Запорізький національний технічний університет

Радіоелектроніка, інформатика, управління

Науковий журнал

Виходить чотири рази на рік
№ 2(45) 2018
Заснований у січні 1999 року.
Засновник і видавець – Запорізький національний технічний університет.
ISSN 1607-3274 (друкований), ISSN 2313-688X (електронний).
Запоріжжя
ЗНТУ
2018

Запорожский национальный технический университет

Радиоэлектроника, информатика, управление

Научный журнал

Выходит четыре раза в год
№ 2(45) 2018
Основан в январе 1999 года.
Основатель и издатель – Запорожский национальный технический университет.
ISSN 1607-3274 (печатный), ISSN 2313-688X (электронный).
Запорожье
ЗНТУ
2018

Zaporizhzhia National Technical University

Radio Electronics, Computer Science, Control

The scientific journal

Published four times per year
№ 2(45) 2018
Founded in January 1999.
Founder and publisher – Zaporizhzhia National Technical University.
ISSN 1607-3274 (print), ISSN 2313-688X (on-line).
Zaporizhzhia
ZNTU
2018

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

Наказом Міністерства освіти і науки України № 1328 від 21.12.2015 р. «Про затвердження рішень Атестаційної колегії Міністерства щодо діяльності спеціалізованих вчених рад від 15 грудня 2015 року» **журнал включений до переліку наукових фахових видань України,** в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата фізико-математичних та технічних наук.

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 CC ліцензією CC-BY.

Опублікованими статтям присвоюється унікальний ідентифікатор цифрового об'єкта DOI.

Журнал входить до наукометричної бази Web of Science.

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'єктивний 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 – Субботін С. О., д-р техн. наук, Україна

Заст. головного редактора – Піза Д. М., д-р техн. наук, Україна

Члени редколегії:

Андроулідакіс Й., д-р філософії, Греція
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
Гімпілевич 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
Камеяма М., д-р техн. наук, Японія
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
Лебедєв 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
Погосов В. В., д-р фіз.-мат. наук, Україна
Рубель О. В., канд. техн. наук, Канада
Хаханов В. І., д-р техн. наук, Україна, редактор розділу з інформатики
Шарпанських О. А., д-р філософії, Нідерланди, редактор розділу з інформатики

Рекомендовано до видання вченою радою ЗНТУ, протокол № 10 від 29.05.2018.

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

© Запорізький національний технічний університет, 2018

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

Приказом Министерства образования и науки Украины № 1328 от 21.12.2015 г. «Об утверждении решений Аттестационной коллегии Министерства относительно деятельности специализированных ученых советов от 15 декабря 2015 года» **журнал включен в перечень научных профессиональных изданий Украины,** в которых могут публиковаться результаты диссертационных работ на соискание ученых степеней доктора и кандидата физико-математических и технических наук.

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 CC лицензией CC-BY.

Опубликованным статьям присваивается уникальный идентификатор цифрового объекта DOI.

Журнал входит в наукометрическую базу Web of Science.

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, ВИНИТИ, 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 нейронечеткие 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 – Субботин С. А., д-р. техн. наук, Украина

Зам. главного редактора – Пиза Д. М., д-р техн. наук, Украина

Члены редколлегии:

Андроулидакис И., д-р философии, Греция
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
Камеяма М., д-р техн. наук, Япония
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
Марковска-Качмар У., д-р наук, Польша
Олещук В. А., канд. физ.-мат. наук, Норвегия, редактор раздела по радиоэлектронике
Онуфриенко В. М., д-р физ.-мат. наук, Украина
Папшицький М., д-р философии, Польша
Погосов В. В., д-р физ.-мат. наук, Украина
Рубель О. В., канд. техн. наук, Канада
Хаханов В. И., д-р техн. наук, Украина, редактор раздела по информатике
Шарпанских А. А., доктор философии, Нидерланды – редактор раздела по информатике

Рекомендовано к изданию ученым советом ЗНТУ, протокол № 10 от 29.05.2018.

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

© Запорожский национальный технический университет, 2018

The scientific journal «Radio Electronics, Computer Science, Control» is published by the Zaporizhzhia National Technical University (ZNTU). since 1999 with periodicity four 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 КВ №6904).

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

By the Order of the Ministry of Education and Science of Ukraine from 21.12.2015 № 1328 "On approval of the decision of the Certifying Collegium of the Ministry on the activities of the specialized scientific councils dated 15 December 2015" **journal is included in the list of scientific specialized periodicals of Ukraine,** where the results of dissertations for Doctor of Science and Doctor of Philosophy in Mathematics and Technical Sciences 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 included into Web of Science.

The journal is abstracted and indexed in leading international and national abstractig 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), Djerelo), 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: S. A. Subbotin, Doctor of Science in Engineering, Ukraine

Deputy Editor-in-Chief: D. M. Piza, Doctor of Science in Engineering, Ukraine

Members of Editorial Board:

I. Androulidakis, Ph. D, Greece
V. M. Bezruk, Doctor of Science in Engineering, Ukraine
Ye. V. Bodyanskiy , Doctor of Science in Engineering, Ukraine, Control 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
M. Kameyama, Doctor of Science, Japan
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
V. V. Pogosov, Doctor of Science in Physics and Mathematics, Ukraine
O. V. Rubel, Ph.D, Canada
A. A. Sharpanskykh, Ph.D, Netherlands, Computer Science section editor
S. N. Vassilyev, Doctor of Science in Physics and Mathematics, Academican of Russian Academy of Sciences, Russia
E. N. Zaitseva, Ph.D, Slovakia

Recommended for publication by the Academic Council of ZNTU, protocol № 10 dated 29.05.2018.

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, Zaporizhzhia, 69063, Ukraine.

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

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

E-mail: rvv@zntu.edu.ua

© Zaporizhzhia National Technical University, 2018

ЗМІСТ

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ....7

Дмитриев В. С.
ВПЛИВ ТЕРМІЧНОЇ ОБРОБКИ НА ВОЛЬТ-АМПЕРНУ ХАРАКТЕРИСТИКУ ТА ІНЖЕКЦІЙНІ ВЛАСТИВОСТІ ГЕТЕРОПЕРЕХОДУ Ag/n-n^+GaAs....7

Жиров Г. Б., Лєнков Є. С.
АЛГОРИТМИ МОДЕЛЮВАННЯ ПРОЦЕСІВ ТЕХНІЧНОГО ОБСЛУГОВУВАННЯ ЗА СТАНОМ СКЛАДНИХ РАДІОЕЛЕКТРОННИХ ОБ'ЄКТІВ....14

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ....22

Lamtyugova S. N, Sidorov M. V., Sytnykova I. V.
METHOD OF NUMERICAL ANALYSIS OF THE PROBLEM OF MASS TRANSFER OF A CYLINDRICAL BODY WITH THE UNIFORM TRANSLATIONAL FLOW....22

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ....30

Бісікало О. В., Грищук Т. В., Ковтун В. В.
ОПТИМІЗАЦІЯ КЛАСИФІКАТОРА АВТОМАТИЗОВАНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ МОВЦЯ КРИТИЧНОГО ЗАСТОСУВАННЯ....30

Dovbysh A. S., Velykodnyi D. V., Protsenko O. B., Zimovets V. I.
OPTIMIZATION OF PARAMETERS OF MACHINE LEARNING OF THE SYSTEM OF FUNCTIONAL DIAGNOSTICS OF THE ELECTRIC DRIVE OF A SHAFT LIFTING MACHINE....44

Kovalevska O. S., Kovalevskyy S. V.
APPLICATION OF ACOUSTIC ANALYSIS IN CONTROL SYSTEMS OF ROBOTIC MACHINE TOOLS....51

Oliinyk A., Subbotin S., Lovkin V., Ilyashenko M., Blagodariov O.
PARALLEL METHOD OF BIG DATA REDUCTION BASED ON STOCHASTIC PROGRAMMING APPROACH....60

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ....73

Alguliyev R. M., Aliguliyev R. M., Imamverdiyev Y. N., Sukhostat L. V.
AN IMPROVED ENSEMBLE APPROACH FOR DOS ATTACKS DETECTION....73

Akhmetov B. B., Lakhno V. A., Malyukov V. P.
MODEL OF INVESTMENT STRATEGIES IN CYBER SECURITY SYSTEMS OF TRANSPORT SITUATIONAL CENTERS....83

Гадецька С. В., Гороховатський В. О.
МЕТОДИ СТРУКТУРНОЇ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ЗАСАДАХ БАЄСОВСЬКОЇ ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ....90

Dadenkov S. A.
THE SIMULATION OF P-PERSISTENT CSMA ALGORITHM OF UNEQUAL RANDOM MULTIPLE ACCESS....98

Конп А. М., Орловский Д. Л.
ПОДХОД К АНАЛИЗУ И ОПТИМИЗАЦИИ МОДЕЛЕЙ БИЗНЕС-ПРОЦЕССОВ В НОТАЦИИ BPMN....108

Miroschnyk M. A., Pakhomov Y. V., Shkil A. S., Kulak E. N., Kucherenko D. Y.
DESIGN AUTOMATION OF EASY-TESTED DIGITAL FINITE STATE MACHINES....117

Panin V. V., Doronin V. V., Aleynikov V. M.
APPLICATION OF THE SYSTEM ANALYSIS OF IMPLEMENTATION OF THE INSTRUMENTAL METHOD OF NAVIGATION ON INLAND WATERWAYS OF UKRAINE....125

Prots'ko I. O., Teslyuk V. M.
DEVELOPMENT OF WFTA BASED ON THE HASHING ARRAY....135

Faure E. V., Shcherba A. I., Kharin A. A.
FACTORIAL CODE WITH A GIVEN NUMBER OF INVERSIONS....143

Федасюк Д. В., Чопей Р. С.
АЛГОРИТМ ПОБУДОВИ ГРАФУ ПОТОКУ КЕРУВАННЯ ЗА ТЕКСТОМ ПРОГРАМИ МОВОЮ С....154

Шитикова Е. В., Табунщик С. С., Табунщик Г. В.
МЕТОД ФОРМИРОВАНИЯ ОБЪЕМА РАБОТ ДЛЯ ПРОГРАММ ИСПЫТАНИЙ НА ОСНОВЕ НЕЧЕТКОГО ВЫВОДА....162

УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ....169

Святный В. А., Иванов А. Ю., Иванов Ю. А.
ИССЛЕДОВАНИЕ ЦИКЛИЧЕСКОГО РАСПИСАНИЯ ДЛЯ МОДЕЛИ СЕТЕВОГО ДИНАМИЧЕСКОГО ОБЪЕКТА РЕАЛЬНОГО ВРЕМЕНИ....169

Firsov S. N., Pishchukhina O. A.
INTELLIGENT SUPPORT OF MULTILEVEL FUNCTIONAL STABILITY OF CONTROL AND NAVIGATION SYSTEMS....177

CONTENTS

RADIO ELECTRONICS AND TELECOMMUNICATIONS.....7

Dmitriev V. S.
THE THERMAL PROCESSING INFLUENCE ON CURRENT-VOLTAGE CHARACTERISTIC AND INJECTION PROPERTIES OF $Ag/n\text{-}n^{+}GaAs$ HETEROJUNCTIONS.....7

Zhyrov G. B., Lenkov E. S.
ALGORITHMS OF SIMULATION OF TECHNICAL SERVICE PROCESSES AS COMPLEX RADIOELECTRONIC OBJECTS.....14

MATHEMATICAL AND COMPUTER MODELING.....22

Lamtyugova S. N, Sidorov M. V., Sytnykova I. V.
METHOD OF NUMERICAL ANALYSIS OF THE PROBLEM OF MASS TRANSFER OF A CYLINDRICAL BODY WITH THE UNIFORM TRANSLATIONAL FLOW.....22

NEUROINFORMATICS AND INTELLIGENT SYSTEMS.....30

Bisikalo O. V., Grischuk T. V., Kovtun V. V.
THE AUTOMATIC SPEAKER RECOGNITION SYSTEM OF CRITICAL USE CLASSIFIER OPTIMIZATION.....30

Dovbysh A. S., Velykodnyi D. V., Protsenko O. B., Zimovets V. I.
OPTIMIZATION OF PARAMETERS OF MACHINE LEARNING OF THE SYSTEM OF FUNCTIONAL DIAGNOSTICS OF THE ELECTRIC DRIVE OF A SHAFT LIFTING MACHINE.....44

Kovalevska O. S., Kovalevskyy S. V.
APPLICATION OF ACOUSTIC ANALYSIS IN CONTROL SYSTEMS OF ROBOTIC MACHINE TOOLS.....51

Oliinyk A., Subbotin S., Lovkin V., Ilyashenko M., Blagodariov O.
PARALLEL METHOD OF BIG DATA REDUCTION BASED ON STOCHASTIC PROGRAMMING APPROACH.....60

PROGRESSIVE INFORMATION TECHNOLOGIES.....73

Alguliyev R. M., Aliguliyev R. M., Imamverdiyev Y. N., Sukhostat L. V.
AN IMPROVED ENSEMBLE APPROACH FOR DOS ATTACKS DETECTION.....73

Akhmetov B. B., Lakhno V. A., Malyukov V. P.
MODEL OF INVESTMENT STRATEGIES IN CYBER SECURITY SYSTEMS OF TRANSPORT SITUATIONAL CENTERS.....83

Gadetska S. V., Gorokhovatsky V. A.
STRUCTURAL CLASSIFICATION IMAGES USING BAYESIAN DECISION MAKING.....90

Dadenkov S. A.
THE SIMULATION OF P-PERSISTENT CSMA ALGORITHM OF UNEQUAL RANDOM MULTIPLE ACCESS.....98

Kopp A. M., Orlovskyi D. L.
AN APPROACH TO BPMN BASED BUSINESS PROCESS MODELS ANALYSIS AND OPTIMIZATION.....108

Miroschnyk M. A., Pakhomov Y. V., Shkil A. S., Kulak E. N., Kucherenko D. Y.
DESIGN AUTOMATION OF EASY-TESTED DIGITAL FINITE STATE MACHINES.....117

Panin V. V., Doronin V. V., Aleynikov V. M.
APPLICATION OF THE SYSTEM ANALYSIS OF IMPLEMENTATION OF THE INSTRUMENTAL METHOD OF NAVIGATION ON INLAND WATERWAYS OF UKRAINE.....125

Prots'ko I. O., Teslyuk V. M.
DEVELOPMENT OF WFTA BASED ON THE HASHING ARRAY.....135

Faure E. V., Shcherba A. I., Kharin A. A.
FACTORIAL CODE WITH A GIVEN NUMBER OF INVERSIONS.....143

Fedasyuk D. V., Chopey R. S.
THE ALGORITHM OF CONSTRUCTING CONTROL FLOW GRAPH BASED ON A PROGRAM WRITTEN IN C.....154

Shytikova O., Tabunshchyk S., Tabunshchyk G.
METHOD FOR FORMING TEST WORKFLOW BASED ON FUZZY INFERENCE.....162

CONTROL IN TECHNICAL SYSTEMS.....169

Svjatnij V. A., Ivanov A. Y., Ivanov Y. A.
CYCLIC SCHEDULE RESEARCH FOR THE MODEL OF REAL-TIME DYNAMIC NETWORK OBJECT.....169

Firsov S. N., Pishchukhina O. A.
INTELLIGENT SUPPORT OF MULTILEVEL FUNCTIONAL STABILITY OF CONTROL AND NAVIGATION SYSTEMS.....177

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

РАДИОЭЛЕКТРОНИКА И ТЕЛЕКОММУНИКАЦИИ

RADIO ELECTRONICS AND TELECOMMUNICATIONS

УДК 621.382.2/3

Дмитриев В. С.

Завідувач навчально-наукової лабораторії кафедри мікроелектронних інформаційних систем Запорізької державної інженерної академії, Запоріжжя, Україна

ВПЛИВ ТЕРМІЧНОЇ ОБРОБКИ НА ВОЛЬТ-АМПЕРНУ ХАРАКТЕРИСТИКУ ТА ІНЖЕКЦІЙНІ ВЛАСТИВОСТІ ГЕТЕРОПЕРЕХОДУ Ag/n-n$^+$GaAs

Актуальність. В наш час дослідження та розробку гетеропереходів проводять у напрямках пошуку нових композицій метал-арсенід галію, розробки технологічних режимів, які змогли б забезпечити відтворюваність параметрів НВЧ приладів з бар'єрами Шотткі.

Срібло у порівнянні з золотом має більшу тепло- та електропровідність, відносно невеликий коефіцієнт дифузії в арсенід галію, що дозволяє зменшити товщину перехідного шару. Перехід до металізації на основі срібла повинен підвищити технічні характеристики виробів. Розробка технологічних режимів виготовлення покращених гетеропереходів до GaAs на основі срібла є актуальною з наукової та практичної точок зору.

Мета роботи – встановлення впливу термічної обробки на реальні вольт-амперні характеристики та величину коефіцієнта інжекції гетеропереходу Ag/n-n$^+$GaAs.

Метод. Метод виготовлення гетеропереходів – термічне випаровування у вакуумі. Розрахунок параметрів гетеропереходу за методом вольт-амперних характеристик.

Результати. Рекомендований режим хімічної обробки GaAs-підкладки. Досліджено вплив термічної обробки на параметри та характеристики гетеропереходів Ag/n-n$^+$GaAs. Досліджено вплив висоти бар'єру на величину коефіцієнта інжекції. Розглянуті та апробовані різні методи визначення висоти бар'єру Шотткі та фактору неідеальності з вольт-амперних характеристик.

Висновки. Встановлено, що підвищення температури відпалу до 803 К дає найбільше значення висоти бар'єру φ_b для гетеропереходу Ag/n-n$^+$GaAs з концентрацією донорів в епітаксійному шарі N_D=2·10^{16} см$^{-3}$. Встановлено, що найбільш точним методом визначення параметрів гетеропереходу з вольт-амперної характеристики є метод direct approximation, який враховує вплив послідовного опору при визначенні фактору неідеальності та висоти бар'єру Шотткі. Розраховані коефіцієнти інжекції γ для гетеропереходів з різною висотою бар'єру Шотткі, отриманих при різних температурах відпалу, мають дуже малі значення, тому інжекцією дірок можна знехтувати.

Ключові слова: бар'єрний перехід, вольт-амперна характеристика, висота бар'єру, фактор неідеальності, коефіцієнт інжекції.

НОМЕНКЛАТУРА

БШ – бар'єр Шотткі;
ВАХ – вольт-амперна характеристика;
НВЧ – надвисока частота;
A^{**} – ефективна стала Річардсона;
C_B – бар'єрна ємність;
$d_{\text{еш}}$ – товщина епітаксійного шару;
k – стала Больцмана;
L_p – дифузійна довжина неосновних носіїв;
N_D – концентрація донорів в епітаксійному шарі;
$n_{\text{підкл.}}$ – концентрація носіїв у підкладці;
q – заряд електрона;
R_c – опір контакту;
R_p – шунтуючий опір;
R_s – опір розтікання;
T – температура оточуючого середовища;
$t_{\text{відп}}$ – температура відпалу;
V – прикладена напруга;
η – фактор неідеальності;
τ_r – час релаксації «гарячих» електронів у металі;
τ_s – стала часу;
$\tau_{\text{відп}}$ – час відпалу;
φ_b – висота бар'єра.

ВСТУП

Арсенід галію привертає увагу дослідників як перспективний матеріал для мікро- та наноелектроніки [1–5]. Різке збільшення долі електронних приладів НВЧ діапазону [1], які виготовляють на основі електронного арсе-

© Дмитриев В. С. 2018
DOI 10.15588/1607-3274-2018-2-1

ніду галію, пов'язане з бурхливим розвитком мобільного зв'язку, використанням високошвидких безпровідних телекомунікаційних засобів та інше.

Срібло у порівнянні з золотом має більшу тепло- та електропровідність, відносно невеликий коефіцієнт дифузії в арсенід галію, що дозволяє зменшити товщину перехідного шару. Перехід до металізації на основі срібла повинен підвищити технічні характеристики виробів на n-GaAs.

До цього часу існують розбіжності в питаннях встановлення причин відхилення ВАХ від ідеальних [6–10]. Припускається [6–10], що реальні вольт-амперні характеристики та інжекційні властивості гетеропереходів метал-напівпровідників залежать від технологічних особливостей їх виготовлення. При виготовленні гетеропереходів до GaAs необхідно враховувати роль поверхневих станів у механізмі випрямлення струму, тому що з ними пов'язано зниження щільності струму насичення. Відсутність у GaAs стабільного природного окислу, а також чутливість його поверхні до впливу різних хімічних речовин, що використовують у технологічних процесах, потребують розробки нових підходів до технологічних процесів їх виготовлення. Серед найбільш ефективних засобів управління структурою межі розділу є відпал тонкоплівкових структур, механізм якого до теперішнього часу ще до кінця не вивчений і не розкритий. Тому, з нашої точки зору, дослідження впливу режимів термообробки на вольт-амперну характеристику і параметри гетеропереходів Ag/n-n$^+$GaAs виділяється у самостійну наукову проблему, рішення якої може привести до суттєвого уточнення уявлень про процеси у гетеропереходах метал-GaAs та дозволить покращити характеристики та параметри НВЧ приладів з бар'єрами Шотткі на GaAs.

1 ПОСТАНОВКА ЗАВДАННЯ

Одними з основних якісних показників гетеропереходів метал-напівпровідник є вольт-амперна характеристика, висота бар'єру φ_b, фактор неідеальності η, коефіцієнт інжекції γ [11], значення яких значною мірою залежать від технології виготовлення, зокрема від режимів термообробки контактів метал-напівпровідник. Точність визначення висоти потенційного бар'єру та фактору неідеальності залежать не тільки від точності вимірювання струму та напруги, але й від методу їх визначення. Відомо також [12], що на характеристики діодів з бар'єрами Шотткі при більших значеннях струму починають істотно впливати неосновні носії.

Постановка задачі: розробити режими термообробки для створення гетеропереходів Ag/n-n$^+$GaAs з покращеними характеристиками та провести експериментальні дослідження.

Метою статті є встановлення впливу термічної обробки на реальні вольт-амперні характеристики та величину коефіцієнту інжекції гетеропереходу Ag/n-n$^+$GaAs.

2 ЛІТЕРАТУРНИЙ ОГЛЯД

Реальні епітаксійні плівки, які використовують в основному для виготовлення випрямляючого контакту метал-напівпровідник, мають нерівномірний розподіл домішок, виникаючий в процесі вирощування плівки. Це призводить до появи механічної напруги та може привести до виникнення дислокацій, дефектів упаковки чи тріщин. Обробка поверхні напівпровідника [13–15] для усунення різних дефектів, видалення чужорідних домішок, отримання мінімального мікрорельєфу призводить до непланарності межі розділу метал-напівпровідник. Електрофізичні параметри реальних поверхонь напівпровідників після обробки існуючими способами стають неоднорідними уздовж поверхні [16].

Реальні характеристики та параметри інжектуючих бар'єрних переходів срібло-арсенід галію залежать від режимів їх термообробки. У відомих роботах [13–15] запропоновані режими хімічної та термічної обробки, при яких значення висоти бар'єру Шотткі для контактів Ag/GaAs знаходяться в інтервалі 0,82 В...0,95 В. Для покращення цих параметрів необхідно розробити нові способи створення контактів.

3 МАТЕРІАЛИ ТА МЕТОДИ

Для дослідження використовувався n-n$^+$GaAs(111)B епітаксійний монокристалічний, $d_{еш}$=2мкм, N_D=2·10^{16} см$^{-3}$, $n_{підкл}$=10^{18} см$^{-3}$.

GaAs-підкладку знежирювали в суміші толуолу і метилового спирту (1:2), обробляли у сірчанокислому травнику H_2SO_4:H_2O_2:H_2O=3:1:1 з подальшою витримкою у діоксиянтарній кислоті HOOC-CH(OH)-CH(OH)-COOH продовж 1,2·10^3 секунд.

За результатами попередніх досліджень [10–13] в якості бар'єрного матеріалу обрано срібло високої чистоти, яке утворює з GaAs сполуки, що не змінюються в робочому діапазоні температур. Контакти виготовляли методом вакуумного випаровування. Для вимірювання вольт-амперних характеристик бар'єрного переходу метал-напівпровідник використовувався чотирьохзондовий метод [17].

При аналізі ВАХ тонкоплівкових структур з бар'єром Шотткі Ag/n-n$^+$GaAs припускалося, що в початкових структурах механізм переносу носіїв заряду при концентрації вільних носіїв в GaAs N_D=2· 10^{16} см$^{-3}$ визначається термоелектронною емісією (Т-емісією). Rhoderick E. H., Williams R. H. [18] представили вольт-амперну характеристику з урахуванням параметрів, які описують стаціонарну ВАХ (висоти бар'єру Шотткі φ_b, фактору неідеальності η, послідовного R_s та шунтуючого R_p опорів), у вигляді:

$$I = I_0 \exp(qV_{pn}/\eta kT)[1-\exp(-qV/kT)] + V_{pn}/R_p \,, \quad (1)$$

де $I_0 = A^{**}T^2 S\exp(-q\varphi_b/kT)$ – струм насичення, $V_{pn} = V - IR_S$ – напруга на переході метал-напівпровідник.

З рівняння (1) були визначені умови [18, 19], при яких: внесок R_s та R_p дуже малий (менше 1%) і ними можна знехтувати. Також обумовлені обмеження, коли внесок бар'єра Шотткі на фоні цих опорів значний (для досягнення похибки менше 1% на фоні похибки вимірювання) [18, 19]. Якщо на ВАХ відсутні ділянки, обмежені цими умовами, то вона не дає інформації про коефіцієнт неідеальності та висоту бар'єра.

Найбільш простим у реалізації є розрахунок параметрів ВАХ за методом, запропонованим Rhoderick E. H. у роботі [18], де струм крізь бар'єр Шотткі описаний виразом:

$$I = I_0 \exp(qV/\eta kT)[1-\exp(-qV/kT)] \,. \quad (2)$$

Після побудови ВАХ у координатах $\ln(I)/[1-\exp(-qV/kT)] \propto V$, виконують апроксимацію

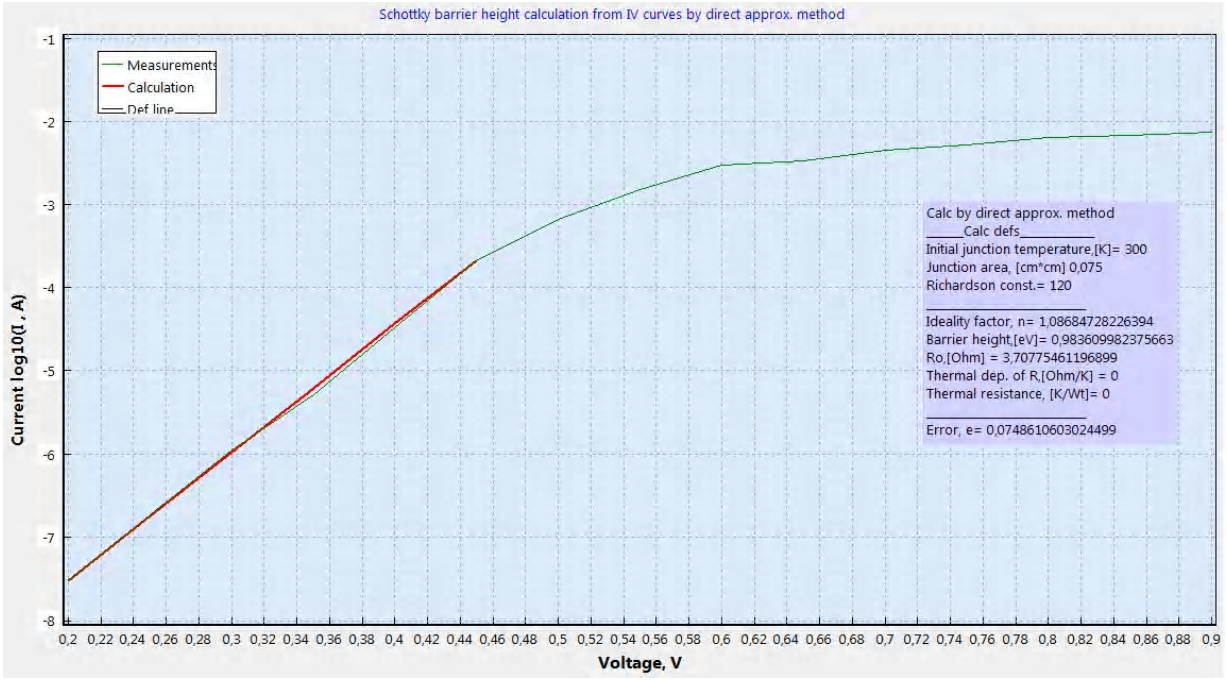
прямої лінійної ділянки $y=a+bx$, отримують коефіцієнти $b=\ln(I_0)$ та $a=q/\eta kT$, звідки розраховують φ_b и η, але при цьому не враховується послідовний опір R_s, що може привести до додаткових помилок при визначенні ділянки ВАХ, де цим впливом нехтують.

У [20–22] для розрахунку параметрів БШ запропоновано метод прямої апроксимації усієї довжини ВАХ (the direct approximation method), описаний формулою (1), недоліком якого є складні розрахунки.

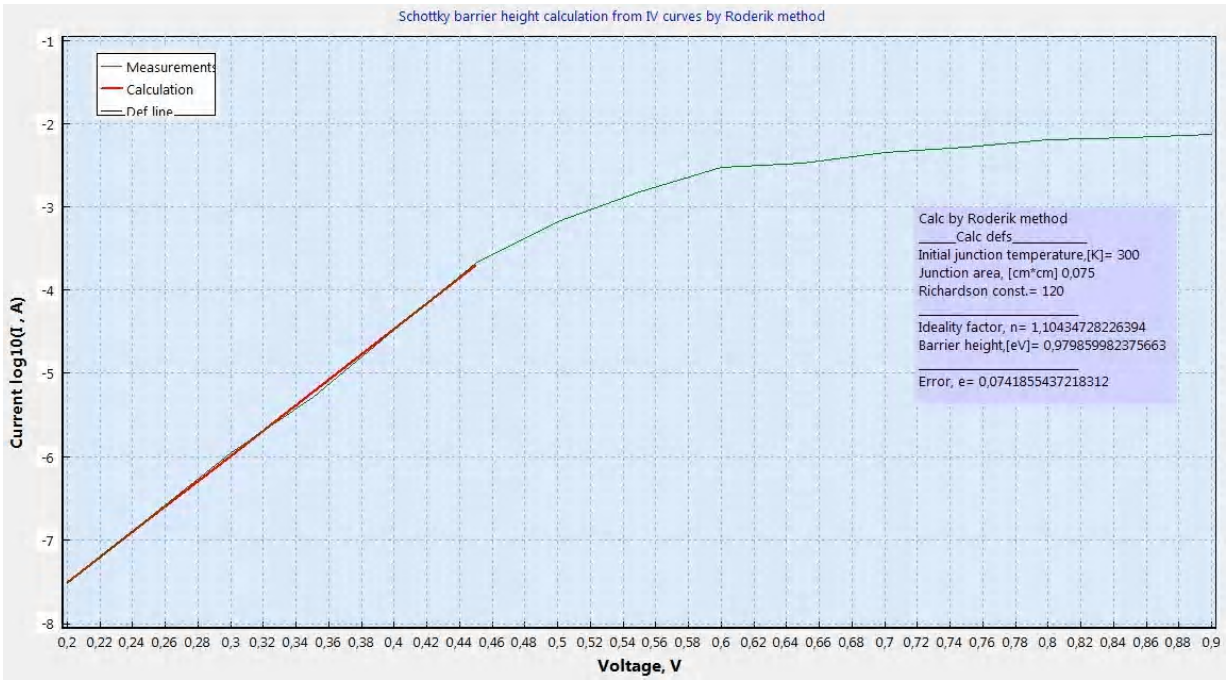
4 ЕКСПЕРИМЕНТИ

При нанесенні на GaAs-підкладку плівки срібла методом вакуумного випаровування за рахунок теплоти конденсації і можливих хімічних реакцій відбувається руйнування гратки в приповерхневому шарі напівпровідника ще до формування суцільної металевої плівки. Наслідки цієї стадії значною мірою визначаються умовами приготування контакту і режимами термообробок. Досліджували вольт-амперні характеристики бар'єрних переходів Ag/n-n$^+$GaAs, виготовлених у вакуумі методом термічного випаровування при температурах відпалу $t_{відп}$=703...853 К та часі відпалу $\tau_{відп}$ =6· 10^2 секунд.

На рис. 1 представлені експериментальні ВАХ контактів Ag/n-n$^+$GaAs, виготовлених при температурному режимі $t_{відп}$=803 К, $\tau_{відп}$ =6·10^2 секунд, до яких проведе-



а



б

Рисунок 1 – Прямі вольт-амперні характеристики бар'єрних переходів Ag/n-n$^+$GaAs

ний розрахунок параметрів з ВАХ (табл. 1) за методом direct approximation (рис. 1а) та методом Rhoderick E.H. (рис. 1б) за допомогою програми IVbarrierCalc2 [19].

5 РЕЗУЛЬТАТИ

Результати розрахунків основних параметрів гетеропереходів Ag/n-n$^+$GaAs, виготовлених при різних температурних режимах, наведені у табл. 1.

Встановлено, що підвищення температури відпалу до 803 К дає найбільше значення висоти бар'єру φ_b для гетеропереходу Ag/n-n$^+$GaAs.

У табл. 2 представлені значення коефіцієнта інжекції γ, розраховані для гетеропереходів з різною висотою бар'єру Шотткі за формулою [11]:

$$\gamma = \frac{qn_i^2 D_p}{N_D L_p A^{**} T^2 \exp(-q\varphi_b / kt)}.$$

6 ОБГОВОРЕННЯ

Порівняльний аналіз отриманих значень фактора неідеальності η (табл. 1), розрахованих двома різними методами, показав, що для бар'єрних переходів на основі срібла до GaAs з N_D=10^{16} см$^{-3}$,' створених при рекомендованому способі термічної обробки ($t_{відп}$=803 К, $\tau_{відп}$=6·10^2 секунд, φ_b =0,98 В), найменше значення η=1,104 отримано у методі розрахунку direct approximation.

У ряді першоджерел [11,18] зазначено, що фактор неідеальності при низькому рівні легування (N_D=10^{15} см$^{-3}$) та температурі 300 К близький до 1, але при зростанні рівня легування (N_D=10^{16} см$^{-3}$) відмінність η від одиниці стає суттєвою.

Дослідження вказують, що для визначення висоти потенційного бар'єру при невеликій протяжності експонентної ділянки ВАХ метод direct approximation є найбільш точним, оскільки він враховує послідовний опір і ділянку ВАХ при $V < kT/q$.

Нелінійний характер залежності φ_b від прикладеної напруги з'являється у області зміщень, обмеженої відрізками між 0,44...074 В (рис. 1), що призводить відповідно до зміни фактора неідеальності η.

Для гетеропереходів Ag/n-n$^+$GaAs з φ_b – 0,78...0,98 В коефіцієнти інжекції γ змінюються у діапазоні 5,28·10^{-12}…1,2·10^{-8}. Розраховані значення коефіцієнта інжекції мають дуже малі значення, тому інжекцією дірок можна знехтувати.

ВИСНОВКИ

Встановлено, що підвищення температури відпалу до 803 К дає найбільше значення висоти бар'єру φ_b для гетеропереходу Ag/n-n$^+$GaAs.

Коефіцієнти інжекції γ для бар'єрів Шотткі Ag/n-n$^+$GaAs (N_D=10^{16} см$^{-3}$) мають дуже малі значення γ, тому інжекцією дірок можна знехтувати.

При рекомендованому способі термічної обробки бар'єрних переходів Ag/n-n$^+$GaAs ($t_{відп}$=803 К, $\tau_{відп}$ = =6·10^2 секунд, φ_b =0,98 В) найменше значення η =1,087 отримано за методом розрахунку direct approximation.

Нелінійний характер залежності φ_b від прикладеної напруги з'являється в області зміщень, обмеженій відрізками між 0,44...074 В, що призводить відповідно й до зміни фактора неідеальності η.

При невеликій протяжності експонентної ділянки ВАХ для визначення висоти потенційного бар'єру метод direct approximation є найбільш точним, оскільки він враховує послідовний опір і ділянку ВАХ при $V < kT/q$.

СПИСОК ЛІТЕРАТУРИ

1. Белоус А. И. Космическая электроника : в 2 кн. Кн. 2 / А. И. Белоус, В. А. Солодуха, С. В. Шведов. – М. : Техносфера, 2015. – 489 с.
2. Малошумящие арсенид-галлиевые усилители при воздействии электромагнитных помех повышенных интенсивностей [Электронный ресурс] / [С. В. Платонов, Н. В. Пермяков, Б. И. Селезнев и др.]. // Вестник Новгородского государственного университета. – 2012. – № 67. – С. 29–32. – Режим доступа к ресурсу: http://www.novsu.ru/file/1010219
3. Контакты Шоттки к высокоомным эпитаксиальным слоям GaAs для детекторов частиц и квантов [Электронный ресурс] /

Таблиця 1 – Результати розрахунку параметрів гетеропереходів Ag/n-n$^+$GaAs

$t_{відп}$, К	Метод розрахунку	η	φ_b, В	R_s, Ом	e
703	*direct approx.*	1,254429775274	0,77321028599789	0,219689495861	0,28632938336540
	Rhoderick	1,203635086926	0,78356534807517	–	0,09680402395094
753	*direct approx.*	1,248802105905	0,80498828767418	0,758142769336	0,31679567002442
	Rhoderick	1,300052105905	0,80048828767418	–	0,44193998683910
803	*direct approx.*	1,086847282263	0,98360998237566	3,707754611968	0,07486106030244
	Rhoderick	1,104347282263	0,97985998233756	–	0,07418554372183
853	*direct approx.*	1,328157574655	0,84888672517418	1,025553047657	0,30231278123006
	Rhoderick	1,339593791431	0,84663494074446	–	0,2906834246612

Таблиця 2 – Вплив φ_b на γ для гетеропереходів Ag/n-n$^+$GaAs

φ_b, В	0,78	0,82	0,86	0,9	0,94	0,98
γ	5,28·10^{-12}	2,48·10^{-11}	1,16·10^{-10}	5,45·10^{-10}	2,56·10^{-9}	1,2·10^{-8}

[Г. И. Кольцов, С. И. Диденко, А. В. Черных и др.] // Физика и техника полупроводников. – 2012. – № 8. – С. 1088–1093. – Режим доступа к ресурсу: http://journals.ioffe.ru/articles/viewPDF/7776

4. Дмитриев В. С. Технологические особенности изготовления усилителя бегущей волны / В. С. Дмитриев, Е. Я. Швец // Сборник материалов 10-й международной молодежной научно-технической конференции «Современные проблемы радиотехники и телекоммуникаций «РТ-2014» / В. С. Дмитриев, Е. Я. Швец. – Севастополь : СевНТУ, 2014. – С. 158. (ISBN 978-617-612-072).
5. Дмитрієв В.С. Інжектуючі бар'єрні переходи на основі арсеніду галія для приладів НВЧ діапазону/ В.С. Дмитрієв // Перспективні технології та прилади. – 2017. – Вип. 10 (1). – С. 50–53, ISSN 2313-5352
6. Herná ndez M. P. Barrier height determination in homogeneous nonideal Schottky contacts / M. P. Herná ndez, C. F. Alonso, J. L. Peñ a // Journal of Physics D: Applied Physics. – 2001. – Vol. 34, №8. – P. 1157–1162.
7. Торхов Н. А. Природа прямых и обратных токов насыщения в контактах металл-полупроводник с барьером Шоттки / Н. А. Торхов // Физика и техника полупроводников. – 2010. – Т. 44, №6. – С. 767–774.
8. Chand S. Origin of non-ideal current-voltage characteristics of metal-semiconductor contact: A numerical study [Електронний ресурс] / S. Chand, L. Kamar. // Indian Journal of Engineering & Materials Sciences. – 2000. – Vol. 7, № 5–6. – P. 268–273. – Режим доступу до ресурсу: http://nopr.niscair.res.in/bitstream/123456789/24409/1/IJEMS%207%285-6%29%20268-273.pdf.
9. Зависимость проводимости от толщины активной области в тонкопленочных диодах Шоттки на GaAs [Електронний ресурс] / [С. А. Зуев, Г. В. Килесса, Э. Э. Асанов и др.]. // Физика и техника полупроводников. – 2016. – Т. 50, № 6. – С. 825–829. – Режим доступу до ресурсу: http://journals.ioffe.ru/articles/viewPDF/43212
10. Божков В. Г. Влияние нелинейной зависимости высоты барьера от смещения на измеряемые параметры контактов с барьером Шоттки [Електронний ресурс] / В. Г. Божков, А. В. Шмаргунов // Материалы 21-й Межд. Крымской конф. «СВЧ техника и телекоммуникационные технологии» (КрыМиКо – 2011). Севастополь, 10–14 сентября 2011. – Севастополь : СевНТУ. – 2011. – С. 246–247. – Режим доступу до ресурсу: http://naukainform.kpi.ua/CriMiCo/Crimico/2011/246_247.pdf
11. Sze S. M. Physics of Semiconductor Devices, 3rd Edition / S. M. Sze, K. K. Ng. – Hoboken : A John Wiley & Sons, Inc., 2007. – 815 p.
12. Dmitriev V. S. The effect of the injected minority carriers on the parameters of Schottky barrier diodes / V. S. Dmitriev, E. Y. Shvets // Зб. тез доповідей міжнар. наук.-техн. конф. «Фундаментальні та прикладні проблеми сучасних технологій». – Тернопіль : ТНТУ, 2015. – С. 10–11.
13. On the evaluation of Schottky barrier diode parameters of Pd, Au and Ag/n-GaAs [Электронный ресурс] / P. Jayavel, J. Kumar, P. Ramasam, R. Premanand. // Indian Journal of Engineering and Materials Sciences. – 2001. – Vol. 7, № 5–6. – P. 340–343. – Режим доступа к ресурсу: http://nopr.niscair.res.in/bitstream/123456789/24425/1/IJEMS%207%285-6%29%20340-343.pdf
14. Weber E. R. Structure and reliability of Metal contacts to GaAs / E. R. Weber, J. Washburn. – Berkeley, CA: Office of naval research AD-A259 184, Department of Materials Science and Mineral Engineering University of California, 1991. – 41 p.
15. Пат. 95094 Україна, МПК H 01 L 29/47. Спосіб виготовлення контактів з бар'єром Шотки на арсеніді галію / Дмитрієв В. С., Дмитрієва Л. Б. ; заявник і патентовласник Запорізька державна інженерна академія. – № u201407081; заявл. 23.06.14; опубл. 10.12.14, Бюл. № 23/2014.
16. Дмитриев В. С. Развитие физической модели неоднородных контактов металл-полупроводник / В. С. Дмитриев // Materialy X mezinarodni vědecko-prakticka konference «Moderni vymoženosti vědy – 2014» / В. С. Дмитриев. – Praha: Publishing House «Education and Science», 2014. – Dil 39 «Technicke vědy». – С. 3–5. – (ISBN 978-966-8736-05-6).
17. Физические методы диагностики в микро- и наноэлектронике / [А. В. Беляев, Н. С. Болтовец, Е. Ф. Венгер и др. ; Под ред. проф. Беляева А. Е., проф. Конаковой Р. В.]. – Харьков : ИСМА, 2011. – 284 с.
18. Rhoderick E. H. Metal-Semiconductor Contacts / E. H. Rhoderick, R. H. Williams. – Oxford : Clarendon Press, 1988. – 252 p.
19. Methods for determination of Schottky barrier height from I–V curves / [Ya. Ya. Kudryk, V. V. Shynkarenko, V. S. Slipokurov et al] // CriMiCo'2014, September 7–13, 2014. – Sevastopol, Crimea. – P. 673–674.
20. Extraction of Schottky diode parameters including parallel conductance using a vertical optimization method [Електронний ресурс] / A. Ferhat-Hamida, Z. Ouennoughi, A. Hoffmann, R. Weiss. // Solid-State Electronics. – 2002. – Vol. 46, № 5. – P. 615–619. – Режим доступу до ресурсу: https://doi.org/10.1016/S0038-1101(01)00337-9
21. Karaboga N. The parameter extraction of the thermally annealed Schottky barrier diode using the modified artificial bee colony [Електронний ресурс] / N. Karaboga, S. Kockanat, H. Dogan // Applied Intelligence. – 2013. –Vol. 38, №3. – P. 279–288. – Режим доступу до ресурсу: https://link.springer.com/article/10.1007%2Fs10489-012-0372-x.
22. Direct extraction of semiconductor device parameters using lateral optimization method [Електронний ресурс] / [A. Ortiz-Conde, Y. Ma, J. Thomson, etc]. // Solid-State Electronics. – 1999. – Vol. 43, № 4. – P. 845–848. – Режим доступу до ресурсу: https://doi.org/10.1016/S0038-1101(99)00044-1

Стаття надійшла до редакції 22.09.2017.
Після доробки 28.12.2017.

Дмитриев В. С.
Заведующий учебно-научной лабораторией кафедры микроэлектронных информационных систем Запорожской государственной инженерной академии, Запорожье, Украина

ВЛИЯНИЕ ТЕРМИЧЕСКОЙ ОБРАБОТКИ НА ВОЛЬТ-АМПЕРНУЮ ХАРАКТЕРИСТИКУ И ИНЖЕКЦИОННЫЕ СВОЙСТВА ГЕТЕРОПЕРЕХОДОВ Ag/n-n$^+$GaAs

Актуальность. В настоящее время исследования и разработку гетеропереходов проводят в направлениях поиска новых композиций металл-арсенид галлия, разработки технологических режимов, которые смогли бы обеспечить воспроизводимость параметров СВЧ приборов с барьерами Шоттки.

Серебро по сравнению с золотом имеет большую тепло- и электропроводность, относительно небольшой коэффициент диффузии в арсенид галлия, позволяющий уменьшить толщину переходного слоя. Переход к металлизации на основе серебра должен повысить технические характеристики изделий. Разработка технологических режимов изготовления улучшенных гетеропереходов к GaAs на основе серебра является актуальной с научной и практической точки зрения.

Цель работы – установление влияния термической обработки на реальные вольт-амперные характеристики и величину коэффициента инжекции гетероперехода Ag/n-n$^+$GaAs.

Метод. Метод изготовления гетеропереходов – термическое испарение в вакууме. Расчет параметров гетероперехода методом вольт-амперных характеристик.

Результаты. Рекомендуется режим химической обработки GaAs-подложки. Исследовано влияние термической обработки на параметры и характеристики гетеропереходов Ag/n-n$^+$GaAs. Исследовано влияние высоты барьера на величину коэффициента инжекции. Рассмотрены и апробированы различные методы определения высоты барьера Шоттки и фактора неидеальности по вольт-амперным характеристикам.

Выводы. Установлено, что повышение температуры отжига до 803 К дает наибольшее значение высоты барьера φ_b для гетероперехода Ag/n-n$^+$GaAs с концентрацией доноров в эпитаксиальных слое N_D=2·10^{16} см$^{-3}$. Установлено, что наиболее точным методом определения параметров гетероперехода по вольт-амперной характеристике является метод direct approximation, учитывающий влияние последовательного сопротивления при определении фактора неидеальности и высоты барьера Шоттки. Рассчитанные коэффициенты инжекции γ для гетеропереходов с разной высотой барьера Шоттки, полученных при различных температурах отжига, имеют очень малые значения, поэтому инжекцией дырок можно пренебречь.

Ключевые слова: барьерный переход, вольт-амперная характеристика, высота барьера**,** фактор неидеальности, коэффициент инжекции.

Dmitriev V. S.

Head of the Educational and Scientific Laboratory of the Microelectronic Information Systems Department of the Zaporizhzhia State Engineering Academy, Zaporizhzhia, Ukraine

THE THERMAL PROCESSING INFLUENCE ON CURRENT-VOLTAGE CHARACTERISTIC AND INJECTION PROPERTIES OF Ag/n-n$^+$GaAs HETEROJUNCTIONS

Context. Nowadays the research and development of heterojunctions are carried out to search new metal-gallium arsenide compositions, the development of technological regimes which could ensure the microwave devices with Schottky barriers parameters reproducibility.

Comparing to gold, silver has a large thermal and electrical conductivity, a relatively small diffusion coefficient to gallium arsenide, which makes it possible to reduce the transition layer thickness. The transition to silver-based metallization should improve the products technical characteristics. The technological regimes development for manufacturing improved silver-based heterojunctions to GaAs is relevant from scientific and practical points of view.

Objective. The goal of the work is to determine the heat treatment influence on real current-voltage characteristics and the injection coefficient value of the Ag/n-n$^+$GaAs heterojunction.

Method. The heterojunctions manufacturing method is vacuum thermal evaporation. The calculation of the heterojunction parameters was made by the current-voltage characteristics method.

Results. The GaAs-substrate chemical treatment regime is recommended. The heat treatment effect on the Ag/n-n$^+$GaAs heterojunctions parameters and characteristics is studied. The barrier height influence on the injection coefficient value was investigated. Various methods for determining the Schottky barrier height and the nonideality factor for the current-voltage characteristics have been examined and tested.

Conclusions. It is established that the annealing temperature increase up to 803 K gives the highest barrier height for the Ag/n-n$^+$GaAs heterojunction with a donor density of N_D=2·10^{16} cm^{-3} into epitaxial layer. It has been established that the most accurate method for the heterojunction parameters determining using the current-voltage characteristic is the direct approximation method, which takes into account the effect of the series resistance in determining the nonideality factor and the Schottky barrier height. The calculated injection coefficients for heterojunctions with different Schottky barrier heights, obtained at different annealing temperatures, have very small values, therefore, hole injection can be neglected.

Keywords: barrier transition, current-voltage characteristic, barrier height, non-ideality factor, injection coefficient.

REFERENCES

1. Belous A. Y., Solodukha V. A., Shvedov S. V. Kosmycheskaya elektronyka : v 2 kn. Kn. 2. Moscow, Tekhnosfera, 2015, 489 p.
2. Platonov S. V., Permyakov N. V., Seleznev B. Y. i dr. Maloshumyashchye arsenyd-hallyevye usylytely pry vozdeystvyy elektromahnytnykh pomekh povyshennykh yntensyvnostey [Elektronny resurs], *Vestnyk Novhorodskoho hosudarstvennoho unyversyteta*, 2012, No. 67, pp. 29–32. Rezhym dostupa k resursu: http://www.novsu.ru/file/1010219
3. Kol'tsov H. Y., Dydenko S. Y., Cherny'kh A. V. i dr. Kontakty Shottky k vysokoomnym epytaksyal'nym sloyam GaAs dlya detektorov chastyts y kvantov [Elektronny resurs], *Fyzyka y tekhnyka poluprovodnykov*, 2012, No. 8, pp. 1088–1093. Rezhym dostupa k resursu: http://journals.ioffe.ru/articles/viewPDF/7776
4. Dmytryev V. S., Shvets E. Ya. Tekhnolohycheskye osobennosty yzhotovlenyya usylytelya behushchey volny, *Sbornyk materyalov 10-y mezhdunarodnoy molodezhnoy nauchno-tekhnycheskoy konferentsyy «Sovremenny'e problemy' radyotekhnyky y telekommunykatsyy «RT–2014»*. Sevastopol', SevNTU, 2014, P. 158. (ISBN 978-617-612-072).
5. Dmytriyev V. S. Inzhektuyuchi bar'yerni perekhody na osnovi arsenidu haliya dlya priladiv NVCH diapazonu, *Prohresivni tekhnolohiyi ta prylady,* 2017, Vyp. 10(1), pp. 50–53, ISSN 2313-5352
6. Hernández M. P., Alonso C. F., Peña J. L. Barrier height determination in homogeneous nonideal Schottky contacts, *Journal of Physics D: Applied Physics*, 2001, Vol. 34, No. 8, pp. 1157–1162.
7. Torkhov N. A. Priroda pryamykh i obratnykh tokov nasyshcheniya v kontaktakh metall-poluprovodnik s baryerom Shottki, *Fizika i tekhnika poluprovodnikov,* 2010, Vol. 44, No. 6, pp. 767–774.
8. Chand S., Kamar L. Origin of non-ideal current-voltage characteristics of metal-semiconductor contact: A numerical study [Електронний ресурс], *Indian Journal of Engineering & Materials Sciences,* 2000, Vol. 7, No. 5–6, pp. 268–273. Режим доступу до ресурсу: http://nopr.niscair.res.in/bitstream/123456789/24409/1/IJEMS%207%285-6%29%20268-273.pdf.
9. Zuyev S. A., Kilessa G. V., Asanov E. E. i dr. Zavisimost provodimosti ot tolshchiny aktivnoy oblasti v tonkoplenochnykh diodakh Shottki na GaAs [Elektronny resurs], *Fizika i tekhnika poluprovodnikov*, 2016, Vol. 50, No. 6, pp. 825–829. Rezhim dostupu do resursu: http://journals.ioffe.ru/articles/viewPDF/43212
10. Bozhkov V. G., Shmargunov A. V. Vliyaniye nelineynoy zavisimosti vysoty baryera ot smeshcheniya na izmeryaemye

parametry kontaktov s baryerom Shottki [Elektronny resurs], *Materialy 21-y Mezhd. Krymskoy konf. "SVCh tekhnika i telekommunikatsionnye tekhnologii" (KryMiKo – 2011). Sevastopol, 10–14 sentyabrya 2011*. Sevastopol, SevNTU, 2011, pp. 246–247. Rezhim dostupu do resursu: http://naukainform.kpi.ua/CriMiCo/Crimico/2011/246_247.pdf

11. Sze S. M., Ng K. K. Physics of Semiconductor Devices, 3rd Edition. Hoboken, A John Wiley & Sons, Inc., 2007, 815 p.
12. Dmitriev V. S., Shvets E. Y. The effect of the injected minority carriers on the parameters of Schottky barrier diodes, *Zb. tez dopovide- mizhnar. nauk.-tekhn. konf. «Fundamental'ni ta prikladni problemi suchasnikh tekhnolohi-»*. Ternopil, TNTU, 2015, pp. 10–11.
13. Jayavel P., Kumar J., Ramasam P., Premanand R. On the evaluation of Schottky barrier diode parameters of Pd, Au and Ag/n-GaAs [Elektronny resurs], *Indian Journal of Engineering and Materials Sciences*, 2001, Vol. 7, No. 5–6, pp. 340–343. Rezhim dostupu do resursu: http://nopr.niscair.res.in/bitstream/123456789/24425/1/IJEMS%207%285-6%29%20340-343.pdf
14. Weber E. R., Washburn J. Structure and reliability of Metal contacts to GaAs. Berkeley, CA: Office of naval research AD-A259 184, Department of Materials Science and Mineral Engineering University of California, 1991, 41 p.
15. Dmstriyev V. S., Dmstriyeva L. B.; zayavnyk i patentovlasnsk Zaporiz'ka derzhavna inzhenerna akademiya. Pat. 95094 Ukrayina, MPK H 01 L 29/47. Sposib vihotovlennya kontaktiv z bar'yerom Shotky na arsenidi haliyu. № u201407081; zayavl. 23.06.14; opubl. 10.12.14, Byul. № 23/2014.
16. Dmytryev V. S. Razvytye fyzycheskoy modely neodnorodnыkh kontaktov metall-poluprovodnyk, *Materialy X mezinarodni v ědecko-prakticka konference «Moderni vymoženosti v ědy-2014»*. Praha, Publishing House «Education and Science», 2014, Dil 39 «Technicke vědy», pp. 3–5. (ISBN 978-966-8736-05-6).
17. Belyaev A. V., Boltovets N. S., Venger Ye. F. i dr. Pod red. prof. Belyaeva A. Ye., prof. Konakovoy R. V. Fizicheskiye metody diagnostiki v mikro- i nanoelektronike. Kharkov, ISMA, 2011, 284 p.
18. Rhoderick E. H., Williams R. H. Metal-Semiconductor Contacts. Oxford, Clarendon Press, 1988, 252 p.
19. Kudryk Ya. Ya., Shynkarenko V. V., Slipokurov V. S., Bigun R. I., Kudryk Ya. Ya. Methods for determination of Schottky barrier height from I–V curves, *CriMiCo'2014, September 7–13*, 2014. Sevastopol, Crimea, pp. 673–674.
20. Ferhat-Hamida A., Ouennoughi Z., Hoffmann A., Weiss R. Extraction of Schottky diode parameters including parallel conductance using a vertical optimization method [Elektronny resurs], *Solid-State Electronics*, 2002, Vol. 46, No. 5, pp. 615–619. Rezhim dostupu do resursu: https://doi.org/10.1016/S0038-1101(01)00337-9
21. Karaboga N., Kockanat S., Dogan H. The parameter extraction of the thermally annealed Schottky barrier diode using the modified artificial bee colony [Elektronny resurs], *Applied Intelligence*, 2013, Vol. 38, No. 3, pp. 279–288. Rezhim dostupu do resursu: https://link.springer.com/article/10.1007%2Fs10489-012-0372-x.
22. Ortiz-Conde A., Ma Y., Thomson J. etc Direct extraction of semiconductor device parameters using lateral optimization method [Elektronny resurs], *Solid-State Electronics*, 1999, Vol. 43, No. 4, pp. 845–848. Rezhim dostupu do resursu: https://doi.org/10.1016/S0038-1101(99)00044-1

УДК 004.052

Жиров Г. Б.[1], Лєнков Є. С.[2]

[1]Канд. техн. наук, старший науковий співробітник, провідний науковий співробітник науково-дослідного центру Військового інституту Київського національного університету імені Тараса Шевченка, Київ, Україна
[2]Канд. техн. наук, старший науковий співробітник наукового центру Військового інституту телекомунікацій та інформатизації, Київ, Україна

АЛГОРИТМИ МОДЕЛЮВАННЯ ПРОЦЕСІВ ТЕХНІЧНОГО ОБСЛУГОВУВАННЯ ЗА СТАНОМ СКЛАДНИХ РАДІОЕЛЕКТРОННИХ ОБ'ЄКТІВ

Актуальність. Розробка складних радіоелектронних об'єктів та проведення їх технічного обслуговування неможлива без розробки моделей та програмних засобів, які б дані процеси оптимізували. Для цього у роботі розроблені відповідні алгоритми, які є частиною програмних засобів, що є актуальною задачею.

Мета. Метою даної роботи є розробка підходів щодо підвищення показників надійності об'єкта та зменшення вартості його експлуатації на основі використання імітаційного статистичного моделювання.

Метод. Метод дослідження базується на ідеї, що технічний стан об'єкта визначається технічним станом його елементів, а для визначення технічного стану елемента використовується поняття визначального параметра. Значення параметра можна контролювати засобами діагностування. В якості моделей відмов елементів використовується *DN*-розподіл.

Результати. Розроблені алгоритми являють складову частину імітаційної статистичної моделі об'єкта. Проведення моделювання дозволяє визначити оптимальні конструктивні характеристики об'єкта та параметри проведення його технічного обслуговування, а критеріями виступають максимізація показників надійності та мінімізація вартості експлуатації об'єкта.

Висновки. Наукова новизна роботи полягає в удосконаленні алгоритмів проведення технічного обслуговування за станом, які на відміну від існуючих використовують прогнозний параметр середньої швидкості деградації визначального параметра елемента на основі моделі трипараметричного експоненціального згладжування.

Практичне значення полягає у розробці алгоритмів, що покладені в основу програмного забезпечення, яке реалізує процес визначення показників надійності та вартості експлуатації об'єкта радіоелектронної техніки від конструктивного складу самого об'єкта та від параметрів проведення технічного обслуговування.

Ключові слова: технічне обслуговування за станом, адаптивне технічне обслуговування, визначальний параметр.

НОМЕНКЛАТУРА

BBD – двійкова схема рішень;

FORM – метод першого порядку другого моменту;

MEMS – метод моделювання структурної надійності;

MMDD – багатостадійна багатозначна схема рішень;

Modelir_ATO – процедура алгоритму моделювання адаптивного ТОС;

Modelir_TOC – процедура алгоритму моделювання ТОС з постійною періодичністю контролю;

РЕТ – радіоелектронна техніка;

ТО – технічне обслуговування;

ТОіР – технічне обслуговування та ремонт;

ТОР – технічне обслуговування за ресурсом;

ТОС – технічне обслуговування за станом;

ТС – технічний стан;

$a_i(t_k)$ – поточне значення середньої швидкості деградації i-го параметра;

$\widetilde{a}_i(t_k)$ – прогнозне значення швидкості деградації i-го параметра;

$C_{ктс}$ – вартість операції контролю ТС;

Count – найбільший індекс списку ListEo;

cto – змінна, в якій накопичується значення вартості ТО;

DN – дифузний немонотонний розподіл;

E – покажчик на дані, що стосуються поточного елемента e_i;

E_o – попередньо визначена множина елементів, що відмовляють;

e_i – i-й елемент об'єкта;

i – параметр циклу;

$L(t_{k-1})$ – згладжена величина на час контролю t_{k-1};

ListEo – список елементів множини E_o;

n – кількість елементів, що обслуговуються;

N_I – кількість реалізацій моделювання;

$R_i(t)$ – величина залишкового ресурсу в момент часу t;

$\overline{R}_i(t)$ – величина середнього залишкового ресурсу в довільний момент часу t;

R_{0i} – випадкова реалізація повного ресурсу i-го елемента;

$S(t_{k-1})$ – коефіцієнт сезонності на час контролю t_{k-1};

$S(t_{k-1-s})$ – коефіцієнт сезонності за цей же період в попередньому сезоні;

t – поточний час;

t_i – запланований (модельний) час відмови i-го елемента;

t_{0i} – час останнього оновлення i-го елемента;

T_k – періодичність проведення контролю;

t_k – час контролю;

Tto_p – змінна, в якій накопичується значення часу проведення наступного ТО;

tto – змінна, в якій накопичується значення тривалості ТО;

$T'_{ср}$ – найменше значення середнього наробітку до відмови;

© Жиров Г. Б., Лєнков Є. С., 2018
DOI 10.15588/1607-3274-2018-2-2

$\widetilde{T}_{\text{ср}\,i}(t)$ – уточнене значення середнього наробітку до відмови *i*-го елемента;

$T_{\text{ср}\,i}$ – середній наробіток до відмови *i*-го елемента;

$\widetilde{T}_{\text{то}}(t)$ – уточнене значення до проведення наступного ТО;

$t_{\text{то}}$ – запланований час наступного ТО;

$T_{\text{е}}$ – час експлуатації об'єкта;

$\widetilde{T}_0$ – модельоване значення середнього наробітку на відмову;

$u_i(t)$ – нормоване значення визначального параметра *i*-го елемента;

u_{TOi} – нормоване граничне значення визначального параметра *i*-го елемента;

α – коефіцієнт згладжування ряду;

β – коефіцієнт згладжування тренда;

γ – коефіцієнт упередження;

Δt – інтервали дискретності накопичення статистики;

λ_i – інтенсивність відмов *i*-го елемента;

$\tau_{\text{то а}}$ – адміністративний час на проведення ТО;

$\tau_{\text{ктс}}$ – середній час контролю ТС;

χ – коефіцієнт згладжування сезонності.

ВСТУП

Експлуатація складних сучасних технічних об'єктів, таких як РЕТ, неможлива без якісно налагодженої системи ТОіР. Така система необхідна для забезпечення необхідного рівня надійності даних об'єктів. Необхідність проведення технічного обслуговування полягає у своєчасній заміні елементів, які знаходяться в перед відмовному стані, а ремонт проводиться з метою відновлення справного або працездатного стану об'єкта, а також відновлення ресурсу всього об'єкта або його частини [1, 2].

Показники надійності та вартості експлуатації об'єктів РЕТ залежать, як від властивостей безвідмовності та ремонтопридатності самих об'єктів, так і від параметрів процесу проведення ТОіР. Для того, щоб мати можливість оптимізувати характеристики об'єкта (на етапі його розробки), а також параметри системи його ТОіР, необхідно попередньо побудувати моделі, які встановлюють зв'язок між параметрами об'єкта, системами та прогнозованими показниками надійності та вартості експлуатації об'єкта РЕТ.

На сьогоднішній день, залежно від критерію, який використовується при визначенні термінів проведення ТО, існують дві основні стратегії організації проведення ТО: ТО за ресурсом (ТОР) і ТО за станом (ТОС). При проведенні ТОР здійснюється контроль за поточним ресурсом об'єкта і в разі, якщо залишковий ресурс знижується до деякого заданого граничного значення, проводиться технічне обслуговування. При проведенні ТОС здійснюється контроль за поточним технічним станом об'єкта, а ТО проводиться у разі, якщо технічний стан об'єкта погіршується до деякого заданого неприпустимого рівня. У свою чергу, при проведенні ТОС можна виділити 2 стратегії: ТОС з постійною періодичністю контролю та ТОС зі змінною періодичністю контролю (адаптивне ТОС).

1 ПОСТАНОВКА ЗАДАЧІ

На сьогоднішній час, існує загальна проблема оптимізації характеристик об'єкта та параметрів системи ТОіР з метою підвищення показників надійності об'єкта та зменшення вартості його експлуатації. Визначати показники надійності складних технічних об'єктів та їх взаємозв'язок з параметрами системи ТОіР на основі статистичних даних є складною задачею. Значно простіше та ефективніше, для встановлення зв'язків між даними параметрами і прогнозованими показниками надійності і вартості, проводити імітаційне статистичне моделювання. Для проведення моделювання необхідні: модель безвідмовності об'єкта та моделі (алгоритми) процесу проведення технічного обслуговування та ремонту.

Таким чином, у статті вирішується актуальна наукова задача щодо будови алгоритмів моделювання процесу проведення технічного обслуговування за станом для різних стратегій. Алгоритми являють собою складові частини загальної імітаційної статистичної моделі складного об'єкта РЕТ та параметрів системи ТОіР.

2 ОГЛЯД ЛІТЕРАТУРИ

При проектуванні складного технічного об'єкта важливо мати різноманітні інструменти (моделі, методики та алгоритми), а також програмне забезпечення для прогнозування показників надійності і вартості експлуатації в залежності від конструктивних особливостей самого об'єкта, а також від параметрів його експлуатації, параметрів технічного обслуговування і ремонту. Питаннями розрахунку і прогнозування надійності, а також створення моделей (в тому числі і імітаційних) займалася велика кількість відомих вчених в тому числі: Креденцер Б. П., Барзилович Є. Ю., Кокс Д., Сміт В ., Каштанов В. О., Ушаков І. О., Барлоу Р., Прошан Ф. та ін.

В останні десятиліття, з метою поліпшення точності, були розроблені нові моделі і методи моделювання надійності складних технічних об'єктів і систем.

В [3] запропонований метод наслідування і оцінки надійності на основі динамічних байесовских мережах з використанням чисельного моделювання. Складні технічні об'єкти представляються типовими динамічними системами. При цьому, на думку авторів, даний метод долає обмеження аналітичного методу і методу багаторівневого синтезу та забезпечує ефективний засіб для оцінки надійності складних динамічних систем та об'єктів.

Недолік байесовских мереж полягає в тому, що вкрай складно отримати апостеріорний розподіл по усім можливим конфігураціям прихованих елементів мережі. Навчання мережі може бути простим, якщо б вдалося отримати достовірну вибірку по апостеріорному розподілу прихованих станів. Однак дана процедура є досить складною, в тому числі і в сенсі «ефекту пояснення», який полягає в тому, що дві причини, які є незалежними, можуть бути залежними для того, хто спостерігає певний ефект, який залежить від обох цих причин. Таким чином, щоб отримати апостеріорний розподіл, навіть у першому прихованому шарі мережі, потрібно інтегрувати по усім можливим конфігураціям прихованих змінних. А така процедура для складних мереж стає практично неможливою.

В [4] розроблено алгоритми для оцінки надійності систем з використанням байесовских мереж при експоненціальному збільшенні кількості інформації, яку необхідно зберегти по мірі збільшення кількості компонентів у системі. Недоліки, які притаманні методам, що описані у [3], так само притаманні і алгоритмам на основі байесовских мереж.

В [5] вивчалися фактори та методи розрахунку надійності складних систем з використанням методу динамічного моделювання. Завдяки симуляції роботи об'єкта та подальшого аналізу можуть бути отримані основні коефіцієнти надійності, однак даний підхід проблематично застосовувати до складних технічних систем і об'єктів.

В [6, 7] зроблений висновок, що традиційний аналіз дерева відмов не завжди достатній при аналізі складних систем. Для подолання даних обмежень пропонується застосовувати динамічні дерева відмов (DFT), а також різні методи до їх аналізу, у тому числі із застосуванням марківського підходу та моделювання за методом Монте-Карло. Перевагою даних методів є достатньо простий процес моделювання та низька обчислювальну складність, однак точність визначення показників надійності може бути недостатньо для вирішення практичних задач.

В [8] запропоновано методологію визначення оптимальних параметрів надійності (швидкості відмов та відновлень) компонентів для системи теплопостачання, які забезпечують необхідний рівень надійності мережі. Методологія вирішення зазначеної проблеми базується на методах теорії гідравлічних схем, показників надійності вузла та моделях випадкових процесів Маркова.

Переваги запропонованої методології полягають у спільній оптимізації надійності компонентів джерел та схем мереж, інтеграції процедур зменшення частот відмов та поліпшення швидкості відновлення, відсутності необхідності проводити ітеративні розрахунки при використанні середніх параметрів надійності компонентів, враховуючи необхідні рівні показників надійності. Недоліком запропонованого підходу є зосередження на визначенні комплексних показників надійності.

Також, на сьогоднішній час, відсутній вітчизняний програмний комплекс по розрахунку показників надійності складних радіоелектронних об'єктів з урахуванням різних стратегій технічного обслуговування і ремонту. Математичною основою такого програмного комплексу повинна бути імітаційна статистична модель самого об'єкта та системи TOiP, як єдиного цілого. Для вирішення наведеної загальної проблеми, необхідно вирішити задачу щодо будови алгоритмів моделювання процесу проведення технічного обслуговування для різних стратегій.

3 МАТЕРІАЛИ ТА МЕТОДИ

Принцип ТО за станом, реалізований в моделі на основі прийняття наступних припущень: для обслуговуваних елементів (E_o – визначена множина елементів, що відмовляють), завжди існують визначальні параметри, які характеризують їх технічний стан; випадковий процес деградації визначальних параметрів, має постійну середню швидкість. Найбільш близько зазначеним припущенням відповідає модель *DN*-розподілу випадкового наробітку до відмови.

Моделювання процесу ТОС з постійною періодичністю контролю. Сутність стратегії ТОС з постійною періодичністю контролю висвітлений у [1, 9]. Використання даної стратегії передбачає при технічному контролі проводити вимірювання визначальних параметрів усіх елементів, що потенційно підлягають обслуговуванню (E_0). Якщо виміряне значення визначального параметра *i*-го елемента $u_i(t)$ перевищує заздалегідь задане значення (значення рівня проведення ТО u_{TOi}), проводиться оновлення (заміна) *i*-го елемента. Кожну відмову елемента можна інтерпретувати як подію, що полягає у перевищенні поточного нормованого значення визначального параметра елемента $u_i(t)$, граничної межі. Якщо припустити, що середній ресурс елемента убуває в часі по лінійному закону, то величина середнього залишкового ресурсу в довільний момент часу t дорівнює $\overline{R}_i(t) = T_{\text{ср}\,i} - t$, де $T_{\text{ср}\,i}$ – середній наробіток до відмови *i*-го елемента.

При моделюванні заздалегідь відомі заплановані випадкові моменти часу відмов елементів та періодичність проведення контролю – T_k. Якщо відмова *i*-го елемента запланована на момент часу $t_i, (t_i > t)$, то його залишковий ресурс в момент часу t дорівнює:

$$R_i(t) = t_i - t, \tag{1}$$

де t_i – запланований (модельний) час відмови *i*-го елемента, t – поточний час.

Залишковому ресурсу $R_i(t)$ можна поставити у відповідність випадкове значення нормованого визначального параметра $u_i(t)$, величина якого визначається:

$$u_i(t) = 1 - \frac{R_i(t)}{R_{0i}}, \tag{2}$$

де R_{0i} – випадкова реалізація повного ресурсу *i*-го елемента. Значення величини R_{0i} визначається виразом:

$$R_{0i} = t_i - t_{0i}, \tag{3}$$

де t_{0i} – час останнього оновлення *i*-го елемента; t_i – запланований час відмови *i*-го елемента.

Після підстановки у вираз (2) формул по визначенню залишкового ресурсу *i*-го елемента $R_i(t)$ (формула 1) та повного ресурсу *i*-го елемента R_{0i} (формула 3), нормоване значення визначального параметра *i*-го елемента $u_i(t)$ можна отримати за виразом (4):

$$u_i(t) = \frac{t - t_{0i}}{t_i - t_{0i}}. \tag{4}$$

Рішення про необхідність проведення ТО *i*-го елемента приймається за критерієм – $u_i(t) \geq u_{TO\,i}$. З урахуванням (4) даний критерій представляється нерівністю:

$$t \geq t_{0i} + u_{TO\,i}(t_i - t_{0i}). \tag{5}$$

Таким чином, в процесі моделювання в момент часу t проведення контролю, необхідно для кожного i-го елемента, що обслуговується, перевірити виконання умови (5), якщо ця умова виконується, імітувати проведення ТО. Імітація ТО повинна включати: генерування нового випадкового значення часу відмови елемента t_i, а час останнього оновлення i-го елемента t_{0i} встановити рівним поточному часу t.

Алгоритм процесу моделювання технічного обслуговування за станом з постійною періодичністю контролю (процедура Modelir_TOC).

Структурна схема алгоритму моделювання процесу ТОС з постійною періодичністю контролю (процедура Modelir_TOC) зображена на рис. 1. Алгоритм виконується в поточний момент модельного часу t, що є часом контролю ТС об'єкта.

Оператор 1 виконує підготовчі операції, встановлює початкові значення змінних: i – параметр циклу, в якому буде здійснюватися перебір всіх елементів $e_i \in E_o$ (елементів списку ListEo); E_o – попередньо визначена множина елементів, що відмовляють; tto і cto – змінні, в яких накопичуються значення тривалості та вартості ТО (результуюча інформація процедури). Оператор 2 формує покажчик E на дані, що відносяться до поточного елемента e_i. Оператор 3 перевіряє, чи є елемент e_i таким, що обслуговується при проведенні ТО. Якщо елемент обслуговується, то виконуються оператори 4–7. Оператор 4 імітує вимір визначального параметра $u_i(t)$ відповідно до виразів (1–4). Оператор 5 перевіряє виконання критерію $u_i(t) \geq u_{TOi}$. Оператор 6 імітує виконання ТО (оновлення елемента e_i) та генерує наступний випадковий час до відмови елемента e_i.

Оператор 7 здійснює накопичення даних у змінних tto і cto. Оператори 8 і 9 забезпечують циклічне виконання операторів 4–7 для всіх обслуговуваних елементів зі списку ListEo. Як тільки параметр циклу i вийде за межі діапазону індексів елементів списку ListEo (i > ListEo.Count – 1), оператор 9 завершує виконання циклу. У змінних tto і cto, які є параметрами процедури, формуються значення тривалості та вартості виконаного технічного обслуговування.

Моделювання процесу ТОС з адаптивною періодичністю контролю. Формалізований опис процесу, що моделюється. Сутність стратегії ТОС з адаптивною періодичності контролю висвітлений у [1, 10, 11]. У поточний момент часу t визначається (вимірюється) середня швидкість деградації визначальних параметрів елементів за виразом:

$$a_i(t) = u_i(t)/(t - t_{0i}). \qquad (6)$$

За «виміряним» значенням $a_i(t)$ розраховуються прогнозні значення швидкості деградації $\tilde{a}_i(t)$, які після кожного вимірювання уточнюються за формулою двопараметричного експоненціального згладжування:

$$\tilde{a}_i(t_k) = \beta \tilde{a}_i(t_{k-1}) + (1-\beta) a_i(t_k), \qquad (7)$$

де $\tilde{a}_i(t_{k-1})$ – попереднє (отримане в момент часу t_{k-1}) значення прогнозної швидкості $\tilde{a}_i(t)$; $a_i(t_k)$ – поточне значення середньої швидкості деградації i-го параметра, виміряне в момент часу t_k; β – коефіцієнт згладжування тренда.

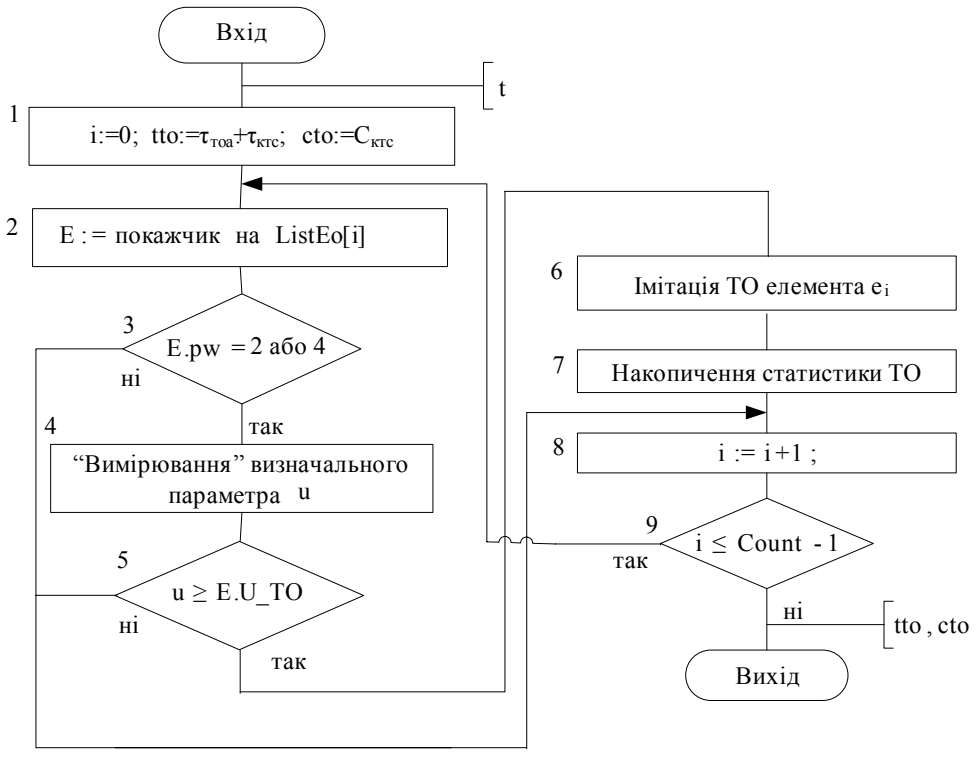


Рисунок 1 – Алгоритм процедури Modelir_TOC

В свою чергу, оцінку середньої швидкості деградації, при виборі моделі трипараметричного експоненціального згладжування, визначається за виразом:

$$\widetilde{a}_i(t_k) = [L(t_{k-1}) + T(t_{k-1})] \times S(t_{k-s});$$

$$L(t_{k-1}) = \alpha \cdot \frac{a_i(t_{k-1})}{S(t_{k-1-s})} + (1-\alpha)[S(t_{k-2-s}) - T(t_{k-2})];$$

$$T(t_{k-1}) = \beta \cdot (S(t_{k-1}) - S(t_{k-2})) + (1-\beta) \cdot T(t_{k-2});$$

$$S(t_{k-1}) = \chi \cdot \frac{a_i(t_{k-1})}{L(t_{k-1})} + (1-\chi) \cdot S(t_{k-1-s}),$$

де α – коефіцієнт згладжування ряду; β – коефіцієнт згладжування тренда; χ – коефіцієнт згладжування сезонності; $L(t_{k-1})$ – згладжена величина на час контролю t_{k-1}; $T(t_{k-1})$ – значення тренду на час контролю t_{k-1}; $S(t_{k-1})$ – коефіцієнт сезонності на час контролю t_{k-1}; $S(t_{k-1-s})$ – коефіцієнт сезонності за цей же період в попередньому сезоні. За уточненими значеннями швидкості $\widetilde{a}_i(t)$ визначаються уточнені оцінки середнього наробітку до відмови елементів:

$$\widetilde{T}_{\text{ср}i}(t) = 1/\widetilde{a}_i(t). \qquad (8)$$

Нове значення часу до наступного ТО $\widetilde{T}_{\text{то}}$ визначається за евристичною формулою:

$$\widetilde{T}_{\text{то}}(t) = \gamma T'_{\text{ср}}, \qquad (9)$$

де $T'_{\text{ср}}$ – найменше значення середнього наробітку до відмови, з отриманих за формулою (8); γ – коефіцієнт упередження, оптимальна величина якого визначається експериментально (за результатами моделювання). Після цього визначається запланований час наступного ТО $t_{\text{то}} = t + \widetilde{T}_{\text{то}}(t)$.

Алгоритм процесу моделювання адаптивного ТОС. Структурна схема алгоритму моделювання адаптивного ТОС (процедура Modelir_ATO), зображена на рис. 2.

4 ЕКСПЕРИМЕНТИ

Специфіка сучасних складних радіоелектронних об'єктів полягає у наявності в їх складі великої кількості різнотипних вузлів, блоків та інших комплектуючих елементів, які зазвичай мають різний рівень надійності, а також різні закони процесів зносу та старіння. Перевірити виграш у показниках надійності та вартості експлуатації для різних стратегій технічного обслуговування, при проведенні натурних експериментів, дуже складно, в першу чергу, через великий час проведення експериментів та відсутності точних теоретичних моделей кожного об'єкта. Між тим авторами були проведені експерименти над складною спеціальною радіоелектронною техні-

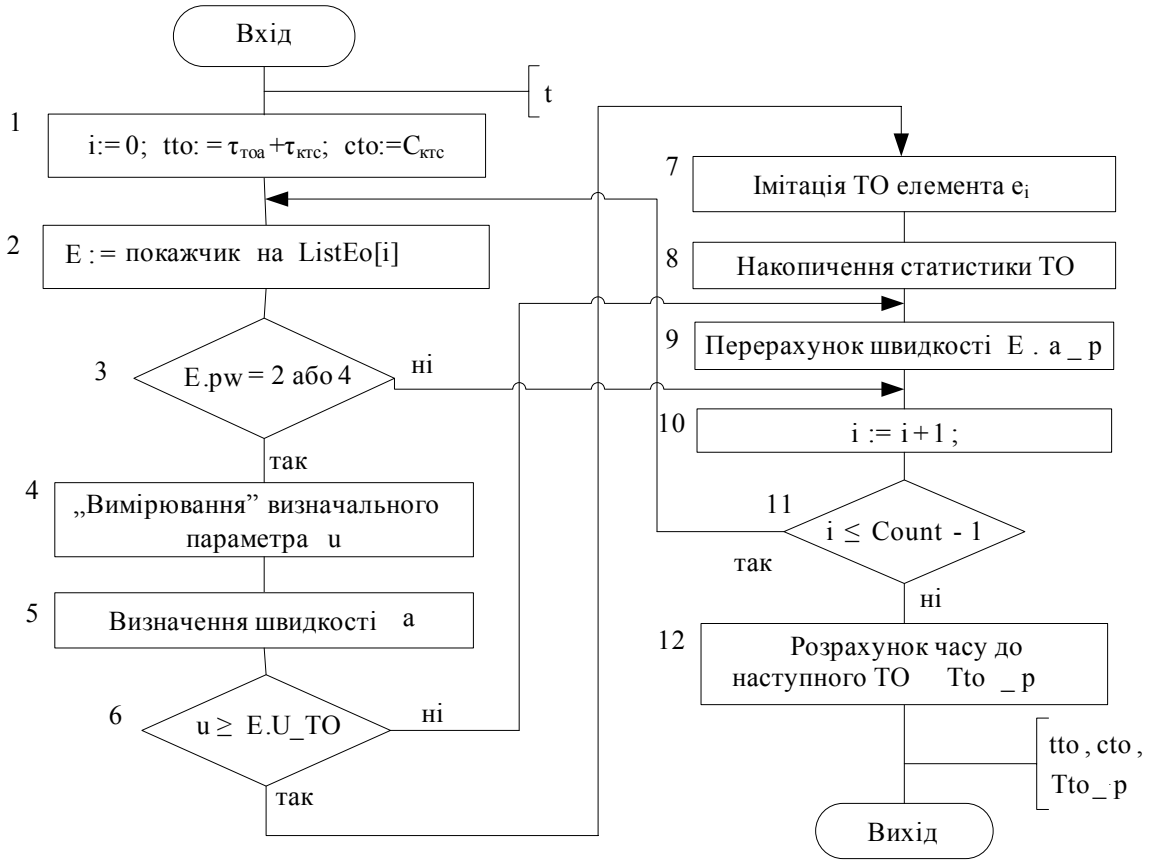


Рисунок 2 – Алгоритм процедури Modelir_ATO

кою. Навіть короткий опис умов проведення експерименту та аналізу результатів, займає значний обсяг інформації (декілька десятків сторінок). Між тим розрахунки підтвердили зазначений вище виграш, при використанні імітаційного статистичного моделювання. Сутність такої перевірки полягає у аналізі поведінки показників надійності при направленому варіюванні основними параметрами технічного обслуговування.

Для наочності та розуміння сутності впливів параметрів процесу ТО на показники надійності та вартості експлуатації радіоелектронного об'єкта у статті показано моделювання таких процесів на простому об'єкті – тестовому об'єкті Test-1.

Конструктивна структура даного об'єкта наведена на рис. 3. Об'єкт Test-1 представляється послідовною структурною схемою надійності та має 6 рівнів вкладеності. Окремі елементи на рисунку позначені окремими колами. Елементи, що включені до множини відновлювальних елементів (E_{o}), відмічені штриховкою.

При проведенні моделювання кількість елементів, що обслуговуються, змінювалась від 1 до 3. Для цього було обрано найменш надійні елементи (12, 132 та 11111). Одночасно змінювались параметри ТО: періодичність контролю $T_{\text{к}}$ – для стратегії ТОС; коефіцієнт упередження γ – для стратегії адаптивного ТОС. Параметри моделювання однакові для різних стратегій: T_{e} = 20 років, Δt = 3 міс. і N_I = 300, де T_{e} – час експлуатації об'єкта; Δt – інтервали дискретності накопичення статистики; N_I – кількість реалізацій моделювання.

5 РЕЗУЛЬТАТИ

За результатами моделювання були побудовані залежності показника безвідмовності об'єкта $\widetilde{T}_0$ від параметрів технічного обслуговування рис. 4.

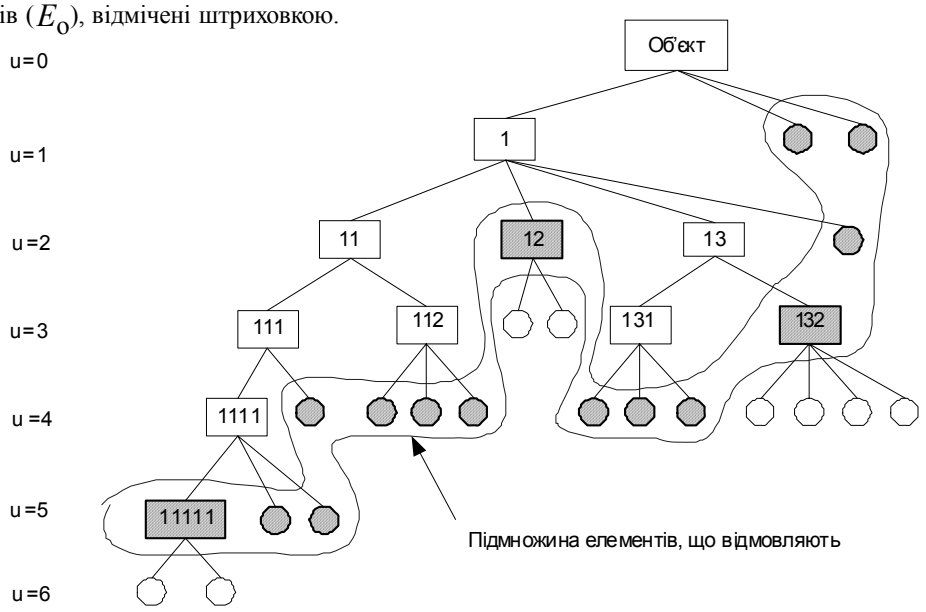
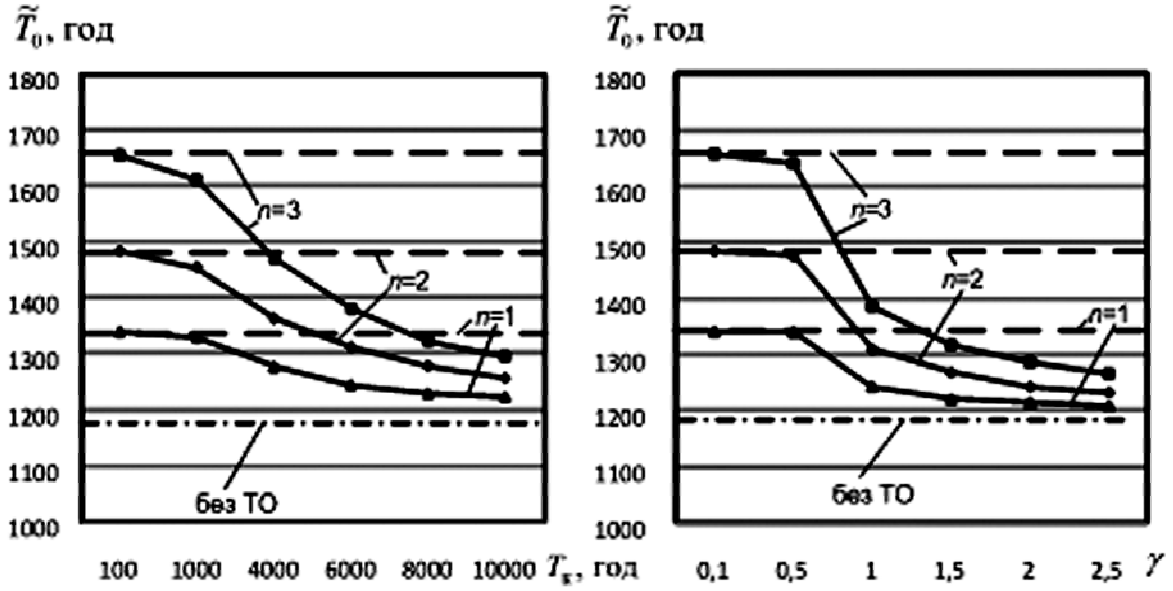


Рисунок 3 – Конструктивна структура тестового об'єкта Test-1



а б

Рисунок 4 – Залежності середнього наробітку на відмову $\widetilde{T}_0$ від параметрів ТО:
а – стратегія ТОС з постійною періодичністю контролю; б – стратегія адаптивне ТОС

Для стратегії ТОС з постійною періодичністю контролю, $T_к$ змінювалась в діапазоні від 100 до 10000 год, для стратегії адаптивного ТОС, коефіцієнт упередження γ змінювався в межах від 0,1 до 2,5, а коефіцієнт згладжування тренда β, дорівнював – 0,5. Для всіх стратегій, рівні u_{TOi} була задані однаковими і складали: $u_{TOi} = 0{,}5$.

Пунктирними лініями на графіках показані рівні безвідмовності (середній наробіток на відмову $\widetilde{T}_0$), що відповідають ідеальному технічному обслуговуванню. Під ідеальним ТО розуміється таке ТО, при якому елементи, що обслуговуються, ніколи не відмовляють. Штрих-пунктирною лінією показаний рівень безвідмовності об'єкта ($\widetilde{T}_0$ = 1187 год.), у якого ТО не проводиться.

6 ОБГОВОРЕННЯ

З аналізу графіків, рис. 4 видно, як зростає рівень безвідмовності об'єкта при збільшенні обсягу ТО (n – кількість елементів, що обслуговується), а також при зменшенні періодичності технічного обслуговування. Періодичність ТО при стратегіях ТО за станом є випадковою, оскільки обслуговування елементів проводяться тільки у тих випадках, якщо їх ТС погіршився до заданого граничного рівня ТО. Для даних стратегій, середня періодичність ТО зменшується, якщо зменшується періодичність контролю $T_к$ (коефіцієнт упередження γ). При невеликій періодичності ТО (при $T_к$ = 100 год і при γ = 0,1) рівень безвідмовності об'єкта $\widetilde{T}_0$ наближається до рівня, що відповідає ідеальному технічному обслуговуванню.

Статистична помилка (95% довірчий інтервал) оцінок $\widetilde{T}_0$ у всіх розрахунках в розглянутих прикладах моделювання ТО для об'єкта Test-1 дорівнює $\varepsilon = 15 \div 16\%$.

За отриманими результатами розрахунків, з урахуванням ТО, видно, що поведінка показника безвідмовності $\widetilde{T}_0$, при варіюванні параметрами ТО, цілком відповідає фізичному змісту модельованого процесу, що дозволяє зробити якісний висновок про адекватність розроблених алгоритмів процесу проведення ТО.

ВИСНОВКИ

У роботі вирішено актуальну наукову задачу щодо розробки алгоритмів моделювання процесів проведення технічного обслуговування за станом складних об'єктів РЕТ для різних стратегій ТО.

Наукова новизна роботи полягає в удосконаленні алгоритмів проведення технічного обслуговування за станом складних об'єктів РЕТ, які на відміну від існуючих використовують прогнозний параметр середньої швидкості деградації визначального параметра елемента на основі моделі трипараметричного експоненціального згладжування, що дозволить підвищити точність визначення часу контролю технічного стану з урахуванням сезонності.

Практичне значення полягає у розробці програмного забезпечення, яке реалізує процес визначення показників надійності та вартості експлуатації об'єкта РЕТ від конструктивного складу самого об'єкта та від параметрів проведення технічного обслуговування. Отримані результати необхідні при проектуванні (модернізації) об'єкта та визначенні оптимальних, для нього, параметрів проведення ТОіР.

СПИСОК ЛІТЕРАТУРИ

1. Прогнозирование надежности сложных объектов радиоэлектронной техники и оптимизация параметров их технической эксплуатации с использованием имитационных статистических моделей / [С. В. Ленков, К.Ф. Боряк, Г. В. Банзак и др.]. – Одесса : Изд-во «ВМВ», 2014. – 256 с.
2. Основы надежности и техническое обеспечение радиоэлектронных средств РТВ ПВО / [А. Н. Буточнов, Б. П. Креденцер, В. Г. Тоценко и др.]. – К. : КВИРТУ ПВО, 1982. – 226 с.
3. Xiao Feng Liang Warship reliability evaluation based on dynamic bayesian networks and numerical simulation / Xiao Feng Liang, Hong Dong Wang, Hong Yi, Dan Li // Ocean Engineering. – 2017. – No. 136. – P. 129–140.
4. Iris Tien Algorithms for Bayesian network modeling and reliability assessment of infrastructure systems / Iris Tien, Armen Der Kiureghian // Reliability Engineering & System Safety. – 2016. – No. 156. – P. 134–147.
5. Zhang Hui Reliability Factors Analysis of MEMS Safety and Arming System / Zhang Hui, Wang Ying, Lou Wenzhong, Liu Fangyi, Wang Fufu // Engineering Materials. – 2014. – Vol. 609–610. – P. 1494–1497. DOI:10.4028/www.scientific.net/KEM.609-610.1494.
6. Hao J. Reliability analysis based on improved dynamic fault tree / J. Hao, L. Zhang, L. Wei // Engineering asset management. – London : Springer, 2011. – P. 283–299. https://doi.org/10.1007/978-1-4471-4993-4_26.
7. Lindhe A. Approximate dynamic fault tree calculations for modelling water supply risks / A. Lindhe, T. Norberg, L. Rosén // Reliability Engineering & System Safety. – 2012. – Volume 106, October 2012. – P. 61–71. DOI: 10.1016/j.ress.2012.05.003
8. Postnikova Ivan V. A Methodology for Optimziation of Component Reliability of Heat Supply Systems / [Ivan V. Postnikova, Valery A. Stennikova, Ekaterina E. Mednikovaa, Andrey V. Penkovskiia] // Energy Procedia. – 2017. – Vol. 105. – P. 3083–3088. https://doi.org/10.1016/j.egypro.2017.03.643
9. Жиров Г.Б. Алгоритмічна модель процесу технічного обслуговування за станом з постійною періодичністю контролю / Г. Б. Жиров, Е. С. Ленков // Сучасна спеціальна техніка: наук-пр. журн. – 2017.– №1(48). – С. 26–29.
10. Жиров Г. Б. Алгоритмічна модель адаптивного технічного обслуговування за станом озброєння і військової техніки / Г. Б. Жиров, Е. С. Ленков / Збірник праць Національної академії Державної прикордонної служби України імені Б. Хмельницького. Серія : військові та технічні науки. –2017. – Вип. 1 (71). – С.368–378.
11. Жиров Г. Б. Усовершенствованная имитационная модель процесса технического обслуживания и ремонта сложного технического объекта / Г. Б. Жиров, Е. С. Ленков // Журнал Харківського національного університету Повітряних Сил імені І. Кожедуба «Системи обробки інформації». – 2017. – Вип. № 3 (149). – С. 14–18.

Стаття надійшла до редації 05.12.2017.
Після доробки 12.02.2018.

Жиров Г. Б.[1], Ленков Е. С.[2]

[1]Канд. техн. наук, с.н.с., ведущий научный сотрудник научно-исследовательского центра Военного института Киевского национального университета имени Тараса Шевченко, Киев, Украина

[2]Канд. техн. наук, старший научный сотрудник научного центра Военного института телекоммуникаций и информатизации, Киев, Украина

АЛГОРИТМЫ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ ПО СОСТОЯНИЮ СЛОЖНЫХ РАДИОЭЛЕКТРОННЫХ ОБЪЕКТОВ

Актуальность. Разработка сложных радиоэлектронных объектов и проведение технического обслуживания невозможна без разработки моделей и программных средств для оптимизации данных процессов. Для этого в работе разработаны соответствующие алгоритмы, которые являются частью программных средств, что является актуальной задачей.

Цель. Целью данной работы является разработка подходов по повышению показателей надежности объекта и уменьшения стоимости его эксплуатации на основе использования имитационного статистического моделирования.

Метод. Метод исследования базируется на идее, что техническое состояние объекта определяется техническим состоянием его элементов, а для определения технического состояния элемента используется понятие определяющего параметра. Значение параметра можно контролировать средствами диагностирования. В качестве моделей отказов элементов используется *DN*-распределение.

Результаты. Разработанные алгоритмы являются составной частью имитационной статистической модели объекта. Проведение моделирования позволяет определить оптимальные конструктивные характеристики объекта и параметры проведения его технического обслуживания, а критерием выступают максимизация показателей надежности и минимизация стоимости эксплуатации объекта.

Выводы. Научная новизна работы заключается в усовершенствовании алгоритмов технического обслуживания по состоянию, которые в отличие от существующих используют прогнозный параметр средней скорости деградации определяющего параметра элемента на основе модели трехпараметрического экспоненциального сглаживания.

Практическое значение состоит в разработке алгоритмов, которые положены в основу программного обеспечения, которое реализует процесс определения показателей надежности и стоимости эксплуатации объекта радиоэлектронной техники от конструктивного состава самого объекта и от параметров проведения технического обслуживания.

Ключевые слова: техническое обслуживание по состоянию, адаптивное техническое обслуживание, определяющий параметр.

Zhyrov G. B.[1], Lenkov E. S.[2]

[1]PhD, Senior Research Fellow, Leading Researcher at the Research Center of Military Institute of Taras Shevchenko National University, Kyiv, Ukraine

[2]PhD, Senior Researcher at the Scientific Center of Military Institute of Telecommunications and Informatization, Kyiv, Ukraine

ALGORITHMS OF SIMULATION OF TECHNICAL SERVICE PROCESSES AS COMPLEX RADIOELECTRONIC OBJECTS

Context. Development of complex radioelectronic facilities and maintenance is impossible without the development of models and software, to optimize these processes. For this, the corresponding algorithms have been developed in the work, which are part of the software, which is an urgent task.

Objective. The purpose of this work is to develop approaches to improve the reliability of the object and reduce the cost of its operation based on the use of simulation statistical modeling.

Method. The method of investigation is based on the idea that the technical state of an object is determined by the technical state of its elements, and the concept of a determining parameter is used to determine the technical state of an element. The value of the parameter can be controlled by means of diagnostics. As the models of element failures, a DN-distribution is used.

Results. The developed algorithms are an integral part of the simulation statistical model of the object. The simulation allows to determine the optimal design characteristics of the object and the parameters of its maintenance, and the criterion is to maximize reliability indicators and minimize the cost of operating the facility.

Conclusions. The scientific novelty of the work is to improve the state maintenance algorithms, which, unlike the existing ones, use the predictive parameter of the average degradation rate of the element's determining parameter based on the three-parameter exponential smoothing model.

Practical value consists in the development of algorithms that form the basis of software that implements the process of determining reliability indicators and the cost of operating an object of radioelectronic equipment from the structural composition of the facility itself and the parameters of maintenance.

Keywords: maintenance according to condition, adaptive maintenance, defining parameter.

