

p-ISSN 1607-3274  
e-ISSN 2313-688X

# Радіоелектроніка Інформатика Управління

Радиоэлектроника  
Информатика  
Управление

Radio Electronics  
Computer Science  
Control

2014/2

ISSN 1607-3274



9 771607 327005 44>



Запорізький національний технічний університет

## Радіоелектроніка, інформатика, управління

Науковий журнал

Виходить двічі на рік

№ 2(31) 2014

Заснований у січні 1999 року.

Засновник і видавець – Запорізький національний технічний університет.

ISSN 1607-3274 (друкований), ISSN 2313-688X (електронний).

Запоріжжя

ЗНТУ

2014

Запорожский национальный технический университет

## Радиоэлектроника, информатика, управление

Научный журнал

Выходит дважды в год

№ 2(31) 2014

Основан в январе 1999 года.

Основатель и издатель – Запорожский национальный технический университет.

ISSN 1607-3274 (печатный), ISSN 2313-688X (электронный).

Запорожье

ЗНТУ

2014

Zaporizhzhya National Technical University

## Radio Electronics, Computer Science, Control

The scientific journal

Published twice per year

№ 2(31) 2014

Founded in January 1999.

Founder and publisher – Zaporizhzhya National Technical University.

ISSN 1607-3274 (print), ISSN 2313-688X (on-line).

Zaporizhzhya

ZNTU

2014

**Науковий журнал «Радіоелектроніка, інформатика, управління»** (скорочена назва – РІУ) видається Запорізьким національним технічним університетом (ЗНТУ) з 1999 р. періодичністю два номери на рік.

Зареєстрований Державним комітетом інформаційної політики, телебачення та радіомовлення 29.01.2003 р. Свідоцтво про державну реєстрацію друкованого засобу масової інформації серія КВ №6904.

**ISSN** 1607-3274 (друкований), **ISSN** 2313-688X (електронний).

Постановою президії ВАК України № 1-05/4 від 26.05.2010 р. журнал **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 (радіофізики).

В журналі безкоштовно публікуються наукові статті англійською, російською та українською мовами.

**Правила оформлення статей** подано на сайті: <http://ric.zntu.edu.ua/information/authors>.

Журнал забезпечує **безкоштовний відкритий он-лайн доступ** до повнотекстових публікацій.

Журнал дозволяє авторам мати авторські права і зберігати права на видання без обмежень. Журнал дозволяє користувачам читати, завантажувати, копіювати, поширювати, друкувати, шукати або посилатися на повні тексти своїх статей. Журнал дозволяє повторне використання його вмісту у відповідності з СС ліцензією CC-BY.

Опублікованими статтями присвоюється унікальний ідентифікатор цифрового об'єкта DOI.

**Журнал реферується та індексується** у провідних міжнародних та національних реферативних журналах і наукометричних базах даних, а також розміщується у цифрових архівах та бібліотеках з безкоштовним доступом у режимі on-line (у т. ч. DOAJ, DOI, CrossRef, EBSCO, eLibrary.ru / РИНЦ, Google Scholar, Index Copernicus, INSPEC, ISSN, Ulrich's Periodicals Directory, WorldCat, ВІНІТІ, Джерело), повний перелік яких подано на сайті: <http://ric.zntu.edu.ua/about/editorialPolicies#custom-0>.

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

**Тематика журналу** містить: радіофізику, мікро-, нано- і радіоелектроніку, апаратне і програмне забезпечення комп’ютерної техніки, комп’ютерні мережі і телекомунікації, теорію алгоритмів і програмування, оптимізацію і дослідження операцій, міжмашинну і людино-машинну взаємодію, математичне і комп’ютерне моделювання, обробку даних і сигналів, управління в технічних системах, штучний інтелект, включаючи системи, засновані на знаннях, і експертні системи, інтелектуальний аналіз даних, розпізнавання образів, штучні нейронні і нейро-нечіткі мережі, нечітку логіку, колективний інтелект і мультиагентні системи, гібридні системи.

Усі статті, пропоновані до публікації, одержують **об’ективний розгляд**, що оцінюється за суттю без урахування раси, статі, віросповідання, етнічного походження, громадянства або політичної філософії автора(ів).

Усі статті проходять двоступінчасте закрите (анонімне для автора) **резензування** штатними редакторами і незалежними рецензентами – провідними вченими за профілем журналу.

## РЕДАКЦІЙНА КОЛЕГІЯ

**Головний редактор** – Погосов В. В., д-р фіз.-мат. наук, Україна

**Заст. головного редактора** – Субботін С. О., д-р. техн. наук, Україна

**Члени редколегії:**

Безрук В. М., д-р техн. наук, Україна  
Бодянський Є. В., д-р техн. наук, Україна, редактор розділу з управління  
Васильєв С. М., д-р фіз.-мат. наук, академік РАН, Росія  
Гімплевич Ю. Б., д-р техн. наук, Україна  
Горбань О. М., д-р фіз.-мат. наук, Великобританія  
Дробахін О. О., д-р фіз.-мат. наук, Україна  
Зайцева О. М., канд. фіз.-мат. наук, Словаччина  
Карпуков Л. М., д-р техн. наук, Україна  
Корніч Г. В., д-р фіз.-мат. наук, Україна, редактор розділу з радіофізики  
Кулік А. С., д-р техн. наук, Україна  
Лебедев Д. В., д-р техн. наук, Україна, редактор розділу з управління  
Левашенко В. Г., канд. фіз.-мат. наук, Словаччина  
Лиснянський А., канд. техн. наук, Ізраїль  
Марковська-Качмар У., д-р наук, Польща  
Олещук В. О., канд. фіз.-мат. наук, Норвегія, редактор розділу з радіоелектроніки  
Онуфрієнко В. М., д-р фіз.-мат. наук, Україна  
Папшицький М., д-р філософії, Польща  
Піза Д. М., д-р техн. наук, Україна  
**[Потапенко Є. М.]** д-р техн. наук, Україна  
Рубель О. І., канд. техн. наук, Канада  
Хаханов В. І., д-р техн. наук, Україна, редактор розділу з інформатики  
Чумаченко В. П., д-р фіз.-мат. наук, Україна, редактор розділу з радіоелектроніки  
Шарпанських О. А., д-р філософії, Нідерланди, редактор розділу з інформатики

Рекомендовано до видання вченого радою ЗНТУ, протокол № 6 від 22.12.2014 р.

Журнал зверстаний редакційно-видавничим відділом ЗНТУ.

**Веб-сайт журналу:** <http://ric.zntu.edu.ua>.

**Адреса редакції:** Редакція журналу «РІУ», Запорізький національний технічний університет, вул. Жуковського, 64, м. Запоріжжя, 69063, Україна.

Тел: (061) 769-82-96 – редакційно-видавничий відділ

Факс: (061) 764-46-62

E-mail: [rvv@zntu.edu.ua](mailto:rvv@zntu.edu.ua)

**Научный журнал «Радиоэлектроника, информатика, управление»** (сокращенное название – РИУ) издается Запорожским национальным техническим университетом (ЗНТУ) с 1999 г. периодичностью два номера в год.

Зарегистрирован Государственным комитетом информационной политики, телевидения и радиовещания 29.01.2003 г. (Свидетельство о государственной регистрации печатного средства массовой информации серия КВ №6904).

**ISSN 1607-3274** (печатный), **ISSN 2313-688X** (электронный).

Постановлением ВАК Украины №1-05/4 от 26.05.2010 г. журнал **внесен в перечень научных специализированных изданий Украины**, в которых могут публиковаться результаты диссертационных работ на получение научных степеней доктора и кандидата технических наук и физико-математических наук (радиофизика).

В журнале бесплатно публикуются научные статьи на английском, русском и украинском языках.

**Правила оформления статей** представлены на сайте: <http://ric.zntu.edu.ua/information/authors>.

Журнал обеспечивает **бесплатный открытый он-лайн доступ** к полнотекстовым публикациям. Журнал разрешает авторам иметь авторские права и сохранять права на издание без ограничений. Журнал разрешает пользователям читать, загружать, копировать, распространять, печатать, искать или ссылаться на полные тексты своих статей. Журнал разрешает повторное использование его содержания в соответствии с СС лицензией CC-BY.

Опубликованным статьям присваивается уникальный идентификатор цифрового объекта DOI.

**Журнал реферируется и индексируется** в ведущих международных и национальных реферативных журналах и наукометрических базах данных, а также размещается в цифровых архивах и библиотеках с бесплатным доступом on-line (в т.ч. DOAJ, DOI, CrossRef, EBSCO, eLibrary.ru / РИНЦ, Google Scholar, Index Copernicus, INSPEC, ISSN, Ulrich's Periodicals Directory, WorldCat, ВИНТИ, Джэрэло), полный перечень которых представлен на сайте: <http://ric.zntu.edu.ua/about/editorialPolicies#custom-0>.

**Журнал распространяется** по Каталогу периодических изданий Украины (подписной индекс – 22914).

**Тематика журнала** включает: радиофизику, микро-, нано- и радиоэлектронику, аппаратное и программное обеспечение компьютерной техники, компьютерные сети и телекоммуникации, теорию алгоритмов и программирования, оптимизацию и исследование операций, межмашинное и человеко-машинное взаимодействие, математическое и компьютерное моделирование, обработку данных и сигналов, управление в технических системах, искусственный интеллект, включая системы, основанные на знаниях, и экспертные системы, интеллектуальный анализ данных, распознавание образов, искусственные нейронные и нейро-нечеткие сети, нечеткую логику, коллективный интеллект и мультиагентные системы, гибридные системы.

Все статьи, предлагаемые к публикации, получают **объективное рассмотрение**, которое оценивается по существу без учета расы, пола, вероисповедания, этнического происхождения, гражданства или политической философии автора(ов).

Все статьи проходят двухступенчатое закрытое (анонимное для автора) **рецензирование** штатными редакторами и независимыми рецензентами – ведущими учеными по профилю журнала.

## **РЕДАКЦИОННАЯ КОЛЛЕГИЯ**

**Главный редактор** – Погосов В. В., д-р физ.-мат. наук, Украина

**Зам. главного редактора** – Субботин С. А., д-р. техн. наук, Украина

**Члены редколлегии:**

Безрук В. М., д-р техн. наук, Украина

Бодянский Е. В., д-р техн. наук, Украина, редактор раздела по управлению

Васильев С. Н., д-р физ.-мат. наук, академик РАН, Россия

Гимпилевич Ю. Б., д-р техн. наук, Украина

Горбань А. Н., д-р. физ.-мат. наук, Великобритания

Дробахин О. О., д-р физ.-мат. наук, Украина

Зайцева Е. Н., канд. физ.-мат. наук, Словакия

Карпуков Л. М., д-р техн. наук, Украина

Корнич Г. В., д-р физ.-мат. наук, Украина, редактор раздела по радиофизике

Кулик А. С., д-р техн. наук, Украина

Лебедев Д. В., д-р техн. наук, Украина, редактор раздела по управлению

Левашенко В. Г., канд. физ.-мат. наук, Словакия

Лиснянский А., канд. техн. наук, Израиль

Марковская-Качмар У., д-р наук, Польша

Олещук В. А., канд. физ.-мат. наук, Норвегия, редактор радиоэлектроники

Онуфриенко В. М., д-р физ.-мат. наук, Украина

Папшицкий М., д-р философии, Польша

Пиза Д. М., д-р техн. наук, Украина

Потапенко Е. М., д-р техн. наук, Украина

Рубель О. И., канд. техн. наук, Канада

Хаханов В. И., д-р техн. наук, Украина, редактор раздела по информатике

Чумаченко В. П., д-р физ.-мат. наук, Украина, редактор раздела по радиоэлектронике

Шарпанских А. А., доктор философии, Нидерланды – редактор раздела по информатике

Рекомендовано к изданию ученым советом ЗНТУ, протокол № 6 от 22.12.2014 г.

Журнал сверстан редакционно-издательским отделом ЗНТУ.

**Веб-сайт журнала:** <http://ric.zntu.edu.ua>.

**Адрес редакции:** Редакция журнала «РИУ», Запорожский национальный технический университет, ул. Жуковского, 64, г. Запорожье, 69063, Украина.

Тел.: +38-061-769-82-96 – редакционно-издательский отдел

Факс: (061) 764-46-62

E-mail: [rvv@zntu.edu.ua](mailto:rvv@zntu.edu.ua)

**The scientific journal «Radio Electronics, Computer Science, Control»** is published by the Zaporizhzhya National Technical University (ZNTU) since 1999 with periodicity two numbers per year.

The journal is registered by the State Committee for information policy, television and radio broadcasting of Ukraine in 29.01.2003. The journal has a State Registration Certificate of printed mass media (series KB №6904).

**ISSN** 1607-3274 (print), **ISSN** 2313-688X (on-line).

By the Decree of the Presidium of All-Ukrainian Certification Commission № 1-05/4 dated 05.26.2010 the journal is **included to the list of the scientific specialized periodicals of Ukraine** where the results of dissertations for D.Sc. and Ph.D in Technical Sciences, Physics and Mathematics may be published.

The journal publishes scientific articles in English, Russian, and Ukrainian free of charge.

The **article formatting rules** are presented on the site: <http://ric.zntu.edu.ua/information/authors>.

The journal provides policy of **on-line open (free of charge) access** for full-text publications. The journal allow the authors to hold the copyright without restrictions and to retain publishing rights without restrictions. The journal allow readers to read, download, copy, distribute, print, search, or link to the full texts of its articles. The journal allow reuse and remixing of its content, in accordance with a CC license CC-BY.

Published articles have a unique digital object identifier (DOI).

**The journal is abstracted and indexed** in leading international and national abstracting journals and scientometric databases, and also placed to the digital archives and libraries with a free on-line access (including DOAJ, DOI, CrossRef, EBSCO, eLibrary.ru / РИНЦ, Google Scholar, Index Copernicus, INSPEC, ISSN, Ulrich's Periodicals Directory, WorldCat, VINITI (All-Russian Institute of scientific and technical information), Djerevo), full list of which is presented on the site: <http://ric.zntu.edu.ua/about/editorialPolicies#custom-0>.

**The journal is distributed** by the Catalogue of Ukrainian periodicals (the catalog number is 22914).

**The journal scope:** radio physics, micro-, nano- and radio electronics, computer hardware and software, computer networks and telecommunications, algorithm and programming theory, optimization and operations research, machine-machine and man-machine interfacing, mathematical modeling and computer simulation, data and signal processing, control in technical systems, artificial intelligence, including knowledge-based and expert systems, data mining, pattern recognition, artificial neural and neuro-fuzzy networks, fuzzy logics, swarm intelligence and multiagent systems, hybrid systems.

All articles proposed for publication receive an **objective review** that evaluates substantially without regard to race, sex, religion, ethnic origin, nationality, or political philosophy of the author(s).

All articles undergo a two-stage **blind peer review** by the editorial staff and independent reviewers – the leading scientists on the profile of the journal.

## EDITORIAL BOARD

**Editor-in-Chief:** V. V. Pogosov, Doctor of Science in Physics and Mathematics, Ukraine

**Deputy Editor-in-Chief:** S. A. Subbotin, Doctor of Science in Engineering, Ukraine

**Members of Editorial Board:**

V. M. Bezruk, Doctor of Science in Engineering, Ukraine

Ye. V. Bodyanskiy , Doctor of Science in Engineering, Ukraine, Control section editor

V. P. Chumachenko, Doctor of Science in Physics and Mathematics, Ukraine, Radio Electronics section editor

O. O. Drobakhin, Doctor of Science in Physics and Mathematics

Yu. B. Gimpilevich, Doctor of Science in Engineering, Ukraine

A. N. Gorban, Doctor of Science in Physics and Mathematics, United Kingdom

V. I. Hahanov, Doctor of Science in Engineering, Ukraine, Computer Science section editor

L. M. Karpukov, Doctor of Science in Engineering, Ukraine

G. V. Kornich, Doctor of Science in Physics and Mathematics, Ukraine, Radio Physics section editor

A. S. Kulik, Doctor of Science in Engineering, Ukraine

D. V. Lebedev, Doctor of Science in Engineering, Ukraine, Control section editor

V. G. Levashenko, Ph.D, Slovakia

A. Lisnianski, Ph.D, Israel

U. Markowska-Kaczmar, Doctor of Science, Poland

V. A. Oleshchuk, Ph.D in Physics and Mathematics, Norway, Radio Electronics section editor

V. M. Onufrienko, Doctor of Science in Physics and Mathematics, Ukraine

M. Paprzycki, Ph.D, Poland

D. M. Piza, Doctor of Science in Engineering, Ukraine

E. M. Potapenko, Doctor of Science in Engineering, Ukraine

O. I. Rubel, Ph.D, Canada

A. A. Sharpanskykh, Ph.D, Netherlands, Computer Science section editor

S. N. Vassilyev, Doctor of Science in Physics and Mathematics, Academician of Russian Academy of Sciences, Russia

E. N. Zaitseva, Ph.D, Slovakia

Recommended for publication by the Academic Council of ZNTU, protocol № 6 dated 22.12.2014.

The journal is imposed by the editorial-publishing department of ZNTU.

**The journal web-site** is <http://ric.zntu.edu.ua>.

**The address of the editorial office:** Editorial office of the journal «Radio Electronics, Computer Science, Control», Zaporizhzhia National Technical University, Zhukovskiy street, 64, Zaporizhzhya, 69063, Ukraine.

Tel.: +38-061-769-82-96 – the editorial-publishing department.

Fax: +38-061-764-46-62

E-mail: [rvv@zntu.edu.ua](mailto:rvv@zntu.edu.ua)

## ЗМІСТ

<b>РАДІОФІЗИКА.....</b>	<b>8</b>
Вовк С. М., Борулько В. Ф. Двойственный метод минимума пространственной протяженности для робастного оценивания параметров дипольных источников излучения.....	8
<b>РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ.....</b>	<b>18</b>
Бондарев Б. Н., Кабак В. С. Статистические характеристики сигналов многоканальных систем с OFDM.....	18
Зіньковський Ю. Ф., Сидорук Ю. К., Туровський А. О. Конструкційна оптимізація хвилеводно-вібраторного випромінювача з повернутою площину поляризації.....	23
Зубков А. М., Щерба А. А. Підвищення ефективності артилерійської розвідки шляхом конструктивно-функціональної інтеграції повітряних та наземних засобів спостереження.....	29
Крищук В. Н., Шило Г. Н., Каспирович Н. А., Огренич Е. В. Оптимизация оребренного канала в системах принудительного воздушного охлаждения РЭА.....	34
Марусенкова Т. А. Метод оцінювання похибки вимірювання магнітного поля 3D-зондом з нелінійними вихідними характеристиками.....	38
Невлюдов И. Ш., Андрусович А. А., Аллахверанов Р. Ю. Мониторинг обеспечения жизненного цикла РЭС при эксплуатации.....	44
Пиза Д. М., Звягинцев Е. А. Анализ влияния переменной базы на качество подавления помех при времязадержке пространственной обработке сигналов.....	51
<b>МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ.....</b>	<b>55</b>
Kozhukhivska O. A., Bidyuk P. I., Kudriachov V. F., Kozhukhivskyj A. D. Optimization of some reinsurance strategies.....	55
<b>НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ.....</b>	<b>64</b>
Subbotin S. A. The instance individual informativity evaluation for the sampling in neural network model synthesis.....	64
Korablev N. M., Ivaschenko G. S. Parallel immune algorithm of short-term forecasting based on model of clonal selection.....	73
Kucherenko Ye. I., Trokhimchuk S. N., Driuk O. D. Knowledge-oriented technologies in highly automated production.....	79
Romanuke V. V. A framework for classifier single training parameter optimization on training two-layer perceptron in a problem of turned 60-by-80-images classification.....	85
Богучарский С. И., Машталир С. В. Анализ текстур в последовательности изображений на основе векторного квантования.....	94
Дмитриченко В. Д., Заковоротный А. Ю. Архитектуры и алгоритмы функционирования нейронных сетей Хемминга и Хебба, способных дообучаться и распознавать новую информацию.....	100
Олійник А. О. Видобування знань на основі дерев розв'язків та стохастичного пошуку.....	110
Тищенко О. К., Плісс І. П., Шкуро К. О. Прогнозуюча нейро-фаззі мережа на основі багатовимірного нео-фаззі-нейрона та її процедура навчання.....	120
<b>ПРОГРЕСІВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ.....</b>	<b>126</b>
Kirichenko L. O., Kobitskaya Yu. A., Habacheva A. Yu. Comparative analysis of the complexity of chaotic and stochastic time series.....	126
Бабичев С. А. Оптимизация процесса предобработки информации в системах кластеризации высокоразмерных данных.....	135
Берко А. Ю., Висоцька В. А., Чирун Л. В. Проектування та впровадження систем електронної контент-комерції.....	143
Довбши А. С., Шелехов І. В., Прилепа Д. В. Інформаційно-екстремальний алгоритм навчання системи діагностування емоційно-психічного стану людини.....	156
<b>ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ.....</b>	<b>164</b>
Дорофеев Ю. И., Любчик Л. М. Робастное подавление возмущений при управлении насосными станциями в системе централизованного водоснабжения.....	164
<b>УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ.....</b>	<b>174</b>
Гурко А. Г., Янчевский И. В. Гарантированное управление движением манипуляционного робота.....	174

---

# CONTENTS

<b>RADIOPHYSICS.....</b>	<b>8</b>
<i>Vovk S. M., Borulko V. F.</i> Dual method of minimum spatial extent for robust estimation of dipole radiationsources.....	8
<b>RADIO ELECTRONICS AND TELECOMMUNICATIONS.....</b>	<b>18</b>
<i>Bondarev B. M., Kabak V. S.</i> Signals statistical characteristics of the multichannel systems with OFDM.....	18
<i>Zinkovskiy Yu. F., Sydoruk Yu. K., Turovskiy A. O.</i> Design optimization of waveguide fed dipole applicator with inclined polarization plane.....	23
<i>Zubkov A. M., Shcherba A. A.</i> Improving of artillery reconnaissance efficiency through constructive-functional integration of air and land surveillance.....	29
<i>Krischuk V. N., Shilo G. N., Kaspyrovych N. A., Ogrenich E. V.</i> Finned channel optimizing in radio electronic devices for forced air cooling systems.....	34
<i>Marusenkova T. A.</i> Method of evaluating inaccuracy of magnetic field measurement using a 3d-probe with nonlinear output characteristics.....	38
<i>Nevliudov I. Sh., Andrusevich A. A., Allahveranov R. U.</i> Monitoring of lifecycle of radio electronic equipment in exploitation.....	44
<i>Piza D. M., Zviahintsev E. A.</i> Analysis of influence of variable base on the quality of jammer suppression with time-space signal processing.....	51
<b>MATHEMATICAL AND COMPUTER MODELLING.....</b>	<b>55</b>
<i>Kozhukhivska O. A., Bidyuk P. I., Kudriachov V. F., Kozhukhivskyj A. D.</i> Optimization of some reinsurance strategies.....	55
<b>NEUROINFORMATICS AND INTELLIGENT SYSTEMS.....</b>	<b>64</b>
<i>Subbotin S. A.</i> The instance individual informativity evaluation for the sampling in neural network model synthesis.....	64
<i>Korablev N. M., Ivaschenko G. S.</i> Parallel immune algorithm of short-term forecasting based on model of clonal selectio...	73
<i>Kucherenko Ye. I., Trokhimchuk S. N., Driuk O. D.</i> Knowledge-oriented technologies in highly automated production.....	79
<i>Romanuke V. V.</i> A framework for classifier single training parameter optimization on training two-layer perceptron in a problem of turned 60-by-80-images classification.....	85
<i>Bogucharskiy S. I., Mashtalir S. V.</i> Image sequences texture analysis based on vector quantization.....	94
<i>Dmitrienko V. D., Zakoverotniy A. Yu.</i> Architecture and algorithms of neural networks hamming and hebb, capable learn and identify new information.....	100
<i>Oliinyk A.</i> Knowledge extraction based on decision trees and stochastic search.....	110
<i>Tyshchenko O. K., Pliss I. P., Shkuro K. O.</i> A forecasting neuro-fuzzy network based on a multidimensional neo-fuzzy neuron and its learning procedure.....	120
<b>PROGRESSIVE INFORMATION TECHNOLOGIES.....</b>	<b>126</b>
<i>Kirichenko L. O., Kobitskaya Yu. A., Habacheva A. Yu.</i> Comparative analysis of the complexity of chaotic and stochastic time series.....	126
<i>Babichev S. A.</i> Optimization of information preprocessing in clustering systems of high dimension data.....	135
<i>Berko A. Y., Vysotska V. A., Chyrun L. V.</i> Design and implementation of electronic content commerce systemsdesign and implementation of electronic content commerce systems.....	143
<i>Dovbysh A. S., Shelehov I. V., Prylepa D. V.</i> Information-extreme algorithm for system diagnostics emotional and mental person's state learning.....	156
<b>THEORY AND METHODS OF AUTOMATIC CONTROL.....</b>	<b>164</b>
<i>Dorofieiev Yu. I., Lyubchyk L. M.</i> Robust disturbances rejection in pumping stations control for centralized drinking water distribution system.....	164
<b>CONTROL IN TECHNICAL SYSTEMS.....</b>	<b>174</b>
<i>Gurko A. G., Yanchevskiy I. V.</i> Guaranteed control of robot-manipulator movement.....	174

## ПАМЯТИ КОЛЛЕГИ

Украинское образование и наука понесли тяжелую утрату – 5 декабря 2014 года на 76 году жизни ушел из жизни ученый и педагог, профессор кафедры электропривода и автоматизации промышленных установок Запорожского национального технического университета, доктор технических наук, профессор, многолетний главный редактор журнала «Электротехника и электроэнергетика», редактор раздела управления в технических системах журнала «Радиоэлектроника, информатика, управление» Потапенко Евгений Михайлович.

Родился Евгений Михайлович 15 декабря 1938 года в г. Тбилиси, в семье военнослужащего. Среднее образование получал в школах Белоруссии, России, Абхазии, Украины.

После окончания в 1962 году Харьковского авиационного института по специальности «Двигатели летательных аппаратов» работал в НПП «Хартрон» (1962–1983), где в качестве разработчика алгоритмов управления принимал участие в создании спутников «Целина Д», «Природа», «Січ», «Океан», из них 10 лет в Харькове и 11 лет в Запорожье. В 1970 году успешно защитил кандидатскую диссертацию на тему «Системы управления космическими аппаратами».

С 1983 года вся последующая жизнь Евгения Михайловича связана с кафедрой электропривода и автоматизации промышленных установок Запорожского национального технического университета, где он зарекомендовал себя как ведущий ученый в области систем управления техническими объектами. Он работает старшим преподавателем, доцентом, а с 1996 года профессором. В 1995 году в Институте кибернетики НАНУ им защищена докторская диссертация на тему «Синтез и анализ робастных систем управления».

Около 20 лет жизни профессора Потапенко Е.М. связаны с работой в редакционных коллегиях научных журналов. Из них более 4 лет Евгений Михайлович – главный редактор журнала «Электротехника и электроэнергетика» (2009–2012 гг.). Являлся членом специализированного ученого совета по защите диссертаций.

Более 30 лет Евгений Михайлович читал лекции по теории автоматического управления, подготовил десятки методических пособий.

Область научных интересов – высокоточные методы управления подвижными объектами и технологическими системами с применением векторного управления, в основном, двигателями переменного тока. Автор около 200 научных работ и более 20 изобретений. Совместно со своими учениками выпустил два учебных пособия и две научных монографии.

Вся профессиональная жизнь профессора Потапенко Е. М. связана с наукой и подготовкой аспирантов, чем он жил и посвящал все свободное время. Под его руководством 5 соискателей защитили кандидатские диссертации, был оппонентом более десяти докторских и большого количества кандидатских диссертаций.



Жизнь Евгения Михайловича неразрывно связана со спортом, общением с природой. Он кандидат в мастера по парусному спорту, перворазрядник по спортивной гимнастике. Имеет разряды по прыжкам в воду и тяжелой атлетике. Прививал здоровый образ жизни своим сыновьям и друзьям по утренней зарядке, которой он активно занимался на восходе солнца, практически каждый день, в течение всей жизни. Он был замечательным сыном, братом, мужем, отцом и дедушкой, всегда заботился о своих близких. Евгений Михайлович мог принести домой полуживых кошек и собак и выхаживать; подвязывать поломанные стволы березок на набережной, в засушливое лето носить воду из Днепра, чтобы поливать их.

Трудовые достижения Евгения Михайловича отмечены медалью «За Трудовую Доблесть» и множеством почетных грамот.

С уходом Евгения Михайловича Потапенко мы потеряли выдающегося ученого в области высокоточных методов управления подвижными объектами и технологическими системами, внесшего заметный вклад в развитие отечественной электротехники и подготовку высококвалифицированных кадров. Светлая память о нем и благодарность от учеников, коллег и друзей сохраняется в наших сердцах.

Коллектив кафедры электропривода и автоматизации промышленных установок Запорожского национального технического университета.

Редакция журналов «Радиоэлектроника, информатика, управление» и «Электротехника и электроэнергетика»

# РАДІОФІЗИКА

## РАДІОФІЗИКА

### RADIOPHYSICS

УДК 621.391.8

Вовк С. М.<sup>1</sup>, Борулько В. Ф.<sup>2</sup>

<sup>1</sup>Канд. физ.-мат. наук, доцент кафедры автоматизированных систем обработки информации,  
Днепропетровский национальный университет, Украина

<sup>2</sup>Канд. физ.-мат. наук, старший научный сотрудник кафедры прикладной и компьютерной радиофизики,  
Днепропетровский национальный университет, Украина

## ДВОЙСТВЕННЫЙ МЕТОД МИНИМУМА ПРОСТРАНСТВЕННОЙ ПРОТЯЖЕННОСТИ ДЛЯ РОБАСТНОГО ОЦЕНИВАНИЯ ПАРАМЕТРОВ ДИПОЛЬНЫХ ИСТОЧНИКОВ ИЗЛУЧЕНИЯ

Рассмотрена задача робастного оценивания параметров дипольных источников излучения по результатам измерений ближнего поля, когда данные измерений искажены аддитивным гауссовским шумом и выбросами. Предложен подход к обработке данных, который основан на принципе достижения минимума пространственной протяженности получаемого решения, а также минимума пространственной протяженности невязки решения. Использована модель источников излучения в виде одномерной решетки идеальных диполей Герца, которые расположены вдоль прямой линии, параллельной линии измерений электрического поля, а электрические моменты диполей перпендикулярны к плоскости измерений. Предложен алгоритм обработки данных и представлены результаты численного моделирования.

**Ключевые слова:** электрический диполь, пространственная протяженность, робастное оценивание, регуляризация.

#### ВВЕДЕНИЕ

Одной из важных задач теории антенн является задача оценивания пространственного положения и амплитуд источников излучения по результатам измерений ближнего поля [1–3]. Эта задача относится к классу обратных задач и является некорректной. В [4–5] мы предложили подход к решению такой задачи в скалярном случае, считая, что источники излучения являются точечными источниками однородных сферических волн. В [6] мы рассмотрели векторный случай, предполагая, что источники излучения являются диполями, а данные измерений искажены аддитивным гауссовским шумом. В данной работе мы предполагаем, что данные измерений искажены как аддитивным гауссовским шумом, так и выбросами, которые появляются из-за сбоев измерительной аппаратуры.

Так как поле системы диполей является суперпозицией индивидуальных полей диполей, то можно сформулировать задачу оценивания параметров дипольных источников излучения как задачу решения системы линейных алгебраических уравнений (СЛАУ). В общем случае матрица такой СЛАУ является прямоугольной; ее горизонтальный размер определяется числом «точек решения»

(т.е. точек, в которых источники излучения могут быть расположены), а вертикальный размер определяется числом «точек измерений» (т. е. точек, в которых измерения были проведены). Такой тип матрицы приводит к необходимости поиска псевдорешения [7–8]. Однако, если матрица СЛАУ является плохо обусловленной, то псевдорешение не дает хорошего решения тогда, когда данные измерений искажены шумом. В этом случае можно использовать метод регуляризации Тихонова [9], который основан на требовании получить решение с минимальной энергией и/или минимальной энергией его производных. Однако данный метод обычно дает гладкое решение и тогда, когда решение не является гладким. Для устранения этого недостатка вместо требования минимума энергии решения мы использовали требование минимума пространственной протяженности решения, а также предполагали, что источники излучения расположены редко в неизвестных узлах заданной дискретной сетки. Кроме того, из-за наличия выбросов в данных мы получали СЛАУ, в которой часть уравнений содержала грубые ошибки. Для устранения этого недостатка мы также требовали минимума пространственной протяженности невязки решения. В целом, комбинация этих двух требова-

ний позволила построить метод робастного оценивания [10], который требовался для оценивания параметров дипольных источников излучения при наличии аддитивного шума и выбросов в данных.

## 1 ПОСТАНОВКА ЗАДАЧИ

Существует много способов описания источников излучения с помощью набора электрических диполей. Дальше мы рассмотрим пример, где одному источнику излучения отвечает один электрический диполь. При этом мы также используем концепцию идеального колеблющегося электрического диполя Герца и исследуем случай среды без потерь. Тогда электромагнитное поле, которое излучает идеальный электрический диполь Герца, описывается следующими выражениями [11]:

$$\vec{E} = \frac{p}{4\pi\epsilon} \left[ \vec{r}_0 \frac{2}{r^2} \left( \frac{1}{r} + ik \right) \cos \theta + \vec{\theta}_0 \frac{1}{r} \left( \frac{1}{r^2} + \frac{ik}{r} - k^2 \right) \sin \theta \right] e^{-ikr}, \quad (1)$$

$$\vec{H} = \vec{\phi}_0 \frac{p\omega}{4\pi} \left( \frac{i}{r^2} - \frac{k}{r} \right) \sin \theta e^{-ikr}, \quad (2)$$

где  $\vec{E}$ ,  $\vec{H}$  – комплексные амплитуды электрического (1) и магнитного (2) полей, соответственно;  $\vec{r}_0$ ,  $\vec{\theta}_0$  и  $\vec{\phi}_0$  – орты сферической системы координат;  $i = \sqrt{-1}$ ;  $k = 2\pi/\lambda$  – волновое число;  $\lambda$  – длина волн;  $\epsilon$  – диэлектрическая проницаемость;  $r$  – расстояние от диполя до точки наблюдения;  $\omega$  – угловая частота;  $\vec{p} = p\vec{z}_0$  – электрический дипольный момент, где единичный вектор  $\vec{z}_0$  указывает ориентацию диполя в пространстве. Пусть электрические диполи расположены вдоль прямой линии (линии источников  $S$ ) и имеют одинаковую ориентацию своих дипольных моментов. Пусть линия измерений  $M$  является параллельной к  $S$  в некоторой плоскости и электрические моменты диполей являются перпендикулярными к этой плоскости (рис. 1). Тогда электрическое поле диполя имеет только компоненту

$$E_0 = \frac{p}{4\pi\epsilon} \frac{1}{r} \left( \frac{1}{r^2} + \frac{ik}{r} - k^2 \right) e^{-ikr}. \quad (3)$$

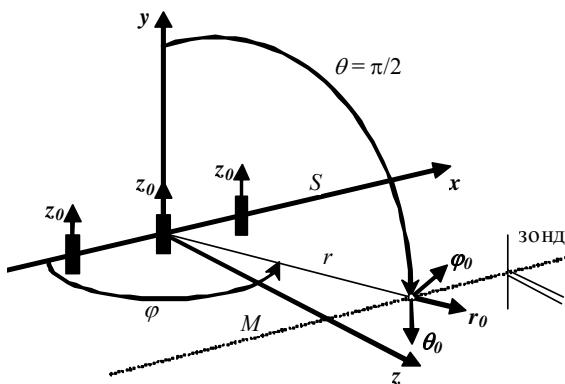


Рисунок 1 – Геометрия задачи

Используем дополнительную систему декартовых координат, в которой ось  $x$  совпадает с  $S$ , ось  $z$  перпендикулярна  $S$  и  $M$ , а ось  $y$  является параллельной вектору  $\vec{z}_0$  электрического дипольного момента. Тогда мы имеем  $E_y = -E_0$ . Пренебрегая влиянием измерительного зонда, на основе (3) можно записать СЛАУ:

$$b_j = \sum_{n=1}^N u_n \left[ k^2 - \frac{ik}{r_{nj}} - \frac{1}{r_{nj}^2} \right] \frac{e^{-ikr_{nj}}}{r_{nj}} + \eta_j;$$

$$r_{nj}^2 = x_{nj}^2 + z^2; x_{nj} = x_n^{(s)} - x_j^{(m)}; j = 1, \dots, J, \quad (4)$$

где  $b_j$  – значения  $E_y$  в «точках измерений»;  $u_n$  – искомые значения поля в «точках решения»;  $\eta_j$  – выборки шума, к которым мы относим как выборки аддитивного шума, так и выбросы, которые характеризуем вероятностью появления и законом распределения их значений;  $z$  – расстояние между  $S$  и  $M$ ;  $x_n^{(s)}$  и  $x_j^{(m)}$  обозначают  $x$ -координаты «точек решения» и «точек измерений», соответственно;  $N$  – число «точек решения»;  $J$  – число «точек измерений». Далее мы предполагаем, что диполи располагаются в некоторых «точках решения» и число диполей намного меньше, чем  $N$ .

В матричной записи СЛАУ (4) имеет вид:  $\mathbf{A}\mathbf{u} = \mathbf{b}$ , где  $\mathbf{A}$  является комплекснозначной матрицей, элементы которой формируются на основе пространственной модели для заданной системы диполей,  $\mathbf{b}$  является комплекснозначным вектором-столбцом, который содержит данные измерений;  $\mathbf{u}$  является комплекснозначным вектором-столбцом, который содержит искомые значения амплитуд источников. Задача оценивания параметров источников излучения заключается в решении СЛАУ (4).

Необходимо отметить, что непосредственное решение СЛАУ (4) в виде:  $\mathbf{u} = \mathbf{A}^{-1}\mathbf{b}$  является возможным только для случая квадратной матрицы  $\mathbf{A}$ , когда число «точек решения»  $N$  равно числу «точек измерений»  $J$ . Но такое решение имеет тот недостаток, что при плохой обусловленности матрицы  $\mathbf{A}$  оно является неустойчивым. Кроме того, если  $\mathbf{A}$  является прямоугольной матрицей, то непосредственное решение в указанном виде является невозможным. В этом случае можно использовать псевдорешение [7–8], например, решение по методу наименьших квадратов  $\mathbf{u} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{b}$ , получаемое как аналитическое решение задачи оптимизации:  $\|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2 \rightarrow \min_{\mathbf{u}}$ , где  $\mathbf{A}^H$  обозначает сопряженное (эрмитово) транспонирование комплекснозначной матрицы  $\mathbf{A}$  и где  $\|\dots\|^2$  обозначает квадрат евклидовой нормы. Но это решение имеет тот же самый недостаток, который проявляется при плохой обусловленности соответствующей матрицы СЛАУ. Здесь отметим, что в обоих указанных случаях число обусловленности матрицы СЛАУ возрастает при уменьшении расстояния между

точками измерений и при увеличении расстояния между точками измерений и точками решений. Более того,

решение  $\mathbf{u} = (\mathbf{A}^H \mathbf{A} + \rho^2 \mathbf{I})^{-1} \mathbf{A}^H \mathbf{b}$ , получаемое аналитически при применении к методу наименьших квадратов стандартной техники квадратичной регуляризации Тихонова:  $\|\mathbf{A}\mathbf{u} - \mathbf{b}\|^2 + \rho^2 \|\mathbf{u}\|^2 \rightarrow \min$ , где  $\rho^2$  является параметром регуляризации Тихонова [9], также имеет недостатки. Этими недостатками являются следующие: 1) решение получается слишком гладким и в тех случаях, когда истинное решение гладким не является; 2) если отдельные элементы данных разрушены выбросами, то получаемое решение может существенно отличаться от истинного решения и приводить к неверным результатам оценивания.

Мы предлагаем решать СЛАУ (4) на основе принципа минимума длительности [12]. Однако здесь мы меняем понятие «длительность» на понятие «пространственная протяженность» [13] и используем это понятие дважды, а именно в исходном пространстве (которым является пространство решения, или источников) и в сопряженном пространстве (которым является пространство невязки решения, или измерений). Поэтому мы назвали предлагаемый подход «двойственный метод минимума пространственной протяженности (ДММПП)» и сформулировали его как задачу оптимизации:

$$f(\mathbf{u}) = D_1[\mathbf{A}\mathbf{u} - \mathbf{b}] + \gamma^2 D_2[\mathbf{u}] \rightarrow \min_{\mathbf{u}}, \quad (5)$$

где  $f(\mathbf{u})$  является целевой функцией; член  $D_1[\mathbf{A}\mathbf{u} - \mathbf{b}]$  обозначает меру «пространственной протяженности» вектора невязки  $\mathbf{A}\mathbf{u} - \mathbf{b}$ ; член  $D_2[\mathbf{u}]$  обозначает меру «пространственной протяженности» вектора решения  $\mathbf{u}$ ;  $\gamma^2$  является параметром внешней регуляризации [4–6].

## 2 МАТЕРИАЛИ И МЕТОДЫ

Согласно (5), основная идея ДММПП заключается в требовании одновременно минимизировать как пространственную протяженность невязки решения, так и пространственную протяженность самого решения. Поэтому ключевым вопросом имплементации ДММПП является вопрос формализации понятия «пространственная протяженность».

Понятие «пространственная протяженность» мы применяем для обозначения некоторой меры носителя пространственно-зависимой функции и, следовательно, как обобщение понятий длина, площадь и объем. При выполнении формализации этого понятия необходимо рассматривать как случай отсутствия, так и случай наличия шума в данных [14].

Для случая отсутствия шума в данных мы используем понятие «строгая пространственная протяженность»,

которое формализуем в трехмерном пространстве в виде функционала

$$D = \iiint \kappa[s(x, y, z)] dx dy dz, \quad (6)$$

где  $s(x, y, z)$  – некоторая функция пространственных независимых переменных  $x, y, z$ , которая описывает объект исследования,  $\kappa[s(x, y, z)] = \begin{cases} 1, & s(x, y, z) \neq 0; \\ 0, & s(x, y, z) = 0 \end{cases}$  – идеальная функция стоимости каждого значения функции  $s(x, y, z)$ .

Для случая наличия шума в данных вместо (6) мы используем понятие «нестрогая пространственная протяженность» (или «квазипротяженность»)

$$D_{\alpha, \beta, \dots} = \iiint \psi[s(x, y, z); \alpha, \beta, \dots] dx dy dz, \quad (7)$$

где  $\psi[s(x, y, z); \alpha, \beta, \dots]$  – такая функция стоимости, которая позволяет получить конечные значения  $D_{\alpha, \beta, \dots}$  приискажении значений  $s(x, y, z)$  шумом и которая при отсутствии шума обеспечивает предельный переход  $D_{\alpha, \beta, \dots} \rightarrow D$  при устремлении параметров  $\alpha, \beta, \dots$  к своим граничным (предельным) значениям.

Дальше мы будем рассматривать только случай одномерной пространственной зависимости, т. е. случай одномерной функции  $s(x)$ . Это, во-первых, упрощает дальнейшее изложение материала, и, во-вторых, отвечает постановке задачи, согласно которой дипольные источники излучения расположены вдоль прямой линии, параллельной линии измерений электрического поля. Учитывая это, для дальнейшей формализации мы предлагаем использовать в качестве функции стоимости следующую функцию:

$$\psi_{BV}[s(x); \alpha, \beta, q] = c[(1 + |s(x)/\alpha|^q)^{\beta/q} - 1], \quad (8)$$

где  $0 < \alpha < \infty$ ,  $-\infty < \beta \leq 1$ ,  $0 < q < \infty$ ,  $c = 1/(1 + 1/|\alpha|^q)^{\beta/q} - 1$  – коэффициент нормировки, который обеспечивает равенство:  $\psi_{BV}[1; \alpha, \beta, q] = 1$ . Тогда, в частности, из (8) при  $q=2$  имеем:

$$\psi_{BV}[s(x); \alpha, \beta, 2] = c[(1 + |s(x)/\alpha|^2)^{\beta/2} - 1]. \quad (9)$$

Отметим, что для конечных значений  $\alpha$  функция (9) при  $\beta = 1$  отвечает функции стоимости для функционала квазипротяженности (7) медианного типа [14]; при  $\beta = 0$  – функции стоимости мириадного типа [15], при  $\beta = -2$  – функции стоимости Демиденко [16–17], а при  $\beta \rightarrow -\infty$  мы получаем в пределе идеальную функцию стоимости. Так как мы априорно предполагаем, что данные измерений искажены шумом и выбросами, то дальнейшее рассмотрение может быть связано только с понятием квазипротяженности и выяснением условий его применимости путем поиска рабочего диапазона параметров  $\alpha, \beta, \dots$

Дальше ми используем формализацию понятия «квазипротяженность» на основе функции (9). Тогда из (5) для дискретного случая мы имеем задачу оптимизации:

$$f(u_1, \dots, u_N) = c_1 \sum_{j=1}^J \left[ (1 + \left| \sum_{n=1}^N a_{jn} u_n - b_j \right|^2 / \alpha_1^2)^{\beta_1/2} - 1 \right] + \\ + \gamma^2 c_2 \sum_{n=1}^N \left[ (1 + |u_n|^2 / \alpha_2^2)^{\beta_2/2} - 1 \right] \rightarrow \min_{u_1, \dots, u_N}, \quad (10)$$

где  $-\infty < \beta_1 \leq 1$ ;  $c_1 = 1 / [(1 + 1 / \alpha_1^2)^{\beta_1/2} - 1]$ ;  $-\infty < \beta_2 \leq 1$ ;  $c_2 = 1 / [(1 + 1 / \alpha_2^2)^{\beta_2/2} - 1]$ ;  $a_{jn}$  есть элемент матрицы  $\mathbf{A}$ ;  $\alpha_1$  и  $\alpha_2$  являются параметрами «внутренней регуляризации». Из (10) можно легко получить задачи, основанные на функционалах квазипротяженности медианного типа с квадратичным слаживанием (полагая  $\beta_1 = \beta_2 = 1$ ), мириадного типа (устремляя  $\beta_1$  и  $\beta_2$  к 0) и демиденковского типа (полагая  $\beta_1 = \beta_2 = -2$ ).

Отметим также, что если  $\alpha_1^2 \gg | \sum_{n=1}^N a_{jn} u_n - b_j |^2$  для

всех  $j$ , то первый член (10) близок к квадратичному члену, и тогда мы имеем задачу [6], которая для  $\alpha_1^2 \gg 1$  может быть записана в виде:

$$f(u_1, \dots, u_N) = \sum_{j=1}^J \left| \sum_{n=1}^N a_{jn} u_n - b_j \right|^2 + \\ + \gamma^2 c_2 \sum_{n=1}^N \left[ (1 + |u_n|^2 / \alpha_2^2)^{\beta_2/2} - 1 \right] \rightarrow \min_{u_1, \dots, u_N}. \quad (11)$$

Это есть задача наименьших квадратов с неквадратичным регуляризирующим членом.

С другой стороны, если  $\alpha_2^2 \gg |u_n|^2$  для всех  $n$ , то второй член (10) стремится к квадратичному члену, который соответствует квадратичной регуляризации (слаживанию) Тихонова [9], что для  $\alpha_2^2 \gg 1$  может быть записано в виде:

$$f(u_1, \dots, u_N) = c_1 \sum_{j=1}^J \left[ (1 + \left| \sum_{n=1}^N a_{jn} u_n - b_j \right|^2 / \alpha_1^2)^{\beta_1/2} - 1 \right] + \\ + \gamma^2 \sum_{n=1}^N |u_n|^2 \rightarrow \min_{u_1, \dots, u_N}. \quad (12)$$

Это есть нелинейная (неквадратичная) задача минимизации с квадратичным регуляризирующим членом.

Наконец, если  $\alpha_1^2 \gg | \sum_{n=1}^N a_{jn} u_n - b_j |^2$  для всех  $j$  и если

$\alpha_2^2 \gg |u_n|^2$  для всех  $n$ , то и первый, и второй члены целевой функции (10) стремятся к квадратичным членам одновременно. В этом случае приближенно мы имеем метод наименьших квадратов с квадратичной регуляризацией Тихонова [9], который (при  $\alpha_1^2 \gg 1$  и  $\alpha_2^2 \gg 1$ ) может быть записан для дискретного случая как решение задачи оптимизации:

$$f(u_1, \dots, u_N) = \sum_{j=1}^J \left| \sum_{n=1}^N a_{jn} u_n - b_j \right|^2 + \gamma^2 \sum_{n=1}^N |u_n|^2 \rightarrow \min_{u_1, \dots, u_N}. \quad (13)$$

Метод, получаемый из (13), может быть также назван «двойственный метод наименьших квадратов», так как он устанавливает, что взвешенная (с помощью  $\gamma^2$ ) сумма квадратов значений невязки решения и сумма квадратов значений решения должна быть минимальна.

Для решения задачи (10) мы использовали численный метод, основанный на методе сопряженных градиентов и заданный с помощью следующей вычислительной схемы:

$$\begin{aligned} \mathbf{u}^{(t+1)} &= \mathbf{u}^{(t)} + h^{(t)} \mathbf{p}^{(t)}, \quad t \geq 0; \\ \mathbf{p}^{(0)} &= -\mathbf{g}^{(0)}; \quad t = 0; \quad \mathbf{p}^{(t)} = -\mathbf{g}^{(t)} + d^{(t-1)} \mathbf{p}^{(t-1)}; \quad t \geq 1; \\ d^{(t-1)} &= \|\mathbf{g}^{(t)}\|^2 / \|\mathbf{g}^{(t-1)}\|^2; \\ h^{(t)} &= \arg \min_h f(\mathbf{u}^{(t)} + h \mathbf{p}^{(t)}), \end{aligned} \quad (14)$$

где  $t$  – номер итерации,  $\mathbf{u}^{(t)}$  – вектор решения на  $t$ -й итерации,  $h^{(t)}$  – шаг вдоль направления спуска  $\mathbf{p}^{(t)}$  на  $t$ -й итерации;  $\mathbf{g}^{(t)}$  – вектор градиента функционала  $f$  на  $t$ -й итерации. Для применения (14) желательно упростить вычисления, обусловленные комплекснозначностью СЛАУ (4). Кроме этого, необходимо построить алгоритм вычисления вектора градиента функционала (10) и алгоритм решения одномерной задачи минимизации для выбора величины шага вдоль направления спуска.

Для упрощения вычислений легко преобразовать комплекснозначную СЛАУ (4) в вещественнозначную СЛАУ. Для этого перепишем (4) в виде:  $\mathbf{A}_0 \mathbf{u}_0 = \mathbf{b}_0$ , где

$\mathbf{A}_0 = \begin{bmatrix} \mathbf{A}^R & -\mathbf{A}^I \\ \mathbf{A}^I & \mathbf{A}^R \end{bmatrix}$  является чисто вещественнозначной блочной матрицей,  $\mathbf{A}^R$  и  $\mathbf{A}^I$  являются вещественной и мнимой частями комплекснозначной матрицы  $\mathbf{A}$ ;  $\mathbf{b}_0 = [\mathbf{b}^R \quad \mathbf{b}^I]^T$  является чисто вещественнозначным вектором измеренных значений, который состоит из вещественной  $\mathbf{b}^R$  и мнимой  $\mathbf{b}^I$  частей известных комплекснозначных данных;  $\mathbf{u}_0 = [\mathbf{u}^R \quad \mathbf{u}^I]^T$  является чисто

вещественнозначным вектором решения, который состоит из вещественной  $\mathbf{u}^R$  и мнимой  $\mathbf{u}^I$  частей комплекснозначного вектора решения.

Далее рассмотрим вопросы вычисления вектора градиента и решения одномерной задачи минимизации для частного случая дипольных источников с вещественнозначными амплитудами.

Если решение  $\mathbf{u}$  является вещественнозначным вектором, то мы имеем вещественнозначную СЛАУ:

$$\mathbf{A}_1 \mathbf{u} = \mathbf{b}_0, \text{ где } \mathbf{A}_1 = \begin{bmatrix} \mathbf{A}^R \\ \mathbf{A}^I \end{bmatrix}; \mathbf{u} = \mathbf{u}^R; \mathbf{b}_0 = \begin{bmatrix} \mathbf{b}^R \\ \mathbf{b}^I \end{bmatrix}. \text{ Из необходи-}$$

мого условия оптимальности задачи (10), для этого частного случая мы имеем систему из  $N$  нелинейных уравнений:

$$\frac{c_1\beta_1}{(\alpha_1)^{\beta_1}} \sum_{j=1}^J \frac{v_j}{(v_j^2 + \alpha_1^2)^{1-\beta_1/2}} a_{jk} + \gamma^2 \frac{c_2\beta_2}{(\alpha_2)^{\beta_2}} \frac{u_k}{(u_k^2 + \alpha_2^2)^{1-\beta_2/2}} = 0;$$

$$v_j = \sum_{n=1}^N a_{jn} u_n - b_j, \quad k = 1, \dots, N, \quad (15)$$

где  $a_{jn}$  – элементы матрицы  $\mathbf{A}_1$ ,  $b_j$  – элементы вектора  $\mathbf{b}_0$ ,  $u_k$  – элементы вещественнозначного вектора  $\mathbf{u}$ . Отметим, что левая часть (15) представляет компоненты вектора градиента функционала (10) для этого частного случая.

Если мы будем запускать на выполнение итерационный процесс (14) с  $\mathbf{u}^{(0)} = 0$  (т. е. с нулевыми начальными значениями  $u_k^{(0)} = 0$  для всех  $k$ ), то согласно (15) мы будем иметь начальный вектор градиента в виде:

$$\mathbf{g}^{(0)} = -(\mathbf{A}_1)^T \mathbf{w}, \quad (16)$$

где вектор  $\mathbf{w}$  состоит из элементов:

$$w_j = \frac{c_1\beta_1}{(\alpha_1)^{\beta_1}} \frac{b_j}{(b_j^2 + \alpha_1^2)^{1-\beta_1/2}}, \quad j = 1, \dots, J, \quad (17)$$

где  $b_j$  – элемент вектора  $\mathbf{b}_0$ . Если  $\alpha_1^2 \gg b_j^2$  и  $\alpha_1^2 \gg 1$ , то из (16)–(17) мы имеем:

$$\mathbf{g}^{(0)} \approx -(\mathbf{A}_1)^T \mathbf{b}_0, \quad (18)$$

где правая часть (18) является вектором градиента для метода наименьших квадратов.

Таким образом, алгоритм вычисления на  $t$ -й итерации вектора градиента функционала (10) состоит из следующих шагов: 1) вычислить компоненты вектора невязки  $\mathbf{v}^{(t)}$  по формуле:  $v_j^{(t)} = \sum_{n=1}^N a_{jn} u_n^{(t)} - b_j$ ;  $j = 1, \dots, J$ ; 2) вычислить компоненты вспомогательного вектора

$\mathbf{w}^{(t)}$  по формуле:  $w_j^{(t)} = \frac{c_1\beta_1}{(\alpha_1)^{\beta_1}} \frac{v_j^{(t)}}{([v_j^{(t)}]^2 + \alpha_1^2)^{1-\beta_1/2}}$ ;

$j = 1, \dots, J$ ; 3) вычислить компоненты вектора градиента  $\mathbf{g}^{(t)}$  по формуле:

$$g_k^{(t)} = \sum_{j=1}^J a_{kj}^T w_j^{(t)} + \gamma^2 \frac{c_2\beta_2}{(\alpha_2)^{\beta_2}} \frac{u_k^{(t)}}{([u_k^{(t)}]^2 + \alpha_2^2)^{1-\beta_2/2}}, \quad k = 1, \dots, N.$$

Одномерная задача минимизации:

$h^{(t)} = \arg \min_h f(\mathbf{u}^{(t)} + h \mathbf{p}^{(t)})$  заключается в выборе величины шага  $h$  вдоль направления спуска  $\mathbf{p}^{(t)}$ . Мы предлагаем решать эту задачу путем использования набора «пробных шагов». Этот набор включает шаги, которые мы условно называли так: «шаги, которые обнуляют невязку решения», «шаги, которые обнуляют решение», и «шаг, который отвечает методу Ньютона».

Шаги, которые обнуляют невязку решения, используются для обнуления некоторого отсчета невязки и определяются формулой:

$$h_j^{discrepancy} = -\left(\sum_{n=1}^N a_{jn} u_n^{(t)} - b_j\right) / \left(\sum_{n=1}^N a_{jn} p_n^{(t)}\right), \quad j = 1, \dots, J. \quad (19)$$

Всего существует  $J$  таких шагов на  $t$ -й итерации.

Шаги, которые обнуляют решение, используются для обнуления некоторого отсчета решения. Эти шаги определяются формулой:

$$h_n^{solution} = -u_n^{(t)} / p_n^{(t)}, \quad n = 1, \dots, N. \quad (20)$$

Всего существует  $N$  таких шагов на  $t$ -й итерации.

Шаги (19) и (20) уменьшают  $f$  в соответствии с (10).

Шаг, который отвечает методу Ньютона, используется для минимизации (10) при условии, что решение находится в окрестности некоторого локального минимума функционала (10). Этот шаг определяется формулой:

$$h^{Newton} = -(\mathbf{g}^{(t)}, \mathbf{p}^{(t)}) / ([\mathbf{A}^T \mathbf{Q}_1 \mathbf{A} + \gamma^2 \mathbf{Q}_2] \mathbf{p}^{(t)}, \mathbf{p}^{(t)}), \quad (21)$$

где  $\mathbf{Q}_1 = diag(\dots, \frac{c_1(\beta_1/2)}{(\alpha_1)^{\beta_1}} \frac{(\alpha_1^2 - (1-\beta_1)[v_j^{(t)}]^2)}{(\alpha_1^2 + [v_j^{(t)}]^2)^{2-\beta_1/2}}, \dots)$  является  $\mathbf{J} \times \mathbf{J}$  диагональной матрицей, зависящей от  $\alpha_1^2$  и от значений вектора невязки  $\mathbf{v}^{(t)}$  на  $t$ -й итерации, и где  $\mathbf{Q}_2 = diag(\dots, \frac{c_2(\beta_2/2)}{(\alpha_2)^{\beta_2}} \frac{(\alpha_2^2 - (1-\beta_2)[u_n^{(t)}]^2)}{(\alpha_2^2 + [u_n^{(t)}]^2)^{2-\beta_2/2}}, \dots)$  являет-

ся  $N \times N$  диагональной матрицей, зависящей от  $\alpha_2^2$  и от значений текущего вектора решения  $\mathbf{u}^{(t)}$  на  $t$ -й итерации. Всего существует только один такой шаг на  $t$ -й итерации.

Алгоритм использования «пробных шагов» следующий. Мы подставляем каждый из этих шагов, рассчитанных с помощью (19)–(21), в одномерную задачу минимизации и выбираем в качестве  $h^{(t)}$  такое значение шага, для которого функционал (10) принимает минимальное значение. Если наилучшее значение имеет один из шагов  $h_j^{discrepancy}$  или  $h_n^{solution}$ , то мы обновляем направление спуска по формуле:  $\mathbf{p}^{(t)} = -\mathbf{g}^{(t)}$ . Если шаг  $h^{Newton}$ , который отвечает методу Ньютона, принимает наилучшее значение  $N$  раз подряд, то мы также обновляем направление спуска. Если наилучшее значение шага равно нулю, то мы останавливаем итерационный процесс.

### 3 ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ

Рис. 2 и рис. 3 представляют результаты численного моделирования прямой и обратной задач для случая одномерной решетки электрических диполей, когда диполи расположены вдоль прямой линии, которая парал-

ельна линии измерений электрического поля, и когда их электрические моменты перпендикулярны к плоскости измерений (рис. 1). Приведенные на рис. 2 и рис. 3 графики (кроме графика на рис. 2а) отображают исходные данные и полученные результаты для одной из 100 реализаций смеси полезного сигнала (которым является реальная и мнимая части электрической компоненты ближнего поля), аддитивного шума и выбросов.

Для моделирования прямой задачи мы использовали (4) со следующими значениями:  $N=J=200$ ;  $\Delta x^{(s)} = \Delta x^{(m)} = \Delta x$ ;  $\Delta x/\lambda = 0,1$ ;  $z/\lambda = 1$ , амплитуды  $u_n = 1$  для  $n=40, 70, 100, 130, 160$  и  $u_n = 0$  для других  $n$  (рис. 2а). Значения  $b_j$  были искажены аддитивным гауссовским шумом и выбросами в виде случайных импульсов. Стандартное отклонение гауссовского шума приблизительно составляло 5 % от максимума абсолютного значения для вещественной (рис. 2б) и для мнимой (рис. 2в) частей электрической компоненты поля, соответственно. Амплитуды выбросов были равномерно распределены в интервале  $[-1, 1]$ , а вероятность их появления была равна 0,1. Это означало, что почти 10 % дискретных отсчетов данных были полностью разрушены, и мы априорно не знали, какие именно дискретные отсчеты данных являлись некорректными.

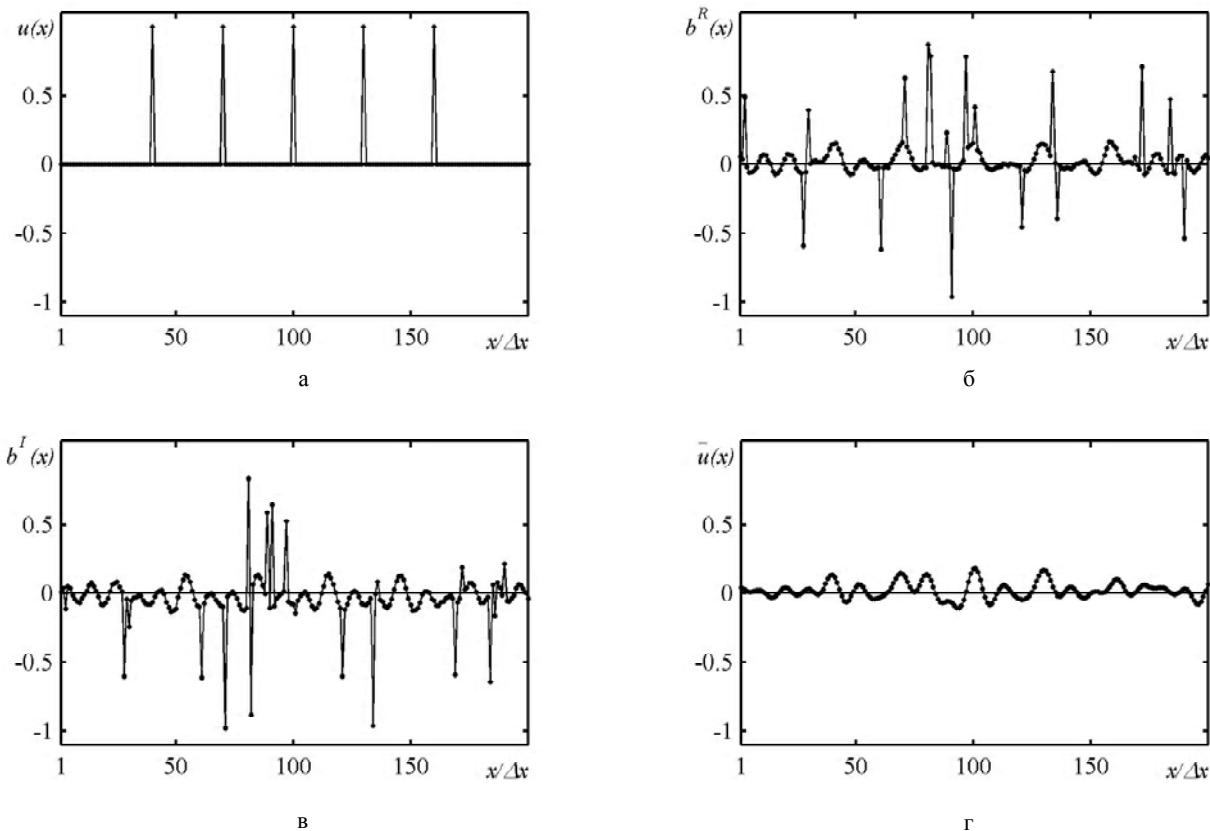


Рисунок 2 – Пространственное распределение электрического поля: а – истинное вещественнонзначное пространственное распределение; б – вещественная и в – мнимая части электрического поля, искаженные аддитивным шумом и выбросами;

г – решение обратной задачи после использования регуляризации Тихонова с оптимальным значением  $\gamma^2 = 1$

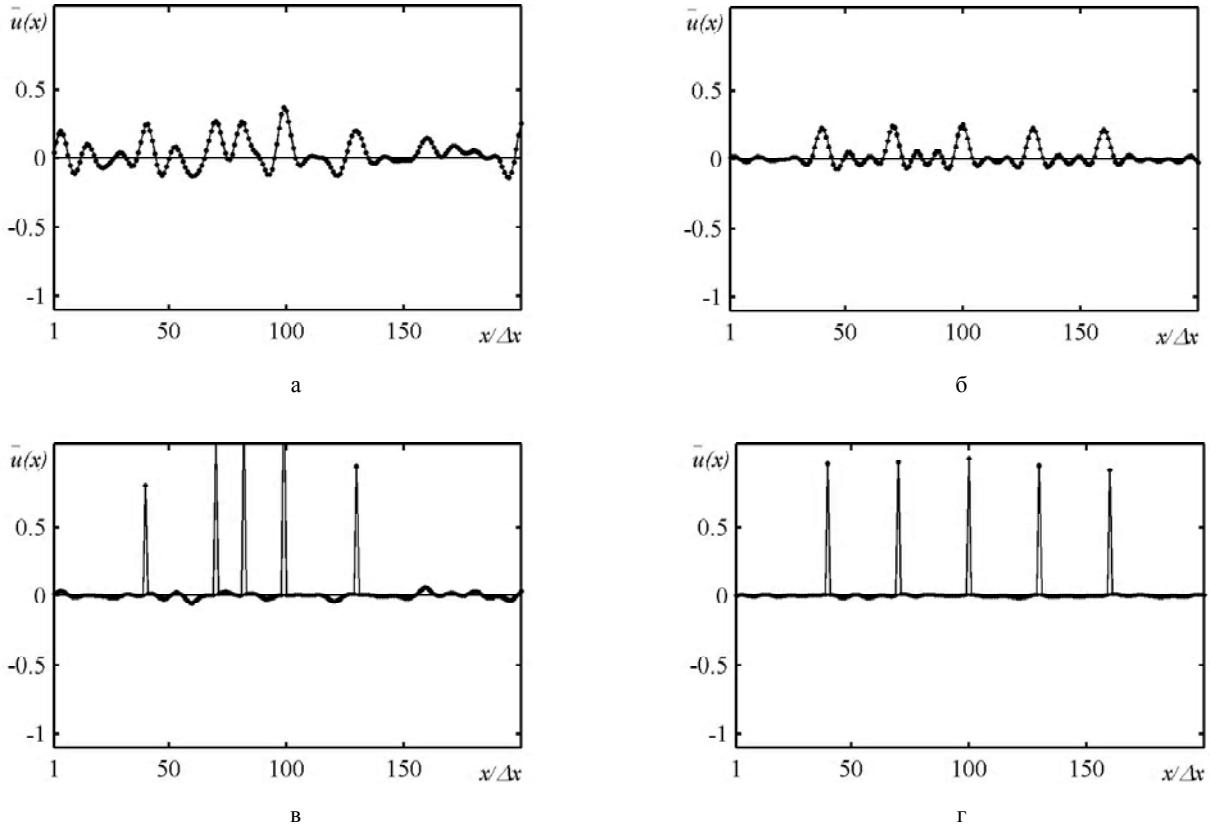


Рисунок 3 – Пространственное распределение после использования ДММПП: а – при  $\alpha_1 = 1, \alpha_2 = 1, \gamma^2 = 0,1$ , когда  $\alpha_1$  и  $\alpha_2$  являются очень большими; б – при  $\alpha_1 = 0,1, \alpha_2 = 1, \gamma^2 = 0,1$ , когда  $\alpha_2$  является очень большим; в – при  $\alpha_1 = 1, \alpha_2 = 0,1, \gamma^2 = 0,1$ , когда  $\alpha_1$  является очень большим; г – при  $\alpha_1 = 0,1, \alpha_2 = 0,1, \gamma^2 = 0,1$ , которые близки к оптимальным значениям

Решение обратной задачи с помощью метода регуляризации Тихонова (13) для оптимального значения  $\gamma^2 = 1$  представлено на рис. 2г. Видно, что это решение не позволяет корректно оценить местоположение и амплитуды источников излучения.

Решение обратной задачи с помощью ДММПП на основе мириадного функционала (т.е. предельного случая функционала (10) при  $\beta_1 \rightarrow 0$  и  $\beta_2 \rightarrow 0$ ) при выборе  $\mathbf{u}^{(0)} = 0$  для различных значений параметров регуляризации представлено на рис. 3.

Рис. 3а показывает решение, полученное при  $\alpha_1 = 1, \alpha_2 = 1, \gamma^2 = 0,1$ , когда  $\alpha_1$  и  $\alpha_2$  имеют большие значения, для которых первый и второй члены функционала (10) являются близкими к квадратичным членам. Видно, что это решение похоже на Тихоновское решение (13), отображенное на рис. 2г, и оно не позволяет корректно оценить местоположение и амплитуды источников излучения.

Рис. 3б показывает решение, полученное при  $\alpha_1 = 0,1, \alpha_2 = 1, \gamma^2 = 0,1$ , когда  $\alpha_2$  имеет большое значение, для которого второй член функционала (10) является близким к квадратичному члену, и, поэтому, когда

мы приближенно имеем задачу (12). Здесь мы видим, что это решение не имеет недостатков, обусловленных наличием выбросов, но оно не удовлетворяет требованию минимальной пространственной протяженности решения, что затрудняет получение оценок значений искомых параметров.

Рис. 3в показывает решение, полученное при  $\alpha_1 = 1, \alpha_2 = 0,1, \gamma^2 = 0,1$ , когда  $\alpha_1$  имеет большое значение, для которого первый член функционала (10) является близким к квадратичному члену, и, поэтому, когда мы приближенно имеем задачу (11). Как и ожидалось, выбросы значений разрушают решение, приводя к неверным результатам оценивания.

Наконец, решение с помощью ДММПП для квазиоптимальных значений параметров  $\alpha_1 = 0,1, \alpha_2 = 0,1, \gamma^2 = 0,1$ , полученное после 144 итераций, представлено на рис. 3г. В этом случае графики полученного (рис. 3г) и истинного (рис. 2а) решений были почти идентичны. Отметим, что полученное решение позволяет корректно оценить местоположение и амплитуды источников излучения путем использования дискретных отсчетов, для которых абсолютное значение амплитуды больше  $\alpha_2$ .

Мы также исследовали поведение максимального значения относительной среднеквадратической ошибки (ОСКО) решения от параметров метода. Табл.1 содержит результаты моделирования, полученные по 100 реализациям смеси полезного сигнала, гауссовского аддитивного шума и выбросов для различных значений внутренних и внешнего параметров регуляризации при применении мириадного функционала. Значения ОСКО были рассчитаны по формуле:

$$\delta = \left[ \sum_{n=1}^N (\bar{u}_n - u_n)^2 / \sum_{n=1}^N u_n^2 \right]^{1/2} \times 100\%, \text{ где } \bar{u}_n \text{ обозначает}$$

дискретный отсчет полученного решения, а  $u_n$  обозначает дискретный отсчет истинного решения. Максимальное значение ОСКО рассчитывалось на основе значений ОСКО, полученных для 100 различных реализаций смеси полезного сигнала, аддитивного шума и выбросов для заданных значений параметров регуляризации, и оно указывало на наихудший (из 100 полученных результатов) результат. Табл. 1 показывает, что для случая на рис. 3г максимальное значение ОСКО было равно 12,68 %. Она также показывает, что существуют такие значения параметров регуляризации, для которых достигается еще меньшее значение максимума ОСКО. Здесь можно видеть, что минимум максимальных значений ОСКО приблизительно равен проценту содержания гауссовского шума во входных данных (около 5 %). Таким образом, табл. 1 показывает степень устойчивости решения и указывает на существование оптимальных значений параметров регуляризации для случая «мириадной» обработки.

Отметим, что большие значения ОСКО были обусловлены различными причинами. Большие значения ОСКО для больших значений  $\alpha_1, \alpha_2$  были получены из-за сдвига глобального минимума относительно истинного решения. Но большие значения ОСКО для малых значений  $\alpha_1, \alpha_2$  были получены из-за нарушения сходимости алгоритма. Это обусловлено тем, что для малых значений  $\alpha_1, \alpha_2$  местоположение глобального минимума является близким к истинному значению, но существует много локальных минимумов. Количество локальных минимумов становится тем больше, чем меньше значения  $\alpha_1, \alpha_2$ . Поэтому процесс сходимости к глобальному минимуму нарушается, а сходимость к локальному минимуму дает неверные результаты.

Таблица 1 – Максимальное значение ОСКО (%) по 100 реализациям при применении мириадного функционала

$\alpha_1^2 = \alpha_2^2$	$\log_{10}(\gamma^2)$				
	-2	-1,5	-1	-0,5	0
0,1	328,56	189,13	100,44	91,17	94,77
0,01	123,69	16,30	12,68	12,27	101,46
0,001	95,05	7,46	5,05	4,50	78,41
0,0001	145,00	117,10	57,47	56,98	141,18

Табл. 2 показывает зависимости  $\delta_{\max}, \delta_{cp}, t_{\max}, \delta$  и  $t$  от параметров  $\beta_1, \beta_2$  для случая использования квазиоптимальных значений  $\alpha_1 = 0,1, \alpha_2 = 0,1, \gamma^2 = 0,1$  при выборе  $q=2$ , где  $\delta_{\max}$  и  $\delta_{cp}$  – максимальное и среднее значения ОСКО, вычисленные по 100 реализациям смеси полезного сигнала, аддитивного шума и выбросов;  $t_{\max}$  – максимальное число выполненных при этом итераций;  $\delta$  – значение ОСКО при отсутствии шума и выбросов, полученное за  $t$  итераций. Из табл. 2 видно, что существует диапазон значений параметров  $\beta_1, \beta_2$ , внутри которого достигаются малые значения как максимального ( $\approx 12,7\%$ ), так и среднего ( $7,5\%$ ) значений ОСКО. Этот диапазон сосредоточен приблизительно вокруг значений  $\beta_1 = \beta_2 = 0$ , которые отвечают использованию мириадных функционалов в задаче (5). Полученные небольшие по величине значения  $t_{\max}$  указывают на возможность использования предложенного алгоритма для обработки данных в реальном масштабе времени.

Правая часть табл. 2 дает представление об ошибке метода для случая отсутствия шума и выбросов при выборе не оптимальных (которые в данном случае должны быть близкими к нулю), а указанных выше квазиоптимальных значений параметров внутренней регуляризации. Видно, что ОСКО  $\delta$  уменьшается при переходе от квазимедианного функционала через мириадный функционал к функционалу Демиденко, что обусловлено ослаблением зависимости соответствующих им функций стоимости от параметров внутренней регуляризации. Однако ни квазимедианный функционал, ни функционал Демиденко для данного случая «не работают» в отличие от мириадного, корневого и квазидемиденковского функционалов. Этот результат был связан с преждев-

Таблица 2 – ОСКО  $\delta$  и число итераций  $t$  при применении мириадного функционала

$\beta_1 = \beta_2$	Данные с шумом и выбросами (100 реализаций)			Данные без шума и выбросов	
	$\delta_{\max}, \%$	$\delta_{cp}, \%$	$t_{\max}$	$\delta, \%$	$t$
-2	71,31	28,11	207	69,88	26
-1	71,65	12,65	228	1,01	243
-1/2	12,60	6,76	233	1,69	298
-1/4	12,61	7,05	214	3,62	350
-1/8	12,63	7,29	193	3,72	532
-1/16	12,65	7,41	154	4,05	343
-1/32	12,66	7,48	144	4,36	310
0	12,68	7,58	187	4,64	118
1/32	12,72	7,66	155	4,77	146
1/16	12,77	7,78	152	5,30	386
1/8	12,80	8,17	227	6,08	281
1/4	73,17	11,66	296	7,83	140
1/2	77,26	70,98	251	74,55	51
1	86,62	85,02	63	85,08	18

ременным остановом вычислений из-за невозможности формирования удачного направления спуска, на что указывает соответствующее малое значение числа итераций  $t$ . Это указывает на возможность и необходимость дальнейшего совершенствования техники решения предложенной задачи оптимизации.

## ВЫВОДЫ

ДММПП является развитием и обобщением традиционного метода наименьших квадратов, который использует регуляризацию Тихонова. ДММПП является устойчивым к аддитивному шуму и выбросам и одновременно позволяет решать задачу оценивания в условиях, когда местоположения источников априорно неизвестны. Показано, что существуют оптимальные значения параметров регуляризации, для которых полученное решение практически совпадает с искомым решением. Численное моделирование подтвердило преимущество предложенного метода по сравнению с методом наименьших квадратов, который использует регуляризацию Тихонова. Предложенная реализация метода, основанная на методе сопряженных градиентов, и оригинальный метод решения одномерной задачи оптимизации на основе техники «пробных шагов» позволяют достичь высокой скорости обработки с возможностью ее выполнения в реальном масштабе времени. Наши дальнейшие исследования направлены на построение адаптивного подхода к выбору параметров регуляризации.

## СПИСОК ЛИТЕРАТУРЫ

- Yaghjian A. D. An Overview of Near-Field Antenna Measurements // IEEE Transactions on Antennas and Propagation. – 1986. – Vol. 34, Jan. – P. 30–44.
- Cano F. Applications of sources reconstruction techniques: Theory and practical results / [F. Cano, M. Sierra-Castaner, S. Burgos, J. L. Besada] // Antennas and Propagation : 4th European Conference EuCAP-2010, Barcelona, 12–16 April 2010 : proceedings. – IEEE, 2010. – P. 1–5.
- Martini E. Reduction of Truncation Errors in Planar Near-Field Aperture Antenna Measurements Using the Gerchberg-Papoulis Algorithm / E. Martini, O. Breinbjerg, S. Maci // IEEE Transactions on Antennas and Propagation. – 2008. – Vol. 56, Nov. – P. 3485–3493.
- Vovk S. M. Using the method of minimum extension to extrapolation of antenna measurements / S. M. Vovk, V. F. Borulko // Mathematical Methods in Electromagnetic Theory : 10th International Conference MMET-2004, Dnipropetrovsk, 14–17 September 2004 : proceedings. – IEEE, 2004. – P. 159–161. DOI: 10.1109/MMET.2004.1396962
- Vovk S. M. Restoration of point sources located along a perpendicular to a plane of measurements by the extension minimum method / S. M. Vovk, V. F. Borulko // Mathematical Methods in Electromagnetic Theory : 11th International Conference MMET-2006, Kharkiv, 26–29 June 2006 : proceedings. – IEEE, 2006. – P. 147–149. DOI: 10.1109/MMET.2006.1689728
- Borulko V. F. Determining sources in the array of electric dipoles / V. F. Borulko, S. M. Vovk // Mathematical Methods in Electromagnetic Theory : International Conference MMET-2010, Kyiv, 6–8 September 2010 : proceedings. – IEEE, 2010. – P. 1–4. DOI: 10.1109/MMET.2010.5611347
- Forsythe G. E. Computer Methods for Mathematical Computations / G. E. Forsythe, M. A. Malcolm, C. B. Moler. – New York, Prentice-Hall, 1977. – 259 p.
- Co T. B. Methods of Applied Mathematics for Engineers and Scientists / T. B. Co. – New York, Cambridge University Press, 2013. – 582 p.
- Tikhonov A. N. Solutions of ill-posed problems / A. N. Tikhonov, V. Y. Arsenin. – New York, Wiley, 1977. – 272 p.
- Huber P. J. Robust statistics / P. J. Huber. – New York : John Wiley and Sons, 1981. – 312 p.
- Neff H. P. Introductory Electromagnetics / H. P. Neff. – New York, Wiley, 1991. – 414 p.
- Vovk S. M. A minimum-duration method for recovering finite signals / Vovk S. M., Borul'ko V. F. // Radioelectronics and Communications. – 1991. – No. 8, vol. 34. – P. 67–69.
- Borulko V. Image Processing by Minimum Extent Method / V. Borulko, S. Vovk // Pattern Recognition and Information Analysis : International Conference PRIA-95, Minsk, 19–21 September 1995 : proceedings. – Szczecin : Wydawnictwo Uczelniane Politechniki Szczecinskiej, 1995. – P. 71–72.
- Vovk S. M. Statement of a problem of definition of linear signals parameters in quasinormed space / S. M. Vovk, V. F. Borul'ko // Radioelectronics and Communications. – 2010. – Vol. 53, No. 7. – P. 367–375.
- Vovk S. M. Myriad and meridian functionals as extreme forms of smoothed quasinorms / S. M. Vovk, V. F. Borulko // Mathematical Methods in Electromagnetic Theory : 14th International Conference MMET-2012, Kyiv, 28–30 August 2012 : proceedings. – IEEE, 2012. – P. 320–323. DOI: 10.1109/MMET.2012.6331268
- Vovk S. Family of generalized Demidenko functionals for robust estimation / S. Vovk, V. Borulko // Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory : 17th International Seminar/Workshop DIPED'12, Tbilisi, 24–27 September 2012 : proceedings. – IEEE, 2012. – P. 151–154.
- Vovk S. Spectrum extrapolation by using combined root-Demidenko family of quasi-duration functionals / S. Vovk, V. Borulko // Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves : 8th International Kharkov Symposium MSMW'13, Kharkov, 23–28 June 2013 : proceedings. – IEEE, 2013. – P. 461–463. DOI: 10.1109/MSMW.2013.6622088

Стаття поступила в редакцію 10.06.2014.  
После доработки 27.10.2014.

Вовк С. М.<sup>1</sup>, Борулько В. Ф.<sup>2</sup>

<sup>1</sup>Канд. фіз.-мат. наук, доцент кафедри автоматизованих систем обробки інформації, Дніпропетровський національний університет, Україна

<sup>2</sup>Канд. фіз.-мат. наук, старший науковий співробітник кафедри прикладної та комп'ютерної радіофізики, Дніпропетровський національний університет, Україна

## ПОДВІЙНИЙ МЕТОД МІНІМУМУ ПРОСТОРОВОЇ ПРОТЯЖНОСТІ ДЛЯ РОБАСТНОГО ОЦІНЮВАННЯ ПАРАМЕТРІВ ДИПОЛЬНИХ ДЖЕРЕЛ ВИПРОМІНЮВАННЯ

Розглянута задача робастного оцінювання параметрів дипольних джерел випромінювання за результатами вимірювань близького поля, коли дані вимірювань спотворені аддитивним гауссівським шумом та викидами. Запропоновано підхід до оброб-

ки даних, який заснований на принципі досягнення мінімуму просторової протяжності отримуваного розв'язку, а також мінімуму просторової протяжності нев'язки розв'язку. Використана модель джерел випромінювання у вигляді одновимірної гратки ідеальних диполів Герца, які розташовані уздовж прямої лінії, паралельної до лінії вимірювань електричного поля, а електричні моменти диполів перпендикулярні до площини вимірювань. Запропоновано алгоритм обробки даних та представлені результати числового моделювання.

**Ключові слова:** електричний диполь, просторова протяжність, робастне оцінювання, регуляризація.

Vovk S. M.<sup>1</sup>, Borulko V. F.<sup>2</sup>

<sup>1</sup>PhD, Associate Professor of the department of Automated Systems of Information Processing, Dnipropetrovsk National University, Ukraine

<sup>2</sup>PhD, Senior Scientist of the department of Applied and Computer Radio Physics, Dnipropetrovsk National University, Ukraine

### DUAL METHOD OF MINIMUM SPATIAL EXTENT FOR ROBUST ESTIMATION OF DIPOLE RADIATION SOURCES

The robust estimation of dipole radiation-source parameters from near-field measurement is considered. The case of data distorted by additive noise and by spikes is investigated. A new approach based on a principle of minimal spatial extent is suggested. This approach states that the solution extent as well as the extent of solution discrepancy should be minimal together. The model of radiation sources described by the one-dimensional array of ideal Hertzian dipoles is used. The task geometry, when dipoles are located along a straight line, which is parallel to the measurement line of electric field, and when electric moments of all dipoles are perpendicular to the measurement plane, is considered. Algorithm of data processing based on conjugate gradient method is proposed. Numerical simulations with the error boundaries are presented.

**Keywords:** electric dipole, spatial extent, robust estimation, regularization.

## REFERENCES

1. Yaghjian A. D. An Overview of Near-Field Antenna Measurements, *IEEE Transactions on Antennas and Propagation*, Vol. 34, Jan. 1986, pp. 30–44.
2. Cano F., Sierra-Castaner M., Burgos S., Besada J. L. Applications of sources reconstruction techniques: Theory and practical results, *Antennas and Propagation, 4th European Conference EuCAP-2010*. Barcelona, 12–16 April 2010, proceedings, IEEE, 2010, pp. 1–5.
3. Martini E., Breinbjerg O., Maci S. Reduction of Truncation Errors in Planar Near-Field Aperture Antenna Measurements Using the Gerchberg-Papoulis Algorithm, *IEEE Transactions on Antennas and Propagation*, 2008, Vol. 56, Nov, pp. 3485–3493.
4. Vovk S. M., Borulko V. F. Using the method of minimum extension to extrapolation of antenna measurements, *Mathematical Methods in Electromagnetic Theory, 10th International Conference MMET-2004*. Dniproptetrovsk, 14–17 September 2004, proceedings, IEEE, 2004, pp. 159–161. DOI: 10.1109/MMET.2004.1396962
5. Vovk S. M., Borulko V. F. Restoration of point sources located along a perpendicular to a plane of measurements by the extension minimum method, *Mathematical Methods in Electromagnetic Theory : 11th International Conference MMET-2006*. Kharkiv, 26–29 June 2006, proceedings, IEEE, 2006, pp. 147–149. DOI: 10.1109/MMET.2006.1689728
6. Borulko V. F., Vovk S. M. Determining sources in the array of electric dipoles, *Mathematical Methods in Electromagnetic Theory : International Conference MMET-2010*. Kyiv, 6–8 September 2010 : proceedings, IEEE, 2010, pp. 1–4. DOI: 10.1109/MMET.2010.5611347
7. Forsythe G. E., Malcolm M. A., Moler C. B. Computer Methods for Mathematical Computations. New York, Prentice-Hall, 1977, 259 p.
8. Co T. B. Methods of Applied Mathematics for Engineers and Scientists. New York, Cambridge University Press, 2013, 582 p.
9. Tikhonov A. N., Arsenin V. Y. Solutions of ill-posed problems. New York, Wiley, 1977, 272 p.
10. Huber P. J. Robust statistics. New York, John Wiley and Sons, 1981, 312 p.
11. Neff H. P. Introductory Electromagnetics, New York, Wiley, 1991, 414 p.
12. Vovk S. M., Borul'ko V. F.: A minimum-duration method for recovering finite signals, *Radioelectronics and Communications*, 1991, vol. 34, No. 8, pp. 67–69.
13. Borulko V., Vovk S. Image Processing by Minimum Extent Method, *Pattern Recognition and Information Analysis : International Conference PRIA-95*. Minsk, 19–21 September 1995 : proceedings, Szczecin, Wydawnictwo Uczelniane Politechniki Szczecinskiej, 1995, pp. 71–72
14. Vovk S. M., Borul'ko V. F. Statement of a problem of definition of linear signals parameters in quasinormed space, *Radioelectronics and Communications*, 2010, Vol. 53, No. 7, pp. 367–375.
15. Vovk S. M., Borulko V. F. Myriad and meridian functionals as extreme forms of smoothed quasinorms, *Mathematical Methods in Electromagnetic Theory, 14th International Conference MMET-2012*. Kyiv, 28–30 August 2012, proceedings, IEEE, 2012, pp. 320–323. DOI: 10.1109/MMET.2012.6331268
16. Vovk S., Borulko V. Family of generalized Demidenko functionals for robust estimation, *Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory : 17th International Seminar, Workshop DIPED '12*. Tbilisi, 24–27 September 2012, proceedings, IEEE, 2012, P. 151–154.
17. Vovk S., Borulko V. Spectrum extrapolation by using combined root-Demidenko family of quasi-duration functionals, *Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves, 8th International Kharkov Symposium MSMW'13*. Kharkov, 23–28 June 2013, proceedings, IEEE, 2013, pp. 461–463. DOI: 10.1109/MSMW.2013.6622088

# РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

## РАДІОЕЛЕКТРОНИКА ТА ТЕЛЕКОММУНИКАЦІИ

### RADIO ELECTRONICS AND TELECOMMUNICATIONS

УДК 621.391

Бондарев Б. Н.<sup>1</sup>, Кабак В. С.<sup>2</sup><sup>1</sup>Канд. тех. наук, доцент, профессор кафедры радиотехники и телекоммуникаций, Запорожский национальный технический университет, Украина<sup>2</sup>Канд. тех. наук, доцент кафедры радиотехники и телекоммуникаций, Запорожский национальный технический университет, Украина

#### СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ СИГНАЛОВ МНОГОКАНАЛЬНЫХ СИСТЕМ С OFDM

Проводится уточнение имеющихся в литературе статистических характеристик (законов распределения, числовых характеристик, пик-фактора) групповых сигналов многоканальных систем передачи информации с использованием частотного уплотнения с ортогональными поднесущими (OFDM). Отмечается не совсем корректное использование при вычислении пик-фактора многоканального сигнала закона распределения Релея для его огибающей. Получены новые результаты для значений пикфактора. Показана допустимость превышения групповым сигналом порогового уровня на одном или нескольких тактовых интервалах, что проявляется как импульсная помеха на этих интервалах во всех каналах. При большом числе каналов это явление практически не будет влиять на помехоустойчивость.

Полученные результаты позволяют снизить на 2–4 дБ требования к динамическому диапазону и линейности амплитудных характеристик аппаратуры многоканальных систем с OFDM.

**Ключевые слова:** уплотнение, средняя мощность, максимальная мощность, пик-фактор, закон распределения вероятностей, групповой сигнал.

#### ВВЕДЕНИЕ

В последние годы все большее применение в многоканальных системах находит частотный метод уплотнения с ортогональными поднесущими (OFDM) и многопозиционными сигналами с квадратурной модуляцией (QAM). К таким системам относятся перспективные системы мобильной связи стандарта LTE, системы цифрового телевидения DVB – стандартов, системы широкополосного доступа WiMax. Число поднесущих частот (частотных каналов) в технологии LTE составляет около 200, в цифровом телевидении – около 2000 или 8000.

Наиболее полное изложение принципов OFDM, достаточно полный перечень первоисточников и краткие исторические сведения приведены в [1]. Впервые этот метод уплотнения был использован в американской системе коротковолновой радиосвязи «Кинеплекс» (1958 г.) [2, 3]. Американский патент на этот метод был выдан в 1970 г. [4]. В 1971 г. [5] было предложено исполь-

зование быстрого преобразования Фурье для операций модуляции и детектирования. Следует отметить также более поздние издания [6...10].

В [11] R. van Nee и A. de Wild впервые ввели понятие пикфактора многоканального OFDM-сигнала (но не совсем корректно). В [12] также рассматривалась задача определения пикфактора. Однако, приводимые результаты вызывают сомнения.

Отметим, что определение пикфактора является лишь одной из задач при решении более общей проблемы – определение требований к линейности амплитудных характеристик аппаратуры и уровня межканальных помех.

#### 1 ПОСТАНОВКА ЗАДАЧИ

Целью настоящей работы являются уточнение и корректировка методики и результатов статистических расчетов пикфактора сигналов в многоканальных системах с OFDM, приводимых в [1, 8, 12].

## 2 МАТЕРИАЛЫ И МЕТОДЫ

Канальные (индивидуальные) сигналы при использовании OFDM и QAM представляют собой отрезки синусоид с различными амплитудами и фазами. При этом несущие частоты соседних каналов различаются на величину ортогонального разноса  $F=1/T_s$ , где  $T_s$  – длительность многопозиционного символа. Величина  $1/T_s$  определяет и так называемую символьную (бодовую) скорость  $R_s$ .

При ортогональном разносе частот между каналами временные функции сигналов (гармонических колебаний) в соседних каналах отличаются на один период колебания. Полное число периодов для радиосигналов измеряется сотнями, тысячами.

Многоканальный сигнал на одном символьном интервале  $T_s$  можно записать в виде

$$S_{\Sigma}(t) = \sum_{k=1}^N a_{kl} \cos(k \frac{2\pi}{T_s} t + \varphi_{kl}), \quad \frac{1}{T_s} = F, \quad (1)$$

а в случае радиосигналов (после модуляции радионесущей)

$$S_{\Sigma}(t) = \sum_{k=1}^N a_{kl} \cos\left[\left(f_0 + \frac{k}{T_s}\right)2\pi t + \varphi_{kl}\right], \quad (2)$$

где  $N$  – число каналов,  $F$  – частота поднесущей в первом канале,  $a_{kl}$  и  $\varphi_{kl}$  – амплитуды и фазы канальных сигналов.

Следовательно, системы с использованием OFDM относятся к системам с параллельным способом передачи канальных сигналов. Поэтому такие системы имеют те же недостатки, что и системы с обычным частотным уплотнением: большие значения пикфактора группового сигнала и повышенные требования к линейности амплитудных характеристик аппаратуры.

Для любого из сигналов можно определить действующее  $U_d$  и максимальное  $U_{\max}$  значения и вычислить значения средней мощности  $P_{\text{ср}}$  и максимальной мощности  $P_{\max}$ . Отношение  $P_{\max}/P_{\text{ср}} = \Pi^2$  называют пик-фактором сигнала. В [1, 8, 12] для пикфактора  $\Pi^2$  используется обозначение PAP, PAPR (peak-to-average-power ratio). Там же вводится понятие (термин) «крест-фактор» (Crest factor, CF), определяемый как  $CF = \sqrt{PAP} = \Pi$ .

Для отрезка синусоиды с постоянной амплитудой и любыми начальными фазами полагают  $P_{\max}/P_{\text{ср}} = \Pi^2 = 2$ .

Для группового (многоканального) сигнала (1) можно применить теорему теории вероятностей о повторении опытов. При этом число опытов нужно полагать равным числу каналов  $N$ , а число возможных событий – числу возможных символов  $M$ , которые можно считать равновероятными с вероятностями, равными  $1/M$ . Вероятности появления на одном временном (символьном) интервале в многоканальных системах определенного сочетания канальных символов будут определяться полиномиальным (многомерным биномиальным) законом распределения вероятностей. При этом вероятностные расчеты будут трудно выполнимыми и практически бесполезными.

Как известно из теории вероятностей, сумма большого числа случайных слагаемых имеет закон распределения, приближающийся к нормальному, с дисперсией (средней мощностью), равной сумме дисперсий (средних мощностей) слагаемых. Следовательно, средняя мощность группового сигнала будет пропорциональна числу частотных каналов  $N$ . Этот же коэффициент будет определять так называемую цену уплотнения (по средней мощности). Определение максимальной мощности группового сигнала в соответствии с выражением (3) представляет собой неразрешимую задачу. Для ее решения необходимо использовать вероятностные (статистические) методы, полагая групповой сигнал случайным процессом с нормальным законом распределения вероятностей.

Для нормального закона распределения известно так называемое правило «3-х сигма», согласно которому значения нормального случайного процесса с вероятностью, равной 0,003, будут находиться за пределами интервала, равного 6-и сигма. Поэтому для нормального случайного процесса часто полагают, что его пикфактор  $\Pi^2 = 3^2 = 9$  (около 10 дБ). Для уменьшения указанной вероятности в теории и технике многоканальной связи (для систем с частотным уплотнением) при передаче речевых (телефонных) сообщений полагают  $\Pi^2 = 16 \dots 25$  (12 … 14 дБ). Отметим, что приведенные цифры отражают тот факт, что для случайных процессов (сигналов) понятие пикфактора имеет вероятностный характер.

Пользуясь соответствующими таблицами для нормального закона распределения, можно определить, что вероятность нахождения случайной величины за пределами интервала шириной 8 сигма (пикфактор равен 16 дБ) равна 0,000063. Следовательно, для многоканальных систем параллельного типа можно ограничиться значением пикфактора группового сигнала  $\Pi^2 = 16$ .

Для приведенных выше расчетов можно воспользоваться приближенным выражением для интегральной функции распределения нормального закона распределения

$$F(x) \approx \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left( 1/x - 1/x^3 + \dots \right). \quad (3)$$

Для вероятности превышения случайной величиной некоторого порога  $x_0$  (с учетом отрицательных значений  $x$ ) можно записать

$$P(|x| > x_0) \approx 0,8 \frac{\sigma}{x_0} e^{-x_0^2/2\sigma^2} = \frac{0,8}{z} e^{-z^2/2}, \quad (4)$$

где  $z^2 = x_0^2 / \sigma^2 = \Pi^2$  – пикфактор для нормального закона распределения,  $\sigma^2 = P_{\text{ср}}$  – дисперсия (средняя мощность).

На рис. 1 кривая 1 показывает зависимость вероятности от выбранного значения пик-фактора  $\Pi^2$  для нормального случайного процесса с нулевым средним значением в соответствии с выражением (4). Это выражение и кривую 1 далее будем считать «эталонными», истинными при сравнении результатов для нормальных случайных процессов, получаемых с использованием разных методов.

Приведенные выше формулы можно применить и к многоканальным системам OFDM с большим числом каналов. Однако, в [1, 11] предложены оригинальные, но не совсем корректные определение понятия «пикфактор» многоканального OFDM-сигнала и способ (методика) его вычисления.

Суть предложенной в [1, 11] методики заключается в следующем. Сначала утверждается, что многоканальный сигнал (1) имеет нормальный закон распределения, нулевое среднее значение и дисперсию (среднюю мощность), равную  $\sigma^2 = P_{\text{cp}}$ . Затем предполагается, что он является узкополосным процессом (2) с релеевским законом распределения его огибающей. Далее для релеевского закона распределения огибающей

$$\omega(\rho) = \frac{\rho}{\sigma^2} e^{-\rho^2/2\sigma^2} \quad (5)$$

рассчитывается вероятностная кривая

$$P(\rho \geq \rho_{\max}) = e^{-\rho_{\max}^2/2\sigma^2} = e^{-z^2}, \quad (6)$$

где величина  $z^2 = \rho_{\max}^2 / 2\sigma^2 = \Pi^2$  ошибочно (в 2 раза меньше) полагается равной пик-фактору сигнала (фактически пикфактору огибающей). Параметр  $\sigma^2$  является параметром (дисперсией) «порождающего» нормального закона. Параметр  $2\sigma^2 = m_2$  для закона Релея является вторым начальным моментом, т. е. полной мощностью случайного процесса, что не совсем соответствует содержанию термина PAPR.

Выражению (6) соответствует кривая 2 на рис. 1. Эта зависимость значительно отличается от «эталонной» зависимости (4) для нормального закона распределения (кривая 1 на рис. 1). Разница в значениях пикфактора составляет около 2-х децибелл.

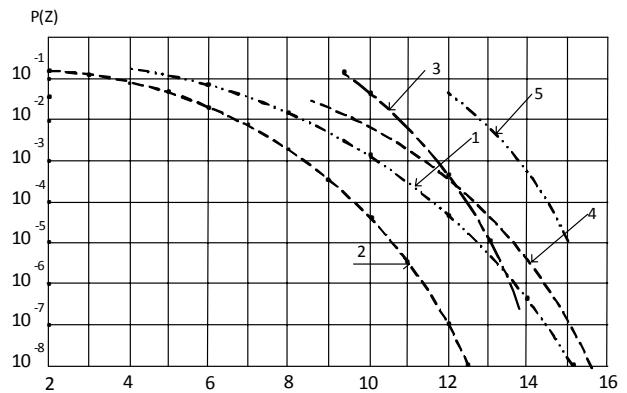
Отметим теперь вторую неточность при выводе выражения (6). Физический смысл понятие «огибающей» имеет для узкополосного процесса, который является двухполлярным (рис. 2). Поэтому необходимо учитывать огибающие как для положительных, так и отрицательных частей случайного процесса. Вероятность превышения огибающей заданного уровня в выражении (6) необходимо удвоить, как это было сделано в выражении (4):

$$P(\rho \geq \rho_{\max}) = 1 - P(\rho \leq \rho_{\max}) = 2e^{-\rho_{\max}^2/2\sigma^2} = 2e^{-z^2}. \quad (7)$$

В соответствии с этим кривые 2 и 3 рис. 1 должны быть приподняты вверх с увеличением вероятностей в 2 раза.

При корректном определении пик-фактора вместо выражений (7) и (6) необходимо записать

$$P(\rho \geq \rho_{\max}) = 2 \cdot e^{-\rho_{\max}^2/2\sigma^2} = 2e^{-z^2/2}. \quad (8)$$



1 – для нормального закона, 2 – для закона Релея при  $\sigma_p^2 = 2\sigma^2$ , 3 – по методике [1], 4 – по закону Релея при  $M_2 = \sigma^2$ , 5 – полученные результаты

Рисунок 1 – Зависимости значений вероятности от выбранного значения пик-фактора сигнала

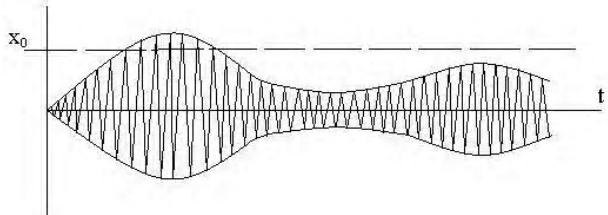


Рисунок 2 – Узкополосный случайный процесс

Зависимость (8) отображается кривой 4 на рис. 1 (без учета коэффициента 2). Отличие кривых 2 и 4 на рис. 1 при использовании закона Релея составляет ровно 3 дБ (в 2 раза), что и следовало ожидать. Отличие эталонной кривой 1 (нормальный закон распределения) и уточненной кривой 4 для релеевского закона составляет около одного децибелла.

Заметим, что результат в выражении (6) можно не удваивать, если рассматривать сигналы в радиочастотном тракте при его работе с отсечкой в режиме В.

Далее в [1], сохраняя смысл термина пик-фактор, производится пересчет вероятностной шкалы с учетом следующего условия. Оно заключается в требовании, чтобы многоканальный сигнал не превышал уровень значения пикфактора на всех « $N$ » временных (тактовых) интервалах в пределах символьного интервала  $T_s$ :

$$P_N(\rho \leq \rho_{\max}) = (1 - e^{-\rho_{\max}^2/2\sigma^2})^N, \quad (9)$$

$$P_N(\rho \geq \rho_{\max}) = 1 - (1 - e^{-\rho_{\max}^2/2\sigma^2})^N. \quad (10)$$

Полученные при этом результаты расчета оказываются отличающимися от результатов, полученных при моделировании. Для корректировки расчетных результатов в выражения (9) и (10) вводится корректирующий коэффициент  $\alpha = 2,8$  [1]

$$P_N(\rho \geq \rho_{\max}) = 1 - (1 - e^{-\frac{\rho_{\max}^2}{2\sigma^2}})^{2N} \approx 2,8Ne^{-\frac{\rho_{\max}^2}{2\sigma^2}} = 2,8Ne^{-Z^2}. \quad (11)$$

Кривая зависимости (11) при  $N = 1024$  показана на рис. 1 (кривая 3). Она совпадает с зависимостью, приводимой в [1] и [8]. Заметим, что в других источниках приводятся кривые в соответствии с (9), что менее наглядно.

Заметим, что в [12] приводится выражение, аналогичное (11), но с коэффициентом 0,64.

Кривая 5 на рис. 1 при  $N = 1024$  аналогична кривой 3, но при вычислениях в соответствии с (10) использовано выражение (4) для нормального закона вместо выражения (6) для закона распределения Релея.

Отметим, что точное и приближенное равенства в (11) имеют разный физический смысл. Точное равенство определяет вероятность нахождения сигнала ниже порога на всех тактовых интервалах. Приближенное равенство в (11) определяет вероятность превышения порога на каком-либо одном из  $N$  тактовых интервалов (при малых значениях экспоненты). Эта вероятность отображается кривой 5 на рис. 1.

Заметим теперь, что превышение заданного уровня только на одном из тактовых интервалов можно рассматривать как появление на этом интервале во всех каналах импульсной помехи малой длительности. Однако это почти не влияет на помехоустойчивость приема. Вероятность одновременного появления такой помехи на нескольких тактовых интервалах можно вычислить с использованием биномиального закона распределения

$$P_N(k) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k}. \quad (12)$$

В соответствии с (12) имеем

$$P_N(1) = N \cdot p (1-p)^{N-1}, \quad P_N(2) = \frac{N(N-1)}{2} p^2 (1-p)^{N-2}, \quad (13)$$

что, например, при  $p = 10^{-5}$  и  $N = 1000$  дает  $P_{1000}(2) \approx 5 \cdot 10^{-5}$ .

Следовательно, условия (10) и (9) при малых значениях по (7) не являются необходимыми. Сравнивая (13) и (4) (кривые 1 и 4 на рис. 1), при  $\frac{x}{\sigma} = 4$  (12 дБ) получим соотношение  $\frac{P(\rho \geq \rho_{\max})}{P(|x| > x_0)} = \frac{2 \cdot x}{0,86} \Rightarrow \frac{8}{0,8} = 10$ .

Достаточным оказывается условие (6) в виде

$$e^{-z^2} \ll 1/N. \quad (14)$$

## ВЫВОДЫ

Показана некорректность имеющихся в литературе результатов для значений пик-фактора групповых сигналов многоканальных систем с OFDM. Получены новые результаты для значений пик-фактора (кривые 1 и 5 на рис. 1). При  $p = 10^{-5}$  для кривой 1 рис. 1 получено значение пикфактора около 13 дБ. Полученные результаты

позволяют уточнить (снизить на 2–4 дБ) требования к динамическому диапазону и линейности амплитудных характеристик аппаратуры многоканальных систем при допущении нескольких превышений уровня на интервале времени, равном длительности символа  $T_s$ , что эквивалентно появлению на этих интервалах не очень опасной импульсной помехи. Показано, что условия (9) и (10) не являются необходимыми.

Предполагается в следующей статье проверить полученные результаты с использованием иного математического аппарата.

## БЛАГОДАРНОСТИ

Работа выполнена в рамках выполнения научно-исследовательской темы Запорожского национального технического университета «Разработка и обоснование рекомендаций по построению систем пространственно-временной обработки сигналов в условиях воздействия комбинированных помех» (№ гос. регистрации 0114U002636).

## СПИСОК ЛИТЕРАТУРЫ

- Van Nee R. OFDM for wireless Multimedia Communication / R. van Nee, R. Prasad. – Boston : Artech House, London, 2000. – 257 p.
- Doelz M. L. Binary Data Transmission Techniques for linear System / M. L. Doelz, E. T. Heald, D. L. Martin // Proc. JRE. – 1957. – Vol. 45, May. – P. 656–661.
- Mosier R. R. Kineplex, a bandwidth efficient binary Transmission System / R. R. Mosier, R. G. Clabaugh // IEEE Trans. – 1958. – Vol. 76, Jan. – P. 723–728.
- Patent No. 3, U. S. 488,4555, filed November 14, Orthogonal Frequency Division Multiplexing, 1966, issued Jan. 6, 1970.
- Weinstein S. B. Data Transmission by Frequency Division Multiplexing using the Discrete Fourier Transform / S. B. Weinstein, P. M. Ebert // IEEE Trans Comm. – 1971. – vol. Com 19. – Oct. – P. 628–637.
- Schulze H. Theory and Application of OFDM and CDMA widebound wireless communication / H. Schulze, Christian Luders. – John Wiley and Sons, 2005. – 420 p.
- Вишневский В. М. Широкополосные беспроводные сети передачи информации / [В. М. Вишневский, А. И. Ляхов, С. Л. Портной, И. В. Шахнович]. – М. : Техносфера, 2005. – 592 с.
- Andrews J. G. Fundamentals of WiMAX: understanding broadband wireless networking / J. G. Andrews, A. Ghosh, R. Muhamed. – Pearson Education, 2007. – 439 p.
- Holma H. LTE for UMTS: OFDMA and SC-FDMA Based Radio Access / H. Holma, A. Toskala. – John Wiley and Sons, 2009. – 433 p.
- Вишневский В. М. Энциклопедия WiMAX. Путь к 4G / В. М. Вишневский, С. Л. Портной, И. В. Шахнович. – М. : Техносфера, 2009. – 472 с.
- Van Nee R. Reducing the peak-to-average power ratio of OFDM / R. Van Nee, A. de Wild // Proc. IEEE Vehicular Technology Conf. (VTC'98). – 1998. – May. – P. 2072–2076.
- Hideki O. On the Distribution of the Peak to Average Power Ratio in OFDM Signals / O. Hideki, J. Hideki // IEEE Trans on Commun. – 2001. – Vol. 49, № 2. – February. – P. 282–289.

Статья поступила в редакцию 18.06.2014.  
После доработки 12.09.2014.

Бондарев Б. М.<sup>1</sup>, Кабак В. С.<sup>2</sup>

<sup>1</sup>Канд. техн. наук, доцент, професор кафедри радіотехніки та телекомунікацій, Запорізький національний технічний університет, Україна

<sup>2</sup>Канд. техн. наук, доцент професор кафедри радіотехніки та телекомунікацій, Запорізький національний технічний університет, Україна

## СТАТИСТИЧНІ ХАРАКТЕРИСТИКИ СИГНАЛІВ БАГАТОКАНАЛЬНИХ СИСТЕМ З OFDM

Проводиться уточнення відомих з літератури статистичних характеристик групових сигналів (законів розподілу ймовірностей, чисельних характеристик, пікфактора) багатоканальних систем передачі з використанням частотного ущільнення з ортогональними піднесучими (OFDM). Відзначаються не зовсім коректне використання під час оцінки пікфактора багатоканального сигналу закону Релея для його обвідної.

Отримані нові результати для значень пік-фактора. Показана припустимість перевищення сигналом порогового рівня на одному або декількох тактових інтервалах, що проявляється як імпульсна завада на цих інтервалах в усіх каналах. Здобуті результати дозволяють знизити вимоги на 2–4 dB до динамічного діапазону і лінійності амплітудної характеристики апаратури багатоканальних систем з OFDM.

**Ключові слова:** ущільнення, середня потужність, максимальна потужність, пік-фактор, поріг, закон розподілу ймовірностей, груповий сигнал.

Bondarev B. M.<sup>1</sup>, Kabak V. S.<sup>2</sup>

<sup>1</sup>Ph.D., Associate Professor, Professor the Department of Radio Engineering and Telecommunications, Zaporizhzhia National Technical University, Ukraine

<sup>2</sup>Ph.D., Associate Professor the Department of Radio Engineering and Telecommunications, Zaporizhzhia National Technical University, Ukraine

## SIGNALS STATISTICAL CHARACTERISTICS OF THE MULTICHANNEL SYSTEMS WITH OFDM

The accurate definition of the found in literature baseband signals statistical characteristics (distribution, numerical results, peak-to-average ratio) in the multichannel systems with orthogonal frequency-division multiplexing (OFDM) is studied. Not strictly correct variable normalization and discrepancies in peak-to-average ratio estimation with using of Rayleigh distribution for envelope are noted.

New results for peak-to-average ratio meanings are obtained. The acceptability of exceeding by threshold value signal on one or several clock intervals is shown, this appears as pulse disturbance on these intervals at all channels.

The obtained results allow to reduce the demands on 2–4 dB to dynamic range and linearity of amplitude characteristic of multichannel systems apparatus with OFDM

**Keywords:** multiplexing, average power, peak power, peak-to-average ratio, threshold, distribution, baseband signal.

## REFERENCES

1. Van Nee R., Prasad R. OFDM for wireless Multimedia Communication. Boston, Artech House, London, 2000, 257 p.
2. Doelz M. L., Heald E. T., Martin D. L. Binary Data Transmission Techniques for linear System, *Proc. JRE*, 1957, vol. 45, May, pp. 656–661.
3. Mosier R. R., Clabaugh R. G. Kineplex, a bandwidth efficient binary Transmission System, *IEEE Trans.*, 1958, Vol. 76, Jan., pp. 723–728.
4. Patent No. 3, U. S. 488,4555, filed November 14, Orthogonal Frequency Division Multiplexing, 1966, issued Jan. 6, 1970.
5. Weinstein S. B., Ebert P. M. Data Transmission by Frequency Division Multiplexing using the Discrete Fourier Transform, *IEEE Trans Comm.*, Vol. Com – 19, 1971, Oct., pp. 628–637.
6. Schulze H., Luders Ch. Theory and Application of OFDM and CDMA widebound wireless communication. John Wiley and Sons, 2005, 420 p.
7. Vishnevskij V. M., Lyaxov A. I., Portnoj S. L., Shaxnovich I. V. Shirokopolosny'e brsprovodny'e seti peredachi informaczii. Moscow, Texnosfera, 2005, 592 p.
8. Andrews J. G., Ghosh Arunabha, Muhamed R. Fundamentals of WiMAX: understanding broadband wireless networking. Pearson Education, 2007, 439 p.
9. Holma H., Toskala A. LTE for UMTS: OFDMA and SC-FDMA Based Radio Access. John Wiley and Sons, Ltd, 2009, 433 p.
10. Vishnevskij V. M., Portnoj S. L., Shaxnovich I. V. E'ncziklopediya WIMAX. Put' k 4G. Moscow, Texnosfera, 2009, 272 p.
11. Van Nee R. , A. de Wild Reducing the peak-to-average power ratio of OFDM, In *Proc. IEEE Vehicular Technology Conf. (VTC'98)*, 1998, May, pp. 2072–2076.
12. Hideki O., Hideki J. On the Distribution of the Peak to Average Power Ratio in OFDM Signals, *IEEE Trans on Commun*, 2001, Vol. 49, No. 2, February, pp. 282–289.

УДК 621.396.677

Зіньковський Ю. Ф.<sup>1</sup>, Сидорук Ю. К.<sup>2</sup>, Туровський А. О.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, професор, Національний технічний університет України «Київський політехнічний інститут», Україна

<sup>2</sup>Доцент, Національний технічний університет України «Київський політехнічний інститут», Україна

<sup>3</sup>Асистент, Національний технічний університет України «Київський політехнічний інститут», Україна

## КОНСТРУКЦІЙНА ОПТИМІЗАЦІЯ ХВИЛЕВОДНО-ВІБРАТОРНОГО ВИПРОМІНЮВАЧА З ПОВЕРНУТОЮ ПЛОЩИНОЮ ПОЛЯРИЗАЦІЇ

Запропоновано конструкцію 5-елементного випромінювача на основі лінійної вібраторно-хвилеводної решітки для використання у складі пристрою опромінення діелектричних матеріалів електромагнітним полем частотою 2,45 ГГц, розраховано основні параметри для досягнення високої рівномірності розподілу поля перед випромінювачем. Показано вплив відстані між вібраторами і поверхнею хвилевода на ширину діаграми розподілу поля, визначено оптимальне розташування зони для розміщення діелектричного матеріалу, що підлягає опроміненню. Для конструкції з оптимізованими параметрами виконано оцінку поляризаційної та амплітудної однорідності поля.

**Ключові слова:** вібратор, хвилевід, близня зона, випромінювач, надвисокі частоти, поляризація.

### ВСТУП

До основних типів хвилеводних антенних решіток відносяться хвилеводно-щільні та хвилеводно-вібраторні відповідно до типу елементів випромінювання, що у них використовуються. Їх перевага порівно з рупорними аналогами полягає у менших розмірах, що особливо важливо для побудови плоских антенних решіток, також з'являється можливість створення розподіленого випромінювача, що живиться від одного джерела. Частіше у діапазоні НВЧ такі антени створюються з використанням мікросмужкової технології, однак для призначень, де має місце випромінювання електромагнітної енергії високої потужності, смужкові лінії стають малопридатними внаслідок виділення у них значної кількості теплоти.

Завдяки кращій технологічності і компактності лінійні решітки переважно створюються на основі щілин, прорізаних у стінці хвилеводу [1, 2]. Водночас лінійні решітки з випромінювачами у вигляді вібраторів, що живляться від прямокутного хвилеводу залишаються малодослідженими.

Хвилеводні антенні решітки знаходять застосування основним чином для спрямованих антен зв'язку або радіолокації. В той же час лише в окремих роботах розглядається близьке поле випромінювання таких антен, наприклад, з точки зору їх застосування для діелектричного нагріву матеріалів [3], генерування плазми [4] тощо.

### 1 ПОСТАНОВКА ПРОБЛЕМИ

Мета роботи полягає у розробленні та дослідженні лінійної решітки випромінювачів для використання у складі мікрохвильового пристрою діелектричного нагріву, опис якого приведено у [6]. Пристрій сформований циліндричною камерою взаємодії, через яку проходить потік речовини, що підлягає опроміненню електромагнітним полем надвисоких частот. Концентрично

навколо камери паралельно її осі розташовані лінійні випромінювачі, що дає змогу забезпечити достатній рівень щільності енергії завдяки просторовому складанню хвиль від генераторів невисокої потужності. Для зниження рівня взаємодії між протилежно розташованими джерелами циліндрична камера оточена спіралеподібним поляризованим дзеркалом, вісь поляризації якого створює кут  $45^\circ$  до основи камери. Таким чином, електромагнітна енергія, що проходить у зону взаємодії з одного боку буде відбиватися від протилежної сторони внаслідок перпендикулярної орієнтації осі поляризації дзеркала з протилежніх сторін. Відповідно, джерела електромагнітної енергії повинні виконуватися у вигляді лінійних решіток, і, при цьому, площа поляризації поля повинна бути повернута під кутом  $45^\circ$  до осі решітки. Крім того, технічні характеристики камери оброблення вимагають забезпечення високої рівномірності напруженості електричного поля в зоні взаємодії.

Дослідження виконувались за допомогою засобів комп'ютерного моделювання електродинамічних структур Ansys HFSS™ і CST MWS™.

### 2 МАТЕРІАЛИ ТА МЕТОДИ

Відомі конструкції щілинних решіток, у яких щілини розміщені в центрі широкої стінки прямокутного хвилевода під кутом до його осі і орієнтовані почергово в протилежні сторони [4, 2]. В таких антенах повертання щілин навколо центру виконується для забезпечення необхідного їх збудження і узгодження з хвилеводом, однак внаслідок симетричності кутів повороту у дальній зоні формується хвиля, поляризована паралельно осі хвилевода. Можлива конструкція хвилеводної антени, яка забезпечує повернуту, відносно осі хвилевода, площину поляризації, у якій щілини орієнтовані паралельно між собою і утворюють певний кут відносно осі хвилевода [2].

В такому випадку щілини розміщаються вздовж широкій стінки хвилевода зі зміщенням відносно його центру почергово у різні сторони (так звані «складні» щілини). Така реалізація забезпечує повернуту поляризацію хвилі у дальній зоні, однак, внаслідок нерівномірного розташування апертур щілин напруженість поля поблизу випромінювача розподіляється нерівномірно. Важливим недоліком щілинних антен є неможливість або значна складність регулювання збудження окремих щілин.

Альтернативою використання щілинних випромінювачів, що дозволить уникнути властивих їм недоліків, є використання випромінюючих елементів у вигляді вібраторів, що живляться від хвилевода. Внаслідок того, що випромінювачі не є інтегрованими у хвилевід це дає змогу регулювати збудження кожного елемента, змінювати кут поляризації випромінювання.

У запропонованій конструкції використаний випромінювач хвилеводно-вібраторного типу (рис. 1) [7], основу якої складає прямокутний хвилевід 4 (рис. 1), що з одного кінця закритий рухомим короткозамикаючим поршнем 5, а іншим кінцем підключається до генератора. Елементи випромінювання енергії утворені вібраторними секціями, розташованими вздовж хвилевода в центрі широкої стінки. Кожна секція складається з відрізка коаксіальної лінії, що з одної сторони входить у хвилевід штирем 3, а з іншої, через щілинний симетричуючий пристрій 2, живить плечі півхвильового вібратора 1. Відбір енергії від хвилевода виконується штирем 3, що є продовженням центрального провідника коаксіальної лінії. Забезпечення повернутого відносно осі хвилевода площини поляризації поля здійснено завдяки повернутому положенню усіх вібраторів на відповідний кут до осі хвилевода, що складає  $45^\circ$ .

Для узгодження хвилеводної решітки із генератором використано реактивний вібратор 6, розташований на відстані  $D_p$  перед першим штирем в центрі широкої стінки і має довжину  $L_p$ , реактивність якого компенсує реактивність усіх штирів.

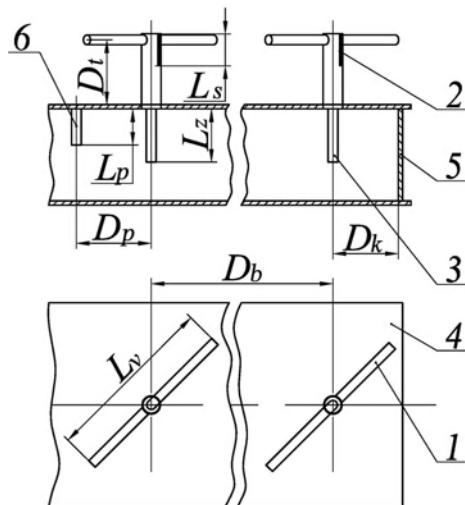


Рисунок 1 – Зовнішній вигляд випромінювача з розрізом

Робоча частота пристрою становить 2,45 ГГц (довжина хвилі  $\lambda_0 \approx 122$  мм), відповідно, обрано тип хвилевода WR-340 (стандарт EIA) з внутрішніми розмірами  $86 \times 43$  мм і товщиною стінки 2 мм.

Для забезпечення високої рівномірності поля поблизу випромінювача всі елементи повинні випромінювати енергію синфазно, а амплітуди збудження кожної секції мають бути розрахованими таким чином, щоб забезпечити максимальну однорідність напруженості електромагнітного поля у заданій зоні; одночасно з цим відстань між вібраторами повинна бути якомога меншою. Мінімальна відстань  $L_b$  між сусідніми елементами може бути досягнута, якщо встановити її приблизно  $\lambda_g/2$  (відповідає довжині хвилі у хвилеводі) з урахуванням поправки на внесену реактивність штирів для компенсування фазового зсуву, однак, при цьому струми у сусідніх коаксіальних відрізках будуть протифазними. Відповідно, для збереження синфазності випромінювання сусідні вібратори взаємно розвернуті на  $180^\circ$ .

Довжина кожного штиря  $L_z$  визначається відповідно до необхідної величини збудження вібратора, що може також регулюватись шляхом зміщення штирів убік від центру або вздовж хвилевода. Останнє можливе завдяки тому, що випромінювач працює у режимі змішаних хвиль і вздовж хвилевода має місце значний коефіцієнт стоячої хвилі по напрязі (КСХн) – положення штирів на по-вздовжній осі впливає на інтенсивність їх збудження за-лежно від того чи він знаходиться у максимумі, чи в мінімумі поля. При цьому необхідно щоб розміщення усіх штирів забезпечувало синфазність їх збудження.

