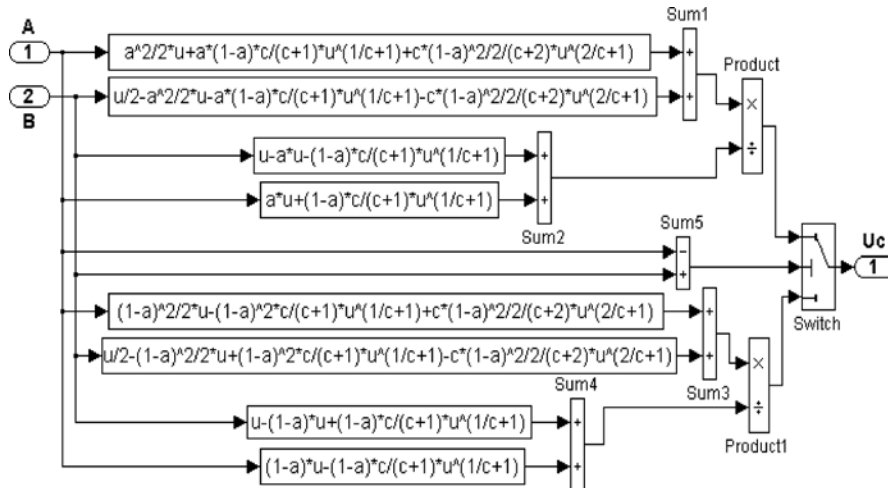


Рис. 4.

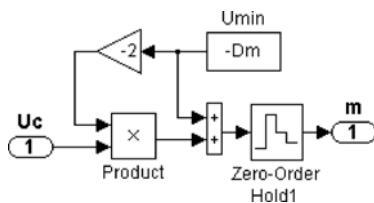


Рис. 5.

в блок сравнения, и рассчитанное на каждом шаге h значение $u_c(t)$ в блоке нормировки выходной переменной преобразуется в выходное напряжение $m(t)$ по формуле $m(t) = m_{\min}[1 - 2u_c(t)]$.

ЗАКЛЮЧЕНИЕ

Изложенные теория и практическая схема нечеткого регулятора с универсальными идентичными входными и выходными функциями принадлежности дают возможность использовать такой регулятор в различных системах автоматического управления и путем настройки параметров регулятора добиваться высокого качества систем управления.

СПИСОК ЛИТЕРАТУРЫ

1. Гостев В. И. Новый метод проектирования одного класса нечетких цифровых регуляторов / Гостев В. И. //

Проблемы управления и информатики. – 2007. – № 6. – С. 73–84.

2. Гостев В. И. Нечеткие регуляторы в системах автоматического управления / Гостев В. И. – К. : Радиоаматор, 2008. – 972 с.
3. Гостев В. И. Фаззи-системы автоматического управления параметрами промышленных объектов: монография / Гостев В. И., Скуртов С. Н. – Нежин : Аспект-Поліграф, 2009. – 432 с.

Надійшла 16.12.2009

Гостев В. И.

ПРОЕКТУВАННЯ НЕЧІТКОГО РЕГУЛЯТОРА З УНІВЕРСАЛЬНИМИ ІДЕНТИЧНИМИ ВХІДНИМИ І ВИХІДНИМИ ФУНКЦІЯМИ ПРИНАЛЕЖНОСТІ

Отримано аналітичні вирази для керуючих впливів на виході нечіткого регулятора при універсальних функціях належності та викладено питання проектування нечіткого регулятора.

Ключові слова: автоматичне керування, нечіткий регулятор, проектування, нечітка логіка, функції належності, Matlab.

Gostev V. I.

DESIGNING OF A FUZZY CONTROLLER WITH UNIVERSAL IDENTICAL INPUT AND OUTPUT MEMBERSHIP FUNCTIONS

Analytical expressions are derived for control actions at the output of a fuzzy controller at universal membership functions. Principles of fuzzy controller designing are stated.

Key words: automatic control, fuzzy controller, design, fuzzy logic, membership functions, Matlab.

ДО ВІДОМА АВТОРІВ

Журнал «Радіоелектроніка, інформатика, управління» (PIU) призначений для публікації найбільш значимих наукових і практичних результатів досліджень учених вищих навчальних закладів і наукових організацій.

Журнал включений у перелік наукових видань України, у яких можуть публікуватися результати дисертаційних робіт на змагання вчених ступенів доктора і кандидата технічних наук і фізико-математичних наук (фізика).

Статті, що опубліковано в журналі, реферуються в реферативних журналах і базах даних ВІНІТІ (Росія) і «Джерело» (Україна). Інтернет-сторінка журналу:

<http://journal.zntu.edu.ua/tic/index.php?page=index>.

Журнал видається два рази в рік і розповсюджується по підписці в Україні. Передплатний індекс 22914.

Для розгляду питання про публікацію статті в редакцію журналу необхідно вислати поштою або представити особисто наступне:

1) рукопис (роздруківку) статті, підписаний на останній сторінці всіма авторами, в двох екземплярах;

2) оригінал експертного висновку про можливість відкритого опублікування статті;

3) супровідний лист-клопотання з організації, де була виконана робота (або лист автора);

4) рецензію від фахівця в даній області з вченим ступенем доктора наук. Підпис рецензента обов'язково мусить бути завіреним.