REFERENCES

1. Lenkov S. V., Boriak K. F., Banzak G. V., Braun V. O., Osypa V. A., Pashkov S. A., Tsytsarev V. N., Berezovskaia Iu. V. Prognozirovanie nadezhnosti slozhnyh obektov radiojelektronnoj tehniki i optimizacija parametrov ih tehnicheskoj jekspluatacii s ispolzovaniem imitacionnyh statisticheskih modelej. Odessa, VMV, 2014 256 p.
2. Butochnov A. N., Kredencer B. P., Tocenko V. G., Cycarev V. N. i dr. *Osnovy nadezhnosti i tehnicheskoe obespechenie radiojelektronnyh sredstv RTV PVO.* Kyiv, KKVIRTU PVO, 1982, 226 p.
3. Xiao Feng Liang, Hong Dong Wang, Hong Yi, Dan Li. (). Warship reliability evaluation based on dynamic bayesian networks and numerical simulation, *Ocean Engineering*, 2017, No. 136, pp. 129–140.
4. Iris Tien, Armen Der Kiureghian Algorithms for Bayesian network modeling and reliability assessment of infrastructure systems. *Reliability Engineering & System Safety*, 2016, Vol. 156, pp. 134–147.
5. Zhang Hui, Wang Ying, Lou Wenzhong, Liu Fangyi, Wang Fufu. Reliability Factors Analysis of MEMS Safety and Arming System. Engineering *Materials,* 2014, Vol. 609–610, pp. 1494–1497. ISSN: 1662-9795. DOI: 10.4028/www.scientific.net/KEM.609-610.1494.
6. Hao J, Zhang L, Wei L. Reliability analysis based on improved dynamic fault tree. *Engineering asset management. London: Springer*, 2011, pp. 283–299. https://doi.org/10.1007/978-1-4471-4993-4_26.
7. Lindhe A, Norberg T, Rosén L. Approximate dynamic fault tree calculations for modelling water supply risks. *Reliability Engineering & System Safety*, 2012, Volume 106, October, pp. 61–71 https://doi.org/10.1016/j.ress.2012.05.003
8. Ivan V. Postnikova, Valery A. Stennikova, Ekaterina E. Mednikovaa, Andrey V. Penkovskiia A Methodology for Optimziation of Component Reliability of Heat Supply Systems *Energy Procedia,* 2017, Volume 105, pp. 3083–3088. https://doi.org/10.1016/j.egypro.2017.03.643
9. Zhirov G. B., Lenkov E. S. Algoritmichna model procesu tehnichnogo obslugovuvannja za stanom z postijnoju periodichnistju kontrolju. *Suchasna specialna tehnika*, 2017, No. 1(48), pp. 26–29.
10. Zhirov G. B., Lenkov E. S. Algoritmichna model adaptivnogo tehnichnogo obslugovuvannja za stanom ozbroєnnja i vijskovoї tehniki. *Zbirnik prac Nacionalnoї akademiї Derzhavnoї prikordonnoї sluzhbi Ukraïni imeni B. Hmelnickogo. Serija: vijskovi ta tehnichni nauki.* 2017, No. 1(71), pp. 368–378.
11. Zhirov G. B., Lenkov E. S. Usovershenstvovannaja imitacionnaja model processa tehnicheskogo obsluzhivanija i remonta slozhnogo tehnicheskogo obekta*, Sistemi obrobki informacii*, 2017, No. 3 (149), pp. 14–18.

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL AND COMPUTER MODELING

UDC 517.9:532.5

Lamtyugova S. N.[1], Sidorov M. V.[2], Sytnykova I. V.[3]

[1]Ph.D, Associate Professor of the Department of Advanced Mathematics of O. M. Beketov National University of Urban Economy in Kharkiv, Kharkiv, Ukraine

[2]Ph.D, Associate Professor, Associate Professor of the Department of Applied Mathematics of Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

[3]Ph.D, Associate Professor of the Department of Advanced Mathematics of O. M. Beketov National University of Urban Economy in Kharkiv, Kharkiv, Ukraine

METHOD OF NUMERICAL ANALYSIS OF THE PROBLEM OF MASS TRANSFER OF A CYLINDRICAL BODY WITH THE UNIFORM TRANSLATIONAL FLOW

Context. The problem of mass transfer of a cylindrical body with a uniform translational flow of a viscous incompressible fluid is examined in the paper.

Objective. The purpose of this work is to develop a new method for numerical analysis of the problem of mass transfer of a cylindrical body with a uniform translational flow, which based on the joint application of the *R*-functions structural method and the Galerkin projection method.

Method. In general case, the problem of stationary mass transfer of a cylindrical body with a viscous incompressible fluid flow is reduced to the solution of the equation of hydrodynamic flow passing a surface and an equation for concentration with corresponding boundary conditions on the surface of the body and far away from it. The geometry of the area, and also the boundary conditions (including the condition at infinity) may be taken into account precisely by using the constructive apparatus of the *R*-functions theory by V. L. Rvachev, the Academician of Ukrainian National Academy of Sciences. In this study, a complete structure of the solution of a linear boundary value problem for the concentration that exactly satisfies the boundary conditions on the boundary and condition at infinity is constructed on the basis of the *R*-functions theory methods, and this made it possible to lead the tasks in the infinite domain to tasks in the finite domain. To solve the linear problem for concentration the numerical algorithm on the basis of Galerkin method is developed.

Results. The computational experiment for the problem of the flow past circular and elliptical cylinders at various Reynolds and Peclet numbers was carried out.

Conclusions. The conducted experiments have confirmed the efficiency of the proposed method of numerical analysis of the problem of mass transfer of a cylindrical body with a uniform translational flow, based on the joint application of the *R*-functions structural method and Galerkin projection method. The prospects for the further research may be to use the developed method for the implementation of iterative methods for solving the task of nonlinear mass transfer, semi-discrete and projection methods for solving the non-stationary tasks, as well as in solving the tasks of optimal management of relevant technological processes.

Keywords: flow task, stationary flow, viscous fluid, stream function, mass transfer, concentration function, *R*-functions method, Galerkin method.

NOMENCLATURE

$c = c(r, \varphi)$ – a function of concentration;

c_0 – the constant concentration given on the boundary $\partial\Omega$ of a streamlined body;

$\partial\Omega$ – the boundary of a streamlined body;

$l = 1,2$ – the number of equation;

$\mathbf{n}$ – the external normal to $\partial\Omega$;

Pe – Peclet number;

r, φ, z – variables of the cylindrical coordinate system;

Re – Reynolds number;

u_l – the solutions of auxiliary problems with homogeneous boundary conditions;

U_∞ – unperturbed fluid velocity at infinity;

$\nu = \mathrm{Re}^{-1}$ – the coefficient of viscosity;

$\{\varphi_k(r, \varphi)\}$ – a complete system of particular solutions of equation $\Delta u = 0$ with respect to the exterior of a cylinder of finite radius;

© Lamtyugova S. N., Sidorov M. V., Sytnykova I. V., 2018
DOI 10.15588/1607-3274-2018-2-3

$\{\tau_j(r,\varphi)\}$ – a complete system of particular solutions of equation $\Delta u = 0$ relative to domain $\{\omega(r,\varphi) < M\}$;

$\{\phi_i(r,\varphi)\}$ – a sequence of functions that is complete with respect to the whole plane;

Φ_1, Φ_2 – undefined components of the solution structure;

$\psi = \psi(r,\varphi)$ – the stream function;

ω – a sufficiently smooth function built by using the constructive apparatus of the *R*-functions theory and describing the geometry of the domain Ω;

$\omega = 0$ – a normalized equation of $\partial\Omega$;

Ω – the flow domain;

Ω_n – the bounded domain $\{0 < \omega(r,\varphi) < M_n\}$.

INTRODUCTION

The problems of calculating viscous flows, which are complicated by mass transfer, are widely used in heat power engineering, chemical and food technologies, geo- and astrophysical studies, and environmental protection. Many processes of chemical technology are associated with fluid movement in process equipment. In the preparation of reagents and in the isolation of reaction products, such operations as leaching, absorption, extraction and distillation play an important role. The laws of hydrodynamics, heat and mass transfer are essential for the entire technological process. Processes of heat and mass transfer are also ones of the most important in the energy sector, as well as in a number of the technological processes in the metallurgical and other industries. In addition, the problems of mass transfer of bodies with a uniform viscous flow underlie the calculation of many technological processes, which are associated with dissolution, extraction, evaporation, precipitation of colloids, etc. Therefore, the development of new, as well as the improvement of existing methods of mathematical modeling and numerical analysis of external stationary problems of hydrodynamics of a viscous incompressible fluid, which take into consideration the mass transfer, is an actual scientific issue.

The object of this study is the stationary hydrodynamic process of flow past bodies by a viscous incompressible fluid, complicated by mass transfer and described by a system of equations with respect to the stream function and concentration.

The subject of this study is a mathematical model of the stationary task of flow past body by a viscous incompressible fluid with allowance for mass transfer, and a method for its numerical analysis.

The purpose of the work is to develop a new method for mathematical modeling of the mass transfer of a cylindrical body with a uniform translational flow which based on the *R*-functions method and Galerkin method.

1 PROBLEM STATEMENT

At small Peclet numbers, to describe the mass transfer process the Oseen approximation (equation 1) is used [1]:

$$\Delta c = \mathrm{Pe}U_\infty\left(\cos\varphi\frac{\partial c}{\partial r} - \frac{\sin\varphi}{r}\frac{\partial c}{\partial\varphi}\right) \text{ in } \Omega. \qquad (1)$$

In the general case, the mass transfer process is described by the equation for the concentration (equation 2) in the form [1, 2]:

$$\Delta c = \frac{\mathrm{Pe}}{r}\left(\frac{\partial\psi}{\partial\varphi}\frac{\partial c}{\partial r} - \frac{\partial\psi}{\partial r}\frac{\partial c}{\partial\varphi}\right) \text{ in } \Omega. \qquad (2)$$

In equations (1) and (2) $\Delta c = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial c}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 c}{\partial\varphi^2}$.

Equations (1) and (2) in a rectangular coordinate system have the following form:

$$\Delta c = \mathrm{Pe}U_\infty\frac{\partial c}{\partial x} \text{ in } \Omega, \qquad (3)$$

$$\Delta c = \mathrm{Pe}\left(\frac{\partial\psi}{\partial y}\frac{\partial c}{\partial x} - \frac{\partial\psi}{\partial x}\frac{\partial c}{\partial y}\right) \text{ in } \Omega, \qquad (4)$$

where $\Delta c = \frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}$.

Equations (1) and (2) should be supplemented by the boundary condition on $\partial\Omega$ and the condition at infinity [1]:

$$c\big|_{\partial\Omega} = c_0, \qquad (5)$$

$$\lim_{r\to+\infty} c = 0. \qquad (6)$$

The stream function $\psi(r,\varphi)$ can be found, for example, as the solution of the following nonlinear task of flow past a cylindrical body by a viscous incompressible fluid [3–6]:

$$\nu\Delta^2\psi = \frac{1}{r}\frac{\partial\psi}{\partial\varphi}\frac{\partial(\Delta\psi)}{\partial r} - \frac{1}{r}\frac{\partial\psi}{\partial r}\frac{\partial(\Delta\psi)}{\partial\varphi} \text{ in } \Omega, \qquad (7)$$

$$\psi\big|_{\partial\Omega} = 0, \left.\frac{\partial\psi}{\partial\mathbf{n}}\right|_{\partial\Omega} = 0, \qquad (8)$$

$$\lim_{r\to+\infty} r^{-1}\psi = U_\infty\sin\varphi, \qquad (9)$$

where $\Delta\psi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\varphi^2}$, $\Delta^2\psi = \Delta(\Delta\psi)$.

The task (1), (5), (6) does not depend on the stream function $\psi(r,\varphi)$, and the solution of the task (2), (5), (6) consists of two steps:

a) the search of the stream function as a solution of the task (7)–(9);

b) the solution of the task for concentration.

2 REVIEW OF THE LITERATURE

Various tasks which are arising in the study of external viscous fluid flows can be investigated theoretically or by means of physical experiment. At present, mathematical modeling and computational experiment are increasingly

being used in the study of hydrodynamic problems. Basically, finite difference method, finite element method, boundary integral element method and others are used for the numerical analysis of such problems. These methods are easy to implement, but do not have the necessary property of universality: when moving to a new area (especially non-classical geometry), it is necessary to generate a new grid, and often to replace complex sections of the boundary with simple ones, composed of, for example, straight line segments. The use of the *R*-functions structural-variational method [7, 8] by V. L. Rvachev, the Academician of Ukrainian National Academy of Sciences, is an alternative to the existing methods of numerical computation of hydrodynamic problems.

The *R*-functions method in computational hydrodynamics was applied in [9–14]. The task of viscous fluid external flow around bodies of revolution in a spherical coordinate system, which is complicated by mass transfer, was solved in [15, 16] with using the *R*-functions method.

In this study, we propose to apply the *R*-functions and Galerkin method for mathematical modeling the problem of mass transfer of a cylindrical body with a uniform translational flow.

3 MATERIALS AND METHODS

The method for solving the task for the stream function, based on the application of *R*-functions, successive approximations and Galerkin method, is described in [17, 18]. Substituting the stream function $\psi(r,\varphi)$ so obtained into equation (2), let us solve the problems (1), (5), (6) and (2), (5), (6) by the *R*-functions method. To do this, using the constructive means of the *R*-functions theory [7, 8] let us construct the structure of the boundary value task solution, i.e. the functions bundle that exactly satisfies the boundary condition and the condition at infinity.

Let us consider a sufficiently smooth function [19]

$$f_M(x)=\begin{cases}1-\exp\dfrac{Mx}{x-M}, & 0\le x<M;\\ 1, & x\ge M \quad (M=\text{const}>0).\end{cases} \tag{10}$$

The function (10) $f_M(x)\in C^{\infty}[0,+\infty)$ and satisfies the conditions:

$$\text{a) } f_M(0)=0\text{; b) } f_M'(0)=1;$$

$$\text{c) } f_M'(x)\ge 0 \ \ \forall x\ge 0\text{; d) } f_M(x)\equiv 1 \ \ \forall x\ge M.$$

Let us denote $\omega_M = f_M(\omega)$, then

$$1)\ \omega_M>0 \text{ in } \Omega;\ 2)\ \omega_M\big|_{\partial\Omega}=0;$$

$$3)\ \left.\frac{\partial\omega_M}{\partial\mathbf{n}}\right|_{\partial\Omega}=-1;\ 4)\ \omega_M\equiv 1, \text{if } \omega_M\ge M.$$

In the problems (3), (5), (6), and (4)–(6) we make the replacement

$$c=c_0(1-\omega_M)+u_l.$$

The choice of such a substitution is due to the fact that the function $c_0(1-\omega_M)$ satisfies the boundary condition (5) and the condition at infinity (6).

Then for u_l, $l=1,2$, we obtain the tasks

$$-\Delta u_l + B_l u_l = F_l \text{ in } \Omega, \tag{11}$$

$$u_l\big|_{\partial\Omega}=0, \tag{12}$$

$$\lim_{r\to+\infty} u_l = 0, \tag{13}$$

where

$$B_1=\mathrm{Pe}U_\infty\frac{\partial u_1}{\partial x},\ B_2=\mathrm{Pe}\left(\frac{\partial\psi}{\partial y}\frac{\partial u_2}{\partial x}-\frac{\partial\psi}{\partial x}\frac{\partial u_2}{\partial y}\right),$$

$$F_1=\Delta\big(c_0(1-\omega_M)\big)-\mathrm{Pe}U_\infty\frac{\partial\big(c_0(1-\omega_M)\big)}{\partial x},$$

$$F_2=\Delta\big(c_0(1-\omega_M)\big)-$$

$$-\mathrm{Pe}\left(\frac{\partial\psi}{\partial y}\frac{\partial\big(c_0(1-\omega_M)\big)}{\partial x}-\frac{\partial\psi}{\partial x}\frac{\partial\big(c_0(1-\omega_M)\big)}{\partial y}\right).$$

Note that $F_l\equiv 0$, $l=1,2$, in the domain $\{\omega(x,y)\ge M\}$.

Let us find the generalized solution u of tasks (11)–(13) as the limit, when $n\to\infty$, of solutions u_n of equations (11), which are considered in a sequence of domains $\{\Omega_n\}$, that is a monotonic exhaustion of an infinite area Ω.

In domains Ω_n we will consider the boundary value tasks

$$-\Delta u_{n,l} + B_l u_{n,l} = F_l \text{ in } \Omega_n, \tag{14}$$

$$u_{n,l}\big|_{\partial\Omega}=0, \tag{15}$$

where functions $u_{n,l}$ are continued by zero outside of Ω_n.

An approximate solution of tasks (14) – (15) for each $n=1,2,...$ according to the Bubnov-Galerkin method, will be sought in the form:

$$u_{n,l,N}=\sum_{j=1}^{N}c_{n,l,j}\phi_j. \tag{16}$$

The numbers $c_{n,l,j}$, $j=1,...,N$, are found from the system of linear algebraic equations

$$\sum_{j=1}^{N}c_{n,l,j}\{[\phi_j,\phi_i]+(K_l\phi_j,\phi_i)\}=(F_l,\phi_i), \tag{17}$$

where

$$i=\overline{1,N},\ l=1,2,$$

$$(F_1,\phi_i)=-\int_{\Omega_n}c_0\nabla(1-\omega_M)\cdot\nabla\phi_i\,d\Omega+$$

$$+\int_{\Omega_n}c_0\mathrm{Pe}U_\infty\cdot(1-\omega_M)\frac{\partial\phi_i}{\partial x}d\Omega,$$

$$(F_2, \phi_i) = -\int_{\Omega_n} c_0 \nabla(1-\omega_M)\cdot\nabla\phi_i d\Omega +$$

$$+\int_{\Omega_n} c_0 \mathrm{Pe}(1-\omega_M)\left(\frac{\partial\psi}{\partial y}\frac{\partial\phi_i}{\partial x}-\frac{\partial\psi}{\partial x}\frac{\partial\phi_i}{\partial y}\right)d\Omega.$$

To construct the coordinate sequence, the complete system of particular solutions of the Laplace equation [20] and the *R*-functions method [7, 8] will be used.

We have proved that for any choice of sufficiently smooth functions Φ_1, Φ_2 and at the requirement that $\Phi_1 \to 0$ as $r \to +\infty$ a function of the form

$$u = \omega_M \Phi_1 + \omega_M (1-\omega_M)\Phi_2$$

exactly satisfies the boundary conditions (12) and (13), that is, it is the structure of the solution of the boundary value problem (11)–(13).

Approximations of the functions Φ_1 и Φ_2 in the domain Ω_n will be seek in the form

$$\Phi_1 = \sum_{k=1}^{m_1} \alpha_k \cdot \varphi_k, \ \Phi_2 = \sum_{j=1}^{m_2} \beta_j \cdot \tau_j,$$

$$\text{where } \{\varphi_k(r,\varphi)\} = \left\{ r^{-n} \begin{matrix}\cos n\varphi \\ \sin n\varphi\end{matrix}, n = 1,2,... \right\},$$

$$\{\tau_j(r,\varphi)\} = \left\{ 1, r^k \begin{matrix}\cos k\varphi \\ \sin k\varphi\end{matrix}, k = 1,2,... \right\}.$$

Then the sequence of functions ,which is complete with relatively to the whole plane, has the form:

$$\{\phi_i(r,\varphi)\} = \{\omega_M(r,\varphi)\varphi_k(r,\varphi),$$

$$\omega_M(r,\varphi)(1-\omega_M(r,\varphi))\tau_j(r,\varphi)\}. \qquad (18)$$

The values of the coefficients α_k $(k = 1,2,..., m_1)$ and β_j $(j = 1,2,..., m_2)$ in accordance with the Bubnov-Galerkin method will be found from the condition of residual orthogonality to the first N $(N = m_1 + m_2)$ elements of the sequence (18), which leads to a system of linear algebraic equations in the form (17).

We have proved the convergence of the Galerkin approximations $u_{n,l,N}$ of the form (16) at $N \to \infty$, $M_n \to \infty$ in the energy norm to the unique generalized solutions of tasks (11)–(13).

4 EXPERIMENTS

A computational experiment was carried out for the tasks of flow past a circular cylinder $x^2 + y^2 = 1$ and an elliptical cylinder $\frac{x^2}{4} + \frac{y^2}{1} = 1$ at $c_0 = 1$, $M = 5, 10$, $m_1 = 10$, $m_2 = 11$ and at various Reynolds and Peclet numbers. The double integrals in the systems for determination $c_{n,l,j}$ were taken approximately by the Gauss formula with 50 nodes for each variable.

5 RESULTS

The concentration lines for a circular cylinder are shown in Fig. 1–3. Fig. 4–6 show the concentration lines for elliptical cylinder.

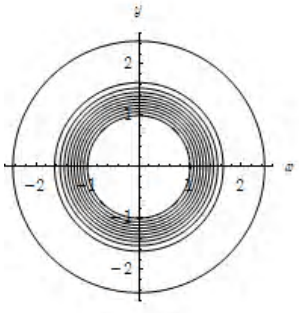
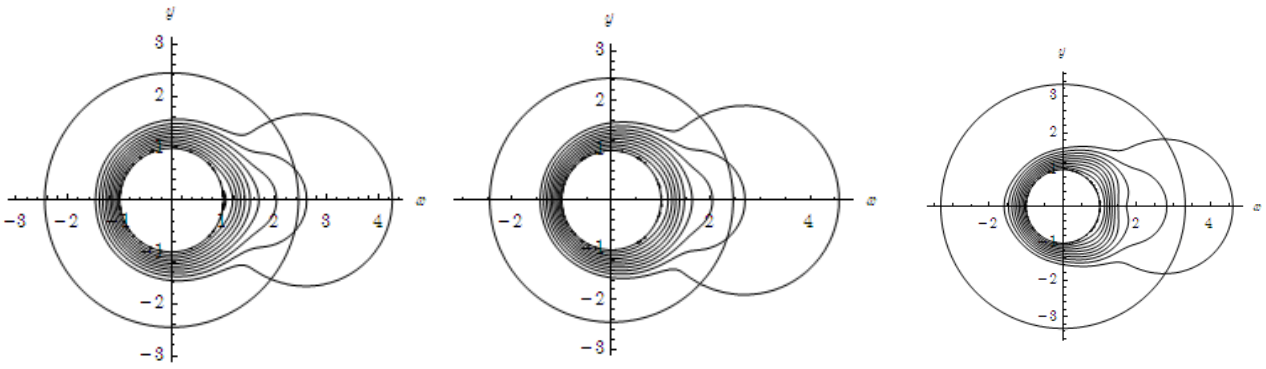


Figure 1 – The concentration lines for a circular cylinder

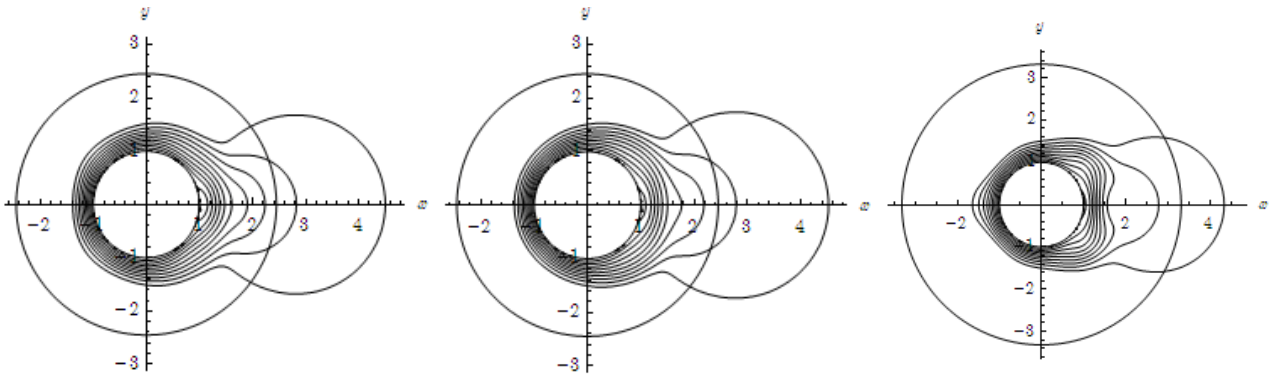
at $M = 5$ and $\mathrm{Pe} = 0{,}01$



a b c

Figure 2 – The concentration lines for a circular cylinder at $\mathrm{Pe} = 10$:

a – $\mathrm{Re} = 0$, b – $\mathrm{Re} = 5$, c – $\mathrm{Re} = 15$



a b c

Figure 3 – The concentration lines for a circular cylinder at $\mathrm{Pe} = 20$:

a – $\mathrm{Re} = 0$, b – $\mathrm{Re} = 5$, c – $\mathrm{Re} = 15$

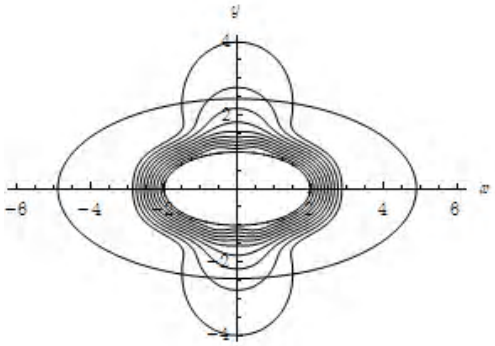


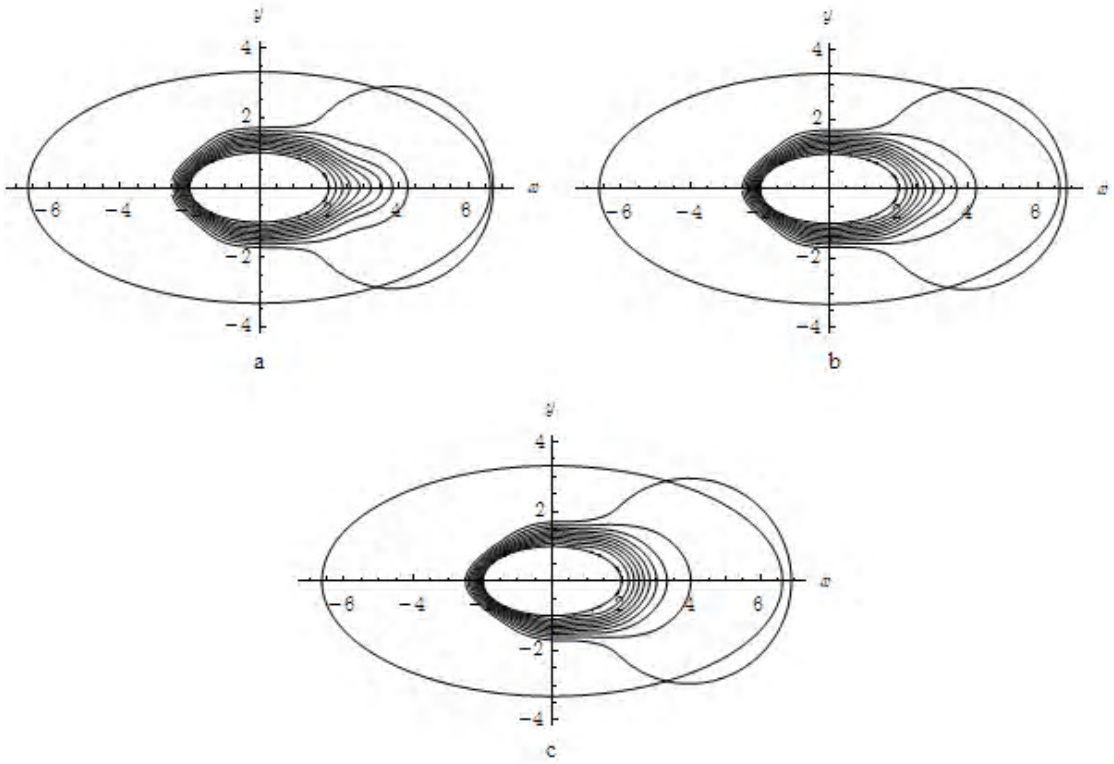
Figure 4 – The concentration lines for elliptical cylinder at $M = 5$ and $\mathrm{Pe} = 0{,}01$

6 DISCUSSION

At small Reynolds and Peclet numbers the substance is transported uniformly, dissolving into the liquid. As the Reynolds and Peclet numbers increase, the particles of substance begin to move with the flow. These results are consistent with the physics of the process. The efficiency of the proposed method for a spherical coordinate system was verified on the problem of flow past a sphere [13], for which an exact solution is known [20].

CONCLUSIONS

The numerical method for calculating the mass transfer of a cylindrical body with a uniform translational flow, which based on the joint application of the *R*-functions method and Galerkin method, is proposed for the first time in this



a b c

Figure 5 – The concentration lines for elliptical cylinder at $\mathrm{Pe} = 10$:

a – $\mathrm{Re} = 5$, b – $\mathrm{Re} = 15$, c – $\mathrm{Re} = 30$

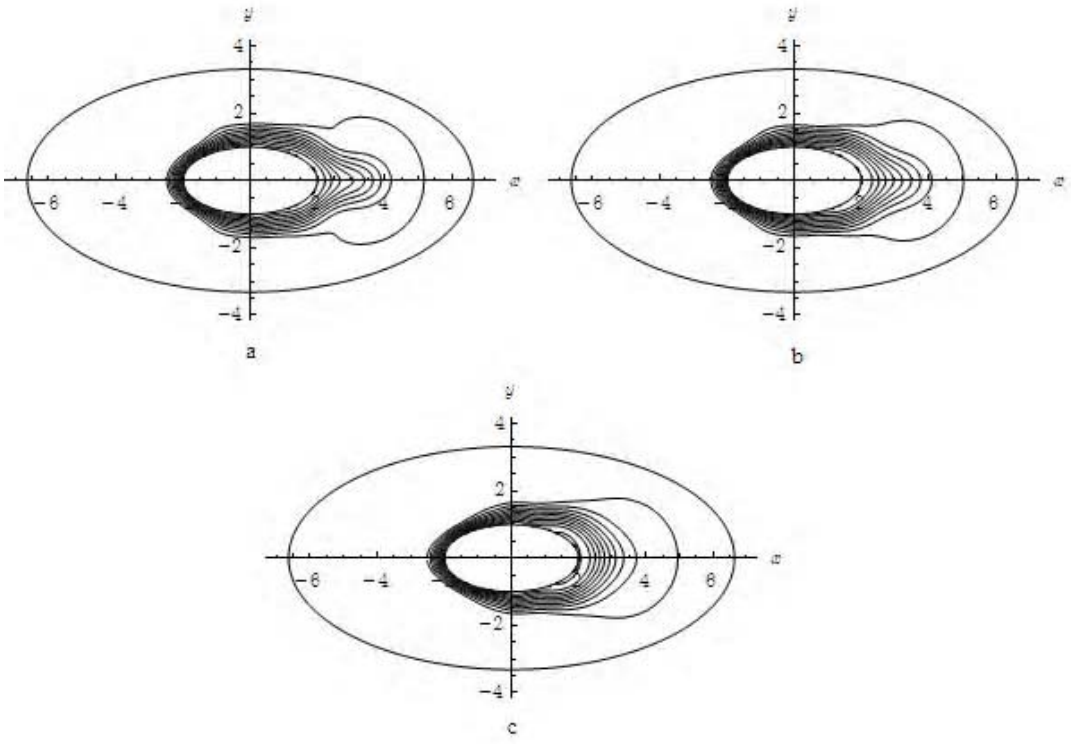


Figure 6 – The concentration lines for elliptical cylinder at $\mathrm{Pe} = 20$:

a – $\mathrm{Re} = 5$, b – $\mathrm{Re} = 15$, c – $\mathrm{Re} = 30$

study. By using the *R*-functions method the structure of the solution of the problems of flow past bodies with allowance for mass transfer, which precisely satisfying the boundary condition on the boundary and the condition at infinity, was constructed, and this made it possible to lead tasks in the infinite domain to tasks in the finite domain. To approximate the uncertain components of the solution structure, the Galerkin method was applied. The stationary problem of flow past a cylindrical body in a cylindrical coordinate system for a circular and elliptical cylinders has been solved numerically for various Reynolds and Peclet numbers.

ACKNOWLEDGMENTS

The work was carried out at the department of Applied Mathematics at Kharkiv National University of Radio Electronics and department of Advanced Mathematics at O. M. Beketov National University of Urban Economy in Kharkiv within the framework of collaborative scientific research conducted by the departments.

REFERENCES

1. Hydrodynamics, Mass and Heat Transfer in Chemical Engineering / A. D. Polyanin, A. M. Kutepov, D. A. Kazenin, A. V. Vyazmin. – CRC Press, 2002. – 408 p.
2. Гупало Ю. П. Массотеплообмен реагирующих частиц с потоком / Ю. П. Гупало, А. Д. Полянин, Ю. С. Рязанцев. – М. : Наука, 1985. – 336 с.
3. Loitsyansky L. G. Mechanics of liquids and gases / L. G. Loitsyansky. – New York: Begell House, 1995. – 803 p.
4. Шкадов В. Я. Течения вязкой жидкости / В. Я. Шкадов, З. Д. Запрянов. – М. : Изд-во Моск. у-та, 1984. – 200 с.
5. Happel J. Low Reynolds number hydrodynamics / J. Happel, H. Brenner. – Netherlands: Martinus Nijhoff, The Hague, 1973. – 559 p.
6. Polyanin A. D. Handbook of Nonlinear Partial Differential Equations. Second Edition / A. D. Polyanin, V. F. Zaitsev. – CRC Press, 2011. – 1912 p.
7. Рвачев В. Л. Теория *R*-функций и некоторые ее приложения / В. Л. Рвачев. – К. : Наук. думка, 1982. – 552 с.
8. Shapiro V. Semi-analytic geometry with R-functions / V. Shapiro // Acta Numerica. – 2007. – Vol. 16. – P. 239–303. DOI: 10.1017/S096249290631001X
9. Колосова С. В. Применение проекционных методов и метода *R*-функций к решению краевых задач в бесконечных областях: дисс. … к. ф.-м. н.: 01.01.07 – Вычислительная математика / С. В. Колосова. – Харьков: ХИРЭ, 1972. – 85 с.
10. Колосова С. В. Применение метода *R*-функций к расчету плоских течений вязкой жидкости / С. В. Колосова, М. В. Сидоров // Вісник Харківського національного університету. Сер. Прикл. матем. і мех. – 2003. – № 602. – С. 61–67.
11. Суворова И. Г. Компьютерное моделирование осесимметричных течений жидкости в каналах сложной формы / И. Г. Суворова // Вестник национального технического университета «ХПИ». – 2004. – № 31. – С. 141–148.
12. Тевяшев А. Д. Об одном подходе к математическому моделированию плоских стационарных течений вязкой несжимаемой жидкости в конечных односвязных областях / А. Д. Тевяшев, Н. В. Гибкина, М. В. Сидоров // Радиоэлектроника и информатика. – 2007. – № 2 (37). – С. 50–57.
13. Максименко-Шейко К. В. Математическое моделирование теплообмена при движении жидкости по каналам с винтовым типом симметрии методом *R*-функций / К. В. Максименко-Шейко // Доповіді Національної академії наук України. – 2005. – № 9. – С. 41–46.
14. Artyukh A. Mathematical modeling and numerical analysis of nonstationary plane-parallel flows of viscous incompressible fluid by *R*-functions and Galerkin method / A. Artyukh, M. Sidorov // Econtechmod. – 2014. – Vol. 3, No. 3. – P. 3–11.

15. Колосова С. В. Об одном методе численного анализа вязких течений, усложненных массообменом (задача обтекания) / С. В. Колосова, С. Н. Ламтюгова, М. В. Сидоров // Радиоэлектроника и информатика. – 2014. – №1 (64). – С. 25–30.
16. Lamtyugova S. N. The *R*-functions method application to solving mass transfer problems / S. N. Lamtyugova // Theoretical and applied aspects of cybernetics. Proceedings of the 2nd international scientific conference of students and young scientists. – Kyiv: Bukrek, 2012. – P. 108–111.
17. Ламтюгова С. Н. Математическое моделирование задач обтекания в цилиндрической системе координат / С. Н. Ламтюгова, М. В. Сидоров // Вісник ХНУ ім. В. Н. Каразіна. Сер. Математичне моделювання. Інформаційні технології. Автоматизовані системи управління. – 2014. – №1105, Вип. 24. – С. 111–121.
18. Lamtyugova S. N. Numerical analysis of the problem of flow past a cylindrical body applying the *R*-functions method and the Galerkin method / S. N. Lamtyugova, M. V. Sidorov // Econtechmod. – 2014. – Vol. 3, No. 3. – P. 43 50.
19. Стрельченко А. Й. Про один метод розв'язування крайових задач / А. Й. Стрельченко, С. В. Колосова, В. Л. Рвачов // Доп. АН УРСР, сер. А. 1972. № 9. С. 837–839.
20. Polyanin A. D. Handbook of Linear Partial Differential Equations for Engineers and Scientists. Second Edition / A. D. Polyanin, V. E. Nazaikinskii. – CRC Press, 2016. – 1632 p.

Article was submitted 04.12.2017.
After revision 15.01.2018.

Ламтюгова С. М.[1], Сидоров М. В.[2], Ситникова Ю. В.[3]

[1]Канд. фіз.-мат. наук, доцент кафедри вищої математики Харківського національного університету міського господарства імені О. М. Бекетова, Харків, Україна

[2]Канд. фіз.-мат. наук, доцент, доцент кафедри прикладної математики Харківського національного університету радіоелектроніки, Харків, Україна

[3]Канд. пед. наук, доцент кафедри вищої математики Харківського національного університету міського господарства імені О. М. Бекетова, Харків, Україна

МЕТОД ЧИСЕЛЬНОГО АНАЛІЗУ ЗАДАЧІ МАСООБМІНУ ЦИЛІНДРИЧНОГО ТІЛА З РІВНОМІРНИМ ПОСТУПАЛЬНИМ ПОТОКОМ

Актуальність. У статті розглядається задача масообміну циліндричного тіла з рівномірним поступальним потоком в'язкої нестисливої рідини.

Мета роботи – розробка нового методу чисельного аналізу задачі масообміну циліндричного тіла з рівномірним поступальним потоком, заснованого на сумісному застосуванні структурного методу *R*-функцій і проекційного методу Гальоркіна.

Метод. У загальному випадку задача про стаціонарний масообмін циліндричного тіла з потоком в'язкої нестисливої рідини зводиться до розв'язання рівняння гідродинамічного обтікання поверхні і рівняння для концентрації з відповідними крайовими умовами на поверхні тіла і далеко від нього. Точно врахувати геометрію області, а також крайові умови (в т.ч. і умову на нескінченності), можна, скориставшись конструктивним апаратом теорії *R* функцій акад. НАН України В. Л. Рвачова. В роботі на основі методів теорії *R*-функцій побудована повна структура розв'язку лінійної крайової задачі для концентрації, яка точно задовольняє крайовим умовам на межі тіла і умові на нескінченності, що дозволило звести задачі в нескінченній області до задач в скінченній області. Для розв'язання лінійної задачі для концентрації розроблено чисельний алгоритм на основі методу Гальоркіна.

Результати. Обчислювальний експеримент проведений для задачі обтікання кругового і еліптичного циліндрів для різних чисел Рейнольдса і Пекле.

Висновки. Проведені експерименти підтвердили працездатність запропонованого методу чисельного аналізу задачі масообміну циліндричного тіла з рівномірним поступальним потоком, заснованого на сумісному використанні структурного методу *R*-функцій і проекційного методу Гальоркіна. Перспективи подальших досліджень можуть полягати у використанні розробленого методу при реалізації ітераційних методів розв'язання задачі нелінійного масообміну, полудіскретних і проекційних методів розв'язання нестаціонарних задач, а також при розв'язанні задач оптимального управління відповідними технологічними процесами.

Ключові слова: задача обтікання, стаціонарна течія, в'язка рідина, функція течії, масообмін, функція концентрації, метод *R*-функцій, метод Гальоркіна.

Ламтюгова С. Н.[1], Сидоров М. В.[2], Ситникова Ю. В.[3]

[1]Канд. физ.-мат. наук, доцент кафедры высшей математики Харьковского национального университета городского хозяйства имени А. Н. Бекетова, Харьков, Украина

[2]Канд. физ.-мат. наук, доцент, доцент кафедры прикладной математики Харьковского национального университета радиоэлектроники, Харьков, Украина

[3]Канд. пед. наук, доцент кафедры высшей математики Харьковского национального университета городского хозяйства имени А. Н. Бекетова, Харьков, Украина

МЕТОД ЧИСЛЕННОГО АНАЛИЗА ЗАДАЧИ МАССООБМЕНА ЦИЛИНДРИЧЕСКОГО ТЕЛА С РАВНОМЕРНЫМ ПОСТУПАТЕЛЬНЫМ ПОТОКОМ

Актуальность. В статье рассматривается задача массообмена цилиндрического тела с равномерным поступательным потоком вязкой несжимаемой жидкости.

Цель работы – разработка нового метода численного анализа задачи массообмена цилиндрического тела с равномерным поступательным потоком, основанного на совместном применении структурного метода *R*-функций и проекционного метода Галеркина.

Метод. В общем случае задача о стационарном массообмене цилиндрического тела с потоком вязкой несжимаемой жидкости сводится к решению уравнения гидродинамического обтекания поверхности и уравнения для концентрации с соответствующими краевыми условиями на поверхности тела и вдали от него. Точно учесть геометрию области, а также краевые условия (в т.ч. и условие на бесконечности), можно, воспользовавшись конструктивным аппаратом теории *R*-функций акад. НАН Украины В. Л. Рвачева. В работе на основании методов теории *R*-функций построена полная структура решения линейной краевой задачи для концентрации, точно удовлетворяющая краевым условиям на границе и условию на бесконечности, что позволило свести задачи в бесконечной области к задачам в конечной области. Для решения линейной задачи для концентрации разработан численный алгоритм на основании метода Галеркина.

Результаты. Вычислительный эксперимент проведен для задачи обтекания кругового и эллиптического цилиндров при различных числах Рейнольдса и Пекле.

Выводы. Проведенные эксперименты подтвердили работоспособность предложенного метода численного анализа задачи массообмена цилиндрического тела с равномерным поступательным потоком, основанного на совместном применении структурного метода *R*-функций и проекционного метода Галеркина. Перспективы дальнейших исследований могут заключаться в использовании разработанного метода при реализации итерационных методов решения задачи нелинейного массообмена, полудискретных и проекционных методов решения нестационарных задач, а также при решении задач оптимального управления соответствующими технологическими процессами.

Ключевые слова: задача обтекания, стационарное течение, вязкая жидкость, функция тока, массообмен, функция концентрации, метод *R*-функций, метод Галеркина.

REFERENCES

1. Polyanin A. D., Kutepov A. M., Kazenin D. A., Vyazmin A. V. Hydrodynamics, Mass and Heat Transfer in Chemical Engineering. CRC Press, 2002, 408 p.
2. Gupalo Ju. P, Poljanin A. D., Rjazancev Ju. S. Massoteploobmen reagirujushhih chastic s potokom. Moscow, Nauka, 1985, 336 p.
3. Loitsyansky L. G. Mechanics of liquids and gases. New York, Begell House, 1995, 803 p.
4. Shkadov V. Ja, Zaprjanov Z. D. Techenija vjazkoj zhidkosti. Moscow, Izd-vo Mosk. u-ta, 1984, 200 p.
5. Happel J., Brenner H. Low Reynolds number hydrodynamics. Netherlands: Martinus Nijhoff, The Hague, 1973, 559 p.
6. Polyanin A. D., Zaitsev V. F. Handbook of Nonlinear Partial Differential Equations. Second Edition. CRC Press, 2011, 1912 p.
7. Rvachev V. L. Teorija R-funkcij i nekotorye ee prilozhenija. Kiev, Nauk. dumka, 1982, 552 p.
8. Shapiro V. Semi-analytic geometry with R-functions, *Acta Numerica*, 2007, V. 16, pp. 239–303. DOI: 10.1017/S096249290631001X
9. Kolosova S. V. Primenenie proekcionnyh metodov i metoda R-funkcij k resheniju kraevyh zadach v beskonechnyh oblastjah: diss. … k. f.-m. n.: 01.01.07. Vychislitel'naja matematika. Har'kov, HIRJe, 1972, 85 p.
10. Kolosova S. V., Sidorov M. V. Primenenie metoda R-funkcij k raschetu ploskih techenij vjazkoj zhidkosti, *Visnyk Harkivs'kogo nacional'nogo universytetu.* Ser. Prykl. matem. i meh., 2003, No. 602, pp. 61–67.
11. Suvorova I. G. Komp'juternoe modelirovanie osesimmetrichnyh techenij zhidkosti v kanalah slozhnoj formy, *Vestnik nacional'nogo tehnicheskogo universiteta «HPI»*, 2004, No. 31, pp. 141–148.
12. Tevjashev A. D., Gibkina N. V., Sidorov M. V. Ob odnom podhode k matematicheskomu modelirovaniju ploskih stacionarnyh techenij vjazkoj neszhimaemoj zhidkosti v konechnyh odnosvjaznyh oblastjah, *Radiojelektronika i informatika*, 2007, No. 2 (37), pp. 50–57.
13. Maksimenko-Shejko K. V. Matematicheskoe modelirovanie teploobmena pri dvizhenii zhidkosti po kanalam s vintovym tipom simmetrii metodom R-funkcij, *Dopovidi Nacional'noi' akademii' nauk Ukrai'ny*, 2005, № 9, pp. 41–46.
14. Artyukh A., Sidorov M. Mathematical modeling and numerical analysis of nonstationary plane-parallel flows of viscous incompressible fluid by *R*-functions and Galerkin method, *Econtechmod*, 2014, Vol. 3, No 3, pp. 3–11.
15. Kolosova S. V., Lamtyugova S. N., Sidorov M. V. Ob odnom metode chislennogo analiza vjazkih techenij, uslozhnennyh massoobmenom (zadacha obtekanija), *Radiojelektronika i informatika*, 2014, No. 1 (64), pp. 25–30.
16. Lamtyugova S. N. The *R*-functions method application to solving mass transfer problems, *Theoretical and applied aspects of cybernetics. Proceedings of the 2nd international scientific conference of students and young scientists*. Kyiv, Bukrek, 2012, pp. 108–111.
17. Lamtyugova S. N., Sidorov M. V. Matematicheskoe modelirovanie zadach obtekanija v cilindricheskoj sisteme koordinat, *Visnyk HNU im. V. N. Karazina*. Ser. Matematychne modeljuvannja. Informacijni tehnologii'. Avtomatyzovani systemy upravlinnja, 2014, №1105, vyp. 24, pp. 111–121.
18. Lamtyugova S. N., Sidorov M. V. Numerical analysis of the problem of flow past a cylindrical body applying the *R*-functions method and the Galerkin method, *Econtechmod*, 2014, Vol. 3, No. 3, pp. 43–50.
19. Strel'chenko A. J., Kolosova S. V., Rvachov V. L. Pro odyn metod rozv'jazuvannja krajovyh zadach, *Dop. AN URSR, ser. A*, 1972, No. 9, pp. 837–839.
20. Polyanin A. D., Nazaikinskii V. E. Handbook of Linear Partial Differential Equations for Engineers and Scientists. Second Edition. CRC Press, 2016, 1632 p.

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

УДК 681.327.12

Бісікало О. В.[1], Грищук Т. В.[2], Ковтун В. В.[3]

[1]Д-р техн. наук, професор, декан факультету комп'ютерних систем і автоматики Вінницького національного технічного університету, Вінниця, Україна

[2]Канд. техн. наук, доцент, доцент кафедри комп'ютерних систем управління Вінницького національного технічного університету, Вінниця, Україна

[3]Канд. техн. наук, доцент, доцент кафедри комп'ютерних систем управління Вінницького національного технічного університету, Вінниця, Україна

ОПТИМІЗАЦІЯ КЛАСИФІКАТОРА АВТОМАТИЗОВАНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ МОВЦЯ КРИТИЧНОГО ЗАСТОСУВАННЯ

Актуальність. Розглянуто питання адаптації згортального нейромережевого класифікатора для використання у автоматизовані системі розпізнавання мовців критичного застосування (АСРМКЗ). Об'єктом дослідження є індивідуальні особливості мовного процесу людини.

Мета роботи. Розроблення заходів по виділенню з мовного сигналу індивідуальних для мовця ознак, підвищення їх інформативності в результаті виконання факторного аналізу, їх візуальне представлення для використання згортального нейромережевого класифікатора та оптимізація його архітектури для потреб АСРМКЗ.

Метод. Запропоновано заходи по оптимізації процедури класифікації мовців АСРМКЗ, для чого теоретично обґрунтовано оптимальний спосіб представлення інформативних ознак і метод підвищення їх інформативності, обґрунтовано вид топологію і заходи для підвищення ефективності процесу розпізнавання мовців. Зокрема, обґрунтовано доцільність використання нормалізованих за потужністю кепстральних коефіцієнтів PNCC для опису фонограм, записаних в умовах шумного оточення, запропоновано використовувати фільтри Габора для представлення інформації, що аналізуватиметься згортальною нейромережею, вибрано оптимальний метод факторного аналізу, а саме, розріджений метод аналізу головних компонент, для зменшення розмірності вектору ознак із збереженням його інформативності, запропоновано удосконалену топологію згортальної нейромережі для АСРМКЗ, у якій фільтри Габора інтегровано у згортальний шар, що дозволяє оптимізувати їх параметри при навчанні нейромережі, і на повнозв'язному шарі використано глибоку нейромережу із bottleneck-шаром, ваги якого після навчання використано як вхідні дані для контрольного GMM/HMM-класифікатора.

Результати. Методи представлення та оптимізації інформативних для розпізнавання мовця ознак, методи їх візуального представлення та удосконалення топології згортальної нейромережі для прийняття рішень на їх основі.

Висновки. Отримані теоретичні результати знайшли емпіричного підтвердження. Зокрема, доведено стійкість удосконаленої згортальної нейромережі до присутності шумів оточення у вхідних фонограмах, яка виявилася вищою за показники звичайної згортальної нейромережі та глибокої нейромережі. При зростанні ВСШ>10 дБ контрольний GMM/HMM-класифікатор виявився ефективнішим за нейромережеві щодо імовірності прального розпізнавання мовців, що можна пояснити ефективністю використаної UBM-моделі, але він є і суттєво ресурсоємнішим. Також емпірично виявлено вікна банку фільтрів Габора, які надавали найбільш варіативну щодо індивідуальних особливостей мовлення, інформацію.

Ключові слова: автоматизована система розпізнавання мовців критичного застосування, обробка сигналів, нейромережа, факторний аналіз.

НОМЕНКЛАТУРА

CNN – згортальна нейронна мережа;
DNN – глибока нейронна мережа;
HMM – приховані марковські моделі;
SPCA – розріджений метод главних компонент;
UBM – універсальна фонова модель;
VAD – детектор мовної активності;
WCCN – операції внутрікласової коваріаційної нормалізації;
АСРМКЗ – автоматизована система розпізнавання мовця критичного застосування;
ВСШ – відношення сигнал/шум;
ІПР – середня імовірність правильного розпізнавання;

© Бісікало О. В., Грищук Т. В., Ковтун В. В., 2018
DOI 10.15588/1607-3274-2018-2-4

$STRF[n_t,k]$, $STRF$ – спектрально-темпоральні рецептивні поля;

LPC – коефіцієнтами лінійного прогнозу;

$MF[m,l]$ – спектральне представлення фонограм;

$MFCC[n]$, $MFCC$ – мел-кепстральні коефіцієнти;

$PLP[n]$, PLP – коефіцієнти перцептивного лінійного прогнозу;

$PNCC[m,l]$, $PNCC$ – нормовані за потужністю кепстральні коефіцієнти;

Σ_x – емпірична коваріаційна матриця;

Λ – множина гаусових сумішей;

φ – фаза;

λ_a, λ_b, λ_t – коефіцієнти забування;

v – домінуючий розріджений головний вектор;

ρ – невід'ємний параметр управління розрідженістю головного вектора;

ω_n, ω_k – темпоральні і спектральні частоти модуляції на різних амплітудних рівнях ΔA; $b(i)$ – коефіцієнти лінійного прогнозу;

D – розмірність простору ознак;

$Gb[x,y]$ – імпульсна передатна характеристика фільтра Габора;

f_0 – частота;

F_d – частота дискретизації;

$F(\ldots)$ – перетворення Фур'є;

$F(Gb)$ – гаусовський сигнал, екстремальне значення якого розташовано на центральній частоті фільтра;

$h[n_t,k]$ – віконна функція Хеннінга;

$h_l[\ldots]$ – віконна функція згладжування;

$h_m[n_t,k]$ – нейрон m-го вектора ознак, який відповідає певному фільтру Габора, рецептивне поле якого утворює матрицю $2K_m+1$ (смуги) на $2N_m+1$ (відліки), орієнтовану на поточну смугу фрейму $x(n,k)$;

im – зображення;

k – частота;

L – кількість фільтрів у банку;

l_1 – штрафна норма;

m, l – номери фрейму та частотної смуги фільтру відповідно;

n – кількість коефіцієнтів;

n_t – момент часу;

$p(x|\lambda)$, GMM – суміш гаусових розподілів;

$O[n_t,k]$ – передатні функції нейронів слухової кори голосного мозку;

S – множина мовців;

$s[n_t,k]$ – комплексна синусоїда;

$Q_f[m,l]$ – значення нижньої обвідної після напвіперіодного випрямляча;

$Q_0[m,l]$ – сигнал після напівперіодного випрямлення;

$Q_{tm}[m,l]$ – сигналу після процедури часового маскування;

$Q_{\text{вх}}[m,l]$ – результат ковзного усереднення $MF[m,l]$ на протязі M фреймів;

$Q_{\text{вих}}[m,l]$ – усереднена за часом потужність;

$\{x_i\}$ – результатами спостережень;

$x[\ldots]$ – короткочасний спектр сигналу.

ВСТУП

Серед існуючих класів автоматизованих систем окреме місце займають так звані критичні системи, які функціонують із високою надійністю і зберігають прогнозований її рівень на протязі всього життєвого циклу автоматизованої системи не залежно від будь-яких зовнішніх обставин. При створені критичних систем віддають перевагу перевіреним та знаним методам та технологіям перед новітніми розробками, які не пройшли всебічної емпіричної перевірки. Ресурсозатратні технології, використання яких для розробки поточних автоматизованих систем є економічно невигідним, допустимі при створенні критичних систем, для яких головним є надійність функціонування.

Актуальна класифікація критичних систем розділяє їх за реакцією на виникнення надзвичайних ситуацій та за галуззю експлуатації. За реакцією на виникнення надзвичайних ситуацій критичні системи розділяють на критичні системи, які мають продовжувати виконання функціональних операцій при виникненні відмов чи впливі непередбачуваних факторів, наприклад літак повинен продовжувати летіти за будь-яких обставин, та критичні системи, які повинні обов'язково безпечно завершувати функціонування не залежно від дії оточуючих факторів, наприклад, потрібно загальмувати потяг щоб перевести його у «безпечний» стан. За галуззю експлуатації критичні системи розділяють на:

– критичні системи збереження життя – це автоматизовані системи, збої у функціонуванні яких можуть привести до загибелі людей, суттєвих пошкоджень навколишнього середовища – це, наприклад, автоматизовані системи управління хімічним виробництвом, літаками, потягами метро, атомних електростанцій;

– критичні системи гарантованого функціонування – ці автоматизовані системи, які створюються із передбаченням гарантованого завершення виконуваної операції, наприклад, системи навігації, управління багажем у аеропортах;

– критичні системи економічного застосування – ці автоматизовані системи, створені з метою уникнення значних матеріальних або репутаційних витрат, що зазвичай забезпечується вчасним припиненням виконання певних операцій інтегрованою критичною системою, наприклад, у системах роботи із клієнтами у банках, інтернет-пошукові системи, ERP системи, системи роботи із біржовими операціями тощо;

– критичні системи інформаційної безпеки – ці автоматизовані системи унеможливлюють втрату конференційної інформації за будь-яких умов.

Автори проводять послідовні дослідження щодо синтезу теоретичних засад та методів для практичного впровадження автоматизованих систем розпізнавання мовця в критичні системи. Запропонований матеріал стосується оптимізації роботи підсистеми класифікації автоматизованої системи розпізнавання мовців критичного застосування (АСРМКЗ).

Об'єктом дослідження є індивідуальні особливості процесу мовотворення людини.

Предметом дослідження – методи виділення інформативних для розпізнавання особи мовця ознак із фонограм мовних сигналів, методи компактифікації їх представлення у факторному просторі, формулювання виду подачі інформативних ознак у відповідності із моделями слухового сприйняття людини і завадостійкі автоматизовані методи прийняття рішень для розпізнавання особи мовця.

Мета дослідження – підвищення стійкості АСРМКЗ до впливу різних видів шумів у вхідних фонограмах.

1 ПОСТАНОВКА ЗАДАЧІ

Формалізуємо постановку задачі розпізнавання мовців так. Нехай $S = \{s_1, s_2, \ldots, s_m\}$ – простір образів, які має розпізнавати АСРМКЗ, а $s_i \in S$ – об'єкт розпізнавання. Використовуючи певні правила сформуємо факторний простір F, причому функція $f(s_i): S \to X$ ставить у відповідність кожній фонограмі із записом мовного сигналу мовця s_i точку $f(s_i)$ у факторному просторі. Кожному i-му мовцеві у факторному просторі відповідає множина точок, кожна з яких відповідає опрацьованій фонограмі, утворюючи кластер $C_i \in X$, при чому кластери різних мовців мають не перетинатися. Вирішувалене правило $r(x): X \to M$ дозволяє із певною імовірністю P_{s_i} стверджувати, що точка $f(s_i)$ у факторному просторі належить s_i мовцеві. Задача авторів – сформувати факторний простір F із чіткими границями кластерів мовців і створити вирішувальне правило R, яке за інформацією про мовців із факторного простору максимізуватиме імовірності правильного розпізнавання P_{s_i} для всіх мовців з множини S.

2 ОГЛЯД ЛІТЕРАТУРИ

Основною специфікою АСРМКЗ є обґрунтовано висока імовірність його правильного розпізнавання при визначених рівнях ВСШ. Існуючі методи підвищення надійності розпізнавання мовців у шумному акустичному середовищі базуються на компенсації відмінностей між фонограмами без шуму («чистими» фонограмами) і фонограмами із шумом за рахунок математичних моделей [1, 2] або ведення додаткових факторів [3–5]. Перший підхід передбачає адаптацію моделі мовців до застосування у шумному середовищі, а другий – на використанні алгоритмів теорії цифрової обробки сигналів для фільтрації шумів із збереженням індивідуальності звучання фонограм або винайденні стійких до шумів та інформативних для розпізнавання особи мовця факторах.

Актуальні роботи, які можна віднести до першого варіанту базуються на моделюванні природної здатності людини якісно виконувати задачу розпізнавання мовця у шумному середовищі описуючи фізіологічні аспекти слухової системи та моделюючи когнітивні функції відповідних відділів кори головного мозку у вигляді нейромереж різної топології. Зокрема, дослідження [6, 7] демонструють можливість застосування параметрів прихованого шару (ваг вхідних зв'язків нейронів обраного прихованого шару) навченої глибокої нейромережі в якості факторів для розпізнавання мовців, і доводять їх більшу інформативність порівняно із традиційними факторами, як то, MFCC, LPC, короткочасною енергією і т. ін. Проте, ці результати отримано для значно схожих навчальних та тестувальних наборів фонограм мовців і не досліджувалися в умовах реального акустичного оточення, характерним для якого є присутність не лише періодичних шумів, притаманних каналам зв'язку, а і стохастичних природних чи техногенних шумів.

Роботи, які можна віднести до другого підходу, демонструють близькі уявлення щодо методів оброблювання мовних сигналів у задачах розпізнавання мови/мовця із подальшою відмінністю у інтерпретації отриманих результатів, яка виконується не автоматично, із незмінним порогом чутливості, що знижує ступінь адаптивності створюваних автоматизованих систем розпізнавання на їх основі, а отже, знижує робастість цих систем. Перспективним є синтез адаптивних методів аналізу мовних сигналів для виділенням індивідуальних особливостей мовлення, які поєднують інформативність із універсальністю та обчислювальною ефективністю.

3 МАТЕРІАЛИ І МЕТОДИ

Застосування згортального нейромережевого класифікатора у АСРМКЗ вимагає інформативного візуального представлення факторів, що описують мовний сигнал. Очевидно, що для цього найкраще підходить спектральне представлення, отримане на виході банку фільтрів

$$MF[m,l] = \sum_{k=0}^{1/2K-1} \left| x\left[m, e^{j\omega_k}\right] h_l\left[e^{j\omega_k}\right] \right|^2, \quad (1)$$

де $\omega_k = 2\pi k / F_d$.

Виділимо з фонограми мовного сигналу множини мел-частотних кепстральних коефіцієнтів [8] і коефіцієнтів лінейного прогнозу [9]. Для кожного фрейму на виході банку фільтрів (1) обрахуємо n MFCC-коефіцієнтів

$$MFCC[n] = L^{-1} \sum_{l=1}^{L} \log\left(MF[m,l] \cos\left[2\pi L^{-1}(l+0{,}5)n\right]\right), \quad (2)$$

де n PLP-коефіцієнтів на основі коефіцієнтів лінійного прогнозу $b(i)$ за рекурсій ним відношенням

$$PLP[n] = -b(n) + n^{-1} \sum_{i=1}^{n-1} (n-i) b(i) PLP[n-i]. \quad (3)$$

MFCC та PLP кепстральні коефіцієнти за рахунок розташування фільтрів у банку за Мел-шкалою є стійки-

ми до лінійних спектральних спотворень і відповідають базовим властивостям моделей слухового сприйняття С. Снефа, О. Гітза, Р Лайона [10], але є усередненним представленням відповідних спектральних характеристик і не дозволяють корегування спектральних характеристик мовного сигналу на частотах, які не входять у критичні смуги слухового сприйняття [11]. Для компенсації цих недоліків Х. Занг, М. Хейтц, И. Брюс и Л. Кані [12] сформулювали модель, яка описує відгук активності слухового нерву на мовний сигнал, супроводжуваний шумом, яка дозволила сформулювати метод представлення мовних сигналів множиною нормалізованих за потужністю кепстральних коефіцієнтів PNCC. Таке представлення дозволяє виконати компенсацію шуму за оцінками середньої у часі потужності, отриманої усередненням на протязі кількох фреймів короткочасної оцінки спектру потужності на виході банку фільтрів (1). Оцінюючи змінний у часі поріг шуму і віднімаючи його від короткочасної оцінки спектра потужності отримаємо чисту мовну складову фонограми:

$$Q_{\text{вих}}[m,l]=\begin{cases}\lambda_a Q_{\text{вих}}[m-1,l]+(1-\lambda_a)Q_{\text{вх}}[m,l],\\ \text{якщо } Q_{\text{вх}}[m,l]\ge Q_{\text{вих}}[m-1,l],\\ \lambda_b Q_{\text{вих}}[m-1,l]+(1-\lambda_b)Q_{\text{вх}}[m-1,l],\\ \text{якщо } Q_{\text{вх}}[m,l]< Q_{\text{вих}}[m-1,l].\end{cases} \quad (4)$$

Зауважимо, що точність обчислення коефіцієнтів PNCC цілком визначається точністю процедури детектування мовної активності (для визначення енергії шуму пауз). Отже, після отримання відкоригованого мовного сигналу до нього застосовують ідеальний лінійний напівперіодний випрямляч, після чого мовний сигнал обробляється у першому каналі із повторним застосуванням фільтру (4) для визначення порогового рівня потужності тільки для фреймів пауз. Одночасно у другому каналі до мовного сигналу застосовується процедура часового маскування

$$Q_{tm}[m,l]=\begin{cases}Q_0[m,l],\\ \text{якщо } Q_0[m,l]\ge \lambda_t Q_P[m-1,l],\\ \mu_t Q_P[m-1,l],\\ \text{якщо } Q_0[m,l]< \lambda_t Q_P[m-1,l].\end{cases} \quad (5)$$

Сформований таким чином пороговий енергетичний детектор мовної активності вибирає для формування вектора PNCC значення з першого (4) або другого (5) каналу за правилом

$$PNCC_n[m,l]=\begin{cases}\max\left(Q_{tm}[m,l],Q_f[m,l]\right), & \text{"мова"},\\ Q_f[m,l], & \text{"пауза"}.\end{cases} \quad (6)$$

Наглядно операції для визначення векторів факторів за (2), (3) і (6) представлено на рис. 1.

Зауважимо, що у процесі отримання PNCC застосовувався банк гаматонних фільтрів [12], центральні частоти яких лінійно розподілені у частотному діапазоні 200– 8000 Гц відповідно до шкали ERB, яку використовують у психоакустиці для моделювання акустичних фільтрів. Перехід від частотної шкали у ERB виконується за формулою $ERBs = 21{,}41\log_{10}(1+f/229)$. До переваг шкали ERB окрім того, що вона адекватно відтворює такі властивості сприйняття як псевдологарифмічне зростання ширини критичної смуги із зростанням частоти і логарифмічний закон сприйняття інтервалів частот, можна віднести нечутливість до биття і інтермодуляціям між сигналом і фоновим шумом. Приклади MFC- та PN-спектрограм при різному відношенні сигнал/шум наведено на рис. 2.

Фильтри Габора [13] відносять до смугових фільтрів, які, переважно, використовують у задачах визначення крайових ефектів зображень, і дозволяють виявити діапазон частот сигнала у визначеному проміжку x і напрямку y. Імпульсна передатна характеристика фільтра Габора є добутком гаусової функції на гармонічну:

$$Gb[x,y]=K\frac{1}{\sqrt{2\pi\sigma_x}}e^{\frac{-x^2}{2\pi\sigma_x^2}}\frac{1}{\sqrt{2\pi\sigma_y}}e^{\frac{-y^2}{2\pi\sigma_y^2}}\cdot\cos(2\pi f_0 x+\varphi). \quad (7)$$

В процесі фільтрації відбувається згортання вхідного сигналу, в якості якого може виступати будь-яке зображення im, наприклад MFC- або PN-спектрограма, і фільтра Gb у просторовій області. Процес згортання у просторовій області при переході до частотного простору замінюють на множення. В результаті $F(im\cdot Gb)=F(im)\cdot F(Gb)$ – множення в частотній області амплітуди частот вхідного сигналу, близькі до частоти відповідного синусоїдального сигналу, підсилюються, а решта – затухають.

На ефективність фільтрів Габора критично впливають значення висоти σ_x й ширини σ_y гаусової компоненти (7). При обробці зображень емпірично виявлено, що оптимальна чіткість фільтру досягається при $\sigma_x=\frac{1}{\sqrt{2\pi f_0}}$ і $\sigma_y=\frac{3}{2}\sigma_x$, отже у подальший дослідженнях як початкові використовуватимуться ці величини.

Протягом останнього десятиліття ряд фізіологічних експериментів на різних видах ссавців показав, що нейрони у первинній слуховій корі чутливі до певних спектрально-темпоральних структур, названих спектрально-темпоральними рецептивними полями STRF [14], які є функціональними дескрипторами лінійної обробки змінних акустичних спектрів слуховою системою. Отримані на основі результатів цих досліджень спектрально-темпоральні ознаки увійшли до моделі STRF, яку почали використовувати у системах розпізнавання мови і мовців. Ряд досліджень [14, 15] дозволили визначити множину фільтрів із спектральною, темпоральною та спектрально-темпоральною модуляцією, які у першому наближенні моделюють шаблони збудження нейронів для типових спектрально- темпоральних складових вхідних сигналів. Зокрема, дослідження роботи первинної слухової кори головного мозку виявили множину одиночних пульсацій із синусоїдально-модульованими спектральними профілями з усталеними частотами у просторах часу і логарифмічної частоти, яка може ефективно описувати рецептивні поля та передатні

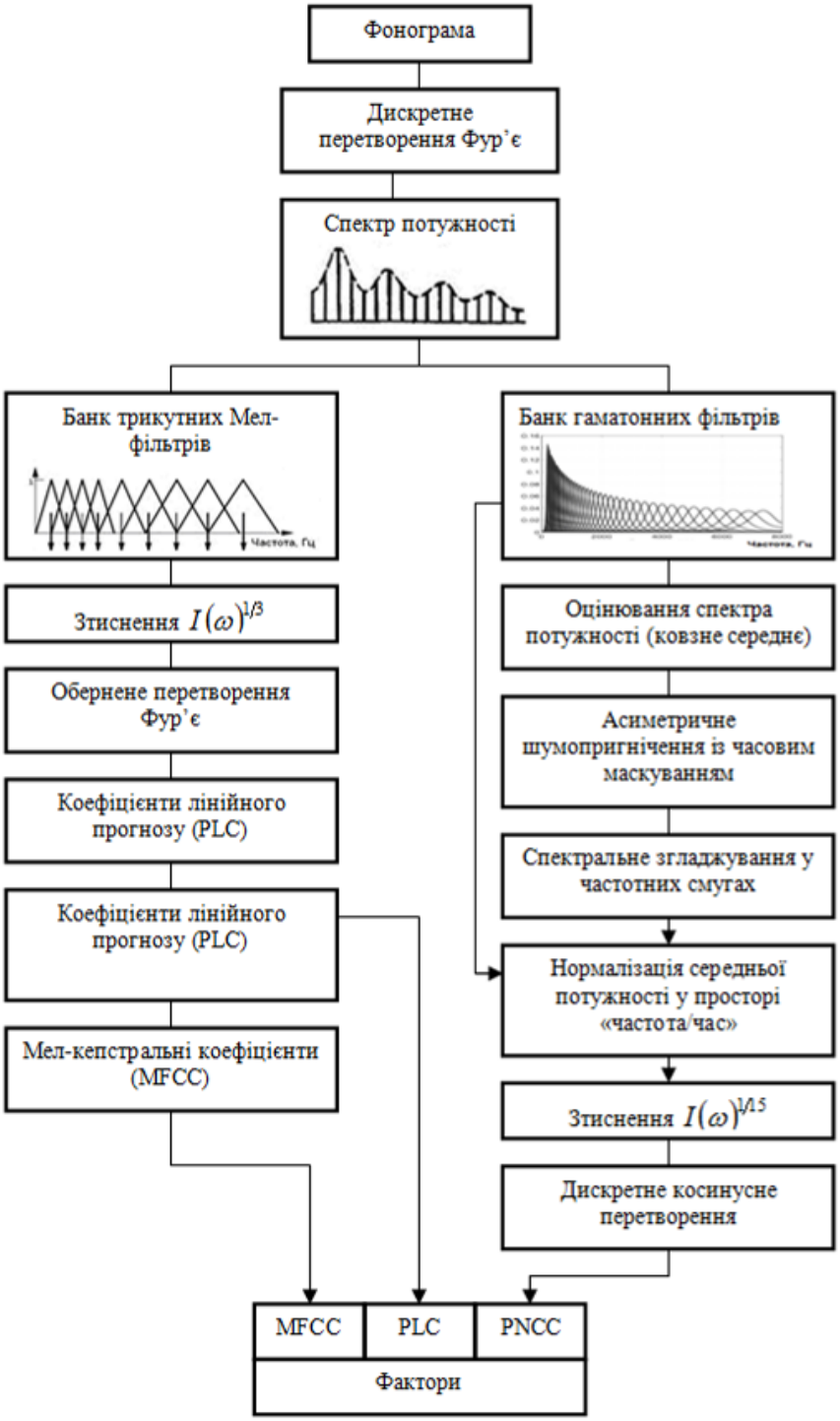


Рисунок 1 – Алгоритм отримання факторів з фонограм

функції нейронів слухової кори голосного мозку $O[n_t,k]=\Delta A\sin\left(2\pi\omega_{n_t}n_t+2\pi\omega_k k+\Phi\right)$. STRF-обробка описується процедурою фільтрації виду $O[n_t,k]=STRF[n_t,k]*x[n_t,k]$, тобто у довільний момент часу n_t і на частоті k реакція нейронів $O[n_t,k]$ описується згортанням STRF та динамічного спектру подразника в околі моменту часу і частоти $x[n_t,k]$. Відповідно, STRF діє як фільтр, який видає піки як реакцію на вхідні сигнали, що за своїми характеристиками наближаються до спектрально-темпоральних ознак, що описують запам'ятовані образи.

Для практичної обробки мовних сигналів зазвичай використовується апроксимація STRF 2D функцією Габора. Для синтезу фільтру Габора, який моделюватимуть STRF, помножимо комплексну синусоїду на віконну функцію Хеннінга:

$$STRF[n_t,k]\cong Gb[n_t,k]=s[n_t,k]\cdot h[n_t,k], \qquad (8)$$

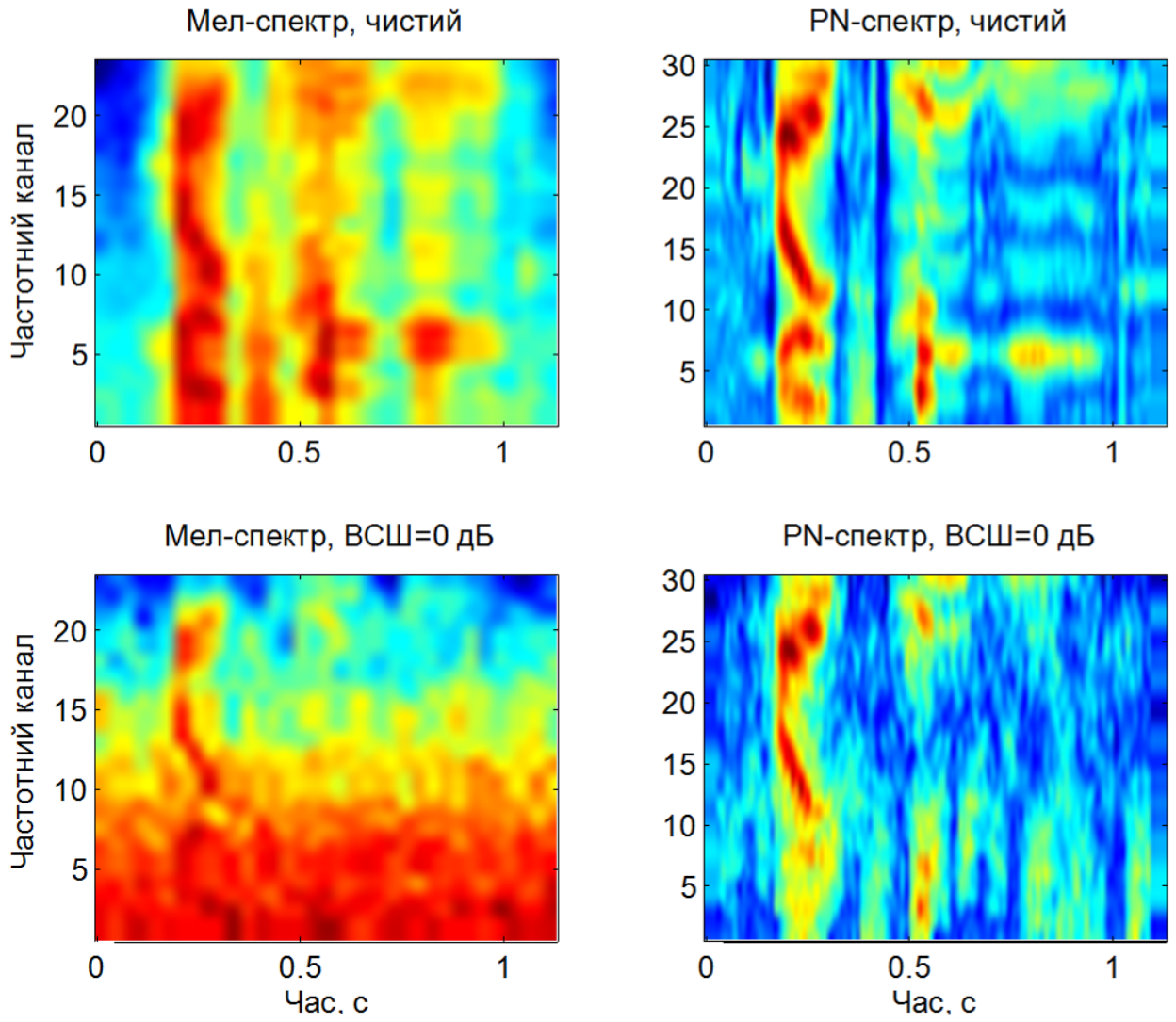


Рисунок 2 – Приклади візуалізація Мел- та PN-спектрів без/з шумом

де комплексну синусоїду, з темпоральною частотою модуляції ω_n і спектральною частотою модуляції ω_k, представлено як $s[n_t,k]=e^{i\omega_n n+i\omega_k k}$, а віконну функцію Хеннінга – як $h[n_t,k]=\left(\frac{1}{2}\left(1-\cos\frac{2\pi n_t}{W_{n_t}+1}\right)\right)\left(\frac{1}{2}\left(1-\cos\frac{2\pi k}{W_k+1}\right)\right)$, де параметри W_n і W_k визначаються як 1.75 цикли відповідної частоти модуляції: $W_{n_t}=1{,}75\frac{2\pi}{|\omega_{n_t}|}$, $W_k=1{,}75\frac{2\pi}{|\omega_k|}$. Для чисто темпоральних або спектральних фільтрів (8) дає нескінченну допоміжну функцію, отже, обмежимося 40 частотними каналами або 99 часовими фреймами, що відповідає максимальній довжині інших фільтрів у відповідних вимірах.

За рахунок налаштовуваних параметрів спектральної та темпоральної частоти модуляції функції Габора матимуть різну інтенсивність і нахил в залежності від кількості коливань у вхідному сигналі. Банк фільтрів Габора, використаний авторами, містить 59 фільтрів, налаштованих на різні темпоральні та спектральні частоти модуляції (темпоральні частоти модуляції ω_n, Гц: 0; 1,9; 3,9; 6,2; 9,9; 15,7; 25; спектральні частоти модуляції ω_k, циклів/октаву: –0,25; –0,1224; –0,06; –0,0293; 0; 0,0293; 0,06; 0,1224; 0,25), які візуалізовано на рис. 3.

При фільтрації PN-спектру створеним банком фільтрів Габора (див. приклад на рис. 4) фільтри із високою частотою модуляції (вузькі фільтри) фільтрують швидкозмінювану частину спектра, а фільтри із низькою модуляційною частотою (широкі фільтри) фільтрують низькочастотну область мовного сигналу. Комплекс вузьких та широких фільтрів, що відповідають різним спектральним частотам модуляції, генерує ознаки, які описують спектральну динаміку вхідного сигналу.

Описаний вище математичний апарат виділення з фонограм факторів та їх представлення після фільтрації Габора є досить ресурсоємним, отже, важливою задачею є здійснення факторного аналізу отриманих з фільтрів ознак для підвищення їх інформативності і зменшення кількості обчислень. Зазвичай у системах розпізнавання мови/мовця застосовують метод главних ком-

понент PCA для отримання лінійної комбінації базових факторів із максимальною варіативністю. В основному, головні вектори є щільними (тобто містять ненульові значення), що ускладнює інтерпретацію результатів PCA. Отже, пропонується використати для оптимізації простору інформативних ознак розріджений метод главних компонент SPCA [16], позбавлений цього недоліку. Цільова функція SPCA отримується із класичної PCA-функції введенням l_1-штрафної норми і набуває виду:

$$\max v^T \Sigma_x v - \rho\left(\|v\|_1^2\right),\ \|v\|_2 = 1. \qquad (8)$$

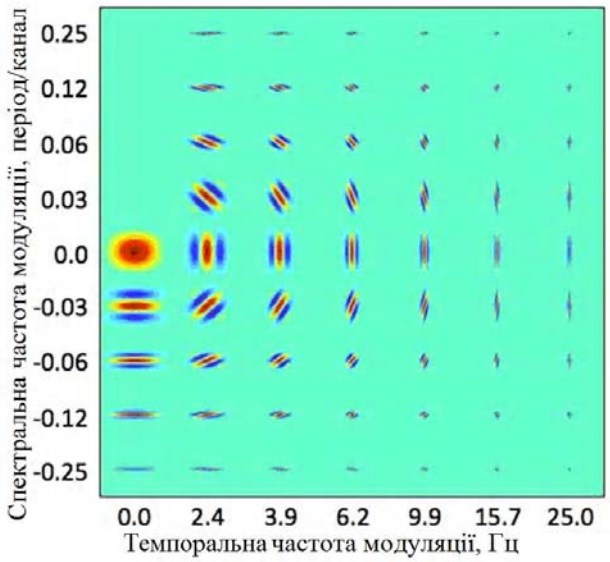


Рисунок 3 – Частотна роздільна здатність створеного банку фільтрів Габора

Перший доданок (8) є цільовою функцією класичного PCA. Вираз (8) можна перетворити на

$$\max Tr\left(\Sigma_x, vv^T\right) - \rho\left(1^T \left|vv^T\right| 1\right),\ \|v\|_2 = 1, \qquad (9)$$

Для підвищення надійності аналізу представимо l_1-штрафну норму у вигляді

$$\rho\left(1^T \left|vv^T\right| 1\right) = \min_U Tr\left(Uvv^T\right),\ -\rho \le U_{ij} \le \rho \qquad (10)$$

і узагальнимо вирази (8)–(10) у вигляді виразу (11)

$$\max \min Tr\left(\left(\Sigma_x + U\right)vv^T\right),\ \|v\|_2 = 1,\ -\rho \le U_{ij} \le \rho, \qquad (11)$$

який, із врахуванням кінцевого результату, представимо як

$$\max \min v^T\left(\Sigma_x + U\right)v,\ \|v\|_2 = 1,\ -\rho \le U_{ij} \le \rho \qquad (12)$$

або

$$\min_U \lambda_{\max}\left(\Sigma_x + U\right),\ -\rho \le U_{ij} \le \rho, \qquad (13)$$

де $\lambda_{\max}\left(\Sigma_x + U\right)$ – максимальне власне значення $\left(\Sigma_x + U\right)$.

Із виразу (13) видно, що для реалізації SPCA необхідно шукати мінімальне з найбільших власних значень коваріаційної і шумової матриць. Для прискорення пошуку U застосуємо алгоритм розширеного метода Лагранжа ALM [16] і обмежимося обчисленням лише головного розрідженого власного вектора, нульовим значенням у якому відповідатимуть ненадійним ознакам.

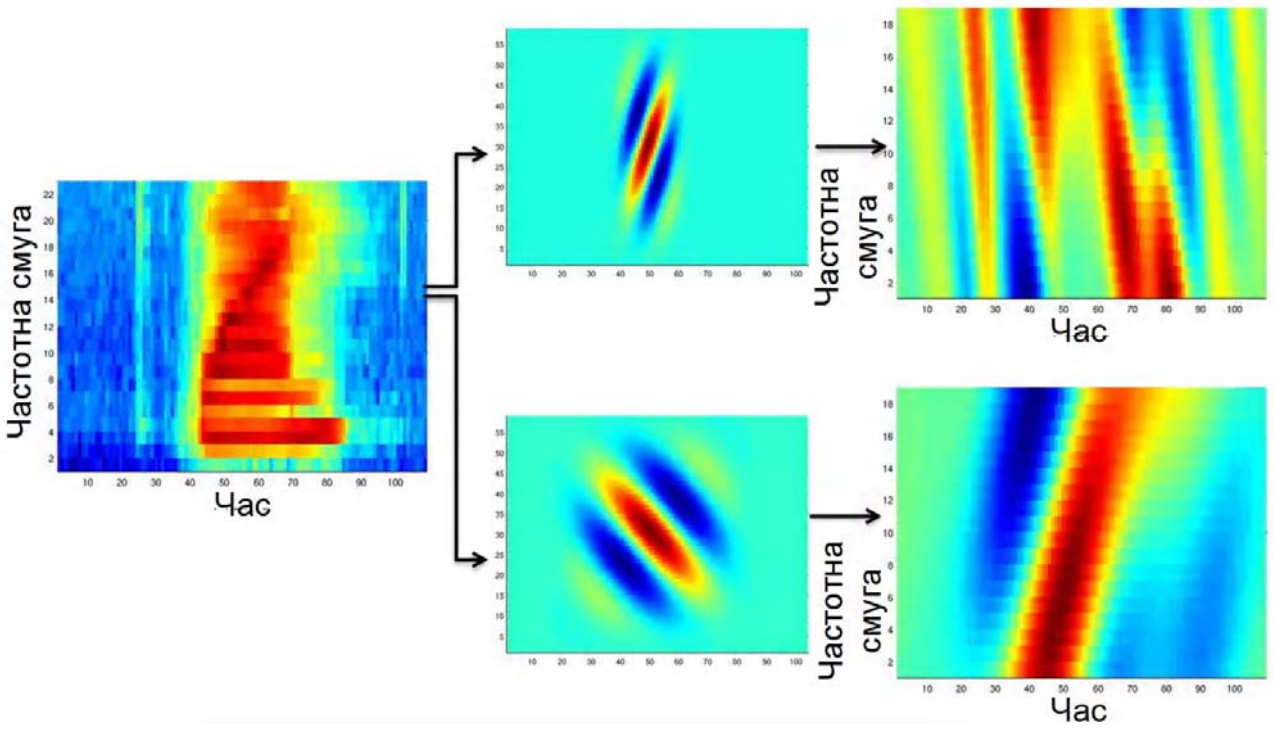


Рисунок 4 – Вище: фільтр із високою темпоральною модуляцією ($\omega_n = 6{,}2$ Гц, $\omega_k = 0{,}03$ цикли/канал) та результат фільтрації; нижче: фільтр із низькою темпоральною модуляцією ($\omega_n = 2{,}4$ Гц, $\omega_k = 0{,}03$ цикли/канал) та результат фільтрації

Отже, для подальшого використання обмежимося лише ознаками, яким відповідають ненульові значення у розрідженому головному векторі.

У створеній АСРМКЗ MFC- та PN-спектрограми, отримані при аналізі вхідних фонограм, після проходження банку фільтрів Габора та оптимізації за допомогою SPCA використовуються в якості вхідних даних для глибокого нейромережевого класифікатора (див. рис. 5), у топологію якого введено вузьке горло, параметри якого після навчання використовуються у HMM/GMM-системі розпізнавання мовця, яку буде описаній згодом.

Маючи візуальні зображення MFC- та PN-спектрограм очевидним є можливість застосування для розпізнавання мовця згортальної нейромережі. Першим (вхідним) шаром згортальної нейромережі є згортальний шар, який виконує процедуру фільтрації вхідного зображення використовуючи фільтр (ядро згортання). Аналіз роботи згортального шару дозволяє передбачити можливість інтеграції 2D фільтрів Габора у топологію згортальної нейромережі, замінивши ядра згортання фільтрами Габора із відповідними часовими і частотними характеристиками. Модифікована таким чином згортальна нейромережа, зображена на рис. 6, використовує коефіцієнти фільтрів Габора як параметри ядра згортання і виконує їх оптимізацію під час навчання, що робить таку нейромережу потенційно більш ефективною за звичайні CNN або DNN нейромережі, де передфільтрація Габора викнується із незмінними параметрами, адже винесена за межі класифікатора.

Автори передбачають, що описана модифікація згортального шару забезпечується застосуванням лінійних функцій активації нейронів шару замість сигмоїдних та встановленням параметру перекривання ядер рівним нулю, що робить згортальний шар передатною функцією фільтрів рецептивних полів. Розміри рецептивних полів для отримання однакових із фільтрами Габора частотно-часових параметрів встановлюємо рівними $2K_m+1$ (смуг) на $2N_m+1$ (відліків) для відповідного фільтру m. Коефіцієнти фільтра Габора інтегруються у вхідні ваги нейромережі W_m згідно відношення

$$h_m[n_t,k]=lin\left(\sum_{i=-N_m}^{N_m}\sum_{j=-K_m}^{K_m}Gb_m(-i,-j)\cdot x(n_t+i,k+j)\right)=$$

$$=Gb_m[n_t,k]*x[n_t,k]. \qquad (14)$$

Ваги зв'язків W_m виконують фільтрацію у рецептивному полі, де індекси коефіцієнтів фільтрів обернені індексам ваг як по вертикалі так і по горизонталі відповідно рівняння (14).

Визначним критерієм для створюваної системи є надійність розпізнавання мовців, яку можна суттєво підвищити використавши поряд із нейромережевою класифікацією інший ефективний метод прийняття рішень для співставлення отриманих результатів, як це показано на рис. 5.

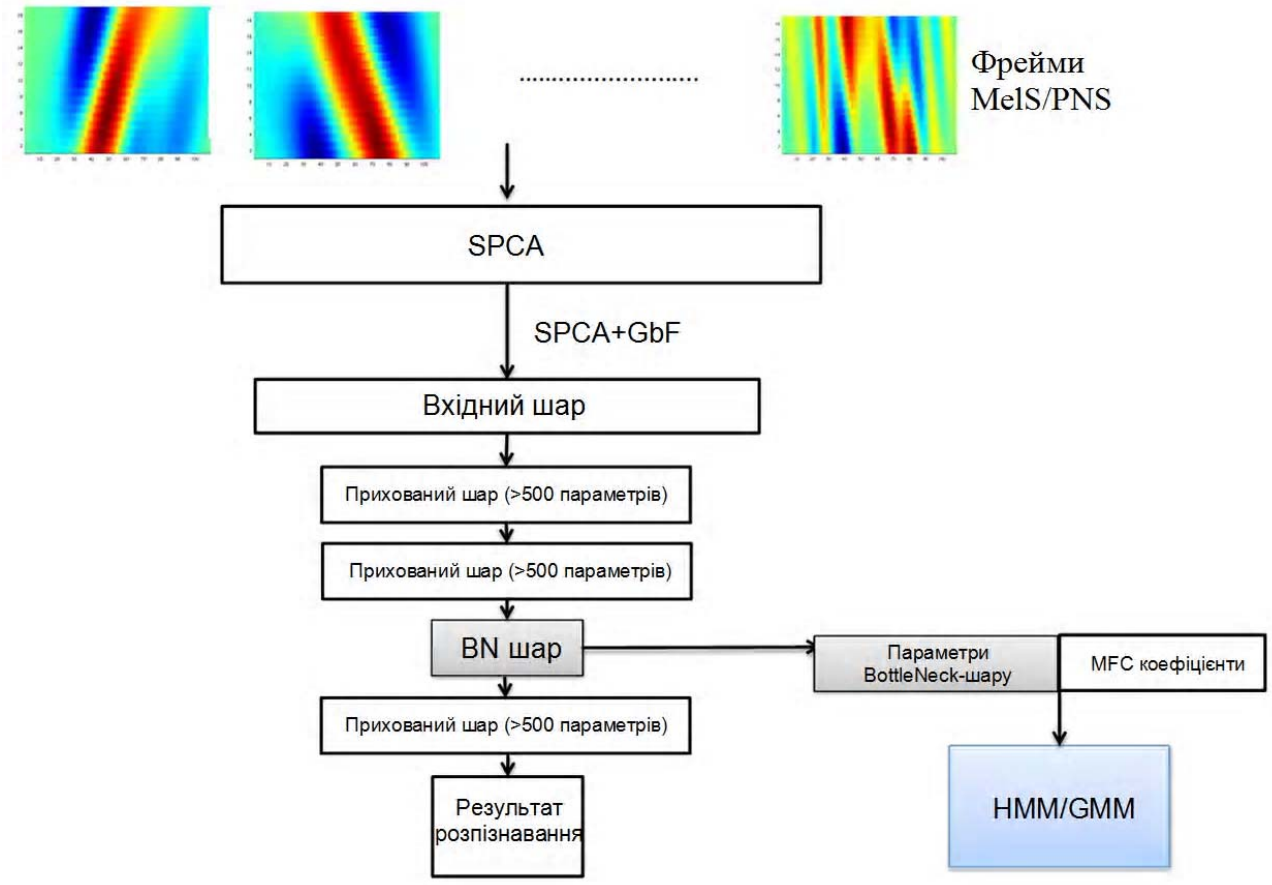


Рисунок 5 – Архітектура АСРМКЗ із глобоким нейромережевим класифікатором і можливістю HMM/GMM аналізу

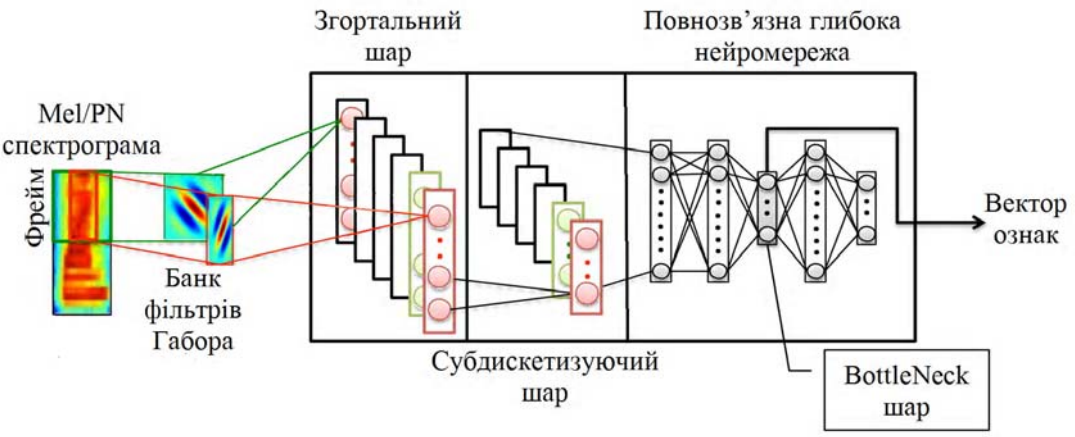


Рисунок 6 – Архітектура модифікованого згортального нейромережевого класифікатора GbCNN АСРМКЗ

У сучасних системах розпізнавання мовця використовують метод апроксимації щільності імовірності в просторі ознак GMM і метод HMM [17]. Гаусова суміш є зваженою сумою M щільності імовірності, яка описується відношенням $p(x|\lambda)=\sum_{i=1}^{M} p_i b_i(x)$, де x – D-вимірний вектор випадкових величин, $b_i(x)$ – функції щільності розподілу складових моделі, p_i – ваги компонентів суміші. Остаточно модель гаусової суміші описується у вигляді вектора $\lambda_i=\{\mu_i,\Sigma_i,p_i\}$, $i=1,\ldots,M$. У нашому випадку, кожному i-му мовцеві відповідає унікальна модель λ_i. Практично процес знаходження оптимальних параметрів λ_i вирішується алгоритмом оцінювання-максимізації [17], робота якого аналогічна алгоритму Баумана-Велча для оцінюванню параметрів прихованих марковських моделей [17]. Так якщо відомий початковий вектор ознак λ_i і обчислений за певними правилами вектор $\bar{\lambda}_i$, то за умови, що $p(x|\bar{\lambda}_i)\geq p(x|\lambda_i)$, вектор $\bar{\lambda}_i$ вважається базовим для повторної ітерації обчислень, які повторюються поки зберігається позитивна динаміка покращення параметрів моделі гаусової суміші i-го мовця. Розширюючи ці міркування на процес розпізнавання мовця, задачу GMM-HMM класифікації можна сформулювати так. Нехай множину мовців $S=\{s_1,s_2,\ldots,s_n\}$ описано множиною гаусових сумішей $\Lambda=\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$, тоді задача класифікації полягає у знаходженні моделі мовця із найбільшим значенням апостеріорної імовірності для визначеної парольної фрази: $S=\arg\max_{1\leq k\leq n}\Pr(\lambda_k|X)=\arg\max_{1\leq k\leq n}\frac{p(x|\lambda_k)\Pr(\lambda_k)}{\Pr(x)}$. Цей вираз за умови рівноімовірної появи мовців $\left(\Pr(\lambda_k)=1/n\right)$, однакового значення $\Pr(x)$ для всіх моделей та за умови незалежності процедур розпізнавання можна переписати так: $S=\arg\max_{1\leq k\leq n}\sum_{i=1}^{M}\log p(x_i|\lambda_k)$, де $p(x_i|\lambda_k)$ – є гаусовою сумішшю, із описаними вище характеристиками.

У сучасних системах розпізнавання мовця використовують метод апроксимації щільності імовірності в просторі ознак GMM і метод HMM [17]. Гаусова суміш є зваженою сумою M щільності імовірності, яка описується відношенням $p(x|\lambda)=\sum_{i=1}^{M} p_i b_i(x)$, де x – D-вимірний вектор випадкових величин, $b_i(x)$ – функції щільності розподілу складових моделі, p_i – ваги компонентів суміші. Остаточно модель гаусової суміші описується у вигляді вектора $\lambda_i=\{\mu_i,\Sigma_i,p_i\}$, $i=1,\ldots,M$. У нашому випадку, кожному i-му мовцеві відповідає унікальна модель λ_i. Практично процес знаходження оптимальних параметрів λ_i вирішується алгоритмом оцінювання-максимізації [17], робота якого аналогічна алгоритму Баумана-Велча для оцінюванню параметрів прихованих марковських моделей [17]. Так якщо відомий початковий вектор ознак λ_i і обчислений за певними правилами вектор $\bar{\lambda}_i$, то за умови, що $p(x|\bar{\lambda}_i)\geq p(x|\lambda_i)$, вектор $\bar{\lambda}_i$ вважається базовим для повторної ітерації обчислень, які повторюються поки зберігається позитивна динаміка покращення параметрів моделі гаусової суміші i-го мовця. Розширюючи ці міркування на процес розпізнавання мовця, задачу GMM-HMM класифікації можна сформулювати так. Нехай множину мовців $S=\{s_1,s_2,\ldots,s_n\}$ описано множиною гаусових сумішей $\Lambda=\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$, тоді задача класифікації полягає у знаходженні моделі мовця із найбільшим значенням апостеріорної імовірності для визначеної парольної фрази: $S=\arg\max_{1\leq k\leq n}\Pr(\lambda_k|X)=\arg\max_{1\leq k\leq n}\frac{p(x|\lambda_k)\Pr(\lambda_k)}{\Pr(x)}$. Цей ви-

раз за умови рівноімовірної появи мовців $\left(\Pr(\lambda_k) = 1/n\right)$, однакового значення $\Pr(x)$ для всіх моделей та за умови незалежності процедур розпізнавання можна переписати так: $S = \arg\max_{1\le k\le n} \sum_{i=1}^{M} \log p(x_i|\lambda_k)$, де $p(x_i|\lambda_k)$ – є гаусовою сумішшю, із описаними вище характеристиками.

4 ЕКСПЕРИМЕНТИ

У наведеному матеріалі автори обгрунтували ряд удосконалень АСРМКЗ, як то, використання факторів на основі PN-представлення мовних сигналів для опису індивідуальності мовлення; використання банку фільтрів Габора у складі АСРМКЗ; застосування SPCA для аналізу факторів опису індивідуальності мовлення; можливість інтеграції фільтрів Габора у топологію згортальної нейромережі; застосування GMM/HMM класифікації на основі ознак bottleneck-шару навченого глибокого нейромережевого класифікатора у складі АСРМКЗ. Далі приведемо поставку емпіричних досліджень адекватності запропонованих удосконалень та оцінювання ефекту від їх впровадження.

В якості бази фонограм для навчання та тестування створеної із застосуванням вищеописаних удосконалень АСРМКЗ використано базу записів із безкоштовної бази даних NOIZEUS [2] – спеціалізованої бази даних Школи інжинірингу та комп'ютерних наук Еріка Джонсона при Університеті Техасу в Далласі, США, яка використовується для дослідження алгоритмів покращення звуку і складається з 30 речень англійської розмовної мови, вимовлених трьома чоловіками та трьома жінками (по 5 на кожного диктора, частота дискретизації записів складає 25 кГц, але задля додавання шуму була зменшена до 8 кГц) та записів типових побутових та техногенних шумів. В ході експерименту АСРМКЗ навчали як фонограмами без додавання шумів, так і фонограмами із додаванням шуму. Навчальна вибірка містила 594 фонограми, де до чистого сигналу додавався штучний шум з рівнями шум/сигнал 0 дБ, 5 дБ, 10 дБ, 15 дБ відповідно. Навчання створеної системи проводилося на фонограмах всіх чотирьох типів відповідно до рівня ВСШ, за умови, що серед навчальної вибірки для кожного із мовців була хоча б одна фонограма із ВСШ = 0 дБ. Фонограми навчальної вибірки використовувалися як вхідні дані для синтезу залежних від статі мовця UBM моделей, повних матриць варіативності, моделей. Для детектування інтервалів мовної активності у фонограмах застосовувався двохканальний VAD алгоритм [18]. Інтервали мовної активності тривалістю 3 секунди розбивалися на фрейми тривалістю 30 мс із 15 мс зсувом, із даних яких екстрагувалися 19 MFCC та PNCC коефіцієнтів, їх енергія, перша і друга їх похідні. До кожної чистої фонограми (із рівнем ВСШ=0) навчальної вибірки підмішувався запис акустичних шумів, вид та рівень ВСШ яких обирався випадково із мовної бази. В результаті на одну чисту фонограму припадало десять із рівнем ВСШ 0, 5, 10 або 15 дБ.

Системи *i*-векторів створеної АСРМКЗ базуються на залежних від статі мовця UBM моделях із 1024 сумішами, навчених на мовному матеріалі бази NOIZEUS, і матрицях повної варіативності із 500 факторами, до яких застосовувалися операції внутрікласової коваріаційної нормалізації WCCN [18] і нормалізації довжини i-векторів [19].

Для реалізації повнозв'язної DNN із чотирма прихованими шарами, серед яких третій – bottleneck-шар, містив 25 нейронів, був використаний фреймворк Caffe [19]. Загальна кількість керованих параметрів нейромережі становила близько $3{,}5\cdot10^6$. Для прискорення навчання нейромережі початкові її параметри отримано за допомогою переднавчання із використанням обмежених машин Больцмана [19], а подальше навчання відбувалося за алгоритмом зворотного розповсюдження помилки із коефіцієнтом швидкості навчання рівним 0,008 і використанням значення перехресної ентропії в якості функції втрат. Цей же класифікатор в подальшому інтегрувався у повнозв'язний шар згортальних нейромереж.

При синтезі CNN на згортальному шарі утворювалося 120 фільтрів, кожен з яких накривав 9 частотних смуг і 15 послідовних відліків вхідних зображень. При створеній згортальної нейромережі із інтегрованими фільтрами Габора GCNN розміри ядер фільтрів у часовому просторі знаходитися у межах від 7 до 99 відліків, а для частотного простору – від 7 до 40 смуг.