Одною з умов оптимізації пристрою було забезпечення максимального узгодження коаксіальних ліній із вібраторними секціями, і, таким чином, розрахунок всього пристрою було розділено на внутрішню та зовнішню задачі.

На першому етапі, в ході вирішення зовнішньої задачі, задавались значення відстаней між сусідніми випромінюючими секціями, визначалися оптимальні параметри вібраторів і симетричних елементів, необхідні амплітуди їх збудження, які забезпечать узгодження кожного вібратора з коаксіальними відрізками, створюючи при цьому рівномірний розподіл поля в області перед ними.

Внутрішня задача полягала у розрахунку прямокутного хвилевода, в центрі широкої стінки якого розташовано ряд хвилеводно-коаксіальних переходів. Розраховувались оптимальні значення відстаней між штирами переходів і довжини кожного з них для забезпечення необхідної амплітуди і синфазності збудження кожного коаксіального відрізу. Одночасно із оптимізацією параметрів штирів відбувався розрахунок розмірів і положення реактивного вібратора для компенсації відбитої від хвилевода хвилі.

### 3 РЕЗУЛЬТАТИ

Характер розподілу поля вздовж хвилевода над вібраторами залежить основним чином від їх кількості, відстані між ними, амплітуди і фази збудження. Розподіл поля у поперечному напрямі визначається більшою мірою па-

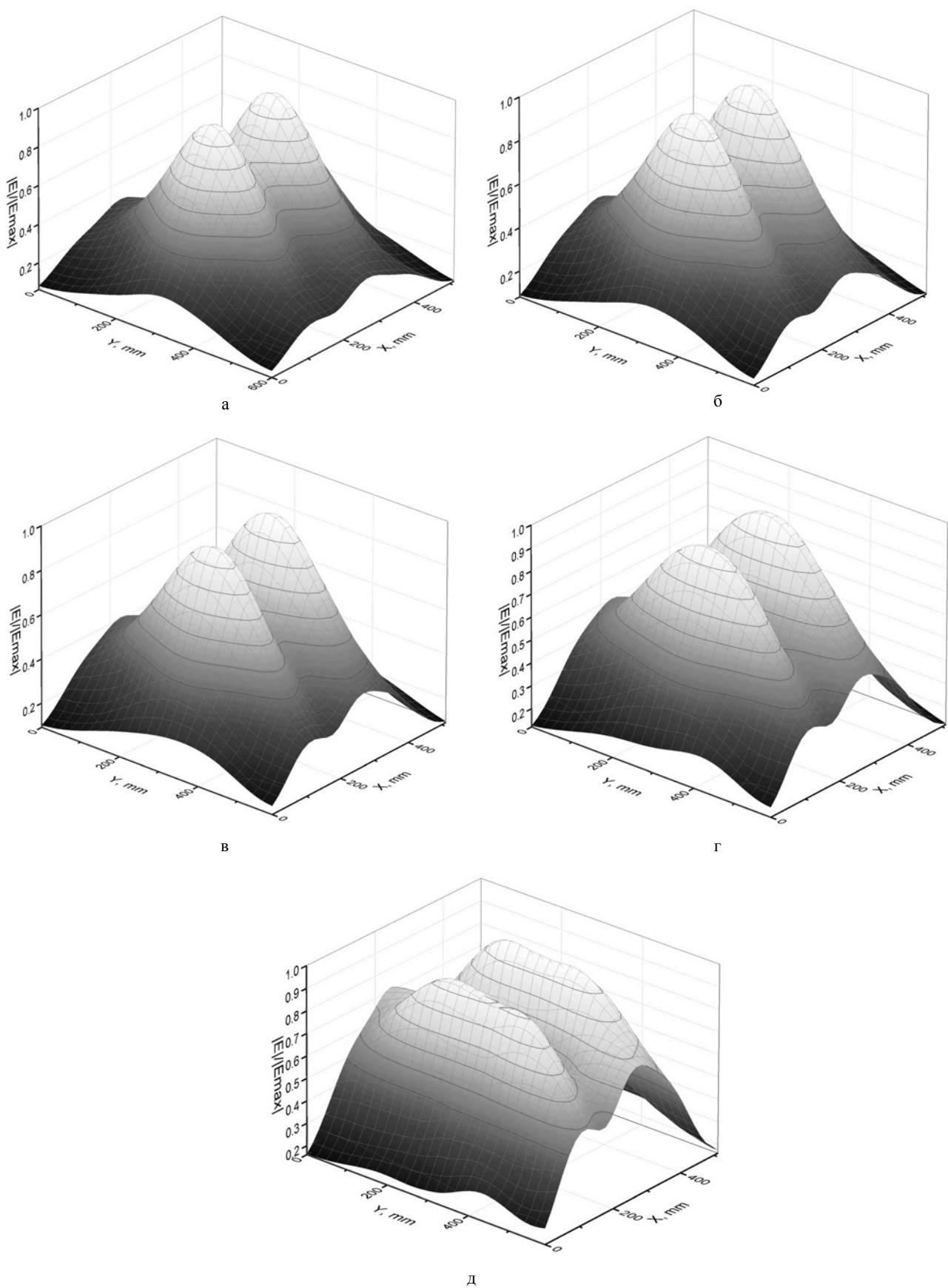


Рисунок 2 – Нормований модуль напруженості електричного поля: а –  $D_t=20$  мм; б –  $D_t=30$  мм; в –  $D_t=40$  мм; г –  $D_t=50$  мм; д –  $D_t=60$  мм

раметрами вібраторів та відстанню від них до поверхні хвилеводу  $D_r$ . Так, при резонансній довжині вібратора і відстані  $D_r$ , близькій до половини довжини хвилі діаграма розподілу поля у ближній зоні розширяється, а при чверті довжини хвилі – звужується. Останнє твердження пояснюється тим, що при відстані від вібратора до хвилевода приблизно рівній  $\lambda_0/4$ , хвилі у перпендикулярному до хвилевода напрямі синфазно складаються із відбитими від хвилевода, внаслідок чого інтенсивність випромінювання енергії у перпендикулярному до хвилевода напрямі зростає. Якщо ж  $D_r$  наближається до величини  $\lambda_0/2$ , то відбиті від металової поверхні хвилі виявляються протифазними і потік енергії у перпендикулярному напрямі зменшується, призводячи до зростання потоку енергії у бокових напрямах.

Для конструкції мікрохвильового пристроя, складовою якого є випромінювач, необхідно забезпечувати вузьку діаграму розподілу напруженості електричного поля у поперечному напрямі – у протилежному випадку хвилі, що випромінюються під значним кутом до нормалі, матимуть значну поляризаційну неузгодженість із поляризованим дзеркалом.

Необхідна кількість вібраторів визначається виходячи із потужності джерела, що живить хвилевід, необхідної щільноти потужності поля у зоні взаємодії і при цьому їх кількість повинна бути достатньою для забезпечення перед випромінювачем зони рівномірного розподілу електромагнітного поля у поздовжньому, відносно хвилевода, напрямі.

У роботі розглядався випромінювач, що складається з п'яти вібраторних елементів.

Шляхом оптимізації параметрів вібраторних секцій, у яких діаметр плечей становив 2 мм, в межах розв'язання зовнішньої задачі, визначено, що для забезпечення величини КСХ на рівні не більше 1,1 довжина усіх вібраторів  $L_v$  повинна становити 50 мм, довжина симетричних щілин  $L_s$  – 23 мм, діаметр внутрішнього провідника та хвильовий опір коаксіальної лінії становив, відповідно, 3 мм та 50 Ом. При цьому, незважаючи на те, що взаємодія центрального вібратора з іншими сильніша, ніж крайніх, розбіжність рівня узгодження кожного з них при однакових параметрах залишається в межах вказаного КСХн.

Важливою проблемою при проектуванні решітки випромінювачів є визначення оптимального розташування зони взаємодії, в межах якої буде забезпечена максимальна рівномірність параметрів поля, і, зокрема його максимальна поляризаційна однорідність. Останній показник важливий для поляризаційного узгодження електромагнітного поля з поляризованим дзеркалом.

Нормований розподіл модуля напруженості електричного поля у площині на відстані  $h = 150$  мм від вібраторів, при різних відстанях від них до хвилевода, зображене на рис. 2, де координата  $x$  орієнтована паралельно, а  $y$  – перпендикулярно осі хвилевода. Збудження вібраторів відбувалось одинаковими за амплітудами і фазами струмами. З рисунків видно, що з віддаленням вібраторів від хвилевода ширина діаграми розподілу поля розширяється, що відповідає вказаному вище твердженню.

При цьому висота розташування площини вібраторів над хвилеводом є основним чинником, що визначає розподіл кута поляризації поля. На рис. 3 приведені характеристики співвідношення між складовою електричного поля паралельною до вібраторів і модулем напруженості електричного поля при різних значеннях  $D_r$  і  $h = 150$  мм, які характеризують ступінь відповідності поляризації поля і поляризованого дзеркала, розміщеного над випромінювачем, у якого вісь поляризації перпендикулярна до вібраторів. Характеристики розраховані у площині, перпендикулярній осі  $x$  в точці максимальної напруженості поля. Очевидно, що однорідність поляризації хвилі зростає з наближенням площини вібраторів до металової площини.

Таким чином, з огляду на покращення амплітудної і поляризаційної однорідності, необхідно встановлювати найменшу висоту розташування вібраторів над металовою основою, яка, однак, обмежена довжиною симетричної щілини, і для даного випадку становить  $D_r = 22$  мм.

Забезпечуючи відповідний розподіл інтенсивності збудження кожного з вібраторів можна змінювати характеристику розподілу напруженості поля у ближній зоні.

Шляхом оптимізації амплітуд збудження кожного елемента було встановлено, що при співвідношенні амплітуд живлення вібраторів як 5:14:10:14:7, відповідно (починаючи від входу), при однакових фазах, досягнуто максимальної рівномірності поля вздовж осі хвилевода. Для цього випадку розподіл напруженості поля, нормований до максимального значення у різних перетинах над випромінювачем з оптимізованими параметрами, зображені на рис. 4. З приведених залежностей можна зробити висновок, що оптимальна зона розташування середовища для опромінення знаходиться на відстані приблизно  $h = 100$  мм від вібраторів. Рівень складової електричного поля, паралельної до плечей вібраторів, для даного випадку, відносно загального модуля напруженості електричного поля, приведено на рис. 5.

В результаті оптимізації були розраховані оптимальні параметри конструкції випромінювача. Визначено, що довжини  $L_s$  штирів для забезпечення зазначених раніше амплітуд збудження вібраторів повинні становити (по-

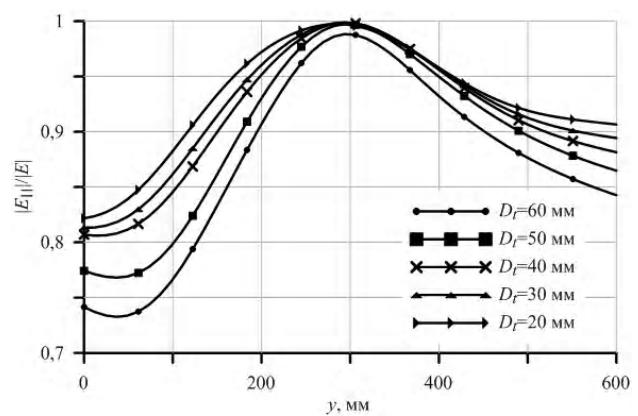


Рисунок 3 – Коефіцієнт поляризаційної однорідності електричного поля

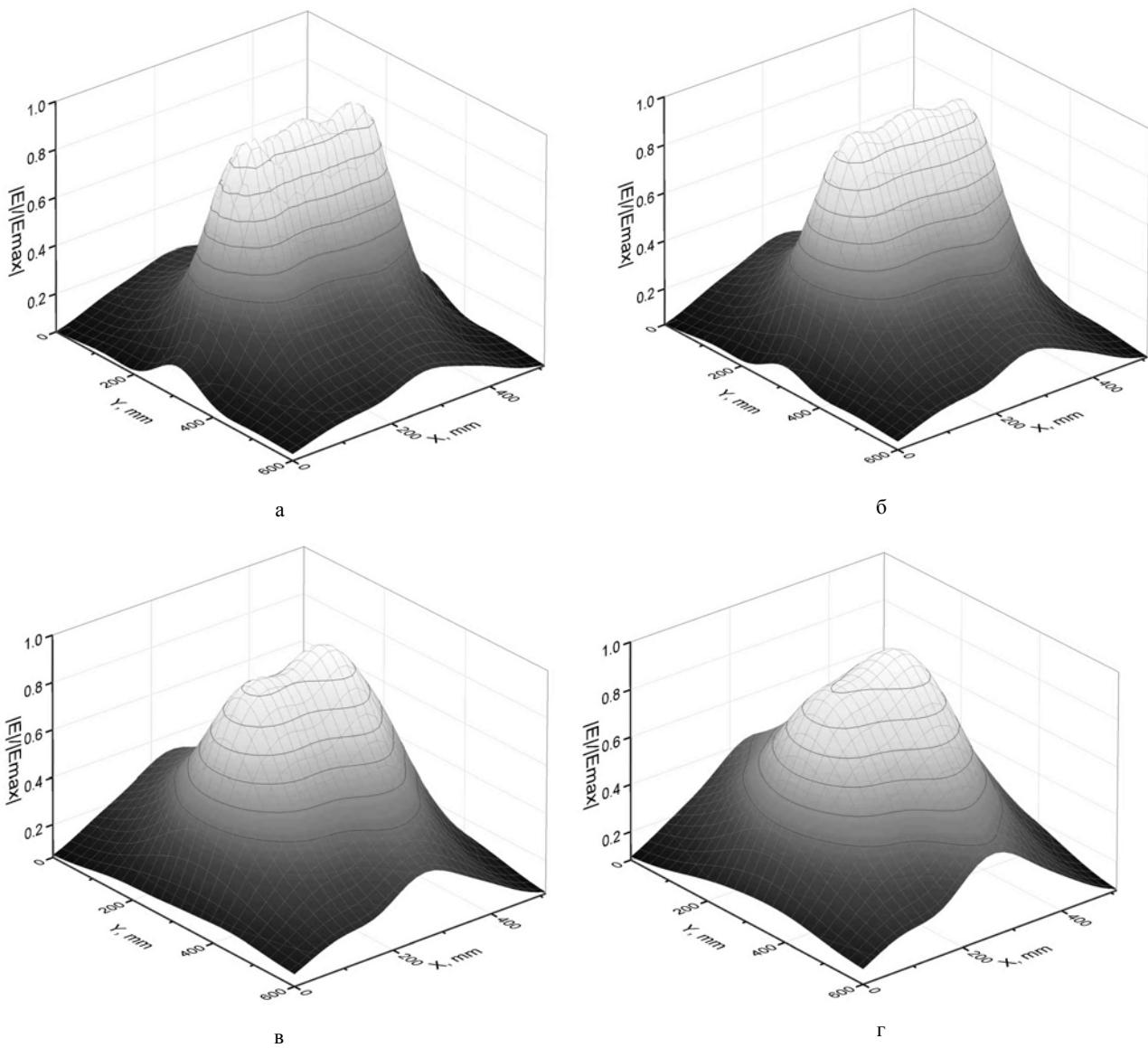


Рисунок 4 – Нормований модуль напруженості електричного поля а –  $D_t = 20$  мм на відстані  $h = 75$  мм; б –  $h = 100$  мм; в –  $h = 150$  мм; г –  $h = 200$  мм від хвилевода

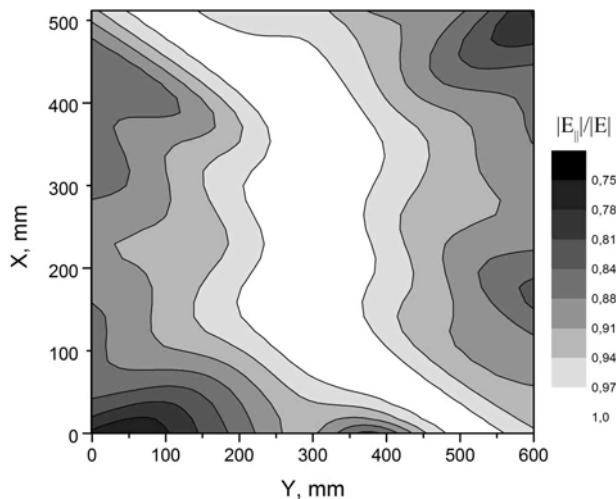


Рисунок 5 – Коефіцієнт поляризаційної однорідності поля при  $h=100$  мм

чинаючи від входу) 13,3 мм, 16 мм, 17,8 мм, 17,6 мм та 17,3 мм, відповідно; відстань між центрами випромінювальних секцій  $D_b$ , для синфазності їх збудження дорівнюють, відповідно, 92 мм, 79 мм, 88 мм, 87 мм.

З метою кількісної оцінки поляризаційної однорідності поля розрахованій показник, що характеризує відносну величину випроміненої енергії, що пройде крізь ідеальне плоске поляризоване дзеркало, розташоване паралельно поверхні хвилевода, вісь поляризації якого орієнтована перпендикулярно осі вібраторів:

$$P = \frac{\iint E_{\parallel}^2(x, y) dx dy}{\iint E^2(x, y) dx dy}. \quad (1)$$

Розрахувавши (1), підставляючи значення, що відповідають рис. 4б та рис. 5, було отримано  $P \approx 0,9$ .

За таких параметрів конструкції було забезпеченено КСХн у всіх коаксіальних відрізках не більше 1,05, а на вході хвилевода не більше 1,1.

## ВИСНОВКИ

Повертання площини поляризації поля у біжній зоні випромінювача можна забезпечити відповідним орієнтуванням вібраторів відносно осі хвилевода. Показано, що для звуження діаграми розподілу поля вібратори повинні розташуватись на мінімальній відстані від поверхні хвилевода. Випромінювач дає змогу забезпечити високий рівень однорідності напруженості електричного поля у напрямі, паралельному осі випромінювача. Для розглянутої конструкції, що складається із 5 вібраторних секцій, повернутих на кут  $45^\circ$  до осі хвилевода, оптимальна зона розташування матеріалу для опромінення знаходиться на відстані 100 мм від площини вібраторів.

## СПИСОК ЛІТЕРАТУРИ

- Orefice M. Design of waveguide-fed series slot arrays / M. Orefice, R. S. Elliott // IEE Proc. – 1982. – Vol. 129, No. 4. – P. 165–169.

Зиньковський Ю. Ф.<sup>1</sup>, Сидорук Ю. К.<sup>2</sup>, Турівський А. А.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, професор, Національний технічний університет України «Київський політехнічний інститут», Україна

<sup>2</sup>Доцент, Національний технічний університет України «Київський політехнічний інститут», Україна

<sup>3</sup>Асистент, Національний технічний університет України «Київський політехнічний інститут», Україна

## КОНСТРУКЦІОННА ОПТИМИЗАЦІЯ ВОЛНОВОДНО-ВІБРАТОРНОГО ИЗЛУЧАТЕЛЯ С ПОВЕРНУТОЙ ПЛОСКОСТЬЮ ПОЛЯРИЗАЦІЇ

Предложена конструкция 5-элементного излучателя на основе линейной вибраторно-волноводной решетки для использования в составе устройства облучения диэлектрических материалов электромагнитным полем частотой 2,45 ГГц, рассчитаны основные его параметры для достижения максимальной равномерности распределения поля перед излучателем. Показано влияние расстояния между вибраторами и поверхностью волновода на ширину диаграммы распределения поля, определено оптимальное расположение зоны для размещения диэлектрического материала, подлежащего облучению. Для конструкции с оптимизированными параметрами выполнена оценка поляризационной однородности поля.

**Ключевые слова:** вибратор, волновод, ближняя зона, излучатель, сверхвысокие частоты, поляризация.

Zinkovskiy Yu. F.<sup>1</sup>, Sydoruk Yu. K.<sup>2</sup>, Turovskiy A. O.<sup>3</sup>

<sup>1</sup>Doctor of Science, Professor, National technical university of Ukraine «Kyiv polytechnic institute», Ukraine

<sup>2</sup>Associate Professor, National technical university of Ukraine «Kyiv polytechnic institute», Ukraine

<sup>3</sup>Assistant, National technical university of Ukraine «Kyiv polytechnic institute», Ukraine

## DESIGN OPTIMIZATION OF WAVEGUIDE FED DIPOLE APPLICATOR WITH INCLINED POLARIZATION PLANE

A design of microwave waveguide fed dipole applicator consisted of 5 elements for use in a device for bulk dielectric materials treatment with microwave electromagnetic field is proposed. In order to provide maximum amplitude and polarization homogeneity of the field distribution before the transmitter the main parameters of the design are calculated. Due to the dipoles being fully matched with feeding coaxial line segments the calculation of the whole structure is divided into internal and external problems. The influence of distance between the dipoles and waveguide's surface on the width of the near field distribution pattern is shown and it is clarified that the reducing the distance will narrowing the field pattern. It was determined that the optimal zone location for placing of the bulk dielectric material to be irradiated is at the distance 100–150 mm from the dipoles.

**Keywords:** vibrator, waveguide, near field, irradiator, microwaves, polarization.

## REFERENCES

- Orefice M., Elliott R. S. Design of waveguide-fed series slot arrays, *IEE Proc.*, 1982, Vol.129, No. 4, pp.165–169.
- Yamaguchi S., Miyashita H., Takahashi T., et. al. Design of inclined and displaced slotted waveguide array antennas with low sidelobe radiation patterns , *Proc. of the 4<sup>th</sup> EuCAP*, 2010, pp. 1–5.
- St-Denis E., Raghavan G., Akyel C., Venkatesh M. A slotted waveguide applicator for continuous flow grain drying, *Journal of microwave power and electromagnetic energy*, 2001, Vol. 36, No. 1, pp. 3–16.
- Design of inclined and displaced slotted waveguide array antennas with low sidelobe radiation patterns / [S. Yamaguchi, H. Miyashita, T. Takahashi, et. al.] // Proceedings of the 4<sup>th</sup> EuCAP. – 2010. – P. 1–5.
- A slotted waveguide applicator for continuous flow grain drying / [E. St-Denis, G. Raghavan, C. Akyel, M. Venkatesh] // *Journal of microwave power and electromagnetic energy*. – 2001. – Vol. 36, No. 1. – P. 3–16.
- Sauve G. Slotted waveguide field applicator for the generation of long uniform plasmas / G. Sauve, M. Moisan, Z. Zakrewski // *Journal of microwave power and electromagnetic energy*. – 1993. – Vol. 28, No. 3. – P. 123–131.
- Meredith R. Engineer's handbook of industrial microwave heating / R. Meredith. – London : IEE Publishing, 1998. – 365 p.
- Пат. Україна 65629 Н05В 6/64. Мікрохвильовий пристрій для передпосівної обробки насіння, сушиння зерна та інших сипучих матеріалів / Сидорук Ю. К.; заяв. 20.05.2011, опубл. 12.12.2011, Бюл. № 23, 2011 р.
- Жук М. С. Проектирование антенно-фидерных устройств / М. С. Жук, Ю. Б. Молочков. – М.-Л. : Энергия, 1966. – 648 с.

Стаття надійшла до редакції 02.04.2014.

УДК 621.396.96

Зубков А. М.<sup>1</sup>, Щерба А. А.<sup>2</sup>

<sup>1</sup>Д-р техн. наук, старший науковий співробітник, провідний науковий співробітник науково-дослідного відділу (ракетних військ та артилерії) Наукового центру Сухопутних військ Академії сухопутних військ імені гетьмана Петра Сагайдачного, Львів, Україна

<sup>2</sup>Ад'юнкт штатний науково-організаційного відділу Академії сухопутних військ імені гетьмана Петра Сагайдачного, Львів, Україна

## ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ АРТИЛЕРІЙСЬКОЇ РОЗВІДКИ ШЛЯХОМ КОНСТРУКТИВНО-ФУНКЦІОНАЛЬНОЇ ІНТЕГРАЦІЇ ПОВІТРЯНИХ ТА НАЗЕМНИХ ЗАСОБІВ СПОСТЕРЕЖЕННЯ

Запропонований та обґрутований метод багатоканальної розвідки рухомих і нерухомих наземних цілей, який знімає обмеження радіогоризонту, інваріантний до типу цілей та оснований на інтеграції відомих наземних та повітряних засобів спостереження в рамках єдиного комплексу інструментальної розвідки.

**Ключові слова:** радіолокаційна станція розвідки наземних рухомих цілей, радіолокаційний комплекс розвідки вогневих позицій, дистанційно пілотований літальний апарат, фазована антenna решітка, мультиспектральна система спостереження.

### НОМЕНКЛАТУРА

ДПЛА – дистанційно пілотований літальний апарат;

ЕПР – ефективна поверхня розсіювання;

РЛК РВП – радіолокаційний комплекс розвідки вогневих позицій;

РЛС РНРЦ – радіолокаційна станція розвідки наземних рухомих цілей;

ФАР – фазована антenna решітка;

$A$  – виграш у дальності артилерійської розвідки;

$C$  – порогове значення;

$D_k$  – ймовірність правильного виявлення у  $k$ -му парціальному каналі;

$D_{\max}$  – дальність прямої видимості;

$F_k$  – ймовірність хибної тривоги у  $k$ -му парціальному каналі;

$H_A$  – висота антени;

$H_{\Pi}$  – висота цілі;

$H_{\text{ДПЛА}}$  – висота ДПЛА;

$h_k$  – ваговий коефіцієнт, який характеризує інформаційний внесок  $k$ -го парціального каналу у виявленні спостережуваної цілі;

$l_{\text{ДПЛА}}$  – відстань до ДПЛА;

$l_{\Pi}$  – відстань до цілі;

$y_k$  –  $l$ -мірний вектор часткових рішень;

$\sigma_{\text{ДПЛА}}$  – ЕПР ДПЛА;

$\sigma_{\text{БП}}$  – ЕПР боеприпасів стріляючої системи.

### ВСТУП

Розвідка наземних цілей в інтересах ракетних військ і артилерії є першим і найважливішим етапом циклу вогневого ураження противника. При цьому необхідно досягнути високої достовірності і точності розвідки цілодобово, за несприятливих погодних умов та наявності завад природного і штучного походження. Основними технічними засобами вирішення цього спектру задач є радіолокаційні засоби. Однак, суттєвим обмеженням ефективності існуючих наземних радіолокаційних засобів

артилерійської розвідки, які ґрунтуються на прямому електромагнітному контакті з ціллю, є [1–2]:

- одноканальність;
- недостатня дальність дії через обмеження радіогоризонту;
- практична працездатність тільки для рухомих цілей (звідси класифікація: РЛС РНРЦ).

Наземні радіолокаційні засоби артилерійської розвідки, які ґрунтуються на дотичному методі визначення місцезнаходження стріляючих наземних об'єктів (радіолокаційні комплекси розвідки вогневих позицій – РЛК РВП) [3–4], забезпечуючи цільову багатоканальність, мають недостатню дальність дії через енергетичні обмеження, пов'язані з малою ЕПР боеприпасів стріляючих систем.

Зняття обмеження радіогоризонту забезпечують засоби спостереження, які встановлюються на ДПЛА [5]. Однак, існуючі схеми доведення розвідувальної інформації з ДПЛА до вогневих засобів характеризуються значною часовою затримкою, що недопустимо для високодинамічних бойових дій.

Тому актуальним є пошук методів і способів забезпечення високоточної артилерійської розвідки на граничних дальностях дії озброєння з одночасним підвищеннем цільової багатоканальністі в діапазоні всіх дальностей ефективного його застосування. При цьому, принципове значення набуває використання потенційних інформаційних можливостей каналів спостереження різних ділянок спектру електромагнітних хвиль [6] шляхом їх конструктивно-функціональної інтеграції.

### 1 ПОСТАНОВКА ЗАДАЧІ

У діапазонах спектру електромагнітних хвиль, які використовуються для розвідки наземних цілей, гранична дальність виявлення, вимірювання координат і розпізнавання цілей  $D_{\max}$  визначається дальністю прямої видимості, яка з урахуванням кривизни земної поверхні складає [7]

$$D_{\max} = 4,12(\sqrt{H_A} + \sqrt{H_{\Pi}}) [\text{км}], \quad (1)$$

де  $H_A$  – висота антени [м],  $H_{\Pi}$  – висота цілі [м].

У радіолокаційному каналі навіть за наявності прямого оптичного контакту з ціллю, максимальна дальльність її спостереження визначається ЕПР. При цьому, діапазон значень типових ЕПР наземних цілей (танк, БМП, БТР, ПТРК, пускові установки тактичних і оперативно-тактичних ракет) коливаються в межах 10–50 м<sup>2</sup>. У РЛК РВП робота по боєприпасам з малою ЕПР ( $\leq 0,1$  м<sup>2</sup>) забезпечується (також в межах прямої оптичної видимості) за рахунок значного енергетичного потенціалу.

Метою статті є обґрунтування методу збільшення дальності та підвищення достовірності багатоцільової артилерійської розвідки, інваріантного до типу і характеристик руху наземних спостережних об'єктів.

## 2 ОГЛЯД ЛІТЕРАТУРИ

Відомі способи збільшення радіогоризонту під час спостереження наземних цілей на основі стіпвідношення (1), які базуються на використанні підйомно-мачтових систем для антенних і приймально-передавальних пристроїв РЛС [8]. Однак, застосування такого підходу суттєво підвищує складність, габаритно-вагові характеристики конструкції і, як наслідок, вартість РЛС РНРЦ з одночасним зниженням її маневреності та скритості. З іншої сторони, відомі способи зняття обмежень радіогоризонту за рахунок застосування ДПЛА зі встановленою на борту апаратурою спостереження наземних цілей і відповідною апаратурою прийому інформації на борту бойової розвідувальної машини, оснащеної комплексом інших засобів інструментальної розвідки (радіолокаційної, теплової, оптичної) [9]. Однак, відокремленість апертурних частин наземних і бортових каналів спостереження не дозволяє реалізувати потенційні можливості артилерійської розвідки у динамічно мінливій фоновоцільовій обстановці на всю глибину по дальності.

У роботах [10–13] показано, що комплексування каналів спостереження різних ділянок спектру електромагнітних хвиль в рамках загальної інформаційно-вимірювальної системи є потужним інструментом підвищення достовірності і точності дистанційного моніторингу наземних об'єктів. У роботі [6] отримані алгоритми багатоспектральної обробки інформації (оптичний + тепловий + радіолокаційний канали) з урахуванням просторово-часового взаємостворення та взаємосинхронізації парціальних спектральних каналів на різних етапах локаційного спостереження (виявлення, оцінка координат, розпізнавання). У роботах [14–15] розроблена методологія і розглянуті варіанти конструктивного комплексування (інтеграції) парціальних спектральних каналів спостереження для забезпечення взаємоузгодженості роботи у просторі та часі. Однак, різка залежність потенціальних можливостей парціальних спектральних каналів оптичного і теплового діапазонів, яка пов'язана зі станом приземного шару атмосфери (погода, пило-димові перешкоди, час доби), обмежує дальльність їх ефективного застосування при установці апаратури на наземних об'єктах, у той час як канали радіодіапазону практично позбавлені цього недоліку.

## 3 МАТЕРІАЛИ ТА МЕТОДИ

Основною науково-технічною ідеєю запропонованого підходу щодо підвищення ефективності артилерійської розвідки за рахунок розширення діапазону дальностей, цільової багатоканальності і зняття обмежень на характер і параметри руху наземних цілей є поєднання переваг високопотенційних багатоканальних РЛК РВП і засобів повітряної розвідки, встановлених на ДПЛА, в рамках інтегрованого комплексу. При цьому, практична реалізація базується на комплексуванні і конструктивно-функціональній інтеграції відомих технічних рішень [3–5, 8].

Практична реалізація методу передбачає (рис. 1):

- застосування одного або декількох просторових каналів спостереження РЛК РВП, які формуються фазованою антенною решіткою (ФАР) у штатному режимі для багатоцільового супроводження боєприпасів стріляючих засобів противника, в якості каналів супроводження ДПЛА;
- розміщення на борту ДПЛА багатоспектральної апаратури розвідки наземних цілей.

Введемо практично виправдані обмеження:

- ДПЛА перебуває близько до зеніту відносно спостережуваної цілі;
- висота ДПЛА  $H_{\text{ДПЛА}}$  значно менша дальності до спостереженої цілі  $l_{\text{Ц}}$ :

$$H_{\text{ДПЛА}} \ll l_{\text{Ц}}; \quad (2)$$

- дальльність до спостережуваної цілі  $l_{\text{Ц}}$  приблизно дорівнює дальності до ДПЛА  $l_{\text{ДПЛА}}$ :

$$l_{\text{Ц}} \approx l_{\text{ДПЛА}}. \quad (3)$$

Тоді, згідно з основним рівнянням радіолокації [7], виграш у дальності артилерійської розвідки у порівнянні з штатним режимом РВП буде складати:

$$A \approx 4 \sqrt{\frac{\sigma_{\text{ДПЛА}}}{\sigma_{\text{БП}}}}, \quad (4)$$

де  $\sigma_{\text{ДПЛА}}$  – ЕПР ДПЛА,  $\sigma_{\text{БП}}$  – ЕПР боєприпасу стріляючої системи.

Слід також зазначити, що низька (у порівнянні з боєприпасами) динаміка польоту ДПЛА сприяє ефективно-енергетичному накопиченню ехо-сигналів [7].

У випадку розміщення на борту ДПЛА багатоспектральної апаратури розвідки наземних цілей оптимальні алгоритми вирішення локаційних задач виявлення, оцінки координат і розпізнавання наземних цілей, отримані в роботі [6], можуть бути адаптовані. Розглянемо на прикладі алгоритму виявлення. У загальному випадку, при комплексуванні парціальних каналів на рівні прийнятих показанально рішень, оптимальний алгоритм має вигляд [6]:

$$\sum_{k=1}^l h_k y_k \geq C, \quad (5)$$

де  $y_k = (y_1, y_2 \dots y_l)$  –  $l$ -мірний вектор часткових рішень;  $y_k = 1$  – рішення про наявність сигналу, прийняте  $k$ -м каналом;  $y_k = 0$  – рішення про відсутність сигналу, прийня-

те  $k$ -м каналом;  $h_k$  – ваговий коефіцієнт, який характеризує інформаційний внесок  $k$ -го парціального каналу у виявленні спостережуваної цілі:

$$h_k = \ln \left( \frac{D_k \times (1 - F_k)}{F_k \times (1 - D_k)} \right), \quad (6)$$

де  $D_k$  – ймовірність правильного виявлення у  $k$ -му парціальному каналі;  $F_k$  – ймовірність хибної тривоги у  $k$ -му парціальному каналі;  $C$  – порогове значення, яке обирається виходячи із дозволного рівня хибної тривоги  $F$  для багатоспектральної системи в цілому.

Без зменшення спільноті розглянемо випадок, коли у виразі (5)  $l=3$ , причому:

- $k=1$  оптичний (денний) канал;
- $k=2$  інфрачервоний (тепловий) канал;
- $k=3$  радіолокаційний канал.

Тоді, за виконання умов (2)–(3), доцільно отримати у виразі (5):

- $h_1=h_2=0$  за наявності метеоопадів і вночі;

–  $h_3=0$  вночі;

–  $h_3=1$  за сприятливих погодних умов, у тому числі вночі.

Ці співвідношення відображають ту обставину, що на малих відстанях, які відповідають висотам польоту тактичних ДПЛА, канали спостереження наземних цілей можна диференціювати за ефективністю з послідуочим відключенням непрацездатних [16]. Слід зазначити, що запропонована методика поширюється на алгоритми оцінки координат і розпізнавання.

Структуру адаптивного багатоспектрального виявляча наземних цілей ілюструє рис. 2.

Адаптивне автоматичне переключення каналів спостереження може бути реалізовано шляхом використання аналізатора завадової обстановки [17].

На рис. 3 наведена загальна структурна схема реалізації запропонованого методу шляхом доповнення штатної апаратури РЛК РВП.

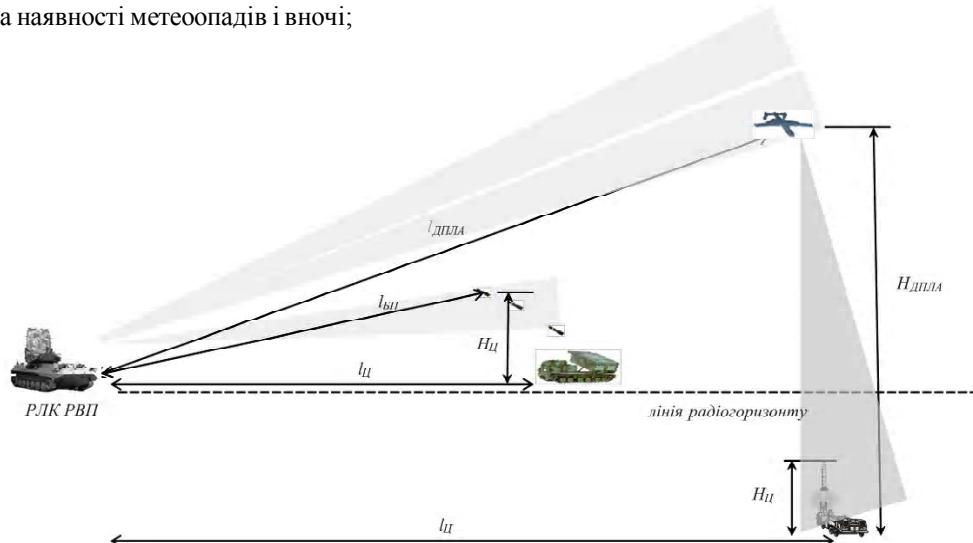


Рисунок 1 – Схема інтеграції штатних цільових каналів РЛК РВП та інформаційних каналів ДПЛА

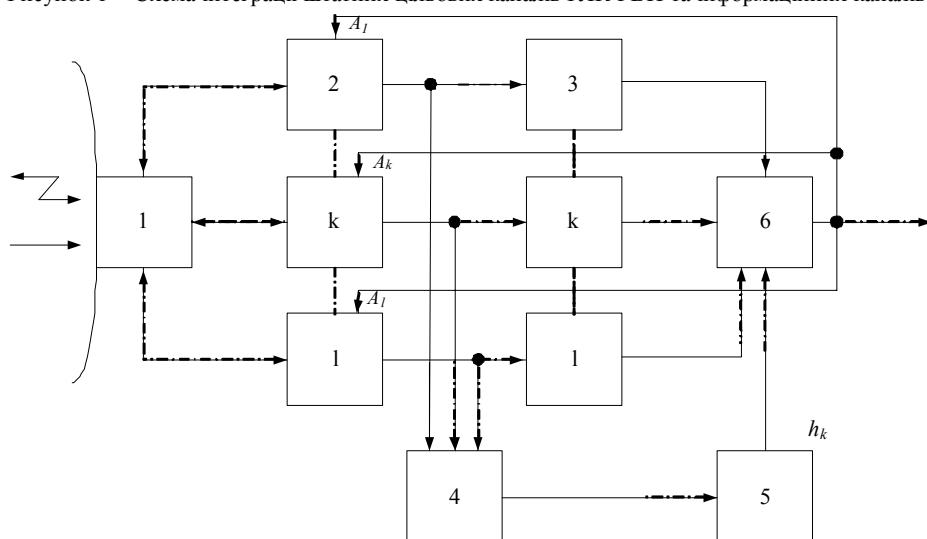


Рисунок 2 – Структурна схема багатоспектрального виявляча: 1 – єдиний діаграмоутворювальний апертурний блок; 2 – приймально-передавальний (приймальний) тракт парціального спектрального каналу; 3 – блок формування вихідних сигналів парціальних каналів; 4 – аналізатор завадової обстановки; 5 – блок розрахунку канальних вагових коефіцієнтів  $h_k$ ; 6 – блок розрахунку вагових сум (5) та прийняття рішень;  $A_1, \dots, A_l$  – сигнали управління адаптацією приймально-передаючих трактів парціальних спектральних каналів

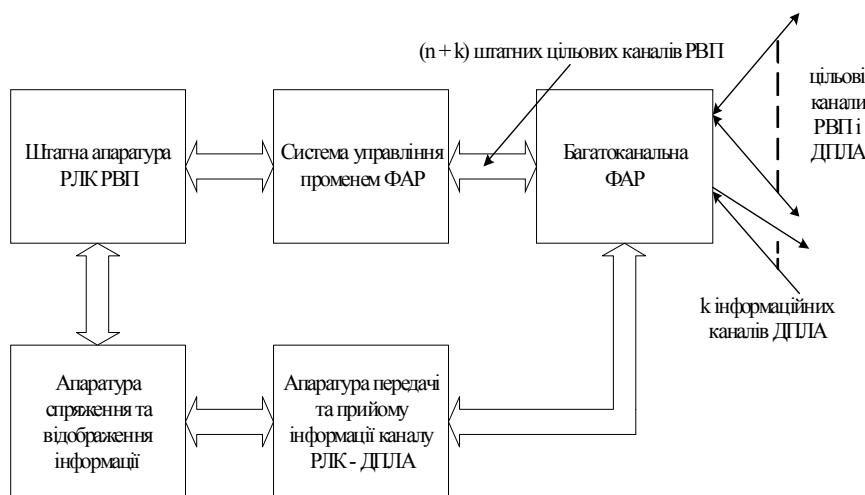


Рисунок 3 – Узагальнена структурна схема інтеграції наземної апаратури

## ВИСНОВКИ

1. Вперше запропонований завадостійкий метод багатоканальної артилерійської розвідки у всьому можливому діапазоні дальностей дії артилерійських і ракетних систем, який забезпечує інваріантність до типу і характеру руху наземних цілей.

2. Обґрунтовані принципи технічної реалізації запропонованого методу з використанням досягнутих характеристик наземних і повітряних засобів інструментальної розвідки на основі їх модернізації.

## СПИСОК ЛІТЕРАТУРИ

1. Тихомиров А. И. Американские РЛС разведки наземных целей / А. И. Тихомиров, А. П. Орлов // Зарубежное военное обозрение. – 1980. – № 2. – С. 42–46.
2. Саврасов В. И. Радиолокационные станции артиллерийской разведки / В. И. Саврасов // Зарубежное военное обозрение. – 1989. – № 8. – С. 26–30.
3. Крупников А. И. Радиолокационные станции контрбатарейной борьбы основных зарубежных стран / А. И. Крупников // Зарубежное военное обозрение. – 2010. – № 12. – С. 32–41.
4. Тымчук В. Ю. Артиллерийские РЛС серии ТРQ: некоторые аспекты построения и работы, уроки модернизации / В. Ю. Тымчук // Артиллерийское и стрелковое вооружение. – 2011. – № 1. – С. 12–19.
5. Кутовий О. П. Тенденції розвитку безпілотних літальних апаратів / О. П. Кутовий // Наука і оборона. – 2000. – № 4. – С. 39–47.
6. Зубков А. Н. Интегрированные многоспектральные поисково-прицельные системы для ракетно-артиллерийского вооружения / А. Н. Зубков, А. А. Щерба // Артиллерийское и стрелковое вооружение. – 2009. – № 1. – С. 14–18.
7. Теоретические основы радиолокации : учебн. пособие для вузов / [А. А. Коростелев, Н. Ф. Клюев, Ю. А. Мельник и др.; под ред. В. Е. Дулевича]. – М. : Сов. радио, 1978. – 608 с.
8. Мобильная РЛС наземной артиллерийской разведки СНАР-15 [Электронный ресурс]. – Режим доступа: <http://www.npostrela.com/tu/products/museum/82/210>.
9. Зайцев Н. А. На передовых позициях. Служебно-боевая разведывательная машина НПО «СТРЕЛА» / Н. А. Зайцев // Военный парад. – 2012. – № 2. – С. 51–55.
10. Зубков А. Н. Интегрированный двухспектральный все-погодный и всесуточный поисково-прицельный комплекс / А. Н. Зубков, В. И. Иванов, Б. М. Казаков // 4-я Международная конференция «Артиллерийские ствольные системы боеприпасы, средства артиллерийской разведки и управления огнем»: сб. трудов – К. : НТЦ АСВ. – 2000. – С. 200–203.
11. Зубков А. Н. Интегрированные многоспектральные системы геомониторинга. Концепция построения / А. Н. Зубков, И. Н. Прудиус // Сб. науч. трудов 3-го Межд. радиоэлектронного форума МРФ'2008. – Харьков : ХНУРЕ. – С. 283–286.
12. Концептуальный облик информационной подсистемы самонаводящихся зенитных управляемых ракет перспективных ЗРК / [А. В. Авласенок, Е. Г. Алексеев, С. П. Литвинов, Ф. Л. Савицкий] // Радиоэлектронника. – 2008. – № 5. – С. 49–54.
13. Авласенок А. В. Современные требования к многоспектральным автоматам сопровождения целей для систем высокоточного оружия и возможные пути их реализации / [А. В. Авласенок, Е. Г. Алексеев, С. П. Литвинов, Ф. Л. Савицкий] // Радиоэлектронника. – 2008. – № 6. – С. 54–61.
14. Зубков А. Н. Радиолокационные средства миллиметрового диапазона для повышения эффективности артиллерийских систем / А. Н. Зубков // Артиллерийское и стрелковое вооружение. – 2005. – № 4 (17). – С. 33–40.
15. Зубков А. Н. Системы радиовидения миллиметрового диапазона. Сопоставление и интеграция с оптическими каналами, результаты эксперимента / А. Н. Зубков // Радиоэлектронника. – 2005. – № 10. – С. 3–10.
16. Зубков А. М. Аналізатор завадової обстановки для адаптивної багатоспектральної системи спостереження / А. М. Зубков, А. В. Д'яков, С. А. Мартиненко, А. А. Щерба // Вісник Вінницького політехнічного інституту. – 2010. – № 4 (91). – С. 68–70.
17. Пат. 94566 Україна, МПК<sup>2006</sup> G01J11/00, G01S13/00. Багатоспектральний виявляч наземних об'єктів / А. М. Зубков, І. Н. Прудиус, А. В. Д'яков, С. А. Мартиненко, Д. О. Миріков, А. А. Щерба. (Україна); заявник Національний університет «Львівська політехніка». – № a201015836; Заявл. 10.02.11; – Опубл. 10.05.11. – Бюл. № 9.

Стаття надійшла до редакції 18.06.2014.

Зубков А. Н.<sup>1</sup>, Щерба А. А.<sup>2</sup>

<sup>1</sup>Д-р техн. наук, старший научный сотрудник, ведущий научный сотрудник научно-исследовательского отдела ракетных войск и артиллерии Научного центра Сухопутных войск Академии сухопутных войск имени гетмана Петра Сагайдачного, Львов, Украина

<sup>2</sup>Адъюнкт штатный научно-организационного отдела Академии сухопутных войск имени гетмана Петра Сагайдачного, Львов, Украина

## ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ АРТИЛЛЕРИЙСКОЙ РАЗВЕДКИ ПУТЕМ КОНСТРУКТИВНО-ФУНКЦИОНАЛЬНОЙ ИНТЕГРАЦИИ ВОЗДУШНЫХ И НАЗЕМНЫХ СРЕДСТВ НАБЛЮДЕНИЯ

Предложен и обоснован метод многоканальной разведки движущихся и неподвижных наземных целей, который снимает ограничения радиогоризонта, инвариантен к типу целей и основан на интеграции известных наземных и воздушных средств наблюдения в рамках единого комплекса инструментальной разведки.

**Ключевые слова:** радиолокационная станция разведки наземных движущихся целей, радиолокационный комплекс разведки огневых позиций, дистанционно пилотируемый летательный аппарат, фазированная антенна решетка, мультиспектральная система наблюдения.

Zubkov A. M.<sup>1</sup>, Shcherba A. A.<sup>2</sup>

<sup>1</sup>Ph.D, Senior Science Master, Laureate of the State Premiums of Ukraine In Science and Technology, Leading Researcher of Research Department (Rocket Troops and Artillery) of Land Forces Scientific Centre Army Academy named after Hetman Petro Sahaydachnyi, Lviv, Ukraine.

<sup>2</sup>Staff Adjunct of Scientific and Organizational Department Army Academy named after Hetman Petro Sahaydachnyi, Lviv, Ukraine

## IMPROVING OF ARTILLERY RECONNAISSANCE EFFICIENCY THROUGH CONSTRUCTIVE-FUNCTIONAL INTEGRATION OF AIR AND LAND SURVEILLANCE

The method of multichannel reconnaissance of movable and immovable ground targets was proposed and substantiated. It removes the radio horizon restrictions; it is invariant to the targets type (design features) and is based on the integration of known surface and air surveillance within a single complex of instrumental reconnaissance. It is based on the use of one or more regular targeted channel of radar system of firing positions reconnaissance with phased antenna array to determine the current coordinates of remotely piloted aircraft equipped with multispectral equipment ground targets surveillance. Thus, in these channels during time division mode, the management information is transmitting on aircraft and generic information is receiving from the board. The proposed approach provides conducting of artillery reconnaissance at any time, in any weather conditions, and the optimal combination of information capabilities of observation devices of different parts of the spectrum of electromagnetic waves without substantial revision of ground and airborne equipment, which is especially important in terms of financial constraints on the development and similar production systems.

**Keywords:** radar of ground moving targets reconnaissance, radar system of firing positions reconnaissance, remotely piloted aircraft, phased antenna array, multispectral surveillance system.

## REFERENCES

1. Tihomirov A. I., Orlov A. P. Amerikanskiye RLS razvedyedki nazemnyh tceley, *Zarubyezhnoye voyennoye obozreniye*, 1980, No. 2, pp. 42–46.
2. Savrasov V. I. Radiolokacionniye stanciy artilleriyskoy razvedki, *Zarubyezhnoye voyennoye obozreniye*, 1989, No. 8, pp. 26–30.
3. Krupnikov A. I. Radiolokacionniye stanciy kontrbatareynoy borby osnovnyh zarubezhnyh stran, *Zarubyezhnoye voyennoye obozreniye*, 2010, No. 12, pp. 32–41.
4. Tymchuk V. Y. Artilleriyskiye RLS serii TPQ: nekotoryye aspekty postroyeniya i raboty, uroki modernizatsii, *Artilleriyskoye i strelkovoye vooruzheniye*, 2011, No. 1, pp. 12–19.
5. Kutowiy O. P. Tendentciyi rozvityku bezpilotnyh litalnyh apparativ, *Nauka i oborona*, 2000, No. 4, pp. 39–47.
6. Zubkov A. N., Shcherba A. A. Integririvannyye mnogospektralnyye poiskovo-pritselniye sistemy dlya raketno-artilleriyskogo vooruzheniya, *Artilleriyskoye i strelkovoye vooruzheniye*, 2009, No. 1, pp. 14–18.
7. Korostelyev A. A., Klyuyev N. F., Melnik Y. A. i dr.; Pod red. V. E. Dulyevicha Teoreticheskiye osnovy radiolokacii: Uchebn. Posobiye dlya vuzov, 2-ye izd., pererab. i dop. Moscow, Sov. radio, 1978, 608 p.
8. Mobilnaya RLS nazemnoy artilleriyskoy razvedki SNAR-15 [Elektronnyy resurs]. Rezhym dostupu: <http://www.npostrela.com/ru/products/museum/82/210>.
9. Zaytcev N. A. Na peredovyh pozitciyah. Sluzhebno-boyevaya razvedyvatelnaya mashyna NPO «STRYELA», *Voyenniy parad*, 2012, No. 2, pp. 51–55.
10. Zubkov A. N., Ivavnov V. I., Kazakov B. M. Intyegrirovanniy dvuhspaktralniy vsyepogodniy i vsyesutochniy poiskovo-
11. Zubkov A. N., Prudyus I. N. Integririvannyye mnogospektralnyye sistemy geomonitoringa. Kontsepciya postruieniya, Sb. naushn. trudov 3-go Myezhd. radioel. foruma MRF'2008. Kharkov, KHNURE, pp. 283–286.
12. Avlasenok A. V., Alyeksyeyev E. G., Lytvynov S. P., Savytckiy F. L. Kontseptualniy oblik informatcyonnoy podsistemy samonavodyashchihsya zenitnyh upravlyayemyh raket perspektivnyh ZRK, *Radoelektronika*, 2008, No. 5, pp. 49–54.
13. Avlasenok A. V., Alyeksyeyev E. G., Lytvynov S. P., Savytckiy F. L. Sovryemennyye tryebovaniya k mnogospektralnym avtomatam soprovozhdyyeniya tcelyey dlya system vysokotochnogo oruzhiya i vozmozhniye puti ikh realizaciyyi, *Radoelektronika*, 2008, No. 6, pp. 54–61.
14. Zubkov A. N. Radiolokacionniye sredstva milimetrovogo diapazona dlya povysheniya efektivnosti artilleriyskikh system, *Artilleriyskoye i strelkovoye vooruzheniye*, 2005, No. 4 (17), pp. 33–40.
15. Zubkov A. N. Sistemy radiovidyeniya milimetrovogo diapazona. Sopostavlyeniye i integraciya sopticheskimi kanalami, rezulaty eksperimenta, *Radoelektronika*, 2005, No. 10, pp. 3–10.
16. Zubkov A. M., D'yakov A.V., Martynenko S. A., Shcherba A. A. Analizator zavadovoyi obstanovky dlya adaptivnoy bagatospektralnoy sistemy sposterezheniya, *Visnyk Vinnytskogo politehnichnogog instytutu*, 2010, No. 4 (91), pp. 68–70.
17. Zubkov A. M., Prudyus I. N., D'yakov A. V., Martynenko S. A., Mymrikov D. O., Shcherba A.A. (Ukrayina) Bagatospektralnyi vyavlyach nazemnyh ob'yektitv: Pat. 94566 Ukrayina, MPK<sup>2006</sup> G01J11/00, G01S13/00; zayavnyk Natcionalnyi universyet «Lvivska politehnika», № a201015836; Zayavl. 10.02.11; Opubl. 10.05.11. Byul. № 9.

УДК 621.396.6.017:004.942

Крищук В. Н.<sup>1</sup>, Шило Г. Н.<sup>2</sup>, Каспирович Н. А.<sup>3</sup>, Гренич Е. В.<sup>4</sup><sup>1</sup>Канд. техн. наук, професор, Запорожский национальный технический университет, Украина<sup>2</sup>Канд. техн. наук, доцент, Запорожский национальный технический университет, Украина<sup>3</sup>Аспирант, Запорожский национальный технический университет, Украина<sup>4</sup>Асистент, Запорожский национальный технический университет, Украина

## ОПТИМИЗАЦІЯ ОРЕБРЕННОГО КАНАЛА В СИСТЕМАХ ПРИНУДИТЕЛЬНОГО ВОЗДУШНОГО ОХЛАЖДЕННЯ РЭА

Рассмотрена возможность использования оребренного канала, как упрощенной модели для анализа тепловых режимов радиоэлектронной аппаратуры кассетного типа с высоким удельным тепловыделением при принудительном воздушном охлаждении.

Исследованы тепловые характеристики оребренного канала. Получены соотношения для оптимальных размеров конструкции оребренного канала, которые позволяют минимизировать тепловое сопротивление. Разработан итерационный алгоритм оптимизации.

Показан пример проектирования оребренного канала с оптимальным тепловым режимом. В результате оптимизации конструкции оребренного канала тепловое сопротивление уменьшилось на 35 %.

**Ключевые слова:** оребренный канал, тепловая модель, тепловое сопротивление, принудительное воздушное охлаждение, системы инженерного анализа, радиоэлектронная аппаратура.

### ВВЕДЕНИЕ

Обеспечение теплового режима является одной из важных задач, возникающих при проектировании радиоэлектронной аппаратуры. Особенно эта задача становится актуальной для мобильной радиоэлектронной аппаратуры кассетного типа, которая характеризуется высоким удельным тепловыделением [1, 2]. В радиоэлектронной аппаратуре кассетного типа используется плотное размещение однотипных блоков. Примером такой конструкции может послужить твердотельный передатчик [3]. Чаще всего такие устройства требуют принудительного воздушного охлаждения и оребрения поверхности наиболее теплонагруженных блоков. Поэтому тепловую модель конструкции можно представить в виде оребренного канала с принудительным воздушным охлаждением.

В литературе рассмотрены методы оптимизации параметров оребрения поверхностей при естественном воздушном охлаждении, позволяющие минимизировать массу [4]. Такие методы основаны на применении упрощенных аналитических формул и не позволяют оптимизировать параметры при принудительном воздушном охлаждении. В качестве еще одного способа оптимизации используется физическое моделирование [5], которое характеризуется значительной трудоемкостью и материальными затратами.

Последние годы для решения задач теплового проектирования используются программные средства инженерного анализа [6, 7]. В радиоэлектронной аппаратуре такой подход позволил значительно повысить точность методов оптимизации конструкций герметичных блоков РЭА [8] и параметров ребристых радиаторов [9] с естественным воздушным охлаждением.

Целью работы является исследование тепловых характеристик оребренного канала при принудительном воздушном охлаждении с использованием средств инженерного анализа и разработка метода оптимизации оребренного канала.

Для достижения поставленной цели необходимо для оребренного канала решить следующие задачи:

- создать тепловую модель,
- исследовать тепловые характеристики,
- определить оптимальные соотношения конструктивных параметров,
- разработать алгоритм оптимизации.

### 1 ТЕПЛОВЫЕ ХАРАКТЕРИСТИКИ ОРЕБРЕННОГО КАНАЛА

Исследования тепловых характеристик модели оребренного канала проводились с использованием модуля инженерного анализа системы автоматизированного проектирования Solid Works. При построении тепловой модели в системе инженерного анализа были приняты следующие ограничения:

- коэффициент теплопроводности материала не зависит от температуры;
- боковые стенки оребренного канала адиабатические;
- для обеспечения принудительного охлаждения на вход подается постоянный воздушный поток на расстоянии 100 мм, который задается объемным расходом воздуха;
- мощность задавалась на внешние поверхности основания оребренного канала.

Тепловая модель оребренного канала состоит из двух устройств. Трехмерная модель канала представлена на рис. 1, где показаны конструктивные элементы канала: 1 – основание, 2 – оребрение. Габаритные размеры устройств:  $H$  – высота канала,  $L$  – длина канала и  $D$  – ширина канала. Стрелкой показано направление воздушного потока.

Площадь поперечного сечения оребренного канала показана на рис. 2, где  $d_p$  – толщина основания канала;  $l_k$  – расстояние между устройствами в канале;  $d_r$  – толщина ребра;  $l_r$  – высота ребра.

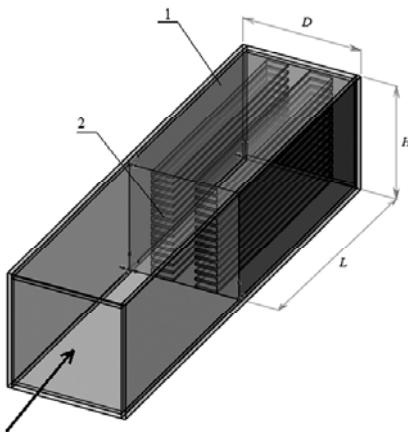


Рисунок 1 – Тепловая модель оребренного канала

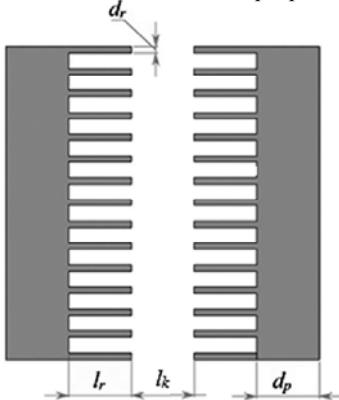


Рисунок 2 – Поперечное сечение оребренного канала

Моделирование проводилось для канала, который имел параметры: длина канала – 100 мм, высота канала – 50 мм, ширина канала – 50 мм, расстояние между устройствами 10 мм. Рассеиваемая мощность – 50 Вт. Температура окружающей среды 50°C.

Исследовали зависимость максимальной температуры канала от параметров оребрения: длина ребра, ширина ребра, количество ребер. Габаритные размеры канала: высота, длина, ширина не изменились.

Результаты моделирования представлялись в виде тепловых характеристик зависимостей теплового сопротивления от параметров оребрения.

Зависимость теплового сопротивления канала от количества ребер представлено на рис. 3, где  $R_t$  – тепловое сопротивление,  $n$  – количество ребер.

Тепловое сопротивление определяется с помощью выражения:

$$R_t = \frac{t_{\max} - t_c}{P}, \quad (1)$$

где  $t_{\max}$  – максимальная температура в канале;  $t_c$  – температура окружающей среды;  $P$  – рассеиваемая мощность.

Как видно из рис. 3, эта тепловая характеристика имеет экстремум, что позволяет провести оптимизацию количества ребер. Для тепловых характеристик такого вида можно использовать математическую модель в виде:

$$R_t(n) = g_1 \cdot n + \frac{g_2}{n}, \quad (2)$$

где  $g_1, g_2$  – коэффициенты модели.

Для определения коэффициентов модели составили систему уравнений в двух точках графика, расположенных по разные стороны от области минимума. Решение системы представляется в виде соотношений:

$$g_1 = (Rt_2 - \frac{g_2}{x_2}) \cdot \frac{1}{x_2},$$

$$g_2 = (Rt_{(k-1)} \cdot x_{(k-1)} - \frac{Rt_2 \cdot (x_{(k-1)})^2}{x_2}) \div (1 - \frac{(x_{(k-1)})^2}{(x_2)^2}). \quad (3)$$

Зависимость теплового сопротивления оребренного канала от толщины ребра представлено на рис. 4. Эта характеристика также имеет экстремум. Математическая модель этой характеристики аналогична виду (2) и имеет подобные выражения для определения коэффициентов.

Предложенные тепловые модели могут быть использованы для оптимизации оребренного канала.

Зависимость теплового сопротивления от длины ребра имеет монотонно-убывающий вид (рис. 5).

В процессе проектирования канала длина ребра выбирается максимально возможной с целью получения наименьшего значения теплового сопротивления.

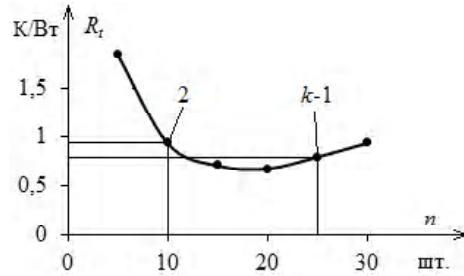


Рисунок 3 – Зависимость теплового сопротивления канала от количества ребер

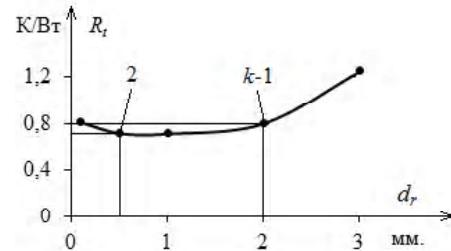


Рисунок 4 – Зависимость теплового сопротивления канала от толщины ребра

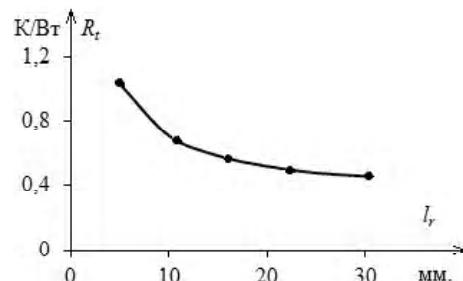


Рисунок 5 – Зависимость теплового сопротивления канала от длины ребра

## 2 ОПТИМИЗАЦІЯ ОРЕБРЕННОГО КАНАЛА

Исследование тепловых характеристик оребренного канала показало, что выбором каналов можно существенно снизить перегревы устройств. Оптимизация параметров оребренного канала производится для целевой функции:

$$R_t(n, d_r) \rightarrow \min. \quad (4)$$

В процессе оптимизации учитывались ограничения на размеры канала:

$$H=H_k, L=L_k, D=D_k$$

где  $H_k, L_k, D_k$  – номинальные значения конструктивных параметров, полученные на этапе проектирования устройства.

При этих ограничениях высота ребер определяется с помощью соотношений

$$l_r = \frac{1}{2}(D - l_k - 2 \cdot d_p). \quad (5)$$

Математическая модель (2) позволяет определить оптимальное количество ребер с помощью соотношения:

$$n = \sqrt{\frac{g_2}{g_1}}. \quad (6)$$

Аналогично определяется оптимальное значение толщины ребра. Это позволяет при оптимизации параметров оребренного канала использовать последовательный итерационный алгоритм, в котором проводится оптимизация одного из параметров, а результаты этой оптимизации используются при оптимизации следующего параметра.

Шаг 1. – Задаются начальные значения: длины ребра, количества ребер и толщины ребра. С помощью моделирования определяется тепловое сопротивление оребрения канала по заданной мощности и допустимому перепаду температур.

Шаг 2. – Моделируются тепловые режимы канала при изменении оптимизируемого параметра на величину  $\pm 20\%$  и определяются коэффициенты  $g_1$  и  $g_2$ .

Шаг 3. – Определяется оптимальное значение параметра.

Шаг 4. – Устанавливается оптимальное значение параметра.

Шаг 5. – Повторяем шаг 2, 3 для следующего параметра.

Шаг 6. – Проверяется условие завершения алгоритма:

$$\frac{Rt^{(k)} - Rt_{\partial}^{(k-1)}}{Rt_{\partial}^k} \leq \varepsilon,$$

где  $k$  – номер итерации;  $\varepsilon$  – погрешность вычислений.

Если условие не выполняется, то переход на шаг 2, иначе – конец алгоритма.

## 3 ЕКСПЕРИМЕНТИ І РЕЗУЛЬТАТИ

Разработанный алгоритм использован для оптимизации оребренного канала, параметры которого приводятся ниже.

Мощность канала  $P = 50$  Вт, температура окружающей среды  $t_c = 50$  °C, коэффициент черноты поверхности теплоотвода 0,91, погрешность вычислений алгоритма оптимизации соответствует  $\varepsilon = 0,05$ .

В качестве начального приближения при оптимизации были выбраны размеры оребренного канала:  $n=10$ ,  $l_r=10$  мм и  $d_r=1$  мм. Результаты оптимизации конструкций оребренного канала приведены в табл. 1.

Проведенная оптимизация показала, что после 9 итераций была достигнута требуемая точность, тепловой режим улучшен, тепловое сопротивление уменьшилось на 35%.

## ВЫВОДЫ

Предложенный метод оптимизации оребренного канала с использованием современных систем автоматического проектирования позволяет реализовать допустимый тепловой режим работы устройства. Процедура оптимизации реализована в виде итерационного алгоритма. Алгоритм обладает достаточно высокой сходимостью.

Сравнение величины теплового сопротивления исходного оребренного канала и полученного в результате оптимизации показало, что оптимизация приводит к улучшению теплового режима, а тепловое сопротивление уменьшилось на 35 %.

## БЛАГОДАРНОСТИ

Работа выполнена в рамках выполнения научно-исследовательской темы Запорожского национального технического университета «Информационные технологии проектирования теплонагруженных радиоэлектронных аппаратов» (номер гос. регистрации 0113U001096).

## СПИСОК ЛИТЕРАТУРЫ

- Ненашев А. П. Конструирование радиоэлектронных средств / А. П. Ненашев. – М. : Высш. шк., 1990. – 432 с.
- Дульнев Р. Н. Тепло- и массообмен в радиоэлектронной аппаратуре / Р. Н. Дульнев. – М. : Высш.шк., 1984. – 247 с.
- Пат. 13704 Україна, МПК G01S 7/36 H04B 15/00. Пристрій для подавлення пасивних завод / Каспирович О. Г., Лаврентьев В. М., Май I. Д., Пресняк I. C.; заявник та патентовласник КП «НВК «Іскра» № 200509636; заяв. 13.10.2005; опубл. 17.04.2006, Бюл. № 4, 2006.
- Ройзен Л. И. Телевой расчет оребренных поверхностей / Л. И. Ройзен, И. Н. Дулькин; под. ред. В. Г. Фастовского. – М. : Энергия, 1977. – 256 с.

Таблица 1 – Параметры оребренного канала

№ итерации	0	1	2	3	4	5	6	7	8	9
$n$ , шт.	10	15	15	18	21	22	22	23	24	24
$d_r$ , мм.	1	0,9	0,7	0,6	0,49	0,4	0,3	0,35	0,27	0,24
$R_t$ , °C/Вт	1,071	0,7086	0,7656	0,7462	0,7252	0,711	0,701	0,688	0,687	0,683

5. Drofenik U. Analysis of Theoretical Limits of Forced-Air Cooling Using Advanced Composite Materials With High Thermal Conductivities / U. Drofenik, A. Stupar, J.W. Kolar // IEEE Transactions on Components, Packaging, and Manufacturing Technology. – 2011. – Vol. 1, № 4. – P. 528–535.
6. Алямовский А. А. Инженерные расчеты в Solid Works Simulation / А. А. Алямовский. – М. : ДМК Пресс, 2010. – 464 с.
7. Гончаров П. С. Практическое использование NX / П. С. Гончаров, М. Ю. Ельцов, С. Б. Коршиков, И. В. Лаптев, В. А. Осиюк. – М. : ИД ДМК Пресс, 2010. – 504 с.
8. Гапоненко Н. П. Оптимизация объема герметичных блоков радиоэлектронной аппаратуры / Н. П. Гапоненко, О. В. Сиротюк, Е. В. Огренич, Ю. А. Лопатка, Е. К. Арешкин // Прикладная радиоэлектроника. – 2012. – Т. 11, № 3. – С. 155–158.
9. Шило Г. Н. Проектирование радиаторов с оптимальными массогабаритными параметрами / Г. Н. Шило, Н. П. Гапоненко, Е. В. Огренич // Технология и конструирование в электронной аппаратуре. – 2011. – № 1–2. – С. 30–33.

Статья поступила в редакцию 21.05.2014.

После доработки 28.05.2014.

Кришук В. М.<sup>1</sup>, Шило Г. М.<sup>2</sup>, Каспирович Н. О.<sup>3</sup>, Огренич Е. В.<sup>4</sup>

<sup>1</sup>Канд. техн. наук, проф., Запорізький національний технічний університет, Україна

<sup>2</sup>Канд. техн. наук, доцент, Запорізький національний технічний університет, Україна

<sup>3</sup>Аспірант, Запорізький національний технічний університет, Україна

<sup>4</sup>Асистент, Запорізький національний технічний університет, Україна

## ОПТИМІЗАЦІЯ ОРЕБРЕННОГО КАНАЛУ В СИСТЕМАХ ПРИМУСОВОГО ПОВІТРЯННОГО ОХОЛОДЖЕННЯ РЕА

Розглянуто можливість використання оребреного каналу, як спрощеної моделі для аналізу теплових режимів радіоелектронної апаратури касетного типу з високим питомим тепловиділенням при примусовому повітряному охолодженні.

Досліджено теплові характеристики оребреного каналу. Отримано співвідношення для оптимальних розмірів конструкції оребреного каналу, які дозволяють мінімізувати тепловий опір. Розроблено ітераційний алгоритм оптимізації.

Наведено приклад проектування оребреного каналу з оптимальним тепловим режимом. В результаті оптимізації конструкції оребреного каналу тепловий опір зменшився на 35 %.

**Ключові слова:** оребрений канал, теплова модель, тепловий опір, примусове повітряне охолодження, системи інженерного аналізу, радіоелектронна апаратура.

Krischuk V. N.<sup>1</sup>, Shilo G. N.<sup>2</sup>, Kaspyrovych N. A.<sup>3</sup>, Ogrenich E. V.<sup>4</sup>

<sup>1</sup>Ph.D, Professor, Zaporozhye National Technical University, Ukraine

<sup>2</sup>Ph.D., Associate Professor, Zaporozhye National Technical University, Ukraine

<sup>3</sup>Postgraduate student, Zaporozhye National Technical University, Ukraine

<sup>4</sup>Assistant, Zaporozhye National Technical University, Ukraine

## FINNED CHANNEL OPTIMIZING IN RADIO ELECTRONIC DEVICES FOR FORCED AIR COOLING SYSTEMS

The finned channel is used as a simplified model for the analysis of thermal conditions of cartridge design electronic equipment with a high power density for forced air cooling. The finned channel thermal model and method of its creating are described by the computer-aided engineering systems. Clearance limitations of the radio electronic units and the manufacturing limitations on width of fins and their numbers are considered during creating the thermal model.

The finned channel thermal characteristics that include the thermal resistance dependence on parameters of fins (width, length, number) are investigated. The approximating functions for these dependencies are proposed.

To minimize the thermal resistance the correlations for finned channel having the optimal sizes are obtained. The iteration optimization algorithm is developed. This approach provides high accuracy. In the algorithm the thermal conditions are calculated by the computer-aided engineering systems.

The example of designing the finned channel with optimal thermal conditions is showed. As a result of optimizing the finned channel thermal resistance decreased by 35 %.

**Keywords:** finned channel, thermal model, thermal resistance, forced air cooling, computer-aided engineering systems, electronic equipment.

## REFERENCES

1. Nenashev A. P. Konstruirovaniye radioelektronnyih sredstv. Moscow, Vyssh. shk., 1990, 432 p.
2. Dulnev R. N. Teplo- i massoobmen v radioelektronnoy apparature, Moscow, Vyssh. shk., 1984, 247 p.
3. Kaspирович О. Г., Lavrentev V. M., May I. D., Presnyak I. S. Pat. 13704 Україна, MPK G01S 7/36 H04B 15/00. Pristly dlya podavleniya pasivnih zavad ; zayavnik ta patentovlasnik KP «NVK «dskra». No. u200509636; zayav.13.10.2005; opubl.17.04.2006, Byul. No. 4, 2006.
4. Royzen L. I., Dulkin I. N. pod. red. V. G. Fastovskogo Teloyov raschet orebrennyih poverhnostey. Moscow, Energiya, 1977, 256 p.
5. Drofenik U., Stupar A., Kolar J. W. Analysis of Theoretical Limits of Forced-Air Cooling Using Advanced Composite Materials With High Thermal Conductivities, *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, 2011, vol. 1, No. 4, pp. 528–535.
6. Alyamovskiy A. A. Inzhenernyie raschetyi v Solid Works Simulation. Moscow, DMK Press, 2010, 464 p.
7. Goncharov P. S., Eltsov M. Yu., Korshikov S. B., Laptev I. V., Osiyuk V. A. Prakticheskoe ispolzovanie NX. Moscow, ID DMK Press, 2010, 504 p.
8. Gaponenko N. P., Sirotyuk O. V., Ogrenich E. V., Lopatka Yu. A., Areshkin E. K. Optimizatsiya ob'ema germetichnyih blokov radioelektronnoy apparaturyi, *Prikladnaya radioelektronika*, 2012, Vol.11, No. 3, pp. 155–158.
9. Shilo G. N., Gaponenko N. P., Ogrenich E. V. Proektirovaniye radiatatorov s optimalnyimi massogabarytnymi parametrami, *Tehnologiya i konstruirovaniye v elektronnoy apparature*, 2011, No. 1–2, pp. 30–33.

## МЕТОД ОЦІНЮВАННЯ ПОХИБКИ ВИМІРЮВАННЯ МАГНІТНОГО ПОЛЯ 3D-ЗОНДОМ З НЕЛІНІЙНИМИ ВИХІДНИМИ ХАРАКТЕРИСТИКАМИ

У статті запропонований метод оцінки придатності результатів вимірювання магнітного поля 3D-зондом, що складається з трьох сенсорів, вихідні характеристики яких залежать від усіх компонентів вектора магнітної індукції, а також їхніх квадратів. Одержані аналітичні залежності, які дозволяють визначити, чи можна 3D-зондом магнітного поля з відомими коефіцієнтами вихідних характеристик забезпечити вимірювання вектора магнітної індукції з відносною похибкою, не більшою за задану.

**Ключові слова:** 3D-зонди магнітного поля, давачі Холла, похибка, вимірювання.

### НОМЕНКЛАТУРА

- PXC – розщеплені холлівські структури;  
 $a_{ij}$  – коефіцієнти при компонентах вектора магнітної індукції у трьох вихідних сигналах 3D-зонда магнітного поля ( $i = \overline{1,3}$ ,  $j = \overline{1,3}$ );  
 $b_{ij}$  – коефіцієнти при квадратах компонент вектора магнітної індукції у трьох вихідних сигналах 3D-зонда магнітного поля ( $i = \overline{1,3}$ ,  $j = \overline{1,3}$ );  
 $B$  – модуль вектора магнітної індукції;  
 $B_x$  – проекція вектора магнітної індукції на вісь абсцис у прямокутній декартовій системі координат, прив'язаній до підкладки 3D-зонда;  
 $B_y$  – проекція вектора магнітної індукції на вісь ординат у прямокутній декартовій системі координат, прив'язаній до підкладки 3D-зонда;  
 $B_z$  – проекція вектора магнітної індукції на вісь аплікат у прямокутній декартовій системі координат, прив'язаній до підкладки 3D-зонда;  
 $B_{xf}$  – значення проекції вектора магнітної індукції на вісь абсцис, знайдене з похибкою;  
 $S_i$  – вихідний сигнал  $i$ -го давача у складі 3D-зонда магнітного поля ( $i = \overline{1,3}$ );  
 $A$  – матриця системи лінійних рівнянь;  
 $\Delta B_x$  – абсолютна похибка величини  $B_x$ ;  
 $\Delta B_y$  – абсолютна похибка величини  $B_y$ ;  
 $\Delta B_z$  – абсолютна похибка величини  $B_z$ ;  
 $\alpha$  – кут між віссю абсцис і вектором магнітної індукції;  
 $\beta$  – кут між віссю аплікат і вектором магнітної індукції;  
 $N$  – допустима відносна похибка, з якою потрібно виміряти величину магнітної індукції;  
 $y$  – допоміжна функція;  
 $q$  – допоміжна змінна, введена для стисlosti запису;  
 $\varphi$  – допоміжна змінна, кут між векторами;  
 $\delta B$  – відносна похибка вимірювання величини  $B$ .

### ВСТУП

Давачі Холла є одними з найбільш поширеніх сенсорів магнітного поля завдяки їхнім численним перевагам, зокрема, здатності вимірювати магнітне поле в широких межах ( $10^{-5} \dots 10^2$  Т), низькому енергоспоживанню ( $<10$  мВт), мінімальним габаритам (1…3 мм) тощо [1, 2]. Традиційні холлівські сенсори характеризуються симетричним розміщенням пар струмових і потенційних електродів. Їхній сигнал лінійно залежить від проекції вектора індукції магнітного поля на нормаль до площини чутливого напівпровідникового шару, що робить традиційні холлівські сенсори простими у калібруванні та використанні. Різновидом давачів Холла є розщеплені холлівські структури (PXC), що відрізняються від традиційних давачів Холла структурною асиметрією [3, 4], зокрема, PXC може мати два струмових і лише один потенційний електрод. За рахунок структурної асиметрії забезпечується можливість одночасного вимірювання усіх компонент вектора магнітної індукції і підвищення просторової роздільної здатності 3D-зондів магнітного поля на основі PXC. Таким чином, PXC володіють перевагами перед традиційними холлівськими сенсорами і мають принципово нові властивості. Однак, окрім переваг, структурна асиметрія PXC породжує також і їхню проблематику – вихідні характеристики PXC у магнітному полі вже понад 500 мТ є нелінійними. Щоб знайти компоненти вектора магнітної індукції за допомогою зондів на основі PXC, слід розв'язати систему нелінійних рівнянь, яка, взагалі кажучи, може мати більше ніж один розв'язок. Метою роботи є визначення вимог, яким повинні відповісти 3D-зонди магнітного поля з вихідними характеристиками, що залежать одночасно від усіх компонент вектора магнітної індукції і їхніх квадратів, щоб забезпечити можливість вимірювання вектора магнітної індукції з похибкою, не більшою за допустиму.

### 1 ПОСТАНОВКА ЗАДАЧІ

У [5] розглянуто 3D-зонд магнітного поля на основі розщеплених холлівських структур, вихідні характеристи-

тики яких у магнітних полях понад 500 мТ описуються функціями вигляду:

$$\begin{aligned} a_{11}B_x + a_{12}B_y + a_{13}B_z + b_{11}B_x^2 + b_{12}B_y^2 + b_{13}B_z^2 &= S_1 \\ a_{21}B_x + a_{22}B_y + a_{23}B_z + b_{21}B_x^2 + b_{22}B_y^2 + b_{23}B_z^2 &= S_2 \quad (1) \\ a_{31}B_x + a_{32}B_y + a_{33}B_z + b_{31}B_x^2 + b_{32}B_y^2 + b_{33}B_z^2 &= S_3, \end{aligned}$$

де  $S_i$  – вихідний сигнал  $i$ -го сенсора у складі 3D-зонда,  $B = \begin{pmatrix} B_x & B_y & B_z \end{pmatrix}$  – вимірювана векторна величина,  $a_{ij}$ ,  $b_{ij}$  ( $i = \overline{1,3}, j = \overline{1,3}$ ) – відомі коефіцієнти.

Рівняння (1) є рівняннями квадрик і, що очевидно з їхнього геометричного змісту, можуть мати більше, ніж один розв'язок (декілька точок перегину квадрик). Неоднозначність розв'язків обмежує та ускладнює застосування 3D-зондів з вихідними характеристиками вигляду (1). Тим не менш, 3D-зонд може бути придатний для використання у випадках, якщо вся множина розв'язків системи (1) знаходиться у межах певної допустимої похибки. У даній роботі поставлена задача знаходження аналітичних залежностей, яким повинні задовольняти коефіцієнти  $a_{ij}$ ,  $b_{ij}$  для забезпечення вимірювання модуля вектора магнітної індукції з відносною похибкою, не більшою за задану.

## 2 ОГЛЯД ЛІТЕРАТУРИ

Як вже відзначалося, зонди на основі розщеплених холлівських структур характеризуються нелінійними характеристиками, причому нелінійність залежить від величини вектора магнітної індукції. Відповідно, результати застосування таких зондів залежать від умов застосування, тобто, для різних величин магнітного поля розкид розв'язків буде різним. Тому для можливості прийняття рішення про коректність результатів доцільно використовувати зазначені зонди у складі інтелектуальних сенсорів [6], «мозком» яких є мікроконтролери.

Можна виділити два принципово різні підходи до використання зондів, для яких результат знаходження шуканої векторної величини на основі їхніх вихідних сигналів є неоднозначним. Один з підходів – дізнатися число розв'язків системи нелінійних рівнянь (зокрема, вигляду (1)). Якщо це число виявиться більшим за 1, – обчислити відстань між розв'язками. Типово для пошуку кількості розв'язків системи нелінійних рівнянь застосовують базис Гребнера [7, 8]. Однак, застосування цього методу в інтелектуальних сенсорах обмежується можливостями мікроконтролерів внаслідок чутливості методу до точності проміжних обчислень. Спосіб відділення коренів для застосування в інтелектуальних сенсорах запропонований у [9]. Знаходження числових областей, що містять розв'язки, дозволяє приблизно оцінити, наскільки розв'язки відрізняються один від одного. Якщо ця відмінність істотна, необхідні додаткові заходи, що дозволили б відкинути зайні розв'язки. Якщо ж усі розв'язки системи нелінійних рівнянь (складених на основі вихідних характеристик зонда та зчитаних сигналів) відрізняються не більше, ніж на допустиму похибку, достатньо

знайти один з розв'язків і провести оцінку, наскільки істинне значення може відрізнятися від знайденого. Можливість виконання цієї умови залежить від коефіцієнтів  $a_{ij}$ ,  $b_{ij}$ . Зокрема, чим меншими є коефіцієнти при квадратах  $B_x^2$ ,  $B_y^2$  і  $B_z^2$  вимірюваної векторної величини (по відношенню до коефіцієнтів при  $B_x$ ,  $B_y$ ,  $B_z$ ), тим менше розв'язок системи рівнянь (1) відрізняється від розв'язку цієї системи рівнянь при  $b_{ij} = 0$  (тобто, від розв'язку системи лінійних алгебраїчних рівнянь).

На практиці коефіцієнти при квадратах  $B_x^2$ ,  $B_y^2$  і  $B_z^2$  у вихідних характеристиках зондів на основі РХС у певному діапазоні магнітних полів є значно меншими за коефіцієнти при компонентах вектора магнітної індукції, що уможливлює існування рішення задачі, сформульованої у статті. Крім того, у кожного з сенсорів на основі РХС типово більш вираженою є чутливість до складової вектора магнітної індукції, перпендикулярної до площини їхнього чутливого шару, ніж чутливість до двох інших компонент.

## 3 МАТЕРІАЛИ ТА МЕТОДИ

Запишемо систему рівнянь (1) у наступному вигляді:

$$\begin{aligned} a_{11}B_x + a_{12}B_y + a_{13}B_z - S_1 - (b_{11}B_x^2 + b_{12}B_y^2 + b_{13}B_z^2), \\ a_{21}B_x + a_{22}B_y + a_{23}B_z - S_2 - (b_{21}B_x^2 + b_{22}B_y^2 + b_{23}B_z^2), \\ a_{31}B_x + a_{32}B_y + a_{33}B_z - S_3 - (b_{31}B_x^2 + b_{32}B_y^2 + b_{33}B_z^2). \quad (2) \end{aligned}$$

Розв'яжемо СЛАР

$$AB = S, \quad (3)$$

де  $A$  – матриця системи (2),  $S = (S_1 \quad S_2 \quad S_3)^T$ . В результаті цього отримуємо вектор-розв'язок  $B$ , який відрізняється від істинного значення вимірюваного вектора. Різниця між обчисленним вектором  $B$  і істинним вектором залежить від величин  $(b_{11}B_x^2 + b_{12}B_y^2 + b_{13}B_z^2)$ ,  $(b_{21}B_x^2 + b_{22}B_y^2 + b_{23}B_z^2)$ ,  $(b_{31}B_x^2 + b_{32}B_y^2 + b_{33}B_z^2)$  – чим вони менші, тим меншою буде похибка. Компонента шуканого вектора  $B_x$  може бути виражена наступним чином:

$$B_{xf} = \frac{S_1(a_{22}a_{33} - a_{23}a_{32}) + S_2(a_{13}a_{32} - a_{12}a_{33}) + S_3(a_{12}a_{23} - a_{13}a_{22})}{|A|}, \quad (4)$$

де  $|A|$  – визначник матриці системи (2).

Абсолютну похибку компоненти вектора  $\Delta B_x$  можна оцінити як абсолютну різницю між значеннями  $B_x$ , обчисленими зі СЛАР (2) і СЛАР (3). Отже:

$$\begin{aligned} \Delta B_x = \frac{1}{|A|} \cdot (b_{11}B_x^2 + b_{12}B_y^2 + b_{13}B_z^2)(a_{22}a_{33} - a_{23}a_{32}) + \\ + \frac{1}{|A|} \cdot (b_{21}B_x^2 + b_{22}B_y^2 + b_{23}B_z^2)(a_{13}a_{32} - a_{12}a_{33}) + \\ + \frac{1}{|A|} \cdot (b_{31}B_x^2 + b_{32}B_y^2 + b_{33}B_z^2)(a_{12}a_{23} - a_{13}a_{22}). \quad (5) \end{aligned}$$

Тут і надалі під позначеннями  $B_x, B_y, B_z$  та  $B$  розуміємо їхні істинні значення. З теорії похибок відомо, що абсолютна похибка суми не перевищує суми абсолютнох похибок складових. Такий підхід для оцінки похибки дасть завищений результат. Тим не менш, спочатку оцінимо абсолютну похибку першого доданка у (5) як функції від  $\alpha$  і  $\beta$  – кутів, що визначають компоненти  $B_x, B_y$  і  $B_z$  вектора  $B$ . Очевидно, що величини  $(a_{22}a_{33} - a_{23}a_{32})$  та  $|A|$  не залежать від  $\alpha$  та  $\beta$  і не впливають на результат, тому шукатимемо максимальне значення функції:

$$y = b_{11}B^2 \cos^2 \alpha \sin^2 \beta + b_{12}B^2 \sin^2 \alpha \sin^2 \beta + b_{13}B^2 \cos^2 \beta. \quad (6)$$

Враховуючи те, що функція  $y$  є періодичною і визначеною для всіх чисел  $\alpha$  і  $\beta$ , шукатимемо її значення у стаціонарних точках.

Зайдемо частинні похідні:

$$\begin{aligned} \frac{\partial y}{\partial \alpha} &= -2b_{11}B^2 \sin^2 \beta \cos \alpha \sin \alpha + 2b_{12}B^2 \sin^2 \beta \sin \alpha \cos \alpha = \\ &= \sin 2\alpha (-b_{11}B^2 \sin^2 \beta + b_{12}B^2 \sin^2 \beta) = \sin 2\alpha \cdot B^2 \sin^2 \beta (b_{12} - b_{11}), \end{aligned}$$

$$\begin{aligned} \frac{\partial y}{\partial \beta} &= 2b_{11}B^2 \cos^2 \alpha \sin \beta \cos \beta + 2b_{12}B^2 \sin^2 \alpha \sin \beta \cos \beta - \\ &- 2b_{13}B^2 \sin \beta \cos \beta = B^2 \sin 2\beta (b_{11} \cos^2 \alpha + b_{12} \sin^2 \alpha - b_{13}), \end{aligned}$$

Розв'яжемо відносно  $\alpha$  і  $\beta$  систему рівнянь:

$$\begin{cases} \sin 2\alpha \cdot B^2 \sin^2 \beta (b_{12} - b_{11}) = 0, \\ B^2 \sin 2\beta (b_{11} \cos^2 \alpha + b_{12} \sin^2 \alpha - b_{13}) = 0, \end{cases}$$

Знайдена множина розв'язків:

1.  $\beta = \pi n, n \in Z$ , значення  $\alpha$  довільне.

2.  $\alpha = \frac{\pi n}{2}, n \in Z, \beta = \frac{\pi k}{2}, k \in Z$ .

3.  $\alpha = \frac{\pi n}{2}, n \in Z$ , значення  $\beta$  – довільне, якщо  $b_{11} = b_{13}$

і  $n$  – парне або  $b_{12} = b_{13}$  і  $n$  – непарне.

4.  $\beta = \frac{\pi n}{2}, n \in Z$ , значення  $\alpha$  – довільне, якщо

$$b_{11} = b_{12} \neq b_{13}.$$

5. Всі значення  $\alpha$  та  $\beta$ , якщо  $b_{11} = b_{12} = b_{13}$ .

Зайдемо значення функції (6) для знайдених стаціонарних точок:

1.  $y = b_{13}B^2$ ;

2. Для парних значень  $n$  і парних значень  $k$ :  $y = b_{13}B^2$ ;

Для парних значень  $n$  і непарних значень  $k$ :  $y = b_{11}B^2$ ;

Для непарних значень  $n$  і парних значень  $k$ :  $y = b_{13}B^2$ ;

Для непарних значень  $n$  і непарних значень  $k$ :

$$y = b_{12}B^2;$$

3. Для непарних значень  $n$   $y = b_{12}B^2 = b_{13}B^2$ , для парних –  $y = b_{11}B^2 = b_{13}B^2$ ;

4. Для непарних значень  $n$   $y = b_{11}B^2 = b_{12}B^2$ , для парних –  $y = b_{13}B^2$ ;

$$5. y = b_{11}B^2 = b_{12}B^2 = b_{13}B^2.$$

Оскільки абсолютну похибку оцінюють як найбільше з абсолютнох значень відхилення між істинним і вимірюваним значенням функції, то перевіряти знайдені стаціонарні точки на наявність у них локального мінімуму або максимуму непотрібно. Вибираємо серед знайдених значень функції у стаціонарних точках максимальне абсолютне значення, враховуючи умови рівності коефіцієнтів  $b_{11}, b_{12}, b_{13}$ .

У будь-якому випадку, максимальне значення функції (6) можна оцінити як  $\max \{|b_{11}|B^2, |b_{12}|B^2, |b_{13}|B^2\}$ , тобто, як максимальний з коефіцієнтів при функціях  $\cos^2 \alpha \sin^2 \beta, \sin^2 \alpha \sin^2 \beta, \cos^2 \beta$ .

Оцінимо максимальне значення функції (5).

Згрупуємо множники при  $B_x^2, B_y^2, B_z^2$ :

$$\begin{aligned} \Delta B_x &= \frac{1}{|A|} \cdot (b_{11}(a_{22}a_{33} - a_{23}a_{32}) + b_{21}(a_{13}a_{32} - a_{12}a_{33}) + \\ &+ b_{31}(a_{12}a_{23} - a_{13}a_{22}))B_x^2 + \\ &+ \frac{1}{|A|} \cdot (b_{12}(a_{22}a_{33} - a_{23}a_{32}) + b_{22}(a_{13}a_{32} - a_{12}a_{33}) + \\ &+ b_{32}(a_{12}a_{23} - a_{13}a_{22}))B_y^2 + \\ &+ \frac{1}{|A|} \cdot (b_{13}(a_{22}a_{33} - a_{23}a_{32}) + b_{23}(a_{13}a_{32} - a_{12}a_{33}) + \\ &+ b_{33}(a_{12}a_{23} - a_{13}a_{22}))B_z^2. \end{aligned} \quad (7)$$

Якщо виразимо  $B_x, B_y$  і  $B_z$  у сферичних координатах, легко побачити, що за аналогією з попереднім випадком (пошуку максимуму функції (6)), максимальним значенням  $\Delta B_x$  буде максимальне серед абсолютнох значень:

$$\begin{aligned} \frac{1}{|A|} \cdot |b_{11}(a_{22}a_{33} - a_{23}a_{32}) + b_{21}(a_{13}a_{32} - a_{12}a_{33}) + \\ + b_{31}(a_{12}a_{23} - a_{13}a_{22})|B_x^2, \end{aligned}$$

$$\begin{aligned} \frac{1}{|A|} \cdot |b_{12}(a_{22}a_{33} - a_{23}a_{32}) + b_{22}(a_{13}a_{32} - a_{12}a_{33}) + \\ + b_{32}(a_{12}a_{23} - a_{13}a_{22})|B_y^2, \end{aligned} \quad (8)$$

Похибки  $\Delta B_y$  і  $\Delta B_z$  знаходимо аналогічно.

Максимальне значення  $\Delta B_y$  можна оцінити як максимальне серед значень:

$$\begin{aligned} \frac{1}{|A|} \cdot & |b_{11}(a_{23}a_{31} - a_{21}a_{33}) + b_{21}(a_{11}a_{33} - a_{13}a_{31}) + \\ & + b_{31}(a_{13}a_{21} - a_{11}a_{23})|B^2, \\ \frac{1}{|A|} \cdot & |b_{12}(a_{23}a_{31} - a_{21}a_{33}) + b_{22}(a_{11}a_{33} - a_{13}a_{31}) + \\ & + b_{32}(a_{13}a_{21} - a_{11}a_{23})|B^2, \quad (9) \\ \frac{1}{|A|} \cdot & |b_{13}(a_{23}a_{31} - a_{21}a_{33}) + b_{23}(a_{11}a_{33} - a_{13}a_{31}) + \\ & + b_{33}(a_{13}a_{21} - a_{11}a_{23})|B^2. \end{aligned}$$

а максимальне значення  $\Delta B_z$  – як максимальне серед значень:

$$\begin{aligned} \frac{1}{|A|} \cdot & |b_{11}(a_{21}a_{32} - a_{22}a_{31}) + b_{21}(a_{12}a_{31} - a_{11}a_{32}) + \\ & + b_{31}(a_{11}a_{22} - a_{12}a_{21})|B^2, \\ \frac{1}{|A|} \cdot & |b_{12}(a_{21}a_{32} - a_{22}a_{31}) + b_{22}(a_{12}a_{31} - a_{11}a_{32}) + \\ & + b_{32}(a_{11}a_{22} - a_{12}a_{21})|B^2, \quad (10) \\ \frac{1}{|A|} \cdot & |b_{13}(a_{21}a_{32} - a_{22}a_{31}) + b_{23}(a_{12}a_{31} - a_{11}a_{32}) + \\ & + b_{33}(a_{11}a_{22} - a_{12}a_{21})|B^2. \end{aligned}$$

Нехай потрібно забезпечити точність обчислень модуля вимірюваної векторної величини, не гіршу за  $N\%$ , тобто, повинно виконуватися співвідношення:

$$\frac{\Delta B}{|B|} \leq N, \text{ що рівноважно } \frac{(\Delta B)^2}{B^2} \leq N^2 \text{ і } (\Delta B)^2 \leq B^2 N^2. \quad (11)$$

Виразимо похибку  $\Delta B$  через координати вектора  $B_x$ ,  $B_y$ ,  $B_z$  та їхні похибки:

$$\Delta B = \sqrt{(B_x + \Delta B_x)^2 + (B_y + \Delta B_y)^2 + (B_z + \Delta B_z)^2} - \sqrt{B_x^2 + B_y^2 + B_z^2}. \quad (12)$$

Піднесемо ліву та праву частину (12) до квадрату:

$$(\Delta B)^2 = (B_x + \Delta B_x)^2 + (B_y + \Delta B_y)^2 + (B_z + \Delta B_z)^2 - 2\sqrt{(B_x + \Delta B_x)^2 + (B_y + \Delta B_y)^2 + (B_z + \Delta B_z)^2} \left( B_x^2 + B_y^2 + B_z^2 \right) + (B_x^2 + B_y^2 + B_z^2). \quad (13)$$

Після алгебраїчних перетворень одержуємо:

$$(\Delta B)^2 = (B_x^2 + 2B_x\Delta B_x + \Delta B_x^2 + B_y^2 + 2B_y\Delta B_y + \Delta B_y^2 + B_z^2 + 2B_z\Delta B_z + \Delta B_z^2) - 2B\sqrt{(B_x + \Delta B_x)^2 + (B_y + \Delta B_y)^2 + (B_z + \Delta B_z)^2} \left( B_x^2 + B_y^2 + B_z^2 \right). \quad (14)$$

Перетворимо (14) наступним чином:

$$\begin{aligned} (\Delta B)^2 = & (B^2 + 2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2) - \\ & - 2B\sqrt{B^2 + 2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} + B^2. \quad (15) \end{aligned}$$

Підставимо (15) у (11):

$$\begin{aligned} (B^2 + 2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2) - \\ - 2B\sqrt{B^2 + 2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} + B^2 \leq (16) \\ \leq N^2 B^2. \end{aligned}$$

Зробимо заміну змінних  $q = \sqrt{B^2 + 2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2}$  і розв'яжемо нерівність

$$q^2 - 2Bq + B^2 (1 - N^2) \leq 0. \quad (17)$$

$$D = (-2B)^2 - 4B^2 (1 - N^2) = 4B^2 N^2,$$

$$q_{1,2} = \frac{2B \pm 2BN}{2} = B(1 \pm N).$$

Розв'язки нерівності:  $B(1 - N) \leq q \leq B(1 + N)$ .

Звідси

$$\begin{aligned} B^2 (1 - N)^2 - B^2 \leq & 2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z + \\ & + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2 \leq B^2 (1 + N)^2 - B^2. \quad (18) \end{aligned}$$

Представимо  $2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z$  як скалярний добуток векторів

$$2\begin{pmatrix} B_x & B_y & B_z \end{pmatrix} \cdot \begin{pmatrix} \Delta B_x & \Delta B_y & \Delta B_z \end{pmatrix} = 2B\sqrt{\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} \cdot \cos\varphi, \quad (19)$$

де  $\varphi$  – кут між векторами.

Найменше значення виразу

$$2B_x\Delta B_x + 2B_y\Delta B_y + 2B_z\Delta B_z + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2. \quad (20)$$

досягається при  $\cos\varphi = -1$ , а найбільше – при  $\cos\varphi = 1$ .

Звідси випливає, що для забезпечення умови (18) найменше можливе значення виразу (20) повинно бути не меншим за  $B^2 N(N - 2)$ , а найбільше – не більшим за  $B^2 N(N + 2)$ :

$$-2B\sqrt{\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2 \geq B^2 N(N - 2), \quad (21)$$

$$2B\sqrt{\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} + \Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2 \leq B^2 N(N - 2).$$

Після множення першої нерівності у (21) на  $-1$  отримуємо систему нерівностей:

$$2B\sqrt{\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} - (\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2) \leq -B^2 N(N - 2). \quad (22)$$

$$2B\sqrt{\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} + (\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2) \leq B^2 N(N + 2).$$

Після додавання лівих і правих частин цих нерівностей, відповідно, одержимо:

$$4B\sqrt{\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} \leq -B^2 N(N-2) + B^2 N(N+2),$$

або, після спрощення,

$$\sqrt{\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2} \leq BN. \quad (23)$$

Піднесемо вираз (23) до квадрату:

$$\Delta B_x^2 + \Delta B_y^2 + \Delta B_z^2 \leq B^2 N^2. \quad (24)$$

Як було показано вище, залежності (8), (9) та (10) дають достатньо строгі оцінки похибок  $\Delta B_x$ ,  $\Delta B_y$ ,  $\Delta B_z$ .

Однак, співвідношення (8), (9) та (10) містять величину  $B^2$ , а ліва частина співвідношення (24) – величину  $B^4$ , яка не є відомою. Після скорочення на  $B^4$  у лівій частині (24) для відомих коефіцієнтів  $a_{ij}$ ,  $b_{ij}$ , буде фіксоване число, у правій – величина  $\frac{N^2}{B^2}$ . Таким чином, підставляючи необхідні значення у формулу (24) і перевіряючи, чи виконується нерівність, приймаємо рішення про достовірність результатів вимірювання вказаної величини магнітної індукції із заданою похибкою. Інше застосування формул – за заданим значенням величини магнітної індукції можна визначити максимальну відносну похибку.

#### 4 ЕКСПЕРИМЕНТИ

Безпосередні експериментальні дослідження є недостатніми внаслідок того, що на практиці складно одержати достатню вибірку 3D-зондів магнітного поля з різними наборами коефіцієнтів вихідних характеристик і потрібно було б провести значну кількість вимірювань. Тому для верифікації одержаних результатів було використано імітаційне моделювання. За допомогою програми, розробленої у середовищі MATLAB, 1000 разів випадковим чином генерувалися значення коефіцієнтів  $a_{ij}$ ,  $b_{ij}$ , ( $i = \overline{1,3}$ ,  $j = \overline{1,3}$ ) вихідних характеристик (1), а для кожного набору коефіцієнтів 50 разів генерувалися випадкові значення компонент вектора магнітної індукції  $B_x$ ,  $B_y$ ,  $B_z$ . За формули (1) розраховувалися значення «вихідних сигналів» давачів  $S_1$ ,  $S_2$ ,  $S_3$ . Для кожного набору випадковим чином згенерованих чисел розраховувалися компоненти вектора магнітної індукції як розв’язки СЛАР (3), без урахування квадратичних складових вихідних характеристик. Далі обчислювалися абсолютні похибки кожної з компонент вектора магнітної індукції (як абсолютна різниця між розв’язком СЛАР (3) та випадковим чином згенерованими компонентами вектора магнітної індукції  $B_x$ ,  $B_y$ ,  $B_z$ ), а також абсолютна  $\Delta B$  і відносна  $\delta B$  похибки самого вектора магнітної індукції  $B$ . Для кожного набору згенерованих чисел за аналітичною залежністю (24) перевірялася можливість вимірювати вектор магнітної індукції з похибкою, не гіршою за задану, і результати зіставлялися з обчисленою відносною похибкою  $\delta B$ .

#### 5 РЕЗУЛЬТАТИ

Для всіх коефіцієнтів  $a_{ij}$ ,  $b_{ij}$ , ( $i = \overline{1,3}$ ,  $j = \overline{1,3}$ ), що задовільняли умові (24), обчислена відносна похибка  $\delta B$  не перевищувала задану відносну похибку. Однак, з-поміж знегерованих випадковим чином коефіцієнтів,  $a_{ij}$ ,  $b_{ij}$ , ( $i = \overline{1,3}$ ,  $j = \overline{1,3}$ ) траплялися такі, що не задовільняли умову (24) і, тим не менш, для жодного з 50 наборів компонент вектора магнітної індукції  $B_x$ ,  $B_y$ ,  $B_z$  відносна похибка  $\delta B$  не перевищувала задану.

#### 6 ОБГОВОРЕННЯ

Порівняльний аналіз результатів імітаційного моделювання з представленими у статті аналітичними розрахунками показали, що оцінки (8), (9) та (10) абсолютних похибок  $\Delta B_x$ ,  $\Delta B_y$ ,  $\Delta B_z$  є завищеними. Іншим можливим джерелом завищення оцінки є спрощення виразу (20), оскільки мало ймовірно, щоб вектори  $(B_x \ B_y \ B_z)$  та  $(\Delta B_x \ \Delta B_y \ \Delta B_z)$  були колінеарні або антиколінеарні. Отже, виконання умови (24) забезпечує можливість вимірювання вектора магнітної індукції із заданою точністю, однак, порушення умови (24) ще не означає відсутність такої можливості. Подальшим розвитком наведених у роботі результатів може бути менш строга і більш складна в обчисленні оцінка придатності 3D-зонда для практичного використання. Якщо ж встановлено, що система рівнянь (1) може мати розв’язки, що відрізняються між собою більше ніж на величину допустимої похибки, шляхом вирішення проблеми може бути структурна модифікація 3D-зонда, при якій до системи (1) додається одне чи більше рівнянь, введення яких обмежує множину розв’язків.

#### ВИСНОВКИ

Запропоновано метод оцінки придатності 3D-зонда магнітного поля на основі давачів Холла, вихідні характеристики яких залежать від усіх компонент вектора магнітної індукції та їхніх квадратів, для вимірювання вектора магнітної індукції з похибкою, що не перевищує задану. Вперше одержано аналітичну залежність (нерівність), яка дозволяє за коефіцієнтами вихідних характеристик 3D-зонда магнітного поля визначити, чи можна забезпечити вимірювання вектора магнітної індукції у заданих межах з допустимою похибкою за допомогою цього 3D-зонда. Якщо коефіцієнти вихідних характеристик сенсорів у складі 3D-зонда, задана величина магнітної індукції та максимальна допустима похибка задовільняють нерівність, результати вимірювання магнітного поля гарантовано перебуватимуть у межах допустимої похибки, інакше – потрібні додаткові дослідження.

Одержано аналітична залежність може бути застосована передусім для задач вимірювання магнітного поля за допомогою сенсорів на основі РХС. Однак, викладений математичний апарат може бути адаптований для сенсорів інших векторних величин.

#### ПОДЯКИ

Роботу виконано за сприяння Лабораторії магнітних сенсорів Національного університету «Львівська політехніка», що люб’язно надала результати експериментальних вимірювань вихідних сигналів дослідних взірців 3D-зондів магнітного поля на основі розщеплених холієвських структур.

## СПИСОК ЛІТЕРАТУРИ

1. Popovic R. S. Hall Effect Devices / R. S. Popovic. – Adam Hilger, Bristol, Philadelphia and New York, 2002. – 420 p.
2. Мікроелектронні сенсорні пристрой магнітного поля / [І. А. Больщакова, М. Р. Гладун, Р. Л. Голяка, З. Ю. Готра та ін.] ; за ред. З. Ю. Готри. – Львів : Вид. Національного університету «Львівська політехніка», 2001. – 412 с.
3. Нові конструкції напівпровідникових тонкоплівкових 3-D сенсорів магнітного поля / [І. А. Больщакова, Р. Л. Голяка, О. Ю. Макідо, Т. А. Марусенкова] // Электроника и связь. – 2009. – № 2–3. – С. 6–10.
4. Popovic R. S. Not-plate-like Hall magnetic sensors and their applications / R. S. Popovic // Sensors and Actuators A: Physical. – 2000. – Vol. 85, Issues 1–3. – P. 9–17.
5. Больщакова І. А. Польова характеристика сенсорів магнітного поля на розчлененіх холівських структурах / І. А. Больщакова, Р. Л. Голяка, Т. А. Марусенкова // Вісник

Марусенкова Т. А.

Канд. техн. наук, асистент, Национальный университет «Львовская политехника», Украина

## МЕТОД ОЦЕНКИ ПОГРЕШНОСТИ ИЗМЕРЕНИЯ МАГНИТНОГО ПОЛЯ 3D-ЗОНДОМ С НЕЛИНЕЙНЫМИ ВЫХОДНЫМИ ХАРАКТЕРИСТИКАМИ

В статье предложен метод оценки применимости результатов измерения магнитного поля с помощью 3D-зонда, состоящего из трех сенсоров, выходные характеристики которых зависят от всех компонентов вектора магнитной индукции и их квадратов. Выведены аналитические зависимости, позволяющие проверить возможность обеспечить измерение вектора магнитного поля с относительной погрешностью не больше заданной с помощью 3D-зонда магнитного поля с известными коэффициентами выходных характеристик.

**Ключевые слова:** 3D-зонды магнитного поля, датчики Холла, погрешность, измерение.

Marusenкова Т. А.

PhD, Assistant, Lviv Polytechnic National University, Ukraine

## METHOD OF EVALUATING INACCURACY OF MAGNETIC FIELD MEASUREMENT USING A 3D-PROBE WITH NONLINEAR OUTPUT CHARACTERISTICS

The work proposes a method allowing to check whether the desired accuracy of magnetic flux-density vector's measurement can be achieved using a magnetic 3D probe based on three sensors with output characteristics dependent on all the magnetic-flux density vector's components and the squared components. As the result of the method we've obtained analytical formulas that allow to predict the 3D-probe's applicability for measuring the magnetic field's magnitude with the acceptable inaccuracy.

**Keywords:** magnetic 3D probes, Hall devices, inaccuracy, measurement.

## REFERENCES

1. Popovic R. S. Hall Effect Devices. Adam Hilger, Bristol, Philadelphia and New York, 2002, 420 p.
2. Bol'shakova I. A., Gladun M. R., Goljaka R. L., Gotra Z. Ju. ta in. Mikroelektronni sensorni prystroi' magnitnogo polja; za red. Z. Ju. Gotry. L'viv. Vyd. Naciona'l'nogo universytetu «L'viv's'ka politehnika», 2001, 412 p.
3. Bol'shakova I. A., Holyaka R. L., Makido O. Yu., Marusenкова Т. А. Novi konstruktsiyi napivprovodnykovykh tonkoplivkovykh 3-D sensoriv mahnitnoho polya, *Elektronika y svyaz'*, 2009, No. 2–3, pp. 6–10.
4. Popovic R. S. Not-plate-like Hall magnetic sensors and their applications, *Sensors and Actuators A: Physical*, 2000, Vol. 85, Issues 1–3, pp. 9–17.
5. Bol'shakova I. A., Holyaka R. L., Marusenкова T. A. Pol'ova kharakterystyka sensoriv mahnitnoho polya na rozchlenenixkh kholliv's'kykh strukturakh, *Elektronika. Visnyk Natsional'noho universytetu «L'viv's'ka politehnika»*, 2010, No. 681, pp. 66–75.
6. Chaudhari M., Dharavath S. Study of Smart Sensors and their Applications, *International Journal of Advanced Research in Computer and Communication Engineering*, 2014, Vol. 3, Issue 1, pp. 5031–5034.
7. Sturmels B. What is a Gröbner Basis? // Notices of the AMS. – 2005. – Vol. 52. – P. 2–3.
8. Adams W. W. Introduction to Gröbner Bases. Graduate Studies in Mathematics, American Mathematical Society, Providence, R. I., 1994.
9. Marusenkova T. Approach To Roots Separation For Solving Nonlinear Equations On ARM Cortex-Based Microcontrollers [Text] / T. Marusenkova, I. Yurchak // «CAD in Machinery Design. Implementation and Educational Issues» : XXII Ukrainian-Polish Conference : prxeedings. – Lviv, 2014. – P. 101–103.

Стаття надійшла до редакції 04.11.2014.

© 2014

УДК 621.37/39.029.3

Невлюдов І. Ш.<sup>1</sup>, Андрусевич А. А.<sup>2</sup>, Аллахверанов Р. Ю.<sup>3</sup><sup>1</sup>Д-р техн. наук, професор, завідувач кафедрою, Харківський національний університет радіоелектроники, Україна<sup>2</sup>Д-р техн. наук, доцент, начальник Криворізького коледжу Національного авіаційного університету, Україна<sup>3</sup>Канд. техн. наук, Харківський національний університет радіоелектроники, Україна

## МОНІТОРИНГ ОБЕСПЧЕНИЯ ЖИЗНЕННОГО ЦИКЛА РЭС ПРИ ЭКСПЛУАТАЦІЇ

В статье рассмотрен подход к мониторингу процессов жизненного цикла радиоэлектронных средств (ЖЦ РЭС) при эксплуатации. Объектом исследования являются процессы жизненного цикла радиоэлектронных средств на этапе эксплуатации радиоэлектронной аппаратуры. Предметом исследования являются методы и модели извлечения информации о процессах жизненного цикла радиоэлектронных средств на этапе эксплуатации. Целью исследований является разработка методов и моделей процессов мониторинга технического ресурса РЭС при ее эксплуатации. Главный научный результат заключается в усовершенствовании модели процесса изменения технического ресурса на стадии эксплуатации РЭС, основанной на термодинамическом описании физико-химического механизма явлений массопереноса и структурных превращений в материалах, формирующих свойства РЭС и значения информативных параметров РЭС, что дало возможность приводить оценку кинетики деградационных процессов и динамики ресурсных характеристик РЭС. На базе модели изменения технического ресурса разработан метод визуализации технического состояния и ресурсных характеристик РЭС, основанный на геометрической интерпретации ситуации, допускающей аналогию описания кинетики деградационных процессов и изменения параметров РЭС, что дало возможность создать математическое и программное обеспечение систем, позволяющих на основе визуализации изменения технического ресурса прогнозировать отказы РЭС. Проведены эксперименты по исследованию свойств предложенного метода. Результаты экспериментов позволяют рекомендовать предложенный метод для использования на практике.

**Ключевые слова:** мониторинг, жизненный цикл, прогнозирование, радиоэлектронные средства.

### НОМЕНКЛАТУРА

$Z(t, T)$  – ресурс объектов;

$T$  – характеристика взаимодействия объектов со средой (нагрузка);

$P(t, T)$  – вероятность безотказной работы;

$\lambda(t, T)$  – интенсивность отказов как скорость расходования ресурса в статистическом смысле, не противоречит его отображению в принятой концепции;

$I_j$  – поток составляющих термодинамической среды (скорость изменения параметров, характеризующих состояние среды, вещества, зарядов, тепла и т. д.), который определяет скорость изменения параметров  $V_j$ ;

$F_{kj}$  – термодинамические силы (градиенты плотности, напряжений, температуры и т. д.);

$M_k$  – линейные феноменологические коэффициенты;

$y_j(0)$  – неравновесные или стационарные значения соответствующих параметров системы в зависимости от рассматриваемого масштаба эволюции;

$L_{jk}$  – кинетические коэффициенты;

$\tau_k$  – времена релаксации соответствующих необратимых процессов в системе;

$u$  – макроскопический параметр системы;

$\Delta u$  – интервал изменения  $u$ ;

$C_i$  – концентрация участвующей в реакции  $i$ -тої компоненты;

$r_i$  – суммарный порядок реакции по  $i$ -тому веществу, он может равняться нулю, быть целым или дробным числом;

$r$  – порядок реакции, он может равняться нулю или единице для линейной и двум для нелинейной модели;

$D_n(\Delta V)$  – дисперсия ошибок на момент  $t_n$ ;

$\delta(t - \tau)$  – дельта-функция Дирака;

$v(t)$  – процесс, обусловленный погрешностями наблюдения;

$D_v$  – дисперсия обусловленная погрешностями наблюдения.

### ВВЕДЕНИЕ

К числу наиболее важных функций мониторинга, реализуемых в настоящее время, относится контроль и прогнозирование состояния РЭС и процессов обеспечения ее жизненного цикла (ЖЦ) [1]. Для сложных систем, в том числе и ЖЦ РЭС, возникает необходимость принятия решений в ситуации отсутствия формальных методов постановки и решения задач, возникающих в ЖЦ РЭС.

В свою очередь в центре внимания мониторинга в период эксплуатации РЭС могут находиться процессы, обеспечивающие и ограничивающие живучесть РЭС.

Очевидно, что в первую очередь представляют интерес деградационные процессы. Главным проявлением ограниченности ресурсных характеристик РЭС являются дефекты, возникающие и развивающиеся в течение эксплуатации РЭС.

В настоящее время для прогнозирования состояния РЭС, обусловленного развитием дефектов предлагаются методы, основанные на интерполяции случайных реализаций временных функций, получаемые в процессе наблюдения за параметрами в начальной и последующих стадиях жизненного цикла РЭС [2]. Однако достоверность такого прогноза незначительна, и не отвечает требованиям при решении практических задач обеспечения надежности РЭС. Предлагается применить метод прогнозирования основанный на принципах распознавания, где может быть задействовано лицо, принимающее решение, которое наблюдает за процессом и принимает решение о состоянии РЭС на основе этих наблюдений и анализа ситуации. Этот подход принят в качестве основного для прогнозирования процессов, происходящих в сложных системах и обладает достоинствами, связанными с возможностями принятия гарантированного решения и качеством прогноза [3].

Поэтому целью исследований является разработка более совершенных методов и моделей мониторинга технического ресурса РЭС полнее отражающих процессы происходящие в РЭС при ее эксплуатации для обеспечения возможности принятия эффективных решений.

Для достижения поставленной цели в рамках данной работы необходимо:

- провести анализ возможных дефектов РЭС, методов и средств их обнаружения;
- разработать методологию моделирования процесса расходования ресурса РЭС и его отображения;
- разработать модель описания деградационных процессов и процесса расходования ресурса РЭС.

## 1 ПОСТАНОВКА ЗАДАЧИ

Для описания процесса расходования ресурса технических систем представляется возможным использование термодинамического подхода [4, 5, 6]. Термодинамический подход, основанный на необратимости физических процессов, объединяет различные точки зрения: уравнения баланса, классическую термодинамическую теорию устойчивости, флуктуационную и эволюционную теорию.

Определение ресурса как функционала от показателей надежности, в частности

$$Z(t, T) = -\ln P(t, T) = \int_0^t \lambda(t, T) dt .$$

С физической точки зрения процесс расходования ресурса является интегральным процессом необратимого изменения (эволюции или деградации) термодинамического состояния каждого объекта, т. е. процессом производства энтропии. Учитывая термодинамический

характер феноменологических моделей надежности РЭС, производство энтропии  $dS / dt$  характеризует все многообразие необратимых физико-химических процессов для реальных условий  $T$  взаимодействия каждого объекта со средой и выступает, таким образом, в качестве интегральной скорости расходования ресурса и может иметь такой же характер как  $\lambda(t, T)$ .

С позиции термодинамического подхода для описания процессов, происходящих в отображаемой среде, сопровождаемых изменением ее внутренних параметров  $V_j$ , характеризующих состояние среды, можно использовать обобщенное уравнение Онзагера [4, 7]

$$I_j = \frac{dV_j}{dt} = \sum_{k=1}^m M_k F_{kj} . \quad (1)$$

Примерами феноменологических коэффициентов могут служить коэффициенты диффузии, теплопроводности и т. п.

Фундаментальным принципом термодинамического подхода является справедливость выражения (1) для всех видов параметров, характеризующих состояние среды (энтропии, количества теплоты и т. д.). В разрабатываемой модели процесса в качестве параметра  $V$ , характеризующего состояние среды, предлагается использовать ВР, которая характеризуется объемом в многомерном или площадью в двухмерном пространствах.

## 2 ОБЗОР ЛИТЕРАТУРЫ

Анализ наиболее характерных дефектов и проявления их при эксплуатации РЭС дает основание предположить, что в основе механизмов возникновения и развития дефектов РЭС лежат преобразования микро- и макроструктуры исходных материалов и химические превращения, происходящие в производстве и эксплуатации РЭС [8, 9, 10]. При этом реализуется схема превращения субмикродефектов в макродефекты.

Анализ возможностей средств обнаружения дефектов и данных по причинам отказов РЭС показывает, что значительная часть дефектов может быть не обнаружена. Поэтому уделяют внимание прогнозированию параметрических отказов в процессе производственных испытаний и технического обслуживания РЭС, в результате принимают решение о техническом состоянии и технологии производства приборов и элементной базы РЭС.

Получение прогнозных результатов при действии переменных внешних факторов повышает эффективность прогноза, делает его действенным инструментом управления производством РЭС.

Однако, современные методы прогнозирования базирующиеся на функциональном анализе, теории рядов, теории экстраполяции и интерполяции, теории вероятностей и математической статистики, теории случайных функций и случайных процессов, корреляционном и спектральном анализе, теории распознавания образов [11–15] не учитывают механизм процессов, приводящих

к изменению состояния РЭС. Достоверность такого анализа информации в ряде случаев не может отвечать требованиям при решении практических задач обеспечения качества РЭС. Рассматриваемый метод, основанный на участии ЛПР, обладает достоинствами связанными с возможностью принятия решений о гарантированном прогнозе отказов РЭС, обусловленных развитием производственных дефектов [8].

### 3 ПРОЦЕССЫ И МЕТОДЫ

Два основных вида процессов расходования ресурса соответствует эволюционным и флюктуационным процессам, происходящим в термодинамической среде. Первый вид имеет характер флюктуаций параметров системы за счет внешних или внутренних факторов. Такой процесс является случайным и может иметь как обратимый, так и необратимый характер. Другой вид связан с эволюционными процессами, т. е. с процессами производства и накопления энтропии объекта. Такой процесс является детерминированным и имеют характер необратимой деградации.

Физическую основу эволюционных процессов составляют неравновесные состояния системы и соответствующие им необратимые процессы. Конечное значение производной энтропии по времени дает основание предположить, что процессы эволюции термодинамических систем носят монотонный релаксационный характер. Этот вывод совпадает с выводами, наблюдаемыми статистической теорией. Уравнения Онзагера при определенных допущениях, в частности при условии стабилизации термодинамических сил могут быть представлены как линейные дифференциальные уравнения во временной области относительно экстенсивных параметров  $y_j$  с постоянными коэффициентами  $L_{jk}$ , решения которых записываются в виде

$$y_j(t) - y_j(0) = \sum_k [L_{jk} \exp(-t/\tau_k)].$$

Очевидно, в оценку ресурса систем включается процесс измерений параметров. Необратимый характер процессов измерения как процессов взаимодействия измерительного прибора и объекта обуславливает определенную взаимосвязь (предельные соотношения) между термодинамическими характеристиками (энергия, энтропия) и информационными (точность, количество информации).

Для флюктуаций, вызванных наблюдениями, существует понятие характерного интервала или минимального масштаба времени, которое определяет масштаб флюктуации

$$\tau = \frac{\Delta y}{\partial y / dt}.$$

Из этого выражения видно, что этот масштаб может изменяться в больших пределах и измерение является частью отображаемых эволюционных и флюктуационных процессов.