5) диск з наступними файлами:

1) електронна версія статті, повністю ідентична роздруківці;

2) відомості про авторів;

3) рисунки у графічному форматі .tif.

Файли з матеріалами статті можна вислати електронною поштою або передати особисто на оптичному диску або USB-накопичувачі.

Вимоги до оформлення статті. Приймаються статті, набрані в редакторі Microsoft Word.

Параметри сторінки:

– розмір паперу – А4 (210x297);

– орієнтація – книжкова;

– шрифт – Times New Roman, розмір – 12 pt;

– міжрядковий інтервал – полуторний;

– верхнє поле – 20 мм, нижнє – 20 мм, ліве – 25 мм, праве – 15 мм.

Сторінки рукопису повинні бути пронумеровані. Не допускаються розбіжності рукопису з електронною версією статті. Текст рукопису не повинен мати рукописних виправлень та позначок.

Послідовність розміщення матеріалу статті:

1) індекс УДК;

2) прізвища й ініціали авторів, назва статті, анотація й ключові слова українською мовою (для громадян України);

3) прізвища й ініціали авторів, назва статті, анотація й ключові слова російською мовою;

4) прізвища й ініціали авторів, назва статті, анотація й ключові слова англійською мовою;

5) текст статті;

6) список літератури.

Текст статті. Приймаються статті російською, українською та англійською мовами. Розмір статті до 0,5 авторського аркуша. У статті слід уникати зайвої деталізації, про-

міжних формул і висновків; не слід наводити відомі факти, повторювати зміст таблиць і ілюстрацій у тексті. Стаття не повинна мати граматичних або інших помилок, а також повинна відповідати тематиці журналу й вимогам ВАК щодо фахових видань. **Структура** тексту статті мусить містити такі необхідні елементи: постановка проблеми в загальному виді і її зв'язок з важливими науковими або практичними завданнями; аналіз останніх досліджень і публікацій, у яких розпочато розв'язання даної проблеми, і на які опирається автор; виділення нерозв'язаних раніше частин загальної проблеми, яким присвячується стаття; формулювання цілей статті (постановка завдання); виклад основного матеріалу дослідження з повним обґрунтуванням отриманих наукових результатів, висновки по даному дослідженню й перспективи подальших досліджень у даному напрямку. Матеріал публікації мусить бути розбитий на підрозділи не більше двох рівнів.

Рисунки розміщуються в тексті й додатково додаються в окремих файлах (формат .tif з роздільною здатністю 150–300 dpi, чорно-білі або у градаціях сірого). Розмір рисунків не повинен перевищувати ширини сторінки (17 см) або ширини колонки (8 см). Написи на рисунках бажано виконувати шрифтом Times New Roman, розмір 10. Рисунки нумерують і підписують унизу.

Формули виконуються за допомогою вбудованого в Word редактора Microsoft Equation. Формули нумерують у круглих дужках праворуч. Формули великого розміру записуються в кілька рядків.

Нумерація рисунків, формул і таблиць наскрізна однакова.

Обсяг **анотації** не повинен перевищувати 40 слів.

Ключові слова наводяться в називному відмінку у кількості до десяти слів.

Список літератури наприкінці статті подається мовою оригіналу і складається в порядку згадування посилань у тексті й відповідно до діючого стандарту на бібліографічний опис. Посилання на літературу в тексті нумеруються послідовно й позначаються цифрою у квадратних дужках.

У **відомостях про авторів** необхідно навести:

1) прізвище, ім'я, по батькові (повністю);

2) учений ступінь;

3) посаду;

4) місце роботи;

5) робочу чи домашню і електронну адресу;

6) робочий, домашній, мобільний телефон.

Статті, які не відповідають зазначеним вимогам, не приймаються до розгляду.

Всі статті проходять закриті рецензування і в разі потреби можуть бути повернуті автору на доробку. Редакція залишає за собою право на літературну редакцію тексту статті без повідомлення автору.

Рукописи й диски не вертаються, коректура та відбитки статей авторам не надсилаються.

Адреса редакції: 69063, м. Запоріжжя, вул. Жуковського, 64, ЗНТУ, редакція журналу «PIU»

Тел.: (061) 7-698-2-96 – редакційно-видавничий відділ

(061) 7-644-6-62 – головний редактор

Факс: (061) 7-642-1-41

E-mail: rvv@zntu.edu.ua

Наукове видання

**Радіоелектроніка,
інформатика,
управління**

№ 2/2010

Науковий журнал

Головний редактор – д-р техн. наук Піза Д. М.

Заст. головного редактора – канд. техн. наук Дубровін В. І.

Комп'ютерне моделювання та верстання

Новікова К. О.

Оригінал-макет підготовлено у редакційно-видавничому відділі ЗНТУ

*Підписано до друку 13.07.2010. Формат 60x84/8.
Папір офс. Різогр. друк. Ум. друк. арк. 19,9.
Тираж 300 прим. Зам. № 1390.*

69063 м. Запоріжжя, ЗНТУ, друкарня, вул. Жуковського, 64