5 РЕЗУЛЬТАТИ

У таблиці 1 наведено набори інформативних ознак, які далі подавалися на класифікатори для прийняття рішень. Набори 1 і 2 утворені значеннями відповідного виду кепстральних коефіцієнтів без додаткової обробки, у наборах 3 і 4 MF- та PN-кепстральні коефіцієнти фільтрувалися банком фільтрів Габора, які накривали частотний діапазон 0-8000 Гц та піддавалися логарифмічній компресії LC. Набори 5 і 6 утворювалися аналогічно наборам 3 і 4, але замість LC-компресії використовувалася нелінійнай компресія енергії PNC, а у наборах 7 і 8 – розріджений метод главних компонент SPCA. Всі набори представляли собою як набори векторів так і спектрограм для подальшого передавання на вхідні

Таблиця 1 – Базові набори інформативних ознак для розпізнавання мовців

№ п/п	Синтезована інформативна ознака	Вид банку фільтрів	Віднімання енергії пауз	Вид компресії	Фільтрація Габора
1	MFCC	Мел	–	–	–
2	PNCC	Гаматон	+	–	–
3	MFC-Gb	Мел	–	LC	+
4	PNC-Gb	Гаматон	+	LC	+
5	MFC-Gb	Мел	–	PNC	+
6	PNC-Gb	Гаматон	+	ONC	+
7	MFC-Gb	Мел	–	SPCA	+
8	PNC-Gb	Гаматон	+	SPCA	+

шари глибокої нейромережі або згортальної нейромережі відповідно. Також за результатами SPCA досліджено інформативність окремих фільтрів Габора на основі оцінювання варіативності отриманих від них інформативних ознак, результати якого представлено на рисунку 7, де темніші області визначають ступінь важливості фільтрів (з позиції зростання варіативності отриманих ознак). Ознаки, отримані з виходів фільтрів з темпоральною частотою модуляції 0, 2,4 і 3,9 Гц (що відповідає приблизно 1, 0,7 та 0,5 с довжини фільтрів), складають понад 90% від загальної розмірності ітогового вектора ознак.

Для навчання HMM/GMM моделей застосовувалися ваги bottleneck-шару відповідних навчених нейромереж. На рис. 8 наведено середні ІПР в залежності від виду класифікатора і рівня ВСШ.

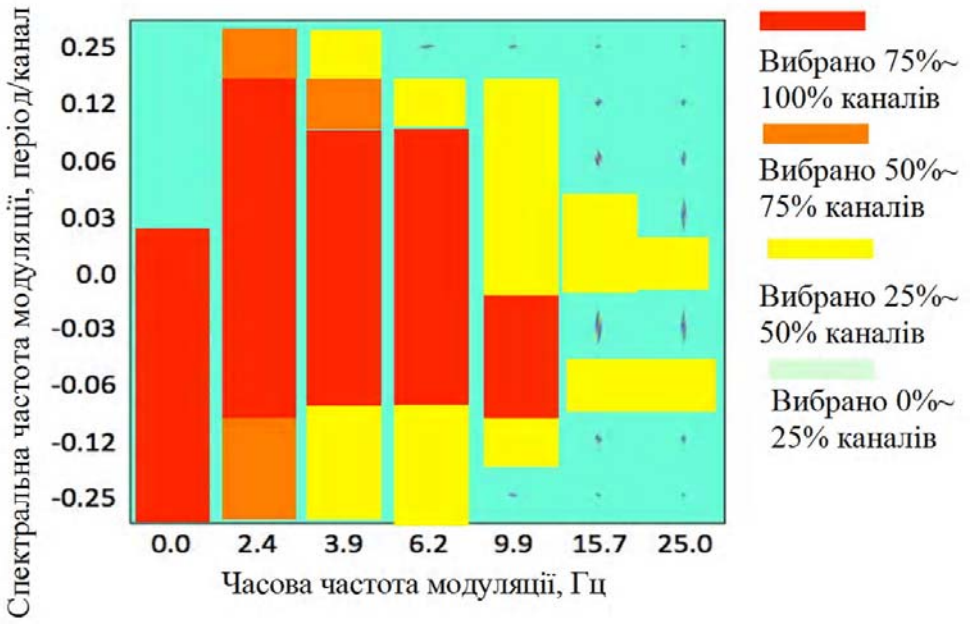


Рисунок 7 – Ефективність фільтрів Габора, оцінювана за результатами SPCA-компресії

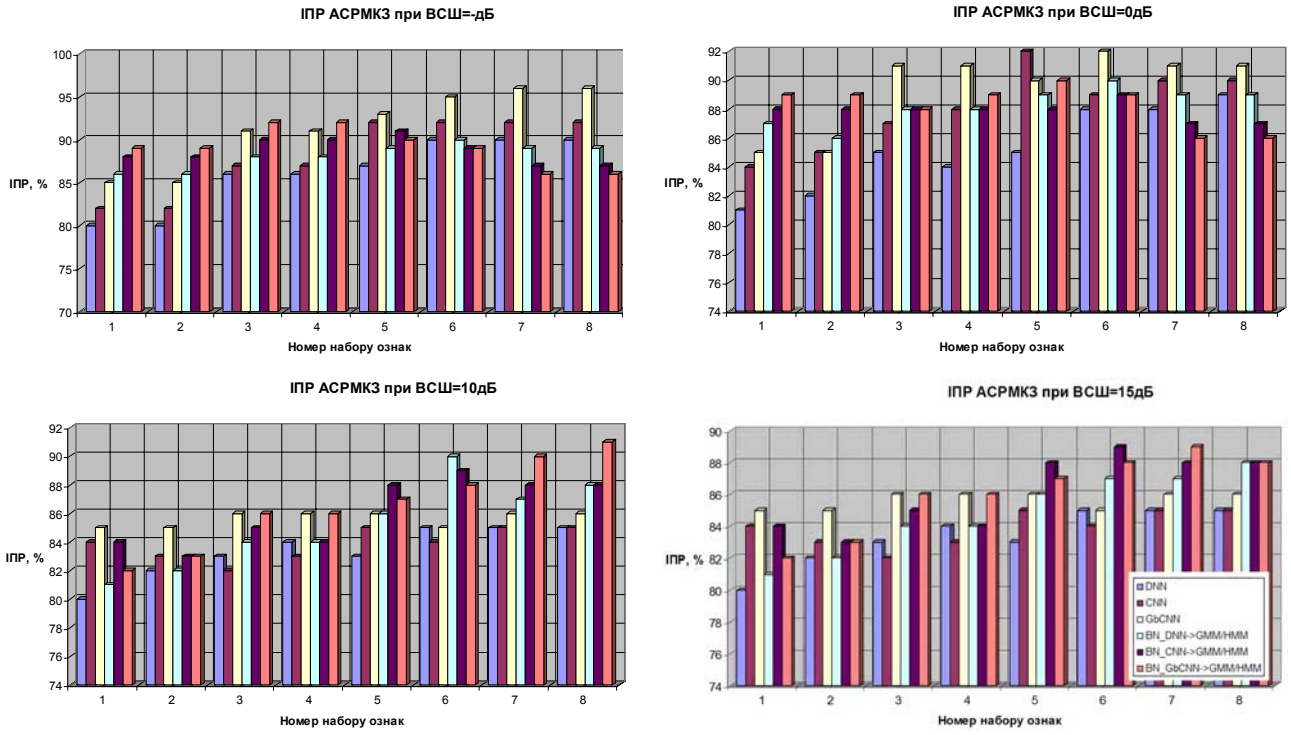


Рисунок 8 – ІПР АСРМКЗ в залежності від виду класифікатора, набору ознак і ВСШ фонограм

6 ОБГОВОРЕННЯ

Результати експериментів, наведені на рис. 8, показали, що запропонована авторами концепція інтеграції фільтрів Габора у згортальний шар CNN дозволяє підвищити як кількісні показники ефективності роботи АСРМКЗ так і підвищити її стійкість до зростання ВСШ, що досягається притаманній CNN адаптивності до мінливості вхідних даних, яка, проте, недостатньо компенсується при високих значеннях ВСШ, що цілком виправдовує застосування контрольного класифікатора, не зважаючи на зростання ресурсоємності ітогової системи.

Як видно з результатів, наведених на рис. 8, із зростанням рівня ВСШ зростає і ефективність HMM/HMM класифікації, що можна пояснити ефективністю роботою UBM-моделі. Також можна відзначити, що набори ознак 7 і 8 демонструють найвищу інформативність, яка зберігається із зростанням ВСШ, що доводить доцільність і адекватність запропонованої процедури факторного аналізу і застосування фільтрації Габора для представлення інформативних ознак.

Використання PNC-ознак також виявляється доцільним при зростанні ВСШ, що обумовлено закладеному у метод розрахунків нормованих за потужністю кепстральних коефіцієнтів заходів для компенсації шумів. Загалом подавання зображень PN-спектрограм на вхід CNN виявилося ефективнішим з MFC- спектрограми із зростанням ВСШ.

Використання розрідженого методу главних компонент дозволило зменшити час на навчання основного і контрольного класифікаторів на 15–17% в залежності від співвідношення «сигнал»/«пауза» у вхідних фонограмах на користь першого класу.

Також наведені на рис. 7 результати доводять більшу інформативність низькочастотної області сигналу, що дозволяє припустити доцільність введення додаткових факторів, які описують індивідуальні особливості мовного процесу у низькочастотній області мовного сигналу, наприклад, характеристик періоду/частоти основного тону.

ВИСНОВКИ

У роботі запропоновано заходи до підвищення стійкості АСРМКЗ до впливу шумів у вхідних фонограмах.

До наукової новизни отриманих результатів можна віднести те, що вперше запропоновано орієнтований на розпізнавання мовця метод інтеграції фільтрів Габора, які імітують роботу слухової системи людини на основі нормалізованих за потужністю кепстральних коефіцієнтів, у вхідний шар згортальної нейронної мережі, що дозволило автоматизовано варіювати чутливістю представлення інформативних для розпізнавання мовця ознак змінюючи параметри фільтрів Габора на етапі навчання нейромережі. Удосконалено нейромережевий класифікатор АСРМКЗ шляхом введення bottleneck-шару у повнозв'язний сегмент згортальної нейронної мережі, що дозволило використовувати його параметри після навчання в якості вхідних даних для контрольного GMM/HMM-класифікатора. Удосконалено спосіб представлення інформативних для розпізнавання мовця ознак у факторному просторі за рахунок застосування розрідженого методу аналізу головних компонент, що дозволяє зменшити довжину вектора інформативних ознак у 2-3 рази із збереженням його інформативності.

Практична цінність отриманих результатів полягає у створенні програмного забезпечення АСРМКЗ, яке реалізує запропоновані наукові результати і дозволяє перевірити їх адекватність у тому числі за даними спеціалізованої бази даних Школи інжинірингу та комп'ютерних наук Еріка Джонсона при Університеті Техасу NOIZEUS.

Проведені дослідження виявили, що при зростанні ВСШ>10 дБ контрольний GMM/HMM-класифікатор виявився ефективнішим за нейромережеві щодо імовірності прального розпізнавання мовців, що можна пояснити ефективністю використаних UBM-моделей, але він є і суттєво ресурсоємнішим. Отже, можливим напрямом подальших досліджень може стати підвищення обчислювальної ефективності GMM/HMM-класифікатора АСРМКЗ.

ПОДЯКИ

Роботу виконано в рамках кафедральній науково-дослідній роботи № 46К4 «Методи моделювання та оптимізації складних систем на основі інтелектуальних технологій» на кафедрі комп'ютерних систем управління Вінницького національного технічного університету за підтримки колективу кафедри і спорідненої кафедри автоматики та інформаційно-вимірювальної техніки ВНТУ.

СПИСОК ЛІТЕРАТУРИ

1. Kalinli O. Noise adaptive training using a vector Taylor series approach for noise robust automatic speech recognition / O. Kalinli, M. L. Seltzer, A. Acero // [Electronic resource]. – Access mode: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/Ozlem_ICASSP09_final.pdf
2. Ковтун В. В. Оцінювання надійності автоматизованих систем розпізнавання мовців критичного застосування / В. В. Ковтун, М. М. Биков, // Вісник Вінницького політехнічного інституту, Вінниця. – 2017. – № 2. – С. 70–76.
3. Kim C. Feature extraction for robust speech recognition based on maximizing the sharpness of the power distribution and on power flooring / C. Kim, R. M. Stern // [Electronic resource]. – Access mode: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.184.9018&rep=rep1&type=pdf
4. Normalized amplitude modulation features for large vocabulary noise-robust speech recognition / [V. Mitra, H. Franco, M. Graciarena, A. Mandal] // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 25–30 March 2012 : proceedings. – Kyoto, Japan: IEEE, 2012. - P. 4117-4120. DOI: 10.1109/ICASSP.2012.6288824.
5. Speech Processing, Transmission and Quality Aspects (STQ). [Electronic resource]. – Access mode: http://www.etsi.org/deliver/etsi_es/201100_201199/201108/01.01.03_60/es_201108v010103p.pdf
6. Graves A. Speech recognition with deep recurrent neural networks / A. Graves, A. R. Mohamed, G. Hinton // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 26–31 May 2013 : proceedings. – Vancouver, BC, Canada : IEEE, 2013. – P. 6645–6649. DOI: 10.1109/ICASSP.2013.6638947
7. Mohamed A. Acoustic modeling using deep belief networks / A. Mohamed, G. Dahl, G. Hinton // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 31 January 2011 : proceedings. – IEEE, 2011. – P. 14–22. DOI: 10.1109/TASL.2011.2109382

8. Davis S. Comparison of parametric representation of monosyllabic word recognition in continuously spoken sentences / S. Davis, P. Mermelstein // [Electronic resource]. – Access mode: http://www.cs.northwestern.edu/~pardo/courses/eecs352/papers/Davis1980-MFCC.pdf
9. Hermansky H. Perceptual Properties of Current Speech Recognition Technology / H. Hermansky, J. Cohen, R. Stern // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 23 July 2013 : proceedings. – IEEE, 2013. – P. 1968–1985. DOI: 10.1109/JPROC.2013.2252316.
10. Virtanen T. Techniques for Noise Robustness in Automatic Speech Recognition / T. Virtanen, R. Singh, B. Raj // John Wiley & Sons, Ltd, Chichester, UK. – 2012. DOI: 10.1002/9781118392683.ch1.
11. Stern R. Hearing is Believing. Biologically inspired methods for robust automatic speech recognition // R. Stern, N. Morgan // [Electronic resource]. – Access mode: https://pdfs.semanticscholar.org/d4a9/a6aa42dcb2011e45a99b0174da6a47777b7a.pdf
12. Kim C. Power-normalized cepstralcoefficients (PNCC) for robust speech recognitions / C. Kim, R. Stern // [Electronic resource]. – Access mode: http://www.cs.cmu.edu/~robust/Papers/OnlinePNCC_V25.pdf
13. Movellan J. Tutorial on Gabor Filters. [Electronic resource] / J. Movellan. – Access mode: http://mplab.ucsd.edu/tutorials/gabor.pdf
14. Mesgarani N. Speech Processing with a Cortical Representation of Audio / N. Mesgarani, S. Shamma // [Electronic resource]. – Access mode: https://pdfs.semanticscholar.org/f1d8/f93cdb64390b3a65f930cee4346c30bd86e4.pdf
15. Morgan N. Using spectro-temporal features to improve AFE feature extraction for automatic speech recognition / N. Morgan, S. Ravuri // [Electronic resource]. – Access mode: https://pdfs.semanticscholar.org/c7c5/04087f2107f0ea9a3cedeeaf5cc0c48c0c92.pdf
16. Berthet Q. Optimal Detection of Sparse Principal Components in High Dimension / Q. Berthet, P. Rigollet / [Electronic resource]. – Access mode: https://arxiv.org/pdf/1202.5070.pdf
17. Оптимізація алфавіту інформативних ознак для автоматизованої системи розпізнавання мовців критичного застосування / [А. О. Береза, М. М. Биков, А. Д. Гафурова, В. В. Ковтун] // Вісник Хмельницького національного університету, серія: Технічні науки, Хмельницький. – 2017. – №3 (249). – С. 222–228.
18. Mak M. W. A study of voice activity detection techniques for NIST speaker recognition evaluations / M. W. Mak, H. B. Yu // [Electronic resource]. – Access mode: https://pdfs.semanticscholar.org/541f/9cfacdac000aadd57cd33b6d86dc96bc3308.pdf
19. Research of neural network classifier in speaker recognition module for automated system of critical use / [Mykola M. Bykov, Viacheslav V. Kovtun, Andrzej Smolarz et al] // SPIE 10445, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017, 1044521; DOI:10.1117/12.2280930.

Стаття надійшла в редакцію 31.12.2017.
Після доробки 22.01.2018.

Бисикало О. В.[1], Грищук Т. В.[2], Ковтун В. В.[3]

[1]Д-р техн. наук, профессор, декан факультета компьютерных систем и автоматики Винницкого национального технического университета, Винница, Украина

[2]Канд. техн. наук, доцент, доцент кафедры компьютерных систем управления Винницкого национального технического университета, Винница, Украина

[3]Канд. техн. наук, доцент, доцент кафедры компьютерных систем управления Винницкого национального технического университета, Винница, Украина

ОПТИМИЗАЦИЯ КЛАССИФИКАТОРА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ ДИКТОРА КРИТИЧЕСКОГО ПРИМЕНЕНИЯ

Актуальность. Рассмотрены вопросы адаптации сверточного нейросетевого классификатора для использования в автоматизированные системы распознавания диктора критического применения (АСРДКП). Объектом исследования является индивидуальные особенности речевого процесса человека.

Цель работы. Разработка средств выделения из речевого сигнала индивидуальных для диктора признаков, повышение их информативности в результате выполнения факторного анализа, их визуальное представление для использования сверточного нейросетевого классификатора и оптимизация его архитектуры для нужд АСРДКП.

Метод. Предложены меры по оптимизации процедуры классификации диктора АСРДКП, для чего теоретически обоснован оптимальный способ представления информативных признаков и метод повышения их информативности, обосновано топологию и меры по повышению эффективности процесса распознавания диктора. В частности, обоснована целесообразность использования нормализованных по мощности кепстральных коэффициентов PNCC для описания фонограмм, записанных в условиях шумного окружения, предложено использовать фильтры Габора для представления информации, которая будет анализироваться сверточной нейросетью, выбран оптимальный метод факторного анализа, а именно, разреженный метод анализа главных компонент, для уменьшения размерности вектора признаков с сохранением его информативности, предложено усовершенствованную топологию сверточной нейросети для АСРДКП, в которой фильтры Габора интегрированы в сверточный слой, что позволяет оптимизировать их параметры в процессе обучения нейросети, и в полносвязном слое использована глубокую нейросеть с bottleneck-слоем, веса которого после обучения выступают в качестве входных данных для контрольного GMM / HMM-классификатора.

Результаты. Методы представления и оптимизации информативных для распознавания диктора признаков, методы их визуального представления и усовершенствование топологии сверточной нейросети для принятия решений на их основе.

Выводы. Полученные теоретические результаты нашли эмпирическое подтверждение. В частности, доказано устойчивость усовершенствованной сверточной нейросети к присутствию шумов во входных фонограммах, которая оказалась выше показателей обычной сверточной нейросети и глубокой нейросети. При росте ОСШ > 10 дБ контрольный GMM / HMM-классификатор оказался эффективнее нейросетевого, что можно объяснить эффективностью использованных UBM-моделей, но он является существенно более ресурсозатратным. Также эмпирически определены параметры окон банка фильтров Габора предоставляющих наиболее вариативную относительно индивидуальных особенностей речи информацию.

Ключевые слова: автоматизированная система распознавания диктора критического применения, обработка сигналов, нейросеть, факторный анализ.

Bisikalo O. V.[1], Grischuk T. V.[2], Kovtun V. V.[3]
[1]Dr.Sc., Professor, Dean of Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, Ukraine
[2]Ph.D., Associate Professor of Computer Control Systems Department, Vinnitsia National Technical University, Vinnitsa, Ukraine
[3]Ph.D., Associate Professor of Computer Control Systems Department, Vinnitsia National Technical University, Vinnitsa, Ukraine

THE AUTOMATIC SPEAKER RECOGNITION SYSTEM OF CRITICAL USE CLASSIFIER OPTIMIZATION

Context. The questions of adapting the convolution neural network classifier use in automatic speaker recognition system of critical use (ASRSCU) are considered. The research object is the individual features of the human speech process.

Objective. Development of means for separating individual features from the speaker's speech signal, increasing their informativeness as a result of the factor analysis, their visual representation for the use of the convolution neural network classifier, and optimizing its architecture for the needs of ASRSCU.

Method. Measures are proposed to optimize the speaker recognition procedure of the ASRSCU, for which the optimal way of informative features representation and the method of increasing their informativeness are theoretically justified, the topology and measures for increasing of the speaker recognition process efficiency are justified. In particular, it is justified the use of power normalized cepstral coefficients (PNCC) for the description of phonograms recorded in noisy environment conditions. We propose to use Gabor filters to represent information that will be analyzed by a convolution neural network, an optimal method of factor analysis (a sparse main components analyzing method) to reduce of the features vector length while preserving its informativeness, an improved topology of the convolution neural network in which the Gabor filters are integrated in to the convolution layer, which allows them to optimize their parameters during the neural network training process, and in a fully connected layer a deep neural network with a bottleneck layer is used, whose weights after training are uses as inputs for the GMM/HMM control classifier.

Results. Methods of representation and optimization of the speaker's individual features, methods for their visual presentation and improvement of the topology of a convolution neural network for making speaker recognition on their basis.

Conclusions. The obtained theoretical results have found empirical confirmation. In particular, the stability of an improved convolution neural network to the noisy input phonograms proved to be higher than the results of an ordinary convolution neural network and a deep neural network. With an SNR increase up to 10 dB, the GMM/HMM classifier is more efficient than the neural network, which can be explained by the efficiency of the used UBM models, but it is much more resource-intensive. Also, the parameters of the Gabor filter bank frames that provide the most variable individual features from the speech signal for speaker recognition are determined empirically.

Keywords*:* automated speaker recognition system of critical use, signal processing, neural network, feature analysis.

REFERENCES

1. Kalinli O., Seltzer M. L., Acero A. Noise adaptive training using a vector Taylor series approach for noise robust automatic speech recognition, [Electronic resource], Access mode: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/Ozlem_ICASSP09_final.pdf
2. Kovtun V. V., Bykov M. M. Otsiniuvannia nadiinosti avtomatyzovanykh system rozpiznavannia movtsiv krytychnoho zastosuvannia, *Visnyk Vinnytskoho politekhnichnoho instytutu, Vinnytsia*, 2017, No. 2, pp. 70–76.
3. Kim C., Stern R. M. Feature extraction for robust speech recognition based on maximizing the sharpness of the power distribution and on power flooring, [Electronic resource]. Access mode: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.184.9018&rep=rep1&type=pdf
4. Mitra V., Franco H., Graciarena M., Mandal A. Normalized amplitude modulation features for large vocabulary noise-robust speech recognition, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 25–30 March 2012 : proceedings. Kyoto, Japan, IEEE, 2012, pp. 4117–4120. DOI: 10.1109/ICASSP.2012.6288824.
5. Speech Processing, Transmission and Quality Aspects (STQ), [Electronic resource]. Access mode: http://www.etsi.org/deliver/etsi_es/201100_201199/201108/01.01.03_60/es_201108v010103p.pdf
6. Graves A., Mohamed A. R., Hinton G. Speech recognition with deep recurrent neural networks, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 26–31 May 2013, proceedings, Vancouver, BC, Canada, IEEE, 2013, pp. 6645–6649. DOI: 10.1109/ICASSP.2013.6638947
7. Mohamed A., Dahl G., Hinton G. Acoustic modeling using deep belief networks, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, *31 January 2011, proceedings*, IEEE, 2011, pp. 14–22. DOI: 10.1109/TASL.2011.2109382
8. Davis S., Mermelstein P. Comparison of parametric representation of monosyllabic word recognition in continuously spoken sentences, [Electronic resource], Access mode: http://www.cs.northwestern.edu/~pardo/courses/eecs352/papers/Davis1980-MFCC.pdf
9. Hermansky H., Cohen J., Stern R. Perceptual Properties of Current Speech Recognition Technology, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 23 July 2013 : proceedings*, *IEEE,* 2013, pp. 1968–1985. DOI: 10.1109/JPROC.2013.2252316.
10. Virtanen T., Singh R., Raj B. Techniques for Noise Robustness in Automatic Speech Recognition, *John Wiley & Sons, Ltd, Chichester, UK,* 2012. DOI: 10.1002/9781118392683.ch1.
11. Stern R., Morgan N. Hearing is Believing. Biologically inspired methods for robust automatic speech recognition, [Electronic resource]. Access mode: https://pdfs.semanticscholar.org/d4a9/a6aa42dcb2011e45a99b0174da6a47777b7a.pdf
12. Kim C., Stern R. Power-normalized cepstralcoefficients (PNCC) for robust speech recognitions, [Electronic resource]. Access mode: http://www.cs.cmu.edu/~robust/Papers/OnlinePNCC_V25.pdf
13. Movellan J. Tutorial on Gabor Filters. [Electronic resource]. Access mode: http://mplab.ucsd.edu/tutorials/gabor.pdf
14. Mesgarani N., Shamma S. Speech Processing with a Cortical Representation of Audio, [Electronic resource]. Access mode: https://pdfs.semanticscholar.org/f1d8/f93cdb64390b3a65f930cee4346c30bd86e4.pdf
15. Morgan N., Ravuri S. Using spectro-temporal features to improve AFE feature extraction for automatic speech recognition, [Electronic resource]. Access mode: https://pdfs.semanticscholar.org/c7c5/04087f2107f0ea9a3cedeeaf5cc0c48c0c92.pdf
16. Berthet Q., Rigollet P. Optimal Detection of Sparse Principal Components in High Dimension, [Electronic resource]. Access mode: https://arxiv.org/pdf/1202.5070.pdf
17. Bereza A. O., Bykov M. M., Hafurova A. D., Kovtun V. V. Optymizatsiia alfavitu informatyvnykh oznak dlia avtomatyzovanoi systemy rozpiznavannia movtsiv krytychnoho zastosuvannia, *Visnyk Khmelnytskoho natsionalnoho universytetu, seriia: Tekhnichni nauky*, Khmelnytskyi, 2017, No. 3(249), pp. 222–228.
18. Mak M. W., Yu H. B. A study of voice activity detection techniques for NIST speaker recognition evaluations, [Electronic resource]. Access mode: https://pdfs.semanticscholar.org/541f/9cfacdac000aadd57cd33b6d86dc96bc3308.pdf
19. Mykola M. Bykov, Viacheslav V. Kovtun, Andrzej Smolarz, Mukhtar Junisbekov, Aliya Targeusizova, Maksabek Satymbekov Research of neural network classifier in speaker recognition module for automated system of critical use, *SPIE 10445, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments* 2017, 1044521; DOI: 10.1117/12.2280930.

UDC 681.518:004.93.1'

Dovbysh A. S.[1], Velykodnyi D. V.[2], Protsenko O. B.[3], Zimovets V. I.[4]

[1]Dr. Sc., Prof., Head of the Computer Science Department of Sumy State University, Sumy, Ukraine
[2]PhD, Head of the Computer Science Department of the Sumy State University, Sumy, Ukraine
[3]PhD, Associate Professor, Associate Professor of Computer Science Department of the Sumy State University, Sumy, Ukraine
[4]Postgraduate student of the Department of Computer Science of Sumy State University, Sumy, Ukraine

OPTIMIZATION OF PARAMETERS OF MACHINE LEARNING OF THE SYSTEM OF FUNCTIONAL DIAGNOSTICS OF THE ELECTRIC DRIVE OF A SHAFT LIFTING MACHINE

Relevance. The actual task of increasing the functional efficiency of machine learning of the system of functional diagnosis of the electric drive of a hoisting mine machine is solved.

The specific objective of this study was to develop a method for the information synthesis of a learning system for the functional diagnosis of the electric drive of a hoisting mine machine, which allows increasing the reliability and efficiency of diagnostic solutions in accordance with the decisive rules built in the process of machine learning.

Method. The method of information-extreme machine learning of the system of functional diagnosis of the electric drive of a mine-hoisting machine is proposed, based on the maximization of the information capacity of the system in the process of its training. Based on the computer-generated optimal learning parameters of the hyperspherical containers of the recognition classes, within the framework of the geometric approach, decisive rules that are practically invariant to the spatiality of the space of diagnostic features are constructed. In addition, increasing the efficiency of machine learning systems is achieved by parallel-sequential optimization of control tolerances for diagnostic features. In this case, the quasi-optimal control tolerances for diagnostic tests obtained during parallel optimization are used as start-ups for their sequential optimization. As a criterion for optimizing the parameters of machine learning, the modified information measure of Kulbak is used, which is a functional of the accuracy characteristics of diagnostic solutions.

Results. The algorithmic and software for machine learning of the system for functional diagnosis of the electric drive of a mine hoisting machine has been developed, which makes it possible to build decisive rules for the adoption of highly reliable diagnostic solutions when the system is in operation.

Conclusions. The results of physical modeling confirm the operability of the proposed method of machine learning and the developed software of the functional diagnosis system of the electric drive of a hoisting mine machine, which allows them to be recommended for solving practical problems of diagnosing and automatic control of traction machines.

Keywords: informationally-extreme intellectual technology, functional control, learning matrix, machine learning, information criterion, electric drive, mine hoisting machine.

NOMENCLATURE

IEI – informationally-extreme intellectual technology;

HMM – hoisting mine machine;

M – number of recognition classes;

N – number of diagnostic features;

n – number of vectors-implementation classes recognition;

x_m – binary averaging vector-implementation, which vertex determines the geometric center of the hyperspherical container of the recognition class X_m^o;

d_m – code distance determines the radius of the hyperspherical container of the recognition class X_m^o, which is restored in the radial basis of the binary space of the diagnostic features;

$\delta_{K,i}$ – parameter, which is equal to half of the symmetrical tolerances at the i-th diagnostic feature;

$d(x_m \oplus x_c)$ – code distance from the X_m^o class center to the nearest (adjacent) X_c^o class center;

$\delta_{H,i}$ – normed (working) tolerance at the i-th diagnostic feature, which defines the parameter $\delta_{K,i}$ domain;

$E_m^{(k)}$ – information criterion for optimization of machine learning parameters, which is calculated at the k-th step of machine learning;

$\hat{\delta}_K$ – quasi-optimal parameter of the control tolerances for the diagnostic features;

G_δ – tolerance range δ_K parameter of the control tolerances for the diagnostic features;

G_E – working (admissible) function domain of the information criterion;

$\{k\}$ – a set of steps in machine learning;

$G_{\delta,i}$ – tolerance range of the $\delta_{K,i}$ parameter of the control tolerances for the i-th diagnostic feature;

L – the number of the sequential optimization procedure runs for the diagnostic features;

$\otimes$ – character of the repeat operation;

$\alpha_m(d)$ – error of the first kind when making classification decisions. It is calculated in the process of restoration of a hyperspherical container of the recognition class X_m^o with radius d;

$\beta_m(d)$ – error of the second kind when making classification decisions. It is calculated in the process of restoration of a hyperspherical container of the recognition class X_m^o with radius d;

10^{-r} – sufficiently small number, which is entered to avoid dividing by zero $(1 < r \le 3)$;

© Dovbysh A. S., Velykodnyi D. V., Protsenko O. B., Zimovets V. I., 2018
DOI 10.15588/1607-3274-2018-2-5

$\delta^*_{K,i}$ – optimal value of the control tolerance parameter;

ρ_m – selection level of the coordinates of the averaged binary vector-implementation of the recognition class;

$A_{KH,i}[k]$ – lower control tolerance for *i*-th diagnostic feature;

$A_{KB,i}[k]$ – upper control tolerance for *i*-th diagnostic feature;

$y_{1,i}$ – the selective mean value of the *i*-th feature of the averaged vector-implementation of the base class X^o_1;

$E^*_m(d)$ – the maximum value of the information criterion for optimizing the parameters of machine learning of the functional control system to recognize the implementation of the class X^o_m;

G_E – working (admissible) domain of the function information criterion of optimization of machine learning parameters;

G_d – tolerance range of geometric parameters of container classes of recognition;

X^o_1 – the base recognition class, for which a system of control tolerances for diagnostic features is determined;

d^*_m – optimal radius of the recognition class container X^o_m;

x_c – binary averaged vector of the nearest neighbor class X^o_c;

$x^{(j)}$ – vector-implementation, which can be recognized;

μ_m – the implementation function that is recognizable to the hyperspherical container of the recognition X^o_m class.

INTRODUCTION

The requirements for functional efficiency and safety of the hoisting mine machine have been significantly increased with the development of deep deposits of minerals, which is a modern trend in the mining industry. In this, the reliability of the HMM considerably depends on the functionality of its electric drive, the output of which from a given technological regime leads to material losses and even man-made accidents. Therefore, organization of the continuous functional diagnostics of the electric drive and its units becomes of great importance. At the same time the use of traditional methods does not provide high reliability and efficiency of the functional diagnostics because of the multidimensionality of diagnostic features and the intersection of functional states of the HMM.

One of the promising ways to increase the functional efficiency of an automated management system of the HMM is to provide it with an intelligent component based on machine learning and image recognition. Thus, the objective of the study is to develop a computer-aided learning algorithm for the system of functional diagnostics of the HMM within the IEI-technology of data analysis, which is based on maximizing the information capacity of the system in the process of its learning.

.

1 PROBLEM STATEMENT

Consider the formalized statement of the information synthesis of the functional diagnostics system of the HMM electric drive. Let the fuzzy alphabet of recognition classes be given in the general case $\{X^o_m \mid m = \overline{1,M}\}$, and each of them characterizes the functional state of units of the electric drive. An educational matrix of the "object-property" type is formed for the given alphabet $\| y^{(j)}_{m,i} \mid i = \overline{1,N};\ j = \overline{1,n} \|$. In addition, a structured vector of parameters that influences the functional efficiency of machine learning of the system of functional diagnostics is given:

$$g_m = < x_m, d_m, \delta_{K,i} >, \quad (1)$$

In addition, the restrictions have been set:

$$d_m \in [0; d(x_m \oplus x_c) - 1],$$

$$\delta_{K,i} \in [0; \delta_{H,i}/2)].$$

While machine learning it is necessary to:

1) optimize parameters of the vector (1) by the average information criterion

$$\overline{E} = \frac{1}{M} \sum_{m=1}^{M} \max_{G_E \cap G_d} E_m(d). \quad (2)$$

2) set decisive rules according to the obtained during the process of machine learning optimal geometry parameters of the containers recognition classes that will guarantee a high probability of making the correct diagnostic decisions.

During the exam, that is, directly in the operational mode of the functional diagnostics, it is necessary to make a classification decision about the affiliation of the current $x^{(j)}$ vector to one of the recognition classes from a given alphabet.

2 REVIEW OF THE LITERATURE

Modern systems and methods for diagnosing the electric drive are divided into two groups [1]. The first group includes methods of test diagnostics, which measure the insulation resistance, leakage currents, the internal resistance of the windings, the tangent angle of the dielectric losses of the windings, etc. Disadvantages of the test diagnostics are the temporary cessation of the operation of the electric motor, the lack of the ability to outage the equipment during operation to prevent its complete failure, lack of control of equipment operating modes, etc.

The second group includes methods of functional diagnostics of the electric drive [2, 3]. Functional diagnostics of the HMM electric drive refers to the so-called poorly formalized processes. It has to do with reasons reasons such as arbitrary initial conditions due primarily to deviations of the lift load from the calculated one, and the intersection of recognition classes due to the ambiguity of the diagnostic features to the functional states. The analysis of modern approaches and methods of functional diagnostics of electromechanical equipment has shown that in order to increase the accuracy and reliability of its work,

it is necessary to use intelligent data analysis technologies [3, 4]. In this, existing methods of intellectual functional diagnosis are based mainly on artificial neural networks [5–7]. The main disadvantage of most known methods of Data Mining technology, including artificial and immune networks, is for their sensitivity to the multidimensionality of the dictionary of diagnostic features and the alphabet of recognition classes. Solving this problem is especially important for functional diagnostics, which in practice needs large volumes of data to be analyzed. In [8] the use of fuzzy methods of control and control of electric drive is considered. It should be noted that the use of fuzzy methods of presentation and withdrawal of new knowledge in the analysis of data is justified in applying a qualitative measurement scale.

One of the most promising ways of analysis and synthesis of the functional diagnostics systems capable to learn is the use of ideas and methods of the so-called IEI-technology, which is based on maximizing the information capability of the system in the process of machine learning [9, 10]. In [11] within the framework of IEI-technology the task of functional diagnostics of a spin-on electric drive was considered. At the same time information-extreme machine learning was carried out at a given optimal in the informational sense of the system of control tolerances for diagnostic features, which did not allow to obtain high functional efficiency of machine learning diagnostic system.

The article deals with the informationally extreme algorithm of machine learning of the functional diagnostic system of HMM electric drive with parallel-sequential optimization of control tolerances for diagnostic features.

3 MATERIALS AND METHODS

The idea of machine learning within the IEI-technology is the transition from the Euclidean to binary space recognition features, where the adaptation of the working binary training matrix to the maximum information ability of the system to recognize the multidimensional vectors of the images implementation is carried out by admissible transformations. In this, the main goal of machine learning to create highly reliable decisive rules is achieved at each step by restoring the training of optimal classes of recognition classes upon the information criterion (2). It is the construction of the decisive rules within the geometric approach which makes them practically not sensitive to the multidimensional diagnostic space and allows evaluating the current functional state of the equipment in real time.

Let's consider the informationally-extreme algorithm of machine learning of the system of functional diagnostics with a hyperspherical classifier, in which reconstruction of classes recognition containers occurs by the parallel-sequential optimization of the system of control tolerances for diagnostic features. In this, the two-cyclic procedure of the parallel optimization of control tolerances for diagnostic features, in which control tolerances are changed simultaneously at each step of machine learning for all features, is initially implemented:

$$\widehat{\delta}_K = \arg\{\max_{G_\delta}\{\max_{G_E \cap G_d} \overline{E}(d)\}\}. \qquad (3)$$

Quasi-optimal control tolerances for diagnostic features, which have been obtained based on the results of parallel optimization, are considered as initial for sequential optimization. This approach allows increasing both the reliability of decision-making and the efficiency of the algorithm of consistent optimization, since the search for the global maximum of the information criterion does not go beyond the boundaries of the function domain. Sequential optimization of the control tolerances for recognition features in the process of machine learning is carried out in accordance with the procedure of approximation of the information optimization criterion global maximum (2) to the limit value in the interval on which a function is defined:

$$\delta^*_{K,i} = \arg \overset{L}{\underset{l=1}{\otimes}} \left\{ \max_{G_{\delta i}} \left[\frac{1}{M} \sum_{m=1}^{M} \max_{G_{Em} \cap G_d} E_m^{(l)}(d_m) \right] \right\},\ i = \overline{1, N}. \qquad (4)$$

Consider the main stages of implementation of the information-extreme learning algorithm of functional control system of the hoisting mine machine drive with parallel-sequential optimization of control tolerances for diagnostic features.

An algorithm (3) for machine learning with parallel optimization of control tolerances is initially implemented. The input data is an array of input training matrix for a given alphabet of recognition classes and a parameter of normalized tolerances field based on the scale of diagnostic features. The main stages of implementing the algorithm:

1) zeroing of the recognition class meter: $m := 0$;

2) increment of the recognition class meter: $m := m + 1$;

3) resetting of the amend steps counter of the tolerance field parameter to zero: $k := 0$.

4) increment of the amend steps counter of the tolerance field parameter: $k := k + 1$.

5) resetting of the amend steps counter of the container radius class: $d := 0$

6) increment of the amend steps counter of the container radius class: $d := d + 1$;

7) calculation of the lower $A_{KH,i}[k]$ and upper $A_{KB,i}[k]$ control tolerances for all diagnostic features, according to the formulas

$$A_{KH,i}[k] = y_{1,i} - \delta[k]\frac{\delta_{Hi}}{100};\ A_{KB,i}[k] = y_{1,i} + \delta[k]\frac{\delta_{Hi}}{100}, \qquad (5)$$

8) the formation of a binary training matrix $\left\| x^{(j)}_{m,i} \right\|$ by the rule

$$x^{(j)}_{m,i} = \begin{cases} 1, & \text{if } A_{KH,i} \le y^{(j)}_{m,i} \le A_{KB,i}; \\ 0, & \text{if else.} \end{cases} \qquad (6)$$

9) calculation of the binary averaged vector x_m for X^o_m class by the rule

$$x_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n}\sum_{j=1}^{n} x^{(j)}_{m,i} > \rho_m; \\ 0, & \text{if else}, \end{cases}$$

where the level of selection of the averaged binary vector-realization coordinates of the X_m^o class recognition by default is equal to $\rho_m = 0{,}5$;

10) pairwise partition of the recognition classes averaged vectors set by the nearest-neighbors method;

11) formation of the training matrix for the pair partition element $\Re_m^{|2|} =< x_m, x_c >$;

12) calculation of the information criterion for optimization of the functional diagnostics system learning parameters, for example, in the form of modification of the information measure of Kulbak [6]

$$E_m(d) = \{1 - [\alpha\ (d) + \beta_m(d)]\} \log_2 \left(\frac{2 - (\alpha_m(d) + \beta_m(d)) + 10^{-r}}{\alpha_m(d) + \beta_m(d) + 10^{-r}} \right); \quad (7)$$

13) if $d < d(x_m \oplus x_c)$, then 6 is fulfilled, otherwise – point14;

14) if $k < \delta_H / 2$, then 4 is fulfilled, otherwise – point 15;

15) if $m \le M$, then 2 is fulfilled, otherwise – point16;

16) calculation of the information optimization criterion (5) of machine learning parameters that is averaged across the alphabet of recognition classes;

17) calculation of the quasi-optimal parameter $\widehat{\delta}$ by the formula (3);

18) calculation of quasi-optimal lower and upper control tolerances for diagnostic features by the formulas (5), respectively:

$$\widehat{A}_{HK,i} = y^*_{1,i} - \widehat{\delta}\frac{\delta_{Hi}}{100}, \quad \widehat{A}_{BK,i} = y^*_{1,i} + \widehat{\delta}\frac{\delta_{Hi}}{100}.$$

19) STOP.

The input data for the algorithm (4) of the sequential optimization of the control tolerances for the diagnostic features is the binary working training matrix that is formed according to the rule (6) for quasi-optimal control tolerances $\{\widehat{A}_{HK,i} \mid i = 1, N\}$, $\{\widehat{A}_{BK,i} \mid i = 1, N\}$, which were obtained at the parallel optimization stage and the normalized tolerances $\{\delta_{H,i}\}$ for the diagnostic features. There is a difference between the implementation of the algorithm (4) and the parallel optimization. It is in sequential determination of the optimal control tolerances for each diagnostic feature. But since the optimal control for the *i*-th feature is determined by non-optimal control tolerances for other diagnostic features, then according to the algorithm (4) several sequences of sequential optimization are carried out until the maximum averaged value of the information criterion remains constant. Thus, the process of learning of the functional diagnostics system is in implementation of the procedure for finding the global maximum of the information criterion function in the working area of its definition and iterative approximation of this maximum to its maximum limit value to construct highly reliable rules.

4 EXPERIMENTS

The implementation of the discussed above algorithm has been carried out on the basis of machine learning of the HMM electric drive functional diagnostics system. The training matrix was formed with the help of "ULYS Systems", which is involved in the design, production and maintenance of products for mine equipment, based on archival data of DTEK Pavlogradvugillia, mine "Heroes of Space", town of Pavlograd. Each line of the training matrix consisted of 51 structured quantitative and categorical diagnostic features that characterized the electrical and temperature characteristics of the nodes on the technological cycles of the HMM. For clarity, diagnostics was carried out for three classes of recognition: class X_1^o described the functional state of the electric drive "Norma" and this class was accepted as the basic one, in relation to which the system of control tolerances was determined. Class X_3^o described high temperatures of the electric motor bearings, and class X_3^o described high temperatures of the engine winding.

Initially, algorithm (3) of the informationa-extreme machine learning for the system of functional diagnostics of the HMM electric drive with parallel optimization of diagnostic features was implemented, in which control tolerances for all features changed simultaneously at each step of the training. To improve the functional efficiency of machine learning, the algorithm for the sequential optimization of control tolerances for diagnostic features was implemented. In this, the results obtained by parallel optimization control tolerances for diagnostic features were considered as the starting point for their consistent optimization. In this, the results obtained by parallel optimization control tolerances for diagnostic features were considered as the starting point for their consistent optimization. Decisive rules were formulated according to the optimal geometric parameters of the hyperspherical containers of recognition classes obtained in the process of informationally-extreme machine learning. Their predicative expression is given by

$$(\forall X_m^o \in \Re^{|M|})(x^{(j)} \in \Re^{|M|})\{if\ [(\mu_m > 0) \,\&\, (\mu_m = \max_{\{m\}}\{\mu_m \mid m = \overline{1, M}\}]\ then\ x^{(j)} \in X_m^o\}. \quad (8)$$

In expression (8), for example, the membership function of the implementation vector to the hyperspherical container of the recognition class, is determined from the formula

$$\mu_m = 1 - \frac{d(x^{(j)} \oplus x_m)}{d_m^*}. \quad (9)$$

Thus, the recognized implementation vector refers to the recognition class from the given alphabet, for which the membership function (9) is maximum.

5 RESULTS

Fig. 1 shows the graph of the information criterion averaged in alphabetical order of the recognition classes (7) versus control tolerances for diagnostic features, which was obtained in the process of diagnostic optimization of the HMM electric drive units according to the algorithm of parallel optimization.

In fig. 1 the shaded area denotes the working (admissible) domain of the function (7), in which the first and the second authenticity exceeds the errors of the first and the second kind respectively.

Figure 2 shows the dynamics of changes in the maximum values of the information criterion averaged in alphabetical order of the recognition classes (7) when carrying out the procedure of the sequential optimization of the control tolerances for the diagnostic features.

Fig. 3 shows the graphs of the information criterion (7) versus the containers radii of recognition classes, constructed based on the results of parallel-sequential optimization of the optimal control tolerances for diagnostic features.

On the results of the functional diagnostic in test mode, that is the actual testing, probability of the right recognition of the implementation vectors for class X_1^o equals $P_t = 0.94$, for class X_2^0 equals $P_t = 0.92$ and for class X_3^0 equals $P_t = 0.89$.

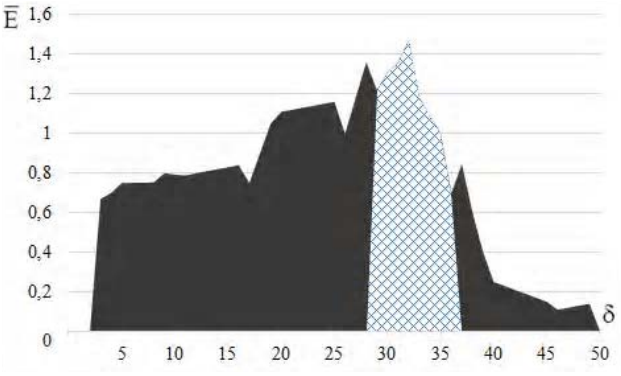


Figure 1 – The graph of information optimization criterion versus parameter of the control tolerances

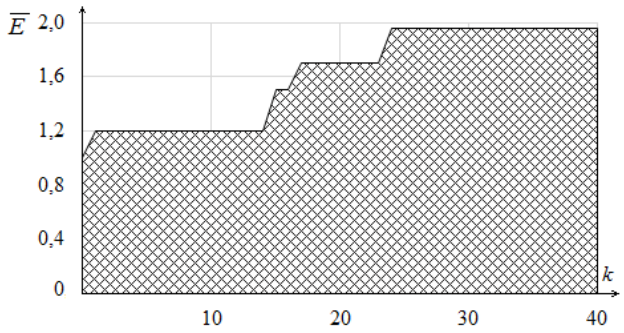


Figure 2 – The graph of maximum information criterion changes from the number of iterations of the sequential optimization of the control tolerances for the diagnostic features

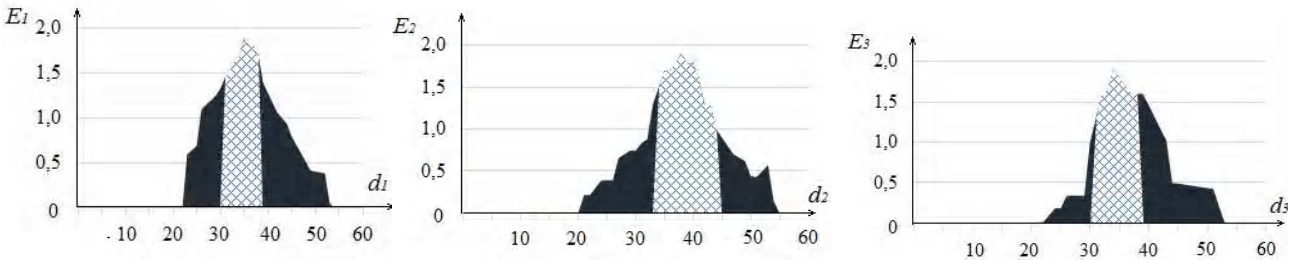
6 DISCUSSION

Analysis of Fig. 1 shows that optimal control tolerances parameter is $\delta^* = \pm 32$ of relative units at the maximum of the information criterion averaged in alphabetical order of the recognition classes (7) $\overline{E}_{\max} = 1.44$, which was obtained in the working domain of the function.

To improve the functional efficiency of machine learning, an algorithm for the sequential optimization of control tolerances for diagnostic features was implemented. In this, the results obtained by parallel optimization control tolerances for diagnostic features were considered as the starting point for their consistent optimization. Sequential optimization of control tolerances for diagnostic signs was carried out according to the procedure (4). Since the number of iterations in a single run of procedure (4) is determined by the number of diagnostic features, the analysis of Fig. 2 shows that already at the first run, the maximum value of the information optimization criterion $\overline{E}^* = 1.96$, which is substantially higher than its value obtained by the parallel optimization of the control tolerances for the diagnostic features (Fig. 1), was reached.

In order to construct highly reliable decisive rules of type (8) in the machine learning process with optimal control tolerances for diagnostic features, it is necessary to determine the optimal geometric parameters of the containers of recognition classes. Analysis of Fig,3 shows that optimal radius of the X_1^o class container equals $d_1^* = 36$ (here and after in code units), for class X_2^0 – $d_2^* = 37$ and for class X_3^0 – $d_3^* = 33$.

It should be noted that the use of constructed geometric approaches to geometric decisive rules significantly reduces the impact of the diagnostic features dictionary multidimensionality and the alphabet of recognition classes on the functional efficiency of machine learning. In this case, there is no need to retrain the system or change its structure as in the application of artificial neural networks. In addition, the construction of decisive rules in the framework of the geometric approach allows to take diagnostic solutions in real time, which is an important condition for functional diagnosis.



a b c

Figure 3 – Graphs of the information criterion (7) versus the containers radii of recognition classes: a – class X_1^0; b – class X_2^0; c – class X_3^0

CONCLUSION

The important scientific and technical task of increasing the functional efficiency of machine learning of the functional diagnostics system of the HMM electric drive is solved.

The scientific novelty of the results obtained is in the development of a new informationally-extreme method of machine learning of the system of the HMM electric drive functional diagnostics with parallel-sequential optimization of control tolerances for diagnostic features of recognition, which allows to formulate high-confidence decisive rules. In this, unlike artificial neural networks, the use of geometric decisive rules formulated within the framework of a geometric approach significantly reduces the influence of the multidimensional diagnostic features dictionary and the alphabet of recognition classes on the functional efficiency of machine learning. In addition, parallel-consecutive optimization of control tolerances for diagnostic recognition features can significantly improve the efficiency of machine learning, since consistent optimization of control tolerances occurs in the working domain of the function information criterion for optimization of machine learning parameters.

Application of the results obtained allows making decisions on the current technical state of the HMM electric drive, to provide timely repair and replacement of its unreliable nodes and to accumulate archival data for solving the problem of predictive diagnosis for planning preventive measures to restore the functional efficiency of the HMM automated control system.

The decisive rules constructed on the results of machine learning are not accurate according to the training matrix. Further studies need to be carried out in order to deepen machine learning by optimizing other spatio-temporal parameters that affect system functional efficiency of the functional diagnostics in further research according to the principle of delayed solutions for making decisive rules accurate on the training matrix.

REFERENCES

1. Сидельников Л. Г. Обзор методов контроля технического состояния асинхронных двигателей в процессе эксплуатации / Л. Г. Сидельников, Д. О. Афанасьев // Вестник ПНИПУ. Геология. Нефтегазовое и горное дело. – 2013. – № 7. – С. 127–137.
2. Bhowmik P. S. Faultdiagnostic and monitoring methods of inductionmotor: a review / S. Pradhan,. M. Prakash // International Journal of Applied Control, Electrical and Electronics. – 2013. – Vol. 1, № 1. – P. 1–18.
3. Szymacski Z. Intelligent, Energy Saving Power Supply and Control System of Hoisting Mine Machine with Compact and Hybrid Drive System / Szymacski Z., Ukiady Zasilania I Sterowania Gyrniczych Maszyn Wyci№gowych Z Napкdem Zintegrowanym Lub Hybrydowym // Archives of Mining Sciences. Retrieved 3 Mar. 2018. – 60 (1). – P. 239–251. DOI:10.1515/amsc-2015-0016
4. Сидоренко В. Н., Мониторинг мощных электроприводов на основе технологии data mining / В. Н. Сидоренко, А. П. Черный // Электротехнические и компьютерные системы. – 2011. – №3.– С. 414–416.
5. 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
6. Petr Dolezel Pattern recognition neural network as a tool for pest birds detection / Petr Dolezel, Pavel Skrabanek, Lumir Gago // *Computational Intelligence (SSCIIEEE Symposium Series on). – 2016*. – P. 1–6. DOI: 10.1109/HSI.2008.4581457
7. Ching-Lien Huang, Tsung-Shin Hsu, Chih-Ming Liu. The Mahalanobis-Taguchi system - Neural network algorithm for data-mining in dynamic environments. Expert Systems with Applications / Ching-Lien Huang, Tsung-Shin Hsu, Chih-Ming Liu // An International Journal. – 2009. – Vol. 36, Issue 3. – P. 5475–5480. DOI: 10.1016/j.eswa.2008.06.120
8. Hasan A. Yousef, Manal Wahba A. Adaptive fuzzy mimo control of induction motors. Expert Systems with Applications / Hasan A. Yousef, Manal Wahba A. // An International Journal. – 2009. – Vol. 36, Issue 3. – P. 4171-4175. DOI 10.1016/j.eswa.2008.04.004
9. Dovbysh A. S., Budnyk M. M., Moskalenko V. V. Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes / A. S. Dovbysh, M. M. Budnyk, V. V. Moskalenko // Journal of automation and information sciences. – New York : Begell House Inc., 2012, v. 44. i. 10. – P. 35–44. DOI: 10.1615/JAutomatInfScien.v44.i10.30
10. Dovbysh A. S. Information-Extreme Method for Classification of Observations with Categorical Attributes / A. S. Dovbysh, V. V. Moskalenko, A. S. Rizhova // Cibernetica and Systems Analysis. – 2016. –Vol. 52, No. 2. – P. 45–52. DOI: 10.1007/s10559-016-9818-1
11. Зимовець В. І. Система функціонального контролю приводу шахтної підйомної машини, що здатна навчатися / В. І. Зимовець, А. С. Чирва, О. І. Марищенко // Журнал інженерних наук. – 2016. – № 2. – С. 15–19.

Article was submitted 10.04.2018.
After revision 25.04.2018.

Довбиш А. С.[1], Великодний Д. В.[2], Проценко О. Б.[3], Зимовець В. І.[4]
[1]Д-р техн. наук, професор, завідувач кафедри комп'ютерних наук Сумського державного університету, Суми, Україна
[2]Канд. фіз.-мат. наук, завідувач секції кафедри комп'ютерних наук Сумського державного університету, Суми, Україна
[3]Канд. фіз.-мат. наук, доцент, доцент кафедри комп'ютерних наук Сумського державного університету, Суми, Україна
[4]Аспірант кафедри комп'ютерних наук Сумського державного університету, Суми, Україна

ОПТИМІЗАЦІЯ ПАРАМЕТРІВ МАШИННОГО НАВЧАННЯ СИСТЕМИ ФУНКЦІОНАЛЬНОГО ДІАГНОСТУВАННЯ ЕЛЕКТРОПРИВОДУ ШАХТНОЇ ПІДЙОМНОЇ МАШИНИ

Актуальність. Розв'язана актуальна задача підвищення функціональної ефективності машинного навчання системи функціонального діагностування електроприводу шахтної підйомної машини.

Мета роботи – розробка методу інформаційного синтезу здатної навчатися системи функціонального діагностування електроприводу шахтної підйомної машини, який дозволяє за побудованими в процесі машинного навчання вирішальними правилами підвищити достовірність та оперативність діагностичних рішень.

Метод. Запропоновано метод інформаційно-екстремального машинного навчання системи функціонального діагностування електроприводу шахтної підйомної машини, що ґрунтується на максимізації інформаційної спроможності системи в процесі її навчання. За отриманими в результаті машинного навчання оптимальними в інформаційному розумінні параметрами гіперсферичних контейнерів класів розпізнавання побудовано в рамках геометричного підходу вирішальні правила, практично інваріантні до багатовимірності простору діагностичних ознак. Крім того, підвищення оперативності машинного навчання системи досягається шляхом паралельно-послідовної оптимізації контрольних допусків на діагностичні ознаки. При цьому отримані в процесі паралельної оптимізації квазіоп-

тимальні контрольні допуски на діагностичні ознаки використовуються як стартові при їх послідовній оптимізації. Як критерій оптимізації параметрів машинного навчання використовується модифікована інформаційна міра Кульбака, яка є функціоналом від точнісних характеристик діагностичних рішень.

Результати. Розроблено алгоритмічне та програмне забезпечення машинного навчання системи функціонального діагностування електроприводу шахтної підйомної машини, яке дозволяє побудувати вирішальні правила для прийняття високо достовірних діагностичних рішень при функціонуванні системи в робочому режимі

Висновки. За результатами фізичного моделювання підтверджено працездатність запропонованого методу машинного навчання і розробленого програмного забезпечення системи функціонального діагностування електроприводу шахтної підйомної машини, що дозволяє їх рекомендувати для розв'язання практичних задач діагностування і автоматичного керування тяговими машинами.

Ключові слова: інформаційно-екстремальна інтелектуальна технологія, функціональне діагностування, машинне навчання, інформаційний критерій, електропривод, шахтна підйомна машина.

Довбыш А. С.[1], Великодный Д. В.[2], Проценко Е. Б.[3], Зимовец В. И.[4]

[1]Д-р техн. наук, профессор, заведующий кафедрой компьютерных наук Сумского государственного университета, Сумы, Украина

[2]Канд. физ.-мат.наук, заведующий секции кафедры компьютерных наук Сумского государственного университета, Сумы, Украина

[3]Канд. физ.-мат.наук, доцент, доцент кафедры компьютерных наук Сумского государственногоуниверситета, Сумы, Украина

[4]Аспирант кафедры компьютерных наук Сумского государственного университета, Сумы, Украина

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ МАШИННОГО ОБУЧЕНИЯ СИСТЕМЫ ФУНКЦИОНАЛЬНОГО ДИАГНОСТИРОВАНИЯ ЭЛЕКТРОПРИВОДА ШАХТНОЙ ПОДЪЕМНОЙ МАШИНЫ

Актуальность. Решена актуальная задача повышения функциональной эффективности машинного обучения системы функционального диагностирования электропривода шахтной подъемной машины.

Цель работы – разработка метода информационного синтеза обучающейся системы функционального диагностирования электропривода шахтной подъемной машины, позволяющего по построенным в процессе машинного обучения решающим правилам повысить достоверность и оперативность диагностических решений.

Метод. Предложено метод информационно-экстремального машинного обучения системы функционального диагностирования электропривода шахтной подъемной машины, основанный на максимизации информационной способности системы в процессе ее обучения. По полученным в результате машинного обучения оптимальным в информационном смысле параметрам гиперсферических контейнеров классов распознавания построены в рамках геометрического подхода решающие правила, практически инвариантные к могомерности пространства диагностических признаков. Кроме того, повышение оперативности машинного обучения системы достигается путем параллельно-последовательной оптимизации контрольных допусков на диагностические признаки. При этом полученные в процессе параллельной оптимизации квазиоптимальные контрольные допуски на диагностические признаки используются в качестве стартовых при их последовательной оптимизации. В качестве критерия оптимизации параметров машинного обучения используется модифицированная информационная мера Кульбака, являющаяся функционалом от точностных характеристик диагностических решений.

Результаты. Разработано алгоритмическое и программное обеспечение машинного обучения системы функционального диагностирования электропривода шахтной подъемной машины, позволяющее построить решающие правила для принятия высокодостоверных диагностических решений при функционировании системы в рабочем режиме.

Выводы. Результаты физического моделирования подтверждают работоспособность предложенного метода машинного обучения и разработанного программного обеспечения системы функционального диагностирования электропривода шахтной подъемной машины, что позволяет их рекомендовать для решения практических задач диагностирования и автоматического управления тяговыми машинами.

Ключевые слова: информационно-экстремальная интеллектуальная технология, функциональное диагностирование, машинное обучение, информационный критерий, электропривод, шахтная подъемная машина.

REFERENCES

1. Sidel'nikov L.G. Obzor metodov kontrolya tekhnicheskogo sostoyaniya asinkhronnykh dvigateley v protsesse ekspluatatsii, *Vestnik PNIPU*, Geologiya. Neftegazovoe i gornoe delo, 2013, No. 7, pp. 127–137.
2. Bhowmik P. S., Pradhan S., Prakash M. Faultdiagnostic and monitoring methods of inductionmotor: a review, *International Journal of Applied Control, Electrical and Electronics*, 2013, Vol. 1, pp. 1–18.
3. Szymacski, Z., Ukiady Zasilania I Sterowania Gyrniczych Maszyn Wyci№gowych Z Napкdem Zintegrowanym Lub Hybrydowym Intelligent, Energy Saving Power Supply and Control System of Hoisting Mine Machine with Compact and Hybrid Drive System. *Archives of Mining Sciences*, 60(1), pp. 239–251. Retrieved 3 Mar. 2018. DOI:10.1515/amsc-2015-0016
4. Sidorenko V. N., Cherniy A. P. Monitoring moshchnykh elektroprivodov na osnove tekhnologii data, *Elektrotekhnicheskie i komp'yuternye sistemy,* 2011, No. 3, pp. 414–416.
5. 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
6. Petr Dolezel, Pavel Skrabanek, Lumir Gago Pattern recognition neural network as a tool for pest birds detection, *Computational Intelligence (SSCIIEEE Symposium Series on), 2016*, pp. 1–6. DOI: 10.1109/HSI.2008.4581457
7. Ching-Lien Huang, Tsung-Shin Hsu, Chih-Ming Liu. The Mahalanobis-Taguchi system – Neural network algorithm for data-mining in dynamic environments. Expert Systems with Applications, *An International Journal,* 2009, Vol. 36, Issue 3, pp. 5475–5480. DOI: 10.1016/j.eswa.2008.06.120
8. Hasan A. Yousef, Manal Wahba A. Adaptive fuzzy mimo control of induction motors. Expert Systems with Applications, *An International Journal,* 2009, Vol. 36, Issue 3, pp. 4171–4175. DOI: 10.1016/j.eswa.2008.04.004
9. Dovbysh A. S., Budnyk M. M., Moskalenko V. V. Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes, *Journal of automation and information sciences*. New York, Begell House Inc., 2012, v. 44. i. 10. pp. 35–44. DOI: 10.1615/JAutomatInfScien.v44.i10.30
10. Dovbysh A. S., Moskalenko V. V., Rizhova A. S. Information-Extreme Method for Classification of Observations with Categorical Attributes, *Cibernetica and Systems Analysis*, 2016, Vol. 52, No. 2, pp. 45–52. DOI: 10.1007/s10559-016-9818-1
11. Zimovec' V. I., Chirva A. S., Marishhenko O. I. Sistema funkcional'nogo kontrolju privodu shahtnoї pidjomnoї mashini, zdatna navchatisja, *Zhurnal inzhenernih nauk,* 2016, No. 2, pp. 15–19.

UDC 621.913

Kovalevska O. S.[1], Kovalevskyy S. V.[2]

[1]PhD, Associate Professor, Associate Professor of the Department of Technology of Mechanical Engineering, Donbas State Engineering Academy, Kramatorsk, Ukraine
[2]Dr. Sc., Professor, Professor of the Department of Technology of Mechanical Engineering, Donbas State Engineering Academy, Kramatorsk, Ukraine

APPLICATION OF ACOUSTIC ANALYSIS IN CONTROL SYSTEMS OF ROBOTIC MACHINE TOOLS

Relevance. The problem of controlling complex technological machines such as machines with mechanisms based on kinematics with parallel structure is given consideration in the article in order to improve accuracy of positioning of actuators, to ensure their dynamic adjustment and optimization of trajectories of displacement of operating elements of the equipment (cutting tools, assembling or controlling instruments). The object of the study is the model of the operating area of a mobile robotic machine tool.

Objective. The goal of the work is to create a concept for controlling a mobile robotic machine tool applying acoustic control on the basis of a reference model based on deep neural networks.

Method. A method of identification and control of a mobile robotic machine tool using spectral description of absorption of acoustic wave with further processing of obtained information is offered. This method allows determining accuracy of positioning of actuators, as well as conducting dynamic adjustment and optimization of trajectories of displacement of operating elements of the equipment. A method of acoustic analysis for precision machining on machine tools with parallel kinematics has been developed.

Results. A neural network reference model has been constructed, which allows to diagnose current characteristics of the state of objects in different conditions, namely mechanism's configuration, mechanism's geometric parameters while running motor-spindle, dynamics of displacement of mechanism's nodes of the experimental stand with variable velocities and load on the drive, as well as temperature changes of the object. The developed neural network models also were tested for adequacy.

Conclusions. The experiments on the study of the dependency between the parameters of the spectrum of the acoustic signal with a given discreteness disturbed by excitatory effect in the form of "white noise" confirmed efficiency of this approach. Prospects for further research may consist in creation of methods for optimal control of complex technological machines to improve accuracy of positioning of actuators and to improve their dynamic settings.

Keywords: acoustic diagnostics, robotic machine tool, neural networks, reference model.

NOMENCLATURE

$\overline{X}$ – a set characterizing the coordinates $\{x_i, y_i, z_i\}$ of the object's actuator;

$\overline{Y}$ – a set characterizing the acoustic spectrum of natural oscillations of the object's system;

$A_i(f_i)$ – amplitude of the *k-th* frequency of the amplitude-frequency response;

Ω – object of modeling;

$\overline{X}'$ – a set of recovered coordinates of the object's actuator;

$\overline{Y}'$ – a set of restored characteristics of the acoustic spectrum of natural oscillations of the object's system;

x_i, y_i, z_i – spatial coordinates of the actuator of the object Ω;

x_k, y_l, z_r – controlled characteristics of the object;

δ – error of recovery $\overline{X}$;

δ' – error of recovery $\overline{Y}$;

W(f) – signal of excitation of the object by the "white noise";

R[W(f)] – reaction of the object to the excitatory action W(f);

x_i', y_i', z_i' – restored spatial coordinates of the object's actuator;

$A'(f_i)$ – restored amplitude-frequency characteristics;

$\overline{X}'_1, \overline{X}'_2 \ldots$ – a set of identifiers of intermediate layers of deep neural network by $\{x_i, y_i, z_i\}$;

$\overline{Y}'_1, \overline{Y}'_2 \ldots$ – a set of identifiers of intermediate layers of deep neural network by $[A_i(f_i)]$.

INTRODUCTION

Perfection of the accuracy and efficiency of mechanical processing technology depends on the quality of control of the actuators of the equipment. Along with increasing technical requirements for actuators, requirements for the control systems of the equipment and processes, in particular for informativeness of the diagnostic channels, are increasing too. In [3, 4, 7, 15, etc.] analysis of the methods and means of analysis of technological equipment and processes of mechanical processing is presented. However, despite a significant progress in creating new tools, methods and means for analysis of objects, topicality of research work in this field is still high [2].

The object of the research is the model of the operating area of a mobile robotic machine tool.

Among actively developing diagnostic methods of objects and processes, it is necessary to allocate the methods of acoustic analysis and creation of control systems on this basis [5, 6, 17].

The subject of the study is a set of methods for analysis of current characteristics of the state of objects in different conditions, namely mechanism's configuration, mechanism's geometric parameters while running motor-

© Kovalevska O. S., Kovalevskyy S. V., 2018
DOI 10.15588/1607-3274-2018-2-6

spindle, dynamics of displacement of mechanism's nodes of the experimental stand with variable velocities and load on the drive, as well as temperature changes of an object.

The aim of the work is to create a concept for controlling a robotic machine tool using acoustic control on the basis of a reference model based on deep neural networks.

1 PROBLEM STATEMENT

Recommendations of work [14] were applied for formalization of mathematical description, in which the spectrum of the signal is represented by spectral characteristics.

In general, these estimates are not dependent and there is a possibility of fluctuations around true value of the spectrum. Using frequency filters with frequency response $S(k)$, modified spectrum $\overline{Y}$ of the object Ω is obtained as a function of its properties $\overline{X}$:

$$\overline{Y} = A_k\left\{f_k\left[S(k)\right]\right\} \to F_y\left(\overline{X}\right).$$

To formulate the problem, the Bayes' theorem [8] was applied. Therefore, two of the main input positions were used: one known, the other one unknown (Fig. 1). The first one is the hypothesis about functional interconnection of frequency response with positioning of the operating element of a robotic machine tool; the other one is statistical data, which, due to transformation of Bayes' theorem, form *a posteriori* information that has an error within the confidence interval.

The input data is the amplitude-frequency response (obtained due to the measurements carried out by spectral analysis of the acoustic signal in the form of a response to the action of the exciting effect of the "white noise"). Amplitudes of the discrete stages of the obtained amplitude-frequency spectrum are determined by the considerations of the volume of calculations and assumption of a decent error.

Computational model allows determining the coordinates of the actuator of a robotic machine tool and can be used for solving direct and inverse tasks. The direct task is to determine the estimate δ of the error, so the sum of quadratic deviations of the restored spectral characteristic is minimized within a given error. The inverse task is solved by direct converters, using the neural network constructed for this purpose in reverse order. Then determine inputs for known coordinates $\overline{X}$ and restore initial values $\overline{Y}$ (Fig. 2).

Evaluation of the neural network should not exceed error of recovery $\mathrm{sum}(\overline{X}-\overline{X}')^2$ (4):

$$\overline{Y} = F_x\left[x_j, y_j, z_j\right], \quad (4)$$

$$\overline{X'} = F_y\left[A_i(f_i)\right],$$

$$i = 1..N,$$

$$j = 1..M,$$

$$M \le N.$$

When solving the task directly the model is built with unknown data, so the required accuracy is obtained through subsequent training.

In the inverse task, knowing the positioning of the recovered coordinates, training of neural network is performed with subsequent restoration of the frequency response spectrum with subsequent restoration of coordinates, which must differ from the inputs not more than the magnitude of error Δ (Fig. 3). Therefore, the optimization task is to minimize errors between the input and the reconstructed positioning coordinates.

As the output data, the concept of frequency response is employed (inputs are amplitudes of discrete degrees of frequency response). The number of such degrees is determined by the degree of discretion and is chosen for reasons of volume of calculations, capacity and required accuracy of the model.

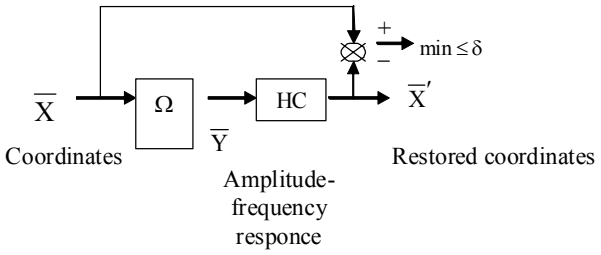


Figure 2 – Concept of neural network model construction

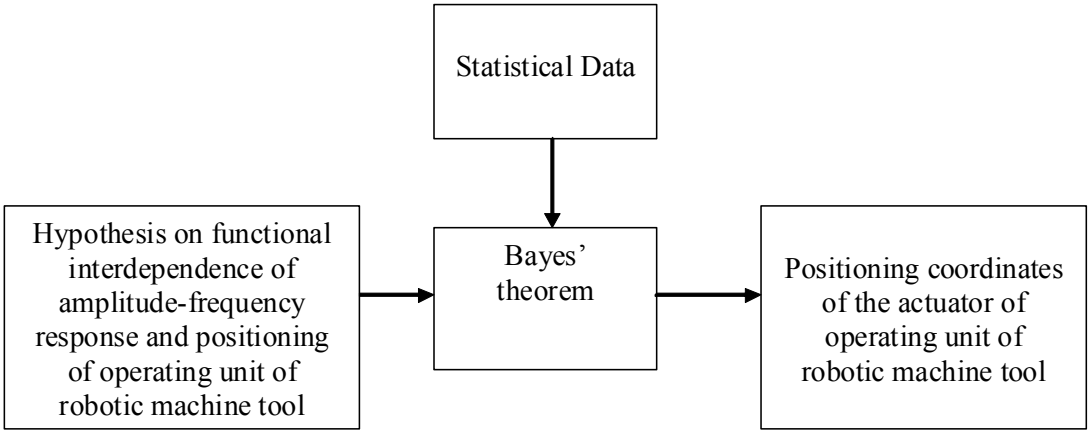


Figure 1 – Representation of the Bayes' approach, taking into account [8]

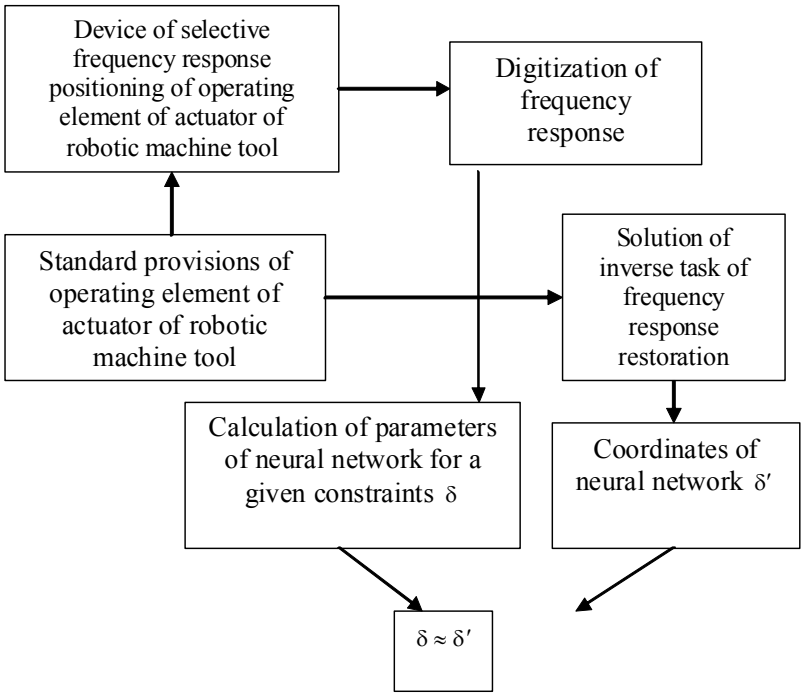


Figure 3 – Information model of the process

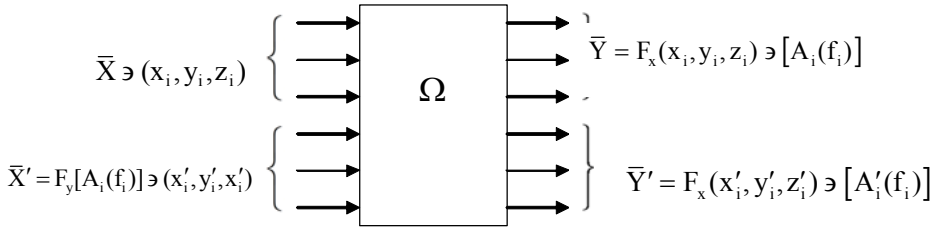


Figure 4 – Modeling of input and recovered coordinates

$$\bar{X} \ni (x_i, y_i, z_i),$$

$$\bar{Y} \ni [A_j(f_j)] = F_x(x_i, y_i, z_i),$$

$$\bar{X}' \ni (x_i', y_i', z_i') = F_y[A_i(f_i)],$$

$$\overline{Y'} \ni [A_i'(f_i)] = F_x[x_i', y_i', z_i'].$$

$$\begin{cases} \sum_i^N [(x_i - x_i')^2 + (y_i - y_i')^2 + (z_i - z_i')^2] \to \min, \\ \sum_j^M [(A_j(f_j) - A_j'(f_j)]^2 \to \min, \\ \left|\bar{X} - \bar{X}'\right| \le \delta, \\ \left|\bar{Y} - \bar{Y}'\right| \le \delta'. \end{cases}$$

For coordinates $\{x_i, y_i, z_i\}$ restrictions are the following:

$$\left|x_i - x_i'\right| \le \delta,$$

$$\left|y_i - y_i'\right| \le \delta,$$

$$\left|z_i - z_i'\right| \le \delta,$$

Under the terms of recovery of amplitude-frequency response within the limits of error δ':

$$\left|A(f_i) - A_i'(f)\right| \le \delta'.$$

Each amplitude $A_i(f_i)$ is endowed by distribution of probabilistic characteristics. For a given input X neural network restores value Y. For a given frequency response specimen a computational model has been created, which allows accurate $\left|\bar{X} - \overline{X'}\right| \le \delta$ recovery $\overline{Y'} = F_x[x_i', y_i', z_i']$ (Fig. 4). For this purpose an optimization task is solved, which contains a target function and has limitations.

2 LITERATURE REVIEW

In [6, 17] authors consider issues related to the analysis of technological equipment. The main task of analysis is to recognize the state of the object in conditions of limited information [1]. The algorithm of recognition is partly based on diagnostic models, which establish relationship between the states of the objects and their reflections in space of assessed indications [18].

Authors [9, 10] propose to use acoustic control for above mentioned algorithms, based on amplitude-frequency response of natural oscillations of objects.

Previously in [12, 18] an assumption was made and partially confirmed, that as an informative source of

diagnostic signal it is necessary to employ amplitude-frequency response of object's own oscillations within the acoustic range.

Among the methods of acoustic control the following are distinguished: the active ones, using radiation and reception of acoustic signals; and the passive ones, based only on reception of acoustic signals [11].

Acoustic control methods are based on interaction of elastic oscillations of controlled products with waves of wide frequency range [13, 16]. For non-destructive testing of multilayer structures these methods are most widely used. The main of them are the low-frequency methods, the ultrasonic method of exposure and, to a lesser extent, the method of reverberation and the acoustic and topographic method, although the possibilities of which are not fully disclosed. Obviously, the main problem lies in methodology and is applied to mathematical apparatus for processing of acoustic signals. Therefore, in this work, a method of acoustic analysis for mechanisms of robotic machine tools is proposed, as well as further processing of data by means of neural networks.

3 MATERIALS AND METHODS

Given the special importance of the issue, the authors of the article offer a comprehensive approach, in which the processes of control and analysis are a procedure for creating a reference model of controlling an object and keeping this model current during the technological operation. The basis of this approach is the acoustic signal spectrum, reflected from the elements of the technological system and the system of its transformation (Fig. 5).

The basis for creation of a reference model is the hypothesis on the informative capabilities of the acoustic signal spectrum as a source of data about the properties and parameters of an object. As shown in [10, 12], the spectrum of eigen-damaged vibrations of an object is the most informative in relation to various properties and parameters of an object. However, informative capabilities of the acoustic signal spectrum for analysis are greatly expanded by excitation of the spectra of forced oscillations by "white noise", emitted by transmitter over the test range.

As for adopted designations: W (f) – is the signal of the object's excitation by the "white noise"; R [W (f)] – is the reaction of the object to excitation by the "white noise"; x_k, x_l…, x_r… – are properties and parameters of the object (configuration, velocity, deformation, effort, strain, acceleration, temperature, etc.) R [W (f)] = F $\{x_k, x_l\, x_r \ldots\}$.

The task of analysis and creation of a reference model of an object is to determine the object's properties and parameters x_k, $x_i \ldots x_r \ldots$ by response R[W(f)] (Fig. 6).

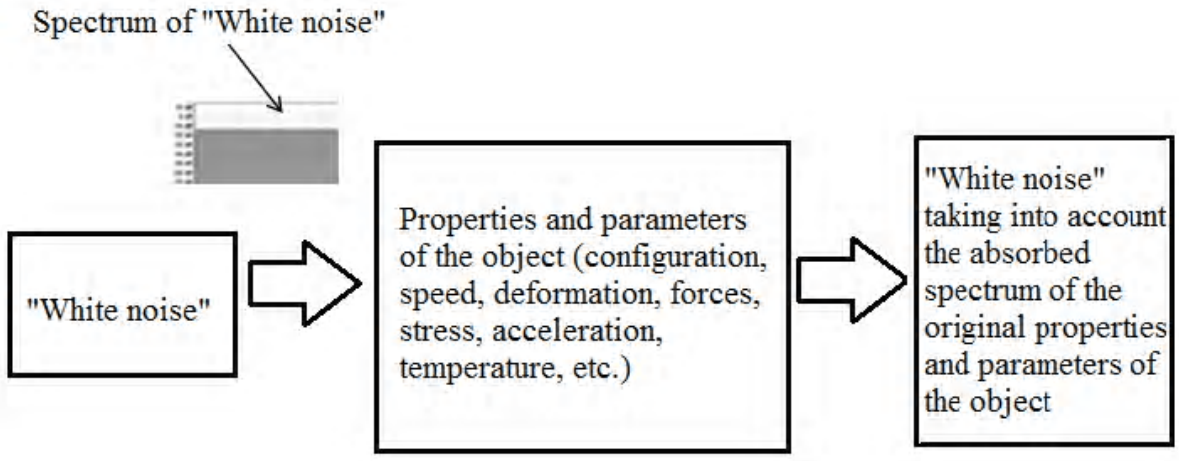


Figure 5 – Structural diagram of creation of a reference model of a mobile robotic machine tool

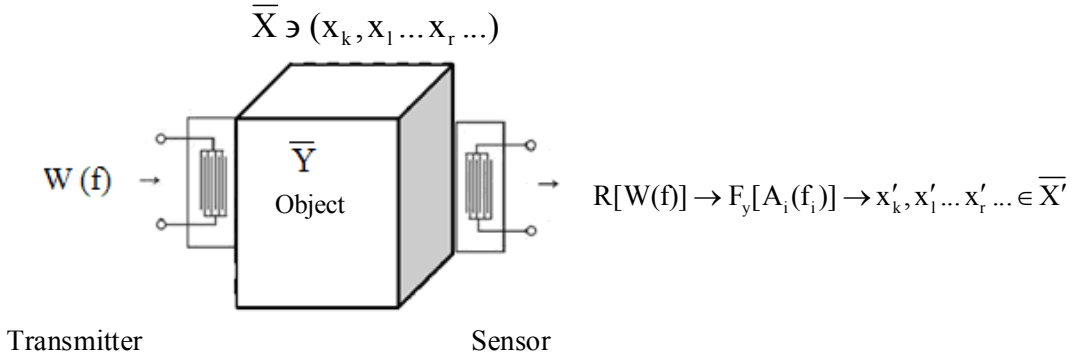


Figure 6 – Information model for determining properties and parameters of an object

Object

$$W(f) \rightarrow \text{Object} \rightarrow R\left[W(f)\right] \rightarrow F_y[A_i(f_i)] \rightarrow \overline{X'}.$$

Transmitter and receiver of the device for diagnostics are reversible piezoelectric elements, where signals are emitted (transmitter) and received (sensor). Since the analysis of an object is carried out relative to the reference signal of the "white noise", this approach permits to normalize initial diagnostic signals with respect to the reference signal.

Each dimension is maximally protected against accidental signal fluctuations. This is achieved by the fact that at each implementation of the elements of the experiment plan, the measurements are averaged over a cycle of 100 successive spectral scans with spectral discreteness of 178.3 from 0 to 20.000 Hz. Owing to the full factor experiment, a randomized combination of factors was employed to create a database, where the factors are the coordinates of positioning of the actuator of the experimental stand.

Acoustic spectrum of response, representing the sum of excited, absorbed and reflected acoustic waves, can be processed by deep neural networks. The result of such processing is a model that integrates the features of multilayer perceptrons and Kohonen Maps. Such an association is possible by virtue of neural networks of cascade configuration and modified neural element.

NeuroPro software package grants an opportunity to apply the resulting model for prognostication of properties and parameters of an object by amplitude and frequency characteristics and thereby ensure functioning of the reference model, which is applied for intelligent control of a robotic machine tool.

After processing of information on positioning in each of the characteristic positions of the executive element, a reference model is obtained, which allows adjusting the coordinates of displacements during processing or assembly by robotic machine tools.

For the research the algorithm of fast calculation of discrete Fourier Transform using the FFT analyzer with digital audio signal input is employed. The analyzer selects successive intervals ("windows") from the signal, by which the spectrum is calculated, which is displayed in the form of a graph of amplitude-frequency dependency (Fig. 7).

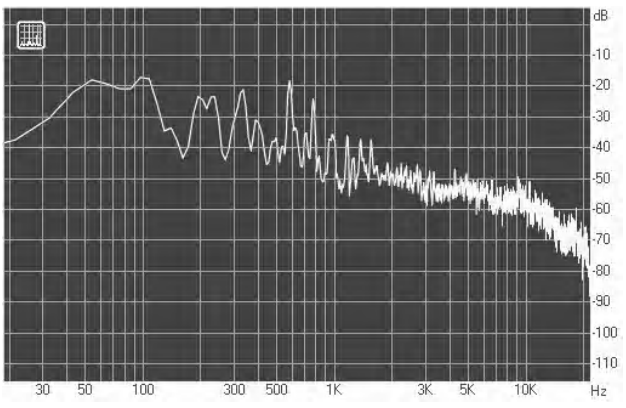


Figure 7 – Display of FFT Analyzer

Similarly to band analyzers, a logarithmic scale is usually employed for the axes of frequency and amplitude. But due to the linear positioning of FFT bands, frequency spectrum may not be sufficiently detailed at low frequencies or excessively oscillate at higher frequencies.

The horizontal line of FFT analyzer displays white noise, which is of equal energy in equal linear frequency intervals.

Parameter *N* is the number of analyzed signal responses, and it is crucial for the spectrum type. The larger *N*, the denser is the frequency grid, by which FFT decomposes the signal, and the more frequency details are visible on the spectrum.

Longer signal areas are analyzed for higher frequency resolution. If the signal within the FFT window changes its properties, the spectrum displays averaged signal information from entire window interval.

To analyze rapid changes in the signal, the length of the window N is assigned a small one. In time the resolution of analysis increases, but frequency decreases. Thus, the resolution of frequency analysis is inversely proportional to the resolution of time, which can be explained by relationship of uncertainty.

In order to create an informative spectrum of an acoustic signal, a hypothesis about its informational content was accepted in case of impact on an object by the influence of "white noise", which, along with detection of active informational frequencies of the spectrum, allows forming single initial conditions during the process of analysis.

Construction of a neural network's reference model for analysis of current characteristics of an object opens with the choice of the analyzed characteristics of an object, determined by the control tasks:

– increase of the accuracy of positioning of the actuator of an object;

– provision of dynamic adjustment of the actuator;

– optimization of the displacement trajectories of the operating actuator.

Solution of these tasks is aimed at improvement of the quality and productivity of technological machines and processes.

4 EXPERIMENTS

Equipment for experimental research:

– an experimental stand in the form of a delta mechanism, which is a group of mechanisms based on parallel kinematics, equipped with a system of numerical software control based on the MEGA 2650 circuit, which provides control programming of the stand using G-codes;

– two sound cards installed on two computers running Windows XP operating system;

– piezoelectric transmitter and sensor.

– an amplifier of the "white noise" signal excitation, 2 watts of power.

Software:

– computer package Spectrolab;

– computer package NeuroPro-0.25;

– computer package Statistica 10.

The plan of the experiments supposed to study the relationship between the parameters of the spectrum of the acoustic signal with a given discretion, disturbed by excited

influence in the form of "white noise". The authors put forward an assumption about a possibility of analysis of characteristics of investigated objects being invariant with respect to the errors made by the hardware part of the diagnostic system. This assumption is based on the fact that the diagnostic signal possesses initial values of the acoustic spectrum formed by the "white noise" by the same information channels and the same means as the main excited diagnostic signal.

Each of measurements *n* possesses maximum protection against accidental signal fluctuations. This is achieved by the fact that at each implementation of the elements from the experiment plan, the measurements were averaged over a cycle of 100 successive scans of spectrum with spectral discreteness of 178.3 from 0 to 20,000 Hz. A full-factor experiment with randomization of combinations of factors allowed formation of a database, where the factors were the positioning coordinates of the actuator of the experimental stand (x_i, y_i, z_i), i=1…n. An example of a fragment of such a database for experimental positions is presented in Table 1.

An example of acoustic spectrum of diagnostic signal for identifying the positioning of the actuator of the experimental stand is presented in Fig. 8.

5 RESULTS

On the experimental basis there were constructed neural network diagnostic models:

– configuration of the mechanism;

– geometric parameters of the mechanism while running motor-spindle;

– dynamics of displacement of nodes of the mechanism of the experimental stand with variable velocity and load on the drive;

– temperature changes of the object.

The created neural network models were tested for adequacy by Fisher's criterion. The results of assessment of these models are presented in Table 2.

Table 1 – Amplitudes of acoustic spectrum by its discrete

No.	Inputs: positioning coordinates, mm			Outputs: amplitude of acoustic spectrum, decibel, according to its discrete, Hz					
	X, mm	Y, mm	Z, mm	172,3	344,6	516,9	…	21882,1	22054,4
1	–52	–30	5	–29,9	–29,76	–46,87	…	–73,68	–81,32
2	52	–30	5	–32,3	–31,94	–46,98	…	–75,34	–83,49
3	0	–60	5	–34,49	–33,72	–45,56	…	–75,61	–83,99
4	0	0	5	–36,68	–36,11	–48,39	…	–76,12	–84,93
…	…	…	…	…	…	…	…	…	…
n–3	–52	–30	250	–54,24	–48,78	–49,53	…	–75,75	–83,77
n–2	52	–30	250	–53,57	–48,33	–48,57	…	–76,55	–84,44
n–1	0	–60	250	–55,13	–48,41	–48,71	…	–76,27	–84,16
n	0	0	250	–54,31	–49,35	–50,21	…	–75,86	–84,02

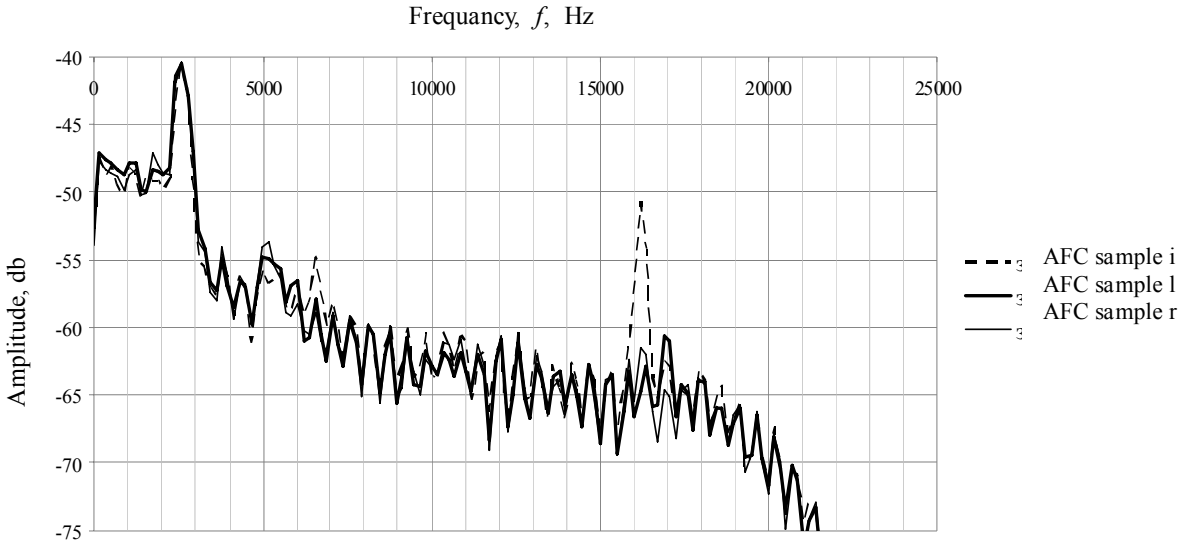


Figure 8 – Examples of spectra of acoustic signals for positioning coordinates of the actuator of an experimental stand

6 DISCUSSION

While training a neural network an input data sample (frequency response parameters) was formed, i.e. the input data relative to the positioning. Within the framework of these input data a neural network model was constructed, for which it was assumed in the very beginning that it would allow establishing a relationship between frequency response and coordinates of the position of the actuator of a robotic machine tool. Yet these data are obtained due to measurements, so there occurs an error (hence there is a statistical probability). These data were obtained before the experiment.

In general, this approach justifies feasibility of implementation of the concept of deep neural network (Fig. 9).

Adequacy of the models is confirmed by repeated excess of calculated value of Fisher's criterion over its table value for a given error probability for the input models – frequency ranges of the acoustic spectrum, constructed for the tuple time of the initial data.

The obtained results of the study confirm the provisions outlined in [12], where the authors state that control of the executive motions of technological machines with mechanisms of parallel structure is a complex task, a solution of which can be done on the basis of solving kinematic problems and dynamics of actuators. As a result, conditions can be created for precise positioning of operating elements of actuators at optimal speeds and accelerations of their displacements on a given trajectory. For these purposes, control systems should be equipped with informative systems of object identification. To confirm this, the results of experimental studies are proposed, which provide a basis for establishment of diagnostic systems of kinematics and dynamics of a mechanism, on the basis of which mobile robotic machine tools are created for application in reconfigurable production systems of mechanical processing.

CONCLUSIONS

During the research an actual task of increasing the accuracy of positioning of actuators of robotic machine tools has been solved.

Table 2 – Verification of mathematical models, constructed according to experimental data, for adequacy by Fisher's criterion under conditions $F_{table} << F_{est}$

Object of the experiment	Input coordinates x_i, y_i, z_i		δ	F_{table}	$F_{est.}$
Configuration of the mechanism	x_o= –75,–50,–25,0,25,50,75		0.01	2.67	$8.6x10^7$
	y_o= –75,–50,–25,0,25,50,75		0.01	2.66	$4.5x10^7$
	z_o= –75,–50,–25,0,25,50,75		0.01	2.66	$1.6x10^8$
	x_1= –75,–50,–25,0,25,50,75		0.01	2.60	$9.6 x10^3$
	y_1= –75,–50,–25,0,25,50,75		0.01	2.60	$3.1 x10^4$
	z_1= –75,–50,–25,0,25,50,75		0.01	2.60	$2.1 x10^4$
Configuration with runnung motor-spindle	x_o= –75,–50,–25,0,25,50,75		0.01	2.63	$1.4 x10^3$
	y_o= –75,–50,–25,0,25,50,75		0.01	2.60	$2.9 x10^8$
	z_o= –75,–50,–25,0,25,50,75		0.01	2.59	$1.8 x10^7$
Trajectory of displacement with variable velocity and load on the drive	$X_0 > X_1$	X_0	0.01	2.60	$5,9 x10^2$
		X_1	0.01	2.60	$6,5 x10^2$
	$Y_0 > Y_1$	Y_0	0.01	2.60	$6,7 x10^2$
		Y_1	0.01	2.60	$6,7 x10^2$
	$Z_0 > Z_1$	Z_0	0.1	2.60	$8,8 x10^2$
		Z_1	0.1	2.60	$8,8 x10^2$
	Displacement velocity		0.01	2.60	$4.7 x10^2$
	Actuator's weight		0.01	2.60	$3.5 x10^2$
Changes in sample temperature within limits 25…90 C^0	Sample 1		0.1 C^0	3.17	$3.7 x10^2$
	Sample 2		0.1 C^0	3.05	$6.9 x10^2$

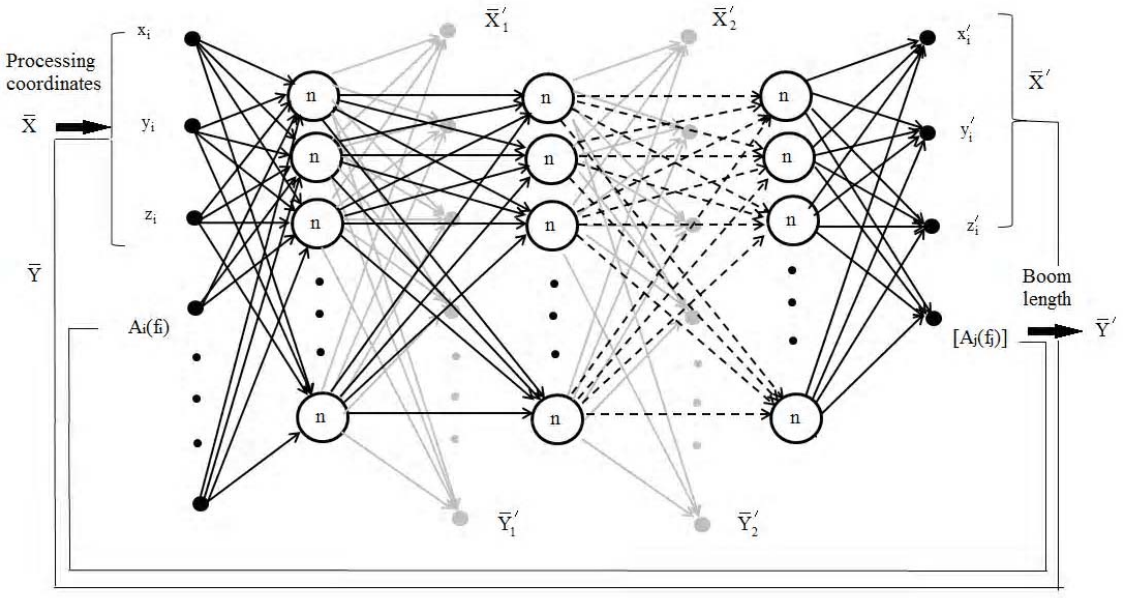


Figure 9 – Deep neural network scheme

The scientific novelty of the work consists in the method of identification and control of a mobile robotic machine tool on the basis of amplitude-frequency response, reflecting absorption of acoustic wave with further processing of information by deep neural networks of cascade architecture.

A new approach is proposed to the control of complex technological machines, such as machines with mechanisms of parallel structure for increasing the accuracy of positioning of actuators, ensuring their dynamic adjustment and optimization of trajectories of displacement of operating elements of the equipment (cutting tools).

Practical value of the obtained results consists in the development of the method of acoustic analysis for mechanisms of robotic machine tools. This allows to expand the range of opportunities for the improvement of the accuracy and performance of their performance.

Prospects for further research may consist in creation of hardware for controlling complex technological machines, such as machines with mechanisms based on parallel kinematics to improve the accuracy of the positioning of actuators, ensuring their dynamic adjustment and optimizing the trajectories of displacements of operating elements of the equipment.

ACKNOWLEDGMENTS

The work has been carried out within the framework of the State Budget Scientific Research of the Donbas State Engineering Academy "Development of neural network models for accelerated technological devising of production of new competitive products" (State Registration Number 0116U003615).

REFERENCES

1. Kovalevskyy S. V. Acoustic Monitoring with Neural Network Analysis / S. V. Kovalevskyy, O. S. Kovalevska // American Journal of Neural Networks and Applications. – 2015. – Vol. 1, No. 2. – P. 39–42. DOI: 10.11648/j.ajnna.20150102.12
2. Tyurin Yu. N. Electric heating using a liquid electrode / Y. N. Tyurin, A. D. Pogrebnjak // Surface and Coatings Technology. – 2001. – P. 293–299.
3. Hohmann M. Production methods and applications for high-quaIity metaI powders and sprayformed products / M. Hohmann, G. Brooks, C. Spiegelhauer. – Produktionsmethoden und Anwendungen fur qualitativ hochwertige Metallpulver und spruhkompaktierte Halbzeuge. Stahl und Eisen, 2005. – 258 p.
4. Yasa E. Application of laser re-melting on Selective laser melting parts. Catholic University of Leuven, Dept. of Mech. Eng, Heverlee, Belgium. Advances in Production Engineering / E. Yasa, J. Kruth // Management. – 6. – 2011. – P. 259–270.
5. David L. A Brief History of Additive / L. David Bourella, J. Joseph Beaman, Jr. a, Ming C. Leub and W. David Rosenc. – Manufacturing and the 2009 Roadmap for Additive Manufacturing: Looking Back and Looking Ahead. RapidTech, 2009. – 328 p.
6. Dutta B. Additive Manufacturing by Direct Metal Deposition / B. Dutta et. al. // Advanced materials & processes. – May. – 2011. – P. 33–36.
7. Fngelo H. C. Powder Metalurge: Science, technology and application / H. C. Fngelo, R. Subramanian. – New Dehli, 2009. – 128 p.
8. Kruglov I. A. Neural networks modeling of multivariable vector functions in ill-posed approximation problems / I. A. Kruglov, O. A. Mishulina // Journal of Computer and System Sciences International, 2013. – Vol. 52. No. 4. – P. 503–518.
9. Ковалевский С. В. Акустический контроль размеров и показателей шероховатости поверхностей деталей с применением нейросетевых моделей / С. В. Ковалевский, Е. С. Ковалевская, А. И. Нагиева // Нейросітьові технології та їх застосування : збірник праць міжнародної наукової конференції. – Краматорськ : ДДМА, 2014. – С. 24–35.
10. Ковалевский С. В. Развитие методов акустической диагностики в машиностроении : монографія / С. В. Ковалевский, Е. С. Ковалевская, В. И. Тулупов. – Краматорск : ДГМА, 2014. – 91 с.
11. Математические основы обработки сигналов. Практикум : учебное пособие / О. С. Вадутов ; Томский политехнический университет. – 3-е изд., испр. и доп. – Томск : Изд-во Томского политехнического университета, 2014. – 102 с.
12. Акустическая диагностика статических и динамических свойств механизмов с параллельной кинематикой // Комплексне забезпечення якості технологічних процесів та систем (КЗЯТПС – 2017): матеріали тез доповідей VII міжнародної науково-практичної конференції (м. Чернігів, 24–27 квіт. 2017 р.) : у 2-х т. / Чернігівський національний технологічний університет [та ін.]; відп. за вип.: Єрошенко Андрій Михайлович [та ін.]. – Чернігів : ЧНТУ, 2017. – Т. 1. – С. 59–60.
13. Осипов Л. А. Обработка сигналов на цифровых процессорах. Линейно-аппроксимирующий метод / Л. А. Осипов. – М. : Горячая линия-Телеком, 2001. – 114 с.
14. Селиванов С. Г. Системотехника инновационной подготовки производства в машиностроении / С. Г. Селиванов, М. Б. Гузаиров. – М. : Машиностроение, 2012. – 568 с.
15. Смирнов В. А. Кинетостатическое моделирование энергоэффективного управления оборудованием с параллельной кинематикой / В. А Смирнов // Вестник ЮУрГУ. Серия «Машиностроение». – 2010. – Вып. 16, № 29. – С. 65–71.
16. Христофоров А. В. Методы анализа спектра сигнала : учебно-методическое пособие к специальному лабораторному практикуму для студентов старших курсов и магистрантов кафедр радиофизического направления / А. В. Христофоров. – Казань, 2004. – 21 с.
17. Жук А. Я. Техническая диагностика. Контроль и прогнозирование: монография / [А. Я. Жук, Г. П. Малышев, Н. К. Желябина, О. М. Клевцов]. – Запорожье : Издательство Запорожской государственной инженерной академии, 2008. – 500 с.
18. Діагностика технологічних систем і виробів машинобудування (з використанням нейромережевого підходу) : монографія / С. В. Ковалевський, О. С. Ковалевська, Є. О. Коржов, А. О. Кошевой ; за заг. ред. д.т.н., проф. С. В. Ковалевського. – Краматорськ : ДДМА, 2016. – 186 с.