Таким образом, при отображении процессов развития дефектов должен учитываться эволюционный и флюктуационный характер процессов.

Статистическая физика дает возможность объяснить кинетику термодинамических параметров среды на основе поведения ансамбля частиц, из которых эта среда состоит. Микроскопическое состояние ансамбля полностью задается каноническими переменными ( $X$ ). С макроскопической же точки зрения состояние вещества определяется весьма ограниченным числом параметров, достаточных для макроскопической характеристики среды. Задание этих параметров, измеряемых в макроскопическом опыте, определяет макроскопическое состояние среды. Макроскопические параметры, в том числе объем ВР являются функциями канонических переменных  $V_k(X)$ , причем  $k = 1, 2, \dots, n$ , где  $n \ll N$ .

Исходя из макроскопических измерений, можно сделать только статистические суждения о значениях микроскопических переменных  $X$ . Таким образом, макроскопически задаваемая система изображается посредством задания плотности вероятности переменных  $\omega(X, t)$ . Зная  $\omega(X, t)$ , можно вычислить статистическое среднее значение  $V(X)$  согласно формуле

$$V = \int \limits_{(X)} V(X) \omega(X, t) dX,$$

а также среднее квадратическое отклонение

$$\Delta V = \sqrt{\overline{(V - \bar{V})^2}},$$

проявляемое как флюктуация наблюдаемой области ВР.

Рассматривая процесс превращения вещества как физико-химическую реакцию, оценивая скорость процесса через изменение объема ВР или ИР в единицу времени, можно предположить, что скорость изменения концентрации  $j$ -той компоненты

$$\frac{dC_j}{dt} = K \prod_i C_i^{r_j}.$$

Для адекватного описания процессов предлагается ограничиваться порядками, равными 0, 1, 2. Константа скорости  $K$  является одним из основных параметров процесса, определяя, по сути, время релаксации и поэтому называется постоянной релаксации.

Для нашего отображения приняты следующие гипотезы:

- двухкомпонентный состав среды;
- зависимость скорости протекания процесса от объема ВР  $W_0 - V$ , где  $W_0$  – начальный объем среды, а  $V$  – объем ВР;

- правило пропорциональности между количеством вещества и его концентрацией при распределении по объему среды, тогда рассматриваемый объем среды и ВР можно воспринимать как концентрацию компонентов, участвующих в реакциях;

– предположение о том, что при протекании реакций термодинамические коэффициенты, также могут зависеть от состояния среды, отражая возможность использования уравнения Онзагера для нелинейных процессов.

Тогда уравнение процесса, объединяющего термодинамический и физико-химический подходы, можно представить в виде

$$\frac{dV}{dt} = K(W_0 - V)^r = f(V). \quad (2)$$

Решение этого уравнения может быть использовано для нахождения детерминированной части временной зависимости объема ВР.

Используя уравнение (2) для выделенных механизмов процессов развития дефектов можно сделать вывод:

– диффузии при  $D$ , пропорциональном концентрации  $W_0 - V$ ,  $r = 2$ , при постоянном  $D$  можно принять  $r = 1$  или  $r = 0$ ;

– различных видов коррозии, с учетом зависимости энергии активации и постоянных реакций от концентрации ИР  $r = 2$ , для линейного описания  $r = 0$  или  $r = 1$ ;

– испарении материала  $r = 0$ .

Так как результатом процесса является  $W_0 = V$ , можно предположить, что  $W_0$  определяется системой равенств

$$\left. \begin{array}{l} W_0 = a^i a^j \\ a^i = x^i(t_n) - x^i_{\min}, \text{ если } \delta x^i < 0 \\ a^i = x^i_{\max} - x^i(t_n), \text{ если } \delta x^i \geq 0 \end{array} \right\}.$$

Отсутствие информации о порядке реакции приводит к задаче его выбора на основе наблюдений. Критерий выбора может быть максимальное значение ошибок  $\Delta V$ , получаемых как разность между наблюдаемым и вычисленным размером области ВР.

На основе интервального оценивания математического ожидания  $M_n(\Delta V)$  ошибок на момент  $t_n$  и того, что  $\max_n \Delta V$  располагается на границе полученного интервала, для одинаковых условий выбора для всех моделей, характеризующихся достоверностью оценок

$$\max_n \Delta V = M_n(\Delta V) + \frac{\sqrt{D_n(\Delta V)}}{\sqrt{n}}.$$

Так как целью такого мониторинга является предсказание, то можно воспользоваться теорией экстраполяции случайных процессов, в частности методами оптимальной фильтрации.

Особенностью решения задачи, связанной с применением оптимальной фильтрации для выделения и предсказания случайной составляющей рассматриваемой модели, является нелинейность эволюционных уравнений. Предлагается линеаризация задачи на основе концепции опорной траектории и разложения нелинейной функции  $f(V)$  в ряд Тейлора. В качестве опорной траек-

тории, достаточно близкой к истинной, используется решения  $V_{on}(t)$  дифференциальных уравнений для различного порядка реакций, тогда оцениванию и предсказанию подлежит случайная аддитивная поправка  $Z$  к детерминированному  $V_{on}$ , которое является решением дифференциальных уравнений и для истинной траектории  $V(t)$ , характеризующей деградационный процесс

$$V(t) = V_{on}(t) + Z(t).$$

Выдвинув гипотезу о нормальности закона распределения, окончательная форма уравнения эволюции примет вид

$$\frac{dV_{on}}{dt} + \frac{dz}{dt} = f(V_{on}) + \frac{df(V)}{dV} \Delta V.$$

$f(V_{on})$  – левая часть дифференциальных уравнений, соответствует система уравнений

$$\frac{dV_{on}}{dt} = f(V_{on}),$$

полагая  $z = \Delta V$

$$\frac{dz}{dt} = \frac{df(V)}{dV} z.$$

Рассматривая решение задачи предсказания случайной составляющей на основе алгоритма экстраполяции Калмана, предусматривающего использование моделей эволюции случайной составляющей и наблюдения в дискретные моменты времени задаваемые номером  $k$ . Начальные условия определяются при  $k = 0$ , при  $k = 1$  выполняется наблюдение, прогнозирование осуществляется для  $k = 2$ . Эти модели определяются уравнениями

$$z(k+1) = \frac{df(V)}{dV} z(k) + w(k),$$

$$g(k) = z(k) + v(k).$$

Здесь  $w(t)$  нормальный белый шум с нулевым средним значением  $\text{cov}\{\omega(t), \omega(\tau)\} = d_w \delta(t - \tau)$ . Можно предположить, что начальные значения случайной составляющей  $z(0)$  и начальная дисперсия  $d_w$  определяются средними значениями  $z(t)$  и дисперсией, полученными в предыдущие моменты наблюдений.

Уравнения экстраполяции имеют вид

$$z(k+1/k) = \frac{df(V)}{dV} z(k/k-1) + K(k+1, k) \cdot [g(k) - z(k/k-1)],$$

с коэффициентом усиления  $K(k+1, k)$ , определяемым из уравнения

$$K(k+1, k) = \frac{df(V)}{dV} D(k/k-1) [D(k/k-1) + D_v]^{-1}.$$

Уравнение дисперсии имеет вид

$$D(k+1/k) = \left( \frac{df(V)}{dV} \right)^2 D(k/k-1) + d_w - \\ - \left( \frac{df(V)}{dV} D(k/k-1) \right)^2 [D(k/k-1) + D_V]^{-1}.$$

Очевидно, полученные выражения по результатам наблюдений  $g(k)$  в момент времени, соответствующий  $k=1$ , позволяют оценить значение  $z(\delta t_p) = z(k+1/k)$  и  $D(\delta t_p) = D(k+1/k)$  через интервал времени прогноза  $\delta t_p$ , соответствующий  $k+1=2$ , т.е. сделать прогноз развития процесса и оценить ошибку прогноза.

#### 4 ЕКСПЕРИМЕНТИ

Основные положения экспериментальных работ включают: методику получения достаточного количества образцов, имеющих дефекты; экспериментальную проверку процесса отображения информации, модели изменения ресурса; сравнение предлагаемых методов прогнозирования с существующими, основанными на анализе случайных реализаций временных функций, получаемых в процессе наблюдения за параметрами РЭС, с применением МНК.

На основе проведенного теоретического обоснования был разработан алгоритм отображения области ВР с учетом изменения ее размеров и формы на интервале предсказания. В результате на экране монитора отображается прямоугольная область, размеры и форма которой определяются в соответствии с предлагаемой моделью, и отображается канва (можно использовать цветовую палитру монитора), ширина которой соответствует интервалу ошибки предсказания при заданной гарантированности. Приведенный алгоритм дает возможность наблюдать на экране монитора ситуацию, возникающую каждый раз, когда после измерения параметров необходимо принимать решение о возможности дальнейшей корректировке технологического процесса производства изделий. Наблюдаемое изображение в достаточной мере характеризует размер ВР и ее расположение относительно границ допустимой области изменения параметров. Характеристика ВР (отображение размытости границ области ВР) дает возможность оценить гарантированную ошибку предсказания.

#### 5 РЕЗУЛЬТАТИ

На рис. 1 приведены результаты отображения изменения параметров в виде графиков временных функций соответствующего процесса. Сплошными линиями приводятся графики уравнений регрессии, полученные с помощью традиционного метода МНК.

На рис. 2 отображена ситуация, когда после ускоренных испытаний в течение 389 часов было принято решение об отказе через  $50 \pm 15$  ч, фактически отказ произошел в интервале 442–447 ч, т.е. в худшем случае через 58 ч. Здесь с помощью разработанной программы отображено 64 прямоугольника, соответствующие элементарной области ВР.

В целом площадь затемненного участка равна части ВР и характеризует динамику изменения ресурса, РЭС. Координаты исходных вершин прямоугольников определялись после измерения параметров в моменты времени, разделенные промежутками 5 и 9 ч. Размеры и форма прямоугольников, дающие представление об области ВР, определяются согласно модели. Ширина отображаемой канвы соответствует интервалу ошибки предсказания при заданной гарантированности и вычислялась по разработанному алгоритму.

На рис. 3 отображена ситуация, характеризуемая стабильностью состояния РЭС. Здесь часть ВР, наблюдае-

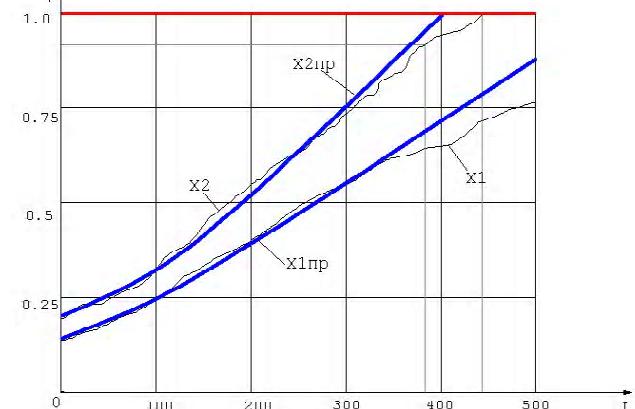


Рисунок 1 – Отображение изменения параметров в виде графиков временных функций

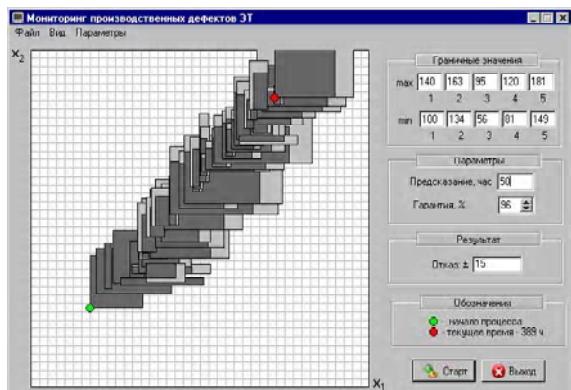


Рисунок 2 – Отображение процесса расходования ресурса РЭС

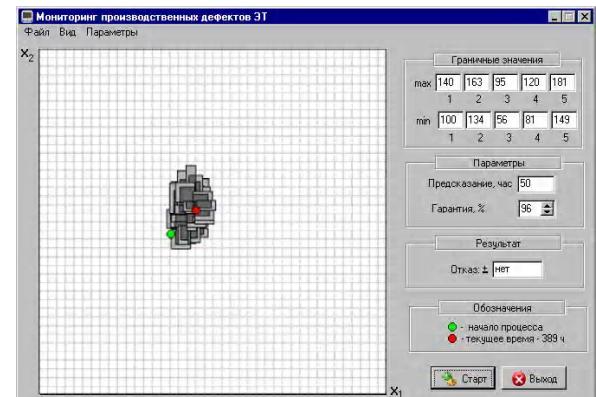


Рисунок 3 – Отображение процесса изменения технического состояния РЭС

мая также в течение 389 ч, занимает небольшую область, что дает основание сделать предположение об отсутствии дефектов и небольшой вероятности параметрических отказов на интервале предсказания 50 ч.

## 6 ОБСУЖДЕНИЕ

Экспериментальные работы и последующий дисперсионный анализ дают основание сделать вывод о том, что на уровне достоверности 0,96 вероятность выхода ошибки за прогнозируемые пределы составила 0,07, интервальная оценка точности прогноза в приведенном примере на уровне достоверности 0,04 составила 20 %, такая же оценка по МНК составила 38 %. В условиях проводимого в работе эксперимента точность прогнозирования была в 1,9 раз выше.

Сравнение отображений процесса, полученного на рис. 1, с рис. 2 и рис. 3 свидетельствует в пользу предложенного метода мониторинга, в части наглядности и достоверности прогнозирования. Что касается гарантии прогнозирования, то кроме количественной характеристики ее можно просматривать на экране монитора в виде вышеупомянутой канвы. Отображение процесса на экране кроме этого характеризует значение наблюдаемых в данный момент времени параметров, техническое состояние и ресурс РЭС, что дает основание в достаточно эффективном совершенствовании мониторинга РЭС.

Аналогичные результаты наблюдались при отображении и предсказании изменения параметров различных изделий РЭС, они находились в рамках установленных возможностей предлагаемой модели.

## ЗАКЛЮЧЕНИЕ

В статье решается актуальная задача разработки теоретических основ мониторинга для обеспечения возможности принятия эффективных решений при поддержке ЖЦ РЭС на базе более совершенных моделей, полнее отражающих происходящие в ЖЦ РЭС процессы.

Научная новизна работы состоит в предложенной модели процесса изменения технического ресурса на стадии эксплуатации РЭС, основанной на термодинамическом описании физико-химического механизма явлений массопереноса и структурных превращений, в материалах, формирующих свойства РЭС и значения информативных параметров РЭС. Это дало возможность проводить оценку кинетики деградационных процессов, динамики ресурсных характеристик РЭС и разработать метод визуализации технического состояния и ресурсных характеристик РЭС, основанный на геометрической интерпретации ситуации, допускающей аналогию описания кинетики деградационных процессов и изменения параметров РЭС.

Практическая ценность полученных результатов определяется разработанными математическими и программными средствами, которые реализованы в соответствии с разработанными теоретическими основами мониторинга жизненного цикла радиоэлектронных средств на этапе эксплуатации.

## БЛАГОДАРНОСТИ

Работа выполнена в Харьковском национальном университете радиоэлектроники (ХНУРЭ) и Криворожском авиационном колледже в рамках госбюджетных тем: №189-1 «Теоретические основы логистических систем в технологии автоматизированного производства радиоэлектронного приборостроения» (ДР № 0105U002739); № 224 «Конструкторско-технологические основы создания перспективных компонентов микроэлектромеханических систем и технологий их производства» (ДР № 0108U002216); № 248-1 «Теоретичні основи мікроелектромеханічних систем, проектування та технології їх виробництва для гнучких інтегрованих систем» (ДР № 0110U002594) на основании приказа Министерства образования и науки Украины № 686 от 22.07.09 г.

## СПИСОК ЛИТЕРАТУРЫ

1. Шептунов С. А. Жизненный цикл продукции / С. А. Шептунов. – М. : Янус-К, 2003. – 244 с.
2. Давыдов П. С. Техническая диагностика радиоэлектронных устройств и систем / П. С. Давыдов. – М. : Радио и связь, 1991. – 256 с.
3. Стрельников В. П. Оценка и прогнозирование надежности электронных элементов систем / В. П. Стрельников, А. В. Федухин. – К. : Логос, 2002. – 486 с.
4. Лоренц Г. А. Статистические теории в термодинамике / Г. А. Лоренц. – М. : РХД, 2001. – 184 с.
5. Андрусович А. А. Гетерогенные модели развития производственных дефектов / А. А. Андрусович, И. Ш. Невлюдов, Ю. М. Роздоловский // Науч.-техн. журнал. Авиационно-космическая техника и технология. – 2003. – № 38. – С. 114–119.
6. Невлюдов И. Ш. Термодинамическая модель жизненного цикла электронной аппаратуры / И. Ш. Невлюдов, А. А. Андрусович, С. В. Сотник // Вестник Академии инженерных наук Украины. – 2007. – № 3 (33). – С. 132–135.
7. Пригожин И. Введение в термодинамику необратимых процессов / И. Пригожин. – М. : РХД, 2001. – 160 с.
8. Мониторинг процессов проектирования, производства и эксплуатации жизненного цикла электронной аппаратуры: монография / А. А. Андрусович, И. Ш. Невлюдов. – Харьков : ФЛП Цуварев А. Ф, 2009. – 272 с.
9. Ефимов И. Е. Микроэлектроника: физические и технологические основы, надежность / И. Е. Ефимов, Ю. И. Горбунов, И. Я. Козырь. – М. : Высшая школа, 1986. – 463 с.
10. Маллер Р. Элементы интегральных схем : пер. с англ. / Р. Маллер, Т. Кейминс. – М.: Мир, 1989. – 630 с.
11. Горелова В. Л. Основы прогнозирования систем / В. Л. Горелова, Е. Н. Мельникова. – М. : Вышш. шк., 1986. – 287 с.
12. Чаплыгин Д. Ю. Имитационная модель динамики отказов и восстановления работоспособности сложных радиоэлектронных систем / Д. Ю. Чаплыгин, П. Б. Абрамов, В. В. Цветков // Математическое моделирование систем обработки информации и управления: Сборник научных трудов. – Воронеж : ин-т МВД России, 2001. – С. 14–19.
13. Воронцов В. Н. Контроль качества и прогнозирование надежности изделий электронной техники по электрофизическим параметрам. докт. техн. наук / В. Н. Воронцов. – Санкт-Петербург : Сев.-Зап. гос. заоч. техн. ун-т, 2002. – 39 с.
14. Радаев Н. Н. Повышение точности прогноза событий, инициирующих чрезвычайные ситуации / Н. Н. Радаев // Измерительная техника. – 2002. – № 5. – С. 13–20.
15. Крищук В. М. Вибір елементів при компенсації зовнішніх впливів / В. М. Крищук, Г. М. Шилу А. О. Намлинський, М. П. Гапоненко // Радіоелектроніка, інформатика, управління. – 2004. – № 2(12). – С. 36–41.

Статья поступила в редакцию 24.10.2014.  
После доработки 07.11.2014.

Невлюдов І. Ш.<sup>1</sup>, Андрусевич А. О.<sup>2</sup>, Аллахверанов Р. Ю.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, професор, завідуючий кафедрою, Харківський національний університет радіоелектроніки, Україна

<sup>2</sup>Д-р техн. наук, доцент, начальник Криворожського коледжу Національного авіаційного університету, Україна

<sup>3</sup>Канд. техн. наук, Харківський національний університет радіоелектроніки, Україна

## МОНІТОРИНГ ЗАБЕЗПЕЧЕННЯ ЖИТТЕВОГО ЦИКЛУ РЕЗ ПРИ ЕКСПЛУАТАЦІЇ

У статті розглянуто підхід до моніторингу процесів життєвого циклу РЕЗ при експлуатації. Об'єктом дослідження є процеси життєвого циклу радіоелектронних засобів на етапі експлуатації радіоелектронної засобів. Предметом дослідження є методи та моделі здобування інформації про процеси життєвого циклу радіоелектронних засобів на етапі експлуатації. Метою дослідження є розробка методів і моделей процесів моніторингу технічного ресурсу РЕЗ при її експлуатації. Основним науковим результатом є удосконалення моделі процесу зміни технічного ресурсу на стадії експлуатації РЕЗ, яка характеризується використанням термодинамічного опису фізико-хімічного механізму явищ масопереносу та структурних перетворень у матеріалах, формуючих властивості РЕЗ і значення інформативних параметрів РЕЗ, що дало можливість проводити оцінку кінетики деградаційних процесів і динаміки ресурсних характеристик РЕЗ. На базі моделі зміни технічного ресурсу розроблено метод візуалізації технічного стану і ресурсних характеристик РЕЗ, заснований на геометричній інтерпретації ситуації, яка допускає аналогію опису кінетики деградаційних процесів і зміни параметрів РЕЗ, що дало можливість створити математичне і програмне забезпечення систем, дозволяючих на основі візуалізації зміни технічного ресурсу прогнозувати відмови РЕЗ. Проведено експерименти по дослідженю властивостей запропонованого методу. Результати експериментів дозволяють рекомендувати запропонований метод для використання на практиці.

**Ключові слова:** моніторинг, життєвий цикл, прогнозування, радіоелектронні засоби.

Nevliudov I. Sh.<sup>1</sup>, Andrusevich A. A.<sup>2</sup>, Allahveranov R. U.<sup>3</sup>

<sup>1</sup>Dr.Sc., Professor, Head of department, Kharkov National University of Radio Electronics, Ukraine

<sup>2</sup>Dr.Sc., Associate Professor, Head of the Krivoy Rog College National Aviation University, Ukraine

<sup>3</sup>PhD., Kharkov National University of Radio Electronics, Ukraine

## MONITORING OF LIFECYCLE OF RADIO ELECTRONIC EQUIPMENT IN EXPLOITATION

The article describes the approach to the monitoring of life cycle processes in the operation of radio electronic equipment (REE). The object of research is the process of the life cycle of electronic resources during the operational phase of electronic equipment. The subject of the study are methods and models of extracting information about the processes of the life cycle of electronic resources during the operational phase. The purpose of research is to develop methods and models of processes for monitoring the service life of the appliance REE. The main scientific result is to improve the process model changes of a technical resource to the operation stage REE based on the thermodynamic description of the physico-chemical mechanism of mass transfer phenomena and structural transformations in materials forming properties of the REE and REE values informative parameters, which enabled the assessment of the result of the kinetics of degradation processes and dynamics of resource characteristics of REE. As part of the main scientific result has been further developed model of changes in the technical resources at the operational stage REE, which is based on the thermodynamic description of physico-chemical mechanism of mass transfer phenomena and structural transformations in materials forming properties of REE and REE values informative parameters, making it possible to assess the kinetics of lead degradation processes and the dynamics of the resource characteristics of REE. Model based on changes in the technical resource developed a method for visualization of technical condition and resource characteristics of REE based on a geometric interpretation of the situation, allowing an analogy describing the kinetics of degradation processes and change the REE, which made it possible to create a mathematical and software systems that allow visualization of the changes on the basis of a technical resource predict failures REE. Experiments were conducted to study the properties of the proposed method. The experimental results allow us to recommend the proposed method for use in practice.

**Keywords:** monitoring, life cycle, forecasting, radio electronic equipment.

## REFERENCES

- Sheptunov S. A. Zhiznennyj cikl produkci. Moscow, Janus-K, 2003, 244 p.
- Davydov P. S. Tehnicheskaja diagnostika radioelektronnyh ustrojstv i sistem. Moscow, Radio i svjaz', 1991, 256 p.
- Strel'nikov V. P., A. V. Feduhin. Ocenka i prognozirovaniye nadezhnosti jelektronnyh elementov sistem. Kiev, Logos, 2002, 486 p.
- Lorenc G. A. Statisticheskie teorii v termodinamike. Moscow, RHD, 2001, 184 p.
- Andrusevich A. A., Nevliudov I. Sh., Rozdolovskij Ju. M. Geterogenyye modeli razvitiya proizvodstvennyh defektov, Nauch.-tehn. zhurnal. Aviacionno-kosmicheskaja tekhnika i tekhnologija, 2003, No. 38, pp. 114–119.
- Nevliudov I. Sh., Andrusevich A. A., Sotnik S. V. Termodinamicheskaja model' zhiznennogo cikla jekletronnoj apparatury, Vestnik Akademii inzhenernyh nauk, 2007, No. 3 (33), pp. 132–135.
- Prigozhin I. Vvedenie v termodinamiku neobratimykh processov. Moscow, RHD, 2001, 160 p.
- Andrusevich A. A., Nevliudov I. Sh. Monitoring processov proektirovaniya, proizvodstva i ekspluatacii zhiznennogo cikla jekletronnoj apparatury. Har'kov, FLP Cuvarev A. F, 2009, 272 p.
- Efimov I. E., Gorbunov Ju. I., Kozyr' I. Ja. Mikroelektronika: fizicheskie i tehnologicheskie osnovy, nadezhnost'. Moscow, Vysshaja shkola, 1986, 463 p.
- Maller R., Kejmins T. Jelementy integral'nyh shem: Per. s angl. Moscow, Mir, 1989, 630 p.
- Gorelova V. L., Mel'nikova E. N. Osnovy prognozirovaniya sistem. Moscow, Vyssh. shk, 1986, 287 p.
- Chaplygin D. Ju., Abramov P. B., Cvetkov V. V. Imitacionnaja model' dinamiki otkazov i vosstanovlenija rabotosposobnosti slozhnyh radioelektronnyh sistem. Matematicheskoe modelirovaniye sistem obrabotki informacii i upravlenija: Sbornik nauchnyh trudov. Voronezh, in-t MVD Rossii, 2001, pp. 14–19.
- Voroncov V. N. Kontrol' kachestva i prognozirovaniye nadezhnosti izdelij jekletronnoj tekhniki po jeklektrofizicheskim parametram. dokt. tehn. nauk. Sankt-Peterburg, Sev.-Zap. gos. zaoch. tehn. un-t, 2002, 39 p.
- Radaev N. N. Povyshenie tochnosti prognoza sobytij, inicirujushhih chrezychchajnye situacii, Izmeritel'naja tekhnika, 2002, No. 5, pp. 13–20.
- Krishhuk V. M., Shilou G. M., Namlians'kij A. O., Gaponenko M. P. Vibir elementiv pri kompensaciї zovnishnih vpliviv, Radioelektronika, informatika, upravlinnja, 2004, No. 2(12), pp. 36–41.

УДК 621.396.96

Пиза Д. М.<sup>1</sup>, Звягинцев Е. А.<sup>2</sup>

<sup>1</sup>Д-р техн. наук, профессор кафедры радиотехники и телекоммуникаций, Запорожский национальный технический университет, Украина

<sup>2</sup>Аспирант кафедры радиотехники и телекоммуникаций, Запорожский национальный технический университет, Украина

## АНАЛИЗ ВЛИЯНИЯ ПЕРЕМЕННОЙ БАЗЫ НА КАЧЕСТВО ПОДАВЛЕНИЯ ПОМЕХ ПРИ ВРЕМЯ-ПРОСТРАНСТВЕННОЙ ОБРАБОТКЕ СИГНАЛОВ

Рассмотрены особенности работы систем помехозащиты когерентно-импульсных радиолокационных станций при дискретной адаптации весовых коэффициентов автокомпенсатора активных помех. Целью исследований является оценка ограничений эффективности автокомпенсатора при совместном воздействии активных и пассивных помех с учетом разности центров антенн основного и компенсационного каналов приема. В статье предложена методика оценки влияния переменной базы на качество подавления активной шумовой помехи для разных режимов работы радиолокационной станции при использовании времязадержки-пространственной обработки сигналов. Проведен анализ межканального фазового сдвига на различных временных интервалах, равных длительности частотной пачки. На примере радара 36Д6 показано, что межканальный фазовый сдвиг в разных режимах работы РЛС при времязадержки-пространственной обработке сигналов не позволяет качественно подавить активную шумовую помеху, действующую по боковым лепесткам диаграммы направленности антенны.

**Ключевые слова:** радар, помехозащита, автокомпенсатор помех, активная шумовая помеха, боковые лепестки диаграммы направленности.

### ВВЕДЕНИЕ

Возможным вариантом построения систем помехозащиты в когерентно-импульсных радиолокационных станциях (РЛС) является время-пространственная обработка сигналов с использованием раздельной двухэтапной процедуры. Принципиальной особенностью построения системы помехозащиты есть то обстоятельство, что при защите главного луча фазовые центры антенн основного и компенсационного каналов совмещены, что снимает ряд ограничений, возникающих при компенсации активной шумовой помехи (АШП). При защите РЛС от АШП, действующей по боковым лепесткам диаграммы направленности антенны (ДНА), использование пространственных отличий между полезным сигналом и помехой реализуется путем разнесенного приема. При этом существует расстояние между фазовыми центрами основной и компенсационной антенн, называемое базой, и так как фазовые центры антенн каналов приема разнесены, значительно усложняется достижение высоких значений коэффициента подавления помех при сканировании антенной системы. Поэтому оценка качества подавления АШП при наличии переменной базы в антенной системе (АС) представляется актуальной.

### 1 ПОСТАНОВКА ЗАДАЧИ

На рис. 1 приведена схема приема АШП на антенную систему, состоящую из двух антенн: основной антенны РЛС  $A_1$  и компенсационной антенны  $A_2$ , фазовые центры которых разнесены на величину базы  $d$ . Расстояния от источника АШП до фазовых центров антенн обозначены как  $R_1$  и  $R_2$ .

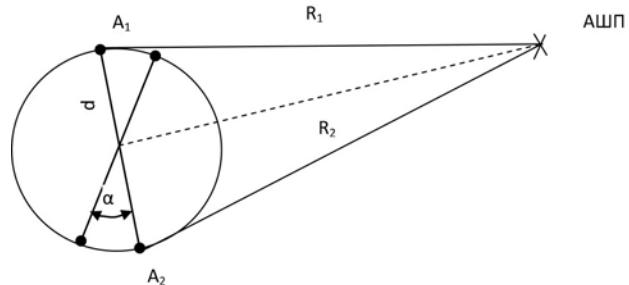


Рисунок 1 – Схема приема АШП на разнесенную antennную систему

Из рис. 1 непосредственно следует, что при изменении угла поворота антенной системы  $\alpha$  от нуля до 90 градусов, проекция базы на направление источника АШП изменяется от нуля до величины, равной  $d$ .

В случае время-пространственной фильтрации сигналов формирование весовых коэффициентов АКП производится после преобразования Фурье

$$\dot{X}(k) = \sum_{n=0}^{N-1} a_n \dot{X}_i(n) \exp\left\{j \frac{2\pi n}{N}\right\}, \quad (1)$$

где  $\dot{X}(k)$  – значение сигналов на выходе  $k$ -того фазового фильтра,  $N$  – количество периодов повторения в частотной пачке;  $\dot{X}_i(n)$  – значение сигнала в  $i$ -тый момент времени в  $n$ -ном периоде повторения;  $a_n$  – коэффициенты весового окна;  $k \in [0, N-1]$ . В дальнейшем будем считать, что количество зондирующих импульсов в излученной

частотной пачке равно количеству точек преобразования Фурье.

В случае, когда обработка сигналов в устройстве фильтрации является пачечной, а адаптация весовых коэффициентов производится дискретно, один раз за частотную пачку по информации 180-градусных фазовых фильтров [1], возникает необходимость оценить ограничения в подавлении АШП при данных условиях.

Необходимо рассмотреть, какие угловые смещения антенной системы имеют место при разных режимах работы РЛС и как изменяется межканальный фазовый сдвиг, вызванный переменной проекцией базы  $d$  на направление источника помех врачающейся антенной системы, что в свою очередь позволит сделать выводы о качестве подавления АШП.

Поэтому представляет интерес оценка потерь, обусловленных изменением межканальных фазовых сдвигов при сканировании антенной системы между интервалами пачечной адаптации весовых коэффициентов пространственного фильтра. Для этого необходимо разработать методику, которая учитывала бы: темп обзора пространства, частоту повторения зондирующих импульсов, количество точек преобразования Фурье.

## 2 МЕТОДИКА РЕШЕНИЯ

Будем считать, что в случае, представленном на рис. 1, при  $R_1=R_2$ , когда помеховые сигналы поступают на антенны  $A_1$  и  $A_2$  одновременно, как межканальный фазовый сдвиг  $\varphi_{\text{МК}}$ , так и задержка  $\tau$  отсутствуют. При этом проекция базы на нормаль к АС равна нулю.

Условимся, что при смещении антенны в процессе сканирования на угол  $\alpha \approx 5^\circ$  активная помеха действует в направлении первых боковых лепестков ДНА РЛИ. При этом нарушается равенство  $R_1=R_2$ , возникает разность хода лучей до фазовых центров антенн  $\Delta R = R_1 - R_2 \neq 0$ . Это приводит изменению межканальных фазовых сдвигов  $\varphi_{\text{МК}}$ , а также к межканальной задержке помеховых сигналов  $\tau$

$$\tau = \frac{d}{c} \sin \alpha, \quad (2)$$

где  $\alpha$  – угловое положение базы относительно равносигнального направления, при котором  $R_1=R_2$ ;  $c$  – скорость света.

Исходными параметрами для оценки потерь, обусловленных изменением межканальных фазовых сдвигов (между интервалами адаптации пространственного фильтра), является следующее: темп обзора 6 или 12 об/мин (соответственно длительность периода обзора – 10 или 5 с), частота повторения импульсов (ЧПИ) – 1500,750 или 375 Гц (соответственно, длительность периода повторения – 0,66; 1,33 или 2,66 мс); количество точек преобразования Фурье – 16, 12 или 8, которые соответствуют периодам повторения: 0,66; 1,33 или 2,66 мс.

Оценка потерь в компенсации АШП может быть произведена путем вычисления межканальных фазовых сдвигов между интервалами адаптации весовых коэффициентов пространственного фильтра при сканировании АС.

В дальнейшем вычисленные значения фазовых сдвигов дают возможность определения потенциальных значений коэффициентов подавления АШП в соответствии с nomogrammами, приведенными в [2, 3].

Для расчета межканальных фазовых сдвигов помеховых сигналов при сканировании АС необходимо учесть темп обзора пространства, а также длительность частотной пачки. При этом длительность частотной пачки  $\tau_{\Pi}$  определяется произведением периода повторения импульсов  $T_{\Pi}$  на их количество  $N$

$$\tau_{\Pi} = T_{\Pi} N. \quad (3)$$

Поскольку темп обзора задается в оборотах за минуту, а при малых смещениях АС в градусах за секунду, то для выполнения расчетов перейдем к безразмерному измерению углов. При этом, например, один оборот АС на  $360^\circ$  может быть представлен безразмерной величиной  $2\pi p$ , где  $p$  – радиан (отношение длины дуги окружности к длине ее радиуса).

## 3 ЕКСПЕРИМЕНТИ

Выполним расчет углового смещения АС на интервале, равном длительности частотной пачки, например, при темпе сканирования один оборот за 5 секунд (12 об/мин). Если полный оборот, равный  $2\pi p$ , АС совершает за 5 с, то за время, равное 1 мс, угловое смещение антенной системы  $\Delta_{MC}$  составляют:

$$\begin{aligned} \Delta_{MC} &= \frac{2\pi p}{T_{\text{об}} \cdot 1000} = \frac{6,28 p}{5 \cdot 1000} = 0,1257 \cdot 10^{-2} p = \\ &= 0,1257 \cdot 10^{-2} \cdot 57,296^\circ = 4,3', \end{aligned}$$

где  $T_{\text{об}}$  – время, в течении которого совершается один оборот АС.

За время, равное периоду повторения импульсов  $T_{\Pi}=0,66$  мс, угловое смещение  $\Delta_{\Pi\Pi}$  АС составит:

$$\Delta_{\Pi\Pi} = \frac{2\pi p}{T_{\text{об}} \cdot T_{\Pi} \cdot 1000} = \frac{6,283 p}{5 \cdot 0,66 \cdot 1000} = 0,8293 \cdot 10^{-3} p = 2,85'.$$

Аналогично выполненный расчет показывает, что за время, равное длительности частотной пачки, при 16-ти точечном преобразовании Фурье, угловое смещение антенной системы  $\Delta_{\Pi}$  составляет 45,6'.

В процессе расчета определены значения углового смещения АС на интервалах  $\Delta_{\Pi}$ , равных длительности частотной пачки для различных режимов работы РЛС 36Д6. Результаты расчетов представлены в табл. 1.

Дальнейший расчет заключался в том, чтобы определить величину изменения проекции базы  $\Delta d$  на направление источника АШП в соответствии с формулой

$$\Delta d = ds \sin \alpha. \quad (4)$$

Предположим, что АШП действует с направления первых боковых лепестков основной антенны и угол  $\alpha$

равен  $5^\circ$ . При этом, в соответствии с (4),  $\Delta d_{\text{п}} = 0,174 \text{ м}$ . Далее, полагая, что начало частотной пачки совпадает с направлением  $\alpha = 5^\circ$ , определим изменение проекции базы к концу пачки  $\Delta d_{\text{п}}$  с учетом смещений АС  $\Delta_{\text{п}}$ , приведенным в табл. 1. При этом изменение проекции базы  $\Delta d_{\text{п}}$  на интервале, равном длительности частотной пачки, составит

$$\Delta d_{\text{п}} = d \sin(5^\circ + \Delta_{\text{п}}) - \Delta d_5 \quad (5)$$

где первое слагаемое определяет проекцию базы в конце пачки, а второе в начале. Например, если ЧПИ=1500 Гц, а темп обзора равен 5 с, то  $\Delta_{\text{п}} = 45,6'$ , при этом  $\Delta d_{\text{п}} = 2 \sin(5^\circ + 45,6') - 0,174 = 0,027 \text{ м}$  или, с учетом длины волны, равной 0,1 м, набег межканального сдвига фазы к концу частотной пачки составит  $97,2^\circ$ .

#### 4 РЕЗУЛЬТАТЫ

Полученные результаты расчета изменения проекции базы  $\Delta d_{\text{п}}$ , а также пересчет изменения проекции базы в межканальный фазовый сдвиг ФМП во всех режимах работы РЛС сведены в табл. 1.

На рис. 2 приведена зависимость несокомпенсированного остатка  $P_0$ , нормированного по мощности АШП, от межканального фазового рассогласования (сдвига) [2, С. 55].

Таблица 1 – Изменение смещения АС, проекции базы  $\Delta d_{\text{п}}$  и межканального фазового сдвига  $\Phi_{\text{МП}}$  на временном интервале, равном длительности частотной пачки, при различных режимах работы РЛС

Частота повторения импульсов, Гц	1500		750		375	
Темп обзора, с	5	10	5	10	5	10
Смещение АС $\Delta_{\text{п}}$ , мин	45,6	22,7	68,4	34,4	90,4	45,8
$\Delta d_{\text{п}}, \text{м}$	0,027	0,013	0,039	0,017	0,055	0,026
$\Phi_{\text{МП}}, \text{град}$	97,2	46,8	140,4	61,2	198	94

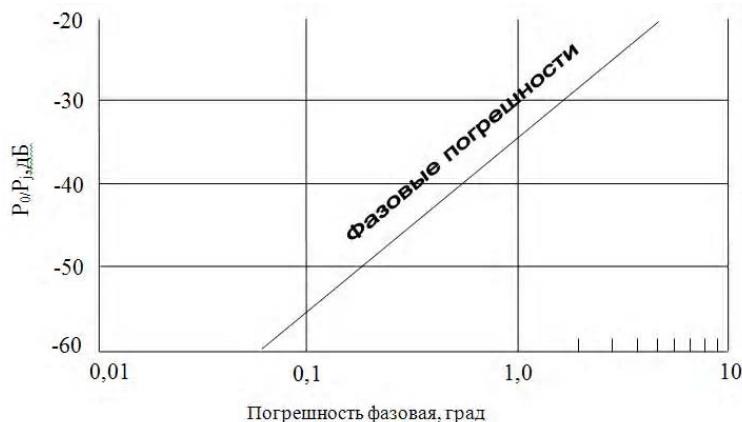


Рисунок 2 – Зависимость отношения  $P_0/P_j$  от фазового межканального рассогласования

Из рис. 2 непосредственно следует, что для реализации коэффициентов подавления порядка 20 дБ межканальный фазовый сдвиг  $\Phi_{\text{МП}}$  должен быть не более  $5^\circ$ . Из табл. 2 следует, что межканальные фазовые сдвиги на интервалах, равных длительности частотной пачки при время-пространственной обработке сигналов, не позволяют реализовать приемлемые значения коэффициента подавления при защите РЛС от АШП, действующей по боковым лепесткам основной антенны РЛС 36Д6.

#### ВЫВОДЫ

Предложена методика для оценки влияния переменной базы на качество подавления активной шумовой помехи для разных режимов работы радиолокационной станции при использовании время-пространственной обработки сигналов. Согласно этой методике проведен анализ межканального фазового сдвига на различных временных интервалах. На примере радара 36Д6 показано, что межканальный фазовый сдвиг в разных режимах работы РЛС при время-пространственной обработке сигналов не всегда позволяет обеспечить высокое качество подавления АШП, действующей по боковым лепесткам ДНА.

В дальнейшем представляется целесообразным провести аналогичную оценку для пространственно-временной обработки сигналов.

#### БЛАГОДАРНОСТИ

Работа выполнена в рамках выполнения научно-исследовательской темы Запорожского национального технического университета «Разработка и обоснование рекомендаций по построению систем пространственно-временной обработки сигналов в условиях воздействия комбинированных помех» (№ гос. регистрации 0114U002636).

#### СПИСОК ЛИТЕРАТУРЫ

- Пат. 48705 Україна МПК G01S 7/36 Спосіб компенсації активної складової комбінованої завади. [Електронний ресурс] / Кононович В. Я. (UA), Кукольницький А. П. (UA), Залевський О. П. (UA), Каспирович О.Г. (UA),

Мейстер Ю. Л. (UA), Денека А. А. (UA); заявник і патентовласник: Казенне підприємство «Науково-виробничий комплекс «Іскра» – заявл. 06.11.2009; опубл. 25.03.2010, бюл. №6, 2010 р. – Режим доступу: <http://base.ukrpatent.org/search.php?action=viewdetails&IdClaim=143619&chapter=description&sid=02b11f4011edf00129a4b672ea5ca5e8>

2. Монзінго Р. А. Адаптивные антенны решетки : Введение в теорию : Пер. с англ. / Р. А. Монзінго, Т. Ц. Міллер. – М. : Радіо і свя́зь, 1986. – 448 с.
3. Піза Д. М. Техническая реализация поляриметра на базе адаптивного фильтра/Д. М. Піза // Радиотехника. – 1977. – Вып. 102. – С. 15–17.

Статья поступила в редакцию 22.09.2014.  
После доработки 29.09.2014.

Піза Д. М.<sup>1</sup>, Звігінцев Є. О.<sup>2</sup>

<sup>1</sup>Д-р техн. наук, професор кафедри радіотехніки та телекомунікацій, Запорізький національний технічний університет, Україна  
<sup>2</sup>Аспірант кафедри радіотехніки та телекомунікацій, Запорізький національний технічний університет, Україна

## АНАЛІЗ ВПЛИВУ ЗМІННОЇ БАЗИ НА ЯКІСТЬ ПРИДУШЕННЯ ЗАВАД ПРИ ЧАСОВО-ПРОСТОРОВІЙ ОБРОБЦІ СИГНАЛІВ

Розглянуті особливості роботи систем завадозахисту когерентно-імпульсності радіолокаційних станцій при дискретній адаптації вагових коефіцієнтів автокомпенсатора активних завад. Метою дослідження являється оцінка обмеження ефективності автокомпенсатора при сумісному впливі активних та пасивних завад з врахуванням розносу фазових центрів антен основного і компенсаційного каналів прийому. У статті запропонована методика оцінки впливу змінної бази на якість придушення активної шумової завади для різних режимів роботи радіолокаційної станції при використанні часово-просторової обробці сигналів. Проведено аналіз міжканального фазового зсуву на різних часових інтервалах, рівних тривалості частотної пачки. На прикладі радару 36Д6 показано, що міжканальний фазовий зсув в різних режимах роботи РЛС при часово-просторовій обробці сигналів не дозволяє якісно придушити активну шумову заваду, діючу по бічним пелюсткам діаграмами спрямованості антени.

**Ключові слова:** радар, завадозахищеність, компенсатор завад, активна шумова завада, бічні пелюстки діаграмами спрямованості.

Piza D. M.<sup>1</sup>, Zviahintsev E. A.<sup>2</sup>

<sup>1</sup>Doctor of Sciences, Professor of the Department of Radio Engineering and Telecommunications, Zaporizhzhya National Technical University, Ukraine

<sup>2</sup>Post-graduate Student, of the Department of Radio Engineering and Telecommunications, Zaporizhzhya National Technical University, Ukraine

## ANALYSIS OF INFLUENCE OF VARIABLE BASE ON THE QUALITY OF JAMMER SUPPRESSION WITH TIME-SPACE SIGNAL PROCESSING

The features of work of noise immunity systems of coherently-pulse radar stations with discrete adaptation of weighting coefficients of an automatic compensator of active noise have been considered. The purpose of research is evaluation of effectiveness restrictions of the automatic compensator under the combined action of active and passive noise with the difference between the antenna centers of the basic and compensation receiving channels. This paper proposes a method for estimating the effect of variable base on the quality of the active suppression of jammer for different modes of radar using time-space signal processing. The analysis was made for the inter-channel phase shift at different time intervals equal to the duration of the frequency of the pack. On the example of radar 36D6 showed that interchannel phase shift in different radar modes of operation with the time-space signal processing does not allow qualitatively suppress jammer exerted on the side lobes of the antenna pattern.

**Keywords:** radar, ECCM, canceller, jamming, side lobes.

## REFERENCES

1. Pat. 48705 Ukrayina MPK9 G01S 7/36 Sposib kompensacii aktyvnoi' skladovoï' kombinovanoï' zavady. [Elektronnyj resurs] / Kononovych V. Ja. (UA), Kukol'nyc'kyj A. P. (UA), Zalevs'kyj O. P. (UA), Kaspyrovych O. G. (UA), Mejster Ju. L. (UA), Deneka A. A. (UA); zjavnyk i patentovlasnyk: Kazenne pidpryjemstvo «Naukovo-vyrobnychij kompleks «Іскра» – zjavl. 06.11.2009; opubl. 25.03.2010, bjur. №6, 2010 – Access: <http://base.ukrpatent.org/searchINV/search.php?action=viewdetails&IdClaim=143619&chapter=description&sid=02b11f4011edf00129a4b672ea5ca5e8>
2. Monzingo R. A., Miller T. C. Adaptivnye antennye reshetki: Vvedenie v teoriyu: Per. s angl. Moscow, Radio i svyaz', 1986, 448 p.
3. Piza D. M. Texnicheskaya realizaciya polyarimetra na baze adaptivnogo fil'tra. Radiotekhnika, 1977, Vyp.102, pp.15–17.

# МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

## МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

### MATHEMATICAL AND COMPUTER MODELLING

UDC 519.766.4

Kozhukhivska O. A.<sup>1</sup>, Bidyuk P. I.<sup>2</sup>, Kudriachov V. F.<sup>3</sup>, Kozhukhivskyj A. D.<sup>4</sup>

<sup>1</sup>PhD, Assistant Professor, Cherkasy University of Technology, Ukraine

<sup>2</sup>Dr. Sc., Professor, Professor, National Technical University of Ukraine «Kyiv Polytechnic Institute», Kyiv, Ukraine

<sup>3</sup>Magistrate at the Cherkassy National University of Technology, Cherkassy, Ukraine

<sup>4</sup>Dr. Sc., Professor, Professor at the Department of Information Technologies, Cherkasy State University of Technology, Ukraine

#### OPTIMIZATION OF SOME REINSURANCE STRATEGIES

The basic purpose of the work is a study of existing approaches to reinsurance directed towards modeling of distribution and minimization of risk for an insurance portfolio, and forming a strategy for its optimal reinsurance using developed decision support system. A method for a search of optimal reinsurance strategy is proposed. For this purpose statistical models were selected that correspond to the structure and volume of portfolio losses as well as the number of these losses. The simulation model for the total insurance losses is developed. While finding an optimal reinsurance strategy it was taken into consideration the dependence of the load coefficient on a specific form of reinsurance. A numerical study of the dependence between optimal reinsurance strategy and the varying load coefficient has been performed. It was established that taking into consideration of the variable load coefficient for specific risk capital values for an insurance company the stop-loss strategy provides worse results than other forms considered. An architecture and the functional layout for decision support system are proposed, and appropriate software was developed in C#. The decision support system functioning has been illustrated on simulated example. The system will provide a useful instrument for a business analytic to support decision making while selecting a strategy for insurance portfolio in specific conditions.

**Keywords:** modeling in reinsurance, optimization of reinsurance, load coefficient, decision support system, choice of reinsurance strategy.

#### NOMENCLATURE

$Y_i$  is random variables that characterize claim sizes to insurance company;

$N$  is a total number of claims to insurance company;

$N$  is a specific number of claims to insurance company;

$P(x)$  is an unconditional probability of  $x$ ;

$\theta$  is a Poisson distribution parameter;

$E$  is a mathematical expectation operator;

$\text{Var}(\cdot)$  is a variance;

$e$  is a base for natural logarithm;

$F(x)$  is a cumulative distribution function for  $x$ ;

$f(x)$  is a distribution function for  $x$ ;

$\Phi(\cdot)$  is a Laplace function;

$N(\cdot)$  is a normal distribution;

$\mu$  is a mean value of random variable;

$\sigma^2$  is variance;

$LN(\cdot)$  is a lognormal distribution;

$k$  is a lognormal distribution parameter;

$c$  is a constant for logistic distribution;

$\alpha$  is a form parameter for logarithmic logistic distribution;

$b$  is an extra scalar parameter for logarithmic logistic distribution;

$X_i$  are random variables that denote independent identically distributed losses;

$F^{*n}$  is an  $n$ -th order convolution for distribution  $F$ ;

$S$  is a collective loss;

$c$  is a proportion coefficient;

$q$  is a proportion coefficient;

$\theta$  is a proportion coefficient;

$G$  is a probability of fulfilling the obligations by insurance company;

$B$  is a total netto-premium;

$C$  is a guarantying capital;  
 $v$  is an insurance sum;  
 $QS$  is a quoted sum;  
 $XL$  is an excedent loss reinsurance;  
 $SL$  is a stop-loss reinsurance;  
 $a_0$  is a limiting predetermined value;  
 $a_1$  is a maximum limiting predetermined value;  
 $L$  is a predetermined priority sum;  
 $L_1$  is a fixed priority value;  
 $R$  is an income for insurance portfolio;  
 $P$  is a total premium;  
 VaR is a value-at-risk;  
 CaR is a capital-at-risk;  
 RoCaR is a return on capital-at-risk;  
 DSS is a decision support system;  
 $\lambda$  is a parameter for Poisson distribution;  
 IC is an insurance company;  
 DM is a decision maker;  
 DSS is a decision support system;  
 CDF is a cumulative distribution function;  
 PDF is a probability density function;

## INTRODUCTION

Modern development of insurance market in Ukraine requires from its participants of appropriate financial stability that is a corner stone for the future successful operation. Taking into consideration existing financial instability of the state as a whole, the solvency indicators for each insurance company (IC) represent a key criterion for their competition ability. This is especially true with respect to the situations taking place in sophisticated quickly changing environment. To smooth the insurance premiums with respect to the risks accepted by an insurer and to balance this way the insurance portfolio as well as to take into consideration potential financial possibilities of an insurer there exist reinsurance institutes. Practical application of the modern reinsurance strategies increases financial stability of insurance operations and their profitability, shifting a part of risks, accepted for insurance, to other insurers.

To develop a new reinsurance program for selected risks it is necessary to find an optimal solution regarding the form and volumes of reinsurance. In other words the problem arises how to take correctly into consideration appropriate number of influencing factors, such as a model of business, financial stability of an enterprise, the enterprise ability to accepting or rejecting the risks, current market conditions and possibilities regarding the ability for risk management solutions. In spite of the fact that insurer and his client both take part in the reinsurance process most of the existing theoretical studies on the subject are devoted to the search of optimal reinsurance form from the insurer (cedent) point of view and stick to some specific reinsurance form taking into consideration selected optimality criterion and calculation of insurance premium principle.

As far as the decisions directed towards reinsurance problems solving are closely linked to financial risks, an appropriate problem statement must be formulated as one that minimizes the risks. That is why modern solutions based

on mathematical risk and decision theory should involve some innovations into information processing and provide a decision maker (DM) with convenient criteria for their usage in the risk management environment.

The search for optimal in various senses insurance strategies is an important direction of research for several decades and it preserves its importance as of today. The first studies regarding the optimal reinsurance were directed towards the search of acceptable loss distribution function between insurer and reinsurer. It was shown that stop-loss reinsurance helps to minimize the variance of respective payments and to maximize expected insurance company income comparing to any other form of reinsurance (loss distribution function) [1–4]. In [5] it was also shown that the quoted reinsurance is optimal in the sense that this is the cheapest way to restrict the variance of non-distributed risks under condition that the load coefficient is increasing with growth of the reinsured part of the loss variance.

The purpose of this work is in studying of some existing approaches to reinsurance directed towards modeling of distribution and risk minimization for an insurance portfolio, and forming a strategy for its optimal reinsurance using developed specialized decision support system (DSS).

## 1 PROBLEM STATEMENT

Let we have the data in the form of identically distributed random variables  $Y_i, i=1,\dots,N$  characterizing claim sizes to insurance company; and  $X_i, i=1,\dots,N$  are independent identically distributed losses of insurance company. Then the problem of searching for optimal reinsurance strategy for an insurance portfolio is to find maximum of expected income value  $E(R)$ , with minimum of CaR such, that goal function  $\text{RoCaR} = E(R)/\text{CaR}$  is maximized. The problem is to be solved under restrictions on the predetermined value of  $a_0$  (the priority sum), and the predetermined maximum value equal to  $a_1$ .

The problem statement also includes statistical models selection for losses of an insurance portfolio, and development of decision support system as a handy computational instrument for decision maker searching for optimal reinsurance strategy.

## 2 REVIEW OF THE LITERATURE

The problem of reinsurance optimization is of interest for insurance companies and researchers what causes development of various solutions directed towards solving this problem. Most of the research works are of theoretical nature. It is assumed in [6] that the expected value premium principle could be applied and it was shown that the stop-loss reinsurance maximizes the expected utility of terminal wealth. In [7] the authors are concerned with the problem of purchasing the best risk protection from insurance company. The question of choosing the risk measure is discussed and several choices of non-symmetric risk measures are examined. Also sufficient conditions for optimality of reinsurance contract are provided within restricted class of admissible contracts. Some explicit forms for optimal

contracts are derived in the case of absolute deviation and truncated risk measure. The article [8] is devoted to theoretical extension of results relevant to introduction of mean-variance premium principles approach. The authors of the research [9] showed that the limited stop-loss and the truncated stop-loss approaches may lead to optimal contracts under fulfilling some criteria. These criteria include maximization of expected utility, stability of insurance company functioning, and the survival probability of a cedent. In research [10] the authors also assume the expected value principle and introduce two classes of optimal reinsurance models hiring minimization of well-known criteria based on value-at-risk (VaR) computing approach, and conditional VaR (CVaR) describing the total risk exposure of an insurer.

All the papers mentioned above made a substantial contribution to the study and development of new approaches for solving the problem of optimal reinsurance. However, this problem cannot be classified as a solved one, many studies in this direction are being continued today. Analysis of the publications shows that most of them have restrictions in the sense that they consider some specific criteria and preselected particular premium payment policy. Some of the criteria maximize expected utility of the final income, and some minimize preselected risk measure in form of loss variance, ruin probability, VaR, CVaR or some others. A cedent is usually concerned about the question: how would optimal reinsurance change from one reinsurance form to another. He needs to determine two following assumptions: – the risk measure that is used to find optimal solution; – the premium principle that is used for calculating the reinsurance premium. These choices lead to alternative optimal solutions based on different strategies.

In our study we consider first the possibilities for statistical modeling of insurance portfolio losses using various distribution types. Also alternative possibilities for selecting the reinsurance strategy are studied including proportional and non-proportional forms. As far as the purpose of decision maker working on reinsurance problems always tries to reduce the value of CaR and to maximize income  $E(R)$ , it will be reasonable to include into optimization criterion the variables mentioned. They are related to each other by another variable called RoCaR. RoCaR maximization provides maximization of expected income  $E(R)$ , and minimization of CaR. Thus, the RoCaR value is suitable for its use as a criterion for searching optimal reinsurance parameter. Finally, a specialized decision support system architecture and software are developed to solve the problem of searching for optimal reinsurance strategy.

### 3 MATERIALS AND METHODS

To analyze the distribution of risk for some insurance portfolio between insurer and reinsurer it is necessary to construct (or select) a model for its total loss. An integrated policy portfolio of risky insurance is as a rule rather heterogeneous even in the case when it can be divided into separate homogeneous parts. Thus here is the problem of

approximating the integrated loss for an arbitrary heterogeneous portfolio. An effective solution for the problem can be known collective risk model. The idea of the collective model is in considering policy portfolio as a source of loss only without taking into consideration the risks related to respective losses. The matter is that initial distributions for a collective model – i.e. distributions of counts and sizes of losses – can be estimated much better (with smaller errors) than the loss distributions for separate homogeneous risk groups.

The collective risk model constructing is based on the following limitations: – the insurance premium is paid at the very beginning of the insurance term, and any other payments are not performed within the period; the claims  $Y_1, Y_2, \dots$ , that are issued to insurance company, are not linked to specific contracts, they are considered as an integrated total risk; the random variables  $Y_1, Y_2, \dots$ , are independent and identically distributed; the total number of claims  $N$  to insurance company and the random variables  $Y_1, Y_2, \dots$  are mutually independent. Very often (e.g. in auto insurance)  $N$  has Poisson distribution

$$P(N=n)=\frac{1}{n!} \theta^n e^{-\theta}, \quad n=0,1,2,\dots,$$

with the parameter  $\theta=E(N)=Var(N)$ . The necessary conditions for the use of this distribution are as follows: – the random variables that characterize number of loss cases in two non-intersecting time intervals are independent; – two and more insurance cases cannot take place simultaneously; – generally the insurance cases can occur at arbitrary moments of time. Usually it is accepted in practice that these conditions are fulfilled for a separate risk and for a portfolio as a whole. For example, in a case of car crash an insurance case is touching for one driver only, and possible accumulation of events we can avoid by applying appropriate insurance strategy that could help to integrate several policies into one risk.

The use of a collective model supposes that during the time period when external factors (say inflation) are changed slightly. In such situations the losses in a specific insurance portfolio are independent and identically distributed. If the assumptions regarding independence of separate losses are usually fulfilled, the assumption relevant to identical distribution does not look realistic because of divergence of insurance premiums. On the other side the losses in a collective model are considered jointly (as an integrated value) on a definite time interval. This fact allows for taking into consideration that these random values belong to the same sample representing a mixture of different distributions for separate losses. In practice each insurance strategy and each portfolio type is characterized with its own (mixed) loss distribution that depends on the sizes of insurance premiums for separate risks and on a type of insurance events. For example, a mean loss caused by industrial

enterprise fire is much higher than a loss caused by the fire in a living house. These two cases are distinguished from the mean loss in auto insurance, and the losses in auto insurance are also different for different types of insurance. However, the practice of insurance shows that the loss structure in all types of insurance are very similar. Say, the number of small losses is usually much larger than the number of large ones. In a strict sense the «concentration of losses» is decreasing with growing of their sizes. The smallest losses are very often also rare and are not substantial from economic point of view. And the quantitative relations between large and small losses as well as the borders between them are different for different insurance types.

In the majority of practical cases the most important point is adequacy of mathematical model to the distribution of losses size in the area of large losses because just large losses make noticeable economic influence. It is reasonable to perform the model search in the family of distributions that have a scalar parameter and contain together with the CDF  $F(x)$  all the distributions of the following type:  $F(x|b)$ ,  $b > 0$ . This is convenient approach in cases when different types of currency are used. The transition to new currency requires changing the scalar parameter only; all other parameters and the PDF  $f(x)=F'(x)$  remain the same.

Very often statistical data is asymmetric and the distributions should be transformed to normal. In fact for asymmetric data with positive asymmetry there exist many examples where application of natural logarithm results in quite acceptable for practical use normal distribution. For example, if  $X$  is a claim size and it is necessary to construct a model for  $Y=\ln(X)$  using the normal distribution:  $Y \sim N(\mu, \sigma^2)$ , then we could use the distribution to build the model for  $X$ . This can be done by the change of variables:  $X=e^Y$ . Thus, we come to the lognormal distribution  $LN(\mu, \sigma^2)$  with scalar parameter  $\mu$ , and the form parameter  $\sigma^2$ :

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{(\ln x - \mu)^2/2\sigma^2};$$

$$F(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right); E(X^k) = \exp(k\mu + k^2\sigma^2/2).$$

The lognormal distribution can be hired as a model for describing loss size in a separate insurance case. Less popular though very similar to normal is logistic distribution that is described by the density function:

$$f(x) = \frac{1}{c\sigma} (1 + \exp((y-\mu)/c\sigma))^{-2} \exp((y-\mu)/c\sigma),$$

and the cumulative distribution function is as follows:

$$F(y) = (1 + \exp((y-\mu)/c\sigma))^{-1}.$$

As a result of logarithmic transform we get the logarithmic logistic distribution:

$$f(x) = \frac{\alpha(x/b)^{\alpha-1}}{b(1+(x/b)^\alpha)^2}; F(x) = 1 - \frac{1}{1+x/b^\alpha};$$

$$E(X^k) = b^k B\left(1+\frac{k}{\alpha}, 1-\frac{k}{\alpha}\right) = b^k \frac{k\pi/\alpha}{\sin(k\pi/\alpha)},$$

where  $b=e^\mu$ ;  $\alpha=1/(c\sigma)$ . The drawback of this distribution is in complexity of its parameters estimation.

Symmetric and defined over all values of real arguments density has Laplace distribution:

$$f(y) = 0.5 \alpha \exp(-\alpha|y-\mu|).$$

The cumulative distribution function consists of the two symmetric with respect to  $\mu$  exponential distributions as follows:

$$F(x) = \begin{cases} 0.5\alpha \exp(\alpha(y-\mu)), & y \leq \mu; \\ 1 - 0.5\alpha \exp(-\alpha(y-\mu)), & y > \mu. \end{cases}$$

After the transform:  $X=e^Y$  we get logarithmic Laplace distribution with form parameter  $\alpha$  and scalar parameter  $b=e^\mu$ :

$$F(x) = \begin{cases} 0.5(x/b)^\alpha, & 0 < x \leq b; \\ 1 - 0.5(x/b)^{-\alpha}, & x > b. \end{cases}$$

PDF for the logarithmic Laplace distribution is defined as follows:

$$f(x) = \begin{cases} \alpha(x/b)^{\alpha-1}/(2b), & 0 < x \leq b; \\ \alpha(x/b)^{-\alpha-1}/(2b), & x > b, \end{cases}$$

$$E(X^k) = b^k \alpha^2 / (\alpha^2 - k^2).$$

Usually the areas of small and medium losses are approximated with two straight lines not well enough, though in the area of high losses such approximation is quite possible even in the cases when the frequency of large losses is somewhat overestimated. The left side ( $x \leq b$ ) of the logarithmic Laplace distribution could be replaced by more appropriate distributions for small losses, say gamma or inverse Gaussian.

Another candidate for the losses approximation on the interval  $(0, b)$  is Pareto distribution with the following CDF:

$F(x) = 1 - (x/b)^{-\alpha}$ ,  $x \geq b$ , where  $b$  is a shift parameter. This function is transformed into zero point distribution defined for all  $x \geq 0$ :  $F(x) = 1 - ((b+x)/b)^{-\alpha}$ , with PDF of the form:

$$f(x) = \frac{\alpha}{b} \left(1 + \frac{x}{b}\right)^{-\alpha-1}.$$

In some forms of insurance the zero point Pareto distribution has a tendency to overestimation of frequency for the largest losses. In such cases it recommended to replace

the transform:  $X = e^Y$ , that allows the transition from exponential to Pareto distribution with the «weaker transform»  $X = Y^Z$ ,  $z > 1$ .

The unbiased exponential distribution  $F(y) = 1 - e^{-\beta y}$  leads to the Weibull distribution in the form:  $F(x) = 1 - \exp(-(x/b)^\alpha)$ ,  $x > 0$ , where  $\alpha = 1/z$ , with the following PDF:

$$f(x) = \frac{\alpha}{b} (x/b)^{\alpha-1} \exp(-(x/b)^\alpha).$$

The given above short review of some possible CDFs and respective PDFs could be used for the losses description in separate insurance cases.

The model describing collective losses is based on a suggestion that random losses in the insurance portfolio are independent for separate insurance cases, belong to the same distribution, and do not depend on their random count on a given time interval under consideration. The last requirement means independence of average size of loss on the number (count) of losses observed. This condition can be violated in some cases, for example, in auto insurance: on icy roads the number of car body damages is growing fast what results in a large number of small losses with the average loss decreasing simultaneously.

Such situations are inevitable in conditions of influence of external factors (climate conditions, demand variations) that influence simultaneously number of cases as well as size of respective losses. If the insurance covers the losses caused by climate factors then the numbers and size of losses cannot be considered as independent. However, the reason for this is in an attempt of performing simultaneous insurance for several possible loss (risk) factors with a single policy. When the risk factors (reasons) are considered separately the number of losses almost never depends on their sizes. Thanks to this fact collective model in most insurance cases is practically applicable. Moreover it was just collective model that initiated the theory of risk and made substantial impact on its further development and success.

Let  $N$  be the number (count) of losses for a given portfolio on a definite time interval (usually one year), and let  $X_1, X_2, \dots, X_N$  are independent identically distributed losses the distribution of which does not depend on  $N$ . Now the integrated losses can be represented in the form:

$$S = X_1 + \dots + X_N.$$

The moments of random variable  $S$  can be found using respective values for  $N$  and  $X$  (the values of  $X_1, X_2, \dots, X_N$  have the same distribution as  $X$ ):

$$\begin{aligned} E(S) &= E_N \left( E \left( \sum_{n=1}^N X_n | N \right) \right) = \\ &= E_N \left( \sum_{n=1}^N E(X_n | N) \right) = E(N)E(X); \end{aligned}$$

More difficulties are encountered when we need to determine the distribution  $G$  for the total loss  $S$  using the distributions for  $N$ , and  $X$ . However, practically there is no other way to find this distribution. Usually it is impossible to perform direct adjustment of the distribution form because of limited data volumes (short samples of annual data). The distribution  $G$  can be expressed via distribution  $p_n = P(N=n)$  for counts  $N$  and the distribution  $F(x) = P(X \leq x)$  for loss  $X$ :

$$G(s) = P(S \leq s) = \sum_{n=0}^{\infty} p_n P(S \leq s | N=n) = \sum_{n=0}^{\infty} p_n F^{*n}(s),$$

where  $F^{*n}$  means  $n$ -th order convolution for distribution  $F$  ( $F^{*0}(x)=0$  with  $x < 0$ , and  $F^{*0}(x)=1$  with  $x > 0$ ). However, such explicit computing of the infinite sum is possible only in very rare cases say when  $N$  has geometric distribution (e.g. negative binomial with the parameter  $\alpha = 1$ ), and  $X$  has exponential distribution. Another possible approach to solving the problem is hiring recursive Panger technique that also requires substantial computational resources. Quite acceptable approach to computing the distribution  $G$  is in simultaneous modeling of the counts number and average loss and generation of appropriate distribution with Monte Carlo technique.

Quite popular is proportional approach to reinsurance when reinsurer takes predetermined part of risk coordinated with insurer. In this case the whole loss  $X$  (loss size for one insurance case or total annual loss) is divided into two parts:  $qX$  and  $(1-q)X$  according to the rule:  $X = qX + (1-q)X$ ,  $0 < q < 1$ . There are different reasons for the risk distribution. Formally the probability of fulfilling the obligations by insurance company  $G(B+C)$  is determined by the total netto-premium  $B$ , distribution of total loss  $S$ , and guarantying capital  $C$ . The client may consider reliability of fulfilling the obligation  $G(B+C)$  as inadequate and can apply for extra protection from some other IC. Consider in short the forms of proportional and non-proportional reinsurance.

Proportional reinsurance forms:

– quoted sum (QS) reinsurance. In such case reinsurer accepts a fixed part of all insurance policies say 50%, i.e. receives 50% of which premium (except for reinsurance commission) and pays 50% for each loss;

– excedent sums reinsurance. With the quoted reinsurance all the risks are split between insurer and reinsurer in some proportion:  $c : (1-c)$ . With the excedent sums reinsurance parts of risks depend on insurance sum  $v$  according to the rule:  $c = c(v) = \min(v_0/v, 1)$ . In other words insurer takes all the risks with the insurance sum  $v \leq v_0$  and splits with reinsurer the risks with the insurance sum that exceeds  $v_0$ . The reinsurer accepts the part of risk that corresponds to the difference between the insurance sum and  $v_0$ . All other details correspond to the quoted reinsurance: for the risk with the insurance sum  $v$  the premium and losses are split between insurer and reinsurer according to the relation:  $c(v)/1-c(v)$ .

Non-proportional reinsurance forms:

– excedent loss (XL) reinsurance. In this case for each loss  $X$  insurer pays the sum  $\min(X, a_0)$  that is limited by the predetermined value of  $a_0$  (the priority sum), and reinsurer pays the following sum:  $\max(X - a_0, 0)$ . Sometimes reinsurance is predetermined in the limits of some maximum value equal to  $a_1$ , i.e.  $\min(\max(X - a_0, a_1), 0)$ . Insurer again is responsible for the part of loss  $\max(X - a_1 - a_0, 0)$  until there exists another reinsurance agreement with priority sum:  $a_1 + a_0$ . Size of the premium that belongs to reinsurer depends on expected count and loss size of that exceeds  $a_0$ , and on the value of  $a_1$ ;

– excedent cumulative loss reinsurance. This form of reinsurance is distinguished from the previous one with the priority sum that depends on the total loss caused by one insurance event (say hurricane or earthquake). This reinsurance form takes into account the possibility for simultaneous occurrence of a large number of small losses that create together substantial sum;

– stop-loss (SL) reinsurance. The stop-loss reinsurance results from development of the excedent loss reinsurance from a separate loss through cumulative to the annual. If the total annual loss  $S$  of insurer (for one insurance form) exceeds predetermined priority sum  $L$ , then reinsurer accepts a part of loss over this priority. Usually not more than some fixed value  $L_1$ . In other words reinsurer takes the loss  $\min(S, L)$  or  $\min(S, L) + \max(S - L - L_1, 0)$ , and reinsurer  $\max(S - L, 0)$  or  $\min(\max(S - L, 0), L_1)$ . The stop-loss form provides insurer with maximum protection when the probability of overriding the limit ( $L_0 + L_1$ ) by total loss is low. Limiting the potential loss of insurer by the value of  $L$ , reinsurer accepts insurance almost completely.

To solve the optimization problem regarding selection of reinsurance form it is necessary to select optimization criterion. Selection the reinsurance form insurer is mostly interested in the following: how much he can reduce insurance portfolio risk and what income he will get from it. Thus, these two variables should be used to construct the optimization criterion. The measure of value-at-risk (VaR) is used very often for determining maximum possible loss that may take place with predetermined probability on a given time interval. This is a popular approach thanks to availability of a set of rather simple estimation techniques. To compute VaR it is necessary to determine the distribution quantile for the total loss. The possibility of determining the data distribution was considered above.

The insurance portfolio income is determined by the formula [11]:  $R = P - S$ , where  $S$  is total loss for insurance portfolio;  $P$  is premium that can be found as follows:  $P = (1 + \theta) E(S)$ , where  $\theta$  is extra value for the insurance portfolio risk;  $E$  is a symbol for mathematical expectation. It is evident that income is a stochastic variable far as it depends on the random loss  $S$ . That is why we should consider expected income:  $E(R) = P - E(S) = \theta E(S)$ .

When the size of insurance premium is known then it is possible to determine the size of insurance company risky capital that it can lose with given probability. This capital is called capital-at-risk (CaR):  $CaR = VaR - P$ . Obviously this value is of substantial importance for insurance company because it shows what capital should the company possess so that to avoid bankruptcy with given probability. As far as the purpose of DM working on reinsurance problems always tries to reduce the value of CaR and to maximize income  $E(R)$ , it will be reasonable to include into optimization criterion the variables mentioned. They are related to each other by another variable called return on CaR (RoCaR):

$$RoCaR = \frac{E(R)}{CaR}.$$

RoCaR maximization provides maximization of expected income  $E(R)$ , and minimization of CaR. Thus, the RoCaR value is suitable for its use as a criterion for searching optimal reinsurance parameter. The reinsurance parameters can take different values for different reinsurance forms but comparing their RoCaR values we can find optimal reinsurance strategy (RoCaR will take maximum value). To perform necessary computing experiments a decision support system was developed that is based on the models and criteria considered.

#### 4 EXPERIMENTS

The DSS should provide DM at insurance company for selection of optimal reinsurance strategy for a given insurance portfolio. The system is based on simulation ideas that provide a possibility for searching optimal solutions for reinsurance problems. DSS includes the following three basic modules: loss estimation, reinsurance module, and final results module (all implemented on the C# programming platform). The simplified system architecture, that corresponds to hierarchical system construction approach, is given in Fig. 1. The DSS like this one are constructed according to the general system analysis principle that suppose availability of functionality completeness, taking into consideration possible uncertainties (like incomplete and noisy data, unknown forms of distributions, parametric uncertainties etc), and control of all computing stages with appropriate statistical criteria.

The loss estimation module computes total loss for an insurance portfolio. To illustrate the DSS functionality (to find the loss value) we used appropriate Monte Carlo simulation technique. The key problem for simulation was in selection of the most suitable probability distribution for the average losses and for the frequency of insurance cases. Decision regarding the choice of the distribution type should be based on the preliminary results of testing available statistical data what has been done in advance.

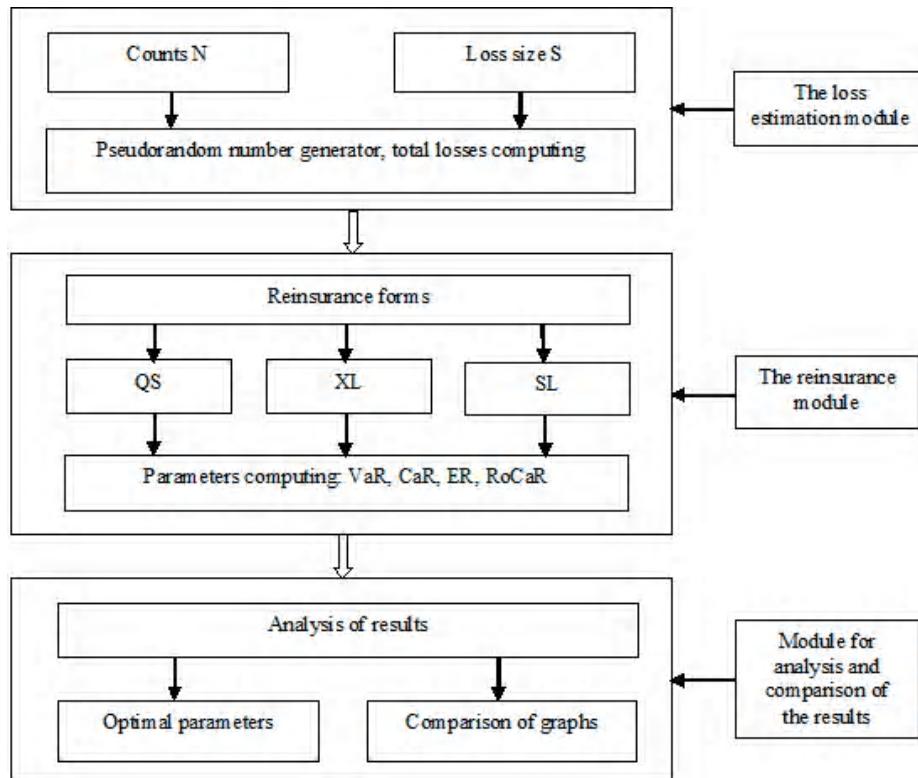


Figure 1 – Architecture of decision support system

This approach helped to determine parameters for the distributions of loss counts,  $N$ , and average loss sizes,  $S$ . The total loss is a random value that depends on the product,  $(N \times S)$ . At each iteration of the simulation algorithm was determined the value of total annual loss. After running definite number of iterations (usually from 2000 to 100000) we got the distribution for total loss. For example, distribution for auto insurance was generated with the following parameters: the losses had logarithmic distribution with parameters:  $\mu=8.3$ , and  $\sigma^2=2.0$ ; counts for the portfolio were distributed according to Poisson law with the parameter  $\lambda=200$  (Fig. 2).

The reinsurance module accepts the values necessary for analysis and search for optimal reinsurance strategy: (1) the probability of overriding CaR value by company's own capital directed for the loss compensation; (2) the size of extra value added for risk for a given portfolio; (3) load coefficient value; (4) the limits for parameters that are necessary for each reinsurance form: quoted, exceedent loss, and stop-loss. This module also contains information regarding dependence of the values VaR,  $E(R)$ , CaR and RoCaR on the parameters for each reinsurance form.

## 5 RESULTS

Consider results of computing for the example of losses having logarithmic distribution and portfolio counts have Poisson distribution. Also the following extra parameters were used:  $\theta=0.15$  is size of extra value added for risk;  $p=0.05$  is the probability of non-solvency. All other necessary parameters are as follows: for the QS form:

$\min(q)=0.25$ ,  $\max(q)=1.0$ ,  $\theta=0.15$ ; for the XL form:  $\min(a_0)=200000$ ,  $\max(a_0)=7000000$ ,  $\theta=0.30$ ; for the ER form:  $\min(L)=7.2$  mil,  $\max(L)=20.0$  mil,  $\theta=0.50$ .

## 6 DISCUSSION

The computations performed with the input data mentioned above show that income  $E(R)$  and risk capital are increasing exponentially with growing of coefficient  $a_0$ , that determines priority insurance sum. The exponent reaches its flat part when  $a_0$  approaches the value of about 6.0 mil for the excedent form of reinsurance. Fig. 3 illustrates dependence of the RoCaR criterion value on  $a_0$ . This curve has implicit maximum where RoCaR accepts the value of about 43% with optimum value of  $a_0=0.54$  mil. Thus, it is easy for a decision maker to compare different alternatives using this value.

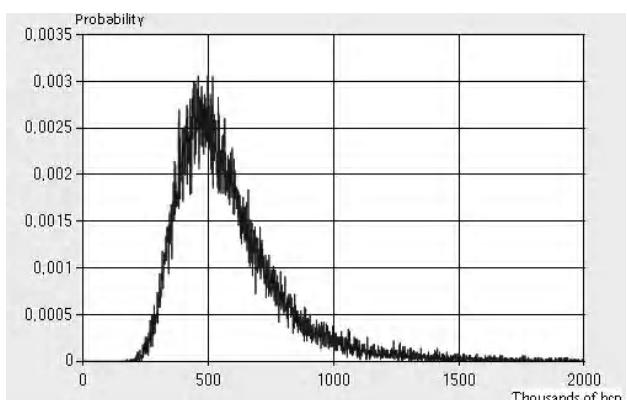


Figure 2 – Distribution for the total portfolio loss

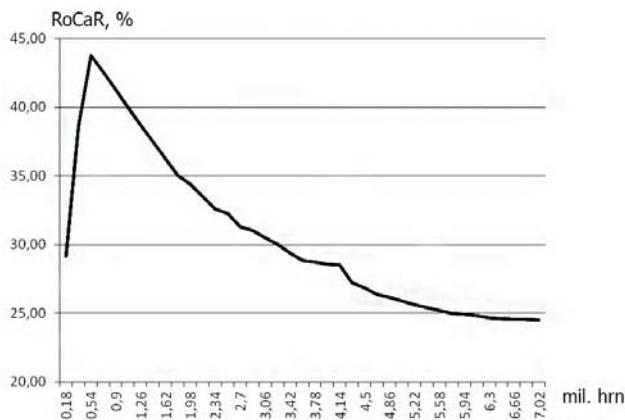


Figure 3 – RoCaR (%) plot versus  $a_0$

The module of results performs comparison of different reinsurance forms as dependent on the company's risk capital and expected income. The final result is the choice of optimal reinsurance form with optimal value of respective parameter using RoCaR as optimality criterion. It was established that the value of risk capital CaR, acceptable for insurance company, can be different for alternative reinsurance forms. It means that optimal reinsurance form should be found in each specific reinsurance case. A substantial help in the search for the optimal strategy provides decision support system as far as it reduces efforts of decision maker regarding sophisticated computations, comparison of results, including their visual representation, and selection of the best one for specific case.

## CONCLUSIONS

Thus, the problem has been solved of determining optimal reinsurance strategy based on application of statistical models that correspond to the structure, size and the number of loss cases for an insurance portfolio. To solve this task appropriate optimization problem statement was formulated.

This is a new solution thanks to the fact that the load coefficient was taken into consideration that depends on the form of reinsurance and influences the premium size. Varying the load coefficient it was established that the stop-loss strategy results in lower quality than the others strategies considered.

The results of computing experiments are useful for practitioners because they help to distinguish between alternative reinsurance strategies and to select an appropriate one in a particular case. Practically useful developments also refer to creating the specialized decision support system architecture, functional layout and software, in C#, for solving the problem of searching for the optimal reinsurance strategy. After adjustment to the special needs of an insurance company the system proposed could serve

as a handy instrument for a decision maker when searching for acceptable reinsurance form.

The future studies should be directed towards expanding the set of statistical and probabilistic models acceptable for the type of DSS considered. More specifically, the models could be in the form of multivariate conditional distributions (say copulas), Bayesian networks as well as alternative optimization procedures. It is easy to expand the DSS developed with new useful functions including automatic (or semiautomatic) comparison of computed alternatives.

## ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project at the National Technical University of Ukraine «Kyiv polytechnic institute»: «Development of information technology for modeling and estimation of financial risks (on the basis of Bayesian approach)» (the state registration number: 0113U000650). The authors are also grateful to the staff of local insurance company Brit Mark for their valuable advices regarding practical issues touching upon reinsurance strategies.

## REFERENCES

1. Borch K. The utility concept applied to the theory of insurance / K. Borch // ASTIN Bulletin. – 1991. – No. 1. – P. 245–255.
2. Klugman S. A. Loss models: from data to decisions / S. A. Klugman, H. H. Panjer, G. E. Willmot. – New York : John Wiley and Sons, 2004. – 688 p.
3. Klugman S. A. Bayesian statistics in actuarial science / S. A. Klugman. – Boston : Kluwer Academic Publishers, 1992. – 243 p.
4. McNeil A. J. Quantitative risk management / A. J. McNeil, R. Frey, P. Embrechts. – New Jersey: Princeton University Press, 2005. – 554 p.
5. Beard R. E. Risk Theory / R. E. Beard, T. Pentikainen, E. Pesonen. – London : Chapman and Hall, 1977. – 191 p.
6. Arrow K. J. Uncertainty and the welfare economics of medical care / K. J. Arrow // American Economic Review. – 1963. – Vol. 53, No. 5. – P. 941–973.
7. Gajek L. Optimal reinsurance under general risk measures / L. Gajek, D. Zagrodny // Insurance: Mathematics and Economics. – 2004. – Vol. 34, No. 2. – P. 227–240.
8. Kaluszka M. Optimal reinsurance under mean-variance premium principles / M. Kaluszka // Insurance: Mathematics and Economics. – 2001. – Vol. 28, No. 1. – P. 61–67.
9. Kaluszka M. An extension of Arrow's result on optimal reinsurance contract / M. Kaluszka, A. Okolewsky // The Journal of Risk and Insurance. – 2008. – Vol. 75, No. 2. – P. 275–288.
10. Cai J. Optimal retention for a stop-loss reinsurance under the Var and CTE risk measures / J. Cai, K. S. Tan // Austin Bulletin. – Vol. 37, No. 1. – P. 93–112.
11. Бондаренко Я. С. Теорія ризику в страхуванні / Я. С. Бондаренко, В. М. Турчин, Є. В. Турчин. – Донецьк : Донецький Національний Університет, 2008. – 112 с.

Article was submitted 22.09.2014.  
After revision 27.10.2014.

Кожуховская О. А.<sup>1</sup>, Бидюк П. И.<sup>2</sup>, Кудряшов В. Ф.<sup>3</sup>, Кожуховский А. Д.<sup>4</sup>

<sup>1</sup>Канд. техн. наук, ст. преподаватель, Черкасский государственный технологический университет, Черкассы, Украина

<sup>2</sup>Д-р техн. наук, профессор, Национальный технический университет Украины «Киевский политехнический институт», Украина

<sup>3</sup>Магистрант, Черкасский государственный технологический университет, Украина

<sup>4</sup>Д-р техн. наук, профессор, Черкасский государственный технологический университет, Украина

## ОПТИМИЗАЦІЯ НЕКОТОРИХ СТРАТЕГІЙ ПЕРЕСТРАХОВАННЯ

Цель работы заключается в исследовании существующих подходов к перестрахованию, направленному на моделирование распределения и минимизацию риска страхового портфеля, а также на формирование стратегии его оптимального перестрахования с использованием системы поддержки принятия решений. Предложен метод определения оптимальной стратегии перестрахования. Для этого выбраны статистические модели, которые соответствуют структуре, объему и количеству убытков страхового портфеля. При определении оптимального варианта перестрахования учтена зависимость коэффициента нагрузки от вида перестрахования. Коэффициент нагрузки учтен при расчете премии, а при сравнении различных форм перестрахования использованы одинаковые значения этого коэффициента. Выполнено численное исследование зависимости оптимальной формы перестрахования от переменного коэффициента нагрузки. Установлено, что учет переменного коэффициента нагрузки при определенных значениях капитала, которым готова рискнуть страховая компания, вариант stop-loss дает худшие результаты, чем другие формы перестрахования. Разработана архитектура, функциональная схема, а также программное обеспечение системы поддержки принятия решений для решения задачи оптимизации перестрахования (программная платформа C#). Продемонстрировано функционирование системы поддержки принятия решений, которая обеспечивает бизнес-аналитика критериями для руководства при принятии решений касательно выбора формы перестрахования страхового портфеля для конкретных условий.

**Ключевые слова:** моделирование в перестраховании, оптимизация перестрахования, коэффициент нагрузки, система поддержки принятия решений, выбор стратегии перестрахования.

Кожухівська О. А.<sup>1</sup>, Бідюк П. І.<sup>2</sup>, Кудряшов В. Ф.<sup>3</sup>, Кожухівський А. Д.<sup>4</sup>

<sup>1</sup>Канд. техн. наук, ст. викладач, Черкаський державний технологічний університет, Україна

<sup>2</sup>Д-р техн. наук, професор, Национальний технічний університет України «Київський політехнічний інститут», Київ, Україна

<sup>3</sup>Магістрант, Черкаський державний технологічний університет, Україна

<sup>4</sup>Д-р техн. наук, професор, Черкаський державний технологічний університет, Україна

## ОПТИМІЗАЦІЯ ДЕЯКИХ СТРАТЕГІЙ ПЕРЕСТРАХУВАННЯ

Мета роботи полягає у дослідженні існуючих підходів до перестрахування, спрямованому на моделювання розподілу і мінімізацію ризику страхового портфеля, а також формування стратегій його оптимального перестрахування з використанням системи підтримки прийняття рішень. Запропоновано метод знаходження оптимальної стратегії перестрахування. Для цього вибрано статистичні моделі, що відповідають структурі, розміру та кількості збитків страхового портфеля, а також побудовано імітаційну модель сукупного страхового збитку. При знаходженні варіанту оптимального перестрахування враховано залежність коефіцієнта навантаження від форми перестрахування. Коефіцієнт навантаження враховано при розрахунку премії, а при порівнянні різних форм перестрахування використано однакові значення цього коефіцієнта. Виконано чисельне дослідження залежності оптимальної форми перестрахування від змінного коефіцієнта навантаження. Встановлено, що врахування змінного коефіцієнта навантаження при певних значеннях капіталу, яким готова ризикнути страхова компанія, варіант stop-loss дає гірші результати ніж інші форми перестрахування. Розроблено архітектуру, функціональну схему, а також програмне забезпечення системи підтримки прийняття рішень для розв'язання задачі оптимізації перестрахування (програмна платформа C#). Продемонстровано функціонування системи підтримки прийняття рішень, яка забезпечує бізнес-аналітика критеріями вибору при прийнятті рішення стосовно вибору форми перестрахування страхового портфеля для конкретних умов.

**Ключові слова:** моделювання у перестрахуванні, оптимізація перестрахування, коефіцієнт навантаження, система підтримки прийняття рішень, вибір стратегії перестрахування.

## REFERENCES

1. Borch K. The utility concept applied to the theory of insurance, *ASTIN Bulletin*, 1991, No. 1, pp. 245–255.
2. Klugman S. A., Panjer H. H., Willmot G. E. Loss models: from data to decisions. New York, John Wiley and Sons, 2004, 688 p.
3. Klugman S. A. Bayesian statistics in actuarial science, Boston, Kluwer Academic Publishers, 1992, 243 p.
4. McNeil A. J., Frey R., Embrechts P. Quantitative risk management. New Jersey, Princeton University Press, 2005, 554 p.
5. Beard R. E., Pentikainen T., Pesonen E. Risk Theory. London, Chapman and Hall, 1977, 191 p.
6. Arrow K. J. Uncertainty and the welfare economics of medical care, *American Economic Review*, 1963, Vol. 53, No. 5, P. 941–973.
7. Gajek L., Zagrodny D. Optimal reinsurance under general risk measures, *Insurance: Mathematics and Economics*, 2004, Vol. 34, No. 2, pp. 227–240.
8. Kaluszka M. Optimal reinsurance under mean-variance premium principles, *Insurance: Mathematics and Economics*, 2001, Vol. 28, No. 1, pp. 61–67.
9. Kaluszka M., Okolewsky A. An extension of Arrow's result on optimal reinsurance contract, *The Journal of Risk and Insurance*, 2008, Vol. 75, No. 2, pp. 275–288.
10. Cai J., Tan K. S. Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures, *Austin Bulletin*, Vol. 37, No. 1, pp. 93–112.
11. Bondarenko Ya. S., Turchyn V. M., Turchyn Ye. V. Teoria ryzyku v strakhuvanni. Donetsk, Donetskyi Natsionalnyi Universytet, 2008, 112 p.

# НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

## НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

### NEUROINFORMATICS AND INTELLIGENT SYSTEMS

UDC 004.93

Subbotin S. A.

*Dr.Sc., Professor, Professor of Department of software tools,  
Zaporizhzhya National Technical University, Ukraine*

#### THE INSTANCE INDIVIDUAL INFORMATIVITY EVALUATION FOR THE SAMPLING IN NEURAL NETWORK MODEL SYNTHESIS

The problem of mathematical support development is solved to automate the sampling at diagnostic and recognizing model building by precedents. The object of study is the process of diagnostic and recognizing neural network model building by precedents. The subject of study is the sampling methods for neural network model building by precedents. The purpose of the work is to increase the speed and quality of the formation process of selected training samples for neural network model building by precedents. The method of training sample selection is proposed which for a given initial sample of precedents and given feature space partition determines the weights characterizing the term and feature usefulness. It characterizes the individual absolute and relative informativity of instances relative to the centers and the boundaries of feature intervals based on the weight values. This allows to automate the sample analysis and its division into subsamples, and, as a consequence, to reduce the training data dimensionality. This in turn reduces the time and provides an acceptable accuracy of neural model training. The software implementing proposed indicators is developed. The experiments to study their properties are conducted. The experimental results allow to recommend the proposed indicators for use in practice, as well as to determine effective conditions for the application of the proposed indicators.

**Keywords:** sample, instance selection, data reduction, neural network, data dimensionality reduction.

#### NOMENCLATURE

$\delta$  is a proportion of the original sample in the training subsample;

$E_{all}$  is a neural network model error for the whole original sample;

$ep_{tr}$  is a number of executed epochs of neural network training;

$E_{tr}$  is a neural network model error for the training sample;

$F()$  is a neural network model structure;

$f()$  is a user criterion characterizing the argument quality relatively to the problem being solved;

$K_{jk}$  is a number of classes, which instances hit the  $k$ -th interval of  $j$ -th feature values;

$N$  is a number of features characterizing original sample;

$N'$  is a number of features in a subsample;

$opt$  is an optimal (desired or acceptable) value of the functional  $f()$  for the problem being solved;

$S$  is a number of instances in the original sample;

$S'$  is a number of instances in a subsample;

$S_{jk}$  is a number of instances in the  $k$ -th term of the  $j$ -th feature;

$t_{tr}$  is a time of neural network model training;

$w$  is a set of controlled (adjusted) parameters of the neural network model;

$x_j^s$  is a value of  $j$ -th input feature  $x_j$ , characterizing the instance  $x^s$ ;

$y^s$  is an output feature value associated with the instance  $x^s$ ;

$y^{s*}$  is a calculated output feature value for the  $s$ -th instance on the neural model output;

$x^s$  is  $s$ -th instance of a sample.

## INTRODUCTION

To automate the decision making in problems of technical and medical diagnosis, as well as in pattern recognition problems it is necessary to have a model of a decision dependence from descriptive features, characterizing an instance to be recognized (an observation of the object or process condition at a certain time). As a rule, due to the lack or inadequacy of expert knowledge in practice such model constructed on the basis of observations or precedents (instances).

The one of the most popular and powerful tools for model building by precedents are artificial neural and neuro-fuzzy networks [1] that can learn by precedents, providing their generalization and extracting knowledge from the data.

The object of study is the process of diagnostic and recognizing neural network model building by precedents.

The process of neural model building is typically time-consuming and highly iterative. This is caused by that training time and accuracy of the neural network model are essentially dependent on the dimensionality and quality of the used training sample. Therefore, to improve the construction speed and quality of neural model it is necessary to reduce the dimension of the sample, providing the preservation of its basic properties.

The subject of study is the sampling methods for neural network model building by precedents.

The known sampling methods [2–23] are highly iterative and low speed, as well as characterized by the uncertainty of quality criteria of formed subsample.

The purpose of the work is to increase the speed and quality of the formation process of selected training samples for neural network model building by precedents.

## 1 PROBLEM STATEMENT

Suppose given the original sample as a set of precedents (instances)  $\langle x, y \rangle$ , where  $x = \{x^s\}$ ,  $x = \{x_j\}$ ,  $x^s = \{x_j^s\}$ ,  $x_j = \{x_j^s\}$ ,  $y = \{y^s\}$ ,  $s = 1, 2, \dots, S$ ,  $j = 1, 2, \dots, N$ .

For a given sample of precedents  $\langle x, y \rangle$  the problem of neural model synthesis can be presented as the problem of finding  $\langle F(), w \rangle$ :  $y^* = F(w, x^s)$ ,  $f(F(), w, \langle x, y \rangle) \rightarrow \text{opt}$ , where the model structure  $F()$  usually specified by the user in practice, and the set of controlled parameters  $w$  is adjusted based on the training sample.

In turn, the problem of subsample formation from a given sample  $\langle x, y \rangle$  is to find such a set of  $\langle x', y' \rangle$ :  $x' \subset \{x^s\}$ ,  $y' = \{y^s | x^s \in x'\}$ ,  $S' < S$ ,  $N' = N$ , wherein  $f(\langle x', y' \rangle, \langle x, y \rangle) \rightarrow \text{opt}$ .

## 2 REVIEW OF THE LITERATURE

The sampling methods for decision-making model building by precedents in [2, 3] are divided into prototype selection methods and prototype construction methods. Here, the prototype means selected subsample relative to the original sample.

The prototype selection methods [4–15] does not modify, but only select the most important instances from the original sample. Depending on the strategy of solution forming these methods are divided into incremental methods

[4, 5] (they successively add instances from the original sample to the subsample) and decremental methods [4–8] (they successively remove instances from the original sample, and obtain a subsample as a result). There are also separated such methods as noise filtering methods [6, 8, 9–11] (they remove instances, which class labels do not equal with most of the neighbor labels), condensation methods [4–7, 12, 13] (this methods add instances from the original sample to the formed subsample, if they bring a new information, but do not add if they have the same class labels as their neighbors), and methods based on stochastic search [12, 14, 15] (they randomly form a subsample from the original sample, considering a set of variants of decisions and selecting the best of them). The common disadvantages of these methods are the high iterativity and big search time, as well as uncertainty in quality criteria selection of formed subsample.

The methods of prototype construction [12, 15–23] based on the original sample build artificial instances, allowing to describe the original sample. Among these methods it is possible to separate the cluster analysis based methods [18, 19, 23] (they replace the original sample by the centers of its clusters), the data squashing methods [17] (they replace the original sample instances by the artificial prototypes having weights obtained on their basis) and the neural network methods (neural network based methods) [16, 20–22] (they train a neural network on the original sample, which is then used for cluster centers extraction as instances of formed subsample). The common disadvantages of these methods is their high iterativity and a big operating time, and the uncertainty in the initial parameter setting. The methods based on a cluster analysis are characterized by disadvantages such as the uncertainty of cluster number, initial parameters, and metric selection for the clustering and training methods. The data squashing methods form prototypes, which are difficult to interpret. The neural network methods have such disadvantages as the difficulty of prototype extraction from the neural network model, the no guarantee of receiving of acceptable neural network model a result of training, the neural network model variability, entailing nonstationarity of constructed prototypes, the orientation on a specific model, the uncertainty in setting the initial parameters of the model and training methods.

Additionally the combined methods are distinguished [3]. They combine the selection and formation of prototypes. The combined methods have the same disadvantages as methods of prototype selection and methods of prototype construction.

Since the prototype construction methods and the hybrid methods related with them are slower than the prototype selection methods, it is advisable to choose the latter as the basis for sampling problem solving.

In order to eliminate the disadvantages of these methods, it is advisable to form a sample without iterative busting of instance combinations by a certain percentage of instance selection from the original sample. This will significantly

reduce the time. Herewith we also need to define the indicators to evaluate the individual instance informativity with regard to their position relatively to the interclass boundaries and to the centers of pseudo-clusters, which. This makes possible to generate a non-random sample, to estimate and guarantee the high quality of selected subsamples.

### 3 MATERIALS AND METHODS

Let's break feature space into rectangular regions limiting the range of values of each feature by its minimum and maximum values. Then the partition projections into feature axis allow to allocate feature intervals of for each of the rectangular block. The intervals can be formed as cluster projections or as a regular grid, or on the basis of class boundaries in sample one-dimensional projections on the feature axes [24].

Then each such interval can be considered as a term and it is possible to evaluate its importance for decision-making on instance belonging to the cluster with the weight of the  $k$ -th term of  $j$ -th feature of  $s$ -th instance  $x^s$  based on a description of the corresponding interval center by the formula (1):

$$w_{C,jk}^s = \exp(-(0.5(r_{jk} - l_{jk}) - x_j^s)^2), \quad (1)$$

as well as the weight of the  $k$ -th term of  $j$ -th feature of the  $s$ -th instance  $x^s$  relatively to the description of the intercluster boundaries determined by the formula (2):

$$w_{B,jk}^s = \exp(-\min((r_{jk} - x_j^s), (x_j^s - l_{jk}))^2). \quad (2)$$

Then the overall significance of the  $k$ -th term of  $j$ -th feature of  $s$ -th instance  $x^s$  relatively to the description of the intercluster boundaries can be estimated using the weight determined by the formula:

$$w_{jk}^s = \max\{w_{C,jk}^s, w_{B,jk}^s\}.$$

Defining for each  $s$ -th instance the term significances, we can also determine the term weights for the whole sample:

$$w_{jk} = \frac{S_{jk}}{SK_{jk}}.$$

Knowing the term significance we can define the feature informativity evaluations by formula (3):

$$w_j = \max_k w_{jk} \quad (3)$$

or by the formula (4):

$$w_j = \frac{1}{k} \sum_{k=1}^{k_j} w_{jk}. \quad (4)$$

It is also possible to use the individual evaluation of the feature informativity in the range [0, 1] defined by the indicators [24].

Based on evaluations of term and feature significance we can determine informativity evaluations for each  $s$ -th sample instance by the formula (5):

$$I_1(x^s) = \frac{\sum_{j=1}^N w_j \sum_{k=1}^{k_j} w_{jk} w_{jk}^s}{\sum_{j=1}^N w_j \sum_{k=1}^{k_j} w_{jk} \max_p \{w_{jk}^p\}} \quad (5)$$

or by the formula (6):

$$I_2(x^s) = \frac{1}{N} \sum_{j=1}^N \left( \frac{w_j}{\max_{i=1,2,\dots,N} \{w_i\} k_j} \sum_{k=1}^{k_j} \left( \frac{w_{jk} w_{jk}^s}{\max_{q=1,2,\dots,k_j} \{w_{jk}\} \left( \max_p \{w_{jk}^p\} \right)} \right) \right). \quad (6)$$

Suggested indicators (5) and (6) provide evaluation of individual informativity of instance  $x^s$  relatively to the initial sample in the range [0, 1]. The greater the value of corresponding indicator, the more valuable is an instance, and vices versa.

If necessary, the estimates (5) and (6) can be further normalized so that they will give not an absolute but relative value of instance significance in the sample (7):

$$I(x^s) = \frac{I(x^s) - \min_p \{I(x^p)\}}{\max_p \{I(x^p)\} - \min_p \{I(x^p)\}}. \quad (7)$$

In this case, the instance with the maximum individual informativity will receive evaluation equal to one, and the instance with minimal informativity will receive evaluation equal to zero. The application of (7) can be useful when it need to simplify the choice of the threshold for separating the sample by the corresponding informativity indicator.

The proposed indicators of evaluation of individual instance significance can be used in the subsample formation from the given original sample by one of the following methods:

1) to form a training subsample of those instances of the original sample, the normalized values of which individual informativity evaluations (7) are greater than some specified threshold;

2) to form a training subsample from the not more than  $S' = \delta S$  instances of the original sample with the greatest individual informativity evaluation values;

3) to form a training subsample from the not more than  $S'/K$  instances of each class of the original sample with the greatest values of individual informativity evaluations;

4) to use a stochastic search based on evolutionary or multi-agent methods, selecting the best in a some sense combination of instances, using information about individual informativity of instances in the search operators to accelerate the search and focusing it on the most promising solutions.

The first method does not obviously determine the number of instances that will fall into the formed sample. The fourth method is iterative and requires the specification and use of quality indicators, the calculation of which can also be time consuming. Therefore, the second and third methods are the most simple applicable in practice and relatively simple from a computational point of view. They are appropriate to examine together with the proposed measures.

#### 4 EXPERIMENTS

The computer program implementing the proposed method, which complements the «Automated system neural network and neuro-fuzzy model synthesis for non-destructive diagnosis and pattern classification on features» (certificate of copyright registration № 35431 from 21.10.2010) was developed to conduct experiments.

The developed software was studied in solving the Fisher Iris classification problem [25]. The initial data sample contains 150 samples characterized by four input features. The output feature determines instance belonging to one of three classes.

On the basis of the original sample the instance informativity evaluations were obtained and subsets of instances as a training samples were selected by the second and third methods.

To study the second method the 25 %, 50 %, 75 % and 100 % (for the control) instances with the greatest values of individual significance was selected from the whole original sample and included to the training set, respectively. To study the third method the 25 %, 50 %, 75 % and 100 % (for the control) instances with the greatest values of individual significance in each class was selected from the original sample and included to the training set, respectively.

Further, for each sample a model based on a two-layer feed-forward neural network was built. It was trained using the Levenberg-Marquardt method [1]. The number of network inputs was determined by  $N$  is the number of features in the corresponding problem. The number of neurons in the second layer of the network corresponds to the number of classes  $K$ . The number neurons of the hidden (first) layer was defined as  $2K$ . All neurons of a network were used the weighted sum as weight (postsynaptic) function and logistic sigmoid as transfer function. The training method parameters were set as follows: the learning rate is 0,01, the allowable number of iterations (epochs) of the method is 1000, the target value of the error function is  $10^{-6}$ .

After neural model training process completion its final characteristics were fixed: the training time  $t_{tr}$  and the number of spent training iterations  $ep_{tr}$ . After training each model was tested separately on the training and the whole original

samples, for each of which the error was determined, respectively,  $E_{tr}$  and  $E_{all}$ . Here each error is the number of instances of corresponding sample for which the estimated value did not match the actual value of the output feature.

#### 5 RESULTS

The fragment of the results of conducted experiments is presented in the table 1. Here we use the following notation for the coding method of sampling: G is a regular grid partition, N is an irregular partition based on class boundaries in one-dimensional sample projections on the feature axis, K is instance selection in each class separately, A is instance selection in the whole sample. Calculated instance informativity indicators are encoded as follows: the first digit codes the  $I$  calculation method: 1 – by the formula (5), 2 – by the formula (6); second digit codes the method of  $w_j$  calculation: 1 – by the formula (3), 2 – by the formula (4)). For each of the experimentally obtained indicators it is also listed a percentage of formed training sample volume relative to the original sample volume. Markers «min», «average» and «max» are designated, respectively, minimum, average and maximum values.

The table 1 shows that the use of the proposed method of instance significance determining allows in practice to select a subsample of smaller volume from of the original sample, enough to construct neural network models with the required accuracy, reducing the time to build models.

Fig. 1 graphically illustrates the instance placement of the original and formed samples in the space of the first two features (the sepal length in cm on the abscissa axis and the sepal width in cm on the ordinate axis are plotted). Here markers «.», «x» and «+» denote the instances of different classes of the initial sample, a marker «o» indicates instances selected to the training set.

It can be seen from the fig. 1 that the proposed method allows to select the most significant instances of the original sample. In this case the obtained results essentially depends on the subsample formation method, the feature space partitioning method and the method of individual instance informativity evaluation.

#### 6 DISCUSSION

As it evident from the table 1, with the increasing of examples number in the formed sample the accuracy is increased (errors for formed training and for the original samples reducing), the training time and the number of training iterations are increased, and vice versa. At the same time a significant reduction of a sample volume to the 25 % of original leads to deterioration the training process characteristics (the time and number of iterations increase) and also to a decrease in accuracy. This can be explained by the fact that instances critical to describe the class separation can not be included to the sample of small volume.

Even a small reduction of the original sample volume in 25 % (up to 75 % of the original sample volume) yielded acceptable accuracy and reduces training time by more than 1.7 times. Reducing the volume of the original sample by a

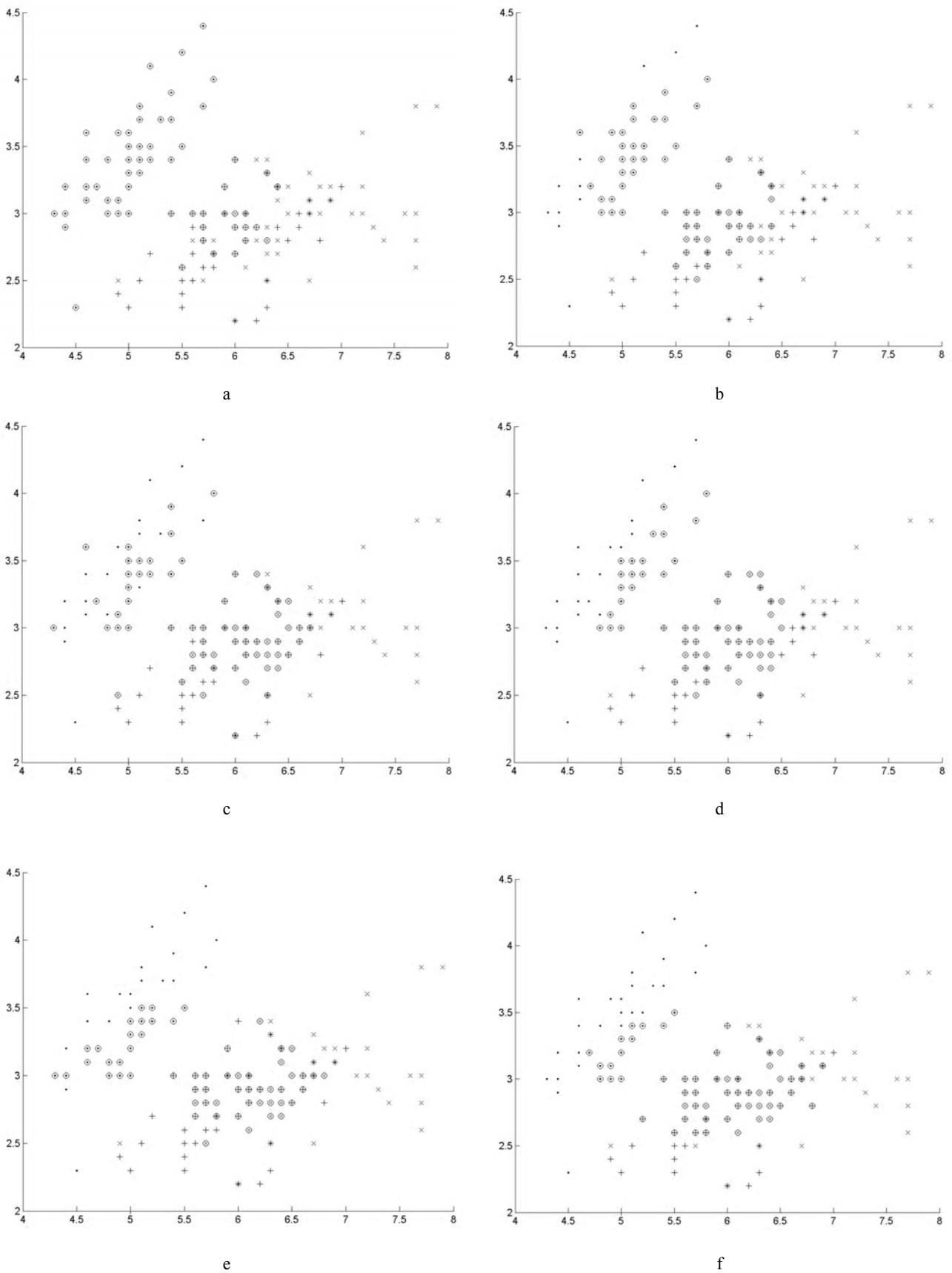


Figure 1 – The example of training samples formation from the original sample for the Fischer iris classification problem in the space of the first two features, where selected 50 % of the instances:

a – I11, G, A; b – I21, G, A; c – I11, G, K; d – I21, G, K; e – I11, N, K; f – I11, N, A

Table 1 – The fragment of experimental results on model building by the formed samples

Method code	<i>I</i> code	<i>t<sub>tr.</sub></i> , s				<i>ep<sub>tr.</sub></i>				<i>E<sub>tr.</sub></i>				<i>E<sub>all</sub></i>					
		25 %	50 %	75 %	100 %	25 %	50 %	75 %	100 %	25 %	50 %	75 %	100 %	25 %	50 %	75 %	100 %		
G	K	I11	0.094	0.218	0.172	0.655	17	51	28	123	0	0	0.006	0	0.064	0.024	0.004	0	
		I12	1.888	0.234	0.343	0.250	630	49	67	33	0	0	0	0	0.091	0.044	0.002	0	
		I21	0.094	0.218	2.371	0.125	21	31	583	19	0	0	0	0	0.058	0.031	0	0	
		I22	0.125	0.172	0.203	0.203	24	35	38	28	0	0	0	0	0.044	0.049	0.022	0	
N	K	I11	0.094	0.546	2.418	1.076	16	137	594	239	0	0	0	0	0.073	0.036	0.004	0	
		I12	0.094	0.250	0.203	1.591	15	55	39	325	0	0.009	0	0	0.044	0.040	0.013	0	
		I21	0.203	0.374	0.203	4.680	50	92	26	1001	0	0	0	0	0.060	0.018	0.013	0	
		I22	0.125	0.140	0.640	2.855	27	25	137	634	0	0	0	0	0.053	0.024	0.024	0	
G	A	I11	0.296	0.421	0.218	0.203	7	101	40	172	0	0.027	0	0	0.444	0.169	0.004	0	
		I12	0.109	0.562	0.312	0.624	9	137	75	114	0	0	0	0	0.229	0.076	0.004	0	
		I21	0.218	2.730	0.343	0.406	68	809	75	90	0.018	0	0	0	0.149	0.044	0.007	0	
		I22	0.109	0.109	0.515	0.577	18	21	114	105	0	0	0	0	0.036	0.064	0.029	0	
N	A	I11	2.964	0.577	3.931	2.964	1001	151	1001	637	0	0	0	0	0.033	0.013	0	0	
		I12	2.980	0.281	0.265	2.714	976	70	62	595	0	0	0	0	0.029	0.033	0.009	0	
		I21	0.125	0.234	0.187	1.326	42	52	45	274	0	0	0	0	0.249	0.009	0.004	0	
		I22	0.640	1.888	0.218	0.468	199	545	32	101	0	0	0	0	0.240	0.009	0.009	0	
min			0.094	0.109	0.172	0.125	7	21	26	19	0	0	0	0	0.029	0.009	0	0	
average			0.635	0.560	0.784	1.295	195	148	185	281	0.001	0.002	0	0	0.119	0.043	0.009	0	
Max			2.980	2.730	3.931	4.680	1001	809	1001	1001	0.018	0.027	0.006	0	0.444	0.169	0.029	0	

half afforded the gain in speed by 2.3 times. This confirms expediency of application of the proposed mathematical support in the neural network model building by precedents.

A method of instance selections in which the subsample is extracted considering the instance significance in the whole original sample (fig. 1a, fig. 1b, fig. 1f), leads to the selection of less informative instances in comparison with the instance selection considering the significance of instances in each class separately (fig. 1c, fig. 1d, fig. 1e). This is because the frequencies of each class instances may be different and in the selection of instances excluding class numbers it is possible to pass a locally important instances. Another cause may be that the instances describing external borders of classes, but do not important for the separation of adjacent classes can be recognized individually as significant, if we ignore their belonging to classes.

We should also note that the method of calculation of individual instance informativity indicators not only quantitatively but also qualitatively effect on the formed sample. It has been established that the indicators I11, defined by formulas (5) and (1), and I12, defined by formulas (5) and (2), respectively, mostly lead to the similar results, which differ significantly from the results for indicators I21, defined by the formulas (6) and (1), and I22, defined by the formulas (6) and (2), respectively. At the same time the indicators I21 and I22 are more resistant to the instance selection method, and the indicators I11 and I12 are the most effective in the instance selection using the instance importance in each class separately.

The significant influence of a feature space partitioning method on the results of the significance evaluation and

selection of instances by the results of conducted experiments can be explained by that the method of irregular partitioning with allocation of class intervals on the axis of each feature [24] allows usually get the best partition in comparison to the regular grid partition method. However, reducing the width of the interval, and correspondingly increasing the number of intervals of each feature axis can improve the latter method results. The selection of the optimum width of the interval is a separate problem that should be carried out taking into account the complexity characteristics of the particular application.

The closest analogue to the proposed method for determining the instance informativity is a set of indicators proposed in [26]. In contrast to the proposed in this paper, the indicators [26] characterize separately instance properties to be informative relatively to the external and internal borders, as well as to the class centers, which is their advantage in the problems of the data visualization and analysis. However, their disadvantages are low speed due to the need to calculate distances between instances, as well as the need and ambiguity of indicator integration to the comprehensive measures of instance informativity.

The advantage of the indicators proposed in this paper is that there is no need to calculate the distances between instances, but disadvantage is the necessary to divide the feature space. However, this disadvantage can be seen as an advantage in the case of large samples: if we use a partition that is simple from a computational point of view (for example, a regular grid) and know the minimum and maximum values of each feature than the computational cost of the proposed indicators will be less than the using of a set [26].

## CONCLUSIONS

The urgent problem of mathematical support development is solved to automate the sampling at diagnostic and recognizing model building by precedents.

The method of training sample selection is firstly proposed. It determines the weights characterizing the term and feature usefulness for a given initial sample of precedents and given feature space partition. It characterizes the individual absolute and relative informativity of instances relative to the centers and the boundaries of feature intervals based on the weight values. This allows to automate the sample analysis and its division into subsamples, and, as a consequence, to reduce the training data dimensionality. This in turn reduces the time and provides an acceptable accuracy of neural model training.

The practical significance of obtained results is that the software realizing the proposed indicators is developed, as well as experiments to study their properties are conducted. The experimental results allow to recommend the proposed indicators for use in practice, as well as to determine effective conditions for the application of the proposed indicators.

## ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Zaporizhzhya National Technical University «Intelligent information technologies of automation of designing, simulation, control and diagnosis of manufacturing processes and systems» (state registration number 0112U005350) and by the international project «Centers of Excellence for young RESearchers» of European Commission (№544137-TEMPUS-1-2013-1-SK-TEMPUS-JPHES).

## REFERENCES

1. Engelbrecht A. Computational intelligence: an introduction / A. Engelbrecht. – Sidney : John Wiley & Sons, 2007. – 597 p. DOI: 10.1002/9780470512517
2. Jankowski N. Comparison of instance selection algorithms I. Algorithms survey / N. Jankowski, M. Grochowski // Artificial Intelligence and Soft Computing : 7th International Conference ICAISC-2004, Zakopane, 7–11 June, 2004 : proceedings. – Berlin : Springer, 2004. – P. 598–603. – (Lecture Notes in Computer Science, Vol. 3070). DOI: 10.1007/978-3-540-24844-6\_90
3. Reinartz T. A unifying view on instance selection / T. Reinartz // Data Mining and Knowledge Discovery. – 2002. – № 6. – P. 191–210. DOI: 10.1023/A:1014047731786
4. Hart P. E. The condensed nearest neighbor rule / P. E. Hart // IEEE Transactions on Information Theory. – 1968. – Vol. 14. – P. 515–516. DOI: 10.1109/TIT.1968.1054155
5. Aha D. W. Instance-based learning algorithms / D. W. Aha, D. Kibler, M. K. Albert // Machine Learning. – 1991. – № 6. – P. 37–66. DOI: 10.1023/A:1022689900470
6. Gates G. The reduced nearest neighbor rule / G. Gates // IEEE Transactions on Information Theory. – 1972, Vol. 18, № 3. – P. 431–433. DOI: 10.1109/TIT.1972.1054809
7. Kibbler D. Learning representative exemplars of concepts: an initial case of study / D. Kibbler, D. W. Aha // Machine Learning : 4th International Workshop, Irvine, 22–25 June 1987 : proceedings. – Burlington : Morgan Kaufmann, 1987. – P. 24–30. DOI: 10.1016/b978-0-934613-41-5.50006-4
8. Wilson D. L. Asymptotic properties of nearest neighbor rules using edited data / D. L. Wilson // IEEE Transactions on Systems, Man, Cybernetics. – 1972. – Vol. 2, № 3. – P. 408–421. DOI: 10.1109/TSMC.1972.4309137
9. Tomek I. An experiment with the edited nearest-neighbor rule / I. Tomek // IEEE Transactions on Systems, Man, and Cybernetics. – 1976. – Vol. 6. – P. 448–452. DOI: 10.1109/TSMC.1976.4309523
10. Jankowski N. Data regularization / N. Jankowski // Neural Networks and Soft Computing : Fifth Conference, Zakopane, 6–10 June 2000 : proceedings. – Czestochowa : Polish Neural Networks Society, 2000. – P. 209–214.
11. Broadley C. E. Addressing the selective superiority problem: automatic algorithm/model class selection / C. E. Broadley // Machine Learning : Tenth International Conference, Amherst, 27–29 June, 1993 : proceedings. – Burlington : Morgan Kaufmann, 1993. – P. 17–24. DOI: 10.1016/b978-1-55860-307-3.50009-5
12. Wilson D. R. Reduction techniques for instancebased learning algorithms / D. R. Wilson, T. R. Martinez // Machine Learning. – 2000. – Vol. 38, № 3. – P. 257–286. DOI: 10.1023/A:1007626913721
13. An algorithm for a selective nearest neighbor decision rule / [G. L. Ritter, H. B. Woodruff, S. R. Lowry, T. L. Isenhour] // IEEE Transactions on Information Theory, 1975. – Vol. 21, № 6. – P. 665–669. DOI: 10.1109/TIT.1975.1055464
14. Domingo C. Adaptive sampling methods for scaling up knowledge discovery algorithms / C. Domingo, R. Gavalda, O. Watanabe // Discovery Science : Second International Conference, DS'99 Tokyo, 6–8 December 1999 : proceedings. – Berlin: Springer, 1999. – P. 172–183. DOI: 10.1007/3-540-46846-3\_16
15. Skalak D. B. Prototype and feature selection by sampling and random mutation hill climbing algorithms / D. B. Skalak // Machine Learning : Eleventh International Conference, New Brunswick, 10–13 July 1994 : proceedings. – Burlington : Morgan Kaufmann, 1994. – P. 293–301. DOI: 10.1016/b978-1-55860-335-6.50043-x
16. Kohonen T. Learning vector quantization / T. Kohonen // Neural Networks. – 1988. – Vol. 1, P. 303 DOI : 10.1016/0893-6080(88)90334-6
17. Likelihood-based data squashing: a modeling approach to instance construction / [D. Madigan, N. Raghavan, W. DuMouchel, M. Nason, C. Posse, G. Ridgeway] // Data Mining and Knowledge Discovery. – 2002. – Vol. 6, № 2. – P. 173–190. DOI: 10.1023/A:1014095614948
18. Support cluster machine / [B. Li, M. Chi, J. Fan, X. Xue] // Machine Learning : 24th International Conference, Corvallis, 20–24 June 2007 : proceedings. – New York, 2007. – P. 505–512. DOI: 10.1145/1273496.1273560
19. Evans R. Clustering for classification: using standard clustering methods to summarise datasets with minimal loss of classification accuracy / R. Evans. – Saarbrücken: VDM Verlag, 2008. – 108 p.
20. Sane S. S. A Novel supervised instance selection algorithm / S. S. Sane, A. A. Ghatol // International Journal of Business Intelligence and Data Mining. – 2007. – Vol. 2, № 4. – P. 471–495. DOI: 10.1504/IJBIDM.2007.016384

21. Reeves C. R. Using genetic algorithms for training data selection in RBF networks / C. R. Reeves, D. R. Bush // Instance Selection and Construction for Data Mining / Eds.: H. Liu, H. Motoda. – Norwell : Kluwer, 2001. – Part VI. – P. 339–356. DOI: 10.1007/978-1-4757-3359-4\_19
22. Suykens J. A. Least squares support vector machine classifiers / J. A. Suykens, J. Vandewalle // Neural Processing Letters. – 1999. – Vol. 9, № 3. – P. 293–300. DOI: 10.1023/A:1018628609742
23. Koskimaki H. Two-level clustering approach to training data instance selection: a case study for the steel industry / H. Koskimaki, I. Juutilainen, P. Laurinen, J. Roning // Neural Networks : International Joint Conference (IJCNN-2008), Hong Kong, 1–8 June 2008 : proceedings. – Los Alamitos: IEEE, 2008. – P. 3044–3049. DOI: 10.1109/ijcnn.2008.4634228
24. UCI machine learning repository [Electronic resource]. – Access mode: <http://archive.ics.uci.edu/ml/datasets/>
25. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition / S. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2013. – Vol. 22, № 2. – P. 97–103. DOI: 10.3103/s1060992x13020082
26. Subbotin S. A. Methods of sampling based on exhaustive and evolutionary search / S. A. Subbotin // Automatic Control and Computer Sciences. – 2013. – Vol. 47, № 3. – P. 113–121. DOI: 10.3103/s0146411613030073

Article was submitted 03.08.2014.

Субботин С. А.

Д-р техн. наук, профессор, Запорожский национальный технический университет, Украина

## ОЦІНІВАННЯ ІНДИВІДУАЛЬНОЇ ІНФОРМАТИВНОСТІ ЕКЗЕМПЛЯРОВ ДЛЯ ФОРМИРОВАННЯ ВИБОРОК ПРИ ПОСТРОЕНИИ НЕЙРОМОДЕЛЕЙ

Решена задача разработки математического обеспечения для автоматизации формирования выборок при построении диагностических и распознающих моделей по прецедентам. Объектом исследования являлся процесс построения диагностических и распознающих нейромоделей по прецедентам. Предмет исследования составляют методы формирования выборок для построения нейромоделей по прецедентам. Цель работы – повысить скорость процесса формирования и качество выделяемых обучающих выборок для построения нейромоделей по прецедентам. Предложен метод формирования обучающих выборок, который для заданной исходной выборки прецедентов и заданного разбиения пространства признаков определяет веса, характеризующие полезность термов и признаков, с учетом значений которых позволяет охарактеризовать индивидуальную абсолютную и относительную информативность экземпляров относительно центров и границ интервалов признаков, что позволяет автоматизировать анализ выборки и ее разделение на подвыборки, и, как следствие, сократить размерность обучающих данных, что, в свою очередь, позволяет сократить время и обеспечить приемлемую точность обучения нейромоделей. Разработано программное обеспечение, реализующее предложенные показатели. Проведены эксперименты по исследованию их свойств. Результаты экспериментов позволяют рекомендовать предложенные показатели для использования на практике, а также определять эффективные условия применения предложенных показателей.

**Ключевые слова:** выборка, отбор экземпляров, редукция данных, нейронная сеть, сокращение размерности данных.

Субботін С. О.

Д-р техн. наук, професор, Запорізький національний технічний університет, Україна

## ОЦІНЮВАННЯ ІНДИВІДУАЛЬНОЇ ІНФОРМАТИВНОСТІ ЕКЗЕМПЛЯРІВ ДЛЯ ФОРМУВАННЯ ВИБОРОК ПРИ ПОБУДОВІ НЕЙРОМОДЕЛЕЙ

Вирішено завдання розробки математичного забезпечення для автоматизації формування вибірок при побудові діагностичних і розпізнавальних моделей за прецедентами. Об'єктом дослідження був процес побудови діагностичних і розпізнавальних нейромоделей за прецедентами. Предмет дослідження становлять методи формування вибірок для побудови нейромоделей за прецедентами. Мета роботи – підвищити швидкість процесу формування та якість виділованих навчальних вибірок для побудови нейромоделей за прецедентами. Запропоновано метод формування навчальних вибірок, який для заданої вихідної вибірки прецедентів і заданого розбиття простору ознак визначає ваги, що характеризують корисність термів і ознак, з урахуванням значень яких дозволяє охарактеризувати індивідуальну абсолютну і відносну інформативність примірників щодо центрів і меж інтервалів ознак, що дозволяє автоматизувати аналіз вибірки і її поділ на підвібірки, і, як наслідок, скоротити розмірність навчальних даних, що, у свою чергу, дозволяє скоротити час і забезпечити прийнятну точність навчання нейромоделей. Розроблено програмне забезпечення, що реалізує запропоновані показники. Проведені експерименти з дослідження їхніх властивостей. Результати експериментів дозволяють рекомендувати запропоновані показники для використання на практиці, а також визначати ефективні умови застосування запропонованих показників.

**Ключові слова:** вибірка, відбір екземплярів, редукція даних, нейронна мережа, скорочення розмірності даних.

## REFERENCES

1. Engelbrecht A. Computational intelligence: an introduction. Sidney, John Wiley & Sons, 2007, 597 p. DOI: 10.1002/9780470512517
2. Jankowski N., Grochowski M. Comparison of instance selection algorithms I. Algorithms survey, *Artificial Intelligence and Soft Computing : 7th International Conference ICAISC-2004, Zakopane, 7–11 June, 2004 : proceedings*. Berlin, Springer, 2004, pp. 598–603. – (Lecture Notes in Computer Science, Vol. 3070). DOI: 10.1007/978-3-540-24844-6\_90
3. Reinartz T. A unifying view on instance selection, *Data Mining and Knowledge Discovery*, 2002, No. 6, pp. 191–210. DOI: 10.1023/A:1014047731786
4. Hart P. E. The condensed nearest neighbor rule, *IEEE Transactions on Information Theory*, 1968. Vol. 14, pp. 515–516. DOI: 10.1109/TIT.1968.1054155
5. Aha D. W., Kibler D., Albert M. K. Instance-based learning algorithms, *Machine Learning*, 1991, No. 6, pp. 37–66. DOI: 10.1023/A:1022689900470

6. Gates G. The reduced nearest neighbor rule, *IEEE Transactions on Information Theory*, 1972, Vol. 18, No. 3, pp. 431–433. DOI: 10.1109/TIT.1972.1054809
7. Kibbler D., Aha D. W. Learning representative exemplars of concepts: an initial case of study, *Machine Learning : 4th International Workshop*, Irvine, 22–25 June 1987, proceedings. – Burlington, Morgan Kaufmann, 1987, pp. 24–30. DOI: 10.1016/b978-0-934613-41-5.50006-4
8. Wilson D. L. Asymptotic properties of nearest neighbor rules using edited data, *IEEE Transactions on Systems, Man, Cybernetics*, 1972, Vol. 2, No. 3, pp. 408–421. DOI: 10.1109/TSMC.1972.4309137
9. Tomek I. An experiment with the edited nearest-neighbor rule, *IEEE Transactions on Systems, Man, and Cybernetics*, 1976, Vol. 6, pp. 448–452. DOI: 10.1109/TSMC.1976.4309523
10. Jankowski N. Data regularization, *Neural Networks and Soft Computing* : Fifth Conference, Zakopane, 6–10 June 2000 : proceedings, Czestochowa, Polish Neural Networks Society, 2000, pp. 209–214.
11. Broadley C. E. Addressing the selective superiority problem: automatic algorithm/model class selection, *Machine Learning*, Tenth International Conference, Amherst, 27–29 June, 1993, proceedings, Burlington, Morgan Kaufmann, 1993, pp. 17–24. DOI: 10.1016/b978-1-55860-307-3.50009-5
12. Wilson D. R., Martinez T. R. Reduction techniques for instancebased learning algorithms, *Machine Learning*, 2000, Vol. 38, No. 3, pp. 257–286. DOI: 10.1023/A:1007626913721
13. Ritter G. L., Woodruff H. B., Lowry S. R., Isenhour T. L. An algorithm for a selective nearest neighbor decision rule, *IEEE Transactions on Information Theory*, 1975, Vol. 21, No. 6, pp. 665–669. DOI: 10.1109/TIT.1975.1055464
14. Domingo C., Gavalda R., Watanabe O. Adaptive sampling methods for scaling up knowledge discovery algorithms, *Discovery Science : Second International Conference*, DS'99 Tokyo, 6–8 December 1999 : proceedings. Berlin. Springer, 1999, pp. 172–183. DOI: 10.1007/3-540-46846-3\_16
15. Skalak D. B. Prototype and feature selection by sampling and random mutation hill climbing algorithms, *Machine Learning*, Eleventh International Conference, New Brunswick, 10–13 July 1994 : proceedings, Burlington, Morgan Kaufmann, 1994, pp. 293–301. DOI: 10.1016/b978-1-55860-335-6.50043-x
16. Kohonen T. Learning vector quantization, *Neural Networks*, 1988, Vol. 1, pp. 303 DOI : 10.1016/0893-6080(88)90334-6
17. Madigan D., Raghavan N., DuMouchel W., Nason M., Posse C., Ridgeway G. Likelihood-based data squashing: a modeling approach to instance construction, *Data Mining and Knowledge Discovery*, 2002, Vol. 6, No. 2, pp. 173–190. DOI: 10.1023/A:1014095614948
18. Li B., Chi M., Fan J., Xue X. Support cluster machine, *Machine Learning*, 24th International Conference, Corvallis, 20–24 June 2007, proceedings. New York, 2007, pp. 505–512. DOI: 10.1145/1273496.1273560
19. Evans R. Clustering for classification: using standard clustering methods to summarise datasets with minimal loss of classification accuracy. Saarbrücken, VDM Verlag, 2008, 108 p.
20. Sane S. S., Ghatal A. A. A Novel supervised instance selection algorithm, *International Journal of Business Intelligence and Data Mining*, 2007, Vol. 2, No. 4, pp. 471–495. DOI: 10.1504/IJBIDM.2007.016384
21. Reeves C. R., Bush D. R. Using genetic algorithms for training data selection in RBF networks, *Instance Selection and Construction for Data Mining*, Eds.: H. Liu, H. Motoda. Norwell, Kluwer, 2001, Part VI, pp. 339–356. DOI: 10.1007/978-1-4757-3359-4\_19
22. Suykens J. A., Vandewalle J. Least squares support vector machine classifiers, *Neural Processing Letters*, 1999, Vol. 9, No. 3, pp. 293–300. DOI: 10.1023/A:1018628609742
23. Koskimaki H., Juutilainen I., Laurinen P., Roning J. Two-level clustering approach to training data instance selection: a case study for the steel industry, *Neural Networks : International Joint Conference (IJCNN-2008)*, Hong Kong, 1–8 June 2008 : proceedings. Los Alamitos, IEEE, 2008, pp. 3044–3049. DOI: 10.1109/ijcnn.2008.4634228
24. UCI machine learning repository [Electronic resource]. – Access mode: <http://archive.ics.uci.edu/ml/datasets/>
25. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition, *Optical Memory and Neural Networks (Information Optics)*, 2013, Vol. 22, No. 2, pp. 97–103. DOI: 10.3103/s1060992x13020082
26. Subbotin S. A. Methods of sampling based on exhaustive and evolutionary search, *Automatic Control and Computer Sciences*, 2013, Vol. 47, No. 3, pp. 113–121. DOI: 10.3103/s0146411613030073

UDC 004.89

Korablev N. M.<sup>1</sup>, Ivaschenko G. S.<sup>2</sup>

<sup>1</sup>Dr.Sc., Professor, Professor of the Department of Electronic Computers, Kharkiv National University or Radio Electronics, Ukraine  
<sup>2</sup>Postgraduate student of the Department of Electronic Computers, Kharkiv National University or Radio Electronics, Ukraine

## PARALLEL IMMUNE ALGORITHM OF SHORT-TERM FORECASTING BASED ON MODEL OF CLONAL SELECTION

The paper studies ways of parallelization of hybrid immune algorithm of short-term forecasting of time series built on the basis of clonal selection model that uses case-based reasoning and the simplest methods of forecasting. There has been analysis performed of two variants of parallelization having different procedure of messaging between the computational nodes. To implement proposed algorithm used MPI.NET technology for messaging systems. To optimize individual computational nodes operations, TPL library for shared memory systems is used. The work presents results of experimental investigations demonstrating efficiency of the proposed approach.

**Keywords:** forecasting, time series, artificial immune systems, clonal selection model, antibody, antigen, affinity, cloning, mutation.

### NOMENCLATURE

AIS is an artificial immune systems;  
CBR is a case based reasoning;  
MPI is a message passing interface;  
TPL is a task parallel library;  
 $N$  is a number of known values of forecasting time series;  
 $N_m$  is a number of forecasting methods, which are used in current model;  
 $N_g$  is a number of iterations during learning of AIS;  
 $k$  is a length of forecasting horizon;  
 $S$  is a number of external factors;  
 $L$  is a length of antibody part which takes part on the affinity determine;  
 $Ab$  is a population of antibodies;  
 $Ab^C$  is a population of clones;  
 $Ab^M$  is a population of memory cell;  
 $Ag$  is a population of antigens;  
 $n$  is a number of antibodies in  $Ab$  ;  
 $n_m$  is a size of population of memory cell;  
 $n'$  is a number of cells selected for cloning and mutation;  
 $n_c$  is a number of clones made by one antibody;  
 $D_{Ab-Ag}$  is a matrix of  $Ab$ - $Ag$  affinity;  
 $D_{Ab-Ab}$  is a matrix of  $Ab$ - $Ab$  affinity;  
 $\sigma_d$  is a threshold coefficient of cell stimulation;  
 $\sigma_S$  is a threshold coefficient of compression;  
 $\sigma_{age}$  is a threshold age of antibodies;  
 $N_{cn}$  is a number of computational nodes;  
 $T$  is a duration of communication operations;  
 $V_{ag}$  is amount of data to transmit one antigen;  
 $V_{ab}$  is amount of data to transmit one antibody;  
 $\alpha$  is a latency of a data network (Hockney model parameter);  
 $\beta$  is a bandwidth of a data network (Hockney model parameter).

### INTRODUCTION

Time series forecasting allows to solve a problem of determination of the future condition of different systems on the basis of analysis of already available historical data. The task of the short-term forecasting is essential to technical and economic systems since the estimation of the future values of a series is used for efficient decision making. Despite the fact that the development of hardware and software provides with more powerful computing platforms due to which it is possible to implement complex algorithms of forecasting, the tasks of economic and technical management makes more severe requirements to the accuracy of forecasting [1]. Taking into account of external factors represented in the form of the related time series allows to improve the accuracy of forecasting significantly. However, a prior list of factors potentially influencing the predicted value is often redundant. That is why a problem of external factors choosing taken into the consideration during the forecasting continues to be relevant [6].

One actively develops approaches to the forecasting of time series on the basis of artificial intelligence techniques such as artificial neural networks and AIS [3] that can be integrated with various approaches and are notable for easy parallelization and abilities to adapt. There are different models based on the principles of immune system: model of clonal selection, model of immune network etc., that can be used to solve forecasting problems [2, 4–6]. These models can use different forecasting techniques that allows to timely take into consideration changes in the basic structure of time series and to compensate for shortcomings of one approaches by means of advantages of the others [2]. Simplicity of the converting to parallel is one of the immune algorithm advantages.

At the present time companies have accumulated a considerable volume of historical data that leads to the significant growth of volumes of input information for forecasting problem. Thus, in case of growing volume of

input data there is a necessity to solve tasks of short-term forecasting under the conditions of limited computational resources and time limit. This is why a relevant way to solve this problem is a parallel programming for multi-core and multiple-processor cluster systems. However, while writing parallel programs algorithms of solving even the simplest problems ceases to be trivial [7].

The object of the research is developing a parallel immune algorithm of short-term forecasting of time series based on the clonal selection model and investigation of its efficiency.

## 1 PROBLEM STATEMENT

One set time series  $Z(t) = z_1, z_2, \dots, z_N$  lengths  $N$  and  $S$  external factors represented in the form of time series  $X_1(t), X_2(t), \dots, X_S(t)$ , the values of which are obtained in the time points as follows  $t_1, t_2, \dots, t_N$ . It is required to determine future values of the series  $Z(t) = z_{N+1}, z_{N+2}, \dots, z_{N+k}$ , where  $k$  is a length of forecasting horizon.

Application of approaches based on different models of AIS allows to achieve good results in solving the problem of short-term forecasting [2, 4–6], but their using characterized by a high computational complexity. One of the ways of efficient organization of computational process may be to use of distributed computing on the multiproCESSOR systems, for example, using MPI technologies. It is required to perform algorithm decomposition and to propose ways of distribution of problems among processors, thereafter to compare the efficiency of the methods proposed by means of evaluation of communication complexity according to Hockney's model [7].

The greatest influence on timing characteristics of immune algorithm has the size of population of antibodies. It is required to determine dependence of AIS learning time on the number of antibodies, thereafter to conduct similar experiments for multiprocessor and multi-core architecture.

## 2 REVIEW OF THE LITERATURE

Proposed in [5] immune algorithm based on the immune network model that uses CBR is successfully applied to solve problems of short-term forecasting, but it demonstrates a poor result for short-memory time series that is defined by the peculiarities of CBR. Partially this problem is solved by the approach [2], based on the application of clonal selection model using heterogeneous antibodies built on the basis of CBR and the simplest forecasting techniques. This approach uses a segmentation of time series and selection for each segment the most efficient forecasting technique. This allows achieving the higher efficiency rather than the application of every predictor individually of the given set for forecasting of the entire basic series. The further development of this approach [6] allows not only to make a selection of the actual external factors, but also it successfully works with the distorted time series, however, this requires a significant increase in the number of antibodies in the population.

Time series forecasting using proposed in [6] immune algorithm is a computationally capacious procedure since,

at first, it is required to train a system on a large training set. To solve this problem it is reasonable to use tools for organization of the distributed computing [7], such as MPI library, and TPL [8]. TPL allows significantly facilitating a writing of a maintainable code that automatically uses the advantages of multi-core systems without changing the application design [8]. This technology will be efficient, if the algorithm includes many cycles that use shared arrays of data.

In the paper presented here the approach proposed in [6] obtained further development using the technologies of parallel programming for multi-core and multiple-processor cluster systems.

## 3 MATERIALS AND METHODS

In this study the model of clonal selection [2, 4, 6] was chosen for parallelization. This model is based on searching of antibodies (decision options), which most fully conform to an antigen (original problem), basing on the knowledge of the affinity (proximity measure between antibody and antigen).

It is possible to present the formal way of AIS model to forecast problem solving [2, 6] in following manner:

$$AIS\_CS = \langle Ab, n', n_c, Ab^C, Ab^M, Ag, D_{Ab-Ag}, \sigma_d, D_{Ab-Ab}, \sigma_s, \sigma_{age} \rangle. \quad (1)$$

Antibody acts like precedent and consists of two parts. The structure of the first part is similar to one of antigen (but it includes sample of one outside factor only), and it represents a set of parameters, describing the problem statement (in our case that are samples of series with known values, including missing values), and it is used to define affinity value. The second part, the length of which is equal to the forecast horizon does not affect to the calculated value of affinity of the antibody and describes the proposed prediction for that sample values of the time series, which is the first part of antibody.

Antigen is combination of values of forecasting time series and attendant series going ahead directly before forecasting values. In that case, forecast problem comes down to a problem of antibody search with the max affinity value, thus, it is search of sample which matches the most to values at the last part of the time series.

In order to use advantages of the way presented at [6], it is required usage of the big size population of antibodies. Overall number of antibodies included in population without missing values of predicted time series can be defined by following way:

$$n = \sum_{i=L_{\min}}^{L_{\max}} ((N - (L_i + k) + 1) N_m S). \quad (2)$$

Using of different possible values of antibody length allows to configure the AIS to forecasting time series [5].

This work represents two methods of immune algorithm parallelization based on island model of the genetic algorithm [9]. Using data parallelism in the task consists in uniform

(except for the control node) distribution of antibodies population among computational nodes and follow-up independently of evolutionary process by every computation node. It helps to use sizeable population of antibodies when using serial version of immune algorithm will exact intolerably much of work time.

General concept for these two parallel realizations of the immune algorithm is using of control process (rank of which is equal 0), which distributes the starting data and forms result after end of work of  $N_{cn}$  computational nodes (with rank more than 0). Control node doesn't have part of antibodies population for cultivation. Size of population on the computational nodes is defined as

$$n_1 = \frac{n}{N_{cn} - 1}. \quad (3)$$

First way of parallelization of the immune algorithm is:

1. Each of the computational nodes generates the unique antibodies populations which have size  $n_1$ .
2. The control node delivers the antigen to computational nodes.
3. Each of the computational nodes selects the antibodies with the most affinity and sends to control node the clone subpopulation which have size  $n_c n'$ .
4. The control node compares the variants of forecast, selects the antibody with the biggest affinity value and forms the result on this iteration.
5. If there is a real value of predicted value, the control node compares the forecast variants of selected antibodies and real value, performs mutation of antibodies of clone population and corrects the coefficients which determine the influence of antibodies type and external factors on the antibodies affinity [6].
6. The control node resolves about inculcation or deletion of clones and being formed the memory cell population  $Ab^M$  for each node.
7. The memory cell population and changed parameters for calculating of the affinity are sent to the initial computational nodes.
8. Steps 2–7 are repeated for each antigen from the training set to achieve one of the stop criteria – carrying out a predetermined number of antibodies generations or reducing forecast error up to the user-specified value.

After the end of the training phase the order of interaction between nodes remains the same. This method of parallelization ensures the saving of unique populations of antibodies on each computational node, having shared between nodes only operations affinity determination and selection of antibodies.

To estimate the time of data transmission between compute nodes, the model proposed by Hockney is used [7]. The volume of data transferred on different iterations of the algorithm may vary significantly. After forming of the population at each iteration control node sends one antigen to each computational node, receives from a population of clones, and returns back to a population of memory cells,

whose size is  $n_m \in [0, n_c n']$ . The volume of control data can be ignored. Then the duration of communication operations in the learning process of AIS for the first version of the parallel algorithm can be expressed as

$$T = (N_{cn} - 1) \left( N_g \left( \alpha + \frac{V_{ag} + V_{ab} n_c n' + V_{ab} n_m}{\beta} \right) \right). \quad (4)$$

Advantages of this approach are:

- unique antibodies in different populations at computational nodes, each node has an independent population of memory cells;
- supervision of algorithm parameters for each node at each iteration.

However, control of process in populations on computational nodes leads to considerable network load, since it occurs a transfer of large amounts of data on each iteration of the algorithm. In order to reduce the communication operations in the second version of the parallel immune algorithm is proposed distribution of the learning process of AIS and population control for individual computational nodes, without the exchange of intermediate results to the control node:

1. On each computational node it forms a unique population of antibodies which size is  $n_1$ .
2. The control node carry out the delivering of training set (antigen set and actual values of forecasting value) to all computational nodes. Every computational node performs step 3–7.
3. Selection and cloning of antibody with the highest affinity for the antigen.
4. Mutation of antibody clones from the population size  $n_c n'$ .
5. Correction of coefficients that determine the impact of realized in the antibody forecasting method and represented the external factor on antibody affinity.
6. Introducing or removing clones from the population regardless of the population of antibodies on other computational nodes.
7. Generates a population of memory cells  $Ab^M$ .
8. Steps 3–7 are repeated for each antigen from the training set to achieve one of the stop criteria.

After completion of the training phase the interaction between the control and the computational nodes is similar to steps 2–7 of the first variant of the algorithm. Computational nodes, having completed the work, send the best result to control node that determines the best individual from the sent options and generates the result of immune algorithm and adjusts the selection parameters on the computational nodes.

The main advantage of this approach is low network load in the learning process – the transfer of data is limited by the teaching selection, consisting of  $m$  antigen:

$$T = (N_{cn} - 1) \left( \alpha + \frac{m V_{ag}}{\beta} \right). \quad (5)$$

The disadvantage of this approach is the possible occurrence of the same antibodies in populations on different computational nodes during the learning process of AIS. Using of suppression will require additional data exchanges between nodes. In addition, there is an irregular loading of nodes at the end of the learning phase: when the best affinity is achieved on one process, there is a standing time, waiting the completion of the rest.

Since the proposed ways of parallelization of the immune algorithm differ mainly in the amount of information exchanges in the learning process of AIS, and further exchange of data is the same, then for the comparison of the approaches is sufficient to determine the duration of data transfer operations. Comparison of the obtained expressions 4–5 shows that the second developed method of parallel computation has considerably lower communication overheads and allows a better scalability as the number of processors increased. Therefore, further studies are carried out with the use of a second approach of the parallelization of the immune algorithm.

#### 4 EXPERIMENTS

To check the efficiency of the parallel algorithm one chose a series of daily values of temperature (over 20000 values) that allows to create an extended training set. For computational experiment in this paper was used cluster, built on the basis of Microsoft Windows Compute Cluster Server 2003 (MS CCS) which supports the MPI.NET library. MPI.NET is an implementation of the MPI library for the platform .NET. This system provides safe and scalable management of the cluster resources, task scheduling, and presents a message passing interface for distributed programming [7]. Before the analysis the parallel version of the immune algorithm is required to determine the maximum size of the population of antibodies, for which it is advisable to use considered in [6], the immune algorithm. Obtained by computational experiment population size is then used in the determination of acceleration for a parallel version of the algorithm.

Table 1 – The dependence of the execution time of the algorithm on the number of antibodies in the population for the serial immune algorithm (Intel Pentium D, 3000 MHz)

Number of antibodies, $10^3$	20	40	80	100	120	140	160	200
Time, sec	3.11	6.63	14.43	18.19	23.82	27.61	32.97	44.85

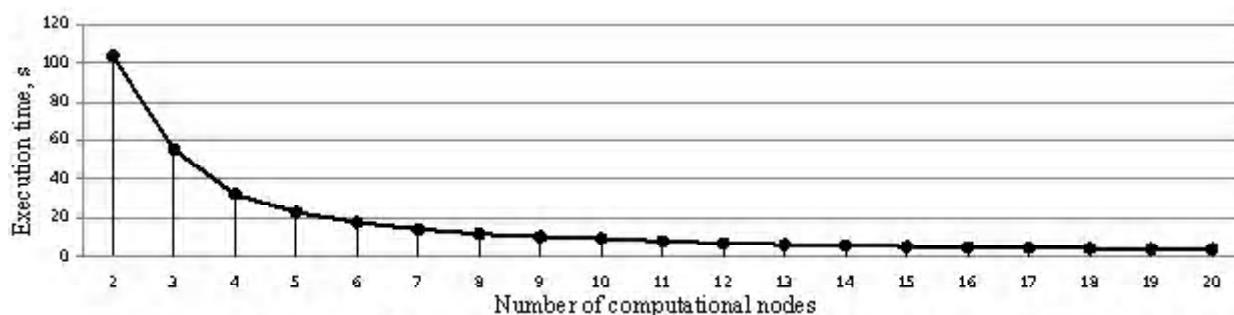


Figure 1 – Dependence of learning time on the number of processors

During the computational experiment was carried out sequential start of parallel algorithm with the same stopping criterion and maximum (for the serial version) population size for different (1–20) number of available computational nodes in the cluster.

For the analysis of efficiency of code optimization for multi-core architecture (used 2-core processor model Intel Pentium D, 3000 MHz) has been made a series of launches of the sequential version of the program adapted to using technology TPL, with different size populations of antibodies.

#### 5 RESULTS

The tab. 1 shows the time of execution of one iteration of the algorithm during the AIS learning depending on the size of the antibodies population.

Fig. 1 illustrates dependence of training time of time series forecasting module on the involved number of processors showing the acceleration close to linear one. When decreasing a number of antibodies in the populations on particular computational nodes a superlinear acceleration is potentially possible.

In the immune algorithm of time series forecasting there are cycles (in particular, determination of affine properties in the populations antibodies and clones) for which a parallel processing is possible, in particular, affinity of every antibody in the population calculates without regard for another ones. As a result of the immune algorithm application adjusted for use of TPL technology an average acceleration has reached 7.5 % (tab. 2).

Fig. 2 illustrates a difference between times of training of the algorithm versions sequential and adjusted for TPL for Intel Pentium D, 3000 MHz dual core processor.

#### 6 DISCUSSION

The results quoted in [6] of the short-term forecasting of time series represented in M3-Competition [10], and distorted series of daily average temperature confirm the effectiveness of the proposed immune algorithm for the short-term

Table 2 – Acceleration in case of use of TPL on the dual core processor, with various numbers of antibodies in the population

Number of antibodies, $10^3$	6	7	8	9	10	11	12	13	14	15
Speedup, %	3.94	4.48	6.0	4.61	14.82	9.49	11.85	7.06	7.04	4.72

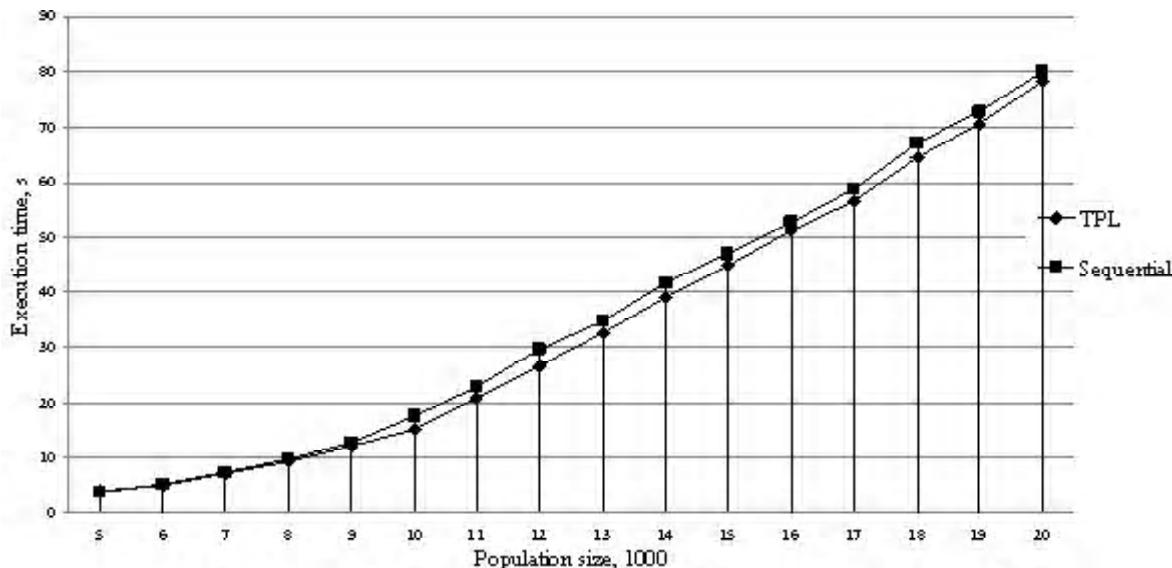


Figure 2 – Dependence of immune algorithm training time on a size of population

forecasting of the distorted time series. However, the growth of a number of missing values requires an increase in the learning sample and the size of antibodies, which leads to the growth of population and requires the use of parallel realization of the immune algorithm, which provides the acceleration close to linear one.

Low efficiency of application of TPL is explained by the fact that in the reviewed modules of the immune algorithm the variables of the most cycles are interdependent that leads to the complication of optimization of the managed code for multi-core processors. Furthermore, in order to get speed gain of parallel program it is necessary that the working time of parallel processes in the areas of code parallelization would exceeds significantly the working hours of generation of parallel flows.

## CONCLUSIONS

The paper studied two ways of parallelization of the immune algorithm on the basis of the proposed in [6] hybrid algorithm of short-term time series forecasting. One used a principle of data paralleling on the basis of an island model of genetic algorithm. Were analyzed the considered methods of parallelization on the basis of consideration of communication complexity.

The feature of the first method of parallelization is ability to control the evolution of AIS on computation nodes by the control node, but this leads to a significant communication load. The peculiarity of the second way of parallelization is a minimization of communication overheads within a cluster by means of organization of independent evolutionary process on individual computational nodes during AIS learning. After the processing of training set an interaction between computational nodes and control node takes place only after every integration of algorithm that

does not allow in future to use the advantages of a parallel version due to high communication overheads. Since the training time of the system significantly exceeds the receipt time of one forecast, this disadvantage can be neglected.

To implement data exchange between the computational nodes was used MPI.NET technology, to accelerate calculations within the populations of individual nodes was used optimization of managed code for multi-core processor using TPL. Using of the distributed computing allows for acceleration close to linear one, while the code optimization for multi-core architecture allows to obtain the acceleration only about 7 %.

Proposed parallel realizations of the immune algorithms have a high speed of computation of forecasting values and an accuracy of forecasting of different time series comparable with other models.

## ACKNOWLEDGEMENTS

This work was performed as part of state budget scientific research work № 273 «Neuro-fuzzy systems for online clustering and classification of data sequences under conditions of distortion by missing and anomalous observations», section № 273-2 «Adaptive techniques and models for data classification and time series prediction in conditions of presence of missing data and outliers, based on artificial immune systems» (state registration number 0113U000361).

## REFERENCES

- Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов / Ю. П. Лукашин. – М. : Финансы и статистика, 2003. – 416 с.
- Кораблев Н. М. Применение модели клonalного отбора, использующей вывод по прецедентам, для прогнозирования временных рядов / Н. М. Кораблев, Г. С. Иванченко // Бионика интеллекта. – № 1(80). – С. 108–111.

3. Dasgupta D. Artificial Immune Systems and Their Applications / D. Dasgupta. – Springer-Verlag, 1999. – 306 p. DOI: 10.1007/978-3-642-59901-9
4. Бидюк П. И. Алгоритм клonalного отбора для прогнозирования нестационарных динамических систем / П. И. Бидюк, В. И. Литвиненко, И. В. Баклан, А. А. Фефелов // Искусственный интеллект. – 2004. – № 4. – С. 89–99.
5. Кораблев Н. М. Применение искусственных иммунных сетей для прогнозирования временных рядов / Н. М. Кораблев, Г. С. Иващенко // Системы обработки информации. – Харьков. – 2012. – № 9(107). – С. 42–45.
6. Кораблев Н. М. Применение модели клonalного отбора для прогнозирования временных рядов, имеющих пропущенные значения / Н. М. Кораблев, Г. С. Иващенко // Электротехнические и компьютерные системы. – Одесса. – 2014. – № 13(89). – С. 170–177.
7. Воеводин В. В. Параллельные вычисления. / В. В. Воеводин, Вл.В. Воеводин – СПб.: БХВ-Петербург, 2002. – 475 с.
8. Timothy G. M. Patterns for Parallel Programming / G. M. Timothy, A. S. Beverly, L. M. Berna. – Addison-Wesley, 2004. – 384 p. DOI: 10.1145/1953611.1953620
9. Монахов О. Г. Параллельный генетический алгоритм для оптимизации торговых стратегий / О. Г. Монахов // Сибирский журнал вычислительной математики. – 2008. – Т. 11, № 4. – С. 423–432.
10. Makridakis S. The M-3 Competition: Results, Conclusions and Implications / S. Makridakis, M. Hibon // International of Forecasting. – Amsterdam : Elsevier, 2000. – № 16. – Р. 451–476. DOI: 10.1016/S0169-2070(00)00057-1

Article was submitted 03.11.2014.

After revision 05.11.2014.

Кораблев Н. М.<sup>1</sup>, Иващенко Г. С.<sup>2</sup><sup>1</sup>Д-р техн. наук, профессор, профессор кафедры электронных вычислительных машин, Харьковский национальный университет радиоэлектроники, Украина<sup>2</sup>Аспирант, Харьковский национальный университет радиоэлектроники, Украина

## ПАРАЛЛЕЛЬНИЙ ІММУННИЙ АЛГОРИТМ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАННЯ НА ОСНОВЕ МОДЕЛІ КЛОНАЛЬНОГО ВІДБОРУ

Рассмотрены способы распараллеливания гибридного иммунного алгоритма краткосрочного прогнозирования временных рядов, построенного на основе модели клonalного отбора, использующей вывод по прецедентам и простейшие методы прогнозирования. Проведен анализ двух способов распараллеливания, отличающихся порядком обмена сообщениями между вычислительными узлами. Для реализации предложенного алгоритма используется технология MPI.NET для систем с передачей сообщений. Для оптимизации работы отдельных вычислительных узлов использована библиотека TPL для систем с общей памятью. Представлены результаты экспериментальных исследований, демонстрирующие эффективность предлагаемого подхода.

**Ключевые слова:** прогнозирование, временной ряд, искусственные иммунные системы, модель клonalного отбора, антитело, антиген, аффинность, клонирование, мутация.

Корабльов М. М.<sup>1</sup>, Иващенко Г. С.<sup>2</sup><sup>1</sup>Д-р техн. наук, профессор, профессор кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки, Україна

## ПАРАЛЛЕЛЬНИЙ ІМУННИЙ АЛГОРИТМ КОРОТКОСТРОКОВОГО ПРОГНОЗУВАННЯ НА ОСНОВІ МОДЕЛІ КЛОНАЛЬНОГО ВІДБОРУ

Розглянуто способи розпаралелювання гібридного імунного алгоритму короткосрокового прогнозування часових рядів, побудованого на основі моделі клonalного відбору, яка використовує висновок за прецедентами та найпростіші методи прогнозування. Проведено аналіз двох способів розпаралелювання, що відрізняються порядком обміну повідомленнями між обчислювальними вузлами. Для реалізації запропонованого алгоритму використовується технологія MPI.NET для систем з передачею повідомлень. Для оптимізації роботи окремих обчислювальних вузлів застосована бібліотека TPL для систем із загальною пам'яттю. Представлені результати експериментальних досліджень, що демонструють ефективність запропонованого підходу.

**Ключові слова:** прогнозування, часовий ряд, штучні імунні системи, модель клonalного відбору, антитіло, антиген, афінність, клонування, мутація.

## REFERENCES

1. Lukashin Yu. P. Adaptivnye metody kratkosrochnogo prognozirovaniya vremennykh ryadov [Adaptive methods of short-term time series forecasting]. Moskow, Finansy i statistika, 2003, 416 p.
2. Korablev N. M., Ivashchenko G. S. Primenenie modeli klonal'nogo otbora, ispol'zuyushchei vydov po pretsedentam, dlya prognozirovaniya vremennykh ryadov [Application of the clonal selection model using the case based reasoning for time series forecasting], *Bionika intellekta*. Kharkov, 2013, pp. 108–111.
3. Dasgupta D. Artificial Immune Systems and Their Applications. Springer-Verlag, 1999, 306 p. DOI: 10.1007/978-3-642-59901-9
4. Bidyuk P. I., Litvinenko V. I., Baklan I. V., Fefelov A. A. Algoritm klonal'nogo otbora dlya prognozirovaniya nestatsionarnykh dinamicheskikh sistem [Clonal selection algorithm for the prediction of nonstationary dynamical systems], *Iskusstvennyi intellekt*, 2004, No. 4, pp. 89–99.
5. Korablev N. M., Ivashchenko G. S. Primenenie iskusstvennykh imunnnykh setei dlya prognozirovaniya vremennykh ryadov [Application of artificial immune networks for time series forecasting], *Sistemi obrobki informatsii*. Kharkov, 2012, No. 9, pp. 42–45.
6. Korablev N. M., Ivashchenko G. S. Model of clonal selection for forecasting time series with missing data [Primenenie modeli klonal'nogo otbora dlya prognozirovaniya vremennykh ryadov, imeyushchikh propushchennye znacheniya], *Elektrotehnicheskie i kom'yuternye sistemy*. Odessa, 2014, No.13 (89), pp. 170–177.
7. Voevodin V. V., Voevodin Vl. V. Parallel'nye vychisleniya [Parallel computing] Sankt-Peterburg, BKhV-Peterburg, 2002, 475 p.
8. Timothy G. M., Beverly A. S., Berna L. M. Patterns for Parallel Programming. Addison-Wesley, 2004, 384 p. DOI: 10.1145/1953611.1953620
9. Monakhov O. G. Parallel'nyi geneticheskii algoritm dlya optimizatsii torgovykh strategii [Parallel genetic algorithm for optimization of trading strategies], *Sibirskii zhurnal vychislitel'noi matematiki*, 2008, Vol. 11, No. 4, pp. 423–432.
10. Makridakis S., Hibon M. The M-3 Competition: Results, Conclusions and Implications / S. Makridakis, M. Hibon // International of Forecasting. Amsterdam, Elsevier, 2000, No. 16, pp. 451–476. DOI: 10.1016/S0169-2070(00)00057-1

UDC 004.942

Kucherenko Ye. I.<sup>1</sup>, Trokhimchuk S. N.<sup>2</sup>, Driuk O. D.<sup>3</sup>

<sup>1</sup>Dr. Sc., Professor, Professor of Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Ukraine

<sup>2</sup>Postgraduate student of Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Ukraine

<sup>3</sup>Postgraduate student of Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Ukraine

## KNOWLEDGE-ORIENTED TECHNOLOGIES IN HIGHLY AUTOMATED PRODUCTION

The expansion of Zadeh-Mamdani method in problems of fuzzy inference on knowledge is considered. A modified method of fuzzy inference is proposed and justified. The proposed method is based on interpretation of components of fuzzy Petri nets as production rules and solving of logical equations in the state space of membership functions of the model, followed by their defuzzification.

The process of perceptron learning as procedure of adjusting the weights and shifts to decrease the difference between target and real signals on its output, using a definite tuning (learning) rule is defined. Modified methods of gradient procedures based on the method of back-propagation for multilayer neural networks are developed.

Application of the proposed approaches based on advanced hybrid models with solving the problems of fuzzy inference and operative informed decision making allowed to reduce the time to identify, locate and eliminate the causes of failure on the set of alternatives, which is confirmed by experiment.

The method appears to be universal in decision-making problems and allows to increase the adequacy of hybrid model and the accuracy of decisions.

**Keywords:** Zadeh-Mamdani, modification, knowledge, rules, inference, Petri net, dichotomy, defuzzification.

### NOMENCLATURE

$\cup$  denotes the union of hybrid model components;  
 $\vee$  denotes the maximum of fuzzy parameters;  
 $\wedge$  denotes the minimum of fuzzy parameters;  
 $*$  denotes the operation of element-wise vectors multiplication;

ANN stands for Artificial Neural Network;  
FKB stands for Fuzzy Knowledge Base;  
HM stands for Hybrid Model;  
PN stands for Petri Net;  
PR stands for Production Rule;

$a_k^{qM}$  is the  $k$ -th element of output vector in  $M$ -th layer for the element from the  $q$ -th sample;

$F$  is a fuzzy incidence function  $F : (P \times T) \cup (T \times P)$ ;  
 $L_k$  are predicates corresponding to space decisions;  
 $L_n$  are predicates corresponding to logical decisions;  
 $L'$  is a subset of production rules;

$M_0$  is a vector of fuzzy initial marking of fuzzy positions of a PN;

$\{M_0\}$  is a set of production rules;

$P$  is a finite set of positions in PN;

$r$  is number of layer in ANN;

$S^{(ANN)}$  is an ANN-based model;

$S^{(k)}$  is a fuzzy state space;

$S^M$  is a number of neurons in the layer;

$T$  is a finite set of transitions in PN;

$\{a_\gamma\}$  is a set of perceptron outputs;

$\{\tau_{i,j}\}$  is a set of discrete time characteristics referred to transitions, positions and components of incidence functions, where  $i \in I$ ,  $j \in J$ .

### INTRODUCTION

It is known that construction of PN and their representation by FKB rules is less time consuming process than usual approaches in computer-aided design. Note that Zadeh-Mamdani approaches are more functional compared to Takagi-Sugeno-Kang method, resolution method and others. This fact has determined the direction of research.

The aim of the research is to develop a HM based on integration of PN and PR on knowledge in problems of fuzzy inference. The work is actual and important for the purpose of making decisions about the state of technological object under uncertainty. This is especially important for automated approach to design the FKB, where an algorithm of object functioning causes problems in its knowledge-based interpretation.

### 1 PROBLEM STATEMENT

Suppose a hybrid model  $S^{(k)}_{HM}$  [1] which operates on the object under parametric and structural uncertainty.

Some part of the data is implemented in the form of fuzzy production rules  $\{if / then_\alpha\}$ ,  $\alpha \in A$ , which produce the FKB.

In problems where the linear approximation is not enough, linear models do not work well, so it is necessary to consider the application of ANN in problems of reproducing the complex dependencies, classification and identification of objects.

For solving the applied problems in the work it is necessary:

- to propose and justify a new hybrid model based on the integration of PN, knowledge-based PR and ANN, as a means of selection of alternatives on the set of PN positions in problems of fuzzy inference;
- to solve the problem of ANN learning using methods based on gradient procedures;
- to propose and justify a modified method of fuzzy inference, based on interpretation of the components of fuzzy PN by production rules, solving the logical equations in a state space of membership functions and classifying the rules followed by defuzzification;
- to confirm the effectiveness of the approach by experiment;
- determine the prospects for further research.

## 2 REVIEW OF THE LITERATURE

Currently, there are many solutions for a wide class of problems, methods and models of intelligent control of production systems [2–4]. The main tools to implement the approaches are extensions of Petri nets [5] and knowledge-oriented methods and models based on fuzzy logic [5, 6], which help to decrease the degree of uncertainty of the solutions. ANN [7, 8] as an approach for modeling processes of different complexity is a universal tool of modeling, classification and pattern recognition.

The advantages of using the extended Petri nets for systems modeling are following:

- concurrency of processes representation;
- modeling of the dynamics of processes in the space of parameters state;
- convenient representation of determinism properties, probabilistic and fuzzy processes.

The main disadvantages of this approach are difficulty of continuous processes representation and existence of conflicts. This requires additional research.

Knowledge-oriented methods based on frame models, production rules, knowledge based on ontological spaces processing have an important advantage – that is, reducing uncertainty degree of processes and objects, as well as possibility of production rules reduction. This increases the performance in problems of inference. The disadvantage is the complexity of knowledge acquisition, tuning of membership functions of fuzzy processes and their interpretation.

The advantage of ANN is their focus on a wide class of problems of data mining in decision-making systems. The disadvantage is the complexity of training the ANN and interpretation of modeling results, determined by the nonlinear nature of the network structure.

In this paper, to implement the knowledge-oriented technologies, it is proposed to use new approaches for expansion and integration of hybrid models [1, 5], which have the advantages of partial models along with significant reduction of their disadvantages.

## 3 MATERIALS AND METHODS

Given PN in the form [1]:

$$S^{(k)} = \langle P, T, F, M_0, L_n, L_k, \{\tau_{i,j}\} \rangle. \quad (1)$$

Components of model  $S^{(k)}$  (1) are defined in a fuzzy state space  $S^{(k)}$ :  $\mu(k) \rightarrow [0, 1]$ . Then HM can be represented as

$$S^{(k)}_{HM} = S^{(k)} \cup \{if / then_\alpha\}, \alpha \in A, \quad (2)$$

where  $\{if / then_\alpha\}$ ,  $\alpha \in A$  is a set of fuzzy PR,  $\alpha \in A$  is a set of indexes of PR,  $\cup$  is a character that defines the union of HM components, expanded functionally for procedures of Zadeh-Mamdani fuzzy inference [9]:

$$y' = \vee x' \wedge \mu(x, y)_{HM}. \quad (3)$$

*Statement 1.* If  $\exists \{p_j\} \subset P, M \subset \{M_0\}, M_{p_j} > 0 | \{p_j\} \neq \emptyset$  for any subset of rules  $L'$ , then fuzzy inference can be implemented on  $L'$ .

The correctness of the statement is obvious, given that the implementation of the rules on  $L'$  reduces the cardinality of the set of rules  $|M_0|$ , and this increases the speed of inference on rules.

Adequacy of fuzzy inference procedures (3) on model (2) is fully defined by interpretation (1). Indeed, let's consider a computational algorithm fragment

$$A \xrightarrow{V} M, \quad (4)$$

for which, using the interpretation rules (1), we obtain (fig. 1).

Then, using the reflection shown in fig. 1, we construct a rule from (2):

$$if x is \mu(x) \rightarrow «small» then y is \mu(y) \rightarrow «average» \quad (5)$$

with terms of the linguistic variables:

$$\mu(x) = \exp(-k_1 x^2) \text{ is «small»,} \quad (6)$$

$$\mu(y) = \exp(-k_2 (x - \beta)^2) \text{ is «average».} \quad (7)$$

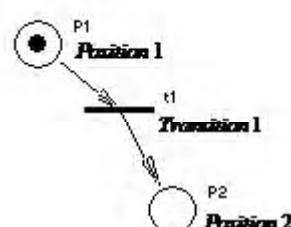


Figure 1 – Reflection of computational algorithm fragment

$$A \xrightarrow{V} M$$

In the paper the problem of learning and tuning the parameters of membership functions  $k_1, k_2, \beta$  (6), (7) is also solved. For this purpose, we used the method of dichotomy [10].

Using fuzzy Mamdani relation [9]:

$$\mu(x, y) = \min(\mu(x), \mu(y)), \quad (8)$$

we construct a matrix of relations, set the vector  $x'$  and find solution (5) based on (3) with further defuzzification based on mass center method. The adequacy of reflection of algorithm (4) is defined by setting in rule (5):

$$x \rightarrow \text{position } p_1, y \rightarrow \text{position } p_2,$$

that is reflection of input  $A$  and output  $M$  respectively of the algorithm (4). Similarly we can show the adequacy of reflection for other algorithm fragments, confirmed by experiment.

Further development of modeling processes, including continuous ones, makes it necessary to use alternative procedures to extend the models (2).

Artificial neural networks (ANN) [11] are extremely powerful modeling tools, which allow to reproduce very complex dependencies. In particular, neural networks are nonlinear by their nature. For many years linear modeling was the main modeling method in most areas, as it had well-designed optimization procedures. However, in problems where the linear approximation is not enough, linear models do not work well.

An important approach for further ANN models representation is perceptron [8], which is the most studied. Neuron used in the perceptron model has a step activation function *hardlim* with strict limitations (fig. 2).

Each element of the perceptron input vector is weighed with appropriate weight  $w_{ij}$ , and their sum is the input of the activation function. Perceptron neuron returns 1, if activation function input  $n > 0$ , or 0, if  $n < 0$ .

Activation function with strict limitations allows the perceptron to classify input vectors, dividing the space of the inputs into 2 areas, as it is shown in fig. 3 for the perceptron with two inputs and shift.

Input space is divided into two areas by separating line  $L$ , which in two-dimensional case is defined by the following equation:

$$w^T p + b = 0. \quad (9)$$

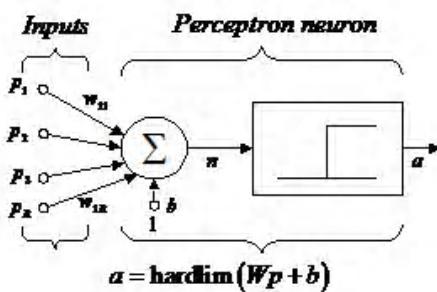


Figure 2 – Perceptron model

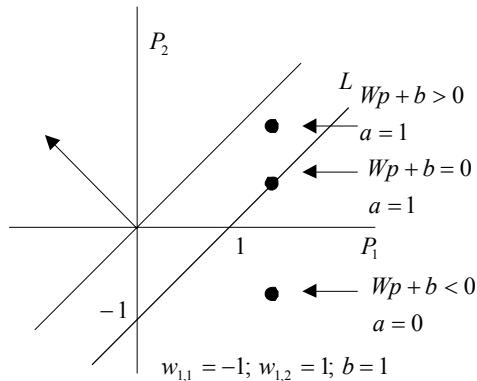


Figure 3 – Input vectors classification

This line is perpendicular to weight vector  $w$  and is shifted by  $b$ . Input vectors above the line  $L$  correspond to the positive potential of the neuron and, therefore, perceptron output for these vectors is equal to 1; vectors below the line correspond to perceptron output equal to 0.

When changing the values of shift and weights, border of line changes its position. Perceptron without shift always forms the separating line passing through the coordinate origin; adding the shift forms the line which doesn't pass through the coordinate origin, as it is shown in fig. 3. If the dimension of the input vector is greater than 2, separating border appears to be a hyperplane.

Perceptron consists of a single layer including 5 neurons, as it is shown in fig. 4; weights  $w_{ij}$  are transfer coefficients from  $j^{th}$  output to  $i^{th}$  neuron. Single-layer perceptron equation is as follows:

$$a = f(Wp + b). \quad (10)$$

Set of perceptron outputs (fig. 4)  $\{a_\gamma\}, \gamma \in H$  is the object of learning, and this appears to be a non-trivial task.

Suppose that model (2) is defined. It is focused on the discrete processes  $P^{(D)}$ . The requirement of the processing of strongly nonlinear continuous processes made it necessary to create a model  $S_\Sigma^{(k)}$ .

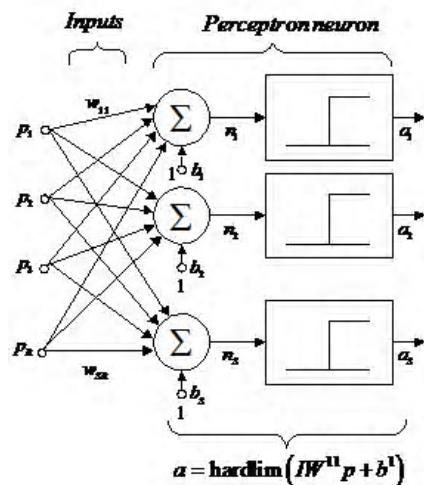


Figure 4 – Perceptron structure

*Statement 2.* If we are given the model (2), its extension by assuming  $\exists p_j \in P \mid p_j : M_{p_j}^{(ANN)} \neq \emptyset$  allows to expand the modeling and reliable decision-making area.

Correctness of statement 2 is obvious, if we consider the functional possibilities of ANN.

Then model  $S_{\Sigma}^{(k)}$  can be represented in the following form:

$$S_{\Sigma}^{(k)} = S^{(k)} \cup S^{(ANN)}.$$

*Statement 3.* Given a model (10), the set of alternatives for any position

$$p_j \in P \mid M_{p_j}^{(ANN)} \neq \emptyset \quad (11)$$

determines the development of processes ( $D$ ) in the set of output transitions

$$p_j \in \{p_i\} \mid \exists t_i \in \{t_j(\text{out})\}, t_i \rightarrow D \quad (12)$$

of the PN (1).

If the network (10) holds (11) and (12), then the problem of learning the network  $S^{(ANN)}$  occurs.

#### 4 EXPERIMENTS

Define the process of perceptron learning as procedure of adjusting the weights and shifts to decrease the difference between target and real signals on its output, using a definite *tuning (learning) rule*. Training procedures are divided into two classes: *supervised learning* and *unsupervised learning*.

Once the initial weights and shifts of the neurons are set by user or using random number generator, the network is ready to start its training procedure [5]. The most important learning methods can be considered the methods based on gradient learning procedures. Neural networks designed to solve practical problems can contain up to several thousands of adjustable parameters, so calculation of the gradient may require a rather high cost of computing resources. Given the specificity of multilayer neural networks, there are developed special methods of calculating the gradient, among which we should highlight the back-propagation method.

The term «back-propagation» refers to the process with which derivatives of the error functional by network parameters can be calculated. This process can be used in combination with different optimization strategies. There are also a lot of variations of back-propagation algorithm.

Consider the expression for the gradient of quality criterion by weighting coefficients for the output layer  $M$  of the network:

$$\frac{\partial J}{\partial w_{ij}^M} = \frac{\partial}{\partial w_{ij}^M} \left( \frac{1}{2} \sum_{q=1}^Q \sum_{k=1}^{S^M} \left( t_k^q - a_k^{qM} \right)^2 \right) = - \sum_{q=1}^Q \sum_{k=1}^{S^M} \left( t_k^q - a_k^{qM} \right) \frac{\partial a_k^{qM}}{\partial w_{ij}^M},$$

$$i = 1, \dots, S^M, j = 0, \dots, S^{M-1}. \quad (13)$$

Functioning rule of layer  $M$  is as follows:

$$a_k^{qM} = f_M \left( \sum_{l=0}^{S^{M-1}} w_{kl}^M a_l^{q(M-1)} \right), m = 1, \dots, S^M. \quad (14)$$

From the equation (13) we receive:

$$\frac{\partial a_k^{qM}}{\partial w_{ij}^M} = \begin{cases} 0, & k \neq i \\ f' \left( n_i^{qM} \right) a_j^{q(M-1)}, & k = i, \\ i = 1, \dots, S^{M-1}, \\ j = 0, \dots, S^{M-1}. \end{cases} \quad (15)$$

Substituting (15) into (13) we obtain:

$$\frac{\partial J}{\partial w_{ij}^M} = - \sum_{q=1}^Q \left( t_i^q - a_i^{qM} \right) f'_M \left( n_i^{qM} \right) a_i^{q(M-1)}.$$

If we denote

$$\Delta_i^{qM} = \left( t_i^q - a_i^{qM} \right) f'_M \left( n_i^{qM} \right), i = 1, \dots, S^M, \quad (16)$$

we receive:

$$\frac{\partial J}{\partial w_{ij}^M} = - \sum \Delta_i^{qM} a_i^{q(M-1)}, i = 1, \dots, S^M, j = 1, \dots, S^{M-1}. \quad (17)$$

After calculating the weights  $w_{ij}^{M-1}$  of layer  $M-1$ , we receive the following general formula:

$$\frac{\partial J}{\partial w_{ij}^r} = - \sum_{q=1}^Q \Delta_i^{q(r-1)} a_j^{q(r-1)}, r = 1, \dots, M,$$

$$i = 1, \dots, S^r, j = 0, \dots, S^{r-1}, \quad (18)$$

where

$$\Delta_i^{qr} = \left( \sum_{k=1}^{S^{r+1}} \Delta_k^{q(r+1)} w_{ki}^{r+1} \right) f'_r \left( n_i^{qr} \right), r = 1, \dots, M-1,$$

$$\Delta_i^{qM} = \left( t_i^q - a_i^{qM} \right) f'_M \left( n_i^{qM} \right), i = 1, \dots, S^M. \quad (19)$$

Fig. 5 shows a diagram of calculations [12] corresponding to the expression (19).

In this diagram symbol  $\ast\ast$  denotes multiplication of vectors  $\Delta$  and  $a^T$ .

#### 5 RESULTS

In our research the system of statements justifying the extension of hybrid models in the form of integration of PN, fuzzy PR and ANN is formulated. The development of such models has allowed to implement the effective management of intellectual processes.

For ANN training gradient procedures on the basis of back-propagation of errors are proposed. In fig. 6 we can see the dependency between the number of performed training iterations and resulting error. This graph is built using function *train* in Matlab.

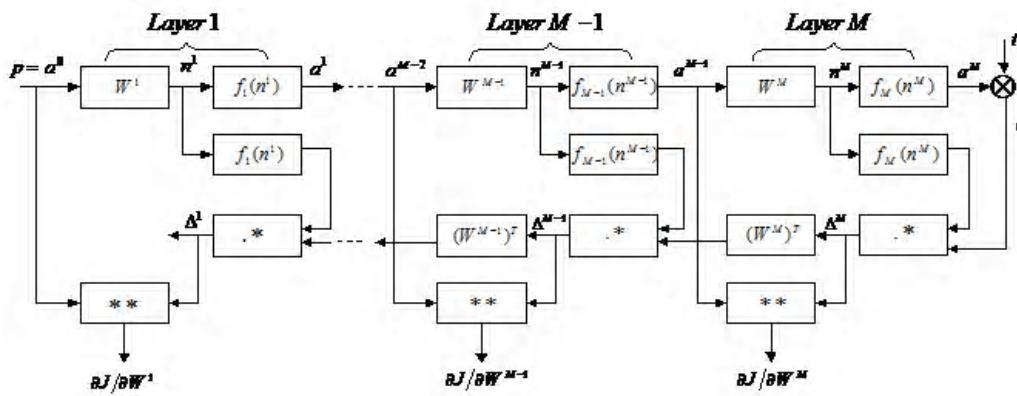


Figure 5 – Diagram of calculations

## 6 DISCUSSION

As noted in earlier publications [13], currently unique high-production machine-assembly areas of machinery enterprises have a modern production base including robot manipulators and cylindrical grinding machines with numerical control of foreign and domestic production equipped with active control systems, which allow realizing the management and control of the technological process. The scheme of the production unit is usually equipped with highly automated equipment with inductive transducer and the electronic measuring system with discrete 0,01; 0,1 microns and error <0,5 %.

In case of failure of the system, serious financial losses can occur, and it will affect the cost  $C$  of final product.

Application of the proposed approaches based on hybrid models (2), (9) with solving the problems of fuzzy inference and operative informed decision making allowed to reduce on the set of alternatives (11), (12) the time  $\tau$  to identify, locate and eliminate the causes of failure to 20 minutes, having the linguistic terms of maintenance stuff professional fitness «average» and higher, which is confirmed by experiment.

## CONCLUSION

Thus, we propose a new hybrid model based on the integration of PN, knowledge-based production rules and ANN as a tool of selection of alternatives on the set of PN

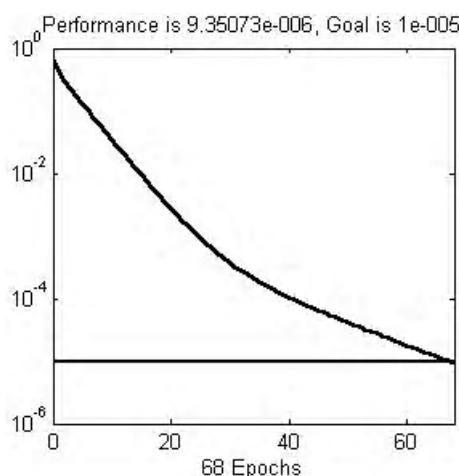


Figure 6 – Dependency between number of iterations and resulting error

positions in fuzzy inference problems. The problem of ANN training by methods based on the gradient procedures is formulated and solved. Modified method of fuzzy inference is proposed and justified. This method is based on the interpretation of the components of fuzzy PN by production rules and logical equations solving in the state space of membership functions of the model, and also rules classification with further defuzzification. Experiment confirmed the effectiveness of the approach. The prospects of further studies are identified.

## REFERENCES

1. Кучеренко Е. И. Гибридные модели и информационные технологии в управлении сложными объектами / Е. И. Кучеренко., С. Н. Трохимчук // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. – Луцьк : ЛНТУ, 2013. – С. 46–51.
2. Конкин Р. В. Методы ранжирования данных с учетом свойств нечетких систем / Р. В. Конкин // Вісник НТУ «ХПІ». Серія Нові рішення в сучасних технологіях. – Х. : НТУ «ХПІ». – 2013. – № 1 (977). – С. 26–30.
3. Jensen R. Computational intelligence and feature selection: rough and fuzzy approaches / R. Jensen, Q. Shen. – Hoboken : John Wiley & Sons, 2008. – 339 p.
4. Jang J. R. ANFIS: Adaptive-network-based fuzzy inference system / J. R. Jang // IEEE transactions on systems and cybernetics. – 1993. – Vol. 23. – P. 665–685. DOI: 10.1109/21.256541.
5. Бодянский Е. В. Нейро-фаззи сети Петри в задачах моделирования сложных систем : монография / Е. В. Бодянский, Е. И. Кучеренко, А. И. Михалев. – Дніпропетровськ : Системні технології, 2005. – 311 с.
6. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition / S. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2013. – Vol. 22, № 2. – P. 97–103. DOI: 10.3103/s1060992x13020082
7. Руденко О. Г. Основы теории искусственных нейронных сетей / О. Г. Руденко, Е. В. Бодянский. – Харьков : ТЕЛЕТЕХ, 2002. – 317 с.
8. Тарасенко О. П. Нейронно-мережні моделі якості : монографія / О. П. Тарасенко, С. М. Трохимчук. – Харків : УПА, 2013. – 115 с.
9. Tsoukalas L. H. Fuzzy and Neural Approaches in Engineering / L. H. Tsoukalas, R. E. Uhrig. – New York : John Wiley&Sons.Inc, 1997. – 587 p.
10. Методи, моделі та інформаційні технології оцінювання станів складних об'єктів / Кучеренко Є. І., Кучеренко В. Є., Гlushenkova I. C., Tvoroshenko I. C. – Харків : ХНАМГ, ХНУРЕ, 2012. – 276 с.

11. Бодянский Е. В. Интеллектуальное управление технологическими процессами : монография / Е. В. Бодянский, Е. И. Кучеренко, А. И. Михалев и др.]. – Днепропетровск : Национальная металлургическая академия Украины, 2013. – 213 с.
12. Бодянский Е. В. Основы теории искусственных нейронных сетей : монография / Е. В. Бодянский. О. Г. Руденко. – Харьков : ТЕЛЕТЕХ, 2002. – 317 с.
13. Кучеренко Е. И. Метод оценивания качества изделий механосборочного производства/ Е. И. Кучеренко, С. Н. Трохимчук // Збірник наукових праць Харківського університету Повітряних Сил. – 2014. – Вип. 2 (39). – С. 183–189.

Article was submitted 13.10.2014.  
After revision 19.10.2014.

Кучеренко Е. И.<sup>1</sup>, Трохимчук С. Н.<sup>2</sup>, Дрюк А. Д.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, профессор Харьковского национального университета радиоэлектроники, Украина

<sup>2</sup>Аспирант Харьковского национального университета радиоэлектроники, Украина

<sup>3</sup>Аспирант Харьковского национального университета радиоэлектроники, Украина

## ЗНАННЯ-ОРИЕНТИРОВАННІ ТЕХНОЛОГІЇ В ВЫСОКОАВТОМАТИЗИРОВАННЫХ ПРОИЗВОДСТВАХ

Рассматривается расширение метода Заде-Мамдани в задачах нечеткого логического вывода, основанного на знаниях. Предложен и обоснован модифицированный метод нечеткого логического вывода, основанный на интерпретации компонент нечетких сетей Петри как правил продукции и решений логических уравнений в пространстве состояний функций принадлежности модели с последующей дефазификацией.

Определен процесс обучения персептрона как процедуры настройки весов и смещений с целью уменьшить разность между желаемым (целевым) и истинным сигналами на его выходе, используя некоторое правило настройки (обучения). Для многослойных нейронных сетей разработаны модифицированные методы градиентных процедур, основанных на методе обратного распространения ошибки.

Применение предлагаемых подходов на основе расширенных гибридных моделей с решением задач нечеткого логического вывода и оперативного принятия обоснованных решений позволило на множестве альтернатив сократить время выявления, локализации и ликвидации причин отказа, что подтверждено экспериментом.

Метод является универсальным в задачах принятия решений и позволяет повысить адекватность гибридных моделей и точность принятия решений.

**Ключевые слова:** Заде-Мамдани, модификация, знания, правила, логический вывод, сети Петри, дихотомия, дефазификация.

Кучеренко Е. И.<sup>1</sup>, Трохимчук С. Н.<sup>2</sup>, Дрюк О. Д.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, профессор Харьковского национального университета радиоэлектроники, Украина

<sup>2</sup>Аспирант Харьковского национального университета радиоэлектроники, Украина

<sup>3</sup>Аспирант Харьковского национального университета радиоэлектроники, Украина

## ЗНАННЯ-ОРИЄНТОВАНІ ТЕХНОЛОГІЇ У ВИСОКОАВТОМАТИЗОВАНИХ ВИРОБНИЦТВАХ

Розглядається розширення методу Заде-Мамдані у задачах нечіткого логічного виведення, що засноване на знаннях. Запропоновані та обґрунтовані модифікований метод нечіткого логічного виведення, заснований на інтерпретації компонент нечітких сіток Петри як правил продукції та розв'язанні логічних рівнянь у просторі станів функцій належності моделі з подальшою дефазіфікацією.

Визначено процес навчання персептрона як процедури налаштування ваг та зміщень з метою зменшити різницю між бажаним (цільовим) та справжнім сигналами на його виході, використовуючи деяке правило налаштування (навчання). Для багатошарових нейронних мереж розроблені модифіковані методи градієнтних процедур, заснованих на методі зворотного поширення помилки.

Застосування запропонованих підходів на основі розширеніх гибридних моделей з розв'язанням задач нечіткого логічного виведення та оперативного прийняття обґрунтovаних рішень дозволило на множині альтернатив скратити час виявлення, локалізації та ліквідації причин відмови, що підтверджено експериментом.

Метод є універсальним у задачах прийняття рішень та дозволяє підвищити адекватність гибридних моделей та точність прийняття рішень.

**Ключові слова:** Заде-Мамдані, модифікація, знання, правила, логічне виведення, сітки Петри, дихотомія, дефазіфікація.

## REFERENCES

1. Kucherenko Ye. I., Trokhimchuk S. N. Gibridnyye modeli i informatsionnyye tekhnologii v upravlenii slozhnymi ob'yektami, *Komp'yuterno-intehrovanii tekhnoloziyi: osvita, nauka, vyrabnytstvo*. Lutsk, LNTU, 2013, pp. 46–51.
2. Konkin R. V. Metody ranzhirovaniya dannykh s uchetom svoystv nechetkikh sistem, *Visnyk NTU «KhPI». Seriya Novi rishennya v suchasnykh tekhnoloziyakh*. Kharkiv, NTU «KhPI», 2013, No. 1 (977), pp. 26–30.
3. Jensen R., Shen Q. Computational intelligence and feature selection: rough and fuzzy approaches. Hoboken, John Wiley & Sons, 2008, 339 p.
4. Jang J. R. ANFIS: Adaptive-network-based fuzzy inference system, *IEEE transactions on systems and cybernetics*, 1993, Vol. 23, pp. 665–685. DOI: 10.1109/21.256541.
5. Bodyanskiy Ye. V., Kucherenko Ye. I., Mikhalev A. I. Neyrofazzi seti Petri v zadachakh modelirovaniya slozhnykh sistem : monohrafiya. Dniproptetrov'sk, Systemni tekhnoloziyi, 2005, 311 p.
6. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition, Optical Memory and Neural Networks (Information Optics), 2013, Vol. 22, No. 2, pp. 97–103. DOI: 10.3103/s1060992x13020082
7. Rudenko O. G., Bodyanskiy Ye. V. Osnovy teorii iskusstvennykh neytronnykh setey. Kharkov, TELETEKH, 2002, 317 p.
8. Tarasenko O. P., Trokhymchuk S. M. Neyronno-merezhnii modeli yakosti :monohrafiya. Kharkiv, UIPA, 2013, 115 p.
9. Tsoukalas L. H., Uhrig R. E. Fuzzy and Neural Approaches in Engineering. New York, John Wiley&Sons.Inc, 1997, 587 p.
10. Kucherenko Ye. I., Kucherenko V. Ye., Hlushenkova I. S., Tvoroshchenko I. S. Metody, modeli ta informatsiyi tekhnoloziyi otsinyuvannya staniv skladnykh ob'yektiv : monografiya. Kharkiv, KhNAME, KhNURE, 2012, 276 p.
11. Bodyanskiy Ye. V., Kucherenko Ye. I., Mikhalev A. I., Filatov V. A., Gasik M. M., Kutsin V. S. Intellektaul'noye upravleniye tekhnologicheskimi protsessami : monografiya. Dnepropetrovsk, Natsional'naya metallurgicheskaya akademiya Ukrainsk, 2013, 213 p.
12. Bodyanskiy Ye. V., Rudenko O. G. Osnovy teorii iskusstvennykh neytronnykh setey : monografiya. Kharkov, TELETEKH, 2002, 317 p.
13. Kucherenko Ye. I., Trokhimchuk S. N. Metod otsenivaniya kachestva izdeliy mekhanosborochnogo proizvodstva, *Zbirnyk naukovykh prats' Kharkiv's'koho universytetu Povitryanykh Syl*, 2014, Iss. 2(39), pp. 183–189.

UDC 004.032.26::004.93+519.8

Romanuke V. V.

Dr. Sc., Professor, Professor of the Department of Applied Mathematics and social Informatics,  
Khmelnytskyi National University, Ukraine

## A FRAMEWORK FOR CLASSIFIER SINGLE TRAINING PARAMETER OPTIMIZATION ON TRAINING TWO-LAYER PERCEPTRON IN A PROBLEM OF TURNED 60-BY-80-IMAGES CLASSIFICATION

A 13-itemed scenario framework for classifier single training parameter optimization is developed. Formally, the problem is to find global extremum (mostly, minimum) of function as a classifier output parameter against its single training parameter. Linking the scenario theory to praxis, the classifier type has been decided on two-layer perceptron. Its input objects are monochrome images of a medium format, having a few thousands independent features. Within the framework, the programming environment has been decided on MATLAB, having powerful Neural Network Toolbox. Keeping in mind the stochasticity of the being minimized function, there is defined statistical  $\epsilon$ -stability of its evaluation by a finite set of data. These data are mined in batch testings of the trained classifier. For exemplification of the scenario framework, there is optimized pixel-to-turn standard deviations ratio for training two-layer perceptron in classifying monochrome 60-by-80-images of the enlarged 26 English alphabet capital letters. The goal is to find a pixel-to-turn standard deviations ratio for the training process in order to ensure minimum of classification error percentage. The optimization relative gain is about a third. The developed framework can be applied also for classifier multivariable optimization, wherein it instructs which item operations shall regard the corresponding multiplicity of variables.

**Keywords:** classifier training parameter optimization, statistical evaluation, optimization scenario, two-layer perceptron, classification error percentage, turned objects classification, monochrome image, pixel-to-turn standard deviations ratio, training set.

### NOMENCLATURE

CEP is a classification error percentage;  
CNN is a convolutional neural network;  
COS is a classifier operation speed;  
GT is a general totality;  
HMLNN is a hierarchical multilayered neural network;  
IMR is an information-memory resources;  
IRGSS is an independently random-generated subsets;  
LIF is a lot of independent features;  
4800-LCRM is a 4800-length-column-reshaped matrices;  
2LP is a two-layer perceptron;  
MI is a monochrome image;  
MOZ is a matrices of ones and zeros;  
PD is a pixel distortion;  
PDTM6080I is a pixel-distorted TM6080I;  
PTSDR is a pixel-to-turn standard deviations ratio;  
4800-250-26-P 2LP is a having 4800 neurons in the input layer, 250 in SHL, and 26 in the output;  
SD is a standard deviation;  
SHL is a single hidden layer;  
NDR is a non-distorted representatives;  
NV is a normal variate;  
TD is a turn distortion;  
TM6080I is a turned monochrome 60-by-80-image;  
ZEUV is a zero expectation and unit variance;  
 $\epsilon$ -SSE is a statistically  $\epsilon$ -stable evaluation;  
 $\mathbf{A}_q$  is a matrix representing the  $q$ -th class NDR;

$\hat{\mathbf{A}}_q(k)$  is a  $60 \times 80$  matrix as the  $q$ -th class representative TM6080I formed at the  $k$ -th stage;

$\mathbf{A}_{\text{PDTM6080I}}^{(k)}$  is a  $4800 \times 26$  matrix of PDTM6080I of all NDR formed at the  $k$ -th stage;

$\mathbf{A}_{\text{TM6080I}}^{(k)}$  is a  $4800 \times 26$  matrix of TM6080I of all NDR formed at the  $k$ -th stage;

$C$  is a number of pure objects' replicas;

$F$  is a number indicating at smoothness in training the perceptron;

$f(\alpha)$  the classifier output parameter influenced with the training parameter  $\alpha$ ;

$G$  is a set of indices;

$\mathbf{I}$  is a identity matrix;

$n$  is a number of the sampled points from the segment of the parameter values;

$N_{\text{classes}}$  is a number of classes;

$N_{\text{features}}$  is a number of the object features;

$p_{\text{EC}}(r)$  is an averaged CEP;

$Q_{\text{pass}}$  is a number of training samples to be passed through 2LP;

$r$  is a variable of PTSDR;

$\mathbf{T}_i$  is a target matrix for the  $i$ -th training subsample;

$\alpha$  is a training parameter influencing significantly on the classifier output parameter;

$\beta(k)$  the angle through which NDR is turned;

- $\sigma_{PD}$  is a SD of PD;  
 $\sigma_{TD}$  is a SD of TD;  
 $\langle \sigma_{PD}^k \rangle$  is a SD of PD at the  $k$ -th stage of PDTM6080I set formation;  
 $\langle \sigma_{TD}^k \rangle$  is a SD of TD at the  $k$ -th stage of TM6080I set formation;  
 $\sigma_{PD}^{(max)}$  is a maximum of SD in PD;  
 $\sigma_{TD}^{(max)}$  is a maximum of SD in TD;  
 $\Xi$  is a matrix of values of NV with ZEUV;  
 $\xi(k)$  is a value of NV with ZEUV, raffled at the  $k$ -th stage of TM6080I set formation.

## INTRODUCTION

The object classification is an important constituent in automation processes [1, 2]. Main hardness in classifying objects arises when the object has LIF, and anyone of objects cannot be ignored for simplification. Usually, a LIF object is a high resolution image after preprocessing [2, 3]. Such images are used in the manufacturing control, technical and medical diagnostics, safety watching systems, astrophysical explorations, and so forth [1, 3, 4]. A hard problem of LIF objects classification lies in COS. More particularly, COS is bettered by reducing the huge IMR which are consumed by the classification system, including the classifier [2, 4, 5]. The classifiers on the basis of HMLNN, CNN, cognitrons and neocognitrons perform perfectly over objects for their classification. At the same time, such classifiers take long periods for the performance and consume gigantic IMR [4, 6, 7]. And vice versa, perceptrons with one or more hidden layers consume far less IMR, while they are much faster. In this way, the perceptron with SHL is frequently used. This is 2LP, and it is the fastest for the most of classification problems, where the object features are distorted by a statistical law having near-normal distribution. However, 2LP cope poorly with classification problems, where object distortions are specific [5, 8]. For images, those specific distortions are skewness, nonlinear scaling, turn, severe rotation, shift, etc. Hence, there is a known duality in the problem of classifying LIF objects. Either we are to better COS and IMR consumption for perfectly performing HMLNN, CNN, cognitrons and neocognitrons, or to optimize some parameters of 2LP to get it trained suitably on specifically distorted LIF objects. And it is plain that the way with 2LP is preferable to bettering COS and IMR consumption with far beyond complicated neuronets.

The article goal is to develop the framework of the scenario type for 2LP classifier single training parameter optimization. At that, the classifier input is fed with LIF objects, making heuristic methods of optimization inapplicable. For achieving the goal, both theoretical and practical phases are to be delineated. Preceding the praxis, theoretical phase is tasked to give a justified succession of jobs whose accomplishment guarantees the classifier influential improvement on its output parameter. The tasking

will be replenished by verifying practically whether that output parameter is bettered over the specifically distorted objects. And the conclusive task is to infer from the development and investigation results.

## 1 PROBLEM STATEMENT

Let, without loss of generality, we have  $\alpha$  and  $f(\alpha)$ , which are both positive. Formally, the decided problem is to find

$$\alpha^* \in \arg \min_{\alpha > 0} f(\alpha) \quad (1)$$

and get 2LP, performing at  $f(\alpha^*)$ . In practicing, classifiers or classification processes mostly have characteristics desired for their minimization. And, naturally, the training parameter  $\alpha$  has its bounds  $\alpha_{\min}$  and  $\alpha_{\max}$ , that is  $\alpha \in [\alpha_{\min}; \alpha_{\max}]$  for  $\alpha_{\max} > \alpha_{\min} > 0$ . Henceforth, the problem (1) is rendered concrete:

$$\alpha^* \in \arg \min_{\alpha \in [\alpha_{\min}; \alpha_{\max}]} f(\alpha). \quad (2)$$

## 2 REVIEW OF THE LITERATURE

2LP in preference to far beyond complicated neural networks has a great many of its parameters. They are divided into the couple of associated parametric groups. The first one is assembled from size of SHL and transfer functions [5]. The second one is assembled from parameters of the training algorithm, methods of the algorithm implementation, and methods of forming training samples [8]. All these parameters and methods are formalized as training parameters [9, 10].

Typically, SHL size and transfer functions of 2LP are set up heuristically, being based on sizes of input and output layers and recent experience [5, 8, 9, 11]. The classifier training parameters optimization is only possible before the classifier is trained. Once 2LP has been trained, the classification process attains its characteristics. Stable estimations of these characteristics can be valued after huge batch testing.

The training process is defined with the set of training samples, feeding the input of 2LP [5, 8, 12]. Parameters of the set of training samples in the training process are specific for a classification problem [5]. To get the classification process improved, training samples are modified accordingly to types of object distortion [10, 13]. The modification has its own parameter or parameters. Sweeping them gives, speaking factually, diverse classifiers. And the best classifier could be chosen among those ones.

The best classifier choice nonetheless requires a framework of successive operations over the classifier training parameters to get them optimized. Before getting started, a parameter to be swept is to be appointed along with the classification process output parameter, which is influenced with the being swept parameter. Hence there is a want of developing the scenario of making the classifier optimization upon its parameter from a raw 2LP up to the classifier having one of its characteristics improved.

### 3 MATERIALS AND METHODS

For solving the problem (1), the function  $f(\alpha)$  evaluation requires defining the object type. The object type is actually defined with its format containing information about the number  $N_{\text{features}}$ . Dimensions do not matter because the object initial mathematical presentation (aggregate of its features) is reshaped into line array.

A convenient object type is MI of a medium format, having a few thousands features. This type and format allows watching and rendering MI distortions readily. The medium format ensures delayless investigation procedures and the classification results acquisition. And those results are going to be powerfully propagated on other objects of medium formats.

Within MATLAB, having powerful Neural Network Toolbox [14, 15] for programming and simulating neural networks, MI bitmap file is coded with ones (white color) and zeros (black color). Therefore the finite subset of GT is of MOZ, containing altogether  $2^{N_{\text{features}}}$  MI. Wholly, GT is infinite, wherein the pixel's value is not necessarily equal to 1 or 0, but it is a real number. Namely within the infinite GT, the specifically distorted objects must be recognized and classified.

After the object type is defined, number  $N_{\text{classes}} \in \mathbb{N} \setminus \{1\}$  is prescribed. And so there is defined GT containing all permissible objects of the defined type. Within the defined GT, NDR of classes are fixed.

Having the object type,  $N_{\text{classes}}$  NDR and GT, the number of neurons in SHL of 2LP is assigned. Rules for SHL neurons number assignation are drilled on experience [5, 11, 16]. Assigning transfer functions in the layers of 2LP is the last item in the configuration of 2LP.

Having had the configuration of 2LP, there is a MATLAB function for training to be assigned. Then the configured 2LP is initialized on MATLAB Neural Network Toolbox simply with the function «feedforwardnet» (the similar function «newff» if for MATLAB versions before R2010).

The function  $f(\alpha)$  evaluation starts with estimating the endpoints of the range  $[\alpha_{\min}; \alpha_{\max}]$ . Having estimated them, this range is to be sampled with a step. This gives a finite subset of points of the training parameter  $\alpha$ . 2LP is trained over this subset, and so the statistics of the trained 2LP is accumulated. By the way, the sampling step can be non-constant if needed, so the subset is

$$\{\alpha_h\}_{h=1}^n \subset [\alpha_{\min}; \alpha_{\max}] \text{ by } n \in \mathbb{N} \setminus \{1, 2\} \\ \text{at } \alpha_1 = \alpha_{\min} \text{ and } \alpha_n = \alpha_{\max}. \quad (3)$$

Really, after the sampling (3) the problem (2) is

$$\alpha^* \in \arg \min_{\alpha \in \{\alpha_h\}_{h=1}^n} f(\alpha) \quad (4)$$

with the function evaluation

$$f(\alpha) = \frac{1}{|G|} \sum_{g \in G} f^{(g)}(\alpha) \quad (5)$$

by the  $g$ -th polyline  $f^{(g)}(\alpha)$ , whose values  $\{f^{(g)}(\alpha)\}_{\alpha \in \{\alpha_h\}_{h=1}^n}$  are taken via testing the trained 2LP in each of  $n$  points  $\{\alpha_h\}_{h=1}^n$ , repeated for  $|G|$  times. Speaking strictly, the batch testings are sufficient for evaluating the function  $f(\alpha)$  adequately if for some  $\varepsilon > 0$

$$|\tilde{f}_1(\alpha) - \tilde{f}_2(\alpha)| < \varepsilon \quad \forall \alpha \in [\alpha_{\min}; \alpha_{\max}] \quad (6)$$

by

$$\tilde{f}_1(\alpha) = \frac{1}{|G_1|} \sum_{g \in G_1 \subset G} f^{(g)}(\alpha),$$

$$\tilde{f}_2(\alpha) = \frac{1}{|G_2|} \sum_{g \in G_2 \subset G} f^{(g)}(\alpha) \quad (7)$$

for the probability  $P(G_1 \cap G_2 \neq \emptyset)$  which is such that

$$\lim_{|G| \rightarrow \infty} P(G_1 \cap G_2 \neq \emptyset) = 0. \quad (8)$$

For wide ranges and insufficiently completed set of indices  $G$ , having its insufficiently great cardinal  $|G|$ , the condition (6) fails. However, the segment  $[\alpha_{\min}; \alpha_{\max}]$  may be narrowed to a subsegment by the following reason. If there is great likelihood of that

$$f(\alpha_1) \leq f(\alpha_2) \text{ for } \alpha_{\min} < \alpha_1 < \alpha_2 < \alpha_{\max} \quad (9)$$

then we can narrow the segment  $[\alpha_{\min}; \alpha_{\max}]$  to  $[\alpha_{\min}; \alpha_0]$  by  $\alpha_0 \in (\alpha_1; \alpha_2)$  or, for instance, to  $[\alpha_{\min}; (\alpha_1 + \alpha_2)/2]$ . Then we proceed in batch testings with re-estimation of the right endpoint  $\alpha_{\max} = \alpha_0$ . Statistical sufficiency is very important here. Thence is an unambiguous definition concerning sufficiency.

**Definition 1.** Function (5) is called  $\varepsilon$ -SSE by the values

$$\left\{ \left. f^{(g)}(\alpha) \right|_{\alpha \in \{\alpha_h\}_{h=1}^n} \right\}_{g \in G} \quad (10)$$

if for fixed  $\varepsilon > 0$  and  $c(\varepsilon) \in \mathbb{N}$  and any pair  $\{G_1, G_2\}$  of IRGSS  $\{G_1 \subset G, G_2 \subset G\}$  of the indices' set  $G$  with

$$\min \{|G_1|, |G_2|\} \geq c(\varepsilon) \quad (11)$$

the inequality

$$|\tilde{f}_1(\alpha) - \tilde{f}_2(\alpha)| < \varepsilon \quad \forall \alpha \in \{\alpha_h\}_{h=1}^n \quad (12)$$

holds by (7) and (8).

So then for approaching to the problem (4) solution, in accordance with Definition 1 we need  $\varepsilon$ -SSE (5) for sufficiently small  $\varepsilon$  and great  $c(\varepsilon)$ . Practically,  $\varepsilon$  can be assigned to

$$\varepsilon \leq \varepsilon_0 \cdot \left( \min_{\alpha \in \{\alpha_h\}_{h=1}^n} f(\alpha) \right) \quad (13)$$

by, say,  $\varepsilon_0 \in \{0.1, 0.05, 0.02, 0.01\}$  (traditional exemplification). The integer  $c(\varepsilon)$  is of the order of hundreds at least. And there predictably thrusts the concept of continuity in  $\varepsilon$ -SSE.

**Definition 2.1.** Function (5) is called continuous  $\varepsilon$ -SSE by the values (10) if  $\forall \varepsilon > 0 \exists \delta > 0$  such that the inequality (12) holds by (7) and

$$P(G_1 \cap G_2 \neq \emptyset) < \delta \quad (14)$$

for IRGSS  $\{G_1 \subset G, G_2 \subset G\}$  by (8).

An alternative definition to Definition 2.1 handles the relationship between cardinals of IRGSS and  $G$ . This one is handy standing off the probability (14) calculation.

**Definition 2.2.** Function (5) is called continuous  $\varepsilon$ -SSE by the values (10) if  $\forall \varepsilon > 0 \exists \delta \in (0; 1]$  such that the inequality (12) holds by (7) and

$$P\left(\frac{|G_1 \cap G_2|}{|G|} < \delta\right) = 1 \quad (15)$$

for IRGSS  $\{G_1 \subset G, G_2 \subset G\}$  by (8).

These two versions of continuity in  $\varepsilon$ -SSE are purely theoretical. However, they drive to solving the problem (4) statistically inasmuch as  $\varepsilon$ -SSE does not determine the minimum point  $\alpha^*$  itself. Note that probability (15) is apparently substituted with the expression  $\frac{|G_1 \cap G_2|}{|G|} < \delta$ .

**Definition 3.** The point  $\alpha^*$  is called solution of the problem (4) with  $\varepsilon$ -SSE (5) by the values (10) if the statement

$$\alpha^* \in \left\{ \arg \min_{\alpha \in \{\alpha_h\}_{h=1}^n} \tilde{f}_1(\alpha) \right\} \cap \left\{ \arg \min_{\alpha \in \{\alpha_h\}_{h=1}^n} \tilde{f}_2(\alpha) \right\} \neq \emptyset \quad (16)$$

holds at (7) by any pair of IRGSS  $\{G_1 \subset G, G_2 \subset G\}$  with the condition (11) for their cardinals.

Verification of that the training parameter  $\alpha$  has been bettered over the specifically distorted objects, having become  $\alpha^*$ , is carried out to get convinced of that

$$f(\alpha^*) < f(\alpha) \quad \forall \alpha \in \{\alpha_h\}_{h=1}^n \setminus \{\alpha^*\}. \quad (17)$$

With the set  $G$  for verification, the function evaluation (5) is suitable for (17).

Entirely, scenario for the single training parameter optimization of 2LP classifier is given in 13 items:

1. Defining the object type.
2. Prescribing the number of classes  $N_{\text{classes}} \in \mathbb{N} \setminus \{1\}$ .
3. Defining GT.
4. Fixing  $N_{\text{classes}}$  NDR of classes within the defined GT.
5. Assigning SHL neurons number.
6. Assigning transfer functions in the layers of 2LP.
7. Assigning a MATLAB function for training.
8. Appointing the parameter  $\alpha$  along with the classification process output parameter  $f(\alpha)$ .
9. Estimating the range  $[\alpha_{\min}; \alpha_{\max}]$  endpoints.
10. Sampling the range  $[\alpha_{\min}; \alpha_{\max}]$  into the finite subset (3).
11. Proceeding in batch testings of the trained 2LP in each of those points of the finite subset (3) in order to obtain statistically stable evaluation (5) for solving the problem (4).
12. If for fixed  $\varepsilon$  and  $c(\varepsilon)$  the  $(n-1)$ -segmented polyline (5) is  $\varepsilon$ -SSE then go to solving the problem (4) and return the result  $\{\alpha^*, f(\alpha^*)\}$  along with verification of that.
13. If the problem (4) has been not solved then go to narrowing the segment  $[\alpha_{\min}; \alpha_{\max}]$  to  $[\alpha_{\min}; \alpha_0]$  by  $\alpha_0 \in (\alpha_1; \alpha_2)$  if possible and recur to the item 10 with  $\alpha_{\max} = \alpha_0$ ; if the narrowing is impossible then recur to the item 11 with the increased  $|G|$ .

#### 4 EXPERIMENTS

Let's exemplify the developed scenario framework in the problem of turned objects classification. This is a widespread problem, when recognizing plane images or their contours the object rotation effect springs up. It is caused with that there cannot be ensured the symmetry just on the moment when the object is captured after it has been traced.

May the format of MI be  $60 \times 80$  and may an MI model be the enlarged English alphabet capital letter by  $N_{\text{classes}} = 26$ . Therefore the finite subset of GT is of  $60 \times 80$  MOZ, containing altogether  $2^{4800}$  MI by  $N_{\text{features}} = 4800$ . Wholly, GT is of matrices whose elements are real numbers. Thus there are 26 capital letters as NDR.

The input layer in 2LP has 4800 neurons, and the output layer has 26 neurons. The size of SHL in 2LP for the problem of classifying TM6080I can be assigned to 250 neurons. Transfer functions are log-sigmoid [5, 8, 10, 12]. This 4800-250-26-P is initialized with «feedforwardnet» («newff») by 1206776 weight and bias values. For training perceptrons

there is the backpropagation algorithm, having many methods of its implementation in MATLAB. One of them is the method of a MATLAB function «traingda» [16], where weight and bias values are updated according to gradient descent with adaptive learning rate [10, 12, 17, 18]. Every 4800-250-26-P shall be trained with the function «traingda», guaranteeing quick training process passage.

In the training process for classifying TM6080I, the training set is formed as addition of matrix of pixel-distorted images and matrix of turned images. Intensity (deepness) of PD is measured via SD  $\sigma_{PD}$  and intensity of TD is measured via SD  $\sigma_{TD}$ , determining the range of left and right angulation due to TD. The parameter here is PTSDR  $r = \sigma_{PD}/\sigma_{TD}$ , and the classification process output parameter  $f(\alpha)$  is CEP. Hence the problem (2) is

$$r^* \in \arg \min_{r \in [r_{\min}; r_{\max}]} p_{EC}(r) \quad (18)$$

by the function  $p_{EC}(r)$ , whose value at PTSDR  $r$  is an averaged CEP over TM6080I or, probably, over PDTM6080I.

PTSDR is constant while training. In the statement  $r = \sigma_{PD}/\sigma_{TD}$  both SD may vary, so the assignment  $r = \sigma_{PD}^{\langle \max \rangle} / \sigma_{TD}^{\langle \max \rangle}$  is formally more correct. Due to a model of TM6080I, an NDR is turned through the angle  $\beta(k) = (180/\pi) \cdot \sigma_{TD}^{\langle k \rangle} \xi(k)$  in degrees around the center point of MI,  $k = \overline{1, F}$  and  $F \in \mathbb{N}$ . SD in the angle  $\beta(k)$  is  $\sigma_{TD}^{\langle k \rangle} = \sigma_{TD}^{\langle \max \rangle} \cdot k / F \forall k = \overline{1, F}$ . From the statement for the angle  $\beta(k)$  by SD  $\sigma_{TD}^{\langle k \rangle}$  it follows that the upper value  $\sigma_{TD}^{\langle \max \rangle} = 0.2$  is enough. Beyond this value MI becomes turned so that its turning is impracticable for 2LP classifier, when rotation angle is about  $\pi/8$  and severer.

In estimating the range  $[r_{\min}; r_{\max}]$  endpoints, note that on  $\sigma_{PD}^{\langle \max \rangle} > 2$  PDTM6080I becomes over-distorted, and on  $\sigma_{PD}^{\langle \max \rangle} < 0.005$  PD is imperceptible. Then  $r_{\min} = 0.025$  and  $r_{\max} = 10$  enclose the range  $[0.025; 10]$  which is to be sampled.

In the training process the input of 4800-250-26-P is fed with the training set

$$\left\{ \mathbf{P}_i^{\langle PDTM6080I \rangle} \right\}_{i=1}^{C+F} = \left\{ \left\{ \mathbf{A} \right\}_{l=1}^C, \left\{ \mathbf{A}_{PDTM6080I}^{\langle k \rangle} \right\}_{k=1}^F \right\} \quad (19)$$

of  $C \in \mathbb{N}$  replicas of all 26 classes NDR and  $F$  matrices of PDTM6080I by targets  $\left\{ \mathbf{T}_i \right\}_{i=1}^{C+F} = \left\{ \mathbf{I} \right\}_{i=1}^{C+F}$ , where  $4800 \times 26$  matrix  $\mathbf{A}$  is formed by concatenating horizontally

4800-LCRM  $\left\{ \mathbf{A}_q \right\}_{q=1}^{26}$  by the  $q$ -th class NDR as  $60 \times 80$  MOZ  $\mathbf{A}_q = \left( a_{uv}^{\langle q \rangle} \right)_{60 \times 80}$ . At the  $k$ -th stage of PDTM6080I set formation,  $4800 \times 26$  matrix

$$\mathbf{A}_{PDTM6080I}^{\langle k \rangle} = \mathbf{A}_{TM6080I}^{\langle k \rangle} + \sigma_{PD}^{\langle k \rangle} \cdot \mathbf{\Xi} \quad (20)$$

by  $\sigma_{PD}^{\langle k \rangle} = \sigma_{PD}^{\langle \max \rangle} \cdot k / F \forall k = \overline{1, F}$ . Concurrently for (20),  $4800 \times 26$  matrix  $\mathbf{A}_{TM6080I}^{\langle k \rangle} = [\bar{a}_{jq}(k)]_{4800 \times 26}$  is formed by concatenating horizontally 4800-LCRM  $\left\{ \hat{\mathbf{A}}_q(k) = \left[ \hat{a}_{uv}^{\langle q \rangle}(k) \right]_{60 \times 80} \right\}_{q=1}^{26}$ . The  $q$ -th class NDR  $\mathbf{A}_q$  becomes TM6080I

$$\mathbf{A}_q(k) = 1 - \mathbf{p}(1 - \mathbf{A}_q, \beta(k), M, S) \quad (21)$$

with MATLAB function «imrotate» in the map  $\mathbf{p}$ , turning the input MI negative  $1 - \mathbf{A}_q$  through the angle  $\beta(k)$ , by the interpolation method handle  $M$  and the handle  $S$  for specifying the size of the returned negative MI. MI is turned in counterclockwise direction if  $\beta(k) > 0$ , and for  $\beta(k) < 0$  MI is turned clockwise; for  $\beta(k) = 0$  MI remains NDR. After all 26 classes NDR have been turned and become TM6080I, the training set (19), formed with (20) by  $\sigma_{PD}^{\langle k \rangle}$  and (21), feeds the input of 4800-250-26-P, passing through 4800-250-26-P with targets  $\left\{ \mathbf{T}_i \right\}_{i=1}^{C+F} = \left\{ \mathbf{I} \right\}_{i=1}^{C+F}$  for  $Q_{\text{pass}} \in \mathbb{N}$  times.

## 5 RESULTS

In the initial probe, PDTM6080I-trained 4800-250-26-P under set  $\{C, F, Q_{\text{pass}}\} = \{2, 8, 10\}$  for the training set (19) is tested with TM6080I at  $\sigma_{TD} \in [0; 0.2]$ . As it has been expected (figure 1), the first approximation over the subset

$$\{0.025, 0.05, \{0, 1h\}_{h=1}^{20}, \{2 + h\}_{h=1}^8\} \subset [0.025; 10] \quad (22)$$

would not have solved the problem (18)

$$r^* \in \arg \min_{r \in [0, 0.025; 10]} p_{EC}(r) \quad (23)$$

because the rough 29-segmented polyline of the function  $p_{EC}(r)$  is not  $\varepsilon$ -SSE. Notwithstanding that PDTM6080I-trained 4800-250-26-P were tested on only TM6080I and  $|G| = 40$  appeared to have been underrated, figure 1 prompts that the range  $[0.025; 10]$  could be narrowed to the subsegment  $[0.025; 1]$ , where  $r^* \in [0.025; 1]$  apparently.

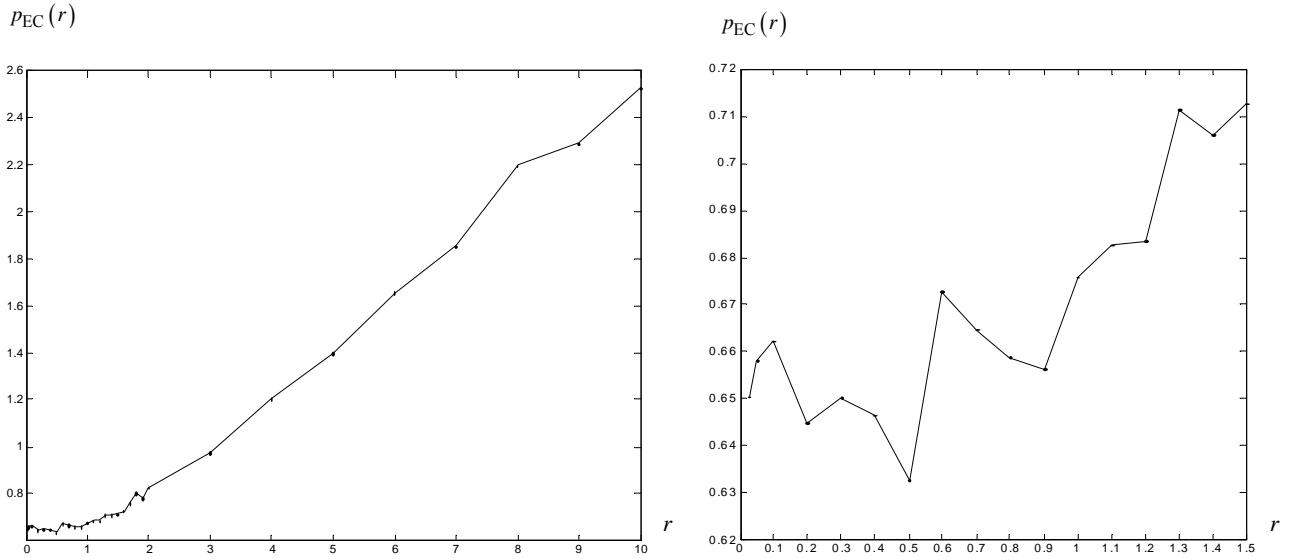


Figure 1 – A rough 29-segmented polyline of the function  $p_{EC}(r)$  over the subset (22), derived from 40 series of 400 batch testings of PDTM6080I-trained 4800-250-26-P (tested on only TM6080I for faster preliminary results)

The second (40 series of 800 batch testings) and third (100 series of 800 batch testings) approximations over the subsets  $\{0.025h\}_{h=1}^{40} \subset [0.025; 1]$  and

$\{0.025h\}_{h=1}^{24} \subset [0.025; 0.6]$  correspondingly do not give  $\varepsilon$ -SSE for the said  $|G| = 40$  and  $|G| = 100$ . Nevertheless, this 0.025-sampling lets feel and see that  $r^* \in [0.025; 0.6]$ , needing though to increase the cardinality of  $G$ . Thanking to that all those 140 polylines of the function  $p_{EC}(r)$  were obtained on 0.025-sampling, the cardinality here could be increased just by superposing those 40 and 100 polylines.

However, neither superposition nor another 200 series of 200 batch testings (appeared to be sufficient to understand classification properties of an PDTM6080I-trained 4800-250-26-P) give  $\varepsilon$ -SSE of the function  $p_{EC}(r)$ . Here, although the point  $r = 0.475$  could have been ventured to accept it as the problem (23) solution, it as well could be refuted.

Herewith, for obtaining  $\varepsilon$ -SSE of the function  $p_{EC}(r)$  at some reasonable  $\varepsilon$  there is  $c(\varepsilon) > 340$ . And the shape of the 23-segmented polyline of the function  $p_{EC}(r)$  becomes more stochastic when extracting a lesser cardinality subset from those 340 polylines. However, such deep stochasticity may be provoked with the current conformation of training and testing. Likely, number of passes  $Q_{pass} = 10$  is too low for expecting stable estimation. Also PTSDR  $r = 5$  for testing is too high, randomizing the output results of estimations additionally.

Nevertheless, there are evidences of that the problem (23) solution is in the segment  $[0.025; 0.5]$  of PTSDR. The further probe with taking  $Q_{pass} = 25$  and testing PDTM6080I-trained 4800-250-26-P with PDTM6080I on  $\sigma_{TD} \in [0; 0.2]$  and  $\sigma_{PD} \in [0; 0.2]$  or  $\sigma_{PD} \in [0; 0.4]$  returns

much the same stochasticity. A corollary is that the upper value of PTSDR for testing does not influence much on stochasticity of  $p_{EC}(r)$ .

Final evaluation on  $|G| = 370$  (figure 2) shows that the problem (23) solution is  $r^* = 0.225$ , having been taken off Definition 3, though. In this situation, the most appropriate conclusion would be that  $r^* \in (0.075; 0.3)$ , where the cavity of the polyline in figure 2 by  $r \in (0.075; 0.3)$  can be seen. And the local minimum at  $r = 0.075$  gives almost the same value of CEP therewith.

Henceforward, the best PDTM6080I-trained 4800-250-26-P by

$$\sigma_{PD}^{\langle \max \rangle} = r^* \cdot \sigma_{TD}^{\langle \max \rangle} = 0.225 \cdot 0.2 = 0.045 \quad (24)$$

under the set  $\{C, F, Q_{pass}\} = \{2, 8, 25\}$  classifies TM6080I at CEP  $p_{EC}(0.225) < 0.28$  and classifies PDTM6080I at CEP  $p_{EC}(0.225) < 0.48$ . Moreover, the best PDTM6080I-trained 4800-250-26-P by (24) under the set  $\{C, F, Q_{pass}\} = \{2, 8, 75\}$  classifies TM6080I at CEP  $p_{EC}(0.225) < 0.19$  and classifies PDTM6080I at CEP  $p_{EC}(0.225) < 0.29$ , and the averaged CEP is  $p_{EC}(0.225) < 0.24$ . As the example visualization, figure 3 shows TM6080I which became recognizable for the trained 2LP classifier after its single training parameter optimization with (24).

The gain of the executed optimization depends on primary heuristics in adjusting PTSDR. Thus, if it were  $r = 1$ , what could have been mostly expected in adjusting off optimization, then the absolute gain is greater than 0.1 %. In relative calculus, it is about a third. For cases with  $r > 1$  the gain is even higher.

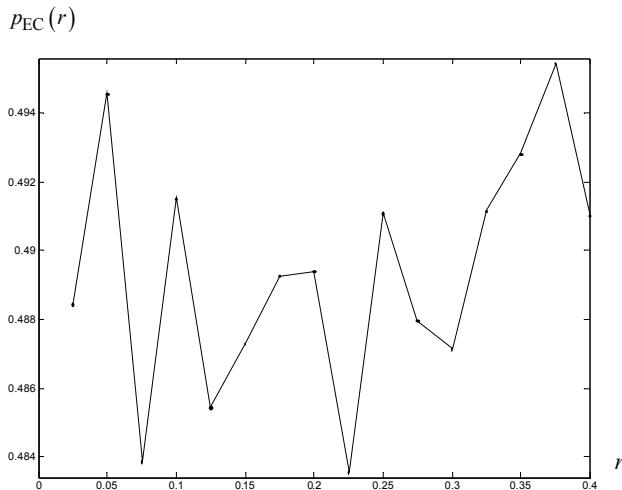


Figure 2 – 0,01-SSE of the function  $p_{EC}(r)$  on  $\sigma_{TD} \in [0; 0,2]$  and  $\sigma_{PD} \in [0; 0,4]$ , derived from 370 series of 400 batch testings of PDTM6080I-trained 4800-250-26-P

Eventually, the 13-itemed scenario framework for classifier single training parameter optimization allowed to have trained 4800-250-26-P on PDTM6080I optimally in order to perform classification TM6080I at the lowest CEP. And PDTM6080I are classified well also. The solution (23) can be applied for training 2LP classifiers in other problems of turned LIF objects classification, where number of features is about 4800 and number of classes is about 26.

## 6 DISCUSSION

The 13-itemed scenario framework suggests a straightforward succession of operations to get 2LP classifier single training parameter optimized. Furthermore, this framework with proper modifications (items 5, 6, and 7) is applicable for optimizing training parameters of HMLNN and CNN. Besides, the classifier input is not necessary to be fed with LIF objects. The gain is either the lowest CEP or getting a classifier output parameter bettered.

The operations from item 1 through item 8 do not require statistics. Otherwise, for narrowing the segment  $[\alpha_{\min}; \alpha_{\max}]$  we need great likelihood of that (9) holds. The greatness is ascertained via series of the trained classifier's batch testings, where the series number is of the order of several tens (hundreds) at least.

The hardest operation is the item 12. For ascertaining statistical  $\varepsilon$ -stability of the polyline (5), the requirement (13) is purely practical. It is clear that for implementing the conditions of Definition 1, the integer  $c(\varepsilon)$  is specified by experience for a few tens of IRGSS  $\{G_1, G_2\}$ . These subsets can be used also in Definition 2.2, before setting up  $G$  and  $\varepsilon$ . The same concerns Definition 3 and its key statement (16).

But sometimes specificity of Definition 3 nonetheless makes it irresolute for a single point solution in (16). Namely that was a cause for accepting the interval solution  $r^* \in (0.075; 0.3)$  in the exemplification with 4800-250-26-P for TM6080I classification.

## CONCLUSIONS

In the problem of 26 TM6080I classification by 2LP, the CEP minimization relative gain is about a third. A best-trained classifier produces CEP not greater than 0.19 % over TM6080I, and not greater than 0.29 % over PDTM6080I. The stated framework hereby provides the CEP minimization by optimizing the 2LP classifier single PDTM6080I-training parameter.

An aggregate of other training parameters (two and more) of the classifier can be considered for optimization as well. Any rigorous framework for such multivariable optimization shall include operations from item 1 through item 4. The operations from item 8 through item 13 shall regard the corresponding multiplicity of variables. Particularly for 2LP, SHL size and integers  $C, F, Q_{\text{pass}}$  could be optimized for CEP minimization. Certainly, optimizing each of them singly is worse for CEP minimization than optimizing an aggregate of them. In defiance of statistical hugeness, multivariable optimization must help in shortening the training process duration on average. And the corresponding framework should have an effective operation item for plotting and extremizing the piecewise linear hypersurface (function of two or more variables) with values of the classifier output parameter.

## ACKNOWLEDGEMENTS

The work is technically supported by the Center of Parallel Computations of Khmelnitskiy National University (Khmelnitskiy, Ukraine).

## REFERENCES

1. Axinte D. A. Approach into the use of probabilistic neural networks for automated classification of tool malfunctions in broaching / D. A. Axinte // International Journal of Machine Tools and Manufacture. – 2006. – Volume 46, Issue 12–13. – P. 1445–1448. DOI: 10.1016/j.ijmachtools.2005.09.017
2. Fukushima K. Increasing robustness against background noise: Visual pattern recognition by a neocognitron / K. Fukushima // Neural Networks. – 2011. – Volume 24, Issue 7. – P. 767–778. DOI: 10.1016/j.neunet.2011.03.017
3. On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images / J. Plaza, A. Plaza, R. Perez, P. Martinez] // Pattern Recognition. – 2009. – Volume 42, Issue 11. – P. 3032–3045. DOI: 10.1016/j.patcog.2009.04.008
4. Siniscalchi S. M. Exploiting deep neural networks for detection-based speech recognition / S. M. Siniscalchi, D. Yu,



Figure 3 – TM6080I by rotation angle not less than  $\pi/12$  of letters, which became recognizable for the best PDTM6080I-trained

4800-250-26-P by (24) under the set  $\{C, F, Q_{\text{pass}}\} = \{2, 8, 75\}$

- L. Deng, C.-H. Lee // Neurocomputing. – 2013. – Volume 106. – P. 148 – 157. DOI: 10.1016/j.neucom.2012.11.008
5. Arulampalam G. A generalized feedforward neural network architecture for classification and regression / G. Arulampalam, A. Bouzerdoum // Neural Networks. – 2003. – Volume 16, Issue 5–6. – P. 561 – 568. DOI: 10.1016/S0893-6080(03)00116-3
  6. Multi-column deep neural network for traffic sign classification / [D. Cireşan, U. Meier, J. Masci, J. Schmidhuber] // Neural Networks. – 2012. – Volume 32. – P. 333–338. DOI: 10.1016/j.neunet.2012.02.023
  7. Fukushima K. Artificial vision by multi-layered neural networks: Neocognitron and its advances / K. Fukushima // Neural Networks. – 2013. – Volume 37. – P. 103–119. DOI: 10.1016/j.neunet.2012.09.016
  8. An efficient hidden layer training method for the multilayer perceptron / [C. Yu, M. T. Manry, J. Li, P. L. Narasimha] // Neurocomputing. – 2006. – Volume 70, Issue 1–3. – P. 525–535. DOI: 10.1016/j.neucom.2005.11.008
  9. Comparing evolutionary hybrid systems for design and optimization of multilayer perceptron structure along training parameters / [P. A. Castillo, J. J. Merelo, M. G. Arenas, G. Romero] // Information Sciences. – 2007. – Volume 177, Issue 14. – P. 2884–2905. DOI: 10.1016/j.ins.2007.02.021
  10. Hoi K. I. Improvement of the multilayer perceptron for air quality modelling through an adaptive learning scheme / K. I. Hoi, K. V. Yuen, K. M. Mok // Computers & Geosciences. – 2013. – Volume 59. – P. 148–155. DOI: 10.1016/j.cageo.2013.06.002
  11. Yuan H. C. A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy / H. C. Yuan, F. L. Xiong, X. Y. Huai // Computers and Electronics in Agriculture. – 2003. – Volume 40, Issue 1–3. – P. 57–64. DOI: 10.1016/S0168-1699(03)00011-5
  12. Kathirvalavakumar T. Neighborhood based modified backpropagation algorithm using adaptive learning parameters for training feedforward neural networks / T. Kathirvalavakumar, S. Jeyaseeli Subavathi // Neurocomputing. – 2009. – Volume 72, Issue 16–18. – P. 3915–3921. DOI: 10.1016/j.neucom.2009.04.010
  13. McLoone S. Improving neural network training solutions using regularisation / S. McLoone, G. Irwin // Neurocomputing. – 2001. – Volume 37, Issue 1–4. – P. 71–90. DOI: 10.1016/S0925-2312(00)00314-3
  14. Kuzmanovski I. Counter-propagation neural networks in Matlab / I. Kuzmanovski, M. Novič // Chemometrics and Intelligent Laboratory Systems. – 2008. – Volume 90, Issue 1. – P. 84–91. DOI: 10.1016/j.chemolab.2007.07.003
  15. Ballabio D. A MATLAB toolbox for Self Organizing Maps and supervised neural network learning strategies / D. Ballabio, M. Vasighi // Chemometrics and Intelligent Laboratory Systems. – 2012. – Volume 118. – P. 24–32. DOI: 10.1016/j.chemolab.2012.07.005
  16. Romanuke V. V. Setting the hidden layer neuron number in feedforward neural network for an image recognition problem under Gaussian noise of distortion / V. V. Romanuke // Computer and Information Science. – 2013. – Volume 6, No. 2. – P. 38–54. DOI: 10.5539/cis.v6n2p38
  17. On-line neural training algorithm with sliding mode control and adaptive learning rate / [A. Nied, S. I. Jr. Seleme, G. G. Parma, B. R. Menezes] // Neurocomputing. – 2007. – Volume 70, Issue 16–18. – P. 2687–2691. DOI: 10.1016/j.neucom.2006.07.019
  18. Yoo S. J. Indirect adaptive control of nonlinear dynamic systems using self recurrent wavelet neural networks via adaptive learning rates / S. J. Yoo, J. B. Park, Y. H. Choi // Information Sciences. – 2007. – Volume 177, Issue 15. – P. 3074–3098. DOI: 10.1016/j.ins.2007.02.009

Article was submitted 23.05.2014.  
After revision 08.10.2014.

Романюк В. В.

Д-р техн. наук, профессор кафедры прикладной математики и социальной информатики, Хмельницкий национальный университет, Украина

### **СТРУКТУРА ДЛЯ ОПТИМИЗАЦИИ ОТДЕЛЬНОГО ОБУЧАЮЩЕГО ПАРАМЕТРА КЛАССИФИКАТОРА НА ПРИМЕРЕ ОБУЧЕНИЯ ДВУХСЛОЙНОГО ПЕРСЕПТРОНА В ЗАДАЧЕ КЛАССИФИКАЦИИ ПОВЕРНУТЫХ ИЗОБРАЖЕНИЙ ФОРМАТА 60-НА-80**

Разрабатывается структура 13-этапного плана для оптимизации отдельного обучающего параметра классификатора. Формально задача состоит в нахождении глобального экстремума (преимущественно минимума) функции, являющейся некоторым выходным параметром классификатора в зависимости от его отдельного параметра обучения. Связывая теорию плана с практикой, типом классификатора избрано двухслойный персептрон. Его входными объектами являются монохромные изображения среднего формата с несколькими тысячами независимых признаков. По представленной структуре программной среды избрано MATLAB, обладающей мощным инструментом Neural Network Toolbox. Учитывая стохастичность минимизируемой функции, определяется статистическая  $\varepsilon$ -стабильность ее оценки по конечному множеству данных. Эти данные достаются по пакетному тестированию обученного классификатора. В качестве примера использования представленной структуры плана оптимизируется соотношение среднеквадратичных отклонений пиксельных искажений и поворотов для обучения двухслойного персептрона с целью классифицировать монохромные изображения формата 60-на-80 увеличенных 26 заглавных букв английского алфавита. Целью является нахождение определенного соотношения среднеквадратичных отклонений пиксельных искажений и поворотов в процессе обучения для того, чтобы обеспечить минимум процента ошибок классификации. Относительный выигрыш такой оптимизации составляет около одной трети. Разработанная структура может быть применена также и для оптимизации классификатора с многими переменными, касательно чего отмечается, какие именно действия плана должны учитывать множественность переменных.

**Ключевые слова:** оптимизация параметра обучения классификатора, статистическое оценивание, план оптимизации, двухслойный персептрон, процент ошибок классификации, классификация объектов с поворотами, монохромное изображение, соотношение среднеквадратичных отклонений пиксельных искажений и поворотов, обучающее множество.

Романюк В. В.

Д-р техн. наук, професор кафедри прикладної математики та соціальної інформатики, Хмельницький національний університет, Україна

## СТРУКТУРА ДЛЯ ОПТИМІЗАЦІЇ ОКРЕМОГО НАВЧАЛЬНОГО ПАРАМЕТРА КЛАСИФІКАТОРА НА ПРИКЛАДІ НАВЧАННЯ ДВОШАРОВОГО ПЕРСЕПТРОНУ В ЗАДАЧІ КЛАСИФІКАЦІЇ ПОВЕРНУТИХ ЗОБРАЖЕНЬ ФОРМАТУ 60-НА-80

Розробляється структура 13-етапного плану для оптимізації окремого навчального параметра класифікатора. Формально задача полягає у знаходженні глобального екстремуму (переважно мінімуму) функції, котра є певним вихідним параметром класифікатора залежно від його окремого параметра навчання. Пов'язуючи теорію плану з практикою, типом класифікатора обрано двошаровий персепtron. Його вхідними об'єктами є монохромні зображення середнього формату з декількома тисячами незалежних ознак. За поданою структурою програмним середовищем обрано MATLAB, котре має потужний інструмент Neural Network Toolbox. Зважаючи на стохастичність функції, що мінімізується, означається статистична  $\varepsilon$ -стабільність її оцінки за скінченою множиною даних. Ці дані дістаються за пакетним тестуванням навченого класифікатора. Для прикладу використання поданої структури плану оптимізується співвідношення середньоквадратичних відхилень піксельних споторен і поворотів для навчання двошарового персептрону з метою класифікувати монохромні зображення формату 60-на-80 збільшених 26 великих літер англійського алфавіту. Ціллю є знаходження певного співвідношення середньоквадратичних відхилень піксельних споторен і поворотів у процесі навчання для того, щоб забезпечити мінімум відсотка помилок класифікації. Відносний виграні такої оптимізації складає близько третини. Розроблена структура може бути застосована також і для оптимізації класифікатора з багатьма змінними, у відношенні чого зазначається, які саме дії плану мають врахувати множинність змінних.

**Ключові слова:** оптимізація параметра навчання класифікатора, статистичне оцінювання, план оптимізації, двошаровий персепtron, відсоток помилок класифікації, класифікація об'єктів з поворотами, монохромне зображення, співвідношення середньоквадратичних відхилень піксельних споторен і поворотів, навчальна множина.

## REFERENCES

1. Axinte D. A. Approach into the use of probabilistic neural networks for automated classification of tool malfunctions in broaching, *International Journal of Machine Tools and Manufacture*, 2006, Volume 46, Issue 12–13, pp. 1445–1448. DOI: 10.1016/j.ijmachtools.2005.09.017
2. Fukushima K. Increasing robustness against background noise: Visual pattern recognition by a neocognitron, *Neural Networks*, 2011, Volume 24, Issue 7, pp. 767–778. DOI: 10.1016/j.neunet.2011.03.017
3. Plaza J., Plaza A., Perez R., Martinez P. On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images, *Pattern Recognition*, 2009, Volume 42, Issue 11, pp. 3032–3045. DOI: 10.1016/j.patcog.2009.04.008
4. Siniscalchi S. M., Yu D., Deng L., Lee C.-H. Exploiting deep neural networks for detection-based speech recognition, *Neurocomputing*, 2013, Volume 106, pp. 148–157. DOI: 10.1016/j.neucom.2012.11.008
5. Arulampalam G., Bouzerdoum A. A generalized feedforward neural network architecture for classification and regression, *Neural Networks*, 2003, Volume 16, Issue 5–6, pp. 561–568. DOI: 10.1016/S0893-6080(03)00116-3
6. Cireşan D., Meier U., Masci J., Schmidhuber J. Multi-column deep neural network for traffic sign classification, *Neural Networks*, 2012, Volume 32, pp. 333–338. DOI: 10.1016/j.neunet.2012.02.023
7. Fukushima K. Artificial vision by multi-layered neural networks: Neocognitron and its advances, *Neural Networks*, 2013, Volume 37, pp. 103–119. DOI: 10.1016/j.neunet.2012.09.016
8. Yu C., Manry M. T., Li J., Narasimha P. L. An efficient hidden layer training method for the multilayer perceptron, *Neurocomputing*, 2006, Volume 70, Issue 1–3, pp. 525–535. DOI: 10.1016/j.neucom.2005.11.008
9. Castillo P. A., Merelo J. J., Arenas M. G., Romero G. Comparing evolutionary hybrid systems for design and optimization of multilayer perceptron structure along training parameters, *Information Sciences*, 2007, Volume 177, Issue 14, pp. 2884–2905. DOI: 10.1016/j.ins.2007.02.021
10. Hoi K. I., Yuen K. V., Mok K. M. Improvement of the multilayer perceptron for air quality modelling through an adaptive learning scheme, *Computers & Geosciences*, 2013, Volume 59, pp. 148–155. DOI: 10.1016/j.cageo.2013.06.002
11. Yuan H. C., Xiong F. L., Huai X. Y. A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy, *Computers and Electronics in Agriculture*, 2003, Volume 40, Issue 1–3, pp. 57–64. DOI: 10.1016/S0168-1699(03)00011-5
12. Kathirvalavakumar T., Jeyaseeli Subavathi S. Neighborhood based modified backpropagation algorithm using adaptive learning parameters for training feedforward neural networks, *Neurocomputing*, 2009, Volume 72, Issue 16–18, pp. 3915–3921. DOI: 10.1016/j.neucom.2009.04.010
13. McLoone S., Irwin G. Improving neural network training solutions using regularisation, *Neurocomputing*, 2001, Volume 37, Issue 1–4, pp. 71–90. DOI: 10.1016/S0925-2312(00)00314-3
14. Kuzmanovski I., Noviu M. Counter-propagation neural networks in Matlab, *Chemometrics and Intelligent Laboratory Systems*, 2008, Volume 90, Issue 1, pp. 84–91. DOI: 10.1016/j.chemolab.2007.07.003
15. Ballabio D., Vasighi M. A MATLAB toolbox for Self Organizing Maps and supervised neural network learning strategies, *Chemometrics and Intelligent Laboratory Systems*, 2012, Volume 118, pp. 24–32. DOI: 10.1016/j.chemolab.2012.07.005
16. Romanuke V. V. Setting the hidden layer neuron number in feedforward neural network for an image recognition problem under Gaussian noise of distortion, *Computer and Information Science*, 2013, Volume 6, No. 2, pp. 38–54. DOI: 10.5539/cis.v6n2p38
17. Nied A., Seleme S. I. Jr., Parma G. G., Menezes B. R. On-line neural training algorithm with sliding mode control and adaptive learning rate, *Neurocomputing*, 2007, Volume 70, Issue 16–18, pp. 2687–2691. DOI: 10.1016/j.neucom.2006.07.019
18. Yoo S. J., Park J. B., Choi Y. H. Indirect adaptive control of nonlinear dynamic systems using self recurrent wavelet neural networks via adaptive learning rates, *Information Sciences*, 2007, Volume 177, Issue 15, pp. 3074–3098. DOI: 10.1016/j.ins.2007.02.009

УДК 516.6

Богучарский С. И.<sup>1</sup>, Машталир С. В.<sup>2</sup><sup>1</sup>Аспирант, Харьковский национальный университет радиоэлектроники, Украина  
<sup>2</sup>Канд. техн. наук, доцент, Харьковский национальный университет радиоэлектроники, Украина

## АНАЛИЗ ТЕКСТУР В ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ВЕКТОРНОГО КВАНТОВАНИЯ

В работе предложен подход к анализу последовательностей изображений (видеоданных). Для этих целей разработаны матричные аналоги существующих нейросетевых подходов, что позволяет учитывать пространственные связи мультимедийной информации и сократить время необходимое на обработку информации за счет введения новой матричной процедуры обучения нейронной сети. В качестве основного признака, по которому производится разбиение исходных данных, были выбраны текстуры, что в свою очередь позволяет повысить точность кластеризации.