Article was submitted 06.02.2018.
After revision 26.02.2018.

Ковалевська О. С.[1], Ковалевський С. В.[2]

[1]Канд. техн. наук, доцент, доцент кафедри технології машинобудування Донбаської державної машинобудівної академії, Краматорськ, Україна

[2]Д-р техн. наук, професор, заведующий кафедри технології машинобудування Донбаської державної машинобудівної академії, Краматорськ, Україна

ЗАСТОСУВАННЯ АКУСТИЧНОГО АНАЛІЗУ В СИСТЕМАХ УПРАВЛІННЯ ВЕРСТАТІВ-РОБОТІВ

Актуальність. Розглянуто задачу управління складними технологічними машинами, такими, як верстати з механізмами на основі кінематики з паралельною структурою для підвищення точності позиціонування виконавчих механізмів, забезпечення їх динамічної настройки і оптимізації траєкторій переміщень робочих органів обладнання (ріжучого інструменту, складального або контролюючого інструменту). Об'єктом дослідження є модель робочої зони мобільного верстата-робота.

Мета роботи – створення концепції управління мобільним верстатом-роботом із застосуванням акустичного контролю на основі еталонної моделі на глибоких нейронних мережах.

Метод. Запропоновано метод ідентифікації та управління мобільним верстатом - роботом з використанням спектрального опису поглинання акустичної хвилі з подальшою обробкою отриманій інформації. Метод дозволяє визначати точність позиціонування виконавчих механізмів, проводити їх динамічну настройку і оптимізацію траєкторій переміщень робочих органів обладнання. Розроблено методику акустичної діагностики забезпечення точності механічної обробки на верстатах з паралельною кінематикою.

Результати. Побудована нейромережева еталонна модель, що дозволяє діагностувати поточну характеристики стану об'єктів в різних умовах, а саме, конфігурацію механізму, геометричні параметри механізму при працюючому мотор-шпинделі, динаміку переміщення вузлів механізму експериментального стенду зі змінною швидкістю і навантаженням на привід, зміни температури об'єкта. Розроблені нейромережні моделі були перевірені на адекватність.

Висновки. Проведені експерименти по дослідженню залежності між параметрами спектра акустичного сигналу із заданою дискретністю, порушеної збудливим впливом у вигляді «білого шуму», підтвердили працездатність даного підходу. Перспективи подальших досліджень можуть полягати в створенні методів оптимального управління складними технологічними машинами для підвищення точності позиціонування виконавчих механізмів, удосконалення їх динамічної настройки.

Ключові слова: акустична діагностика, верстат-робот, нейронні мережі, еталонна модель.

Ковалевская Е. С.[1], Ковалевский С. В.[2]

[1]Канд. техн. наук, доцент, доцент кафедры технологии машиностроения Донбасской государственной машиностроительной академии, Краматорск, Украина

[2]Д-р техн. наук, профессор, профессор кафедры технологии машиностроения Донбасской государственной машиностроительной академии, Краматорск, Украина

ПРИМЕНЕНИЕ АКУСТИЧЕСКОГО АНАЛИЗА В СИСТЕМАХ УПРАВЛЕНИЯ СТАНКОВ-РОБОТОВ

Актуальность. Рассмотрена задача управления сложными технологическими машинами, такими, как станки с механизмами на основе кинематики с параллельной структурой для повышения точности позиционирования исполнительных механизмов, обеспечения их динамической настройки и оптимизации траекторий перемещений рабочих органов оборудования (режущего инструмента, сборочного или контролирующего инструмента). Объектом исследования является модель рабочей зоны мобильного станка-робота.

Цель работы – создание концепции управления мобильным станком-роботом с применением акустического контроля на основе эталонной модели на глубоких нейронных сетях.

Метод. Предложен метод идентификации и управления мобильным станком-роботом с использованием спектрального описания поглощения акустической волны с последующей обработкой полученной информации. Метод позволяет определять точность позиционирования исполнительных механизмов, проводить их динамическую настройку и оптимизацию траекторий перемещений рабочих органов оборудования. Разработана методика акустической диагностики обеспечения точности механической обработки на станках с параллельной кинематикой.

Результаты. Построена нейросетевая эталонная модель, позволяющая диагностировать текущие характеристики состояния объектов в различных условиях, а именно, конфигурацию механизма, геометрические параметры механизма при работающем мотор - шпинделе, динамику перемещения узлов механизма экспериментального стенда с переменной скоростью и нагрузкой на привод, изменения температуры объекта. Разработанные нейросетевые модели были проверены на адекватность.

Выводы. Проведенные эксперименты по исследованию зависимости между параметрами спектра акустического сигнала с заданной дискретностью, возбужденного воздействием в виде «белого шума», подтвердили работоспособность данного подхода. Перспективы дальнейших исследований могут заключаться в создании методов оптимального управления сложными технологическими машинами для повышения точности позиционирования исполнительных механизмов, совершенствование их динамической настройки.

Ключевые слова: акустическая диагностика, станок-робот, нейронные сети, эталонная модель.

REFERENCES

1. Kovalevskyy S. V., Kovalevska O. S. Acoustic Monitoring with Neural Network Analysis, *American Journal of Neural Networks and Applications,* 2015, Vol. 1, No. 2, pp. 39–42. DOI: 10.11648/j.ajnna.20150102.12
2. Tyurin Yu. N., Pogrebnjak A. D. Electric heating using a liquid electrode, *Surface and Coatings Technology*, 2001, pp. 293–299.
3. Hohmann M., Brooks G., Spiegelhauer C. Production methods and applications for high-quality metal powders and sprayformed products. Produktionsmethoden und Anwendungen fur qualitativ hochwertige Metallpulver und spruhkompaktierte Halbzeuge. Stahl und Eisen, 2005, 258 p.
4. Yasa E., Kruth J. Application of laser re-melting on Selective laser melting parts. Catholic University of Leuven, Dept. of Mech. Eng, Heverlee, Belgium. Advances in Production Engineering, *Management,* 6, 2011, pp. 259–270.
5. David Bourella L., Joseph Beaman J., Jr. a, Ming C. Leub and W. David Rosenc. A Brief History of Additive. Manufacturing and the 2009 Roadmap for Additive Manufacturing: Looking Back and Looking Ahead. RapidTech, 2009, 328 p.
6. Dutta B. et. al. Additive Manufacturing by Direct Metal Deposition, *Advanced materials & processes,* May, 2011, pp. 33–36.
7. Fngelo H. C. Subramanian R. Powder Metalurge: Science, technology and application. New Dehli, 2009, 128 p.
8. Kruglov I. A., Mishulina O. A. Neural networks modeling of multivariable vector functions in ill-posed approximation problems, *Journal of Computer and System Sciences International*, 2013, Vol. 52, No. 4, pp. 503–518.
9. Kovalevskiy S. V., Kovalevskaya E. S., Nagieva A. I. Akusticheskiy kontrol razmerov i pokazateley sherohovatosti poverhnostey detaley s primeneniem neyrosetevyih modeley, *NeyrositovI tehnologiyi ta yih zastosuvannya : zbirnik prats mizhnarodnoyi naukovoyi konferentsiyi.* Kramatorsk, DDMA, 2014, pp. 24–35.
10. Kovalevskiy S. V., Kovalevskaya E. S., Tulupov V. I. Razvitie metodov akusticheskoy diagnostiki v mashinostroenii: monografIya. Kramatorsk, DGMA, 2014, 91 p.
11. Vadutov O. S. Matematicheskie osnovyi obrabotki signalov. Praktikum: uchebnoe posobie; Tomskiy politehnicheskiy universitet. 3-e izd., ispr. i dop. Tomsk, Izd-vo Tomskogo politehnicheskogo universiteta, 2014, 102 p.
12. Eroshenko AndrIy Mihaylovich Akusticheskaya diagnostika staticheskih i dinamicheskih svoystv mehanizmov s parallelnoy kinematikoy, *Kompleksne zabezpechennya yakostI tehnologIchnih protsesIv ta sistem (KZYaTPS – 2017): materIali tez dopovIdey VII mizhnarodnoYi naukovo-praktichnoyi konferentsiyi (m. Chernigiv, 24–27 kvit. 2017 r.) : u 2-h t. ChernIgIvskiy natsIonalniy tehnologIchniy unIversitet [ta in.]; vidp. za vip.: [ta in.*]. Chernigiv, ChNTU, 2017, Vol. 1, pp. 59–60.
13. Osipov L. A. Obrabotka signalov na tsifrovyih protsessorah. Lineyno-approksimiruyuschiy metod. Moscow, Goryachaya liniya-Telekom, 2001, 114 p.
14. Selivanov S. G., Guzairov M. B. Sistemotehnika innovatsionnoy podgotovki proizvodstva v mashinostroenii. Moscow, Mashinostroenie, 2012, 568 p.
15. Smirnov V. A. Kinetostaticheskoe modelirovanie energoeffektivnogo upravleniya oborudovaniem s parallelnoy kinematikoy, *Vestnik YuUrGU. Seriya «Mashinostroenie»*, 2010, Vyp. 16, No. 29, pp. 65–71
16. Hristoforov A. V. Metodyi analiza spektra signala. Uchebno-metodicheskoe posobie k spetsialnomu laboratornomu praktikumu dlya studentov starshih kursov i magistrantov kafedr radiofizicheskogo napravleniya. Kazan 2004, 21 p.
17. Zhuk A. Ya., Malyishev G. P., Zhelyabina N. K., Klevtsov O. M. Tehnicheskaya diagnostika. Kontrol i prognozirovanie: monografiya. Zaporozhye, Izdatelstvo Zaporozhskoy gosudarstvennoy inzhenernoy akademii, 2008, 500 p.
18. Kovalevs'kyy S. V., Kovalevs'ka O. S., Korzhov Ye. O., Koshevoy A. O.; za zah. red. d.t.n., prof. S. V. Kovalevs'koho Diahnostyka tekhnolohichnykh system i vyrobiv mashynobuduvannya (z vykorystannyam neyromerezhevoho pidkhodu) : monohrafiya. Kramators'k, DDMA, 2016, 186 p.

UDC 004.272.26: 004.93

Oliinyk A.[1], Subbotin S.[2], Lovkin V.[3], Ilyashenko M.[4], Blagodariov O.[5]

[1]PhD., Associate Professor, Associate Professor of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine
[2]Dr.Sc, Professor, Head of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine
[3]PhD, Associate Professor, Associate Professor of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine
[4]PhD., Associate Professor, Associate Professor of Computer Systems and Networks Department,
[5]Postgraduate student of Department of Software Tools, Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine

PARALLEL METHOD OF BIG DATA REDUCTION BASED ON STOCHASTIC PROGRAMMING APPROACH

Context. The task of automation of big data reduction in diagnostics and pattern recognition problems is solved. The object of the research is the process of big data reduction. The subject of the research are the methods of big data reduction.

Objective. The research objective is to develop parallel method of big data reduction based on stochastic calculations.

Method. The parallel method of big data reduction is proposed. This method is based on the proposed criteria system, which allows to estimate concentration of control points around local extrema. Calculation of solution concentration estimates in the developed criteria system is based on the spatial location of control points in the current solution set. The proposed criteria system can be used in stochastic search methods to monitor situations of excessive solution concentration in the areas of local optima and, as a consequence, to increase the diversity of the solution set in the current population and to cover the search space by control points in a more uniform way during optimization process.

Results. The software which implements the proposed parallel method of big data reduction and allows to select informative features and to reduce the big data for synthesis of recognition models based on the given data samples has been developed.

Conclusions. The conducted experiments have confirmed operability of the proposed parallel method of big data reduction and allow to recommend it for processing of data sets for pattern recognition in practice. The prospects for further researches may include the modification of the known feature selection methods and the development of new ones based on the proposed system of criteria for control points concentration estimation.

Keywords: data sample, pattern recognition, feature selection, parallel computing, informativeness criterion, stochastic programming approach.

ABBREVIATIONS

CMES is a Canonical Method of Evolutionary Search;

GMDH is a Group Method of Data Handling;

MARF is a Method of alternately Adding and Removing of Features;

MMDCA is a Multiagent Method with Direct Connection between Agents;

MMICA is a Multiagent Method with Indirect Connection between Agents;

PCA is a Principal Component Analysis;

PMBDR is a Parallel Method of Big Data Reduction.

NOMENCLATURE

$d(\chi_k, \chi_u)$ is a distance between points χ_k and χ_u of the search space XS;

$d(iter)$ is an average distance between all solutions on the current iteration;

g_{mk} is a m-th coordinate of the k-th solution;

$\overline{g_{mClc}}$ is the m-th coordinate of the c-th cluster center;

$Inform_k$ is an information about the k-th solution;

$InformLI(\chi_k)$ is a flag, which represents presence of solution χ_k in the solution set $R(iter) = \{\chi_1, \chi_2, ..., \chi_{N\chi}\}$ on the last search iteration $iter$;

$InformM(\chi_k)$ is a list of methods, which were used for estimation of solution χ_k;

M is a number of features in the sample of observations S;

$N_{\chi j}$ is a number of control points, which were investigated at the j-th process;

$N(iter)$ is a number of unique sampling points $Xe \in XS$, estimated in the process of feature selection till the current iteration $iter$ inclusively;

$N(XS)$ is a number of discrete space points XS;

P is a set of features (attributes) of observations in the given sample;

p_{qm} is a value of the m-th feature (attribute) of the q-th observation;

Q is a number of observations in the given sample of observations S;

Q_{ic} is a number of incorrectly recognized observations;

$rand[0;1]$ is a randomly generated number from the interval $[0;1]$;

S is a sample of observations (training sample);

$V(p_m)$ is an informativeness of a feature p_m;

$V(\chi_k)$ is a value of objective function of the k-th solution;

t_q is a value of output parameter of the q-th observation;

T is a set of output parameter values;

χ_k is the k-th solution, which corresponds to the k-th investigated control point Xe_k in the search space: $\chi_k \to Xe_k$.

© Oliinyk A., Subbotin S., Lovkin V., Ilyashenko M., Blagodariov O., 2018
DOI 10.15588/1607-3274-2018-2-7

INTRODUCTION

The investigation of complex technical objects and processes is connected with the necessity of big data processing, particularly with the search of feature set which describes investigated objects and processes in the best way [1–6]. The elimination of non-informative or insignificant features for diagnostic and recognition model synthesis process will allow to reduce model synthesis time, amount of processed data and complexity of the model which was built, but also to improve approximation and generalization abilities of the model [7–14].

As is well known [15–18], feature selection process is a highly iterative and resource-demanding procedure, which makes difficult to execute it in practice for solving of the tasks, where data processing should be performed without significant time delays (in on-line mode). Therefore the development of highly productive data reduction methods based on parallel computing is an actual task.

The object of the research is the process of big data reduction. The subject of the research are the methods of big data reduction. The research objective is to develop PMBDR based on stochastic calculations.

1 PROBLEM STATEMENT

Suppose we have data sample $S=<P,T>$, which consists of Q observations. Every observation is characterized by values of input attributes p_{q1}, p_{q2}, …, p_{qM} and output parameter t_q, where p_{qm} is a value of the m-th input feature of the q-th observation ($q=1, 2, ..., Q$, $m=1, 2, ..., M$); M is a total number of input features in the sample of observations S. Then the problem of informative feature selection can be ideally [1, 7, 19–21] stated as searching for the feature combination P^* from the initial data sample $S=<P,T>$ with minimum value of the given criterion of feature set quality estimation: $V(P^*)=\min\limits_{Xe\in XS} V(Xe)$, where Xe is a member of the set XS; $V(Xe)$ is a criterion of estimation of significance of feature set Xe; XS is a set of all possible feature combinations, which are obtained from the initial feature set P.

2 REVIEW OF THE LITERATURE

At present different methods are used for data reduction by means of informative feature selection. The most frequently used methods are the following ones.

Method of complete enumeration [1, 2, 7] estimates each control point Xe from all possible (2^M-1) control points in the search space XS. Because of complete enumeration of all possible solutions $Xe\in XS$, this method allows to find solution P^*, which has optimal value of objective function $V(P^*)=\min\limits_{Xe\in XS} V(Xe)$. As computing complexity of this method $O\left(2^M\right)$ significantly depends on input feature number M of training sample $S=<P,T>$, this method can be used for selection of features from small data samples. It substantially troubles and makes it impossible to apply this method for possessing of big data samples.

Heuristic methods [2, 7] (method of sequentially feature adding, method of sequentially feature removing) use greedy search strategy, which sequentially add (remove) features to the current feature set. Such approach is more simple in comparison with complete enumeration and demands less computing and time costs. But combinations of features P^*, selected by such methods, are generally characterized by unacceptable values of optimality criterion $V(P^*)$, because heuristic methods investigate very limited areas of search space. As a result feature combinations have optimal (or acceptable) value of objective criterion $V(P^*)$. Computing complexity of such methods is proportional to square of feature number M of input sample $S=<P,T>$: $O\left(M^2\right)$. Therefore application of such approach, when features are selected from big data samples, is also difficult.

Methods of stochastic search are based on application of probabilistic procedures for processing of control points $Xe\in XS$ and generally work with some solution set $R(iter)=\{\chi_1,\chi_2,...,\chi_{N\chi}\}$ on every iteration. Every k-th solution $\chi_k\in R(iter)$ corresponds to the k-th control point Xe_k in the search space XS on the investigated iteration $iter$: $\chi_k\to Xe_k$. Such methods can use evolutionary, multiagent or other approaches of computational intelligence as mathematical basis. Methods of stochastic search during the given number of iterations $Iter$ process $Iter\cdot N_\chi$ control points (where N_χ is a number of solutions, which are processed on every iteration of stochastic search). Therefore computing complexity $O(Iter\cdot N_\chi)$ of this approach does not depend directly on number of features M in the input sample. It allows to apply it for big data reduction. But such methods are given to recirculation in the areas of local optima (during search process some set of control points χ_k is concentrated around local extrema areas). It reduces its application efficiency and raises search time. Therefore expansion of the investigated areas of search space XS is based on usage of big number N_χ on control points χ_k, which are investigated on every iteration. This approach is not effective too because of the low diversity of solutions in the set $R(iter)$. Besides usage of big number of control points N_χ on every iteration increases search time.

Approach, which performs ranking of features p_m according to the values of individual significance $V(p_m)$ regarding output parameter T, can be used for feature selection also. Such approach is computationally simple (its computing complexity is $O(M)$), but it does not take

into account interdependence of features. Therefore in practice when features are interdependent, this approach doesn't allow to select feature sets, which have optimal or acceptable values of criterion of estimation of group informativeness $V(P^*)$.

Thus shortcomings of the existing feature selection methods cause necessity of the development of new method, which is based on stochastic approach and highly productive computings and is free from the described shortcomings.

3 MATERIALS AND METHODS

As was mentioned, application of the known methods of feature selection in practice for big data processing is difficult due to high iterativity and big amount of computings [1–12]. Besides it search strategies, used for feature selection, are also not enough effective for investigation of different areas of search space. Thus greedy strategy, which is used in heuristic methods of feature selection [1, 2, 7], allows to investigate very small part of search space, because in such cases well-defined, determinate action sequence is used, and this sequence performs very limited analysis of feature space (during optimization of objective criterion $V(P^*)$ small number of sampling points is investigated). Method of complete enumeration also applies well-defined action sequence, which investigates all points of search space, and because of significant time costs its application is impossible when there is significant number of features M in initial set $S =< P,T >$.

In stochastic methods (evolutionary, multiagent, etc. [1, 7, 12]) strategies, based on probabilistic search and examination of randomly selected points *Xe* of search space *XS*, are used. It allows to investigate the greater part of search space in comparison with deterministic methods. But methods, which use stochastic strategy, are subjected to recirculation in the areas of local extrema (if local optima areas are found on some iteration, then solutions are subsequently concentrated around such areas). Regardless mechanisms of local extrema leaving (for example, usage of mutation operator in evolutionary search methods or procedure of agent restarting in agent-oriented methods of computational intelligence), concentration of some solutions (control points) around local extrema areas is present on the following search iterations too. It reduces search efficiency (the same areas of feature space are investigated), raises time of its execution on computing machine, and in some cases does not allow to find acceptable solution.

Therefore for elimination of the presented defects in the developed parallel stochastic method of feature selection it is proposed to use combination of different strategies of stochastic search (methods based on evolutionary and multiagent approaches [7, 12]), which should be implemented at different nodes of parallel system. Application of different strategies, based on probabilistic approach, will allow to significantly extend search space coverage in comparison with the existing methods [1, 2, 7]. Application of parallel computings will allow to reduce search time and, as consequence, raise practical threshold of applicability of feature selection methods for big data processing.

In the proposed parallel method of big data reduction during initialization phase at the main core Pr_0 data reduction process is started, input data is read from user (data sample $S =< P,T >$, parameters of method, etc.).

Then feature selection methods are allotted between cores $Pr_1, Pr_2, ..., Pr_{NPr-1}$ of computing system, and also access to the input sample $S =< P,T >$ is passed. At that it is proposed to apply one core Pr_{NPr-2}, Pr_{NPr-1} for low iterative methods (based on decision trees and associative rules) correspondingly. Between the rest of cores $Pr_1, Pr_2, ..., Pr_{NPr-3}$ more complex data reduction methods, which are based on evolutionary and multiagent approaches, are uniformly allotted. Then for example, in the case of system with 24 cores, feature selection methods are allotted between cores of computing system in the following way: Pr_0 – main process, Pr_1 – Pr_6 – feature selection based on evolutionary search with feature grouping [21], Pr_7 – Pr_{11} – feature selection based on evolutionary method with feature clusterization [22], Pr_{12} – Pr_{16} – feature selection based on multiagent search with direct connection between agents [23], Pr_{17} – Pr_{21} – feature selection based on multiagent search with indirect connection between agents [23], Pr_{22} – feature selection based on decision trees [21], Pr_{23} – feature selection based on associative rules [24].

After that at every node $Pr_1, Pr_2, ..., Pr_{NPr-1}$ feature reduction process for the sample $S =< P,T >$ is performed. To raise space search coverage uniformity during feature selection, different methods of stochastic search are used at different nodes of parallel system. For these purposes it is proposed to use the following methods:

– evolutionary search with feature grouping [21] is based on usage of prior information about feature significance during feature selection process. As prior information for evolutionary search, estimations $V(p_m)$ of individual informativeness of features p_m, which are calculated at the method initialization stage, are used in evolutionary operators of crossover and mutation;

– evolutionary method with feature clusterization [22] as the previous method uses estimations $V(p_m)$ of individual informativeness of features p_m for evolutionary optimization. In addition to estimations $V(p_m)$, information about location of features p_m in observation space is also used. It allows to group features during search process and to form control points Xe_k using features, which are located distantly in feature space, eliminating in such a way combinations with interdependent features from consideration;

– multiagent method with direct connection between agents [23] is based on application of agent technologies of computational intelligence without usage of heuristic search procedures, applies agent approach for data exchange, allowing to investigate search space areas with

perspective control points in more detail. This method can be efficiently applied for feature selection during classification model synthesis (when output parameter has discrete values);

– multiagent method with indirect connection between agents [23] applies evolutionary operators of crossover and mutation at agent simulation phase, allowing to investigate search space more efficiently in comparison with the known multiagent methods and to reduce search time. This method allows to select feature combination with the highest significance when features are interdependent, is not subjected to recirculation in local optima, does not use greedy search strategy and does not make additional demands for objective function shape;

– feature selection method based on decision tress [21] which estimates informativeness of feature set Xe_k using decision tress which are synthesized during search process. Method allows to estimate individual and group informativeness of features p_m of training sample $S=<P,T>$ using structure of synthesized tree, performs phases of addition of root features and tree truncation. Such method is not highly iterative and resource-demanding, so it can be applied for finding of combination of the most significant features, when time and computing resources are limited, or can use small number of nodes Pr_j, when parallel systems are applied;

– feature selection method based on associative rules [24] can be efficiently used for informative feature selection from data samples $S=<P,T>$, generated based on transactional data sets $D=\{T_1,T_2,...,T_{N_D}\}$, where every element (transaction) T_j, $j=1,2,...,N_D$ contains information about some interrelated events, objects or processes. At that transactions T_j of data set *D* represent list from some element set. In the feature selection method based on associative rules [24] estimation of feature informativeness $V(p_m)$ is performed using information about interest level of extracted association sets (associative rules).

During feature selection in the proposed PMBDR processes $Pr_1,Pr_2,...,Pr_{NPr-1}$ can exchange signals with the main process Pr_0. At that signal Sgn_{inj} about completion of feature selection on the *j*-th process Pr_j is received from processes $Pr_1,Pr_2,...,Pr_{NPr-1}$ by the main process Pr_0, when one of the given stopping criteria is satisfied. For such purposes the following criteria can be used: $Crit_1$ – successful finding of combination of features P^*, which satisfies the given minimal acceptable search conditions (for example: $V(P^*)\le V_{\min}$, where $V_{\min}$ is a minimal acceptable value of feature set optimality criterion, which was set by user at initialization phase); $Crit_2$ – maximum acceptable number of search iterations; $Crit_3$ – maximum acceptable number of objective function value computing. The other criteria can be also used as stopping criteria.

Signals Sgn_{outj} about the necessity of feature selection procedure completion on the specific process Pr_j are received by the processes $Pr_1,Pr_2,...,Pr_{NPr-1}$ from the main process Pr_0. Signals Sgn_{outj} can be forwarded by the main process in the following situations:

– if signal Sgn_{inj} about successful search completion, when criterion $Crit_1$ is satisfied, is received from any process $Pr_1,Pr_2,...,Pr_{NPr-1}$. In this case the further search at the other processes loses meaning, because acceptable solution is found at the process Pr_j;

– if signal Sgn_{inj} about search completion, when criterion $Crit_2$ or $Crit_3$ is satisfied, is received from the set of processes $Pr_1,Pr_2,...,Pr_{NPr-1}$ (for example, not less than from the half of processes $Pr_1,Pr_2,...,Pr_{NPr-1}$). In this case the further feature selection procedure at the remaining processes is not advisable, because of idle time of the bigger part of computational system nodes, and current information is sent from processes $Pr_1,Pr_2,...,Pr_{NPr-1}$ to the process Pr_0;

– if maximum acceptable search time $Crit_4$ is reached, at every process $Pr_1,Pr_2,...,Pr_{NPr-1}$ current search iteration is finished and information about set of investigated control points $Xe\in XS$ and corresponding values of objective function $V(Xe)$ is sent to the main process.

During search process information $Inf_k=<Xe_k,V(Xe_k)>$ about points $Xe\in XS$ of search space XS which were investigated at every core $Pr_1,Pr_2,...,Pr_{NPr-1}$ is saved. It allows to estimate spatial location of solutions and its movement during search process. Besides it, such approach allows not to perform iterative estimation (calculation of values of objective function $V(Xe)$) of solutions $Xe\in XS$, which were estimated on the previous iterations, reducing search time in such a way.

Processes $Pr_1,Pr_2,...,Pr_{NPr-1}$ during data reduction procedure realization can efficiently exchange information $Inf_k=<Xe_k,V(Xe_k)>$ between each other. It allows to organize parallel search (similarly to island model [6, 7]) at group of processors, which are used for implementation of the same feature selection method, and also not to investigate iteratively control points which have been already estimated.

When feature selection procedure is finished at the nodes $Pr_1,Pr_2,...,Pr_{NPr-1}$, phase of collection and distribution of current information about optimization process is performed. At this phase information $InformPr_j$ about sets of investigated control points is received by the main process Pr_0 from the processes

$Pr_1, Pr_2, ..., Pr_{NPr-1}$. Such information contains coordinates of control points in search space, corresponding values of objective function, and also secondary information about methods, for which these points were estimated):

$$InformPr_j = \{Inform_1, Inform_2, ..., Inform_{N\chi j}\},$$

$$Inform_k = <\chi_k, V(\chi_k), InformLI(\chi_k), InformM(\chi_k)>.$$

After information $InformPr_j$ is received from all processes $Pr_1, Pr_2, ..., Pr_{NPr-1}$, it is combined on the main process Pr_0: $Inform = \bigcup_{j=1}^{N_{Pr}} InformPr_j$. It is significant that during combination of sets $InformPr_j$ situations, when the same solution χ_k is presented in different sets, can happen. In this case list of all methods, where solution χ_k took part, is saved to variable $InformM(\chi_k)$. Value of objective function $V(\chi_k)$ is chosen as the best from estimations, which were obtained at different processes. Different values of objective function $V(\chi_k)$ for the same point χ_k of search space can appear, because of general usage of errors of models, which were built based on feature set, which corresponds to the point χ_k, as objective function. At that artificial neural networks or other models of computational intelligence, can be used as such models. Training of such models is performed using probabilistic procedures, explaining possible differences in estimations $V(\chi_k)$ for the same values of χ_k.

After information $InformPr_j$ about sets of investigated control points χ_k is received, its concentration is estimated for the current solution set $R(iter) = \{\chi_1, \chi_2, ..., \chi_{N\chi}\}$ around local extrema $\nu_{conc}(iter)$ at the main process Pr_0. Calculation of estimations of solution concentration around local extrema $\nu_{conc}(iter)$ is performed for the purpose of defining of uniformity of coverage of search space XS during feature selection process. If there are situations when the majority of solutions χ_k is grouped in small areas of local optima, it is proposed to add extra control points, located outside of local extrema, to new solution set $R(iter+1)$.

For estimation of solution concentration $\nu_{conc}(iter)$ the current solution set $R(iter) = \{\chi_1, \chi_2, ..., \chi_{N\chi}\}$ should be divided into groups (clusters) $Cl(iter) = \{Cl_1, Cl_2, ..., Cl_{NCl}\}$, depending on its spatial location. For this purpose well-known cluster analysis methods should be applied [7].

Then for estimation of solution concentration around local extremum, the following criteria are calculated:

1) average distance $dC(Cl_c)$ between solutions in the specific cluster (1):

$$dC(Cl_c) = \frac{2}{|Cl_c|(|Cl_c|-1)} \sum_{k=1}^{|Cl_c|} \sum_{u=k+1}^{|Cl_c|} d(\chi_k, \chi_u),\ \chi_k, \chi_u \in Cl_c, \quad (1)$$

where distance $d(\chi_k, \chi_u)$ between points χ_k and χ_u of the search space XS, which belong to the cluster Cl_c, is calculated using expression (2):

$$d(\chi_k, \chi_u) = \frac{1}{M} \sum_{m=1}^{M} |g_{mk} - g_{mu}|; \quad (2)$$

2) dispersion $DC(Cl_c)$ of the solution χ_k within the cluster Cl_c represents average distance from the center $\overline{\chi_c}$ to solutions χ_k, belonging to the cluster Cl_c (3):

$$DC(Cl_c) = \frac{1}{|Cl_c|} \sum_{\chi_k \in Cl_c} d(\chi_k, \overline{\chi_c}), \quad (3)$$

where distance $d(\chi_k, \overline{\chi_c})$ between solution χ_k and center of the c-th cluster $\overline{\chi_c} = \{\overline{g_{1Clc}}, \overline{g_{2Clc}}, ..., \overline{g_{MClc}}\}$ is calculated using expression (4):

$$d(\chi_k, \chi_c) = \sqrt{\sum_{m=1}^{M} (g_{mk} - \overline{g_{mClc}})^2}, \quad (4)$$

where the m-th coordinate $\overline{g_{mClc}}$ of the c-th cluster center is calculated using formula (5):

$$\overline{g_{mClc}} = \frac{1}{|Cl_c|} \sum_{\chi_k \in Cl_c} g_{mk}. \quad (5)$$

The lower values of the criteria $dC(Cl_c)$ and $DC(Cl_c)$ corresponds to the higher grouped solutions, located in the c-th cluster Cl_c;

3) average cluster distance $dC(iter)$ between solutions on the current search iteration $iter$ (6):

$$dC(iter) = \frac{\sum_{c=1}^{N_{Cl}} |Cl_c| dC(Cl_c)}{\sum_{c=1}^{N_{Cl}} |Cl_c|} = \frac{1}{N_\chi} \sum_{c=1}^{N_{Cl}} |Cl_c| dC(Cl_c). \quad (6)$$

Criterion $dC(iter)$ characterizes average distance between different control points on the current iteration $iter$ within central cluster;

4) average cluster dispersion $DC(iter)$ of solutions on the current search iteration $iter$ (7):

$$DC(iter) = \frac{1}{N_\chi} \sum_{c=1}^{N_{Cl}} |Cl_c| DC(Cl_c). \quad (7)$$

The lower values of the criteria $dC(iter)$ and $DC(iter)$ corresponds to the solutions (control points), which are higher grouped around local optima on the current iteration $iter$;

5) coefficient of solution concentration on the current iteration (8):

$$\nu_{conc}(iter)=\frac{dC(iter)}{d(iter)}, \quad (8)$$

where average distance $d(iter)$ between all solutions on the current iteration is calculated using expression (9):

$$d(iter)=\frac{2}{N_\chi(N_\chi-1)}\sum_{k=1}^{N_\chi}\sum_{u=k+1}^{N_\chi}d(\chi_k,\chi_u),\ \chi_k,\chi_u\in R(iter). \quad (9)$$

Using estimates of solution dispersion $DC(iter)$, coefficient of solution concentration on the current iteration can be calculated using expression (10):

$$\nu_{conc}(iter)=\frac{DC(iter)}{D(iter)}, \quad (10)$$

where dispersion $D(iter)$ of solutions on the current iteration can be calculated using expression (11):

$$D(iter)=\frac{1}{N_\chi}\sum_{k=1}^{N_\chi}d\left(\chi_k,\overline{\chi}\right), \quad (11)$$

where distance $d\left(\chi_k,\overline{\chi}\right)$ between solution χ_k and central solution $\overline{\chi}$ on iteration $iter$ is calculated using expression (12):

$$d\left(\chi_k,\overline{\chi}\right)=\sqrt{\sum_{m=1}^{M}\left(g_{mk}-\overline{g_m}\right)^2}, \quad (12)$$

where m-th coordinate $\overline{g_m}$ of central solution $\overline{\chi}$ is calculated using expression (13):

$$\overline{g_m}=\frac{1}{N_\chi}\sum_{k=1}^{N_\chi}g_{mk}, \quad (13)$$

Value of criterion $\nu_{conc}(iter)$ belongs to the interval $(0;1)$. The closer the value of this criterion is to 1, the lower grouped solutions are (correspondingly, search space is covered by control points in more uniform way). The values of criterion $\nu_{conc}(iter)$, which are close to zero, evidence significant solution concentration around local extrema.

6) maximum number of control points $\chi_k\in Cl_c$, grouped within one local extremum (in the area of the cluster Cl_c):

$$N_{\max Ncl}(iter)=\max_{c=1,2,...,N_{Cl}}\left(\left|Cl_c\right|\right). \quad (14)$$

The bigger value criterion $N_{\max Ncl}(iter)$ has, the bigger number of solutions is grouped within one local extremum and correspondingly the lower location uniformity solutions have in the search space XS on iteration $iter$.

When decisions about excessive concentration of control points around some areas of local extrema are made, it is proposed to use integral criterion $\nu_{conc}(iter)$ and also criterion $N_{\max Ncl}(iter)$ of maximum number of control points $\chi_k\in Cl_c$, which are grouped within one local extremum, in the developed data reduction method.

If even one criterion has value which is over the given threshold ($\nu_{conc}(iter)>\nu_{concThr}$ or $N_{\max Ncl}(iter)>N_{\max NclThr}$), the decision about excessive concentration of control points within local extrema areas is made. It is proposed to add extra control points $R(iter+1)$, which are located outside of local extrema, to the current solution set to raise uniformity of search space coverage. The number of extra control points is proposed to set equal to the number of solutions in the set $R(iter)$.

For this purpose average values $\overline{g_m}$ of m-th coordinates of central solution $\overline{\chi}$ are calculated using expression (13). Values $\overline{g_m}$ demonstrate local concentration of solutions in projection of m-th feature axis. The closer the value $\overline{g_m}$ is to 1, the bigger the number of solutions χ_k characterizes m-th feature as informative. Similarly when $\overline{g_m}\to 0$, m-th feature is considered as non-informative in solution set $R(iter)=\{\chi_1,\chi_2,...,\chi_{N\chi}\}$.

Then using calculated values $\overline{g_m}$ and randomly generated numbers $rand[0;1]$, new solutions $\chi_a=\{g_{1a},g_{2a},...,g_{Ma}\}$ should be found. m-th coordinate g_{ma} of these solutions can be calculated using expression (15):

$$g_{ma}=\begin{cases}1, & rand[0;1]>\overline{g_m};\\ 0, & rand[0;1]\le\overline{g_m}.\end{cases} \quad (15)$$

Thus m-th coordinate g_{ma} of new control point χ_a will have bigger probability of possessing the value $g_{ma}=b$, if lower number of solutions χ_k in the set $R(iter)=\{\chi_1,\chi_2,...,\chi_{N\chi}\}$ has the same value of the m-th coordinate. Such approach will allow to generate new solutions χ_a, which are significantly distant from the current solution set $R(iter)=\{\chi_1,\chi_2,...,\chi_{N\chi}\}$, in such a way reducing solution concentration during data reduction process and raising uniformity of search space coverage.

It is proposed to use prior information on individual informativeness of features, in such a way allowing to secure values of the m-th coordinates, corresponding to the features and having significant effect on output parameter values. For this purpose at the initialization stage it is proposed to calculate values of individual informativeness $V(p_m)$ of features p_m, which characterize correlation between feature p_m and output parameter T.

Values of pair correlation coefficient, feature entropy, sign correlation criterion can be used as estimates of $V(p_m)$ [7, 21]. If prior information on individual significance of features is used, expression (15) can be modified in the following way:

$$g_{ma} = \begin{cases} 1, \ rand[-1;1] > (\overline{g_m} - V(p_m)); \\ 0, \ rand[-1;1] \le (\overline{g_m} - V(p_m)). \end{cases} \quad (16)$$

Such approach allows to raise probability of generation of new solutions χ_a with genes g_{ma}, corresponding to the highly informative features p_m. At that probabilistic approach maintains possibility of generation of solutions χ_a, which are remotely situated from the current set of control points $R(iter) = \{\chi_1, \chi_2, ..., \chi_{N\chi}\}$.

After the generation of the necessary number of additional control points χ_a, sets $R(iter) = \{\chi_1, \chi_2, ..., \chi_{N\chi}\}$ and $R_a(iter) = \{\chi_{a1}, \chi_{a2}, ..., \chi_{aN\chi}\}$ are united into the set $R(iter+1)$.

After that data reduction procedure is restarted at the nodes $Pr_1, Pr_2, ..., Pr_{NPr-1}$. At that new initial sets of solutions $R_j(iter+1)$ for the corresponding feature selection methods are formed based on the set $R(iter+1)$. The sets $R_j(iter+1)$ are formed at the nodes Pr_j in the following way. At the beginning solutions $\chi_{elit,j}$ with the highest values of objective function $V(\chi_{elit,j}) = \max\limits_{\chi_k \in R_j(iter)} (V(\chi_k))$ on the previous iteration are selected. Thus elite solutions $\chi_{elit,j}$ with the best values of objective function V are automatically transferred into the next population $R_j(iter+1)$, enabling usage of results which were got on the previous iterations and approaching of new initial search points to optimal ones. Number of elite solutions $\chi_{elit,j}$, which are automatically transferred into the next search iteration, is set by user at method initialization stage, and generally is equal to 2–5% of total number of solutions, which are used at the separate node of computation system. Then solutions χ_k are randomly chosen from the set $R(iter+1)$. The overall number of solutions is set according to the requirements of feature selection method, which is used at the *j*-th node Pr_j of computation system.

It is significant that besides the values of coordinates of control points χ_k, mathematical support, which is used at nodes Pr_j, has access to information about all solutions which were estimated earlier and its corresponding values of objective functions $<\chi_k, V(\chi_k)>$, allowing to avoid recurrent estimation of solutions which were estimated earlier and to reduce search time.

Then using initial sets of control points $R_j(iter+1)$ at the nodes $Pr_1, Pr_2, ..., Pr_{NPr-1}$, data reduction procedures are realized.

The described process should be continued till one of the following stopping criteria will be achieved: $Crit_1$ – successful finding of combination of features P^*, which satisfies the given minimal acceptable search conditions; $Crit_5$ – exceeding of total maximum permissible search time on the parallel system; $Crit_6$ – maximum permissible number of restart of data reduction procedure at the nodes $Pr_1, Pr_2, ..., Pr_{NPr-1}$.

Thus the proposed PMBDR proposes to use different strategies of stochastic search, based on evolutionary and multiagent approaches and realized at different nodes of parallel system. Usage of different strategies, based on probabilistic approach, allows to considerably extend coverage of search space. It is proposed to add control points, which are located outside of local optima, to the current solution set in the proposed method for raising of search space coverage uniformity during search process. Application of parallel computing in the proposed method makes it possible to reduce search time and, as consequence, to raise practical threshold of feature selection methods applicability for big data processing.

The criteria system, which enables to estimate concentration of control points around local extrema, was proposed. Calculation of solution concentration estimates in the developed criteria system is based on the spatial location of control points in the current solution set. The proposed criteria system can be used in stochastic search methods to monitor situations of excessive solution concentration in the areas of local optima and, as a consequence, to increase the diversity of the solution set in the current population and to cover the search space by control points in a more uniform way during optimization process.

4 EXPERIMENTS

For experimental investigation of the efficiency of the proposed method application for feature selection and pattern recognition problems solving, the vehicle recognition task [21], which is characterized by the data sample containing 10000 observations, was used. Every sample observation presents vehicle image and is formed by values of 26 features and 1 output parameter which defines, if observation belongs to the considered class.

At the beginning of the pattern recognition problem solution process concerning the considered task, feature selection methods were applied. These methods allowed to get informative feature set, which was considered as the most informative. It allowed to solve feature selection problem on the one hand, and on the other hand to select informative feature set, which then was used for model synthesis realization. Every such model was then used for vehicle recognition based on the classification with two classes: if observation belongs to the corresponding class (motorcyclist, passenger car, truck, bus, minivan or an object which is not recognized) or no. That is totally 5 such models were synthesized.

The following methods besides PMBDR, proposed in the paper, were considered as feature selection methods: PCA, GMDH, CMES, MARF, MMICA, MMDCA.

Let's consider criteria, which were used for investigation of the obtained results of feature selection problem solution.

Number of features k, which formed informative feature set, was considered as basic estimation criterion. Taking into account that the problem was solved for 5 variants of the problem statement separately, the number of selected features was presented as rounded average value, as well as interval of these values (minimum and maximum values, that is the lowest and the largest cardinal number of the set of informative features selected as a result of the corresponding method application).

Taking into account shortcomings which should be eliminated in the proposed method, it is necessary not only to estimate obtained results for the given conditions, but also to estimate the depth of coverage of search space XS. Therefore the corresponding criterion $\nu_{Deep}(iter, XS)$ should be calculated using expression (17):

$$\nu_{Deep}(iter, XS) = \frac{N(iter)}{N(XS)}. \quad (17)$$

As a set XS presents collection of all possible feature combinations p_m ($m = 1, 2, ..., M$), obtained from the initial feature set P ($|P| = M$), quantity $N(iter)$ can be calculated in the following way (18):

$$N(XS) = |XS| = 2^M - 1. \quad (18)$$

The number of unique sampling points, estimated on the current iteration, can be considered as alternative to the criterion presented above. It makes possible to demonstrate convergence of the method in absolute representation and in particular to compare it with methods, which are characterized by finding of local optima instead of global ones.

The results of the feature selection phase directly influence on the quality of pattern recognition solutions, therefore the following criteria were set as investigation criteria for the obtained pattern recognition problem solutions:

– recognition error E, which is defined in the following way:

$$E = \frac{Q_{ic}}{Q}. \quad (19)$$

– method operating time T_c, which is needed by method to achieve an acceptable solution.

The software based on the proposed method was written in C language using MPI and CUDA libraries: data exchange between the core and the rest of the cluster nodes was performed using multiple MPI exchange functions (Bcast, Gather, Scatter, Reduce).

For realization of parallel computing in experimental investigation, hardware of Software Tools Department of Zaporizhzhia National Technical University was used.

For investigation of PBMDR, evolutionary search with feature grouping, multiagent method with indirect and direct connection between agents and also method of feature selection based on associative rules were used at different nodes of parallel system as the most suitable for the considered task solving based on the preliminary comparison.

5 RESULTS

Table 1 presents results of informative feature set selection based on feature selection methods, expressed as interval and average value of cardinal number of such set (for alternatives of the forecasted recognition class).

Table 1 – Number of features which were selected by feature selection methods during vehicle recognition

№	Feature selection method	Values of comparison criteria		
		K_{min}	K_{max}	K
1	PCA	12	13	12,4
2	GMDH	11	12	11,2
3	CMES	10	11	10,4
4	MARF	11	13	12,2
5	MMICA	10	11	10,2
6	MMDCA	10	11	10,4
7	PMBDR	10	11	10,2

The dependence of number of unique sampling points on the current iteration number for CMES is presented in the Figure 1.

The analogous presentation of the number of unique sampling points, investigated on the current iteration, for MMDCA is showed in the Figure 2.

The change of unique sampling points number during execution of parallel method of big data reduction is presented in the Figure 3.

In the Figure 4 the diagram, which presents distribution of vehicle recognition error level during investigation of pattern recognition problem depending on the feature selection method which was applied on the corresponding stage, is presented.

The diagram, which presents ratio of vehicle recognition operation time on application of different feature selection methods, is presented in the Figure 5.

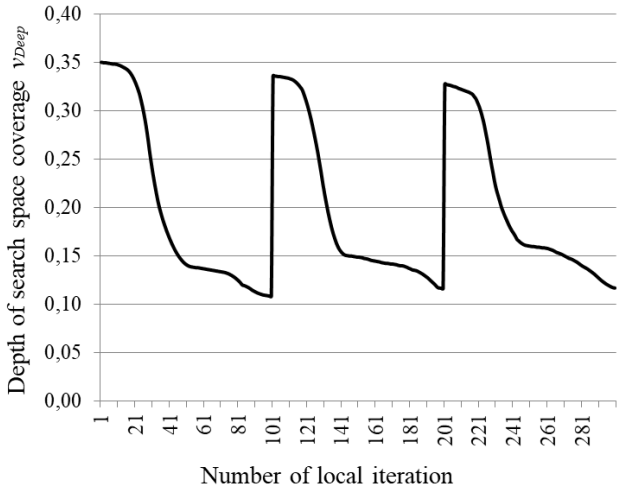


Figure 1 – Graph of dependence between number of unique sampling points and number of CMES iteration

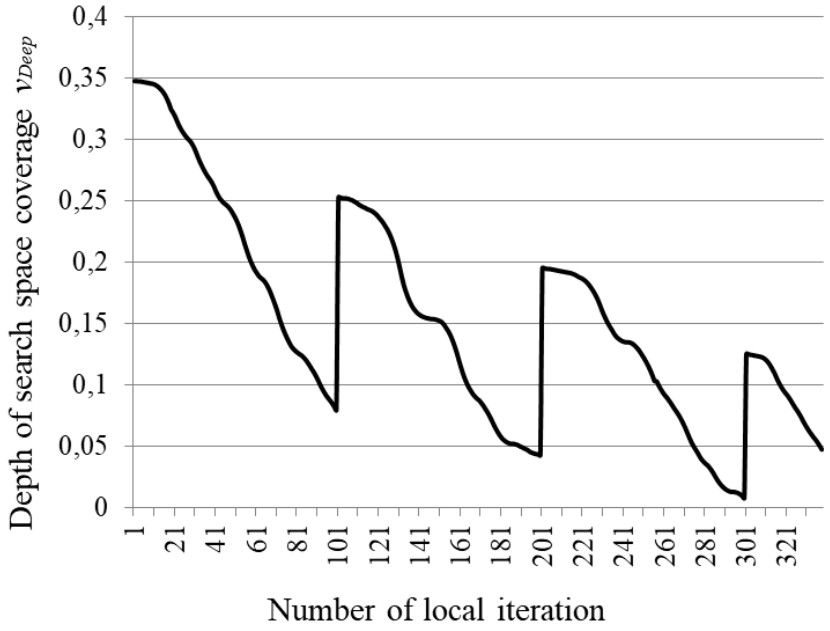


Figure 2 – Graph of dependence between number of unique sampling points and number of MMDCA iteration

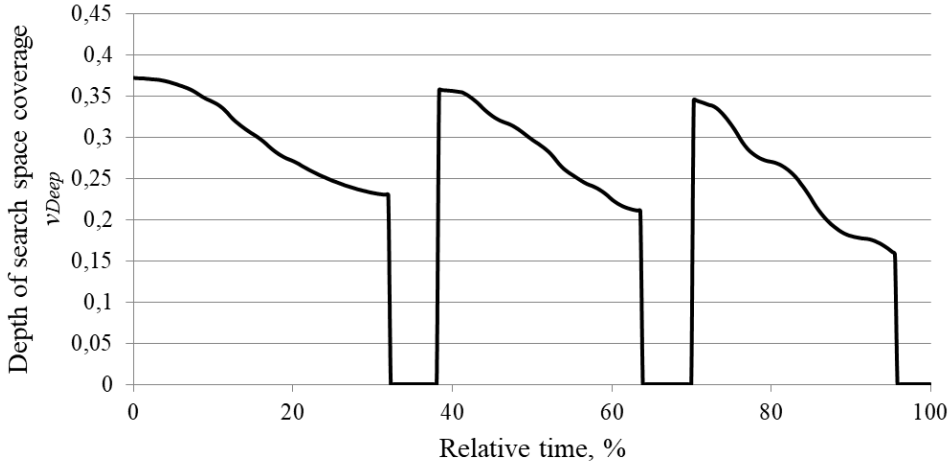


Figure 3 – Graph of dependence between number of unique sampling points and number of iteration of parallel method of big data reduction

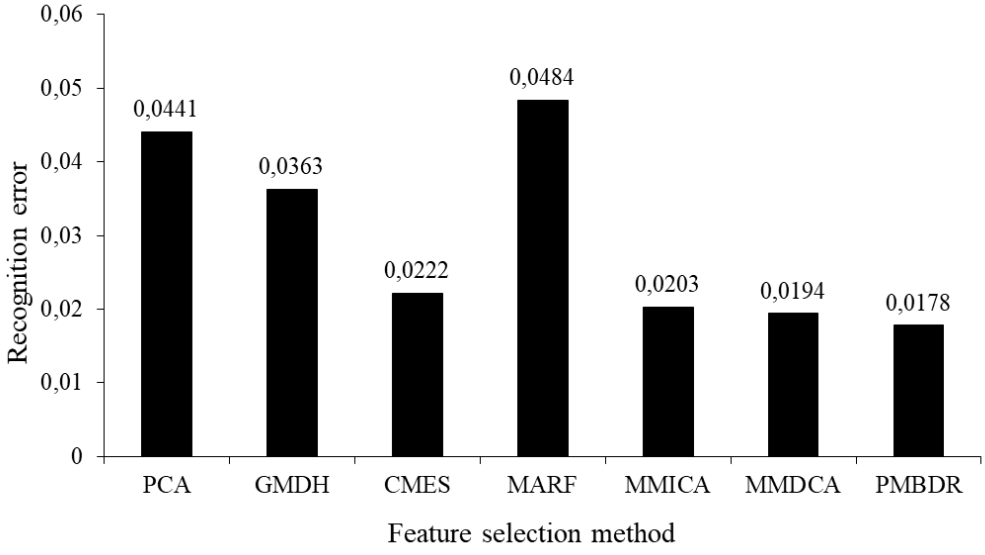


Figure 4 – Diagram of vehicle recognition error distribution

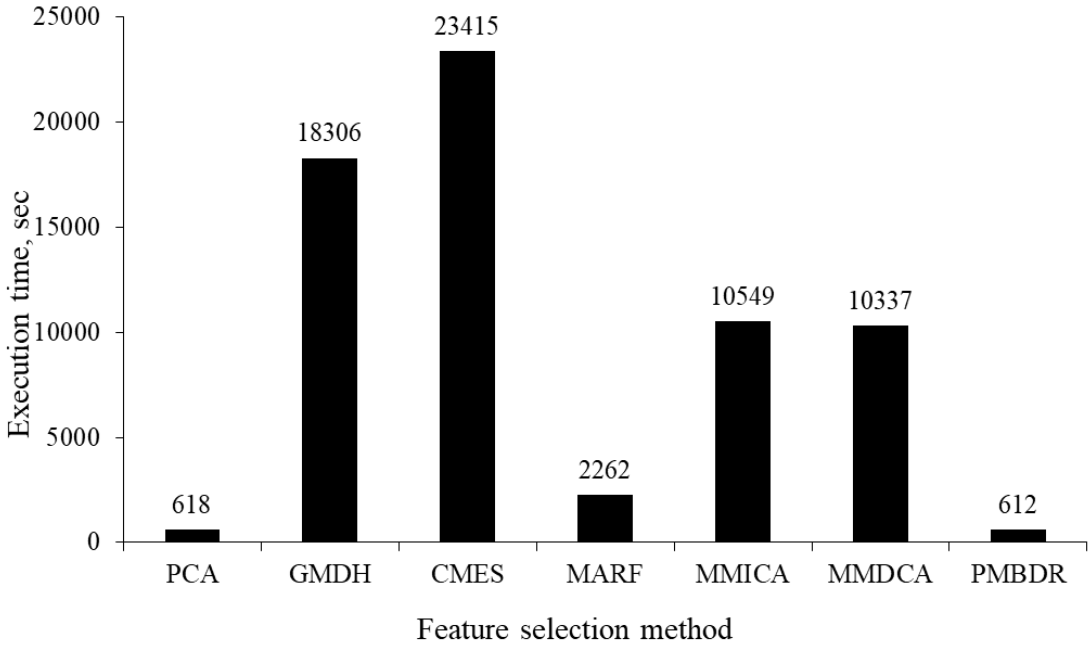


Figure 5 – Diagram of vehicle recognition operation time distribution

6 DISCUSSION

The results of experiments, shown in the Table 1, demonstrated that the set with the lowest number of informative features was selected by MMICA and PMBDR (10.2 on average). MMDCA and CMES were characterized by almost the same set (10.4).

Figures 1–3 show the dependence of the search space coverage depth on the current iteration number for 3 methods with the best recognition results (Table 1, Figures 4–5).

As can be seen from the Figure 1, during CMES execution process high initial values of search space coverage decrease rapidly (2.36 times during 30 iterations), leading to shortcoming when significant coverage of the search space is implemented only at the initial stage, so almost half of iterations is performed over a set of unique points that cover only 15% of the search space. At the same time the initial coverage of the search space does not decrease significantly after iteration set is repeated, that is the repeated implementation of iterations begins with almost the same set of unique points (33–35 %).

Fig. 2 demonstrates the same presentation of the search space for MMDCA. In this case, there is no quick reduce of the number of unique points as it was in the evolutionary search (search space coverage is decreased till 15% during 3/4 of iterations). However, this method leads to the following situation: when iteration set is repeated, the initial coverage of the search space is constantly reduced, and more than 30% of unique points is considered only during the first 27 iterations, leading to the fact that at the last stage of this method implementation a small set of points (compared with initial set) is considered.

PMBDR (Fig. 3) actually allowed to inherit positive characteristics of search space investigation obtained by the methods considered above. Besides it PMBDR extends these advantages, adding extra control points. To represent the depth of search space coverage using parallel computing additional indicator, relative time, was used. It is caused by the fact, that realization of iterations of different strategies at different nodes of parallel system lasts different time, and so it is necessary to normalize this presentation for illustration of overall coverage of search space. The relative time is expressed as a percentage ratio of current time to total running time of the entire parallel system.

As can be seen from the Figure 3, when iteration set is repeated, the initial search space is comparable each time (34.5–37.2 %). If realization of each such repetition is considered separately, it is noticeable that the coverage depth does not reduce as quickly as in the evolutionary search, as a result reducing probability of falling into local optima. Every repetition of iteration set ends with reduction of search space depth to 0, because computations on the main core Pr_0 are realized at that period of time, thus search is not performed.

The results of vehicle recognition problem solving (recognition error and execution time), presented in the Figures 4 and 5, demonstrated, that the best values corresponded to the PMBDR, proposed in the paper.

The developed method allowed to get recognition error of 0.0178, which is 8.2% and 12.3 % more accurate than MMDCA and MMICA correspondingly, 19.8% more accurate than CMES. Thus the best recognition results in terms of accuracy were demonstrated by the methods, which selected feature sets with the lowest cardinal number for the given task.

At the same time, the proposed method proved to be the best in terms of execution time, the value of which was 612 sec. PCA demonstrated comparable speed of work: it has performed recognition process 6 sec. faster. However, its recognition error was almost 2.5 times higher than

recognition error of the proposed method. The next result in these terms showed MARF, which made it possible to perform recognition 3.7 times slower than the proposed method, but was characterized by the largest (among the considered methods) recognition error (0.0484). MMDCA and MMICA, having recognition error which is comparable with the proposed method, realized recognition 17.24 and 16.89 times slower.

Thus it can be argued that the proposed parallel method of large data reduction allows to effectively solve the informative features selection problem, which leads to an effective solution of the problem of pattern recognition, besides in comparison with the other methods of informative feature selection the proposed method is implemented faster with the lowest recognition error.

CONCLUSIONS

In this paper the actual task of automation of feature informativeness estimation process in diagnostics and pattern recognition problems was solved.

Scientific novelty of the paper is in the proposed parallel method of big data reduction. This method is based on the proposed criteria system, which allows to estimate concentration of control points around local extrema. Calculation of solution concentration estimates in the developed criteria system is based on the spatial location of control points in the current solution set. The proposed criteria system can be used in stochastic search methods to monitor situations of excessive solution concentration in the areas of local optima and, as a consequence, to increase the diversity of the solution set in the current population and to cover the search space by control points in a more uniform way during optimization process.

Practical significance of the paper consists in the solution of practical problems of pattern recognition. Experimental results showed that the proposed method allowed to select informative feature set and it could be used in practice for solving of practical tasks of diagnostics and pattern recognition.

ACKNOWLEDGMENTS

The work was performed as part of the research work "Methods and means of decision-making for data processing in intellectual recognition systems" (number of state registration 0117U003920) of software tools department of Zaporizhzhia National Technical University and was partially supported by the international project "Internet of Things: Emerging Curriculum for Industry and Human Applications" (ALIOT, registration number 573818-EPP-1-2016-1-UK-EPPKA2-CBHE-JP) funded by the Erasmus+ programme of the European Union.

REFERENCES

1. Jensen R. Computational intelligence and feature selection: rough and fuzzy approaches / R. Jensen, Q. Shen. – Hoboken: John Wiley & Sons, 2008. – 339 p. DOI: 10.1002/9780470377888.
2. Lee J. A. Nonlinear dimensionality reduction / J. A. Lee, M. Verleysen. – New York : Springer, 2007. – 308 p. DOI: 10.1007/978-0-387-39351-3.
3. Mulaik S. A. Foundations of Factor Analysis / S. A. Mulaik. – Boca Raton, Florida : CRC Press. – 2009. – 548 p.
4. Oliinyk A. Production rules extraction based on negative selection / A. Oliinyk // Radio Electronics, Computer Science, Control. – 2016. – Vol. 1. – P. 40–49. DOI: 10.15588/1607-3274-2016-1-5.
5. McLachlan G. Discriminant Analysis and Statistical Pattern Recognition / G. McLachlan. – New Jersey : John Wiley & Sons, 2004. – 526 p. DOI: 10.1002/0471725293.
6. Bow S. Pattern recognition and image preprocessing / S. Bow. – New York : Marcel Dekker Inc., 2002. – 698 p. DOI: 10.1201/9780203903896.
7. Encyclopedia of machine learning / [eds. C. Sammut, G. I. Webb]. – New York : Springer, 2011. – 1031 p. DOI: 10.1007/978-0-387-30164-8.
8. A comparison of approaches to large-scale data analysis / [A. Pavlo, E. Paulson, A. Rasin, D. J. Abadi et al] // International Conference on Management of Data. – 2009. – P. 165–178. DOI: 10.1145/1559845.1559865.
9. The model for estimation of computer system used resources while extracting production rules based on parallel computations / [A. A. Oliinyk, S. Yu. Skrupsky, V. V. Shkarupylo, S. A. Subbotin] // Радіоелектроніка, інформатика, управління. – 2017. – № 1. – С. 142–152. DOI: 10.15588/1607-3274-2017-1-16.
10. Sulistio A. Simulation of Parallel and Distributed Systems: A Taxonomy - and Survey of Tools / A. Sulistio, C. S. Yeo, R. Buyya // International Journal of Software Practice and Experience. Wiley Press. – 2002. – P. 1–19.
11. Shin Y.C. Intelligent systems : modeling, optimization, and control / C. Y. Shin, C. Xu. – Boca Raton: CRC Press, 2009. – 456 p. DOI: 10.1201/9781420051773.
12. Oliinyk A. A. Information Technology of Diagnosis Model Synthesis Based on Parallel Computing / [A. A. Oliinyk, S. A. Subbotin, S. Yu. Skrupsky et al] // Радіоелектроніка, інформатика, управління. – 2017. – № 3. – С. 139–151.
13. Kira K. A practical approach to feature selection / K. Kira, L. Rendell // Machine Learning : International Conference on Machine Learning ML92, Aberdeen, 1–3 July 1992 : proceedings of the conference. – New York : Morgan Kaufmann, 1992. – P. 249–256. DOI: 10.1016/B978-1-55860-247-2.50037-1.
14. Shitikova O. V. Method of Managing Uncertainty in Resource-Limited Settings / O. V. Shitikova, G. V. Tabunshchyk // Радіоелектроніка, інформатика, управління. – 2015. – № 2. – С. 87–95. DOI: 10.15588/1607-3274-2015-2-11.
15. Guyon I. An introduction to variable and feature selection / I. Guyon, A. Elisseeff // Journal of machine learning research. – 2003. – № 3. – P. 1157–1182.
16. Hyvarinen A. Independent component analysis / A. Hyvarinen, J. Karhunen, E. Oja. – New York: John Wiley & Sons, 2001. – 481 p. DOI: 10.1002/0471221317.
17. Oliinyk A. A. Parallel multiagent method of big data reduction for pattern recognition / A. A. Oliinyk, S. Yu. Skrupsky, V. V. Shkarupylo, O. Blagodariov // Радіоелектроніка, інформатика, управління. – 2017. – № 2. – С. 82–92.
18. Bezdek J. C. Pattern Recognition with Fuzzy Objective Function Algorithms / J. C. Bezdek. – N.Y. : Plenum Press, 1981. – 272 p. DOI: 10.1007/978-1-4757-0450-1.
19. Oliinyk A. Parallel computing system resources planning for neuro-fuzzy models synthesis and big data processing / A. Oliinyk, S. Skrupsky, S. Subbotin, O. Blagodariov, Ye. Gofman // Радіоелектроніка, інформатика, управління. – 2016. – № 4. – С. 61–69. DOI: 10.15588/1607-3274-2016-4-8.
20. Zaigham Mahmood Data Science and Big Data Computing: Frameworks and Methodologies / Zaigham Mahmood // Springer International Publishing. – 2016. – P. 332. DOI: 10.1007/978-3-319-31861-5.
21. Субботін С. О. Неітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережних моделей : мо-

нографія / С. О. Субботін, А. О. Олійник, О. О. Олійник ; під заг. ред. С. О. Субботіна. – Запоріжжя : ЗНТУ, 2009. – 375 с.
22. Subbotin S. Entropy Based Evolutionary Search for Feature Selection / S. Subbotin, A. Oleynik // The experience of designing and application of CAD systems in Microelectronics : IX International Conference CADSM-2007, 20–24 February 2007 : proceedings of the conference. – Lviv, 2007. – P. 442–443. DOI: 10.1109/CADSM.2007.4297612.
23. Oliinyk A. O. Agent technologies for feature selection / A. O. Oliinyk, O. O. Oliinyk and S. A. Subbotin // Cybernetics and Systems Analysis. – 2012. – Vol. 48, Issue 2. – P. 257–267. DOI: 10.1007/s10559-012-9405-z.
24. Oliinyk A. Training Sample Reduction Based on Association Rules for Neuro-Fuzzy Networks Synthesis / A. Oliinyk, T. Zaiko, S. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2014. – Vol. 23, № 2. – P. 89–95. DOI: 10.3103/S1060992X14020039.

Article was submitted 25.03.2018.
After revision 17.04.2018.

Олійник А. О.[1], Субботін С. О.[2], Льовкін В. М.[3], Ильяшенко М. Б.[4], Благодарьов О. Ю.[5]

[1]Канд.техн.наук, доцент, доцент кафедри програмних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

[2]Д-р техн. наук, професор, завідувач кафедри програних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

[3]Канд.техн.наук, доцент, доцент кафедри програних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

[4]Канд. техн. наук, доцент, доцент кафедри комп'ютерних систем та мереж, Запорізький національний технічний університет, Запоріжжя, Україна

[5]Аспірант кафедри програних засобів, Запорізький національний технічний університет, Запоріжжя, Україна

ПАРАЛЕЛЬНИЙ МЕТОД РЕДУКЦІЇ ВЕЛИКИХ ДАНИХ НА ОСНОВІ СТОХАСТИЧНОГО ПРОГРАМУВАННЯ

Актуальність. Вирішено задачу автоматизації задача автоматизації процесу редукції великих даних при діагностуванні та розпізнаванні образів. Об'єкт дослідження – процес редукції великих даних. Предмет дослідження – методи редукції великих даних.

Мета роботи полягає в створенні паралельного методу редукції даних на основі стохастичних обчислень..

Метод. Запропоновано паралельний метод редукції великих даних. Даний метод ґрунтується на запропонованій системі критеріїв, що дозволяють оцінювати концентрованість контрольних точок близько локальних екстремумів. Обчислення оцінок концентрованості рішень в розробленій системі критеріїв засноване на просторовому розташуванні контрольних точок в поточній множині рішень. Запропонована система критеріїв може використовуватися в методах стохастичного пошуку для відстеження ситуацій надмірної концентрації рішень в областях локальних оптимумів, і, як наслідок, для підвищення різноманітності множини рішень в поточній популяції і більш рівномірного покриття простору пошуку контрольними точками в процесі оптимізації.

Результати. Розроблено програмне забезпечення, яке реалізує запропонований паралельний метод редукції великих даних і дозволяє виконувати відбір інформативних ознак і скорочення великих вибірок даних при синтезі розпізнавальних моделей.

Висновки. Проведені експерименти підтвердили працездатність запропонованого паралельного методу редукції великих даних і дозволяють рекомендувати його для використання на практиці при обробці масивів великих даних для розпізнавання образів. Перспективи подальших досліджень можуть полягати в модифікації існуючих і розробки нових методів відбору ознак на основі розробленої системи критеріїв оцінювання концентрованості контрольних точок близько локальних екстремумів.

Ключові слова: вибірка даних, розпізнавання образів, відбір ознак, паралельні обчислення, критерій інформативності, стохастичний підхід.

Олейник А. А.[1], Субботин С. А.[2], Левкин В. Н.[3], Ильяшенко М. Б.[4], Благодарев А. Ю.[5]

[1]Канд. техн. наук, доцент, доцент кафедры программных средств, Запорожский национальный технический университет, Запорожье, Украина

[2]Д-р техн. наук, професоор, заведующий кафедрой программных средств, Запорожский национальный технический университет, Запорожье, Украина

[3]Канд. техн.наук, доцент, доцент кафедры программных средств, Запорожский национальный технический университет, Запорожье, Украина

[4]Канд. техн. наук, доцент, доцент кафедры компьютерных систем и сетей,Запорожский национальный технический университет, Запорожье, Украина

[5]Аспирант кафедры программных средств,Запорожский национальный технический университет, Запорожье, Украина

ПАРАЛЛЕЛЬНЫЙ МЕТОД РЕДУКЦИИ БОЛЬШИХ ДАННЫХ НА ОСНОВЕ СТОХАСТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Актуальность. Решена задача автоматизации процесса редукции больших данных при диагностировании и распознавании образов. Объект исследования – процесс редукции больших данных. Предмет исследования – методы редукции больших данных.

Цель работы заключается в создании параллельного метода редукции данных на основе стохастических вычислений.

Метод. Предложен параллельный метод редукции больших данных. Данный метод основывается на предложенной системе критериев, позволяющих оценивать концентрированность контрольных точек около локальных экстремумов. Вычисление оценок концентрированности решений в разработанной системе критериев основано на пространственном расположении контрольных точек в текущем множестве решений. Предложенная система критериев может использоваться в методах стохастического поиска для отслеживания ситуаций чрезмерной концентрации решений в областях локальных оптимумов, и, как следствие, для повышения разнообразия множества решений в текущей популяции и более равномерного покрытия пространства поиска контрольными точками в процессе оптимизации.

Результаты. Разработано программное обеспечение, которое реализует предложенный параллельный метод редукции больших данных и позволяет выполнять отбор информативных признаков и сокращение больших выборок данных при синтезе распознающих моделей.

Выводы. Проведенные эксперименты подтвердили работоспособность предложенного параллельного метода редукции больших данных и позволяют рекомендовать его для использования на практике при обработке массивов больших данных для распознавания образов. Перспективы дальнейших исследований могут заключаться в модификации существующих и разработки новых методов отбора признаков на основе разработанной системы критериев оценивания концентрированности контрольных точек около локальных экстремумов.

Ключевые слова: выборка данных, распознавание образов, отбор признаков, параллельные вычисления, критерий информативности, стохастический подход.

REFERENCES

1. Jensen R., Shen Q. Computational intelligence and feature selection: rough and fuzzy approaches. Hoboken, John Wiley & Sons, 2008, 339 p. DOI: 10.1002/9780470377888.
2. Lee J. A., Verleysen M. Nonlinear dimensionality reduction. New York, Springer, 2007, 308 p. DOI: 10.1007/978-0-387-39351-3.
3. Mulaik S. A. Foundations of Factor Analysis. Boca Raton, Florida, CRC Press, 2009, 548 p.
4. Oliinyk A. Production rules extraction based on negative selection, *Radio Electronics, Computer Science, Control*, 2016, Vol. 1, pp. 40–49. DOI: 10.15588/1607-3274-2016-1-5.
5. McLachlan G. Discriminant Analysis and Statistical Pattern Recognition. New Jersey, John Wiley & Sons, 2004, 526 p. DOI: 10.1002/0471725293.
6. Bow S. Pattern recognition and image preprocessing. New York, Marcel Dekker Inc., 2002, 698 p. DOI: 10.1201/9780203903896.
7. eds. Sammut C., Webb G. I. Encyclopedia of machine learning. New York, Springer, 2011, 1031 p. DOI: 10.1007/978-0-387-30164-8.
8. Andrew Pavlo, Paulson E., Rasin A., Abadi D. J., DeWitt D. J. A comparison of approaches to large-scale data analysis, *International Conference on Management of Data*, 2009, pp. 165–178. DOI: 10.1145/1559845.1559865.
9. Oliinyk A. A., Skrupsky S. Yu., Shkarupylo V. V., Subbotin S. A. The model for estimation of computer system used resources while extracting production rules based on parallel computations, *Radio Electronics, Computer Science, Control,* 2017, No. 1, pp. 142–152. DOI: 10.15588/1607-3274-2017-1-16.
10. Sulistio A., Yeo C. S., Buyya R. Simulation of Parallel and Distributed Systems: A Taxonomy – and Survey of Tools, *International Journal of Software Practice and Experience. Wiley Press*, 2002, pp. 1–19.
11. Shin Y. C., Xu C. Intelligent systems : modeling, optimization, and control. Boca Raton, CRC Press, 2009, 456 p. DOI: 10.1201/9781420051773.
12. Oliinyk A. A., Subbotin S. A., Skrupsky S. Yu., Lovkin V. M., Zaiko T. A. Information Technology of Diagnosis Model Synthesis Based on Parallel Computing, *Radio Electronics, Computer Science, Control*, 2017, No. 3, pp. 139–151.
13. Kira K., Rendell L. A practical approach to feature selection, *Machine Learning : International Conference on Machine Learning ML92, Aberdeen, 1–3 July 1992 : proceedings of the conference*. New York, Morgan Kaufmann, 1992, pp. 249–256. DOI: 10.1016/B978-1-55860-247-2.50037-1.
14. Shitikova O. V., Tabunshchyk G. V. Method of Managing Uncertainty in Resource-Limited Settings, *Radio Electronics, Computer Science, Control*, 2015, No. 2, pp. 87–95. DOI: 10.15588/1607-3274-2015-2-11.
15. Guyon I., Elisseeff A. An introduction to variable and feature selection, *Journal of machine learning research,* 2003, No. 3, pp. 1157–1182.
16. Hyvarinen A., Karhunen J., Oja E. Independent component analysis. New York, John Wiley & Sons, 2001, 481 p. DOI: 10.1002/0471221317.
17. Oliinyk A. A., Skrupsky S. Yu., Shkarupylo V. V., Blagodariov O. Parallel multiagent method of big data reduction for pattern recognition, *Radio Electronics, Computer Science, Control*, 2017, No. 2, pp. 82–92.
18. Bezdek J. C. Pattern Recognition with Fuzzy Objective Function Algorithms. N.Y., Plenum Press, 1981, 272 p. DOI: 10.1007/978-1-4757-0450-1.
19. Oliinyk A., Skrupsky S., Subbotin S., Blagodariov O., Gofman Ye. Parallel computing system resources planning for neuro-fuzzy models synthesis and big data processing*, Radio Electronics, Computer Science, Control*, 2016, Vol. 4, pp. 61–69. DOI: 10.15588/1607-3274-2016-4-8.
20. Zaigham Mahmood Data Science and Big Data Computing: Frameworks and Methodologies, *Springer International Publishing*, 2016, pp. 332. DOI: 10.1007/978-3-319-31861-5.
21. Subbotin S., Oliinyk A., Oliinyk O. Noniterative, evolutionary and multi-agent methods of fuzzy and neural network models synthesis : monograph. Zaporizhzhya, ZNTU, 2009, 375 p. (In Ukrainian).
22. Subbotin S., Oleynik A. Entropy Based Evolutionary Search for Feature Selection, *The experience of designing and application of CAD systems in Microelectronics : IX International Conference CADSM-2007, 20–24 February 2007 : proceedings of the conference*. Lviv, 2007, pp. 442–443. DOI: 10.1109/CADSM.2007.4297612.
23. Oliinyk A. O., Oliinyk O. O. and Subbotin S. A. Agent technologies for feature selection, *Cybernetics and Systems Analysis*, 2012, Vol. 48, Issue 2, pp. 257–267. DOI: 10.1007/s10559-012-9405-z.
24. Oliinyk A., Zaiko T., Subbotin S. Training Sample Reduction Based on Association Rules for Neuro-Fuzzy Networks Synthesis, *Optical Memory and Neural Networks (Information Optics),* 2014, Vol. 23, No. 2, pp. 89–95. DOI: 10.3103/S1060992X14020039.

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

PROGRESSIVE INFORMATION TECHNOLOGIES

UDC 004.056.5

Alguliyev R. M.[1], Aliguliyev R. M.[2], Imamverdiyev Y. N.[3], Sukhostat L. V.[4]

[1]Full member of Azerbaijan National Academy of Sciences, Dr.Sc, Professor, Director Institute of Information Technologies, Azerbaijan National Academy of Sciences, Baku, Azerbaijan
[2]Corresponding member of Azerbaijan National Academy of Sciences, Dr.Sc, Head of Department, Institute of Information Technology, Azerbaijan National Academy of Sciences, Baku, Azerbaijan
[3]PhD, Associate Professor, Head of Department Information Technology, Azerbaijan National Academy of Sciences, Baku, Azerbaijan
[4]PhD, Senior Researcher, Institute of Information Technology, Azerbaijan National Academy of Sciences, Baku, Azerbaijan

AN IMPROVED ENSEMBLE APPROACH FOR DOS ATTACKS DETECTION

Context. The task of using the ensemble of classifiers to detect DoS attacks in large arrays of network traffic data is solved to withstand attacks on the network.

Objective of this paper is to build an ensemble of classifiers that surpasses single classifiers in terms of accuracy.

Method. To achieve the formulated goal an algorithm, that indicates the probability of belonging to certain classes, which return a vector of classification scores for each point, is proposed. The peculiarity of the proposed approach is that for each point from the dataset, the predicted class label corresponds to the maximum value among all scores obtained by classification methods for a given point. As classifiers, decision trees, k-nearest neighbors algorithm, support vector machines with various kernel functions, and naпve Bayes are considered. A comparative analysis of the proposed approach with single classifiers is considered using the following metrics: accuracy, precision, recall, and F-measure.

Results. The experiments have been performed in R 3.4.1 on the NSL-KDD dataset of network attacks, which was divided into three classes (DoS, normal network behavior and other types of attack).

Conclusions. The conducted experiments have confirmed the efficiency of the proposed approach. The most accurate result showed an ensemble of five classifiers. The development of techniques for attacks detection based on an ensemble of classifiers avoids the problems inherent in most approaches since it is capable of detecting both known and new attacks with high accuracy. It can be concluded that the proposed approach for network attacks detection is of practical significance. In order to further study the attacks detection in network traffic, studies will be performed on real Big data sets.

Keywords: network security, network attacks, DoS, classification, ensemble of classifiers, Big data.

NOMENCLATURE

KDD – Knowledge Discovery and Data Mining;
DoS – Denial of Service Attack;
DDoS – Distributed Denial of Service Attack;
U2R – Users to Root Attack;
R2L – Remote to Local Attack;
Probe – Probing Attack;
TCP – Transmission Control Protocol;
IP – Internet Protocol;
TP – True Positive;
FP – False Positive;
TN – True Negative;
FN – False Negative;
BN – Naïve Bayes;
DT – decision tree;
SVM – support vector machines;
KNN – k-nearest neighbors algorithm;
ANN – artificial neural network;
WOAR-SVM – weighted one-against-rest SVM;
RBPBoost – Resilient Back Propagation Boosting;
MAP – maximum a posteriori probability;
RBF – radial based function kernel function;
Polynom – polynomial kernel function;
Linear – linear kernel function;
$X(t)$ – feature vector measured at time t;
w – normal vector perpendicular to the hyperplane;
b – offset of the hyperplane;
r – distance from the data point to the separating hyperplane;

© Alguliyev R. M., Aliguliyev R. M., Imamverdiyev Y. N., Sukhostat L. V., 2018
DOI 10.15588/1607-3274-2018-2-8

$P(H)$ – a priori probability of each class without information on the variable x;

$P(H \mid x)$ – a posteriori probability of the variable x over the possible classes;

$P(x \mid H)$ – conditional probability of x at the likelihood H;

C – set of class labels;

x_i – dataset of points;

n – number of data points;

m – number of classifiers;

M – classification methods;

k – number of classes;

A – classification scores for each point of the dataset;

P – vector of ensemble scores.

INTRODUCTION

Big data analysis in intrusion detection and in solving network security problems is attracting increasing attention, as it facilitates the study of large amounts of complex and disparate data and detects network intrusions and contributes to the fight against cyber-attacks [1].

Network attacks are one of the causes of the abnormal phenomena observed in the work of the network equipment, as well as traffic transmission over the network. Anomalies of network traffic may result in incorrect operation of a single channel or entire network segments, lead to a denial of service in this network equipment. Network attacks are constantly changing because attackers use individual approaches. It is also affected by changes in software and hardware.

The solution of anomaly detection problem is not trivial since anomalies nature itself is changeable. Providing a comprehensive definition of abnormal or normal behavior in the context of a computer network is quite subtle [2, 3]. Another reason is that some anomaly detection methods require labels of normal and abnormal behaviors that are difficult to obtain [4, 5]. In addition, choose the right tool for anomalies detection is not easy. In [6], the study has shown that the intended tool may well be suitable for only one type of abnormality, but not for all. It is a very realistic assumption, that the selection of anomaly detection method is not simple when anomaly types are not known a priori. In addition, the network scale is the problem: when it detects anomalies it needs to consider the distribution of the tasks implementation process between multiple network servers in order to increase the overall performance and the system ability to work in case of failure of its individual elements, taking into account the size of the growing networks [7, 8].

Vulnerabilities in the communication protocol stack (TCP/IP) result in intentional or unintentional distributed denial of service (DDoS) attacks. DDoS attacks can be detected using existing machine learning methods.

The latest research works have been based on binary classification methods, which can distinguish between two states (“normal” or “abnormal”). In the event of a conflict between binary classifiers, the final solution is achieved by comparing their accuracy. An alternative solution is possible by developing an ensemble of classifiers. A new approach is required to combine such classifiers in the ensemble.

The research objective is to construct an ensemble of classifiers that surpasses single classifiers in terms of accuracy in order to detect DoS attacks.

1 PROBLEM STATEMENT

To solve the task of DoS attacks detection in network traffic, an ensemble of classifiers that surpasses single classifiers in terms of accuracy is suggested in this paper.

Let us denote the following notations (Table 1): $x_i \in R^n \ (i = \overline{1,n})$ is the point from the dataset, where n is the total number of points in the input dataset, $M = \{M_1, M_2, ..., M_m\}$ is the set of classification methods, subsequently combined into an ensemble, m is the number of classification methods, a_{ij} is the classification score for the point x_i, k is the number of classes.

Table 1 – Evaluation of classifiers for each point

Data point	Classifiers	Ensemble score
X	$M_1 \ \dots \ M_m$	
x_1	$a_{11} \ \dots \ a_{1m}$	$\max\limits_j(a_{1j})$
$\vdots$	$\vdots \ \ddots \ \vdots$	$\vdots$
x_n	$a_{n1} \ \dots \ a_{nm}$	$\max\limits_j(a_{nj})$

It is necessary to obtain a vector of ensemble scores on the basis of single classifiers to each data point in order to improve the classification accuracy.

2 REVIEW OF THE LITERATURE

A number of studies and review articles have been devoted to the intrusion detection technology [9, 10] or data mining for specific applications [11]. Since the introduction of intrusion detection principles by Denning in 1987, a large number of reactive protection systems have been developed [12–14].

Intrusion detection methods can be divided into three categories: single, hybrid, and assemblies [15]. Support vector machines (SVMs) and artificial neural networks (ANNs) are the most popular approaches among single classifiers. Several classifiers are combined to a higher goal of a significant increase in efficiency of the classifier known as an ensemble of classifiers [16]. The majority votes, bagging and boosting are some common strategies for combining classifiers [17]. Although it is known that the disadvantages of classifiers’ components accumulate in the ensemble of classifiers, but it works very effectively in varying combinations. Thus, the researchers become more and more interested in the application of the ensemble of classifiers every day.

The important cybersecurity problems for mathematical and statistical solutions have been shown in [18]. A method to improve the detection accuracy by an ensemble of two-layer SVM based on rotation forest was presented in [19]. The experiments were conducted on the KDD CUP 1999 dataset. The output of ensemble network was made by majority voting. The second layer result is used to focus on two classes “normal” and “attack”.

Classification accuracy has been improved by combining opinions from multiple experts into one using an ensemble approach in [20]. The ensemble construction method uses PSO generated weights to create the ensemble of classifiers with better accuracy for intrusion detection. This work was based on binary classification methods, which can distinguish between two states. In case of conflict between binary classifiers final decision has been reached by comparing their accuracy.

In [21] a multistep framework based on machine learning techniques to create an efficient classifier was introduced. A novel fuzzy weighting method for ensemble classifiers was proposed. Thus, adding the fuzzy weighted combiner can tag weights to classifiers related to their cost and performance.

Architecture of intelligent false alarm filter by employing a method of voted ensemble selection aiming to maintain the accuracy of false alarm reduction was proposed in [22]. The experiment was conducted using SVM, decision tree, and k-nearest neighbor (KNN) machine learning algorithms. The proposed method was validated on a real dataset.

The paper [23] aims to identify multiclass SVM models best suited to the intrusion detection task. A new approach (WOAR-SVM) based on a set of optimal, or near-optimal, weight coefficients, which define the relationship between the decision rules of the binary SVM classifiers was developed.

A generic architecture for automated DDoS attack detection and response system for the collaborative environment using machine learning were proposed in [24]. The main objective of this paper was to minimize the cost of classification errors of the intrusion detection. The proposed classification algorithm, RBPBoost, was achieved by combining ensemble of classifier outputs and Neyman Pearson cost minimization strategy, for final classification decision.

3 MATERIALS AND METHODS

More information about intrusion detection can be obtained by data classification methods. Theoretically, classification algorithms can achieve high performance, i.e. they can minimize the number of false alarms and maximize detection accuracy. One of the most attractive features of the algorithms is the ability to distinguish normal from abnormal behavior [8]. In the context of the intrusion detection, the classification algorithm is typically a map which adapts to the network invisible abnormalities [25].

Formally, each data instance is a feature vector X measured at time t and denoted as $X(t)$.

The classification algorithm is aimed to train the function that maps all the samples to their own states. To achieve its purpose, they use a set of data instances within the network. This set is known as the training dataset. Some algorithms learn a mapping function by the use of labeled training sets, where each sample in the training set is marked as one of the states. These algorithms are called supervised learning algorithms. The purpose of the use of these algorithms is to achieve high classification accuracy.

The most popular supervised machine learning methods include SVM, decision trees, Bayesian networks, KNN algorithm, etc.

The aim of SVM is to classify data points X of an n-dimensional space using $(n-1)$-dimensional hyperplane. The hyperplane satisfies $w^T x + b = 0$, where X is a set of points, w is a normal vector perpendicular to the hyperplane and b is an offset of the hyperplane $w^T x + b = 0$ from the initial point along the direction of w.

The distance from the data point to the separating hyperplane $w^T x + b = 0$ can be calculated as $r = (w^T x + b)/\|w\|$. The closest to the hyperplane data points are called support vectors. The distance between support vectors is known as margin. Linear SVM can be achieved by quadratic optimization:

$$\arg\min_{w,b}\left(\frac{1}{2}\|w\|^2\right),$$
$$y(w^T x + b) \geq 1. \quad (1)$$

SVM can find accurately linear, nonlinear and complex classification boundaries, even with a small amount of training sample.

SVMs are widely used for transmission of various type of data by switching the kernel function. The most used kernel functions include linear, polynomial, radial based function and sigmoid.

However, choosing the kernel function and fit the relevant parameters by SVM are still in the procedure of trial and error. SVM is fast, but its duration is increased four times when the data size of the sample is doubled. Unfortunately, the root of SVM algorithms is in binary classification. To solve the problems of multi-class classification several SVM for binary classes can be combined by the classification of each class or classification of each class pair.

A decision tree (DT) is a tree-structure model, which has leaves that represent classes or solutions, and branches that represent conjunctions of features that lead to those classifications.

Tree-structure classification of an input vector is performed by bypassing the tree from the root node to the end with a leaf. Each tree node computes inequality on the basis of one of the input variables. Each leaf is assigned to a particular class. Each inequality, which is used to divide the input space is based only on one of the input variables. Linear DTs are like binary DTs, except for the fact that inequality calculated at each node has a random linear form, which may depend on several variables. DT depends on the rules of "if-then", but does not require any parameters and metrics. This simple and interpretable structure allows decision trees to solve the problem for different types of attributes. DTs can also manage the missing values or noisy data. However, they can not guarantee the optimal accuracy, unlike other machine learning techniques. Although decision trees are easy to learn and implement, they are not often used for intrusion detection. This is due to the fact that finding the smallest decision tree is NP-hard.

Bayesian network classifier is based on Bayes' rule, which gives the hypothesis H of classes and data x

$$P(H \mid x) = \frac{P(x \mid H)P(H)}{P(x)}, \quad (2)$$

where $P(H)$ represents the a priori probability of each class without information on the variable x, $P(H \mid x)$ is a posteriori probability of the variable over the possible classes, is the conditional probability of x at this likelihood H. Bayesian network nodes are represented with random variables and arcs representing probabilistic relationships between variables and conditional probabilities. Node always calculates the posterior probabilities, giving proof of inheritance for the selected nodes.

Naïve Bayes (NB) is a simple Bayesian network model, which assumes that all variables are independent. It is necessary to find the maximum likelihood hypothesis, which defines the class label for the test data x, for classification by NB.

NB classifier can be resolved by the hypothesis of maximum a posteriori probability (MAP) for data as follows:

$$\arg\max_{c_j \in C} P(x \mid c_j) P(c_j), \tag{3}$$

where x is an observable data, and $C = \{c_j\}$ is a set of class labels.

Naïve Bayes is effective for tasks with a logical conclusion. However, Naıve Bayes is based on the strong assumption of variables independence.

Numbers of nearest neighbors k and distances measures are key components of the KNN algorithm. Selection of the number k should be based on cross-validation. By increasing the number k, the effect of noise in the data during classification is reduced, and this can erase the difference between the classes. In practice, k have to be less than the square root of the total number of training samples.

In the case of multiclass classification, KNN method is based on measuring the distance from one data sample to each trained sample [26]. The k-smallest distances are calculated, and the most common class based on these KNNs is considered to be the label of the output class.

KNN does not require training parameters. It is easy to implement, but it requires a lot of memory and time.

The proposed algorithm, that indicates the probability of belonging to certain classes, returns a vector of classification scores for each data point.

The peculiarity of the proposed approach is that for each point from the dataset, the predicted class label corresponds to the maximum value among all scores obtained by clustering methods for a given point.

The algorithm of the proposed approach for network attacks detection based on an ensemble of classifiers is presented below:

Input: $x_i \in R^n$: dataset of points

n: number of data points

m: number of classifiers

$M = \{M_1, M_2, ..., M_m\}$: classification methods

k: number of classes

Output:

$A = \{a_{ij}\}_{n \times m}$: classification scores for each point of the dataset

P: vector of ensemble scores

for i=1 to n do

for j=1 to do

Calculate the value of a_{ij} for M_i

End

$$P_i = \max_j a_{ij}$$

End

It is required to increase the accuracy of DoS attacks detection by using an ensemble of classifiers.

4 EXPERIMENTS

For the experiments was considered NSL-KDD dataset of network attacks [27], built on the basis of KDD-99 database on the initiative of the American Association for the Defense Advanced Research Projects Agency (DARPA) [28]. A dataset of connections was collected to conduct the research in the field of intrusion detection, which covers a wide range of intrusions simulated in a medium that mimics the US Air Force network.

Statistical analysis showed that there are important issues in the databases that highly affect the performance of the systems, as well as lead to a very bad evaluation of anomaly detection approaches. The considered database NSL-KDD has the following advantages:

1. No redundant records in the training set, so that the classifier does not show any bias results.

2. Duplicate records are not present in the test set. It contains some of the attacks that are not present in the training set.

3. A number of records selected from each difficulty levels group are inversely proportional to the number of records in the original dataset KDD.

The training set consists of 21 different attacks of 37 present in the test set. In addition, the number of records in the training (125973 samples) set and test set (22544 samples) of NSL-KDD is acceptable. This advantage makes it accessible for experiments on comprehensive data without the need to randomly select a small portion of data. Consequently, the evaluation of the results of various research projects is consistent and comparable.

All attacks in NSL-KDD are divided into four groups:

– DoS attacks include: “neptun”, “back”, “smurf”, “pod”, “land”, and “teardrop”;

– U2R (Users to Root Attack) attacks include: “buffer_overflow”, “loadmodule”, “rootkit”, and “perl”;

– R2L (Remote to Local Attack) attacks include: “warezclient”, “multihop”, “ftp_write”, “imap”, “guess_passwd”, “warezmaster”, “spy”, and “phf”;

– Probe (Probing Attack) attacks include the following types of attack: “portsweep”, “satan”, “nmap”, and “ipsweep”.

The main objectives put forward in network intrusion detection include recognition of rare types attack, increasing the accuracy of suspicious activity detection, as well as increasing the efficiency of real-time intrusion detection models. Each record has 41 attributes, which describe various features.

To evaluate the performance of classifiers, the following metrics are used: Accuracy, Recall, Precision, and F-measure. For any classification algorithm, four classification cases are possible, and this helps to understand the difference between the following metrics: True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) results.

Classification accuracy can be defined as the proportion of the correct results, which is achieved by the classifier:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (4)$$

Precision shows how much of the objects identified by the classifier as positive are really positive:

$$precision = \frac{TP}{TP + FP}. \quad (5)$$

Recall shows which part of the positive objects was selected by the classifier:

$$recall = \frac{TP}{TP + FN}. \quad (6)$$

F-measure is a metric that combines the recall and precision:

$$F - measure = \frac{2 \times recall \times precision}{recall + precision}. \quad (7)$$

Error rate (or Misclassification error) measures the ratio of incorrectly classified samples over the total number of classified samples:

$$Error\ rate = \frac{FP + FN}{TP + TN + FP + FN}. \quad (8)$$

5 RESULTS

It is the experiments were conducted using Windows® 10–64 bits operating system platform with core i7 processor 2.5 GHz, 8.0 GB RAM. The proposed approach was evaluated on R 3.4.1. NSL-KDD dataset was divided into 3 classes (DoS, Normal and Other attacks (U2R, R2L, and Probe)).

A comparative analysis of the proposed approach with single classifiers has been carried out. The ensemble of classifiers consisted of combinations of DT, SVM with various kernel functions, NB and KNN.

Accuracy, Recall, Precision, and F-measure were considered as evaluation metrics. The classification results are shown in Tables 2–5. From these tables, ranks (shown in square brackets) were obtained for each metric of each class, the results of which are shown in Tables 6–8.

It can be concluded from Table 2 that the highest accuracy (92.33%) of DoS attacks detection was achieved for the ensemble of five classifiers (DT+KNN+SVM(Polynom)+ +NB+SVM(Linear)), which exceeded the result of the single classifier (KNN) by 4.12%.

Despite the fact that NB shows the lowest result (80.45%), when adding it to the ensemble of classifiers, the accuracy of the proposed approach increased and amounted to 92.19% for four classifiers (DT+KNN+SVM(Polynom)+NB), and for six classifiers (DT+KNN+SVM(Polynom)+NB+SVM(Linear)+ +SVM (RBF)) – 91.89%.

Comparison of the recall and precision values for DoS attacks detection is shown in Tables 3 and 4, respectively.

Table 5 presents the F-measure results, where the performance of classification methods in general, combining the Recall and Precision values, is evaluated.

For the DoS class, the ensemble DT+KNN+SVM(Polynom)+NB+SVM(Linear) showed the best rank [29] according to the three metrics Accuracy, Precision and F-measure, according to the metric Recall the ensemble DT+KNN+SVM(Polynom)+NB demonstrates the best result. From Table 6, the worst result showed the NB method under three metrics (out of four), and the DT method had the worst rank by one metric (Recall).

From Table 7, the ensemble DT+KNN+ +SVM(Polynom)+NB showed the best results for the Normal class under the three metrics, and the worst result showed the NB method according to Accuracy, Recall, and F-measure.

Table 2 – Comparison of the classification accuracy of the proposed algorithm with other classifiers

Class / Method	DoS	Normal	Other attacks
DT	86.32% [9]	77.19% [9]	64.25% [10]
KNN	88.21% [5]	79.67% [5]	65.80% [8]
SVM(Linear)	87.21% [7]	77.96% [8]	66.22% [7]
SVM(Polynom)	86.64% [8]	79.50% [6]	68.31% [5]
SVM(RBF)	87.25% [6]	78.84% [7]	65.38% [9]
NB	80.45% [10]	71.25% [10]	66.50% [6]
DT+KNN+SVM(Polynom)	90.74% [4]	84.77% [4]	74.53% [4]
DT+KNN+SVM(Polynom)+NB	92.19% [2]	88.63% [1]	82.76% [3]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)	92.33% [1]	88.58% [3]	83.35% [1]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)+SVM(RBF)	91.89% [3]	88.60% [2]	82.93% [2]

Table 3 – Comparison of classification methods in terms of Precision

Class / Method	DoS	Normal	Other attacks
DT	95.94% [7]	63.17% [10]	88.15% [6]
KNN	95.88% [8]	65.87% [6]	86.05% [8]
SVM(Linear)	96.65% [5]	64.37% [9]	84.47% [9]
SVM(Polynom)	96.14% [6]	65.70% [7]	88.12% [7]
SVM(RBF)	85.07% [9]	64.92% [8]	96.17% [2]
NB	73.03% [10]	74.78% [4]	41.80% [10]
DT+KNN+SVM(Polynom)	96.74% [4]	71.95% [5]	94.41% [5]
DT+KNN+SVM(Polynom)+NB	97.94% [2]	77.74% [1]	95.76% [4]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)	99.98% [1]	76.81% [3]	100% [1]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)+SVM(RBF)	97.91% [3]	77.60% [2]	96.06% [3]

Table 4 – Comparison of classification methods in terms of Recall

Class / Method	DoS	Normal	Other attacks
DT	74.19% [10]	97.33% [8]	29.75% [10]
KNN	78.08% [5]	97.62% [5]	33.28% [8]
SVM(Linear)	75.72% [7]	96.19% [9]	34.42% [7]
SVM(Polynom)	74.77% [8]	97.50% [7]	38.23% [6]
SVM(RBF)	76.00% [6]	97.57% [6]	32.54% [9]
NB	74.51% [9]	57.07% [10]	58.51% [4]
DT+KNN+SVM(Polynom)	82.86% [4]	98.62% [4]	49.99% [5]
DT+KNN+SVM(Polynom)+NB	85.28% [1]	98.63% [3]	66.44% [3]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)	84.67% [2]	100% [1]	66.70% [2]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)+SVM(RBF)	84.66% [3]	98.77% [2]	66.72% [1]

Table 5 – Comparison of classification methods in terms of F-measure

Class / Method	DoS	Normal	Other attacks
DT	83.67% [9]	76.62% [9]	44.48% [10]
KNN	86.07% [5]	78.66% [5]	48.00% [8]
SVM(Linear)	84.91% [6]	77.13% [8]	48.91% [6]
SVM(Polynom)	84.12% [8]	78.50% [6]	53.33% [5]
SVM(RBF)	84.90% [7]	77.97% [7]	47.07% [9]
NB	73.76% [10]	64.74% [10]	48.76% [7]
DT+KNN+SVM(Polynom)	89.27% [4]	83.20% [4]	65.37% [4]
DT+KNN+SVM(Polynom)+NB	91.17% [2]	86.95% [1]	78.45% [3]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)	91.69% [1]	86.88% [3]	80.02% [1]
DT+KNN+SVM(Polynom)+NB+SVM (Linear)+SVM(RBF)	90.81% [3]	86.92% [2]	78.74% [2]

The best rank for Other attacks class is obtained for ensemble DT+KNN+SVM(Polynom)+NB+SVM(Linear) according to the Accuracy, Precision, and F-measure metrics. The ensemble DT+KNN+SVM(Polynom)+NB+ +SVM(Linear)+SVM(RBF) showed the best result under one metric (Recall) (Table 8).

6 DISCUSSION

From Tables 6–8, we can conclude:

1) The best results among the methods for network attacks detection showed the ensemble DT+KNN+SVM(Polynom)+NB.

2) The methods KNN and DT (Table 6), NB and SVM (Linear) (Table 7) and SVM (RBF) and SVM (Linear) have the equal ranks (Table 8).

3) The proposed approach with different combinations of classifiers is superior to single classifiers.

To show a comparison of methods more clearly, we demonstrate this in Fig. 1, where the error rates for each classifier are shown. For each classifier, the error rates were computed. In the figure, the red color indicates the classification error rates for DoS attacks, the green color indicates the error rates for other types of attack and the blue color shows the error rates for "normal" state. Single classifiers (DT, NB, SVM, and KNN) were compared with various ensembles of classifiers using the proposed approach.

Table 6 – The resultant rank of the methods for DoS class

Method	The number of times the method is in the S^{th} rank S =										Resultant rank	
	1	2	3	4	5	6	7	8	9	10		
DT+KNN+SVM(Polynom)+NB+SVM (Linear)	3	1	0	0	0	0	0	0	0	0	3.9	1
DT+KNN+SVM(Polynom)+NB	1	3	0	0	0	0	0	0	0	0	3.7	2
DT+KNN+SVM(Polynom)+NB+SVM (Linear)+SVM(RBF)	0	0	4	0	0	0	0	0	0	0	3.2	3
DT+KNN+SVM(Polynom)	0	0	0	4	0	0	0	0	0	0	2.8	4
SVM(Linear)	0	0	0	0	1	1	2	0	0	0	1.9	5
SVM(RBF)	0	0	0	0	0	2	1	0	1	0	1.6	6
SVM(Polynom)	0	0	0	0	0	1	0	3	0	0	1.4	7
KNN	0	0	0	0	3	0	0	1	0	0	0.9	8
DT	0	0	0	0	0	0	1	0	2	1	0.9	9
NB	0	0	0	0	0	0	0	0	1	3	0.5	10

Table 7 – The resultant rank of the methods for Normal class

Method	The number of times the method is in the S^{th} rank S =										Resultant rank	
	1	2	3	4	5	6	7	8	9	10		
DT+KNN+SVM(Polynom)+NB	3	0	1	0	0	0	0	0	0	0	3.8	1
DT+KNN+SVM(Polynom)+NB+SVM (Linear)+SVM(RBF)	0	4	0	0	0	0	0	0	0	0	3.6	2
DT+KNN+SVM(Polynom)+NB+SVM (Linear)	1	0	3	0	0	0	0	0	0	0	3.4	3
DT+KNN+SVM(Polynom)	0	0	0	3	1	0	0	0	0	0	2.7	4
KNN	0	0	0	0	3	1	0	0	0	0	2.3	5
SVM(Polynom)	0	0	0	0	0	2	2	0	0	0	1.8	6
SVM(RBF)	0	0	0	0	0	1	2	1	0	0	1.6	7
NB	0	0	0	1	0	0	0	0	0	3	1.0	8
SVM(Linear)	0	0	0	0	0	0	0	2	2	0	1.0	9
DT	0	0	0	0	0	0	0	1	2	1	0.8	10

Table 8 – The resultant rank of the methods for Other attacks class

Method	The number of times the method is in the S^{th} rank S =										Resultant rank	
	1	2	3	4	5	6	7	8	9	10		
DT+KNN+SVM(Polynom)+NB+SVM (Linear)	3	1	0	0	0	0	0	0	0	0	3.9	1
DT+KNN+SVM(Polynom)+NB+SVM (Linear)+SVM(RBF)	1	2	1	0	0	0	0	0	0	0	3.6	2
DT+KNN+SVM(Polynom)+NB	0	0	3	1	0	0	0	0	0	0	3.1	3
DT+KNN+SVM(Polynom)	0	0	0	2	2	0	0	0	0	0	2.6	4
SVM(Polynom)	0	0	0	0	2	1	1	0	0	0	2.1	5
NB	0	0	0	1	0	1	1	0	0	1	1.7	6
SVM(RBF)	0	1	0	0	0	0	0	0	3	0	1.5	7
SVM(Linear)	0	0	0	0	0	1	2	0	1	0	1.5	8
KNN	0	0	0	0	0	0	0	4	0	0	1.2	9
DT	0	0	0	0	0	1	0	0	0	3	0.8	10

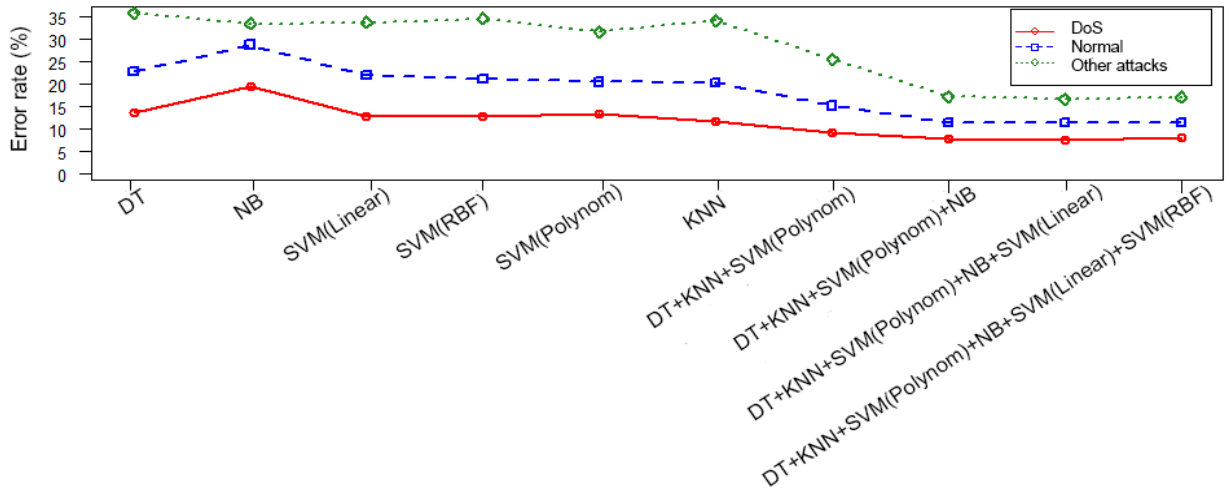


Figure 1 – Error rates for the classifiers on NSL-KDD dataset

KNN, SVM(Linear), and SVM(RBF) gave the lowest error rates among the single classifiers for DoS class (11.79%, 12.79%, and 12.75%, respectively). The lowest result for the DoS class showed the NB method and amounted to 19.55%. Despite the fact that the NB classifier has the highest error rate, in combination with other classifiers in the proposed approach, it showed a good result.

CONCLUSIONS

At present, the processing and analysis of Big data are important for ensuring information security. Intrusion detection is one of the serious problems in the field of network security. In this study, in order to resist attacks on the network, the ensemble of classifiers was successfully applied. The ensemble improves recognition accuracy by combining various single classifiers. The ensemble of classifiers consisted of combinations of DT, SVM with various kernel functions, NB and KNN algorithms.

In general, the considered classification methods showed high accuracy of DoS attacks detection. The most accurate result was shown by an ensemble of five classifiers (DT+KNN+SVM(Polynom)+NB+SVM(Linear)). It can be concluded that the proposed approach for network attacks detection is of practical significance.

ACKNOWLEDGEMENTS

This work was supported by the Science Development Foundation under the President of the Republic of Azerbaijan – Grant № EƏF-KETPL-2-2015-1(25)-56/05/1.

REFERENCES

1. Aliguliyev R. M. Multidisciplinary problems of big data in information security / R. M. Aliguliyev, Y. N. Imamverdiyev, M. S. Hajirahimova // Proceedings of the II International scientific and practical conference Information Security and Computer Technologies. – 2017. – P. 10-11.
2. Nallaivarothayan H. An evaluation of different features and learning models for anomalous event detection / H. Nallaivarothayan, D. Ryan, S. Denman, S. Sridharan, C. Fookes // Proceedings of the International Conference on Digital Image Computing: Techniques and Applications. – 2013. – P. 1–8. DOI: 10.1109/dicta.2013.6691480
3. Xie M. Anomaly detection in wireless sensor networks: a survey / M. Xie, S. Han, B. Tian, S. Parvin // Journal of Network and Computer Applications. – 2011. – Vol. 34. – P. 1302–1325. DOI: 10.1016/j.jnca.2011.03.004
4. Davis J. J. Data preprocessing for anomaly based network intrusion detection: a review / J. J. Davis, A. J. Clark // Computers & Security. – 2011. – Vol. 30. – P. 353–375. DOI: 10.1016/j.cose.2011.05.008
5. Fiorea U. Network anomaly detection with the restricted boltzmann machine / U. Fiorea, F. Palmierib, A. Castiglionec, A. D. Santis // Neurocomputing. – 2013. – Vol. 122. – P. 13–23. DOI: 10.1016/j.neucom.2012.11.050
6. Chandola V. Anomaly detection: a survey / V. Chandola, A. Banerjee, V. Kumar // ACM Computing Surveys. – 2009. – Vol. 41, № 3. – P. 1–58. DOI: 10.1145/1541880.1541882
7. Anceaume E. Anomaly characterization in large scale networks / E. Anceaume, Y. Busnel, E. L. Merrer, R. Ludinard, J. Marchand, B. Sericola // Proceedings of the 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks. – 2014. – P. 68–79. DOI: 10.1109/dsn.2014.23
8. Dua S. Data mining and machine learning in cybersecurity / S. Dua, X. Du. – Boca Raton, FL: CRC Press, 2011. – 256 p. DOI: 10.1201/b10867
9. Catania C. A. Automatic network intrusion detection: current techniques and open issues / C. A. Catania, C. G. Garino // Computers and Electrical Engineering. – 2012. – Vol. 38, № 5. – P. 1062–1072. DOI: 10.1016/j.compeleceng.2012.05.013
10. Ahmed M. A survey of network anomaly detection techniques / M. Ahmed, A. Mahmood, J. Hu // Journal of Network and Computer Applications. – 2016. – Vol. 60. – P. 19–31. DOI: 10.1016/j.jnca.2015.11.016
11. Wu S. X. The use of computational intelligence in intrusion detection systems: a review / S. X. Wu, W. Banzhaf // Applied Soft Computing. – 2010. – Vol. 10, № 1. – P. 1–35. DOI: 10.1016/j.asoc.2009.06.019
12. Chandola V. Data mining for cyber security / V. Chandola, E. Eilertson, L. Ertoz, G. Simon, V. Kumar. – New York: Springer, 2006. – 159 p. – (Data Warehousing and Data Mining Techniques for Computer Security.). DOI: 10.1007/978-0-387-47653-7
13. Lee W. A framework for constructing features and models for intrusion detection systems / W. Lee, S. J. Stolfo // ACM Transactions on Information and System Security. – 2000. – Vol. 3, № 4. – P. 227–261. DOI: 10.1145/382912.382914
14. Mahoney M. V. Learning nonstationary models of normal network traffic for detecting novel attacks / M. V. Mahoney, P. K. Chan / / Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining. – 2002. – P. 376–386. DOI: 10.1145/775047.775102
15. Hodge V. A survey of outlier detection methodologies / V. Hodge, J. Austin // Artificial Intelligence Review. – 2004. – Vol. 22, № 2. – P. 85–126. DOI: 10.1023/b:aire.0000045502.10941.a9
16. Farid D. M. Adaptive intrusion detection based on boosting and naive bayesian classifier / D. M. Farid, M. Z. Rahman, C. M. Rahman // International Journal of Computer Applications. – 2011. – Vol. 24, № 3. – P. 12–19. DOI: 10.5120/2932-3883

17. Laurentys C. A. A novel artificial immune system for fault behavior detection / C. A. Laurentys, R. M. Palhares, W. M. Caminhas // Expert Systems with Applications. – 2011. – Vol. 38. – P. 6957–6966. DOI: 10.1016/j.eswa.2010.12.019
18. Meza J. Mathematical and statistical opportunities in cybersecurity / J. Meza, S. Campbell, D. Bailey // Paper LBNL-1667E, Lawrence Berkeley National Laboratory, Berkeley, CA. – 2009. – P. 1–11. DOI: 10.2172/950976
19. Lin L. SVM ensemble for anomaly detection based on rotation forest / L. Lin, R. Zuo, S. Yang, Z. Zhang // Proceedings of the 3rd International Conference on Intelligent Control and Information Processing. – 2012. – P. 150–153. DOI: 10.1109/icicip.2012.6391455
20. Aburomman A. A novel SVM-kNN-PSO ensemble method for intrusion detection system / A. A. Aburomman, M. B. I. Reaz // Applied Soft Computing. – 2016. – Vol. 38. – P. 360–372. DOI: 10.1016/j.asoc.2015.10.011
21. Masarat S. A novel framework, based on fuzzy ensemble of classifiers for intrusion detection systems / S. Masarat, H. Taheri, S. Sharifian // Proceedings of the 4th International Conference on Computer and Knowledge Engineering. – 2014. – P. 165–170. DOI: 10.1109/iccke.2014.6993345
22. Meng Y. Enhancing false alarm reduction using voted ensemble selection in intrusion detection / Y. Meng, L. F. Kwok, W. Li // International Journal of Computational Intelligence Systems. – 2013. – Vol. 6, № 4. – P. 626–638. DOI: 10.1080/18756891.2013.802114
23. Aburomman A. A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems / A. A. Aburomman, M. B. I. Reaz // Information Sciences. – 2017. – Vol. 414. – P. 225–246. DOI: 10.1016/j.ins.2017.06.007
24. Kumar P. A. R. Distributed denial of service attack detection using an ensemble of neural classifier / P. A. R. Kumar, S. Selvakumar // Computer Communications. – 2011. – Vol. 34, № 11. – P. 1328–1341. DOI: 10.1016/j.comcom.2011.01.012
25. Nguyen N.T. Advances in multimedia and network information system technologies / N. T. Nguyen, A. Zgrzywa, A. Czyzewski. – Berlin: Springer-Verlag, 2010. – 318 p. DOI: 10.1007/978-3-642-14989-4
26. Zhang M.-L. A k-nearest neighbor based algorithm for multi-label classification / M.-L. Zhang, Z.-H. Zhou // Proceedings of IEEE International Conference on Granular Computing. – 2005. – P. 718–721. DOI: 10.1109/grc.2005.1547385
27. Aggarwal P. Analysis of KDD dataset attributes-class wise for intrusion detection / P. Aggarwal, S. K. Sharma // Procedia Computer Science. – 2015. – Vol. 57. – P. 842–851. DOI: 10.1016/j.procs.2015.07.490
28. McHugh J. Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by Lincoln laboratory / J. McHugh // ACM Transactions on Information and System Security. – 2000. – Vol. 3, № 4. – P. 262–294. DOI: 10.1145/382912.382923
29. Aliguliyev R. M. Performance evaluation of density-based clustering methods / R. M. Aliguliyev // Information Sciences. – 2009. – Vol. 179. – P. 3583–3602. DOI: 10.1016/j.ins.2009.06.012

Article was submitted 29.12.2017.
After revision 15.01.2018.

Алгуліев Р. М.[1], Алигуліев Р. М.[2], Імамвердієв Я. Н.[3], Сухостат Л. В.[4]

[1]Академік Національної Академії Наук Азербайджану, д-р техн. наук, проф., директор Інституту Інформаційних Технологій Національної Академії Наук Азербайджану, Баку, Азербайджан

[2]Член-кореспондент Національної Академії Наук Азербайджану, д-р техн. наук, зав. відділом, Інститут Інформаційних Технологій Національної Академії Наук Азербайджану, Баку, Азербайджан

[3]Канд. техн. наук, доцент, зав. відділом, Інститут інформаційних технологій Національної Академії Наук Азербайджану, Баку, Азербайджан

[4]Канд. техн. наук, старший науковий співробітник, Інститут інформаційних технологій Національної Академії Наук Азербайджану, Баку, Азербайджан

ПОКРАЩЕНИЙ ПІДХІД ВИЯВЛЕННЯ DOS АТАК З ВИКОРИСТАННЯМ АНСАМБЛЮ

Актуальність. Розв'язана задача застосування ансамблю класифікаторів для виявлення DoS атак у великих масивах даних мережевого трафіку, щоб протистояти атакам в мережі.

Мета роботи полягає в побудові ансамблю класифікаторів, що перевершує поодинокі класифікатори з точки зору точності.

Метод. Для досягнення поставленої мети в роботі запропоновано алгоритм, який вказує на ймовірність приналежності до певних класів, який повертає вектор оцінок класифікаторів для кожної точки. Особливість запропонованого підходу полягає в тому, що для кожної точки з набору даних передбачена мітка класу відповідає максимальному значенню серед всіх оцінок, отриманих методами класифікації для даної точки. Як класифікаторів були розглянуті дерева рішень, алгоритм *k*-найближчих сусідів, машини опорних векторів з різними ядерними функціями і наївний байесовский класифікатор. Порівняльний аналіз запропонованого підходу з розглянутими одиничними класифікаторами проводиться за наступними метрик: точність, повнота, «влучність» і *F*-міра.

Результати. Експерименти були проведені на мові R 3.4.1 на наборі даних мережевих атак NSL-KDD, який був розбитий на три класи (DoS, «нормальну» поведінку мережі та інші типи атак).

Висновки. Проведені експерименти підтвердили працездатність запропонованого підходу. Найбільш точний результат показав ансамбль з п'яти класифікаторів. Розробка техніки виявлення атак, заснованої на застосуванні ансамблю класифікаторів, дозволяє уникнути проблем, характерних для більшості підходів, оскільки він здатний з високою точністю виявити як відомі, так і нові атаки. Можна зробити висновок про практичну значимість запропонованого підходу до виявлення атак в мережі. З метою подальшого вивчення виявлення атак в мережевому трафіку будуть проведені дослідження на реальних наборах даних великої розмірності.

Ключові слова: інформаційна безпека, мережеві атаки, DoS, класифікація, ансамбль класифікаторів, Big data.

Алгулиев Р. М.[1], Алыгулиев Р. М.[2], Имамвердиев Я. Н.[3], Сухостат Л. В.[4]

[1]Академик Национальной Академии Наук Азербайджана, д-р техн. наук, проф., директор Института Информационных Технологий Национальной Академии Наук Азербайджана, Баку, Азербайджан

[2]Член-корреспондент Национальной Академии Наук Азербайджана, д-р техн. наук, зав. отделом, Институт Информационных Технологий Национальной Академии Наук Азербайджана, Баку, Азербайджан

[3]Канд. техн. наук, доцент, зав. отделом, Институт информационных технологий Национальной Академии Наук Азербайджана, Баку, Азербайджан

[4]Канд. техн. наук, старший научный сотрудник, Институт информационных технологий Национальной Академии Наук Азербайджана, Баку, Азербайджан

УЛУЧШЕННЫЙ ПОДХОД ОБНАРУЖЕНИЯ DOS АТАК С ПРИМЕНЕНИЕМ АНСАМБЛЯ

Актуальность. Решена задача применения ансамбля классификаторов к обнаружению DoS атак в больших массивах данных сетевого трафика, чтобы противостоять атакам в сети.

Цель работы заключается в построении ансамбля классификаторов, превосходящего единичные классификаторы с точки зрения точности.

Метод. Для достижения поставленной цели в работе предложен алгоритм, указывающий на вероятности принадлежности к определенным классам, возвращающий вектор оценок классификаторов для каждой точки. Особенность предложенного подхода состоит в том, что для каждой точки из набора данных предсказанная метка класса соответствует максимальному значению среди всех оценок, полученных методами классификации для данной точки. В качестве классификаторов были рассмотрены деревья решений, алгоритм *k*-ближайших соседей, машины опорных векторов с различными ядерными функциями и наивный байесовский классификатор. Сравнительный анализ предложенного подхода с рассмотренными единичными классификаторами проводится по следующим метрикам: точность, полнота, «меткость» и *F*-мера.

Результаты. Эксперименты были проведены на языке R 3.4.1 на наборе данных сетевых атак NSL-KDD, который был разбит на три класса (DoS, «нормальное» поведение сети и другие типы атак).

Выводы. Проведенные эксперименты подтвердили работоспособность предложенного подхода. Наиболее точный результат показал ансамбль из пяти классификаторов. Разработка техники обнаружения атак, основанной на применении ансамбля классификаторов, позволяет избежать проблем, характерных для большинства подходов, поскольку он способен с высокой точностью обнаружить как известные, так и новые атаки. Можно сделать вывод о практической значимости предложенного подхода к обнаружению атак в сети. В целях дальнейшего изучения обнаружения атак в сетевом трафике будут проведены исследования на реальных наборах данных большой размерности.

Ключевые слова: информационная безопасность, сетевые атаки, DoS, классификация, ансамбль классификаторов, Big data.

REFERENCES

1. Aliguliyev R. M., Imamverdiyev Y. N., Hajirahimova M. S. Multidisciplinary problems of big data in information security, *Proceedings of the II International scientific and practical conference Information Security and Computer Technologies*, 2017, pp. 10–11.
2. Nallaivarothayan H. Ryan D., Denman S., S. Sridharan, C. Fookes An evaluation of different features and learning models for anomalous event detection, *Proceedings of the International Conference on Digital Image Computing: Techniques and Applications*, 2013, pp. 1–8. DOI: 10.1109/dicta.2013.6691480
3. Xie M., Han S., Tian B., Parvin S. Anomaly detection in wireless sensor networks: a survey, *Journal of Network and Computer Applications*, 2011, Vol. 34, pp. 1302–1325. DOI: 10.1016/j.jnca.2011.03.004
4. Davis J. J., Clark A. J. Data preprocessing for anomaly based network intrusion detection: a review, *Computers & Security*, 2011, Vol. 30, pp. 353–375. DOI: 10.1016/j.cose.2011.05.008
5. Fiorea U., Palmierib F., Castiglionec A., Santis A. D. Network anomaly detection with the restricted boltzmann machine, *Neurocomputing*, 2013, Vol. 122, pp. 13–23. DOI: 10.1016/j.neucom.2012.11.050
6. Chandola V., Banerjee A., Kumar V. Anomaly detection: a survey, *ACM Computing Surveys*, 2009, Vol. 41, No. 3, pp. 1–58. DOI: 10.1145/1541880.1541882
7. Anceaume E., Busnel Y., Merrer E. L., Ludinard R., Marchand J., Sericola B. Anomaly characterization in large scale networks, *Proceedings of the 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, 2014, pp. 68–79. DOI: 10.1109/dsn.2014.23
8. Dua S., Du X. Data mining and machine learning in cybersecurity. Boca Raton, FL, CRC Press, 2011, 256 p. DOI: 10.1201/b10867
9. Catania C. A., Garino C. G. Automatic network intrusion detection: current techniques and open issues, *Computers and Electrical Engineering*, 2012, Vol. 38, No. 5, pp. 1062–1072. DOI: 10.1016/j.compeleceng.2012.05.013
10. Ahmed M., Mahmood A., Hu J. A survey of network anomaly detection techniques, *Journal of Network and Computer Applications*, 2016, Vol. 60, pp. 19–31. DOI: 10.1016/j.jnca.2015.11.016
11. Wu S. X., Banzhaf W. The use of computational intelligence in intrusion detection systems: a review, *Applied Soft Computing*, 2010, Vol. 10, No. 1, pp. 1–35. DOI: 10.1016/j.asoc.2009.06.019
12. Chandola V., Eilertson E., Ertoz L., Simon G., Kumar V. Data mining for cyber security. New York: Springer, 2006, 159 p. (Data Warehousing and Data Mining Techniques for Computer Security.). DOI: 10.1007/978-0-387-47653-7
13. Lee W., Stolfo S. J. A framework for constructing features and models for intrusion detection systems, *ACM Transactions on Information and System Security*, 2000, Vol. 3, No. 4, pp. 227–261. DOI: 10.1145/382912.382914
14. Mahoney M. V., Chan P. K. Learning nonstationary models of normal network traffic for detecting novel attacks, *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2002, pp. 376–386. DOI: 10.1145/775047.775102
15. Hodge V., Austin J. A survey of outlier detection methodologies, *Artificial Intelligence Review*, 2004, Vol. 22, No. 2, pp. 85–126. DOI: 10.1023/b:aire.0000045502.10941.a9
16. Farid D. M., Rahman M. Z., Rahman C. M. Adaptive intrusion detection based on boosting and naive bayesian classifier, *International Journal of Computer Applications*, 2011, Vol. 24, No. 3, pp. 12–19. DOI: 10.5120/2932-3883
17. Laurentys C. A., Palhares R. M., Caminhas W. M. A novel artificial immune system for fault behavior detection, *Expert Systems with Applications*, 2011, Vol. 38, pp. 6957–6966. DOI: 10.1016/j.eswa.2010.12.019
18. Meza J., Campbell S., Bailey D. Mathematical and statistical opportunities in cybersecurity, *Paper LBNL-1667E, Lawrence Berkeley National Laboratory, Berkeley, CA*, 2009, pp. 1–11. DOI: 10.2172/950976
19. Lin L. Zuo R., Yang S., Zhang Z. SVM ensemble for anomaly detection based on rotation forest, *Proceedings of the 3rd International Conference on Intelligent Control and Information Processing*, 2012, pp. 150–153. DOI: 10.1109/icicip.2012.6391455
20. Aburomman A., Reaz M. B. I. A novel SVM-kNN-PSO ensemble method for intrusion detection system, *Applied Soft Computing*, 2016, Vol. 38, pp. 360–372. DOI: 10.1016/j.asoc.2015.10.011
21. Masarat S., Taheri H., Sharifian S. A novel framework, based on fuzzy ensemble of classifiers for intrusion detection systems, *Proceedings of the 4th International Conference on Computer and Knowledge Engineering*, 2014, pp. 165–170. DOI: 10.1109/iccke.2014.6993345
22. Meng Y., Kwok L. F., Li W. Enhancing false alarm reduction using voted ensemble selection in intrusion detection, *International Journal of Computational Intelligence Systems*, 2013, Vol. 6, No. 4, pp. 626-638. DOI: 10.1080/18756891.2013.802114
23. Aburomman A. A., Reaz M. B. I. A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems, *Information Sciences,* 2017, Vol. 414, pp. 225–246. DOI: 10.1016/j.ins.2017.06.007
24. Kumar P. A. R., Selvakumar S. Distributed denial of service attack detection using an ensemble of neural classifier, *Computer Communications*, 2011, Vol. 34, No. 11, pp. 1328–1341. DOI: 10.1016/j.comcom.2011.01.012
25. Nguyen N. T., Zgrzywa A., Czyzewski A. Advances in multimedia and network information system technologies. Berlin, Springer-Verlag, 2010, 318 p. DOI: 10.1007/978-3-642-14989-4
26. Zhang M.-L., Zhou Z.-H. A k-nearest neighbor based algorithm for multi-label classification, *Proceedings of IEEE International Conference on Granular Computing,* 2005, pp. 718–721. DOI: 10.1109/grc.2005.1547385
27. Aggarwal P., Sharma S. K. Analysis of KDD dataset attributes-class wise for intrusion detection, *Procedia Computer Science*, 2015, Vol. 57, pp. 842–851. DOI: 10.1016/j.procs.2015.07.490
28. McHugh J. Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory, *ACM Transactions on Information and System Security*, 2000, Vol. 3, No. 4, pp. 262–294. DOI: 10.1145/382912.382923
29. Aliguliyev R. M. Performance evaluation of density-based clustering methods, *Information Sciences*, 2009, Vol. 179, pp. 3583–3602. DOI: 10.1016/j.ins.2009.06.012

UDC 004.33(035)

Akhmetov B. B.[1], Lakhno V. A.[2], Malyukov V. P.[3]

[1]PhD, Associate Professor, Rector of the Caspian State University of Technologies and Engineering named after Sh. Yessenov, Aktau, Kazakhstan
[2]Dr. Sc., Professor, Head of Department of cybersecurity, European University, Kyiv, Ukraine
[3]Dr. Sc., Associate Professor, Professor, Department of Informatics Systems, Mathematics Disciplines, European University, Kyiv, Ukraine

MODEL OF INVESTMENT STRATEGIES IN CYBER SECURITY SYSTEMS OF TRANSPORT SITUATIONAL CENTERS

Actuality. The actual task of finding the optimal strategy for control the procedure of mutual financial investments to the situation center for cyber security on transport.

The aim of the work – the development of a model for a decision support system on the continuous mutual investment in a cyber security situational center, which differs from the existing ones by solving a bilinear differential quality game with several terminal surfaces.

Method. In order to achieve the goal there was used a discrete-approximation method for solving a bilinear differential quality game with dependent motions. Application of this method in the developed decision support system, unlike existing ones, gives concrete recommendations ft choosing control decisions in the investment process.

Results – the proposed model gives concrete recommendations at choosing strategies in the investment process at the creation of a protected situational center. In the course of the computational experiment, there was considered a new class of bilinear differential games that allowed adequately to describe the process of investing in cyber security means of situational transport centers in Kazakhstan and Ukraine.

Conclusions. For the first time, there was proposed a model describing the process of mutual investment based on the solution of bilinear equations and a differential quality game with several terminal surfaces. Considered the peculiarity of the differential game on the example of mutual investment in the means of cyber security of the situational transport center. In this case, the right-hand side of the system of differential equations is represented in the form of bilinear functions with arbitrary coefficients. The model allows to predict the results of investment and to find strategies for managing the investment process in the protection and cyber security systems of the situational transport center.

Keywords: decision support system, model, bilinear functions, mutual investment, cyber security, situational transport center.

NOMENCLATURE

CICS – critically important computer systems;
CS – cyber security;
ISS – information security system;
DSS – decision support system;
SC – situational center;
a_1 – coefficient determining the interest payment for the financial resource of the second investor (2nd) to the first investor (1st);
a_2 – coefficient determining the interest payment for the financial resource of the first investor (1nd) to the second investor (1st);
g^* – coefficient determining the beam balance;
k_d – the dollar exchange rate in relation to the national currency (hryvnia, tenge, etc.);
M_0 – terminal set, which determines the end of investment by the 2nd investor;
N_0 – terminal set, which determines the end of investment by the 1st investor;
r_1 – coefficient determining the share of return of the financial resource of the 2nd investor to the 1st one;
r_2 – coefficient determining the share of return of the financial resource of the 1st investor to the 2st investor;
R_+^2 – positive ortant;
t – time parameter;
u – control parameter of the 1st investor;
u_* – optimal strategy of the 1st investor;
ν – control parameter of the 2nd investor;
W_1 – a set of preferences for the 1st investor;
W_2 – a set of preferences for the 2nd investor;
z_1 – the value of the financial resource of the 1st investor;
z_2 – the value of the financial resource of the 2nd investor;
α_1 – the growth rate of the financial resource of the 1st investor;
α_2 – the growth rate of the financial resource of the 2nd investor
β_1 – coefficient determining the share of repayment of the debt of the 1st investor to the 2nd investor;
β_2 – coefficient determining the share of repayment of the debt of the 2nd investor to the 1st investor.

INTRODUCTION

Nowadays, many projects in the field of transport control are developing in the direction of creating large situational cents (SC) that provide solutions to specific tasks, in particular, the protection of critically important computer systems (CICS) of the transport industry. Investing in innovative projects, for example, in the field of cyber security (CS) and information security, is characterized by a high degree of uncertainty and riskiness [1, 2]. In work [1] there is noted that many investors in the CS sector face to chaotic and uncertainty choosing an effective investment strategy and method of risk control. Many enterprises and companies engaged in servicing CICS, in particular, on transport, spending more on information security systems (ISS) and CS, do not feel confident that the chosen investment strategy

© Akhmetov B. B., Lakhno V. A., Malyukov V. P., 2018
DOI 10.15588/1607-3274-2018-2-9

makes the infrastructure of information and communication systems really safe [2].

The last one, in particular, is connected with the lack of practically applicable models, taking into account the ways of finding different variants of strategies in mutual financial investment of CS objects [3, 4]. Similar problems arose at choosing mutual investment strategies in the SC for Kazakhstan, Ukraine, and other countries actively engaged in the development of ISS for CICS. In this regard, it is urgent to develop new models for intellectualized computer decision support systems (DSS) that adequately describe and evaluate investment strategies in the CS CICS.

The object of the research is the process of mutual financial investment in the situational center for transport cyber security.

The aim of this work was to develop a model for decision support systems for continuous mutual investment in a cyber security situational center, which differs from existing ones by solving a bilinear differential quality game with several terminal surfaces.

1 PROBLEM STATEMENT

There is considered the model for the decision support systems on continuous mutual investment in means of the CS of the SC of transport. The model is based on solving a bilinear differential quality game with two terminal surfaces. Mathematical statement: there are two players that control a dynamic system, given by a system of bilinear differential equations with dependent motions. The sets of strategies of players U, V are determined, respectively. Two terminal surfaces are defined. The aim of the first player is to put the dynamic system using its control strategies on the terminal surface M_0, despite the actions of the second player. The aim of the second player is to put the dynamic system through its control strategies to the terminal surface, no matter how the first player acts. The solution is to find a set of initial states of objects and their strategies that allow objects to put the system to that or another surface. The solution of such systems of differential equations can not be found using the Cauchy formula. For the right-hand sides of differential equations with arbitrary coefficients of results on bilinear differential games, particulaliy such differential equations are used in our model, it was not obtained.

2 LITERATURE REVIEW

The problem of effective financial investment in the CS systems and the protection of CICS is one of the most important in the security sphere of almost all states [1, 2]. A large number of researches have been devoted to this subject [3, 4]. However, it should be noted that the authors [4–6] did not offer any real recommendations on the development of mutual financial investment strategies in the CS CICS [4–6]. A separate direction of the researches in this area was the work devoted to the application of various expert systems [7–9] and DSS [10, 11] for the selection of rational investment strategies in the ISS and CS. The approaches developed by the authors [12, 13] and [7, 8, 14] at solving this problem do not allow to find effective recommendations and investment control strategies in the CS. This circumstance necessitates the development of new models for DSS, which would provide an opportunity to find optimal strategies for mutual financial investment in the means of protecting the CICS. This can be achieved, in particular, by applying methods of the differential quality games theory with several terminal surfaces [11, 15]. It should be noted that for such differential games the approaches outlined in works [1, 4, 5] are not applicable, because within the framework of decisions the enemy player can apply measurable controls, and in the work under consideration there are allowed any controls including non-measurable functions.

As the analysis of the latest researches in this area showed the problem of further development of DSS models in the tasks of continuous mutual investment in the ISS and CS CICS, in particular, the situational center on transport, remains relevant.

3 MATERIALS AND METHODS

An investor from a country, where the US dollar is in cash, having free capital, is trying to choose the most preferable options for his placement. In order to do this, he chooses a counterparty, i.e. an object for investing his own funds, for example, in Kazakhstan, where the tenge is in cash. The object of investment are ISS and CS of SC on transport, as one of the components of the critical infrastructure of the state. The investor wants to increase his capital, the counterparty – to improve their financial and economic level.

If such an objective for an investor arises regularly, it is advisable to use DSS in the decision-making process related to investing. The formalization of the investment process is given under the assumption that the investor is an investment company, a counterparty is a financial structure in another country. We call this structure a financial corporation.

An investment company, having some free resources (its investment capital), increases them by α_1 times. Then it is necessary to decide what part of these resources it will invest in active operations. These operations are the allocation of resources in the investment projects of the corporation in the sphere of CS of the SC and in the repayment of the debt that the company has at that time. We will consider that the corporation does the same with respect to this investment company. The interaction between the company and the corporation will be described under the following assumptions: 1) the investment company controls financial resources z_1 estimated in dollars (USD); 2) the corporation controls financial resources z_2 estimated in national currency (for example, tenge (Kazakhstan) or hryvnia (Ukraine), 3) during the interaction, the ratio of the dollar to the national currency (dollar exchange rate) remains constant.

After that, the investment company and the counterparty identified the share of resources allocated for mutual active operations, their interaction is determined through the following system of differential equations:

$$dz_1(t)/dt = -z_1(t) + \alpha_1(t)\cdot z_1(t) + \left[\left(1-\beta_1(t)\right)\cdot\left(a_1(t)+r_1(t)\right)-1\right]\times$$
$$\times u(t)\cdot\alpha_1(t)\cdot z_1(t) + \left[1-\left(a_2(t)+r_2(t)\right)\cdot\left(1-\beta_2(t)\right)\right]\times$$
$$\times v(t)\cdot\alpha_2(t)\cdot z_2(t)/k_d; \quad (1)$$

$$dz_2(t)/dt = -z_2(y) + \alpha_2(t)\cdot z_2(t) + [(1-\beta_2(t))\cdot(a_2(t)+r_2(t))-1]\times \\ \times v(t)\cdot\alpha_2(t)\cdot z_2(t) + [1-(a_1(t)+r_1(t))\cdot(1-\beta_1(t))]\times \\ \times u(t)\cdot\alpha_1(t)\cdot z_1(t)/k_d. \quad (2)$$

Therefore, at time t the company's value $dz_1(t)/dt$ (in dollars – \$) will be equal to the sum of the following values: $\alpha_1(t)\cdot z_1(t)$; percents $a_1(t)\cdot(1-\beta_1(t))\cdot u(t)\cdot\alpha_1(t)\cdot z_1(t)$ for invested financial resources $(1-\beta_1(t))\cdot u(t)\cdot\alpha_1(t)\cdot z_1(t)$ of the company; $r_1(t)\cdot(1-\beta_1(t))\cdot u(t)\cdot\alpha_1(t)\cdot z_1(t)$ – share of the "returned" investment resource $(1-\beta_1(t))\cdot u(t)\cdot\alpha_1(t)\cdot z_1(t)$ of the company; "unreturned" asset (investment) – $\{(1-r_2(t))\cdot(1-\beta_2(t))/k_d\}\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$ of the corporation (in \$); $\{\beta_2(t)/k_d\}\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$ – the repayment of the corporation debt to the company, from which (out of the amount) there is deducted: financial resource $z_1(t)$; the value $u(t)\cdot\beta_1(t)\cdot\alpha_1(t)\cdot z_1(t)$, allocated to repay the debt that the company had at time t to the corporation, as well as the amount $u(t)\cdot(1-\beta_1(t))\cdot\alpha_1(t)\cdot z_1(t)$, allocated for investment at time t; $\{a_2(t)\cdot(1-\beta_2(t))/k_d\}\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$ – interest payment for investment resources $\{(1-\beta_2(t))/k_d\}\cdot v(t)\cdot\alpha_2\cdot z_2(t)$ of the corporation.

The value $dz_2(t)/dt$ (national currency, tenge, hryvnia, etc.) at time t will be equal to the sum of the following terms of the quantities: $\alpha_2(t)\cdot z_2(t)$; percents $a_2(t)\cdot(1-\beta_2(t))\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$ for invested financial resources $(1-\beta_2(t))\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$ of the corporation; $r_2(t)\cdot(1-\beta_2(t))\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$ – share of the "returned" investment resource of the company $(1-\beta_2(t))\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$; "unreturned" asset – $(1-r_1(t))\cdot(1-\beta_1(t))\cdot u(t)\cdot k_d\cdot\alpha_1(t)\cdot z_1(t)$ of the corporation; the values $u(t)\cdot\beta_1(t)\cdot k_d\cdot\alpha_1(t)\cdot z_1(t)$, characterized the repayment of the company debt from which there is deducted: financial resource $z_2(t)$; the sum $v(t)\cdot\beta_2(t)\cdot\alpha_2(t)\cdot z_2(t)$, allocated to repay the debt of the corporation to the company at time t; $(1-\beta_2(t))\cdot v(t)\cdot\alpha_2(t)\cdot z_2(t)$ – the corporation's funds for conducting active operations (investments); $a_1(t)\cdot(1-\beta_1(t))\cdot k_d\cdot u(t)\cdot\alpha_1(t)\cdot z_1(t)$ – interest payment for investment resources $(1-\beta_1(t))\cdot u(t)\cdot\alpha_1(t)\cdot z_1(t)$ of the company.

Interaction ends when the conditions are fulfilled:

$$(z_1(t), z_2(t)) \in M_0, \quad (3)$$

$$(z_1(t), z_2(t)) \in N_0. \quad (4)$$

It is assumed that

$$M_0 = \{(z_1, z_2): (z_1, z_2) \in R_+^2,\ z_1 > 0,\ z_2 = 0\}, \quad (5)$$

$$N_0 = \{(z_1, z_2): (z_1, z_2) \in R_+^2,\ z_1 = 0,\ z_2 > 0\}. \quad (6)$$

The following options are considered: 1) loss of investment resources of the corporation, but the investment company multiplied its capital (for example, with a decrease of the national currency exchange rate); 2) loss of the capital of the investment company, but the corporation multiplied its capital (for example, with improper selection of hardware and software ISS and CS). If the conditions (3) and (4) are not fulfilled, the investor interaction continues.

Further we accept: the investment company – the player (I); the corporation – the player (II). Interaction of players is considered within the framework of a positional differential game with complete information [8, 11, 13]. Within the framework of this scheme, the interaction "generates" two tasks – from the point of view of the first confederate player and from the point of view of the second confederate player. Because of symmetry, it suffices to consider one of them. For example, from the point of view of the first confederate player. For this, we will define the pure strategies of the first confederate player. Denote by $T = [0, M]$ – the time segment, i.e. a set characterizing the area of the time parameter; M – a positive real number.

Definition. The pure strategy of the first confederate player is the function $u: T\cdot[0,1]\cdot[0,1] \to [0,1]$, which designate the state of information (position) $(t, (z_1(0), z_2(0)))$ as the value $u(t, (z_1(0), z_2(0))): 0 \le u(t, (z_1(0), z_2(0))) \le 1$.

The pure strategy of the first confederate player is the function (rule) which designate the state of information at the moment t the value $u(t, (z_1(0), z_2(0)))$, that determines the amount of the resource (capital) value of the first player that he allocated to "invest" the second player, in our case ISS and CS SC on transport. With regard to the knowledge of the enemy player (within the framework of the positional game scheme), there are made no assumptions, which is equivalent to the fact that the enemy player chooses his control effect $u(t)$ based on any information. After defining the strategies in task 1 it is necessary to determine the set of preference for the first player. Considering the fact that for the description of the proposed approach it is enough to confine with a qualitative description, the set of preferences W_1 of the first player. We assume that W_1 – a set of such initial resources $(z_1(0), z_2(0))$ of players that have the property: for such initial states there is a strategy of the first player, which, for any realizations of the strategy of the second player, "leads", at one of the moments of time t, the state of the system $(z_1(t), z_2(t))$ when will be fulfilled the condition (3). Moreover, the second player does not have a strategy that can "lead" to the fulfillment of condition (4), at one of the preceding moments of time t. The strategy of the

first player with this property is called optimal. Suppose that for any moment of time t the following conditions are fulfilled: $\alpha_1(t)=\alpha_1$; $\alpha_2(t)=\alpha_2$; $\beta_1(t)=\beta_1$; $\beta_2(t)=\beta_2$; $r_1(t)=r_1$; $r_2(t)=r_2$. Accept: $q_1=(1-\beta_1)\cdot(a_1+r_1)-1$; $q_2=(1-\beta_2)\cdot(a_2+r_2)-1$. The following cases are possible: a) $q_1\geq 0; q_2\geq 0$; b) $q_1\prec 0; q_2\prec 0$; c) $q_1\succ 0; q_2\leq 0$; d) $q_1\leq 0; q_2\succ 0$. Different ratios of other interaction parameters are possible, for example, α_1, α_2. We give the solution of the game, i.e. the sets of preferences and optimal strategies for the first player.

For the case a) we will obtain:

$$W_1=\left\{(z_1(0),z_2(0)):(z_1,z_2)\in \operatorname{int} R_+^2,\ z_2(0)<w^*\cdot z_1(0)\right\} \quad (7)$$

for

$$w^*=\left\{-[q_2\cdot\alpha_2+\alpha_2-q_1\cdot\alpha_1-\alpha_1]/[2q_2\cdot\alpha_2]+\right.$$
$$\left.2\sqrt{\{[q_2\cdot\alpha_2+\alpha_2-q_1\cdot\alpha_1-\alpha_1]/2q_2\cdot\alpha_2\}^2+(q_1\cdot\alpha_1)/(q_2\cdot\alpha_2)}\right\}; \quad (8)$$

$u_*(z_1,z_2)=\left\{1,\ z_2<w\cdot z_1,(z_1,z_2)\in \operatorname{int} R_+^2\right\}$ and is not defined otherwise}.

In cases b) and c) a set W_1 is empty.

In case c) and $\alpha_2>\alpha_1+q_1\cdot\alpha_1$ we will obtain

$$W_1=\left\{(z_1(0),z_2(0)):(z_1(0),z_2(0))\in \operatorname{int} R_+^2,\ z_2(0)<\delta\cdot z_1(0)\right\} \quad (9)$$

for $\delta=(q_1\cdot\alpha_1)/(\alpha_2-q_1\cdot\alpha_1-\alpha_1)$;

$u_*(z_1,z_2)=\left\{1,\ z_2<\delta\cdot z_1,(z_1,z_2)\in \operatorname{int} R_+^2\right\}$, and is not defined otherwise}.

In case c) and $\alpha_2\leq\alpha_1+q_1\cdot\alpha_1$ we obtain:

$$W_1=\operatorname{int} R_+^2,\ u_*(z_1,z_2)=\left\{1,(z_1,z_2)\in \operatorname{int} R_+^2\right\}, \quad (10)$$

and is not defined otherwise.

The task 2 is solved symmetrically (from the point of view of the second confederate player).

4 EXPERIMENTS

A computational experiment was performed in the PTC Mathcad 4 environment. The model was also implemented as a separate software module on the Java language, 4. The module is available for cross-platform use as an independent software product, or as a plug-in to the DSS "SSDMI" [15].

The axis z_1 – corresponds to the currency "tenge" (or another national currency, for example, hryvnia). The axis z_2 – corresponds to the currency "dollar". The areas of preference from the point of view of the second player are "adjacent" to the areas of preference of the first player. These areas are divided among themselves by the balance beams. The beams have the following property: if a pair of states ($z_1(0)$, $z_2(0)$) belongs to the balance beam, then the players have strategies that allow them to be on the balance beam for all subsequent moments of time t. In the course of the computational experiment, in the space of variables (z_1,z_2) there are balance beams. Therefore, if the interaction starts from these states, then players have strategies that allow them to stay on the balance beam. Consequently, for given values $(z_1(0),z_2(0))$, we can find a relation to the parameters interaction, at which the pair $(z_1(t),z_2(t))$ will be on the balance beam. If the initial states (resources) are not on the interaction balance beam, then during the computational experiment you can try to change the interaction parameters. It is required for the initial resources to be on a balance beam. The last will allow for players to continue the interaction as long as they want.

5 RESULTS

Fig. 1–3 shows the results for the 3 test calculations of the investment estimation in the CS system of the Situational Center for Railway Transport of the Republic of Kazakhstan.

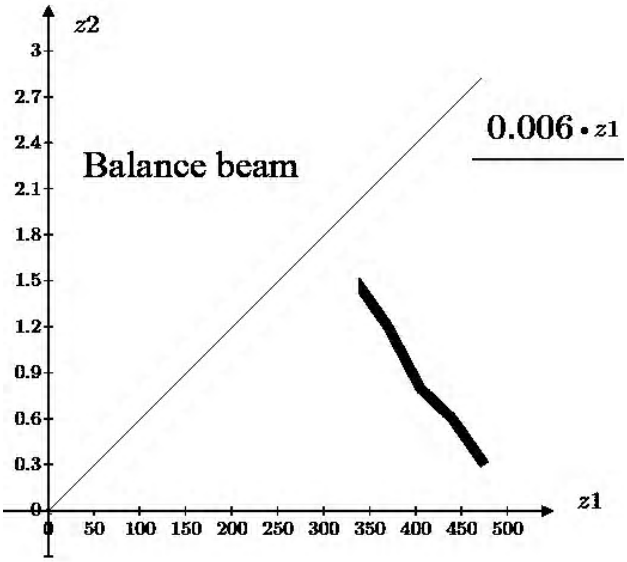


Figure 1 – Computational experiment results 1

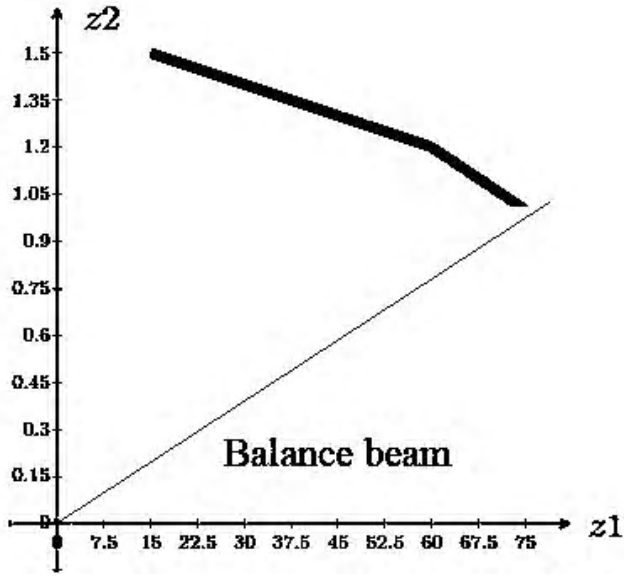


Figure 2 – Computational experiment results 2

The first calculation for the case when the start point $(z_1(0), z_2(0))$ belongs to the preference area of the 1st player ("Kazakhstan", the variable x denotes the amount of tenge, the variable y denotes the number of dollars, the exchange rate is 1 $ – 335 "tenge"). The second calculation is for the case when the start point $(z_1(0), z_2(0))$ belongs to the preference area of the 2nd players (the counterpart of "Kazakhstan"). The third calculation is for the case when the start point $(z_1(0), z_2(0))$ belongs to the "balance beam", i.e. to the beam which shares the areas of the players' preference. The results of the computational experiment are presented in Table 1.

Fig. 4 presents a form for interpretation the results of decision-making support received using the DSS "SSDMI". The balance beam is shown with markers for the choice of strategies for mutual investment in the means of the CS SC of the railway transport of the Republic of Kazakhstan. Calculations: $1-g^* = 0.006$; $2-g^* = 0.012$; $3-g^* = 0.018$.

The obtained results demonstrate the effectiveness of the proposed approach. During the testing of the model in the PTC Mathcad 4 environment, as well as in the DSS "SSDMI", there was established the correctness of the results.

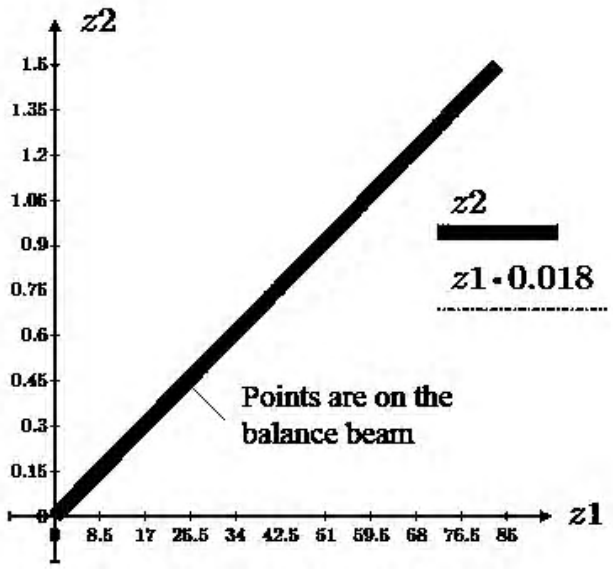


Figure 3 – Computational experiment results 3

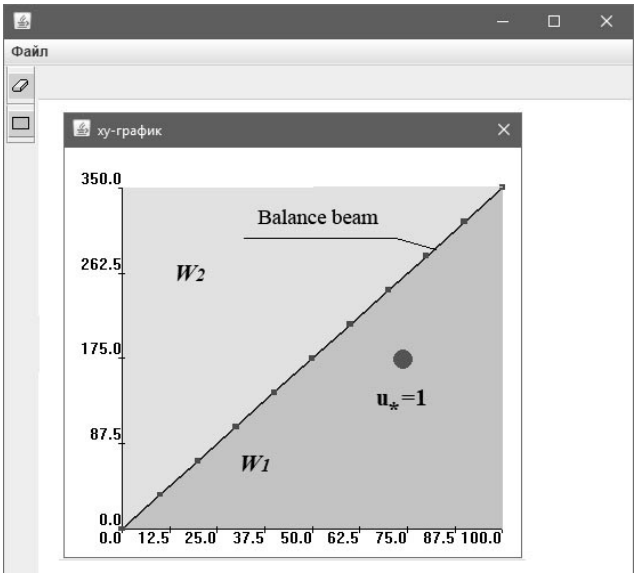


Figure 4 – General view of the DSS module for the selection of strategies for mutual investment the means of CS SC

Approbation of DSS "SSDMI" was also carried out for real investment projects in the field of cyber security of Ukraine and Kazakhstan [14–16].

6 DISCUSSION

The prediction data obtained during the computational experiment and practical approbation of the DSS "SSDMI" module did not always coincide with the actual data [14, 15]. This is a certain disadvantage of the approach outlined in the article. We can only try to reduce the discrepancies, in particular, through the intellectual analysis of data in DSS, and also improving the toolkit for prediction the investment processes in the CS means. It is also necessary to take into account the specifics of the investment objects, because even at the similarity of procedures there are specific features at investing in the CS means, for example, the situational center of aviation and railway transport. If this toolkit is reasonably replenished with the above-mentioned IT, then an attempt will be made more efficient for the investment process. This approach allows to do this, because it allows to select the parameters variables for the balance. The comparison of the results of the computational experiment and the practical data of investment project evaluation is compared with the results described in [5, 9, 13, 14]. In the

Table 1 – The results of the computational experiment on the choice of the mutual investment strategy in the means of the CS SC of transport (on the example of the Republic of Kazakhstan)

Calculations	Modeling results
1	(Z_1 (0), Z_2 (0)) = (335, 1.5); (Z_1 (1), Z_2 (1)) = (370, 1.2); (Z_1 (2), Z_2 (2)) = (405, 0.8); (Z_1 (3), Z_2 (3)) = (440, 0.6); (Z_1 (4), Z_2 (4)) = (475, 0.3)
2	(Z_1 (0), Z_2 (0)) = (75, 1.0); (Z_1 (1), Z_2 (1)) = (60, 1.2); (Z_1 (2), Z_2 (2)) = (45, 1.3); (Z_1 (3), Z_2 (3)) = (30, 1.4); (Z_1 (4), Z_2 (4)) = (15, 1.5)
3	(Z_1 (0), Z_2 (0)) = (83.33, 1.5); (Z_1 (1), Z_2 (1)) = (66.67,1.2); (Z_1 (2), Z_2 (2)) = (55.56, 1.0); (Z_1 (3), Z_2 (3)) = (44.44, 0.8); (Z_1 (4), Z_2 (4)) = (27.77, 0.5)

course of the comparative analysis it was found that the proposed toolkit allows the participants of the investment process significantly to improve the effectiveness and predictability of their activities, on average by 12–16% [15].

APPRECIATION

The work was carried out within the framework of the grant competition for scientific and and technical projects for 2018–2020 of the Republic of Kazakhstan, registration number AP05132723 “Development of adaptive expert systems in the field of cyber security of critically important information objects”.

CONCLUSIONS

There was proposed a model for the decision support system module for mutual investment in cyber security systems of a situational transport center. Unlike the existing solutions, the proposed model gives concrete recommendations at choosing strategies in the investment process of the creation of a protected situational center. If the forecast is unsatisfactory, it is possible to flexibly adjust the parameters of the investment process in the CS systems in order to achieve an acceptable financial result by the parties.

The scientific novelty of the results obtained in the article is that there was firstly considered a new class of bilinear differential games that allowed to describe adequately the process of investing in the means of the CS of situational transport centers in Kazakhstan and Ukraine. A distinctive feature of this approach, in particular, tested during the computational experiment, is the use of a solution based on a bilinear differential quality game with several terminal surfaces. The peculiarity of this differential game is that the right-hand side of the differential equations system is a bilinear function with arbitrary coefficients.

The practical significance of the obtained results is that a cross-platform software product was developed in the form of an independent module for the decision support system – SSDMI. The “SSDMI” module implements the proposed model, based on the application of methods of the differential games theory. The SSDMI module allows to reduce the discrepancies between the forecast data and the real return on investment in the funds of the CS of CICS.

Further perspectives of the development of this research is the transfer of the accumulated experience at the development of the DSS on the Android platform. This will increase the mobility of decision-making on mutual investment.

REFERENCES

1. Externalities and the magnitude of cyber security underinvestment by private sector firms: a modification of the Gordon-Loeb model / [L. A. Gordon, M. P. Loeb, W. Lucyshyn et al.] // Journal of Information Security. – 2015. – Vol. 6, № 1. – P. 24–30. DOI: 10.4236/jis.2015.61003
2. Decision support approaches for cyber security investment / [A. Fielder, E. Panaousis, P. Malacaria et al.] // Decision Support Systems. – 2016. – Vol. 86. – P. 13–23. DOI: org/10.1016/j.dss.2016.02.012
3. Meland P. H. Mitigating risk with cyberinsurance / P. H. Meland, I. A. Tondel, B. Solhaug // IEEE Security & Privacy. – 2015. – Vol. 13, № 6. – P. 38 – 43. DOI: 10.1109/MSP.2015.137
4. Smeraldi F. How to spend it: optimal investment for cyber security / F. Smeraldi, P. Malacaria // 1st International Workshop on Agents and CyberSecurity, Paris, France, 06–08 May 2014 : Proceedings. – New York : ACM, 2014. – P. 8. DOI: 10.1145/2602945.2602952
5. Cyber-investment and cyber-information exchange decision modeling / [D. K. Tosh, M. Molloy, S. Sengupta et al.] // High Performance Computing and Communications IEEE 7th International Symposium on Cyberspace Safety and Security (CSS), New York, 24–26 August 2015 : Proceedings. – New York : IEEE, 2015. – P. 1219–1224.
6. Alpcan T. Modeling dependencies in security risk management / T. Alpcan, N. Bambos // Risks and Security of Internet and Systems Fourth International Conference, Toulouse, France, 19–22 October 2009 : Proceedings. – France : IEEE, 2009. – P. 113–116. DOI : 10.1109/CRISIS.2009.5411969
7. Peppard J. The Strategic Management of Information Systems: Building a Digital Strategy / J. Peppard, J. Ward. – U.K. : Wiley, 2016. – 504 p.
8. Malyukov V.P. A differential game of quality for two groups of objects / V.P. Malyukov // Journal of Applied Mathematics and Mechanics. – 1991. – Vol. 55, № 5. – P. 596–606.
9. Cassidy A. A practical guide to information systems strategic planning / A. Cassidy. – New York : CRC press, 2016. – 395 p.
10. Bonczek R. H. Foundations of decision support systems / R. H. Bonczek, C. W. Holsapple, A. B. Whinston. – USA : Academic Press, 2014. – 412 p.
11. Malyukov V.P. Discrete-approximation method for solving a bilinear differential game / V.P. Malyukov // Cybernetics and Systems Analysis. – 1993. – Vol. 29, № 6. – P. 879–888.
12. Assessing survivability to support power grid investment decisions / [A. Koziolek, A. Avritzer, S. Suresh et al.] // Reliability Engineering & System Safety. – 2016. – Vol. 155. – P. 30–43. DOI: 10.1016/j.ress.2016.05.015
13. Lakhno V. Development of a Support System for Managing the Cyber Security of Information and Communication Environment of Transport / V. Lakhno, Al. Petrov, Ant. Petrov // Information Systems Architecture and Technology : 38th International Conference on Information Systems Architecture and Technology (ISAT 2017), Wroclaw, 17–19 September 2017 : proceedings. – Wroclaw : Springer, 2017. – P. 113–127. DOI: 10.1007/978-3-319-67229-8_11
14. Lakhno V. A. Development of a support system for managing the cyber security / V. A. Lakhno // Radio Electronics, Computer Science, Control. – 2017. – № 2. – P. 109–116. DOI: 10.15588/1607-3274-2017-2-12
15. Development of the decision making support system to control a procedure of financial investment / [V. Lakhno, V. Malyukov, N. Gerasymchuk et al.] // Eastern-European Journal of Enterprise Technologies. – 2017. – Vol. 6, № 3. – P. 34–41. DOI: 10.15587/1729-4061.2017.119259
16. Designing a decision support system for the weakly formalized problems in the provision of cybersecurity / [B. Akhmetov, V. Lakhno, Y. Boiko et al.] // Eastern-European Journal of Enterprise Technologies. – 2017. – Vol. 1, № 2. – P. 4–15. DOI: 10.15587/1729-4061.2017.90506

Article was submitted 30.12.2017.
After revision 25.01.2018.

Ахметов Б. Б.[1], Лахно В. А.[2], Малюков В. П.[3]

[1]Канд. техн. наук, доцент, ректор Каспійського державного університету технологій та інжинірингу ім. Ш. Есенова, Актау, Казахстан

[2]Д-р техн. наук, професор, зав. кафедри кібербезпеки та управління захистом інформаційних систем, Європейський університет, Київ, Україна

[3]Д-р ф.-м. наук, доцент, професор кафедри інформаційних систем та математичних дисциплін, європейський університет, Київ, Україна

МОДЕЛЬ СТРАТЕГІЙ ІНВЕСТУВАННЯ В СИСТЕМИ КИБЕРБЕЗПЕКИ СИТУАЦІЙНИХ ЦЕНТРІВ ТРАНСПОРТУ

Актуальність. Реалізована актуальна задача пошуку оптимальної стратегії управління процедурою взаємного фінансового інвестування в ситуаційний центр з кібербезпеки на транспорті.

Мета роботи – розробка моделі для системи підтримки рішень по безперервному взаємному інвестуванню в ситуаційний центр з кібербезпеки, яка відрізняється від наявних рішень, застосуванням білінійної диференційної гри з кількома термінальними поверхнями.

Метод. Для досягнення мети був застосований дискретний-апроксимаційний метод вирішення білінійної диференціальної гри якості із залежними рухами. Використання даного методу у розробленій системі підтримки рішень, на відміну від наявних, дає конкретні рекомендації при виборі управлінських рішень в інвестиційному процесі.

Результати. Запропонована модель дає конкретні рекомендації при виборі стратегії в інвестиційному процесі побудови захищеного ситуаційного центру. В ході обчислювального експерименту розглянуто новий клас білінійних диференційних ігор, що дозволило адекватно описати процес інвестування в засоби кібербезпеки ситуаційних центрів транспорту в Казахстані та Україні.

Висновки. Вперше запропонована модель, що описує процес взаємного інвестування, що базується на вирішенні білінійних відношень та диференційній грі з кількома термінальними поверхнями. Розглядається особливість диференціальної гри на прикладі взаємного інвестування в засоби кібербезпеки ситуаційного центру транспорту. При цьому права частина системи диференціальних рівнянь, представлена у вигляді білінійних функцій з будь-якими коефіцієнтами. Модель дозволяє прогнозувати результати інвестицій та пошуку стратегій управління інвестиційним процесом у системах захисту та кібербезпеки ситуаційного центру транспорту.

Ключові слова: система підтримки вирішення, модель, білінійні функції, взаємне інвестування, кібербезпека, ситуаційний центр транспорту.

Ахметов Б. Б.[1], Лахно В. А.[2], Малюков В. П.[3]

[1]Канд. техн. наук, доцент, ректор Каспийского государственного университета технологий и инжиниринга им. Ш. Есенова, Актау, Казахстан

[2]Д-р техн. наук, профессор, зав. кафедры кибербезопасности и управления защитой информационных систем, Европейский университет, Киев, Украина

[3]Д-р ф.-м. наук, доцент, профессор кафедры информационных систем и математических дисциплин, Европейский университет, Киев, Украина

МОДЕЛЬ СТРАТЕГИЙ ИНВЕСТИРОВАНИЯ В СИСТЕМЫ КИБЕРБЕЗОПАСНОСТИ СИТУАЦИОННЫХ ЦЕНТРОВ ТРАНСПОРТА

Актуальность. Решена актуальная задача нахождения оптимальной стратегии управления процедурой взаимного финансового инвестирования в ситуационный центр по кибербезопасности на транспорте.

Цель работы – разработка модели для системы поддержки решений по непрерывному взаимному инвестированию в ситуационный центр по кибербезопасности, отличающаяся от существующих решением билинейной дифференциальной игры качества с несколькими терминальными поверхностями.

Метод. Для достижения цели был использован дискретно-апроксимационный метод решения билинейной дифференциальной игры качества с зависимыми движениями. Применение данного метода в разработанной системе поддержки решений, в отличие от существующих, дает конкретные рекомендации при выборе управленческих решений в инвестиционном процессе.

Результаты. Предложенная модель дает конкретные рекомендации при выборе стратегий в инвестиционном процессе построения защищенного ситуационного центра. В ходе вычислительного эксперимента рассмотрен новый класс билинейных дифференциальных игр, который позволил адекватно описать процесс инвестирования в средства кибербезопасности ситуационных центров транспорта в Казахстане и Украине.

Выводы. Впервые предложена модель, описывающая процесс взаимного инвестирования, основанная на решении билинейных уравнений и дифференциальной игры качества с несколькими терминальными поверхностями. Рассмотрена особенность дифференциальной игры на примере взаимного инвестирования в средства кибербезопасности ситуационного центра транспорта. При этом правая часть системы дифференциальных уравнений представлена в виде билинейных функций с произвольными коэффициентами. Модель позволяет спрогнозировать результаты инвестирования и найти стратегии управления инвестиционным процессом в системы защиты и кибербезопасности ситуационного центра транспорта.

Ключевые слова: система поддержки решений, модель, билинейные функции, взаимное инвестирование, кибербезопасность, ситуационный центр транспорта.

REFERENCES

1. Gordon L. A., Loeb M. P., Lucyshyn W. et al. Externalities and the magnitude of cyber security underinvestment by private sector firms: a modification of the Gordon-Loeb model, *Journal of Information Security*, 2015, No. 6(1), pp. 24–30. DOI: 10.4236/jis.2015.61003
2. Fielder A., Panaousis E., Malacaria P. et al. Decision support approaches for cyber security investment, *Decision Support Systems*, 2016, No. 86, pp. 13–23. DOI: org/10.1016/j.dss.2016.02.012
3. Meland P. H., Tondel I. A., Solhaug B. Mitigating risk with cyberinsurance, *IEEE Security & Privacy*, 2015, No. 13(6), pp. 38–43. DOI: 10.1109/MSP.2015.137
4. Smeraldi F., Malacaria P. How to spend it: optimal investment for cyber security, *1st International Workshop on Agents and CyberSecurity*, Paris, France, 06–08 May 2014 : proceedings. New York, ACM, 2014, pp. 8. DOI: 10.1145/2602945.2602952
5. Tosh D. K., Molloy M., Sengupta S. et al. Cyber-investment and cyber-information exchange decision modeling, *High Performance Computing and Communications (HPCC), IEEE 7th International Symposium on Cyberspace Safety and Security (CSS)*, New York , 24–26 August 2015 : proceedings. New York, IEEE, 2015, pp. 1219–1224.
6. Alpcan T., Bambos N. Modeling dependencies in security risk management, *Risks and Security of Internet and Systems (CRiSIS), Fourth International Conference,* Toulouse, France, 19–22 October 2009: proceedings. France, IEEE, 2009. pp. 113–116. DOI : 10.1109/CRISIS.2009.5411969
7. Peppard J., Ward J. The Strategic Management of Information Systems: Building a Digital Strategy, U.K. Wiley, 2016, 504 p.
8. Malyukov V.P. A differential game of quality for two groups of objects, *Journal of Applied Mathematics and Mechanics*, 1991, Vol. 55, No.5, pp. 596 – 606.
9. Cassidy A. A practical guide to information systems strategic planning, New York, CRC press, 2016, 395 p.
10. Bonczek R. H., Holsapple C. W., Whinston A. B. Foundations of decision support systems, USA, Academic Press, 2014, 412 p.
11. Malyukov V.P. Discrete-approximation method for solving a bilinear differential game, *Cybernetics and Systems Analysis*, 1993,Vol. 29, No. 6, pp. 879–888.
12. Koziolek A., Avritzer A., Suresh S. et al. Assessing survivability to support power grid investment decisions, *Reliability Engineering & System Safety*, 2016, Vol. 155, pp. 30–43. DOI: 10.1016/j.ress.2016.05.015
13. Lakhno V., Petrov Al., Petrov Ant. Development of a Support System for Managing the Cyber Security of Information and Communication Environment of Transport, *Information Systems Architecture and Technology* : *38th International Conference on Information Systems Architecture and Technology (ISAT 2017)*, Wroclaw, 17–19 September 2017 : proceedings. Wroclaw, Springer, 2017, pp. 113–127. DOI: 10.1007/978-3-319-67229-8_11
14. Lakhno V. A. Development of a support system for managing the cyber security, *Radio Electronics, Computer Science, Control*, 2017, No. 2, pp. 109–116. DOI: 10.15588/1607-3274-2017-2-12
15. Lakhno V., Malyukov V., Gerasymchuk N. et al. Development of the decision making support system to control a procedure of financial investment, *Eastern-European Journal of Enterprise Technologies*, 2017, Vol. 6, No. 3, pp. 24–41. DOI: 10.15587/1729-4061.2017.119259
16. Akhmetov B., Lakhno V., Boiko Y. et al. Designing a decision support system for the weakly formalized problems in the provision of cybersecurity, *Eastern-European Journal of Enterprise Technologies*, 2017, Vol. 1, No. 2, pp. 4–15. DOI: 10.15587/1729-4061.2017.90506

УДК 004.932.2:004.93'1

Гадецька С. В.[1], Гороховатський В. О.[2]

[1]Канд. фіз.-мат. наук, доцент, доцент кафедри вищої математики, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

[2]Д-р техн. наук, професор, професор кафедри інформатики, Харківський національний університет радіоелектроніки, Харків, Україна

МЕТОДИ СТРУКТУРНОЇ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ЗАСАДАХ БАЄСОВСЬКОЇ ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ

Актуальність. Забезпечення результативності та багатофункціональності сучасних систем комп'ютерного зору вимагає створення різноманіття ефективних методів інтелектуального оброблення візуальної інформації. Розвиток систем структурного розпізнавання безпосередньо пов'язаний як із побудовою нових ефективних методів, так і з необхідністю створення дієвого механізму оцінювання результативності таких методів для довільних прикладних зразків візуальних даних. Одним із засобів, що базуються на статистичних характеристиках структурних даних, є апарат баєсовської теорії прийняття рішень. Обчислення апостеріорних ймовірностей віднесення опису візуального об'єкта до множини еталонів дає можливість як безпосередньо здійснювати розпізнавання на їх підставі, так і попередньо оцінити результативність процедур порівняння чи обчислення релевантності описів стосовно конкретної прикладної бази зображень. Особливу увагу приділяють вивченню структури множини дескрипторів зображень, що безпосередньо впливає на показники функціонування систем розпізнавання.

Мета. Вивчення можливості та особливостей застосування статистичної теорії розпізнавання щодо механізму прийняття рішень та оцінювання ефективності у виді ймовірностей віднесення опису об'єкта до класу, а також порівняння отриманих результатів обчислень із експериментальними даними комп'ютерного моделювання.

Метод. Запропоновано метод розпізнавання на основі застосування кластерних характеристик бази еталонних зображень із використанням баєсовської теорії прийняття рішень. Підсумком дослідження є створення механізму розпізнавання та оцінювання результативності процедур обчислення релевантності структурних описів.

Результати. Головним результатом статті є підтвердження фундаментального зв'язку методів порівняння з еталоном та статистичного підходу у розпізнаванні образів стосовно структурних описів у вигляді множини характерних ознак зображень, які результативно представлено кластерним виглядом. Запропонований у роботі більш простий в аспекті обчислювальних витрат статистичний підхід на підставі баєсовських оцінок може застосовуватися для попередніх розрахунків ефективності розпізнавання без проведення затратних експериментів з програмного моделювання.

Засвідчено ефективність розробленого методу обчислення ймовірнісних оцінок розпізнавання для прикладних баз зображень. Результат класифікації продемонстрував універсальність та коректність застосування методу, кожний із тестових об'єктів у декількох розглянутих базах зображень розпізнаний правильно.

Здійснено порівняння отриманих кількісних результатів обчислень із експериментальними даними комп'ютерного моделювання.

Висновки. У проведеному дослідженні запропоновано метод структурної класифікації зображень на основі кластерного подання опису засобами баєсовської теорії прийняття рішень. Основна ідея застосування належного математичного апарату полягає у віднесенні аналізованого об'єкту до еталону, що має найбільше значення апостеріорної ймовірності. Розроблений метод забезпечує достатній рівень розрізнення зображень, що підтвердили описані розрахунки та результати моделювання. Впроваджено механізм оцінювання результативності аналізованих методів структурного розпізнавання в межах прикладної бази зображень.

Наукова новизна дослідження полягає у синтезі нового методу структурного розпізнавання зображень та попереднього оцінювання ефективності шляхом застосування засобів баєсовської теорії прийняття рішень і побудови класифікаційних висновків у просторі кластер-еталон.

Практична значущість роботи – отримання прикладних розрахункових моделей для застосування методів структурного розпізнавання і підтвердження їх результативності в конкретних прикладах базах зображень.

Ключові слова: комп'ютерний зір, структурне розпізнавання зображень, множина структурних ознак, дескриптори SURF, релевантність описів, кластерне подання опису, формула Баєса, апостеріорна ймовірність віднесення до класу, критерій близькості опису у базі еталонів.

НОМЕНКЛАТУРА

СО – структурна ознака;

SURF – Speeded up robust features (прискорене виділення стійких особливостей);

ORB – Oriented FAST and Rotated BRIEF (орієнтований FAST и повернений BRIEF);

Z – база еталонних описів;

Z^j – еталон j-го класу;

J – число класів об'єктів;

M – розбиття на множині Z;

M_i – кластер з номером i;

k – число класів (кластерів) СО;

h – вектор подання еталону;

H – матриця кластерного опису бази зображень;

h_i^j – елемент вектора кластерного подання: i - номер кластеру, j – номер класу;

m – множина центрів кластерів;

s_j – кількість елементів в еталоні Z^j;

O – опис розпізнаваного візуального об'єкту;

o_l – l-й елемент опису розпізнаваного об'єкту;

s – кількість елементів у кластерному описі об'єкта;

ρ – метрика на множині СО;

δ – поріг значущості для метрики ρ;

© Гадецька С. В., Гороховатський В. О., 2018

DOI 10.15588/1607-3274-2018-2-10

r_j – відстань (ступінь релевантності) між кластеризованими векторними описами об'єкту та j-го еталону;

ε – поріг значущості для мінімуму релевантності;

$P\left(O / Z^j\right)$ – апріорна ймовірність належності об'єкта до j-го еталону;

$P\left(Z^j\right)$ – ймовірність появи еталону Z^j;

C_n^k – число комбінацій з n по k;

γ_d – сукупний критерій близькості опису класу з номером d у базі еталонів;

$\lambda_{\text{кор}}$ – коригуючий коефіцієнт.

ВСТУП

Фундаментальне завдання сучасних систем комп'ютерного зору зводиться до формування корисних висновків та реалізації результативних управляючих рішень стосовно об'єктів реального світу на основі аналізу їх зображень [1–8]. Практичними перевагами структурних методів аналізу та розпізнавання зображень в цих системах є можливість побудови інтелектуальних рішень на основі змінюваних підмножин елементів структурного опису, що забезпечує потрібну завадостійкість [1–3]. Розвиток систем структурного розпізнавання безпосередньо пов'язаний як із побудовою нових ефективних підходів, так і з необхідністю створення дієвого механізму оцінювання результативності методів для конкретних прикладних зразків даних. Одним із таких засобів є апарат баєсовської теорії прийняття рішень [4, 5].

Об'єкт дослідження – методи структурного розпізнавання зображень у системах комп'ютерного зору.

Предмет дослідження – застосування баєсовської теорії прийняття рішень для побудови методів розпізнавання та попереднього оцінювання результативності структурного розпізнавання в межах прикладної бази зображень із використанням кластерної моделі опису розпізнаваних візуальних об'єктів.

Мета роботи – вивчення можливості та особливостей застосування статистичної теорії розпізнавання щодо механізму прийняття рішень та отримання значень показників ефективності як ймовірностей віднесення опису об'єкта до класу, а також порівняння отриманих результатів обчислень із експериментальними даними комп'ютерного моделювання.

1 ПОСТАНОВКА ЗАДАЧІ

Формальною основою структурного розпізнавання є база описів зображень еталонів (алфавіт класів), що подана у вигляді сукупності СО як множини $Z = \{Z^j\}_{j=1}^J$ [1–3].

Виконаємо на множині Z деяке розбиття $Z = M = \{M_i\}_{i=1}^k$, $M_i \cap M_d = \varnothing, M_i \neq \varnothing$. Вважаємо $z \in M_i$ еквівалентними між собою. Маємо дві системи класів $\forall z \in Z$: $\{Z^j\}$ – для зображень еталонів та $\{M_i\}$ – для кластерів СО. Розбиттям $\{M_i\}_{i=1}^k$ опис Z^j еталону трансформується у вектор цілих чисел

$$h[Z^j] = \left(h_1^j, h_2^j, ...h_i^j, ...h_k^j\right),$$

$$h_i^j = card\{z \mid z \in Z^j \;\&\; z \in M_i\}. \quad (1)$$

Подання (1) – це образ еталону в кластерному вигляді, а множина векторів (1) бази із J зразків має вигляд матриці $H[Z] = \{\{h_i^j\}_{j=1}^J\}_{i=1}^k$. Рядок матриці H відображає кластерний опис окремого еталону, а стовпець – зміст кластера із числа рівноцінних елементів різних еталонів.

Виникає нагальна необхідність щодо побудови рішень та оцінювання ефективності методів структурного розпізнавання в аспекті кластерного подання даних. Привабливою є ідея застосування математичного апарату статистичної теорії розпізнавання, що ґрунтується на баєсовській теорії прийняття рішень для попереднього розрахунку результативності без проведення трудомістких експериментальних досліджень.

2 ОГЛЯД ЛІТЕРАТУРИ

Сучасні структурні методи оброблення візуальної інформації зводяться до оцінювання подібності дескрипторних описів аналізованого і еталонного зображень, що представлені у вигляді множин СО [1–3, 6–8]. Значення отриманої оцінки відображає ступінь релевантності двох інформаційних одиниць, які є описами об'єктів у побудованому просторі кількісних ознак.

Множина СО у вигляді сукупності числових векторів формується за допомогою спеціальних детекторів, найбільш поширеними серед яких на сьогодні вважають SURF і ORB [6–8]. Одним із шляхів продуктивного співставлення базової множини СО з аналізованими об'єктами є побудова описів еталонів на підставі кластеризації, що значно знижує розмірність досліджуваних даних за рахунок формування їх групованої просторової структури [2, 3, 8]. Результатом кластеризації є образ об'єкта (1), який можна розглядати як певну комбінаторну конфігурацію, організовану з елементів базової множини центрів кластерів $m = \{m_1, m_2, ..., m_k\}$, $m_i \in M_i$, на підставі сформованої кластерної структури.

Практичне впровадження методів структурного розпізнавання потребує застосування механізму оцінювання їх результативності для випробуваних прикладних зразків даних. Одним із таких засобів є апарат баєсовської теорії прийняття рішень, що є фундаментальним статистичним інструментарієм розпізнавання образів [4, 5]. Баєсовський класифікатор вирішує задачу дискримінантного аналізу і функціонує в термінах апостеріорної ймовірності віднесення об'єкта до класу [9, 10]. Ймовірнісний підхід також дає можливість здійснити попереднє оцінювання результативності структурного розпізнавання в межах прикладної бази зображень.

3 МАТЕРІАЛИ І МЕТОДИ

Під розпізнаванням розуміємо відображення множини описів об'єктів $\{O\}$ в скінченну множину номерів еталонів $\{1,...,J\}$, що здійснюється шляхом віднесення опису $O = \{o_l\}$ невідомого візуального об'єкта до одного з елементів еталонної множини $Z = \{Z^j\}_{j=1}^J$ (рис. 1). Розпізнавання базуємо на теоретико-множинному описі даних, що передбачає кластерну модель на підставі статистичного подання (1).

Рисунок 1 – Схема структурного розпізнавання

Процес розпізнавання будемо здійснювати порівняно за двома підходами:

1) шляхом зіставлення інтегральних характеристик образу виду (1) [2, 8];

2) застосуванням баєсовського класифікатора [9, 10].

Розглянемо формально сутність обох методів класифікації більш детально.

Метод 1.

1. Для побудови кластерного опису для $Z=\{Z^J\}_{j=1}^J$, що відображає статистичний розподіл еталонів за кластерами і не залежить від кількості СО в еталоні, нормуємо матрицю $H[Z]=\{\{h_i^j\}_{j=1}^J\}_{i=1}^k$ в рядках, отримуємо:

$$\alpha^i[Z^j]=\{\{h_i^{j^*}\}_{j=1}^J\}_{i=1}^k=\{\{h_i^j/s_j\}_{j=1}^J\}_{i=1}^k. \qquad (2)$$

2. Віднесемо кожний елемент $o_l\in O$ об'єкта до кластеру $M_i\subseteq M$ у відповідності з конкурентним правилом:

$$o_l\to M_i \mid \arg\min\rho(o_l,m_d)=i. \qquad (3)$$

Для фільтрації помилкових елементів, які можуть бути отримані в результаті реалізації (3), виконаємо верифікацію мінімуму m_i: $\rho(o_l,m_l)\le\delta$. Якщо нерівність не виконується, елемент o_l не відносимо до жодного з кластерів.

3. За результатом виконання кроку 2 формуємо опис (1) об'єкта $O=\{o_l\}$: $O=(h_1,h_2,...,h_k)^o$.

4. Обчислюємо ступінь r_j релевантності нормованих кластерних описів як відстань $r_j=\beta(\alpha^i[O],\alpha^i[Z^j])$ для об'єкта та рядків матриці еталонів у просторі R^k.

5. Віднесемо об'єкт O до класу $d\in\{1,...,J\}$ за правилом: $d=\arg\min_j r_j$. Перевіримо цінність мінімуму релевантності r_d: $r_d\le\varepsilon$, де ε – встановлений поріг. У разі порушення умови клас об'єкта вважаємо невизначеним через відсутність значущої відповідності у наявному еталонному просторі.

Метод 2.

1. Виходячи з кластерних описів еталонів $Z=\{Z^J\}_{j=1}^J$ і об'єкта $O=\{o_l\}$, обчислимо за формулою Баєса апостеріорні ймовірності $P(Z_j/O)$ віднесення об'єкта послідовно до кожного з описів $h[Z^j]$, $j\in\{1,...,J\}$ [4, 5]:

$$P(Z_j/O)=\frac{P(O/Z^j)\cdot P(Z^j)}{\sum_{j=1}^J P(O/Z^j)\cdot P(Z^j)}. \qquad (4)$$

Появу еталонів для спрощення аналізу вважаємо рівноймовірною.

Дискретний характер проблеми віднесення об'єкта $O=(h_1,h_2,...,h_k)^o$ до еталону Z^j в кластерному поданні призводить до обчислення набору ймовірностей узагальненого гіпергеометричного розподілу [11]:

$$P(O/Z^j)=\prod_{i=1}^k C_{h_i^j}^{h_i}/C_{s_j}^{s}. \qquad (5)$$

Зауважимо, що формула (5) має сенс лише для невід'ємних цілих значень параметрів, які задовольняють умовам:

$$h_i\le h_i^j,\ s_j\le s,\ i=1,...,k,\ j=1,...,J. \qquad (6)$$

Обмеження (6) у практичних розрахунках можуть потребувати певного корегування вхідних даних, яке можна здійснювати, в першу чергу, шляхом пропорційного змінювання характеристик досліджуваного об'єкта. Коригування кластерного подання об'єкта при невиконанні обмежень (6) пропонуємо здійснювати з округленням результату до цілого числа за формулою:

$$h_i'=\lambda_{\text{кор}}\cdot h_i,\ i=1,...,k,\ \lambda_{\text{кор}}=\min_i\frac{\min_j h_i^j}{h_i}.$$

Розрахунок здійснюється для $i\in\{1,...,k\}$, для яких порушена умова (6).

Вказані корегування загалом можуть вплинути на числові результати розрахунків, але принципово не змінюють рішення щодо факту найімовірнішого віднесення об'єкту до одного з еталонів. Відмітимо також, що при застосуванні формули (5) можливе виникнення складнощів при невеликих значеннях h_i^j. Це може означати надмірно дрібне розбиття на кластери, якого можна запобігти шляхом їх укрупнення.

2. Віднесемо об'єкт O до класу $d\in\{1,...,J\}$ за правилом: $d=\arg\max_j P(Z_j/O)$.

Як бачимо, перший метод обчислює ступінь релевантності шляхом дефініції подібності кластерних описів зображень об'єкта і еталонів, а другий – на підставі визначення значень ймовірності віднесення опису до відповідного класу. Оцінювання результативності процедур структурного розпізнавання в межах конкретної прикладної бази зображень здійснимо на підставі порівняння відповідних числових результатів – матриці відстаней між еталонами у методі 1 та ймовірностей віднесення еталону до класу у методі 2.

4 ЕКСПЕРИМЕНТИ

Розглянемо приклад реалізації методу 2 за допомогою експериментальної бази з числом еталонів 4 в поданні із 10 кластерів, до якої у [8] було застосовано метод 1. Проаналізуємо особливості використання методу 2, а також проведемо порівняльний аналіз результатів, одержаних за обома методами.

Вхідне кластерне подання мало вигляд, наведений у табл. 1 [8].

Процес знаходження апостеріорних ймовірностей $P(Z_j / O)$ віднесення об'єкта до описів $h[Z^j]$, $j \in \{1,...,J\}$ включає застосування формули (5) та передбачає виконання умов (6). Продемонструємо реалізацію підходу, взявши в якості об'єктів самі еталони.

Так, кластерне представлення об'єкта O^1, яким вибрано еталон Z^1, має перевищення кількості елементів у кластерах $M_2, M_3, M_5, M_6, M_8, M_{10}$ у порівнянні з кластерним поданням еталонів Z^2, Z^3, Z^4. Аналізуючи співвідношення між кількостями елементів відповідних кластерів, отримуємо найменший з коефіцієнтів пропорційності, рівний 0,21, що приводить до відкорегованого кластерного опису об'єкта O^1, що задовольняє умові (6):

$$O^1_{\text{кор}} = (2,2,2,2,3,3,1,2,2,2).$$

Таке коригування забезпечує можливість застосування традиційного підходу щодо розрахунку апостеріорних ймовірностей (5). При цьому вираз (5) обчислюється безпосередньо і, наприклад, для еталона Z^1 приймає вигляд:

$$P(O^1_{\text{кор}} / Z^j) = C_7^2 C_9^2 C_9^2 C_9^2 C_{14}^3 C_{15}^3 C_6^1 C_{11}^2 C_{11}^2 C_9^2 / C_{100}^{21}.$$

Обираючи об'єктами еталони Z^2, Z^3, Z^4, аналогічно отримуємо відкориговані кластерні описи:

$$O^2_{\text{кор}} = (3,4,3,6,1,5,5,3,4,3), \quad O^3_{\text{кор}} = (5,2,1,4,3,6,2,3,4,3),$$
$$O^4_{\text{кор}} = (2,2,2,3,3,3,2,3,3,2).$$

Таблиця 1 – Кількість СО еталонів в кластерному представленні

	Кластер									
Еталон	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
Z^1	7	9	9	9	14	15	6	11	11	9
Z^2	8	11	8	16	3	14	13	9	11	7
Z^3	15	7	3	13	9	18	7	8	11	9
Z^4	10	8	8	13	13	12	7	11	11	7

Вважаючи еталони рівноймовірними, що у нашому випадку означає $P(Z^j) = 0{,}25$, розраховуємо за формулою (4) апостеріорні ймовірності віднесення $O^1_{\text{кор}}$, $O^2_{\text{кор}}$, $O^3_{\text{кор}}$, $O^4_{\text{кор}}$ послідовно до кожного з описів $h[Z^j]$, $j \in \{1,...,4\}$. Результати обчислень занесемо у табл. 2.

У відповідності до методу 2 (крок 2) згідно з баєсовською теорією відносимо аналізований об'єкт до еталону, який має найбільш $O^3_{\text{кор}}$ є значення апостеріорної ймовірності. Результат класифікації підтверджує коректність роботи розробленого методу 2, оскільки кожний із тестових об'єктів розпізнаний правильно.

Зауважимо, що отримані значення ймовірностей мають суттєві відмінності: перший і четвертий еталони розпізнано не так впевнено (із значно нижчими показниками ймовірностей), ніж другий і третій. Ці ж висновки, що підкреслюють особливості розглянутої експериментальної бази даних, отримано нами у роботі [8], що підкреслює близькість обох методів. Зауважимо, що схожість теоретичного підґрунтя статистичних методів та методів порівняння з еталоном для бінарних зображень доведено у [4].

Адекватність проведених розрахунків можна також підтвердити значенням сукупного критерію близькості кожного з векторних описів у аналізованій базі еталонів [1]:

$$\gamma_d = \sum_{\substack{j=1 \\ j \neq d}}^{J} \sum_{i=1}^{k} \left| h_i^j - h_i^d \right|. \qquad (7)$$

Чим менше значення γ_d, тим ближче цей еталон до сукупності решти еталонів. Відповідно, і класифікується він гірше. У нашому випадку за таблицею 1 критерій (7) має значення $\gamma_1 = 82$, $\gamma_2 = 104$, $\gamma_3 = 96$, $\gamma_4 = 70$. Це свідчить про значнішу наближеність першого і четвертого еталонів до сукупності еталонів загалом, ніж другого і третього, що можна вважати додатковим поясненням відмінностей значень ймовірностей у таблиці 2. Як бачимо, еталон з номером 4 різниться в базі з найменшою ефективністю. Критерій (7) аналогічно матриці відстаней [4] має універсальне призначення щодо оцінювання ефективності довільних систем кількісних ознак.

Зауважимо, що кластерне подання еталонів і досліджуваного об'єкту можна розглядати як мультимножини, зважаючи, що центри кластерів виступають базовими елементами мультимножини. Отже, критерій (7) з цієї точки зору має сенс відстані між фіксованою мульти-

Таблиця 2 – Апостеріорні ймовірності віднесення коригованих описів об'єктів

	Об'єкт			
Еталони	$O^1_{кор}$	$O^2_{кор}$	$O^3_{кор}$	$O^4_{кор}$
Z^1	0,536	0,001	0,014	0,389
Z^2	0,005	0,993	0,001	0,014
Z^3	0,069	0,001	0,927	0,063
Z^4	0,389	0,005	0,058	0,533

множиною та сукупністю мультимножин бази еталонів, що може безпосередньо застосовуватися в прикладних задачах розпізнавання візуальних об'єктів [3].

Тепер зупинимося на результатах розрахунків щодо реалізації запропонованого методу 2 на прикладі експериментальної бази зображень гербів міст України (Дніпро, Київ, Харків, Кременчук), сформованої на підставі застосування детектору SURF [2]. Кластерне подання має вигляд табл. 3.

Результати обчислень, що представлені у табл. 4, свідчать про достатньо високий рівень ефективності розпізнавання, що можна пояснити апріорними суттєвими відмінностями між еталонами аналізованої бази у використовуваному просторі ознак.

Суттєва відмінність між еталонами в табл. 4 підтверджується також розрахунками за критерієм (7), який у ситуації різної кількості елементів в еталонах пропонуємо застосовувати через перехід до нормованого кластерного представлення [2]. Дійсно, критерій (7) приймає значення $\gamma_1 = 134,\ \gamma_2 = 150,\ \gamma_3 = 154,\ \gamma_4 = 155$, що не дає підстав стверджувати про наявність явно вираженої близькості жодного з еталонів до сукупності решти еталонів бази.

Проаналізуємо тепер результати застосування розробленого методу 2 до експериментальної бази морських ссавців [8], структурний опис якої отримано за допомогою детектора SURF, вхідне кластерне подання бази містить табл. 5. Результати обчислень наведемо в табл. 6.

Таблиця 3 – Кількість СО еталонів у кластерному представленні

Еталони	Кластери							
	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8
Z^1	40	75	24	9	5	61	5	34
Z^2	10	91	38	27	5	66	4	63
Z^3	102	113	20	17	4	14	3	69
Z^4	83	51	57	10	5	41	5	61

Таблиця 4 – Апостеріорні ймовірності розпізнавання коригованих об'єктів

Еталони	Об'єкт			
	$O^1_{\text{кор}}$	$O^2_{\text{кор}}$	$O^3_{\text{кор}}$	$O^4_{\text{кор}}$
Z^1	0,9999	0,00002	0,0001	0,0026
Z^2	0,0000	0,99998	0,0000	0,0000
Z^3	0,0000	0,00000	0,9992	0,0001
Z^4	0,0001	0,00000	0,0007	0,9973

Таблиця 5 – Кількість СО еталонів для бази морських ссавців

Еталони	Кластери				
	M_1	M_2	M_3	M_4	M_5
Z^1	10	11	3	7	8
Z^2	19	14	0	20	16
Z^3	7	12	4	26	5
Z^4	8	4	3	6	7

Таблиця 6 – Апостеріорні ймовірності розпізнавання коригованих об'єктів

Еталони	Об'єкт			
	$O^1_{\text{кор}}$	$O^2_{\text{кор}}$	$O^3_{\text{кор}}$	$O^4_{\text{кор}}$
Z^1	0,565	0,026	0,007	0,080
Z^2	0,253	0,952	0,090	0,130
Z^3	0,008	0,000	0,900	0,000
Z^4	0,174	0,021	0,003	0,789

Як бачимо із розрахункової табл. 6, результат класифікації реальних зображень також підтверджує коректність роботи розробленого методу 2. Зауважимо, що відповідні обчислення для 1-го, 3-го і 4-го еталонів тут здійснювалися без урахування значень 3-го кластера, що пов'язано із необхідністю дотримання обмежень (6) для коректного застосування формули (5). Тут можна вважати, що цей кластер містить настільки незначну у порівнянні з іншими кількість елементів, що не може принципово вплинути на остаточний результат щодо правильного розпізнавання об'єкту.

Зауважимо, що обчислення зіставних значень критерію (7) в умовах різного числа СО в описах еталонів та можливих нульових значень у поданні окремих еталонів потребує додаткового коригування вхідних параметрів h_i^j для виразу (7). Таке попереднє коригування може бути виконане шляхом вирівнювання кількості СО еталонів фільтрацією чи випадковим вибором, виключенням із аналізу окремих несуттєвих за значенням кластерів чи нормуванням через приведення цих параметрів до фіксованого діапазону.

Із проведених розрахунків для змодельованих та реальних описів візуальних об'єктів можна зробити узагальнюючий умовивід про універсальні властивості запропонованого методу 2 щодо прийняття рішення про клас об'єктів на підставі їх структурного опису, трансформованого до кластерного вигляду. Ключовим критерієм, який можна застосувати для верифікації розрахунків, є значення (7).

5 РЕЗУЛЬТАТИ

Провідним результатом статті є підтвердження та перевірка розрахунком фундаментального зв'язку методів порівняння з еталоном та статистичного підходу у розпізнаванні образів стосовно структурних описів у вигляді множини характерних ознак зображень, які результативно представлено у кластерним апаратом. Ці підходи при розпізнаванні можуть бути використані одночасно або для контролю результату прийняття рішень. З іншого боку, запропонований у роботі більш простіший в аспекті обчислювальних витрат статистичний підхід на підставі баєсовських оцінок може застосовуватися для попередніх розрахунків ефективності розпізнавання без проведення затратних експериментів з програмного моделювання.

Засвідчено ефективність розробленого методу обчислення ймовірнісних оцінок якості розпізнавання для прикладних баз зображень. Результат класифікації продемонстрував коректність застосування методу, оскільки кож-

ний із тестових об'єктів у декількох розглянутих базах зображень розпізнаний правильно.

Здійснено порівняння отриманих кількісних результатів обчислень із експериментальними даними комп'ютерного моделювання.

6 ОБГОВОРЕННЯ

Порівнюючи результати розпізнавання за двома аналізованими підходами – шляхом визначення інтегральних характеристик для структурного опису образу та застосуванням баєсовського класифікатора на підставі кластерного подання – отримано порівнянні результати класифікації. Числовими результатами є матриці відстаней між еталонами у першому методі та ймовірності віднесення еталону до класу – у другому.

Дискретна природа постановки проблеми і, відповідно, обчислення набору ймовірностей узагальненого геометричного розподілу в процесі застосування формули Баєса у розробленому методі 2 призводить до накладення певних обмежень щодо значень параметрів образу, задоволення яких може потребувати додаткового коригування вхідних даних – як на етапі формування кластерної структури еталонів і аналізованого об'єкту, так і на етапі роботи із сформованими структурними описами.

Проведені розрахунки для змодельованих і реальних прикладів візуальних об'єктів показали, що баєсовський класифікатор чуттєво реагує на різновиди опису і здійснює правильне розпізнавання із різним ступенем впевненості, що пояснюємо особливостями використаних конкретних експериментальних баз даних. Для підтвердження адекватності результатів класифікації застосовано критерій близькості осібного векторного опису у аналізованій базі еталонів.

ВИСНОВКИ

У проведеному дослідженні з використанням баєсовського апарату прийняття рішень вирішена задача класифікації зображень на базі структурного опису як множини характерних ознак, представленого у кластерному вигляді. Для цього запропоновано модифікований метод структурної класифікації. Сенс застосування баєсовського підходу полягає у віднесенні аналізованого об'єкту до еталону, що має найбільше значення апостеріорної ймовірності. Розроблений метод забезпечує достатній рівень розрізнення зображень, що підтвердили описані розрахунки та результати моделювання. Впроваджено також механізм оцінювання результативності методів структурного розпізнавання в межах прикладної бази зображень .

Наукова новизна дослідження полягає у синтезі нового методу структурного розпізнавання зображень та попереднього оцінювання ефективності шляхом застосування засобів баєсовської теорії прийняття рішень і побудови класифікаційних вирішень у просторі кластер-еталон.

Практична значущість роботи – отримання прикладних розрахункових моделей для застосування та оцінювання ефективності методів структурного розпізнавання і підтвердження їх результативності в конкретних прикладах базах зображень.

ПОДЯКИ

Робота виконана в рамках держбюджетної НДР Харківського національного університету радіоелектроніки «Розробка гібридних систем і методів обчислювального інтелекту для обробки потоків нечіткої інформації в умовах нестаціонарності та невизначеності» (номер держ. реєстрації 0116U002539).

СПИСОК ЛІТЕРАТУРИ

1. Гороховатский В. А. Структурный анализ и интеллектуальная обработка данных в компьютерном зрении / В. А. Гороховатский. – Х. : Компания СМИТ, 2014. – 316 с.
2. Гороховатский В. А. Структурное распознавание изображений с применением моделей интеллектуальной обработки и самоорганизации признаков / В. А. Гороховатский, А. В. Гороховатский, А. Е. Берестовский // Радиоэлектроника, информатика, управление. – 2016. – №3 (38). – С. 39–46.
3. Гороховатский В. А. Формализм мультимножеств в задачах структурного распознавания и поиска в базах видеоданных / В. А. Гороховатский, Ю. А. Куликов // Искусственный интеллект. – 2012. – № 1. – С. 76–85.
4. Duda R.O. Pattern classification / R. O. Duda, P. E. Hart, D. G. Stork. – 2ed., Wiley, 2000. – 738 p.
5. Шлезингер М. И. Математические средства обработки изображений / М. И. Шлезингер. – К. : Наукова думка, 1989. – 200 с.
6. Bay H. Surf: Speeded up robust features / H. Bay, T. Tuytelaars, L.Van Gool // Computer Vision: Ninth European Conference on Computer Vision, Graz, 7–13 May, 2006: proceedings. – Berlin : Springer, 2006. – P. 404–417.
7. Karami E. Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images / E. Karami, S. Prasad, M. Shehata. – Режим доступу https://www.researchgate.net/publication/292157133_
8. Гороховатский В. А. Классификация изображений на основе кластерного представления структурных описаний / В. А. Гороховатский, В. С. Столяров // Бионика интеллекта. – 2016. – № 2 (87). – С.83–87.
9. Вапник В. Н. Теория распознавания образов (статистические проблемы обучения) / В. Н. Вапник., А. Я. Червоненкис. – М. : Наука, 1974.– 416 с.
10. Гонсалес Р. Цифровая обработка изображений/ Р. Гонсалес, Р. Вудс; [пер. с англ.]. – М. : Техносфера, 2005. – 1070 с.
11. Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1 / В. Феллер; [пер. с англ.]. – М. : Мир, 1984.– 528 с.
12. Shapiro, L. and Stockman, G., (2001), Computer vision, Prentice Hall. – 625 p.
13. Szeliski R. Computer Vision: Algorithms and Applications / R. Szeliski. – London : Springer, 2010. – 979 p.

Стаття надійшла до редакції 06.11.2017.
Після доробки 18.12.2017.

Гадецкая С. В.[1], Гороховатский В. А.[2]

[1]Канд. физ.-мат. наук, доцент, доцент кафедры высшей математики, Национальный технический университет «Харьковский политехнический институт», Харьков, Украина

[2]Д-р техн. наук, профессор, профессор кафедры информатики, Харьковский национальный университет радиоэлектроники, Харьков, Украина

МЕТОДЫ СТРУКТУРНОЙ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ БАЙЕСОВСКОЙ ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ

Актуальность. Обеспечение результативности и многофункциональности современных систем компьютерного зрения требует создания разнообразия эффективных методов интеллектуальной обработки визуальной информации. Развитие систем структурного

распознавания непосредственно связано как с построением новых эффективных методов, так и с необходимостью создания действенного механизма оценивания результативности таких методов для конкретных прикладных образцов визуальных данных. Одним из средств, базирующихся на статистических характеристиках структурных данных, является аппарат байесовской теории принятия решений. Вычисление апостериорных вероятностей отнесения описания визуального объекта к множеству эталонов дает возможность как непосредственно производить распознавание с их помощью, так и предварительно оценить результативность процедур сравнения или вычисления релевантности описаний относительно конкретной прикладной базы изображений. Особенное внимание уделяют изучению структуры множества дескрипторов изображений, что непосредственно влияет на показатели функционирования систем распознавания.

Цель. Изучение возможности и особенностей применения статистической теории распознавания в механизме принятия решений и оценивание эффективности в виде вероятностей отнесения описания объекта к классу, а также сравнение полученных результатов вычислений с экспериментальными данными компьютерного моделирования.

Метод. Предложено метод распознавания на основе применения кластерных характеристик базы эталонных изображений с использованием байесовской теории принятия решений. Итогом исследования является создание механизма распознавания и оценивания результативности процедур вычисления релевантности структурных описаний.

Результаты. Главным результатом статьи является подтверждение фундаментальной связи методов сравнения с эталоном и статистического подхода в распознавании образов относительно структурных описаний в виде множества характерных признаков изображений, представленных кластерным описанием. Предложенный в работе более простой в аспекте расчетных затрат статистический подход на основе байесовских оценок может применяться для предварительных расчетов эффективности распознавания без проведения затратных экспериментов по программному моделированию.

Доказана эффективность разработанного метода вычисления вероятностных оценок для прикладных баз изображений. Результат классификации продемонстрировал универсальность и корректность применения метода, каждый из тестовых объектов в нескольких рассмотренных базах изображений распознан правильно.