**Ключевые слова:** изображение, текстура, кластеризация, матричные алгоритмы, нейронные сети.

### НОМЕНКЛАТУРА

$M \times N$  – размерность изображения;  
 $m \times n$  – размер анализируемого блока изображения;  
 $x(k)$  – вектор входного сигнала;  
 $w_q^L$  – центры классов при разбиении;  
 $\text{Sp}(\cdot)$  – след матрицы;  
 $\eta(k)$  – параметр шага обучения;  
 $D(x(kn), w_q^L(kn))$  – расстояние между входными данными и центрами кластеров;  
 $\mu((x(k+1), w_q^L(k+1))$  – оценка уровня принадлежности данных кластеру;  
 $A$  – матрица коэффициентов размерности  $(m \times m)$ ;  
 $C$  – матрица коэффициентов размерности  $(n \times n)$ ;  
 $\xi(k) = \{\xi_{i_1 i_2}(k)\}$  – матричный белый шум с ограниченной дисперсией;  
 $h$  – порядок авторегрессии;  
 $E_A^{k+1}, E_C^{k+1}$  – критерии оптимизации;

$B(k)$  – блочная матрица с информацией о матричной-пространственно-временной авторегрессионной модели.

### ВВЕДЕНИЕ

Стремительное развитие информационных технологий естественным образом приводит к появлению больших объемов мультимедийной информации. При этом качество получаемой информации все улучшается в связи с усовершенствованием средств ее получения. В свою очередь появляется все больше задач, связанных с анализом мультимедиа информации. И если ранее исследователь мог ограничиваться в своих задачах одним или несколькими статическими изображениями, то в настоящее время для получения более эффективных алгоритмов анализа используются либо последовательности, связанных каким-либо образом, изображений, например, полученные через равные интервалы времени, либо

вообще видеоданные, которые всегда можно представить в виде последовательности кадров, а соответственно – последовательности изображений.

При этом одним из вполне естественных подходов по анализу видеопоследовательностей представляется использование математического аппарата анализа временных рядов. Они вполне могут применяться для задачи разбиения видеоданных на однородные по содержимому сегменты. С другой стороны эту же задачу можно интерпретировать как многомерную кластеризацию (сегментацию) множества изображений по некоторым признакам. Наряду с информацией о цвете также можно использовать текстурные признаки для проведения сегментации. Анализу текстур последовательностей изображений и посвящена данная работа.

### 1 ПОСТАНОВКА ЗАДАЧИ

Пусть обрабатываемое изображение задано в виде  $M \times N$  – числовой матрицы, которая должна быть разбита в процессе анализа на однородные в смысле текстур классы (кластеры, сегменты). Исходная матрица разбивается на блоки размера  $(m \leq M) \times (n \leq N)$ , каждый из которых в рамках традиционного подхода [12] преобразуется в  $(mn \times 1)$  – векторы, общее число которых равняется  $MN(mn)^{-1}$ . Именно этот набор и подается на входы классифицирующей нейронной сети.

### 2 АНАЛИЗ ЛИТЕРАТУРЫ

Анализ текстур – классификация и сегментация является важным направлением в общей теории распознавания образов и обработки изображений (Video Mining), а для его реализации к настоящему времени сложился целый ряд подходов, имеющих в своей основе различный математический аппарат. Здесь, прежде всего, следует отметить традиционный статистический подход [1, 2] методы стохастической оптимизации [3], геометрические методы [4], марковские модели [5–8], алгоритмы вейвлет-обработки сигналов [10] и, наконец,

достаточно широкий класс искусственных нейронных сетей [9–12]. При этом авторы работы [12] подчеркивают, что нейросетевой подход может в значительной мере объединить в себе все вышеупомянутые, благодаря универсальным аппроксимирующими и фильтрующими свойствам, способности к настройке своих характеристик непосредственно в процессе обработки, восстановлению нелинейных разделяющих гиперповерхностей сложной формы. Возможности нейросетевого подхода были продемонстрированы в процессе решения набора тестовых задач классификации текстур с помощью трехслойного персептрона с 75-ю нейронами в первом скрытом слое, числом нейронов в выходном слое равным количеству возможных классов и обучаемого с помощью RPROP – алгоритма, являющегося по сути модификацией стандартного алгоритма обратного распространения ошибок ( $\delta$ -правила обучения с учителем) [11]. При этом входной сигнал с окна, образованного 25 пикселями предварительно векторизуется с учетом того, что каждому пикселу ставится в соответствие три цвета. Большое количество настраиваемых синаптических весов многослойного персептрона (только в первом слое их число достигает  $75^2$ ), а также низкая скорость обучения на основе  $\delta$ -правила заставляют искать иные, альтернативные многослойным архитектуры нейронных сетей, содержащие меньшее число настраиваемых параметров. Кроме того, при анализе изображений удобнее обрабатывать входную информацию не в форме вектора, а непосредственно подавать ее на обработку в виде окна-матрицы, т.е. перейти от обработки многомерных сигналов к анализу двумерных полей [13–15].

### 3 МАТЕРИАЛЫ И МЕТОДЫ

**Классификация текстур на основе матричного векторного квантования.** Нами предлагается вместо  $(mn \times 1)$  – векторов-образов представлять входные сигналы в форме  $(m \times n)$  – матриц  $x(k) = \{x_{i_1, i_2}(k)\}$ ,  $i_1 = 1, 2, \dots, m; i_2 = 1, 2, \dots, n; k = 1, 2, \dots, MN(mn)^{-1}$  в случае  $m < M, n < N$  или текущее дискретное время или номер изображения в базе данных при  $m = M, n = N$ . Кроме того, вместо традиционного многослойного персептрона мы предлагаем использовать нейронную сеть обучаемого векторного квантования (LVQ) [16] в матричной ее модификации [17]. Нейронные сети обучаемого векторного квантования, введенные Т. Кохоненом, основаны на идеях конкурентного обучения и имеют однослойную архитектуру, полностью совпадающую с архитектурой популярной самоорганизующейся карты. В основе этих сетей лежит техника векторного квантования, состоящая в компактном представлении больших массивов информации, заданных в виде набора векторов, а в нашем случае набора  $(m \times n)$  – матриц  $x(k)$ , в форме ограниченного набора центров классов  $w_q^L, q = 1, 2, \dots$ , достаточно хорошо в смысле прямой метри-

ки разбивающих пространство входов на однородные области (текстуры), при этом  $w_q^L$  также является  $(m \times n)$  – матричный, а  $p$  – это возможное число классов в обучающей выборке.

Процесс обучения с учителем в матричном VLQ происходит следующим образом. Для каждого предварительно пронормированного образа-матрицы  $x(k+1)$  такого, что  $\text{Sp}(x(k+1)x^T(k+1)) = 1$ , определяется нейрон-победитель, синаптические веса которого  $w_q^{L*}(k)$  соответствуют центру определенного класса-текстуры. Иначе говоря, победителем является нейрон с минимальным расстоянием до входной матрицы, т. е.

$$\begin{aligned} \text{Sp}((x(k+1) - w_q^{L*}(k))(x(k+1) - w_q^{L*}(k))^T) &< \\ &< \text{Sp}((x(k+1) - w_q^L(k))(x(k+1) - w_q^L(k))^T), \\ &\forall q = 1, 2, \dots, p. \end{aligned}$$

Поскольку обучение является контролируемым, принадлежность матрицы  $x(k+1)$  к конкретному классу известна, что позволяет проанализировать две возможные ситуации, возникающие в обучаемом векторном квантовании:

- 1) входной образ  $x(k+1)$  и нейрон-победитель  $w_q^{L*}(k)$  относятся к одному и тому же классу;
- 2) входной образ и нейрон-победитель относятся к разным классам.

Тогда матричная модификация правила обучения LVQ может быть записана в виде

$$w_q^L(k+1) = \begin{cases} w_q^{L*}(k) + \eta(k+1)(x(k+1) - w_q^{L*}(k)), \\ \text{если } x(k+1) \text{ и } w_q^{L*}(k) \text{ принадлежат к одному классу;} \\ w_q^{L*}(k) - \eta(k+1)(x(k+1) - w_q^{L*}(k)), \\ \text{если } x(k+1) \text{ и } w_q^{L*}(k) \text{ принадлежат разным классам;} \\ w_q^L(k) \text{ для нейронов, не победивших в момент } k+1, \end{cases} \quad (1)$$

где  $\eta(k+1)$  – параметр шага обучения, монотонно уменьшающийся в процессе настройки.

Данное правило имеет ясный физический смысл: если нейрон-победитель и предъявленный образ относятся к одному классу, то центр  $w_q^{L*}(k)$  подтягивается к  $x(k+1)$ , перемещая при этом границу между классами. В противном же случае центр класса  $w_q^{L*}(k)$  отталкивается от  $x(k+1)$ , отталкивая при этом и соответствующую границу.

Особенностью задачи, связанной с обработкой изображений, является то, что в одном окне  $x(k+1)$  одновременно может оказаться сразу несколько текстур относящихся к разным классам. В этом случае процедуру обучения (1) следует дополнить оценкой уровня принадлежности образа  $x(k+1)$  к разным классам

$q = 1, 2, \dots, p$ . Вводя расстояние между  $x(k+1)$  и каждым из центров  $w_q^L(k+1) \forall q$  в виде

$$D(x(kn), w_q^L(kn)) = (\text{Sp}((x(k+1) - w_q^L(k+1))(x(k+1) - w_q^L(k+1))^T))^{\frac{1}{2}},$$

можно записать оценку уровня принадлежности в виде

$$\begin{aligned} \mu((x(k+1), w_q^L(k+1))) &= \frac{D^{-1}(x(k+1), w_q^L(k+1))}{\sum_{q=1}^p D^{-1}(x(k+1), w_q^L(k+1))} = \\ &= \frac{(\text{Sp}((x(k+1) - w_q^L(k+1))(x(k+1) - w_q^L(k+1))^T))^{\frac{1}{2}}}{\sum_{q=1}^p (\text{Sp}((x(k+1) - w_q^L(k+1))(x(k+1) - w_q^L(k+1))^T))^{\frac{1}{2}}}. \end{aligned} \quad (2)$$

Используя оценку (2), несложно произвести нечеткую классификацию образа, содержащего сразу несколько текстур.

**Классификация текстур в последовательности изображений.** Традиционно задача кластеризации текстур решается в «статическом» варианте, когда анализ подвергается единственное и неизменяемое изображение. Именно такой задаче посвящено подавляющее число публикаций по данной проблематике. В то же время достаточно интересной представляется задача классификации текстур в последовательности изображений, где сегменты этой последовательности изменяются во времени, могут появляться новые и исчезать старые текстуры. Такие текстуры называются «временными» (temporal) и изучены они явно недостаточно.

Так, в [18] предлагается каждый пиксель изображения рассматривать как линейную комбинацию предыдущих его состояний, т. е. в форме авторегрессионной модели. Вектором же признаков, на основе которого производится классификация, являются коэффициенты этих моделей. В [19] рассмотрена, так называемая, пространственно-временная авторегрессионная модель (STAR), куда включена предыстория не только конкретного пикселя, но и его окружения, заданного в форме окна. STAR – модель, описывающая пространственно-временную эволюцию текстур, полностью описывается вектором своих коэффициентов, который подается на классифицирующую систему – искусственную нейронную сеть.

С тем, что бы перейти от векторного описания к матричному, для описания эволюции двумерного поля, заданного в виде  $(m \times n)$  – матрицы,  $x(k)$ , воспользуемся моделью, введенной в [20], вида

$$x(k) = Ax(k-1)C + \xi(k). \quad (3)$$

Поскольку описанию (3) соответствует авторегрессия первого порядка, можно ввести в рассмотрение матричную модель авторегрессии произвольного порядка  $h$  вида

$$\begin{aligned} x(k) &= \sum_{d=1}^h A_d x(k-d) C_d + \xi(k) = \\ &= A(x(k-1) : x(k-2) : \dots : x(k-h)) C + \xi(k) = \tilde{Ax}(k-1) C + \xi(k), \end{aligned} \quad (4)$$

$$\text{где } A = (A_1 : A_2 : \dots : A_h), \quad C = \begin{pmatrix} C_1 \\ \vdots \\ C_2 \\ \vdots \\ \vdots \\ C_h \end{pmatrix},$$

$$\tilde{x}(k-1) = \begin{pmatrix} x(k-1) & & 0 \\ & x(k-2) & \\ 0 & & x(k-h) \end{pmatrix} - (hm \times hn) -$$

матрица.

Несложно заметить, что модель (4) является обобщением выражения (3) и совпадает с ним при  $h = 1$ .

Для определения  $(m \times hm)$  и  $(hn \times n)$  – матриц коэффициентов  $A$  и  $C$  можно воспользоваться адаптивным оптимальным градиентным алгоритмом идентификации матричных моделей [21] минимизирующим критерий

$$\begin{aligned} E_A^{k+1} &= \text{Sp}(x(k+1) - A(k)\tilde{x}(k)C(k))(x(k+1) - A(k)\tilde{x}(k)C(k))^T = \\ &= \text{Sp}V_A(k+1)V_A^T(k+1) = \|V_A(k+1)\|^2, \\ E_C^{k+1} &= \text{Sp}(x(k+1) - A(k+1)\tilde{x}(k)C(k))(x(k+1) - A(k+1)\tilde{x}(k)C(k))^T = \\ &= \text{Sp}V_C(k+1)V_C^T(k+1) = \|V_C(k+1)\|^2. \end{aligned}$$

и имеющим вид

$$\left\{ \begin{aligned} A(k+1) &= A(k) + (\text{Sp}V_A(k+1)C^T(k)\tilde{x}^T(k)\tilde{x}(k) \times \\ &\times C(k)V_A^T(k+1))(\text{Sp}V_A(k+1)C^T(k)\tilde{x}^T(k) \times \\ &\times (k)\tilde{x}(k)C(k)C^T(k)\tilde{x}^T(k)\tilde{x}(k)C(k)V_A^T(k+1))^T \times \\ &\times V_A(k+1)C(k)\tilde{x}^T(k), \\ C(k+1) &= C(k) + (\text{Sp}V_C^T(k+1)A(k+1)\tilde{x}(k)\tilde{x}^T(k) \times \\ &\times A^T(k+1)V_C(k+1))(\text{Sp}A(k+1)\tilde{x}(k)\tilde{x}^T(k) \times \\ &\times (k)A^T(k+1)V_C(k+1)V_C^T(k+1)A(k+1)\tilde{x}(k)\tilde{x}^T(k) \times \\ &\times (k+1)\tilde{x}^T(k)A^T(k+1)V_C(k+1)). \end{aligned} \right. \quad (5)$$

Можно также заметить, что фактически алгоритм (5) является матричным обобщением популярного алгоритма Уидроу-Хоффа [11], используемом при обучении искусственных нейронных сетей.

Объединяя матрицы коэффициентов  $A(k+1)$ ,  $C(k+1)$ , получаем  $(m+hn) \times (hm+n)$  блочную матрицу  $B(k+1)$ :

$$B(k+1) = \begin{pmatrix} A(k+1) & \vdots & O \\ \cdots & \vdots & \cdots \\ O & \vdots & C(k+1) \end{pmatrix},$$

содержащую в себе полную информацию о матричной-пространственно-временной авторегрессионной модели. Именно эта матрица подается на вход матричной VLQ-системы. При этом алгоритм обучения приобретает вид

$$w_q^L(k+1) = \begin{cases} w_q^{L*}(k) + \eta(k+1)(B(k+1) - w_q^{L*}(k)); \\ \text{если } B(k+1) \text{ и } w_{q_1 q_2}^{L*}(k) \text{ принадлежат одному классу,} \\ w_q^{L*}(k) - \eta(k+1)(B(k+1) - w_q^{L*}(k)); \\ \text{если } B(k+1) \text{ и } w_q^{L*}(k) \text{ принадлежат разным классам,} \\ w_q^L(k) \text{ для нейронов, не победивших в момент } k+1. \end{cases}$$

Несложно также записать и оценку уровней принадлежности в случае проведения нечеткой классификации:

$$\mu(B(k+1), w_q^L(k+1)) = \frac{D^{-1}(B(k+1), w_q^L(k+1))}{\sum_{q=1}^p D^{-1}(B(k+1), w_q^L(k+1))} =$$

$$= \frac{(Sp((B(k+1) - w_q^L(k+1))(B(k+1) - w_q^L(k+1))^T))^{\frac{1}{2}}}{\sum_{q=1}^p (Sp((B(k+1) - w_q^L(k+1))(B(k+1) - w_q^L(k+1))^T))^{\frac{1}{2}}}.$$

#### 4 ЭКСПЕРИМЕНТЫ

Экспериментальный анализ предложенной классифицирующей нейронной сети проводился с видеопоследовательностями, состоящими из 2250 изображений в формате RGB с прогрессивной разверткой и разрешением 640 на 416 точек. Для элиминации влияния любого возможного сжатия использовались изображения в формате .BMP. Иначе говоря, в контейнерах .AVI были представлены 90-секундные видеоряды (всего 10), для которых интерактивно определены границы сегментов, т.е. однородных по «содержанию» последовательностей видеокадров. Рис. 1 иллюстрирует левую и правую гра-

ницы сегмента, состоящего из 75 видеокадров (лесные пожары, National Geographic), для которых текстурные признаки являются максимально репрезентативными. Для различных видео число сегментов составляло 20–25 семантически однородных подпоследовательностей. При этом длина обучающей выборки варьировалась от 40 до 350 фрагментов. В вычислительном эксперименте число фрагментов составляло: 256 (блоки –  $40 \times 26$  элементов), 64 (блоки –  $80 \times 52$  элемента), 16 (блоки –  $120 \times 104$  элемента).

На рис. 2 показано представление среднего изображения рис. 1 16 блоками.

#### 5 РЕЗУЛЬТАТЫ

Результаты экспериментов позволяют утверждать, что для случаев четкой визуальной границы между темпоральными сегментами классифицирующая нейронная сеть обеспечивает 100 % определение сегментов при любых размерах фрагментов. Однако увеличение числа фрагментов нередко приводит к тому, что при «плавном» (например, медленном панорамировании или искусственных эффектах типа «растворение») переходе от сегмента к сегменту возникают ложные «блоковые» границы, не соответствующие семантике видео. Данный эффект иллюстрируется рис. 3 (циunami 11 сентября 2011 г. в Японии, NHK), на котором приведен сегмент с шагом 50 видеокадров. С одной стороны, избежать ошибок можно добавлением признаков или, используя пирамидальное представление изображений, стратифицированным огрублением (понижением разрешающей способности) изображений. С другой стороны, при рациональных размерах фрагментов получаем инструмент для более детальной сегментации видео, когда во внимание принимаются локальные изменения сюжета. В целом, классифицирующая нейронная сеть на базе анализа текстур при блочном представлении видео является достаточно перспективным инструментарием для систем поиска видеоданных по их содержанию.

#### ВЫВОДЫ

В статье предложен подход к кластеризации статических и динамических текстур в последовательности изображений с помощью обучаемого векторного квантования. Введен алгоритм обучения нейронной сети в матричном варианте, что позволяет сократить время обработки и анализировать изображение в целом. Развиваемый подход отличается численной простотой и быстродействием.



Рисунок 1 – Пример границ сегментов видеопоследовательностей



Рисунок 2 – Представление изображения блоками



Рисунок 3 – Пример ложных границ сегментов

## БЛАГОДАРНОСТИ

Работа выполнена в рамках госбюджетной темы Харьковского национального университета радиоэлектроники «Нейро-фаззи системы для текущей кластеризации и классификации последовательностей данных в условиях их искривленности отсутствующими и аномальными наблюдениями» (№ ДР 0113U000361).

## СПИСОК ЛІТЕРАТУРЫ

1. Derin H. Modeling and segmentation of noisy and textured images using Gibbs random fields / H. Derin, H. Elliot // IEEE Trans. on Pattern Analysis and Machine Intelligence. – 1987. – Vol. 9. – P. 39–55.
2. Varma M. A statistical approach to texture classification from single images / M. Varma, A. Zisserman // Int. J. of Computer Vision. – 2005. – Vol. 62. – P. 61–81.
3. Liu L. Texture classification from random features / L. Liu, P. Fieguth // IEEE Trans. on Pattern Analysis and Machine Intelligence. – 2012. – No. 34. – P. 574–586.
4. Chen Y. Q. Statistical geometrical features for texture classification / Y. Q. Chen, M. S. Nixon, D. W. Thomas // Pattern Recognition. – 1995. – No. 28. – P. 537–552.
5. Cross G. R., Jain A.K. Markov random field texture models / G. R. Cross, A. K. Jain // IEEE Trans. on Pattern Analysis and Machine Intelligence. – 1983. – No. 5. – P. 25–39.
6. Chellappa R. Classification of textures using Gaussian – Markov random fields / R. Chellappa, S. Chatterjee // IEEE Trans. on Acoust., Speech, Signal Processing. – 1985. – No. 33. – P. 959–963.
7. Ghahramani Z. An introduction to hidden Markov models and Bayesian networks / Z. Ghahramani // Int. J. of Pattern Recognition and Artificial Intelligence. – 2001. – No. 15. – P. 9–42.
8. Beal M. J. The infinite hidden Markov model / M. J. Beal, Z. Ghahramani, C. E. Rasmussen // Neural Information Processing Systems. – 2001. – P. 577–584.
9. Neural Networks for Signal Processing / ed. B. Kosko. – New Jersey : Prentice Hall, 1992. – 399 p.
10. Rojas R. Neural Network: A Systematic Introduction / R. Rojas. – Berlin : Springer-Verlag, 1996. – 502 p.
11. Хайкин С. Нейронные сети: Полный курс / С. Хайкин. – М. : Издательский дом Вильямс, 2006. – 1104 с.
12. Raveane W. Texture classification with neural networks / W. Raveane, M.A.G. Arrieta ; eds. Omatsu et. al // Distributed Computing&Artificial Intelligence. – Berlin, Springer, 2013. – P. 305–332.
13. Caldas-Pinto J. R. Self-tuning filters and predictors for two-dimensional systems / J. R. Caldas-Pinto, P. E. Wellstead // Int. J. of Control. – 1985. – Vol. 42, No. 2. – P. 457–505.
14. Wellstead P. E. Two-dimensional adaptive prediction, smoothing and filtering / P. E. Wellstead, G. R. Wagner, J. R. Caldas-Pinto // IEE Proc. – 1987. – No. 3. – P. 253–267.
15. Bodyanskiy Ye. Discrete adaptive identification and extrapolation of two-dimensional fields / Ye. Bodyanskiy, I. Pliss, V. Timofeev // Pattern Recognition and Image Analysis. – 1995. – No. 3. – P. 410–416.
16. Kohonen T. Self-organizing maps / T. Kohonen. – Berlin : Springer Verlag, 1995. – 362 p.

17. Богучарский С.И. Обработка изображений на основе комбинированного векторного квантования с использованием полуkontroliruemogo обучения / Богучарский С. И. // Вестник НТУ «ХПИ». – 2014. – № 6 (1049). – С. 20–29.
18. Mao J. Texture classification and segmentation using multiresolution simultaneous autoregressive models / J. Mao, A. K. Jain // Pattern Recognition. – 1992. – No. 25. – P. 173–188.
19. Ardizzone E. Using temporal texture for content-based video retrieval / E. Ardizzone, A. Capra, M. La Cascia // Journal of Visual Languages and Computing. – 2000. – No. 11. – P. 241–252.
20. Кунцевич В.М. О решении задачи двумерной дискретной фильтрации // Автоматика и телемеханика. – 1987. – № 6. – С. 68–78.
21. Бодянский Е.В. О решении задачи управления матричным объектом в условиях неопределенности / Е. В. Бодянский, И. П. Плисс // Автоматика и телемеханика. – 1990. – No. 2. – С. 175–178.

Статья поступила в редакцию 17.09.2014.

Посля доработки 20.10.2014.

Богучарський С. І.<sup>1</sup>, Машталір С. В.<sup>2</sup>

<sup>1</sup>Аспірант, Харківський національний університет радіоелектроніки, Україна

<sup>2</sup>Канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, Україна

## АНАЛІЗ ТЕКСТУР В ПОСЛІДОВНОСТЯХ ЗОБРАЖЕНЬ НА БАЗІ ВЕКТОРНОГО КВАНТУВАННЯ

В роботі запропонованій підхід до аналізу послідовностей зображень (відеоданих). Для цього розроблені матричні аналоги існуючих нейромережевих підходів, що дозволяє враховувати просторові зв’язки мультимедійної інформації та скоротити час необхідний на обробку інформації за рахунок введення нової матричної процедури навчання нейронної мережі. В якості базової ознаки, за якою проводиться розбиття вихідних даних, було обрано текстури, що в свою чергу дозволяє підвищити точність класифікації.

**Ключові слова:** зображення, текстура, класифікація, матричні алгоритми, нейронні мережі.

Bogucharskiy S. I.<sup>1</sup>, Mashtalir S. V.<sup>2</sup>

<sup>1</sup>Postgraduate student, Kharkiv National University of Radio Electronics, Ukraine

<sup>2</sup>Ph.D., Associate Professor, Kharkiv National University of Radio Electronics, Ukraine

## IMAGE SEQUENCES TEXTURE ANALYSIS BASED ON VECTOR QUANTIZATION

The approach for the image sequence (videodata) analysis is proposed. For this purpose, the matrix analogs of existing neural network approaches is developed. This allows to takes into account the spatial relationships of multimedia information, and to reduce the time of information processing through the introduction of a new neural network matrix training procedures.

The texture is selected as a basic characteristic for original data partition, which, in turn, improves the clustering accuracy.

**Keywords:** image, texture, clustering, matrix algorithm, neural networks.

## REFERENCES

1. Derin H., Elliot H. Modeling and segmentation of noisy and textured images using Gibbs random fields, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1987, No. 9, pp. 39–55.
2. Varma M., Zisserman A. A statistical approach to texture classification from single images, *Int. J. of Computer Vision*, 2005, No. 62, pp. 61–81.
3. Liu L., Fieguth P. Texture classification from random features, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012, No. 34, pp. 574–586.
4. Chen Y. Q., Nixon M. S., Thomas D. W. Statistical geometrical features for texture classification, *Pattern Recognition*, 1995, Vol. 28, pp. 537–552.
5. Cross G. R., Jain A. K. Markov random field texture models, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1983, No. 5, pp. 25–39.
6. Chellappa R., Chatterjee S. Classification of textures using Gaussian-Markov random fields, *IEEE Trans. on Acoust., Speech, Signal Processing*, 1985, Vol. 33, pp. 959–963.
7. Ghahramani Z. An introduction to hidden Markov models and Bayesian networks, *Int. J. of Pattern Recognition and Artificial Intelligence*, 2001, No. 15, pp. 9–42.
8. Beal M. J., Ghahramani Z., Rasmussen C. E. the infinite hidden Markov model, *Neural Information Processing Systems*, 2001, pp. 577–584.
9. Neural Networks for Signal Processing / Ed. by B. Kosko, Englewood Cliffs, New Jersey, Prentice Hall, 1992, 399 p.
10. Rojas R. Neural Network : A Systematic Introduction. Berlin, Springer-Verlag, 1996, 502 p.
11. Haikin S. Neironnije seti: Polniy kurs. Moscow, Izdatelskiy dom «Wiley», 2006, 1104 p.
12. Raveane W., M.A.G. Arrieta ; Eds. Omatsu et. al. Texture classification with neural networks, *Distributed Computing & Artificial Intelligence*. Berlin, Springer, 2013, pp. 305–332.
13. Caldas-Pinto J. R., Wellstead P. E. Self-tuning filters and predictors for two-dimensional systems, *Int. J. of Control*, 1985, Vol. 42, No. 2, pp. 457–505.
14. Wellstead P. E., Wagner G. R., Caldas-Pinto J. R. Two-dimensional adaptive prediction, smoothing and filtering, *IEE Proc*, 1987, No. 3, pp. 253–267.
15. Bodyanskiy Ye, Pliss I., Timofeev V., Discrete adaptive identification and extrapolation of two-dimensional fields, *Pattern Recognition and Image Analysis*, 1995, No. 3, pp. 410–416.
16. Kohonen T. Self-organizing maps. Berlin, Springer Verlag, 1995, 362 p.
17. Bogucharskiy S. I. Obrabotka izobrazenij na osnove kombinirovannogo vektornogo kvantovaniya s ispolzovaniem polukontroliruemogo obucheniya, *Vestnik NTU «KhPI»*, 2014, No. 6 (1049), pp. 20–29.
18. Mao J., Jain A. K. Texture classification and segmentation using multiresolution simultaneous autoregressive models, *Pattern Recognition*, 1992, No. 25, pp. 173–188.
19. Ardizzone E., Capra A., La Cascia M. Using temporal texture for content-based video retrieval, *Journal of Visual Languages and Computing*, 2000, No. 11, pp. 241–252.
20. Kunzevich V. M. O reshenii zadachi dvumernoy diskretnoy filtracii, *Avtomatika i telemehanika*, 1987, No. 6, pp. 68–78.
21. Bodyanskiy Ye. V., Pliss I. P. O reshenii zadachi upravleniya matrichnim obyektom v usloviyah neopredelennosti, *Avtomatika i telemehanika*, 1990, No. 2, pp. 175–178.

УДК 004.383.8.032.26

Дмитриенко В. Д.<sup>1</sup>, Заковоротный А. Ю.<sup>2</sup><sup>1</sup>Д-р техн. наук, профессор кафедры вычислительной техники и программирования, Национальный технический университет «Харьковский политехнический институт», Украина<sup>2</sup>Канд. техн. наук, доцент, доцент кафедры вычислительной техники и программирования Национальный технический университет «Харьковский политехнический институт», Украина

## АРХИТЕКТУРЫ И АЛГОРИТМЫ ФУНКЦИОНИРОВАНИЯ НЕЙРОННЫХ СЕТЕЙ ХЕММИНГА И ХЕББА, СПОСОБНЫХ ДООБУЧАТЬСЯ И РАСПОЗНАВАТЬ НОВУЮ ИНФОРМАЦІЮ

Решена задача дообучения классических дискретных нейронных сетей Хемминга и Хебба без потерь запомненной ранее информации. Объектом исследования являются процессы распознавания и классификации образов в системах, построенных на основе искусственных нейронных сетей. Предметом исследования являются архитектуры и алгоритмы функционирования искусственных нейронных сетей. Цель работы: разработка стабильно-пластичных нейронных сетей Хемминга и Хебба. Разработаны архитектуры и алгоритмы функционирования дискретных стабильно-пластичных нейронных сетей Хемминга и Хебба, которые не только могут дообучаться в процессе функционирования, но и распознавать новую информацию. Новые сети могут стать альтернативой дискретным нейронным сетям адаптивной резонансной теории. Разработанный подход по дообучению может быть обобщен и на другие нейронные сети. Проведены экспериментальные исследования разработанных алгоритмов функционирования искусственных нейронных сетей. Результаты экспериментов подтверждают правильность предложенного подхода.

**Ключевые слова:** распознавание и классификация образов, стабильно-пластичные нейронные сети, нейронная сеть Хемминга, нейронная сеть Хебба, адаптивная резонансная теория.

### НОМЕНКЛАТУРА

- АРТ – адаптивная резонансная теория;
- АРТ-1 – дискретная нейронная сеть адаптивной резонансной теории;
- НС – нейронная сеть;
- $A_p$  –  $p$ -й входной сигнал нейронов;
- $a$  – число компонент, в которых совпадают входной и эталонный вектора в нейронной сети Хемминга;
- $a_{\min}$  – минимально допустимое значение числа совпадающих компонент входного и эталонного векторов в сети Хемминга;
- $F$  – обобщенная функция числа нейронов весов связей нейронной сети и входного изображения;
- $F_i$  – обобщенная функция;
- $g_A(U_{\text{вх}})$  – функция активации  $A$ -нейронов сети Хемминга;
- $g_Y(U_{\text{вх}})$  – функция активации  $Y$ -нейронов сети Хемминга;
- $g_Z(U_{\text{вх}})$  – функция активации  $Z$ -нейронов сети Хемминга;
- $k_1$  – константа;
- $L$  – число дообучений нейронной сети Хебба;
- $m$  – число  $Y$ -нейронов сети Хебба;
- $M$  – множество из  $m$  пар входных распознаваемых изображений и соответствующих им выходных изображений нейронной сети Хебба;
- $n$  – количество входных нейронов сети Хемминга;

$n_d$  – число нейронов в  $d$ -й группе  $Y$ -нейронов сети Хебба;

$N$  – число групп распознающих нейронов сети Хебба;

$N$  – число нейронов сети;

$N_{\text{пп}}$  – предельное число эпох обучения сети Хебба;

$N_{\text{об}}$  – число обученных нейронов в сети АРТ;

$N_{\text{изб}}$  – число избыточных нейронов в сети АРТ;

$q$  – число  $Z$ -нейронов, хранящих эталонные изображения в сети Хемминга;

$R_x$  – расстояние по Хеммингу между входным и эталонным векторами;

$S^k$  –  $k$ -е входное изображение нейронной сети Хебба;

$S_i^k$  –  $i$ -я биполярная компонента  $k$ -го входного изображения сети Хебба;

$S_i$  –  $i$ -й входной нейрон сети Хемминга;

$S^*$  – входное изображение;

$Y^*$  – вектор выходных сигналов нейронной сети;

$Y^p$  – выходной вектор, сопоставленный  $p$ -му изображению;

$Y_m^p$  –  $m$ -я компонента выходного вектора  $Y^p$ ;

$y_m^k$  –  $m$ -я компонента выходного вектора сети для изображения  $S^k$ ;

$t$  – время;

$t^k$  – вектор выходных сигналов нейронной сети Хебба, соответствующий входному изображению  $S^k$ ;

$t_m^k$  –  $m$ -я компонента вектора входных сигналов  $t^k$ ;  
 $U_{\text{вх}} A_p$  – входной сигнал нейрона  $A_p$  сети Хемминга;  
 $U_{\text{вых}} A_p$  – выходной сигнал нейрона  $A_p$  сети Хемминга;  
 $U_{\max}$  – константа;  
 $U_{\text{вх}} S_i$  – входной сигнал элемента  $S_i$  нейронной сети Хемминга;  
 $U_{\text{вых}} S_i$  – выходной сигнал элемента  $S_i$  нейронной сети Хемминга;  
 $U_{\text{вх}} Y_j$  – выходной сигнал нейрона  $Y_j$  сети Хебба;  
 $U_{\text{вых}} Y_j$  – выходной сигналы нейрона  $Y_j$  сети Хебба;  
 $U_{\text{вых}} Y_k^d$  – выходной сигнал нейрона  $Y_k^d$   $d$ -й группы нейронов сети Хебба;  
 $U_{\text{вх}} Z_p$  – входной сигнал нейрона  $Z_p$  сети Хемминга;  
 $U_{\text{вых}} Z_m$  – выходной сигнал нейрона  $Z_m$  сети Хемминга;  
 $U_{\text{вх}} \Sigma_d$  – входной сигнал нейрона  $\Sigma_d$  сети Хебба;  
 $U_{\text{вых}} \Sigma_d$  – выходной сигнал нейрона  $\Sigma_d$  сети Хебба;  
 $W$  – множество весов связей нейронной сети;  
 $W_{\text{об}}$  – число обученных весов связей нейронной сети АРТ;  
 $W_{\text{изб}}$  – число избыточных весов связей нейронной сети АРТ;  
 $W_{ij}^A$  – вес связи между нейронами  $A_i$  и  $A_j$  в сети Хемминга;  
 $W_{ip}$  – вес связи от  $i$ -го нейрона входного  $S$ -слоя к  $p$ -му нейрону  $Z$ -слоя сети Хемминга;  
 $W_{ij}$  – вес связи от  $i$ -го нейрона  $S$ -слоя к  $j$ -му нейрону  $Y$ -слоя сети Хебба;  
 $W_{ij}(\text{new})$  – новый вес связи нейронной сети Хебба;  
 $W_{ij}(\text{old})$  – предыдущий вес связи нейронной сети Хебба;  
 $Y^*$  – вектор выходных сигналов нейронной сети;  
 $Z_p$  –  $p$ -й элемент слоя  $Z$ -нейронов сети Хемминга;  
 $\varepsilon$  – константа;  
 $\Sigma$  – дополнительный суммирующий нейрон сети Хемминга для обнаружения новой информации на входе сети.

## ВВЕДЕНИЕ

Системы распознавания, классификации и управления на основе искусственных НС широко применяются для решения самых различных задач. Успешное применение НС при решении разнообразных задач обусловлено их универсальностью, возможностью замены различных традиционных проблем, возникающих при решении конкретных прикладных задач, обучением НС на примерах, возможностью самообучения и разработки НС под решаемую задачу, и т. д.

При наличии достаточно полной информации об объекте исследования, возможно использование целого ряда различных видов НС, подходящих для создания распознавающих, классифицирующих и управляющих систем.

Однако при разработке систем для реальных объектов зачастую возникают ситуации, когда информация об объекте далеко не полная или будет уточняться в процессе функционирования объекта. Это обстоятельство резко сокращает круг НС, которые целесообразно использовать в таких ситуациях, поскольку обучение новому образу в общем случае приведет к полному переобучению НС. Также ограничение на применяемые сети накладывает и нечувствительность сетей к новым данным во входной информации. Все вышеперечисленное привело к разработке принципиально новых конфигураций НС использующих в своей основе АРТ и решающих проблему стабильности-пластичности, то есть проблему восприятия и запоминания новой информации при сохранении уже имеющейся.

Указанное достоинство сетей АРТ достигается за счет введения в архитектуру НС избыточных нейронов, которые используются только для дообучения сети. Введение избыточных нейронов может позволить решить проблему дообучения или проблему стабильности-пластичности и других НС.

Целью статьи является разработка стабильно-пластичных нейронных сетей Хемминга и Хебба.

## 1 ПОСТАНОВКА ЗАДАЧИ

Пусть некоторая нейронная сеть с  $n$  входными и  $p$  выходными нейронами обучена распознаванию множества входных изображений  $S = \{S^1 = (S_1^1, \dots, S_n^1), S^2 = (S_1^2, \dots, S_n^2), \dots, S^k = (S_1^k, \dots, S_n^k)\}$  каждому из которых на выходе нейронной сети соответствует свой выходной вектор  $Y^P = (Y_1^P, \dots, Y_m^P)$ ,  $p = \overline{1, k}$ . Тогда распознавание некоторого входного изображения  $S^* = (S_1^*, \dots, S_n^*)$  сводится к вычислению соотношения  $Y^* = F(N_H, W, S^*)$ .

Поскольку при обучении нейронной сети, как правило, определяется все множество весов связей  $\{W_{11}, W_{12}, \dots\}$  всех нейронов, то при необходимости дообучить нейронную сеть распознаванию еще одного изображения  $S^{k+1}$  приходится выполнять полное и трудоемкое обучение сети распознаванию множества изображений  $S = \{S^1, \dots, S^k, S^{k+1}\}$ . Этот недостаток известных архитектур нейронных сетей привел к разработке нейронных сетей АРТ, в которых для запоминания новой информации вводились избыточные нейроны, благодаря которым не нужно менять веса связей, хранящих информацию об уже запомненных изображениях. Кроме того, сети АРТ обладают еще одним важным свойством – они могут обнаруживать на своих входах новые изображения, которых нет в памяти сети, а затем могут запоминать их как представителей новых классов. Фактически распознавание некоторого входного изображения  $S^*$  в сетях АРТ выполняется с помощью соотношения  $Y^* = F_1(N_{\text{об}}, N_{\text{изб}}, W_{\text{об}}, W_{\text{изб}}, S^*)$ .

Ставится задача определения функций  $F_1$  для нейронных сетей Хемминга и Хебба, которые превратили бы эти сети в стабильно-пластичные, т.е. сети, способные дообучаться и, кроме того, распознавать на своих входах новую информацию. Для этого необходимо модифицировать архитектуру и алгоритмы функционирования этих нейронных сетей, а также провести экспериментальные исследования разработанных модификаций нейронных сетей Хемминга и Хебба.

## 2 ОБЗОР ЛІТЕРАТУРЫ

Проблема стабильности-пластичности является одной из самых сложных и трудно решаемых. Характер восприятия получаемой информации постоянно связан с решением дилеммы, является ли некоторый образ «новой» информацией, и, следовательно, реакция на него должна быть поисково-познавательной, с сохранением этого образа в памяти, либо этот образ является вариантом «старой», уже запомненной информации, и в этом случае реакция должна соответствовать ранее накопленному опыту. Специальное запоминание этого образа в последнем случае не требуется. Таким образом, такое восприятие одновременно пластично, адаптировано к новой информации, и при этом оно стабильно, то есть не разрушает память о старых образах [1–7].

Не все нейронные сети приспособлены к решению этой задачи. Так, например, многослойный персептрон, обучающийся по методу обратного распространения, запоминает весь пакет обучающей информации, при этом образы обучающей выборки предъявляются в процессе обучения многократно. Попытки затем обучить персептрон новому образу приведут к модификации синаптических связей с неконтролируемым разрушением памяти о предыдущих образах. Следовательно, персептрон не способен к запоминанию дополнительной информации, ему необходимо полное переобучение сети.

Аналогичная ситуация имеет место в сетях Хебба и сети Кохонена, обучающихся на основе самоорганизации. Сети Хебба, Кохонена, Хемминга и др. всегда выдают положительный результат при классификации. Тем самым, эти нейронные сети не в состоянии отделить новые образы (изображения) от искаженных или зашумленных версий старых образов.

Исследования по проблеме стабильности-пластичности привели к созданию нового научного направления – адаптивной резонансной теории и НС на ее основе. Принцип работы сетей АРТ заключается в следующем: нейронной сетью производится сопоставление входных изображений с изображениями известных классов, которые получены на этапе обучения сети и хранятся в ее весах связей. Этап сопоставления входных изображений с известными классами изображений осуществляется с определенной точностью, задаваемой специальным параметром сходства. Если параметр сходства входного изображения с одним из известных прототипов изображений превышает заданный порог, то в режиме обучения прототип данного класса модифицируется, чтобы стать более похожим на предъявленное изображение, а входное

изображение относится к данному классу изображений. В свою очередь, если параметр сходства входного изображения не достигает заданного порога при анализе всех запомненных классов изображений, то на основе входного изображения создается новый класс изображений, что становится возможным благодаря избыточным (нераспределенным) нейронам, входящим в структуру нейронной сети. Таким образом, новые образы могут создавать новые классы, без искажения запомненной ранее информации [8–12].

Указанное достоинство сетей АРТ достигается за счет введения в архитектуру нейронных сетей избыточных нейронов, которые используются только для дообучения сети. Введение избыточных нейронов может позволить решить проблему дообучения или проблему стабильности-пластичности и других нейронных сетей. Поэтому в данной работе разработаны архитектуры и алгоритмы функционирования дискретных стабильно-пластичных нейронных сетей Хемминга и Хебба, которые не только могут дообучаться, но и распознавать новую информацию. Новые сети могут стать альтернативой дискретным нейронным сетям адаптивной резонансной теории.

## 3 МАТЕРИАЛЫ И МЕТОДЫ

Стабильность дискретных нейронных сетей АРТ-1 [10] обеспечивается наличием дополнительных нейронов, на которые и записывается новая информация. При этом каждый распознавающий нейрон хранит или будет хранить информацию в весах своих связей только об одном изображении. Подобный слой нейронов имеет и сеть Хемминга (рис. 1), где каждый  $Z_p$  ( $p = 1, \dots, m$ ) нейрон в весах своих связей хранит одно изображение, являющееся представителем своего класса.

Сеть Хемминга имеет  $n$  входных нейронов  $S_1, \dots, S_n$ , воспринимающих бинарные компоненты входных изображений  $S^q = (S_1^q, \dots, S_n^q)$ ,  $q = 1, \dots, L$ . Входные и выходные сигналы  $S$ -элементов при входном изображении  $S^q$  определяются соотношениями  $U_{\text{вх}} S_i = S_i^q$ ,  $U_{\text{вых}} S_i = U_{\text{вх}} S_i$ ,  $i = 1, \dots, n$ .

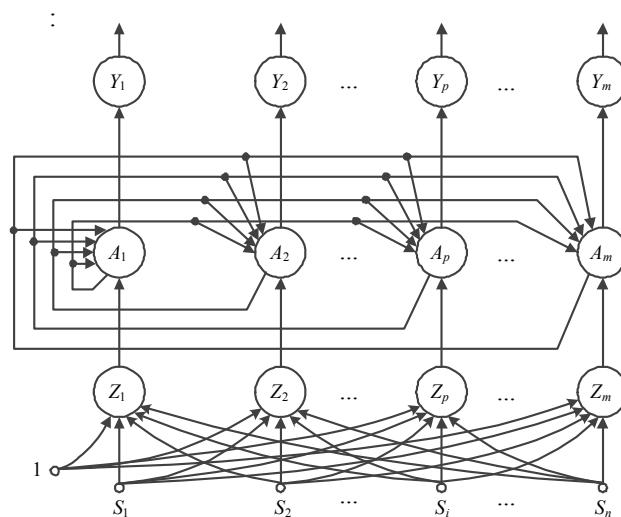


Рисунок 1 – Нейронная сеть Хемминга

При предъявлении входного изображения  $S^q$  каждый  $Z$ -нейрон рассчитывает свой входной сигнал в соответствии с выражением вида (1):

$$U_{\text{вх} Z_p} = \frac{n}{2} + \sum_{i=1}^n W_{ip} S_i^q, p = 1, \dots, m. \quad (1)$$

Функция активации  $Z$ -нейронов описывается соотношением (2):

$$g_Z(U_{\text{вх}}) = \begin{cases} 0, & \text{если } U_{\text{вх}} \leq 0, \\ k_1 U_{\text{вх}}, & \text{если } 0 < k_1 U_{\text{вх}} \leq U_{\text{max}}, \\ U_{\text{max}}, & \text{если } k_1 U_{\text{вх}} > U_{\text{max}}. \end{cases} \quad (2)$$

С помощью функции активации (2) определяются выходные сигналы нейронов  $Z$ -слоя:  $U_{\text{вых} Z_1}, \dots, U_{\text{вых} Z_m}$ , которые являются входными сигналами нейронов  $A_p$  ( $p = 1, \dots, m$ ). Функция активации нейронов  $A_p$  ( $p = 1, \dots, m$ ) и веса их связей задаются соотношениями:

$$g_A(U_{\text{вх}}) = \begin{cases} U_{\text{вх}}, & \text{если } U_{\text{вх}} > 0; \\ 0, & \text{если } U_{\text{вх}} \leq 0, \end{cases}$$

$$W_{ij}^A = \begin{cases} 1, & \text{если } i = j; \\ -\varepsilon, & \text{если } i \neq j, \quad i, j = 1, \dots, m, \end{cases}$$

где  $0 < \varepsilon \leq 1/m$ .

Слой  $A$ -нейронов функционирует циклически, динамика элементов описывается итерационным выражением:

$$U_{\text{вх} A_p}(t+1) = q_A(U_{\text{вых} A_p}(t) - \varepsilon \sum_{j=1, j \neq p}^m U_{\text{вых} A_j}(t)), \\ p = 1, \dots, m, \quad t = 0, 1, 2, \dots,$$

при начальных условиях:  $U_{\text{вх} A_p}(0) = U_{\text{вых} Z_p}$ ,  $p = 1, \dots, m$ .

Если среди входных сигналов нейронов  $A_1, \dots, A_m$  окажется один наибольший сигнал  $U_{\text{вых} A_p}$ , то в результате итерационного процесса только один  $A$ -нейрон останется с выходным сигналом, большим нуля. Поскольку выходные сигналы  $A$ -элементов поступают на входы  $Y$ -нейронов, которые имеют функцию активации вида:

$$g_Y(U_{\text{вх}}) = \begin{cases} 1, & \text{если } U_{\text{вх}} > 0, \\ 0, & \text{если } U_{\text{вх}} \leq 0, \end{cases}$$

то в результате на выходе сети Хемминга только один нейрон  $Y_p$  окажется с единичным выходным сигналом. Единичный выход этого нейрона и нулевые всех остальных и будут указывать на то, что предъявленное изображение  $S^q$  наиболее близко, в смысле расстояния Хемминга, к эталонному изображению, хранящемуся в весах связей нейрона  $Z_p$ . Если в слой  $Z$ -нейронов сети

Хемминга ввести дополнительные  $Z$ -нейроны (и соответствующие им  $A$ - и  $Y$ -нейроны), способные запоминать новую информацию (эталонные изображения новых классов), то такая сеть также сможет дообучаться и будет обладать свойствами стабильности и пластичности, то есть свойствами дискретной НС АРТ-1.

Еще одно достоинство нейронных сетей АРТ – возможность выделять новую информацию. Это достигается тем, что входное изображение сравнивается по величине параметра сходства с эталонными изображениями, хранящимися в весах связей распознающих нейронов. Если входное изображение относится к новому классу, то при сравнении с любым эталонным изображением, хранящимся в весах связей распределенных распознающих нейронов, получается параметр сходства меньше минимально допустимого. Поэтому в результате такого сравнения все распределенные распознающие  $Y$ -нейроны будут заторможены, что и является признаком того, что входное изображение несет новую информацию, которой нет в памяти нейронной сети АРТ.

В нейронной сети Хемминга сравнение изображений (или bipolarных векторов) выполняется по числу компонент, в которых входное и эталонное изображения совпадают:  $a = n - R_x$ . Поэтому по аналогии с НС АРТ можно задать минимально допустимое значение числа совпадающих компонент  $a_{\min}$ , при которых могут сравниваться входное и эталонное изображения, а при  $a < a_{\min}$  входное изображение и эталонное несравнимы или, другими словами, входное изображение несет новую информацию по сравнению с изображением, хранящимся в весах связей нейронной сети. Учесть значение  $a_{\min}$  в архитектуре НС несложно – достаточно ввести дополнительное смещение на нейроны  $Z$ -слоя (точнее – изменить величину смещения  $n/2$ ). При этом выходные сигналы  $Z$ -нейронов будут описываться модифицированным соотношением (1):

$$U_{\text{вых} Z_p} = \frac{n}{2} - a_{\min} + 1 + \sum_{i=1}^n W_{ip} U_{\text{вых} S_i} = a - a_{\min} + 1, \\ p = 1, \dots, q. \quad (3)$$

Из (3) и (2) следует:

$$U_{\text{вых} Z_p} = \begin{cases} 0, & \text{если } a < a_{\min}, \\ 1, & \text{если } a = a_{\min}, \\ k_1 U_{\text{вх} Z_p}, & \text{если } a > a_{\min} \text{ и } 0 < k_1 U_{\text{вх} Z_p} \leq U_{\text{max}}, \\ U_{\text{max}}, & \text{если } a > a_{\min} \text{ и } k_1 U_{\text{вх} Z_p} > U_{\text{max}}. \end{cases}$$

Таким образом, если на выходе всех распределенных  $Z$ -нейронов выходные сигналы равны нулю ( $U_{\text{вых} Z_p} = 0$ ,  $p = 1, \dots, q$ ), то входное изображение по принятому критерию близости изображений не является похожим ни на одно из известных эталонных изображений и несет новую информацию. Для аппаратного определения новизны входного изображения в архитектуру нейронной сети Хемминга достаточно ввести один

нейрон  $\Sigma$ , суммирующий выходные сигналы всех нейронов  $Z$ -слоя. Если после предъявления входного изображения на выходе нейрона  $\Sigma$  нулевой выходной сигнал, то это означает, что входное изображение по расстоянию Хемминга находится далеко от эталонных изображений, хранящихся в весах связей нейронов  $Z$ -слоя, и является представителем нового класса изображений, который можно запомнить нераспространенным нейроном сети Хемминга. Если после предъявления входного изображения выходной сигнал нейрона положителен, то это означает, что входное изображение является представителем известного класса изображений, хранящегося в памяти нейронной сети.

Результаты, полученные по сети Хемминга, могут быть обобщены и на другие нейронные сети. Проиллюстрируем это на примере нейронной сети Хебба. Архитектура этой сети приведена на рис. 2 [10]. Основу архитектуры составляет группа из  $m$  бинарных или биполярных нейронов  $Y_1, \dots, Y_m$ . Сеть может запоминать до  $2^m$  черно-белых изображений. Однако, применение этой сети для запоминания и распознавания  $2^m$  (или близких к  $2^m$  чисел) различных изображений во многих случаях приводит к неразрешимым проблемам адаптации весов связей нейросети. Поэтому часто рекомендуют использовать данную сеть для запоминания только  $m$  различных классов изображений, задавая каждое из них, единичным выходным сигналом на выходе только одного нейрона  $Y$ -слоя (выходные сигналы остальных ( $m - 1$ ) элементов должны принимать значение « $-1$ » для биполярных нейронов или « $0$ » для бинарных).

Нейронная сеть обучается с помощью алгоритма на основе правила Хебба. Для биполярных нейронов алгоритм обучения имеет следующие основные шаги.

*Шаг 1.* Задается множество  $M = \{(S^1, t^1), \dots, (S^m, t^m)\}$ , состоящее из  $m$  пар (входное изображение  $S^k = (S_1^k, \dots, S_n^k)$ , необходимый вектор выходных сигналов  $t^k = (t_1^k, \dots, t_m^k)$ ,  $k = 1, \dots, m$ ), задаются нулевые выходные сигналы нейронов  $Y$ -слоя:  $U_{\text{вых}Y_j} = 0$ ,  $j = 1, \dots, m$ . Задается предельное число эпох  $N_{\text{пр}}$  обучения сети. Инициируются веса связей нейронов:  $W_{ij} = 0$ ,  $i = 0, 1, \dots, n$ ,  $j = 1, \dots, m$ .

*Шаг 2.* Начинается обучение нейронной сети правильной реакции на каждое входное изображение. Последовательно каждая пара  $(S^k, t^k)$ ,  $k = 1, \dots, m$ , проверяется на правильность реакции нейронной сети на входное изображение. Если полученный выходной вектор сети  $(y_1^k, \dots, y_m^k)$  отличается от заданного  $t^k = (t_1^k, \dots, t_m^k)$ , то

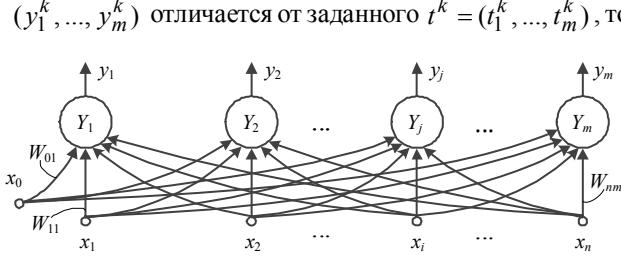


Рисунок 2 – Нейронная сеть Хебба

выполняются шаги 3–5 алгоритма. Если выходной вектор сети совпадает с требуемым, то осуществляется переход к следующей паре. Если выходной вектор совпадает с заданным и проверяемая пара является последней, то переход на шаг 7 алгоритма.

*Шаг 3.* Инициируется множество входных нейронов:

$$x_0 = 1, \quad x_i = S_i^k, \quad i = 1, \dots, n.$$

*Шаг 4.* Инициируется множество выходных нейронов сети:  $U_{\text{вых}Y_j} = t_j^k$ ,  $j = 1, \dots, m$ .

*Шаг 5.* Корректируются веса связей нейронов сети по правилу Хебба:  $W_{ij}(\text{new}) = W_{ij}(\text{old}) + x_i t_j^k$ ,  $i = 0, 1, \dots, n$ ,  $j = 1, \dots, m$ .

*Шаг 6.* Если скорректированы веса связей для последней пары  $(S^m, t^m)$ , то переход на шаг 7 алгоритма. В противном случае – на шаг 2 алгоритма.

*Шаг 7.* Проверяется условие останова, то есть правильность функционирования сети при предъявлении каждого входного изображения. Если сеть функционирует правильно, то переход на шаг 8 алгоритма. Если сеть функционирует неправильно, то проверяется достижение предельного числа эпох обучения сети  $N_{\text{пр}}$ . Если  $N_{\text{пр}}$  не достигнуто, по переход на шаг 2 алгоритма, в противном случае – переход на следующий шаг алгоритма.

*Шаг 8.* Останов.

Для обеспечения стабильности в нейронных сетях АРТ при запоминании новой информации используются дополнительные нейроны, введение которых позволяет запоминать новую информацию, и не разрушать уже имеющуюся. Этот же эффект может быть достигнут и в нейронной сети Хебба при введении дополнительных распознающих нейронов.

Когда нейронная сеть Хебба обучается распознаванию множества из  $n_0$  ( $n_0 \geq 3$ ) изображений  $M_0 = \{S^1, \dots, S^{n_0}\}$ , каждое из которых кодируется на выходах  $Y$ -нейронов  $(Y_1, \dots, Y_{n_0})$  одним единичным сигналом и  $(n_0 - 1)$  сигналами « $-1$ ». Суммарный выходной сигнал нейронов  $Y_1, \dots, Y_{n_0}$ , правильно распознающих одно из изображений, относящихся к множеству  $M_0$ , в этом случае равен:

$$\sum_{i=1}^{n_0} U_{\text{вых}Y_i} = -(n_0 - 2). \quad (4)$$

Если при предъявлении нейронной сети изображения  $S^*$  соотношение (4) не выполняется, то это является признаком появления на выходе нейронной сети новой информации (представителя нового класса изображений). При накоплении некоторого множества таких изображений  $M_1 = \{S^{n_0+1}, S^{n_0+2}, \dots, S^{n_0+n_1}\}$ , где  $n_1 \geq 3$ , могут быть обучены с помощью алгоритма Хебба следующие  $n_1$  нейронов сети. Признаком того, что некоторое входное изображение  $S^*$  относится к множеству  $M_1$  являются соотношения (5):

$$\sum_{i=1}^{n_0} U_{\text{вых}Y_i}(S^*) \neq -(n_0 - 2);$$

$$\sum_{i=1}^{n_1} U_{\text{вых}} Y_i^0(S^*) = -(n_1 - 2). \quad (5)$$

Появление изображений, для которых не выполняются соотношения (4) и (5) указывает на необходимость следующего дообучения нейронной сети. Число  $N$  шагов дообучения практически не ограничено.

На рис. 3 приведена архитектура нейронной сети Хебба, которая может дообучаться  $N$  раз. Сеть состоит из слоя входных нейронов  $x_1, \dots, x_n$  и  $Y$ -слоя распознающих нейронов. Нейроны  $Y$ -слоя разбиты на  $(N + 1)$  группу, каждая из которых может содержать разное число нейронов (но не меньше трех) и формироваться в процессе функционирования сети.

Первая группа нейронов  $Y_1^0, \dots, Y_{n_0}^0$  используется для запоминания исходной информации в виде множества изображений  $M_0$ . Вторая и последующие группы нейронов  $(Y_1^1, \dots, Y_{n_1}^1), \dots, (Y_1^N, \dots, Y_{n_N}^N)$  используются для обучения нейронной сети по мере накопления новой информации и последовательного запоминания множеств изображений  $M_1, M_2, \dots, M_N$ . В каждой группе нейронов выходные сигналы  $Y$ -нейронов поступают на входы управляющего нейрона, имеющего функцию активации вида (6):

$$U_{\text{вых}} \Sigma_d = \begin{cases} 0, & \text{если } U_{\text{вх}} \Sigma_d = \sum_{k=1}^{n_d} U_{\text{вых}} Y_k^d = -(n_d - 2); \\ 1, & \text{если } U_{\text{вх}} \Sigma_d \neq -(n_d - 2), d = 0, 1, \dots, N. \end{cases} \quad (6)$$

В начале режима распознавания обнуляются выходные сигналы  $\Sigma$ -нейронов и блокируется работа нейронов  $Y_1^1, Y_2^1, \dots, Y_{n_1}^1$ .

При подаче на вход нейронной сети, которая дообучалась  $L$  раз ( $1 < L < N$ ), некоторого изображения  $S^*$  вначале активируются нейроны первой группы  $Y_1^0, \dots, Y_{n_0}^0$ , выходные сигналы которых поступают на входы управляющего нейрона  $\Sigma_0$ . Если выполняется соотношение (7):

$$U_{\text{вх}} \Sigma_0 = \sum_{k=1}^{n_0} U_{\text{вых}} Y_k^0 = -(n_0 - 2), \quad (7)$$

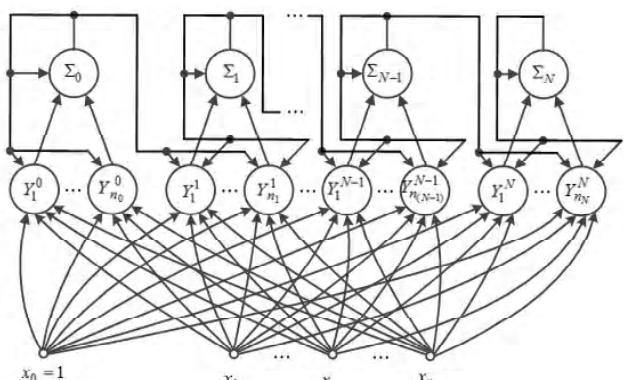


Рисунок 3 – Архитектура нейронной сети Хебба, которая может дообучаться  $N$  раз

то  $U_{\text{вых}} \Sigma_0 = 0$  и входное изображение  $S^*$  относится к множеству изображений  $M_0$ , а конкретный класс изображений определяют выходные сигналы нейронов  $Y_1^0, \dots, Y_{n_0}^0$ .

Если соотношение (7) не выполняется, то  $U_{\text{вых}} \Sigma_0 = 1$ , что указывает на то, что входное изображение  $S^*$  не принадлежит множеству  $M_0$ . Единичный выходной сигнал нейрона  $\Sigma_0$  по цепи обратной связи фиксирует единичный выходной сигнал нейрона  $\Sigma_0$  и поступает на входы нейронов первой группы, блокируя их. Одновременно этот же единичный сигнал управляющего нейрона поступает на входы распознающих нейронов  $Y_1^1, \dots, Y_{n_1}^1$  второй группы и переводит их в активное состояние. Выходные сигналы нейронов  $Y_1^1, \dots, Y_{n_1}^1$  поступают на входы управляющего нейрона  $\Sigma_1$ , с помощью которого определяется принадлежность изображения  $S^*$  к множеству  $M_1$ . Если выполняется соотношение (8):

$$U_{\text{вх}} \Sigma_1 = \sum_{k=1}^{n_1} U_{\text{вых}} Y_k^1 = -(n_1 - 2), \quad (8)$$

то изображение  $S^*$  принадлежит множеству  $M_1$ , а конкретный класс изображений определяют выходные сигналы нейронов  $Y_1^1, \dots, Y_{n_1}^1$ . Если равенство (8) не выполняется, то изображение  $S^*$  не принадлежит множеству  $M_1$  и  $U_{\text{вых}} \Sigma_1 = 1$ . Сигнал  $U_{\text{вых}} \Sigma_1$  блокирует нейроны  $Y_1^1, \dots, Y_{n_1}^1$  и подключает к распознаванию нейроны третьей группы и т. д. В результате этого процесса входное изображение либо будет отнесено к одному из известных множеств изображений  $M_0, M_1, M_2, \dots, M_L$ , либо будет воспринято как новое изображение, которое необходимо запомнить сетью на  $(L + 1)$ -м дообучении НС.

#### 4 ЭКСПЕРИМЕНТЫ

Выполним экспериментальные исследования разработанного принципа обеспечения стабильности-пластичности НС Хебба. Для этого сначала произведем обучение НС. Продемонстрируем работоспособность предложенного подхода на НС Хебба, которая состоит из восьми  $Y$ -нейронов (рис. 4), имеющих следующую функцию активации:

$$U_{\text{вых}} Y_j = \begin{cases} 1, & \text{если } U_{\text{вх}} Y_j \geq 0; \\ -1, & \text{если } U_{\text{вх}} Y_j < 0. \end{cases}$$

Обучим изображениями  $S^1, S^2, S^3$  (рис. 5) нейроны  $Y_1, Y_2, Y_3$ . Нумерация, элементов изображений отображена на изображении  $S^0$  рис. 5. Затем продемонстрируем процесс работы обученной НС Хебба при предъявлении на ее входы изображений  $S^1, \dots, S^9$  (рис. 5).

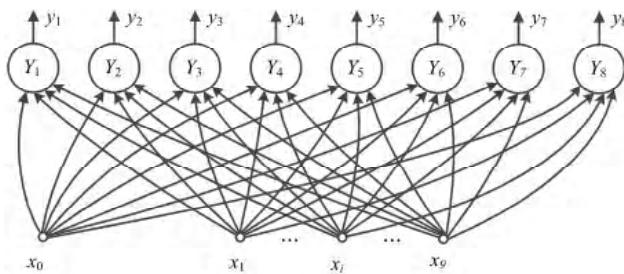


Рисунок 4 – Нейронна сеть Хебба з дополнювальними запомінаючими нейронами

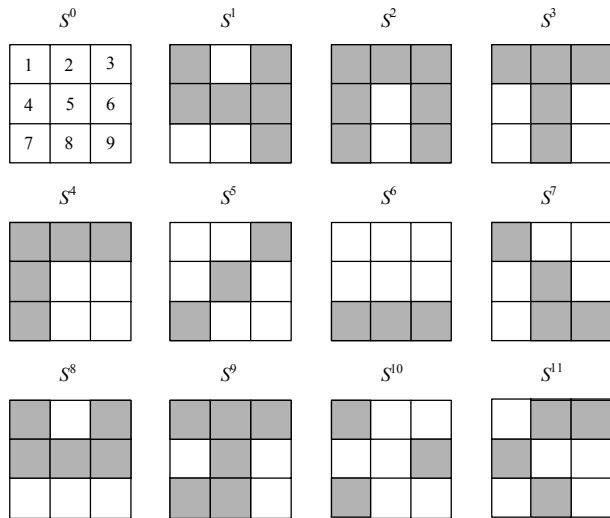


Рисунок 5 – Изображення, запоминаємі нейронною сіткою Хебба

Для обучения нейронной сети изображениями  $S^1, S^2, S^3$  необходимо выполнить следующие шаги:

*Шаг 1.* Задается множество  
 $M = \{(S^1, t^1), (S^2, t^2), (S^3, t^3)\} = \{((1, -1, 1, 1, 1, 1, 1, -1, -1,$

Таблица 1 – Результаты расчетов весов связей нейронов  $Y_1, Y_2, Y_3$  НС Хебба после обучения изображениями  $S^1, S^2, S^3$

Вес связи	$W_{01}$	$W_{11}$	$W_{21}$	$W_{31}$	$W_{41}$	$W_{51}$	$W_{61}$	$W_{71}$	$W_{81}$	$W_{91}$	$W_{02}$	$W_{12}$	$W_{22}$	$W_{32}$	$W_{42}$
Значение	-1	-1	-3	-1	1	1	1	-1	-1	1	-1	-1	1	-1	1
Вес связи	$W_{52}$	$W_{62}$	$W_{72}$	$W_{82}$	$W_{92}$	$W_{03}$	$W_{13}$	$W_{23}$	$W_{33}$	$W_{43}$	$W_{53}$	$W_{63}$	$W_{73}$	$W_{83}$	$W_{93}$
Значение	-3	1	3	-1	-1	-1	-1	1	-1	-3	1	-3	-1	3	-3

Таблица 2 – Результаты распознавания изображений  $S^1, \dots, S^9$

Нейроны	Выходные сигналы нейронов при предъявлении изображений								
	$S^1$	$S^2$	$S^3$	$S^4$	$S^5$	$S^6$	$S^7$	$S^8$	$S^9$
$Y_1$	1	-1	-1	-1	1	1	1	1	-1
$Y_2$	-1	1	-1	1	-1	1	-1	-1	-1
$Y_3$	-1	-1	1	-1	1	1	1	-1	1
Результат распознавания	+	+	+	+	-	-	-	+	+

Таблица 3 – Результаты расчетов весов связей нейронов  $Y_4, Y_5, Y_6$  после обучения изображениями  $S^5, S^6, S^7$

Вес связи	$W_{04}$	$W_{14}$	$W_{24}$	$W_{34}$	$W_{44}$	$W_{54}$	$W_{64}$	$W_{74}$	$W_{84}$	$W_{94}$	$W_{05}$	$W_{15}$	$W_{25}$	$W_{35}$	$W_{45}$
Значение	-1	-1	1	3	1	1	1	1	-3	-3	-1	-1	1	-1	1
Вес связи	$W_{55}$	$W_{65}$	$W_{75}$	$W_{85}$	$W_{95}$	$W_{06}$	$W_{16}$	$W_{26}$	$W_{36}$	$W_{46}$	$W_{56}$	$W_{66}$	$W_{76}$	$W_{86}$	$W_{96}$
Значение	-3	1	1	1	1	-1	3	1	-1	1	1	1	-3	1	1

Таблица 4 – Результаты распознавания изображений  $S^5$ ,  $S^6$ ,  $S^7$ ,  $S^{10}$  и  $S^{11}$

Нейроны	Выходные сигналы нейронов при предъявлении изображений				
	$S^5$	$S^6$	$S^7$	$S^{10}$	$S^{11}$
$Y_4$	1	-1	-1	1	1
$Y_5$	-1	1	-1	1	1
$Y_6$	-1	-1	1	-1	-1
Результат распознавания	+	+	+	-	-

Таблица 5 – Результаты расчетов весов связей нейронов  $Y_7$ ,  $Y_8$  после обучения изображениями  $S^{10}$

Вес связи	$W_{07}$	$W_{17}$	$W_{27}$	$W_{37}$	$W_{47}$	$W_{57}$	$W_{67}$	$W_{77}$	$W_{87}$	$W_{97}$
Значение	0	2	-2	-2	-2	0	2	2	-2	0
Вес связи	$W_{08}$	$W_{18}$	$W_{28}$	$W_{38}$	$W_{48}$	$W_{58}$	$W_{68}$	$W_{78}$	$W_{88}$	$W_{98}$
Значение	0	-2	2	2	2	0	-2	-2	2	0

## 6 ОБСУЖДЕНИЕ

В табл. 1 приведены результаты расчетов весов связей нейронов  $Y_1$ ,  $Y_2$ ,  $Y_3$ , после обучения сети изображениям  $S^1$ ,  $S^2$ ,  $S^3$ . Эти веса связей используются при расчете данных для следующих таблиц.

В табл. 2 представлены результаты предъявления обученной нейронной сети изображений  $S^1, \dots, S^9$ . Знаки «+» и «-» в табл. 2 означают, что соответствующее изображение распознано и не распознано, соответственно.

Из анализа табл. 2 следует, что входные изображения  $S^1, S^2, S^3, S^4, S^8, S^9$  вызывают реакцию сети, при которой при предъявлении любого из этих изображений на выходе нейронной сети наблюдается один положительный единичный сигнал и два отрицательных единичных сигнала, то есть сеть относит входное изображение к одному из известных классов. Анализ данных табл. 2 изображений показывает, что произошло правильное распознавание всех шести изображений, несмотря на то, что три из них имеют дефекты ( $S^4, S^8, S^9$ ).

Нетрудно увидеть, что при отнесении входного изображения к одному из трех классов, представители которых хранятся в весах связей нейронов  $Y_1$ ,  $Y_2$ ,  $Y_3$ , суммарный выходной сигнал этих нейронов равен минус единице:  $\sum_{j=1}^3 U_{\text{вых}} Y_j = -1$ .

Если входное изображение не относится к классам изображений хранящихся в весах связей нейронов  $Y_1$ ,  $Y_2$ ,  $Y_3$ , то имеем  $\sum_{j=1}^3 U_{\text{вых}} Y_j \neq -1$ , что служит признаком

новой информации (нового изображения) на входном слое нейронов и необходимости обучения следующей группы  $Y$ -нейронов распознавать эту входную информацию. При этом обученные распознавающие нейроны должны быть заблокированы в режиме распознавания и заблокированы или игнорироваться в режиме обучения следующей группы нейронов.

Нейроны  $Y_4$ ,  $Y_5$ ,  $Y_6$  несложно обучить распознаванию изображений  $S^5$ ,  $S^6$ ,  $S^7$  алгоритмом Хебба. Результат обучения приведен в табл. 3. При этом множество M1 из обучающих пар задавалось следующим образом:

$$M = \{(S^5, t^5), (S^6, t^6), (S^7, t^7)\} = \{((-1, -1, 1, -1, 1, -1, 1, -1, -1), (1, -1, -1)), ((-1, -1, -1, -1, -1, 1, 1, 1), (-1, 1, -1)), ((1, -1, -1, -1, 1, -1, 1, 1), (-1, -1, 1))\}.$$

В табл. 4 приведены результаты распознавания изображений  $S^5$ ,  $S^6$ ,  $S^7$ ,  $S^{10}$  и  $S^{11}$ . Знаки «+» и «-» в табл. 4 означают, что соответствующее изображение распознано и не распознано, соответственно.

Из анализа табл. 4 следует, что нейроны  $Y_4$ ,  $Y_5$ ,  $Y_6$  относят изображения  $S^5$ ,  $S^6$ ,  $S^7$  к известным классам, эталоны которых хранятся в весах связей этих нейронов. Изображения  $S^{10}$  и  $S^{11}$  нейронами  $Y_4$ ,  $Y_5$ ,  $Y_6$  распознаются как не принадлежащие к этим классам, это же наблюдается и при предъявлении изображений  $S^{10}$ ,  $S^{11}$  и первой группе нейронов  $Y_1$ ,  $Y_2$ ,  $Y_3$ . Таким образом, для распознавания изображений  $S^{10}$ ,  $S^{11}$  необходимо обучить еще два  $Y$ -нейрона:  $Y_7$ ,  $Y_8$ . В табл. 5 приведены результаты расчетов весов связей нейронов  $Y_7$ ,  $Y_8$  после предъявления изображений  $S^{10}$ ,  $S^{11}$ .

Нейроны  $Y_7$ ,  $Y_8$  будут распознавать указанные изображения после того, как при попытке их распознавания будут заблокированы сначала нейроны  $Y_1$ ,  $Y_2$ ,  $Y_3$ , а затем нейроны  $Y_4$ ,  $Y_5$ ,  $Y_6$ .

## ВЫВОДЫ

В работе решена актуальная задача разработки новых классов стабильно-пластичных нейронных сетей на основе нейронных сетей Хемминга и Хебба.

Научная новизна работы состоит в следующем:

- впервые для нейронной сети, использующей при распознавании входного изображения расстояние Хемминга, предложена архитектура и алгоритмы функционирования, обеспечивающие, в отличии от сети Хемминга, свойство дообучения нейронной сети и распознавание новой информации на своих входах;

– вперше для нейронної мережі, обучаючоїся з допомогою правила Хебба, предложена архітектура та алгоритми функціонування, що забезпечують, в порівнянні з відомою нейронною мережею Хебба, стабільність-пластичність нейронної мережі та розпізнавання нової інформації на своїх входах;

– вперше для нейронної мережі, обучаючоїся з допомогою правила Хебба, предложена модульна архітектура, що забезпечує можливість многократного дообучення нейронної мережі.

Практическа цінність отриманих результатів полягає в тому, що розроблені архітектури нейронних мереж та алгоритми їх функціонування відкривають можливість для створення систем розпізнавання, класифікації, управління та підтримки приняття рішень для об'єктів в умовах існування априорної неопределеності, коли планирується оперативне дообучення розроблених систем по мере накопичення інформації про об'єкти.

Перспектива дальніших дослідженням полягає в розробці нових класів стабільно-пластичних нейронних мереж, в частності, перцепtronів з будь-яким числом шарів нейронів, що мають неперервні функції активування.

## БЛАГОДАРНОСТИ

Робота виконана в рамках науково-дослідницьких робіт кафедри «Висуточні методи та програмування» Національного технічного університету «Харківський політехнічний інститут» в згідності з госбюджетною прикладною темою за планом Міністерства освіти та науки України «Розробка інтелектуальних систем підтримки приняття рішень для діагностики, управління та оптимізації техніческих та біотехніческих об'єктів» (ГР № 0113U000449). Автори висловлюють свою вдячність МОН України та НТУ «ХПІ» за фінансову підтримку проводимих дослідженій.

Дмитренко В. Д.<sup>1</sup>, Заковоротний О. Ю.<sup>2</sup>

<sup>1</sup>Д-р техн. наук, професор, професор кафедри обчислювальної техніки та програмування Національного технічного університету «Харківський політехнічний інститут», Україна

<sup>2</sup>Канд. техн. наук, доцент, доцент кафедри обчислювальної техніки та програмування Національного технічного університету «Харківський політехнічний інститут», Україна

## АРХІТЕКТУРИ ТА АЛГОРІТМИ ФУНКЦІОНАВАННЯ НЕЙРОННИХ МЕРЕЖ ХЕММІНГА І ХЕББА, ЗДАТНИХ ДОНАВЧАТИСЯ Й РОЗПІЗНАВАТИ НОВУ ІНФОРМАЦІЮ

Вирішена задача донавчання класичних дискретних нейронних мереж Хеммінга та Хебба без втрат інформації, що вже була запам'ятована. Об'єктом дослідження є процеси розпізнавання та класифікації образів в системах, що побудовані на основі штучних нейронних мереж. Предметом дослідження є архітектури та алгоритми функціонування штучних нейронних мереж. Мета роботи: розробка стабільно-пластичних нейронних мереж Хеммінга та Хебба. Розроблені архітектури та алгоритми функціонування дискретних стабільно-пластичних нейронних мереж Хеммінга та Хебба, які не тільки можуть донавчатися в процесі функціонування, а й розпізнавати нову інформацію. Нові мережі можуть стати альтернативою дискретним нейронним мережам адаптивної резонансної теорії. Розроблений підхід по донавчанню може бути узагальнений й на інші нейронні мережі. Проведено експериментальні дослідження розроблених алгоритмів функціонування штучних нейронних мереж. Результати експериментів підтверджують правильність запропонованого підходу.

**Ключові слова:** розпізнавання та класифікація образів, стабільно-пластичні нейронні мережі, нейронна мережа Хеммінга, нейронна мережа Хебба, адаптивна резонансна теорія.

## СПИСОК ЛИТЕРАТУРЫ

- Suzuki K. Artificial Neural Networks: Architectures and Applications / K. Suzuki. – Publisher: InTech, 2013. – 256 p.
- Bianchini M. Handbook on Neural Information Processing (Intelligent Systems Reference Library) / M. Bianchini. – Publisher : Springer, 2013. – 499 p.
- Cirrincione M. Power Converters and AC Electrical Drives with Linear Neural Networks (Energy, Power Electronics, and Machines) / M. Cirrincione, M. Pucci, G. Vitale. – Publisher : CRC Press, 2012. – 631 p.
- Галушкин А. И. Нейронные сети. Основы теории / А. И. Галушкин. – М.: Горячая линия. – Телеком, 2012. – 496 с.
- Капля В. И. Системы искусственного интеллекта : учебное пособие. – Волгоград : ИУНЛ ВолгГТУ. – 2011. – 97 с.
- Russell S. Artificial Intelligence: A Modern Approach, Third Edition / S. Russell, P. Norvig. – Publisher : Prentice Hall, 2010. – 1152 p.
- Девятков В. В. Системы искусственного интеллекта / Гл. ред. И. Б. Федоров. – М.: Изд-во МГТУ им. Н. Э. Баумана, 2001. – 352 с.
- Carpenter G. A. Massively parallel architecture for self-organising neural pattern recognition machine/G. A. Carpenter, S. A. Grossberg // Computing, Vision, Graphics and Image Processing. – 1987. – Vol. 37. – P. 54–115.
- Grossberg S. Competitive learning: From interactive activation to adaptive resonance / S. Grossberg // Cognitive Science. – 1987. – Vol. 11. – P. 23–63.
- Fausett L. Fundamentals of Neural Networks. Architectures, Algorithms and Applications / L. Fausett. – New Jersey : Prentice Hall Int., Inc., 1994. – 461 p.
- Дмитренко В. Д. Нейросетевое устройство направленных ассоциаций / В. Д. Дмитренко, А. Ю. Заковоротный, Хавина И. П. // Научные ведомости. Серия : История, политология, экономика, информатика. – Белгород : БГУ, 2010. – № 7(78), Вып. 14/1. – С. 110–119.
- Дмитренко В. Д. Ассоциативная нейронная сеть АРТ / В. Д. Дмитренко, А. Ю. Заковоротный, В. А. Бречко // Сборник трудов Международной молодежной конференции «Прикладная математика, управление и информатика». – Белгород : ИД «Белгород», 2012. – Т. 1. – С. 115–118.

Статья поступила в редакцию 18.11.2014.  
После доработки 17.10.2014.

Dmitrienko V. D.<sup>1</sup>, Zakovorotnyi A. Yu.<sup>2</sup>

<sup>1</sup>Dr. Sc., Professor, Professor of department Computer Science and Programming, National Technical University «Kharkiv Polytechnic Institute», Ukraine

<sup>2</sup>PhD, Associate Professor, Associate Professor of department Computer Science and Programming, National Technical University «Kharkiv Polytechnic Institute», Ukraine

## ARCHITECTURE AND ALGORITHMS OF NEURAL NETWORKS HAMMING AND HEBB, CAPABLE LEARN AND IDENTIFY NEW INFORMATION

The problem of the classical discrete neural networks Hamming and Hebb lossless previously stored information additional training. The object of research is the process of recognition and classification of images on systems that are based on artificial neural networks. The subject of research is the architecture and algorithms of artificial neural networks. Objective: To develop a stable and plastic neural networks Hamming and Hebb. The architecture and algorithms of discrete stable and plastic neural networks Hamming and Hebb, which not only can be trained during functioning, but also to recognize the new information. New networks can be an alternative to discrete neural network adaptive resonance theory. The developed approach for training can be generalized to other neural networks. Experimental investigations of the functioning of the developed algorithms of artificial neural networks. The experimental results confirm the validity of the proposed approach.

**Keywords:** recognition and classification of images, stable and plastic neural networks, Hamming neural network, Hebb neural network, adaptive resonance theory.

## REFERENCES

1. Suzuki K. Artificial Neural Networks: Architectures and Applications. Publisher, InTech, 2013, 256 p.
2. Bianchini M. Handbook on Neural Information Processing (Intelligent Systems Reference Library). Publisher, Springer, 2013, 499 p.
3. Cirrincione M., Pucci M., Vitale G. Power Converters and AC Electrical Drives with Linear Neural Networks (Energy, Power Electronics, and Machines). Publisher, CRC Press, 2012, 631 p.
4. Galushkin A. I. Nejronnye seti. Osnovy teorii. Moscow, Gorjachaja linija, Telekom, 2012, 496 p.
5. Kaplja V. I. Sistemy iskusstvennogo intellekta: uchebnoe posobie. Volgograd, IUNL VolgGTU, 2011, 97 p.
6. Russell S., Norvig P. Artificial Intelligence: A Modern Approach, Third Edition. Publisher, Prentice Hall, 2010, 1152 p.
7. Devyatkov V. V. Gl. red. I. B. Fjodorov. Sistemy iskusstvennogo intellekta. Moscow, Izd-vo MGTU im. N. Je. Baumana, 2001, 352 p.
8. Carpenter G. A., Grossberg S. A. Massively parallel architecture for self-organising neural pattern recognition machine, *Computing, Vision, Graphics and Image Processing*, 1987, Vol. 37, P. 54–115.
9. Grossberg S. Competitive learning: From interactive activation to adaptive resonance, *S Cognitive Science*, 1987, Vol. 11, pp. 23–63.
10. Fausett L. Fundamentals of Neural Networks. Architectures, Algorithms and Applications. New Jersey, Prentice Hall Int., Inc., 1994, 461 p.
11. Dmitrienko V. D., Zakovorotnyi A. Y., Havina I. P. Nejrosetevoe ustroystvo napravlennyh associacij, *Nauchnye vedomosti. Serija: Istorija, politologija, jekonomika, informatika*. Belgorod, BGU, 2010, No. 7(78), Vyp. 14/1, pp. 110–119.
12. Dmitrienko V. D., Zakovorotnyi A. Yr., Brechko V. A. Asociativnaja nejronnaja set' ART, *Sbornik trudov Mezhdunarodnoj molodezhnoj konferencii «Prikladnaja matematika, upravlenie i informatika»*. Belgorod, ID «Belgorod», 2012, Vol. 1, pp. 115–118.

УДК 004.93

Олійник А. О.

Канд. техн. наук, доцент, Запорізький національний технічний університет, Україна

## ВИДОБУВАННЯ ЗНАНЬ НА ОСНОВІ ДЕРЕВ РОЗВ'ЯЗКІВ ТА СТОХАСТИЧНОГО ПОШУКУ

Вирішено завдання розробки математичного забезпечення для автоматизації видобування набору знань у вигляді продукційних правил з навчальних вибірок даних. Об'єктом дослідження є процес побудови моделей неруйнівного контролю якості. Предмет дослідження становлять методи видобування продукційних правил для синтезу моделей контролю якості. Мета роботи: підвищити ефективність процесу виявлення продукційних правил для побудови моделей контролю якості на основі навчальних вибірок. Запропоновано стохастичний метод синтезу дерев розв'язків, який використовує інформацію про інформативність ознак, складність синтезованого дерева, а також точність його розпізнавання, що дозволяє на початковому етапі формувати множину деревовидних структур, що характеризуються простою ієрархією і невисокою помилкою розпізнавання, в процесі пошуку створювати нові множини рішень з урахуванням інформації про значущість ознак та інтерпретабельність створюваних дерев, що, у свою чергу, забезпечує можливість побудови дерев розв'язків з невеликою кількістю елементів (вузлів та зв'язків між ними) і прийнятною точністю розпізнавання, а також видобування на його основі найбільш цінних екземплярів. Розроблено програмне забезпечення, що реалізує запропонований метод. Проведено експерименти з дослідження властивостей запропонованого методу. Результати експериментів дозволяють рекомендувати запропонований метод для використання на практиці.

**Ключові слова:** вибірка, дерево розв'язків, модель контролю якості, продукційне правило, стохастичний пошук.

### НОМЕНКЛАТУРА

- $Br(\chi_k)$  – гіллястість дерева розв'язків  $\chi_k$ ;
- $Cl_c$  –  $c$ -й кластер;
- $d_{ab}$  –  $b$ -й вузол  $a$ -го рівня синтезованого дерева розв'язків;
- $E$  – похибка моделі контролю якості;
- $G(\chi_k)$  – значення цільової функції  $k$ -го розв'язку при стохастичному пошуку;
- $Gener(\chi_k)$  – узагальнення дерева розв'язків  $\chi_k$ ;
- $GenerM(\chi_k)$  – узагальнення моделі дерева розв'язків  $\chi_k$ ;
- $Idist(\chi_k)$  – внутрілистова відстань дерева розв'язків  $\chi_k$ ;
- $Int(\chi_k)$  – інтерпретовність дерева розв'язків  $\chi_k$ ;
- $M$  – кількість атрибутів;
- $N_{branch}(\chi_k)$  – кількість гілок дерева розв'язків  $\chi_k$ ;
- $N_{er}$  – кількість неправильно розпізнаних спостережень вибірки  $S$ ;
- $N_{level}(\chi_k)$  – кількість рівнів дерева розв'язків  $\chi_k$ ;
- $N_{node}(\chi_k)$  – кількість вузлів дерева розв'язків  $\chi_k$ ;
- $N_\chi$  – кількість розв'язків на кожній ітерації стохастичного пошуку;
- $P$  – набір характеристик (ознак) спостережень;
- $p_{mab}$  – ознака-перевірка у вузлі  $d_{ab}$  дерева розв'язків;

- $p_{qm}$  – значення  $m$ -го атрибуту  $q$ -го спостереження;
- $p_{TRab}$  – граничне значення ознаки-перевірки вузла  $d_{ab}$ ;
- $Q$  – кількість спостережень;
- $R\sigma_m$  – значення рангу  $m$ -ї ознаки  $p_m$ ;
- $RB$  – база правил;
- $rule_r$  –  $r$ -те правило бази правил;
- $S$  – навчальна вибірка;
- $T$  – множина значень вихідного параметру;
- $t_q$  – значення вихідного параметру  $q$ -го спостереження;
- $V_m$  – оцінка індивідуальної інформативності  $m$ -ї ознаки;
- $VE(p_m)$  – ентропія ознаки  $p_{qm}$ ;
- $\sigma_{mc}$  – ширина розкиду значень ознаки  $p_m$  в  $c$ -му кластері;
- $\chi_k$  –  $k$ -й розв'язок стохастичного пошуку.

### ВСТУП

Побудова автоматизованих систем неруйнівного контролю якості пов'язана з необхідністю синтезу моделей прийняття рішень [1]. Як базис для побудови таких моделей ефективно можуть використовуватися нейро-нечіткі мережі [2–4], які є гібридною моделлю обчислювально-го інтелекту, що характеризується високою інтерпретовністю та сполучає у собі властивості систем, заснованих на знаннях, і однорідних обчислювальних структур.

Процес синтезу таких моделей пов'язаний з необхідністю видобування правил на основі заданих вибірок даних. Однак вибірки даних, що описують результати вимірювань характеристик реальних технічних об'єктів і процесів можуть містити дубляж інформації, зокрема, надлишкові для прийняття рішень ознаки й екземпляри [5, 6]. Крім того, можливі ситуації, при яких у вихідній вибірці кількість екземплярів одного класу істотно відрізняється від кількості екземплярів іншого класу (при використанні традиційного навчання екземпляри одного класу можуть пригнічувати екземпляри іншого класу) [5–7]. Отже, застосування відомих методів видобування продукційних правил для синтезу моделей контролю якості на основі нейро-нечітких мереж у деяких випадках є недоцільним.

Тому у цій роботі пропонується на основі заданих вибірок даних синтезувати дерева розв'язків і видобувати на їх основі продукційні правила, що дозволить виділяти найцінніші екземпляри, ранжурувати ознаки за значущістю й, отже, усувати деяку надлишковість інформації, а також скоротити простір пошуку й час синтезу нейро-нечітких моделей контролю якості.

Однак відомі методи синтезу дерев розв'язків [8–11] передбачають використання «жадібного» підходу, що не дозволяє в процесі побудови таких моделей повторно розглядати ознаки  $P_m$ , за якими вже було виконано розбиття. Це може привести до низьких узагальнюючих властивостей синтезованої моделі, внаслідок її складності, а, отже, до надлишкового числа правил, витягнутих з неї, що зробить систему правил більше громіздкою та менш інтерпретованою.

Тому в цій роботі для побудови дерев розв'язків пропонується використовувати інтелектуальний стохастичний пошук [11–13], що дозволяє досліджувати різні області пошукового простору й не використовує жадібну стратегію.

Метою роботи є створення методу видобування знань у вигляді продукційних правил на основі дерев розв'язків і стохастичного пошуку.

## 1 ПОСТАНОВА ЗАДАЧІ

Нехай задана множина спостережень  $S = \langle P, T \rangle$ , де  $P$  – набір характеристик (ознак) спостережень,  $T$  – множина значень вихідного параметру. Набори значень  $P$  і  $T$  можуть бути подані у вигляді матриці (1) та вектора (2), відповідно:

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} & \dots & p_{1M} \\ p_{21} & p_{22} & \dots & p_{2m} & \dots & p_{2M} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{q1} & p_{q2} & \dots & p_{qm} & \dots & p_{qM} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{Q1} & p_{Q2} & \dots & p_{Qm} & \dots & p_{QM} \end{pmatrix}, \quad (1)$$

$$T = (t_1 \ t_2 \ \dots \ t_q \ \dots \ t_Q)^{\text{tr}}. \quad (2)$$

Тоді задача видобування продукційних правил полягає в пошуку такого набору правил  $RB = \{rule_1, rule_2, \dots, rule_{NR}\}$ , при якому забезпечується прийнятне значення заданого критерію якості  $G$ , де кожне  $r$ -е правило  $rule_r$  являє собою імплікацію антецедента (набору ознак  $P_m$  та їх граничних значень  $p_{TR}$ ) та консеквента (значення  $t_q$  вихідного параметру  $T$  при виконанні умов, поданих в антецеденті). Як цільовий критерій  $G$  при видобуванні продукційних правил можуть бути використані, наприклад:

– похибка розпізнавання (у задачах з дискретним виходом  $T$ ) [1–3, 14], що обчислюється за формулою:  $E = N_{er} / Q$ ;

– середньоквадратична похибка (у випадку, коли вихідний параметр  $T$  може приймати дійсні значення з деякого діапазону  $T \in [t_{\min}; t_{\max}]$ ) [1–3, 14], що розраховується за формулою:  $E = \frac{1}{Q} \sum_{q=1}^Q (t_q - t_{q \text{ mod}})^2$ , де  $t_{q \text{ mod}}$  – значення вихідного параметру  $q$ -го спостереження, розраховане за набором правил  $RB$ .

## 2 ОГЛЯД ЛІТЕРАТУРИ

Основні підходи до формування бази правил на основі вибірок даних  $S$  для синтезу моделей контролю якості на основі нейро-нечітких систем полягають у наступному [14, 15]:

– копіювання навчальної вибірки в базу знань – для кожного екземпляра навчальної вибірки формується окреме правило. Перевагою даного методу є простота та висока швидкість роботи, недоліком – відсутність узагальнюючих властивостей і громіздкість одержуваної мережі;

– оптимізація кількості продукційних правил – знаходження такого значення кількості продукційних правил  $NR$ , при якій значення помилки  $E$  є мінімальним, для чого при різних значеннях  $NR$  навчають мережу і вимірюють значення помилки, після чого оптимізують функцію  $E(NR)$  за параметром  $S$ . Недоліком даного методу є дуже високі вимоги до обчислювальних ресурсів, обумовлені необхідністю заново навчати мережу на кожному кроці;

– спільна оптимізація ваг мережі та кількості продукційних правил шляхом вирішення багатоекстремальної оптимізаційної задачі або автоматичне визначення числа кластерів у навчальній вибірці та встановлення центрів функцій приналежності в їхні центри на основі кластер-аналізу;

– скорочення (редукція) правил. При цьому підході виключаються суперечливі правила, які взаємно компенсуються, а також одне з двох співпадаючих правил, як такі, що не несуть нової інформації. При скороченні видаляються ті продукційні правила, вплив яких на точність виявляється мінімальним після оцінки індивідуального

внеску кожного продукційного правила у вихідний сигнал мережі, одержуваної шляхом використання ортогонального методу найменших квадратів. Істотним недоліком методів скорочення є необхідність спочатку працювати зі свідомо надлишковою за розміром базою знань, що обумовлює в ряді випадків повільну роботу методів.

– нарощування (конструювання) правил: формується початкова база продукційних правил (вона може бути і порожньою), що потім послідовно поповнюється нечіткими правилами. Недоліком даного методу є відсутність явного зв'язку між процедурою додавання продукційних правил і точністю апроксимації, що повинна визначатися окремо.

Наявність зазначених недоліків обумовлює необхідність розробки нових методів побудови бази продукційних правил. Тому у цій роботі пропонується видобувати продукційні правила на основі дерев розв'язків [8–11], побудованих на основі заданих вибірок даних  $S$ .

Проте відомі методи ідентифікації дерев рішень, зокрема ID3, CART, CHAID, QUEST, C4.5 [8–11], мають певні недоліки, пов'язані з великою обчислювальною складністю, проблемами формування дерева рішень (ріст дерева, відсікання частини дерева) і т. ін. [8–10]. Крім того, такі методи використовують жадібну стратегію пошуку: якщо ознака була обрана один раз, і за нею виконано розбиття на підмножини, то метод не може повернутися назад і вибрати інший атрибут, який привів би до кращого розбиття, внаслідок чого в результаті часто синтезу-

ються дерева розв'язків, що не забезпечують прийнятий рівень апроксимації [9]. Тому в даній роботі запропоновано для синтезу дерев розв'язків використовувати інтелектуальний стохастичний пошук [11–14], що дозволяє досліджувати різні області пошукового простору і не використовує жадібну стратегію.

### 3 МАТЕРІАЛИ ТА МЕТОДИ

Для синтезу дерев розв'язків пропонується використовувати інформацію про інформативність ознак, складність синтезованого дерева, а також точність його розпізнавання. Це дозволить на початковому етапі формувати множину деревоподібних структур, що характеризуються простою ієархією й невисокою похибкою розпізнавання, у процесі пошуку створювати нові множини розв'язків з урахуванням інформації про значущість ознак й інтерпретованість створюваних дерев, що, у свою чергу, забезпечить можливість побудови дерева розв'язків з невеликою кількістю елементів (вузлів і зв'язків між ними) і прийнятною точністю розпізнавання, а також видобування на його основі найцінніших екземплярів.

На етапі ініціалізації при синтезі дерев розв'язків пропонується генерувати початкову множину розв'язків  $R^{(0)} = \{\chi_1^{(0)}, \chi_2^{(0)}, \dots, \chi_{N\chi}^{(0)}\}$ , де  $N\chi$  – кількість розв'язків у множині  $R^{(0)}$ .

Кожен  $k$ -й розв'язок  $\chi_k$  являє собою структуру, що відповідає певному дереву розв'язків (рис. 1).

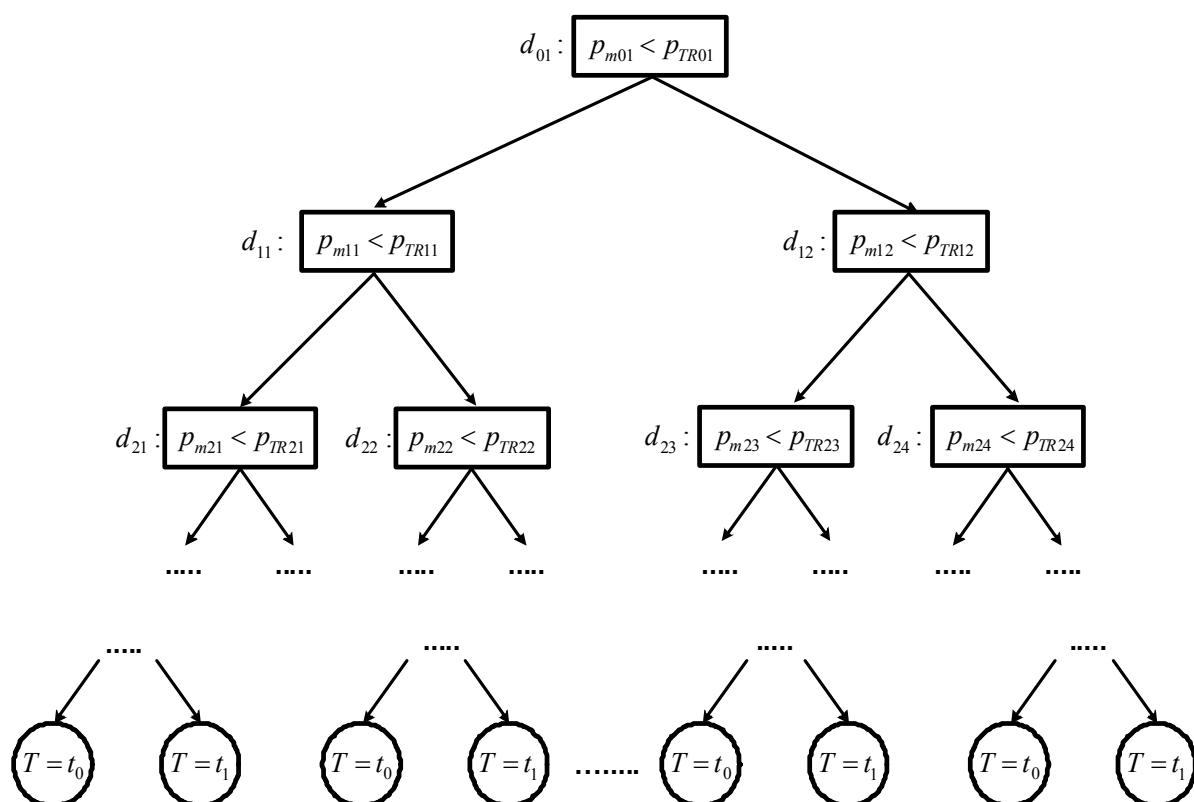


Рисунок 1 – Схематичне подання структури  $\chi_k$  при синтезі дерев розв'язків

На рис. 1 позначення  $T = t_a$  відповідає значенню  $t_a$  вихідного параметра  $T$  при проході від кореня дерева до одного з його листів (кінцевих вузлів, що містять значення вихідного параметра при виконанні умов, що знаходяться у вищестоячих вузлах-батьках).

Як видно з рис. 1, при синтезі кожного дерева розв'язків  $\chi_k$  необхідно визначити його структуру, що представляє собою набір взаємозалежних вузлів  $d_{ab}$ , які містять інформацію про ознаку  $p_{mab}$ , за якою відбувається розбиття, її граничне значення  $p_{TRab}$ , а також посилання на лівого  $d_{(a+1)(2b-1)}$  й правого  $d_{(a+1)(2b)}$  нашадків:  $d_{ab} = \langle p_{mab}, p_{TRab}, d_{(a+1)(2b-1)}, d_{(a+1)(2b)} \rangle$ . Отже, для побудови дерева розв'язків  $\chi_k$  потрібно сформувати множину взаємозалежних вузлів-перевірок  $d_{ab}$ , визначивши для кожного з них ознаку-перевірку  $p_{mab}$  та її граничне значення  $p_{TRab}$ .

При виборі ознаки-перевірки  $p_{mab}$  для вузла  $d_{ab}$  будемо використовувати апріорну інформацію про значущість ознак  $V_m$ . Оскільки дерева розв'язків ефективно застосовуються, як правило, для вирішення задач розпізнавання, які характеризуються кінцевою кількістю класів вихідного параметру  $T$ , для оцінювання інформативності  $V_m$  ознак доцільно використовувати характеристики, що дозволяють оцінювати значущість ознаки  $p_m$  стосовно вихідного параметра  $T$ , який приймає дискретні значення  $t_q$  з кінцевою множиною. Як такі характеристики можуть бути використані наступні критерії [6, 8, 14, 15]:

– ентропія ознаки  $VE(p_m)$  – критерій, що відображає ступінь невизначеності стану об'єкта [8, 14], розраховується за формулою:

$$VE(p_m) = - \sum_{n=1}^{N_{int}(p_{mn})} \left( \rho(p_{mn}) \sum_{l=1}^{N_{int}(T)} \rho(p_{mn}, T_l) \log_2 \rho(p_{mn}, T_l) \right),$$

де  $\rho(p_{mn}) = \frac{N(p_{mn})}{M}$  – ймовірність того, що значення

ознаки  $p_m$  екземплярів вибірки  $S$  потрапить до  $n$ -го інтервалу діапазону її зміни;  $N(p_{mn})$  – кількість екземплярів вибірки  $S$ , значення  $m$ -ї ознаки яких, належать  $n$ -му інтервалу діапазону її зміни;  $N_{int}(p_{mn})$  – кількість інтервалів, на які розбивається діапазон значень  $m$ -ї ознаки  $p_m$ ;  $N_{int}(T)$  – кількість інтервалів, на які розбивається діапазон значень вихідного параметра  $T$ ;

$\rho(p_{mn}, T_l) = \frac{N(p_{mn}, T_l)}{N(p_{mn})}$  – умовна ймовірність того, що

значення вихідного параметра  $T$  потрапить в  $l$ -й інтервал  $T_l$  за умови, що  $m$ -а ознака  $p_m$  потрапить в  $n$ -й інтервал  $p_{mn}$ ;  $N(p_{mn}, T_l)$  – кількість екземплярів вибірки  $S$ , значення вихідного параметра  $T$  яких належать  $l$ -му інтервалу діапазону його зміни  $T_l$  за умови, що значення їх  $m$ -ї ознаки належить  $n$ -му інтервалу  $p_{mn}$ ;

– теоретико-інформаційний критерій  $VT(p_m)$  – передбачає використання кількості інформації, що одержує система в процесі розпізнавання об'єктів у результаті використання оцінюваної ознаки:

$$VT(p_m) = \sum_{l=1}^{N_{int}(T)} \sum_{n=1}^{N_{int}(p_{mn})} \rho(p_{mn}, T_l) \log_2 \frac{\rho(p_{mn}, T_l)}{\rho(p_{mn}) \rho(T_l)},$$

де  $\rho(T_l)$  – ймовірність того, що значення вихідного параметра  $T$  потрапить в  $l$ -й інтервал  $T_l$  діапазону зміни його значень.

Однак застосування таких критеріїв передбачає, що відомими є всі значення всіх екземплярів у навчальній вибірці  $S$ , що є не завжди можливим при вирішенні практичних задач діагностування й управління якістю продукції (це обумовлено можливими проблемами при вимірюванні параметрів деяких реальних технічних об'єктів або процесів). У випадку, якщо деякі екземпляри навчальної вибірки  $S$  містять пропущені значення ознак  $p_m$  або вихідного параметра  $T$ , пропонується використовувати наступний підхід до оцінювання індивідуальної значущості  $V_m$ . Оцінювати інформативність ознак будемо виходячи з їхньої значущості для визначення границь кластерів – груп компактно розташованих екземплярів у просторі ознак. Для цього за допомогою методів кластерного аналізу [7, 14] пропонується виявляти групи екземплярів (кластери) виходячи з їх геометричного розташування.

Будемо вважати, що ознака  $p_m$  є тим важливішою для кластера  $Cl_c$  ( $c = 1, 2, \dots, N_{cl}$ ), чим меншою є ширина розкиду  $\sigma_{mc}$  її значень у цьому кластері, що розраховується

за формулою:  $\sigma_{mc} = \sqrt{\sum_{q=1}^{N_{inst}(Cl_c)} (p_{mq} - \bar{p}_{mc})^2}$ , де  $\bar{p}_{mc}$  –

середнє значення  $m$ -ї ознаки  $p_m$  в  $c$ -му кластері;  $N_{inst}(Cl_c)$  – кількість екземплярів  $c$ -го кластеру  $Cl_c$ .

Потім найбільш значущій ознаці  $c$ -го кластера (ознаці з мінімальним значенням величини  $\sigma_{mc}$ ) будемо ставити у відповідність найбільше значення рангу  $R\sigma_{mc} = M$ , наступній за зростанням величини  $\sigma_{mc}$  ознаці привласнимо значення  $R\sigma_{mc} = M - 1$  і т. д. При однакових значеннях величини  $\sigma_{mc}$  ознакам ставляться у відповідність середні значення рангів  $R\sigma_{mc}$ . Ознакам, що характеризуються низькими значеннями індивідуальної інформативності для екземплярів  $c$ -го кластеру ( $\sigma_{mc} < \sigma_{mc \min}$ ), ставиться у відповідність нульове значення величини  $R\sigma_{mc}$ :  $R\sigma_{mc} = 0$ .

Значення загального рангу  $R\sigma_m$   $m$ -ї ознаки  $p_m$  по всіх  $N_{cl}$  кластерах визначимо як суму значень рангів  $R\sigma_{mc}$  за

формулою:  $R\sigma_m = \sum_{c=1}^{N_{cl}} R\sigma_{mc}$ . Оцінку індивідуальної інформативності  $V_m$   $m$ -ї ознаки визначимо за формулою:

$$V_m = \frac{R\sigma_m}{\max_{n=1,2,\dots,M} \{R\sigma_n\}}.$$

Таким чином, ознака з максимальним значенням рангу  $R\sigma_m$  є найбільш інформативною і характеризується одиничним значенням критерію  $V_m$ . Використання запропонованого критерію дозволяє ранжирувати ознаки, виходячи з їхньої значущості для опису границь компактного розташування екземплярів, що, у свою чергу, дозволяє оцінювати індивідуальну інформативність ознак у вибірках, у яких деякі екземпляри містять пропущені значення ознак або вихідного параметра.

Після оцінювання інформативності  $V_m$  кожної ознаки  $p_m$  вибірки  $S$  відбувається формування дерев  $\chi_k$  у вигляді відповідних структур даних. Ознаки  $p_m$  з високими оцінками індивідуальної інформативності  $V_m$  є більш значущими (істотно впливають на вихідний параметр  $T$ ), отже, у запропонованому методі такі ознаки будуть мати більшу ймовірність відбору як ознаки-перевірки  $p_{mab}$  для відповідного вузла  $d_{ab}$ .

Нехай оцінка індивідуальної інформативності  $m$ -ї ознаки  $p_m$  дорівнює  $V_m$ . Тоді нормована оцінка індивідуальної інформативності ознак  $V_{m\_norm}$  у вибірці  $S$  може

бути розрахована за формулою:  $V_{m\_norm} = \frac{V_m - V_{min}}{V_{max} - V_{min}}$ ,

де  $V_{min}$  та  $V_{max}$  – мінімальне та максимальне значення інформативності  $V_m$  ознак  $p_m$  у вибірці  $S$ , відповідно.

Вибір ознак  $p_m$  для використання у вузлі-перевірці  $d_{ab}$  будемо виконувати виходячи зі значення величини  $V_{m\_norm}$ . Для цього будемо послідовно переглядати ознаки  $p_1, p_2, \dots, p_M$  та порівнювати значення величини  $V_{m\_norm}$  з  $rand[0; 1]$  – випадково згенерованим числом в інтервалі  $[0; 1]$ . У випадку, якщо виконується умова  $V_{m\_norm} \geq rand[0; 1]$ , ознака  $p_m$  вважається добре поділяючою екземпляри на класи та включається в поточний вузол  $d_{ab}$  дерева розв'язків  $\chi_k$  як ознака-перевірка.

Далі для ознаки-перевірки  $p_{mab}$  у вузлі  $d_{ab}$  виконується визначення граничного значення  $p_{TRab}$ . Для цього розраховуються значення похибки розпізнавання  $E_{ab}$  екземплярів  $S_{ab}$ , що попадають у вузол  $d_{ab}$ , при різних значеннях  $p_{TRab} \in [p_{mab\ min}; p_{mab\ max}]$ , і вибирається таке значення  $p_{TRab}$ , при якому значення похибки розпізнавання  $E_{ab}$  буде найменшим. У випадку, якщо знайдено значення  $p_{TRab}$ , при якому похибка розпізнавання знаходиться в припустимих межах  $E_{ab} < E_{TR}$ , тоді вважається, що ознака  $p_{mab}$  при граничному значенні  $p_{TRab}$  дозволяє здійснювати прийнятне розбиття множини екземплярів  $S_{ab}$  на класи  $t_0$  та  $t_1$ , отже, її нащадками будуть вузли, що представляють собою листи дерева розв'язків  $\chi_k$  – кінцеві вузли, що містять значення вихідного параметра  $T = t_0$  та  $T = t_1$ . Якщо в результаті розбиття  $p_{mab} < p_{TRab}$  в одну з частин дерева ввийдуть екземпляри  $S_{ab}$  з єдиним значенням вихідного параметра (наприклад,  $T = t_1$ ), то нащадок, що відповідає даній умові, стає листом, і розбиття триває тільки для другого нащадка вузла  $d_{ab}$ .

Формування дерева розв'язків  $\chi_k$  триває доти, поки не буде досягнуто прийнятної точності розпізнавання ( $E < E_{min}$ ) або інші критерії, що характеризують складність дерева (досягнення максимально припустимої кількості вузлів, рівнів, гілок та ін.). Analogічним чином на етапі ініціалізації запропонованого методу формується  $N_\chi$  дерев розв'язків  $\chi_k$ .

Потім виконується оцінювання якості синтезованих дерев розв'язків  $\chi_k, k = 1, 2, \dots, N_\chi$ . Для цього пропонується використовувати цільову функцію  $G = G(\chi_k)$ , що враховує інтерпретовність дерева  $Int(\chi_k)$  і його розпізнаючі властивості (похибку розпізнавання  $E(\chi_k)$ ), і може бути визначена за формулою:  $G(\chi_k) = \gamma_1 Int(\chi_k) + \gamma_2 E(\chi_k)$ , де  $\gamma_1$  та  $\gamma_2$  – коефіцієнти, що дозволяють урахувати важливість критеріїв  $Int(\chi_k)$  та  $E(\chi_k)$ , відповідно.

Для оцінювання інтерпретовності  $Int(\chi_k)$  дерев розв'язків пропонується використовувати такі критерії:

– ширина дерева – може бути визначена як кількість його гілок  $N_{branch}(\chi_k)$  або кількість вузлів  $N_{node}(\chi_k) = N_{branch}(\chi_k) + 1$ ;

– глибина дерева – визначається як кількість його рівнів  $N_{level}(\chi_k)$ ;

– гіллястість дерева  $Br(\chi_k)$  – пропонується обчислювати як відношення кількості вузлів  $N_{node}(\chi_k)$  дерева  $\chi_k$  до максимально можливої кількості вузлів дерева глибини  $N_{level}(\chi_k)$ :  $Br(\chi_k) = \frac{N_{node}(\chi_k)}{\max node(N_{level}(\chi_k))}$ ,

$\max node(N_{level}(\chi_k))$  – максимально можлива кількість вузлів дерева  $\chi_k$  глибини  $N_{level}(\chi_k)$  – величина, що може бути визначена за формулою:

$$\max node(N_{level}(\chi_k)) = \sum_{c=1}^{N_{level}(\chi_k)} 2^{c-1} = 2^{N_{level}(\chi_k)} - 1;$$

– узагальнення рішень  $Gener(\chi_k)$  – відношення кількості листів (вузлів-рішень)  $N_{leaf}(\chi_k)$  дерева  $\chi_k$  до кількості екземплярів  $Q$  навчальної вибірки  $S$ :

$$Gener(\chi_k) = \frac{N_{leaf}(\chi_k)}{Q}.$$

– внутрілиста відстань  $Idist(\chi_k)$  між екземплярами вибірки, що потрапили в конкретний лист (екземплярами, які відповідають конкретним умовам, поданим у вигляді шляхів від кореня дерева до його листів) – чим менше дана відстань, тим вище компактність рішень у відповідних листах, і, отже, тим кращим є розбиття, що виконується деревом  $\chi_k$ . Критерій  $Idist(\chi_k)$  пропонується обчислювати за формулою:

$Idist(\chi_k) = \frac{1}{N_{leaf}(\chi_k)} \sum_{c=1}^{N_{leaf}(\chi_k)} Idist_c$ , де  $Idist_c$  – середня відстань між екземплярами вибірки  $S$ , що попадають в  $c$ -й лист дерева  $\chi_k$  – величина, що розраховується за формулою:

$$Idist_c = \frac{1}{N_{inst}(Leaf_c)} \sum_{q=1}^{N_{inst}(Leaf_c)} \sqrt{\sum_{m=1}^M (p_{mq} - \overline{p_{mc}})^2},$$

де  $N_{inst}(Leaf_c)$  – кількість екземплярів вибірки  $S$ , що попадають в  $c$ -й лист  $Leaf_c$  дерева  $\chi_k$ ;  $\overline{p_{mc}}$  – середнє значення  $m$ -ї ознак екземплярів, що попадають в  $c$ -й лист дерева  $\chi_k$ . З метою приведення значень показника  $Idist(\chi_k)$  до одного інтервалу при аналізі різних вибірок даних, як значення ознак  $p_{mq}$  рекомендується використовувати нормовані значення;

– узагальнення моделі дерева  $GenerM(\chi_k)$  – відношення кількості всіх настроюваних параметрів  $N_{param}(\chi_k)$  моделі на основі дерева розв’язків до розмірності вибірки  $S$ :  $GenerM(\chi_k) = \frac{N_{param}(\chi_k)}{Q \cdot M}$ , де

$N_{param}(\chi_k)$  – кількість настроюваних параметрів дерева  $\chi_k$  – визначається за формулою:  $N_{param}(\chi_k) = 2N_{node}(\chi_k)$ , оскільки кожен  $d_{ab}$  вузол характеризується двома параметрами  $p_{mab}$  та  $p_{TRab}$ .

Як критерій інтерпретовності  $Int(\chi_k)$  можна використовувати один із запропонованих вище критеріїв ( $N_{branch}(\chi_k)$ ,  $N_{node}(\chi_k)$ ,  $N_{level}(\chi_k)$ ,  $Br(\chi_k)$ ,  $Idist(\chi_k)$ ) або їх комбінацію.

Після оцінювання якості синтезованих дерев розв’язків  $\chi_k$ ,  $k = 1, 2, \dots, N_\chi$  виконується перевірка критеріїв зупинення стохастичного пошуку. Як такі критерії можуть бути використані: досягнення прийнятного значення цільової функції  $G(\chi_k)$ , перевищення максимально припустимої кількості ітерацій  $N_I$ , відсутність істотних покращень значення цільової функції  $G(\chi_k)$  протягом заданої кількості ітерацій.

При невиконанні критеріїв зупинення виконується оператор відбору розв’язків  $\chi_k$  для формування нової множини  $R^{(i)} \rightarrow R^{(i+1)} = \{\chi_1^{(i+1)}, \chi_2^{(i+1)}, \dots, \chi_{N_\chi}^{(i+1)}\}$ . Для

цього із множини  $R^{(i)}$  відбираються розв’язки  $\chi_k^{(i)}$  з метою створення набору розв’язків  $RP^{(i)}$ , допущених до відтворення, – розв’язків, на основі яких буде згенеровано новий набір  $R^{(i+1)}$ .

Для відбору розв’язків  $\chi_k^{(i)}$  у множину  $RP^{(i)}$  кожній структурі  $\chi_k^{(i)}$  ставиться у відповідність інтервал  $GI(\chi_k) \in [GI_{min}(\chi_k); GI_{max}(\chi_k)]$ . Величини  $GI_{min}(\chi_k)$  та

$GI_{max}(\chi_k)$  обчислюються в такий спосіб:  $GI_{min}(\chi_k) = GI_{max}(\chi_{k-1})$ ,  $GI_{max}(\chi_k) = GI_{min}(\chi_k) + G_O(\chi_k)$ , де  $GI_{min}(\chi_k)$  й  $GI_{max}(\chi_k)$  – мінімальне та максимальне значення в інтервалі  $GI(\chi_k)$ , відповідно;  $GI_{min}(\chi_1) = 0$  – мінімальне значення в інтервалі  $GI(\chi_1)$  першого розв’язку  $\chi_1$  в популяції  $R^{(i)}$ ;  $G_O(\chi_k)$  – відносне значення цільової функції  $G(\chi_k)$  розв’язку  $\chi_k$  в множині  $R^{(i)}$ , визначається за формулою:  $G_O(\chi_k) = \frac{GM - G(\chi_k)}{\sum_{K=1}^{N_\chi} (GM - G(\chi_K))}$ , де

$GM = \max_{k=1,2,\dots,N_\chi} \{G(\chi_k)\}$  – максимальне значення цільової функції в множині хромосом  $R^{(i)} = \{\chi_1^{(i)}, \chi_2^{(i)}, \dots, \chi_{N_\chi}^{(i)}\}$ .

Як видно з наведених вище формул,  $GI_{min}(\chi_0) = 0$  і  $GI_{max}(\chi_{N_\chi}) = 1$ . Отже, кожному розв’язку  $\chi_k$  ставиться у відповідність деякий інтервал залежно від значення його цільової функції  $G(\chi_k)$ : чим вище значення  $G(\chi_k)$ , тим ширше інтервал  $[GI_{min}(\chi_k); GI_{max}(\chi_k)]$ . У сукупності інтервали  $GI(\chi_k)$ ,  $k = 1, 2, \dots, N_\chi$  утворюють інтервал  $GI(R^{(i)}) \in [0;1]$ .

Після цього генерується випадкове число  $rnd = rand[0;1]$  з інтервалу  $[0;1]$ . У множину допущених до відтворення розв’язків  $RP^{(i)}$  заноситься розв’язок  $\chi_k$ , що попадає в інтервал  $GI(\chi_k)$  якого відповідає випадково згенероване число  $rnd$ :  $rnd \in [GI_{min}(\chi_k); GI_{max}(\chi_k)]$ . Таким чином, чим більше ширина діапазону  $GI(\chi_k)$ , тим вище ймовірність розв’язку  $\chi_k$  бути відібраним для відтворення. Генерація випадкових чисел  $rnd$  і відбір розв’язків  $\chi_k$  для відтворення триває доти, поки не буде повністю сформовано множину  $RP^{(i)}$ ,  $|RP^{(i)}| = N_\chi$ .

Потім на основі розв’язків  $\chi_k$  із множини  $RP^{(i)}$  виконується створення нових розв’язків  $\chi_k^{(i+1)}$ . Для цього із множини  $RP^{(i)}$  вибираються випадковим чином два розв’язки  $\chi_{parent1}$  й  $\chi_{parent2}$ . Після чого в кожному з розв’язків-батьків вибираються вузли  $d_{abparent1}$  й  $d_{abparent2}$  (при цьому ознака-перевірка  $p_{mabparent2}$  вузла  $d_{abparent2}$  не повинна бути ідентичною кожній з ознак, що знаходяться у вузлах дерева  $\chi_{parent1}$  по напрямку від його кореня до вузла  $d_{abparent1}$ , включно, аналогічна умова діє й при виборі вузла  $d_{abparent1}$ ), по яких буде відбуватися обмін між відповідними частинами дерев розв’язків

$\chi_{parent1}$  і  $\chi_{parent2}$ . У результаті такого обміну створюються два нові розв'язки  $\chi_{child1}$  й  $\chi_{child2}$ . При цьому перший розв'язок-нащадок  $\chi_{child1}$  є ідентичним розв'язку  $\chi_{parent1}$  за винятком частини (піддерева), вихідним вузлом якої є  $d_{abparent1}$ . Замість вузла  $d_{abparent1}$  в розв'язку  $\chi_{child1}$  знаходиться вузол  $d_{abparent2}$ , після якого розташовується відповідна частина дерева  $\chi_{parent2}$ . Аналогічно формується дерево-нащадок  $\chi_{child2}$ . Після цього виконується перерахунок граничних значень  $PTrab$  ознак-перевірок  $p_{mab}$  від вузла  $d_{abparent2}$  ( $d_{abparent1}$ ) дерева  $\chi_{child1}$  ( $\chi_{child2}$ ) до відповідних кінцевих вузлів, що містять значення вихідного параметру. У випадку, якщо в результаті таких перетворень на нижніх рівнях дерев  $\chi_{child1}$  і  $\chi_{child2}$  виявляються вузли, перевірки в який пов'язані з ознаками, що вже зустрічаються на більш високих рівнях дерева при проході від кореня до відповідних листів, то такі ознаки-перевірки у вузлах низьких рівнів замінюються на ознаки-перевірки наступного рівня відповідного дерева  $\chi_{child1}$  або  $\chi_{child2}$ . Такий підхід дозволяє для однакових батьків створювати множини нащадків. Створення нових розв'язків  $\chi_k^{(i+1)}$  за допомогою описаного вище підходу триває доти, поки не буде сформовано  $N_{cross} = \beta N_\chi$  розв'язків, де  $\beta$  – коефіцієнт, що визначає значущість формування нової множини розв'язків за допомогою описаної вище процедури схрещування.