Выполнено сравнение полученных количественных результатов вычислений с экспериментальными данными компьютерного моделирования.

Выводы. В проведенном исследовании предложено метод структурной классификации изображений на основе кластерного представления описания средствами байесовской теории принятия решений. Основная идея применения соответствующего математического аппарата состоит в отнесении анализируемого объекта к эталону, который имеет наибольшее значение апостериорной вероятности. Разработанный метод обеспечивает достаточный уровень различения изображений, что подтвердили описанные расчеты и результаты моделирования. Внедрено механизм оценивания результативности анализируемых методов структурного распознавания в рамках прикладной базы изображений.

Научная новизна исследования состоит в синтезе нового метода структурного распознавания изображений и предварительного оценивания эффективности с помощью применения средств байесовской теории принятия решений и построения классификационных выводов в пространстве кластер-эталон.

Практическая ценность работы – получение прикладных расчетных моделей для применения методов структурного распознавания и подтверждение их результативности в конкретных прикладных базах изображений.

Ключевые слова: компьютерное зрение, структурное распознавание изображений, множество структурных признаков, дескрипторы SURF, релевантность описаний, кластерное представление описания, формула Байеса, апостериорная вероятность отнесения к классу, критерий близости описания в базе эталонов.

Gadetska S. V.[1], Gorokhovatsky V. A.[2]

[1]PhD in Physics and Mathematics, Associate Professor, Associate Professor of the Department of Higher Mathematics, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine

[2]Dr. Sc. Sciences, Professor, Professor of the Department of Computer Science, National University of Radio Electronics, Kharkiv, Ukraine

STRUCTURAL CLASSIFICATION IMAGES USING BAYESIAN DECISION MAKING

Relevance. Ensuring the effectiveness and multifunctionality of modern computer vision systems requires the creation of a variety of effective methods for intellectual processing of visual information. The development of systems of structural recognition is directly connected with the construction of new effective methods as well as the need to create a mechanism for assessing the effectiveness of these methods for specific applications of visual data. Bayesian decision theory is one of the tools, based on the statistical characteristics of structural data. The calculation of a posteriori probabilities of assigning a description of a visual object to a set of etalons makes it possible to directly perform the process of recognition as well as preliminary evaluate the effectiveness of procedures for comparing or calculating the relevance of descriptions with respect to a specific application image database. Special attention is paid to the study of the structure of the set of descriptors, which directly affects the functioning of recognition systems.

Goal. Investigation of the possibility and peculiarities of the application of the statistical recognition theory in the decision-making mechanism and the evaluation of effectiveness in the form of the probabilities of classifying an object description as class. Comparison of the results of computations with experimental computer modeling data.

Method. A method of recognition based on the application of cluster characteristics of the image base using the Bayesian decision theory is proposed. The result of investigation is the creation of a mechanism for evaluating the effectiveness of procedures for calculating the relevance of descriptions with respect to the application image database.

Results. The main result of the paper is the confirmation of the fundamental relationship between methods of comparison with etalons and the statistical approach in pattern recognition with respect to structural descriptions in the form of a set of characteristic features of images represented by a cluster description. The statistical approach based on Bayesian estimates, which is simpler in sense of estimated costs, can be used for preliminary calculations of recognition efficiency without costly experiments on software modeling.

The effectiveness of the developed method for calculating probabilistic estimates for applied image bases is proved. The result of the classification demonstrated the universality and correctness of the application of the method, each of the test objects in several of the examined image bases was correctly recognized.

The obtained numerical results of the computations are compared with the experimental data of computer modeling.

Conclusions. In the conducted research the method of structural classification of images on the basis of a cluster representation of the description by means of Bayesian decision theory is proposed. The basic idea of applying the corresponding mathematical approach is in assigning the analyzed object to an etalon that has the greatest value of a posteriori probability. The developed method provides a sufficient level of discrimination of images, which was confirmed by the described calculations and simulation results, is offered. Mechanism for evaluating the effectiveness of the analyzed methods of structural recognition within the framework of the applied image database has been introduced.

The scientific novelty of the research consists in the synthesis of a new method of structural recognition of images and preliminary estimation of efficiency by using the means of Bayesian decision theory and constructing classificatory solutions in the space of a cluster-etalon.

The practical significance of the work is the obtaining of applied computational models for the application of the methods of structural recognition and confirmation of their effectiveness in specific applied image bases.

Keywords: computer vision, structural image recognition, set of structural attributes, SURF descriptors, relevance of descriptions, cluster representation of description, Bayesian formula, posteriori probability of classifying, proximity criterion of description in the base of etalons.

REFERENCES

1. Gorokhovatskiy V. A. Strukturnyy analiz i intellektual'naya obrabotka dannykh v komp'yuternom zrenii. Moscow, Kompaniya SMIT, 2014, 316 p.
2. Gorokhovatskiy V. A., Gorokhovatskiy A. V., Berestovskiy A. Ye. Strukturnoye raspoznavaniye izobrazheniy s primeneniyem modeley intellektual'noy obrabotki i samoorganizatsii priznakov, *Radio Electronics, Computer Science, Control,* 2016, No. 3 (38), pp. 39–46.
3. Gorokhovatskiy V. A., Kulikov YU. A. Formalizm mul'timnozhestv v zadachakh strukturnogo raspoznavaniya i poiska v bazakh videodannykh, *Iskusstvennyy intellekt,* 2012, No. 1, pp. 76–85.
4. Duda R. O., Hart P. E., Stork D. G. Pattern classification. 2ed., Wiley, 2000, 738 p.
5. Shlezinger M. I. Matematicheskiye sredstva obrabotki izobrazheniy. Kiev, Naukova dumka, 1989, 200 p.
6. Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features, *Computer Vision: Ninth European Conference on Computer Vision, Graz, 7–13 May, 2006:* proceedings. Berlin, Springer, 2006, pp. 404–417.
7. Karami E., Prasad S., Shehata M.. Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images. Rezhim dostupa https://www.researchgate.net/publication/292157133_
8. Gorokhovatskiy V. A., Stolyarov V. S. Klassifikatsiya izobrazheniy na osnove klasternogo predstavleniya strukturnykh opisan, *Bionika intellekta*, 2016, No. 2 (87), pp. 83–87.
9. Vapnik V. N., Chervonenkis A. YA. Teoriya raspoznavaniya obrazov (statisticheskiye problemy obucheniya). Moscow, Nauka, 1974, 416 p.
10. Gonsales R., Vuds R.; [Per. s angl.]. Tsifrovaya obrabotka izobrazheniy. Moscow, Tekhnosfera, 2005, 1070 p.
11. Feller V. [Per. s angl.]. Vvedeniye v teoriyu veroyatnostey i yeye prilozheniya. T. 1. Moscow, Mir, 1984, 528 p.
12. Shapiro L. and Stockman G. Computer vision, Prentice Hall, 2001, 625 p.
13. Szeliski R. Computer Vision: Algorithms and Applications. London: Springer, 2010, 979 p.

UDC 004.057.4:004.051

Dadenkov S. A.

Ph.D., Associate Professor at the Department of «Automatics and Telemechanics», Perm National Research Polytechnic University, Perm, Russia

THE SIMULATION OF P-PERSISTENT CSMA ALGORITHM OF UNEQUAL RANDOM MULTIPLE ACCESS

Context. The relevance of the simulation of p-persistent CSMA algorithm of unequal random multiple access is caused by importance of accounting personal node priorities for execution of adequate assessment of probable and time response characteristics of information transfer in designing widespread fieldbus-networks LonWorks, BacNet, etc.

The aim of this work is to obtain analytical expressions for estimating probabilities of successful and failed node access to the channel, the transmission time information in the network with equal and unequal rivalry of nodes based on p-persistent and predictive p-persistent CSMA algorithms of random multiple access.

Method. The method of probability theory has been used there to solve problems. The analysis of equal and unequal multiple access p-persistent CSMA nodes in the saturation mode of the channel by the transmission with a constant number of node-rivals has been executed. A method of estimating the time of successful transmission of information by the node through an average number of cycles of rivalry and probabilistic characteristics of the access nodes has been proposed. Expressions to estimate the necessary probabilistic and time characteristics were obtained: the average length of transmission and time of information transmission by the node, the effective bandwidth. The proposed method mainly differs from analogs by the possibility of characteristics estimation of information transmission given the manifold priorities of the network nodes, which improves the accuracy of the performed calculations.

Results. The analytical expressions have been received, quantification and analysis of probabilistic and temporal characteristics of information transfer in a network with equal and unequal random multiple p-persistent CSMA access has been performed.

Conclusions. The simulation results of equal and unequal rivalry network channel illustrate the importance of taking into account individual priorities of the nodes in the network. The feasibility of using access priorities is the need to transfer important information on the background of universal low-priority data. The significance of the results is to perform an adequate quantitative assessment of the characteristics of information transmission in a channel with equal and unequal parties.

Keywords: equal and unequal rivalry, random multiple access, probabilistic and temporal characteristics, industrial network, LonWorks, fieldbus, predictive p-persistent CSMA.

NOMENCLATURE

CSMA – an algorithm of unequal random carrier-sense multiple access;

BL – predicted load on channel (backlog);

c, C – nominal and effective carrier capacity of channel in the network;

$d_к$, d_y – an average number of access slots with collision and at successful transmission;

d_{y}^{l} – an average number of access slots for a successful node transfer with *l*-priority;

D – a probability function of selection by the node with priority $r \neq l$ of the winning node, of slot $>s$;

E – an average number of access attempts until successful transmission;

I – an function that determines whether a node with priority *l*, slot $s < l$;

k – an access number leading to successful transmission;

l, L – number of priority access slots of the node and a plurality of slots of network nodes;

ln – information packet length in bits;

m, *M* – number of messages in the node and the set of messages in the network nodes;

n – number of nodes in the network;

$p_к$, $p_{\text{к}}^{s}$ – probability of unsuccessful transmission (with collision) in the network, on the slot with the number *s*;

p_{y}, p_{y1} – probability of successful transmission in the network, by one node;

$p_{\text{y1}}^{l}, p_{\text{y1}}^{l,s}$ – probability of successful node access with *l*-priority, on the slot with number *s*;

$p_{\text{y1}}^{l,s'}$ – conditional probability of winning a node with priority *l*, on the slot number *s*;

s, S – the current number of the access slot of the node and a plurality of network nodes slots;

T – an average time of successful information transfer by one network node;

T^{l} – an average time of successful transfer of the information packet by the node with *l*-priority;

T_m – an average time of successful transmission of m information messages from the node queue;

$T_{\text{п}}$ – an average duration of successful transmission of the information packet;

W, W_{base} – the current and base width of access window (race)

X, Y – sets of parameters and characteristics of observations;

β_1 – a minimum inter-packet access interval;

β_2 – duration of the slot (priority/random);

β_3 – a time of transmission of one bit information over the network channel;

τ – duration of a burst (packet) cycle;

$\tau_{\text{к}}$, τ_{y} – duration of unsuccessful (with collision) and successful burst cycle;

τ_0 – a time of node delay in race for access to channel.

© Dadenkov S. A., 2018

DOI 10.15588/1607-3274-2018-2-11

INTRODUCTION

The wide prevalence of random multiple access algorithms in industrial fieldbus networks is caused by the potentially high efficiency of transfer channel capacity usage and low time of information delivery in the transmission of sporadic traffic, as compared with deterministic algorithms. A variety of modifications of the probabilistic algorithms of unequal random carrier-sense multiple access CSMA is caused by different efficiency of information transmission depending on: channel occupancy, the number of nodes in the network, regular/sporadic traffic and other parameters. Such access algorithms are widespread: CSMA/CA (CAN, KNX/EIB, DeviceNet networks), 1-CSMA/CD (EtherNet/IP, ModBus/TCP), predictive *p*-Persistent CSMA (BacNet, LonWorks). Literature analysis [1–6] shows the high efficiency of *p*-persistent CSMA access algorithms adapted or self-adapting (predictive) to the changing network channel load. The problem of using this algorithm in an industrial network is caused by several factors. Firstly, the small elaboration of the methods and the low accuracy of the quantitative assessment of the information delivery performance and warranty do not allow develop real-time networks. Secondly, the methods of analyzing the characteristics of networks with unequal participants, which limits the effectiveness of the algorithm in the transfer of technological information of varying importance (priority): controlling, regulation, signaling, measurement, data, are practically not worked out. This significantly reduces the efficiency of the *p*-persistent CSMA algorithm. The solution of the problems actualizes the study of the issues of upgrading of the known and creating new models of the algorithm and methods for analyzing the characteristics of information transmission in the network.

The object of the research is the algorithm of unequal random multiple access *p*-Persistent CSMA nodes to the shared transmission medium.

The subject of the study is models and methods for evaluation the probabilistic and temporal characteristics of information transfer in the network with the analyzed algorithm.

The aim of this work was to develop a method for evaluating the characteristics of information transmission in the network with the algorithm *p*-Persistent CSMA, characterized by a set of parameters {*c*, *n*, *M*, *L*, *W*, *BL*, *ln*, β_1, β_2, β_3}: the probability of successful p_y and unsuccessful $p_к$ access of nodes to the channel, the average time *T* of successful information transmission, the effective carrier channel capacity *C*.

1 PROBLEM STATEMENT

Let there are a complete set of observations <*X, Y*>, where *X* is a set of observation parameters, *Y* is a set of characteristics. The set of parameters and characteristics of the observations can be represented as arrays:

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1a} & \dots & x_{1A} \\ x_{21} & x_{22} & \dots & x_{2a} & \dots & x_{2A} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{q1} & x_{q2} & \dots & x_{qa} & \dots & x_{qA} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{Q1} & x_{Q2} & \dots & x_{Qa} & \dots & x_{QA} \end{pmatrix},$$

$$Y = \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1b} & \dots & y_{1B} \\ y_{21} & y_{22} & \dots & y_{2b} & \dots & y_{2B} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ y_{q1} & y_{q2} & \dots & y_{qb} & \dots & y_{qB} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ y_{Q1} & y_{Q2} & \dots & y_{Qb} & \dots & y_{QB} \end{pmatrix}.$$

Since the parameters of $\{x_q\}$ observations take discrete values, the evaluation of particular probability characteristics of $\{y_q\}$ can be performed using the probability multiplication theorem. Since the set of observations is finite and formed by a complete group of events, and observations are equally probable, the evaluation of the general probabilistic characteristics p_q for the groups of observations can be performed in several ways. The first method is based on the classical definition of probability, according to which the probability of an event is the relation of the number of favorable elementary outcomes to the total number of all possible outcomes. The second method is based on the use of the formula of total probability and Bayes′ rule that allow to calculate absolute *p* and conditional *p′* probabilities of events. This method is preferable because of the complexity of direct calculation of probabilities in solving problem. In this case, the evaluation of the average values of the observation characteristics can be obtained by superposition of their individual values, given the conditional probabilities of the outcomes:

$$y_b = \sum_q p_q^{'} \cdot y_{qb}.$$

2 REVIEW OF THE LITERATURE

The analysis of networks with different algorithms of random multiple access in the literature is execute mainly with the use of classical Queuing systems with repeated calls (as a result of loss or lack of response). Such works include publications of famous Russian authors: A. A. Nazarov [7, 8], I. I. Khomyakov [9], S. N. Stepanov [10]. Because of the specificity of access algorithms, the results of the work are mainly applicable to the analysis of networks with 1-persistent CSMA/CD access algorithm used in Industrial Networks Industrial Ethernet, ModbusTCP, EtherNet / IP, HSE, ProfiNet, Interbus-TCP / IP, etc. Taking into account the functioning peculiarities of the studied protocols *p*-persistent CSMA with and without loading is performed in the works of the authors Moshe Kam [1], M. Miśkowicz [2, 3], P. Buchholz and J. Plonnigs [4], S. A. Dadenkov's [5]. The analysis is performed using the apparatus of probability theory and discrete Markov chains.

The research of *p*-persistent protocols of the CSMA family in [1–3] was performed in the mode of channel saturation with information transmission by a constant number of network nodes. The models and methods of analysis constructed as a result of researches allow to execute an assessment of the top level of information transmission delay in the network. The models take into

account the set of algorithm parameters [6]. To obtain more accurate estimates of probabilistic and temporal characteristics of the network with sporadic nature of the information load in the works [4, 5], were suggested the models considering the dynamics of changes not only in the Protocol parameter of the width of the access algorithm window, which is characteristic of the works [1–3], but also the number of messages and nodes-rivals for the communication channel. From the literature [6] there is well known the possibility of using the priorities of access nodes for the information transmission of different importance. The small elaboration of this issue and the absence of models and methods for assessing the characteristics of networks with unequal access of nodes limit the possibility of applying the procedure in practice. This actualizes the solution of the tasks set in the work.

3 MATERIALS AND METHODS

The algorithm of random multiple access *p*-persistent CSMA is characterized by pseudo-constant level of persistence *p* nodes of network, determined by the ability to start transfer to one of the temporary access slots. The predictive *p*-persistent CSMA Protocol is characterized by a variable persistence level, changing according to the load predicted for the communication channel. Regulation of probability p of beginning of transfer by a node in a wide range [0.0625..0.000976] allows to reduce probability of collision at increase of loading on the network channel and to reduce time of access at decrease in loading of the channel. Load prediction is carried out by distributed evaluation by the nodes, where "response" network traffic expected to transfer according to the data of the field delta_backlog package.

Access nodes to the channel according to the algorithm *p*-persistent CSMA is performed in burst (packet) cycles. Cycles are divided into free and busy packets. A busy packet cycle represents the channel access time and packet transfer time (fig. 1).

The algorithm of access to the channel includes the following steps (Fig. 1) [6]:

1. Testing channel activity. The node monitors the state of the channel and moves to the next step when there is no activity in the channel.

2. Minimum inter-packet access interval. The node waits for the time inter-packet interval β_1 to determine that the channel is free.

3. Priority access interval. The node waits for a fixed number of priority time slots *l*, which each duration is β_2. A lower number of access slots correspond to the highest priority. Priorities for access are used only after employing burst cycles [6].

4. Random access interval. The node waits for a delay *T*, that is equal to a random number *s* of time slots of β_2 duration. The number of slots is choosing according to the uniform distribution of the range [0...*W* – 1], where the time slot length $W=W_{base} \bullet BL$, W_{base} – the base access slot length (16), *BL* (backlog) – the load predicted on channel [1..64]. The access slot length determines the adaptive (for predictive) probability level of node beginning transfer $p=1/W$.

5. Testing channel activity and transmission package. The node starts transmission if the channel is free and suspends it on the next packet cycle otherwise. Packet transmission delay is to product of the packet length *ln* in bits and the transfer time of one information bit β_3, inversely proportional to the carrier capacity of channel. When several nodes choose an equal number of temporary (priority and random) slots, the transfer results in a collision that damages the packets. The packets involved in the collision, depending on the service transmission will be lost or require retransmission. The algorithm with prediction at collision increases the load level, and in a successful and free cycle decreases by 1 [6].

Let us analyze the communication rivalry of nodes according to the algorithm of multiple access *p*-persistent CSMA in saturation mode, but taking into account the priorities of nodes not previously analyzed in the known works. Saturation mode is characterized at every moment by competition for the channel of *n* nodes, each of which has an individual number of priority slots *l* and contains $m \geq 1$ messages for transfer.

$$T = \tau_0 + E \cdot \tau, \tag{1}$$

where *E* – is the average number of access attempts for a successful transmission; τ – is the average duration of the burst (packet) cycle, regardless of the type (successful/ unsuccessful); delay $\tau_0 = 0$ for saturation mode, while the stochastic transfer $\tau_0 \leq \tau - \beta_1$. The number of access attempts *E* is unique for each node and depends on the number of priority access slots l_i of nodes-rivals per channel.

The probability of successful transmission can be determined by the sum of the probability of winning of *n* nodes with different priorities *l* belonging to the set *L*:

$$p_y = \sum_{i=1}^{n} p_{y1}^{l_i} = \sum_{\forall l \in L} n_l \cdot p_{y1}^{l}, \tag{2}$$

where p_{y1}^{l} is the probability of successful access and transmission by one node with *l*-priority.

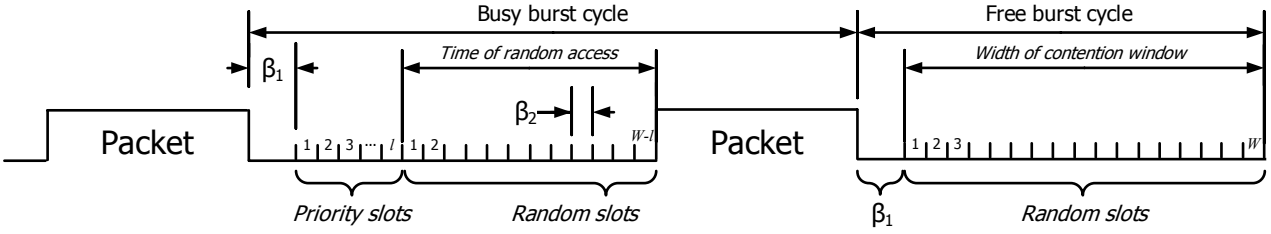


Figure 1 – The structure of burst cycles

In a channel with no or equal priority of nodes, the probability of successful transmission does not depend on the priority and is determined only by the number of nodes-rivals:

$$p_{\text{y}} = n \cdot p_{y1}. \quad (3)$$

The probability of an unsuccessful (collision) in saturation mode of the network channel:

$$p_{\kappa} = 1 - p_{y}. \quad (4)$$

The probability of successful transmission of the node p_{y1}^{l} with an l-priority passing on an access attempt (in a burst cycle) with a number k is determined by the product of event probabilities:

– the node hadn't been winning the rivalry before attempt with the number k;

– the node won the rivalry on the access attempt k:

$$p_{y1}^{l}(k) = (1 - p_{y1}^{l})^{k-1} \cdot p_{y1}^{l}. \quad (5)$$

In that case the average number of transmission attempts by the node E with the l-priority:

$$E^{l} = \sum_{k=1}^{\infty} k \cdot p_{y1}^{l}(k) = \sum_{k=1}^{\infty} k \cdot \left(1 - p_{y1}^{l}\right)^{k-1} \cdot p_{y1}^{l}. \quad (6)$$

Multiplying both sides of the expression (6) by the value, we obtain:

$$(1 - p_{y1}^{l}) \cdot E^{l} = \sum_{k=1}^{\infty} k \cdot (1 - p_{y1}^{l})^{k} \cdot p_{y1}^{l}. \quad (7)$$

The expression difference (6) and (7) determine the average number of attempts of node access with an l-priority:

$$E^{l} = \left[1 + (1 - p_{y1}^{l}) + (1 - p_{y1}^{l})^{2} + \ldots\right]. \quad (8)$$

The expression (8) contains an infinite geometric series, so:

$$E^{l} = (p_{y1}^{l})^{-1}. \quad (9)$$

The average number of attempts before a successful transmission by nodes without regard to the priority level of a node can be defined as the arithmetic mean:

$$E = \frac{1}{n} \cdot \sum_{\forall l \in L} \frac{n_l}{p_{y1}^{l}}. \quad (10)$$

The average number of attempts before a successful transmission by nodes without regard to the priority level of a node can be defined as the arithmetic mean:

$$E = p_{y1}^{-1}. \quad (11)$$

To assess the probability of successful access of the node with an l-priority in the burst cycle with the time slot length W, let us analyze rivalry of n nodes with different priorities in the field of access with the number of time-slots equal to the sum of the number of priority l and random d slots (for example, the slot width $W = W_{\text{base}} = 16$) (fig. 2).

The duration of the rivalry between the nodes for access to the channel is limited by the number of time slots W free burst cycle, regardless of the priorities of the nodes. This limits the number of random slots that can be used to organize successful transmission of information by nodes: $\{l+1 \ldots W - l_{\min}\}$, where l is the number of priority slots of the node, $l_{\min}$ is the minimum number of priority slots of the node of analyzed network. Thus, the width of the random access range does not change and remains equal to $W - l_{\min}$ (for convenience in work W). In the access field of nodes with the number of priority slots greater than 0, access slots appear, the use of which cannot lead to a successful transmission. These slots are the initial slots $s \leq 1$ and the final access slots $s > W - l_{\min}$ (fig. 2, l=4, $l_{\min}$=0). This should be taken into account when calculating the probabilities of successful access of nodes with different priorities. Further analysis of the mechanism of slot access is advisable to perform for slots, the number of which starts with 1. Therefore, if the value of $l_{\min}$ exceeds 0, the shift "to the left" is necessary and renumbering of priorities of all nodes to the value of $l_{\min}$ according to the rule $l = (l_{\text{old}} - l_{\min} + 1)$, l_{old} – the old number of access slots of the node. The value of the $l_{\min}$ shift must then be taken into account as a constant component of the number of network channel access slots.

The probability of a node's successful access to a channel is the sum of the probability $p_{y1}^{l,s}$ of winning a rivalry on each slot s from the width of the access window $[1 \ldots W]$:

$$p_{y1}^{l} = \sum_{s=1}^{W} p_{y1}^{l,s}. \quad (12)$$

The probability $p_{\text{y1}}^{l,s}$ selecting by the node a slot number s, provided that the node wins on this slot, is determined by the joint execution of the next conditions:

– the winning node has a choice of slot number $s \geq l$;

– the winning node with priority l selects one slot from W;

– other nodes with priority l select one of the $(W - s)$ of later slots;

– other nodes with priorities other than l select one of the later slots:

$$p_{\text{y1}}^{l,s} = I(s \geq l) \cdot \frac{1}{W} \cdot \left(\frac{W - s}{W}\right)^{n_l - 1} \cdot \prod_{\forall r \in (L \setminus \{l\})} D(s > r), \quad (13)$$

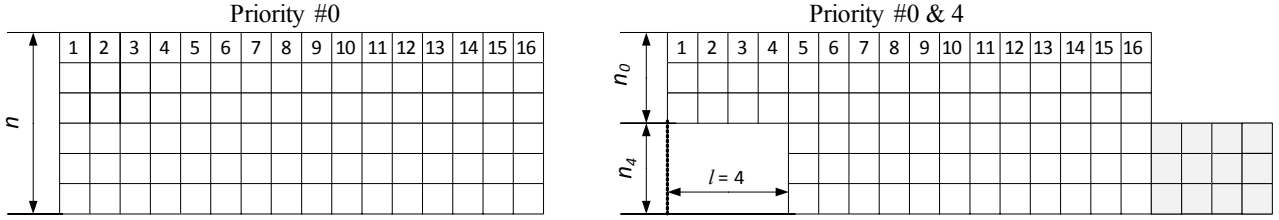


Figure 2 – Fields of equal and unequal rivalry of nodes for access to the channel

where the indicator function $I(s \ge l)$ considers the inability of choice by node with priority l of a slot with absolute number $s < l$ (fig. 2), probability function $D(s > r)$ of selecting by nodes with priorities r, different from l winning node, of slot with number > s:

$$I(\cdot) = \begin{cases} 1, \text{ condition is satisfied } (\cdot), \\ 0, \text{condition } (\times) \text{ is not satisfied;} \end{cases}$$

$$D(\cdot) = \begin{cases} \left(\frac{W-s+r}{W}\right)^{n_r}, \text{ condition is satisfied } (\cdot), \\ 1, \text{ condition } (\cdot) \text{ is not satisfied.} \end{cases} \quad (14)$$

The expression (3) to assess the probability of successful transmission in the channel takes the form:

$$p_{\text{y}} = \sum_{\forall l \in L} \left[n_l \cdot \sum_{s=1}^{W} \left[I(s \ge l) \cdot \frac{1}{W} \cdot \left(\frac{W-s}{W}\right)^{n_l - 1} \cdot \prod_{\forall r \in (L \setminus \{l\})} D(s > r) \right] \right]. \quad (15)$$

Particular type of the expression of the probability of a successful transmission regardless of the type of priority:

$$p_{\text{y}} = \frac{n}{W} \cdot \sum_{s=1}^{W} \left[\left(\frac{W-s}{W}\right)^{n-1} \right]. \quad (16)$$

The average number of burst cycles (9) before a successful transmission by node with l-priority, considering (12):

$$E^l = \left[\sum_{s=1}^{W} \left[I(s \ge l) \cdot \frac{1}{W} \cdot \left(\frac{W-s}{W}\right)^{n_l - 1} \cdot \prod_{\forall r \in (L \setminus \{l\})} D(s > r) \right] \right]^{-1}. \quad (17)$$

The expression (17), without regard to the priorities of the nodes-rivals:

$$E = \left[\frac{n}{W} \cdot \sum_{s=1}^{W} \left[\left(\frac{W-s}{W}\right)^{n-1} \right] \right]^{-1}. \quad (18)$$

The average length of the packet cycle in (1) at the width of the access window W is the sum of the lengths of the successful and failed packet cycles, taking into account their probabilities:

$$\tau = p_{\text{к}} \cdot \tau_{\text{к}} + p_{\text{y}} \cdot \tau_{\text{y}}, \quad (19)$$

where τ_{y}, $\tau_{\text{к}}$ – average length of successful and unsuccessful packet cycles,

$$\tau_{\text{y}} = \beta_1 + [l_{\min} + d_{\text{y}} - 1] \cdot \beta_2 + \tau_{\text{п}},$$

$$\tau_{\text{к}} = \beta_1 + [l_{\min} + d_{\text{к}} - 1] \cdot \beta_2 + \tau_{\text{п}}. \quad (20)$$

where $l_{\min}$ is the minimum number of priority access slots of a network node, d_{y} is the average number of slots in which a node wins a rivalry; $d_{\text{к}}$ is the average number of slots at the collision; $\tau_{\text{п}}$ is the packet transmission duration $\tau_{\text{п}} = \ln \cdot \beta_3$.

The average number of slots d_{y}^l selected by the winning node with an l-priority in a successful packet cycle can be estimated as the sum of the probabilities of selecting each slot s, provided that the node wins the rivalry on that slot:

$$d_{\text{y}}^l = \sum_{s=1}^{W} s \cdot p_{\text{y1}}^{l,s'}, \quad (21)$$

where $p_{\text{y1}}^{l,s'}$ is the conditional probability of winning a node with priority l on slot s:

$$p_{\text{y1}}^{l,s'} = \frac{p_{\text{y1}}^{l,s}}{p_{\text{y1}}^{l}}. \quad (22)$$

Then the average number of slots on successful access:

$$d_{\text{y}} = \sum_{\forall l \in L} \frac{n_l}{n} \cdot d_{\text{y}}^l = \sum_{\forall l \in L} \frac{n_l}{n} \cdot \sum_{s=1}^{W} s \cdot p_{\text{y1}}^{l,s'} =$$

$$= \sum_{\forall l \in L} \frac{n_l}{n \cdot p_{\text{y1}}^{l}} \cdot \sum_{s=1}^{W} s \cdot p_{\text{y1}}^{l,s} = \sum_{\forall l \in L} \frac{n_l}{n} \cdot \sum_{s=1}^{W} (s-1) \cdot p_{\text{y1}}^{l,s} \cdot \quad (23)$$

Without regard to the priorities of nodes in the expression (23) takes the form:

$$d_{\text{y}} = \sum_{s=1}^{W} (s-1) \cdot p_{\text{y1}}^{l,s} =$$

$$= \sum_{s=1}^{W} \left[(s-1) \cdot I(s \ge l) \cdot \frac{1}{W} \cdot \left(\frac{W-s}{W}\right)^{n_l - 1} \cdot \prod_{\forall r \in (L \setminus \{l\})} D(s > r) \right]. \quad (24)$$

The average number of slots in an unsuccessful access is determined similarly (21):

$$d_{\text{к}} = \sum_{s=1}^{W} s \cdot p_{\text{к}}^{s'}, \quad (25)$$

where $p_{\text{к}}^{s'}$ is the conditional probability of collision at the width of the window W on the slot with the number s:

$$p_{\text{к}}^{s'} = \frac{p_{\text{к}}^{s}}{p_{\text{к}}}, \quad (26)$$

where $p_{\text{к}}^{s}$ is the absolute probability of collision at the width of the window W on the slot with the number s.

The collision on slot s with the width of the access window W occurs under joint conditions:

– $i \ge 2$ nodes choose equal number of access slots, with probability: $\frac{1}{W^i}$;

– the other n-i nodes select the larger s slot from the remaining ones, with the probability $\prod^{n-i} \frac{W-s+l}{W}$, where l is the individual priority got into collision;

– number of all possible combinations C_n^i of i nodes got into collision of n nodes-rivals: $C_n^i = \frac{n!}{(n-i)! \cdot i!}$;

– the possibility of participating in collision of different numbers of i=[2..n] nodes-rivals for the network channel.

Thus, an expression to evaluate the probability of a collision on a slot with a number s:

$$p_{\text{к}}^{s} = \sum_{i=2}^{n} C_n^i \cdot \frac{1}{W^i} \cdot \prod^{n-i} \frac{W-s+l}{W}. \quad (27)$$

The application of the proposed expression (27) in practice is complicated by the realization of the procedure of search of various combinations of i-nodes in collision and (n–i) nodes not in collision with different levels of priorities l. Therefore, an expression can be used to evaluate this value, regardless of the priorities of the rivalry-nodes:

$$p_{\text{к}}^{s} = \sum_{i=2}^{n} C_n^i \cdot \frac{(W-s)^{n-i}}{W^n}. \quad (28)$$

The expression (28) allows to obtain the upper estimation of probability of collision on slot s for the scenario of rivalry of n nodes with different levels of priorities. To define more precisely the value of the probability of collision should be taken into account for each slot s access the maximum number n of rivalry-nodes.

The expression (19) for estimating the duration of a burst cycle takes the form:

$$\tau = \beta_1 + [d_{\text{к}} \cdot p_{\text{к}} + d_{\text{у}} \cdot p_{\text{у}} + l_{\min} - 1] \cdot \beta_2 + \tau_{\text{п}}. \quad (29)$$

Average time of successful transmission of one information packet by a node with l-priority, taking into account delay of messages transmission of other n–1 nodes:

$$T^l = \tau_{\text{о}} + E^l \cdot \tau = \tau_{\text{о}} + \frac{\tau}{p_{\text{у}1}^l}. \quad (30)$$

Average time of successful transmission of a message in a channel, regardless of the priority of the node:

$$T = \tau_{\text{о}} + \frac{\tau}{n} \cdot \sum_{\forall l \in L} \frac{n_l}{p_{\text{у}1}^l}. \quad (31)$$

For the channel network, regardless of or equal to the priorities of the nodes, the expression (31):

$$T = \tau_{\text{о}} + \frac{\tau \cdot n}{p_{\text{у}}} = \tau_{\text{о}} + \tau \cdot W \cdot \left[\sum_{s=1}^{W} \left(\frac{W-s}{W} \right)^{n-1} \right]^{-1}. \quad (32)$$

The average transmission time of any node message with queue $m \geq 1$:

$$T_m = \frac{\sum_{i=1}^{m} i \cdot T}{m}. \quad (33)$$

The average length of successful transmission through the communication channel, taking into account possible collision, is determined by (32), but taking into account the full probability of successful access:

$$T_{\text{п}} = \frac{\tau}{p_{\text{у}}}. \quad (34)$$

The effective network capacity, which shows the number of packets successfully transmitted per unit of time, inversely proportional to the duration of transmission:

$$C = T_{\text{п}}^{-1}. \quad (35)$$

The performed analysis allowed us to obtain analytical expressions for estimating the main probabilistic and temporal characteristics of information transmission in the network channel with the algorithm of unequal random multiple access p-Persistent CSMA in the conditions of unequal rivalry between n nodes with individual l-priorities, and the width of the access window.

Packet transmission over the network with the predictive p-persistent CSMA access algorithm is accompanied by a dynamic change in the width of contention window W, and hence the transmission characteristics. Analytical models of dynamic change of window width and methods of estimation of characteristics of information transmission in saturation mode are offered in works [1–3]. The work [5] proposes a multidimensional discrete Markov chain describing the sporadic process of formation and transmission of information messages by nodes, considering the change not only in the width of the access window, but also the number of rival nodes per channel and messages for transmission. The use of the proposed expressions for assessing the characteristics of unequal rivalry in these models will improve the accuracy of calculations. The ongoing study of access priorities at the constant width of the contention window W is aimed at evaluating the effectiveness of the mechanism of unequal access of the algorithm p-persistent CSMA.

4 EXPERIMENTS

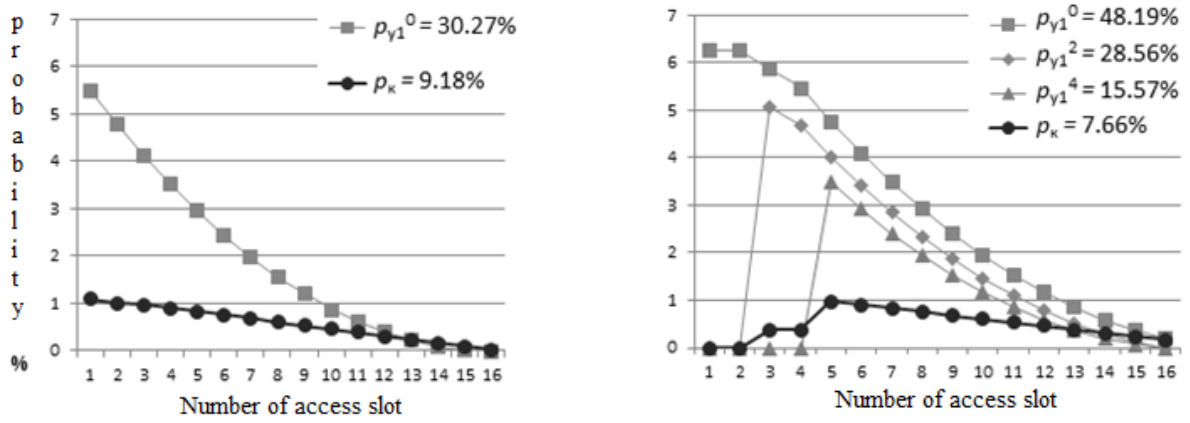
Study of the effectiveness of the unequal access of the algorithm p-persistent CSMA is performed in the saturation mode of the network channel. The estimation and comparative analysis of probabilities of successful and unsuccessful access of nodes to the channel, time of information transmission at equal and unequal rivalry in a network with different number of nodes are carried out. The initial data of the experiments and evaluate the characteristics shown in table 1.

Table 1 – Initial data and estimated characteristics

№ Experiment	Number of node n_l		Window width W	Characteristics
1	n=3	n_0=3	16	$p_{\text{у}1}^{l,s}$, $p_{\text{у}1}^{l,s'}$, $p_{\text{у}1}^{l}$, $p_{\text{у}}^{l}$, $p_{\text{к}}^{s}$, $p_{\text{к}}$, $T_{\text{п}}$
		n_0=1 n_2=1 n_4=1		
2	n=9	n_0=9		
		n_0=3 n_2=3 n_4=3		

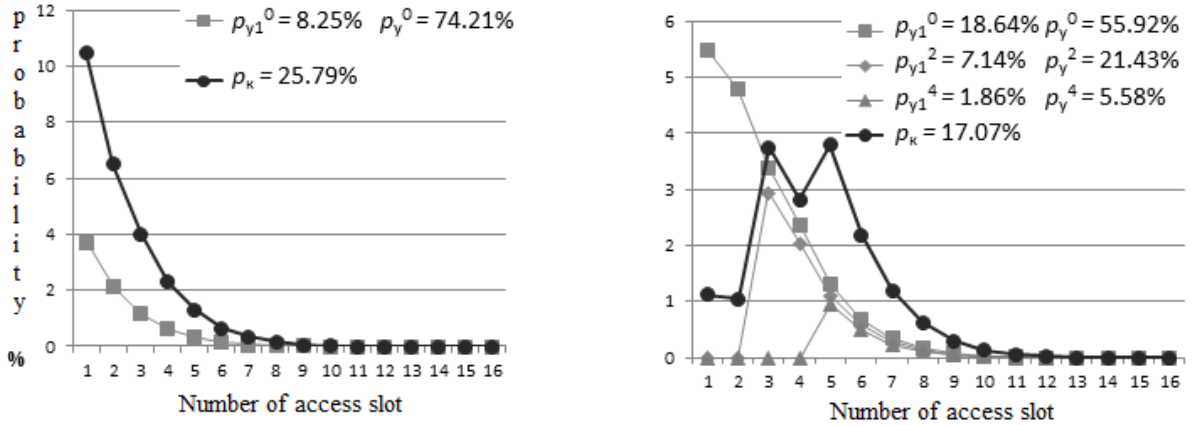
5 RESULTS

The results of evaluation of probabilistic characteristics of access $p_{y1}^{l,s}$, $p_{y1}^{l,s'}$, p_{y1}^{l}, p_{y}^{l}, $p_{к}^{s}$, $p_{к}$ for experiment 1 are shown in figure 3, and for experiment 2 in figure 4 and 5. The graphs illustrate the dependences of the probabilities of successful and unsuccessful node access to the channel with equal and unequal access from the time slot number *s* of the rivalry window *W*. The legends of the graphs contain quantitative values of the probability of access to the channel by one and all nodes with a certain priority, the probability of collision. The analysis of the results will assess the potential efficiency of using access algorithm priorities to reduce the probability of collision, as well as reducing the probability of successful transmission for nodes with low priorities.



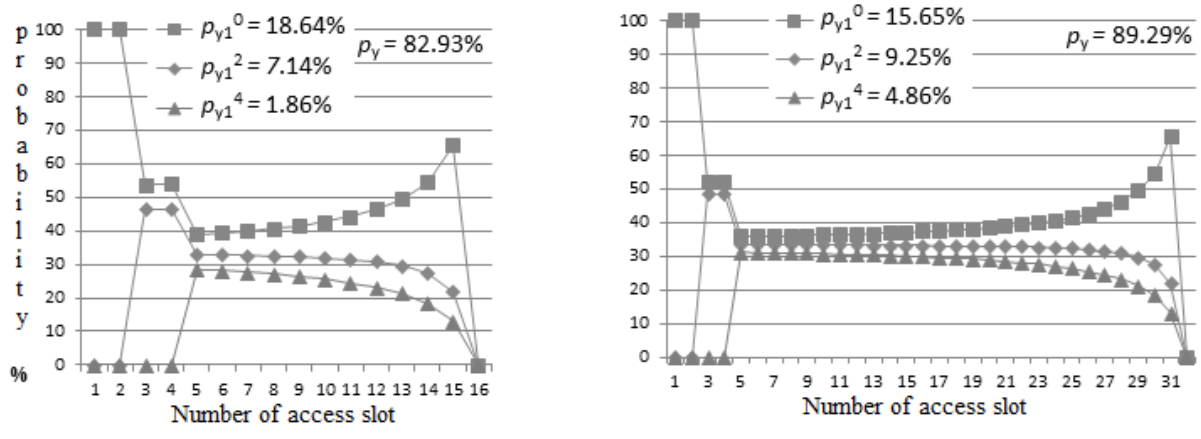
a b

Figure 3 – Graphs of the probabilities of successful and unsuccessful access (experiment 1):
a – equitable rivalry; b – unequal rivalry



a b

Figure 4 – Graphs of the probabilities of successful and unsuccessful access (experiment 2):
a – equitable rivalry; b – unequal rivalry



a b

Figure 5 – Graphs of conditional probability of winning in the rivalry on access slots

The results of estimating the average time of successful $T_п$ transmission in a network channel with equal and unequal nodes (experiment 1 and 2) are given in table 2. The analysis of taken results will allow to estimate the potential increase of the effective capacity C from the channel by reducing the probability of collision between the $p_к$ and unequal access.

Table 2 – Average transmission time (ms) in a channel with equal and unequal nodes

n	Rivalry	W=16	W=32	W=48	W=64	W=80	W=96
3	Equal	2,95	3,52	4,15	4,81	5,47	6,14
	Unequal	3,21	3,82	4,47	5,13	5,80	6,47
9	Equal	3,08	2,94	3,10	3,31	3,55	3,80
	Unequal	3,03	3,16	3,37	3,61	3,85	4,11

6 DISCUSSION

A comparative analysis of probabilities of successful and unsuccessful information transmission at equal and unequal rivalry between the node slots of the window (fig. 3, fig. 4) illustrates a reduction in the probability of a collision $p_к$ in a channel using access priorities. For experiment 1, the relative decrease in the probability of collision was $\approx$16% (from 9.18% to 7.66%, fig. 3), but with the increase in the number of nodes (experiment 2, fig. 4) reached $\approx$33% (from 25.79% to 17.07%). Reducing of collision probability in unequal rivalry achieved by the redistribution of probability of access in a time slots of rivalry window. Separation of access of rival nodes on initial (1–4) slots, which account for the greatest probability of collision (fig. 3a, fig. 4a), contributes to the organization of successful transmission nodes with high priority (fig. 3b, fig. 4b, fig. 5). The displacement of the rivalry nodes from the initial slots leads to the appearance of probability peaks of collision on slots s = 3, 5 with an increasing number of nodes-rivals on the channel (fig. 3b, fig. 4b). The tendency to decrease the probability of a collision with an increasing number of the access slot is maintained due to reducing the probability of a collision of packets on later slots.

Predictable feature of unequal rivalry is essential difference of probabilities of successful nodes access to the channel, hence time of information transmission. Analysis of graphical dependencies (fig. 3b, fig. 4b) illustrates a small difference in the probability of successful access of nodes with different priorities on the slots of the contention window W. However, the analysis of the absolute values of the access probabilities p_{y1}^{l} of l-priority nodes shows the opposite. In the first and second experiments, the probability ratios for nodes with extreme priorities (0 and 4) are 3 (48.19:15.57) and 10 (18.64:1.86), respectively, against 1 for equal competition. So that, the access time for successful transmission of low-priority nodes is also 3 and 10 times different from high-priority nodes. The difference in the probability of successful access $p_{y1}^{l,s'}$ of nodes with different priorities for slots is clearly illustrated in figure 5. A significant difference in access probabilities was found on the initial and final slots. However, the degree of significance of the final slots of the access is low due to the smallness of the absolute value of the successful transmission probability (<1%) (fig. 3b, fig. 4b). On the middle slots of the contention window (fig. 5) there is a low differentiation of access probabilities of nodes with different priorities. As the width of the rivalry window W increases, access probabilities "blur" (fig. 5), which reduces the differences between high- and low-priority nodes and without priority nodes. There is also an abrupt reduction in the difference between unsuccessful equal and unequal access. The ratio of probabilities of successful access of high- and low-priority nodes decreases rapidly to 2 and below. The obtained result at W=16 is due to the high ratio of the number of priority access slots l=4 and the window width.

The disadvantage of reducing the probability of collision with a small width of the window W, besides to differentiating the probability of access of unequal nodes, is the increase the number of temporary access slots for successful transmission of information compared to equal rivalry. For example, for experiment 2 (W=16), the number of successful/ unsuccessful access slots, with equal and unequal rivalry is 2.09/2.30 and 3.28/4.46, accordingly. The increase in the number of slots leads to an increase in the delay of access and information transmission. However, the increase in the number of slots of unequal access is accompanied by a decrease in the probability of collision. The results of the assessment of the opposite influence of given characteristics of unequal access on the time of information transmission $T_п$ are compared with the indicators characteristic of equal access in table 2 at different widths of the contention window W. The analysis of the obtained results evidences the privileged influence on the transmission time of the number of access slots. The achieved reduction in transmission time by reducing the probability of a collision does not exceed the increase time access due to the increased number of time slots. A minor exception is observed only in the second experiment at W=16. Thus, the transmission time of the information packet on the network channel is less for equal access. This ensures more efficient use of the bandwidth of the channel with the transmission of a large amount of useful information per unit of time. All the conclusions obtained in the work are also valid for the rivalry of nodes with a minimum number of priority slots other than 0. Comparing the unequal rivalry of nodes with the same rivalry, but with the priorities of nodes shifted by l, illustrates the equality of the absolute probabilistic characteristics of access, but a significant increase in the number of access slots, which increases the delay in the information transmission.

The analysis of the experimental results illustrates the expedience of using access priorities of the p-persistent CSMA algorithm to organize cycles of important information transmission against the background of the transfer of generally significant low-priority data. In this case, a reduction in the probability of collision is achieved and the time of information transmission is less than with equal rivalry. Unequal rivalry can't be used to increase the effective use of the bandwidth of the channel, due to the increase in the number of slots and, accordingly, the access time.

CONCLUSIONS

The article analyzes the algorithm of access to the transmission medium p-persistent CSMA. The review of publications and literature sources revealed the relevance

and insufficient elaboration of the problem of quantitative estimation of probabilistic and temporal characteristics of information transmission in networks with unequal access on the basis of the analyzed algorithm. As a result of the analysis, analytical expressions are obtained for assessing the probability of successful/unsuccessful access and the time of information transmission considering the diverse priorities of nodes. Experiments are performed and the comparative analysis of equal and unequal rivalry of knots for the network channel is executed. The conclusions about the importance of taking into account the priorities in the network with the method of access are drawn. The expediency of using access priorities for organization of cycles of important information transmission against the background of general-purpose low-priority data transfer is noted. Obtaining adequate quantitative estimates and conclusions of the use of priorities in the sporadic nature of the information load requires additional research, for example, using the model [5]. The practical significance of the obtained results lies in the possibility of quantitative assessment and design of network channels with unequal random multiple access type *p*-persistent CSMA with specified probabilistic and temporal characteristics of information transmission.

ACKNOWLEDGEMENTS

The work is executed at financial support of the Ministry of education and science of the Russian Federation (RFMEFI58317X0062) in the boundaries of the BRICS project "Hybrid process of manufacturing parts for the aerospace industry: modelling, software development and verification". The author thanks the scientific advisor professor Kon E. L. For the help and useful discussion of the work.

REFERENCES

1. Moshe Kam. Collision Resolution Simulation for Distributed Control Architectures using LonWorks / Moshe Kam // IEEE International Conference on Automation Science and Engineering. – Edmonton, Canada : IEEE, 2005. – P. 319–326. DOI: 10.1109/COASE.2005.1506789
2. Miśkowicz M. Analysis of Mean Access Delay in Variable-Window CSMA / M. Miśkowicz // Sensors. Schweiz: Molecular Diversity Preservation International, 2007. – Vol. 7. – P. 3535–3559.
3. Miśkowicz M. Access delay in LonTalk MAC protocol / M. Miśkowicz // Computer Standards and Interfaces. Nederland: Elsevier Science Publishing Company, 2009. – P. 548–556. DOI: 10.1016/J.CSI.2008.03.025
4. Buchholz P. Analytical analysis of access-schemes of the CSMA type / P. Buchholz, J. Plonnigs // Proc. of IEEE International Workshop on Factory Communication Systems WFCS'2004, Wien, 2004. – P. 127–136.
5. Даденков С. А. Подход к построению аналитической модели информационно-управляющей сети LonWorks на основе нейрочипов / С. А. Даденков, Е. Л. Кон // Нейрокомпьютеры: разработка, применение. – 2013. – № 11. – С. 64–69.
6. Дитрих Д. LON-технология, построение распределенных приложений / Д. Дитрих, Д. Лой, Г. Ю. Швайнцер // Пер. с нем. под ред. О. Б. Низамутдинова. – Пермь : Звезда, 1999 – 242 с.
7. Назаров А. А. Исследование сети связи, управляемой адаптивным протоколом случайного множественного доступа, в условиях критической загрузки / А. А. Назаров, Д. Ю. Кузнецов // Проблемы передачи информации. – 2004. – № 3. – С. 69–80.
8. Назаров А. А. Общий подход к исследованию марковских моделей сетей передачи данных, управляемых статическими протоколами случайного множественного доступа / А. А. Назаров, С. А. Цой // Автоматика и вычислительная техника. – 2004. – № 4. – С. 73–85.
9. Хомичков И. И. Исследование моделей локальной сети с протоколом случайного множественного доступа / И.И. Хомичков // Автоматика и телемеханика.– 1993.–№ 12.–С.89–90.
10. Степанов С. Н. Асимптотический анализ моделей с повторными вызовами в области больших потерь / С.Н. Степанов // Проблемы передачи информации. – 1993. – № 3. – С. 54–75.

Article was submitted 21.12.2017.
After revision 25.01.2018.

Даденков С. А.

Канд. техн. наук, доцент кафедри «Автоматика і телемеханіка» Пермського національного дослідницького політехнічного університету, г. Пермь, Росія

МОДЕЛЮВАННЯ АЛГОРИТМУ НЕРІВНОПРАВНИХ СЛУЧАЙНОГО МНОЖИННОГО ДОСТУПУ P-PERSISTENT CSMA

Актуальність моделювання алгоритму нерівноправного випадкового множинного доступу p-persistent CSMA викликана важливістю обліку індивідуальних пріоритетів вузлів для виконання адекватної оцінки імовірнісних і тимчасових характеристик передачі інформації при проектуванні поширених fieldbus-мереж LonWorks, BacNet і ін.

Мета роботи полягає в отриманні аналітичних виразів для оцінки ймовірностей успішного і невдалого доступу вузлів до каналу, часу передачі інформації в мережі з рівноправним і нерівноправним суперництвом вузлів на основі алгоритмів випадкового множинного доступу p-persistent і predictive p-persistent CSMA.

Метод. Для розв'язання задачі використано апарат теорії ймовірностей. Виконано аналіз множинного рівноправного і нерівноправного доступу p-persistent CSMA вузлів в режимі насичення каналу передачею інформації постійним числом вузлів-суперників. Запропоновано спосіб оцінки часу успішної передачі інформації вузлом через середня кількість циклів суперництва і імовірнісних характеристик доступу вузлів. Отримані вирази для оцінки необхідних імовірнісних і тимчасових характеристик: середньої тривалості передачі і часу передачі інформації вузлом, ефективної пропускної здатності каналу. Запропонований спосіб переважно відрізняється від аналогів можливістю оцінки характеристик передачі інформації з урахуванням різнорідних пріоритетів вузлів мережі, що підвищує точність виконуваних розрахунків.

Результати. Отримано аналітичні вирази і виконана кількісна оцінка і аналіз імовірнісних і тимчасових характеристик передачі інформації в мережі з рівноправним і нерівноправним випадковим множинним доступом p-persistent CSMA.

Висновки. Результати моделювання рівноправного і нерівноправного суперництва в мережевому каналі ілюструють важливість врахування індивідуальних пріоритетів вузлів в мережі. Доцільність використання пріоритетів доступу полягає в необхідності передачі важливої інформації на тлі загальнозначущих фонових даних. Значимість результатів полягає у виконанні адекватної кількісної оцінки характеристик передачі інформації в каналі з рівноправними і нерівноправними учасниками.

Ключові слова: рівноправне і нерівноправне суперництво, випадковий множинний доступ, імовірнісні і тимчасові характеристики, промислова мережа, LonWorks, fieldbus, predictive p-persistent CSMA.

Даденков С. А.
Канд. техн. наук, доцент кафедры «Автоматика и телемеханика» Пермского национального исследовательского политехнического университета, г. Пермь, Россия

МОДЕЛИРОВАНИЕ АЛГОРИТМА НЕРАВНОПРАВНОГО СЛУЧАЙНОГО МНОЖЕСТВЕННОГО ДОСТУПА *P*-PERSISTENT CSMA

Актуальность моделирования алгоритма неравноправного случайного множественного доступа *p*-persistent CSMA вызвана важностью учета индивидуальных приоритетов узлов для выполнения адекватной оценки вероятностных и временных характеристик передачи информации при проектировании распространенных fieldbus-сетей LonWorks, BacNet и др.

Цель работы состоит в получении аналитических выражений для оценки вероятностей успешного и неудачного доступа узлов к каналу, времени передачи информации в сети с равноправным и неравноправным соперничеством узлов на основе алгоритмов случайного множественного доступа *p*-persistent и predictive *p*-persistent CSMA.

Метод. Для решения задачи использован аппарат теории вероятностей. Выполнен анализ множественного равноправного и неравноправного доступа *p*-persistent CSMA узлов в режиме насыщения канала передачей информации постоянным числом узлов-соперников. Предложен способ оценки времени успешной передачи информации узлом через среднее количество циклов соперничества и вероятностных характеристик доступа узлов. Получены выражения для оценки необходимых вероятностных и временных характеристик: средней продолжительности передачи и времени передачи информации узлом, эффективной пропускной способности канала. Предложенный способ преимущественно отличается от аналогов возможностью оценки характеристик передачи информации с учетом разнородных приоритетов узлов сети, что повышает точность выполняемых расчетов.

Результаты. Получены аналитические выражения и выполнена количественная оценка и анализ вероятностных и временных характеристик передачи информации в сети с равноправным и неравноправным случайным множественным доступом *p*-persistent CSMA.

Выводы. Результаты моделирования равноправного и неравноправного соперничества в сетевом канале иллюстрируют важность учета индивидуальных приоритетов узлов в сети. Целесообразность использования приоритетов доступа заключается в необходимости передачи важной информации на фоне общезначимых низкоприоритетных данных. Значимость результатов заключается в выполнении адекватной количественной оценки характеристик передачи информации в канале с равноправными и неравноправными участниками.

Ключевые слова: равноправное и неравноправное соперничество, случайный множественный доступ, вероятностные и временные характеристики, промышленная сеть, LonWorks, fieldbus, predictive *p*-persistent CSMA.

REFERENCES

1. Moshe Kam. Collision Resolution Simulation for Distributed Control Architectures using LonWorks, *IEEE International Conference on Automation Science and Engineering. Edmonton, Canada: IEEE*, 2005, pp. 319–326. DOI: 10.1109/COASE.2005.1506789
2. Miśkowicz M. Analysis of Mean Access Delay in Variable-Window CSMA, *Sensors. Schweiz: Molecular Diversity Preservation International*, 2007, Vol. 7, pp. 3535–3559.
3. Miśkowicz M. Access delay in LonTalk MAC protocol, *Computer Standards & Interfaces. Nederland: Elsevier Science Publishing Company*, 2009, pp. 548–556. DOI: 10.1016/J.CSI.2008.03.025
4. Buchholz P., Plonnigs J. Analytical analysis of access-schemes of the CSMA type, *Proc. of IEEE International Workshop on Factory Communication Systems WFCS'2004, Wien*, 2004, pp. 127–136.
5. Dadenkov S. A., Kon E. L. Podxod k postroeniyu analiticheskoj modeli informacionno-upravlyayushhej seti LonWorks na osnove nejrochipov, *Nejrokomp'yutery: razrabotka, primenenie*, 2013, No. 11, pp. 64–69.
6. Ditrix D. Loj D., Yu G. Shvajncer LON-texnologiya, postroenie raspredelennyx prilozhenij. Per. s nem. Pod red. O. B. Nizamutdinova. Perm', Zvezda, 1999, 242 p.
7. Nazarov A. A., Kuznecov D. Ju. Issledovanie seti svjazi, upravljaemoj adaptivnym protokolom sluchajnogo mnozhestvennogo dostupa, v uslovijah kriticheskoj zagruzki, *Problemy peredachi informacii*, 2004, No. 3, pp. 69–80.
8. Nazarov A. A., Coj S. A. Obshhij podhod k issledovaniju markovskih modelej setej peredachi dannyh, upravljaemyh staticheskimi protokolami sluchajnogo mnozhestvennogo dostupa, *Avtomatika i vychislitel'naja tehnika*, 2004, No. 4, pp. 73–85.
9. Homichkov, I.I. Issledovanie modelej lokal'noj seti s protokolom sluchajnogo mnozhestvennogo dostupa, *Avtomatika i telemehanika*, 1993, No. 12, pp. 89–90.
10. Stepanov, S.N. Asimptoticheskij analiz modelej s povtornymi vyzovami v oblasti bol'shih poter', *Problemy peredachi informacii,* 1993, No. 3, pp. 54–75.

УДК 004.042:004.94

Копп А. М.[1], Орловский Д. Л.[2]

[1]Аспирант кафедры программной инженерии и информационных технологий управления Национального технического университета «Харьковский политехнический институт», Харьков, Украина
[2]Канд. техн. наук, доцент, доцент кафедры программной инженерии и информационных технологий управления Национального технического университета «Харьковский политехнический институт», Харьков, Украина

ПОДХОД К АНАЛИЗУ И ОПТИМИЗАЦИИ МОДЕЛЕЙ БИЗНЕС-ПРОЦЕССОВ В НОТАЦИИ BPMN

Актуальность. В настоящее время все больше организаций представляют знания о своей деятельности с точки зрения бизнес-процессов. В то же время, построение моделей бизнес-процессов носит субъективный характер, что приводит к возникновению ошибок, связанных с компетентностью аналитика, либо, к примеру, с недостаточно подробным описанием предметной области. Таким образом, перед тем, как модели бизнес-процессов будут использованы для сбора, хранения и распространения организационных знаний, они должны быть проанализированы с точки зрения наличия различного рода недостатков.

Цель. Целью работы является разработка подхода к анализу и оптимизации моделей бизнес-процессов, используемых для представления организационных знаний.

Метод. В работе было рассмотрено и дополнено формальное представление моделей бизнес-процессов на основе ориентированных графов. Были рассмотрены и формализованы правила построения моделей бизнес-процессов. На основе характеристики структурной избыточности графа, а также коэффициента сбалансированности моделей бизнес-процессов, был предложен коэффициент структурного соответствия моделей бизнес-процессов правилам построения. С учетом формализованных правил построения и коэффициента структурного соответствия, был предложен формальный подход, предназначенный для формирования рекомендаций по совершенствованию моделей бизнес-процессов.

Результаты. С использованием разработанного программного обеспечения, которое реализует предложенный подход к анализу и оптимизации моделей бизнес-процессов, были выполнены контрольные расчеты для моделей бизнес-процессов закупки и продажи продукции. Для данных моделей были определены ошибки построения и сформированы рекомендации по их устранению. Преобразованные модели лишены ошибок и в дальнейшем могут быть использованы для представления организационных знаний.

Выводы. Предложенный подход позволяет определять ошибки, допущенные при построении моделей бизнес-процессов, и формировать рекомендации по их устранению.

Ключевые слова: моделирование бизнес-процессов, нотация моделирования, анализ, оптимизация.

НОМЕНКЛАТУРА

BPM – Business Process Management;
BPMN – Business Process Model and Notation;
XML – eXtensible Markup Language;
XPDL – XML Process Definition Language;
BPEL – Business Process Execution Language;

S – множество вершин, представляющих собой стартовые события;

A – множество вершин, представляющих собой задачи;

G – множество вершин, представляющих собой шлюзы;

F – множество вершин, представляющих собой конечные события;

$E \subseteq (V \times V)$ – множество направленных ребер;

L – множество меток вершин;

$l : V \to L$ – функция разметки вершин;

$p : E \to [0,1]$ – функция, сопоставляющая ребрам значения вероятности;

x_j – булева переменная, принимающая значение 0, в случае невыполнения j-го правила построения модели бизнес-процесса;

r – количество правил построения моделей бизнес-процессов;

$\Delta|S|$ – целочисленные изменения в количестве вершин, представляющих собой стартовые события;

$\Delta|F|$ – целочисленные изменения в количестве вершин, представляющих собой конечные события;

$\Delta \deg^+(v_i)$ – целочисленные изменения в значениях полустепеней захода i-й вершины графа BG;

$\Delta \deg^-(v_i)$ – целочисленные изменения в значениях полустепеней исхода i-й вершины графа BG;

CSC' – значение коэффициента структурного соответствия с учетом изменений.

ВВЕДЕНИЕ

В связи с тем, что все больше организаций представляют знания о своей деятельности с точки зрения бизнес-процессов, возникает необходимость использования информационных технологий, позволяющих моделировать, внедрять и анализировать бизнес-процессы [1].

В качестве таких инструментов выступают системы класса BPM, которые предоставляют средства моделирования, проектирования, внедрения, анализа и непрерывного совершенствования бизнес-процессов в организации [2]. Таким образом, BPM-системы и их расширения служат для сбора, хранения и распространения организационных знаний, в то время как для представления знаний служат модели бизнес-процессов [3].

© Копп А. М., Орловский Д. Л., 2018
DOI 10.15588/1607-3274-2018-2-12

Одной из распространенных нотаций моделирования бизнес-процессов в BPM-системах является BPMN. Данная нотация предоставляет набор графических инструментов, интуитивно понятных для бизнес-пользователей, но также способных представлять сложную семантику бизнес-процесса. В то же время, доступность нотации для использования нетехническими пользователями может стать причиной возникновения случайных или системных ошибок, допущенных при построении моделей бизнес-процессов в нотации BPMN, связанных с компетентностью автора [4, 5].

Целью работы является разработка подхода к анализу и оптимизации моделей бизнес-процессов в нотации BPMN, используемых системами класса BPM для представления организационных знаний.

1 ПОСТАНОВКА ЗАДАЧИ

Для моделирования бизнес-процессов в нотации BPMN используются следующие основные категории элементов [6]:

1) объекты потока (flow objects) – задача (Task), подпроцесс (Sub-Process), шлюз (Gateway), событие (Event);

2) соединяющие объекты (connecting objects) – поток последовательности (Sequence Flow), поток сообщений (Message Flow), ассоциация (Association);

3) дорожки (swimlanes) – пул (Pool), полоса (Lane);

4) артефакты (artifacts) – объект данных (Data Object), текстовая аннотация (Text Annotation), группа (Group).

В общем виде, модель бизнес-процесса в нотации BPMN является ориентированным графом BG, который описывается кортежем $BG = (V, E, L, l, p)$, причем множество вершин данного графа имеет вид $V = S \cup A \cup G \cup F$ [7].

Для выполнения анализа и последующего совершенствования модели BPMN, граф BG необходимо построить на основе описания бизнес-процесса в формате XPDL, который является международным стандартом и используется для обмена моделями бизнес-процессов между различными информационными системами большинством крупных поставщиков BPM-систем [8, 9].

Построение моделей бизнес-процессов носит субъективный характер, что приводит к возникновению ошибок, связанных с компетентностью аналитика, недостаточно подробным описанием предметной области и т.д. [10]. Таким образом, перед тем, как модель бизнес-процесса будет использована BPM-системой для представления организационных знаний, она должна быть проанализирована с точки зрения наличия недостатков.

2 ОБЗОР ЛИТЕРАТУРЫ

Типичные ошибки, допускаемые пользователями при построении моделей бизнес-процессов в нотации BPMN, рассматриваются в работах [11, 12]. Наиболее часто встречающиеся ошибки построения моделей бизнес-процессов в нотации BPMN, авторы работы [11] выделили в «анти-шаблоны (anti-patterns)», демонстрирующие неверное использование объектов потока и соединяющих объектов:

1) для бизнес-процесса не определено конечное событие (End Event);

2) промежуточное событие (Intermediate Event) исключено из потока бизнес-процесса (Process Flow);

3) для разделения потока бизнес-процесса на несколько маршрутов не используется шлюз;

4) для бизнес-процесса не определено стартовое событие (Start Event);

5) задача не связана с остальным бизнес-процессом;

6) задача не приводит к завершению бизнес-процесса;

7) при моделировании межпроцессного взаимодействия (модель сотрудничества или модель хореографии в спецификации BPMN 1.x) у бизнес-процессов отсутствует стартовое и/или конечное событие.

Для формального описания моделей бизнес-процессов в нотации BPMN с целью их дальнейшего анализа и совершенствования используется математический аппарат теории графов [12–14] и сетей Петри [9, 15–17]. Это не удивительно, поскольку в BPMN бизнес-процессы описываются с помощью объектов потока и соединяющих объектов в формате ориентированных графов [18]. Применение сетей Петри было рассмотрено для анализа моделей бизнес-процессов в работе [19]. В работах [13, 14] анализ моделей бизнес-процессов в нотации BPMN осуществляется с использованием взвешенных ориентированных графов. Кроме того, в работах [15-17] рассматривается применение технологий интеллектуального анализа процессов (Process Mining) для анализа полученных сетей Петри.

В исследовании [20] рассматривается подход к анализу моделей бизнес-процессов в нотации BPMN, предусматривающий вычисление и анализ следующих метрик:

1) длина процесса (process length);

2) объем процесса (process volume);

3) сложность процесса (process difficulty);

4) сложность интерфейса (interface complexity) задачи;

5) коэффициент сетевой сложности;

6) уровень взаимосвязанности (connectivity level) задач процесса.

Помимо [20], метрики сложности бизнес-процесса, описываемого с помощью нотации BPMN, рассматриваются также в исследовании [21]. Однако метрики, предназначенные для оценки сложности моделируемого бизнес-процесса, не отражают особенности построения моделей в нотации BPMN и, следовательно, не могут быть использованы для их анализа с точки зрения наличия недостатков.

В работах [22, 23] рассматривается подход к количественному анализу моделей бизнес-процессов в нотации IDEF0, обобщаемый также и для анализа диаграмм потоков данных, предполагающий расчет коэффициента сбалансированности. Согласно данному подходу, в рамках одной диаграммы не желательны ситуации, когда для различных блоков значительно отличается число соединенных с ними дуг, что, как правило, является признаком ошибок при проектировании и эксплуатации бизнес-процесса [22].

Особенности применения коэффициента сбалансированности для анализа диаграмм потоков данных были рассмотрены в работе [10]. Несмотря на свои преимущества, данный подход также не может быть применен

для анализа моделей бизнес-процессов в нотации BPMN, поскольку он не учитывает особенности различных объектов потока (задачи, подпроцессы, шлюзы, события) и правила их соединения.

3 МАТЕРИАЛЫ И МЕТОДЫ

Для соответствия нотации BPMN, описание графа BG, предложенного в работе [7], необходимо дополнить множеством вершин I, представляющих собой промежуточные события. Таким образом, множество вершин графа BG примет вид $V = S \cup A \cup G \cup F \cup I$.

Основными элементами языка XPDL, с помощью которых описывается бизнес-процесс, являются деятельности (Activity) и переходы (Transition) [18, 24]. Тогда, для построения ориентированного графа BG, необходимо отразить деятельности бизнес-процесса, описанного с помощью XPDL, в вершины графа $v \in V$, а переходы – в направленные ребра графа $(v, w) \in E$. Согласно стандарту XPDL, деятельности используются для представления задач (подпроцессов), событий и шлюзов. Поэтому, при построении множества вершин V графа BG на основе описания процесса в формате XPDL, необходимо учитывать следующие правила:

1) если деятельность содержит элемент Implementation, вершина будет представлять собой задачу (подпроцесс) $v \in A \subset V$;

2) если деятельность содержит элемент Event, содержащий элемент StartEvent, вершина будет представлять собой стартовое событие $v \in S \subset V$;

3) если деятельность содержит элемент Event, содержащий элемент EndEvent, вершина будет представлять собой конечное событие $v \in F \subset V$;

4) если деятельность содержит элемент Event, содержащий элемент IntermediateEvent, вершина будет представлять собой промежуточное событие $v \in I \subset V$;

5) если деятельность содержит элемент Route, вершина будет представлять собой шлюз $v \in G \subset V$.

Для построения множества направленных ребер $(v, w) \in E$ графа BG необходимо анализировать атрибуты переходов «From» и «To», содержащие значения идентификаторов «Id» деятельностей, связанных данными переходами (рис. 1). Поскольку стандарт XPDL основывается на XML, идентификаторы деятельностей и переходов также являются атрибутами данных элементов. Так, например, в случае перехода со значениями «A1» и «A2» для атрибутов «From» и «To» соответственно, ребро графа будет направлено от вершины, соответствующей деятельности со значением идентификатора «A1», к вершине, соответствующей деятельности со значением идентификатора «A2» (рис. 1).

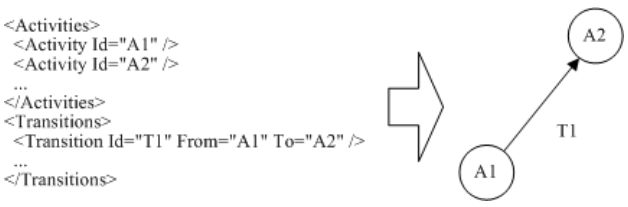


Рисунок 1 – Пример построения фрагмента графа BG

Одним из способов представления ориентированного графа является матрица смежности. Поскольку модель бизнес-процесса в нотации BPMN является ориентированным графом [7, 13–14], для осуществления ее анализа необходимо построить соответствующую матрицу смежности. Необходимо отметить, что в одном документе XPDL могут содержаться определения нескольких бизнес-процессов. Таким образом, в результате применения предлагаемого подхода, для каждого бизнес-процесса будет построен соответствующий ориентированный граф BG_i, $i \in \overline{1,m}$, где m – количество определенных в документе XPDL бизнес-процессов, а также матрица смежности данного графа. Матрица смежности графа BG, описывающего модель бизнес-процесса в нотации BPMN, является квадратной матрицей PM размера n, причем $n = |V|$. Кроме того, нотация BPMN не допускает связывание объектов потока более чем одним соединяющим объектом, поэтому матрица смежности PM будет являться бинарной матрицей.

Для выявления ошибок [11, 12], допускаемых при построении моделей бизнес-процессов в нотации BPMN, необходимо проверить выполнение следующих правил:

1) $|F| > 0$ – для бизнес-процесса должно быть определено конечное событие;

2) $\forall v_i \in I : \sum_{j=1}^{n} PM(i,j) > 0 \wedge \forall v_i \in I : \sum_{j=1}^{n} PM(j,i) > 0$,

$i \in \overline{1,n}$ – промежуточное событие не должно быть исключено из потока бизнес-процесса;

3) $\forall v_i \in A : \sum_{j=1}^{n} PM(i,j) = 1 \wedge \forall v_i \in G : \sum_{j=1}^{n} PM(i,j) > 1$,

$i \in \overline{1,n}$ – для разделения потока бизнес-процесса на несколько маршрутов должен использоваться шлюз;

4) $|S| > 0$ – для бизнес-процесса должно быть определено стартовое событие;

5) $\forall v_i \in A : \sum_{j=1}^{n} PM(j,i) > 0 \wedge \forall v_i \in A : \sum_{j=1}^{n} PM(i,j) > 0$,

$i \in \overline{1,n}$ – задача должна быть связана с остальным бизнес-процессом и приводить к его завершению.

Поскольку один документ XPDL может содержать определения m бизнес-процессов, проверку выполнения правил (1–5) необходимо осуществлять для каждого i-го бизнес-процесса на основе ориентированного графа BG_i и его матрицы смежности PM_i, где $i \in \overline{1,m}$.

Так, например, с учетом правил (1) и (2), при моделировании межпроцессного взаимодействия (хореография/сотрудничество), для каждого i-го бизнес-процесса должно быть определено хотя бы одно стартовое и одно конечное событие, т.е. $\forall BG_i : |S| > 0 \wedge |F| > 0$, $i \in \overline{1,m}$.

Таким образом, в случае наличия ошибок построения модели бизнес-процесса, одно или несколько рассмотренных выше правил построения моделей бизнес-процессов в нотации BPMN (1–5) не будут выполняться. Другими словами, если в модели бизнес-процесса присутствуют ошибки, результатом вычисления функции $y(x_1, x_2, \ldots, x_r)$ будет значение 0:

$$y(x_1, x_2, \ldots, x_r) = x_1 \wedge x_2 \wedge \ldots \wedge x_r = \bigwedge_{j=1}^{r} x_j.$$

На основе изложенного выше материала, сформируем общую схему анализа моделей бизнес-процессов в нотации BPMN (рис. 2).

На основе характеристики структурной избыточности [25], отражающей превышение общего числа связей над минимально необходимым числом связей, а также коэффициента сбалансированности [22, 23], был сформирован коэффициент структурного соответствия модели бизнес-процесса в нотации BPMN правилам построения:

$$CSC = \left(\operatorname{sgn}\left(|S|\right)-1\right)+\left(\operatorname{sgn}\left(|F|\right)-1\right)+\sum_{v_i \in I}\frac{\operatorname{sgn}\left(\sum_{j=1}^{n} PM(i,j)\right)-1}{|I|}+$$

$$+\sum_{v_i \in I}\frac{\operatorname{sgn}\left(\sum_{j=1}^{n} PM(j,i)\right)-1}{|I|}+\sum_{v_i \in A}\frac{\operatorname{sgn}\left(\sum_{j=1}^{n} PM(j,i)\right)-1}{|A|}-$$

$$-\sum_{v_i \in A}\operatorname{sgn}\left(\left|1-\sum_{j=1}^{n} PM(i,j)\right|\right).$$

Значение коэффициента $CSC < 0$ свидетельствует о наличии ошибок построения [11, 12] анализируемой модели бизнес-процесса в нотации BPMN. Таким образом, значения данного коэффициента могут изменяться в интервале $\left[-5-|A|,0\right]$.

Для формирования рекомендаций по устранению обнаруженных ошибок построения модели бизнес-процесса в нотации BPMN, необходимо определить изменения в соответствующем ей ориентированном графе BG, т.е. решить следующую задачу:

$$CSC' = \left(\operatorname{sgn}\left(|S|+\Delta|S|\right)-1\right)+\left(\operatorname{sgn}\left(|F|+\Delta|F|\right)-1\right)+$$

$$+\sum_{v_i \in I}\frac{\operatorname{sgn}\left(\sum_{j=1}^{n} PM(i,j)+\Delta\deg^{-}(v_i)\right)-1}{|I|}+$$

$$+\sum_{v_i \in I}\frac{\operatorname{sgn}\left(\sum_{j=1}^{n} PM(j,i)+\Delta\deg^{+}(v_i)\right)-1}{|I|}+$$

$$+\sum_{v_i \in A}\frac{\operatorname{sgn}\left(\sum_{j=1}^{n} PM(j,i)+\Delta\deg^{+}(v_i)\right)-1}{|A|}-$$

$$-\sum_{v_i \in A}\operatorname{sgn}\left(\left|1-\sum_{j=1}^{n} PM(i,j)+\Delta\deg^{-}(v_i)\right|\right) \to \max,$$

$$|S|+\Delta|S| > 0,\ |F|+\Delta|F| > 0,$$

$$\forall v_i \in I: \sum_{j=1}^{n} PM(i,j)+\Delta\deg^{-}(v_i) > 0,$$

$$\sum_{j=1}^{n} PM(j,i)+\Delta\deg^{+}(v_i) > 0,\ i \in \overline{1,n},$$

$$\forall v_i \in A: \sum_{j=1}^{n} PM(i,j)+\Delta\deg^{-}(v_i) = 1,$$

$$\sum_{j=1}^{n} PM(j,i)+\Delta\deg^{+}(v_i) > 0,\ i \in \overline{1,n}.$$

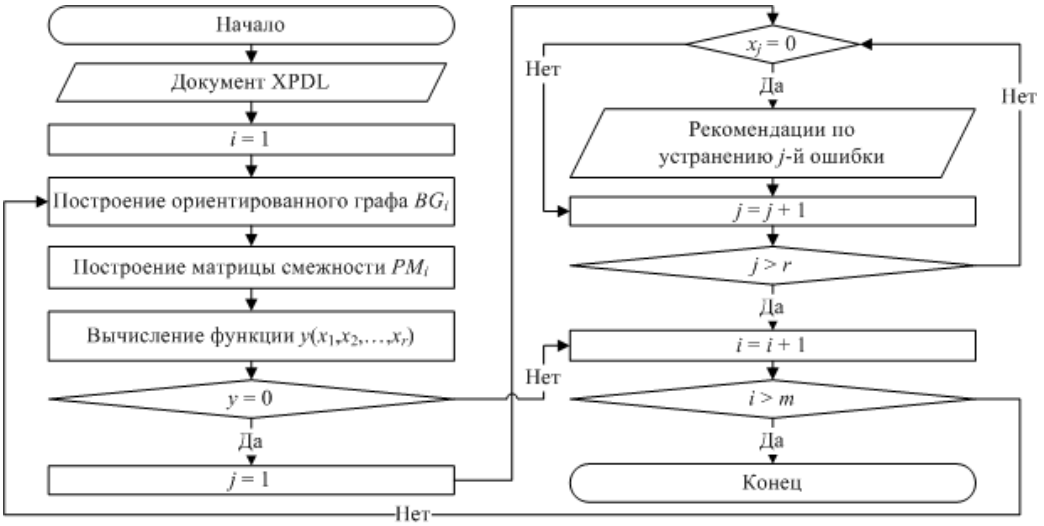


Рисунок 2 – Общая схема анализа моделей бизнес-процессов в нотации BPMN

Применение предложенного подхода позволит определять ошибки, допущенные при построении моделей бизнес-процессов в нотации BPMN, а также формировать рекомендации по их устранению на основе правил (1–5).

4 ЭКСПЕРИМЕНТЫ

В качестве примера рассмотрим диаграмму BPMN, описывающую бизнес-процессы торговой организации (рис. 3). Для построения диаграмм BPMN использовалось бесплатное средство моделирования Bizagi Process Modeler, входящее в состав системы управления бизнес-процессами Bizagi BPM Suite.

На диаграмме представлены бизнес-процессы «Закупка продукции» и «Продажа продукции», для каждого из которых были построены и проанализированы соответствующие ориентированные графы BG_1 и BG_2, а также их матрицы смежности PM_1 и PM_2.

Реализация предложенного подхода была осуществлена с использованием платформы .NET и языка программирования C#. Разработанный инструмент позволяет загружать определения бизнес-процессов, выгруженные из BPM-систем в формате XPDL, определять значения коэффициента структурного соответствия, ошибки построения моделей бизнес-процессов в нотации BPMN, а также формировать рекомендации по их устранению (рис. 4).

5 РЕЗУЛЬТАТЫ

Анализ моделей бизнес-процессов (рис. 3) при помощи разработанного программного средства позволил обнаружить следующие ошибки построения (рис. 5):

1) для бизнес-процесса «Закупка продукции» не определено конечное событие;

2) задача «Прием продукции» не приводит к завершению бизнес-процесса «Закупка продукции»;

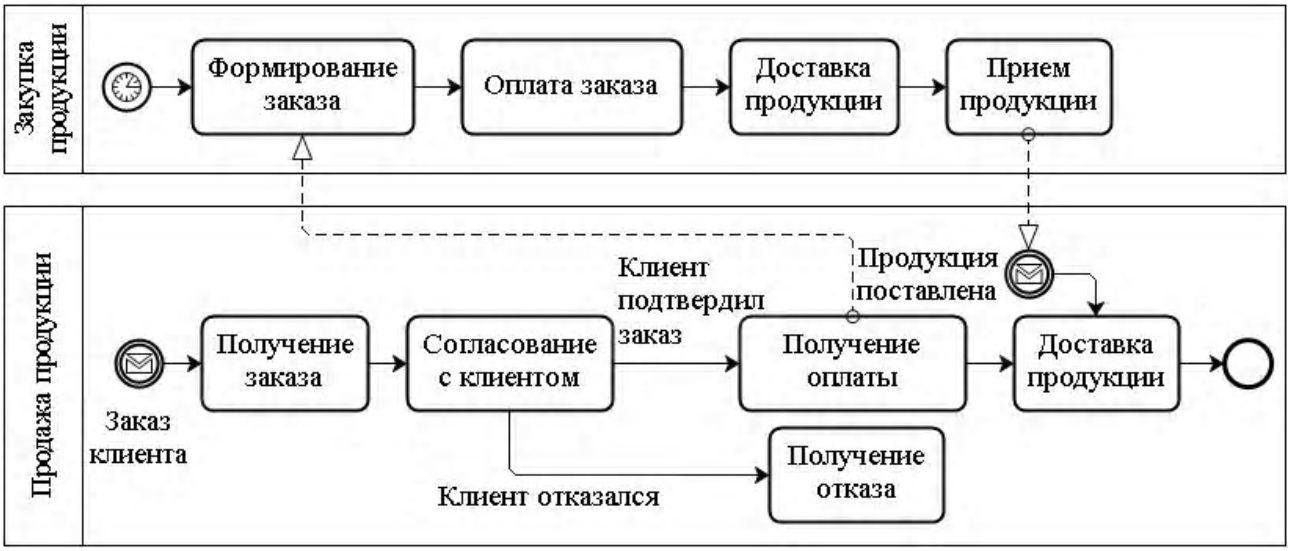


Рисунок 3 – Исходная диаграмма BPMN

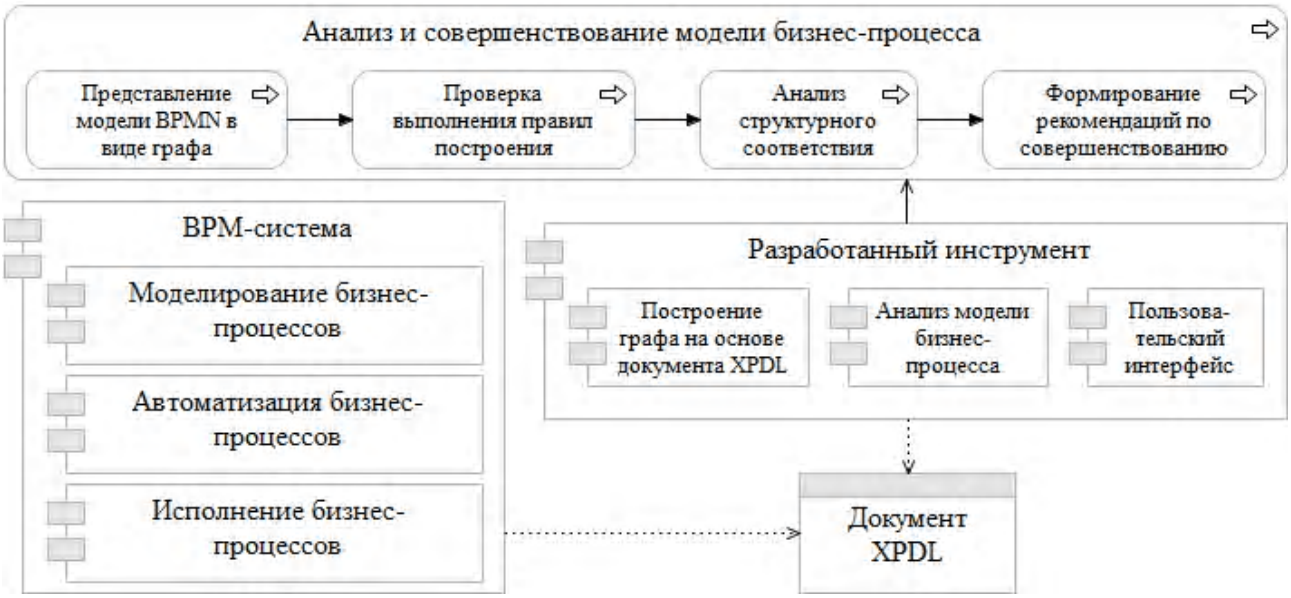


Рисунок 4 – Взаимодействие разработанного инструмента с BPM-системой

3) для бизнес-процесса «Продажа продукции» не определено стартовое событие;

4) промежуточное событие «Продукция поставлена» исключено из потока бизнес-процесса «Закупка продукции»;

5) промежуточное событие «Заказ клиента» исключено из потока бизнес-процесса «Продажа продукции»;

6) для разделения потока бизнес-процесса «Продажа продукции» на несколько маршрутов не используется шлюз;

7) задача «Получение отказа» не приводит к завершению бизнес-процесса «Продажа продукции».

Наличие ошибок, допущенных при построении моделей рассматриваемых бизнес-процессов, подтверждают результаты вычисления функции $y(x_1, x_2, \ldots, x_r)$ (табл. 1).

Значения коэффициента структурного соответствия моделей бизнес-процессов «Закупка продукции» и «Продажа продукции» правилам построения (1–5) моделей BPMN составили $CSC_1 = -2$ и $CSC_2 = -4$ соответственно.

На основе рекомендаций (табл. 2), которые были получены в результате применения предложенного подхода, исходная диаграмма (рис. 3) была преобразована с учетом правил построения (1–5) моделей бизнес-процессов в нотации BPMN.

В результате выполненных преобразований, значения коэффициента структурного соответствия моделей бизнес-процессов «Закупка продукции» и «Продажа продукции» правилам построения (1–5) моделей BPMN составили $CSC'_1 = 0$ и $CSC'_2 = 0$ соответственно.

Модели бизнес-процессов, представленные на преобразованной диаграмме BPMN (рис. 6), лишены ошибок [11, 12] и в дальнейшем могут быть использованы системами класса BPM для сбора, хранения и распространения организационных знаний.

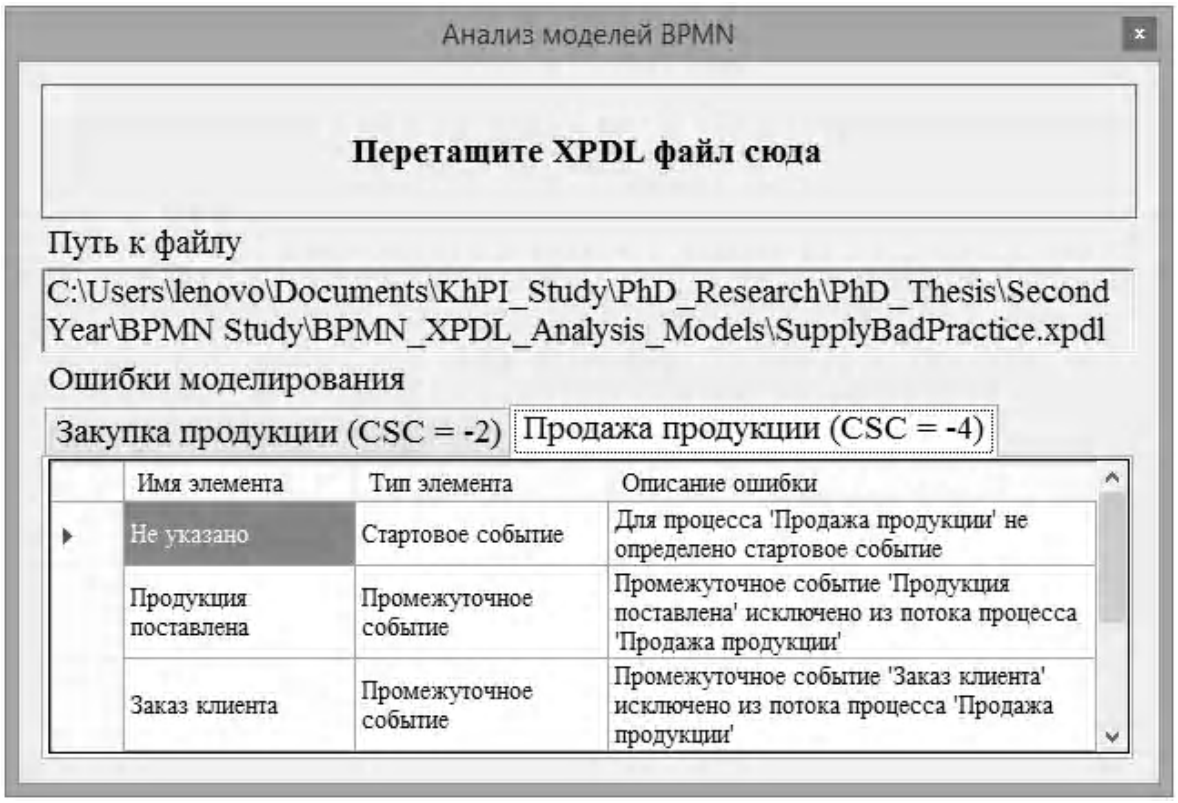


Рисунок 5 – Результаты анализа диаграммы BPMN с использованием разработанного программного средства

Таблица 1 – Результаты анализа моделей бизнес-процессов

Модель бизнес-процесса	x_1	x_2	x_3	x_4	x_5	$y(x_1, x_2, \ldots, x_r)$
Закупка продукции	0	1	0	1	0	$0 \wedge 1 \wedge 0 \wedge 1 \wedge 0 = 0$
Продажа продукции	1	0	0	0	0	$1 \wedge 0 \wedge 0 \wedge 0 \wedge 0 = 0$

Таблица 2 – Рекомендации по устранению ошибок моделей бизнес-процессов

Модель бизнес-процесса	$\Delta\lvert S\rvert$	$\Delta\lvert F\rvert$	$\Delta\deg^+(v_i)$	$\Delta\deg^-(v_i)$
Закупка продукции	0	1	–	$\Delta\deg^-(v_5) = 1$
Продажа продукции	1	0	$\Delta\deg^+(v_1) = 1$ $\Delta\deg^+(v_6) = 1$	$\Delta\deg^-(v_3) = -1$ $\Delta\deg^-(v_5) = 1$

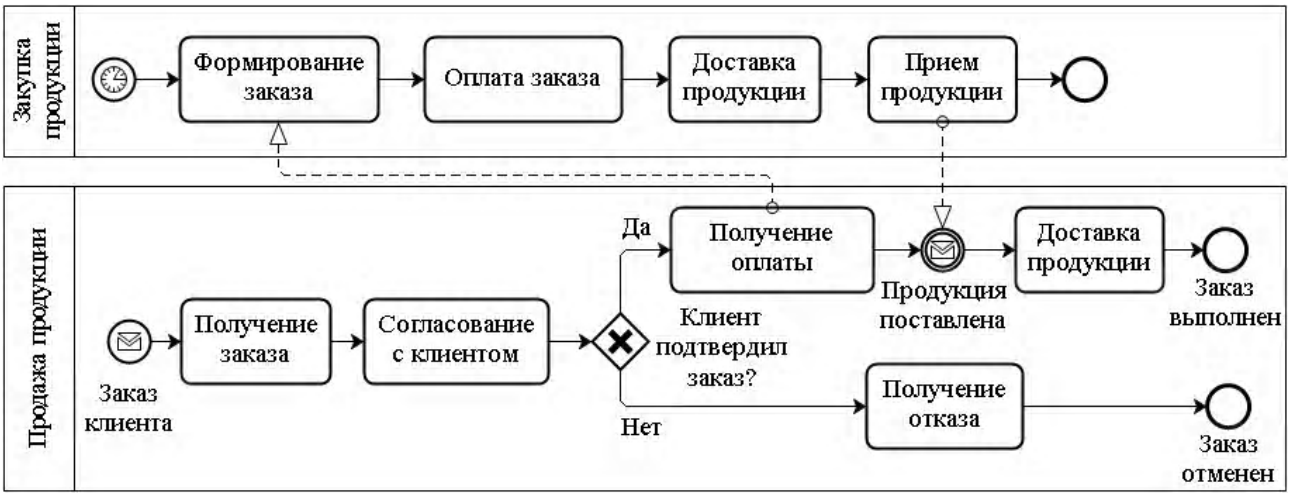


Рисунок 6 – Преобразованная диаграмма BPMN

6 ОБСУЖДЕНИЕ

Применение предложенного подхода, по сравнению с подходами, предполагающими анализ метрик сложности бизнес-процессов [20, 21] и коэффициента сбалансированности моделей бизнес-процессов [22, 23], позволяет учитывать основные особенности построения моделей бизнес-процессов в нотации BPMN. Однако для анализа моделей бизнес-процессов с помощью предложенного подхода требуется выполнение значительно большего объема вычислений. В то же время, благодаря формализации правил построения моделей бизнес-процессов в нотации BPMN (1–5), предложенный подход позволяет количественно оценивать степень соответствия моделей данным правилам, определять ошибки [11, 12], связанные с неверным использованием объектов потока и соединяющих объектов, а также формировать рекомендации по их устранению.

В отличие от существующих подходов, основанных на применении теории графов и сетей Петри [7, 9, 12–18], предложенный подход предполагает построение ориентированного графа, описание которого [7] было дополнено множеством промежуточных событий, на основе документа в формате XPDL, который является международным стандартом обмена моделями бизнес-процессов между различными информационными системами.

В сравнении с разработанным программным средством, реализующим предложенный подход, анализ исходной диаграммы (рис. 3) в Bizagi Process Modeler продемонстрировал отсутствие каких-либо ошибок построения моделей бизнес-процессов. Инструментальные средства (например, Enterprise Architect, Camunda BPM, Activiti Modeler и др.) осуществляют анализ диаграмм BPMN лишь с точки зрения синтаксиса [6], что не позволяет определить ошибки построения моделей бизнес-процессов, в отличие от предложенного подхода.

Недостатком предложенного подхода является использование только формата XPDL, что существенно сужает возможности интеграции разработанного инструмента с другими BPM-системами. Таким образом, в дальнейших исследованиях необходимо рассмотреть использование форматов BPMN 2.0 и BPEL, которые также основаны на XML.

ВЫВОДЫ

Предложен подход к анализу и оптимизации моделей бизнес-процессов в нотации BPMN, используемых системами класса BPM для представления организационных знаний.

Рассмотрено и дополнено формальное представление моделей бизнес-процессов в нотации BPMN с помощью ориентированных графов [7], построение которых предложено осуществлять на основе описаний моделей бизнес-процессов в формате XPDL. Рассмотрены и формализованы правила построения моделей бизнес-процессов в нотации BPMN.

Научная новизна полученных результатов состоит в том, что дальнейшее развитие получил коэффициент сбалансированности моделей бизнес-процессов [21, 22], на основе которого был предложен коэффициент структурного соответствия моделей BPMN правилам построения. Кроме того, предложен формальный подход к формированию рекомендаций по совершенствованию моделей бизнес-процессов в нотации BPMN.

Практическая значимость полученных результатов заключается в том, что на основе предложенного подхода было разработано программное средство, позволяющее определять ошибки, допущенные при построении моделей бизнес-процессов в нотации BPMN, и формировать рекомендации по их устранению. При помощи разработанного программного средства были выполнены контрольные расчеты для моделей бизнес-процессов закупки и продажи продукции.

БЛАГОДАРНОСТИ

Работа выполнена в рамках диссертационного исследования на тему «Модели и информационные технологии процессного управления бизнес-структурами предприятия», проводимого во время обучения в аспирантуре Национального технического университета «Харьковский политехнический институт» по специальности 122 «Компьютерные науки». Результаты исследований также используются в учебном процессе по дисциплине «Инжиниринг и реинжиниринг бизнес-систем» при подготовке студентов образовательно-квалификационного уровня «магистр» по специальности 122 «Компьютерные науки».

СПИСОК ЛИТЕРАТУРЫ

1. Yan Z. A Framework for Business Process Model Repositories [Text] / Z. Yan, P. W. P. J. Grefen // Business Process Management Workshops. – 2010. – Vol. 66. – P. 559–570. DOI: 10.1007/978-3-642-20511-8_51
2. Wasilewski A. Business process management suite (BPMS) market changes 2009 – 2015 [Text] / A. Wasilewski // Information Systems in Management. – 2016. – Vol. 5. – P. 585–592.
3. Curko K. The Role of Business Process Management Systems and Business Intelligence Systems in Knowledge Management [Text] / K. Curko, V. B. Vuksic, A. Loncar // International Journal of Computers And Communications. – 2009. – Vol. 3. – №. 2. – P. 17–24.
4. White S. A. BPMN modeling and reference guide: understanding and using BPMN [Text] / S. A. White, D. Miers. – Future Strategies Inc., 2008. – 225 p.
5. Kluza K. Overview of BPMN model equivalences: towards normalization of BPMN diagrams [Text] / K. Kluza, K. Kaczor // 8th Workshop on Knowledge Engineering and Software Engineering (KESE2012) at the at the biennial European Conference on Artificial Intelligence (ECAI 2012): August. – 2012. – Vol. 28. – P. 38–45.
6. Jordan D. Business Process Model and Notation (BPMN) Version 2.0 [Text] / D. Jordan, J. Evdemon. – Object Management Group, 2011. – 538 p.
7. Braghetto K. R. From business process model and notation to stochastic automata network [Text] / K. R. Braghetto, J. E. Ferreira, J. M. Vincent // Relatório Técnico RTMAC-2011-03, Universidade São Paulo. – 2011. – Vol. 7. – P. 73.
8. Shapiro R. Workflow Management Coalition Workflow StandardProcess Definition Interface-XML Process Definition Language [Text] / R. Shapiro, M. Martin. – The Workflow Management Coalition, 2008. – 87 p.
9. Transforming XPDL to Petri nets [Text] / [H. Zha, Y. Yang, J. Wang et al.] // International Conference on Business Process Management. – Springer, Berlin, Heidelberg, 2007. – P. 197–207. DOI: 10.1007/978-3-540-78238-4_21
10. Корр А. Разработка подхода к анализу и оптимизации диаграмм потоков данных [Text] / A. Kopp, D. Orlovskyi // ScienceRise. – 2017. – Vol. 7. – P. 33–42. DOI: 10.15587/2313-8416.2017.107048.
11. Rozman T. Analysis of most common process modeling mistakes in BPMN process models [Text] / T. Rozman, G. Polancic, R. V. Horvat // 2008 BPM and Workflow Handbook. – 2008. – P. 1–15.
12. BPMN Symbols: The 10 most common mistakes to avoid [Electronic resource]. – Access mode: https://www.heflo.com/blog/process-modeling/bpmn-symbols/
13. Guiding the Creation of Choreographed Processes with Multiple Instances Based on Data Models [Text] / [M. T. Gómez-López, J. M. Pérez-Álvarez, A. J. Varela-Vaca, R. M. Gasca et al.] // International Conference on Business Process Management. – Springer, Cham, 2016. – P. 239–251. DOI: 10.1007/978-3-319-58457-7_18
14. Ontology-Based Heuristics for Process Behavior: Formalizing False Positive Scenarios [Text] / [J. Roa, E. Reynares, M. L. Caliusco, P. Villarreal et al.] // International Conference on Business Process Management. – Springer, Cham, 2016. – P. 106–117. DOI: 10.1007/978-3-319-58457-7_8
15. Dijkman R. M. Semantics and analysis of business process models in BPMN [Text] / R. M. Dijkman, M. Dumas, C. Ouyang // Information and Software technology. – 2008. – Vol. 50. – № 12. – P. 1281–1294. DOI: 10.1016/j.infsof.2008.02.006
16. Dijkman R. M. Formal semantics and analysis of BPMN process models using Petri nets [Text] / R. M. Dijkman, M. Dumas, C. Ouyang // Queensland University of Technology, Tech. Rep. – 2007. – P. 1–30.
17. Kalenkova A. A. Discovering, analyzing and enhancing BPMN models using ProM [Text] / A. A. Kalenkova, W. van der Aalst, M. de Leoni // Proceedings of the BPM Demo sessions 2014 Co-located with 12th International Conference on Business Process Management (BPM2014) Eindhoven, The Netherlands, September 10, 2014. – CEUR-WS. org, 2014. – P. 36–40.
18. Mapping from BPMN-Formed Business Processes to XPDL Business Processes [Text] / [M. Jung, H. S. Kim, M. H. Jo, K. H. Tak, H. S. Cha, J. H. Son] // ICEB. – 2004. – P. 422–427.
19. Копп А. М. Об одном подходе к решению задачи оптимизации структуры бизнес-процессов предприятия [Текст] / А. М. Копп, Д. Л. Орловский // Вісник НТУ «ХПІ». Серія: Системний аналіз, управління та інформаційні технології. – 2015. – № 58. – С. 102–108.
20. Kluza K. Proposal of square metrics for measuring business process model complexity [Text] / K. Kluza, G. J. Nalepa // Computer Science and Information Systems (FedCSIS), 2012 Federated Conference on. – IEEE, 2012. – P. 919–922.
21. Cardoso J. Business process control-flow complexity: Metric, evaluation, and validation [Text] / J. Cardoso // International Journal of Web Services Research. – 2008. – Vol. 5. – №. 2. – P. 49.
22. Анализ предметной области оборота рабочей конструкторской документации [Электронный ресурс]. – Режим доступа: http://www.kb-ametist.com/rus/18_publish/10.pdf
23. Замятина О. М. Моделирование систем : учебное пособие [Текст] / О. М. Замятина. – Томск: Изд-во ТПУ, 2009. – 204 с.
24. Van der Aalst W. M. P. Patterns and xpdl: A critical evaluation of the xml process definition language [Text] / W. M. P. van der Aalst // BPM Center report BPM-03-09, BPMcenter. org. – 2003. – P. 1–30.
25. Широков Л. А. Исследование систем управления : учебное пособие [Текст] / Л. А. Широков. – М. : МГИУ, 2010. – 168 с.

Статья поступила в редакцию 02.12.2017.
После доработки 10.01.2017.

Копп А. М.[1], Орловський Д. Л.[2]

[1]Аспірант кафедри програмної інженерії та інформаційних технологій управління Національного технічного університету «Харківський політехнічний інститут», Харків, Україна

[2]Канд. техн. наук, доцент, доцент кафедри програмної інженерії та інформаційних технологій управління Національного технічного університету «Харківський політехнічний інститут», Харків, Україна

ПІДХІД ДО АНАЛІЗУ ТА ОПТИМІЗАЦІЇ МОДЕЛЕЙ БІЗНЕС-ПРОЦЕСІВ В НОТАЦІЇ BPMN

Актуальність. В даний час все більше організацій представляють знання про свою діяльність з точки зору бізнес-процесів. У той же час, побудова моделей бізнес-процесів має суб'єктивний характер, що призводить до виникнення помилок, пов'язаних з компетентністю аналітика, або, наприклад, з недостатньо докладним описом предметної області. Таким чином, перед тим, як моделі бізнес-процесів будуть використані для збору, зберігання та поширення організаційних знань, вони повинні бути проаналізовані с точки зору наявності різного роду недоліків.

Мета. Метою роботи є розробка підходу до аналізу та оптимізації моделей бізнес-процесів, які використовуються для представлення організаційних знань.

Метод. У роботі було розглянуто та доповнено формальне представлення моделей бізнес-процесів на основі орієнтованих графів. Були розглянуті та формалізовані правила побудови моделей бізнес-процесів. На основі характеристики структурної надмірності графа, а також коефіцієнта збалансованості моделей бізнес-процесів, був запропонований коефіцієнт структурної відповідності моделей бізнес-процесів правилам побудови. З урахуванням формалізованих правил побудови та коефіцієнта структурної відповідності, було запропоновано формальний підхід, призначений для формування рекомендацій щодо вдосконалення моделей бізнес-процесів.

Результати. З використанням розробленого програмного забезпечення, яке реалізує запропонований підхід до аналізу та оптимізації моделей бізнес-процесів, були виконані контрольні розрахунки для моделей бізнес-процесів закупівлі і продажу продукції. Для даних моделей були визначені помилки побудови та сформовано рекомендації щодо їх усунення. Перетворені моделі позбавлені помилок та в подальшому можуть бути використані для представлення організаційних знань.

Висновки. Запропонований підхід дозволяє визначати помилки, допущені при побудові моделей бізнес-процесів, та формувати рекомендації щодо їх усунення.

Ключові слова: моделювання бізнес-процесів, нотація моделювання, аналіз, оптимізація.

Kopp A. M.[1], Orlovskyi D. L.[2]

[1]Postgraduate student of Department of Software Engineering and Management Information Technologies, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

[2]Ph.D., Associate Professor, Associate Professor of Department of Software Engineering and Management Information Technologies, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

AN APPROACH TO BPMN BASED BUSINESS PROCESS MODELS ANALYSIS AND OPTIMIZATION

Context. Today, most organizations represent knowledge about their activities from a business processes point of view. At the same time, business process modeling is a subjective activity, which leads to errors associated with an analyst competence or i.e. with a poorly detailed subject area description. Thus, business process models should be analyzed for various kinds of shortcomings before they used for collecting, storing, and sharing of the organizational knowledge.

Objective. The objective of the study is to develop an approach to the analysis and optimization of business process models used to represent the organizational knowledge.

Method. A formal representation of business process models, based on directed graphs, has been considered in this study. Business process models design rules have been considered and formalized. A coefficient of business process models structural conformity to the design rules, which is based on the graphs redundancy characteristic and the business process models balancing coefficient, has been proposed. A formal approach to forming business process models enhancement recommendations has been proposed with considering the formalized design rules and the coefficient of structural conformity.

Results. Calculations for models of products purchase and sale business processes, using developed software which is implemented the proposed approach to business process models analysis and optimization, have been performed. The design shortcomings of the considered models have been defined and the enhancement recommendations formed. Transformed models are error free and further can be used to represent organizational knowledge.

Conclusions. The proposed approach allows to define the business process models design shortcomings and to develop recommendations for their elimination.

Keywords: business process modeling, modeling notation, analysis, optimization.

REFERENCES

1. Yan Z., Grefen P. W. P. J. A Framework for Business Process Model Repositories [Text], *Business Process Management Workshops*, 2010, Vol. 66, pp. 559–570. DOI: 10.1007/978-3-642-20511-8_51
2. Wasilewski A. Business process management suite (BPMS) market changes 2009 – 2015 [Text], *Information Systems in Management*, 2016, Vol. 5, pp. 585–592.
3. Curko K., Vuksic V. B., Loncar A. The Role of Business Process Management Systems and Business Intelligence Systems in Knowledge Management [Text], *International Journal of Computers And Communications*, 2009, Vol. 3, No. 2, pp. 17–24.
4. White S. A., Miers D. BPMN modeling and reference guide: understanding and using BPMN [Text]. *Future Strategies Inc.,* 2008, 225 p.
5. Kluza K., Kaczor K. Overview of BPMN model equivalences: towards normalization of BPMN diagrams [Text], *8th Workshop on Knowledge Engineering and Software Engineering (KESE2012) at the at the biennial European Conference on Artificial Intelligence (ECAI 2012): August*, 2012, Vol. 28, pp. 38–45.
6. Jordan D., Evdemon J. Business Process Model and Notation (BPMN) Version 2.0 [Text]. Object Management Group, 2011, 538 p.
7. Braghetto K. R., Ferreira J. E., Vincent J. M. From business process model and notation to stochastic automata network [Text], *Relatório Técnico RTMAC-2011-03, Universidade São Paulo,* 2011, Vol. 7, pp. 73.
8. Shapiro R., Martin M. Workflow Management Coalition Workflow StandardProcess Definition Interface-XML Process Definition Language [Text]. The Workflow Management Coalition, 2008, 87 p.
9. Zha H., Yang Y., Wang J. et al. Transforming XPDL to Petri nets [Text], *International Conference on Business Process Management.* Springer, Berlin, Heidelberg, 2007, pp. 197–207. DOI: 10.1007/978-3-540-78238-4_21
10. Kopp A., Orlovskyi D. Razrabotka podhoda k analizu i optimizacii diagramm potokov dannyh [Text], *ScienceRise,* 2017, Vol. 7, pp. 33–42. DOI: 10.15587/2313-8416.2017.107048.
11. Rozman T., Polancic G., Horvat R. V. Analysis of most common process modeling mistakes in BPMN process models [Text], *2008 BPM and Workflow Handbook*, 2008, pp. 1–15.
12. BPMN Symbols: The 10 most common mistakes to avoid [Electronic resource]. – Access mode: https://www.heflo.com/blog/process-modeling/bpmn-symbols/
13. Gómez-López M. T., Pérez-Álvarez J. M., Varela-Vaca A. J., Gasca R. M. et al. Guiding the Creation of Choreographed Processes with Multiple Instances Based on Data Models [Text], *International Conference on Business Process Management*. Springer, Cham, 2016, pp. 239–251. DOI: 10.1007/978-3-319-58457-7_18
14. Roa J., Reynares E., Caliusco M. L., Villarreal P. et al.] Ontology-Based Heuristics for Process Behavior: Formalizing False Positive Scenarios [Text], *International Conference on Business Process Management.* Springer, Cham, 2016, pp. 106–117. DOI: 10.1007/978-3-319-58457-7_8
15. Dijkman R. M., Dumas M., Ouyang C. Semantics and analysis of business process models in BPMN [Text], *Information and Software technology*, 2008, Vol. 50, No. 12, pp. 1281–1294. DOI: 10.1016/j.infsof.2008.02.006
16. Dijkman R. M., Dumas M., Ouyang C. Formal semantics and analysis of BPMN process models using Petri nets [Text], *Queensland University of Technology, Tech. Rep*, 2007, pp. 1–30.
17. Kalenkova A. A., van der Aalst W., M. de Leoni Discovering, analyzing and enhancing BPMN models using ProM [Text], *Proceedings of the BPM Demo sessions 2014 Co-located with 12th International Conference on Business Process Management (BPM2014) Eindhoven, The Netherlands, September 10*, 2014, CEUR-WS. org, 2014, pp. 36–40.
18. Jung M., Kim H. S., Jo M. H., Tak K. H, Cha H. S., Son J. H. Mapping from BPMN-Formed Business Processes to XPDL Business Processes [Text], *ICEB,* 2004, pp. 422–427.
19. Kopp A. M., Orlovskyi D. L. Ob odnom podhode k resheniju zadachi optimizacii struktury biznes-processov predprijatija [Tekst], *Visnyk NTU «HPI». Serija: Systemnyj analiz, upravlinnja ta informacijni tehnologii',* 2015, No. 58, pp. 102–108.
20. Kluza K., Nalepa G. J. Proposal of square metrics for measuring business process model complexity [Text], *Computer Science and Information Systems (FedCSIS), 2012 Federated Conference on.,IEEE,* 2012, pp. 919–922.
21. Cardoso J. Business process control-flow complexity: Metric, evaluation, and validation [Text], *International Journal of Web Services Research,* 2008, Vol. 5, No. 2, pp. 49.
22. Analiz predmetnoj oblasti oborota rabochej konstruktorskoj dokumentacii [Jelektronnyj resurs]. Rezhim dostupa: http://www.kb-ametist.com/rus/18_publish/10.pdf
23. Zamjatina O. M. Modelirovanie sistem: uchebnoe posobie [Tekst]. Tomsk, Izd-vo TPU, 2009, 204 p.
24. Van der Aalst W. M. P. Patterns and xpdl: A critical evaluation of the xml process definition language [Text], *BPM Center report BPM-03-09, BPMcenter. org*, 2003, pp. 1–30.
25. Shirokov L. A. Issledovanie sistem upravlenija: uchebnoe posobie [Tekst]. Moscow, MGIU, 2010, 168 p.

УДК 681.326

Miroschnyk M. A.[1], Pakhomov Y. V.[2], Shkil A. S.[3], Kulak E. N.[3], Kucherenko D. Y.[3]

[1]Dr. Sc., Professor, Head of the Department of Information Technologies, Ukrainian State University of Railway Transport, Kharkiv, Ukraine
[2]Assistant of Department of Gas and Heating Systems Operation, Kharkiv National University of Urban Economy, Kharkiv, Ukraine
[3]Ph.D, Associate Professor of Design Automation Department, Kharkiv National University of Radioelectronics, Kharkiv, Ukraine

DESIGN AUTOMATION OF EASY-TESTED DIGITAL FINITE STATE MACHINES

Context. The relevance of the work is to provide minimal additional hardware costs during design automation of easy-tested digital devices, which are represented by models of control finite state machines on hardware description languages.

Objective. To develop procedures of models' constructing of easy-tested control finite state machines on hardware description languages and estimate hardware costs for different methods of hardware redundancy introduction to HDL-models of finite state machines.

Method. The introduction to HDL-models of control finite state machines, which are presented in the form of the FSM template, hardware redundancy (additional fragments of the HDL-code), providing the forcing setting of finite state machine into an arbitrary state without the use of synchronizing sequences. For implementation of this approach, the method of FSM's state table extending is applied, which ensures the mode of bypassing of all nodes of FSM' state diagram in the diagnostic mode.

Results. Simulation of extended VHDL-models of the control FSM using Active-HDL confirmed the operability of this approach. Synthesis of these models using CAD XILINX ISE confirmed the receipt of testable structures and showed the minimum hardware costs for the method associated with the extension of the state table, in comparison with the organization of the shift register in the Scan Path mode.

Conclusions. The task of computer-aided design of testable control finite state machine on the basis of application of FSM' setting methods into given state is solved in the work. The optimal way of the setting organization into an arbitrary state of the control FSM is to expand the state table, which improves the controllability of FSM' states and leads to the structure' transformation of their HDL-models into easy-tested ones.

The scientific novelty of the work is the transformation of control FSM' models on hardware description languages, which is realized by introduction of the additional symbol to the state table, providing the settings of the FSM into an arbitrary state without the use of synchronizing sequences.

The practical significance of obtained results is to confirm the optimality, in terms of additional hardware costs, of the setting method of the control FSM into an arbitrary state by introducing the additional symbol into the state table.

Keywords: control finite state machine, state table, shift register, scanned path, hardware description language, CAD, Active-HDL, XILINX ISE.

TERMINOLOGY

A – state alphabet of FSM;
CLK – synchronization signal;
FSM – finite state machine;
HDL – hardware description language;
MUX – multiplexor;
TDI – test data input;
VHDL – very high speed integrated circuits HDL;
X – input alphabet of FSM;
Y – output alphabet of FSM;
$\delta(a_i, x_k) = a_j$ – transition function of FSM;
$\lambda(a_i, x_k) = y_\alpha$ – output function of FSM;
DE – diagnostic experiment;
CP – combinational part;
PLD – programmable logic device;
CAD – computer-aided design system;
ST – state table;
DD – digital device.

INTRODUCTION

Computer-aided design of digital devices (DD) based on specification which is represented in the form of hardware description languages (HDL). Structural methods of diagnosis, which are oriented to the detection of constant faults, are not effective. At the same time take into consideration that modern DD have a big size which makes the application of functional methods practically impossible.

Specialized data processing and control DD as a rule described by finite state machine (FSM). The forms of FSM representation are state table (ST) and state diagram (SD).

Due to the complexity of the diagnostic experiments (DE) under FSM, various methods of testability assurance were proposed. Namely, modification of FSM models providing for the introduction of hardware redundancy, both at structural level (additional inputs and outputs to ensure the simplicity of DE) and at functional level (additions and changes in the functional description of the FSM, namely state diagram).

Thanks to this approach testable FSM becomes easy-tested FSM. For easy-tested FSM, tasks of test diagnosis are simply solved within the limits of the established design costs. By reducing the cost of one or more main factors, determining the complexity of the test diagnosis, you can make easy-tested FSM.

1 PROBLEM STATEMENT

At a high-level design of digital control devices based on finite state machines, the form of specification' representation is the state table or the state diagram.

The form of FSM description using VHDL language is the FSM template, i.e. the code where the transitions function and outputs function are collected into separate processes, and the assignment of the new state is carried out in a special process associated with synchronization.

During design of testable control FSM, the hardware redundancy, which provides easy-testability, is advisable

© Miroschnyk M. A., Pakhomov Y. V., Shkil A. S., Kulak E. N., Kucherenko D. Y., 2018
DOI 10.15588/1607-3274-2018-2-13

to introduce at the initial stage of design (HDL coding). An easy-tested FSM is a finite state machine for which it is possible to construct a diagnostic experiment of minimum length by ensuring that the machine will go to any state for a minimum number of cycles.

Thus, a functional model of an abstract FSM in the form of a ST or a SD is given and on its basis a VHDL model is constructed in the form of an FSM template. It is necessary to consider different ways of introducing hardware redundancy into the VHDL model to provide easy- testability, and to choose the optimal method in terms of additional hardware costs. Hardware redundancy in VHDL-models is provided by adding to the HDL-code additional conditional operators that guarantee the construction of a scanned path in the memory part of the FSM, which is confirmed by the results of automated synthesis. In terms of additional hardware costs the optimal method is that, which provides minimal additional hardware costs by Quine' assessment of the circuit gate equivalence, which is synthesized into FPGA chip in automatic mode by the CAD tools.

2 LITERATURE REVIEW

Testability increasing methods by introducing hardware redundancy into the circuit implementation are sufficiently developed and widely used in the design. In the classical work about the development and classification of structural methods of testable design [1], the concepts of "controllability" and "observability" were introduced as the basis for structural testability evaluation. The principles of shift registers organization in the memory part of DD and the construction of a scanned path based on them are described. In addition, modifications of known structural methods for constructing tests using scanning methods are proposed. Practical recommendations on the use of structural methods to increase the testability of DD and the construction of diagnostic algorithms based on them are outlined in [2, 3].

In the classical article [4], methods of built-in self-test (BIST) are considered as the basis for ensuring the testability of designed devices. The application of BIST technology in designing of digital systems on programmable logic devices (PLD) is shown. In [5], ideas for testable design based on Boundary Scan technology using the JTAG port were further developed. Standards of boundary scanning IEEE 1149.x, IEEE 1500 and IEEE 1532, as well as the standard of testable design IEEE 1687 are considered. Technologies of testable design and testing of digital-analog circuits are proposed in [6]. Methods of functional decomposition of circuits that provide automated construction of testable digital-analog structures are considered.

Functional methods of the testable design of DD are considered in detail in [7]. For the digital automata presented in the form of ST, concepts of the diagnosed and definitely diagnosed classes of FSM are introduced and methods for bringing state tables of FSMs to the specified classes are proposed. Procedures for carrying out of DE with FSM using setting, synchronizing, diagnostic and transfer sequences are considered. In addition, the idea of improving the testability of the FSM is justified by introducing hardware redundancy by expanding the input alphabet, the output alphabet and the states alphabet. The problems of DE organization by bypassing all nodes and arcs of the state diagram are considered in [8]. Estimates of the length and completeness of DE with different variants of the diagram bypassing are given. The procedures for increasing the FSM' testability by equivalent transformations and expanding the input alphabet are proposed.

In [9], structural models of FSM are considered that allow us to use the values of output variables as codes of internal states. This approach was used during the synthesis of Mealy FSM on PLD, which allowed to reduce the cost of implementation by 1.5-2 times. Also, the proposed method of state coding allows to reduce the power consumption of the circuit implementations of FSMs [10].

In [11], the methods of testable design of DD based on the PLD and the organization of DE in the networks of cellular automata are described. The problems of automating the DE realization under HDL-models of FSM in the form of a FSM template are considered in [12]. The procedures for the DE organization in the verification system of CAD for PLD are considered and procedures for localization of design errors in HDL models are proposed.

Thus, the actual task is the development of procedures for design automation of the of easily-tested FSM, which are represented in hardware description languages, the comparison additional hardware costs for structural and functional methods of increasing testability based on comparison of synthesis results of testable HDL-models using CAD system.

3 MATERIALS AND METHODS

A finite state machine as an abstract model of DD with memory is determined by the six parameters $W =< X, A, Y, \delta, \lambda >$, where $X = \{x_1, x_2, ..., x_m\}$ – input alphabet; $A = \{a_1, a_2, ..., a_n\}$ – set of FSM' states; $Y = \{y_1, y_2, ..., y_r\}$ – output alphabet; $\delta(a_i, x_k) = a_j$ – the transition function, which is defined as $\lambda(z_i, x_k) = y_\alpha$ – the output function of the FSM.

The circuit is testable if the following parameters:

– generating procedures of test sets;

– estimating of their effectiveness;

– implementation of the test diagnostic [1] subject to the compliance financial costs, time costs in specified limits.

DE under FSM is a process of setting the input sequences, observing the corresponding output sequences, and making conclusions which is based on these observations. Experiments are classified by 3 groups: the identification of the FSM' states, the identification of the FSM' input sequence and FSM identification with n states which is different from all other FSM with the same number of states. The diagnostic experiment, during which all nodes of the diagram are bypassed, is used to detect design errors in models of control FSMs, which are represented by the state table or by the state diagram.

Thus, the easy-tested FSM is FSM which can be set to any state less than for n cycles, where n is the number of FSM states without using synchronizing sequences [13].

Let's consider conditions which are necessary to meet to create easy-tested FSMs based on FSM template:

– opportunity to switch FSM to the testing mode and vice versa at any cycle of the machine operation;

– setting FSM to any state in (n–1) clock cycles in the testing mode, where n is the number of FSM' states; Hamiltonian cycle can be organized for any of states;

– opportunity to build the easy-tested FSM in an automated mode by CAD tools.

The base of structural testable design is the method of scan path which allows to solve the problem of testability by reducing the complexity of the circuit structure.

The idea of scan path method is following:

1) first of all, we have to check memory elements;

2) next step is to check the combinational part (CP) for the possibility of setting internal variables to any state (regardless of their previous state).

A simple technique for creating the scan path is placing a multiplexer 2-to-1 just ahead of each flip-flop. The multiplexer control line, connected to scan mode selection signal. Scanning mode gives the opportunity to set memory units to any state by putting a sequence of signals to the input of the scanned data.

Let's consider the VHDL-model of Mealy FSM, which is represented by the state table (fig. 1a) and the state diagram (fig. 1b).

In the future (unless otherwise stated) we will consider FSMs whose controlled by generator of synchronization pulses (CLK), i.e. FSM has not only inputs (x_1, . . . , x_n), but also CLK input, which is used to set synchronization sequence. Therefore, the input signal on the transition (a_m, a_s) corresponding to the path $a_m X(a_m, a_s) – Y(a_m,a_s)a_s$ will be not $X(a_m,a_s)$, but the conjunction $CLK \cdot X(a_m, a_s)$. Thus, for $X(a_m, a_s) = 1$ (unconditional transition), the FSM operates only from the synchronization pulse. Usually CLK signal is not added to arcs of state diagram.

Let's consider a structural method to organize the scanning path in HD- models of finite state machines by manually setting the shift register directly in the model.

The difference between this model and FSM template is codes of FSM' states have to be set by force.

The fragment of the VHDL-model of Mealy FSM in the form of FSM template [12] with a shift register in the memory part is shown in Fig. 2. The signal A = 1 set the standard mode of FSM operation, A = 0 – the shift mode in the register from the external input TDI (test data input) in the direction to the higher discharges. FSM states in this model are coded forcibly according tp the number of state (during states' internal signals declaration) using the minimum number of flip-flops (3 in this case) to store the code of the state.

```
architecture FSM_MX of FSM_Mealy_MX is
signal  State, NextState: STD_LOGIC_vector (2
downto 0);
signal a1: STD_LOGIC_vector (2 downto 0):=″001";
signal a2: STD_LOGIC_vector (2 downto 0):=″010";
signal a3: STD_LOGIC_vector (2 downto 0):=″011";
signal a4: STD_LOGIC_vector (2 downto 0):=″100";
signal a5: STD_LOGIC_vector (2 downto 0):=″101";
begin
Sreg0_CurrentState: process (Clk, Reset)
begin
     if Reset='1' then State <= a1;
     elsif Clk'event and Clk = '1' then
          if A='1' then State <= NextState;
     else State <= State(1 downto 0) & TDI;
          end if;
     end if;
end process;
```

Figure 2 – The fragment of the VHD-model of Mealy FSM with shift register (FSM_Mealy_MX)

The time diagram of Mealy FSM with the shift register in the memory part and the external input TDI (FSM_Mealy_MX) is shown in Fig. 3.

Let's analyze the result of simulation:

– signal *State* presented in decimal form;

– from 0 ns till 500 ns, the FSM operates in the normal mode (A=1);

– from 500 ns till 2000 ns, the FSM operates in the shift mode (A=0).

After synthesis of Mealy FSM with shift register and input TDI (FSM_MX) Xilinx ISE will create a circuit, the fragment of which is shown in Fig. 4.

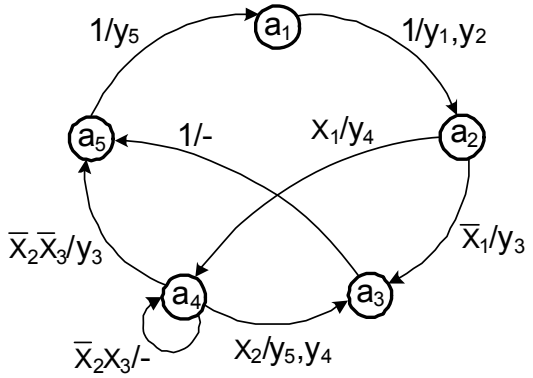
The Boolean equation for trigger D2 is:

$$D_2 = \{[(x_3 \vee \overline{x_3})\, y_3\, \overline{x_2} \vee (x_1\, y_1 \vee y_2)]\, A \vee Q_2\, \overline{A}\} \overline{\mathrm{Re}\, set}.$$

Actually we got multiplexor relatively to the control signal A, which fully corresponds to the circuit structure with the path scanning.

a	1	x_1	$\overline{x_1}$	x_2	$\overline{x_2}x_3$	$\overline{x_2 x_3}$
a_1	a_2/ y_1,y_2					
a_2		a_3/y_3	a_4/y_4			
a_3	a_5/–					
a_4				a_3/ y_4,y_5	a_4/–	a_5/ y_3
a_5	a_1/ y_5					

a



b

Figure 1 – The state table and the state diagram of Mealy FSM

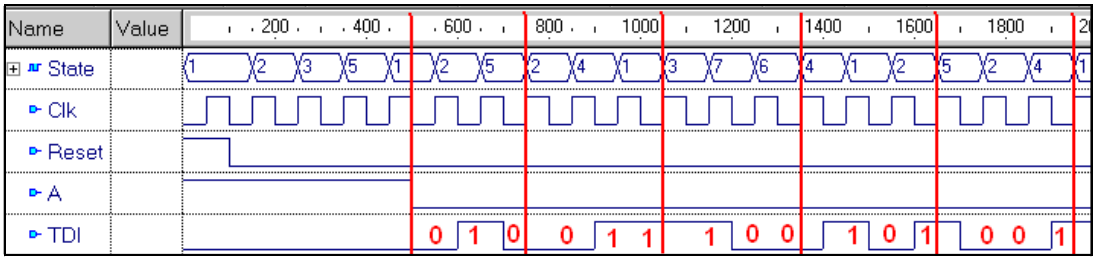


Figure 3 – Timing diagram of Mealy FSM from 500 ns till 2000 ns with shift register (FSM_Mealy_MX)

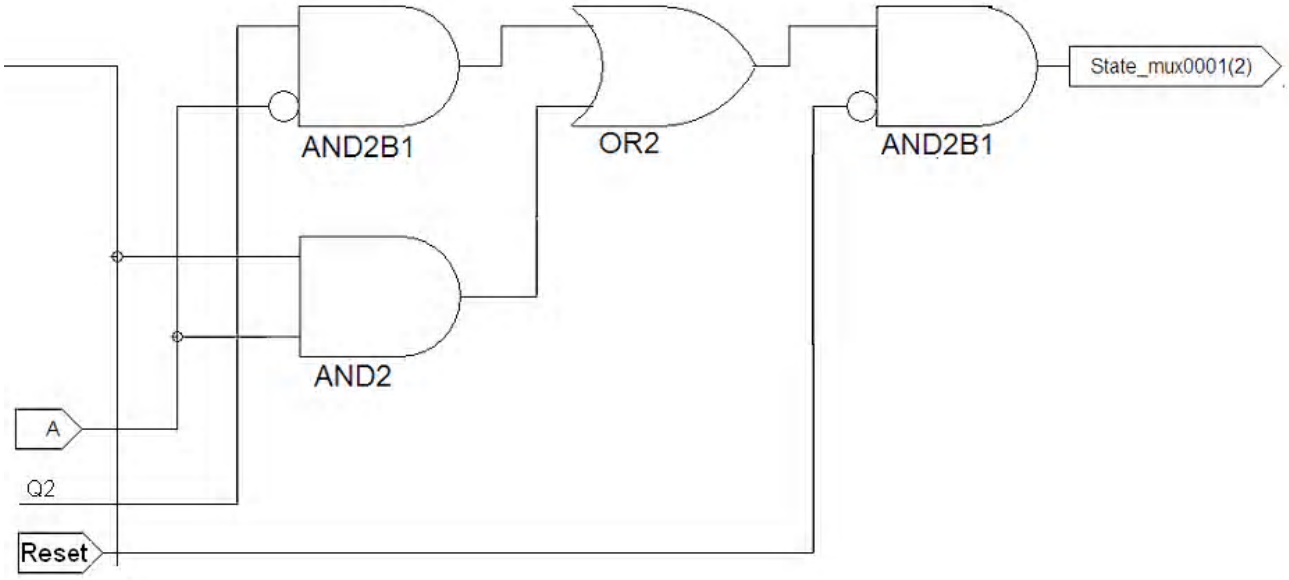


Figure 4 – The circuit' fragment of Mealy FSM with the shift register (FSM_Mealy_MX)

Rotate shift register is used to implement Hamiltonian cycle of diagram's bypassing. Operation modes of Mealy FSM with the shift register are following: signal A = 1 corresponds to the mode of the "normal" operation of the FSM (parallel load), A = 0 – the mode of rotate shift to the left. The FSM encoding style is "one-hot" code (unitary code): a1 (00001), a2 (00010), a3 (00100), a4 (01000), a5 (10000). Thus, for each state there is a trigger. Such model allows transferring into new state during each new clock cycle, and moreover, to bypass all states during the number of cycles equal to the number of states in a shift mode.

```
if A='1'   then State <= NextState;
else   State <= State(3 downto 0) & State(4);
end if;
```

The proposed methodology for converting FSM into easy-tested by introducing shift registers solves assigned problem, but also has disadvantages.

1. TDI input is used for organizing the BoundaryScan structure, but in the standard operating mode of the FSM, the control of this input is not provided.
2. During FSM scanning an open shift register creates difficulties for organizing a Hamiltonian cycle.
3. Using arbitrary encoding style for FSM's states, a shift register reduces the speed in scanning mode.
4. Although this approach effectively solves the problem of the Hamiltonian cycle construction of diagram's bypassing, but due to unitary state coding leads to significant hardware costs and performance decreasing.

Thus, the proposed methodology does not solve all assigned tasks. The solution lies in the field of functional transformation of state table by expanding the input alphabet to ensure testability.

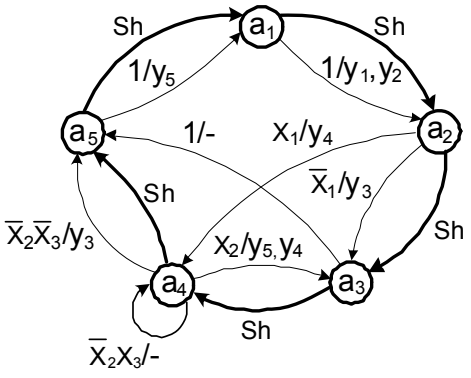
During a functional approach to designing (analyzing) of digital devices which specified by the finite state machine $W =< X, A, Y, \delta, \lambda >$, increasing the testability of an FSM is possible only by expanding the input alphabet X, the alphabet of states A, or the output alphabet Y [7]. On the one hand, this leads to the introduction of hardware redundancy into the circuit of DD, and on the other hand it ensures the preservation of the algorithm for the operation of the FSM, which is usually specified by the state table or state diagram.

In [14] authors of this paper proposed and theoretically justified the introduction of the additional column (symbol) in the state table of FSM by providing for this symbol the transition function of the state diagram of the shift register, which allows you to set the FSM to any given state.

Let's extend the ST of Mealy FSM (Fig. 1) by adding the column Sh (Shift). If Sh=1, the FSM operates in the setting mode in any given state, and if Sh=0, the FSM realizes the given algorithm. The expanded ST and the modified SD are shown in fig. 5.

a	1	x_1	$\overline{x_1}$	x_2	$\overline{x_2}x_3$	$\overline{x_2}\overline{x_3}$	Sh
a_1	a_2/ y_1,y_2						a_2/–
a_2		a_3/ y_3	a_4/ y_4				a_3/–
a_3	a_5/–						a_4/–
a_4				a_3/ y_4,y_5	a_4/–	a_5/y_3	a_5/–
a_5	a_1/ y_5						a_1/–

a



b

Figure 5 – Extended ST and modified SD of Mealy FSM with Sh signal

4 EXPIREMENTS

To test the effectiveness of the proposed method of increasing the testability of the FSM by expanding the input alphabet and introducing an additional column into the state diagram, let's consider the modification of the VHDL-model shown in fig. 5.

The fragment of VHDL architecture of the easy-tested Mealy FSM with the additional signal Sh is shown in Fig. 6. The main advantage of this model is the possibility of transition from the settings mode to the operation mode of the FSM at any moment of time.

```
– Formation of combinational part of FSM – the
description of states' transitions
Sreg0_NextState: process (State, Sh, x1, x2,
x3)
begin
y1 <= '0'; y2 <= '0'; y3 <= '0'; y4 <= '0'; y5
<= '0';
  case State is
when a1 => if Sh = '1' then NextState <= a2;
else NextState <= a2; y1 <= '1'; y2 <= '1';
end if;
when a2 => if Sh = '1' then NextState <= a3;
elsif x1='1' then NextState <= a4; y4<='1';
else NextState <= a3; y3 <= '1';
end if;
when a3 => if Sh = '1' then NextState <= a4;
else NextState <= a5;
end if;
when a4 => if Sh = '1' then NextState <= a5;
elsif x2='1' then NextState<=a3; y4<='1'; y5<=
'1';
elsif x3='1' then NextState <= a4;
else NextState <= a5; y3 <= '1';
end if;
when a5 => if Sh = '1' then NextState <= a1;
else  NextState <= a1; y5 <= '1';
end if;
when others => NextState <= a1;
end case;
end process;
```

Figure 6 – The architecture of the VHDL-model of the easy-tested Mealy FSM (FSM_Mealy_Sh)

The simulation result of the VHDL model of the easy-tested FSM (FSM_Mealy_Sh), which provides the bypass of the state diagram on the Hamiltonian cycle, is shown in Fig. 7. If Sh = 0 (from 0 ns till 400 ns), the FSM operates in the standard mode, if Sh = 1 (from 400 ns till 950 ns) – bypassing of all states cyclically (Hamilton cycle).

Let's note one feature of this waveform is the absence of output signals in the bypassing mode of transitions of state diagram (Hamiltonian cycle). This is due to the fact that the output signals y_i of Mealy FSM are associated with the arcs of the state diagram, and the introduction of additional arcs Sh doesn't provide the presence of any output signal.

Of course, when constructing of the VHDL model of the easy-tested FSM, each arc Sh could be matched with an additional output signal, but this would lead to an expansion of the output alphabet and to additional hardware costs, which does not correspond to the formulation of the problem.

The result of the VHDL model synthesis with the signal Sh (FSM_Mealy_Sh) is shown in Fig. 8. Matching signals in the CAD XILINX ISE: Sh – In0, x1 – In1, x2 – In2, x3 – In3, Reset – Rst, Clk – Clk, Q2 " FSM_FFd2 (internal signal from the output of flip-flop D2), Q3 " FSM_FFd3 (internal signal from the output of flip-flop D3), D2 " FSM_FFd2:In (internal signal to the input of flip-flop D2).

The complete equation of the excitation function for discharge D2 will be:

$$D_2 = Q_2Q_3x_3\overline{x_2}\,\overline{Sh} \vee \overline{Q_2}Q_3x_1\overline{Sh} \vee Q_2Q_3x_2\overline{Sh} \vee$$
$$\vee\, \overline{Q_2}Q_3\overline{x_1}\,\overline{Sh} \vee (Q_2\overline{Q_3} \vee \overline{Q_2}Q_3)Sh\,.$$

From this expression it is simple to do the conclusion that with respect to the signal Sh, the multiplexor was received. This is proves the synthesis of easy-tested circuit with has possibility of shift register implementation in the memory part of the FSM.

5 RESULTS

Let's consider the different variants of the circuit synthesis of the easy-tested FSM in terms of additional hardware costs, which can be estimated by Quine's criterion. Quine costs are defined as the total number of inputs of all gates in the circuit, since the number of inputs of the gate is proportional to the number of transistors in it.

Analyses of results are shown in Table 1. As can be seen from the table, in terms of hardware costs the optimal

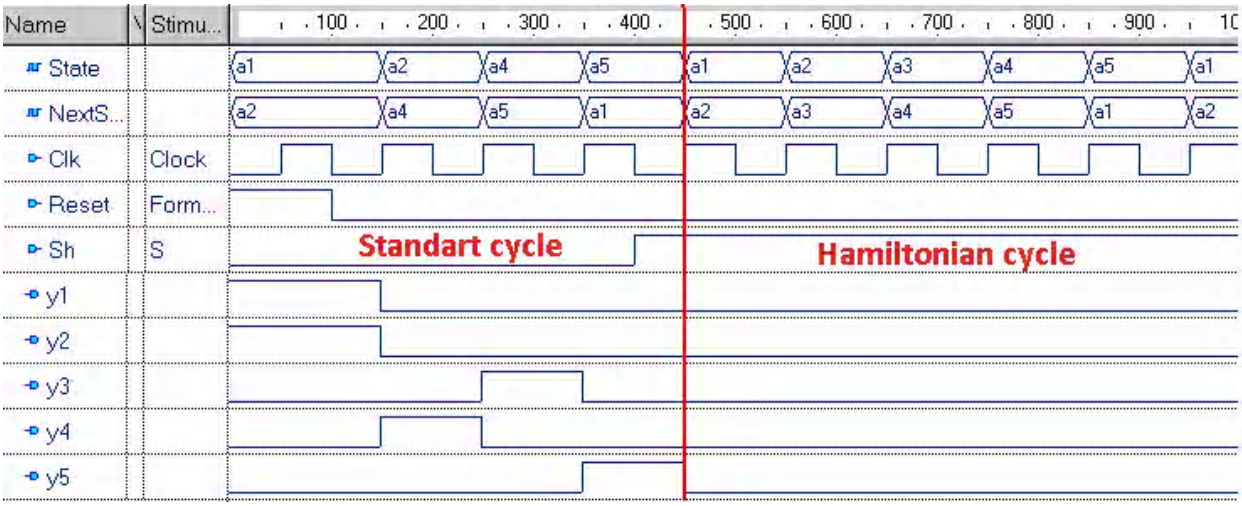


Figure 7 – The time diagram of Mealy FSM with the signal Sh (FSM_Mealy_Sh)

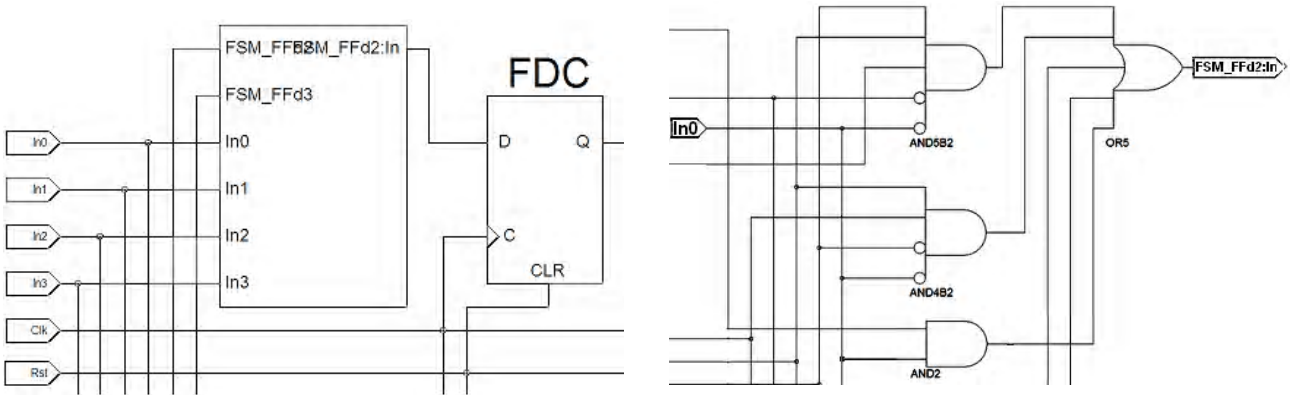


Figure 8 – The fragment of the synthesized circuit of Mealy FSM with the signal Sh (FSM_Mealy_Sh)

Table 1 – Hardware costs for circuits of easy-tested Mealy FSM

№	Type of FSM	Project name in Active-HDL	Number of flip-flops	Total costs by Quine
1	Mealy, standard	FSM_Mealy	3	110
2	Mealy with shift register	FSM_Mealy_MX	3	137
3	Mealy with shift reg., one-hot encoding style	FSM_MX_unit	5	193
4	Mealy with additional column in state table	FSM_Mealy_Sh	3	130

is the introduction of an additional column in the ST of the FSM, which increases the controllability of the FSM' states. Hardware costs in this case increase by 20–25%, depending on the type of machine.

6 DISCUSSION

This work is the first in a series of works related to the development of functional methods of automated testable design of the DD, which is represented in the form of FSM using hardware description languages. In existing CAD systems based on hardware description languages, such as Active-HDL or Riviera from Aldec Inc., there are software modules with a visual interface for automated generation of HDL models of finite state machines based on the SD (State Diagram Editor) and the conduction of diagnostic experiment on bypassing of nodes and arcs of the SD (ASFTEST) [12]. It would be advisable to supplement these modules with procedures for constructing easy-tested HDL-models of FSM.

Thus, the prospect of these researches consists in the development of a software module with a visual interface for entering the state diagram of the FSM, which will implement the proposed ideas related to the expansion of the input alphabet of the ST. This module can be integrated into simulation system of HDL-models Active-HDL and Riviera from Aldec Inc. If we talk about further theoretical researches, they can be associated with increasing testability by expanding the output alphabet and the states alphabet in models of FSM, and also by developing methods of automating testable encoding of FSM's states. A separate area of research can be the development of designing methods for easy-tested FSM, which are optimal from the point of view of time parameters of conducting DE.

CONCLUSIONS

The introduction of hardware redundancy underlies the method of automated design of the easy-tested FSM, which is the subject of this article. The form of FSM model in VHDL is FSM template. Additional fragments of the VHDL code ensure the forced setting of the FSM into an arbitrary state without synchronizing sequences. The shift register

in the memory part of the FSM was considered for path scanning. Diagnosis mode involves the mode of bypassing of all nodes of the FSM' state diagram, which is provided by the proposed method state table expansion.

Due to the application of this method, the controllability of the FSM states, and consequently the testability, is increased. Active-HDL confirms the operability of this approach by simulation of the extended VHDL-models of the FSM.

CAD XILINX ISE confirmed the receipt of testable structures and showed the minimum hardware costs for the method associated with the extension of the state table, in comparison with the organization of the shift register in the Scan Path mode.

The scientific novelty of the work consists of the further development of the testability increasing method of FSM, by expanding the input alphabet, which gave the possibility to automatize of the design process of easy-tested FSM using HDL.

The practical significance of the obtained results is to develop the procedures for adding redundancy and expanding the input alphabet of HDL-models by adding additional conditional operators in the HDL-code, which is conduct to an arbitrary state.

Developed procedures can be applied during the development of an additional program module for CAD, which will automatically generate the HDL code of the easy-tested FSM.

ACKNOWLEDGMENTS

The work was carried out within the framework of the state budget research topic of the Kharkov National University of Radioelectronics "Cyber-physical system Smart cloud management of transport" (state registration number 0115U-000712 from 04.03.2015, order of the Ministry of Education of Ukraine № 105 from 09.02.15) on the basis of the Design automation department.

REFERENCES

1. Беннеттс Р. Дж. Проектирование тестопригодных логических схем : пер. с англ. / Р. Дж. Беннеттс. – М. : Радио и связь, 1990. – 176 с.
2. Пархоменко П. П. Основы технической диагностики (Оптимизация алгоритмов диагностирования, аппаратурные средства) / П. П. Пархоменко, Е. С. Согомонян ; под ред. П. П. Пархоменко. – М. : Энергия, 1981. – 320 с.
3. Горяшко А. П. Проектирование легко тестируемых дискретных устройств: идеи, методы, реализация / А. П. Горяшко // Автоматика и телемеханика. – 1984. – № 7. – С. 5–35.
4. Stroud C. E. A designer's guide to built-in self-test / Charles E. Stroud. – Kluwer Academic Publishing, 2002. – 319 p.
5. Городецкий А. Введение в технологии JTAG и DFT. Тестирование в технологиях граничного сканирования и тестопригодное проектирование / Ами Городецкий. – Palmarium Academic Publishing, Germany, 2012. – 308 с.
6. Mosin S. Methodology to Design-For-Test Automation for Mixed Signal Integrated Circuit / S. Mosin // Proceedings of the International Symposium. EWDTS'2013, September 27–30, Rostov on Don, Russia, 2013. – P. 178–183.
7. Тоценко В. Г. Алгоритмы технического диагностирования цифровых устройств / В. Г. Тоценко. – М. : Радио и связь, 1985. – 240 с.
8. Богомолов А. М. Контроль и преобразования дискретных автоматов / А. М. Богомолов, И. С. Грунский, Д. В. Сперанский. – К. : Наукова думка, 1975. – 175 с.
9. Solov'ev V. V. Minimization of mealy finite-state machines by using the values of the output variables for state assignment / V. V. Solov'ev // Journal of Computer and Systems Sciences International. – January 2017.– Volume 56, Issue 1. – P. 96–104.
10. Solov'ev V. V. Minimization of Power Consumption of Finite State Machines by Splitting Their Internal States / V. V. Solov'ev, T. N. Grzes // Journal of Computer and Systems Sciences International.– 2015. – Vol. 54, No. 3. – P. 367–374.
11. Мирошник М. А. Проектирование диагностической инфраструктуры вычислительных систем и устройств на ПЛИС: монография / М. А. Мирошник. – Х. : ХУПС, 2012. – 188 с.
12. Шкиль А.С. Автоматизация поиска ошибок проектирования в HDL-моделях конечных автоматов / А. С. Шкиль, Г. П. Фастовец, А. С. Серокурова // АСУ и приборы автоматики. – 2014. – Вып. 168. – С. 43–52.
13. Miroshnyk M. A. Design automation of testable finite state machines / [М. А. Miroshnyk, Д. Е. Кучеренко, Ю. В. Пахомов и др.] // 15th IEEE EAST-WEST DESIGN & TEST SYMPOSIUM (EWDTS-2017). Харьковский национальный университет радиоэлектроники. – 2017. – P. 203–208
14. Бережная М. А. Синхронизирующие последовательности в конечных детерминированных автоматах / М. А. Бережная // Вестник НТУ «ХПИ». – 2008. – № 57. – С. 7–15.
15. Методы синтеза легкотестируемых цифровых автоматов / [М. А. Мирошник (Бережная), Ю. В. Пахомов А. С. Гребенюк, И. В. Филиппенко] // Інформаційно-керуючі системи на залізничному транспорті. – 2016. – № 5. – С. 28–39.

Article was submitted 22.09.2017.
After revision 06.11.2017.

Мірошник М. А.[1], Пахомов Ю. В.[2], Шкіль О. С.[3], Кулак Е. М.[3], Кучеренко Д. Ю.[3]

[1]Д-р техн. наук, професор, зав. кафедрою інформаційних технологій Українського державного університету залізничного транспорту, Харків, Україна

[2]Асистент кафедри експлуатації газових і теплових систем Харківського національного університету міського господарства, Харків, Україна

[3]Канд. техн. наук, доцент кафедри автоматизації проектування обчислювальної техніки Харківського національного університету радіоелектроніки, Харків, Україна

АВТОМАТИЗАЦІЯ ПРОЕКТУВАННЯ ЛЕГКОТЕСТОВАНИХ ЦИФРОВИХ АВТОМАТІВ

Актуальність роботи полягає в забезпеченні мінімальних додаткових апаратурних витрат при автоматизованому проектуванні легкотестованих цифрових пристроїв, представлених моделями кінцевих автоматів, що управляють, на мовах опису апаратури.

Мета роботи. Розробити процедури побудови моделей легкотестованих автоматів, що управляють, на мовах опису апаратури і оцінити апаратурні витрати для різних способів надання апаратурної надмірності в HDL-моделі автоматів.

Метод. Внесення в HDL-моделі кінцевих автоматів, що управляють, представлених у формі автоматного шаблону, апаратурної надмірності (додаткових фрагментів HDL-коду), які забезпечують примусову установку автомата в довільний стан без використання синхронізуючих послідовностей. Для реалізації цього підходу застосований метод розширення таблиці переходів-виходів автомата, який забезпечує режим обходу усіх вершин графа переходів автомата (станів) в режимі діагностування.

Результати. Моделювання розширених HDL-моделей автомата, що управляє, засобами Active-HDL підтвердило працездатність цього підходу. Синтез цих моделей інструментальними засобами автоматизованого проектування XILINX ISE підтвердив отримання тестопридатних структур і показав мінімальні апаратурні витрати для методу, пов'язаного з розширенням таблиці переходів-виходів, в порівнянні з організацією зсувного регістра в режимі Scan Path.

Висновки. У роботі вирішено завдання автоматизованого проектування тестопридатних автоматів, що управляють, на основі застосування методів установки автоматів в заданий стан. Оптимальним з точки зору апаратурних витрат способом організації

установки в довільний стан автоматів, що управляють, є розширення таблиці переходів-виходів, яке підвищує керованість станів автомата і призводить до перетворення структури їх HDL-моделей в легкотестовані.

Наукова новизна роботи полягає у подальшому розвитку методів підвищення тестопридатності кінцевих автоматів за рахунок розширення вхідного алфавіту в HDL-моделях у формі автоматного шаблону, що дало можливість автоматизувати процес проектування легкотестованих автоматів з використанням мов опису апаратури.

Практична цінність отриманих результатів полягає у розробці процедур внесення надлишковості і розширення вхідного алфавіту в HDL-моделях кінцевих автоматів у формі автоматного шаблону шляхом внесення додаткових умовних операторів у HDL-код, які забезпечують встановлення автомату у довільний стан. Розроблені процедури можуть бути застосовані при розробці додаткового програмного модуля САПР цифрових пристроїв, який буде в автоматизованому режимі формувати HDL-код легкотестованого кінцевого автомату.

Ключові слова: автомат, що управляє, таблиця переходів-виходів, зсувний регістр, сканований шлях, мова опису апаратури, САПР, Active-HDL, XILINX ISE.

Мирошник М. А.[1], Пахомов Ю. В.[2], Шкиль А. С.[3], Кулак Э. Н.[3], Кучеренко Д. Е.[3]

[1]Д-р техн. наук, профессор, зав. кафедрой информационных технологий Украинского государственного университета железнодорожного транспорта, Харьков, Украина

[2]Ассистент кафедры эксплуатации газовых и тепловых систем Харьковского национального университета городского хозяйства, Харьков, Украина

[3]Канд. техн. наук, доцент кафедры автоматизации проектирования вычислительной техники Харьковского национального университета радиоэлектроники, Харьков, Украина

АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ ЛЕГКОТЕСТИРУЕМЫХ ЦИФРОВЫХ АВТОМАТОВ

Актуальность работы состоит в обеспечении минимальных дополнительных аппаратурных затрат при автоматизированном проектировании легкотестируемых цифровых устройств, представленных моделями конечных управляющих автоматов на языках описания аппаратуры.

Цель работы. Разработать процедуры построения моделей легкотестируемых управляющих автоматов на языках описания аппаратуры и оценить аппаратурные затраты для разных способов введения аппаратурной избыточности в HDL-модели автоматов.

Метод. Внесение в HDL-модели конечных управляющих автоматов, представленных в форме автоматного шаблона, аппаратурной избыточности (дополнительных фрагментов HDL-кода), обеспечивающих принудительную установку автомата в произвольное состояние без использования синхронизирующих последовательностей. Для реализации данного подхода применен метод расширения таблицы переходов-выходов автомата, который обеспечивает режим обхода всех вершин графа переходов автомата (состояний) в режиме диагностирования.

Результаты. Моделирование расширенных HDL-моделей управляющего автомата средствами Active-HDL подтвердило работоспособность данного подхода. Синтез данных моделей инструментальными средствами автоматизированного проектирования XILINX ISE подтвердил получение тестопригодных структур и показал минимальные аппаратурные затраты для метода, связанного с расширением таблицы переходов-выходов, по сравнению с организацией сдвигового регистра в режиме Scan Path.

Выводы. В работе решена задача автоматизированного проектирования тестопригодных управляющих автоматов на основе применения методов установки автоматов в заданное состояние. Оптимальным с точки зрения аппаратурных затрат способом организации установки в произвольное состояние управляющих автоматов является расширение таблицы переходов-выходов, которое повышает управляемость состояний автомата и приводит к преобразованию структуры их HDL-моделей в легкотестируемые.

Научная новизна работы состоит в дальнейшем развитии метода повышения тестопригодности конечных автоматов за счет расширения входного алфавита в HDL-моделях в форме автоматного шаблона, что дало возможность автоматизировать процесс проектирования легкотестируемых автоматов с использованием языков описания аппаратуры.

Практическая ценность полученных результатов заключается в разработке процедур внесения избыточности и расширения входного алфавита в HDL-моделях конечных автоматов в форме автоматного шаблона путем внесения дополнительных условных операторов в HDL-код, которые обеспечивают установку автомата в произвольное состояние. Разработанные процедуры могут быть применены при разработке дополнительного программного модуля САПР цифровых устройств, который будет в автоматизированном режиме формировать HDL-код легкотестируемого конечного автомату.

Ключевые слова: управляющий автомат, таблица переходов-выходов, сдвиговый регистр, сканируемый путь, язык описания аппаратуры, САПР, Active-HDL, XILINX ISE.

REFERENCES

1. Bennetts R. G. Proektirovanie testoprigodnyh logicheskih shem: per. s angl. Moscow, Radio i svjaz', 1990, 176 p.
2. Parhomenko P. P., Sogomonjan E. S.; pod red. Parhomenko P. P. Osnovy tehnicheskoj diagnostiki (Optimizacija algoritmov diagnostirovanija, apparaturnye sredstva). Moscow, Jenergija, 1981, 320 p.
3. Gorjashko A. P. Proektirovanie legko testiruemyh diskretnyh ustrojstv: idei, metody, realizacija, *Avtomatika i telemehanika*, 1984, No. 7, pp. 5–35.
4. Stroud C. E. A designer's guide to built-in self-test. Kluwer Academic Publishing, 2002, 319 p.
5. Gorodetsky A. Introduction to JTAG and DFT technology. Testing in edge scanning technologies and testable design. Palmarium Academic Publishing, Germany, 2012, 308 p.
6. Mosin S. Methodology to Design-For-Test Automation for Mixed Signal Integrated Circuit, *Proceedings of the International Symposium. EWDTS'2013, September 27–30, Rostov on Don*. Russia, 2013, pp.178–183.
7. Tocenko V. G. Algoritmy tehnicheskogo diagnostirovanija cifrovyh ustrojstv. Moscow, Radio i svjaz', 1985, 240 p.
8. Bogomolov A. M., Grunskij I. S., Speranskij D. V. Kontrol' i preobrazovanija diskretnyh avtomatov. Kiev, Naukova dumka, 1975, 175 p.
9. Solov'ev V. V. Minimization of mealy finite-state machines by using the values of the output variables for state assignment, *Journal of Computer and Systems Sciences International*, 2017, January, Volume 56, Issue 1, pp. 96–104.
10. Solov'ev V. V., Grzes T. N. Minimization of Power Consumption of Finite State Machines by Splitting Their Internal States, *Journal of Computer and Systems Sciences International*, 2015, Vol. 54, No. 3, pp. 367–374.
11. Miroshnik M. A. Proektirovanie diagnosticheskoj infrastruktury vychislitel'nyh sistem i ustrojstv na PLIS: monografija. Xar'kov, HUPS, 2012, 188 p.
12. Shkil' A. S., Fastovec G. P., Serokurova A. S. Avtomatizacija poiska oshibok proektirovanija v HDL-modeljah konechnyh avtomatov, *ASU i pribory avtomatiki*, 2014, Vol. 168, pp. 43–52.
13. Miroshnyk M. A., Kucherenko D. E., Paxomov Yu. V., German E'. E., Shkil' A. S., Kulak E'. N. Design automation of testable finite state machines, *15th IEEE EAST-WEST DESIGN & TEST SYMPOSIUM (EWDTS-2017)*. Xar'kovskij nacional'nyj universitet radioe'lektroniki, 2017, pp. 203–208.
14. Berezhnaja M. A. Sinhronizirujushhie posledovatel'nosti v konechnyh determinirovannyh avtomatah, *Vestnik NTU "HPI"*, 2008, No. 57, pp. 7–15.
15. Miroshnik (Berezhnaja) M. A., Pahomov Ju. V., Grebenjuk A. S., Filippenko I. V. Metody sinteza legkotestiruemyh cifrovyh avtomatov, *Informacijno-kerujuchi sistemi na zaliznichnomu transporti*, 2016, No. 5, pp. 28–39.

UDC 629.05

Panin V. V.[1], Doronin V. V.[2], Aleynikov V. M.[3]

[1]Dr. Sc., Professor, Rector, State University of Infrastructure and Technologies, Kyiv, Ukraine
[2]PhD, Professor at the Department of Technical Systems and Control Processes in Navigation, State University of Infrastructure and Technologies, Kyiv, Ukraine
[3]Postgraduate student, State University of Infrastructure and Technologies, Kyiv, Ukraine

APPLICATION OF THE SYSTEM ANALYSIS OF IMPLEMENTATION OF THE INSTRUMENTAL METHOD OF NAVIGATION ON INLAND WATERWAYS OF UKRAINE

Context. The task of applying the scientific approach to the introduction of a modern navigation system on inland waterways of Ukraine is considered to solve, first of all, weakly structured problems characterized by the presence of an uncertainty factor and containing both formalized and unformulated elements that are not translated into the language of mathematics.

Objective – development of the methodology of system analysis at the stage of introduction of the instrumental method of navigation on inland waterways of Ukraine, which guarantees high quality of modeling in accordance with the exact criteria of navigation and economy.

Method. An approach to data processing is proposed, which is based on the study of the properties of the system with her further formalization. A model of the layout of the system based on electronic chart systems is developed, characterizing the fulfillment of the assigned tasks with an increased probability of safe navigation. The features of the state of the system using a mathematical model in the form of an undirected graph are considered. An algorithm for realizing the functional stability of the system is proposed. The results of the experiments make it possible to recommend the proposed method for introducing the instrumental method of navigation on inland waterways of Ukraine.

Results. The parameters of experimental research of the functioning of the stages of the functional stability of the coastal infrastructure and ship navigation equipment have been developed.

Conclusions. In the process of research a set of procedures that contain elements of scientific novelty are used and are aimed at formulating a problem situation, proposing a set of alternatives that are compared with performance criteria using the system stability factors.

Keywords: safety of navigation, electronic chart system, instrumental navigation method, algorithm, multiparametric optimization, graph, linguistic analysis, modeling.

NOMENCLATURE

$A = \{a_i\}$ – a set of system elements;

$C_с$, $C_о$, $C_в$ – complexity of the system: system, own, mutual;

D – a set of elements of the system and the links between them;

D_{min} – minimum distance to dangers;

def ($pE–A$) – power polynomial of Laplace variable p of order n;

E – deviation of the changed state of the system from the calculated one;

$F_в$ – function of system outputs;

Fc – transition function of system state;

$f(x; y)$ – function of position of vessel motion;

$F(x)$ – the component that determines the random nature of changes in system performance;

h_{10} – zero depth at the first water measuring point;

k_1, k_2 – coefficients of motion and safety during its turning circle;

M – mean–square error of vessel location;

$P_{бп}$ – probability of safe navigation of water transport;

p – probability of events x;

($pE–A$), $S(p)$ – matrices: characteristical, adjoint;

Q – a set of links between elements of the system;

$S(\varepsilon)$ – the area of infinitely small deviation during vessel motion;

T_t – evolution operator;

V – vector space x_0, y_0;

V_4 – actual length of the lock–chamber;

V_5 – dimensions of lock approach canal;

$V_{5–2}$ – radius of the turning circle

V_6, $V_{6–1}$ – the width of navigable channel under bridges and the height of a bridge above water;

V_7 – the smallest distance of the wires of electric power lines from the water level;

$v^{+(-)}, e^{+(-)}, f^{+(-)}$ –increased (reduced) values {V, E, F} of the constituent elements of the system;

υ – system memory parameter;

$X(t)$ – function of inputs of the system (control algorithm);

x_0, y_0 – display coordinates at calculated state of the system;

(x_0^1, y_0^1) – display coordinates at changed state of the system;

x_i – actual value of the adjustable parameter;

Δx – inconsistency between x_{req} i x_i;

X_{req} – required value of the adjustable parameter;;

$Y(t)$ – function of system outputs;

$y_{01вп}$ – the value of water level at the first water measuring point;

x, Y – values on the scale from the domain Y;

a, b – any value in the plurality of linear admissible transformations, a> 0;

Z_{im} – specific properties of the system element;

Z_i – state of the system element;

Z, $Z(t)$ – state of the system;

Z_1, Z_2 – state of subsystems;

© Panin V. V., Doronin V. V., Aleynikov V. M., 2018
DOI 10.15588/1607-3274-2018-2-14

$Z_{сд}$ – a set of admissible states of the system;

Δ_3 – draft of the vessel;

Δ_{4-1} – margin of locks width V_{4-1};

Δ_6, Δ_{6-1} – maximum deviation of motion under bridges and height margin;

Δ_7 – Minimum margin at the height of wires of power lines;

$\delta y_{01вп}$ – deviation from zero depth at the first water measuring point;

Ω – vessel motion.

INTRODUCTION

In this century, river transport becomes one of the most important elements of transportation, and its development is a powerful catalyst for economic growth and increase of the state's defense capability. An effective mechanism of operation of water vehicles in the current conditions of intellectualizing decision-making system is the transition to an navigation instrumental procedure instead of existing pilot method and further improving of motion safety of water transport on inland waterways.

Particularly noteworthy are the most important applicative results of the proposed method, such as the possibility of automated display of an individual navigable lane for the vessel, depending on the chosen draft and the actual water level, and in the future, the refusal of installing coastal and floating structures of navigation equipment and the refusal of pilot principle of ship navigation.

The system of providing the navigation instrumental procedure is a coherent practice for collecting, integrating, exchanging, presenting and analyzing information on vessels and coastal services by means of information technologies to increase the safety of navigation and environmental protection on inland waterways of Ukraine. This is a system for the integrated use of information that provides and maintains electronic cartographic means of navigation [1]. Thus, there is combining of information from numerous sources to find more accurate and reliable data about the situation than the results obtained from these sources separately. In addition, it should be noted that the introduction of navigation instrumental procedure will allow to shut down the ship navigation on the fairway channels in particular periods and follow only selected vessels quickly.

In this case, system analysis is used to solve, firstly, semistructured problems of implementing the system of providing the navigation instrumental procedure, namely problems, the composition of the elements and relationships of which are only partially established; the problems appear, as a rule, in situations that are characterized by the presence of an uncertainty factor and contain both formalized and non–formalized elements that are not translated into the language of mathematics. There are no universal methods and methods for system analysis. What is common to all methods of system analysis is determining the regularities of system operation, opting for system structure and choosing the best option by solving the problems of decomposition, analysis and synthesis of the system under investigation.

The system implementation of providing the navigation instrumental procedure is considered as a system object, taking into account its consistent nature, that is, not as a whole, but as a complex of interconnected constituent elements, their properties and processes. Besides this, system analysis is the most constructive of the directions of system researches, which involves methods and models of theory of systems for their practical application.

The aim of this work is to develop a method of system analysis at the stage of system implementation of providing navigation instrumental procedure on inland waterways of Ukraine, which guarantees high-quality modeling in accordance with precise navigation criteria.

1 PROBLEM STATEMENT

In order to end of purpose, it is necessary to consistently solve a number of tasks related to the implementation of divided information and telecommunication system. The most important step in solving the problems is to develop a method of system analysis.

Let suppose that we have an initial sample of setting the problem of system analysis of the current pilot navigation method, which is as follows: $A = \{a_k\}$, $Z_i = \{z_{kn}\}$, $Q = \{q_{kn}\}$, $D = \{A, Q\}$. This does not correspond to the regularities of the operation of a complex object with the current navigation instrumental procedure through existing local, sectoral tasks and principles of ensuring its life cycle. The solution of the problem is to determine complex problems of the system object and processes, which differ from a simple sum of properties of elements with the same type of multilevel communications.

Initial sampling for system analysis for providing the navigation instrumental procedure has the following form: $A = \{a_i\}$, $Z_i = \{z_{ij}\}$, $Q = \{q_{ij}\}$, $D = \{A, Q\}$, $k \neq i, j \neq m$, $Z(t_1) \rightarrow Z(t_2) \rightarrow Z(t_3)$ ...,$Z(t) = F$c $[\boldsymbol{X}(t)]$.

It is necessary to solve the problem of system operation in a continuous mode $f(x_0 - 0) = f(x_0 + 0) = f(x_0)$, $\forall x \in R$ and taking into account the signs of dynamic system $x(t) = T_t\, x(t_0)$ in the conditions of significant uncertainty and the impact of the environment on it.

The evaluation method of task solving for functioning of navigation instrumental procedure with a continuous output is criteria method; according to this method, each individual alternative is estimated by a specific number (criterion, target function). Comparison of alternatives is reduced to comparison of corresponding numbers, that is, for the entire set of alternatives $X = \{x_n\}$, the target function $Z = f(x) \Rightarrow$ max or min is introduced. The values of $Z_{min} < Z < Z_{max}$ alternatives are expressed in terms of scalar, vector, and plural quantities. The main type of scales is the scale of intervals, since this type contains scales that are accurate to the set of positive linear admissible transformations of the form $f(x) = ax + b$:

$$\frac{x_1 - x_2}{x_3 - x_4} = \frac{f(x_1) - f(x_2)}{f(x_3) - f(x_4)} = \text{const}.$$

2 LITERATURE REVIEW

It happened that the problem of the instrumental procedure support system use on the inland waterways of

Ukraine, which is being investigated, has not been fully studied yet, some of its issues are investigated superficially and fleetingly. As the analysis showed, a scientifically grounded approach of this direction was considered only in the form of hypotheses without theoretical substantiation.

When applying the functional stability of the instrumental procedure support system, which is considered in this article, various variants of the concept of stability are taken into account-Lyapunov stability, stability in probability, Lagrange stability, etc. However, this concept refers not to this object, but only to one of its properties. There are plenty of methods for analyzing stability developed for different systems: methods of Lyapunov, Hurwitz, Mykhailov, Popov, and others. In the classical theory of stability, the criteria for determining the stability of the system are developed [6, 7].

The concept of functional stability for a dynamic system was first defined in works of professor O. A. Mashkov [2]. The theory of developing of stable distributed automated control systems is proposed by professor O. V. Barabash. [3]. This theory is based on the estimation of the connectivity of the graphs of network. The proposed approach is used in the study of the state of the system using a mathematical model in the form of a non-oriented graph, which allowed to consider the features of system safety and develop an algorithm for implementing its functional stability.

There are also other approaches to providing functional stability for special purpose navigation system (Professor Yu. V. Kravchenko) [11]. It should be noted that this classical approach to the theory of stability for modern systems, based on the active use of methods of artificial intelligence, does not meet the requirements concerning adequacy and reliability of estimation of stability of such systems' operation. Due to this, the dependence of the models and methods of the theory of functional stability on the domain of their application is taken into account in the system analysis.

Besides this, the imperfection of regulatory standards allowed the designer to independently develop and implement algorithms of implementing of functional capability of electronic and cartographic equipment. As the linguistic analysis showed, this resulted in significant differences in the design of information and control panels, various sub-levels of the menu, formats of vector electronic map display. According to the data set out in the UNECE Resolutions, in addition to the mandatory requirements for the equipment of river electronic cartographic systems ("must be"), there are some which are not obligatory ("may be") without a specific indication of minimum requirements for displaying such information, which greatly complicated the operation of various types of systems and conducting system analysis.

The problem connected with the system of providing the navigation instrumental procedure also calls into question the need of detailed depth mapping on electronic river maps, since the precise indication of the position of a fairway channel is much more important for navigation safety. The publications about the update of the displayed depth in the system electronic map of rivers with the navigation instrumental procedure did not solve the problem and the same problem was considered only at the level of suggestions and hypotheses.

In scientific works of L. L. Vahushchenko "Ship navigation and information systems", "Modern information technologies in navigation" [1, 4], D. A. Haharskyi "Electronic cartographic systems in modern navigation" [5], the problem of using the detailed array of depth with the navigation instrumental procedure consisted only in taking into account the corrections using the cartographic editor of the river electronic mapping system that is, manually without a systematic scientifically grounded approach to the periodicity of updating the existing depths on the river electronic map, its automation and display.

An important problem is the choice of structure to represent the system of providing the navigation instrumental procedure, taking into account its features and capabilities. On the basis of the comparative analysis of structures of system theory proposed by academicians P. K. Anokhin, L von Bertalanfi, A. A. Bohdanov [6, 7], information estimates of the degree of wholeness and the coefficient of use of system components, which are interpreted as estimates of stability in making structural elements freeness are used.

3 MATERIALS AND METHODS

First, consider the circular diagram of the stages of increasing the efficiency of water vehicles operation on inland waterways in Ukraine (Fig. 1).

From the circular scheme it follows that the increase of efficiency of operation of water transport on inland waterways of Ukraine is consistent. The system of providing the navigation method is considered to be a set of objects having integral properties and regularities, as well as a set of connectivities between objects and their properties.

Fig. 2 shows the key components of the navigation instrumental procedure on inland waterways in Ukraine. According to this:

– the operation of navigation equipment and equipment of the coastal infrastructure belongs to the divided hierarchical intellectualized system of management of a complex technical object;

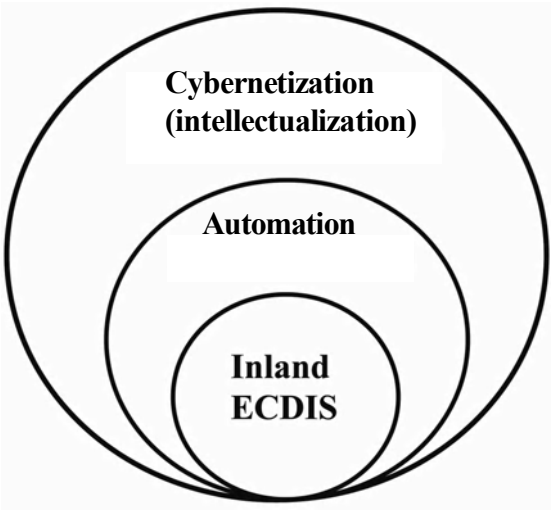


Figure 1 – The stages of increasing the efficiency of water transport exploitation

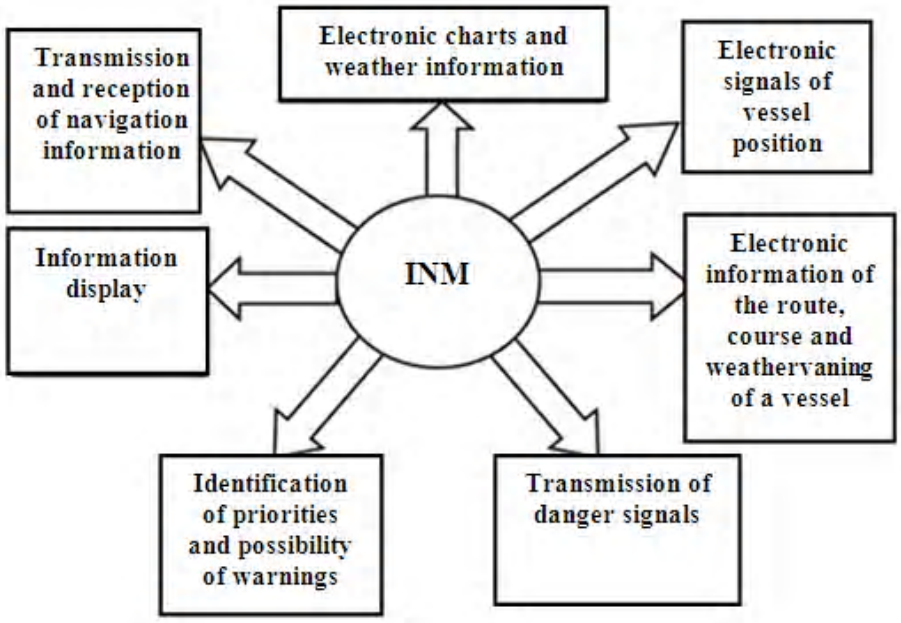


Figure 2 – Key components of the instrumental navigation method

– the real process of operation of navigation equipment system and equipment of coastal infrastructure should correspond to the calculated one;

– the stability of such a method is based on the analysis of a number of factors that are currently not taken into account due to the use of approximate models. It should be noted that the problem of determining the functional stability of such a complex organizational system remains open today.

The system should be optimized according to the set of criteria [8] characterizing the implementing of set tasks with the probability of safe navigation of water transport

$$P_{sm} = 1 - \exp\,(D_{min}/M)^2 \geq 95\%. \quad (1)$$

Fig. 3 shows the layout of the system.

In our situation, the system for providing the navigation instrumental procedure is a complex entity, which is a set of interconnected elements. That is, an element is an indivisible part of the system that is independent. The indivisibility of the element is considered as inappropriate account of its internal structure within the framework of the model. The element itself is characterized only by its external manifestations in the form of links and interconnections with other elements. Plurality A of system elements can be described as [9]

$$A = \{a_i\},\ i = 1, ..., n. \quad (2)$$

Each element is characterized by specific properties Z_{i1}, …, Z_{im}, which uniquely determine it in this system. The set of all properties m of the element a_i will be defined as the state of the element Z_i.

$$Z_i = (Z_{i1}, Z_{i2}, Z_{i3}, \ldots, Z_{ik}, \ldots, Z_{im}). \quad (3)$$

$$Q = \{q_{ij}\},\ i,j = 1, ..., n.,\ D = \{A, Q\}, \quad (4)$$

$$Z = (Z_1, Z_2, Z_3, \ldots, Z_k, \ldots, Z_m), \quad (5)$$

Bilateral constraints cannot be greater than n & $(n–1)$.

Possible states of such a real system form a certain set of admissible states of the system (subspace $Z_{СД}$). The system has the properties of equilibrium, stability and is able to move from one state to another [10]

$$Z(t_1) \rightarrow Z(t_2) \rightarrow Z(t_3)\ \ldots, Z(t) = [Z(t-1), y(t), x(t)]. \quad (6)$$

Let us consider the dependencies of the state of the system on the functions of inputs and outputs.

$$\boldsymbol{X} = (x_1, x_2, x_3, \ldots, x_k, \ldots, x_r) \Rightarrow Z(t) = F\text{c}\,[\boldsymbol{X}(t)]$$

$$Z(t) = F\text{c}\,[\boldsymbol{X}(t), Z(t-1), Z(t-2), ...]. \quad (7)$$

Fig. 4 shows the feedback of individual elements of the system

$$Y(t) = F_{\text{в}}\,[\boldsymbol{X}(t)]. \quad (8)$$

For the $F_{\text{в}}$ function, the real system has signs of static.

At the initial stage of the study, the object was studied with its further formalization. At this stage, a two-fold task appeared. On the one hand, this is the formalization of the object of the system research, on the other hand, the process of studying the system, the process of setting and solving the problem are formalized too. Features of the system:

1) the system depends not only on the functions of inputs $X(t)$, but also on the functions of states $Z(t-1)$, $Z(t-2)$, that is, it has signs of the dynamical system

$$Y(t) = F_{\text{в}}\,[\boldsymbol{X}(t), Z(t)\,, Z(t-1), Z(t-2), (Z - \upsilon)]; \quad (9)$$

2) functioning of the system is carried out in a continuous mode

$$\frac{dZ(t)}{dt} = F_C\left[X(t), Z(t)\right], \qquad Y(t) = F_B\left[X(t), Z(t)\right]; * \quad (10)$$

3) $\forall\ Z_{im} \in Z \rightarrow Z_{im} = \neg\ Z,\ Z = Z_1\ \&\ Z_2;$

Z_{im}	Z	$Z_{im} = \neg$ Z	Z_1	Z_2	$Z_1 \& Z_2$
0	0	0	0	0	0
0	1	1	0	1	0
1	1	0	1	0	0
1	0	1	1	1	1

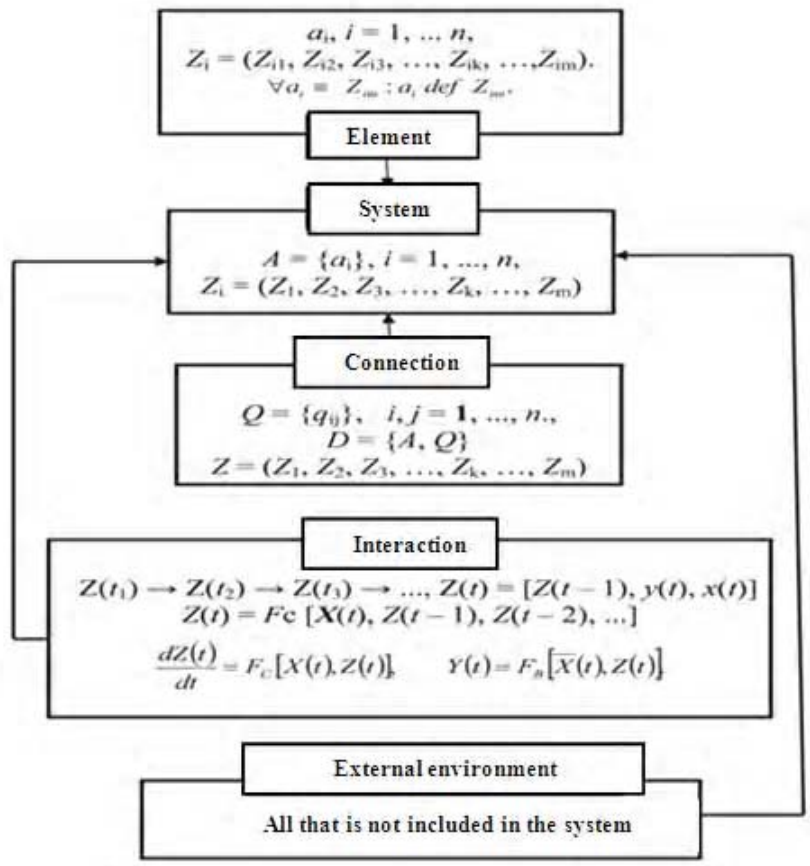


Figure 3 – The layout scheme of the system providing instrumental navigation method

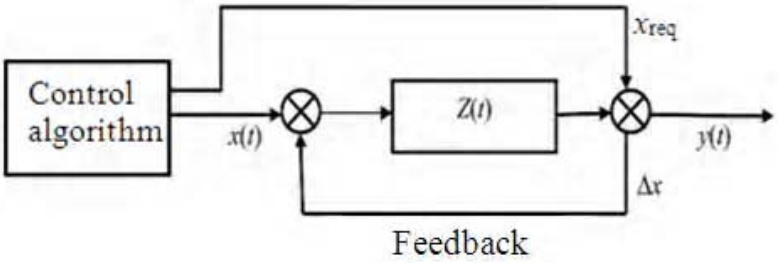


Figure 4 – Scheme of the subsystem with a single feedback

4) the system operates in conditions of significant uncertainty and effect of the environment, which causes the random nature of changes of its indicators

$$pF(x) = \sum_{i=1}^{n} p_i(x_i) = 1, \; \int_{-\infty}^{\infty} p(x)dx = F(\infty) - F(-\infty) = 1 - 0 = 1;$$

5) the system carries out a purposeful selecting of its behavior. The functional stability of the system characterizes the deviation of the main functions from the display coordinates [11]

$$\forall E > 0 \Rightarrow \Delta_{x,y} > 0, \; V\left[f(x_0, y_0), f(x_0^1, y_0^1)\right] < \Delta_{x,y} \Rightarrow$$
$$\Rightarrow V\left\{f(x,y)[t,(x_0,y_0)], f(x,y)\left[t,(x^1,y^1)\right]_0\right\} < E, \quad \forall t \in (o, \infty).$$

The stabilization process is characterized by the output from the domain of a given tube $W(t)$. The area $W(\tau)$ depends on $\tau = t_2 - t_1$ the duration of the transient (Fig. 5). Instead of

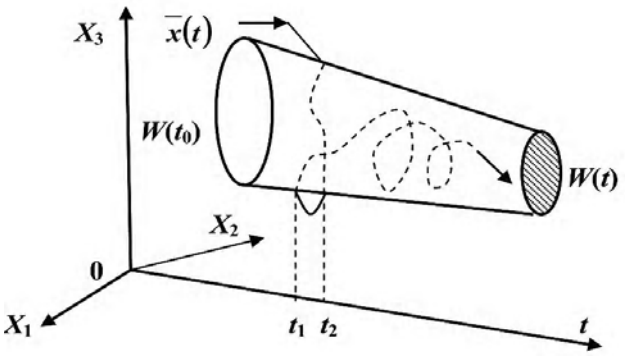


Figure 5 – Parameter variation of system parameters

inappropriate state of the system, the reaction in the form of a set of control influences $u(t)$ that belong to a set of hyper-real numbers U^* is carried out and aimed at neutralizing the consequences in a minimum time τ:

$$\left\{\int_{(\Omega)} X(x_1, x_2, t_2)d\omega > W(t_0)\right\} \Rightarrow \left\{u(t) \in U^*[\delta[w(t), w(\tau)]]\right\}.$$

The constituent $\delta[w(t), w(\tau)] = \sup\{p[z, w(t) \cap w(\tau)],\ z \in w(t) \cup w(\tau)\}$ characterizes the distance p in the processes $\Omega(x, \tau, t) \in S$ (ε) $[w(t)] < \varepsilon$. Condition: $|t - \tau| < \gamma \Rightarrow \delta\ [\omega(\tau), \omega(\tau)] < \varepsilon.$

The processes of parameter stabilization are characterized as the stages of ensuring the functional stability of the corresponding subsystems, which work properly in the given intervals of variation of parameters.

An important task of this system analysis is the problem of decision making. The problem of decision making is connected with the choice of a certain alternative of the development of the system in a variety of uncertainties. With the use of modern mathematical apparatus it appeared impossible to rely only on strict mathematical methods. Therefore, informal procedures are widely used in system analysis: combination of formal and informal methods of analysis and synthesis.

Let's dwell on choosing the structure for the presentation of the system, taking into account its features and capabilities. The information estimates of the degree of wholeness α and the coefficient of the use of components of the β system are used, which are interpreted as stability estimates when making structure elements independent. These estimates are received from the ratio that determines the relationship between the system, eigen and mutual complexity of this system [12]

$$\begin{cases} C_C = C_0 + C_B, \\ \alpha = -\dfrac{C_B}{C_0},\ \beta = \dfrac{C_C}{C_0},\ \beta = 1 - \alpha. \end{cases} \quad (11)$$

The functional component of a multicomponent system is characterized by:

– its own complexity. In this case, C_o is the total complexity of system elements without connections between them (navigation tools, bridges, power lines, locks);

– the system complexity of C_c characterizes the content of the system as a whole;

– the mutual complexity of $C_в$ characterizes the degree of interconnection of the elements in the system;

– the estimate of α describes the stability and controllability of the system. A *minus* sign in the expression α is introduced for to be positive, since $C_o > C_c$;

– β – coefficient describes the use of system elements.

We use a mathematical model in the form of a non–oriented graph due to logical connections in the system are two-sided. There are no loops and multiple edges in the graph. A set of vertices V corresponds to a set of constituent elements of the power system n, and a set of edges (E, F) corresponds to a set of functional connections between the constituent elements of the system. The specific state of this system characterizes $\langle v, e, f \rangle$:

$$G(V, \langle E, F \rangle),\ V = \{v_i\},\ \langle E, F \rangle = \{\langle e, f \rangle_{ij}\},\ i, j = 1,2,...n.$$

$$v \subseteq V,\ |v| \le |V|,\ e \subseteq E,\ |e| \le |E|,\ f \subseteq F,\ |f| \le |F|.$$

$$|v^+| = |V|.\quad |e^+| = |E|.\quad |f^+| = |F|.\quad |v^-| < |V|.$$

$$|e^-| < |E|.\quad |f^-| < |F|.$$

Signs of functional SZIMN stability are the connectivity of the graph {K = 1} and the presence of bridges in its composition $N_L \ge 1$, as well as the loops of connection

$$N_V \ge 1\ \{K = 1\} \wedge \lfloor \{N_V \ge 1\} \vee \{N_L \ge 1\} \rfloor \quad (12)$$

It should be noted that (12) significantly reduces the functional stability of the divided information system, since all routes of information transmission from the vertices of one subgraph to the top of another will include this bridge or loop of convergence. Therefore, in order to bring the system into a functionally stable state, it is necessary to introduce reserve lines into the structure so that there are no bridges or junction loops in the structure. Several independent and alternative information transmission routes will appear.

The functional stability of the system is characterized by criteria [11] regarding the dimensions of the navigable channel $v_1^- < k_1 \cdot V_1$; corner radius of the navigable channel on roll tacks $v_2^- < k_2 \cdot V_2$; depth of the navigable channel $v_3^- < V_3 - \Delta_3$; dimensions of the bridges $v_6^- < V_6 - 2 \cdot \Delta_6$, $v_{6-1}^- < V_{6-1} - \Delta_{6-1}$; height of the wires of power lines $v_7^- < V_7 - \Delta_7$; the accuracy of the vessel's coordinates $e_1^- \le$ 10 м (with a probability of 0.95); safe distances v_8^+, taking into account the edges in a non-oriented graph (12). Figure 6 shows the algorithm of system stability.

The developed method revealed the need to supply the system with a new inclusion element – an automated display of depths and isobaths for the actual water level on the electronic map. The application of the "black box" is a subsystem, which is isolated and connected to the external environment through a set of inputs $y_{1(t)}$ and outputs $y_{k(t)}$. Outputs of the "black box" model describe the performance of the subsystem, and inputs describe resources and constraints (Fig. 7):

$$h(t) = \underbrace{\frac{1}{\det(pE - A)}}\, S(p) \cdot B + \underbrace{y_{01BП} + h_{10}}_{const} + \left\| (\delta y_{01BП})_{i,1} \right\| \cdot \quad (13)$$

The expression (13) is the result of intelligent processing of the data flow from water measuring points using symbolical and formalized Laplace transformations while navigating a vessel and it means a complete expression for obtaining differential corrections h (t) of the depths indicated on the river electronic map. Thus, the uncertainty of situational dynamics in the n–dimensional space of vessel motion is eliminated. The conceptual model for obtaining differential corrections is depicted in Fig. 8.

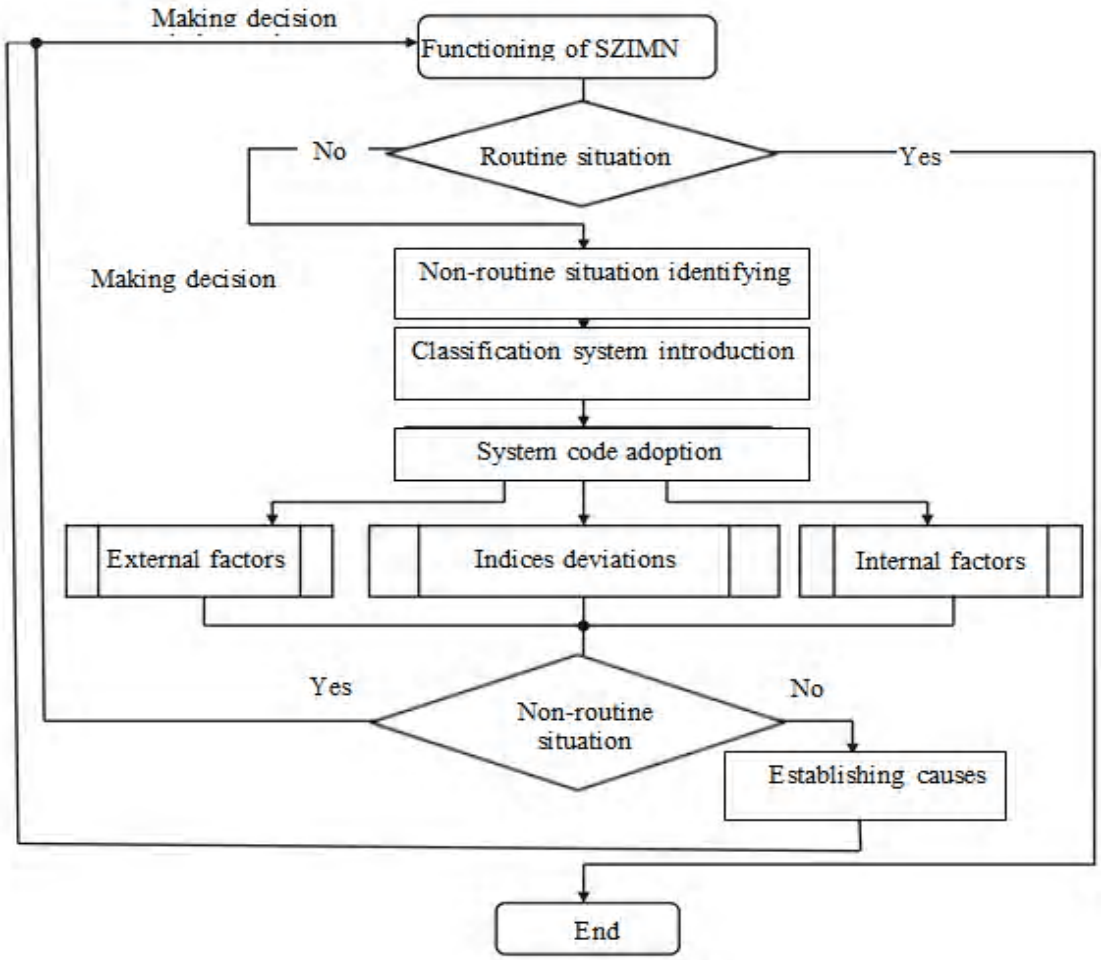


Figure 6 – Algorithm for the implementation of functional stability of the system

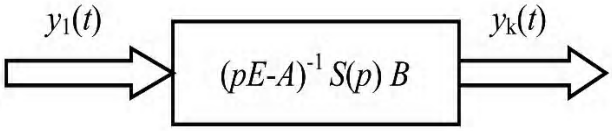
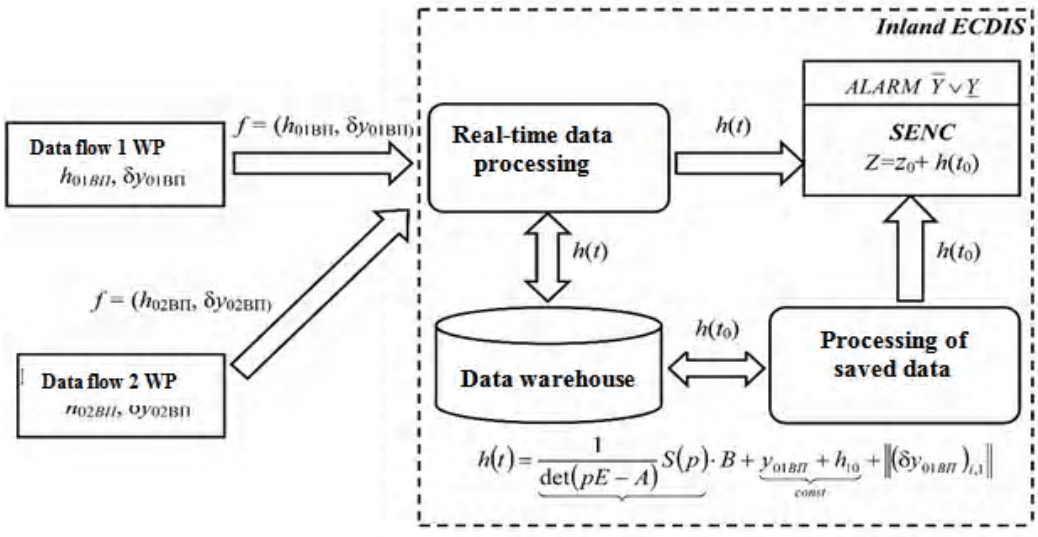


Figure 7 – Scalar transmission function of the subsystem of automated depth display



* ВП – WP

Figure 8 – Conceptual model for obtaining differential corrections

4 EXPERIMENTS

The developed technique enabled to consider the proposed system as two functionally interconnected components of subsystems F_1 and F_2 (river electronic chart system and coastal infrastructure). The result of the interaction of the two subsystems F_1 and F_2 is performing of the tasks with a probability of safe navigation $P_{бп} \geq 95\%$ using expression (1) for specific conditions of water traffic. Besides this, the developed technique revealed the need to supply this system with a new inclusion element – an automated display of depths and isobaths for the actual water level on the river electronic map. Methods of collective generating of ideas and expert methods are used for its analysis.

Experimental studies of modeling of functioning process of the stages of functional stability of the coastal infrastructure and the studies of navigation equipment were carried out on the simulator of SeeMYENC River Electronic Chart System. Accordingly, when modeling the situation, certain elements of the system were set to nonoptimal values. In this case, the optimality of the system as a whole was determined. The calculation of the margin of stability with the capacity to be rebuilt in accordance with external conditions for the effective achievement of overall goal was carried out using the formulas (1, 11). The stability area in the vector space was calculated using the information from the river electronic maps. The fundamental point of modeling is that the removal of the consequences of non–routine situations was carried out by redistributing existing resources, including changing the mode of operation of the subsystems. Possible parameter variation was controlled within specified limits of the "operating tube" (Fig. 5). Observations obtained with this method were used to get a sample for solving the problem of ensuring the safety of water transport using expression (1). There is one fact of quitting the "tube" when positioning in the subsystem F_1.

5 RESULTS

The results of the experiments carried out are presented in table 1.

6 DISCUSSION

The existing visual (pilot) navigation method was not based on solving of tasks applying a scientific approach to the introduction of modern navigation system on inland waterways of Ukraine. Hence, the methods of system analysis were simple. Thus, the system analysis of the pilot navigation method was based on its consideration not as a system object, but as a simple sum of properties of the elements with the same type of multilevel communications [13]. This does not correspond to the regularities of complex object operation in a modern way of navigation through existing local, branch tasks and principles of ensuring its life cycle. The stability of functioning of the modern navigation method is based on the analysis of a number of factors that are currently not taken into account due to the use of approximate models [11]. It should be noted that the problem of determining the functional stability of such a complex organizational system as the system of providing the navigation instrumental procedure on inland waterways of Ukraine remains open today.

The article shows us a scientific approach to the introduction of a modern navigation system in the form of solving weakly structured problems, in the first instance, that are characterized by the presence of uncertainty factor and contain both formalized and non–formalized elements; the article also used information estimates of the degree of wholeness and the coefficient of use of components of the system, which are interpreted as stability estimates when making structure elements independent. Thus, combining of information from numerous sources takes place, in order to find more accurate and reliable data about the situation than the results obtained from these sources separately.

The proposed method of system analysis provides greater accuracy of forecasting, eliminates significant ambiguity, vagueness and unpredictability of situations of extreme, risky nature. Thus, the theoretical incompleteness and multi–valued possibility of various interpretations in practice leads to dangerous statistics of disasters, accidents, unwanted events with passengers, crews, cargos, as well as contamination of the ecological system [10]. However, the proposed method requires more time spent on sampling than the method of system analysis for the visual (pilot) navigation method, and the success / effectiveness of its application depends on its requirements for computing resources. Subject, theme, task of introducing the proposed method will be more understandable when recording the full description of the whole object using single language in the form of

$$\left\{A = \{a_i\}_{i=\overline{1,n}}, Z = \{z_{ij}\}_{i-\overline{1,n}; j=\overline{i,m}}, Q = \{q_{ijk}\}_{i=\overline{1,n}; j=\overline{1,m}; k=\overline{1,l}}\right\}$$

CONCLUSIONS

The task of applying the scientific approach to the implementation of the modern navigation system on inland waterways of Ukraine on the basis of the river electronic cartographic systems is solved in the paper. The most important stage in the implementation of the modern navigation system is the complex implementation of the system analysis of the problem situation, taking into account the specific character of water transport traffic. In the process of research during the process of system analysis, a set of

Table 1 – Experiment results

ENC	α (11)	β (11)	δy (12), *m*	$h(t)$ (12), *m*	Δ_3, *m*	P_{sm} (1), %	Дата, час *t*
UA7N007.000	0,79	0,21	0,2	4.6	3,2	0,97	24.07.2017, 10–00
UA8N0028.000	0,73	0,27	0	5,2	3,2	0,96	25.07.2017, 19–00
UA8N0082.000	0,67	0,33	0,2	5,4	3,6	0,96	27.07.2017, 11–00
UA7N0036.000	0,80	0,29	0,1	6,2	3,6	0,97	28.07.2017, 17–00

procedures with elements of scientific novelty was used: the definition of an overall aim, the recommendation of a set of alternatives to achieve these aims, which are compared to some or other performance criteria.

The scientific novelty of the obtained results is that as a result of decomposition, when applying the proposed system analysis, new elements and subsystems were first justified and implemented: such as the functional stability of the navigation system (estimates, signs, criteria) with a set of permissible states, the automated display of real depths on the system electronic card.

The practical significance of the obtained results is that the introduction of the proposed method will enable refusal of the installation of coastal and floating means of navigation equipment, will provide differentiated mapping of depths on the electronic map, depending on an actual water level, and refuse the pilot principle of navigation on the inland waterways of Ukraine. In addition, it should be noted that the introduction of the navigation instrumental procedure will enable to quickly shut down navigation on the fairway channels in particular periods and follow only selected vessels.

Prospects for further research. It should be noted that the final conclusion about the correctness of the system analysis of such a complex system can only be made on the basis of the results of its practical application. The final result will depend not only on how thorough and theoretically grounded the methods used in the analysis are, but also on how accurately and qualitatively the received recommendations are realized. It should also be borne in mind that the situation, where researches are initially conducted and then their results are put into practice, happens relatively rare in practice, and it is only in those cases when it comes to relatively simple systems. In the study of such a complex system, it varies over time, both by itself and under the influence of researches. That is, the problem is gradually solved in the course of an active research.

ACKNOWLEDGEMENT

The work was carried out within the scope of the research theme “Methods of ensuring the safety of water transport with the use of the detailed depicting of depths in river electronic-mapping systems”, which is approved by the decision of the Scientific Council of the State University of Infrastructure and Technologies (registration number ID: 64940 27.08.2016 (398– 1)).

REFERENCES

1. Вагущенко Л. Л. Современные информационные технологии в судовождении: учебное пособие / Л. Л. Вагущенко. – Одесса : ОНМА, 2013. – 135 с.
2. Машков О. А. Розвиток теорії функціональної стійкості, як стійкості функціонала якості складної системи / О. А. Машков, В. Р. Косенко. – Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту : Міжнародна наукова конференція ISDMCI’2016, Херсон, 24–28 травня 2016 р. : матеріали конференції. – Херсон: Видавництво ПП Вишемирський В. С., 2016. – С. 93–104.
3. Барабаш О. В. Функціональна стійкість – властивість складних технічних систем / О. В. Барабаш, Ю. В. Кравченко // Збірник наукових праць НАОУ. – 2002. – Бюл. № 40. – С. 225–229.
4. Вагущенко Л. Л. Судовые навигационно-информационные системы: учебное пособие / Л. Л. Вагущенко. – Одесса : ОНМА, 2004. – 302 с.
5. Гагарский Д. А. Электронные картографические системы в современном судовождении: учебное пособие / Д. А. Гагарский. – Санкт-Петербург: ГМА имени адмирала Макарова, 2007. – 61 с.
6. Анохин П. К. Избранные труды: философские аспекты теории систем : монография / П. К. Анохин. – Москва : «Наука», 1978.– 190 с.
7. Берталанфи Л. Фон. Общая теория систем: критический обзор : монография / Берталанфи Л. Фон. – Москва : «Прогресс», 1969. – 82 с.
8. Доронін В. В. Системна технологія розв’язку оперативних задач навігації для синтезу законів експлуатації водного транспорту / В. В. Доронін. – Харків : Системи обробки інформації. – 2015. – № 10 (135). – С. 186 – 191.
9. Бронштейн И. Н. Справочник по математике: коллективная монография / И. Н. Бронштейн, К. А. Семендяев. – Москва : «Наука», 1980. – 974 с.
10. Панін В. В. Структурне моделювання та символьні перетворення для управління рухом транспортних засобів: колективна монографія / [В. В. Панін, Г. Л. Баранов, А. М. Носовський и др.]. – Київ : МОН України, 2014. –310 с.
11. Doronin V. V. Application of evaluation criteria of functional sustainability instrumental method of navigation on Ukraine’s inland water ways / V. V. Doronin, M. V. Aleinikov, V. M. Aleinikov // Intellectual Systems for Decision Making and Problems of Computational Intelligence: XIII International Scientific Conference (ISDMCI’2017). – Kherson: PP Vyshemyrskyi V. S., 2017. – P. 179–180.
12. Волкова В. Н. Системный анализ и его применение в АСУ : коллективная монография / В. Н. Волкова, А. А. Денисов. – Ленинград : ЛПИ, 1983. – 83 с.
13. Симоненко С. В. Навігаційне забезпечення мореплавства: виробнично-практичний посібник / С. В. Симоненко, М. Ф. Голодов. – Київ : ДУ «Держгідрографія», 2014. – 264 с.