Другим способом формування нових розв'язків  $\chi_k^{(i+1)}$  є мутація, що припускає виконання деяких змін над структурою  $\chi_k$ , відібраною із множини  $RP^{(i)}$ . У розробленому стохастичному методі синтезу дерев розв'язків оператор мутації пропонується виконувати в такий спосіб. Спочатку з множини  $RP^{(i)}$  відбирається розв'язок  $\chi_{mutated}$ , у якому випадковим чином вибирається мутуючий вузол  $d_{abmutated}$ , потім у цьому вузлі виконується заміна ознаки-перевірки  $p_{mab}$  на іншу, що не знаходиться у вузлах дерева  $\chi_{mutated}$  по напрямку від його кореня до вузла  $d_{abmutated}$ , включно. Після визначення нової ознаки-перевірки, виконується перерахунок її граничного значення  $PTrab$  й подальше переформатування дерева, починаючи від вузла  $d_{abmutated}$  (аналогічно етапу ініціалізації відбувається побудова фрагмента дерева від вузла  $d_{abmutated}$  доти, поки не будуть досягнуті відповідні критерії зупинення процесу формування дерева). Формування нових розв'язків  $\chi_k^{(i+1)}$  за допомогою мутації триває, поки не буде сформовано  $N_{mutation} = \gamma N_\chi$  розв'язків, де  $\gamma$  – коефіцієнт, що визначає значущість створення нової множини розв'язків за допомогою мутації.

У нову множину розв'язків  $R^{(i+1)}$  заносяться  $N_{cross}$  і  $N_{mutation}$  розв'язків, згенерованих за допомогою схрещування й мутації, а також  $N_{elite} = \alpha N_\chi$ , елітних розв'язків  $\chi_k^{(i)}$  із множини  $R^{(i)}$ , що характеризуються най-

рачими значеннями цільової функції  $G(\chi_k^{(i)})$  в популяції  $R^{(i)}$ , де  $\alpha$  – коефіцієнт, що визначає значущість включення елітних особин у нову множину.

Потім виконується оцінювання значень цільової функції  $G$  для розв'язків  $\chi_k^{(i+1)}$ :  $G = G(\chi_k^{(i+1)})$ ,  $k = 1, 2, \dots, N_\chi$  і формування нової множини  $R^{(i+2)} = \{\chi_1^{(i+2)}, \chi_2^{(i+2)}, \dots, \chi_{N_\chi}^{(i+2)}\}$ . Даний процес триває до виконання критеріїв зупинення.

Результатом виконання стохастичного методу синтезу дерев розв'язків є дерево  $\chi_{opt}$  з мінімальним значенням цільової функції  $G(\chi_{opt}) = \min_{k=1, 2, \dots, N_\chi} \{G(\chi_k)\}$ .

Після цього на основі синтезованого дерева розв'язків  $\chi_{opt}$  виконується видобування правил  $RB$ , що представляють собою найцінніші екземпляри. Для цього обробляється кожний шлях від кореня дерева  $\chi_{opt}$  до листа, у результаті чого будеться відповідне правило, що узагальнює інформацію, подану в деякій множині екземплярів вихідної вибірки  $S$ . Використовуючи такий підхід, видобувається  $N_{RDT}$  правил, загальна кількість яких дорівнює кількості листів (кінцевих вузлів, що містять значення вихідного параметра) синтезованого дерева розв'язків  $\chi_{opt}$ .

Таким чином, розроблений стохастичний метод синтезу дерев розв'язків використовує інформацію про інформативність ознак, складність синтезованого дерева, а також точність його розпізнавання, що дозволяє на початковому етапі формувати множину деревоподібних структур, яка характеризується простою ієархією й невисокою похибкою розпізнавання, у процесі пошуку створювати нові множини розв'язків з урахуванням інформації про значущість ознак й інтерпретовність генерованих дерев, що, у свою чергу, забезпечує можливість побудови дерева розв'язків із невеликою кількістю елементів (вузлів і зв'язків між ними) та прийнятною точністю розпізнавання, а також видобування на його основі найцінніших екземплярів.

#### 4 ЕКСПЕРИМЕНТИ

Виконаємо експериментальне дослідження розробленого стохастичного методу синтезу дерев розв'язків. Для цього порівняємо його з відомими аналогами – методом C4.5 [8–10], методом CART [8–10], а також еволюційним методом побудови дерев розв'язків, запропонованім в [16].

З метою експериментального порівняння запропонованого й відомих методів на мові C# розроблено програмне забезпечення, що дозволяє на основі заданої вибірки даних  $S = \langle P, T \rangle$  виконувати побудову дерев розв'язків за допомогою різних методів. За допомогою розробленого програмного забезпечення вирішувалася задача прийняття рішень при неруйнівному контролі якості кузовів автотранспортних засобів [17].

При виготовленні автотранспортних засобів важливим етапом є неруйнівний контроль якості кузовів. Виявлення некондиційних виробів (кузовів) на ранніх стадіях виготовлення автомобіля дозволить зменшити витрати на усунення дефектів, і, отже, зменшити собівартість виробництва.

У процесі виробництва кузовів автотранспортних засобів на кожному етапі їхнього виготовлення вимірюються деяка група параметрів – контрольних точок, розташованих на кузові та навісних вузлах:

- перша група – контрольні точки на чорному (незабарвленому) кузові;
- друга група – контрольні точки на навісних вузлах (дверях і капоті);
- третя група – зазори й сполучення між навісними вузлами й кузовом.

До аналізованих на першому й другому етапах параметрів (першої й другої групи) відносяться відхилення від номінальних значень контрольних точок. Як правило, більшість таких параметрів знаходитьться в областях допуску, і, отже, не впливають на якість кузова транспортного засобу.

Вимірювання параметрів третьої групи пов’язано з необхідністю установки навісних вузлів (дверей і капота) на пофарбований кузов. Однак при установці навісних вузлів на кузов можуть виникнути деякі деформації, обумовлені відхиленнями номінальних розмірів чорного кузова, дверей і капота від еталонних розмірів (при цьому відхилення кожного з вимірюваних параметрів першої й другої групи може знаходитись в межах допуску), а також іншими факторами, що виникають при складанні. Такі деформації приводять до утворення зазорів і сполучень між навісними вузлами й кузовом, є досить частими й помітними для покупців продукції. Усунення таких недоліків пов’язано з необхідністю розбирання кузова й навісних вузлів, а також з повторним фарбуванням і складанням, що ускладнює й здорожує процес виготовлення якісних виробів.

Тому актуальною є задача побудови моделей залежностей показників третьої групи від параметрів першої й другої груп.

Виявлено, що найбільш важливими параметрами перших двох груп є 18 точок, розташованих в області порога кузова й в областях кріплення петель для складання кузова й навісних вузлів [17]. При цьому в шести точках фіксуються відхилення по двох координатах (третя координата є базовою, внаслідок чого відхилення по даній координаті є нульовим), в інших дванадцяти вимірюються всі три координати, отже, вибірка даних містила значення 48 вхідних параметрів. Також виділено 16 істотних

параметрів третьої групи (зазори й сполучення між дверима й порогом, капотом і крилом, передніми й задніми дверима й ін.).

Таким чином, необхідно синтезувати 16 моделей залежностей параметрів третьої групи від 48 вхідних ознак (параметрів першої й другої груп). Вихідна вибірка містила інформацію про 172 вироби. Нижче наведено результати побудови однієї з моделей. Для інших параметрів отримано аналогічні результати.

Як критерій оцінювання інтерпретовності  $Int(\chi_k)$  дерев розв’язків  $\chi_k$  при використанні стохастичного й еволюційного методів використовувалося узагальнення розв’язків  $Gener(\chi_k) = N_{leaf}(\chi_k)/Q$ , оскільки кількість листів  $N_{leaf}$  відповідає загальній кількості правил  $N_{RDT}$ , що ви добуваються з дерева  $\chi_k$ , і, отже, визначає структуру базу правил, синтезованої на основі дерева розв’язків  $\chi_k$ . Коефіцієнти, що дозволяють врахувати важливість критеріїв  $Int(\chi_k)$  і  $E(\chi_k)$ , вибиралися рівними 0,5 (кожний критерій  $Int(\chi_k)$  і  $E(\chi_k)$  в процесі експериментів мав однакову значущість), отже, цільова функція  $G(\chi_k) = \gamma_1 Int(\chi_k) + \gamma_2 E(\chi_k)$  при стохастичному і еволюційному пошуку визначалася в такий спосіб:  $G(\chi_k) = 0,5 Gener(\chi_k) + 0,5 E(\chi_k)$ .

## 5 РЕЗУЛЬТАТИ

Результати експериментів по дослідження різних методів синтезу дерев розв’язків при побудові діагностичної моделі якості кузовів автомобілів наведено в таблиці 1.

## 6 ОБГОВОРЕННЯ

Як видно з таблиці 1, запропонований стохастичний метод і еволюційний метод синтезу дерев розв’язків дозволяють будувати більш прийнятні діагностичні моделі на основі дерев розв’язків (значення критерію  $G$  нижче в порівнянні з моделями, синтезованими на основі методів CART [8–10] і C4.5 [8–10]), оскільки не використовують жадібну стратегію. Розроблений стохастичний метод забезпечив побудову дерева розв’язків, що характеризується незначною кількістю структурних елементів ( $N_{level} = 7$ ,  $N_{node} = 73$ ,  $N_{leaf} = 37$ ), а також високими апроксимаційними (помилка навчальної вибірки склада  $E = 0,016$ ) і узагальнюючими (помилка на тестових даних  $E_t = 0,028$ ) властивостями, що досягається за рахунок використання в процесі синтезу інформації про інформативність ознак, складність синтезованого дерева та точність його розпізнавання. Скорочення часу пошуку ( $t = 7,31$ ) в по

Таблиця 1 – Результати експериментів

Метод	$N_{level}$	$N_{node}$	$N_{leaf}$	$Gener$	$GenerM$	$Br$	$G$	$Idist$	$t$ , мс	$E$	$E_t$
CART [8–10]	8	95	48	0,279	0,023	0,37	0,166	0,012	15,12	0,052	0,093
C4.5 [8–10]	8	103	52	0,302	0,025	0,4	0,174	0,078	14,82	0,046	0,072
Еволюційний метод [16]	7	83	42	0,244	0,02	0,65	0,131	0,0041	7,42	0,018	0,031
Стохастичний метод	7	73	37	0,215	0,018	0,57	0,116	0,0032	7,31	0,016	0,028

рівнянні з відомими методами забезпечені за рахунок формування на етапі ініціалізації запропонованого методу множини деревоподібних структур, що характеризуються простою ієрархією та невисокою помилкою розпізнавання, а також створення в процесі пошуку нових множин розв'язків із урахуванням інформації про значущість ознак та інтерпретовність генерованих дерев.

Низьке значення критерію узагальнення розв'язків  $Gener = 0,215$  свідчить про високі узагальнюючі здатності дерева: вибірку з 172 екземплярів перетворено в дерево розв'язків, з якого, у свою чергу виділено  $N_{leaf} = 37$  продукційних правил.

Дерева, синтезовані за допомогою еволюційного [16] і запропонованого стохастичного методів є більш гіллястими (значення критеріїв становлять  $Br = 0,65$  і  $Br = 0,62$ , відповідно) у порівнянні з деревами, побудованими за допомогою методів CART [8–10] і C4.5 [8–10] ( $Br = 0,37$  і  $Br = 0,4$ ), що свідчить про більш компактне розташування вузлів.

Таким чином, результати порівняльного аналізу показали, що запропонований стохастичний метод синтезу дерев розв'язків не уступає по якості побудови деревоподібних моделей прийняття рішень відомим методам, і забезпечує можливість побудови дерев розв'язків з невеликою кількістю структурних елементів і прийнятною точністю розпізнавання.

## ВИСНОВКИ

У роботі вирішено актуальну задачу автоматизації видобування знань у вигляді набору продукційних правил з навчальних вибірок даних.

Наукова новизна роботи полягає в тому, що запропоновано стохастичний метод синтезу дерев розв'язків, що використовує інформацію про інформативність ознак, складність синтезованого дерева, а також точність його розпізнавання, що дозволяє на початковому етапі формувати множину деревоподібних структур, яка характеризуються простою ієрархією та невисокою помилкою розпізнавання, у процесі пошуку створювати нові множини розв'язків із урахуванням інформації про значущість ознак і інтерпретовність генерованих дерев, що, у свою чергу, забезпечує можливість побудови дерева розв'язків з невеликою кількістю елементів (вузлів і зв'язків між ними) та прийнятною точністю розпізнавання, а також видобування на його основі найцінніших екземплярів.

Запропоновано систему критеріїв оцінювання моделей на основі дерев розв'язків, що містить у собі критерій оцінювання апроксимаційних властивостей (помилка розпізнавання) і інтерпретовності (ширина, глибина, гіллястість, узагальнення рішень, узагальнення моделі, внутрілисрова відстань) синтезованого дерева. Розроблену систему критеріїв можна використовувати для автоматизації аналізу властивостей і порівняння моделей на основі дерев розв'язків при вирішенні задач неруйнівного контролю якості.

Практична цінність отриманих результатів полягає в тому, що: розроблено програмне забезпечення, яке реа-

лізує запропонований метод і дозволяє виконувати побудову моделей контролю якості на основі дерев розв'язків, а також видобувати продукційні правила з вибірок даних; вирішено практичне завдання прийняття рішень при неруйнівному контролі якості кузовів автотранспортних засобів.

Перспективи подальших досліджень полягають у заснованні запропонованого підходу до видобування знань у вигляді набору продукційних правил з навчальних вибірок даних при синтезі нейро-нечітких моделей для вирішення практичних задач неруйнівного контролю якості.

## ПОДЯКИ

Роботу виконано в рамках держбюджетної науково-дослідної теми Запорізького національного технічного університету «Інтелектуальні інформаційні технології автоматизації проектування, моделювання, керування та діагностування виробничих процесів і систем» (номер державної реєстрації 0112U005350) за підтримки міжнародного проекту «Centers of Excellence for young RESearchers» (CERES) програми «Tempus» Європейської Комісії (реєстраційний номер 544137-TEMPUS-1-2013-1-SK-TEMPUS-JPHES).

## СПИСОК ЛІТЕРАТУРИ

1. Ding S. X. Model-based fault diagnosis techniques: design schemes, algorithms, and tools / S. X. Ding. – Berlin: Springer, 2008. – 473 p.
2. Rutkowski L. Flexible neuro-fuzzy systems : structures, learning and performance evaluation / L. Rutkowski. – Boston : Kluwer, 2004. – 276 p.
3. Нейро-фаззи сети Петри в задачах моделирования сложных систем/[Е. В. Бодянский, Е. И. Кучеренко, А. И. Михалев] – Днепропетровск : Системные технологии. – 2005. – 311 с.
4. Jang J. R. ANFIS: Adaptive-network-based fuzzy inference system / J. R. Jang // IEEE transactions on systems and cybernetics. – 1993. – Vol. 23. – P. 665–685. DOI: 10.1109/21.256541.
5. Mulaik S. A. Foundations of Factor Analysis / S. A. Mulaik. – Boca Raton, Florida: CRC Press. – 2009. – 548 p.
6. Jensen R. Computational intelligence and feature selection: rough and fuzzy approaches / R. Jensen, Q. Shen. – Hoboken: John Wiley & Sons, 2008. – 339 p.
7. Abonyi J. Cluster analysis for data mining and system identification / J. Abonyi, B. Feil. – Basel : Birkhäuser, 2007. – 303 p.
8. Rokach L. Data Mining with Decision Trees. Theory and Applications / L. Rokach, O. Maimon. – London : World Scientific Publishing Co, 2008. – 264 p. DOI: 10.1142/9097.
9. Quinlan J. R. Induction of decision trees / J. R. Quinlan // Machine Learning. – 1986. – No. 1. – P. 81–106. DOI: 10.1007/BF00116251.
10. Classification and regression trees / L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone. – California : Wadsworth & Brooks, 1984. – 368 p.
11. Интеллектуальные информационные технологии проектирования автоматизированных систем диагностирования и распознавания образов : монография / [С. А. Субботин, Ан. А. Олейник, Е. А. Гофман, С. А. Зайцев, Ал. А. Олейник ; под ред. С. А. Субботина. – Харьков] : ООО «Компания Смит», 2012. – 317 с.
12. Yu X. Introduction to Evolutionary Algorithms (Decision Engineering) / X. Yu, M. Gen. – London: Springer, 2010. – 418 p. DOI: 10.1007/978-1-84996-129-5.
13. Gen M. Genetic algorithms and engineering design / M. Gen, R. Cheng. – New Jersey: John Wiley & Sons, 1997. – 352 p. DOI: 10.1002/9780470172254

- 14 Computational intelligence in fault diagnosis / eds.: V. Palade, C.D. Bocaniala, L. Jain. – London: Springer, 2006. – 362 p. DOI: 10.1007/978-1-84628-631-5.
- 15 Субботін С. О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : навч. посібник / С. О. Субботін. – Запоріжжя : ЗНТУ, 2008. – 341 с.
- 16 Гофман Е. А. Еволюціонний метод синтеза дерев'євих рішень / Е. А. Гофман, А. А. Олейник, С. А. Субботін // Штучний інтелект. – 2011. – № 2. – С. 6–14.

Олейник А. А.

Канд. техн. наук, доцент, Запорожський національний технічний університет, Україна

## ІЗВЛЕЧЕНІЙ ЗНАНИЙ НА ОСНОВЕ ДЕРЕВЬЄВ РЕШЕНИЙ И СТОХАСТИЧЕСКОГО ПОИСКА

Решена задача розробки математичного обсягування для автоматизації вивчення знань в виде набора продукційних правил из обучаючих виборок даних. Объектом исследования являлся процесс построения моделей неразрушающего контроля качества. Предмет исследования составляют методы извлечения продукционных правил для синтеза моделей контроля качества. Цель работы: повысить эффективность процесса извлечения продукционных правил для построения моделей контроля качества по обучающим выборкам. Предложен стохастический метод синтеза деревьев решений, который использует информацию об информативности признаков, сложности синтезируемого дерева, а также точности его распознавания, что позволяет на начальном этапе формировать множество древовидных структур, характеризующихся простой иерархией и невысокой ошибкой распознавания, в процессе поиска создавать новые множества решений с учетом информации о значимости признаков и интерпретабельности генерируемых деревьев, что, в свою очередь, обеспечивает возможность построения дерева решений с небольшим количеством элементов (узлов и связей между ними) и приемлемой точностью распознавания, а также извлечение на его основе наиболее ценных экземпляров. Разработано программное обеспечение, реализующее предложенный метод. Проведены эксперименты по исследованию свойств предложенного метода. Результаты экспериментов позволяют рекомендовать предложенный метод для использования на практике.

**Ключові слова:** вибірка, дерево рішень, модель контролю якості, продукційне правило, стохастичний пошук.

Oliynyk A.

Ph.D., Associate Professor, Zaporizhzhya National Technical University, Ukraine

## KNOWLEDGE EXTRACTION BASED ON DECISION TREES AND STOCHASTIC SEARCH

The problem of mathematical support development is solved to automate the extraction knowledge as production rules from the training data samples. The object of study is the process of constructing models of non-destructive quality control. The subject of study are methods of production rules extraction for synthesis of quality control models. The purpose of the work is to improve the efficiency of the process of production rules extraction for constructing models of quality control based on training samples. The stochastic method for the decision trees synthesis is proposed, which uses information about the informativeness of features, the complexity of the synthesized tree, as well as the accuracy of its recognition, which allows to form on the initial stage a set of tree structures, characterized by a simple hierarchy and low error recognition, in the process of search to create a new set of solutions with taking into account information about the significance of the features and interpretability of generated trees, which, in turn, provides the possibility of constructing a decision tree with a small number of elements (nodes and branches between them), and an acceptable recognition accuracy and retrieval based on it the most valuable instances. The software implementing proposed method is developed. The experiments to study the properties of the proposed method are conducted. The experimental results allow to recommend the proposed method for use in practice.

**Keywords:** sample, decision tree, model of quality control, production rule, stochastic search.

## REFERENCES

1. Ding S. X. Model-based fault diagnosis techniques: design schemes, algorithms, and tools. Berlin, Springer, 2008, 473 p.
2. Rutkowski L. Flexible neuro-fuzzy systems : structures, learning and performance evaluation. Boston, Kluwer, 2004, 276 p.
3. Bodjanskiy E. V., Kucherenko E. I., Mihalev A. I. Nejro-fazzi seti Petri v zadachah modelirovaniya slozhnyh system, Dnepropetrovsk, Sistemnye tehnologii, 2005, 311 p.
4. Jang J. R. ANFIS: Adaptive-network-based fuzzy inference system, *IEEE transactions on systems and cybernetics*, 1993, Vol. 23, pp. 665–685. DOI: 10.1109/21.256541
5. Mulaik S. A. Foundations of Factor Analysis. Boca Raton, Florida, CRC Press, 2009, 548 p.
6. Jensen R., Shen Q. Computational intelligence and feature selection: rough and fuzzy approaches. Hoboken, John Wiley & Sons, 2008, 339 p.
7. Abonyi J., Feil B. Cluster analysis for data mining and system identification, Basel, Birkhäuser, 2007, 303 p.
8. Rokach L., Maimon O. Data Mining with Decision Trees. Theory and Applications. London, World Scientific Publishing Co, 2008, 264 p. DOI: 10.1142/9097.
9. Quinlan J. R. Induction of decision trees, *Machine Learning*, 1986, No. 1, pp. 81–106. DOI: 10.1007/BF00116251.
10. Breiman L., Friedman J. H., Olshen R. A., Stone C. J. Classification and regression trees California, Wadsworth & Brooks, 1984, 368 p.
11. Subbotin S. A., Olejnik An. A., Gofman E. A., Zajcev S. A., Olejnik Al. A.; pod red. S. A. Subbotina. Intellektual'nye informacionnye tehnologii proektirovaniya avtomatizirovannyh sistem diagnostirovaniya i raspoznavaniya obrazov : monografija. Har'kov, OOO «Kompanija Smit», 2012, 317 p.
12. Yu X., Gen M. Introduction to Evolutionary Algorithms (Decision Engineering). London, Springer, 2010, 418 p. DOI: 10.1007/978-1-84996-129-5.
13. Gen M., Cheng R. Genetic algorithms and engineering design. New Jersey, John Wiley & Sons, 1997, 352 p. DOI: 10.1002/9780470172254.
- 14 Palade V., Bocaniala C. D., Jain L., eds. Computational intelligence in fault diagnosis. London, Springer, 2006, 362 p. DOI: 10.1007/978-1-84628-631-5.
- 15 Subbotin S. O. Podannja j obrabka znan' u sistemah shtuchnogo intelektu ta pidtrimki prijnattja rishen' : navch. posibnik, Zaporizhzhya, ZNTU, 2008, 341 p.
- 16 Gofman E. A., Olejnik A. A., Subbotin S. A. Jevoljucionnyj metod sinteza derev'ev reshenij, *Shtuchniy intelekt*, 2011, No. 2, pp. 6–14.
- 17 Gofman E. A., Olejnik A. A., Subbotin S. A. Ispol'zovanie derev'ev reshenij dlja diagnostirovaniya avtotransportnyh sredstv, *Informacionnye upravlyajuushie sistemy i kompjuternyy monitoring : II Mezhdunarodnaja nauchno-tehnicheskaja konferencija IUS i KM-2011*, Doneck, 11–13 aprilja 2011 g. : materialy konferencii. Doneck, DonNTU, 2011, Vol. 1, pp. 159–163.

Стаття надійшла до редакції 01.09.2014.

Після доробки 26.09.2014.

Тищенко О. К.<sup>1</sup>, Плісс І. П.<sup>2</sup>, Шкуро К. О.<sup>3</sup><sup>1</sup>Канд. техн. наук, старший науковий співробітник Проблемної НДЛ АСУ, Харківський національний університет радіоелектроніки, Україна<sup>2</sup>Канд. техн. наук, старший науковий співробітник, провідний науковий співробітник Проблемної НДЛ АСУ, Харківський національний університет радіоелектроніки, Україна<sup>3</sup>Канд. техн. наук, науковий співробітник Проблемної НДЛ АСУ, Харківський національний університет радіоелектроніки, Україна

## ПРОГНОЗУЮЧА НЕЙРО-ФАЗЗІ МЕРЕЖА НА ОСНОВІ БАГАТОВИМІРНОГО НЕО-ФАЗЗІ-НЕЙРОНА ТА ЇЇ ПРОЦЕДУРА НАВЧАННЯ

У статті запропоновано прогнозуючу нейро-фаззі мережу, що призначена для вирішення задач екстраполяції багатовимірних нестационарних стохастичних та хаотичних часових рядів за умов короткої навчальної вибірки. В основі мережі полягає багатовимірний нео-фаззі-нейрон із спеціально організованим вхідним шаром та сплайн-функціями належності. Введена система забезпечує високу якість апроксимації у сенсі середньоквадратичної похибки та високу швидкість збіжності за рахунок використання процедури навчання другого порядку. Розроблено програмне забезпечення, що реалізує запропоновану архітектуру нейро-фаззі мережі. Проведено ряд експериментів з дослідження властивостей запропонованої мережі. Результати експериментів підтвердили придатність запропонованої архітектури до розв'язання задач Data Mining та більш високу точність у порівнянні з традиційними прогнозуючими нейро-фаззі системами.

**Ключові слова:** обчислювальний інтелект, багатовимірний нео-фаззі-нейрон, процедура навчання, прогнозування часових рядів, функція належності.

### НОМЕНКЛАТУРА

- $MN$  – багатовимірний нео-фаззі-нейрон;
- $MNS$  – багатовимірний нео-фаззі-синапс;
- $a$  – стан системи Лоренца;
- $\dot{a}$  – похідна від  $a$ ;
- $b$  – параметр системи Лоренца;
- $c_i$  – вектор центроїдів функцій належності  $i$ -го спостереження;
- $c_{i,j}$  –  $j$ -та компонента вектора  $c_i$ ;
- $h$  – кількість функцій належності в нелінійному синапсі;
- $m$  – кількість виходів системи;
- $n$  – кількість входів системи;
- $N$  – загальна кількість спостережень у вибірці;
- $P$  – коваріаційна матриця;
- $r$  – параметр системи Лоренца;
- $t$  – стан системи Лоренца;
- $\dot{t}$  – похідна від  $t$ ;
- $u$  – стан системи Лоренца;
- $\dot{u}$  – похідна від  $u$ ;
- $W$  – матриця синаптичних ваг;
- $w_{mhn}$  – синаптична вага  $n$ -го входу системи  $h$ -ої функції належності  $m$ -го нейрона;
- $w_{lj}^{[i]}$  –  $l$ -та синаптична вага  $i$ -го входу системи для  $j$ -го центроїду;
- $x(k)$  – вектор вхідних значень;
- $\hat{x}(k)$  – вектор вихідних значень (прогнозів);
- $x_i(k)$  –  $i$ -а компонента вектора вхідних значень;

$\hat{x}_i(k)$  –  $i$ -та компонента вектора вихідних значень (прогнозів);  
 $z^{-1}$  – елемент чистого запізнювання;  
 $\alpha$  – параметр забування;  
 $\mu_l^{[i]}$  –  $l$ -та функція належності  $i$ -го входу системи;  
 $\mu_{ij}$  – функція належності  $i$ -го спостереження до  $j$ -го центроїду;  
 $\mu(k)$  – вектор функцій належності;  
 $\sigma$  – параметр системи Лоренца.

### ВСТУП

Задача оброблення багатовимірних часових рядів досить часто виникає у багатьох технічних, медико-біологічних, соціально-економічних дослідженнях, де від якості отримуваних рішень істотно залежить точність синтезованих результатів. Треба відзначити, що у багатьох реальних завданнях часові ряди характеризуються високим рівнем нелінійності (можливо навіть хаотичності) та нестационарності своїх параметрів, наявністю нерегулярних трендів, стрибків та аномальних викидів. Зрозуміло, що традиційні методи аналізу часових послідовностей, що базуються на регресійному, кореляційному, спектральному та інших подібних підходах, що мають на увазі априорну наявність достатньої вибірки спостережень, виявляються неефективними. Альтернативою традиційним статистичним методам може слугувати математичний апарат обчислювального інтелекту й, перш за все, штучні нейронні мережі та нейро-фаззі системи завдяки своїм універсальним апроксимуючим властивостям. Разом з тим з апроксимуючих властивостей зовсім не

випливають екстраполюючі, оскільки урахування давньої передисторії для побудови, наприклад, прогнозувальної моделі може погіршити якість прогнозу. У зв'язку з цим, обробляючи істотно нестационарні процеси, треба відмовитись від процедур навчання, заснованих на зворотному поширенні похибок або методі найменших квадратів, щоб скористатися процедурами, що ґрунтуються на локальних критеріях та «короткій» пам'яті на зразок алгоритму Качмажа-Уідроу-Хоффа. При цьому застосовані методи навчання мають забезпечувати не тільки високу швидкодію, але й фільтрувальні властивості для придушення стохастичної «шумової» компоненти в оброблюваному сигналі. У зв'язку з цим синтез спеціалізованих гіbridних систем обчислювального інтелекту, що призначаються для розв'язання задач обробки істотно нестационарних багатовимірних часових рядів за умов невизначеності та забезпечують разом з високою швидкістю навчання фільтрацію завад, є досить цікавою та перспективною задачею.

Незважаючи на велику кількість наукових робіт з цієї тематики [1–3], все ще існує задача обробки багатовимірних часових рядів, викликана потребою у покращенні якості отриманих результатів та збільшенні швидкості оброблення даних. У зв'язку з цим актуальною є задача розробки методів аналізу багатовимірних часових рядів на основі нейро-фаззі мереж для послідовного оброблення нелінійних багатовимірних часових рядів, що здатні функціонувати за умов дефіциту априорної інформації та забезпечують можливість обробки цих рядів із короткою навчальною вибіркою, а також характеризуються підвищеною швидкістю навчання. Згідно з усім сказаним вище, робота є актуальну, що і визначає перспективність отримання як теоретичних, так і практичних результатів.

## 1 ПОСТАНОВКА ЗАДАЧІ

Нехай задано багатовимірний нелінійний нестационарний стохастичний часовий ряд, який можна записати у вигляді вектора  $x(k) = (x_1(k), \dots, x_n(k))^T$  – вхідні дані для запропонованої обчислювальної системи.

Тоді задача дослідження полягає в online прогнозуванні багатовимірних нелінійних нестационарних стохастичних або хаотичних часових рядів за умов невизначеності та короткої навчальної вибірки на основі введеній нейро-фаззі мережі, в основі якої полягає багатовимірний нео-фаззі-нейрон та адаптивний метод його навчання, що забезпечують високу якість апроксимації та екстраполяції, а також підвищену швидкість збіжності за рахунок функцій належності спеціального вигляду та використання швидкодіючого алгоритму навчання.

Результатом роботи запропонованої нейро-фаззі мережі є вектор прогнозів  $\hat{x}(k+1) = (\hat{x}_1(k+1), \hat{x}_2(k+1), \dots, \hat{x}_m(k+1))^T$ .

У якості цільової функції може бути використана середньоквадратична похибка  $E = \frac{1}{N} \sum_{k=1}^N (x(k) - \hat{x}(k))^2$ , яку потрібно мінімізувати.

## 2 ОГЛЯД ЛІТЕРАТУРИ

Методи, що широко використовуються зараз для обробки істотно нестационарних багатовимірних часових рядів, мають низку недоліків, а саме: радіально-базисна нейронна мережа схильна до «прокльону розмірності» [1–3], а адаптивна нейро-фаззі система нечіткого виведення [4] має громіздку архітектуру та досить повільно навчається за рахунок використання алгоритму зворотного поширення похибок. А це веде до збільшення кількості нейронів і кількості нечітких правил у базі знань, звідки випливає необхідність збільшення обсягу навчальної вибірки для налаштування подібної системи. Як було зазначено вище, нейро-фаззі системи, що використовують для навчання алгоритм зворотного поширення похибок, характеризуються низькою швидкістю, що робить їх застосування при обробці даних у послідовному режимі не дуже вдалим [4–6]. Уникнути цих недоліків можна за рахунок використання гіbridних систем, що поєднують як теорію штучних нейронних мереж, яка дозволяє отримати універсальні апроксимувальні властивості і здатність навчатися, так і теорію нечіткої логіки, яка дає можливість наділяти систему лінгвістичного інтерпретованістю [7–9].

Відомі дослідження архітектури нео-фаззі-предиктора [10] та архітектури на основі багатовимірного нео-фаззі-нейрона [11, 12], але слід відзначити їх декотру громіздкість та не завжди високу швидкість оброблення даних у режимі online.

У цій роботі запропоновано використовувати архітектуру на основі нео-фаззі-нейронів [13–16], що характеризуються високою швидкістю навчання за рахунок лінійної залежності виходів від настоюваних ваг; обчислювальною простотою; здатністю знаходження глобального мінімуму критерія навчання у реальному часі та реалізацією лінгвістичних «if-then» правил.

## 3 МАТЕРІАЛИ ТА МЕТОДИ

На рис. 1 зображено архітектуру нейро-фаззі мережі. На вхід системи подається багатовимірний сигнал  $x(k) = (x_1(k), \dots, x_n(k))^T$ . Слід зазначити, що елементи чистого запізнювання використовуються для шару функцій належності. Далі сигнал від них надходить до синаптичних ваг знову через шар елементів чистого запізнювання. Сумування отриманих сигналів відбувається досить незвичним чином (рис. 1).

Традиційні нео-фаззі-нейрони були запропоновані Т. Ямакавою та його колегами [13–16]. Вони є нейронними моделями з нелінійними синапсами. Вихід нелінійного синапсу нейрона формується шляхом сумування вихідних сигналів синапсів, які представлені нелінійними функціями. Вони можуть апроксимувати нелінійне відображення «вхід-вихід» одним нейроном.

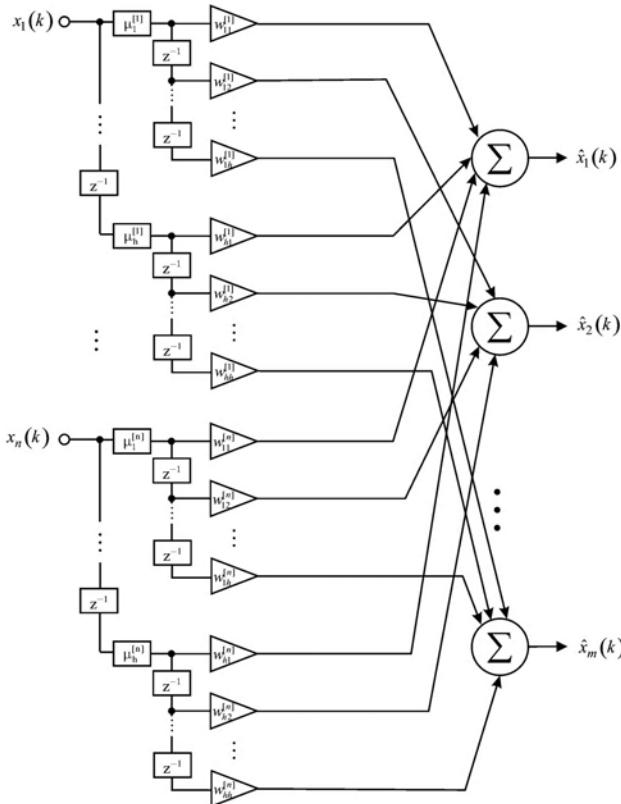


Рисунок 1 – Архітектура запропонованої прогнозуючої нейро-фаззі мережі

Автори нео-фаззі-нейрона використовували традиційні трикутні конструкції, що відповідають умовам розбиття Руспіні у якості функцій належності:

$$\mu_{ij}(x_i) = \begin{cases} \frac{x_i - c_{i,j-1}}{c_{ij} - c_{i,j-1}}, & \text{якщо } x_i \in [c_{i,j-1}, c_{ij}]; \\ \frac{c_{i,j+1} - x_i}{c_{i,j+1} - c_{ij}}, & \text{якщо } x_i \in [c_{ij}, c_{i,j+1}]; \\ 0, & \text{у протилежному випадку.} \end{cases}$$

де  $c_{ij}$  обрані досить випадково (зазвичай рівно розподілені) параметри центрів функцій належності на інтервали  $[0,1]$ , де  $0 \leq x_i \leq 1$ . Такий вибір функцій належності гарантує, що вхідний сигнал  $x_i$  активує тільки дві сусідні функції, а їх сума буде завжди рівною одиниці.

Зрозуміло, що такі конструкції, як поліноміальні гармонічні функції, вейвлети, ортогональні функції, тощо також можуть бути використані в якості функцій належності в нелінійних синапсах.

Відзначимо, що трикутні функції належності забезпечують кусково-лінійну апроксимацію, яка призводить до погіршення точності результатів. З метою мінімізації цього ефекту можна збільшити кількість синаптичних ваг і, таким чином, значно ускладнити архітектуру і алгоритм

навчання. Кубічні сплайні можна використати як функції належності, аби позбутися вищезазначених проблем:

$$\mu_{ij}(x_i) = \begin{cases} 0,25 \left( 2 + 3 \frac{2x_i - c_{ij} - c_{i,j-1}}{c_{ij} - c_{i,j-1}} - \left( \frac{2x_i - c_{ij} - c_{i,j-1}}{c_{ij} - c_{i,j-1}} \right)^3 \right), & x \in [c_{i,j-1}, c_{ij}]; \\ 0,25 \left( 2 - 3 \frac{2x_i - c_{i,j+1} - c_{ij}}{c_{i,j+1} - c_{ij}} + \left( \frac{2x_i - c_{i,j+1} - c_{ij}}{c_{i,j+1} - c_{ij}} \right)^3 \right), & x \in (c_{ij}, c_{i,j+1}]; \\ 0, & \text{у протилежному випадку.} \end{cases}$$

Такі кубічні сплайні також задовольняють умовам розбиття Руспіні та покращують характеристики апроксимації процесу нечіткого висновування. З іншого боку, використання кубічних сплайнів забезпечує гладку поліноміальну апроксимацію та дозволяє моделювати нестационарні сигнали з високою точністю результатів.

Архітектура типового нео-фаззі-нейрона в якості складової частини багатовимірного нейрона обчислювальної системи є надлишковою, бо вектор вхідних сигналів  $x(k)$  подається до однотипних нелінійних синапсів нео-фаззі-нейронів, кожен нейрон яких генерує на вихіді сигнал  $\hat{x}_d(k)$ ,  $d = 1, 2, \dots, m$ . У результаті компоненти вектору вихідів  $\hat{x}(k) = (\hat{x}_1(k), \hat{x}_2(k), \dots, \hat{x}_m(k))^T$  обчислюються незалежно. Цю надлишковість системи можна обійти, використовуючи багатовимірний нео-фаззі-нейрон [11, 12], архітектура якого є модифікацією системи, запропонованої у [10]. Структурними елементами такої обчислювальної системи є складені нелінійні синапси  $MNS_i$ , при чому кожен синапс містить  $h$  функцій належності  $\mu_j^{[i]}$  та  $mh$  настроюваних синаптичних ваг  $w_{jl}^{[i]}$ . Хоча багатовимірний нео-фаззі-нейрон має  $mh$  синаптичних ваг, але тільки  $hn$  функцій належності, що в  $m$  разів менше, ніж якщо система була б сформована з традиційних нео-фаззі-нейронів.

Запропонована нейро-фаззі система складається з багатовимірних нео-фаззі-синапсів. Вводячи до розгляду  $(hn \times 1)$ -вектор функцій належності  $\mu(k) = (\mu_{11}(x_1(k)), \mu_{21}(x_1(k)), \dots, \mu_{h1}(x_1(k)), \dots, \mu_{1l}(x_1(k)), \dots, \mu_{hl}(x_1(k)))^T$  та  $(m \times hn)$ -матрицю синаптичних ваг

$$W = \begin{pmatrix} w_{111} & w_{112} & \cdots & w_{1li} & \cdots & w_{1hn} \\ w_{211} & w_{212} & \cdots & w_{2li} & \cdots & w_{2hn} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{m11} & w_{m12} & \cdots & w_{mli} & \cdots & w_{mhn} \end{pmatrix},$$

вихідний сигнал  $MN_j$  у  $k$ -й момент часу можна записати у вигляді

$$\hat{x}(k) = W\mu(k).$$

Навчання багатовимірного нео-фаззі-нейрона може бути реалізовано за допомогою матричної модифікації експоненційно-зваженого рекурентного методу найменших квадратів у формі [17–21]:

$$\begin{cases} W(k+1) = W(k) + \frac{(\hat{x}(k+1) - W(k)\mu(k+1))(\mu(k+1))^T P(k)}{\alpha + (\mu(k+1))^T P(k)\mu(k+1)}; \\ P(k+1) = \frac{1}{\alpha} \left( P(k) - \frac{P(k)\mu(k+1)(\mu(k+1))^T P(k)}{\alpha + (\mu(k+1))^T P(k)\mu(k+1)} \right), 0 < \alpha \leq 1, \end{cases}$$

де  $\hat{x}(k+1) = (\hat{x}_1(k+1), \hat{x}_2(k+1), \dots, \hat{x}_m(k+1))^T$ .

#### 4 ЕКСПЕРИМЕНТИ

Аби продемонструвати ефективність запропонованої нейро-фаззі мережі та її процедури навчання, у якості тестової вибірки було обрано багатовимірний ряд на основі диференційних рівнянь моделі Лоренца:

$$\begin{cases} \dot{a} = \sigma(u - a), \\ \dot{u} = a(r - t) - u, \\ \dot{t} = au - bt. \end{cases} \quad (1)$$

Диференційне рівняння (1) можна записати у дискретному часі в рекурентній формі:

$$\begin{cases} a(k+1) = a(k) + \sigma(u(k) - a(k))dt, \\ u(k+1) = u(k) + (ra(k) - a(k)t(k) - u(k))dt, \\ t(k+1) = t(k) + (a(k)u(k) - bt(k))dt, \end{cases} \quad (2)$$

де  $\sigma = 10, r = 28, b = \frac{8}{3}, dt = 0,001$ .

Атрактор Лоренца є фрактальної структурою, що відповідає довгостроковій поведінці генератора Лоренца. Сам генератор є тривимірною динамічною системою, що демонструє хаотичний потік даних, а також відомий своєю лемніскатною формою. На рис. 2 можна побачити, як розвивається стан динамічної системи (три змінні тривимірної системи) з часом.

Вибірка даних, що отримана за допомогою формули (2), містила 10000 спостережень, серед яких 7000 спостережень – навчальна вибірка, 3000 – перевірочна. Синаптичні ваги ініціалізувалися як нулі, також можна було це робити випадковим чином. На рис. 2 на осіх розташовані значення параметрів системи Лоренца.

Таблиця 1 – Результати експериментів

Нейро-фаззі мережа	Кількість настроюваних параметрів	Час навчання	Помилка навчання	Помилка тестування
Адаптивний нео-фаззі-предиктор	60	6,2 с	0,0189	0,0231
Архітектура на основі багатовимірних нео-фаззі-нейронів	45	4,1 с	0,0154	0,0171
Запропонована нейро-фаззі мережа	27	3,9 с	0,0121	0,0132

#### 5 РЕЗУЛЬТАТИ

Результати експериментів з дослідження різних методів прогнозування хаотичного часового ряду Лоренца наведено в таблиці 1.

#### 6 ОБГОВОРЕННЯ

Як можна побачити із таблиці 1, запропонована нейро-фаззі мережа та процедура її навчання досить непогано показали себе при розв'язанні тестової задачі. Слід відзначити меншу кількість настроюваних параметрів (у півтора рази у порівнянні з архітектурою на основі багатовимірних нео-фаззі-нейронів та більше, ніж у два рази, якщо порівнювати з адаптивним нео-фаззі-предиктором). Слід також відзначити найменший час серед усіх систем, але час не суттєво відрізняється від часу, продемонстрованого архітектурою на основі багатовимірних нео-фаззі-нейронів. Якщо ж говорити про якість прогнозів, то запропонована мережа продемонструвала найкращі результати серед нейро-фаззі мереж (перевага над конкурентами склала у випадку з адаптивним нео-фаззі-предиктором 27% та 56% у випадку з другою архітектурою).

Таким чином, результати порівняльного аналізу показали, що запропонована архітектура нейро-фаззі мережі за якістю прогнозів не поступається іншим відомим нейро-фаззі методам, а за кількістю настроюваних параметрів навіть переважає їх. Запропонована процедура навчання дозволяє системі швидше налаштовувати параметри у порівнянні з аналогами.

#### ВИСНОВКИ

У роботі розв'язано актуальну задачу прогнозування багатовимірних нелінійних нестационарних стохастичних часових рядів.

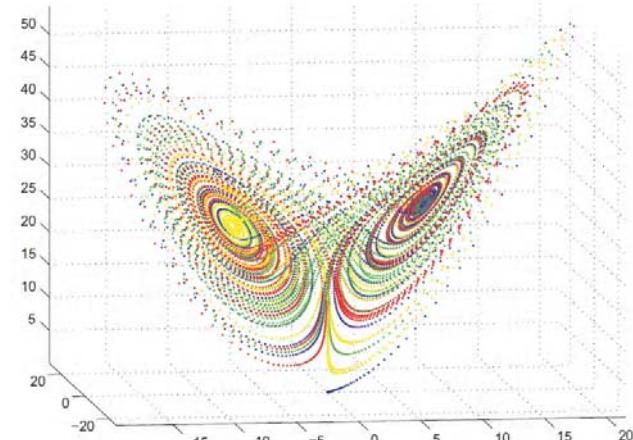


Рисунок 2 – Хаотичний часовий ряд

Наукова новизна роботи полягає у тому, що запропоновано прогнозуючу нейро-фаззі мережу на основі багатовимірних нео-фаззі-нейронів та введено процедуру навчання для оброблення багатовимірних даних у режимі online. У якості функцій належності використовуються кубічні сплайні, що забезпечують високу точність апроксимації та екстраполяції у порівнянні з відомими нейро-фаззі системами прогнозування. Введена система може бути використана для прогнозування суттєво нестационарних часових рядів за умов короткої навчальної вибірки.

Практична цінність отриманих результатів полягає в тому, що цю мережу можна використовувати при розв'язанні широкого кола задач Data Mining та, перш за все, прогнозування.

Перспективи подальших досліджень полягають у тому, щоб застосувати деякі інші функції належності та протестувати інші типи вузлів для нейро-фаззі мережі.

## ПОДЯКИ

Роботу виконано в рамках держбюджетної науково-дослідної теми Харківського національного університету радіоелектроніки «Гібридні системи обчислювально-го інтелекту для адаптивної обробки викривлених даних за умов ап'єрної та поточної невизначеності» (номер державної реєстрації 0113U000361).

## СПИСОК ЛІТЕРАТУРИ

1. Cichocki A. Neural Networks for Optimization and Signal Processing / A. Cichocki, R. Unbehauen. – Stuttgart : Teubner, 1993. – 526 p.
2. Haykin S. Neural Networks. A Comprehensive Foundation / S. Haykin. – Upper Saddle River: Prentice Hall, 1999. – 842 p.
3. Schalkoff R. J. Artificial Neural Networks / R. J. Schalkoff. – N. Y. : The McGraw-Hill Comp., 1997. – 528 p.
4. Jang J.-S. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence / J.-S. Jang, C.-T. Sun, E. Mizutani. – Upper Saddle River : Prentice Hall, 1997. – 640 p.
5. Osowski S. Sieci neuronowe do przetwarzania informacji / S. Osowski. – Warszawa : Oficyna Wydawnicza Politechniki Warszawskiej, 2006. – 422 p.
6. Du K.-L. Neural Networks and Statistical Learning / K.-L. Du, M. N. S. Swamy. – London: Springer-Verlag, 2014. – 815 p.
7. Rutkowski L. Computational Intelligence. Methods and Techniques / L. Rutkowski. – Berlin : Springer-Verlag, 2008. – 514 p.
8. Mumford C. L. Computational Intelligence / C. L. Mumford, L. C. Jain. – Berlin : Springer-Verlag, 2009. – 725 p.
9. Kruse R. Computational Intelligence. A Methodological Introduction / [R. Kruse, C. Borgelt, F. Klawonn et al]. – Berlin : Springer-Verlag, 2013. – 488 p.
10. Бодянський Е. В. Адаптивний фільтр-предиктор многомерних существенно нестационарных временных рядов / Е. В. Бодянський, О. А. Романюк, О. С. Удовенко // Системи обробки інформації. – Харків, 2009. – Вип. 4 (78). – С. 23–28.
11. Бодянський Є. Прогнозування багатовимірних нестационарних часових рядів на основі адаптивної нео-фаззі-моделі / Є. Бодянський, О. Тищенко, Д. Копаліані // Вісник НУ «Львівська політехніка» «Комп'ютерні науки та інформаційні технології». – 2012. – № 744. – С. 312–118.
12. Bodyanskiy Ye. A multivariate non-stationary time series predictor based on an adaptive neuro-fuzzy approach / Ye. Bodyanskiy, O. Tyshchenko, W. Wojcik // Elektronika – konstrukcje, technologie, zastosowania. – 2013. – № 8. – P. 10–13.
13. Yamakawa T. A neo-fuzzy neuron and its application to system identification and prediction of the system behavior / T. Yamakawa, E. Uchino, J. Miki, H. Kusanagi // IIZUKA-92: 2-nd International Conference on Fuzzy Logic and Neural Networks: proceedings. – Iizuka, Japan, 1992. – P. 477–483.
14. Uchino E. Soft computing based signal prediction, restoration and filtering / E. Uchino, T. Yamakawa // Intelligent Hybrid Systems: Fuzzy Logic, Neural Networks and Genetic Algorithms / ed.: Da Ruan. – Boston : Kluwer Academic Publishers, 1997. – P. 331–349.
15. Miki J. Analog implementation of neo-fuzzy neuron and its on-board learning / J. Miki, T. Yamakawa // Computational Intelligence and Applications / ed.: N.E. Mastorakis. – Piraeus : WSEAS Press, 1999. – P. 144–149.
16. Takagi T. Fuzzy identification of systems and its application to modeling and control / T. Takagi, M. Sugeno // IEEE Transactions on System, Man and Cybernetics. – 1985. – 15. – P. 116–132.
17. Bodyanskiy Ye. Anadaptive learning algorithm for a neuro-fuzzy network / Ye. Bodyanskiy, V. Kolodyazhnyi, A. Stephan // Computational Intelligence : Theory and Applications / ed.: 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. Kolodyazhnyi // Integrated Computer-Aided Engineering. – 2003. – 10. – № 4. – P. 399–409.
19. Bodyanskiy Ye. Adaptive nonlinear control using neo-fuzzy model / Ye. Bodyanskiy, V. Kolodyazhnyi // Sinergies Between Information Processing and Automation / eds.: O. Sawodny, P. Sharff. – Aachen : Shaker Verlag, 2004. – P. 122–127.
20. Bodyanskiy Ye. A self-training robust neo-fuzzy controller with constraints on control actions / Ye. Bodyanskiy, I. Kokshenev, V. Kolodyazhnyi, P. Otto // IWK-2005: 50 Internationales Wissenschaftliches Kolloquium: Tagungsband. – TU Ilmenau, 2005. – P. 125–126.
21. An adaptive learning algorithm for a neo-fuzzy neuron / [Ye. Bodyanskiy, I. Kokshenev, V. Kolodyazhnyi, P. Otto] // EUSFLAT-2003: 3rd International Conference of European Union Society for Fuzzy Logic and Technology: proceedings. – Zittau, Germany, 2003. – P. 375–379.

Стаття надійшла до редакції 07.11.2014.

Після доробки 11.11.2014.

Тищенко А. К.<sup>1</sup>, Плисс И. П.<sup>2</sup>, Шкуро К. А.<sup>3</sup>

<sup>1</sup>Канд. техн. наук, старший научный сотрудник Проблемной НИЛ АСУ, Харьковский национальный университет радиоэлектроники, Украина

<sup>2</sup>Канд. техн. наук, старший научный сотрудник, ведущий научный сотрудник Проблемной НИЛ АСУ, Харьковский национальный университет радиоэлектроники, Украина

<sup>3</sup>Канд. техн. наук, научный сотрудник Проблемной НИЛ АСУ, Харьковский национальный университет радиоэлектроники, Украина

## ПРОГНОЗИРУЮЩАЯ НЕЙРО-ФАЗЗІ СЕТЬ НА ОСНОВЕ МНОГОМЕРНОГО НЕО-ФАЗЗІ-НЕЙРОНА И ЕЕ ПРОЦЕДУРА ОБУЧЕНИЯ

В статье предложена прогнозирующая нейро-фаззі сеть, предназначенная для решения задач экстраполяции многомерных нелинейных нестационарных стохастических и хаотических временных рядов в условиях короткой обучающей выборки. В основе

сети лежит многомерный нео-фаззи-нейрон с организованным специальным образом входным слоем и сплайн-функциями при- надлежности. Введенная система обеспечивает высокое качество аппроксимации в смысле среднеквадратичной ошибки и высокую скорость сходимости за счет использования процедуры обучения второго порядка. Разработано программное обеспечение, реализующее предложенную архитектуру нейро-фаззи сети. Проведен ряд экспериментов по исследованию свойств предложенной сети. Результаты экспериментов подтвердили применимость предложенной архитектуры при решении задач Data Mining и более высокую точность по сравнению с традиционными прогнозирующими нейро-фаззи системами.

**Ключевые слова:** вычислительный интеллект, многомерный нео-фаззи-нейрон, процедура обучения, прогнозирование временных рядов, функция принадлежности.

Tyshchenko O. K.<sup>1</sup>, Pliss I. P.<sup>2</sup>, Shkuro K. O.<sup>3</sup>

<sup>1</sup>Ph.D., Senior Researcher at Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Ukraine

<sup>2</sup>Ph.D., Senior Researcher, Leading Researcher at Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Ukraine

<sup>3</sup>Ph.D., Researcher at Control Systems Research Laboratory, Kharkiv National University of Radio Electronics, Ukraine

## A FORECASTING NEURO-FUZZY NETWORK BASED ON A MULTIDIMENSIONAL NEO-FUZZY NEURON AND ITS LEARNING PROCEDURE

A forecasting neuro-fuzzy network designated for solving extrapolation tasks of multidimensional nonlinear non-stationary stochastic and chaotic time series under conditions of a short learning sample is proposed in the paper. The network is built with the help of a multidimensional neo-fuzzy neuron with an input layer which is organized in a special manner and a spline membership function. The proposed system provides high approximation quality in terms of a mean square error and high convergence speed on account of using the second-order learning procedure. A software that implements the proposed neuro-fuzzy network has been developed. A number of experiments has been held in order to research the system's properties. Experimental results prove the fact that the developed architecture could be used in Data Mining tasks and the fact that the proposed neuro-fuzzy network has higher accuracy compared to traditional forecasting neuro-fuzzy systems.

**Keywords:** computational intelligence, multidimensional neo-fuzzy neuron, learning procedure, time series prediction, membership function.

## REFERENCES

1. Cichocki A., Unbehauen R. Neural Networks for Optimization and Signal Processing. Stuttgart, Teubner, 1993, 526 p.
2. Haykin S. Neural Networks. A Comprehensive Foundation. Upper Saddle River, Prentice Hall, 1999, 842 p.
3. Schalkoff R. J. Artificial Neural Networks. N.Y, The McGraw-Hill Comp., 1997, 528 p.
4. Jang J.-S., Sun C.-T., Mizutani E. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Upper Saddle River: Prentice Hall, 1997, 640 p.
5. Osowski S. Sieci neuronowe do przetwarzania informacji. Warszawa, Oficyna Wydawnicza Politechniki Warszawskiej, 2006, 422 p.
6. Du K.-L., Swamy M.N.S. Neural Networks and Statistical Learning. London, Springer-Verlag, 2014, 815 p.
7. Rutkowski L. Computational Intelligence. Methods and Techniques. Berlin, Springer-Verlag, 2008, 514 p.
8. Mumford C. L., Jain L. C. Computational Intelligence. Berlin, Springer-Verlag, 2009, 725 p.
9. Kruse R., Borgelt C., Klawonn F., Moewes C., Steinbrecher M., Held P. Computational Intelligence. A Methodological Introduction. Berlin, Springer-Verlag, 2013, 488 p.
10. Bodyanskiy Ye. V., Romanyuk O. A., Udovenko O. S. Adaptive filter-predictor for multidimensional non-stationary time series, *Systemy obrobky informacii*. Kharkiv, 2009, Vol. 4 (78). – P. 23–28.
11. Bodyanskiy Ye., Tyshchenko O., Kopaliani D. Multidimensional non-stationary time-series prediction with the help of an adaptive neo-fuzzy model, *Bulletin of the National university «Lvivska politehnika»*, No. 744. Lviv, Lvivska politehnika Publishing, 2012, P. 312–320.
12. Bodyanskiy Ye., Tyshchenko O., Wojcik W. A multivariate non-stationary time series predictor based on an adaptive neuro-fuzzy approach, *Elektronika – konstrukcje, technologie, zastosowania*, 2013, No. 8, pp. 10–13.
13. Yamakawa T., Uchino E., Miki J., Kusanagi H. A neo-fuzzy neuron and its application to system identification and prediction of the system behavior, *IIZUKA-92: 2-nd International Conference on Fuzzy Logic and Neural Networks: proceedings*. Iizuka, Japan, 1992, pp. 477–483.
14. Uchino E., Yamakawa T. ed.: Da Ruan. Soft computing based signal prediction, restoration and filtering, *Intelligent Hybrid Systems: Fuzzy Logic, Neural Networks and Genetic Algorithms*. Boston, Kluwer Academic Publishers, 1997, pp. 331–349.
15. Miki J., Yamakawa T. ed.: N. E. Mastorakis Analog implementation of neo-fuzzy neuron and its on-board learning, *Computational Intelligence and Applications*. Piraeus, WSES Press, 1999, pp. 144–149.
16. Takagi T., Sugeno M. Fuzzy identification of systems and its application to modeling and control, *IEEE Transactions on System, Man and*, 1985, 15, pp. 116–132.
17. Bodyanskiy Ye., Kolodyazhnyi V., Stephan A. ed.: Reusch B. An adaptive learning algorithm for a neuro-fuzzy network, *Computational Intelligence: Theory and Applications*. Berlin-Heidelberg-New-York, Springer, 2001, pp. 68–75.
18. Otto P., Bodyanskiy Ye., Kolodyazhnyi V. A new learning algorithm for a forecasting neuro-fuzzy network, *Integrated Computer-Aided Engineering*, 2003, 10, No. 4, pp. 399–409.
19. Bodyanskiy Ye. Adaptive nonlinear control using neo-fuzzy model /Ye. Bodyanskiy, V. Kolodyazhnyi// Sinergies Between Information Processing and Automation / eds.: O. Sawodny, P. Sharff. – Aachen: Shaker Verlag, 2004. – P. 122–127.
20. Bodyanskiy Ye., Kokshenev I., Kolodyazhnyi V., Otto P. A self-training robust neo-fuzzy controller with constraints on control actions, *IWK-2005: 50 Internationales Wissenschaftliches Kolloquium*, Tagungsband, TU Ilmenau, 2005, pp. 125–126.
21. Bodyanskiy Ye., Kokshenev I., Kolodyazhnyi V., Otto P. An adaptive learning algorithm for a neo-fuzzy neuron, *EUSFLAT-2003: 3rd International Conference of European Union Society for Fuzzy Logic and Technology: proceedings*. Zittau, Germany, 2003, pp. 375–379.

# ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

## ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

## PROGRESSIVE INFORMATION TECHNOLOGIES

UDC 519.2:004.9

Kirichenko L. O.<sup>1</sup>, Kobitskaya Yu. A.<sup>2</sup>, Habacheva A. Yu.<sup>3</sup>

<sup>1</sup>Dr. Sc., Associate Professor, Professor of Department of Applied Mathematics,  
Kharkiv National University of Radio Electronics, Ukraine

<sup>2</sup>Postgraduate of Department of Applied Mathematics, Kharkiv National University of Radio Electronics, Ukraine  
<sup>3</sup>Student of Department of Applied Mathematics, Kharkiv National University of Radio Electronics, Ukraine

### COMPARATIVE ANALYSIS OF THE COMPLEXITY OF CHAOTIC AND STOCHASTIC TIME SERIES

The new approach to the recognition mechanism of the time series generating process based on the results of the entropy and the recurrent analysis is proposed. The comparative analysis of the realizations properties of chaotic and stochastic processes with different correlation structure was carried out. It is shown that the derived set of information characteristics allows to distinguish the realizations of deterministic chaotic and fractal random processes. Depending on complexity measures of time series of process parameters were obtained. The information characteristics dependencies from the process parameters were obtained. The results of bioelectric signals and financial time series study are presented.

**Keywords:** time series, measures of complexity, approximate entropy, recurrence plot, pseudo-phase space, embedding dimension.

#### NOMENCLATURE

$A$  is a control parameter;  
 $ApEn$  is approximate entropy;  
 $C_m(\varepsilon)$  is correlation integral;  
 $Det$  is measure of determinism;  
 $F(t)$  is a  $m$ -dimensional pseudo-phase space;  
 $l_i$  is a length of the  $i$ -th diagonal line;  
 $K$  is a set of characteristics of recurrence and entropy analysis;  
 $N$  is a total number of points in the pseudo-phase space;  
 $N_l$  is a number of diagonal lines;  
 $n_{i,m}(\varepsilon)$  is a number of vectors, that similar vector  $P_m(i)$ ;  
 $P(l)$  is a frequency distribution of the diagonal lines lengths;  
 $RP_{i,j}$  is recurrence plot;  
 $RR$  is measure of recurrence;  
 $x(t)$  is a point of time series;

$x_i$  is a point in the reconstructed pseudo-phase space;  
 $\varepsilon$  is a neighborhood size;  
 $\phi$  is a autoregressive coefficient;  
 $\xi(t)$  is a uncorrelated white noise;  
 $\sigma_0$  is a diffusion coefficient;  
 $\tau$  is a delay period;  
 $\Theta(\cdot)$  is a Heaviside function.

#### INTRODUCTION

Most dynamical systems are «complex systems», which implies the ladder structure with nonlinear feedback. These include the processes inherent in the human body and nature, informational, physical, technical and social processes. In practice, they are represented by time series, which are a projection of the internal and external relations of the dynamical system. One of the objectives of time series analysis is to extract information from the series and infer the properties and mechanism of the process that generates the series.

Mathematical models of complex systems exhibiting irregular dynamics are both random and deterministic chaotic processes. Identification of the mechanism generating process based on characteristics obtained by time series is a daunting task. There are many approaches to the study of time series based on traditional statistical analysis, and the methods of nonlinear chaotic dynamics.

The object of study is the deterministic chaotic and stochastic fractal processes in the technical, economic and biological systems. The subject of study is the time series of a random type and the estimation methods of their characteristics. The purpose of the work is the following: based on the results recurrence and entropy analysis of fractal time series to identify the mechanism of generating process (deterministic or stochastic).

## 1 PROBLEM STATEMENT

Suppose given a time series of an irregular type  $X = \{x(t)\}, t=1, \dots, N$ . Let this time series have fractal properties. Let we have obtained the set of qualitative and quantitative characteristics  $K = \{K_i\}, i=1, \dots, m$  of the resulting recurrence and entropy analysis. Need to find out whether the process of generating this series is chaotic deterministic or random. For this is necessary to conduct a comparative analysis of modelling time series of various types and determine the set of characteristics  $K' \subset K$  for which the differences are significant

## 2 REVIEW OF THE LITERATURE

Most methods of chaotic dynamics used for time series analysis, based on the reconstruction space of single realization using the procedure Packard-Takens [1–4]. The reconstruction of the pseudo-phase space allows us to compute the embedding dimension, which is the main means of distinguishing chaotic and random processes [3, 5]. This approach allows us to well distinguish between chaotic dynamics and uncorrelated random noise, however, because this method is based on the estimation of the fractal dimension and detection autocorrelation relations, it has no effect for the fractal random processes having long dependence [6, 7].

In [8] proposed a method that extends the capabilities of nonlinear time series analysis, based on the fundamental property of dissipative dynamical systems – recurrence states. This method of analysis, based on the representation of process properties in the form of geometric structures, is a means for detection the hidden dependencies in the observed processes [9–12]. The method of recurrence plots is widely used for the analysis of stochastic time series of different nature [6, 13–16]. One of the characteristics of the complexity of the system behavior is entropy. Entropy methods of time series analysis are also used a reconstruction phase space [3, 7, 17, 18]. One of the characteristics that demonstrate the complexity of the time series dynamics is the approximate entropy of similarity introduced in [7].

## 3 MATERIALS AND METHODS

Consider the basic features of the recurrence and entropy analysis. The main idea of the application of nonlinear dynamics methods to the analysis of the realizations of a dynamical system is that the basic structure, which contains all the information about the system, namely, an attractor of a system, can be reconstructed by measuring only single component of this system [1, 3, 19]. Reconstruction phase space attractor is reduced to the construction of the pseudo-phase space. Widely used procedure Packard-Takens allows to restore the phase trajectory of a dynamical system from single realization:

$$F(t) = [x(t), x(t+\tau), \dots, x(t+m\tau)]. \quad (1)$$

One of the most common methods used in practice to determine the existence of chaotic determinacy and estimate the fractal dimension of the attractor is to study the properties of the correlation integral  $C_m(\varepsilon)$  and behavior of the correlation dimension  $d_C(m)$  depending on the dimension  $m$  of the pseudo-phase space. The correlation integral  $C_m(\varepsilon)$  is a probability that a pair of points on the reconstructed attractor in  $m$ -dimensional space is within a distance of  $\varepsilon$  each other:

$$C_m(\varepsilon) = \frac{1}{N^2} \lim_{\varepsilon \rightarrow 0} \sum_{i,j=1}^N \Theta(\varepsilon - \|x_i - x_j\|). \quad (2)$$

Dependence the correlation integral on  $\varepsilon$  at small  $\varepsilon$  obeys a power law, i.e.  $\lim_{\varepsilon \rightarrow 0} C_m(\varepsilon) = a\varepsilon^{d_C}$ . By increasing the dimension of the pseudo-phase space  $m$  correlation dimension  $d_C(m)$  increases too. However, for deterministic chaotic time series correlation dimension will ultimately be saturate with its true value. Value  $m$  at which  $d_C(m)$  will stop changing, is the embedding dimension. For uncorrelated stochastic realizations embedding dimension increases with the dimension of the pseudo-phase space  $m$ .

Recurrence plot is a projection of the  $m$ -dimensional pseudo-phase space onto the plane [12, 14, 17]. Let point  $x_i$  corresponds to the phase trajectory  $x(t)$  describing the dynamic system in the  $m$ -dimensional space at a time  $t = i$ , for  $i = 1, \dots, N$ , then the recurrence plot  $RP$  is array of pixels, where a nonzero element of the coordinates  $(i, j)$  corresponding to the case where the distance between  $x_j$  and  $x_i$  is smaller  $\varepsilon$ :

$$RP_{i,j} = \Theta(\varepsilon - \|x_i - x_j\|), \quad x_i, x_j \in R^m, i, j = 1, \dots, N. \quad (3)$$

The states  $x_j$  are recurrence if they are contained into the  $m$ -dimensional neighborhood of point  $x_i$  with size  $\varepsilon$ . Arbitrarily chosen recurrence point does not contain useful information about the state of the system at time moments  $i$  and  $j$ , only the totality of recurrence points allows you to restore the system properties. Analysis of the plot topology allows us to classify the observed processes: homogeneous processes with independent random values, processes with

slowly varying parameters, periodic or oscillating processes corresponding to nonlinear systems, etc.

Numerical analysis of recurrence plots allows us to calculate the measure of complexity structures of recurrence plots, such as a measure of recurrence and determinism etc. The measure of recurrence  $RR$  shows the density of recurrence points:

$$RR = \frac{1}{N^2} \sum_{i,j}^N RP_{i,j}. \quad (4)$$

Measure of determinism  $Det$  is a characteristic of predictability process and equal to the ratio of the number of points in diagonal lines to the total number of recurrence points:

$$Det = \sum_{l=l_{\min}}^N P(l) / \sum_{i,j}^N RP_{i,j}. \quad (5)$$

Approximate entropy  $ApEn$  is the statistics of time series regularity that defines the possibility of its forecasting. Time series that contain a many of duplicate values, have a relatively small value, and for less predictable process  $ApEn$  value is larger. Methods of estimating the approximate entropy  $ApEn$  considered in [7, 17].

Consider a time series  $\{x(t)\}$ ,  $t = 1, \dots, N$ . Let the vector  $P_m(i)$  is subsequence values  $\{x_i, x_{i+1}, \dots, x_{i+m}\}$  length of  $m$ . Two vectors  $P_m(i)$  и  $P_m(j)$  will be similar, if the following condition:

$$|x_{i+k} - x_{j+k}| < \varepsilon, \quad 0 \leq k < m.$$

For each  $i = 1, \dots, N-m+1$  value  $C_{im}(\varepsilon)$  is calculated:

$$C_{im}(\varepsilon) = \frac{n_{i,m}(\varepsilon)}{N-m+1}. \quad (6)$$

Approximate entropy  $ApEn$  determined by the formula

$$ApEn(m, \varepsilon) = \ln \frac{C_m(\varepsilon)}{C_{m+1}(\varepsilon)}, \quad (7)$$

$$\text{where } C_m(\varepsilon) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} C_{im}(\varepsilon).$$

Consider the basic model types of data needed to conduct research and their statistical properties. As input data have been chosen realizations of deterministic chaotic systems and realizations of stochastic processes having different correlation structure: uncorrelated noise, autoregressive processes with short-term dependence and fractal processes with long-term memory.

Chaos is a complex dynamics of deterministic systems in steady state. The main feature of such systems is sensitive dependence to arbitrarily small changes in initial conditions. If  $d_0$  is the initial distance between two points, then for short time  $t$  later the distance between the trajectories, which start from these points, becomes

$d(t) = d_0 e^{\lambda t}$ , where the value of  $\lambda$  is the Lyapunov exponent. This leads to the loss of deterministic predictability and the need to introduce probabilistic characteristics to describe the dynamics of chaotic systems.

Iterated maps  $x_{n+1} = f(C, x_n)$ , where  $C$  is control parameter, are the most simple and intuitive mathematical chaotic models [1, 3]. For a wide class of nonlinear functions  $f$  the sequence  $\{x_n\}_{n=0}^\infty$  is chaotic. In the case of dissipative map the orbits  $\{x_n\}_{n=0}^\infty$  lead to an attractor having a fractal structure.

Logistic map is the most famous example of chaotic maps. This one-dimensional quadratic map is defined as follows:

$$x_{n+1} = Ax_n(1-x_n) \quad A \in (0.4] \text{ and } x_n \in [0, 1]; \quad (8)$$

Diagram of the Lyapunov exponent  $\lambda$  is given in the upper part of fig. 1. Chaotic dynamics ( $\lambda > 0$ ) is observed when the parameter  $A > A^* = 3.569\dots$ . The regions of chaos alternate with «windows of stability» in which the dynamics becomes periodic. At the bottom of fig. 1 shows the time realizations of logistic map for parameter values  $A = 3, 7$  and  $A = 3, 9$ . The corresponding Lyapunov exponents are equal  $\lambda = 0, 37$  and  $\lambda = 0, 5$ .

Autoregressive process of 1st order was chosen as processes with short-term dependence [5]:

$$X(t) = \phi X(t-1) + \xi(t), \quad |\phi| < 1 \quad (9)$$

Autoregressive coefficient value  $\phi$  characterizes the degree of the autocorrelation process. Fig. 2 shows the realizations of the autoregressive process of the different values of coefficient  $\phi$ .

At present it has been generally accepted, that many stochastic processes in nature and in engineering exhibit a long-range dependence and fractal structure [20, 21]. Stochastic process  $X(t)$  is self-similar with self-similarity parameter  $H$ , if the process  $a^{-H} X(at)$  is described by the same finite-dimensional distributions that  $X(t)$ . One of the most famous and simple models of stochastic dynamics that have fractal properties, is the fractional Brownian motion (FBM).

Gaussian process  $X(t)$  with a parameter  $H$ ,  $0 < H < 1$  called FBM if its increments  $\Delta X(\tau) = X(t+\tau) - X(t)$  have a distribution of the form:

$$P(\Delta X < x) = \frac{1}{\sqrt{2\pi}\sigma_0\tau^H} \cdot \int_{-\infty}^x \text{Exp}\left[-\frac{z^2}{2\sigma_0^2\tau^{2H}}\right] dz. \quad (10)$$

FBM with the parameter  $H = 0, 5$  coincides with the classical Brownian motion. Parameter  $H$  called the Hurst exponent, is the degree of self-similarity. Along with this property, the index characterizes the measure of long-term dependence of a stochastic process, i.e. that autocorrelation function  $r(k)$  decreases as a power law:  $r(k) \sim k^{-\beta}$ ,  $k \rightarrow \infty$ , where  $0 < \beta < 1$  and  $H = 1 - (\beta/2)$ . Fig. 3 shows the realizations of the FBM for the values  $H = 0, 3, 0, 5, 0, 8$ .

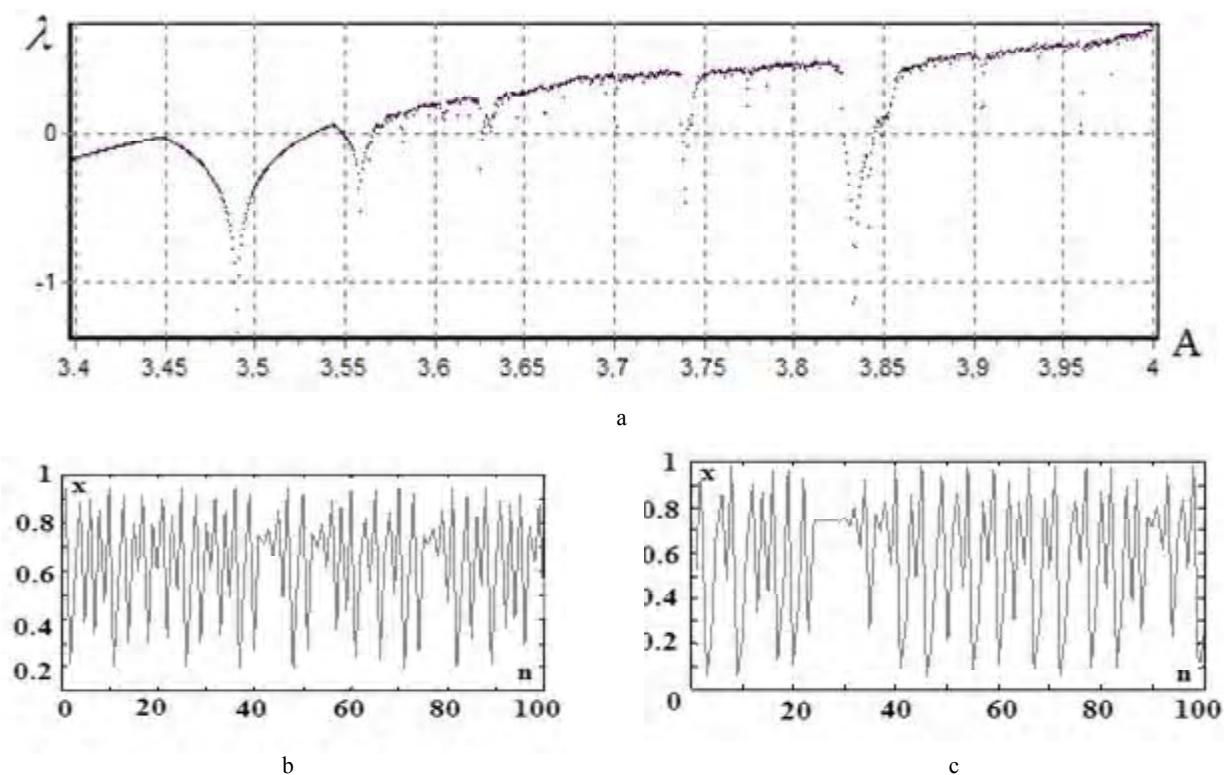


Figure 1 – Diagram of the Lyapunov exponent and realizations for logistic map: a – diagram of the Lyapunov exponent, b – realization with  $A=3.7$ , c – realization with  $A=3.9$

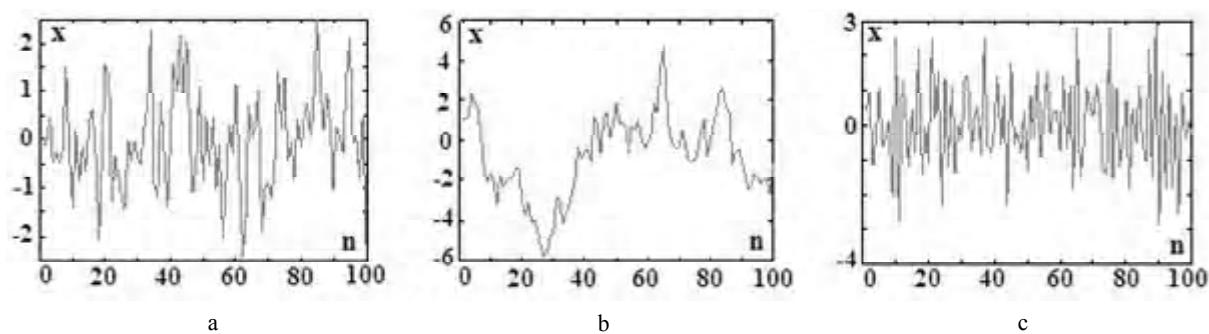


Figure 2 – The realizations of autoregressive process for different  $\phi$ : a –  $\phi = 0.2$ , b –  $\phi = 0.9$ ; c –  $\phi = -0.7$

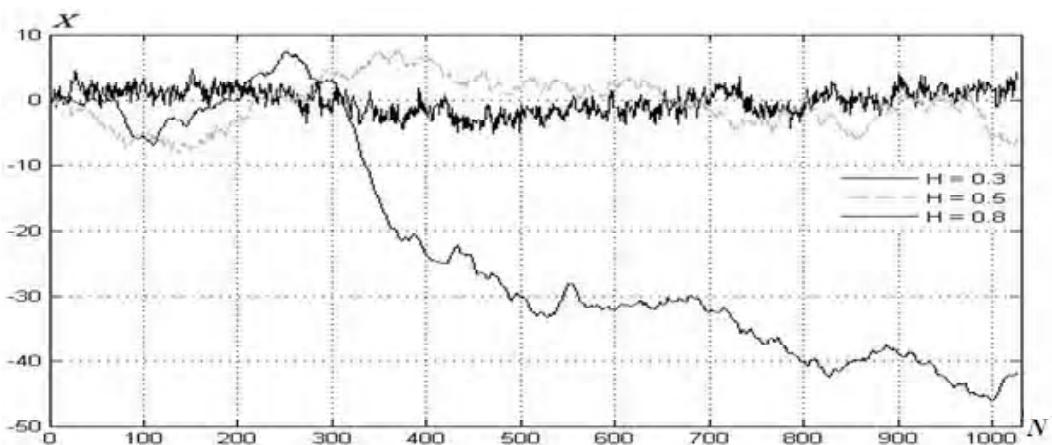


Figure 3 – FBM for different values  $H$

## 4 EXPERIMENTS

For carrying out numerical experiments investigated time realizations  $X = \{x_i\}$ ,  $i = 1, \dots, N$ , were generated according to (8–10) for various  $N$ . It was first performed the reconstruction of pseudo-phase space and estimation of embedding dimension. For this procedure were used time realizations of  $N = 10000$ . Smaller length of time series usually leads to large errors [1–3].

For carrying out entropy and recurrence analysis were used time realizations of  $N = 1000$ . This length is sufficient for good visualization of recurrence plots and small errors in the quantitative characteristics (about  $\pm 0.005$ ). For each of the generated realization estimates of approximate entropy and a number of the recurrent characteristics were obtained. Estimation procedure was carried out for realizations of every type processes  $M = 100$  times and each estimate was averaged over the  $M$  values.

For clarity, we first considered the example of a completely different process on complexity: a periodic motion and uncorrelated white noise. For chaotic processes the realizations with different values of Lyapunov exponents were investigated. For autoregressive processes the autoregressive coefficient value was changed. The realizations of FBM were generated for different values of the Hurst exponent.

After examining the results of the analysis the modeling realizations the entropy and recurrence analysis of real time series such as bioelectrical signals and financial series was performed.

## 5 RESULTS

Consider the reconstruction of pseudo-phase space and estimation of embedding dimension. Fig. 4 shows the typical dependence of the correlation dimension  $d_C(m)$  on the dimension  $m$  of pseudo-phase space constructed in accordance with (1) for the realizations of an autoregressive process, the logistic map and FBM.

The comparative entropy and recurrence analysis of chaotic realizations and realizations of the stochastic processes having different correlation structure was carried out. Fig. 5 shows the recurrence plots for the sum of two sinusoids and independent values of a normal random variable. Table 1 shows the corresponding values of the measures of recurrence  $RR$ , determinism  $Det$  and approximate entropy  $ApEn$ .

Recurrence plots for realizations of map (8), autoregression and FBM with different values of parameters are presented on fig. 6. In the case of logistic map the Lyapunov exponents are equal to  $\lambda = 0.37, 0.5, 0.69$  according to parameter values.

Table 2 shows the means of recurrence  $RR$ , determinism  $Det$  and approximate entropy  $ApEn$  corresponding to the plots above.

Table 1 – Quantitative characteristics of complexity of sinusoid and uncorrelated noise

	$RR$	$Det$	$ApEn$
Sinusoid	0.18	0.998	0.03
Uncorrelated noise	0.0003	0.025	1.7

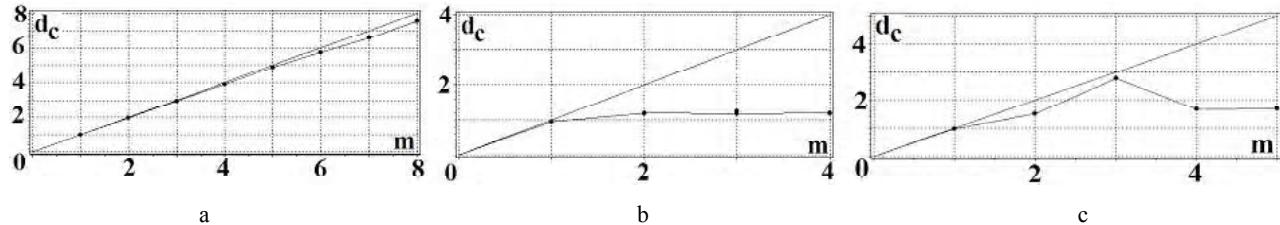


Figure 4 – Dependence of  $d_C(m)$  on  $m$  for different processes: a – autoregression, b – chaotic map, c – FBM

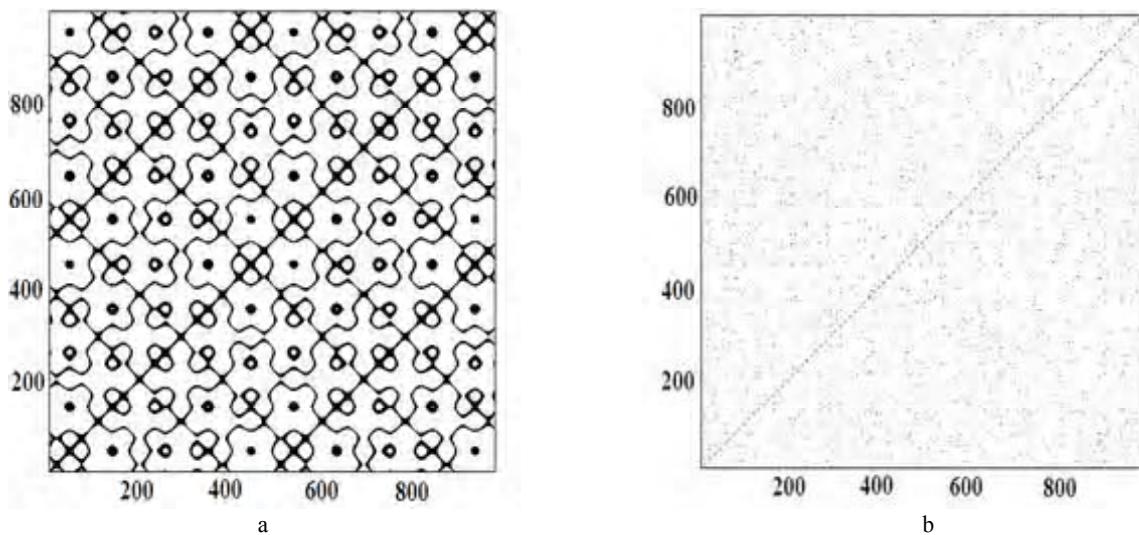


Figure 5 – Recurrence plots for sinusoid and noise: a – sum of sinusoids; b – uncorrelated noise

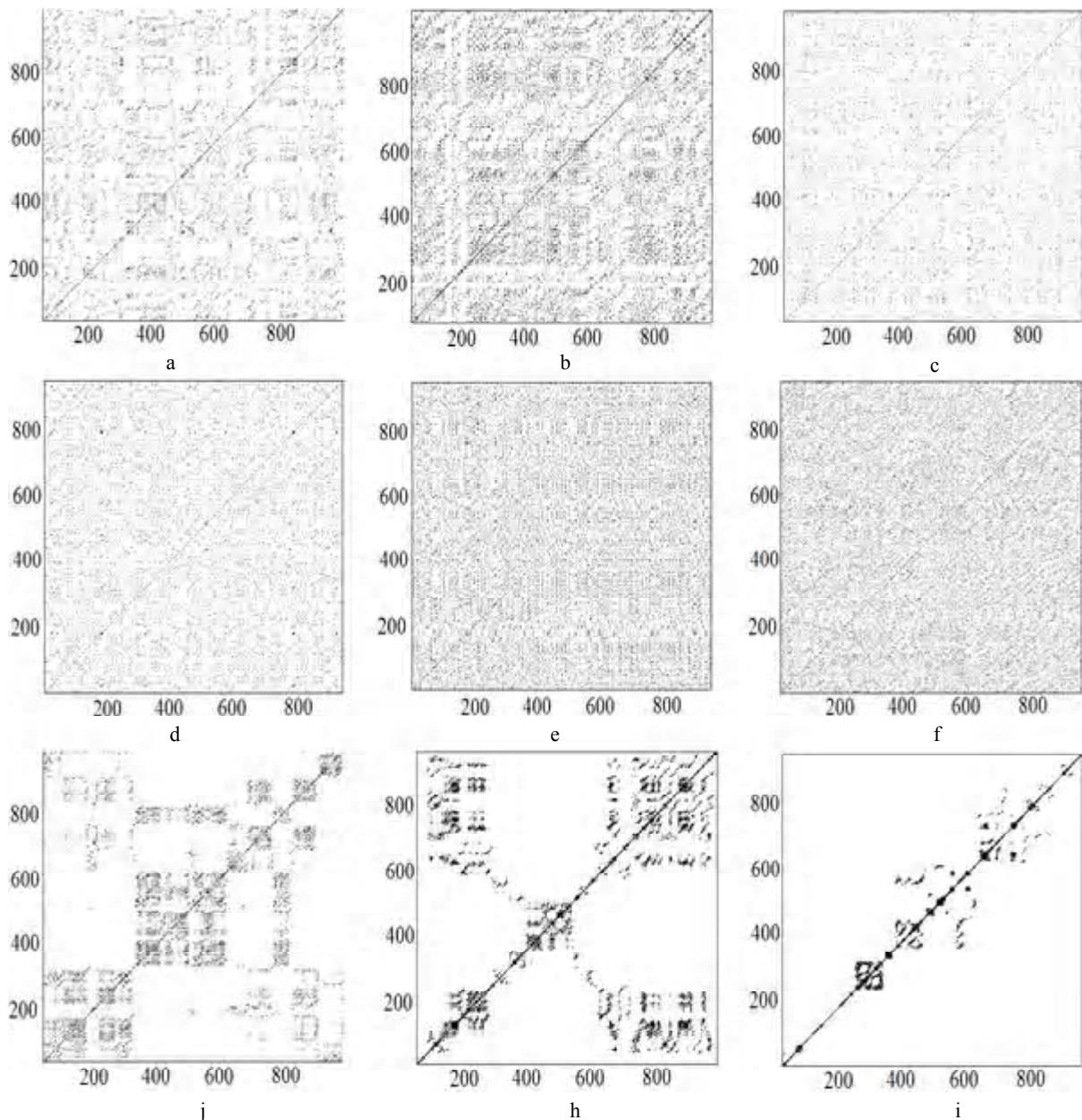


Figure 6 –Recurrence plots for under study realizations: a – logistic map with  $A=3.7$ , b –logistic map with  $A=3.9$ , c – logistic map with  $A=4$ , d – autoregression with  $\phi = 0.3$ , e –autoregression with  $\phi = 0.6$ , f – autoregression with  $\phi = 0.9$ , j – FBM with  $H = 0.3$ , h – FBM with  $H = 0.6$ , i – FBM with  $H = 0.9$

Table 2 – Quantitative characteristics of complexity of realizations

Logistic map				Autoregression				FBM			
$A$	$RR$	$Det$	$ApEn$	$\varphi$	$RR$	$Det$	$ApEn$	$H$	$RR$	$Det$	$ApEn$
3.7	0.008	0.1	0.93	0.3	0.0003	0.03	1.72	0.3	0.02	0.55	0.47
3.9	0.004	0.07	1.2	0.6	0.0005	0.05	1.65	0.6	0.02	0.87	0.21
4	0.002	0.05	0.86	0.9	0.002	0.13	1.25	0.9	0.01	0.95	0.12

In this work the time series corresponding to a various complex dynamical systems: bioelectrical signals and financial series were considered. In particular, the RR-intervals series were investigated. RR-interval is the time interval between adjacent teeth of electrocardiogram and it equals to the duration of the cardiac cycle. Initial data for

the research in this paper were obtained on a specialized website [24] containing an extensive medical database. As an example of financial series, the dynamics of change in the currency pair EUR/RUB for 2004–2006 was examined. Fig. 7 shows the time series and recurrence plots of data described above.

Quantitative recurrence and entropy characteristics obtained from the time series are presented in Table 3.

## 6 DISCUSSION

Numerical analysis shows that the realizations of the random and deterministic chaotic motion may have similar statistical characteristics [3, 22, 23]. Reconstruction of pseudo-phase space and estimations of embedding dimension detected essential differences in the structure of chaotic realizations and realizations autoregressive processes with short-term dependence. However, the embedding dimension, evaluated for the FBM realizations with a long-term dependence, is also limited [6, 7]. The estimation results presented in fig. 4 confirm that the construction of the pseudo-phase space and the estimation of embedding dimension cannot be a reliable tool for distinguishing between chaotic and stochastic fractal realizations and fitting of appropriate mathematical models.

Carried out recurrent analysis detected strong differences in visual topology and the numerical characteristics of realizations of the above processes. It is obvious that the characteristics of chaotic and random processes must be located within the range of characteristic values calculated for the periodic and completely random trajectories, see fig.

Table 3 – Quantitative characteristics of complexity of time series

	<i>RR</i>	<i>Det</i>	<i>ApEn</i>
RR-intervals	0.05	0.61	1.07
EUR/RUB	0.08	0.85	0.17

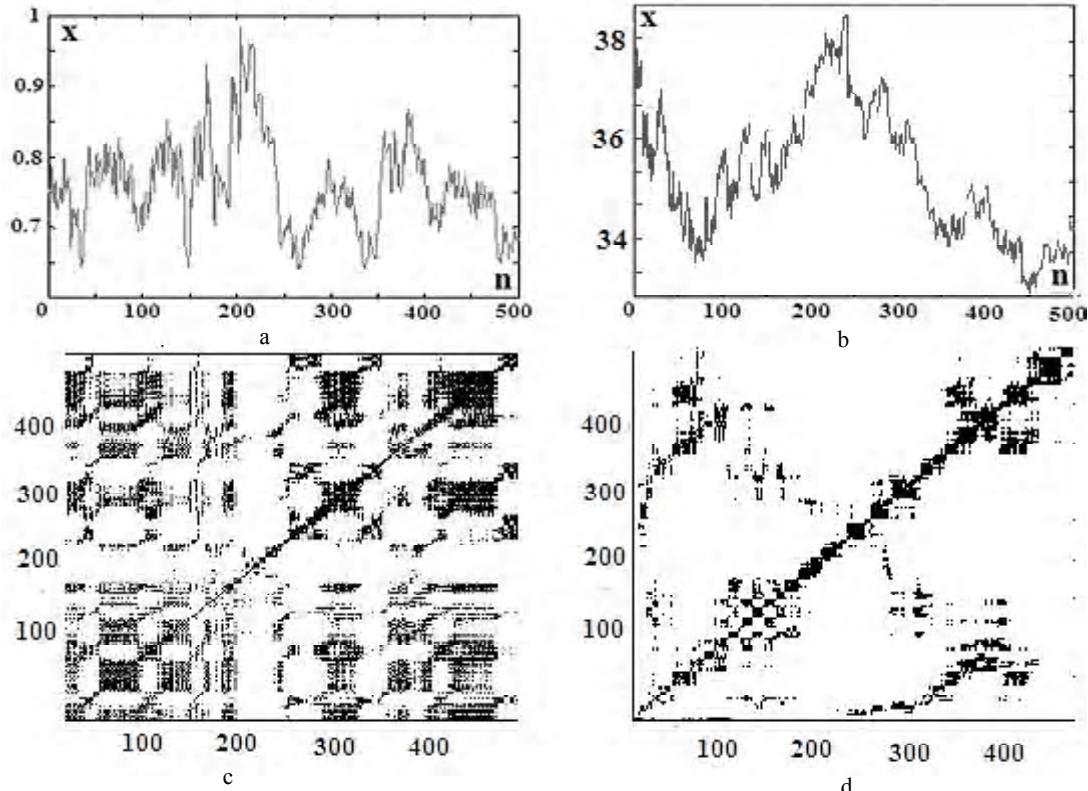


Figure 7 – Recurrence plots of real time series: a – series of RR-intervals, b – series of EUR/RUB, c – recurrence plots of RR-intervals, d – recurrence plots of EUR/RUB

5 and tab. 1. It can be noted for chaotic realizations that greater value of Lyapunov exponent corresponds to a greater randomness of the system, which is clearly evident on recursive plots: the existence of some plot structure replaced uniform filling (top of Fig. 6). In the case of autoregressive process (middle Fig. 6) it is necessary to note the lack of plot structure and uniform filling regardless of the autocorrelation degree. The recurrent plots of FBM have the specific structure, which depends on the Hurst exponent value (bottom of Fig. 6). With the increasing exponent  $H$  the range of values, i.e. plot filling, decreases.

As regards the quantitative characteristics, the research has shown that the most informative recurrent characteristics are the indexes of recurrence and determinism. The values *RR* and *Det* are measure of regularity, therefore in each case they decrease, when randomness or uncorrelation of realizations increase. The entropy *ApEn* is measure of unpredictability therefore it increases with uncorrelation. Value ranges of characteristics are quite different for various processes. This allows us to identify the generating process by the set of characteristics.

Based on the results of qualitative and quantitative analysis can be propose for modeling realizations RR-intervals to use deterministic chaotic systems, while the mathematical modeling of S&P500 series should be based on self-similar stochastic processes. For a correct choice of the model in the first case the estimation of such characteristics as the Lyapunov exponent, invariant measure

distribution, etc. is necessary, and in the second case – the estimation of fractal characteristics.

## CONCLUSION

Using the results of the recurrent and entropy analysis to distinguish deterministic chaotic and fractal random processes was first proposed in this work. It is shown that the set of characteristics such as indexes of recurrence and determinism, approximate entropy and recurrence plot allows to identify the type of process that generated the time series. The dependences of information complexity measures of time series from the parameters of the processes were obtained. Thus it is possible to choose the mathematical model of process has a certain correlation and recursive structure for the simulation and forecasting. It is shown that series of RR-intervals corresponds to a chaotic process and S&P500 series has the structure corresponding to a fractal Brownian motion. Further studies propose the calculation of confidence intervals for estimates of the characteristics, the analysis of the short time series and investigation a large number of real time series a various complex dynamical systems.

## ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Kharkiv National University of Radioelectronics «Evolutionary hybrid systems of computational intelligence with variable structure for data mining» (state registration number 0108U009112).

## REFERENCES

1. Mun F. Chaotic Vibrations / F. Mun. – New York : Wiley, 1987. – 309 p.
2. Peitgen H. Chaos and Fractals. New Frontiers of Science / H.-O. Peitgen, H. Jorgens, D. Saupe. – New York : Springer, 2004. – 895 p.
3. Shuster H. G. Deterministic Chaos: An Introduction, second revised edition / H. G. Shuster. – New York : VCH Publishers, 1988. – 248 p.
4. Takens F. Detecting strange attractors in turbulence / F. Takens, D. A. Rand, L.-S. Young // Dynamical Systems and Turbulence. – New York : Springer-Verlag, 1981. – Vol. 898. – P. 366–381.
5. Ширяев А. Н. Основы стохастической финансовой математики. Том 1: Факты. Модели / А. Н. Ширяев. – М. : Фазис, 1998. – 512 с.
6. Dabi-Prashad O. Investigation of Time Series of Original Values of Currency Rates Measured on Small Time Frames on FOREX Using Methods of Chaos Theory / O. Dabi-Prashad, L. Kirichenko // Radioelectronics & Informatics. – 2009. – No. 4. – P. 18–24.
7. Pincus S. M. Approximate entropy as a measure of system complexity / S. M. Pincus // Proceedings of the National Academy of Sciences. – 1991. – Vol. 88. – P. 2297–2301.
8. Eckmann J. P. Recurrence Plots of Dynamical Systems / J. P. Eckmann, S. O. Kamphorst, D. Ruelle // Europhysics Letters. – 1987. – No. 4. – P. 973–977.
9. March T. K. Recurrence plot statistics and the effect of embedding / T. K. March, S. C. Chapman, R. O. Dendy // Physica D: Nonlinear Phenomena. – Netherlands : Elsevier, 2005. – P. 171–184.
10. Marwan N. Recurrence-plots-based measures of complexity and application to heart-rate-variability data / [Marwan N., Wessel N., Meyerfeldt U.] // Physical Review E. – 2002. – Vol. 66. – P. 855–874.
11. Соловйов В. М. Моделювання явищ самоорганізації в фінансово-економічних системах / В. М. Соловйов, В. Д. Дербенцев, О. Д. Шарапов // Моделювання та інформаційні системи в економіці: Міжвідомчий наук. збірник. – К. : КНЕУ, 2003. – Вип. 70. – С. 74–85.
12. Владимирский Э. И. Синергетические методы управления хаотическими системами / Э. И. Владимирский, Б. И. Исмайлова. – Баку : ЭЛМ, 2011. – 240 с.
13. Iwanski J. S. Recurrence plots of experimental data: To embed or not to embed? / J. S. Iwanski, E. Bredley // Chaos: An Interdisciplinary Journal of Nonlinear Science. – 1998. – Vol. 8, No. 4. – P. 861–871.
14. Recurrence plots for the analysis of complex system / [N. Marwan, M. Romano, M. Thiel, J. Kurths] // Physics Reports. – 2007. – Vol. 438. – P. 237–329.
15. Zbilut J. P. Recurrence quantification based Liapunov exponent for monitoring divergence in experimental data / J. P. Zbilut, J.-M. Zaldívar-Comenges, F. Stozzi // Physics Letter A. – 2002. – Vol. 297. – P. 173–181.
16. Васюта К. С. Рекуррентный анализ процессов в телекоммуникационных системах / К. С. Васюта // Наукові записки УНДІЗ. – Київ : ДП УНДІЗ, 2008. – № 6(8). – С.90–96.
17. Дербенцев В. Д. Синергетичні та еконофізичні методи дослідження динамічних та структурних характеристик економічних систем: монографія / В. Д. Дербенцев, О. А. Сердюк, В. М. Соловйов, О. Д. Шарапов. – Черкаси : Брама-Україна, 2010. – 287 с.
18. Joshua S. Physiological time-series analysis using approximate entropy and sample entropy / S. Joshua, J. Richman, R. Moorman // American Journal of Physiology – Heart and Circulatory Physiology. – 2000. – Vol. 278, No. 6. – P. 2039–2049.
19. Packard N.H. Geometry from a time series / N. H. Packard, J. P. Crutchfield, J. D. Farmer, R. S. Shaw // Physical Review Letters. – 1980. – Vol. 45. – P. 712–716.
20. Cronover R. M. Introduction to fractals and chaos / R. M. Cronover. – New York : John and Bartlett Publishers, 2000. – 306 p.
21. Feder J. Fractals / J. Feder. – New York : Plenum, 1987. – 283 p.
22. Comparative analysis of pseudorandom number generation in the up-to-date wireless data communication / [Kirichenko L., Tsekhmistro R., Krug O., Storozhenko A.] // Telecommunications and Radio Engineering. – 2011. – No. 70 (4). – P. 325 333.
23. Кириченко Л. О. Сравнительный рекуррентный анализ хаотических и случайных процессов / Л. О. Кириченко, Ю. А. Кобицкая, А. Ю. Хабачева // Физико-технологические проблемы радиотехнических приборов, средств телекоммуникаций,nano- и микроэлектроники: Черновцы, 24–26 октября 2013 г.: материалы конференции. – Черновцы: ЧНУ им. Ю. Федьковича. 2013 – С. 21.
24. PhysioNet the research resource for complex physiologic signals. [Electronic resource]. – Access mode: [www.physionet.org](http://www.physionet.org).

Article was submitted 12.09.2014.  
After revision 03.11.2014.

Кириченко Л. О.<sup>1</sup>, Кобицкая Ю. А.<sup>2</sup>, Хабачева А. Ю.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, доцент, профессор кафедры прикладной математики, Харьковский национальный университет радиоэлектроники, Украина

<sup>2</sup>Аспирант кафедры прикладной математики, Харьковский национальный университет радиоэлектроники, Украина

<sup>3</sup>Магистрант кафедры прикладной математики, Харьковский национальный университет радиоэлектроники, Украина

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ СЛОЖНОСТИ ХАОСТИЧЕСКИХ И СТОХАСТИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ

Предложен новый подход к распознаванию механизма процесса, породившего временной ряд, базирующийся на результатах энтропийного и рекуррентного анализа. Проведен сравнительный анализ свойств реализаций хаотических и стохастических процессов, имеющих различную корреляционную структуру. Показано, что полученное множество характеристик информационной сложности позволяет различать реализации детерминированных хаотических и фрактальных случайных процессов. Получены зависимости информационных характеристик от параметров процессов. Приведены результаты исследования биоэлектрических сигналов и финансовых рядов.

**Ключевые слова:** временной ряд, мера сложности, энтропия подобия, рекуррентная диаграмма, псевдо-фазовое пространство, размерность вложения.

Кіріченко Л. О.<sup>1</sup>, Кобицкая Ю. О.<sup>2</sup>, Хабачова А. Ю.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, доцент, профессор кафедри прикладної математики, Харківський національний університет радіоелектроніки, Україна

<sup>2</sup>Аспірант кафедри прикладної математики, Харківський національний університет радіоелектроніки, Україна

<sup>3</sup>Магістрант кафедри прикладної математики, Харківський національний університет радіоелектроніки, Україна

## ПОРІВНЯЛЬНИЙ АНАЛІЗ СКЛАДНОСТІ ХАОСТИЧНИХ ТА СТОХАСТИЧНИХ ЧАСОВИХ РЯДІВ

Запропоновано новий підхід до розпізнавання механізму процесу, що генерує часовий ряд, який базується на результатах ентропійного і рекуррентного аналізу. Проведено порівняльний аналіз властивостей реалізацій хаотичних та стохастичних процесів, що мають різну кореляційну структуру. Показано, що отримана множина характеристик інформаційної складності дозволяє розрізняти реалізації детермінованих хаотичних і фрактальних випадкових процесів. Отримано залежності інформаційних характеристик від параметрів процесів. Наведено результати дослідження біоелектрических сигналів і фінансових рядів.

**Ключові слова:** часовий ряд, міра складності, ентропія подібності, рекуррентна діаграма, псевдо-фазовий простір, розмірність вкладення.

## REFERENCES

1. Mun F. Chaotic Vibrations. New York, Wiley, 1987, 309 p.
2. Peitgen H.-O., Jorgens H., Saupe D. Chaos and Fractals. New Frontiers of Science. New York, Springer, 2004, 895 p.
3. Shuster H. G. Deterministic Chaos: An Introduction, second revised edition. New York, VCH Publishers, 1988, 248 p.
4. Takens F., Rand D. A., Young L.-S. Detecting strange attractors in turbulence, *Dynamical Systems and Turbulence*. New York, Springer-Verlag, 1981, Vol. 898, pp. 366–381.
5. Shiryaev A. N. Osnovy stohasticheskoy finansovoy matematiki. Tom 1: Fakty. Modeli. Moskva, Fazis, 1998, 512 p.
6. Dabi-Prashad O., Kirichenko L. Investigation of Time Series of Original Values of Currency Rates Measured on Small Time Frames on FOREX Using Methods of Chaos Theory, *Radioelectronics & Informatics*, 2009, No. 4, pp. 18–24.
7. Pincus S. M. Approximate entropy as a measure of system complexity, *Proceedings of the National Academy of Sciences*, 1991, Vol. 88, pp. 2297–2301.
8. Eckmann J. P., Kamphorst S. O., Ruelle D. Recurrence Plots of Dynamical Systems, *Europhysics Letters*, 1987, No. 4, pp. 973–977.
9. March T. K. Chapman S. C., Dendy R. O. Recurrence plot statistics and the effect of embedding, *Physica D: Nonlinear Phenomena*. Netherlands: Elsevier, 2005, pp. 171–184.
10. Marwan N., Wessel N., Meyerfeldt U., Schirdewan A., Kurths J. Recurrence-plots-based measures of complexity and application to heart-rate-variability data, *Physical Review E*, 2002, Vol. 66, pp. 855–874.
11. Soloviov V. M., Derbencov V. D., Sharapov O. D. Modeljuvannja javyshh samoorganizacij' v finansovo-ekonomichnyh sistemah. *Modeljuvannja ta informacijni sistemy v ekonomici: Mizhvidomchyy nauk. zbirnyk*. Kyiv, KNEU, 2003, Vyp.70, pp. 74–85.
12. Vladimirovskij Je. I., Ismajlov B. I. Sinergeticheskie metody upravlenija haoticheskimi sistemami. Baku, JeLM, 2011, 240 p.
13. Iwanski J. S., Bredley E. Recurrence plots of experimental data: To embed or not to embed? Chaos: An Interdisciplinary Journal of Nonlinear Science, 1998, Vol. 8, No. 4, pp. 861–871.
14. Marwan N., Romano M., Thiel M., Kurths J. Recurrence plots for the analysis of complex system, *Physics Reports*, 2007, Vol. 438, pp. 237–329.
15. Zbilut J. P., Zaldívar-Comenges J.-M., Stozzi F. Recurrence quantification based Liapunov exponent for monitoring divergence in experimental data, *Physics Letter A*, 2002, Vol. 297, pp. 173–181.
16. Vasylta K. S. Rekurrentnyj analiz processov v telekommunikacionnyx sistemakh, *Naukovyi zapiski UNDIZ*. Kiiv, DP UNDIZ, 2008, No. 6(8), pp. 90–96.
17. Derbencov V. D. Serdyuk O. A., Soloviov V. M., Sharapov O. D. Sinergetichni ta ekonofofizichni metodi doslidzhennya dinamichnix ta strukturnix xarakteristik ekonomichnix sistem: monografiya. Cherkasi, Brama-Ukraina, 2010, 287 p.
18. Joshua S., Richman J., Moorman R. Physiological time-series analysis using approximate entropy and sample entropy, *American Journal of Physiology – Heart and Circulatory Physiology*, 2000, Vol. 278, No. 6, pp. 2039–2049.
19. Packard N. H., Crutchfield J. P., Farmer J. D., Shaw R. S. Geometry from a time series, *Physical Review Letters*, 1980, Vol. 45, pp. 712–716.
20. Cronover R. M. Introduction to fractals and chaos. New York, John and Bartlett Publishers, 2000, 306 p.
21. Feder J. Fractals. New York, Plenum, 1987, 283 p.
22. Kirichenko L., Tsekhmistro R., Krug O., Storozhenko A. Comparative analysis of pseudorandom number generation in the up-to-date wireless data communication, *Telecommunications and Radio Engineering*, 2011, No. 70 (4), pp. 325–333.
23. Kirichenko L. O., Kobickaja Ju. A., Habacheva A. Ju. Sravnitel'nyj rekurrentnyj analiz haoticheskikh i sluchajnyh processov, Fiziko-tehnologicheskie problemy radiotekhnicheskikh priborov, sredstv telekomunikacij, nano- i mikroelektroniki: Chernovcy, 24–26 oktyabrya 2013 g.: materialy konferencii. Chernovcy, ChNU im. Ju. Fed'kovicha, 2013, P. 21.
24. PhysioNet the research resource for complex physiologic signals. [Electronic resource]. – Access mode: [www.physionet.org](http://www.physionet.org).

Канд. техн. наук, доцент, доцент кафедры информатики и компьютерных наук, Херсонский национальный технический университет, Украина

## ОПТИМИЗАЦИЯ ПРОЦЕССА ПРЕДОБРАБОТКИ ИНФОРМАЦИИ В СИСТЕМАХ КЛАСТЕРИЗАЦИИ ВЫСОКОРАЗМЕРНЫХ ДАННЫХ

Представлена методика выбора оптимального метода нормализации при построении кластерной структуры объектов, отличительной особенностью которых является высокая размерность признакового пространства. В качестве основного критерия оценки качества предобработки данных использовался критерий энтропия Шеннона и относительное изменение энтропии в процессе трансформации данных. Понижение размерности признакового пространства исследуемых объектов производилось при помощи компонентного анализа. Построена модель системы кластеризации с использованием алгоритма нечеткой кластеризации fuzzy C-means, при помощи которой произведена оценка качества кластеризации при использовании различных методов предобработки данных. Показано, что для исследуемых данных наилучшим методом нормализации является метод десятичного масштабирования, при котором энтропия обработанного сигнала принимает наименьшее значение, при этом в процессе трансформации данных компонентным анализом относительное изменение энтропии не превышает допустимых норм.

**Ключевые слова:** Кластеризация, размерность признакового пространства, нормализация, энтропия.

### НОМЕНКЛАТУРА

ДНК – дезоксирибонуклеиновая кислота;  
ALL – острая лимфобластная лейкемия;  
AML – острая миелоидная лейкемия;  
 $A_i$  – аппроксимирующие коэффициенты на  $i$ -м уровне вейвлет-декомпозиции сигнала;  
СКМ система компьютерной математики;  
 $D_i$  – диагональные детализирующие коэффициенты на  $i$ -м уровне вейвлет-декомпозиции сигнала;  
 $E_{0i}$  – энтропия нормализованных признаков  $i$ -го объекта;  
 $E_i$  – энтропия выделенных главных компонент  $i$ -го объекта;  
 $\sigma(i)$  – стандартное отклонение значений в  $i$ -й строке;  
 $H_i$  – горизонтальные детализирующие коэффициенты на  $i$ -м уровне вейвлет-декомпозиции сигнала;  
 $i$  – количество строк в массиве данных;  
 $j$  – количество признаков в  $i$ -й строке;  
 $K$  – положительная константа для согласования размерности;  
 $m$  – количество признаков  $i$ -го объекта;  
 $n$  – количество объектов;  
 $p_i$  – вероятность осуществления  $i$ -го события;  
sym4 – вейвлет симплет-4;  
 $V_i$  – вертикальные детализирующие коэффициенты на  $i$ -м уровне вейвлет-декомпозиции сигнала;  
 $XN(i, j)$  – нормализованное значение  $j$ -го элемента в  $i$ -й строке;  
 $X(i, j)$  – исходное значение  $j$ -го элемента в  $i$ -й строке;  
 $X_{\max}(i)$  – максимальное значение в  $i$ -й строке;

$X_{\min}(i)$  – минимальное значение в  $i$ -й строке;  
 $\bar{X}_{us}(i)$  – усеченное среднее в  $i$ -й строке (отброшены 2 % максимальных и минимальных значений);  
 $\bar{X}_{op}$  – среднее значение элементов в априори выбранном опорном векторе исходного массива данных;  
 $X(op, j)$  – значение  $j$ -го элемента в опорном векторе массива;  
 $\hat{M}(i, j)$  – значение функции регрессии, соответствующее  $j$ -му признаку  $i$ -й строки массива.

### ВВЕДЕНИЕ

В настоящее время во многих областях научных исследований возникает необходимость в разработке моделей и систем кластеризации объектов сложной природы, отличительной особенностью которых является высокая размерность признакового пространства. Одним из основных методов, используемых в данное время для сокращения размерности пространства признаков, является метод главных компонент, позволяющий существенно уменьшить количество признаков, характеризующих объект при максимальном сохранении полезной информации об объекте. Основной проблемой, возникающей при использовании данного метода, является его высокая чувствительность к методам предобработки данных, вследствие чего характер разбиения объектов на кластеры будет существенно различаться в зависимости от комбинации используемых методов предобработки. Одним из ключевых методов предобработки данных является их нормализация, в результате которой происходит приведение исходных данных к требуемому ди-

пазону и распределению. Вследствие этого возникает актуальная проблема создания методики обоснованного выбора метода нормализации высокоразмерных данных сложной природы, позволяющей при дальнейшем сокращении размерности признакового пространства максимизировать объективность процесса кластеризации исследуемых объектов.

## 1 ПОСТАНОВКА ЗАДАЧИ

При построении модели исходные данные представляются в виде матрицы, строками в которой являются исследуемые объекты, а столбцами – признаки, характеризующие соответствующий объект (1):

$$X = \{x_{ij}\}, i = 1, \dots, n, j = 1, \dots, m. \quad (1)$$

Выбор набора признаков осуществляется в результате сужающего отображения:

$$\{X^m\} \xrightarrow{F} \{X^k\}, k < m, \quad (2)$$

при котором достигается экстремум некоторого критерия качества  $J_X(F)$ .  $F$  в (2) представляет собой функционал преобразования множества  $\{X^m\}$  в множество  $\{X^k\}$ ,  $k$  – размерность нового признакового пространства. Каждому отображению (2) ставится в соответствие некоторое значение критерия  $J_X(F)$ . Результатом полученных отображений является функция  $g(J_X)$ . Задача сокращения размерности признакового пространства заключается в нахождении такого отображения, при котором достигается экстремум функции  $g(J_X)$ . В данной работе для решения поставленной задачи в качестве функции  $g(J_X)$  предлагается использовать зависимость критерия энтропии Шеннона от используемого метода нормализации [1], определяемую по формуле (3):

$$H = -K \cdot \sum_{i=1}^n p_i \ln(p_i). \quad (3)$$

Значения критерия энтропии вычисляются для оригинального ненормализованного сигнала, после нормализации, и для главных компонент исследуемых объектов. В соответствии с принципом максимума энтропии [2]: «Для данных знаний в форме ограничений существует только одно распределение, удовлетворяющее этим ограничениям, которое можно выбрать с помощью процедуры, удовлетворяющей «аксиомам согласованности». Это уникальное распределение определяется максимизацией энтропии». В соответствии с вышеизложенным предлагается следующая методика проведения эксперимента:

- фильтрация исходных данных с целью минимизации шумовой компоненты исследуемых данных;
- вычисление энтропии Шеннона ненормализованного сигнала;

- нормализация данных различными методами нормализации;
- вычисление энтропии нормализованных данных;
- вычисление главных компонент каждой матрицы нормализованных данных;
- вычисление энтропии векторов главных компонент исследуемых объектов;
- кластеризация тестового множества исследуемых объектов методом fuzzy C-means;
- построение графиков полученных результатов, их интерпретация и анализ.

## 2 ОБЗОР ЛІТЕРАТУРИ

Анализ публикаций по обозначенной проблеме [3–7] показывает, что большинство методов и алгоритмов кластеризации, используемых в настоящее время в различных областях человеческой деятельности, ориентированы на небольшую размерность вектора признаков исследуемого объекта (не более 1000). В [8, 9] процесс кластеризации представляется в виде модели, что позволило перенести в теорию кластерного анализа все основные методы теории самоорганизации моделей на основе метода группового учета аргументов, а именно: многоэтапность поиска лучшей кластеризации; критериальный подход для оценки качества кластеризации; использование методов формирования признакового пространства и формирования кластеров; выбор мер сходства между объектами, кластерами и объектом и кластером. В последнее время с развитием методов биоинформатики созданы базы данных клеток биологических объектов, особенностью которых является высокая размерность ( $\approx 80000$ ) и высокая степень зашумленности, определяемая биологическими и технологическими факторами, возникающими в процессе подготовки и проведения эксперимента по их формированию [10, 11]. Вследствие этого возникает необходимость в разработке эффективных для данного типа данных методов их предобработки и сокращения размерности признакового пространства без существенной потери информации об исследуемых объектах. Методам нормализации высокоразмерных данных биологической природы посвящены работы [12, 13]. В общем случае все методы нормализации можно разделить на две подгруппы: методы, использующие опорное множество эталонного объекта и методы, использующие всю совокупность исследуемых данных. В первом случае методы нормализации подразделяются на линейные и нелинейные, во втором случае используют метод циклической локальной регрессии, метод контрастов и квантильную нормализацию. В работе [13] представлены результаты сравнительного анализа существующих методов нормализации на примере выборок данных, взятых с двух микромассивов ДНК. Результаты их исследований показали, что методы, не использующие эталон, дают лучшее качество нормализации данных. При попарной нормализации более эффективным оказался метод квантильной нормализации. Кроме того, метод квантильной нормализации ока-

зался самым быстрым из используемых методов. Однако, несмотря на достигнутые успехи в данной предметной области существует ряд нерешенных или частично решенных проблем.

К нерешенным частям общей проблемы относится отсутствие эффективной методики выбора комплекса методов предобработки информации и сокращения размера признакового пространства в системах кластеризации высокоразмерных данных сложной биологической природы.

Целью статьи является разработка методики выбора оптимального метода нормализации на этапе предварительной обработки высокоразмерных данных сложной биологической природы с последующим сокращением размерности признакового пространства и кластеризацией исследуемых объектов, основным критерием оценки качества обработки информации в которой является энтропия Шеннона.

### 3 МАТЕРИАЛЫ И МЕТОДЫ

В соответствии с предложенной методикой проведения эксперимента первым этапом является фильтрация данных, целью которой является минимизация шумовой компоненты исследуемых данных. Фильтрация данных осуществлялась при помощи вейвлетов, представляющих собой локализованные в пространстве функции, способные отслеживать и нужным образом обрабатывать локальные особенности исследуемых объектов. Эффективность использования вейвлетов для предварительной обработки высокоразмерных данных обусловлена характером распределения признакового пространства исследуемых объектов. Например, в случае анализа микромассивов ДНК признаками являются уровни экспрессии соответствующих генов, при этом гены на микромассиве распределяются так, что различающиеся гены с различным уровнем экспрессии будут чередоваться с частотой, существенно ниже частоты фоновой компоненты освещенности и частоты фоновых помех. Уровень экспрессии генов можно оценить по интенсивности освещенности в той или иной точке. В процессе вейвлет-декомпозиции из исходного сигнала выделяется его низкочастотная компонента, при этом сохраняются локальные особенности исходных данных. Основная идея вейвлет-декомпозиции сигнала заключается в следующем: исходное пространство интенсивностей освещенностей, соответствующее различным уровням экспрессии генов, разлагается на систему подпространств так, что каждое последующее подпространство является вложенным в предыдущее (4):

$$I \subset I_1 \subset I_2 \subset I_3 \subset \dots \subset I_n. \quad (4)$$

Полученные подпространства при этом должны удовлетворять свойству (5):

$$\bigcap_{i=0}^n I_i = \{0\}. \quad (5)$$

Каждое из полученных подпространств разлагается, в свою очередь, с использованием вейвлет-функции и масштабирующей функции на пространства аппроксимирующих и детализирующих коэффициентов (6):

$$I_i = \{A_i, H_i, V_i, D_i\}. \quad (6)$$

В случае, если объект характеризуется одним вектором признаков, результатом вейвлет-декомпозиции будет вектор аппроксимирующих и один вектор детализирующих коэффициентов. Аппроксимирующие коэффициенты несут информацию о низкочастотной составляющей сигнала, детализирующие коэффициенты содержат информацию о высокочастотной составляющей сигнала. В большинстве случаев полезный сигнал содержится в низкочастотной составляющей, а детализирующие коэффициенты несут информацию о локальных особенностях сигнала. Шумовая составляющая представляет собой высокочастотную компоненту, которая так же содержится в детализирующих коэффициентах. Вследствие вышесказанного можно сделать вывод, что обработав нужным образом детализирующие коэффициенты, и восстановив сигнал из аппроксимирующих коэффициентов и обработанных детализирующих коэффициентов, получаем матрицу признаков исследуемых объектов при минимальном уровне шумовой компоненты, что способствует повышению точности решения поставленной задачи.

В [14] представлены результаты исследований по определению типа вейвлета, уровня вейвлет-декомпозиции и значений соответствующих коэффициентов в системе фильтрации хроматограмм. В качестве основного критерия оценки качества фильтрации использовался критерий энтропии. В данном случае информацией являлась шумовая компонента сигнала, вследствие чего оптимальный набор параметров фильтра определялся из условия минимума энтропии. С учетом результатов, полученных в [14], при построении модели кластеризации объектов на этапе фильтрации использовался вейвлет симплет-4 при четвертом уровне вейвлет декомпозиции с мягкой обработкой детализирующих коэффициентов и значением порогового коэффициента, равного 4.

Целью этапа нормализации является приведение эмпирических данных микромассивов и их характеристик к одноковому диапазону и распределению, чаще всего к нормальному. При проведении исследований использовались следующие методы нормализации данных:

– минимаксная нормализация (7):

$$XN(i, j) = \frac{X(i, j) - X_{\min}(i)}{X_{\max}(i) - X_{\min}(i)}, \quad (7)$$

– десятичное масштабирование (8):

$$XN(i, j) = \frac{X(i, j)}{10^k}, \quad (8)$$

где параметр  $k$  подбирается так, чтобы максимальное значение элемента в массиве было меньше 1;

– нормализация при помощи стандартного отклонения (9):

$$XN(i, j) = \frac{X(i, j) - \bar{X}_{us}(i)}{\sigma(i)}, \quad (9)$$

– метод линейной нормализации, предложенный разработчиками Affymetrix применительно к анализу микромассивов ДНК (10):

$$XN(i, j) = \frac{\bar{X}_{op}}{\bar{X}_{us}(i)} X(i, j); \quad (10)$$

– нелинейная нормализация с использованием логистической передаточной функции (11):

$$XN(i, j) = \frac{1}{1 + \exp\left(-\frac{X(i, j) - \bar{X}_{us}(i)}{\sigma(i)}\right)}; \quad (11)$$

– метод контрастов, в котором предполагается что между векторами  $M$  и  $A$  существует линейная регрессионная зависимость. Векторы  $M$  и  $A$  для каждого элемента массива данных вычисляются по формулам (12) и (13):

$$M(i, j) = \log_2\left(\frac{X(i, j)}{X(op, j)}\right); \quad (12)$$

$$A(i, j) = \log_2(X(i, j) \cdot X(op, j)). \quad (13)$$

Нормирующая поправка вычисляется по формуле (14):

$$\delta M(i, j) = M(i, j) - \hat{M}(i, j). \quad (14)$$

Нормализованное значение вычисляется по формуле (15):

$$XN(i, j) = 2^{\left(A(i, j) + \frac{\delta M(i, j)}{2}\right)}; \quad (15)$$

– квантильная нормализация основана на предположении, что признаки, характеризующие  $n$ -объектов, имеют одинаковые распределения данных. Тогда график квантилей в  $n$ -мерном пространстве будет представлять собой линию, лежащую вдоль диагонали, координаты которой представлены вектором:  $\left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)$ . Нормализованные значения признаков исследуемых векторов представляют собой проекции точки  $n$ -мерного графика квантилей на данную диагональ.

Оценка качества предобработки признаков исследуемых объектов оценивалось посредством анализа характера изменения критерия «Энтропия Шеннона» (3). Если под текущим событием понимать значение признака  $j$  в векторе  $i$   $X(i, j)$ , то вероятность осуществления данного события  $p(i, j) = X(i, j)^2$ . Тогда учитывая, что как мера количества информации в исследуемом векторе энтропия является величиной безразмерной, формулу (3) для  $i$ -й строки можно представить как (16):

$$H_i = -\sum_{j=1}^c X(i, j)^2 \cdot \ln(X(i, j)^2). \quad (16)$$

В соответствии с концепцией связи энтропии и информации о состоянии системы максимальное количество информации соответствует минимуму энтропии, при этом полная информация соответствует нулевой энтропии. Тогда естественно предположить, что наиболее качественная предобработка признаков объекта может быть выбрана из условия минимума энтропии Шеннона, что соответствует максимальной информации об исследуемых объектах, при этом в процессе проведения компонентного анализа изменение энтропии должно быть минимальным, что свидетельствует о минимальной потере информации в процессе выделения главных компонент.

#### 4 ЕКСПЕРИМЕНТИ

Моделирование процесса кластеризации исследуемых объектов производилось с использованием программной платформы KNIME-2.10.1, которая является свободно доступным программным продуктом и предназначена для обработки данных различной природы и извлечения с них информации, а также для моделирования всевозможных процессов и систем. Структурная блок-схема используемой модели кластеризации представлена на рисунке 1. В качестве экспериментальной базы для проведения исследований использовалась база данных больных лейкемией [15], представляющая собой массив размером  $72 \times 7131$ . Каждая строка содержит информацию об уровне экспрессии генов больных клеток отдельного человека. Объекты делятся на два класса. Один класс представляет собой пробы клеток, взятых у больных острой миелоидной лейкемией, а другой представляет клетки больных острой лимфобластной лейкемией.

Фильтрация исходных данных, их нормализация каждым из вышеописанных методов и выделение главных компонент производилось с использованием СКМ MATLAB. Этот выбор определялся тем обстоятельством, что платформа KNIME имеет ограничения на размерность признакового пространства исследуемых объек-

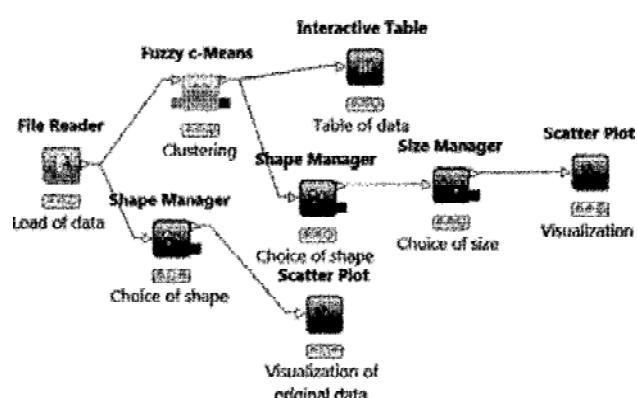


Рисунок 1 – Структурная блок-схема модели кластеризации объектов

тов. Далее матрица главных компонент загружалась в программную среду KNIME и подвергалась обработке с целью кластеризации исследуемых объектов. В качестве алгоритма кластеризации был выбран алгоритм нечеткой кластеризации C-means. Этот выбор определяется характером распределения объектов в пространстве. Вариации признаков исследуемых объектов в силу их природных особенностей не позволяют однозначно отнести объект к тому или иному кластеру. Имеет смысл только утверждение о степени принадлежности объекта тому или иному кластеру, вследствие чего для анализа такого типа данных наиболее целесообразны нечеткие методы кластеризации, одним из которых является метод C-means.

## 5 РЕЗУЛЬТАТЫ

Диаграмма распределения процентного содержания корректно распределенных объектов в зависимости от

используемого метода нормализации представлена на рис. 2. На рис. 3 представлены графики распределения энтропии исследуемых объектов для оригинального и нормализованного различными методами сигналов. На рис. 4 показаны аналогичные графики для главных компонент исследуемых объектов.

Для оценки степени изменения энтропии в процессе вычисления главных компонент исследуемого объекта рассчитывалось относительное изменение энтропии в процессе трансформации исследуемых векторов признаков объектов (17):

$$\frac{dE_i}{E_{0i}} = \frac{|E_i - E_{0i}|}{|E_{0i}|}. \quad (17)$$

Графики относительного изменения энтропий нормализованного и трансформированного векторов при использовании различных методов нормализации представлены на рис. 5.

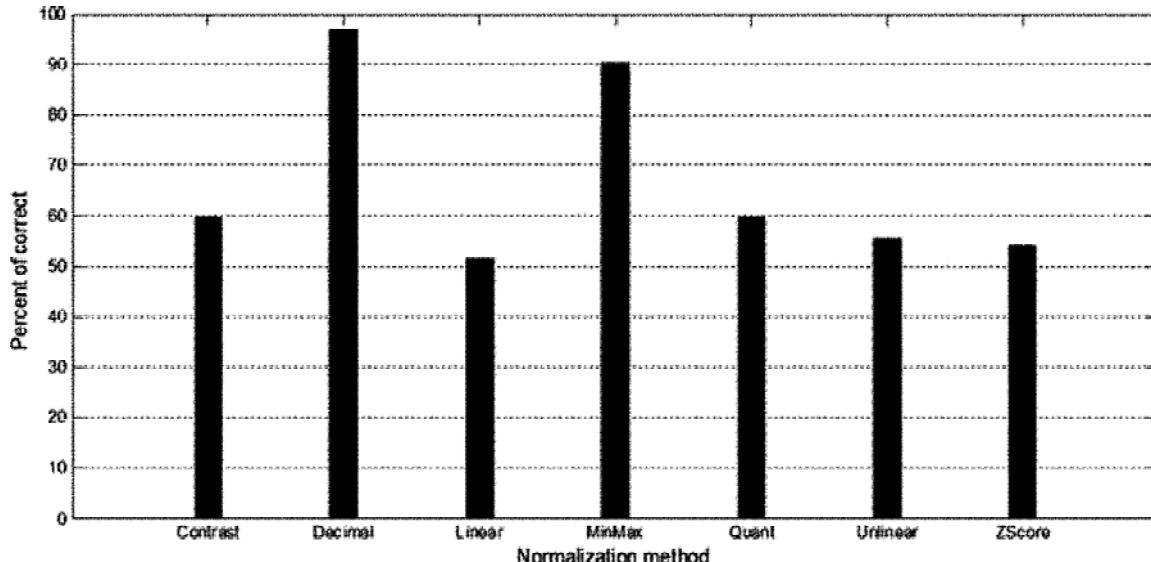


Рисунок 2 – Диаграмма распределения процентов корректно распределенных по кластерам объектов в зависимости от метода нормализации

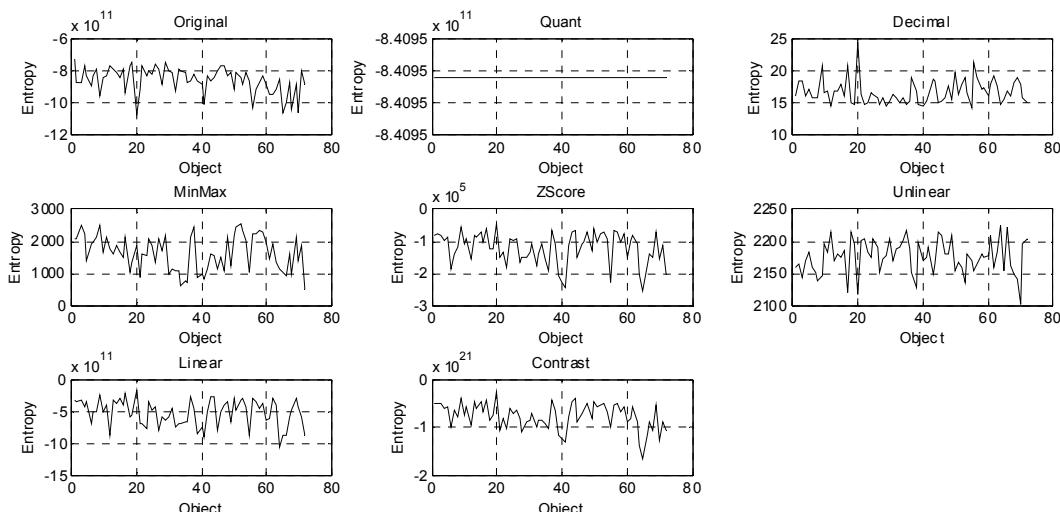


Рисунок 3 – Графики энтропий оригинального и нормализованного различными методами сигналов: Entropy – значение энтропии Шеннона, Object – исследуемый объект

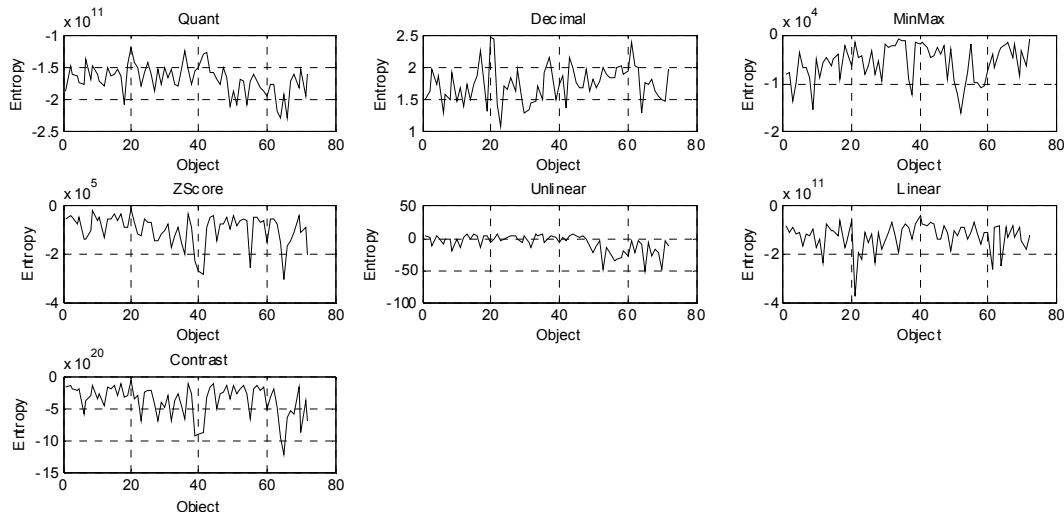
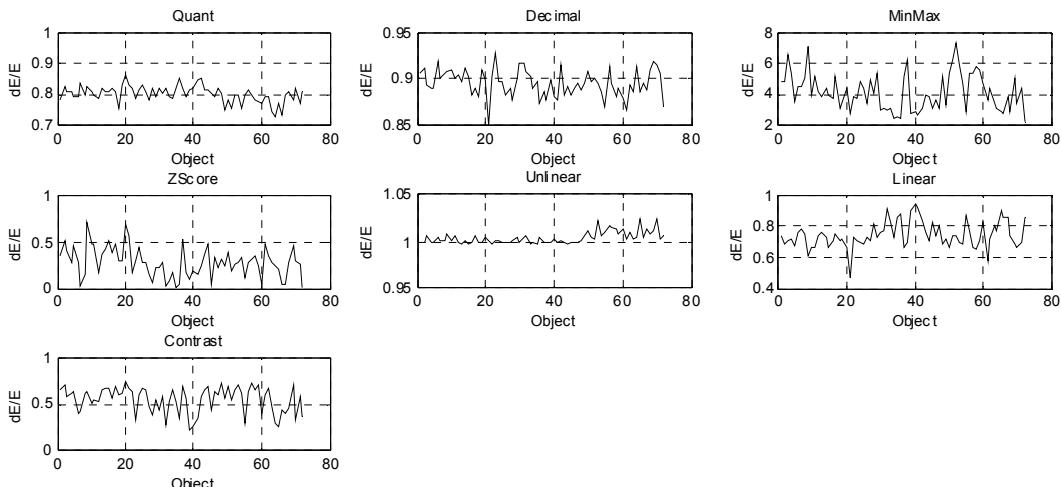


Рисунок 4 – Графики ентропій главних компонент исследованных объектов

Рисунок 5 – Графики относительного изменения энтропии при использовании различных методов нормализации:  $dE/E$  – относительное изменение энтропии; Object – исследуемый объект

## 6 ОБСУЖДЕНІ

Аналіз рис. 3 позволяє сделать вывод, что наименьшую ентропию имеют вектора, нормализованные методами decimal-scaling, min-max и unlinear. При этом метод десятичного масштабирования оказывается по критерию ентропии более предпочтительным. Однако абсолютное значение критерия ентропии не является объективным, поскольку диапазон варьирования значений признаков исследуемых объектов изменяется в зависимости от используемого метода нормализации, что оказывает непосредственное влияние на абсолютное значение критерия ентропии.

О величине потери информации вследствие трансформации вектора (вычислении главных компонент) можно судить по относительному критерию изменения ентропии. Из анализа графиков, изображенных на рис. 3–5 следует, что при использовании методов квантильной нормализации, десятичного масштабирования, Z-масштабирования, линейной нормализации и нормали-

зации методом контрастов происходит приблизительно одинаковая потеря информации. Однако по сумме двух критериев метод десятичного масштабирования является наиболее предпочтительным для нормализации исследуемого типа сигналов. При использовании минимаксного и нелинейного методов нормализации наблюдается более существенное возрастание ентропии, а значит и большая потеря информации об исследуемом объекте, что подтверждается результатами кластеризации, представленными на рис. 2. На рис. 6 показаны диаграммы распределения объектов по кластерам при использовании decimal-scaling и min-max методов нормализации. Анализ распределения объектов по кластерам, представленного на рис. 6, подтверждает полученные результаты о более высокой эффективности нормализации методом decimal-scaling по сравнению с другими используемыми методами при кластеризации объектов, отличительной особенностью которых является высокая размерность признакового пространства.

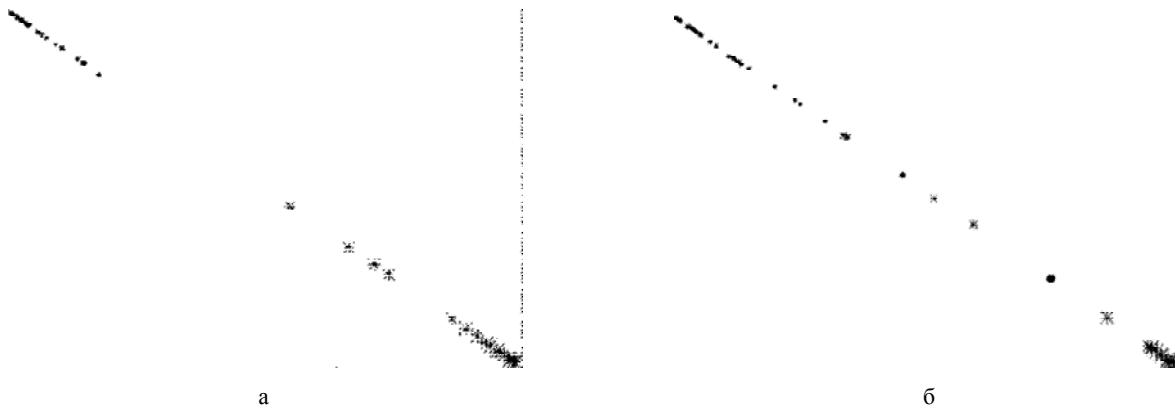


Рисунок 6 – Результаты разбиения объектов на кластеры при использовании: а – decimal-scaling нормализации; б – min-max нормализации

## ВЫВОДЫ

В работе получила дальнейшее развитие технология использования критерия «Энтропия Шеннона» в качестве информационного критерия для оценки количества полезной информации в исследуемых данных, отличительной особенностью которых является высокая размерность пространства признаков. Предложен комплексный подход по оценке качества обработки информации на основе критерия энтропии и относительного изменения энтропии в процессе трансформации данных с использованием компонентного анализа. Как показали исследования, метод главных компонент имеет высокую чувствительность к методам нормализации данных на этапе их предобработки, вследствие чего высокую актуальность приобретают исследования по оптимизации использования методов предобработки высокоразмерных данных с целью повышения качества их кластеризации.

Практический интерес представляет применение предложенной технологии для обработки микромассивов ДНК, объектами которых являются клетки больного органа, а признаками – уровень экспрессии генов, определяющих состояние соответствующей клетки. Особенностью исследуемых данных является высокий уровень шума и высокая размерность пространства признаков, что ограничивает использование традиционных методов обработки информации. Результаты исследований, показали, что для данных биологической природы, представляющих собой уровни экспрессии генов больных клеток, наименьшая потеря информации наблюдается при использовании decimal-scaling метода нормализации данных. Количество правильно распределенных по кластерам объектов в этом случае составляет 97%. Исследования также показали, что в случае отсутствия априорной информации о принадлежности объекта тому или иному кластеру оптимальную комбинацию методов предобработки можно выбрать на основании критерия Энтропия Шеннона. Значения данного критерия для наиболее качественно предобработанного сигнала является минимальным, что соответствует максимальному количеству сохраненной информации. Кроме того, относительное изменение энтропии в процессе трансформации сигнала компонентным анализом в случае применения метода decimal-scaling не

превышает допустимого значения, что свидетельствует о незначительной потере информации в процессе трансформации данных.

Перспективными направлениями дальнейших исследований является нахождение аналитического выражения комплексного критерия оценки качества предобработки информации, одну из составляющих которого будет составлять энтропия Шеннона.

## БЛАГОДАРНОСТИ

Работа выполнена в рамках госбюджетной научно-исследовательской темы Херсонского национального технического университета «Исследование искусственных иммунных систем и методов мультифрактального анализа в задачах идентификации и классификации биологических сигналов» (номер государственной регистрации 0109U009003).

## СПИСОК ЛИТЕРАТУРЫ

1. Shannon C. E. A mathematical theory of communication / C. E. Shannon // Bell System Technical Journal. – 1948. – Vol. 27. – P. 379–423, 623–656.
2. Shore J. E. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy / J. E. Shore, R. W. Johnson // IEEE Transactions on Information theory. – 1980. – Vol. IT-26. – P. 26–37.
3. Data Analysis of Bio-Medical Data Mining using Enhanced Hierarchical Agglomerative Clustering / [Krishnaiah J. V., Chandra Sekar D. V., Ramchand K., Rao H.] // International Journal of Engineering and Innovative Technology. – 2012. – Vol. 2, Issue 3. – P. 43–49.
4. Liang J. Computational analysis of microarray gene expression profiles: clustering, classification, and beyond / J. Liang, S. Kachalo // Chemometrics and Intelligent Laboratory Systems. – 2002. – No. 62. – P. 199–216.
5. Rezankova H. Cluster analysis of economic data / H. Rezankova // Statistica. – 2014. – No. 94(1). – P. 73–86.
6. Li Y. Text document clustering based on frequent word meaning sequences / Y. Li, S. M. Chung, J. D. Holt // Data & Knowledge Engineering. – 2008. – No. 64(1). – P. 381–404.
7. Jain A. K. Data clustering: A review / A. K. Jain, M. N. Murty, P. J. Flynn // ACM Computing Surveys. – 1999. – Vol. 31, No. 3. – P. 264–323.
8. Ивахненко А. Г. Объективная кластеризация на основе теории самоорганизации моделей / А. Г. Ивахненко // Автоматика. – 1987. – № 5. – С. 6–15.

9. Ивахненко А. Г. Алгоритмы метода группового учета аргументов (МГУА) при непрерывных и бинарных признаках / А. Г. Ивахненко. – К. : Институт кибернетики АН Украины, 1992. – 49 с.
10. Ивахно С. С. Методы кластеризации в программе Microarraytool для анализа данных ДНК-микроарреев / С. С. Ивахно, А. И. Корнелюк, О. П. Минцер // Медична інформатика та інженерія. – 2008. – № 3. – С. 33–40.
11. Ивахно С. С. Огляд технологій та аналіз даних / С. С. Ивахно, О. И. Корнелюк // Український біохімічний журнал. – 2004. – № 2 (76). – С. 5–19.
12. Bioinformatics and Computational Biology. Solutions Using R and Bioconductor / [R. Gentleman, V. Carey, W. Huber et al.]. – New York : Springer, 2005. – 473 p.
13. A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Variance and Bias / [Bolstad B. M., Irizarry R. A., Astrand M., Speed T. P.] // Bioinformatics. – 2003. – Vol. 19. – P. 185–193.
14. Фільтрація хроматограмм с помошью вейвлет-аналіза с использованием критерия энтропии / [С. А. Бабичев, Н. И. Бабенко, А. А. Дидақ и др.] // Системные технологии. – 2010. – № 6 (71). – С. 17–22.
15. Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring / [T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, et al.] // Science. – 1999. – Vol. 286. – P. 531–537.

Статья поступила в редакцию 24.10.2014.  
После доработки 04.11.2014.

Бабічев С. А.

Канд. техн. наук, доцент, доцент кафедри інформатики та комп’ютерних наук, Херсонський національний технічний університет, Україна

## ОПТИМІЗАЦІЯ ПРОЦЕСУ ПЕРЕДОБРОБКИ ІНФОРМАЦІЇ У СИСТЕМАХ КЛАСТЕРИЗАЦІЇ ВИСОКОРОЗМІРНИХ ДАНИХ

Представлено методику вибору оптимального методу нормалізації при побудові кластерної структури об'єктів, відмінною особливістю яких є висока розмірність простору ознак. Як основний критерій оцінки якості передобробки даних використовувався критерій ентропія Шенонна і відносна зміна ентропії у процесі трансформації даних. Зниження розмірності простору ознак досліджуваних об'єктів здійснювалося за допомогою компонентного аналізу. Побудовано модель системи кластеризації з використанням алгоритму нечіткої кластеризації fuzzy C-means, за допомогою якої зроблено оцінку якості кластеризації при використанні різних методів передобробки даних. Показано, що для досліджуваних даних найкращим методом нормалізації є метод десяткового масштабування, при якому ентропія обробленого сигналу приймає найменше значення, при цьому в процесі трансформації даних компонентним аналізом відносна зміна ентропії не перевищує дозволених норм.

**Ключові слова:** кластеризація, розмірність простору ознак, нормалізація, ентропія.

Babichev S. A.

PhD., Associate Professor, Department of Informatics and Computer Science, Kherson National Technical University, Kherson, Ukraine

## OPTIMIZATION OF INFORMATION PREPROCESSING IN CLUSTERING SYSTEMS OF HIGH DIMENSION DATA

The methodic of choice of optimal normalization method for object cluster structure of creation, with high dimension of feature space, is shown. The Shannon entropy criterion and entropy relative change were used as main criterions of estimating the data preprocessing quality during the data transformation. Decreasing of feature space dimension of tested objects was realized by component analysis. Model of system clustering with the use of fuzzy C-means algorithm was constructed, which the help of whith the estimate of clustering quality was established by the use of different data preprocessing methods. It's shown that the best normalization method for tested data is decimal-scaling method, by which the entropy of processed signal gets minimal significance, and relative change of entropy doesn't exceed permissible norms during the process of data transformation by component analysis.

**Keywords:** clustering, the feature space dimension, normalization, entropy.

## REFERENCES

1. Shannon C. E. A mathematical theory of communication, *Bell System Technical Journal*, 1948, Vol. 27, pp. 379–423, 623–656.
2. Shore J. E., Johnson R. W. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, *IEEE Transactions on Information theory*, 1980, Vol. IT-26, pp. 26–37.
3. Krishnaiah J. V., Chandra Sekar D. V., Ramchand K., Rao H. Data Analysis of Bio-Medical Data Mining using Enhanced Hierarchical Agglomerative Clustering, *International Journal of Engineering and Innovative Technology*, 2012, Vol. 2, Issue 3, pp. 43–49.
4. Liang J., Kachalo S. Computational analysis of microarray gene expression profiles: clustering, classification, and beyond, *Chemometrics and Intelligent Laboratory Systems*, 2002, No. 62, pp. 199–216.
5. Rezankova H. Cluster analysis of economic data, *Statistica*, 2014, No. 94(1), pp. 73–86.
6. Li Y. Text document clustering based on frequent word meaning sequences, *Data & Knowledge Engineering*, 2008, No. 64(1), pp. 381–404.
7. Jain A. K., Murty M. N., Flynn P. J. Data clustering: A review, *ACM Computing Surveys*, 1999, Vol.31, No. 3, pp. 264–323.
8. Ivahnenko A. G. Objektivnaja klasterizacija na osnove teorii samoorganizacii modelej, *Avtomatika*, 1987, No.5, pp. 6–15.
9. Ivahnenko A. G. Algoritmy metoda gruppovogo ucheta argumentov (MGUA) pri nepreryvnih i binarnyh priznakah. Kiev, Institut kibernetiki AN Ukrayny, 1992, 49 p.
10. Ivahno S. S., Korneljuk A. I., Mincer O. P. Metody klasterizacii v programme Microarraytool dlja analiza dannyh DNK-mikroarreev, *Medichna informatika ta inzhenerija*, 2008, No. 3, pp. 33–40.
11. Ivahno S. S., Korneljuk O. I. Ogljad tehnologij ta analiz danih, *Ukrainskij biohimichnij zhurnal*, 2004, No. 2(76), pp. 5–19.
12. Gentleman R. Carey V., Huber W., Irizarry R., Dudoit S. Bioinformatics and Computational Biology. Solutions Using R and Bioconductor. New York, Springer, 2005, 473 p.
13. Bolstad B. M., Irizarry R. A., Astrand M., Speed T. P. A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Variance and Bias, *Bioinformatics*, 2003, Vol. 19, pp. 185–193.
14. Babichev S. A., Babenko N. I., Didyk A. A., Litvinenko V. I., Fefelov A. A., Shkurdoda S. V. Fil'tracija hromatogramm s pomoshsh'ju vejlvet-analiza s ispol'zovaniem kriterija jentropii, *Sistemnye tehnologii*, 2010, No. 6(71), pp. 17–22.
15. Golub T. R., Slonim D. K., Tamayo P., Huard C., at al. Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring, *Science*, 1999, Vol. 286, pp. 531–537.

УДК 004.738.5

Берко А. Ю.<sup>1</sup>, Висоцька В. А.<sup>2</sup>, Чирун Л. В.<sup>3</sup>

<sup>1</sup>Д-р. техн. наук, професор, завідувач кафедри «Загальна екологія та екоінформаційні системи» Національного університету «Львівська політехніка», Україна

<sup>2</sup>Асистент кафедри «Інформаційні системи та мережі», Національний університет «Львівська політехніка», Україна

<sup>3</sup>Канд. техн. наук, доцент кафедри «Програмне забезпечення», Національний університет «Львівська політехніка», Україна

## ПРОЕКТУВАННЯ ТА ВПРОВАДЖЕННЯ СИСТЕМ ЕЛЕКТРОННОЇ КОНТЕНТ-КОМЕРЦІЇ

Запропоновано метод опрацювання комерційного контенту як етап життєвого циклу контенту в системах електронної комерції. Метод опрацювання комерційного контенту описує процеси формування інформаційних ресурсів в системах електронної контент-комерції та спрощує технологію управління комерційним контентом. У цій роботі проаналізовано основні проблеми електронної контент-комерції та функціональних сервісів опрацювання комерційного контенту. Запропонований метод дає можливість створити засоби опрацювання інформаційних ресурсів в системах електронної контент-комерції та реалізувати підсистеми формування, управління та супроводу комерційного контенту.

**Ключові слова:** інформаційний ресурс, комерційний контент, контент-аналіз, контент-моніторинг, контентний пошук, система електронної контент-комерції.

### НОМЕНКЛАТУРА

ІТ – інформаційні технології;

CMS – Content Management System (система керування вмістом);

СЕКК – система електронної контент-комерції;

$N$  – кількість джерел контенту;

$x_i$  –  $i$ -те джерело інформації;

$X(x_i)$  – вхідний набір контенту (множина контенту  $x$   $i$ -того джерела);

$X_i$  – набір контенту як результат відбору з  $i$ -го джерела;

$U_i$  – множини умов формування вхідного набору контенту;

$\langle X_i, U_i \rangle$  – множини значень контенту у заданому форматі і множини умов;

$Verification(\langle X_i, U_i \rangle)$  – оператор верифікації контенту;

$Qualification(\langle X_i, U_i \rangle)$  – оператор кваліфікації контенту;

$Conversion(\langle X_i, U_i \rangle)$  – оператор перетворення контенту;

$Downloading(\langle X_i, U_i \rangle)$  – оператор завантаження контенту;

$\langle X, U \rangle$  – загальний інформаційний ресурс;

$S$  – система електронної контент-комерції;

$X = \{x_1, x_2, \dots, x_{n_X}\}$  – множина контенту з різних джерел;

$Q = \{q_1, q_2, \dots, q_{n_Q}\}$  – множина інформаційних запитів користувачів;

$C = \{c_1, c_2, \dots, c_{n_C}\}$  – множина комерційного контенту;

$V = \{v_1, v_2, \dots, v_{n_V}\}$  – множина умов супроводу контенту та зовнішніх впливів середовища на систему;

$H = \{h_1, h_2, \dots, h_{n_H}\}$  – множина умов формування та управління контентом;

$Z = \{z_1, z_2, \dots, z_{n_Z}\}$  – множина компонентів інформаційного ресурсу;

$T = \{t_1, t_2, \dots, t_{n_T}\}$  – час транзакцій управління контентом;

$Y = \{y_1, y_2, \dots, y_{n_Y}\}$  – колекція вихідних характеристик роботи системи;

$y_j = \{a_1, a_2, \dots, a_g\}$  – колекція статистичних даних функціонування системи;

$a_1$  – кількість відвідувань за період часу  $\Delta t$ ;

$a_2$  – середній час відвідування інформаційного ресурсу (хв:с) за період часу  $\Delta t$ ;

$a_3$  – показник відмовлень (%) за період часу  $\Delta t$ ;

$a_4$  – досягнута мета пошуку;

$a_5$  – динаміка контенту (%);

$a_6$  – кількість всього переглядів сторінок;

$a_7$  – кількість переглядів сторінок за одне відвідування;

$a_8$  – нові відвідування (%);

$a_9$  – абсолютно унікальні відвідувачі;

$a_{10}$  – джерело трафіка у % тощо;

$Function$  – оператор формування вихідних даних системи;

$Formation$  – оператор формування контенту;

$Management$  – оператор управління контентом;

$Support$  – оператор супроводу контенту;

$u_f = \{u_1(x_i), \dots, u_{n_U}(x_i)\}$  – множина умов формування контенту.

### ВСТУП

Активний розвиток Інтернет сприяє зростанню потреб в отриманні оперативних даних виробничого/стратегічного характеру і реалізації нових форм інформаційного обслуговування через сучасні ІТ електронної комерції [1, 6, 23–25]. Документована інформація, підготовлена відповідно до потреб користувачів, є комерційним контентом [1, 6–7, 12–16, 23–25]. Сьогодні електронна комерція є об'єктивною реальністю та перспективним бізнес-процесом [1]. Інтернет є бізнес-середови-

щем, а комерційний контент є товаром з найбільшим попитом і продажами в ньому та основним об'єктом процесів електронної контент-комерції. Комерційний контент можна відразу замовити, оформити, оплатити та отримати on-line як товар. Через Інтернет продають весь спектр комерційного контенту – наукові та публіцистичні статті, музику, книги, фільми, фото, програмне забезпечення тощо. Відомими корпораціями, які реалізують електронну контент-комерцію, є Google через Play Market, Apple – Apple Store, Amazon – Amazon.com. Більшість рішень та досліджень проведено на рівні конкретних проектів. СЕКК побудовані за закритим принципом як разові проекти. Сучасні СЕКК орієнтовані на реалізацію комерційного контенту, який створений за межами системи. Проектування, створення, впровадження та супровід СЕКК неможливе без використання сучасних методів та інформаційних технологій формування, управління та супроводу комерційного контенту.

СЕКК передбачає використання інформаційних технологій для взаємодії торгівельних компаній з роздрібними покупцями, забезпечення повного циклу продажу контенту [1–2]. У СЕКК склад учасників набагато ширший: крім користувачів (продавців, покупців) містить низку фінансових установ (банк-емітент, банк продавця/покупця, банк-екваєр), комп’ютерних центрів тощо [1]. Користувачами виступають переважно фізичні особи, а також установи, соціальні заклади, інші види споживачів (юридичні особи). Продавцями в СЕКК є різні організаційні форми контент-торгівлі. Комуникаційну мережу формуєть провайдери, сервери, процесингові центри тощо. Систему доставки становить Інтернет. Усі складові елементи взаємодіють у системі взаємозв’язків. У цьому гарантія стійкості й надійності СЕКК. Важливими елементами СЕКК є організаційні форми електронної контент-торгівлі, що мають єдину спрямованість – забезпечення процесу роздрібної купівлі-продажу, але відрізняються складом, структурою, призначеннем в СЕКК [1–2].

Існує певне протиріччя між активним розвитком і поширенням ІТ та СЕКК з одного боку, та порівняно не-значним обсягом наукових досліджень з цієї тематики та їх локальністю з іншого [1]. Це протиріччя породжує проблему стимулювання інноваційного розвитку сектору електронної контент-комерції через створення і запровадження відповідних новітніх прогресивних ІТ, що негативно впливає на темпи зростання цієї частини ринку. В межах загальної проблеми актуальною є задача розроблення науково обґрунтованих методів опрацювання інформаційних ресурсів електронної контент-комерції та побудови на їх основі технологічних програмних засобів для створення, поширення і сталого розвитку СЕКК. В роботі проведено дослідження з метою визначення закономірностей, особливостей та залежностей у процесах опрацювання інформаційних ресурсів в СЕКК.

Метою статті є розроблення методів та технологічних засобів опрацювання інформаційних ресурсів в системах електронної контент-комерції для збільшення обсягів реалізації комерційного контенту постійним користувачам.

## 1 ПОСТАНОВКА ЗАДАЧІ

Основною задачею проекту створення СЕКК є розроблення архітектури інформаційного ресурсу, яку формують за зворотною реакцією користувачів відповідно до типу поширення комерційної діяльності. Контент (від англ. content – зміст, вміст, суть) в галузі інформаційних технологій є формалізованими відомостями і знаннями, розміщеними у середовищі системи і, на відміну від даних, без детальної специфікації їх властивостей, способів формалізації і впорядкування [1]. Інформаційний ресурс СЕКК – це множина даних з набором властивостей, які є об'єктом дій технології перетворення їх в комерційний контент [1]. Результат застосування однієї технології може бути інформаційним ресурсом іншої. Виникає проблема створення єдиного концептуального опису всього інформаційного ресурсу для стабільного підтримання зовнішніх/внутрішніх позначень контенту відповідно до їх завдань, вимог та змін [1, 3–5]. Необхідно класифікувати інформаційні ресурси для дослідження їх природних, технологічних та споживчих якостей з метою виявлення характерних та специфічних властивостей, а також закономірностей та особливостей їх формування та застосування. За основу взято основні властивості контенту в СЕКК як синтаксис, структура та семантика та обрано основні фактори класифікації як способи подання контенту, структурування та призначення ресурсу, доступу до ресурсу.

Нехай існує деяка попередньо визначена множина  $N$  первинних джерел контенту з фіксованим або змінним складом. Кожне джерело інформації  $x_i$ , де  $i = \overline{1, N}$ , формує деяку множину значень, що містять відомості/знання/факти з предметної області СЕКК. Результатом звернення певних технологічних засобів СЕКК до джерела  $x_i$  є генерування множини значень  $X(x_i)$ , яка сприймається і подається у визначеній формі. В процесі відбору і фіксації генерованих значень, згідно технологічних особливостей системи, згенеровану кожним джерелом інформації множину значень перетворюють на вхідний набір контенту, вказаного формату –  $X_i$ , де  $i = \overline{1, N}$ . Кожен набір контенту подається у вигляді структурованих, слабкоструктурованих даних або даних без вказаного описання структури.

Структурування контенту передбачає формування для кожного набору опису його складу, способів поєднання елементів та їх впорядкування – множини умов  $U_i$ , де  $i = \overline{1, N}$ . Кожен набір контенту є поєднанням множини значень у заданому форматі і множини умов –  $\langle X_i, U_i \rangle$ , у випадку формування вхідного набору контенту без опису структури  $U_i = \emptyset$ . Отриманий контент перед збереженням в базі даних проходить процедуру верифікації та валідації для підтвердження його формальної і змістової коректності/релевантності щодо вимог системи. При не відповідності зазначенім критеріям частина контенту вилучається з подальшого застосування. Відфільтрований контент форматують та зберігають,

після чого відповідні відомості і знання стають доступними для застосування користувачами СЕКК. Формально таку послідовність технологічних процесів подано як ланцюжок при  $i = \overline{1, N}$ :

$$\begin{aligned} x_i \rightarrow X(x_i) \rightarrow X_i \rightarrow \langle X_i, U_i \rangle \rightarrow \text{Verification}(\langle X_i, U_i \rangle) \rightarrow \\ \rightarrow \text{Qualification}(\langle X_i, U_i \rangle) \rightarrow \text{Conversion}(\langle X_i, U_i \rangle) \rightarrow \\ \rightarrow \text{Downloading}(\langle X_i, U_i \rangle) \rightarrow \langle X, U \rangle. \end{aligned} \quad (1)$$

## 2 ОГЛЯД ЛІТЕРАТУРИ

Основні класи користувачів/персонажів інформаційного ресурсу (клієнти, керівники робочих груп і адміністратори) визначають дизайн інформаційного ресурсу і процес ухвалення рішень (рис. 1) [1]. Порівняння типів поширення комерційної діяльності.

1. *Web-вітріна* має суттєві недоліки: дозволяє організувати тільки торгівлю на замовлення; не дозволяє налагодити торгівлю через Інтернет; не зменшує операційні витрати й на утримання штату; неповоротке рішення з погляду управління і недостатньо гнучке з погляду організації маркетингових акцій; імідж компанії гірший, ніж у компанії з СЕКК; малоефективна і нерентабельна організація торгівлі [1].

2. На створення CMS необхідна більша кількість разових витрат порівняно з Web-вітринами, але вони ефективніші, оскільки використання CMS рентабельніше щодо обороту [8–11]. Існує альтернатива самостійному створенню громіздкої системи – це орендування у спеціалізованих компаній: витрати рівномірно розподіляються в часі.

3. СЕКК обов'язково містить Web-вітрину з каталогом контенту (з можливістю пошуку) і необхідними інтерфейсними елементами для введення реєстраційної інформації, формування замовлення, здійснення платежів через Інтернет, оформлення доставки (e-mail або on-line), одержання інформації про компанію-продавця і on-line допомоги. Реєстрація покупця відбувається при оформленні замовлення або при входженні в систему [1, 6, 23–25].

Процес опрацювання інформаційних ресурсів складається з таких підсистем: формування контенту (збирання/створення, систематизація та поширення комерційного контенту), управління комерційним контентом (опрацювання, аналіз та подання комерційного контенту), супровід комерційного контенту (структурування, модерація та узагальнення комерційного контенту) [1–25]. На рис. 2 подана структура підсистеми опрацювання інформаційних ресурсів [3–4]. Визначивши пріоритетні тематики контенту та їх зв'язок із основними класами користувачів, створюють архітектуру інформаційного ресурсу (табл. 2), ієархію контенту, його способи подання і взаємодії з ним кожного класу користувачів (на приклад, теми та питання порядку денного або сесії, планування цих питань) [1]. Основні властивості інформаційних ресурсів в СЕКК [1–5].

1. *Неоднорідність*. Наявність складових різного походження, змісту і формату.

2. *Узгодженість*. Відсутність суперечливих або протилежних значень контенту.

3. *Доступність формату*. Доступність для всіх користувачів на основі стандартизованих методів, засобів та інтерфейсів.

4. *Відкритість*. Здатність до взаємодії, обміну значеннями та спільногом використання з зовнішніми ресурсами.

5. *Динамічність*. Швидка актуалізація, відповідно до умов системи чи середовища.

6. *Масштабованість*. Можливість зміни логічного/фізичного обсягу контенту (величин/понять та їх позначення).

7. *Контрольованість*. Фіксація зміни/використання контенту та його впливу на СЕКК.

Перетворення різнопідвидів за природою, змістом та походженням даних в узгоджений централізований інформаційний ресурс є однією з важливих проблем побудови та функціонування СЕКК [1]. При цьому важливими завданнями є забезпечення інформаційних потреб проблемно-орієнтованих елементів системи, підтримання

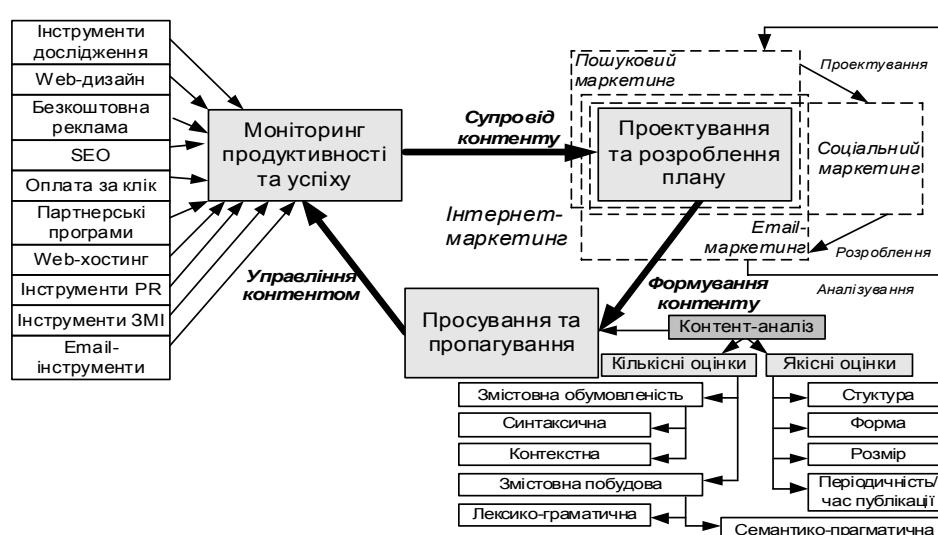


Рисунок 1 – Схема визначення дизайну інформаційного ресурсу та етапів ухвалення рішень

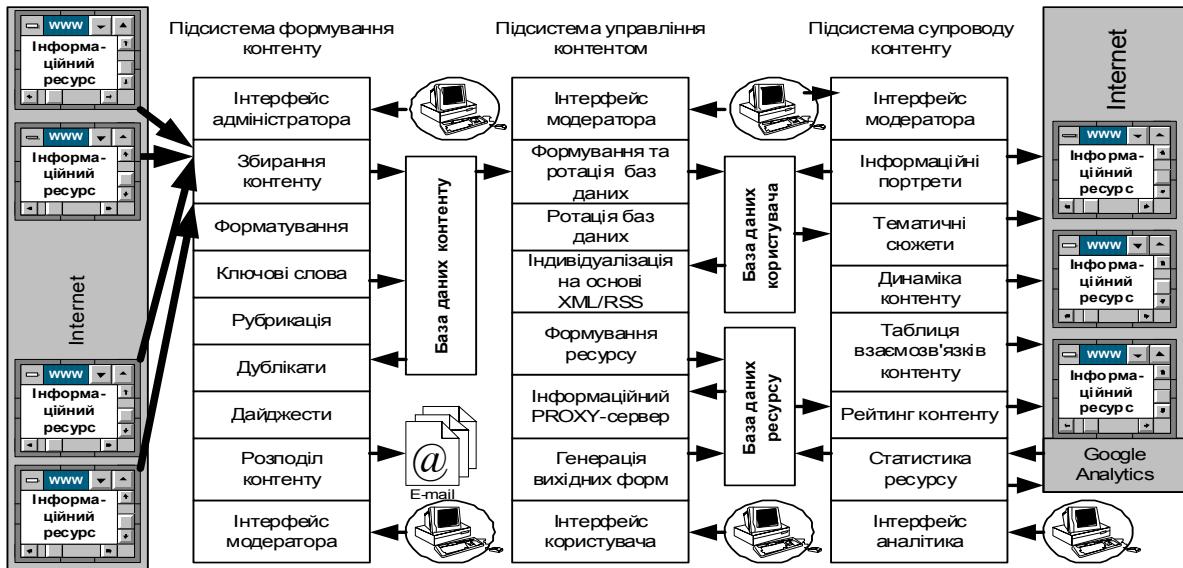


Рисунок 2 – Структура підсистеми опрацювання інформаційних ресурсів в СЕКК [3–4]

доступу до даних різних категорій користувачів, дотримання правил цілісності та несуперечності даних, мінімізація та контроль надлишку даних, здатність до розвитку та зміни внутрішньої організації інформаційного ресурсу, дотримання вимог якості та ефективності даних [1–5]. Важливим є забезпечення інваріантності середовища СЕКК до модифікації інформаційних ресурсів у таких змінах: способів подання, форматів, та внутрішньої орга-нізації контенту; середовища зберігання контенту, фізичних одиниць зберігання, технічних засобів; вимог користувачів, появи нових вимог та категорій користувачів; порядку розподілу контенту та способів доступу користувачів [1–6, 8–11, 23–15].

### 3 МАТЕРІАЛИ І МЕТОДИ

Модель електронної контент-комерції подано як

$$S = \langle X, Q, C, V, H, Function, T, Z, Y \rangle. \quad (2)$$

Процес роботи системи управління контентом  $S$  описують оператором

$$y_j(t_p + \Delta t) = Function(x_i, q_d, c_r, v_l, h_k, t_p, z_w). \quad (3)$$

Основними етапами процесу опрацювання інформаційних ресурсів в СЕКК є формування, управління та супровід контенту, які мають наступні зв'язки: контент  $\rightarrow$  формування контенту  $\rightarrow$  база даних  $\rightarrow$  управління контентом  $\rightarrow$  інформаційний ресурс або запит користувача  $\rightarrow$  управління контентом  $\rightarrow$  інформаційний ресурс  $\rightarrow$  супровід контенту  $\rightarrow$  база даних. Модель СЕКК подано як

$$S = \langle X, Q, Formation, H, C, V, Management, Support, Z, T, Y \rangle. \quad (4)$$

Оператор формування комерційного контенту *Formation* – відображення контенту в новий стан, який відрізняється від попереднього появою нової частини кон-

тенту, яка доповнює попередній стан. Оператор управління комерційним контентом *Management* – відображення комерційного контенту в новий стан, який відрізняється від попереднього відповідно значеннями визначальних параметрів (актуальність, повнота релевантність, автентичність, достовірність), що задоволяють наперед визначенім вимогам. Оператор супроводу комерційного контенту *Support* – відображення комерційного контенту в колекцію значень, які утворюють як результат аналізу, моніторингу, оцінювання взаємодії з користувачем, пошуковими системами та іншими інформаційними ресурсами, що є основою для прийняття рішень щодо формування та управління контентом. Етап формування контенту описано оператором *Formation* вигляду

$$c_r = Formation(u_f, x_i, t_p). \quad (5)$$

Контент подано наступним чином:

$$c_r = \left\{ \bigcup_f u_f \mid (x_i \in X) \wedge (\exists u_f \in U), U = U_{x_i} \vee U_{\overline{x_i}}, i = \overline{1, m}, f = \overline{1, n} \right\}. \quad (6)$$

Етап управління контентом описано оператором *Management* вигляду

$$z_w = Management(q_d, c_r, h_k, t_p), \quad (7)$$

де  $H = \{h_1(c_{i+1}, q_d), \dots, h_{n_H}(c_{i+n_H}, q_d)\}$ . Управління контентом подано як

$$z_w = \left\{ \bigcup_{k=1}^{n_H} h_k(c_{i+1}, q_d) \mid \begin{cases} (c_{i+k} \in C) \wedge (q_d \in Q) \wedge (h_k \in H_q), \\ H = H_{q_d} \vee H_{\overline{q_d}}, i = \overline{1, n_C}, d = \overline{1, n_Q}, k = \overline{1, n_H} \end{cases} \right\}. \quad (8)$$

Етап супроводу контенту описано оператором *Support* вигляду

$$y(t_p + \Delta t) = Support(v_l, h_k, c_r, z_w, t_p, \Delta t), \quad (9)$$

де  $v_l = (v_1(q_i, h_k, c_r, z_w, t_p), \dots, v_{n_V}(q_i, h_k, c_r, z_w, t_p))$ .

Вихідні дані реалізовано

$$y_j = \left\{ \bigcup_{l=1}^{n_V} \left| \begin{array}{l} (\exists q_d \in Q) \wedge (\exists z_w \in Z) \wedge (\forall v_l \in V) \wedge (\forall (c_r \wedge q_d) \in h_k), \\ V = V_{q_d} \vee V_{z_w}, d = \overline{l, n_Q}, l = \overline{1, n_V}, w = \overline{1, n_Z}, r = \overline{1, n_C}, k = \overline{1, n_H} \end{array} \right. \right\}. \quad (4)$$

Загальні принципи формування інформаційних ресурсів СЕКК (рис. 3) визначають порядок і способи відбору інформації із первинних джерел, її фіксації, фільтрування, перетворення до визначеного формату для формування контенту і розміщення в базі даних.

I. Підсистему формування контенту реалізують у вигляді контент-моніторингу зі збирання контенту з джерела за визначеними методами (рис. 4–5).

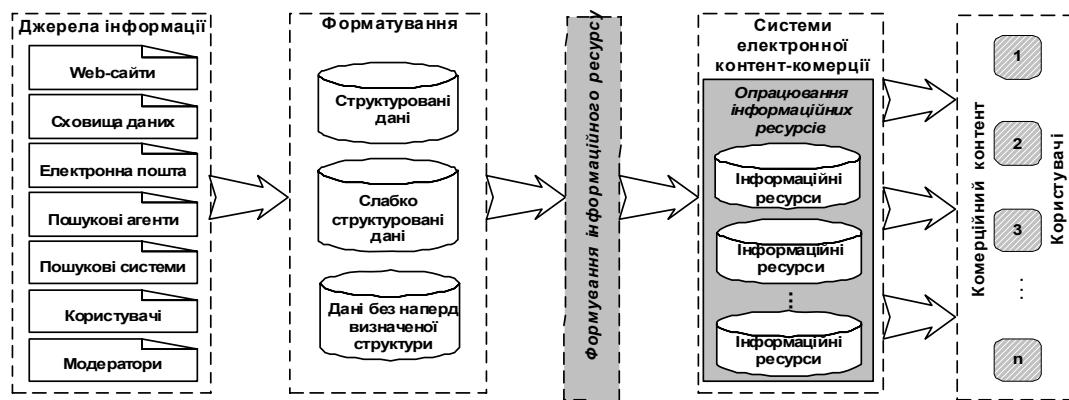


Рисунок 3 – Порядок формування і використання інформаційних ресурсів в СЕКК



Рисунок 4 – Основні вимоги до систем управління контентом для побудови СЕКК

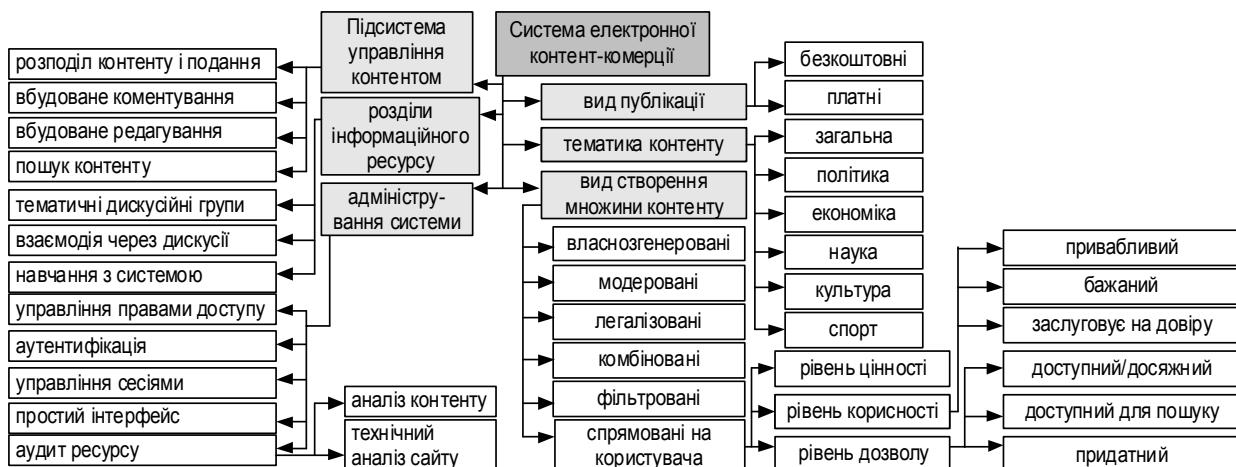


Рисунок 5 – Основні компоненти систем управління контентом для побудови СЕКК

Підсистема забезпечує постійне поповнення бази даних контентом, ефективний одночасний доступ до бази даних користувачами, зручні засоби пошуку необхідного контенту. Підсистема дає змогу економити на ресурсах адміністрування та ; підтримувати анонімність; не перевантажувати Інтернет-трафік; самостійного сканувати Інтернет.

**ІІ. Підсистема управління контентом** складається із таких відношень: *користувач → опрацювання контенту → база даних → аналіз контенту → база даних → подання контенту → користувач*. Аналіз контенту та інформаційного ресурсу (рис. 8), їхнє моделювання та подання є одним з найбільш інформативних методів кількісного вивчення динаміки окремих тематичних потоків комерційного контенту.

За зміною величин управління контентом визначать швидкість розвитку окремих тематичних напрямів або всього контентного простору. Стійкі статистичні зв'язки між окремим контентом свідчать про кореляції окремих тематик, про ефективність посилань на публікації попередників, більш ранні цитування, републікації тощо. Механізми, що базуються на узагальненіх методах кластерного аналізу, виявляють повідомлення в потоках контенту, що формують навколо себе нові тематичні напрямами. Кластерний аналіз, теорія фракталів і автомодельних процесів при їхньому коректному застосуванні кількісно оцінюють ступінь зв'язку в тематичних контентних потоках.

**ІІІ. Генерація сторінок за запитом.** Підсистема подання (рис. 9) генерує сторінку із змістом при запиті на нього на основі контенту з бази даних. У режимі діалогового доступу до бази даних забезпечується перегляд, пошук і відображення контенту, надається можливість звертання до оригіналів контенту в Інтернет. Контент в базі даних змінюється за допомогою модуля редактування. Сторінки наново створюються сервером при кожному запиті, що збільшує навантаження на ресурси системи. Навантаження знижується при використанні засобів кешування в сучасних Web-серверах (рис. 10). Задача повнотекстового пошуку у великих масивах контенту є неефективною. Проблему точності вирішує пошук в анотованому контенті. Замість пошуку у повному контенті доцільно вести пошук за анотаціями – пошукових образах контенту.

Квазіреферат для великого контенту є утворенням, що віддалено нагадує вихідний контент та часто не сприймається людиною. Пошуковий образ такого контенту із зваженими ключовими словами приводить до цілком адекватних результатів при повнотекстовому пошуку. Квазіреферат будеться з фрагментів контенту із більшими ваговими значеннями.

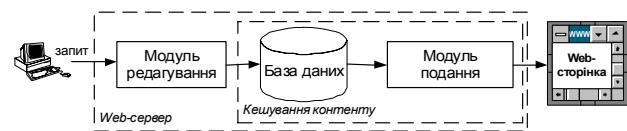


Рисунок 9 – Схема генерації сторінок за запитом

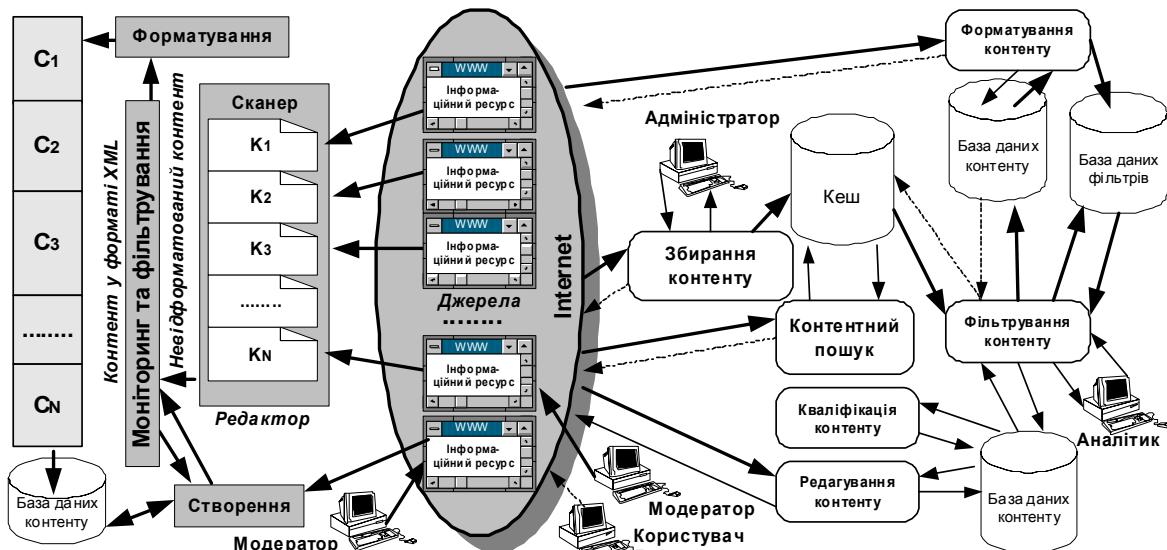


Рисунок 7 – Схема формування контенту в системах електронної контент-комерції

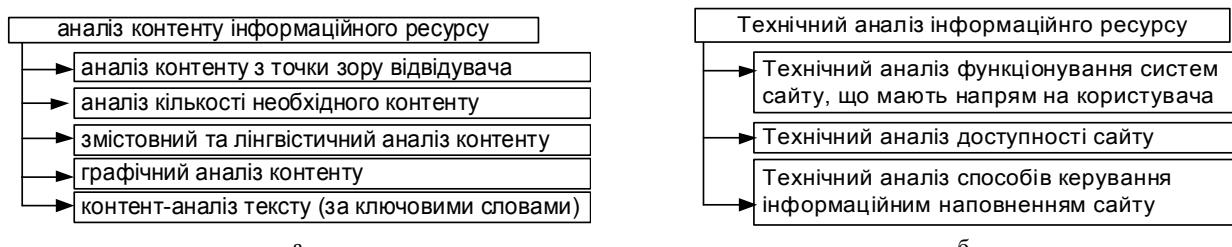


Рисунок 8 – Методи аналізу: а – контенту; б – інформаційного ресурсу

**2. Генерація сторінок при редагуванні** – це набір статичних сторінок (рис. 11), які безпосередньо редагує модератор прямо в базі даних системи електронної контент-комерції (рис. 10б). В такій підсистемі відсутня інте-

рактивність між відвідувачем та контентом інформаційного ресурсу в системі електронної контент-комерції.

**3. Змішаний тип** включає особливості перших двох типів підсистем управління комерційним контентом (рис. 12).

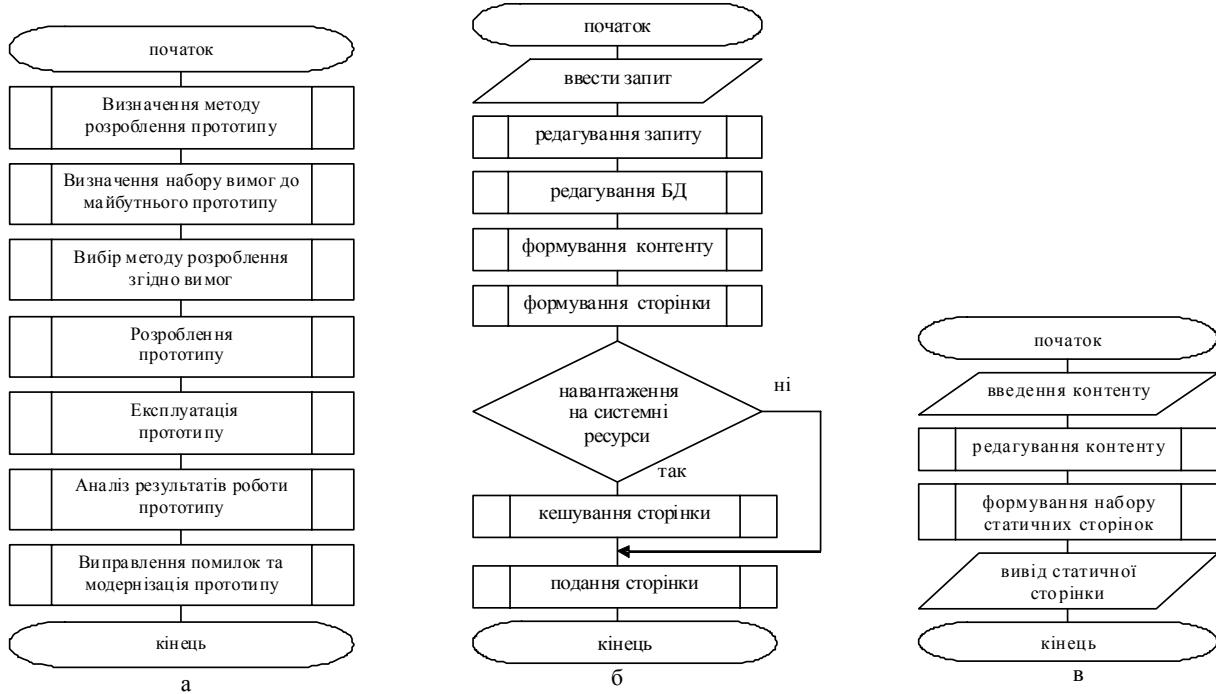


Рисунок 10 – Процес: а – створення прототипу; генерації сторінок б – за запитом; в – при редагуванні

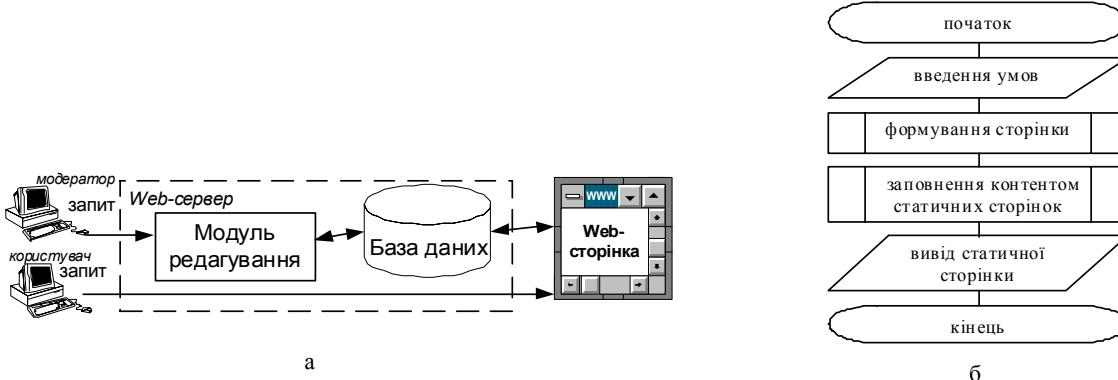


Рисунок 11 – Редагування в СЕКК : а – схема генерування сторінок при редагуванні в СЕКК; б – процес заповнення сторінок

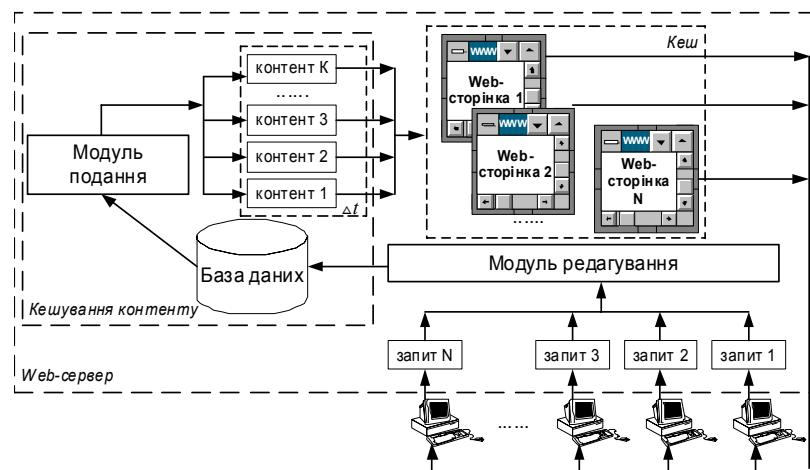


Рисунок 12 – Схема кешування згенерованих сторінок змішаного типу

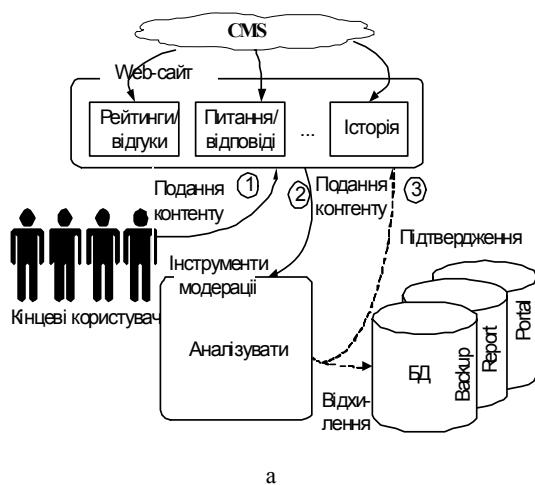
Переваги кешування сторінок з контентом: підсистема подання генерує сторінку один раз; сторінка в кеші існує час  $\Delta t$  поки є актуальним контент; готова сторінка швидше завантажується з кешу; кеш оновлюється періодично вручну/автоматично (по закінченню деякого терміну часу  $\Delta t$  або при модифікації розділів інформаційного ресурсу або контенту). Види реалізації процесу генерації сторінок змішаного типу: шляхом кешування; через формування інформаційних блоків (редагування інформаційного ресурсу та збереження блоків або формування сторінок через збирання сторінки із інформаційних блоків при запиті відповідного контенту користувачем). Підсистему управління контентом змішаного типу складніше реалізувати, але вона більш оперативно працює з користувачем через гнучкий інтерфейс, надаючи необхідний комерційний релевантний контент без надлишку інформації.

**III. Підсистема супроводу контенту** – це комплексні послідовні процеси: Користувач → структурування контенту → база даних → модерація контенту → база даних → узагальнення контенту → модератор. Актуальність розроблення загальної архітектури модуля супроводу контенту полягає у необхідності отримувати оперативні і об'єктивні оцінювання рівня конкуренції на сегменті фінансового ринку контенту; оцінювати рівень конкурентів та міри їх конкурентоспроможності на фінансовому ринку з розповсюдженням контенту. Причиною актуальності є швидкий темп росту попиту в поширенні контенту; постійна взаємодія суспільства і науковців з новітніми інформаційними технологіями; можливість збільшення прибутку існуючих/потенційних клієнтів, збільшення рейтингу; покращення бізнесу через розуміння клієнта. Проблемою супроводу контенту є відсутність загальних підходів для автоматичного аналізу функціонування системи електронної контент-комерції. З метою залучення більшої кількості клієнтів на Web-сторінки СЕКК додають опції аналізу контенту (рейтинги, відгуки і коментарі користувачів) про контент (рис. 13). Сайт із відгуками про контент привертає більшу кількість

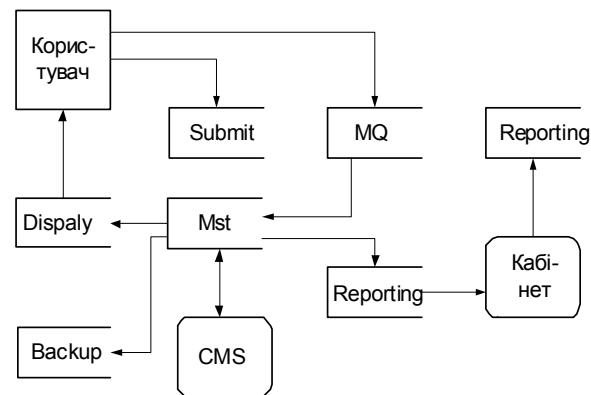
відвідувачів, але ця функціональність не включає модерації та призводить до надмірної інформації на Web-вітрині. Кінцевий споживач змушений сегментувати конструктивну інформацію, що вимагає зусиль та відштовхуває потенційного клієнта. Контент, який потрапляє на інформаційний ресурс, не фільтрується, тобто взаємодія користувачів із порталом є односторонньою.

Підсистема супроводу контенту виконує процес підвищення рейтингу СЕКК із інтерактивним інтерфейсом та вирішує наступні задачі: збирання маркетингової інформації з поширення контенту; підтримка діалогу з кінцевим споживачем контенту; формування каталогу комерційного контенту та послуг СЕКК; інформаційна підтримка кінцевого споживача контенту; формування віртуального кошику клієнта комерційного контенту; реєстрування кінцевого споживача комерційного контенту; Опрацювання замовлення на комерційний контент. Процес проектування СЕКК є інтерактивним і протікає від аналізу до створення прототипу і пробних випробувань, починаючи з формування специфікацій, верстки, формування контенту та його подальше розміщення згідно структури інформаційного ресурсу. Необхідно сконцентрувати рішення на бізнес-цілях і потребах кінцевих користувачів. До визначення функціональних вимог і початку процесу розроблення аналізуються потреби кінцевих користувачів за допомогою листів опитування, альтернатив проектування і прототипів різного ступеня готовності, збирається цінна інформація, одночасно викликаючи у користувачів відчуття участі в процесі проектування та завойовується їх довіра. Для визначення інформаційної архітектури комерційного контенту завдяки користувачам визначають інформацію про робочі групи, минулі/майбутні конференції та всіх членів співтовариства.

По зворотній реакції користувачів стає також зрозуміло, що необхідна чітка, ясна і проста архітектура інформаційного ресурсу. Є три класи користувачів (або персонажів), що приходять на інформаційний ресурс: клієнти, керівники робочих груп і адміністратори, які визначають дизайн інформаційного ресурсу і процес



а



б

Рисунок 13 – Підсистема супроводу контенту: а – структура та б – інформаційна модель

ухвалення рішень. Визначивши, який комерційний контент є важливим, і як він пов'язаний з основними класами користувачів, створюють архітектуру інформаційного ресурсу, його ієрархію, способи подання і способи взаємодії кожного класу користувачів з цією інформацією. Наприклад, контент конференцій містить питання порядку денного або сесії, планування цих питань, теми і питання конференції. Під час аналізу формують додаткові функціональні можливості СЕКК. Наприклад, для підтримки активного співовариства додають форми дискусій і коментарів по вмісту, підтримку контекстної зворотної реакції та інтерактивну взаємодію, використання унікального, але нейтрального бренду або візуального ідентифікатора.

#### 4 ЕКСПЕРИМЕНТИ

Для детального аналізу функціонування СЕКК типу Інтернет-газета та Інтернет-журнал було розроблено та

впровадження вісім різних систем відповідно з підтриманням для кожної з них різною кількістю етапів життєвого циклу контенту. Тобто для різних реалізованих систем були були розроблені не всі компоненти або взагалі не були розроблені підсистеми опрацювання інформаційних ресурсів як формування, управління та супровід контенту. В табл. 1 подано список реалізованих СЕКК із вказанням наявності реалізованих підсистем опрацювання інформаційних ресурсів з підтриманням життєвого циклу комерційного контенту.

#### 5 РЕЗУЛЬТАТИ

У табл. 2 подані результати роботи розроблених систем згідно з Google Analytics за період часу 2010–2014 р.

#### 6 ОБГОВОРЕННЯ

Аналіз результатів супроводу комерційного контенту дозволяє визначити причин формування цільової аудиторії

Таблиця 1 – Реалізація опрацювання інформаційних ресурсів в розроблених СЕКК

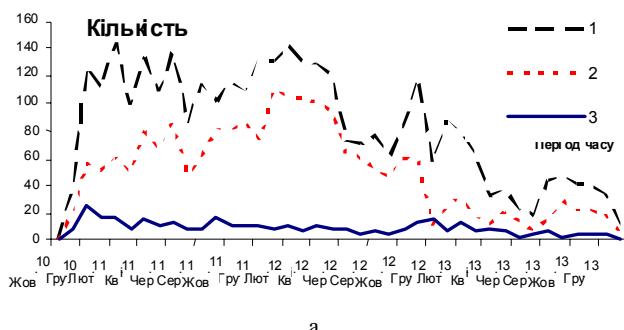
№	Інформаційний ресурс	Адреса ресурсу	Тип	Наявність підсистеми		
				формування	управління	супровід
1	Фотогалерея Висоцьких	fotoghalereja-vysocjkykh.com	журнал	+/-	+	+/-
2	Вголос	vgолос.com.ua	газета	+	+	+
3	Татьяна	tatjana.in.ua	журнал	-	+/-	+/-
4	Прес-Тайм	presstime.com.ua	газета	+/-	+	+
5	AutoChip	www.autochip.vn.ua	журнал	-	+	+/-
6	Курси валют	kursyvalyut.com	газета	+	+	-
7	Добрий ранок	dobryjranok.com	газета	+/-		+/-
8	Інформація для бізнесу	goodmorningua.com	газета	+/-	+/-	-
9	Львівська школа № 3	зсп3львів.in.ua	газета	-	-	-
10	Victana	victana.lviv.ua	журнал	-	+/-	+

Таблиця 2 – Результати роботи систем за період часу з 10.2010 р. по 03.2014 р

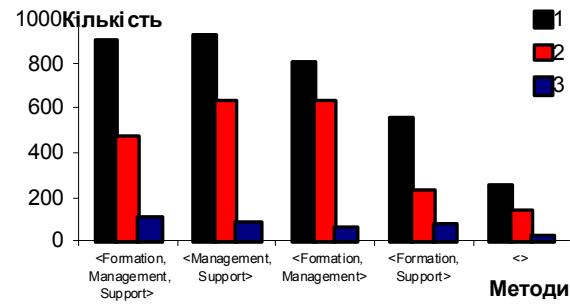
Характеристика	Номер інформаційного ресурсу									
	1	2	3	4	5	6	7	8	9	10
Формування контенту (%)	30	100	10	40	20	90	70	60	0	50
Управління контентом (%)	90	100	50	80	30	40	60	20	0	70
Супровід контенту (%)	30	100	10	40	20	50	80	60	0	70
Унікальність контенту (%)	100	80	100	70	30	20	50	40	100	60
Відвідування	4865	5997052	1381	3654456	9606	20132	8724	25	7	3138
Середня тривалість відвідування (хв:с)	4:41	2:14	3:56	2:04	1:51	1:02	2:27	8:12	0:46	4:15
Показник відмов (%)	56,14	71,90	53,15	83,08	55,67	82,92	68,15	48,0	97,02	32,92
Показник досягнутих конверсій	7,83	0	0	0	12,51	0	0	0	0	0
Унікальні відвідувачі	3215	2501402	728	1501202	7105	16586	4996	7	5	1345
Перегляди сторінок	22071	11588861	5464	769923	2490	31982	18892	81	12	1813
Сторінки/відвідування	4,54	1,93	3,96	1,67	2,59	1,59	2,17	3,24	1,67	5,78
% нових відвідувань	65,45	41,68	52,57	39,88	73,88	82,39	57,23	28,0	97,32	42,86
% повторних відвідувань	34,55	58,32	47,43	60,12	26,12	17,61	42,77	72,0	2,68	57,14
Джерело – not set/non-organic search results	91,14	52,48	62,49	42,46	88,13	58,22	26,984	23	6,72	31,90
Джерело – Organic search (%)	4,67	36,10	22,23	31,22	8,91	34,83	24,34	0	0,13	26,04
Джерело – прямий трафік (%)	2,14	11,20	7,53	26,12	2,35	5,34	48,73	77	93,12	27,88
Джерело – інші сайти (%)	1,99	0,20	3,19	0	0,58	1,55	0,04	0	0,02	6,37
Джерело – соц.мережі (%)	0,06	0,02	4,56	0	0,02	0,06	0,01	0	0,01	7,81

торії за набором характеристик функціонування СЕКК. Регулюючи тематичний набір комерційного контенту, його унікальність, оперативність його формування та адекватне управління ним згідно індивідуальних потреб постійного користувача, можна моделювати межі цільової соціальної аудиторії та кількість унікальних відвіду-

вачів з пошукових систем. На рис. 15–17 подані результати роботи розроблених систем у вигляді графіків, з яких випливає, що при наявності всіх етапів життєвого циклу контенту суттєво збільшується обсяг відвідувань та унікальних користувачів.

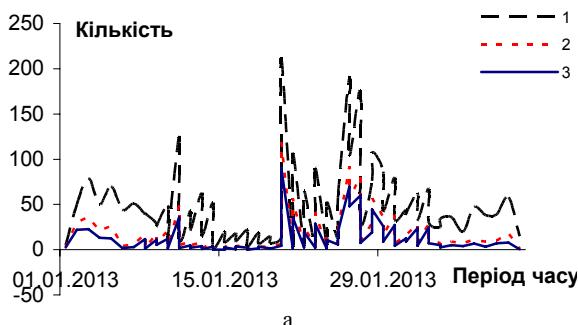


а

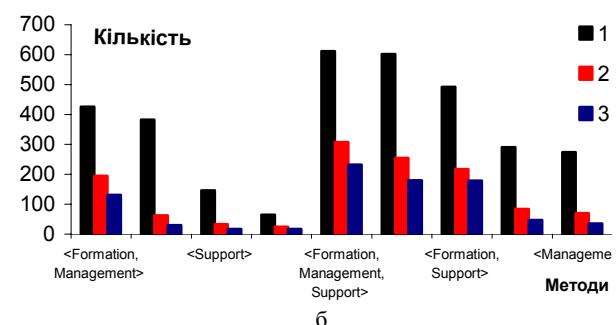


б

Рисунок 14 – Статистичний аналіз функціонування «Фотогалерея Висоцьких»: а – залежність реалізації комерційного контенту від відвідування інформаційного ресурсу [fotogalereja-vysocjkykh.com](http://fotogalereja-vysocjkykh.com); б – співвідношення відвідування інформаційного ресурсу [fotogalereja-vysocjkykh.com](http://fotogalereja-vysocjkykh.com) і реалізації комерційного контенту в залежності від застосування методів опрацювання комерційного контенту, де 1 – кількість всіх відвідувань інформаційного ресурсу, 2 – кількість відвідувань постійних користувачів, 3 – кількість реалізації контенту постійним користувачам СЕКК

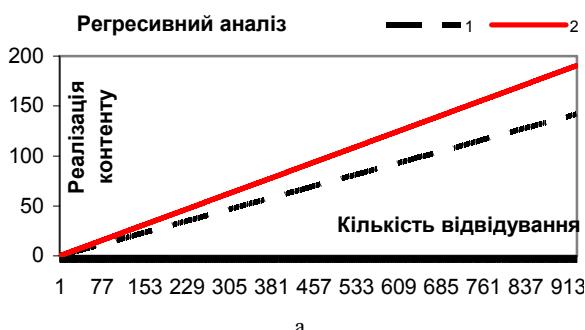


а

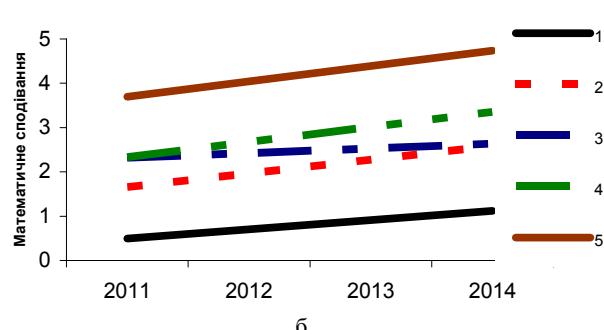


б

Рисунок 15 – Статистичний аналіз функціонування «Victana»: а – статистичний розподіл відвідування інформаційного ресурсу [victana.lviv.ua](http://victana.lviv.ua) та реалізації комерційного контенту; б – залежність відвідування інформаційного ресурсу [victana.lviv.ua](http://victana.lviv.ua) та реалізації комерційного контенту від підключення програмних засобів опрацювання інформаційного контенту, де 1 – кількість всіх відвідувань інформаційного ресурсу, 2 – кількість відвідувань постійних користувачів, 3 – кількість реалізації контенту постійним користувачам СЕКК



а



б

Рисунок 16 – а – Регресивний аналіз збільшення продажів комерційного контенту для інформаційного ресурсу [kursyvalyut.com](http://kursyvalyut.com) (1 – без використання методів, 2 – з використанням методів); б – Математичне сподівання реалізації комерційного контенту через інформаційний ресурс (1 – <>, 2 – <Management>, 3 – <Support>, 4 – <Formation>, 5 – <Formation, Management, Support>)

Сервіс ведення статистики відвідувань інформаційного ресурсу дозволяє оцінити збільшення обсягів продажу комерційного контенту від прямо пропорційної залежності збільшення кількості відвідувань інформаційного ресурсу, кількості постійних користувачів, перспективності маркетингових заходів (рис. 16).

Наявність підсистем формування, управління та супроводу комерційного контенту в системах електронної контент-комерції сприяє збільшенню обсягів реалізації комерційного контенту постійному користувачу на 9 %, активного залучення унікальних відвідувачів, потенційних користувачів та розширення меж цільової та регіональної аудиторії на 11 %, переглянутих сторінок на 12 %, часу відвідування інформаційних ресурсів на 7 %.

## ВИСНОВКИ

У статті розв'язана задача розроблення методів та програмних засобів формування, управління та супроводу інформаційних продуктів у вигляді теоретично обґрунтованої концепції шляхом автоматизації опрацювання інформаційних ресурсів в СЕКК для збільшення обсягів продажу контенту постійному користувачу, активного залучення потенційних користувачів та розширення меж цільової аудиторії.

Наукова новизна роботи полягає в тому, що в процесі теоретичних та експериментальних досліджень отримано такі нові наукові результати. Вперше розроблено методи формування, управління та супроводу комерційного контенту, які базуються на принципах уніфікованості, універсальності і загальності через лінгвістичний аналіз контенту, що дало змогу реалізувати стандартизовані програмні засоби опрацювання інформаційних ресурсів в системах електронної контент-комерції. Удосконалено структуру СЕКК, яка на основі модулів опрацювання інформаційних ресурсів та шляхом розподілу процесів формування, управління та супроводу комерційного контенту, що дає можливість реалізації етапів його життєвого циклу. Отримали подальший розвиток моделі життєвого циклу комерційного контенту, які завдяки оперативної організації процесу опрацювання інформаційних ресурсів в СЕКК підвищують ефективність та рентабельність функціонування аналогічних систем.

Практична цінність отриманих результатів полягає в тому, що розроблено та впроваджено програмні засоби для формування, управління та супроводу комерційного контенту з метою досягнення більшого ефекту від роботи на рівні власника (підвищення рентабельності, зростання інтересу користувачів) та користувача (зрозумілість, спрощення інтерфейсу, уніфікація процесів опрацювання інформаційних ресурсів та розширення вибору функціональних можливостей) систем електронної контент-комерції.

Перспективи подальших досліджень полягають у заснованні запропонованого підходу для розроблення рекомендацій з проектування структури СЕКК шляхом деталізації етапів опрацювання інформаційних ресурсів

та реалізації підсистем формування, управління та супроводу комерційного контенту, що дасть змогу ефективно просто підтримувати життєвий цикл комерційного контенту на рівні розробника систем (зменшення ресурсів та часу на розроблення, покращення якості роботи систем).

## ПОДЯКИ

Роботу виконано в рамках наукових досліджень кафедри інформаційних систем та мереж Національного університету «Львівська політехніка» на тему «Дослідження, розроблення і впровадження інтелектуальних розподілених інформаційних технологій та систем на основі ресурсів баз даних, сковищ даних, просторів даних та знань з метою прискорення процесів формування сучасного інформаційного суспільства». Результати досліджень здійснювались у рамках держбюджетної науково-дослідної роботи на тему «Розробка методів, алгоритмів і програмних засобів моделювання, проектування та оптимізації інтелектуальних інформаційних систем на основі Web-технологій “ВЕБ”» (№ держреєстрації 0102U001171).

## СПИСОК ЛІТЕРАТУРИ

1. Берко А. Системи електронної контент-комерції / А. Берко, В. Висоцька, В. Пасічник. – Львів : НУЛП, 2009. – 612 с.
2. Клифт顿 Б. Google Analytics: профессиональный анализ посещаемости веб-сайтов / Б. Клифттон. – М. : Вильямс, 2009. – 400 с.
3. Основы моделирования и оценки электронных информационных потоков / [Д. Ландэ, В. Фурашев, С. Брайчевский, О. Григорьев]. – К. : Инжиниринг, 2006. – 348 с.
4. Ландэ Д. Основы интеграции информационных потоков: монография / Д. Ландэ. – К. : Инжиниринг, 2006. – 240 с.
5. Математична лінгвістика / [В. Висоцька, В. Пасічник, Ю. Щербина, Т. Шестакевич]. – Л. : Новий Світ-2000, 2012. – 359 с.
6. Berko A. Features of information resources processing in electronic content commerce / A. Berko, V. Vysotska, L. Chyrun // Applied Computer Science. ACS journal. – Volume 10, Number 2. – Poland, 2014. – P. 5–19.
7. CM Lifecycle Poster [Electronic resource] / Content Management Professionals. – Retrieved 20 July 2010. – Access mode: <http://www.cmprosold.org/resources/poster/>.
8. EMC. Content Management Interoperability Services. Appendices. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 17 p.
9. EMC. Content Management Interoperability Services. Part I. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 76 p.
10. EMC. Content Management Interoperability Services. Part II – REST protocol binding. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 79 p.
11. EMC. Content Management Interoperability Services. Part II – SOAP protocol binding. Version 0.5 / EMC, IBM, Microsoft. – Hopkinton : EMC, 2008. – 37 p.
12. Hackos J. Content Management for Dynamic Web Delivery / J. Hackos. – Hoboken : Wiley, 2002. – 432 p.
13. Halvorson K. Content Strategy for the Web / K. Halvorson. – Reading : New Riders Press, 2009. – 192 p.

14. McGovern G. Content Critical / G. McGovern, R. Norton. – Upper Saddle River : FT Press, 2001. – 256 p.
15. McKeever S. Understanding Web content management systems: evolution, lifecycle and market / S. McKeever // Industrial Management & Data Systems (MCB UP). – 2003. – № 103 (9). – P. 686–692.
16. Nakano R. Web content management: a collaborative approach / R. Nakano. – Boston: Addison Wesley Professional, 2002. – 222 p.
17. Papka R. On-line News Event Detection, Clustering, and Tracking : thesis for the degree doctor of philosophy / R. Papka. – Amherst : Massachusetts University, 1999. – 154 p.
18. Woods R. Defining a Model for Content [Electronic resource] / R. Woods. – 2010. – Access mode: [http://www.contentmanager.net/magine/article\\_785\\_defining\\_a\\_model\\_for\\_content\\_governance.html](http://www.contentmanager.net/magine/article_785_defining_a_model_for_content_governance.html).
19. Rockley A. Managing Enterprise Content: A Unified Content Strategy / A. Rockley. – Reading : New Riders Press, 2002. – 592 p.
20. Stone W. R. Plagiarism, Duplicate Publication and Duplicate Submission: They Are All Wrong! / W.R. Stone // IEEE Antennas and Propagation, 2003. – Vol. 45. – № 4. – P. 47–49.
21. Sullivan D. Invisible Web Gets Deeper [Electronic resource] / D. Sullivan // Search Engine Report. – 2002. – Access mode: <http://searchenginewatch.com/sereport/article.php/2162871>.
22. The Content Management Possibilities Poster [Electronic resource] / Metatutorial Services, Inc. – Retrieved 20 July 2010. – Access mode: <http://metatutorial.com/pagea.asp?id=poster>.
23. Vysotska V. Web Content Processing Method for Electronic Business Systems / V. Vysotska, L. Chyrun [Electronic resource] // International Journal of Computers & Technology. – 2013. – Vol 12, № 2. – P. 3211–3220. – Access mode: <http://cirworld.com/index.php/ijct/article/view/3299>.
24. Vysotska V. Designing features of architecture for e-commerce systems [Electronic resource] / V. Vysotska, L. Chyrun // MEST Journal. – Vol. 2, No.1. – 2014. – P. 57–70. – Access mode: [http://mest.meste.org/MEST\\_1\\_2014/R\\_06.pdf](http://mest.meste.org/MEST_1_2014/R_06.pdf).
25. Vysotska V. Analysis and evaluation of risks in electronic commerce / V. Vysotska, I. Rishnyak, L. Chyrun // Experience of Designing and Applications of CAD Systems in Microelectronics (CAD Systems in Microelectronics, CADSM '07) : 9 th International Conference, Lviv, 24 February 2007 : proceedings. Lviv, NU «Lviv'ska Polytechnika, 2007. – P. 332–333.

Стаття надійшла до редакції 17.10.2014.  
Після доробки 04.11.2014.

Берко А. Ю.<sup>1</sup>, Высоцкая В. А.<sup>2</sup>, Чирун Л. В.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, профессор, заведующий кафедрой «Общая экология и екоинформационные системы», Национальный университет «Львовская политехника», Украина

<sup>2</sup>Ассистент кафедры «Информационные системы и сети», Национальный университет «Львовская политехника», Львов, Украина

<sup>3</sup>Канд. техн. наук, доцент кафедры «Программное обеспечение», Национальный университет «Львовская политехника», Львов, Украина.

## ПРОЕКТИРОВАНИЕ И ВНЕДРЕНИЕ СИСТЕМ ЭЛЕКТРОННОЙ КОНТЕНТ-КОММЕРЦИИ

Предложен метод обработки коммерческого контента как этап жизненного цикла контента в системах электронной коммерции. Метод обработки коммерческого контента описывает процессы формирования информационных ресурсов в системах электронной контент-коммерции и упрощает технологию управления коммерческим контентом. В этой работе проанализированы основные проблемы электронной контент-коммерции и функциональных сервисов обработки коммерческого контента. Предложенный метод позволяет создать средства обработки информационных ресурсов в системах электронной контент-коммерции и реализовать подсистемы формирования, управления и сопровождения коммерческого контента.

**Ключевые слова:** информационный ресурс, коммерческий контент, контент-анализ, контент-мониторинг, контентный поиск, система электронной контент-коммерции.

Berko A. Y.<sup>1</sup>, Vysotska V. A.<sup>2</sup>, Chyrun L.V.<sup>3</sup>

<sup>1</sup>Dr.Sc., professor, Head of Department of Environment and Ecosystems, Lviv Polytechnic National University, Ukraine

<sup>2</sup>Assistant Information Systems and Networks Department, Lviv Polytechnic National University, Ukraine

<sup>3</sup>PhD, Associate professor of Software department, Lviv Polytechnic National University, Ukraine

## DESIGN AND IMPLEMENTATION OF ELECTRONIC CONTENT COMMERCE SYSTEMS

The method of commercial content processing as the content life cycle stage in electronic commerce systems is proposed. The method of commercial content processing describes the information resources forming in electronic content commerce systems and automation technology that simplifies the commercial content management. The main problems of electronic content commerce and functional services of commercial content processing are analyzed. The proposed method gives an opportunity to create an instrument of information resources processing in electronic commerce systems and to implement the subsystem of commercial content formation, management and support.

**Keywords:** information resources, commercial content, content analysis, content monitoring, content search, electronic content commerce systems.

## REFERENCES

1. Berko A., Vysotska V., Pasichnyk V. Systemy elektronnoyi kontent-komertsii. L'viv, NULP, 2009, 612 p.
2. Clifton B. Google Analytics: professionalny analiz poseschaemosti web-saytov. Moscow, Vilyams, 2009, 400 p.
3. Lande D., Furashev V., Braychevsky S., Grigorev O. Osnovy modelirovaniya i otsenki elektronnyh informatsionnyh potokov, Kiev, Inzhiniring, 2006, 348 p.
4. Lande D. Osnovy integratsii informatsionnyh potokov: monografiya. Kiev, Inzhiniring, 2006, 240 p.
5. Vysotska V., Pasichnyk V., Scherbyna J., Shestakevych T. Matematychna linhvistyka. L'viv, Novyy Svit-2000, 2012, 359 p.
6. Berko A., Vysotska V., Chyrun L. Features of information resources processing in electronic content commerce, *Applied Computer Science. ACS journal*, Volume 10, No. 2, Poland, 2014, pp. 5–19.
7. CM Lifecycle Poster [Electronic resource], Content Management Professionals. Retrieved 20 July 2010. Access mode: <http://www.cmprosold.org/resources/poster/>.
8. EMC. Content Management Interoperability Services. Appendices. Version 0.5 / EMC, IBM, Microsoft. Hopkinton, EMC, 2008, 17 p.
9. EMC. Content Management Interoperability Services. Part I. Version 0.5 / EMC, IBM, Microsoft. Hopkinton, EMC, 2008, 76 p.
10. EMC. Content Management Interoperability Services. Part II – REST protocol binding. Version 0.5 / EMC, IBM, Microsoft. Hopkinton, EMC, 2008, 79 p.
11. EMC. Content Management Interoperability Services. Part II – SOAP protocol binding. Version 0.5 / EMC, IBM, Microsoft. Hopkinton, EMC, 2008, 37 p.
12. Hackos J. Content Management for Dynamic Web Delivery. Hoboken, Wiley, 2002, 432 p.
13. Halvorson K. Content Strategy for the Web. Reading, New Riders Press, 2009, 192 p.
14. McGovern G., Norton R. Content Critical. Upper Saddle River, FT Press, 2001, 256 p.
15. McGovern G., Norton R. Understanding Web content management systems: evolution, lifecycle and market, *Industrial Management & Data Systems (MCB UP)*, 2003, No. 103 (9), pp. 686–692.
16. Nakano R. Web content management: a collaborative approach. Boston, Addison Wesley Professional, 2002, 222 p.
17. Papka R. On-line News Event Detection, Clustering, and Tracking : thesis for the degree doctor of philosophy. Amherst, Massachusetts University, 1999, 154 p.
18. Woods R. Defining a Model for Content [Electronic resource]. 2010, Access mode: [http://www.contentmanager.net/magazine/article\\_785\\_defining\\_a\\_model\\_for\\_content\\_governance.html](http://www.contentmanager.net/magazine/article_785_defining_a_model_for_content_governance.html).
19. Rockley A. Managing Enterprise Content: A Unified Content Strategy. Reading, New Riders Press, 2002, 592 p.
20. Stone W. R. Plagiarism, Duplicate Publication and Duplicate Submission: They Are All Wrong!, *IEEE Antennas and Propagation*, 2003, Vol. 45, No. 4, pp. 47–49.
21. Sullivan D. Invisible Web Gets Deeper [Electronic resource], *Search Engine Report*. – 2002. – Access mode: <http://searchenginewatch.com/sereport/article.php/2162871>.
22. The Content Management Possibilities Poster [Electronic resource]. Metatorial Services, Inc, Retrieved 20 July 2010, Access mode: <http://metatorial.com/pagea.asp?id=poster>.
23. Vysotska V., Chyrun L. Web Content Processing Method for Electronic Business Systems [Electronic resource], *International Journal of Computers & Technology*, 2013, Vol 12, No. 2, pp. 3211–3220, Access mode: <http://cirworld.com/index.php/ijct/article/view/3299>.
24. Vysotska V., Chyrun L. Designing features of architecture for e-commerce systems [Electronic resource], *MEST Journal*, Vol. 2, No. 1, 2014, pp. 57–70, Access mode: [http://mest.meste.org/MEST\\_1\\_2014/R\\_06.pdf](http://mest.meste.org/MEST_1_2014/R_06.pdf).
25. Vysotska V., Rishnyak I., Chyrun L. Analysis and evaluation of risks in electronic commerce, Experience of *Designing and Applications of CAD Systems in Microelectronics (CAD Systems in Microelectronics, CADSM '07) : 9th International Conference*. Lviv, 24 February 2007, proceedings, NU «Lviv'ska Polytechnika», 2007, pp. 332–333.

УДК 681.518:004.93.1

Довбиш А. С.<sup>1</sup>, Шелехов І. В.<sup>2</sup>, Прилепа Д. В.<sup>3</sup><sup>1</sup>Д-р техн. наук, професор, зав. каф. комп’ютерних наук, Сумський державний університет, Україна<sup>2</sup>Канд. техн. наук, доцент кафедри комп’ютерних наук, Сумський державний університет, Україна<sup>3</sup>Аспірант кафедри комп’ютерних наук, Сумський державний університет, Україна

## ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНИЙ АЛГОРИТМ НАВЧАННЯ СИСТЕМИ ДІАГНОСТУВАННЯ ЕМОЦІЙНО-ПСИХІЧНОГО СТАНУ ЛЮДИНИ

Розглядається метод розпізнавання емоційно-психічного стану людини за зображеннями її обличчя. При цьому запропоновано для одержання додаткової інформації використовувати такі фрагменти зображення обличчя людини, які включають зони очей і носу. Формування вхідного математичного опису системи розпізнавання здійснюється шляхом аналізу лівопівкульних та правопівкульних зображень обличчя людини. При цьому попередньо для стабільного емоційно-психічного стану особи формується навчальна матриця яскравості зображення, як всього обличчя, так і його відповідних фрагментів. Машинне навчання здійснюється в рамках інформаційно-екстремальної інтелектуальної технології, що ґрунтуються на максимізації інформаційної спроможності системи розпізнавання. Як критерій функціональної ефективності машинного навчання використано модифіковану інформаційну міру Кульбака, яка розглядається як функціонал від точнісних характеристик двохальтернативних рішень. За побудованою категорійною моделлю у вигляді відображення множин, задіяних в процесі навчання, розроблено інформаційно-екстремальний алгоритм оптимізації геометричних параметрів гіперсферичних контейнерів класів розпізнавання, що відновлюються в процесі навчання в радіальному базисі простору ознак. За результатами фізичного моделювання доведено, що фрагменти обличчя є достатньо інформативними при розпізнаванні емоційно-психічного стану людини.

**Ключові слова:** інформаційно-екстремальна інтелектуальна технологія, комп’ютеризована система діагностування, психодіагностування, навчання, критерій функціональної ефективності.

### НОМЕНКЛАТУРА

ІЕЛ-технологія – інформаційно-екстремальна інтелектуальна технологія;

СППР – система підтримки прийняття рішень;

КФЕ – критерій функціональної ефективності;

$m$  – номер поточного класу розпізнавання;

$M$  – потужність алфавіту класів розпізнавання;

$n$  – кількість векторів-реалізацій (далі просто реалізації);

$x_m$  – еталонна реалізація, визначає геометричний центр контейнера класу  $X_m^o$ ;

$d_m$  – радіус контейнера класу  $X_m^o$ ;

$E_m$  – інформаційний КФЕ навчання системи розпізнавання реалізації класу  $X_m^o$ ;

$G_E$  – робоча (допустима) область визначення функції КФЕ;

$G$  – простір вхідних факторів;

$T$  – множина моментів часу зняття інформації;

$\Omega$  – простір ознак розпізнавання;

$Z$  – простір можливих емоційно-психологічних станів людини;

$Y$  – вибіркова множина (вхідна навчальна матриця);

$\tilde{\mathfrak{R}}^{[2]}$  – нечітке розбиття простору ознак розпізнавання;

$I^{[l]}$  – множина допустимих гіпотез;

$l$  – кількість статистичних гіпотез;

$\mathfrak{J}^{[q]}$  – множина точнісних характеристик, де  $q = l^2$ ;

$E$  – терм-множина значень інформаційного КФЕ;

$\tilde{L}^{[2]}$  – множина діагнозів емоційно-психологічних станів людини;

$\theta$  – оператор, який буде розбиття  $\tilde{\mathfrak{R}}^{[2]}$  простору ознак на класи розпізнавання;

$\Psi$  – оператор, що перевіряє основну статистичну гіпотезу про належність реалізації  $x_{m,l}^{(j)}$  класу  $X_m^o$ ;

$\gamma$  – оператор, що формує множину точнісних характеристик для заданої системи оцінок рішень;

$\varphi$  – оператор, що обчислює множину значень інформаційного КФЕ;

$r$  – оператор, що реалізує ітераційний процес оптимізації геометричних параметрів розбиття  $\tilde{\mathfrak{R}}^{[2]}$ ;

$U$  – оператор, що регламентує процес навчання і дозволяє оптимізувати параметри його планування;

$G_d$  – радіус контейнерів класів розпізнавання;

$i$  – номер ознаки розпізнавання фрагмента зображення;

$j$  – номер реалізації класів фрагмента зображення обличчя пацієнта;

$D_1^{(k)}$  – перша достовірність, обчислена на  $k$ -му кроці навчання;

$D_2^{(k)}$  – друга достовірність, обчислена на  $k$ -му кроці навчання;

$\alpha^{(k)}$  – помилка першого роду, обчислена на  $k$ -му кроці навчання;

$\beta^{(k)}$  – помилка другого роду, обчислена на  $k$ -му кроці навчання;

$10^{-r}$  – мале число для уникнення поділу на нуль.

## ВСТУП

Технологія психодіагностування та корекції привертала увагу спеціалістів з давніх часів. Комп'ютерні технології психодіагностування за зображенням обличчя та його фрагментів дозволяють підвищити достовірність оцінки психічного стану людини та функціональну ефективність його корекції [1, 2].

Відомі методи розпізнавання зображень [3, 4] носять в основному модельний характер, оскільки не враховують перетин класів розпізнавання, що характеризують різні психічні стани пацієнта та довільні початкові умови формування зображень. Тому основним шляхом підвищення функціональної ефективності системи психодіагностування за зображеннями обличчя пацієнта є надання її властивості адаптивності на основі машинного навчання. Одним із перспективних підходів до вирішення цієї проблеми є застосування ідей і методів IEI-технології, яка основана на максимізації інформаційної спроможності системи підтримки прийняття рішень в процесі її навчання [5]. У працях [1, 6] запропоновано в рамках IEI-технології алгоритм навчання системи психодіагностування за зображенням обличчя, але асимптотична достовірність розпізнавання виявилася недостатньо високою.

В статті розглядається інформаційно-екстремальний алгоритм навчання СППР для відео-комп'ютерної психодіагностики та корекції, з використанням фрагментів зображення обличчя людини.

## 1 ПОСТАНОВКА ЗАДАЧІ

Розглянемо формалізовану постановку задачі інформаційного синтезу для розпізнавання емоційно-психічного стану людини за фрагментами зображення її обличчя. Нехай дано алфавіт  $\{X_m^o \mid m = \overline{1, M}\}$  класів розпізнавання, які характеризують різні психологічні стани пацієнта, і навчальну матрицю яскравості зображення обличчя пацієнта  $\|y_{m,i}^{(j)}\|$ ,  $i = \overline{1, N}$ ,  $j = \overline{1, n}$ . При цьому рядок матриці  $\{y_{m,i}^{(j)} \mid i = \overline{1, N}\}$  визначає  $j$ -ту реалізацію, а стовпчик  $\{y_{m,i}^{(j)} \mid j = \overline{1, n}\}$  – навчальну вибірку значень  $i$ -ї ознаки. Відомий структурований вектор параметрів навчання СППР  $g = \langle x_m, d_m \rangle$ . Необхідно на етапі навчання СППР оптимізувати координати вектора  $g$  шляхом пошуку глобально-го максимуму усередненого за алфавітом класів розпізнавання інформаційного КФЕ навчання системи

$$\bar{E} = \frac{1}{M} \sum_{m=1}^M \max_{G_E} E_m, \quad (1)$$

На етапі екзамену, тобто безпосереднього діагностуванні, необхідно прийняти рішення про належність реалізації, що розпізнається, до одного із класів заданого алфавіту.

## 2 ОГЛЯД ЛІТЕРАТУРИ

Розпізнавання певної емоції по обличчю людини спочатку розглядалося як суто медична проблема [7, 8]. Практична цінність цих праць полягає в структуризації елементів обличчя та їх класифікації залежно від психоемоційного стану людини, що дозволяє формувати вхідний математичний опис комп'ютерних систем психодіагностування.

В праці [9, 10] запропоновано комп'ютерна емоційно-орієнтована технологія, яка встановлює відповідність фрагментів обличчя різним емоційним станам людини. При цьому, для розпізнавання емоцій використовується такі традиційні методи, як байесівський класифікатор, штучні нейронні мережі та нечіткі методи подання, та виведення знань. Основним недоліком розглянутих методів розпізнавання емоційного стану людини є їх модельність, оскільки вони не враховують довільні початкові умови формування образів та вплив неконтрольованих факторів. Крім того, авторами не розглядається зв'язок емоційного і психічного станів людини. Загальним основним недоліком методів є застосування структурних або дистанційних мір схожості, що зменшує достовірність класифікації за умов апріорно нечіткого розбиття простору ознак на класи розпізнавання.

В праці [11] розглянуто застосування теоретико-інформаційного підходу до розпізнавання та відстеження положення обличчя та його структурних елементів. При цьому, інформаційна міра використовується як критерій різноманітності об'єктів, але не розглядалася задача інформаційного синтезу системи розпізнавання психоемоційного стану людини за обличчям.

Таким чином, з метою усунення визначених недоліків перспективним є створення в рамках теоретико-інформаційного підходу систем психодіагностування на основі ідей машинного навчання та розпізнавання образів, що дозволить надати їм властивість адаптивності за умов апріорної невизначеності.

## 3 МАТЕРІАЛИ І МЕТОДИ

Категорійну модель процесу навчання системи психодіагностування в рамках IEI-технології подамо у вигляді діаграми відображення множин. При цьому вхідний математичний опис СППР представимо у вигляді структури

$$\Delta_B = \langle G, T, \Omega, Z, Y; \Pi, \Phi, f \rangle,$$

де  $\Pi : G \times T \times \Omega \rightarrow Z$  – оператор переходів, що відбуває механізм зміни емоційно-психологічних станів;  $\Phi : G \times T \times \Omega \times Z \rightarrow Y$  – оператор оброблення зображення (формування вхідної навчальної матриці  $Y$ );  $f : Y \rightarrow X^P$  – оператор відображає вхідну навчальну матрицю яскравості  $Y$  зображення обличчя в бінарну матрицю  $X^{|P|}$ , яка враховує значення яскравості виділених на зображенні заданих  $P$  фрагментів.

Діаграму відображень множин у процесі навчання СППР за інформаційно-екстремальним алгоритмом з оптимізацією геометричних параметрів гіперсферичних контейнерів класів розпізнавання показано на рис. 1.

На діаграмі (рис. 1) правий контур операторів реалізує в рамках IEI-технології базовий алгоритм навчання СППР, основною задачею якого є оптимізація геометричних параметрів розбиття  $\tilde{\mathfrak{R}}^{[2]}$  при заданій системі контрольних допусків на ознаки розпізнавання.

#### 4 ЕКСПЕРИМЕНТИ

Розглянемо модифікований алгоритм навчання СППР для психодіагностування за фрагментами зображення обличчя пацієнта в рамках IEI-технології, який відновлює в радіальному базисі простору ознак гіперсферичні контейнери класів розпізнавання. При цьому оптимальний радіус контейнера класу розпізнавання визначається як

$$d_m^* = \arg \max_{G_E \cap G_d} \bar{E}. \quad (2)$$

Вхідною інформацією для навчання СППР є багатовимірна навчальна матриця яскравості пікселів рецепторного поля зображення обличчя  $\|y_{m,I}^{(J)}\| \mid m = \overline{1, M}; I = \overline{1, N}; J = \overline{1, n} \|$ ; система полів контрольних допусків  $\{\delta_i\}$  на ознаки розпізнавання і рівні селекції  $\{\rho_m\}$  координат еталонних двійкових векторів-реалізацій класів розпізнавання. Крім того, задається структура координат вершин прямокутних фрагментів зображення обличчя пацієнта  $\langle i_B, i_E, j_B, j_E \rangle$ . При цьому потужність словника ознак розпізнавання, сформованого за фрагментом зображенень, дорівнює  $i_E - i_B$  ознак, а кількість реалізацій  $- j_E - j_B$ .

Розглянемо основні етапи реалізації базового алгоритму навчання:

1. Формується по фрагменту зображення бінарна навчальна матриця  $\|x_{m,i}^{(j)}\|^{(p)}$ , елементи якої дорівнюють

$$x_{m,i}^{(j)} = \begin{cases} 1, & \text{якщо } y_{m,i}^{(j)} \in \delta_i; \\ 0, & \text{якщо } y_{m,i}^{(j)} \notin \delta_i. \end{cases} \quad (3)$$

де  $i = i_B \dots i_E$ ,  $j = j_B \dots j_E$ .

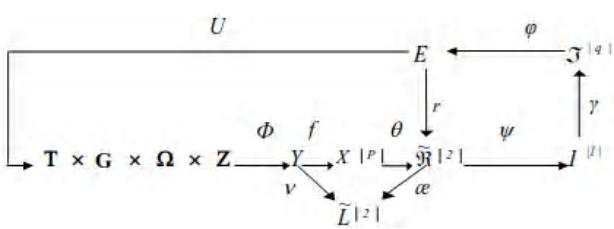


Рисунок 1 – Категорійна модель навчання СППР

2. Формування масиву еталонних двійкових векторів  $\{x_{m,i} \mid m = \overline{1, M}\}$ , елементи якого визначаються за правилом:

$$x_{m,i} = \begin{cases} 1, & \text{якщо } \frac{1}{n} \sum_{j=1}^n x_{m,i}^{(j)} > \rho_m; \\ 0, & \text{інакше.} \end{cases} \quad (4)$$

3. Для кожного класу  $X_m^{(p)}$  здійснюється пошук найближчого «сусіда»  $X_l^{(p)}$ , для якого виконується умова  $l = \arg \min_{\{t\}} d(x_m \oplus x_t)$ , де  $\{t\} = \{\overline{1, m-1}; \overline{m+1, M}\}$ .

4. Оптимізація кодової відстані  $d_m^{(p)}$  відбувається за рекурентною процедурою  $d_m(k) = [d_m(k-1) + h \mid d_m(k) \in G_m^d]$ . При цьому приймається  $E_m^{(p)}(0) = 0$ .

3. Для кожного класу  $X_m^{(p)}$  здійснюється пошук найближчого «сусіда»  $X_l^{(p)}$ , для якого виконується умова  $l = \arg \min_{\{t\}} d(x_m \oplus x_t)$ , де  $\{t\} = \{\overline{1, m-1}; \overline{m+1, M}\}$ .

4. Оптимізація кодової відстані  $d_m^{(p)}$  відбувається за рекурентною процедурою  $d_m(k) = [d_m(k-1) + h \mid d_m(k) \in G_m^d]$ . При цьому приймається  $E_m^{(p)}(0) = 0$ .

5. Процедура закінчується при знаходженні максимуму КФЕ в робочій області визначення його функції

$$E_m^{*(p)} = \max_{\{d\}} E_m, \quad (5)$$

де  $\{d\} = \{d_1, \dots, d_{\max}\} \in [0; d(x_m \oplus x_l) - 1]$  – множина радіусів концентрованих гіперсфер, центр яких визначається вершиною еталонного вектора  $x_m^{(p)} \in X_m^{(p)}$ . При цьому множина  $\{d\}$  є так само множиною кроків навчання системи розпізнавання.

Таким чином, основною функцією базового алгоритму навчання у рамках IEI-технології є обчислення на кожному кроці навчання інформаційного КФЕ і організація пошуку його глобального максимуму в робочій області визначення функції критерію з метою визначення оптимальних геометричних параметрів контейнерів класів розпізнавання.

Як критерій оптимізації параметрів навчання у рамках IEI-технології може розглядатися будь-яка статистична інформаційна міра, яка є функціоналом від точнісних характеристик. Широкого використання в алгоритмах навчання набула модифікація інформаційної міри Кульбака [12]:

$$E_m^{(k)} = 0,5 \log_2 \left( \frac{D_1^{(k)} + D_2^{(k)} + 10^{-r}}{\alpha^{(k)} + \beta^{(k)} + 10^{-r}} \right) * \left[ (D_1^{(k)} + D_2^{(k)}) - (\alpha^{(k)} + \beta^{(k)}) \right]. \quad (6)$$

## 5 РЕЗУЛЬТАТИ

Формування навчальної матриці здійснювалося за відомою технологією відео-комп'ютерної психодіагностики та корекції, яка базується на порівнянні лівопівкульного та правопівкульного портретів для оцінки психологічного стану людини. Вхідний математичний опис СППР було сформовано за графічними даними, наведеними в праці [2]. На рис. 2 показано портрети особи з нестабільним станом, які характеризують клас  $X_1^o$ , сформований із лівопівкульних (рис. 2a) і клас  $X_2^o$ , сформований із правопівкульних портретів (рис. 2b). Відповідно сформовано портрети особи із стабільним емоційно-психологічним станом – клас  $X_3^o$  (рис. 2c) і клас  $X_4^o$  (рис. 2d).

Результат оптимізації за критерієм (6) геометричних параметрів контейнерів класів розпізнавання, що відновлюються в радіальному базисі простору ознак за ітераційною процедурою базового алгоритму навчання, для повних портретів (рис. 2) показано на рис. 3.

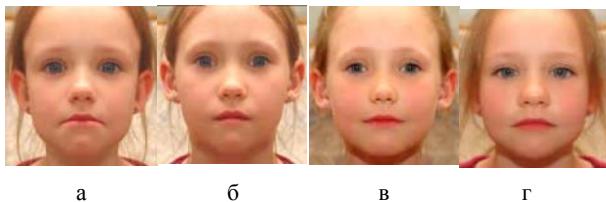


Рисунок 2 – Портрети, які використані при психодіагностиці: а – клас  $X_1^o$ ; б – клас  $X_2^o$ ; в – клас  $X_3^o$ ; г – клас  $X_4^o$

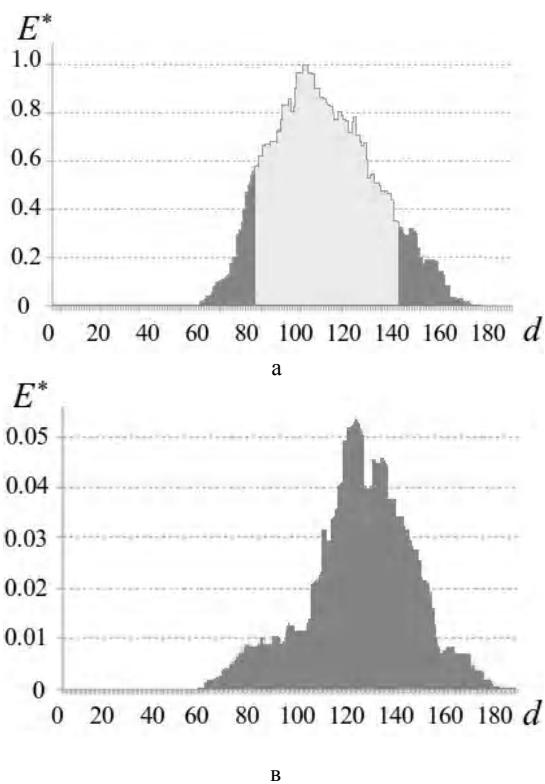


Рисунок 3 – Графіки залежності КФЕ (6) від радіусів контейнерів класів розпізнавання для повного портрету: а – клас  $X_1^o$ ; б – клас  $X_2^o$ ; в – клас  $X_3^o$ ; г – клас  $X_4^o$

На рис. 3 світлим кольором показано робочі (допустимі) області визначення функції КФЕ, в яких перша та друга достовірності перевершують відповідно помилки першого та другого роду. Аналіз рис. 3 показує, що незначні максимальні значення КФЕ характерні для особи зі стабільним емоційно-психологічним станом, що пояснюється схожістю лівопівкульних і правопівкульних портретів. Крім того, через близькість центрів класів  $X_3^o$  і  $X_4^o$  на рис. 3 і рис. 3г відсутні робочі області.

З метою дослідження інформативності фрагментів обличчя було сформовано навчальну матрицю яскравості для зображень фрагментів, які характеризували зону очей та носа. На рис. 4 показано портрети особи з нестабільним станом, які характеризують клас  $X_1^o$ , сформований із лівопівкульних (рис. 4a) і клас  $X_2^o$ , сформований із правопівкульних портретів (рис. 4b). Відповідно сформовано портрети особи із стабільним емоційно-психологічним станом – клас  $X_3^o$  (рис. 4c) і клас  $X_4^o$  (рис. 4d).

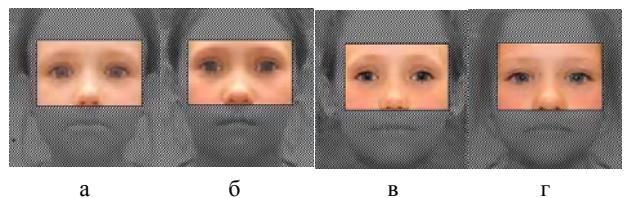
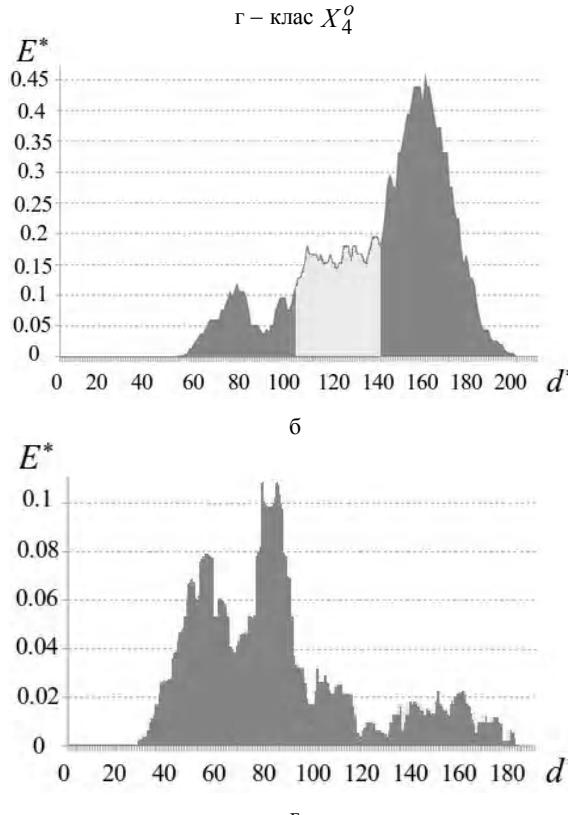


Рисунок 4 – Фрагменти портрета обличчя, які включають зону очей та носу: а – клас  $X_1^o$ ; б – клас  $X_2^o$ ; в – клас  $X_3^o$ ; г – клас  $X_4^o$



Графіки залежності динаміки зміни КФЕ при формуванні вирішальних правил для портретів пацієнта до та після проведення психокорекції показано на рис. 5.

Аналіз рис. 5а і рис. 5б, одержаних за ліво- та право-півкульним портретом особи з нестабільним емоційно-психологічним станом, вказує на можливість створення вирішальних правил, перша та друга достовірності яких перевищують 0,5. Результати навчання для особи зі стабільним емоційно-психологічним станом за аналогічними алгоритмами, які подано на рис. 5в і рис. 5г, показують неможливість формування релевантних вирішальних правил, що свідчить про високий ступінь схожості ліво- та право-півкульних портретів. Таким чином, запропонований алгоритм навчання дозволяє будувати вирішальні правила, здатні визначати різноманітність між право- і лівопівкульними портретами обличчя.

На наступному етапі, з метою дослідження інформативності фрагментів обличчя було сформовано навчальну матрицю яскравості для зображень фрагментів, які характеризували зону носа (рис. 6).

Графіки залежності динаміки зміни КФЕ при формуванні вирішальних правил для портретів пацієнта до та після проведення психокорекції показано на рис. 7.

Аналіз рис. 7а і рис. 7б, одержаних за ліво- та право-півкульним портретом особи з нестабільним емоційно-психологічним станом, вказує на можливість створення вирішальних правил. Результати навчання для особи зі стабільним емоційно-психологічним станом за аналогічними алгоритмами, які подано на рис. 7в і рис. 7г, показують неможливість формування релевантних вирішальних правил, що свідчить про високий ступінь схожості ліво- та право-півкульних портретів. Таким чином, запропонований алгоритм навчання дозволяє будувати вирішальні правила, здатні визначати різноманітність між право- і лівопівкульними портретами обличчя.

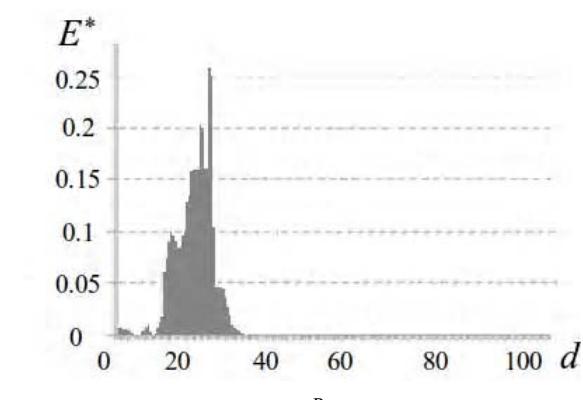
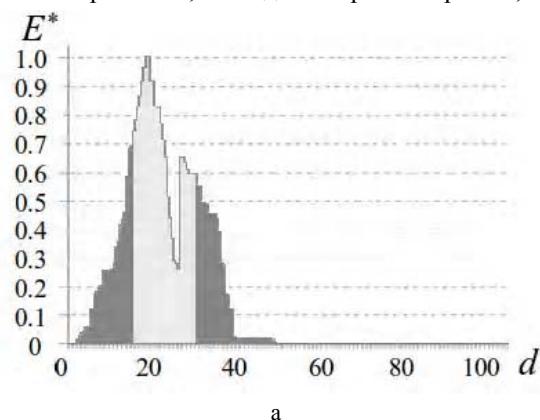


Рисунок 5 – Графіки залежності КФЕ (6) від радіусів контейнерів класів розпізнавання для фрагментів портрету, які включають зону очей та носу: а – клас  $X_1^o$ ; б – клас  $X_2^o$ ; в – клас  $X_3^o$ ; г – клас  $X_4^o$

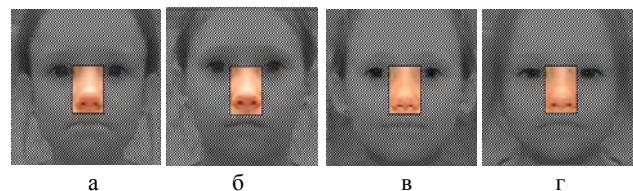


Рисунок 6 – Фрагменти портрета, що включають зону носа:

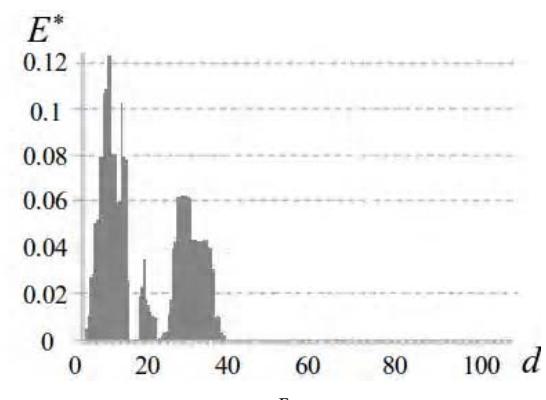
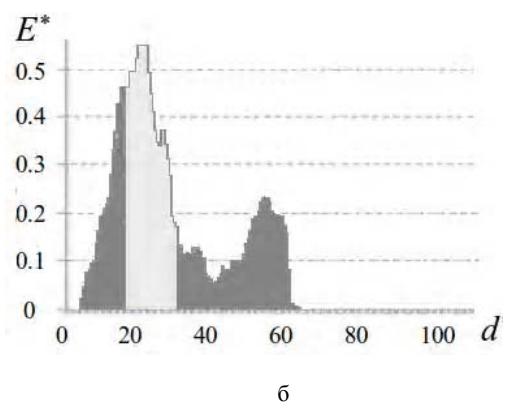
а – клас  $X_1^o$ ; б – клас  $X_2^o$ ; в – клас  $X_3^o$ ; г – клас  $X_4^o$

зують неможливість формування релевантних вирішальних правил, оскільки значення КФЕ є близькими до нуля, що свідчить про високий ступінь схожості ліво- та право-півкульних портретів. Таким чином, запропонований алгоритм навчання дозволяє будувати вирішальні правила, здатні визначати різноманітність між право- і лівопівкульними портретами обличчя.

## 6 ОБГОВОРЕННЯ

Аналогічний аналіз зображень пацієнта, що були отримані на різних етапах психокорекції, дозволяє відслідкувати динаміку зміни КФЕ при формуванні вирішальних правил протягом всього періоду лікування (рис. 8).

Аналіз рис. 8а показує, що на другому етапі лікування відбулося погіршення психологічного стану. На це вказує збільшення значення КФЕ при формуванні вирішальних правил. Після цього КФЕ, а отже і відмінність між право-півкульним і лівопівкульним портретами, поступово зменшувалася і на четвертому кроці відповідні вирішальні правила не було сформовано. Аналіз рис. 8б показує, що на другому етапі лікування не відбулося



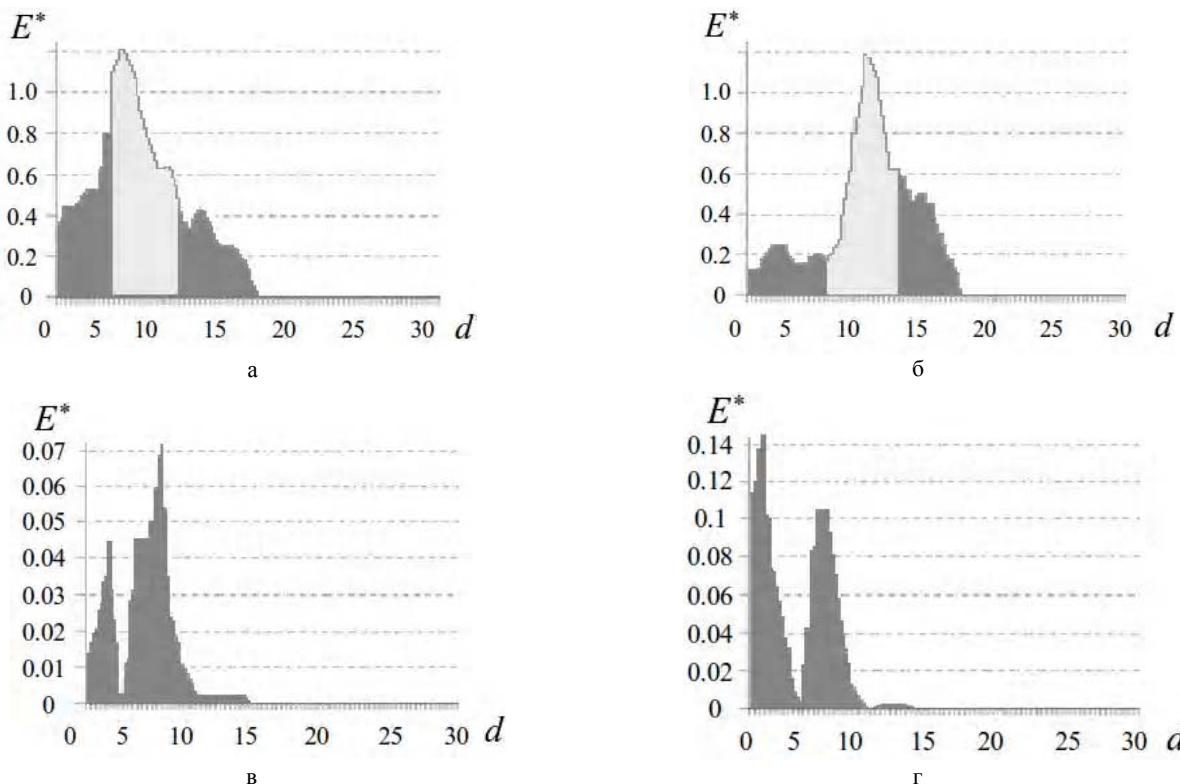


Рисунок 7 – Графіки залежності КФЕ (6) від радіусів контейнерів класів розпізнавання для фрагментів портрету, які включають зону носу: а – клас  $X_1^O$ ; б – клас  $X_2^O$ ; в – клас  $X_3^O$ ; г – клас  $X_4^O$

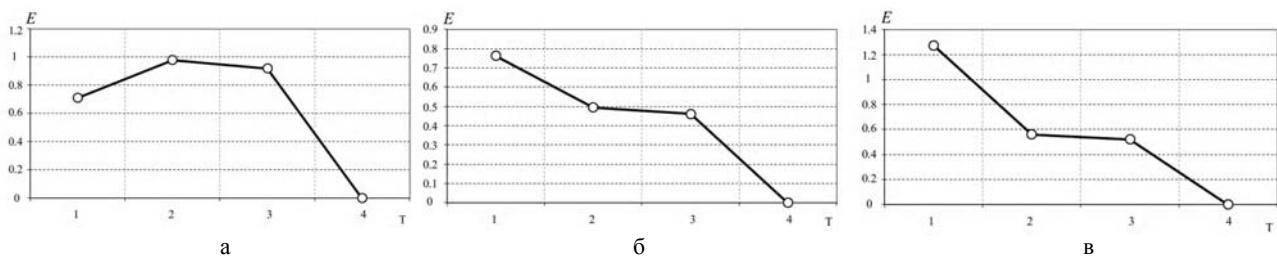


Рисунок 8 – Графіки зміни КФЕ (6) при формуванні вирішальних правил протягом всього періоду лікування:  
 а – повний портрет; б – зона очей і носу; в – зона носу

погіршення психологічного стану, але на третьому етапі мало місце певне «гальмування» процесу психокорекції. На це вказують незначні зміни значення КФЕ при формуванні вирішальних правил на третьому етапі лікування відносно другого етапу. На четвертому кроці аналогічно до рис. 8а відповідні вирішальні правила не було сформовано. Аналіз рис. 8в показує, що аналогічно до рис. 8б, на другому етапі лікування не відбулося погіршення психологічного стану, але на третьому етапі мало місце певне «гальмування» процесу психокорекції. На це вказують незначні зміни значення КФЕ при формуванні вирішальних правил на третьому етапі лікування відносно другого етапу. На четвертому кроці аналогічно до рис. 8а–8б відповідні вирішальні правила не було сформовано.

Узагальнена траекторія психокорекції за динамікою зміни КФЕ для двох різних частин портрету подана на рис. 9.

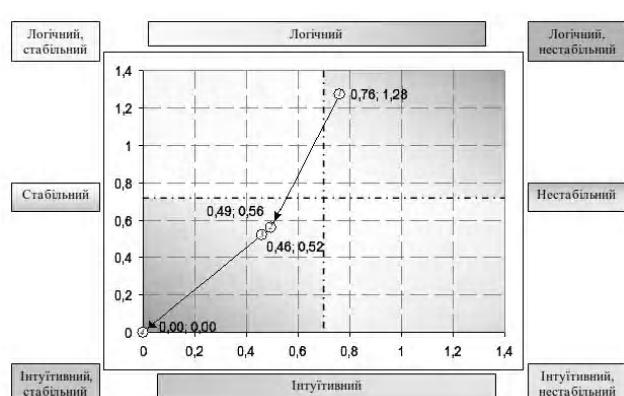


Рисунок 9 – Траекторія психокорекції за динамікою зміни КФЕ

На рис. 9 траєкторія психокорекції формується на площині Е<sub>1</sub>Е<sub>2</sub>. При цьому дана площа розбита на чотири області, що відповідають об'єктивній типології психологочних станів людини [2]. В верхній правій частині показана область логічних дестабільних психологічних станів, правій нижній – інтуїтивних дестабільних, лівій верхній – логічних стабільних, лівій нижній – інтуїтивних стабільних. Використання такого способу відображення процесу лікування дозволяє спостерігати за коректністю вжитих засобів, ефективністю окремих етапів та лікування в цілому.

## ВИСНОВКИ

1. В роботі запропоновано модифікацію технології відео-комп'ютерної психодіагностики та корекції, що застосовується в Центрі Екології Людини НІКАР при Інституті Проблем Керування Російської Академії Наук шляхом її інтеграції з інформаційно-екстремальною інтелектуальною технологією.

2. Для сформованих навчальних матриць на базі фрагмента ліво- та право півкульних портретів запропоновано алгоритм обчислення інформаційної міри між цими портретами з метою визначення їх різноманітності. Порівняння максимальних значень інформаційних мір, отриманих при оптимізації параметрів навчання системи за портретами людини до та після проведення психокорекції показало, що використання частини портрета пацієнта для психоаналізу є більш ефективним, оскільки портрет може бути нестабільним, наприклад, через вікові особливості чи різні специфічні обмеження.

3. Показана можливість використання лише частини зображення для психодіагностики та об'єднання результатів психодіагностики за окремими фрагментами портрету, що дозволяє відобразити процес лікування у вигляді траєкторії психокорекції на площині об'єктивних типологій психологічних станів людини.

## ПОДЯКИ

Робота виконана в рамках держбюджетних науково-дослідницьких тем «Розроблення науково-методичних основ та інформаційних засобів аналізу і синтезу інтелектуальних систем, що самонавчаються» (№ 52.17.02-01.12/13 ЗП), «Інтелектуальна система діагностування та прогнозування перебігу і наслідків патологічного про-

цесу» (№ 52.17.02-02.13\14 ЗП) на базі лабораторії інтелектуальних систем кафедри комп'ютерних наук Сумського державного університету.

## СПИСОК ЛІТЕРАТУРИ

- Шелехов І. В. Оптимізація параметрів навчання комп'ютеризованої системи діагностування емоційно-психічного стану людини / І. В. Шелехов, Д. В. Прилепа // Радіоелектронні і комп'ютерні системи. – 2014. – №1(65). – С. 161–167.
- Ануашвили А. Н. Объективная психология на основе волновой модели мозга / А.Н. Ануашвили. – М. : Экон-Информ, 2008. – 292 с.
- Анисимов Б. В. Распознавание и цифровая обработка изображений / Б. В. Анисимов, В. Д. Курганов, В. К. Злобин. – М. : Высшая школа, 1983. – 256 с.
- Лютгер Дж. Ф. Искусственный интеллект. Стратегии и методы решения сложных проблем / Дж. Ф. Лютгер. – М. : Вильямс, 2003. – 864 с.
- Довбиш А. С. Основи проектування інтелектуальних систем : навчальний посібник / А. С. Довбиш. – Суми : СумДУ, 2009. – 171 с.
- Шелехов І. В. Комп'ютеризована системи діагностування емоційно-психічного стану людини за фрагментами зображення обличчя / І. В. Шелехов, Д. В. Прилепа // Advanced information systems and technologies : матеріали третьої міжнародної науково-практичної конференції (Суми, 14–16 травня 2014 р.), AIST. – 2014. – С. 118–119.
- Meneghini F. Clinical Facial Analysis Elements, Principles, and Techniques / F. Meneghini, P. Biondi. – Berlin : Springer, 2012. – 219 р.
- George R. M. Facial Geometry: Graphic Facial Analysis for Forensic Artists / R. M. George. – Charles C Thomas Publisher, 2007. – 82 р.
- Petta P. Emotion-Oriented Systems. The Humaine Handbook / P. Petta, C. Pelachaud, R. Cowie. – Berlin : Springer, 2011. – 816 р.
- André E. Experimental methodology in emotion-oriented computing / E. André // IEEE Pervasive Computing. – 2011. – Vol. 10 (3). – P. 54–57.
- Colmenarez A. J. Facial Analysis from Continuous Video with Applications to Human-Computer Interface / A. J. Colmenarez, Z. Xiong, T-S. Huang. – Berlin : Springer Science, 2004. – 134 р.
- Довбиш А. С. Інтелектуальні інформаційні технології в електронному навчанні / А. С. Довбиш, А. В. Васильєв, В. О. Любчак. – Суми : СумДУ, 2014. – 172 с.

Стаття надійшла до редакції 04.11.2014.

Після доробки 26.11.2014.

Довбыш А. С.<sup>1</sup>, Шелехов И. В.<sup>2</sup>, Прилепа Д. В.<sup>3</sup>

<sup>1</sup>Д-р техн. наук, профессор, зав. каф. комп'ютерных наук, Сумський державний університет, Україна

<sup>2</sup>Канд. техн. наук, доцент кафедри комп'ютерных наук Сумського державного університета, Україна

<sup>3</sup>Аспірант кафедри комп'ютерных наук, Сумський державний університет, Україна

## ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНИЙ АЛГОРИТМ ОБУЧЕННЯ СИСТЕМИ ДІАГНОСТИКИ ЕМОЦІОНАЛЬНО-ПСИХІЧЕСКОГО СОСТОЯННЯ ЧЕЛОВЕКА

Рассматривается метод распознавания эмоционально-психического состояния человека по изображению его лица. При этом предложено для получения дополнительной информации использовать такие фрагменты изображения лица человека, которые включают зоны глаз и носа. Формирование входного математического описания системы распознавания осуществляется путем анализа левополушарных и правополушарных изображений лица человека. При этом предварительно для стабильного эмоционально-психического состояния лица формируется учебная матрица яркости изображения, как всего лица, так и его соответствующих фрагментов. Машинное обучение осуществляется в рамках информационно-экстремальной интеллектуальной технологии,

основанной на максимизации информационной способности системы распознавания. Как критерий функциональной эффективности машинного обучения использовано модифицированную информационную меру Кульбака, которая рассматривается как функционал от точностных характеристик двухальтернативных решений. По построенной категорийной модели в виде отображения множеств, задействованных в процессе обучения, разработан информационно-экстремальный алгоритм оптимизации геометрических параметров гиперсферических контейнеров классов распознавания, которые восстанавливаются в процессе обучения в радиальном базисе пространства признаков. По результатам физического моделирования доказано, что фрагменты лица являются достаточно информативными при распознавании эмоционально-психического состояния человека.

**Ключевые слова:** информационно-экстремальная интеллектуальная технология, компьютеризированная система диагностирования, психодиагностика, обучения, критерий функциональной эффективности.

Dovbysh A. S.<sup>1</sup>, Shelekhov I. V.<sup>2</sup>, Prylepa D. V.<sup>3</sup>

<sup>1</sup>Dr.Sc., Professor, head of department of Computer Science, Sumy State University, Ukraine

<sup>2</sup>Ph.D., Associate Professor of department of Computer Science Sumy State University, Ukraine

<sup>3</sup>Postgraduate student of department of Computer Science Sumy State University, Ukraine

## INFORMATION-EXTREME ALGORITHM FOR SYSTEM DIAGNOSTICS EMOTIONAL AND MENTAL PERSON'S STATE LEARNING

A method of the emotional and mental person's state recognition using facial image is considered. The fragments with eye and nose areas of the image are prompted for additional information obtaining. A forming of the input mathematical description of pattern recognition system by analyzing the left hemisphere and right hemisphere images of the human face is proposed. The preliminary process of forming training matrix by image brightness for a stable emotional and mental person's state uses the whole images as well as corresponding fragments. Machine learning in the framework of information-extreme intellectual technologies is based on maximizing the information capacity of the recognition system. A criterion for the functional efficiency of machine learning uses a modified information measure Kullback as a functional of the accuracy characteristics of the two alternative solutions. An information-extreme algorithm for optimization geometrical parameters of recovering in radial basis of the feature space during the learning process hyperspherical containers of recognition classes is developed by the categorical model of mapping involved in the learning process sets. Physical modeling results proved that the fragments of facial image are quite informative for the emotional and mental person's state recognition.

**Keywords:** information-extreme intelligent technology, computer diagnostic system, psychodiagnostics, learning, the criterion of the functional efficiency.

## REFERENCES

1. Shelekhov I. V., Prylepa D. V. Optimization of the parameters of learning computerized diagnostic systems emotional-mental state of the person, *Radioelectronic and computer systems*, 2014, No. 1 (65), pp. 161–167.
2. Anuashvili A. N. Objective psychology based on the wave model of the brain. Moscow, Econ-Inform, 2008, 292 p.
3. Anisimov B. V., Kurganov V. D., Zlobin V. K. Recognition and digital image processing. Moscow, 1983, 256 p.
4. Lugger, J.F. Artificial Intelligence. Strategies and methods for solving complex problems. Moscow, St. Petersburg, Kiev, Publishing Williams, 2003, 864 p.
5. Dovbysh A. S. Foundation of intelligent systems designing: tutorial. Sumy, Publisher SSU, 2009, 171 p.
6. Shelekhov I. V., Prylepa D. V. Computerized system of diagnosing emotional and mental state of a person with fragments of facial image, *Advanced information systems and technologies Proceedings of the Third International Scientific-Practical Conference* (Sumy, 14–16 May 2014), AIST, 2014, pp. 118–119.
7. Meneghini F., Biondi P. Clinical Facial Analysis Elements, Principles, and Techniques. Berlin, Springer, 2012, 219 p.
8. George R. M. Facial Geometry: Graphic Facial Analysis for Forensic Artists, Charles C Thomas Publisher, 2007, 82 p.
9. Petta P., Pelachaud C., Cowie R. Emotion-Oriented Systems. The Humaine Handbook. Berlin, Springer Science & Business Media, 2011, 816 p.
10. Andr E. Experimental methodology in emotion-oriented computing, *IEEE Pervasive Computing*, 2011, Vol.10 (3), pp. 54–57.
11. Colmenarez A. J., Xiong Z., Huang T-S. Facial Analysis from Continuous Video with Applications to Human-Computer Interface, Berlin, Springer, 2004, 134 p.
12. Dovbysh A. S., Vasiliev A.V., Liubchak V. A. Intelligent information technologies in e-learning. Sumy, SSU, 2014, 172 p.

# ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ

# ТЕОРИЯ И МЕТОДЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

# THEORY AND METHODS OF AUTOMATIC CONTROL

УДК 681.5.01

Дорофеев Ю. И.<sup>1</sup>, Любчик Л. М.<sup>2</sup><sup>1</sup>Канд. техн. наук, доцент кафедры системного анализа и управления, Национальный технический университет «Харьковский политехнический институт», Украина<sup>2</sup>Д-р техн. наук, профессор, заведующий кафедрой компьютерной математики и математического моделирования, Национальный технический университет «Харьковский политехнический институт», Украина

## РОБАСТНОЕ ПОДАВЛЕНИЕ ВОЗМУЩЕНИЙ ПРИ УПРАВЛЕНИИ НАСОСНЫМИ СТАНЦИЯМИ В СИСТЕМЕ ЦЕНТРАЛИЗОВАННОГО ВОДОСНАБЖЕНИЯ

Предложен подход к решению задачи синтеза стабилизирующего робастного управления запасами воды в системе подачи и распределения воды крупного города. Математическая модель системы подачи и распределения воды представлена в виде нелинейной дискретной модели в пространстве состояний с запаздыванием. Предложена методика факторизации матриц модели, описывающих влияние нелинейных термов, которая позволила представить заданные структурные ограничения в виде линейных матричных неравенств. Для подавления влияния возмущений, моделирующих изменения неизвестного, но ограниченного внешнего спроса, одновременно с обеспечением устойчивости замкнутой системы, применена методика инвариантных эллипсоидов, которая позволила сформулировать задачу в терминах линейных матричных неравенств, а синтез управления свести к последовательности задач одномерной выпуклой оптимизации и полуопределенного программирования. В качестве примера рассмотрен фрагмент системы подачи и распределения воды города Харьков.

**Ключевые слова:** система подачи и распределения воды, управление запасами воды, метод инвариантных эллипсоидов, прямой метод Ляпунова, линейное матричное неравенство, задача полуопределенного программирования.

### НОМЕНКЛАТУРА

- ЛМН – линейное матричное неравенство;  
 СПРВ – система подачи и распределения воды;  
 $A_i$  – матрица динамики;  
 $B_i$  – матрицы влияния управлений;  
 $C$  – матрица выходов;  
 $D$  – допустимое множество значений внешних возмущений;  
 $D_i$  – диаметр трубы, исходящей из резервуара  $i$ ;  
 $E$  – матрица влияния внешних возмущений;  
 $R_i$  – радиус резервуара  $i$ ;  
 $U$  – допустимое множество значений управляющих воздействий;  
 $X$  – допустимое множество значений состояний;

- $d(k)$  – вектор внешних возмущений в момент времени  $k$ ;  
 $d^{\max}$  – вектор верхних граничных значений объемов потребляемой воды и утечек;  
 $d^{\min}$  – вектор нижних граничных значений объемов потребляемой воды и утечек;  
 $g$  – ускорение свободного падения;  
 $h_i(k)$  – напор (давление) воды в узле  $i$  в момент времени  $k$ ;  
 $k$  – номер дискретного момента времени;  
 $m$  – количество управляемых гидравлических элементов (насосов и вентиляй);  
 $n$  – количество резервуаров;  
 $q$  – количество секторов потребления;  
 $q_{in,i}(k)$  –  $i$ -й входящий поток воды в момент времени  $k$ ;

$q_{out,j}(k)$  –  $j$ -й исходящий поток воды в момент времени  $k$ ;

$u(k)$  – вектор управляющих воздействий в момент времени  $k$ ;

$u^{\max}$  – вектор, определяющий максимальные мощности насосов;

$v_i(k)$  – объем воды в резервуаре  $i$  в момент времени  $k$ ;

$x^{\min}$  – вектор, определяющий минимально допустимые объемы воды в резервуарах;

$x^{\max}$  – вектор, определяющий объемы резервуаров;

$z_i$  – высота расположения центра тяжести узла  $i$ ;

$\Delta t$  – период дискретизации;

$\Lambda$  – целочисленная переменная, кратная  $\Delta t$ , которая определяет максимальное значение интервала времени, необходимого для транспортировки воды между резервуарами сети;

$\alpha_K$  – коэффициент Кориолиса.

## ВВЕДЕНИЕ

Система водоснабжения представляет собой жизненно необходимую составляющую существования любого города. Управление системами водоснабжения крупных городов по мере их роста становится предметом пристального внимания специалистов в области автоматического управления. Проблема устойчивого развития и управления системой водоснабжения для города Харьков стоит особенно остро. Предпосылками являются территориальная расположность города в относительно маловодном районе, большое водопотребление населением и промышленностью, групповое водоснабжение целого ряда прилежащих к мегаполису городов, высокая изношенность инженерных коммуникаций, возросшие требования к качеству водоснабжения со стороны населения.

Современные системы водоснабжения, представляющие собой комплекс сооружений для подъема, очистки и подачи потребителям воды, с точки зрения их функционирования можно представить в виде трех относительно независимых по характеру и критериям управления подсистем [1]: источники водоснабжения и водоприемные сооружения, очистные сооружения, СПРВ.

Наиболее дорогостоящей и сложной среди указанных подсистем является СПРВ, которая имеет специфические характеристики, усложняющие ее анализ и управление:

– СПРВ имеет распределенную крупномасштабную архитектуру;

– резервуары и исполнительные элементы имеют эксплуатационные ограничения;

– для подачи воды в высотные здания необходимо обеспечить определенное давление воды, значение которого нелинейно зависит от объема перекачиваемой воды, что приводит к необходимости использования нелинейных моделей;

– транспортирование воды требует определенных временных затрат, что приводит к необходимости учета запаздываний.

Гидравлические элементы, из которых состоят СПРВ, делятся на две категории: управляемые и неуправляемые. Управляемыми являются насосные станции и вентили. Неуправляемыми элементами являются трубы, резервуары, а также узлы, которые соответствуют точкам сети, где водные потоки объединяются или разделяются. Еще одним элементом моделей СПРВ являются секторы потребления, которые позволяют моделировать объемы спроса на воду со стороны потребителей.

Для обеспечения непрерывного и эффективного функционирования СПРВ необходимо создание запасов питьевой воды. Управление запасами воды заключается в определении объемов воды, которую необходимо закачать в резервуары, и моментов времени, когда необходимо это делать. В системах управления запасами воды предусматривается периодический контроль состояния через равные промежутки времени, а момент принятия решения о пополнении запаса обычно совпадает с началом каждого периода.

В настоящее время в большинстве случаев управление СПРВ осуществляется с помощью эвристических правил и «исторически сложившихся» стратегий, которые являются результатом многолетнего эмпирического опыта. Безусловно, эти стратегии в целом являются адекватными, однако они не гарантируют оптимальной политики управления в современных крупномасштабных сетях. В связи с этим весьма актуальной является задача разработки оптимальных стратегий управления системами водоснабжения, в основе которых лежат математические модели работы сети и методы современной теории автоматического управления.

Целью настоящей работы является синтез робастной стабилизирующей стратегии управления запасами воды в системе централизованного водоснабжения, гарантирующей полное и своевременное удовлетворение неизвестных, но ограниченных объемов потребления воды при условии минимизации затрат на производство и транспортировку воды с учетом заданных эксплуатационных ограничений.

## 1 ПОСТАНОВКА ЗАДАЧИ

В работе [2] представлены различные подходы к построению математических моделей СПРВ. В настоящей работе в качестве переменных состояний модели рассматриваются объемы воды в резервуарах, измеряемые в  $m^3$ , и величины давления (напора) воды в исходящих из резервуаров потоках, измеряемые в м. Предполагается, что структура СПРВ известна, а состояния доступны непосредственному измерению. В качестве управляющих переменных выступают объемы воды, перекачиваемые через управляемые гидравлические элементы в текущем периоде, измеряемые в  $m^3$ . Внешними возмущениями являются объемы потребляемой воды и объемы утечек за период времени  $\Delta t$ , также измеряемые в  $m^3$ .

Для любого узла сети можно записать уравнение баланса массы в виде разностного уравнения, учитывающего все входящие и исходящие потоки:

$$v_i(k+1) = v_i(k) + \sum_i q_{in,i}(k) - \sum_j q_{out,j}(k), \quad i = \overline{1, n}.$$

Связь между объемом перекачиваемой воды и напором воды в потоке определяется уравнением Бернули, которое выражает закон сохранения механической энергии в жидкости и может быть записано для любых двух произвольных узлов сети  $i$  и  $j$  по ходу потока:

$$h_i(k) + \frac{8a_K\rho}{\pi^2 D_i^4} u_i(k)^2 + \rho g z_i = h_j(k) + \frac{8a_K\rho}{\pi^2 D_j^4} d_j(k)^2 + \rho g z_j, \quad (1)$$

где  $u(k) \in \mathbf{R}^m$ ;  $d(k) \in \mathbf{R}^q$ .

Напор воды в потоке, исходящем из резервуара, имеющего форму цилиндра, зависит от объема воды и радиуса резервуара:

$$h_i(k) = \frac{1}{\pi R_i^2} v_i(k). \quad (2)$$

Чтобы избавиться в уравнениях (1) от постоянных величин  $z$  введем переменные  $\bar{h}_j(k) = h_j(k) + \rho g(z_j - z_i)$  и запишем уравнения с учетом (2) в следующем виде:

$$\bar{h}_j(k) = \frac{1}{\pi R_i^2} v_i(k) + F_i u_i^2(k) + H_j d_j^2(k), \quad (3)$$

где  $F_i \in \mathbf{R}^{n \times m}$ ,  $F_i = \frac{8a_K\rho}{\pi^2 D_i^4} \cdot \text{diag}(1, 0, \dots, 0)$ , ...,  $F_m = \frac{8a_K\rho}{\pi^2 D_m^4} \cdot \text{diag}(0, \dots, 0, 1)$ ,  $H_j \in \mathbf{R}^{n \times q}$ ,  $H_1 = -\frac{8a_K\rho}{\pi^2 D_1^4} \cdot \text{diag}(1, 0, \dots, 0)$ , ...,  $H_q = -\frac{8a_K\rho}{\pi^2 D_q^4} \cdot \text{diag}(0, \dots, 0, 1)$ .

Представим вектор  $u2(k) = \{u_i^2(k)\}_{i=1}^m$ , элементами которого являются значения управляющих воздействий в квадрате, в следующем виде:

$$u2(k) = e_1^u u^T(k) \cdot E_1^u \cdot u(k) + \dots + e_m^u u^T(k) \cdot E_m^u \cdot u(k),$$

где  $e_i^u \in \mathbf{R}^{m \times 1}$ :  $e_1^u = (1, 0, \dots, 0)^T$ , ...,  $e_m^u = (0, \dots, 0, 1)^T$ ,  $E_i^u \in \mathbf{R}^{m \times m}$ :  $E_1^u = \text{diag}(1, 0, \dots, 0)$ , ...,  $E_m^u = \text{diag}(0, \dots, 0, 1)$ .

Аналогичным образом представим вектор  $d2(k) = \{d_j^2(k)\}_{j=1}^q$ , элементами которого являются значения внешних возмущений в квадрате:

$$d2(k) = e_1^d d^T(k) \cdot E_1^d \cdot d(k) + \dots + e_q^d d^T(k) \cdot E_q^d \cdot d(k),$$

где  $e_i^d \in \mathbf{R}^{q \times 1}$ :  $e_1^d = (1, 0, \dots, 0)^T$ , ...,  $e_q^d = (0, \dots, 0, 1)^T$ ,  $E_i^d \in \mathbf{R}^{q \times q}$ :  $E_1^d = \text{diag}(1, 0, \dots, 0)$ , ...,  $E_q^d = \text{diag}(0, \dots, 0, 1)$ .

Тогда нелинейные слагаемые в уравнениях (3) могут быть представлены в виде кусочно-непрерывных нелинейных функций: первая зависит от дискретного времени и вектора управлений  $h_u(k, u) : \mathbf{R}^{m+1} \rightarrow \mathbf{R}^n$ , вторая –

от дискретного времени и вектора возмущений  $h_d(k, d) : \mathbf{R}^{q+1} \rightarrow \mathbf{R}^n$ :

$$h_u(k, u) = \sum_{i=1}^m F_i \cdot u2(k) = \sum_{i=1}^m F_i \cdot \Sigma_u(k),$$

$$h_d(k, d) = \sum_{j=1}^q H_j \cdot d2(k) = \sum_{j=1}^q H_j \cdot \Sigma_d(k),$$

где  $\Sigma_u(k) = \sum_{i=1}^m e_i^u u^T(k) E_i^u u(k)$ ,  $\Sigma_d(k) = \sum_{j=1}^q e_j^d d^T(k) E_j^d d(k)$ .

Для моделирования запаздываний, связанных с транспортировкой воды от источников до резервуаров, а затем до потребителей, используется модель дискретной задержки, поскольку предполагается, что значения интервалов времени, требуемых для транспортировки воды по каждому участку сети, известны и не меняются в процессе функционирования.

Тогда математическая модель СПРВ может быть представлена в виде нелинейной дискретной модели в пространстве состояний с запаздыванием следующего вида:

$$x(k+1) = A_l x(k) + \sum_{t=0}^{\Lambda} B_t u(k-t) + Ed(k) + h_u^x(k, u) + h_d^x(k, d), \quad (4)$$

где  $x(k) = [v(k)^T, \bar{h}(k)^T]^T \in \mathbf{R}^{2n}$  – составной вектор состояний модели;  $A_l \in \mathbf{R}^{2n \times 2n}$ ;  $E \in \mathbf{R}^{2n \times q}$ ;  $B_t \in \mathbf{R}^{2n \times m}$ ,  $t = \overline{0, \Lambda}$ ;  $h_u^x : \mathbf{R}^{m+1} \rightarrow \mathbf{R}^{2n}$ ,  $h_d^x : \mathbf{R}^{q+1} \rightarrow \mathbf{R}^{2n}$ ,  $h_u^x = \begin{bmatrix} 0_{n \times m} \\ \sum_{i=1}^m F_i \end{bmatrix} \cdot \Sigma_u(k)$ ,  $h_d^x = \begin{bmatrix} 0_{n \times q} \\ \sum_{j=1}^q H_j \end{bmatrix} \cdot \Sigma_d(k)$ .

Очевидно, что структура сети определяется матрицами  $A_l$ ,  $B_t$ ,  $E$ , методика построения которых изложена в работе [3].

В процессе функционирования системы должны выполняться эксплуатационные ограничения:

$$x(k) \in X = \left\{ x \in \mathbf{R}^n : x^{\min} \leq x \leq x^{\max} \right\},$$

$$u(k) \in U = \left\{ u \in \mathbf{R}^m : 0 \leq u \leq u^{\max} \right\}, \quad (5)$$

где векторы  $x^{\min}$ ,  $x^{\max}$  и  $u^{\max}$  считаются заданными.

Предполагается, что векторы внешних возмущений удовлетворяют ограничениям:

$$d(k) \in D = \left\{ d \in \mathbf{R}^q : d^{\min} \leq d \leq d^{\max} \right\}$$

где векторы  $d^{\min}$  и  $d^{\max}$  предполагаются известными.

Преобразуем модель (4) к стандартному виду без запаздываний с помощью расширения вектора состояний [4] путем включения в него векторов предыдущих управляющих воздействий, определяющих объемы воды,

находящейся в процессе транспортировки. Вектор состояний расширенной модели будет иметь вид:

$$\xi(k) = [v(k)^T, \bar{h}(k)^T, u(k-1)^T, u(k-2)^T, \dots, u(k-\Lambda)^T]^T.$$

Тогда уравнения расширенной модели сети примут вид:

$$\begin{aligned} \xi(k+1) &= A\xi(k) + Bu(k) + Gd(k) + h_u^{ex}(k, u) + h_d^{ex}(k, d), \\ x(k) &= C\xi(k), \end{aligned} \quad (6)$$

где  $h_u^{ex}: \mathbf{R}^{m+1} \rightarrow \mathbf{R}^N$ ,  $h_d^{ex}: \mathbf{R}^{q+1} \rightarrow \mathbf{R}^N$ ,  $N = 2n + \Lambda m$ ;  $h_u^{ex}(k, u) = F\Sigma_u(k)$ ,  $h_d^{ex}(k, d) = H\Sigma_d(k)$ ; матрицы модели имеют соответствующую блочную структуру:

$$A = \begin{bmatrix} A_1 & B_1 & \dots & B_{\Lambda-1} & B_\Lambda \\ 0_{m \times 2n} & 0_{m \times m} & \dots & 0_{m \times m} & 0_{m \times m} \\ 0_{m \times 2n} & I_{m \times m} & \dots & 0_{m \times m} & 0_{m \times m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{m \times 2n} & 0_{m \times m} & \dots & I_{m \times m} & 0_{m \times m} \end{bmatrix}, \quad B = \begin{bmatrix} B_0 \\ I_{m \times m} \\ 0_{m \times m} \\ \vdots \\ 0_{m \times m} \end{bmatrix},$$

$$G = \begin{bmatrix} E \\ 0_{m \times q} \\ 0_{m \times q} \\ \vdots \\ 0_{m \times q} \end{bmatrix}, \quad C^T = \begin{bmatrix} I_{2n \times 2n} \\ 0_{m \times 2n} \\ \vdots \\ 0_{m \times 2n} \end{bmatrix}, \quad F = \begin{bmatrix} 0_{n \times m} \\ \sum_{i=1}^m F_i \\ 0_{\Lambda m \times m} \end{bmatrix}, \quad H = \begin{bmatrix} 0_{n \times q} \\ \sum_{j=1}^q H_j \\ 0_{\Lambda m \times q} \end{bmatrix}.$$

Предполагается, что пара матриц  $(A, B)$  является стабилизируемой, а нелинейные функции являются ограниченными и удовлетворяют квадратичным неравенствам [5]:

$$\begin{aligned} h_u^{ex T}(k, u)h_u^{ex}(k, u) &\leq \alpha_u^2 \Sigma_u^T(k)F^T F \Sigma_u(k) \leq \\ &\leq \alpha_u^2 u^T(k) \left( \sum_{i=1}^m e_i^u u^T(k) E_i^u \right)^T F^T F \left( \sum_{i=1}^m e_i^u u^T(k) E_i^u \right) u(k) \leq \\ &\leq \alpha_u^2 u^T(k) \left( \sum_{i=1}^m e_i^u u^{\max} E_i^u \right)^T F^T F \left( \sum_{i=1}^m e_i^u u^{\max} E_i^u \right) u(k) \leq \\ &\leq \alpha_u^2 u^T(k) \Sigma_u^{\max T} F^T F \Sigma_u^{\max} u(k), \end{aligned}$$

$$h_d^{ex T}(k, d)h_d^{ex}(k, d) \leq \alpha_d^2 d^T(k) \Sigma_d^{\max T} H^T H \Sigma_d^{\max} d(k),$$

$$\text{где } \Sigma_u^{\max} = \sum_{i=1}^m e_i^u u^{\max} E_i^u, \quad \Sigma_d^{\max} = \sum_{j=1}^q e_j^d d^{\max} E_j^d,$$

$\alpha_u > 0$ ,  $\alpha_d > 0$  – некоторые скаляры.

Для системы (6) рассматривается задача синтеза робастной относительно неопределенных, но ограниченных внешних воздействий  $d(k) \in D$ , стратегии управле-

ния запасами, обеспечивающей для любого начального состояния  $x(0) \in X$  минимизацию критерия качества, который гарантирует безопасность работы системы, а также асимптотическую робастную устойчивость замкнутой системы при условии выполнения заданных ограничений на состояния и управления (5).

## 2 ОБЗОР ЛИТЕРАТУРЫ

Выбор модели управления запасами определяется характером спроса со стороны внешних потребителей. В настоящее время для синтеза стратегии управления запасами с заданной моделью спроса широко применяется метод прогнозирующего управления [6], основанный на решении оптимизационной задачи, в которой последовательность управлений вычисляется в каждый момент времени так, чтобы минимизировать некоторый критерий качества, определяемый прогнозируемыми состояниями системы на некотором временном горизонте. После определения последовательности управляющих воздействий только первый ее элемент используется для управления в текущий момент времени. Затем измеряется новое состояние системы и процедура повторяется, т.е. реализуется принцип отступающего горизонта. Указанный подход позволяет строить управление как в программном виде, так и в виде обратной связи, а также учитывать ограничения на состояния и управления.

Однако, на практике, как правило, отсутствует информация для построения адекватной модели внешнего спроса, которая необходима для синтеза прогнозирующего управления. Одним из подходов к решению задачи управления запасами в условиях неопределенности спроса является использование концепции «неизвестных, но ограниченных» воздействий [7]. При этом соответствующая модель спроса характеризуется интервальной неопределенностью.

В последнее десятилетие сформировался новый подход к рассматриваемой проблематике, основанный на концепции инвариантных множеств. Инвариантные множества широко используются в различных задачах теории автоматического управления в динамических системах при наличии неопределенностей [8]. Среди различных форм инвариантных множеств особо выделяются эллипсоиды вследствие их простой структуры и прямой связи с квадратичными функциями Ляпунова.

В рамках метода инвариантных эллипсоидов [9] в качестве технического средства используется математический аппарат ЛМН [10]. После того, как были развиты вычислительные методы, основанные на идеях выпуклой оптимизации, и для их реализации были разработаны соответствующие алгоритмы и программное обеспечение, ЛМН стали рассматриваться в качестве общего метода анализа и синтеза линейных систем как в непрерывном, так и в дискретном случае. Однако в большинстве работ, посвященных задаче подавления ограниченных внешних возмущений, техника ЛМН применяется для подавления возмущений, ограниченных в какой-либо норме. Тогда как спецификой задач управлений запасами воды является неограниченность значений

переменных и наличие эксплуатационных ограничений, что приводит к необходимости учета несимметричных ограничений на значения внешнего спроса, состояний и управляющих воздействий.

### 3 МАТЕРИАЛЫ И МЕТОДЫ

Согласно теореме об аппроксимации произвольных выпуклых множеств, не обладающих свойством симметрии относительно начала координат [11], множество  $D$  может быть аппроксимировано эллипсоидом:

$$E(d^*, P_d) = \left\{ d \in \mathbf{R}^q : (d(k) - d^*)^\top P_d^{-1} (d(k) - d^*) \leq 1 \right\},$$

матрица которого  $P_d$  и вектор  $d^*$ , определяющий координаты центра, определяются на основании известных граничных значений объемов потребляемой воды:

$$P_d = \text{diag} \left( \frac{q^2}{4} (d_1^{\max} - d_1^{\min})^2, \dots, \frac{q^2}{4} (d_q^{\max} - d_q^{\min})^2 \right),$$

$$d^* = \frac{1}{2} (d^{\min} + d^{\max}).$$

Аналогично выполним аппроксимацию множества  $X$  эллипсоидом максимального объема

$$E(x^*, P_x) = \left\{ x \in \mathbf{R}^{2n} : (x(k) - x^*)^\top P_x^{-1} (x(k) - x^*) \leq 1 \right\}, \quad (7)$$

у которого вектор  $x^*$ , определяющий координаты центра, определим позднее, а матрица эллипса  $P_x$  вычисляется на основании векторов  $x^{\min}$  и  $x^{\max}$ :

$$P_x = \text{diag} \left( \frac{1}{4} (\min \{x_1^* - x_1^{\min}, x_1^{\max} - x_1^*\})^2, \dots, \frac{1}{4} (\min \{x_n^* - x_n^{\min}, x_n^{\max} - x_n^*\})^2 \right).$$

Построим закон управления в виде нестационарной линейной обратной связи по сигналу рассогласования между наличным и страховым уровнями запаса воды:

$$u(k) = K(k) (\xi(k) - \xi^*), \quad (8)$$

где  $K(k) \in \mathbf{R}^{m \times N}$  – нестационарная матрица коэффициентов обратной связи в момент  $k$ .

Значения элементов вектора  $\xi^*$ , определяющего уровень страховых запасов воды, вычисляются на основании средних значений объемов потребления с помощью продуктивной модели Леонтьева:

$$\xi^* = \left[ v^{*T}, 0_{1 \times n}, \underbrace{v^{*T}, \dots, v^{*T}}_{\Lambda} \right]^T,$$

$$v^* = (I - \Pi)^{-1} d^{\text{mean}},$$

$$d_j^{\text{mean}} = \begin{cases} \frac{q}{2} (d_j^{\max} + d_j^{\min}), & j = \overline{1, q}, \\ 0, & j = \overline{q+1, n}, \end{cases} \quad (9)$$

где  $\Pi$  – матрица смежности ориентированного графа, описывающего модель сети.

Значения элементов вектора  $x^*$ , определяющего координаты центра эллипса (7) будут равны

$$x^* = \left[ v^{*T}, 0_{1 \times n} \right]^T.$$

Расширенную модель замкнутой системы для управления (8) представим в виде:

$$\begin{aligned} \xi(k+1) &= A_f(k) (\xi(k) - \xi^*) + A \xi^* + G (d(k) - d^*) + \\ &\quad + G d^* + h_u^{ex}(k, u) + h_d^{ex}(k, d), \\ x(k) &= C \xi(k), \quad A_f(k) = A + BK(k). \end{aligned} \quad (10)$$

Оптимальное управление запасами воды в СПРВ сводится к задаче построения стратегии управления потоками воды от источников к потребителям с целью полного и своевременного удовлетворения спроса на воду при условии оптимизации критерии качества, обеспечивающих безопасность и устойчивость работы системы.

Для безопасного функционирования системы водоснабжения и минимизации утечек, а также поддержания высокого качества воды ее объемы в резервуарах и величины напора не должны превышать максимально возможных и не опускаться ниже минимально допустимых, то есть быть приближенными к значениям, определяющим размер страховых запасов. Для математического выражения данного критерия в виде квадратичной функции в случае бесконечного временного горизонта используется следующее выражение:

$$J_1^\infty(k) = \sum_{k=0}^{\infty} (\xi(k) - \xi^*)^\top R_\xi (\xi(k) - \xi^*)$$

где  $R_\xi \in \mathbf{R}^{N \times N}$  – положительно определенная диагональная весовая матрица, элементы которой определяют штрафы за отклонение текущих значений уровня воды от страховых.

Экономические затраты обусловлены затратами на химическую обработку воды и на электроэнергию, необходимую для перекачивания воды с помощью насосов. Критерий, учитывающий стоимость производства и транспортировки воды, может быть описан следующим выражением:

$$J_2^\infty(k) = \sum_{k=0}^{\infty} u^\top(k) R_u u(k),$$

где  $R_u \in \mathbf{R}^{m \times m}$  – положительно определенная диагональная весовая матрица.

Регулирование мощности насосов должно осуществляться плавно, чтобы избежать нежелательных переходных процессов в герметичных трубах, что может привести к их повреждению. Для обеспечения эффекта сглаживания скачков управляющих воздействий предлагается критерий следующего вида:

$$J_3^\infty(k) = \sum_{k=1}^{\infty} \Delta u^T(k) R_\Delta \Delta u(k),$$

$$\Delta u(k) = u(k) - u(k-1),$$

где  $R_\Delta \in \mathbf{R}^{m \times m}$  – положительно определенная диагональная весовая матрица.

В результате критерий качества синтезируемой системы управления будет иметь вид:

$$J^\infty(k) = J_1^\infty(k) + J_2^\infty(k) + J_3^\infty(k). \quad (11)$$

Стабилизирующие алгоритмы управления, как правило, основаны на оценивании верхнего граничного значения критерия качества с помощью функции Ляпунова. Определим квадратичную функцию Ляпунова, построенную на решениях системы (10):

$$V(\xi(k) - \xi^*) = (\xi(k) - \xi^*)^T P(k) (\xi(k) - \xi^*), \quad (12)$$

где  $P(k) = P^T(k) \in \mathbf{R}^{N \times N}$  – симметрическая положительно определенная матрица.

В соответствии с прямым методом Ляпунова вычислим первую разность функции (12) по  $k$  в силу системы (10) и потребуем выполнения следующего неравенства:

$$V(\xi(k+1) - \xi^*) - V(\xi(k) - \xi^*) \leq -J^\infty(k). \quad (13)$$

Если неравенство (13) выполняется  $\forall k \geq 0$ , то функция  $V(\xi(k) - \xi^*)$  определяет верхнее граничное значение критерия (11) [12]:

$$\max_{d(k) \in E(d^*, P_d)} J^\infty(k) \leq V(\xi(k) - \xi^*). \quad (14)$$

Управляющие воздействия  $u(k)$  будем искать из условия минимизации верхней оценки критерия (11), что позволяет упростить решение задачи. Тогда в соответствии с (14) найдем управляющие воздействия  $u(k)$  из условия минимизации функции Ляпунова:

$$u(k) = \arg \min_{u(k) \in U} V(\xi(k) - \xi^*).$$

В соответствии с методом инвариантных эллипсоидов [8] их можно рассматривать как характеристику влияния внешних возмущений на траектории динамической системы. Для оценки степени влияния возмущений  $d(k)$  на выходы  $x(k)$  удобно использовать инвариантные эллипсоиды вида:

$$E(\xi^*, P(k)) = \left\{ \xi \in \mathbf{R}^N : (\xi - \xi^*)^T P^{-1}(k) (\xi - \xi^*) \leq 1 \right\}, \quad (15)$$

которые аппроксимируют множества достижимости замкнутой системы (10) при действии возмущений  $d(k) \in E(d^*, P_d)$ . В соответствии с [12] введем матричную переменную:

$$Q(k) = \gamma(k) P^{-1}(k), \quad (16)$$

где  $\gamma(k)$  – положительный скаляр.

Тогда задача синтеза робастного стабилизирующего управления заключается в вычислении в каждый дискретный момент времени  $k$  матрицы коэффициентов обратной связи  $K(k)$ , которая стабилизирует замкнутую систему (10) и обеспечивает минимизацию по критерию следа эллипса (15). Соответствующий результат представлен следующей теоремой.

**Т е о р е м а.** Рассмотрим систему (6), замкнутую с помощью закона управления (8) при ограничениях (5), и пусть нестационарные матрицы  $P(k) = \gamma(k) Q^{-1}(k)$  и

$$K(k) = Y(k) Q^{-1}(k) \quad (17)$$

получены в результате решения оптимизационной задачи:

$$\min_{Q(k), Y(k), a_u, a_d, \tau_1, \tau_2, \tau_3} \gamma(k) \quad (18)$$

при ограничениях на матричные переменные  $Q(k) = Q^T(k) \in \mathbf{R}^{N \times N}$ ,  $Y(k) \in \mathbf{R}^{m \times N}$  и скалярные параметры  $\gamma(k)$ ,  $a_u$ ,  $a_d$ ,  $\tau_1$ ,  $\tau_2$ ,  $\tau_3$ :

$$\gamma(k) > 0, \quad \tau_1 \geq 0, \quad \tau_2 \geq 0, \quad \tau_3 \geq 0, \quad (19)$$

$$Q(k) \succ 0, \quad \begin{bmatrix} 1 \\ (\xi(k) - \xi^*)^T & Q(k) \end{bmatrix} \begin{bmatrix} (\xi(k) - \xi^*)^T \\ Q(k) \end{bmatrix} \succeq 0, \quad (20)$$

$$\left[ \begin{array}{ccccccccc} Q(k) & Y^T(k)R_A K & 0 & 0 & 0 & 0 & 0 & \Theta^T(k) & 0 \\ K^T R_A Y(k) & -\gamma(k)K^T R_A K & 0 & 0 & 0 & 0 & 0 & 0 & Q(k)R_\xi^{1/2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & (A-I)^T & Y^T(k)R_\Sigma^{1/2} \\ 0 & 0 & 0 & -\gamma(k)\tau_1 I & 0 & 0 & 0 & (A-I)^T & Y^T(k)\Sigma_u^{\max T} F^T \\ 0 & 0 & 0 & 0 & -\gamma(k)\tau_2 I & 0 & 0 & I & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I & \\ 0 & 0 & 0 & 0 & 0 & 0 & \Sigma_d^{\max T} H^T & I & \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\gamma(k)}{\tau_2 \alpha_d^2} I & 0 & 0 \\ \Theta(k) & 0 & A-I & I & I & 0 & 0 & I & 0 \\ \Theta(k) & 0 & A-I & I & I & 0 & 0 & G & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & Q(k) & \gamma(k)G P_d^{1/2} \\ R_\xi^{1/2} Q(k) & 0 & 0 & 0 & 0 & 0 & 0 & \gamma(k)P_d^{1/2} G^T & 0 \\ R_\Sigma^{1/2} Y(k) & 0 & 0 & 0 & 0 & 0 & 0 & \gamma(k)\tau_3 I & 0 \\ F \Sigma_u^{\max T} Y(k) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \gamma(k)I \\ \end{array} \right] \succeq 0, \quad (21)$$

где  $\Theta(k) = A(k)Q(k) + BY(k)$ ,  $K = K(k-1)$ ,  $R_\Sigma = R_u + R_\Delta$ ,

$$\begin{bmatrix} \gamma(k)P_x & \gamma(k)C \\ \gamma(k)C^T & Q(k) \end{bmatrix} \succeq 0, \quad (22)$$

$$Y(k)(\xi(k) - \xi^*) \succeq 0, \quad (23)$$

$$\begin{bmatrix} Y^T(k) \left( (u^{\max})^+ \right)^T (\xi(k) - \xi^*)^T & Y^T(k) \left( (u^{\max})^+ \right)^T (\xi(k) - \xi^*)^T \\ (\xi(k) - \xi^*)^T (u^{\max})^+ Y(k) & Q(k) \end{bmatrix} \succeq 0, \quad (24)$$

где «+» – псевдообращение Мура-Пенроуза.

Если задача (18)–(24) имеет решение, то система (6), замкнутая с помощью закона управления (8), для любого начального состояния  $x(0) \in X$  и неопределенного, но ограниченного, внешнего возмущения  $d(k) \in E(d^*, P_d)$  является асимптотически робастно устойчивой при ограничениях (5).

Доказательство теоремы аналогично приведенному в работе [13].

Таким образом, задача синтеза робастного стабилизирующего управления запасами воды в СПРВ, описываемой нелинейной дискретной моделью (4) в пространстве состояний, при наличии транспортных запаздываний в условиях неизвестного, но ограниченного внешнего спроса сводится к задаче минимизации линейной функции (18) при ограничениях в виде ЛМН (19)–(24), которая при фиксированных значениях скалярных параметров  $\alpha_u$ ,  $\alpha_d$ ,  $\tau_1$ ,  $\tau_2$ ,  $\tau_3$  является задачей полуопределенного программирования.

Если оптимизационная задача (18)–(24) имеет решение, то найденное в соответствии с (8), (17) управляющее воздействие применяется для управления СПРВ в текущий момент времени. Затем измеряется новое значение вектора состояний и, в соответствии с принципом отступающего горизонта, задачи полуопределенного программирования и одномерной выпуклой оптимизации

решаются в реальном времени для вычисления нового значения матрицы коэффициентов обратной связи и, соответственно, нового управляющего воздействия.

#### 4 ЕКСПЕРИМЕНТИ

В качестве примера рассмотрим фрагмент СПРВ города Харьков, изображенный на рис. 1. Он состоит из наземного источника, трех насосных станций ( $u_2, u_4, u_5$ ), двух подкачивающих насосов ( $u_1, u_3$ ), пяти резервуаров ( $v_1, \dots, v_5$ ) и четырех секторов потребления ( $d_1, \dots, d_4$ ).

В табл. 1, 2 и 3 приведены значения параметров рассматриваемого фрагмента СПРВ.

Технологическая матрица  $P$  совпадает с матрицей смежности графа  $G = (\{1, 2, 3, 4, 5\}, \{(2, 1), (4, 3), (5, 2), (5, 4)\})$ , описывающей модель сети.

Выберем период дискретизации равным 1 час. Пусть заданы значения интервалов, требуемых для перекачивания воды  $\Lambda_1 = \Lambda_2 = \Lambda_3 = \Lambda_4 = 1$ ,  $\Lambda_5 = 3$ , а также граничные значения объемов потребления и утечек воды  $d^{\min} = [90, 60, 70, 60]^T$ ,  $d^{\max} = [150, 100, 110, 120]^T$ . Моделирование осуществлялось в течение 24 периодов, что соответствует суткам. С целью учета суточных колебаний

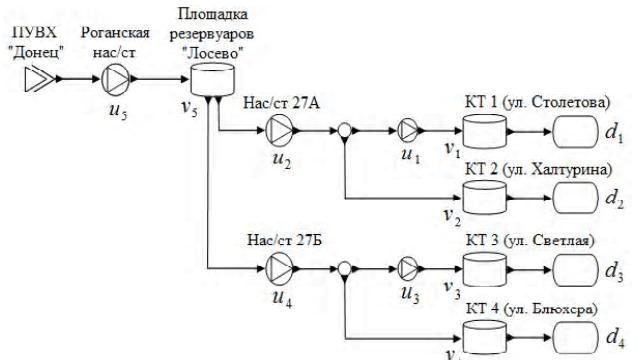


Рисунок 1 – Фрагмент системи подачі і розподілення води міста Харків

Таблица 1 – Характеристики резервуаров

Название	Переменная	Минимальный объем воды (м <sup>3</sup> )	Максимальный объем воды (м <sup>3</sup> )	Относительная высота $z$ (м)
КТ 1	$v_1$	50	500	156,40
КТ 2	$v_2$	50	500	120,98
КТ 3	$v_3$	50	500	144,91
КТ 4	$v_4$	50	500	146,97
«Люсево»	$v_5$	300	3000	180,08

Таблица 2 – Границные значения напора воды в системе

Название	Переменная	Минимальный напор воды (м)	Максимальный напор воды (м)
КТ 1	$h_1$	3,5	5,5
КТ 2	$h_2$	3,0	5,0
КТ 3	$h_3$	3,5	5,5
КТ 4	$h_4$	3,0	7,0

Таблица 3 – Характеристики насосов

Название	Переменная	Минимальная мощность (м <sup>3</sup> /час)	Максимальная мощность (м <sup>3</sup> /час)	Относительная высота $z$ (м)
Насос КТ 1	$u_1$	0	150	156,40
Нас/ст 27 А	$u_2$	0	200	160,50
Насос КТ 3	$u_3$	0	150	144,91
Нас/ст 27 Б	$u_4$	0	200	162,38
Роганская нас/ст	$u_5$	0	400	178,08

ний значения объемов потребляемой воды моделировались как нормально распределенные случайные величины с изменяющимся средним значением:

$$d_1 \rightarrow N(m_1, 7), \quad m_1 \in [60, 120], \quad d_2 \rightarrow N(m_2, 4), \\ m_2 \in [30, 70], \quad d_3 \rightarrow N(m_3, 6), \quad m_3 \in [50, 100], \\ d_4 \rightarrow N(m_4, 3), \quad m_4 \in [20, 60].$$

Значения страховых уровней запаса воды, вычисленные в соответствии с (9), равны  $v^* = [364,588,276,446,1034]^T$ . Выберем начальные условия  $v(0) = [400, 500, 400, 500, 2000]^T$  и значения элементов диагональных весовых матриц  $r_\xi = 1,5 \cdot 10^{-7}$ ,  $r_u = 1,0 \cdot 10^{-5}$ ,  $r_\Delta = 1,5 \cdot 10^{-6}$ .

## 5 РЕЗУЛЬТАТЫ

Моделирование выполнено в среде MATLAB. Численное решение задачи (18)–(24) получено с помощью свободно распространяемого пакета CVX [14]. Графики переходных процессов для узла 1 при  $\alpha_u = \alpha_d = 0,15$ ,  $\tau_1 = \tau_2 = \tau_3 = 5,0 \cdot 10^{-3}$  и стохастическом внешнем спросе приведены на рис. 2. На рис. 3 представлена фазовая траектория в пространстве состояний, определенном для резервуаров 1, 2 и 3.

## 6 ОБСУЖДЕНИЕ

Значения весовых матриц  $R_\xi$ ,  $R_u$ ,  $R_\Delta$  были выбраны из условия разрешимости оптимизационной задачи (18)–(24) и определяют размер эллипсоида, аппроксимирующего множество достижимости замкнутой системы (10) при действии возмущений  $d(k) \in E(d^*, P_d)$ .

Анализ полученных результатов позволяет сделать вывод об эффективности предложенного подхода к решению задачи синтеза робастной стабилизирующей стратегии управления запасами воды в СПРВ, поскольку значения объемов воды в резервуарах, а также напора воды в трубах в условиях стохастически изменяющихся объемов потребления воды находятся в пределах заданных ограничений, а объемы перекачиваемой воды не превышают максимальных возможностей насосов.

## ВЫВОДЫ

В работе предложен подход к решению задачи построения робастного стабилизирующего управления запасами воды в централизованной системе подачи и распределения воды. Основными особенностями рассмотренной задачи являются: нелинейность математической модели; неопределенность, но ограниченность внешнего спроса на воду; наличие запаздываний, связанных с транспортировкой воды; наличие эксплуатационных ограничений на значения состояний и управлений.

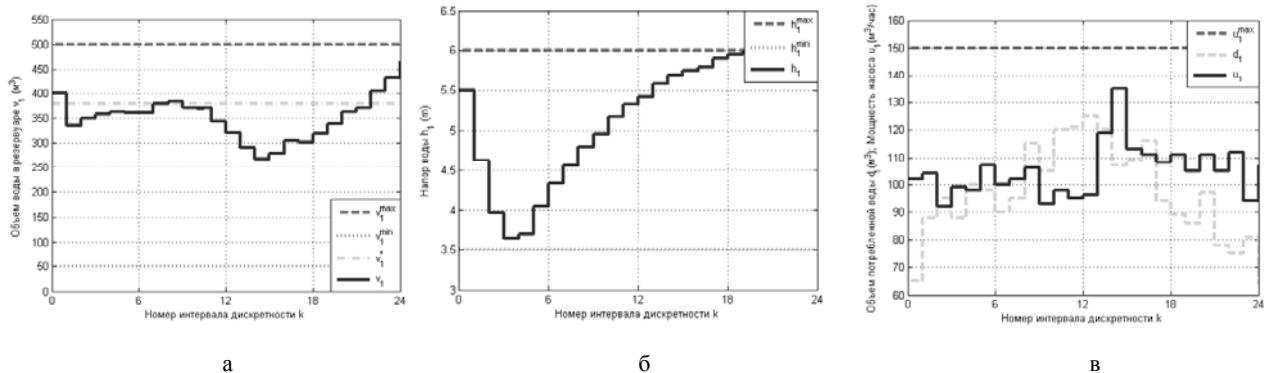


Рисунок 2 – Графики переходних процесів в системі подачі и распределения воды:  
а – значення об'єму води в резервуаре 1; б – значення напору води в потоці, юходящем из резервуара 1;  
в – значення об'єму потребленої води в секторе 1 и об'єму води, перекачаної насосом 1

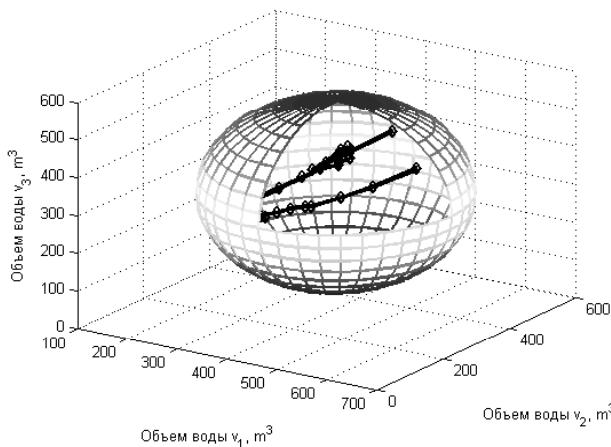


Рисунок 3 – Фазова траєкторія и інваріантний елліпсоїд, отриманий на останньому кроці, для резервуарів 1, 2 і 3

Для падавлення впливу возмущень, моделюючих змінення зовнішнього спроса, одночасно з обезпеченням устойчивості замкнутої системи, применена методика інваріантних елліпсоїдів, яка позволяла сформулювати задачу в термінах лінійних матричних неравенств, а синтез управління свести до послідовності задач одномерної випуклої оптимізації та полуопределеної програмування.

Полученное управление зависит от выбранного значения уровня страховых запасов воды. В рамках предложенного подхода возможен выбор оптимальных значений страховых запасов, поскольку полученное решение задачи синтеза робастного управления задает, фактически, алгоритмическую зависимость между уровнем страховых запасов и оптимальным значением критерия качества.

## СПИСОК ЛИТЕРАТУРЫ

- Петров В. А. Управление региональными системами водоснабжения / В. А. Петров. – Харьков : Основа, 1999. – 320 с.
- Brdys M. A. Operational control of water systems: structures, algorithms, and applications / M. A. Brdys, B. Ulanicki. – London: Prentice Hall, 1994. – 364 p.
- Дорофеев Ю. И. Построение математических моделей управляемых сетей поставок с учетом запаздываний потоков / Ю. И. Дорофеев, А. А. Никульченко // Системні
- Blanchini F. Feedback control of production-distribution systems with unknown demand and delays / F. Blanchini, R. Pesenti, F. Rinaldi and W. Ukovich // IEEE Transaction on Robotics and Automation, Special Issue on Automation of Manufacturing Systems. – 2000. – Vol. RA-16. – No. 3. – P. 313–317.
- Stipanovic D. M. Robust stability and stabilization of discrete-time non-linear systems: the LMI approach / D. M. Stipanovic, D. D. Siljak // International Journal of Control. – 2001. – Vol. 74. – P. 873–979.
- Bemporad A. Robust model predictive control: a survey / A. Bemporad, M. Morari // Lecture Notes in Control and Information Sciences. – 1999. – Vol. 245. – P. 207–226.
- Bertsekas D. P. Recursive state estimation for a set-membership description of uncertainty / D. P. Bertsekas, I. Rhodes // IEEE Trans. Automat. Control. – 1971. – Vol. 16. – P. 117–128.
- Blanchini F. Set theoretic methods in control / F. Blanchini, S. Miani. – Boston : Birkhauser, 2007. – 504 p.
- Хлебников М. В. Оптимизация линейных систем при ограниченных внешних возмущениях (техника инвариантных эллипсоидов) / М. В. Хлебников, Б. Т. Поляк, В. М. Кунцевич // Автоматика и телемеханика. – 2011. – № 11. – С. 9–59.
- Linear matrix inequalities in system and control theory / [S. Boyd, El. Ghaoui, E. Feron, V. Balakrishnan]. – Philadelphia: SIAM, 1994. – 187 p.
- Черноуско Ф. Л. Оцінювання фазового состояния динаміческих систем. Метод елліпсоїдів / Ф. Л. Черноуско. – М. : Наука, Главная редакция физ.-мат. литературы, 1988. – 320 с.
- Kothare M. V. Robust constrained model predictive control using linear matrix inequalities / M. V. Kothare, V. Balakrishnan, M. Morari // Automatica. – 1996. – Vol. 32 № 10. – P. 1361–1379.
- Lyubchyk L. M. Robust model predictive control of constrained supply networks via invariant ellipsoids technique / L. M. Lyubchyk, Y. I. Dorofieiev, A. A. Nikulchenko // Proc. IFAC Conf. on Manufacturing Modelling, Management and Control MIM’2013. – 2013. – P. 1618–1623.
- Grant M. CVX: Matlab software for disciplined convex programming, version 2.0 beta [Electronic resource] / M. Grant, S. Boyd // Acces mode: <http://cvxr.com/cvx>.

Статья поступила в редакцию 09.09.2014.  
После доработки 23.09.2014.

Дорофеєв Ю. І.<sup>1</sup>, Любчик Л. М.<sup>2</sup>

<sup>1</sup>Канд. техн. наук, доцент, Національний технічний університет «Харківський політехнічний інститут», Україна

<sup>2</sup>Д-р техн. наук, професор, Національний технічний університет «Харківський політехнічний інститут», Україна

## РОБАСТНЕ ПРИДУШЕННЯ ЗБУРЕНЬ ПРИ УПРАВЛІННІ НАСОСНИМИ СТАЦІЯМИ В СИСТЕМІ ЦЕНТРАЛІЗОВАНОГО ВОДОПОСТАЧАННЯ

Запропоновано підхід до вирішення задачі побудови стабілізуючого робастного управління запасами води в системі подачі і розподілу води великого міста. Математична модель системи подачі і розподілу води подана у вигляді нелінійної дискретної моделі в просторі станів із запізненням. Запропоновано методику факторизації матриць моделі, що описують вплив нелінійних термів, яка дозволила подати задані структурні обмеження у вигляді лінійних матричних нерівностей. Для придушення впливу збурень, що моделюють зміни невідомого, але обмеженого зовнішнього попиту, одночасно із забезпеченням стійкості замкнутої системи, застосована методика інваріантних еліпсоїдів, яка дозволила сформулювати задачу в термінах лінійних матричних нерівностей, а синтез управління звести до послідовності задач одновимірної опуклою оптимізації та напіввизначеного програмування. Як приклад розглянуто фрагмент системи водопостачання міста Харків.

**Ключові слова:** система подачі і розподілу води, управління запасами води, метод інваріантних еліпсоїдів, прямий метод Ляпунова, лінійна матрична нерівність, задача напіввизначеного програмування.

Dorofieiev Yu. I.<sup>1</sup>, Lyubchik L. M.<sup>2</sup>

<sup>1</sup>PhD, Associate Professor, National Technical University «Kharkiv Polytechnic Institute», Ukraine

<sup>2</sup>Dr. Sc., Professor, National Technical University «Kharkiv Polytechnic Institute», Ukraine

## ROBUST DISTURBANCES REJECTION IN PUMPING STATIONS CONTROL FOR CENTRALIZED DRINKING WATER DISTRIBUTION SYSTEM

An approach to solving the problem of stabilizing robust water inventory control synthesis for drinking water distribution system of a large city is proposed. The mathematical model of drinking water distribution system is presented in the form of a nonlinear discrete state-space model with time-delay. The technique of model matrix factorization describing the influence of nonlinear terms, which allowed to introduce structural constraints in the form of linear matrix inequalities, is proposed. To suppress the disturbances influence simulating unknown but bounded external demand, while ensuring robust stability of the closed-loop system, is used the invariant ellipsoids technique, which allowed to formulate the control problem in terms of linear matrix inequalities. As a result the control synthesis problem is reduced to a sequence of one-dimensional convex optimization problems and semi-definite programming. As an example, a fragment of the Kharkiv drinking water distribution system is consider.

**Keywords:** drinking water distribution system, water inventory control, invariant ellipsoids method, Lyapunov's direct method, linear matrix inequality, semi-definite programming.

## REFERENCES

1. Petrosov V. A. Upravlenie regional'nyimi sistemami vodosnabzhenija. Khar'kov, Osnova, 1999, 320 p.
2. Brdys M. A., Ulanicki B. Operational control of water systems: structures, algorithms, and applications. London, Prentice Hall, 1994, 364 p.
3. Dorofieiev Yu. I., Nikulchenko A. A. Postroenie matematicheskikh modelej upravliaemih setej postavok s uchetom zapazdivaniy potokov, *System Research & Information Technologies*, 2013, No. 1, pp. 16–27.
4. Blanchini F., Pesenti R., Rinaldi F., Ukovich W. Feedback control of production-distribution systems with unknown demand and delays, *IEEE Transaction on Robotics and Automation, Special Issue on Automation of Manufacturing Systems*, 2000, Vol. RA-16, No. 3, pp. 313–317.
5. Stipanovic D. M., Siljak D. D. Robust stability and stabilization of discrete-time non-linear systems: the LMI approach, *International Journal of Control*, 2001, Vol. 74, pp. 873–979.
6. Bemporad A., Morari M. Robust model predictive control: a survey, *Lecture Notes in Control and Information Sciences*, 1999, Vol. 245, pp. 207–226.
7. Bertsekas D. P., Rhodes I. Recursive state estimation for a set-membership description of uncertainty, *IEEE Trans. Automat. Control*, 1971, Vol. 16, pp. 117–128.
8. Blanchini F., Miani S. Set theoretic methods in control. Boston, Birkhauser, 2007, 504 p.
9. Hlebnikov M. V., Poljak B. T., Kuncevich V. M. Optimizacija linejnih sistem pri ogranicennih vnesnih vozmuschenijah (tehnika invariantnih ellipsoidov), *Automation and Remote Control*, 2011, № 11, pp. 9–59.
10. Boyd S., Ghaoui El., Feron E., Balakrishnan V. Linear matrix inequalities in system and control theory, Philadelphia, SIAM, 1994, 187 p.
11. Chernous'ko F. L. Ocenivanie fazovogo sostojaniya dinamicheskikh sistem. Metod ellipsoidov. Moscow, Nauka, 1988, 320 p.
12. Kothare M. V., Balakrishnan V., Morari M. Robust constrained model predictive control using linear matrix inequalities, *Automatica*, 1996, Vol. 32(10), pp. 1361–1379.
13. Lyubchik L. M., Dorofieiev Y. I., Nikulchenko A. A. Robust model predictive control of constrained supply networks via invariant ellipsoids technique, Proc. IFAC Conf. on Manufacturing Modelling, Management and Control MIM'2013, 2013, pp. 1618–1623.
14. Grant M., Boyd S. CVX: Matlab software for disciplined convex programming, version 2.0 beta. Acces mode: <http://cvxr.com/cvx>.

# УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

## УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

### CONTROL IN TECHNICAL SYSTEMS

---

---

УДК 681.514

Гурко А. Г.<sup>1</sup>, Янчевский И. В.<sup>2</sup>

<sup>1</sup>Канд. техн. наук, доцент, доцент кафедры автоматизации и компьютерно-интегрированных технологий Харьковского национального автомобильно-дорожного университета, Украина

<sup>2</sup>Д-р физ.-мат. наук, доцент, профессор кафедры теории машин и механизмов, Харьковский национальный автомобильно-дорожный университет, Украина

#### ГАРАНТИРОВАННОЕ УПРАВЛЕНИЕ ДВИЖЕНИЕМ МАНИПУЛЯЦИОННОГО РОБОТА

---

Рассматривается задача управления движением манипуляционного робота при наличии в его модели параметрической неопределенности и действии неконтролируемых возмущений. Предложен робастный регулятор, основанный на методе вычисляемого момента, в котором для вычисления дополнительного управления, обеспечивающего компенсацию действия неопределенных факторов, используется гарантированный подход, согласно которому значения этих факторов известны лишь с точностью до принадлежности некоторым множествам. Для оптимизации управления использован квадратичный критерий качества. Эффективность предложенного решения подтверждено результатами численного моделирования.

**Ключевые слова:** манипуляционный робот, регулятор с вычисляемым моментом, гарантированное управление, множество возможных состояний.

#### НОМЕНКЛАТУРА

РВМ – регулятор с вычисляемым моментом;

$\mathbb{R}$  – множество действительных чисел;

$n$  – количество звеньев манипуляционного робота;

$p$  – количество измеряемых выходных координат системы;

$K$  – момент квантования;

$N$  – количество моментов квантования;

$E$  – единичная матрица;

$\mathbf{q}^d, \dot{\mathbf{q}}^d, \ddot{\mathbf{q}}^d$  –  $(n \times 1)$ -векторы желаемых обобщенных координат, их скоростей и ускорений;

$\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}$  –  $(n \times 1)$ -векторы фактических обобщенных координат, их скоростей и ускорений;

$D(\mathbf{q})$  – симметричная положительно определенная  $(n \times n)$ -матрица инерции робота;

$F(\mathbf{q}, \dot{\mathbf{q}})$  –  $(n \times 1)$ -вектор, включающий все центробежные и кориолисовы компоненты, силы трения и силы гравитации;

$\mathbf{M}$  –  $(n \times 1)$ -вектор действующих на звенья робота управляющих моментов;

$\mathbf{M}_L$  –  $(n \times 1)$ -вектор действующих на звенья робота возмущающих моментов;

$e, \dot{e}, \ddot{e}$  – ошибки слежения по положению, по скорости и по ускорению, соответственно;

$K_p, K_v$  –  $(n \times n)$ -матрицы коэффициентов обратных связей по положению и по скорости;

$\eta$  – неопределенность;

$(\hat{\cdot})$  – оценка параметра;

$(\tilde{\cdot})$  – разность между действительным значением параметра и его оценкой;

$a \in \mathbb{R}^n$  – основное управляющее воздействие;

$\delta a \in \mathbb{R}^n$  – дополнительное управляющее воздействие;

$x \in \mathbb{R}^{2n}$  – вектор координат состояния;

$y \in \mathbb{R}^p$  – вектор результатов измерений;

$v$  – вектор шумов измерений;

$g, \omega$  – заданные функции;

**V** – функция Ляпунова;  
**J** – целевой функционал;  
 $\Omega$  – информационное множество;  
 $\partial\Omega$  – совокупность граничных элементов множества  $\Omega$ .

## ВВЕДЕНИЕ

Манипуляционные роботы с быстродействующими и высокоточными системами управления движением крайне востребованы при автоматизации различных технологических операций в самых разных отраслях промышленности, в медицине и в научных исследованиях. Как правило, управление движением таких роботов осуществляется в условиях неопределенности. Эта неопределенность может быть связана с упрощениями, принимаемыми при построении моделей манипуляционных роботов, неполнотой информации о значениях таких параметров как массы и моменты инерции звеньев, силы трения в сочленениях, возможные внешние возмущения и шумы измерений. Кроме того, некоторые характеристики, например, масса объекта, с которым оперирует манипулятор, могут значительно изменяться в процессе его функционирования. Указанные неопределенности оказывают существенное влияние на эффективность работы системы управления, и даже стать причиной потери системой устойчивости.

Для обеспечения инвариантности системы управления роботом по отношению и имеющимся факторам неопределенности разрабатываются различные регуляторы [1–19], в основе которых лежат разнообразные подходы. Однако, большинство из них не нашло практического применения в силу различных недостатков, таких как, например, сложность реализации регулятора, большое количество итераций при расчете управления, необходимость измерения всех координат вектора состояния и т. д.

Целью данной работы является разработка эффективного и относительно простого с вычислительной точки зрения регулятора для управления движением манипуляционного робота в условиях неопределенности.

## 1 ПОСТАНОВКА ЗАДАЧИ

Представим динамическую модель  $n$ -звенного манипулятора в виде [13]:

$$\mathbf{D}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{F}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{M}. \quad (1)$$

В классической постановке задачи управления манипулятором широко используется линеаризация обратной связи [12, 13], в результате которой получают следующий закон управления:

$$\mathbf{M} = \mathbf{D}(\mathbf{q})\mathbf{a} + \mathbf{F}(\mathbf{q}, \dot{\mathbf{q}}). \quad (2)$$

В этом уравнении, при использовании РВМ, значение **a** определяется выражением:

$$\mathbf{a} = \ddot{\mathbf{q}}^d + \mathbf{K}_v \dot{\mathbf{e}} + \mathbf{K}_p \mathbf{e}, \quad (3)$$

где матрицы  $\mathbf{K}_v$  и  $\mathbf{K}_p$  подлежат определению.

Сопоставив уравнения (1)–(3) получаем линейное уравнение для ошибки  $\dot{\mathbf{e}} + \mathbf{K}_v \dot{\mathbf{e}} + \mathbf{K}_p \mathbf{e} = 0$ , которое устойчиво при положительно определенных матрицах  $\mathbf{K}_v$  и  $\mathbf{K}_p$ .

Так как при синтезе управления значения входящих в (2) параметров известны приближенно, то приходится пользоваться их оценками:

$$\mathbf{M} = \hat{\mathbf{D}}(\mathbf{q})\mathbf{a} + \hat{\mathbf{F}}(\mathbf{q}, \dot{\mathbf{q}}), \quad (4)$$

На основании (4) и (1) следует:

$$\ddot{\mathbf{q}} = \mathbf{a} - \boldsymbol{\eta}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{a}), \quad (5)$$

где  $\boldsymbol{\eta} = (\mathbf{E} - \mathbf{D}^{-1}(\mathbf{q})\hat{\mathbf{D}}(\mathbf{q}))\mathbf{a} - \mathbf{D}^{-1}(\mathbf{q})\tilde{\mathbf{F}}(\mathbf{q}, \dot{\mathbf{q}})$ . В дальнейшем для простоты записи функциональные зависимости будем опускать. Тогда, с учетом (3) и (5), получим

$$\ddot{\mathbf{e}} = \ddot{\mathbf{q}}^d - \mathbf{a} + \boldsymbol{\eta}. \quad (6)$$

Таким образом, задача управления движением манипуляционного робота по заданной траектории сводится к поиску закона управления **a**, который бы компенсировал ошибку в нелинейной нестационарной системе (6).

## 2 ОБЗОР ЛИТЕРАТУРЫ

Для борьбы с неопределенностями **η** разрабатываются различные схемы адаптивных и робастных регуляторов [1–9], среди которых эффективным является адаптивный регулятор со скользящим режимом [2, 5, 6]. Обзоры существующих решений по применению такого типа регулятора к управлению манипуляторами можно найти в [6, 9, 10]. Вместе с тем следует отметить, что недостатком, присущим таким регуляторам, является необходимость борьбы с дребезгом фазовой траектории вдоль линии переключения [10, 11].

Распространенным подходом к управлению манипуляционными роботами при отсутствии неопределенностей является использование РВМ [12, 13]. Достоинства РВМ: потенциально высокая точность слежения за заданной траекторией, малые коэффициенты обратной связи, малые затраты энергии [1]. Однако на эффективность применения РВМ для управления манипуляционными роботами существенно влияют неопределенностии, поэтому исследования по совершенствованию РВМ с целью обеспечения его робастности по отношению к неопределенностям являются актуальными. Среди публикаций последних лет следует отметить [1, 9–19]. В частности, в [12] описан регулятор с компенсатором, в закон управления которого добавлены динамические фильтры ошибок по положению и по скорости. В книге [13] приведены РВМ, в которых вместо точных значений параметров математической модели робота используются их оценки. Это позволяет уменьшить объем вычислений в режиме реального времени. Компенсация же ошибок управления, вызванных отклонениями оценочных значений математической модели от номинальных, осуществляется путем вычисления дополнительных управляющих воздействий.

В [14] рассмотрено два непараметрических регрессионных метода для уточнения параметров модели робота в реальном времени. Результаты этой оценки ис-

пользовались для управления манипулятором при помощи РВМ. В результате реализована адаптивная система управления. В [15] описан метод робастного управления манипулятором, основанный на РВМ и прямом методе Ляпунова. В работе [11] описан модифицированный РВМ, который, фактически, является вариантом регулятора со скользящим режимом.

В [10] предложены две адаптивные схемы управления манипуляционным роботом, которые сочетают в себе РВМ и компенсаторы неопределенностей, использующие аппарат нечеткой логики. Комбинации РВМ с регуляторами на нечеткой логике для оценки и компенсации неопределенностей описаны также в [9, 11, 16, 17]. Ф. Пилтан с соавторами использовал математические средства нечеткой логики для настроек коэффициентов РВМ [18, 19].

Все приведенные системы помимо очевидных достоинств обладают и некоторыми недостатками, например, вычислительной сложностью, зависимостью от субъективных мнений экспертов [9] и проч.

При управлении динамическими системами в условиях неопределенности развивается так называемый игровой подход, при котором предполагается, что неопределенные факторы известны с точностью до принадлежности заданным информационным множествам. В процессе управления строятся новые множества, представляющие собой гарантированную оценку состояния системы, а управляющее воздействие отыскивается из условия минимизации некоторого функционала, тогда как неопределенные факторы стремятся его максимизировать [20–24]. Полученное управляющее воздействие гарантирует достижение цели управления при любых допустимых реализациях факторов неопределенности. Представляется перспективным разработка системы управления движением манипуляционным роботом на основе указанного подхода.

### 3 МАТЕРИАЛЫ И МЕТОДЫ

При синтезе гарантированного управления в качестве информационных множеств рассматриваются или выпуклые многогранники, или аппроксимирующие их с целью упрощения вычислений эллипсоиды. Однако в последнем случае появляется дополнительная погрешность аппроксимации. В [23] авторами настоящей работы предложено применение математических средств теории R-функций, особенностью которых является относительная простота и гибкость с точки зрения автоматизации программирования, для построения эволюционирующих во времени информационных множеств в виде многогранников. В [24] приведена процедура синтеза управления, основанная на введенном понятии дна функции Ляпунова, позволяющая существенно минимизировать объем вычислений при определении оптимального управления динамическим объектом в условиях неопределенности. В настоящей работе представлено развитие полученных в [23, 24] результатов на новый класс объектов, в частности, на системы управления движением манипуляционных роботов.

Рассмотрим дискретную систему управления движением манипуляционного робота. Пусть система подвержена действию неопределенных внешних возмущений  $\eta_k$ , в качестве априорной информации относительно которых известна принадлежность некоторым компактным множествам  $\Omega_k^{\Pi} (\eta_k \in \Omega_k^{\Pi}, k=0,1,\dots,N)$ . Каждая  $\eta_k$  может принимать любое значение из соответствующего множества  $\Omega_k^{\Pi}$  с равной вероятностью.

Для системы (6) найдем закон управления  $a_k$ , гарантирующий ее асимптотическую устойчивость и обеспечивающий оптимальный в смысле выбранного критерия качества результат слежения за заданной траекторией для любого значения  $\eta_k$ .

Примем

$$\dot{a}_k = \ddot{q}_k^d + K_v \dot{e}_k + K_p e_k + \delta a_k. \quad (7)$$

Введем вектор состояния:

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{e}_k \\ \dot{\mathbf{e}}_k \end{bmatrix}.$$

Тогда ошибка движения замкнутой системы может быть записана в виде линейного разностного уравнения:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}_k\{\eta_k - \delta a_k\}, \quad k=0,1,\dots,N-1, \quad (8)$$

где  $\mathbf{A}$  и  $\mathbf{B}$  – матрицы размерности  $(2n \times 2n)$  и  $(2n \times n)$ , соответственно:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{E} \\ -K_p & -K_v \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{E} \end{bmatrix}.$$

В результате измерения вектора  $\mathbf{x}_k$  формируется вектор результатов измерений  $\mathbf{y}_k$ :

$$\mathbf{y}_k = \mathbf{g}_k(\mathbf{x}_k, \mathbf{v}_k), \quad k=1,2,\dots,N-1. \quad (9)$$

Целью управления примем минимизацию функционала вида:

$$J_k(\mathbf{x}_k, \delta a_k) = V_k(\mathbf{x}_{k+1}) + \omega_k(\mathbf{x}_k, \delta a_k), \quad (10)$$

где заданная функция  $\omega_k$  определяет, например, затраты на реализацию управления.

Для корректной постановки задачи минимизации функционала (10) необходимо доопределить информацию о неопределенностях  $\eta_k$ . Так как значения  $\eta_k$  могут с равной вероятностью принимать любое значение из множества  $\Omega_k^{\Pi}$ , то приходится рассматривать такие значения, которые максимизируют критерий (10).

Ограниченнность информации о  $\eta_k$  и  $\mathbf{v}_k$  в виде соответствующих множеств приводит к тому, что в результате измерения (9) выходных координат системы, информация о значении текущих координат состояния робота получается в виде множественной оценки  $\mathbf{x}_k \in \Omega_k^r$ . Для расчета управления необходимо получение точечной оценки  $\mathbf{x}_{k+1}^u \in \Omega_{k,k+1}^w$  состояния системы из информационного множества возможных состояний  $\Omega_{k,k+1}^w$  использу-

зуют точку, которая либо обеспечивает максимум некоторому функционалу качества [20], либо удовлетворяет условию минимума ошибки в наихудшем случае (чебышевский центр). Ориентация на значение  $\mathbf{x}^u$ , максимизирующее функционал, в частном случае (10), предпочтительнее, поскольку такой подход позволяет учитывать не только минимум ошибки, но и затраты ресурсов на реализацию управления, а также ограничения на их значения. Таким образом, задача синтеза системы управления сводится к известной из теории дифференциальных игр минимаксной задаче:

$$\min_{\delta \mathbf{a}_k \in \Omega_k^u} \max_{\eta_k \in \Omega_k^\eta} \max_{v_k \in \Omega_k^v} \max_{x_k \in \Omega_k^r} \mathbf{J}_k(\mathbf{x}_k, \delta \mathbf{a}_k). \quad (11)$$

Очевидно, что решение задачи (11) гарантирует определенный результат, зависящий от вида критерия  $\mathbf{J}_k$ , при любых допустимых реализациях факторов неопределенности.

Описание множества возможных состояний системы управления манипуляционным роботом будем производить в соответствии со следующим алгоритмом.

1. Пусть в произвольный момент квантования  $k$  имеется оценка состояния системы в виде  $\mathbf{x}_k \in \Omega_k^r$ . Производится преобразование:

$$\Omega_{k,k+1}^f = A\Omega_k^r, \quad (12)$$

где  $\Omega_{k,k+1}^f$  множество, представляющее собой прогноз возможных состояний  $\mathbf{x}_{k+1}^f \in \Omega_{k,k+1}^f$  системы в  $[k+1]$ -й момент, в которые она должна перейти в свободном движении между  $k$ -м и  $[k+1]$ -м моментами квантования из состояния  $\mathbf{x}_k \in \Omega_k^r$ .

2. Строится новое множество  $\Omega_{k,k+1}^w$  с учетом влияния внешних возмущений  $\eta_k$  на элементы множества  $\Omega_{k,k+1}^f$  путем трансформации (размытия) множества  $\Omega_{k,k+1}^f$  при помощи преобразования:

$$\Omega_{k,k+1}^w = \Omega_{k,k+1}^f \bigcup \mathbf{B}_k \partial \Omega_k^\eta. \quad (13)$$

Множество  $\Omega_{k,k+1}^w$ , таким образом, представляет собой прогноз возможных состояний системы в  $[k+1]$ -й момент под влиянием неопределенности  $\eta_k$ .

3. Находится значение  $\mathbf{x}_{k+1}^u \in \Omega_{k,k+1}^w$  состояния системы, которое используется для расчета управляющего воздействия  $\delta \mathbf{a}_k$ , минимизирующего функционал качества (11).

4. Производится перемещение множества  $\Omega_{k,k+1}^w$  найденным управлением  $\delta \mathbf{a}_k$ . При этом формируется новое информационное множество  $\Omega_{k+1}^u$ , представляющее собой априорную оценку состояния системы, в которые она перейдет к моменту  $[k+1]$  после применения управления  $\delta \mathbf{a}_k$  и условиях наличия неопределенности  $\eta_k$ .

5. Осуществляется новое измерение выходных координат системы с целью получения апостериорной оценки  $\mathbf{x}_{k+1} \in \Omega_{k+1}^r$  состояния системы в момент квантования  $[k+1]$ , где

$$\Omega_{k+1}^r = \Omega_{k+1}^u \bigcap \Omega_{k+1}^v. \quad (14)$$

Далее приведенная выше процедура итерационно повторяется. Структурная схема изложенной системы управления манипуляционным роботом приведена на рис. 1.

#### 4 ЭКСПЕРИМЕНТЫ

Для проверки эффективности предложенного решения выполнено моделирование системы управления движением двухзвенным плоским манипуляционным роботом, динамика которого описывается уравнением (1) со следующими параметрами [17]:  $\mathbf{F}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{G}(\mathbf{q})$ ,

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}; \quad \mathbf{D}(\mathbf{q}) = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix},$$

$$\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} -C_{12}\dot{q}_2 & -C_{12}(\dot{q}_1 + \dot{q}_2) \\ C_{12}\dot{q}_1 & 0 \end{bmatrix}, \quad \mathbf{G}(\mathbf{q}) = \begin{bmatrix} G_1 g \\ G_2 g \end{bmatrix},$$

$$\mathbf{M} = \begin{bmatrix} M_1 - M_{L1} \\ M_2 - M_{L2} \end{bmatrix},$$

где  $D_{11} = (m_1 + m_2)l_1^2 + m_2l_2^2 + 2m_2l_1l_2 \cos q_2$ ,  $D_{12} = D_{21} = m_2l_2^2 + m_2l_1l_2 \cos q_2$ ;  $D_{22} = m_2l_2^2$ ,  $C_{12} = m_2l_1l_2 \sin q_2$ ;  $G_1 = (m_1 + m_2)l_1 \cos q_2 + m_2l_2 \cos(q_1 + q_2)$ ,  $G_2 = m_2l_2 \cos(q_1 + q_2)$ ;  $g = 9,81 \text{ м/с}^2$ ;  $M_{Li}$  ( $i=1,2$ ) – внешние возмущения.

Параметры звеньев манипуляционного робота приведены в табл. 1.

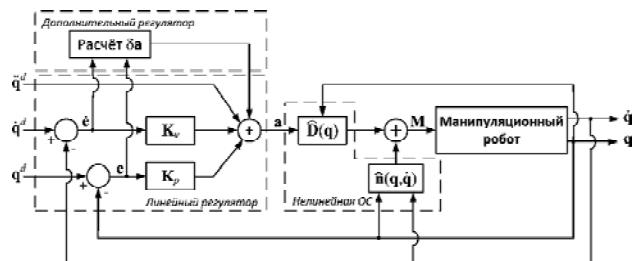


Рисунок 1 – Структурная схема системы управления манипуляционным роботом

Таблица 1 – Параметры звеньев манипуляционного робота и внешнее возмущение

№ звена, $i$	Масса, $m_i$ кг	Длина, $l_i$ м	Внешние возмущения, $M_{Li}$ Нм
1	2,74	0,115	$\sin(2t)$
2	2,01	0,130	$2\cos(4t)$

Желаемые траектории  $q_i^d$   $i=1,2$  движения звеньев описывается выражением:

$$q_i^d = b_i(1-e^{-2t^3}) + c_i(1-e^{-2t^3})\sin(\omega_i t), \text{рад}, \quad (16)$$

где  $b_1=\pi/4$  рад,  $b_2=\pi/3$  рад,  $c_1=\pi/9$  рад,  $c_2=\pi/6$  рад,  $\omega_1=4$  рад/с и  $\omega_2=3$  рад/с.

Движение начиналось при начальных условиях  $q_1(0)=0,25$  рад,  $q_2(0)=-0,3$  рад,  $\dot{q}_1(0)=\ddot{q}_1(0)=0$  ( $i=1,2$ ). В процессе моделирования измерялись только значения обобщенных координат  $q_i$  ( $i=1,2$ ) с неизвестным, но ограниченным по значению, случайным шумом  $|v_i| \leq 0,00872$  рад, скорости изменения обобщенных координат не измерялись. На звенья робота действовали возмущения в виде внешних моментов  $M_{Li}$  ( $i=1,2$ ), при этом регулятору была известна лишь интервальная оценка указанных возмущений (табл. 1). В качестве цели управления было принято решение задачи (12), где

$$\mathbf{V}_k = \mathbf{x}_{k+1}^T \mathbf{P} \mathbf{x}_{k+1}; \boldsymbol{\omega}_k = \delta \mathbf{a}_k^T \mathbf{R} \delta \mathbf{a}_k;$$

$$\mathbf{P} = \begin{bmatrix} 5,242 & 1,120 \\ 1,120 & 1,858 \end{bmatrix}; \mathbf{R} = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,5 \end{bmatrix}.$$

С целью достижения независимого управления звеньями робота матрицы  $\mathbf{K}_p$  и  $\mathbf{K}_v$  выбраны диагональными со следующими значениями ненулевых элементов:  $\{\mathbf{K}_p\}_{11}=\{\mathbf{K}_p\}_{22}=100$ ,  $\{\mathbf{K}_v\}_{11}=\{\mathbf{K}_v\}_{22}=20$ . Период квантования был принят 0,05 с. Построение эволюционирующих во времени множеств осуществлялось при помощи R-функций.

## 5 РЕЗУЛЬТАТИ

Результаты моделирования представлены на рис. 2–4. Рис. 2 отображает эффективность отработки системой управления желаемых траекторий. Здесь time – время, tracking error – ошибка управления, desired trajectory – желаемая траектория, actual trajectory – реальная траектория.

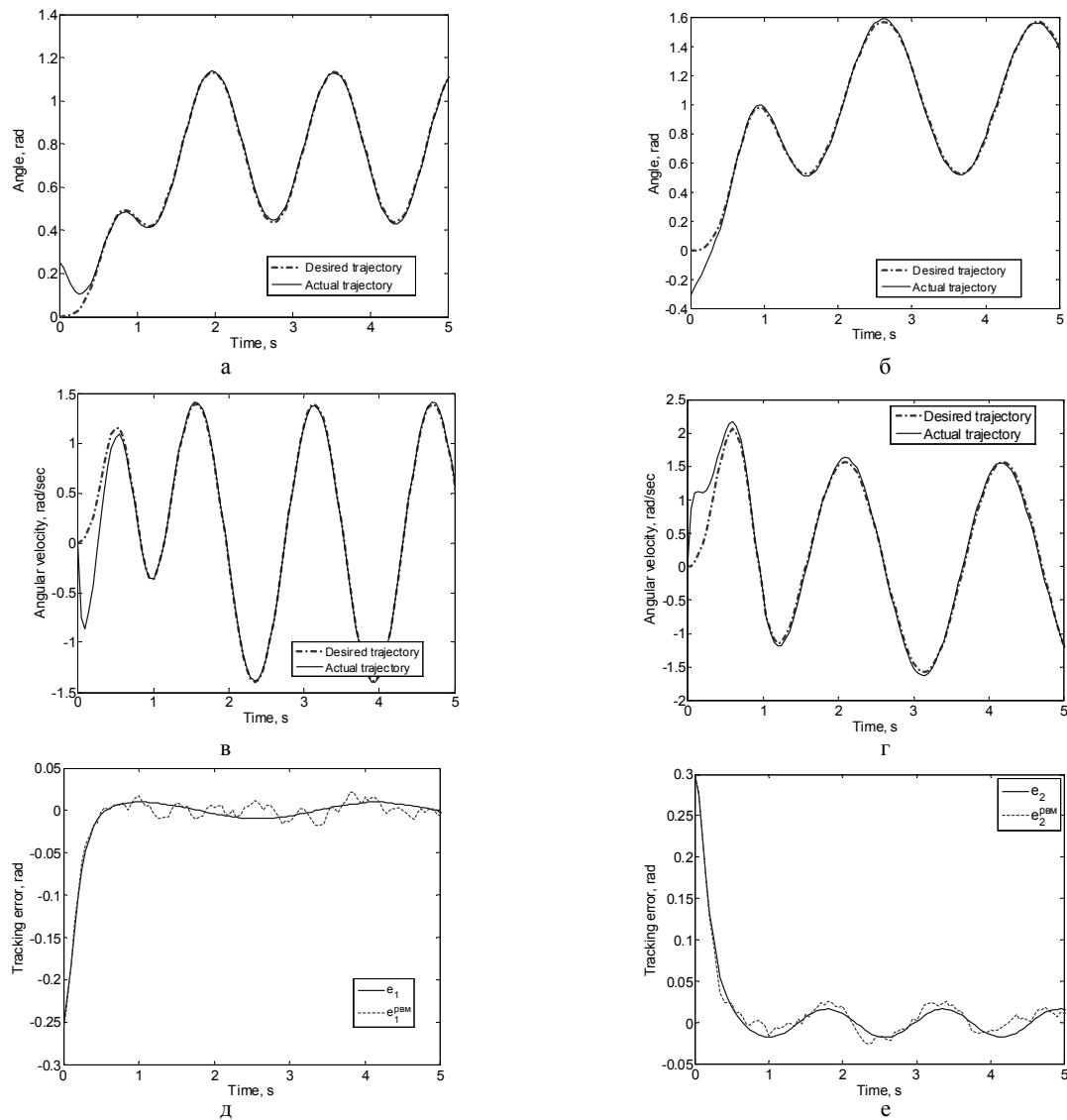


Рисунок 2 – Отработка желаемых траекторий движения: а – угол поворота первого звена (Angle); б – угол поворота второго звена (Angle); в – угловая скорость первого звена (Angular velocity); г – угловая скорость второго звена (Angular velocity); д – ошибки управления для первого звена (Tracking error); е – ошибки управления для второго звена (Tracking error)

Для сравнения на рис. 2д,е вместе с кривыми ошибок  $e_i$  ( $i=1,2$ ) слежения за желаемой траекторией робота с предлагаемым регулятором приведены кривые ошибок  $e_i^{\text{PBM}}$  ( $i=1,2$ ) робота с классическим РВМ. Действующие на звенья манипуляционного робота управляющие и возмущающие моменты приведены на рис. 3.

На рис. 4 приведены спрогнозированные множества  $\Omega^r$  возможных состояний системы и истинные состояния  $x^t$  для момента  $k = N$ . Построение множеств  $\Omega^r$  производилось при помощи выражений (12)–(14).

## 6 ОБСУЖДЕНИЕ

Результаты моделирования показывают, что предлагаемый регулятор обеспечивает высокое качество управления движением (рис. 2): большая начальная ошибка отработки заданной траектории, вызванная ненулевыми начальными условиями, компенсируется за 0,5 с. После этого момента желаемые и фактические траектории достаточно близки несмотря на действие неопределенных внешних возмущений и шумов измерений. Значения абсолютной ошибки по положению не превышают 0,0095 рад и 0,01705 рад для первого и второго звена, соответственно, что составляет 2 % и 2,3 % (рис. 2д,е). Из рис. 2д,е также легко видеть преимущество предлагаемого регулятора перед классическим РВМ.

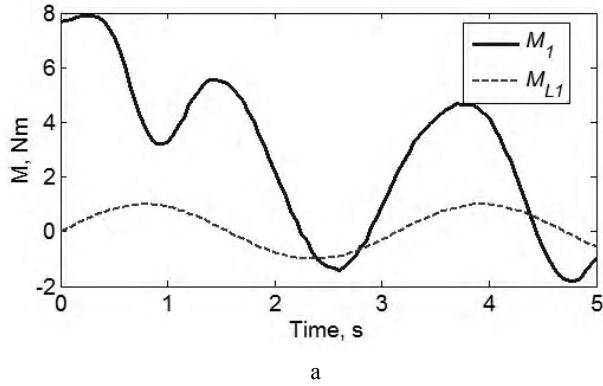
Как видно из рис. 3, регулятор обеспечивает плавное изменение управляющих моментов, что положительно сказывается на динамике исполнительных приводов робота.

Рис. 4 иллюстрирует высокую точность прогноза возможных значений координат состояния системы. Очевидно, что уменьшение меры множества  $\Omega^r$  повысит качество управления, поэтому на этапе проектирования регулятора необходим тщательный анализ возможных факторов неопределенности и интервалов их возможных изменений. Обеспечить более высокое качество управления можно также путем коррекции меры множеств на каждом шаге квантования по результатам наблюдений за функционированием системы, т.е. приданием регулятору адаптивных свойств.

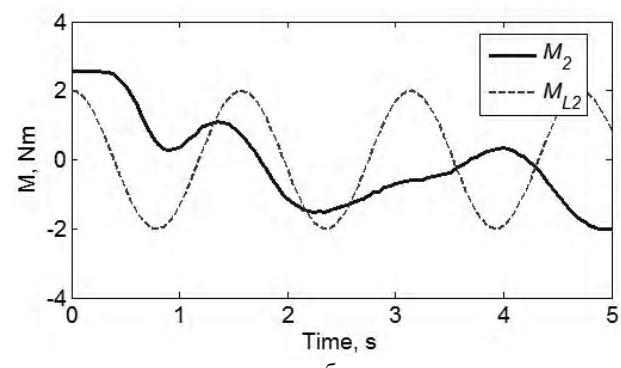
Следует также отметить, что двухзвеный манипулятор в качестве объекта при моделировании выбран исключительно из соображений наглядности, и на практике количество  $n$  звеньев манипуляционного робота ограничивается лишь вычислительной мощностью управляющей ЭВМ.

## ВЫВОДЫ

Рассмотрена актуальная задача управления движением манипуляционного робота в условиях неопределенности. Для ее решения разработан новый регулятор, основанный на игровом подходе в рамках концепции гарантированного управления. Данный регулятор

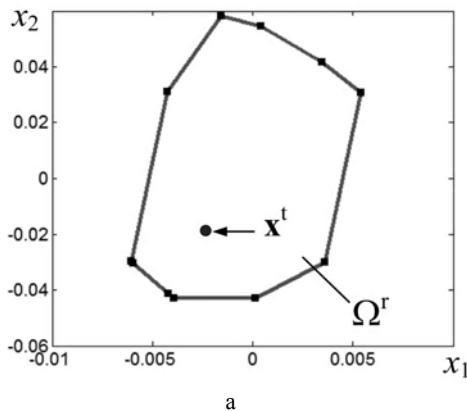


a

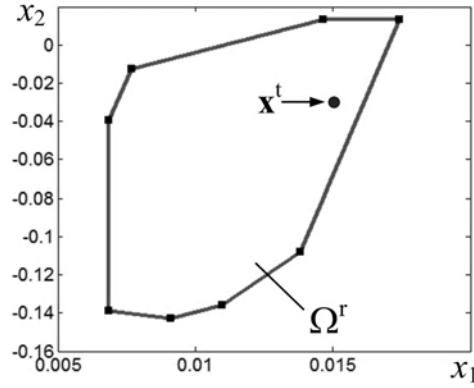


б

Рисунок 3 – Управляющий  $M$  и возмущающий  $M_L$  моменты: а – для первого звена; б – для второго звена



а



б

Рисунок 4 – Множества  $\Omega^r$  возможных состояний и истинные состояния  $x^t$  системы:  
а – для первого звена; б – для второго звена

предполагает наличие двух контуров: первый контур рассчитывает основное управляющее воздействие при помощи метода вычисляемого момента, а назначение второго – определение дополнительного управления, компенсирующего действие неопределенных факторов на основе дифференциальной игры с квадратичным функционалом качества.

Практическая ценность предложенного регулятора заключается в гарантированном получении близких к оптимальным законов управления при любых допустимых реализациях факторов неопределенности, а также в относительно невысоких требованиях к вычислительным ресурсам.

Недостатком регулятора является то, что действующие на систему факторы далеко не всегда будут стремиться обеспечить максимальное значение критерия качества. Повышение качества управления возможно за счет введения дополнительного контура адаптации, уточняющего в процессе управления оценки неопределенных параметров. С решением этой задачи связаны наши дальнейшие исследования.

## БЛАГОДАРНОСТИ

Работа выполнена в рамках выполнения госбюджетной научно-исследовательской работы «Разработка моделей и методов синтеза системы автоматического управления рабочим оборудованием экскаватора» (номер государственной регистрации 0113U000175).

## СПИСОК ЛІТЕРАТУРЫ

- Peng W. Computed torque control-based composite nonlinear feedback controller for robot manipulators with bounded torques / W. Peng, Z. Lin, J. Su // IET Control Theory & Applications. – 2009. – Vol. 3, № 6. – P. 701–711. DOI: 10.1049/iet-cta.2008.0259
- Kuo T. C. Trajectory control of a robotic manipulator utilizing an adaptive fuzzy sliding mode / T. C. Kuo // World Academy of Science, Engineering and Technology. – 2010. – Vol. 65. – P. 913–917.
- Yan R. Adaptive learning tracking control of robotic manipulators with uncertainties / R. Yan, K. Tee, H. Z. Li // Journal of Control Theory and Applications. – 2010. – Vol. 8, № 2. – P. 160–165. DOI: 10.1007/s11768-010-0010-2
- Mohanty A. Indirect adaptive robust control of hydraulic manipulators with accurate parameter estimates / A. Mohanty, B. Yao // IEEE Transactions on Control Systems Technology. – 2011. – Vol. 19, № 3. – P. 567–575. DOI: 10.1109/TCST.2010.2048569
- Neural network-based sliding mode adaptive control for robot manipulators / [T. Sun, H. Pei, Y. Pan, H. Zhou, Zhang C.] // Neurocomputing. – 2011. – Vol. 74, № 14. – P. 2377–2384. DOI: 10.1016/j.neucom.2011.03.015
- Position control of robot manipulator: design a novel SISO adaptive sliding mode fuzzy PD fuzzy sliding mode control / [F. Piltan, N. Sulaiman, S. Allahdadi, M. Dialame, A. Zare] // International Journal of Artificial Intelligence and Expert System. – 2011. – Vol. 2, № 5. – P. 208–228.
- Yang Z. J. Decentralized adaptive robust control of robot manipulators using disturbance observers / Z. J. Yang, Y. Fukushima, P. Qin // IEEE Transactions on Control Systems Technology. – 2012. – Vol. 20, № 5. – P. 1357–1365. DOI: 10.1109/TCST.2011.2164076
- Robust control of robot manipulators based on uncertainty and disturbance estimation / [J. P. Kolhe, M. Shaheed, T. S. Chandar, S. E. Talole] // International Journal of Robust and Nonlinear Control. – 2013. – Vol. 23, № 1. – P. 104–122. DOI: 10.1002/rnc.1823
- Abbaszadeh E. Control of robot manipulator with uncertain dynamics by adaptive fuzzy controller / E. Abbaszadeh, M. Haddad-Zarif [Electronic resource] // Journal of Intelligent Automation Systems. – 2014. – Access mode: [http://jias.shahroodut.ac.ir/article\\_225\\_0.html](http://jias.shahroodut.ac.ir/article_225_0.html).
- Adaptive fuzzy computed-torque control for robot manipulator with uncertain dynamics / [Y. Chen, G. Y. Ma, S. X. Lin, J. Gao] // International Journal of Advanced Robotic Systems. – 2012. – Vol. 9. – P. 201–209. DOI: 10.5772/54643
- Acho L. Computed-torque-plus-compensation-plus-chattering controller of robot manipulators [Electronic resource] / L. Acho, Y. Vidal, P. Francesc // Contemporary Robotics – Challenges and Solutions; ed. A. D. Rodi. – InTech, 2009. – Access mode: <http://www.intechopen.com/books/contemporary-robotics-challenges-and-solutions/computed-torque-plus-compensation-plus-chattering-controller-of-robot-manipulators> DOI: 10.5772/7803
- Kelly R. Control of Robot Manipulators in Joint Space / R. Kelly, V. Santibanez, A. Loria. – Berlin : Springer, 2005. – 452 p.
- Springer handbook of robotics / [eds. B. Siciliano, O. Khatib]. – Springer, 2008. – 1611 p.
- Nguyen-Tuong D. Computed torque control with nonparametric regression models / D. Nguyen-Tuong, M. Seeger, J. Peters // American Control Conference, Seattle, USA 11–13 June 2008 : proceedings. – Los Alamitos : IEEE, 2008. – P. 212–217. DOI: 10.1109/ACC.2008.4586493
- Robotics: modelling, planning and control / [B. Siciliano, L. Sciavicco, L. Villani, G. Oriolo]. – Springer. – 2009. – 632 p.
- Chen Y. Adaptive fuzzy and robust  $H_{\infty}$  compensation control for uncertain robot / Y. Chen, G. Ma, S. Lin, J. Gao // International Journal of Robotics and Automation. – 2013. – Vol. 2, № 4. – P. 174–188. DOI: 10.11591/ijra.v2i4.2763
- Robust adaptive inverse dynamics control for uncertain robot manipulator / [Y. Chen, G. Mei, G. Ma, S. Lin, J. Gao] // International Journal of Innovative Computing, Information and Control. – 2014. – Vol. 10, № 2. – P. 575–587.
- Design of model free adaptive fuzzy computed torque controller: applied to nonlinear second order system / [F. Piltan, N. Sulaiman, A. Jalali, F. D. Narouei] // International Journal of Robotics and Automation. – 2011. – Vol. 2, № 4. – P. 232–244.
- Design novel nonlinear controller applied to robot manipulator: design new feedback linearization fuzzy controller with minimum rule base tuning method / F. Piltan, M. Keshavarz, A. Badri, A. Zargari // International Journal of Robotics and Automation. – 2012. – Vol. 3, № 1. – P. 1–18.
- Управление в условиях неопределенности: гарантированные результаты в задачах управления и идентификации : монография / В. М. Кунцевич. – К. : Наук. думка, 2006. – 264 с.
- Kurzhanski A. B. Optimization of output feedback control under set-membership uncertainty / A. B. Kurzhanski, P. Varaiya // Journal of optimization theory and applications. – 2011. – Vol. 151, № 1. – P. 11–32. DOI: 10.1007/s10957-011-9861-z
- Kurzhanski A. B. Output feedback guaranteed tracking control through finite observers / A. B. Kurzhanski, P. A. Tochilin // Conference on Decision and Control: 52nd international conference, Florence, Italy, 10–13 December: proceedings. – IEEE, 2013. – P. 4448–4453. DOI: 10.1109/CDC.2013.6760574
- Гурко А. Г. Оценивание вектора состояния динамической системы в условиях неопределенности / А. Г. Гурко, В. М. Колодяжный // Кибернетика и системный анализ. – 2013. – № 6. – С. 23–31.
- Гурко А. Г. Управление дискретной системой при ограниченных возмущениях / А. Г. Гурко, И. Ф. Еременко // Проблемы управления и информатики. – 2011. – № 6 – С. 1–11.

Статья поступила в редакцию 23.09.2014.

После доработки 07.10.2014.

Гурко О. Г.<sup>1</sup>, Янчевський І. В.<sup>2</sup>

<sup>1</sup>Канд. техн. наук, доцент, доцент кафедри автоматизації та комп’ютерно-інтегрованих технологій Харківського національного автомобільно-дорожнього університету, Україна

<sup>2</sup>Д-р фіз.-мат. наук, доцент, професор кафедри теорії машин та механізмів Харківського національного автомобільно-дорожнього університету, Україна

## ГАРАНТОВАНЕ УПРАВЛІННЯ РУХОМ МАНІПУЛЯЦІЙНОГО РОБОТА

Розглядається задача управління рухом маніпуляційного робота при наявності в його моделі параметричної невизначеності та дії неконтрольованих збурень. Запропоновано робастний регулятор, що заснований на методі обчислюваного моменту, в якому для обчислення додаткового управління, що забезпечує компенсацію дії невизначених факторів, використовується гарантований підхід, згідно якого значення цих факторів відомі лише з точністю до принадлежності деяким множинам. Для оптимізації управління використаний квадратичний критерій якості. Ефективність запропонованого рішення підтверджено результатами чисельного моделювання.

**Ключові слова:** маніпуляційний робот, регулятор з обчислювальним моментом, гарантоване управління, множина можливих станів.

Gurko A. G.<sup>1</sup>, Yanchevskiy I. V.<sup>2</sup>

<sup>1</sup>PhD, Associate professor, Associate Professor of department of automation and computer-integrated technologies, Kharkiv National Automobile and Highway University, Ukraine

<sup>2</sup>Dr.Sc., Associate Professor, Professor of department of theory of machines and mechanisms, Kharkiv National Automobile and Highway University, Ukraine

## GUARANTEED CONTROL OF ROBOT-MANIPULATOR MOVEMENT

The problem of a robot-manipulator movement under parametric uncertainties and uncontrolled disturbances has been considered. The robust controller based on the computed torque control using to calculate additional control that enables compensation of uncertain factors action a guaranteed approach, according to which values of these factors are known only to the limits of a certain set membership, has been proposed. The quadratic performance criterion for control optimization has been used. The efficiency of the proposed solution is confirmed with numerical simulation.

**Keywords:** robot-manipulator, computed torque control, guaranteed control, set of probable states.

## REFERENCES

1. Peng W., Lin Z., Su J. Computed torque control-based composite nonlinear feedback controller for robot manipulators with bounded torques, *IET Control Theory & Applications*, 2009, Vol. 3, No. 6, pp. 701–711. DOI: 10.1049/iet-cta.2008.0259
2. Kuo T. C. Trajectory control of a robotic manipulator utilizing an adaptive fuzzy sliding mode, *World Academy of Science, Engineering and Technology*, 2010, Vol. 65, pp. 913–917.
3. Yan R., Tee K., Li H. Z. Adaptive learning tracking control of robotic manipulators with uncertainties, *Journal of Control Theory and Applications*, 2010, Vol. 8, No. 2, pp. 160–165. DOI: 10.1007/s11768-010-0010-2
4. Mohanty A., Yao B. Indirect adaptive robust control of hydraulic manipulators with accurate parameter estimates, *IEEE Transactions on Control Systems Technology*, 2011, Vol. 19, No. 3, pp. 567–575. DOI: 10.1109/TCST.2010.2048569
5. Sun T., Pei H., Pan Y., Zhou H., Zhang C. Neural network-based sliding mode adaptive control for robot manipulators, *Neurocomputing*, 2011, Vol. 74, No. 14, pp. 2377–2384. DOI: 10.1016/j.neucom.2011.03.015
6. Piltan F., Sulaiman N., Allahdadi S., Dialame M., Zare A. Position control of robot manipulator: design a novel SISO adaptive sliding mode fuzzy PD fuzzy sliding mode control, *International Journal of Artificial Intelligence and Expert System*, 2011, Vol. 2, No. 5, pp. 208–228.
7. Yang Z. J., Fukushima Y., Qin P. Decentralized adaptive robust control of robot manipulators using disturbance observers, *IEEE Transactions on Control Systems Technology*, 2012, Vol. 20, No. 5, pp. 1357–1365. DOI: 10.1109/TCST.2011.2164076
8. Kolhe J. P., Shaheed M., Chandar T. S., Talole S. E. Robust control of robot manipulators based on uncertainty and disturbance estimation, *International Journal of Robust and Nonlinear Control*, 2013, Vol. 23, No. 1, pp. 104–122. DOI: 10.1002/rnc.1823
9. Abbaszadeh E., Haddad-Zarif M. Control of robot manipulator with uncertain dynamics by adaptive fuzzy controller, *Journal of Intelligent Automation Systems*, 2014, [Electronic resource], Access mode: [http://jias.shahroodut.ac.ir/article\\_225\\_0.html](http://jias.shahroodut.ac.ir/article_225_0.html).
10. Chen Y., Ma G. Y., Lin S. X., Gao J. Adaptive fuzzy computed-torque control for robot manipulator with uncertain dynamics, *International Journal of Advanced Robotic Systems*, 2012, Vol. 9, pp. 201–209. DOI: 10.5772/54643
11. Acho L., Vidal Y., Francesc P. editor A.D. Rodi. Computed-torque-plus-compensation-plus-chattering controller of robot manipulators, *Contemporary Robotics – Challenges and Solutions*, InTech, 2009, [Electronic resource], Access mode: <http://www.intechopen.com/books/contemporary-robotics-challenges-and-solutions/>
12. Kelly R., Santibanez V., Loria A. Control of Robot Manipulators in Joint Space. Springer, 2005, 452 p.
13. Springer handbook of robotics. [eds. B. Siciliano, O. Khatib], Springer, 2008, 1611 p.
14. Nguyen-Tuong D., Seeger M., Peters J. Computed torque control with nonparametric regression models, *American Control Conference*. Seattle, USA 11–13 June 2008, proceedings, IEEE, 2008, pp. 212–217. DOI: 10.1109/ACC.2008.4586493
15. Siciliano B., Sciavicco L., Villani L., Oriolo G. Robotics: modelling, planning and control. Springer, 2009, 632 p.
16. Chen Y., Ma G., Lin S., Gao J. Adaptive fuzzy and robust H<sub>∞</sub> compensation control for uncertain robot, *International Journal of Robotics and Automation*, 2013, Vol. 2, No. 4, pp. 174–188. DOI: 10.11591/ijra.v2i4.2763
17. Chen Y., Mei G., Ma G., Lin S., Gao J. Robust adaptive inverse dynamics control for uncertain robot manipulator, *International Journal of Innovative Computing, Information and Control*, 2014, Vol. 10, No. 2, pp. 575–587.
18. Piltan F., Sulaiman N., Jalali A., Narouei F. D. Design of model free adaptive fuzzy computed torque controller: applied to nonlinear second order system, *International Journal of Robotics and Automation*, 2011, Vol. 2, No. 4, pp. 232–244.
19. Piltan F., Keshavarz M., Badri A., Zargari A. Design novel nonlinear controller applied to robot manipulator: design new feedback linearization fuzzy controller with minimum rule base tuning method, *International Journal of Robotics and Automation*, 2012, Vol. 3, No. 1, pp. 1–18.
20. Kuncevich V. M. Управление в условиях неопределенности: гарантированные результаты в задачах управления и идентификации: монография, Kiev, Nauk. dumka, 2006, 264 p.
21. Kurzhanski A. B., Varaiya P. Optimization of output feedback control under set-membership uncertainty, *Journal of optimization theory and applications*, 2011, Vol. 151, No. 1, pp. 11–32. DOI: 10.1007/s10957-011-9861-z
22. Kurzhanski A. B., Tochilin P. A. Output feedback guaranteed tracking control through finite observers, *Conference on Decision and Control, 52nd international conference*, Florence, Italy, 10–13 December: proceedings, IEEE, 2013, pp. 4448–4453. DOI: 10.1109/CDC.2013.6760574
23. Gurko A. G., Kolodjazhnyj V. M. Оценивание вектора состояния динамической системы в условиях неопределенности, *Kibernetika i sistemnyj analiz*, 2013, No. 6, pp. 23–31.
24. Gurko A. G., Eremenko I. F. Управление дискретной системой при органической возмущенности, *Problemy upravlenija i informatiki*, 2011, No. 6, pp. 1–11.

*Наукове видання*

**Радіоелектроніка,  
інформатика,  
управління**

**№ 2/2014**

**Науковий журнал**

Головний редактор – д-р фіз.-мат. наук В. В. Погосов  
Заст. головного редактора – д-р техн. наук С. О. Субботін

Комп'ютерне моделювання та верстання  
Редактор англійських текстів

С. В. Зуб  
С. О. Субботін

Оригінал-макет підготовлено у редакційно-видавничому відділі ЗНТУ

Свідоцтво про державну реєстрацію  
КВ № 6904 від 29.01.2003.

*Підписано до друку 23.01.2015. Формат 60×84/8.  
Папір офс. Різогр. друк Ум. друк. арк. 21,16.  
Тираж 300 прим. Зам. № 06.*

*69063, м. Запоріжжя, ЗНТУ, друкарня, вул. Жуковського, 64*

*Свідоцтво суб'єкта видавничої справи  
ДК № 2394 від 27.12.2005.*