Article was submitted 12.09.2017.
After revision 25.11.2017.

Панін В. В.[1], Доронін В. В.[2], Алейніков В. М.[3]

[1]Д-р техн. наук, професор, ректор, Державний університет інфраструктури та технологій, Київ, Україна

[2]Канд. техн. наук, професор кафедри Технічних систем і процесів управління в судноводінні, Державний університет інфраструктури та технологій, Київ, Україна

[3]Аспірант, Державний університет інфраструктури та технологій, Київ, Україна

ЗАСТОСУВАННЯ СИСТЕМНОГО АНАЛІЗУ ВПРОВАДЖЕННЯ ІНСТРУМЕНТАЛЬНОГО МЕТОДУ НАВІГАЦІЇ НА ВНУТРІШНІХ ВОДНИХ ШЛЯХАХ УКРАЇНИ

Актуальність. Вирішено завдання застосування наукового підходу щодо впровадження сучасної системи навігації на внутрішніх водних шляхах України у вигляді вирішення в першу чергу слабоструктурованих проблем, що характеризуються наявністю фактора невизначеності і містять як формалізовані, так і неформалізовані елементи, що не перекладаються на мову математики.

Мета роботи – розробка методики системного аналізу на етапі впровадження інструментального методу навігації на внутрішніх водних шляхах України, яка гарантує високу якість моделювання згідно з точнісними критеріями судноплавства і економічності.

Метод. Запропонований підхід до обробки даних, який заснований на вивченні властивостей системи з подальшою її формалізацією. Розроблена модель компонування системи на базі електронно-картографічних систем, що характеризує виконання поставлених завдань з підвищеною ймовірністю безпечного плавання. Розглянуто особливості стану системи з використанням математичної моделі у вигляді неорієнтованого графа. Запропонований алгоритм реалізації функціональної стійкості системи. Результати експериментів дозволяють рекомендувати запропонований метод для впровадження інструментального методу навігації на внутрішніх водних шляхах України.

Результати. Відпрацьовані показники експериментальних досліджень процесу функціонування етапів функціональної стійкості берегової інфраструктури і суднового навігаційного обладнання.

Висновки. У процесі дослідження використаний комплекс процедур, що мають елементи наукової новизни і спрямовані на формулювання проблемної ситуації, висунення множини альтернатив, які зіставляються з критеріями ефективності із застосуванням коефіцієнтів стійкості системи.

Ключові слова: безпека руху, електронно-картографічна система, інструментальний метод навігації, алгоритм, багатопараметрична оптимізація, граф, лінгвістичний аналіз, моделювання.

Панин В. В.[1], Доронин В. В.[2], Алейников В. М.[3]

[1]Д-р техн. наук, профессор, ректор, Государственный университет инфраструктуры и технологий, Киев, Украина

[2]Канд. техн. наук, профессор кафедры Технических систем и процессов управления в судовождении, Государственный университет инфраструктуры и технологий, Киев, Украина

[3]Аспирант, Государственный университет инфраструктуры и технологий, Киев, Украина

ПРИМЕНЕНИЕ СИСТЕМНОГО АНАЛИЗА ВНЕДРЕНИЯ ИНСТРУМЕНТАЛЬНОГО МЕТОДА НАВИГАЦИИ НА ВНУТРЕННИХ ВОДНЫХ ПУТЯХ УКРАИНЫ

Актуальность. Решена задача применения научного подхода по внедрению современной системы навигации на внутренних водных путях Украины в виде решения в первую очередь слабоструктурированных проблем, характеризующихся наличием фактора неопределенности и содержащих как формализованные, так и неформализованные элементы, которые не переводятся на язык математики.

Цель работы – разработка методики системного анализа на этапе внедрения инструментального метода навигации на внутренних водных путях Украины, гарантирующей высокое качество моделирования согласно точностных критериев судоходства и экономичности.

Метод. Предложен подход к обработке данных, который основан на изучении свойств системы с дальнейшей ее формализацией. Разработана модель компоновки системы на базе электронно-картографических средств, которая характеризует выполнение поставленных задач с повышенной вероятностью безопасного плавания. Рассмотрены особенности состояния системы с использованием математической модели в виде неориентированного графа. Предложен алгоритм реализации функциональной стойкости системы. Результаты экспериментов позволяют рекомендовать предложенный метод для внедрения инструментального метода навигации на внутренних водных путях Украины.

Результаты. Отработаны показатели экспериментального исследования процесса функционирования этапов функциональной стойкости береговой инфраструктуры и судового навигационного оборудования.

Выводы. В процессе исследования использован комплекс процедур, которые содержат элементы научной новизны и направлены на формулирование проблемной ситуации, выдвижение множества альтернатив, которые сопоставляются с критериями эффективности с применением коэффициентов стойкости системы.

Ключевые слова: безопасность движения, электронно-картографическая система, инструментальный метод навигации, алгоритм, многопараметрическая оптимизация, граф, лингвистический анализ, моделирование.

REFERENCES

1. Vahushchenko L. L. Modern Information Technologies in Navigation: Textbook. Odessa, ONMA, 2013, P. 135.
2. Mashkov O. A., Kosenko V. R. Development of the theory of functional stability as a stability of the functional quality of a complex system, *Intelligent decision making and intelligent computing problems: International Scientific Conference ISDMCI'2016, Kherson, May 24–28, 2016: Conference materials*. Kherson, Publishing house PE Vyshemyrskyi V.S., 2016, P. 93–104.
3. Barabash O. V., Kravchenko Yu. V. Functional stability – the property of complex technical systems, Collection of scientific works of NAOU, 2002, Bull. No. 40, P. 225–229.
4. Vagushchenko L. L. Ship navigation and information systems, Textbook. Odessa, NAOU, 2004, P. 302.
5. Haharski D. A. Electronic cartographic systems in modern navigation : Textbook. St. Petersburg, HMA named after Admiral Makarov, 2007, P. 61.
6. Anokhin P. K. Selected Works: Philosophical Aspects of Systems Theory : Monograph. Moscow, Science, 1978, P. 190.
7. Bertalanfi L. Von. General theory of systems: a critical review: monograph. Moscow, "Progress", 1969, P. 82.
8. Doronin V. V. Systemic technology of the solution of operational tasks of navigation for the synthesis of the laws of operation of water transport. Kharkiv, Systems of information processing, 2015, No. 10 (135), P. 186–191.
9. Bronstein I. N., Semendiaiev K. A. Reference book on math: a collective monograph. Moscow, "Science", 1980, P. 974.
10. Panin V. V., Baranov G. L., Nosovskyi A. M., Tykhonov I. V. Vasko S. M. Structural modeling and symbolic transformations for controlling the motion of vehicles: a collective monograph. Kyiv, MES of Ukraine, 2014, P. 310.
11. Doronin V. V., Aleinikov M. V, Aleinikov V. M. Application of Evaluation Criteria for Functional Sustainability in Ukraine Inland Waterways, *Intellectual Systems for Decision Making and Problems of Computational Intelligence: XIII International Scientific Conference (ISDMCI'2017)*, Kherson, PP Vyshemyrskyi V. S., 2017, P. 179–180.
12. Volkova V. N., Denysov A. A. System analysis and its application in the automated control system: a collective monograph. Leningrad, LPI, 1983, P. 83.
13. Symonenko S. V., Holodov M. F.– Navigation equipment: production and practical textbook. Kyiv, GA "Derzhhidrographia", 2014, P. 264.

UDC 004.421.2:517.443

Prots'ko I. O.[1], Teslyuk V. M.[2]

[1]PhD., Associate Professor of Informatic Systems and Technologies Department, Lviv National Polytechnic University, Lviv, Ukraine
[2] Dr.Sc, Professor of Automatic Systems of Control Department, Lviv National Polytechnic University, Lviv, Ukraine

DEVELOPMENT OF WFTA BASED ON THE HASHING ARRAY

Context. A method of efficient computation of DFT using cyclic convolutions for sizes of integer power of two has been considered. The further development of Winograd Fourier transform algorithm based on a hashing array has been proposed. The research object is the process of the reformulation the basis matrix of DFT into the block-cyclic structures. The research subject lays in the technique of the reformulation the basis matrix of DFT for sizes of integer power of two into the block-cyclic structures.

Objective. The purpose of the work is the analysis of the structure specifics the left-circulant submatrices of the basis square matrix W_N for sizes of transform $N = 2^i$ using the hashing arrays.

Method. The article considers a technique for the efficient computation of DFT using cyclic convolutions for sizes of integer power of two, which is based on the cyclic decomposition of substitution. A hashing array has been proposed for the compressed description of the block-cyclic structure of discrete basis matrix and for the efficient computation of DFT for sizes of integer power of two.

Results. A generalized block-cyclic structure of discrete basis matrix for the efficient computation of DF using cyclic convolutions for sizes of an integer power of two based on the hashing arrays has been determined. The proposed technique is relevant for concurrent programming of DFT and for its implementation in parallel systems.

Conclusions. A general block-cyclic structure of basis matrix of DFT is regularly formed with an increase in the value of the exponent of two and is recommended for use in practice when developing the efficient means of DFT. The prospects for further research will include the formation of block-cyclic structure of basis matrix of DFT for arbitrary sizes.

Keywords: fast Fourier transform, Winograd Fourier transform algorithm, hashing array, block-cyclic structure, cyclic convolution.

NOMENCLATURE

CC – cyclic convolution;
DFT – discrete Fourier transform;
FFT – fast Fourier transform;
LC – left-circulant;
WFTA – Winograd Fourier Transform Algorithm
$a_{k,n}$ – arguments of the functions of the discrete exponential basis;
$D(n)$ – hashing array;
$D'(n)$ – simplified hashing array;
$d_{i,j}$ – element of a subarray of a hashing array;
$d'_{i,j}$ – simplified element of a subarray of a hashing array;
g – primitive root;
k – number of subarrays of a hashing array;
L_i – number of elements in a subarray
m – number of multiplications;
N – length of transform;
n – size of the hashing array;
p – prime number;
s – number of additions;
W_N exponential basis of discrete Fourier transform
$X(k)$ – output data of discrete Fourier transform;
$X1(k)$ – real part of output data of discrete Fourier transform;
$X2(k)$ – imaginary part of output data of discrete Fourier transform;
$x(n)$ – input data of a finite length N of a discrete signal;
$Zc(k,n)$ – matrice of signs of the cosine functions of the discrete exponential basis;
$Zs(k,n)$ – matrice of signs of the sine functionsof the discrete exponential basis;
Ψ – set of the substitutions.

INTRODUCTION

DFT is one of the main concepts in information technology. According to this concept a periodic or an irregular signal is presented in the form of combinations of pure frequencies, which are widely used in different electronic systems. Efficient algorithms of DFT – FFT are widely used in many fields of science and technique. There are numerous variations of FFT algorithms that use properties of periodicity and symmetry of the basis of DFT [1, 2].

Another direction of development of the efficient algorithms is a possibility of computation of FFT through cyclic convolutions [3, 4]. In 1968 C. M. Rader proposed the possibility of effective computation of DFT of prime length N via cyclic convolution of length N–1 [5]. Further development was presented by Winograd Fourier Transform Algorithm [6, 7] especially for transformation of the size of power of p – prime number. Winograd algorithms use the data reindexing, where the specific reordering is based on Chinese remainder theorem, the properties of the direct product of matrices and fast cyclic convolution algorithms. Algorithms of DFT using convolutions are collected and thoroughly considered in the books [8, 9].

The purpose of the work is the analysis of the structure specifics the left-circulant submatrices of the basis square matrix W_N for sizes of transform $N = 2^i$ using the hashing arrays.

1 PROBLEM STATEMENT

Initial data of the efficient algorithms WFTA are: x(n) – input data of a finite length $N = 2i$ ($i > 2$; n=0,1,…, N–1); m – numbers of real multiplications, s – number of real additions for computation $X(n)$ output data of DFT through fast cyclic convolutions. However, WFTA algorithms lead to some irregular structures of basis matrix of DFT. This complicates the application of algorithm into actual software or hardware implementation and accordance the number of operations m, s.

The task of research is to develop a reformation of basis matrix of DFT to the regular left-circulant structures. The efficient technique of the reformation the basis matrix of

© Prots'ko I. O., Teslyuk V. M., 2018
DOI 10.15588/1607-3274-2018-2-15

DFT for the computation through the cyclic convolutions use the hashing array $D(n)$ with the parameters: n – a size of an hashing array; k – a number of a subarrays of a hashing array; Li – a number of elements in a subarray. The values k, Li define the minimum amount and sizes of the cyclic convolutions and the correspondence of the number of operations m, s of the efficient algorithm.

2 LITERATURE REVIEW

WFTA has important theoretical result for computational complexity of DFT algorithm. It took some effort to translate the theoretical algorithm into actual software or hardware implementation. Therefore, these algorithms are researched and updated by many authors [10–13].

The developed in detail of WFTA by S. Zohar in [14, 15] the algorithm is presented by a series of sequential tables, which are convenient, compact, graphical representations of the sequences of arithmetic operations. The one-out-of-nine part of the paper [14, 15] sufficiently proves the basic DFT of algorithms for N = 8, 16 sizes. In the developed tables for N = 8, 16 there are the complications due to the fact that $N = 2^i$ ($i > 2$) has no primitive roots. In the developed of WFTA by S. Zohar there is a complication due to the fact that $N = 2^i$ ($i > 2$) has no primitive roots. Non-primitive elements of the cyclic group generate only a part of the set.

In the paper [14, 15] has been considered a modification of the relabeling scheme to handle this with the case $N = 16 = 2^4$. The primitive root $g = 3$ generates half of elements in the interval (l, N) of these with powers (mod N). Applying this ($r = g^{\rho} \bmod N$), and ($s = g^{\delta} \bmod N$) for other half with elaborate definition ∂ the author gets the following matrix as shown in Table 1.

In view of the complexity of the present case, the Table 1 is obtained in two stages. As a result, for S. Zohar of DFT the algorithm of size N = 16 needs m = 36 numbers of real multiplications or m = 20 numbers of real multiplications (with excluded from count multiplications by 1 and j) and s = 148 number of real additions.

There fore the DFT of the size of $N = 2^i$ ($i > 2$) has no primitive roots. Non-primitive elements of the cyclic group generate only a part of the set. Thus, the efficient technique of the reformulation the basis matrix of DFT into the block-cyclic structures for the computation through the cyclic convolutions needs to use another forms.

3 MATERIALS AND METHODS

DFT is based on the sum of the product of the input signal values and the complex trigonometric functions. DFT directly is computed by the equation given below

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N{}^{nk}, \; k = 0,1,..., N-1, \qquad (1)$$

where $W_N = \exp(-j\, 2\pi /N)$; $X(k)$, $x(n)$ output and input data of a finite length N the discrete signals. The discrete exponential basis W_N of the DFT (1) may be represented by the real and the imaginary parts

$$X(k) = X1(k) + x(0) - jX2(k), \quad k = 0,1,..., N-1, \qquad (2)$$

where

$$X1(k) = \sum_{n=1}^{N-1} x(n) \cos(2\pi kn / N), \quad k = 0,1,..., N-1, \qquad (3)$$

$$X2(k) = \sum_{n=1}^{N-1} x(n) \sin(2\pi kn / N), \quad k = 1,..., N-1. \qquad (4)$$

All the N of DFT components $X(k)$ can be obtained after computation of $X1(k)$ and $X2(k)$ parts (3).

The efficient FFT algorithm using cyclic convolution is based on the decomposition [16] of the matrices of the real $X1(k)$ and the imaginary $X2(k)$ parts of DFT taken separately. Let us analyze the arguments of the functions the discrete exponential basis W_N,

$$a_{k,n} = k\, n\, \alpha_N, \quad k = 1,..., N-1, \qquad (5)$$

and especially the integer (k n) components, without $\alpha_N = 2\pi / N$.

Table 1 – The LC structure of basis matrix of arguments DFT for size N = 16

s^r	**0:**	**4:**	**8:**	**12:**	**2:**	**6:**	**10:**	**14:**	**15:**	**13:**	**7:**	**5:**	**1:**	**3:**	**9:**	**11:**
0:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4:	0	0	0	0	8	8	8	8	12	4	12	4	4	12	4	12
8:	0	0	0	0	0	0	0	0	8	8	8	8	8	8	8	8
12:	0	0	0	0	8	8	8	8	4	12	4	12	12	4	12	4
2:	0	8	0	8	4	12	4	12	14	10	14	10	2	6	2	6
6:	0	8	0	8	12	4	12	4	10	14	10	14	6	2	6	2
10:	0	8	0	8	4	12	4	12	6	2	6	2	10	14	10	14
14:	0	8	0	8	12	4	12	4	2	6	2	6	14	10	14	10
15:	0	12	8	4	14	10	6	2	1	3	9	11	15	13	7	5
13:	0	4	8	12	10	14	2	6	3	9	11	1	13	7	5	15
7:	0	12	8	4	14	10	6	2	9	11	1	3	7	5	15	13
5:	0	4	8	12	10	14	2	6	11	1	3	9	5	15	13	7
1:	0	4	8	12	2	6	10	14	15	13	7	5	1	3	9	11
3:	0	12	8	4	6	2	14	10	13	7	5	15	3	9	11	1
9:	0	4	8	12	2	6	10	14	7	5	15	13	9	11	1	3
11:	0	12	8	4	6	2	14	10	5	15	13	7	11	1	3	9

Basis functions W_N are periodic of N. Therefore, according to the periodic property the matrix of arguments $X_a(k,n)$ of the basic functions contains the components

$$X_a(k,n)=\left[(k\,n)\bmod N\right]=\left\lfloor d_{k,n}\right\rfloor,\quad k=1,\ldots,N-1. \quad (6)$$

each parts (3, 4) respectively is equal to the matrix of arguments

$$X_a(k,n)=\begin{bmatrix}1,\ 2,\ 3,\ \ldots,\ (1\,(N-1))\bmod N\\ 2,\ 4,\ 6,\ \ldots,\ (2(N-1))\bmod N\\ \ldots\\ ((N-1)\,1)\bmod N,\ldots,((N-1)(N-1))\bmod N\end{bmatrix}. \quad (7)$$

Arguments $d_{k,n}$ of the basic functions of DFT is the decomposition of group representations $<N-1,\bullet>$ for prime N, where $\circ-$ is the group operation (6). In accordance with Cayley's theorem [17] the use of the substitutions $\{\Psi_1, \Psi_2, \Psi_3,\ldots, \Psi_{N-1},\}$ of according rows/columns the matrix (7), where algebraic structure $<\Psi, \circ>$ with $\circ$ – the substitution operation, is isomorphic to structure $<N-1,\bullet>$. In result, the $D(n)$ cyclic decomposition of substitution ψ_i is formed. Using of cycle notation, we can write in the form

$$D(n)=(d_{11},d_{12},\ldots,d_{1L1})(d_{21},d_{22},\ldots,d_{2L2})\ldots(d_{kL1},d_{kL2},\ldots,d_{kLk}), \quad (8)$$

where k – a number of a subarrays, d_{ij} – an element of a subarray, L_i – a number of elements in a subarray, n – a size of an array. The form (8) is a short representation of the set of left-circulant submatrices (LC) in the structure of basis matrix of DFT and will be called hashing array $D(n)$.

Let us consider the hashing array $D(n)$ for indexing rows/columns of the matrix of the arguments (7) of DFT for sizes $N = 2^i$. The properties of symmetry and periodicity of DFT basis leads to decrease of values representing components of $d_{k,n}$ of LC submatrices with the supplement of the respective submatrices $Zc(k,n)$ and $Zs(k,n)$ of signs, which contain the value of elements $+1, -1, 0$ (indicate short $+, -, 0$). The simplified matrix components of the arguments $d'_{k,n}$ is determined for sizes $N=2^i$ through a sequence (9, 10) of operations:

$$d'_{k,n}=N-(d_{k,n}\bmod N),\ \text{ if }\ (d_{k,n}\bmod N)>N/2; \quad (9)$$

$$d'_{k,n}=\begin{cases}N/4-(d_{k,n}\bmod N-N/2),\\ \qquad \text{if}\quad N/8<(d_{k,n}\bmod N-N/2\}<N/4;\\ N/4-[N/2-(d_{k,n}\bmod N-N/2)],\\ \qquad \text{if}\quad 3N/8<(d_{k,n}\bmod N-N/2\}<N/2;\\ d_{k,n},\ \text{otherwise.}\end{cases} \quad (10)$$

$$Zc[k,n]=\begin{cases}+1,\ \text{if}\quad 3N/4<c_{k,n}<N/4;\\ 0,\ \text{if}\quad c_{k,n}=N/4,3N/4;\\ -1,\ \text{if}\quad N/4<c_{k,n}<3N/4.\end{cases} \quad (11)$$

$$Zs[k,n]=\begin{cases}+1,\ \text{if}\quad 0<c_{k,n}<N/2;\\ 0,\ \text{if}\quad c_{k,n}=0,N/2;\\ -1,\ \text{if}\quad N/2<c_{k,n}<N.\end{cases} \quad (12)$$

The subarrays of $D(n)$ reproduce Hankel circular submatrices in the structure of a basis square matrix W_N what leads to computation of cyclic convolutions $Z(n)D'(n)$ and input data $x(n)$. For example, DFT for the size $N = 16$ have the hashing array:

$$D(15)=(8)\,(4,12)\,(2,6)\,(10,14)\,(1,3,9,11)\,(15,13,7,5).$$

Using the property of symmetry of DFT the simplified hashing array $D'(8)$ contains the simplified components $d'_{k,n}$, that supplements by respective components of signs $Zc(k,n)$ and $Zs(k,n)$ components:

$$D'(7)=(0)\,(4)\,(2,2)\,(1,3,1,3),$$

$$Zc(7)=(1)\,(0)\,(+,-)\,(+,+,-,-),\ Zs(7)=(0)\,(+)\,(+,+)\,(+,+,-,-).$$

Applying hashing array $D(15)$ we reduce and define the following matrices of simplified arguments of basis with addition of sign for the cosine part as shown in Tables 2, 3, 4.

In summary, the developed algorithm of DFT for size $N = 16$ leads to the computation of one four-point cyclic convolution and two two-point cyclic convolutions. This needs $m = 8$ number of real multiplications for real input data

Table 2 – The matrix of simplified arguments of basis for the cosine/sine part, $N = 16$

k^n	0:	4:	2:	6:	1:	3:	9:	11:
0:	0	0	0	0	0	0	0	0
4:	0	0	8	8	4	4	4	4
2:	0	8	4	4	2	2	2	2
6:	0	8	4	4	2	2	2	2
1:	0	4	2	2	1	3	1	3
3:	0	4	2	2	3	1	3	1
9:	0	4	2	2	1	3	1	3
11:	0	4	2	2	3	1	3	1

Table 3 – Complementing matrix of sign for the cosine part

k^n	0:	4:	2:	6:	1:	3:	9:	11:
0:	1	1	1	1	1	1	1	1
4:	1	1	–1	–1	0	0	0	0
2:	1	–1	0	0	+	–	+	–
6:	1	–1	0	0	–	+	–	+
1:	1	0	+	–	+	+	–	–
3:	1	0	–	+	+	–	–	+
9:	1	0	+	–	–	–	+	+
11:	1	0	–	+	–	+	+	–

Table 4 – Complementing matrix of sign for the sine part

k^n	0:	4:	2:	6:	1:	3:	9:	11:
0:	0	0	0	0	0	0	0	0
4:	0	0	0	0	+1	–1	+1	–1
2:	0	0	+1	–1	+	+	+	+
6:	0	0	–1	+1	+	+	+	+
1:	0	+1	+	+	+	+	–	–
3:	0	–1	+	+	+	–	–	+
9:	0	+1	+	+	–	–	+	+
11:	0	–1	+	+	–	+	+	–

$x(n)$ or $m = 16$ for complex input data $x(n)$ before $m = 20$ of real multiplications for S. Zohar of DFT the algorithm with excluded from count the multiplications by 1 and j. As result, the comparison of the Table 1 and Table 2 of the left-circulant structure of basis matrix of arguments DFT for size $N = 16$ indicates that basis in Table 2 consists entirely of block-cyclic matrix structures.

4 RESULTS

Let us consider the specifics of the structure the LC submatrices of the basis square matrix W_N for sizes of transform $N = 2^i$. Being based on substitutions of matrix rows/columns of data components of arguments the basic functions of DFT are formed by the hashing arrays $D(n)$. The examples of the hashing arrays $D(n)$ and $D'(n)$, $Zc(n)$, $Zs(n)$ (excluded (0) elements) are presented. For N=32 the hashing array consists of k=4 subarrays:

$$D(31) = (1,3,9,27,17,19,25,11)(5,15,13,7,21,31,29,23)(2,6,18,22)(10,30,26,14)(4,12)(20,28)(8,24)(16).$$

The simplified hashing array with the decrease values of elements $d'_{k,n}$ of the arguments (11) correspond to: $D'(31) = (1,3,9,5,15,13,7,11)(5,15,13,7,11,1,3,9)(2,6,14,10)(10,2,6,14)(4,12)(12,4)(8,8)(16)$.

The reduced simplified hashing array with the decrease values of elements $d'_{k,n}$ of the arguments is represented:

$$D'(15) = (1,3,9,5,15,13,7,11)(2,6,14,10)(4,12)(8)(16),$$

$$D'(15) = (1,3,7,5,\ 1,3,7,5)(2,6,2,6)(4,4)(8)(0),$$

and complemented with according arrays of signs

$$Zc(15) = (+,+,-,+,-,-,+,-)(+,+,-,-)(+,-)(-1),$$

$$Zs(15) = (+,+,+,-,-,-,-,+)(+,+,-,-)(+,+)(+)(0).$$

The left-circulant structures of basis matrix of arguments of DFT for size $N = 32$ for the real or the imaginary part is presented in Table 5.

For $N = 64$ the hashing array consists of $k = 5$ subarrays:

$$D(63) = (1,3,9,27,17,51,25,11,33,35,41,59,49,19,57,43)(5,15,45,7,21,63,61,55,37,47,13,39,53,31,29,23)(2,6,18,54,34,38,50,22)(10,30,26,14,42,62,58,46)(4,12,36,44)(20,60,52,28)(8,24)(40,56)(16,48)(32).$$

The reduced simplified hashing array with the decrease values of elements $d'_{k,n}$ of the arguments is presented:

$$D'(31) = (1,3,9,27,17,13,25,11,31,29,23,5,15,19,7,21)(2,6,18,10,30,26,14,22)(4,12,28,20)(8,24)(16)(32),$$

$$D'(31) = (1,3,9,5,15,13,7,11,1,3,9,5,15,13,7,11)(2,6,14,10,2,6,14,10)(4,12,4,12)(8,8)(0),$$

$$Zc(31) = (+,+,+,-,-,+,-,+,-,-,-,+,+,-,+,-)(+,+,-,+,-,-,\ +,-)(+,+,-,-)(+,-)(-1),$$

$$Zs(31) = (+,+,+,+,+,-,+,+,-,-,-,-,-,+,-,-)(+,+,+,-,-,-,\ -,+)(+,+,-,-)(+,+)(0).$$

The hashing array $D'(n)$ of DFT for size $N=2^i$ ($i>2$) could be described in general form:

$$D'(2^{i-1}-1) = D'_1\{2^{i-1}\}\ D'_2\{2^{i-2}\}\ldots D'_{k-1}\{2^1\}\ D'_k\{2^0\}, \quad (13)$$

where $D'_j\{\text{number of elements of the arguments}\}$ (j=1,...,k) of hashing subarrays. Therefore, the general structure of basis matrix of DFT for the real or the imaginary part can be represented in form of Table 6.

Analysis of the structure of submatrices in basis matrix of DFT in Table 5, 6 defines, that some LC submatrices have first row in the form of the reiteration the group of elements, because the simplified hashing subarray $D'_i(n)$ has the form:

$$D'_j(n_k) = (d'_{kL1}, d'_{kL2}, \ldots, d'_{kLj},\ d'_{kL1}, d'_{kL2}, \ldots, d'_{kLj}), \quad (14)$$

and respectively we have the reduction of the size of cyclic convolution in twice.

The conversion of DFT into left-circulant structures is performed on the basis of the hashing arrays. As a result, block cyclic structures of basis matrix of DFT for sizes $N = 2^n$ can be described by the versions of the hashing arrays as shown in Table 7.

Table 5 – The LC structure of basis matrix of arguments DFT for size $N = 32$

s^r	0:	1:	3:	9:	5:	15:	13:	7:	11:	2:	6:	14:	10:	4:	12:	8:
0:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4:	0	1	3	7	5	1	3	7	5	2	6	2	6	4	4	8
8:	0	3	7	5	1	3	7	5	1	6	2	6	2	4	4	8
12:	0	7	5	1	3	7	5	1	3	2	6	2	6	4	4	8
2:	0	5	1	3	7	5	1	3	7	6	2	6	2	4	4	8
6:	0	1	3	7	5	1	3	7	5	2	6	2	6	4	4	8
10:	0	3	7	5	1	3	7	5	1	6	2	6	2	4	4	8
14:	0	7	5	1	3	7	5	1	3	2	6	2	6	4	4	8
15:	0	5	1	3	7	5	1	3	7	6	2	6	2	4	4	8
13:	0	2	6	2	6	2	6	2	6	4	4	4	4	8	8	16
7:	0	6	2	6	2	6	2	6	2	4	4	4	4	8	8	16
5:	0	2	6	2	6	2	6	2	6	4	4	4	4	8	8	16
1:	0	6	2	6	2	6	2	6	2	4	4	4	4	8	8	16
3:	0	4	4	4	4	4	4	4	4	8	8	8	8	16	16	0
9:	0	4	4	4	4	4	4	4	4	8	8	8	8	16	16	0
11:	0	8	8	8	8	8	8	8	8	16	16	0	0	16	16	0

Table 6 – The left-circulant general structure of basis matrix DFT for size $N = 2^n$

$D'_1\{2^{n-1}\}$	$D'_2\{2^{n-2}\}$	…	$D'_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
			$D_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
		…	$D'_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
			$D'_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
	$D'_2\{2^{n-2}\}$	…	$D'_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
			$D'_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
		…	$D'_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
			$D'_{n-1}\{2^1\}$	$D'_n(2^0)$
				$D'_n(2^0)$
$D'_2\{2^{n-2}\}$		$D'_2\{2^{n-2}\}$		
…	…	…	…	
$D_{n-1}\{2^1\}$	D_{n-1} D_{n-1} D_{n-1}	D_{n-1}	D_{n-1} D_{n-1} D_{n-1}	$D_{n-1}\{2^1\}$
D_n … … … … … … …		… … … … … … …		D_n

Table 7 – The versions of hashing arrays $D(2^n–1)$ of DFT for sizes $N = 2^n$

The versions	Hashing arrays $D(2^n\text{-}1)$
n–2) $D(2^n–1)$	$D_1\{2^{n-2}\}$ $D_2\{2^{n-2}\}$ $D_3\{2^{n-3}\}$ $D_4\{2^{n-3}\}$… $D_{k-2}\{2^1\}$ $D_{k-1}\{2^1\}$ $D_k\{2^0\}$;
n–3) $D(2^n–1)$	$D_1\{2^{n-3}\}$ $D_2\{2^{n-3}\}$ $D_3\{2^{n-3}\}$ $D_4\{2^{n-3}\}$ $D_5\{2^{n-4}\}D_6\{2^{n-4}\}…D_{k-2}\{2^1\}D_{-1}\{2^1\}D_k\{2^0\}$;
…	…
1) $D(2^n–1)$	$D_1\{2^1\}D_2\{2^1\}...D_{k-2}\{2^1\}D_{k-1}\{2^0\}...D_k\{2^0\}$;

In case, DFT for sizes $N = 32 = 2^5$ has such versions of hashing arrays:

3) $D(2^5–1)$=
=(1,5,25,29,17,21,9,13)(3,15,11,23,19,31,27,7)(2,10,18,26)(6,30,22,14)(4,20)(8)(24)(12,28) (16).

The structure of processing module is shown in Fig.1, that computation of cosine or sine part of DFT of size $N = 2^n$ corresponds to the Table 6 and defines the half of output data.

The matrix structure of DFT for size $N = 2^n$ defines serial-parallel combination in the U of the results of cyclic convolutions (CC) for sizes $N = 2^i$, i=2,3,…,n–2 (Fig. 1). The implementations of efficient computation (CC) use availability of the fast convolution algorithms [3].

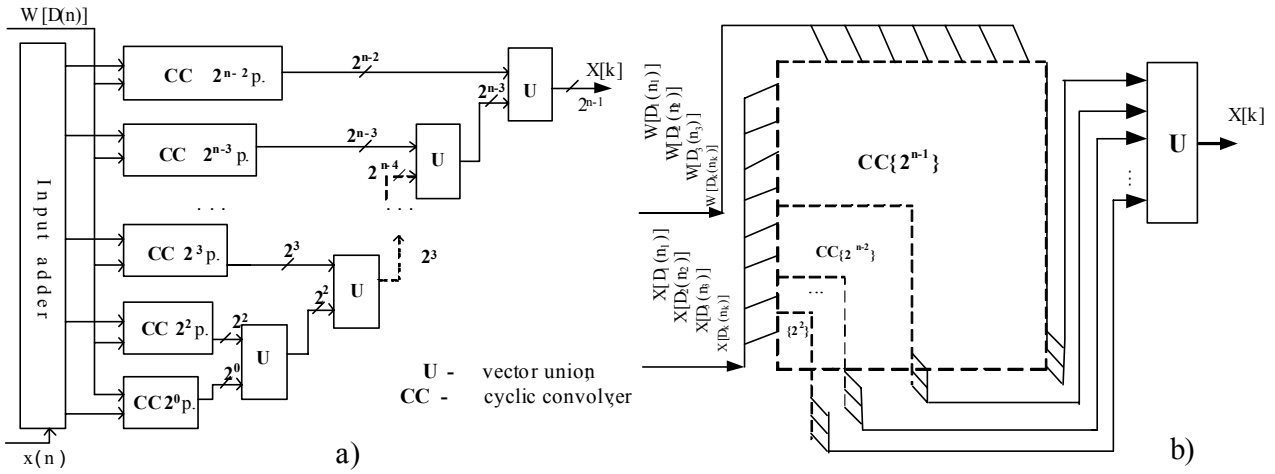


Figure 1 – Structure of processing module for a) concurrency and b) sequential computation of cosine or sine parts of DFT for size $N = 2^n$

The performances of the computation using general block-cyclic structure of basis matrix DFT is an advantage of the proposed developed WFTA. The representation of the structure of basis matrix of DFT via hashing arrays covers the whole structure and is not studied separately as the developed of WFTA by S. Zohar.

The developed algorithm of DFT for size N = 32, 64, 128 leads to the computation with the numbers of real multiplications m = 24, 68, 198 for real input data $x(n)$ with excluded from count multiplications by 1 and j. In case, we use the fast cyclic convolution algorithms with a minimum number of multiplications. The minimum known the numbers of multiplications have the so-called split-radix algorithm of DFT for an arbitrary fixed $N=2^n$ [18]. The numbers of multiplications of our solution are less in the comparison the split-radix FFT with the numbers m = 34, 98, 258 represented in [19] as shown in Fig. 2.

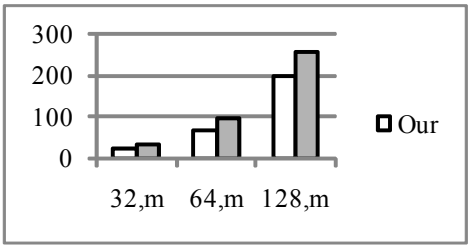


Figure 2 – Comparison of our solution and [19] by the numbers of multiplications of DFT for size N =32, 64, 128

5 DISCUSSION

The computation of real or the imaginary parts of DFT traditionally is executed during three sequential stages: pre-processing, processing, post processing. The computation of cyclic convolutions includes all multiplications of algorithm DFT using of cyclic convolutions on processing stage. The number of arithmetic operations on the basis of this approach is largely dependent on the choice of fast cyclic convolution algorithm (with a minimum number of multiplications or the balance of the operations the additions with the multiplications, and so on).

The proposed algorithm is more flexible in the comparison with the conventional algorithms as it can be applied for realization with length of any power of two. The analysis on the level of simplified hashing array $D'(n)$ with supplement of respective subarrays of $Zc(n)$, $Zs(n)$ signs reduces the amount of computation of cyclic convolutions. The formulation of DFT into left-circulant structures has been found to be very efficient for hardware implementation using VLSI technology [20].

Separate computations of cyclic convolutions (Fig.1), on which we structure the basis matrix of DFT according to the hashing arrays $D(2^n-1)$, and the combinations of convolutions results make the proposed technique important for concurrent programming and for its implementation in parallel systems.

CONCLUSIONS

The article represents the formulation technique of the efficient algorithms of DFT based on the cyclic convolutions for sizes of integer power of two. The developed algorithms are more general than the Rader-Winograd algorithms. The article is further development the algorithm of WFTA by S. Zohar for length of power of two. The number of arithmetic operations on the basis of this technique is defined of fast cyclic convolution algorithms. We have good results in the comparison with of well known algorithms of FFT in case using the fast convolutions with a minimum number of multiplications.

ACKNOWLEDGMENTS

The work was performed as part of research work "Intelligent Information Technologies of Multi-level Management of the Energy Efficiency in the Region" (number of state registration 0117U004450) of Automatic Systems of Control Department of Lviv National Polytechnic University.

REFERENCES

1. Duhamel P. Fast Fourier Transform: A tutorial Review and a State of the Art / P. Duhamel, M. Vitterli // Signal Processing. – 1990. – Vol.19. – P. 259-299. DOI: 10.1016/0165-1684(90)90158-U.
2. Chu E. Inside the FFT black box. Serial and Parallel Fast Fourier Transform Algorithms / E. Chu, A. George. – Boca Raton: CRC Press LLC, 2000.
3. Tolimiery R. Algorithms for Discrete Fourier Transform and Convolution / R. Tolimiery, M. An, C. Lu. – New York : Springer-Verlag, (s.ed.), 1997. – 267 p. DOI: 10.1007/978-1-4757-2767-8.
4. Prots'ko I. Becoming of Discrete Harmonic Transform Using Cyclic Convolutions / I. Prots'ko, R. Rykmas // American Journal of Circuits, Systems and Signal Processing. – 2015. – Vol. 1. – P.114–119.
5. Rader C. M. Discrete Fourier Transforms When the Number of Data Samples is Prime / C. M. Rader // Proceedings of IEEE. – 1968. – Vol. 56. – P. 1107–1108. DOI: 10.1109/PROC.1968.6477.
6. Winograd S. On computing the discrete Fourier transform / S. Winograd // Proceedings National Academy of Science USA, Mathematics. – 1976. – Vol. 73, No. 4. – P. 1005–1006. DOI: 10.1073/pnas.73.4.1005.
7. Winograd S. On computing the discrete Fourier transforms / S. Winograd // Mathematics of Computation. – 1978. – Vol. 32. – P. 175–199. DOI: 10.1090/S0025-5718-1978-0468306-4.
8. Blahut R. E. Fast algorithms for signal processing / R. E. Blahut. – Cambridge : University Press, 2010. DOI: 10.1017/CBO9780511760921.
9. Nussbaumer H. J. Fast Fourier Transform and Convolution Algorithms / H. J. Nussbaumer. – Berlin, Heidelberg : Springer-Verlag, 1982. DOI: 10.1007/978-3-642-81897-4.
10. Lu C. Extension of Winograd Multiplicative Algorithm to Transform Size $N = p^2q$, p^2q^r and Their Implementation. Proceedings of International Conference on Acoustic / C. Lu, R. Tolimieri. – Speech, Signal Processing (ICASSP 89). Scotland, 1989.
11. Silverman H. F. An introduction to Programming the Winograd Fourier Transform algorithm (WFTA) / H. F. Silverman // IEEE Transactions on Acoustic, Speech, Signal Processing (ASSP). – 1977. – Vol. 25, No. 2. – P. 152–165. DOI: 10.1109/TASSP.1977.1162924
12. Patterson R. W. Fixed Point Error Analysis of Winograd Fourier Transform Algorithms / R. W. Patterson, J. H. McClellan // IEEE Transactions on Acoustic, Speech, Signal Processing (ASSP). – 1978. – P. 447–455. DOI: 10.1109/TASSP.1978.1163134.
13. Lavoie P. A high-speed CMOS implementation of the Winograd Fourier transform algorithm / P. Lavoie // IEEE Transactions of Signal Processing. – 1996. – Vol. 44, No. 8. – P. 2121–2126. DOI: 10.1109/78.533738

14. Zohar S. Faster Fourier Transformation: The Algorithm of S. Winograd / S. Zohar // Jet Propulsion Laboratory, JPL Publication 78–104, under NASA Contract No. NAS7-100. – 1979. – P. 1–93.
15. Zohar S. Winograd's discrete Fourier transform algorithm // In book Two-dimensional Digital Signal Processing. Transforms and Median Filters / T. S. Huang. – Berlin, Heidelberg, New York : Springer-Verlag, 1981. – P. 89–152.
16. Prots'ko I. The generalized technique of computation the discrete harmonic transforms / I. Prots'ko // Proceedings of the IVth International Conference MEMSTECH'2008, Polyana, 21–24 may 2008. – P. 101–102.
17. Thomas W. J. Abstract Algebra Theory and Applications / W. J. Thomas. – Stephen F. Austin State University, 2009.
18. Johnson S. G. A modified split-radix FFT with fewer arithmetic operations / S. G. Johnson, M. Frigo // IEEE Transactions of Signal processing. – 2007. – Vol. 55, No. 1. – P. 111–119. DOI: 10.1109/TSP.2006.882087
19. Duhamel P. Implementation of "Split-Radix" FFT Algorithms for Complex, Real, and Real-Symmetric Data / P. Duhamel // IEEE Transactions on Acoustic, Speech, and Signal Processing. – 1986. – Vol. 34, No. 2. – P. 285–295. DOI: 10.1109/TASSP.1986.1164811
20. ASIC Implementation of High Speed Processor for Calculating Discrete Fourier Transformation using Circular Convolution Technique / [P. Saha, A. Banerjee, A. Dandapat, P. Bhattacharyya] // WSEAS Transaction on Circuits and Systems. – 2011. – Vol. 10. – P. 278–288.

Article was submitted 10.11.2017.
After revision 25.12.2017.

Процько І. О.[1], Теслюк В. М.[2]

[1]Канд. техн. наук, доцент кафедри інформаційних систем і технологій, Національний університет «Львівська політехніка», Львів, Україна

[2]Д-р техн. наук, професор кафедри автоматизованих систем управління, Національний університет «Львівська політехніка», Львів, Україна

РОЗВИТОК АВПФ НА ОСНОВІ ТВІРНОГО МАСИВУ

Актуальність. Розглянуто метод ефективного обчислення ДПФ з використанням циклічних згорток для обсягів цілої степені два. Проаналізовано подальший розвиток Алгоритму Віноградa перетворення Фур'є на основі твірного масиву. Об'єктом дослідження є процес реформулювання базисної матриці ДПФ в блочно-циклічні структури. Предмет дослідження полягає в методі реформулювання базисної матриці ДПФ обсягів цілої степені два в блочно-циклічні структури.

Мета роботи полягає в проведенні аналізу особливостей розміщення ліво-циркулянтних підматриць в структурі базисної квадратної матриці W_N для обсягів перетворення $N = 2^i$ на основі використання твірних масивів.

Метод. У статті розглядається методика ефективного обчислення ДПФ з використанням циклічних згорток для обсягів рівних цілій степені два, що базується на циклічному розкладі підстановки. Запропоновано застосування твірного масиву для стислого опису і обчислення блочно-циклічної структури дискретної базисної матриці ДПФ обсягів цілої степені два.

Результати. Визначено загальну блочно-циклічну структуру дискретної базисної матриці для ефективного обчислення ДПФ з використанням циклічних згорток для обсягів цілої степені два на основі твірних масивів. Запропонована методика актуальна для паралельного програмування ДПФ та реалізації в паралельних обчислювальних системах.

Висновки. Загальна блочно-циклічна структура базисної матриці ДПФ регулярно нарощується зі збільшенням значення показника цілої степені два і рекомендується для використання на практиці при розробці ефективних засобів ДПФ. Перспективи подальших досліджень включатимуть формування блочно-циклічних структур базисної матриці ДПФ для довільних розмірів.

Ключові слова: швидке перетворення Фур'є, алгоритм Віноградa перетворення Фур'є, твірний масив, блочно-циклічна структура, циклічна згортка.

Процько И. Е.[1], Теслюк В. М.[2]

[1]Канд. техн. наук, доцент кафедры информационных систем и технологий, Национальный университет "Львівська политэхника", Львив, Украина

[2]Д-р техн. наук, професор кафедры автоматизованных систем управления, Национальный университет "Львівська политэхника", Львив, Украина

РАЗВИТИЕ АВПФ НА ОСНОВАНИИ ОБРАЗУЮЩЕГО МАССИВА

Актуальность. Рассмотрен метод эффективного вычисления ДПФ с использованием циклических сверток для длин равных целой степени два. Проанализировано дальнейшее развитие алгоритма Винограда преобразования Фурье на основании образующего массива. Об'єктом исследования есть процесс реформулирования базисной матрицы ДПФ в блочно-циклические структуры. Предметом исследования есть метод реформулирования базисной матрицы ДПФ для длин равных целой степени два в блочно-циклические структуры.

Цель работы заключается в проведении анализа особенностей размещения лево-циркулянтных подматриц в структуре базисной квадратной матрицы W_N для длин преобразования $N = 2^i$ с использованием образующих массивов.

Об'єктом дослідження є процес реформулювання базисної матриці ДПФ в блочно-циклічні структури. Предмет дослідження полягає в методі реформулювання базисної матриці ДПФ обсягів цілої степені два в блочно-циклічні структури.

Метод. В статье рассматривается методика эффективного вычисления ДПФ с использованием циклических сверток для длин равних целой степени два, основанная на циклическом разложении подстановки. Предложено использовать образующий массив для сжатого описания и вычисления блочно-циклической структуры дискретной базисной матрицы ДПФ длин равних целой степени два.

Результаты. Определена обобщенная блочно-циклическая структура дискретной базисной матрицы для эффективного вычисления ДПФ с использованием циклических сверток для длин равных целой степени два на основании образующих массивов. Предложенная методика актуальна для параллельного программирования ДПФ и реализации в параллельных вычислительных системах.

Выводы. Общая блочно-циклическая структура базисной матрицы ДПФ регулярным образом формируется с увеличением значения показателя целой степени два и рекомендуется для использования на практике при разработке эффективных средств ДПФ. Перспективы дальнейших исследований будут включать формирование блочно-циклических структур базовой матрицы ДПФ для произвольных размеров преобразования.

Ключевые слова: быстрое преобразование Фурье, алгоритм Винограда преобразования Фурье, образующий массив, блочно-циклическая структура, циклическая свертка.

REFERENCES

1. Duhamel P., Vitterli M. Fast Fourier Transform: A tutorial Review and a State of the Art, *Signal Processing*, 1990, Vol.19, pp. 259–299. DOI: 10.1016/0165-1684(90)90158-U.
2. Chu E., George A. Inside the FFT black box. Serial and Parallel Fast Fourier Transform Algorithms. Boca Raton, CRC Press LLC, 2000.
3. Tolimiery R., An M., Lu C. Algorithms for Discrete Fourier Transform and Convolution. New York, Springer-Verlag, (s.ed.), 1997, 267 p. DOI: 10.1007/978-1-4757-2767-8.
4. Prots'ko I., Rykmas R. Becoming of Discrete Harmonic Transform Using Cyclic Convolutions, *American Journal of Circuits, Systems and Signal Processing,* 2015, Vol. 1, pp. 114–119.
5. Rader C. M. Discrete Fourier Transforms When the Number of Data Samples is Prime, *Proceedings of IEEE*, 1968, Vol. 56, pp. 1107–1108. DOI: 10.1109/PROC.1968.6477.
6. Winograd S. On computing the discrete Fourier transform, *Proceedings National Academy of Science USA, Mathematics*, 1976, Vol. 73, No 4, pp. 1005–1006. DOI: 10.1073/pnas.73.4.1005.
7. Winograd S. On computing the discrete Fourier transforms, *Mathematics of Computation*, 1978, Vol. 32, pp. 175–199. DOI: 10.1090/S0025-5718-1978-0468306-4.
8. Blahut R. E. Fast algorithms for signal processing. Cambridge, University Press, 2010. DOI: 10.1017/CBO9780511760921.
9. Nussbaumer H. J. Fast Fourier Transform and Convolution Algorithms. Berlin, Heidelberg, Springer-Verlag, 1982. DOI: 10.1007/978-3-642-81897-4.
10. Lu C., Tolimieri R. Extension of Winograd Multiplicative Algorithm to Transform Size $N = p^2q$, p^2q^r and Their Implementation. Proceedings of International Conference on Acoustic, Speech, Signal Processing (ICASSP 89). Scotland, 1989.
11. Silverman H. F. An introduction to Programming the Winograd Fourier Transform algorithm (WFTA), *IEEE Transactions on Acoustic, Speech, Signal Processing (ASSP)*, 1977, Vol. 25, No. 2, pp. 152–165. DOI: 10.1109/TASSP.1977.1162924
12. Patterson R. W., McClellan J. H. Fixed Point Error Analysis of Winograd Fourier Transform Algorithms, *IEEE Transactions on Acoustic, Speech, Signal Processing (ASSP)*, 1978, pp. 447–455. DOI: 10.1109/TASSP.1978.1163134.
13. Lavoie P. A high-speed CMOS implementation of the Winograd Fourier transform algorithm, *IEEE Transactions of Signal Processing*, 1996,Vol. 44, No. 8, pp. 2121–2126. DOI: 10.1109/78.533738
14. Zohar S. Faster Fourier Transformation: The Algorithm of S. Winograd, *Jet Propulsion Laboratory, JPL Publication 78-104*, under NASA Contract No. NAS7-100, 1979, pp. 1–93.
15. Zohar S. Winograd's discrete Fourier transform algorithm, *In book Two-dimensional Digital Signal Processing. Transforms and Median Filters*. Berlin, Heidelberg, New York, Springer-Verlag, 1981, pp. 89–152.
16. Prots'ko I. The generalized technique of computation the discrete harmonic transforms, *Proceedings of the IVth International Conference MEMSTECH'2008, Polyana*, 21–24 may 2008, pp.101–102.
17. Thomas W. J. Abstract Algebra Theory and Applications. Stephen F. Austin State University, 2009.
18. Johnson S. G., Frigo M. A modified split-radix FFT with fewer arithmetic operations, *IEEE Transactions of Signal processing*, 2007, Vol. 55, No. 1, pp. 111–119. DOI: 10.1109/TSP.2006.882087
19. Duhamel P. Implementation of "Split-Radix" FFT Algorithms for Complex, Real, and Real-Symmetric Data, *IEEE Transactions on Acoustic, Speech, and Signal Processing*, 1986, Vol. 34, No. 2, pp. 285–295. DOI: 10.1109/TASSP.1986.1164811
20. Saha P., Banerjee A., Dandapat A., Bhattacharyya P. ASIC Implementation of High Speed Processor for Calculating Discrete Fourier Transformation using Circular Convolution Technique, *WSEAS Transaction on Circuits and Systems*, 2011, Vol. 10, pp. 278–288.

UDC 004.056

Faure E. V.[1], Shcherba A. I.[2], Kharin A. A.[3]

[1]PhD, Associate Professor, Associate Professor of Department of Information Security and Computer Engineering, Cherkasy State Technological University, Cherkasy, Ukraine
[2]PhD, Associate Professor, Head of Department of Applied Mathematics, Cherkasy State Technological University, Cherkasy, Ukraine
[3]Post-graduate student of Department of Information Security and Computer Engineering, Cherkasy State Technological University, Cherkasy, Ukraine

FACTORIAL CODE WITH A GIVEN NUMBER OF INVERSIONS

Context. Factorial coding with data recovery by permutation provides complex information protection from unauthorized reading and errors in communication channel and has the property of self-synchronization. At the same time, such coding does not allow to detect all low-weight errors that leads to a relatively small increase of reliability.

The purpose of this work is to develop and study the method of factorial coding with a given number of inversions aimed at increasing the reliability of information transmission, by introducing additional redundancy by choosing a class of permutations that satisfy the given criterion.

Method. The main idea of the proposed method is to introduce artificial redundancy by reducing the cardinality of used permutations. Such an approach makes it possible to select from a whole set of permutations a class that possesses the necessary, pre-assigned properties. It was suggested to use the correspondence of the number of permutation inversions to a given class of residues as a sign of belonging to the permutation class in use. A theoretical evaluation of code parameters was performed.

Results. Signal-code constructions for the factorial code with a given number of inversions under the order of permutations $M = 8$ are constructed. For each of the possible classes, the cardinality, code rate, estimates of the probability of undetected error and the relative transmission rate for systems with decision feedback and independent bit errors were investigated. It is shown that the code parameters are not invariant with respect to the selected class of residues for a given modulus. The structural schemes of encoding and decoding devices are developed.

Conclusions. The method of factorial coding with data recovery by permutation has been developed. The use of permutations with a number of inversions that belongs to the selected class of residues allowed increasing the reliability of transmission in exchange for the loss of code rate.

Keywords: factorial coding, permutation, inversion, redundancy, class of residues, reliability of transmission, relative transmission rate.

NOMENCLATURE

FCDR – Factorial Code with Data Recovery by Permutation;

FCGNI – Factorial Code with a Given Number of Inversions;

DF – decision feedback;

SCC – signal-code construction;

$A(x)$ – data word in a binary form;

$B_M(q,R)$ – class of permutations with inversion numbers belonging to the class of residues $\overline{R}_q$;

$f_j(i)$ – number of i-bit errors that can transform the j-th permutation of SCC into any other permutation of the same construction;

$i_M(q,R)$ – amount of information carried by one permutation of the class $B_M(q,R)$;

k – number of binary symbols in information block;

$k_M(q,R)$ – length of binary data word to be converted into a permutation of the class $B_M(q,R)$;

l_r – number of bits for encoding a single symbol of permutation;

M – permutation order;

n – number of bits in a data block;

$N_M(\omega)$ – number of different permutations of the order with a number of inversions equal to ω;

$P_{ud}(FCGNI, p_0)$ – probability of FCGNI undetected errors;

p_0 – bit error probability;

Q – probability of codeword error-free reception;

$R_{FCDR}(x)$ – FCDR codeword in a binary form;

$\overline{R}_q$ – residue class modulo q;

r_{FCDR} – number of bits in FCDR codeword;

α – code redundancy factor (by cardinality);

ν_0 – relative transmission rate;

ν_1 – code rate;

ν_2 – dynamic component of transmission rate loss due to retransmission of data in error;

$W_M(q,R)$ – permutations class cardinality;

ω – number of inversions in permutation.

INTRODUCTION

Rapid growth of modern society computerization level leads to the need to ensure the joint work of many computers including in open communication networks. Particularly acute is the need to ensure safe operation in open computer

© Faure E. V., Shcherba A. I., Kharin A. A., 2018
DOI 10.15588/1607-3274-2018-2-16

networks in the exchange of financial information, as well as information of national importance, such as information from the state's law enforcement agencies. All this leads to the need for complex information security that provides for the following tasks:

– the task of protection against information unauthorized reading (through its encryption);

– the task of ensuring information integrity that includes:

– information protection against errors due to noise in communication channels;

– protection against obtrusion of false information.

All these tasks are solved in existing computer and telecommunication systems and networks. However, each of them is solved independently from each other, by different means created at different times, and based on an individual mathematical basis. A consequence of this approach is an increase of the introduced redundancy. This entails a reduction of data channel bandwidth and leads to an increase of the load of computational basis used to realize the assigned tasks.

Therefore, the development of methods and tools for complex information security based on a single procedure for its processing that ensure a high level of protection and transmission reliability is an actual problem.

To date, a relatively large volume of extensive research on factorial coding methods has been performed from the viewpoint of providing complex information protection. The analysis of works [1–9] reveals the properties of factorial codes and allows planning further research in this scientific direction. In particular, the paper [5] is of exceptional interest. It examined the factorial code with data recovery by permutation (FCDR).

The most important property of FCDR is its ability to provide complex data protection at acceptable values of transmission rate loss. At the same time, FCDR does not detect an even number of bit errors that lead to the transformation of one permutation into another. In channels of acceptable quality (where $\lambda = np_0 \le 1$), the probability of FCDR erroneous decoding is determined by the intensity of low-weight errors. As shown in [1], this probability decreases rapidly with increasing of error weight. Therefore, eliminating low-weight errors from the set of FCDR non-detectable errors will reduce the undetected error probability.

The purpose of this work is to increase the reliability of information transmission when using FCDR. This purpose can be achieved by introducing additional redundancy by choosing a class of permutations that satisfy the given criterion.

1 PROBLEM STATEMENT

Suppose that the information from the source goes to the encoder input in blocks $A(x)$ of k bits. In this case, the cardinality of information words set is 2^k. FCDR realizes a bijective transformation of the set of information words $A(x)$ into an allowed set of permutations of order M $\left(M! \ge 2^k\right)$. Reducing the cardinality of information words of used permutations while preserving their order M makes it possible to choose from $M!$ permutations only those that satisfy the preassigned set of properties.

The task of developing a method of information coding based on factorial coding with data recovering by permutation consists in determining the effective criterion for choosing the class of used permutations, which allows one to vary flexibly the parameters of reliability of information transmission, and also in the analysis of the structure of allowed set of permutations and its influence on undetected error probability and relative transmission rate.

2 REVIEW OF THE LITERATURE

According to [5], the permutation π after the symbols encoding by a uniform binary code is represented as a data block $R_{FCDR}(x)$ of r_{FCDR} bits, where

$$r_{FCDR} = l_r \cdot M, \tag{1}$$

l_r is a number of bits for encoding a single symbol of permutation.

For uniform binary code

$$l_r = \lceil \log_2 M \rceil, \tag{2}$$

where $\lceil a \rceil$ is the smallest integer greater than or equal to a (function 'ceiling' [10] of a).

The length of FCDR code combination n_{FCDR} corresponds to the number of bits in the representation of the permutation π of the order M in a binary form: $n_{FCDR} = r_{FCDR}$.

A necessary (but not sufficient) condition for the bijectivity of the mapping between the set $\{A(x)\}$ and the set of allowed permutations $\{\pi\}$ is the condition of choosing an order M according to the inequality $M! \ge 2^k$. Since $M! = 2^k$ is possible only for $k = 1$, for practical applications we shall proceed from the condition $M! > 2^k$. It follows that

– the maximum length of a binary information word to be converted into a permutation of the order M is given by

$$k_{\max} = \lfloor \log_2 M! \rfloor; \tag{3}$$

– $M! - 2^k$ permutations are not used by the source, and FCDR is a redundant code. Code redundancy factor (by cardinality) is defined in [6] as:

$$\alpha = M! / 2^k. \tag{4}$$

When the value of the number of bits of the information word is chosen according to (3), the minimum redundancy factor (by cardinality) is provided for the given M. In this case,

$$1 < \alpha < 2. \tag{5}$$

An error-detecting factorial code that satisfies the condition (5) will be called a factorial code with natural redundancy. The value of redundancy factor (by cardinality) is $\alpha_{nat} = M! / 2^{k_{\max}}$.

It was shown in [6] that equality (4) implies that the decrease in the length of the information word per Δk bits for a fixed M leads to the increase in the redundancy factor (by cardinality) by a factor of $2^{\Delta k}$. A code with $k < k_{\max}$ for a given M will be called a factorial code with introduced redundancy. In this case, the value of redundancy factor (by cardinality) is $\alpha_{intr} = 2^{\Delta k}$, where $\Delta k = k_{\max} - k$.

As a result of research performed in [5, 8], it is shown that FCDR provides protection against unauthorized reading of data if the law of transformation of source words into permutation is kept in secret. Meanwhile, the protection of information from errors in communication channel is provided by the properties of permutations. The most important of these properties is that the permutation π of order M is a sequence of symbols of the set $\{0;1;2;\ldots;M-1\}$. The position of the symbols in permutation is determined by the information word. In addition, every symbol in the permutation is used only once. Checking the occurrence of each of the symbols only once at the receiving station ensures that all odd number of bit errors and part of even number of bit errors are detected in the received block [5]. This check will be referred to as the data block validation.

The advantage of FCDR is also that it has the property of self-synchronization. This eliminates the need for a combination of frame synchronization to be entered in the data block.

The disadvantage of FCDR is that this code does not detect even number of bit errors, which lead to the transformation of one permutation into another.

In [6] it is shown that if for a given k the value of M is calculated according to the condition $(M-1)! < 2^k < M!$, then inequality $\alpha > 2$ can take place. Then it is possible to insert additional check bits into information part before forming the checksum. This allows to increase the transmission reliability while maintaining the permutation order and the code rate. On the other hand, if the value of M is given in a data transmission system and $k = k_{\max} = \lfloor \log_2 M! \rfloor$, the transmission reliability can be increased due to artificially introduced redundancy. This redundancy is provided by reducing the size of the data block at the encoder input and introducing additional check bits. The results presented in [6] for $k \le 1024$ and $p_0 = 10^{-3}$ indicate an increase in the energy gain due to the application of the proposed method by up to 1.6 dB. At the same time, the method proposed in [6] does not allow to detect some errors, including 2-bit errors, that lead to the transformation of one permutation into another (permutation of the permutation symbols).

3 MATERIALS AND METHODS

The proposed coding method provides the artificial redundancy by reducing the cardinality of used permutations. Consequently, the size of the source word also decreases (in comparison with its maximum possible value): $k = k_{\max} - \Delta k$, $\Delta k > 0$. We consider it obvious that an increase in redundancy should lead to an increase in the transmission reliability.

All the decisions made in this work are oriented to the creation of data transmission systems operating over a communication channel with independent bit errors. Primarily, this applies to block transmission systems where error correction is performed by retransmission of data in error (binary symmetric decision feedback (DF) systems). Such systems include most general and special purpose systems operating over wired telephone channels, radio relay, microwave, satellite and fiber-optic communication channels.

The main idea of the work is to use a number of inversions in permutation as a sign of its belonging to an allowed set. So, if we use only even (odd) permutations to transfer source words, then:

– the cardinality of the allowed set of source words will be halved (exactly half of the permutations of the set of $M!$ permutations is even (odd));

– the decoder will detect all the errors that lead to the transformation of a permutation into another permutation and change its parity. Accordingly, the decoder will not detect errors that lead to the transformation of a permutation into another permutation and do not change its parity.

Note that the transformation of one permutation into another is equivalent to a permutation of its symbols and can be represented as a product of transpositions. Therefore, the transformation of one FCDR codeword into another can be represented as a finite number of consecutive transpositions. Each transposition applied to a permutation changes its parity. Therefore, the consistent application of an odd number of transpositions changes the parity of a permutation. The consistent application of an even number of transpositions does not change the parity of a permutation. This means that the use of only even (odd) permutations allows to detect errors that lead to a transformation equivalent to an odd number of consecutive transpositions over the codeword.

Note also that if the source uses only even (odd) permutations, the decoder will detect all 2-bit errors in a FCDR codeword. This is because a 2-bit error is not detected by FCDR if and only if it generates a transposition of symbols in the codeword. Since the transposition changes the parity of a permutation, it will be detected by the decoder. In this case, an undetected decoding error can be estimated using the expression (2) from [5], where instead of the estimate $f_{per}(2) \le l_r \cdot M/2$ it should be used $f_{per}(2) = 0$.

Thus, it is possible to increase the transmission reliability of systems with FCDR by reducing the cardinality of permitted permutations and detecting their transpositions in the process of transportation.

In turn, reducing the cardinality of the set of permutations – information carriers – by half leads to a need to reduce the block length by one bit and, accordingly, to reduce the code rate. Therefore, the use of permutations for information transferring, in which a number of inversions satisfies the specified requirements, allows exchanging the code rate for the transmission reliability.

In view of the foregoing, the proposed code will be called a factorial code with a given number of inversions (FCGNI). We define the rule for choosing an allowed set of permutations based on the number of inversions. For this, we consider the theoretical basis for constructing FCGNI.

A number of inversions $\omega = inv(\pi)$ in the permutation π of the order M satisfies the condition $0 \le \omega \le 0.5 \cdot M \cdot (M-1)$. Each number ω of possible inversions corresponds to a frequency number $N_M(\omega)$. Due to the definition of frequency numbers, the next expression is valid:

$$\sum_{\omega=0}^{0.5M(M-1)} N_M(\omega) = M!,$$

and besides $N_M(0) = N_M(0.5M(M-1)) = 1$.

In addition, the frequency number $N_M(\omega) = 0$ for $\omega < 0$ and $\omega > 0.5M(M-1)$.

It is shown in [11] that the distribution of inversions on all permutations of a fixed length coincides with the distribution of their major index. That is, the number $N_M(\omega)$ of permutations of order M with ω inversions is the same as the number of permutations of order M with major index equal to ω. These numbers are known as Mahonian numbers. A stronger result is valid: the number of permutations of order M with major index k and ω inversions is the same as the number of permutations of order M with major index ω and k inversions, that is, the two statistics are equidistributed.

According to [12, 13], the following recurrence relation is valid:

$$N_{M+1}(\omega) = N_M(\omega - M) + N_M(\omega + 1 - M) + \\ + \ldots + N_M(\omega) = \sum_{i=\omega-M}^{\omega} N_M(i), \quad (6)$$

wherein $N_0(\omega) = 0$ if $\omega \ge 1$.

In accordance with [12, 14, 15], the frequency numbers are coefficients in the expansion $\prod_{i=0}^{M-1} \left(1 + x + \ldots + x^i\right)$, i.e.

$$\prod_{i=0}^{M-1} \left(1 + x + \ldots + x^i\right) = \sum_{\omega=0}^{0.5M(M-1)} N_M(\omega) \cdot x^{\omega}.$$

Online Encyclopedia of Integer Sequences (OEIS) [16] contains the sequence of McMahon numbers $N_M(\omega)$ for $M \in [1,50]$.

Properties of frequency numbers $N_M(\omega)$.

1. For $M \ge 2$ and $\omega : 2 \le \omega \le C_M^2 - 1$ all the frequency numbers $N_M(\omega) \ge M - 1$.

2. The symmetry property for $M \ge 2$: $N_M(\omega) = N_M\left(C_M^2 - \omega\right)$, $N_M(1) = M - 1$.

3. $\sum_{\omega=0}^{0.5M(M-1)} (-1)^{\omega} N_M(\omega) = 0$.

4. $\sum_{\omega=0}^{0,5M(M-1)} \omega N_M(\omega) = \frac{1}{2} C_M^2 M! = \sum_{\pi} inv(\pi)$.

5. If $\omega_1 < \omega_2 \le \left[0,5 \cdot C_M^2\right]$, then $N_M(\omega_1) < N_M(\omega_2)$.

6. If the number C_M^2 is even, i.e. $C_M^2 = 2l$, then $\max_{\omega} N_M(\omega) = N_M(l)$. If the number C_M^2 is odd, i.e. $C_M^2 = 2l + 1$, then $\max_{\omega} N_M(\omega) = N_M(l) = N_M(l+1)$.

6. The recurrence formula (6) can be reduced to the following form:

$$N_{M+1}(\omega) = N_{M+1}(\omega - 1) + N_M(\omega) - N_M(\omega - 1 - M). \quad (7)$$

Remark. If $\omega < M + 1$, formula (7) has the form $N_{M+1}(\omega) = N_{M+1}(\omega - 1) + N_M(\omega)$, which corresponds to the expressions (I) from [12] and (9) from [17].

We define the residue modulo q of a number of inversions in a permutation of the order M: $R = |\omega|_q$, where $2 \le q \le 0,5 \cdot M \cdot (M-1)$ and $0 \le R \le q - 1$.

The set of various permutations π of the order M with a number of inversions of the residue class $\overline{R}_q$ forms a subset (class) of permutations $B_M(q,R) = \left\{\pi : |inv(\pi)|_q = R\right\}$. Depending on values of q and R, cardinalities $W_M(q,R)$ of classes $B_M(q,R)$ are calculated as follows:

$$W_M(q,R) = \sum_{j=0}^{\left[(0.5 \cdot M \cdot (M-1) - R)/q\right]} N_M(\omega = jq + R). \quad (8)$$

It was shown in [17] that if $q \le M$, then the cardinalities of classes $B_M(q,R)$ are invariant with respect to R and are equal to

$$W_M(q,R) = M!/q. \quad (9)$$

Note that the cardinalities $W_M(q,R)$ of the classes $B_M(q,R)$ can differ significantly for the constant modulus q value. Therefore, in order to maximize the code rate, it is necessary to use the permutation classes of maximum

cardinality for information transmission. In this case, the amount of information carried by one permutation will be equal to

$$i_M(q,R) = \log_2\left(W_M(q,R)\right), \quad (10)$$

and

$$k_M(q,R) \le \lfloor i_M(q,R) \rfloor. \quad (11)$$

At the same time, it does not follow from this that the use of the permutation class $B_M(q,R)$ of maximum cardinality provides the greatest energy efficiency. By energy efficiency, we mean the difference in signal levels at the input of the FCGNI receiver and some other receiver of information, providing the same probability of error-free reception.

All classes $B_M(q,R)$ for $q \le M$ have an equal cardinality. On other hand, the cardinalities of classes $B_M(q,R)$ where $M < q \le 0.5 \cdot M \cdot (M-1)$ are not equal and should be subject to experimental evaluation.

The above theoretical basis of information factorial coding determines the way of FCGNI constructing – selecting for information transmission only those permutations, which numbers of inversions belong to a selected class of residues $\overline{R}_q$. As shown above, if $q = 2$, the code detects all 2-bit errors and a part of errors with higher weight. Selecting a modulus q other than two $\left(2 < q \le 0.5 \cdot M \cdot (M-1)\right)$ will result in a different distribution of detectable and undetectable errors. This requires an informed choice of class $B_M(q,R)$. Note that if $q : |q|_2 = 0$, the code will detect all odd permutations and a part of even permutations.

We will accept $K = k_M(q,R)$. Then the probability of undetected error

$$P_{ud}(FCGNI, p_0) = 2^{-K} \sum_{j=1}^{2^K} \sum_{i=2}^{n} f_j(i)\, p_0^i (1-p_0)^{n-i}. \quad (12)$$

An integral indicator of data transmission quality is a relative transmission rate ν_0 [18]. It is calculated for a DF system as follows:

$$\nu_0 = \nu_1 \nu_2, \quad (13)$$

where $\nu_1 = k/n$.

According to [18], for the simplest DF system

$$\nu_2 = Q + P_{ud}(FCGNI, p_0),$$

where $Q = (1-p_0)^n$.

The structural scheme of the FCGNI encoder is shown in Fig. 1.

From a data source, an information word of k bits is entered in the buffer register 1. Then a source word is transmitted to the block 2 for forming a permutation with a number of inversions of a given class of residues. This permutation, after encoding its symbols with a binary code, is transmitted through the modulator 3 to a receiving station via a communication channel.

Generation of a permutation with a number of inversions of a given class of residues can be performed, for example, by creating a table that consists of 2^k rows. In each of them one of the permutations with a number of inversions $\omega : |\omega|_q = R$ is written.

The number of different tables that can be constructed is equal to

$$\mu(M,q,R,k) = C^{2^k}_{W_M(q,R)} \cdot k!. \quad (15)$$

At the same time, it is created a table that links permutations with information words for decoder. If these tables are kept secret, then the prerequisites for information encryption are created, and the expression (15) determines the key space cardinality. With this approach, the transmitting part of the system with FCGNI is a ROM where tables are stored. A table is selected by a session key. A permutation is selected by a k-bit source word that defines an address of a cell storing this permutation.

The structural scheme of the FCGNI decoder is shown in Fig. 2.

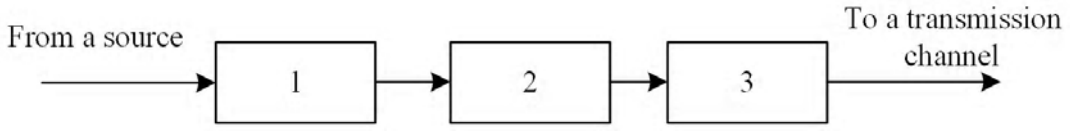


Figure 1 – The structural scheme of the FCGNI encoder

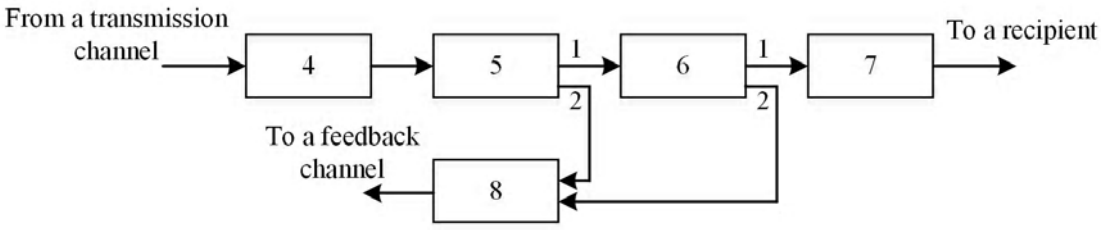


Figure 2 – The structural scheme of the FCGNI decoder

The demodulator 4 allocates data block (which can be affected by interference) from a received sequence (signal and noise mixture). Then the received data block comes to the block of permutation estimation 5. In block 5, the permutation correctness is evaluated – it is verified that each of the symbols $\{0,1,2,...,M-1\}$ is contained only once and that the permutation belongs to the allowed part of permutation set.

If the correctness condition is not satisfied, the permutation estimation block 5 erases the received data block and, by the bus 2, generates a command of retransmission of data in error for the request generation block 8. The request generation block 8 transmits this command to the transmitting station via a feedback channel.

If the correctness condition is satisfied, the permutation estimation block 5, by the bus 1, delivers the received permutation to the block 6 of estimation a number of permutation inversions. If the number of inversions in the received permutation does not belong to a given class of residues, then the block 6 of estimation a number of permutation inversions, by the bus 2, generates a command of retransmission of data in error for the request generation block 8.

If the received permutation has retained the number of inversions belonging to a given class of residues, then, by bus 1 of the block 6 of estimation a number of permutation inversions, the received permutation is transmitted to the permutation decoding block 7. This block performs the inverse transformation of the permutation into an information word using the corresponding table. The decoded permutation from the output of the block 7 is given to a recipient.

Thus, FCGNI detects:

– all errors that transform the transmitted permutation into non-permutation;

– all errors that transform the transmitted permutation into a permutation with a number of inversions that does not belong to a given class of residues.

4 EXPERIMENTS

In Table 1, for $M=8$ $(n=24)$ there are shown the results of experimental estimates of the maximum cardinality $W_{\max}$ of the permutation class $B_M(q,R)$; the value of R at which it is achieved; the value of $k_{\max}=\left[i_M(q,R)\right]$; the code rate $\nu_{1\max}$ corresponding to $k_{\max}$.

Remark. The EC symbols mean that all the classes $B_M(q,R)$ for a given q have an equal cardinality that is defined by (9).

We estimate the probability of undetected error (12), the code rate and the relative transmission rate (13). To do this, we randomly select from each class $B_8(q,R)$ the 2^K permutations that form the signal-code construction (SCC).

5 RESULTS

Fig. 3 shows the graph of dependence of the estimated probability of FCGNI undetected error on the modulus q for $p_0=10^{-3}$.

The dotted line in the graph indicates the estimated probability of FCDR undetected error for $M=8$, $P_{ud}\left(FCDR,10^{-3}\right)=1{,}18\cdot10^{-5}$. In addition, it is presented the graph for the values of R from Table 1 providing the maximum cardinality of the class $B_8(q,R)$.

Table 1 – Classes $B_8(q,R)$ of maximum cardinality for $q\in[2,28]$

q	2	3	4	5	6	7	8	9	10
$W_{\max}$	20160	13440	10080	8064	6720	5760	5040	4522	4184
R	EC	EC	EC	EC	EC	EC	EC	5	4
$k_{\max}$	14	13	13	12	12	12	12	12	12
$\nu_{1\max}$	0,583	0,540	0,540	0,500	0,500	0,500	0,500	0,500	0,500
q	11	12	13	14	15	16	17	18	19
$W_{\max}$	3988	3890	3850	3838	3836	3836	3836	3836	3836
R	3	2	1	0	14	14	14	14	14
$k_{\max}$	11	11	11	11	11	11	11	11	11
$\nu_{1\max}$	0,458	0,458	0,458	0,458	0,458	0,458	0,458	0,458	0,458
q	20	21	22	23	24	25	26	27	28
$W_{\max}$	3836	3836	3836	3836	3836	3836	3836	3836	3836
R	14	14	14	14	14	14	14	14	14
$k_{\max}$	11	11	11	11	11	11	11	11	11
$\nu_{1\max}$	0,458	0,458	0,458	0,458	0,458	0,458	0,458	0,458	0,458

If values of R are chosen to ensure the maximum transmission reliability, the graph of dependence of the estimated probability of FCGNI undetected error on the modulus q for $p_0 = 10^{-3}$ will take the form shown in Figure 4. In addition, it is presented the graph for R providing the maximum transmission reliability of the class $B_8(q, R)$.

Note that in the general case, the classes with the maximum transmission reliability and the classes with the minimum cardinality are not the same for a given $q \in [2, 28]$.

A comparison of code rates providing the maximum SCC cardinality and the maximum reliability on $q \in [2, 28]$ are shown in Fig. 5.

The graph of dependence of the estimated relative transmission rate for FCGNI on the value of q for $M = 8$ and $p_0 = 10^{-3}$ is shown in Fig. 6. For each $q \in [2, 28]$ it is shown the estimated maximum relative transmission rate. The values of R in this case are given in Table 2.

The dashed-dot line on the graph indicates the estimated FCDR relative transmission rate for $M = 8$, $\nu_0\left(FCDR, 10^{-3}\right) = 0,610$.

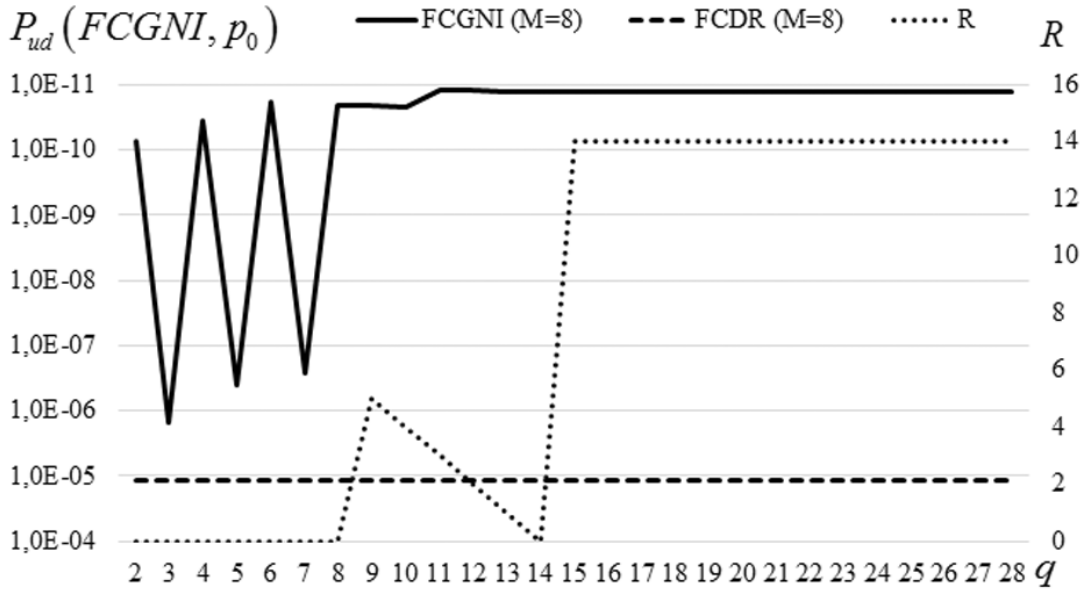


Figure 3 – Graph of dependence of the estimated probability of FCGNI undetected error on the modulus q for $p_0 = 10^{-3}$ (maximum code rate in the class $B_8(q, R)$)

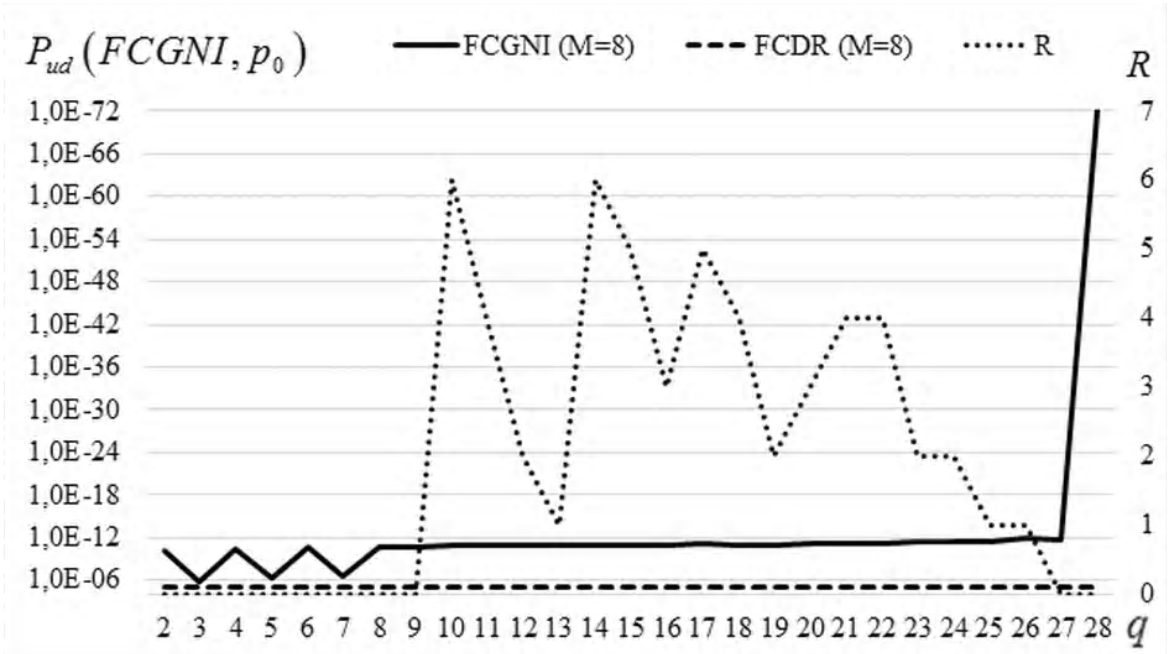


Figure 4 – Graph of dependence of the estimated probability of FCGNI undetected error on the modulus q for $M = 8$ and $p_0 = 10^{-3}$ (maximum transmission reliability in the class $B_8(q, R)$)

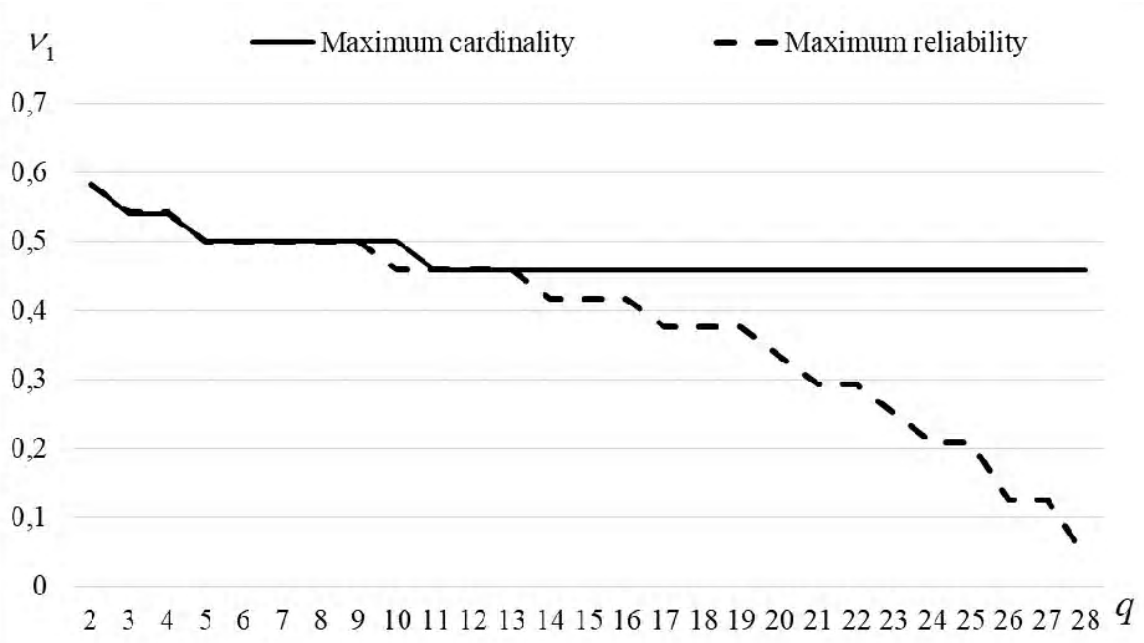


Figure 5 – Graphs of code rates providing the maximum SCC cardinality and the maximum reliability on $q \in [2,28]$ for $M = 8$ and $p_0 = 10^{-3}$

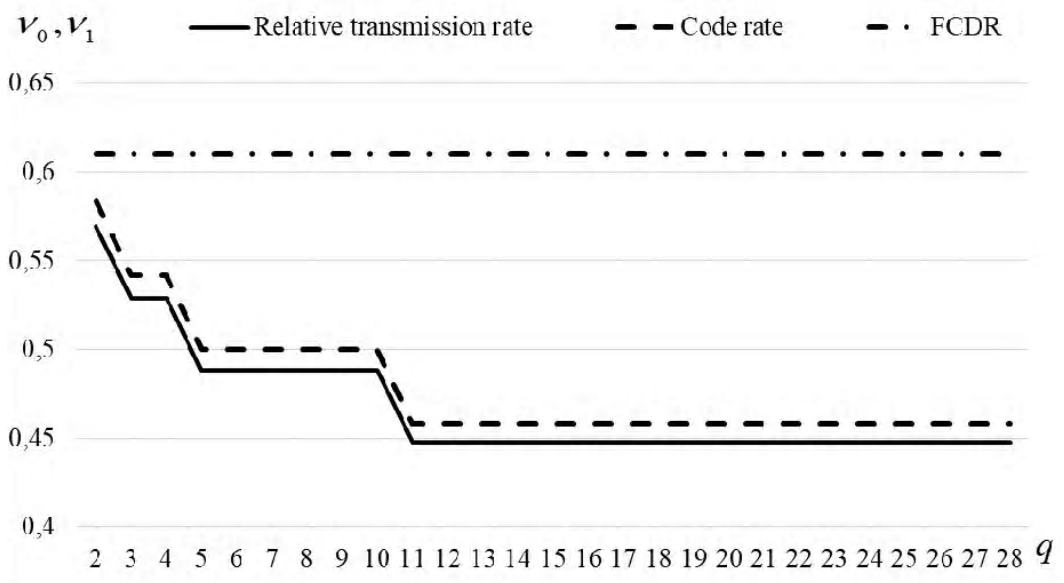


Figure 6 – Graph of dependence of the estimated FCGNI relative transmission rate on $q \in [2,28]$ for $M = 8$ and $p_0 = 10^{-3}$ (maximum relative transmission rate in the class $B_M(q,R)$)

Table 2 – Values of R for the maximum relative transmission rate for $q \in [2,28]$

q	2	3	4	5	6	7	8	9	10
R	ER	ER	ER	ER	ER	ER	ER	ER	3–5
q	11	12	13	14	15	16	17	18	19
R	ER	ER	ER	0–5,9–13	0–4,9–14	0–2,10–15	0,1,10–16	0,10–17	10–18
q	20	21	22	23	24	25	26	27	28
R	10–18	10–18	10–18	10–18	10–18	10–18	10–18	10–18	10–18

Remark. The ER symbols mean that all the classes $B_M(q,R)$ for a given q have an equal relative transmission rate.

6 DISCUSSION

The graph of Fig. 6 indicates that for $M = 8$ and $p_0 = 10^{-3}$ the relative transmission rate reaches the maximum values at the maximum values of the code rate. This is explained by the fact that $Q \gg P_{ud}\left(FCGNI, p_0\right)$ and, according to (14), $\nu_2 \approx Q$. Therefore, it is advisable to use $B_8\left(2, R\right)$, $R \in \{0;1\}$ to maximize the relative transmission rate for $M = 8$ and $p_0 = 10^{-3}$. For others M and $p_0 = 10^{-3}$, the values of q and R can vary.

It also follows from Figure 6 that FCDR exceeds FCGNI in the relative transmission rate for $M = 8$ and $p_0 = 10^{-3}$. Additional studies show that for $M = 8$ FCGNI has a gain in this indicator at $p_0 > 0.1$. At the same time, Figures 3 and 4 indicate that FCGNI is aimed at reducing the probability of undetected error (for example, for $q = 2$ – by more than five orders in comparison with FCDR). This can be useful in systems with high requirements for the probability of "false alarm".

Note that the presented coding method has a disadvantage inherent in all block ciphers. Identical plaintext blocks are converted into identical ciphertext blocks. Therefore, in order to eliminate the statistical redundancy and reduce the probability of identical blocks appearance, it is advisable to compress the plaintext before transformation.

CONCLUSIONS

The problem of increasing the reliability of information transmission using FCDR is solved.

The scientific novelty of the work is as follows. The method of factorial coding with data recovery by permutation has been further developed by using in SCC permutations with a number of inversions from a given class of residues. This makes it possible to reduce the number of undetected errors and, accordingly, to increase the transmission reliability.

The practical significance of the obtained results lies in the experimental estimates of FCGNI parameters, as well as in the developed structural schemes of encoding and decoding devices, which make it possible to carry out their practical implementation.

It is shown that by choosing the appropriate class of residues for the number of inversions ω in a permutation π, one can select a desired value of energy gain in exchange for the loss of the relative transmission rate. For example, the choice of the class $B_8\left(2,1\right)$ provides for $M = 8$ and $p_0 = 10^{-3}$ the probability of undetected error $P_{ud}\left(FCGNI, p_0\right) = 7.24 \cdot 10^{-11}$ and the code rate $\nu_1 = 0.583$. For comparison, the same indicators for FCDR – $P_{ud}\left(FCDR, p_0\right) = 1.18 \cdot 10^{-5}$, $\nu_1 = 0.625$. At the same time, the task of developing the principles for choosing a permutations set of a given class that forms an optimal structure of SCC is the subject of further research.

ACKNOWLEDGMENTS

The work was carried out as a result of joint research of the Department of Information Security and Computer Engineering and the Department of Applied Mathematics of Cherkasy State Technological University on the topic "Models, methods and means of complex information protection based on factorial coding" within the framework of the research work "Synthesis of cryptographic transformation operations with given characteristics" (state registration number 0116U008714) of the Department of Information Security and Computer Engineering of Cherkasy State Technological University.

The authors are grateful to the Associate Professor of the Department of Information Security and Computer Engineering of Cherkasy State Technological University Ph.D., Associate Professor Shvydkyi Valerii Vasylovych for comprehensive support in writing the work and useful discussions of the results.

REFERENCES

1. Фауре Э. В. Контроль целостности информации на основе факториальной системы счисления/ Э. В. Фауре, В. В. Швыдкий, А. И. Щерба // Journal of Baku Engineering University – Mathematics and Computer Science. – 2017. – Vol. 1, № 1. – P. 3–13.
2. Фауре Э. В. Комбинированное факториальное кодирование и его свойства / Э. В. Фауре, В. В. Швыдкий, В. А. Щерба // Радіоелектроніка, інформатика, управління. – 2016. – № 3. – С. 80–86. DOI: 10.15588/1607-3274-2016-3-10.
3. Пат. 107655 Україна, МПК G06F21/64 (2013.01), H04L1/16 (2006.01). Спосіб контролю цілісності інформації / Рудницький В. М., Фауре Е. В., Швидкий В. В., Щерба А. І. ; заявник та патентовласник Черкаський державний технологічний університет. – № a201505937 ; заявл. 16.06.2015; опубл. 24.06.2016, Бюл. № 12.
4. Пат. 107657 Україна, МПК H03M13/09 (2006.01), H04K1/06 (2006.01), G09C1/06 (2006.01). Спосіб комбінованого кодування інформації / Рудницький В. М., Фауре Е. В., Швидкий В. В., Щерба А. І. ; заявник та патентовласник Черкаський державний технологічний університет. – № a201508148 ; заявл. 17.08.2015; опубл. 24.06.2016, Бюл. № 12.
5. Фауре Э. В. Факториальное кодирование с восстановлением данных / Э. В. Фауре // Вісник Черкаського державного технологічного університету. – 2016. – № 2. – С. 33–39. DOI: 10.24025/bulletinchstu.v1i2.82932.
6. Фауре Э. В. Метод повышения эффективности факториального кодирования с восстановлением данных / Э. В. Фауре // Вісник Черкаського державного технологічного університету. – 2016. – №4. – С. 57–61.
7. Фауре Э. В. Факториальное кодирование с несколькими контрольными суммами / Э. В. Фауре // Вісник Житомирського державного технологічного університету. – 2016. – №3. – С. 104–113. DOI: 10.26642/tn-2016-3(78)-104-113.
8. Пат. 117004 Україна, МПК H03M13/09 (2006.01), H04L1/16 (2006.01), G04C1/06 (2006.01). Спосіб факторіального кодування з відновленням даних / Фауре Е. В., Харін О. О., Швидкий В. В., Щерба А. І. ; заявник та патентовласник Черкаський державний технологічний університет. – №u201613641; заявл.30.12.2016; опубл.12.06.2017, Бюл.№11.
9. Фауре Э. В. Факториальное кодирование с исправлением ошибок / Э. В. Фауре // Радіоелектроніка, інформатика, управління. – 2017. – № 3. – С. 130–138. DOI: 10.15588/1607-3274-2017-3-15.

10. Graham R. L. Concrete mathematics: a foundation for computer science / Ronald L. Graham, Donald E. Knuth, Oren Patashnik. – 2nd ed. – Reading: Addison-Wesley, 1994. – 657 p. ISBN13: 978-0201558029, ISBN10: 0201558025.
11. MacMahon P. A. The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects / P. A. MacMahon // American Journal of Mathematics. – 1913. – № 35 (3). – P. 281–322. DOI: 10.2307/2370312.
12. Comtet L. Advanced Combinatorics / Louis Comtet. – Dordrecht: D. Reidel Publishing Company, 1974. – 343 p. DOI: 10.1007/978-94-010-2196-8.
13. Moritz R. H. A Coin-Tossing Problem and Some Related Combinatorics / Roger H. Moritz, Robert C. Williams // Mathematics Magazine. – 1988. – №1. – Vol. 61. – P. 24–29. DOI: 10.2307/2690326.
14. Mendes A. A Note on Alternating Permutations / Anthony Mendes // The American Mathematical Monthly. – 2007. – Vol. 114, № 5. – P. 437–440.
15. Stanley R. P. Enumerative Combinatorics / Richard P. Stanley. – V.1. – 2nd ed. – New York : Cambridge University Press, 2011. – 725 p. ISBN13: 978-1107602625, ISBN10: 1107602629.
16. The on-line encyclopedia of integer sequences. A008302. – Режим доступа : http://oeis.org/A008302.
17. Кнут Дональд Э. Искусство программирования. В 7 т. Т.3. Сортировка и поиск, 2-е изд. / Дональд Эрвин Кнут ; пер. с англ. под ред. В. Т. Тертышного и И. В. Красикова. – М. : ООО «И. Д. Вильямс», 2007. – 832 с. ISBN: 978-5-8459-0082-1.
18. Финк Л. М. Теория передачи дискретных сообщений / Л. М. Финк Изд. 2-е. – М. : Советское радио, 1970. – 728 с.

Article was submitted 31.12.2017.
After revision 22.01.2018.

Фауре Э. В.[1], Щерба А. И.[2], Харин А. А.[3]

ФАКТОРІАЛЬНІ КОДИ З ЗАДАНИМ ЧИСЛОМ ІНВЕРСІЙ

[1]Канд. техн. наук, доцент, доцент кафедри інформаційної безпеки та комп'ютерної інженерії Черкаського державного технологічного університету, Черкаси, Україна

[2]Канд. фіз.-мат. наук, доцент, зав. кафедри прикладної математики Черкаського державного технологічного університету, Черкаси, Україна

[3]Аспірант кафедри інформаційної безпеки та комп'ютерної інженерії Черкаського державного технологічного університету, Черкаси, Україна

Актуальність. Факторіальне кодування з відновленням даних за перестановкою забезпечує комплексний захист інформації від несанкціонованого читання і помилок в каналі зв'язку і має властивість самосинхронізації. У той же час, таке кодування не дозволяє виявити всі помилки малої кратності, що призводить до порівняно невеликого показника підвищення достовірності.

Метою цієї роботи є розробка та дослідження методу факторіального кодування з заданим числом інверсій, спрямованого на підвищення достовірності передавання інформації за рахунок введення додаткової надлишковості шляхом вибору класу перестановок, які відповідають заданому критерію.

Метод. Основна ідея пропонованого методу полягає в штучному внесенні надлишковості за рахунок зменшення потужності використовуваних перестановок. Такий підхід дозволяє з усієї множини перестановок виділити клас, який володіє необхідними, наперед заданими, властивостями. У якості ознаки приналежності до використовуваного класу перестановок у роботі запропоновано використовувати відповідність числа їх інверсій заданому класу лишків. Виконано теоретичну оцінку параметрів коду.

Результати. Побудовано сигнально-кодові конструкції для факторіального коду з заданим числом інверсій для порядку перестановок $M = 8$. Для кожного з можливих класів досліджено потужність, швидкість коду, оцінку ймовірності невиявленої помилки і відносної швидкості передавання для систем з вирішальним зворотним зв'язком і незалежними бітовими помилками. Показано, що параметри коду не є інваріантними по відношенню до вибраного класу лишків для заданого модуля. Розроблено структурні схеми пристроїв кодування та декодування.

Висновки. Отримав подальший розвиток метод факторіального кодування з відновленням даних за перестановкою, який за рахунок використання перестановок, число інверсій у яких належить обраному класу лишків, дозволив підвищити достовірність передавання в обмін на втрату швидкості коду.

Ключові слова: факторіальне кодування, перестановка, інверсія, надлишковість, клас лишків, достовірність передавання, відносна швидкість передавання.

Фауре Э. В.[1], Щерба А. И.[2], Харин А. А.[3]

[1]Канд. техн. наук, доцент, доцент кафедры информационной безопасности и компьютерной инженерии Черкасского государственного технологического университета, Черкассы, Украина

[2]Канд. физ.-мат. наук, доцент, зав. кафедры прикладной математики Черкасского государственного технологического университета, Черкассы, Украина

[3]Аспирант кафедры информационной безопасности и компьютерной инженерии Черкасского государственного технологического университета, Черкассы, Украина

ФАКТОРИАЛЬНЫЕ КОДЫ С ЗАДАННЫМ ЧИСЛОМ ИНВЕРСИЙ

Актуальность. Факториальное кодирование с восстановлением данных по перестановке обеспечивает комплексную защиту информации от несанкционированного чтения и ошибок в канале связи и обладает свойством самосинхронизации. В то же время, такое кодирование не позволяет обнаружить все ошибки малой кратности, что приводит к сравнительно небольшому показателю повышения достоверности.

Целью данной работы является разработка и исследование метода факториального кодирования с заданным числом инверсий, направленного на повышение достоверности передачи информации за счет введения дополнительной избыточности путем выбора класса перестановок, удовлетворяющих заданному критерию.

Метод. Основная идея предлагаемого метода состоит в искусственном внесении избыточности за счет уменьшения мощности используемых перестановок. Такой подход позволяет из всего множества перестановок выделить класс, обладающий необходимыми, наперед заданными, свойствами. В качестве признака принадлежности к используемому классу перестановок в работе предложено использовать соответствие числа их инверсий заданному классу вычетов. Выполнена теоретическая оценка параметров кода.

Результаты. Построены сигнально-кодовые конструкции для факториального кода с заданным числом инверсий при порядке перестановок $M = 8$. Для каждого из возможных классов исследованы мощность, скорость кода, оценки вероятности необнаруженной

ошибки и относительной скорости передачи для систем с решающей обратной связью и независимыми битовыми ошибками. Показано, что параметры кода не являются инвариантными по отношению к выбранному классу вычетов при заданном модуле. Разработаны структурные схемы устройств кодирования и декодирования.

Выводы. Получил дальнейшее развитие метод факториального кодирования с восстановлением данных по перестановке, который за счет использования перестановок, число инверсий в которых принадлежит выбранному классу вычетов, позволил повысить достоверность передачи в обмен на потерю скорости кода.

Ключевые слова: факториальное кодирование, перестановка, инверсия, избыточность, класс вычетов, достоверность передачи, относительная скорость передачи.

REFERENCES

1. Faure E'. V., Shvydkii V. V., and Shcherba A. I. Kontrol' celostnosti informacii na osnove faktorial'noj sistemy schisleniya, *Journal of Baku Engineering University, Mathematics and Computer Science,* 2017, Vol. 1, No. 1, pp. 3–13.
2. Faure E'. V., Shvydkii V. V., and Shcherba V. A. Kombinirovannoe faktorial'noe kodirovanie i ego svoistva, Radioelektronika, informatika, upravlinnya, 2016, No. 3, pp. 80–86. DOI: 10.15588/1607-3274-2016-3-10.
3. Rudnyc'kyj V. M., Faure E'. V., Shvydkyj V. V., and Shherba A. I. Pat. 107655 Ukrai'na, MPK G06F 21/64 (2013.01), H04L 1/16 (2006.01). Sposib kontrolju cilisnosti informacii'; zajavnyk ta patentovlasnyk Cherkas'kyj derzhavnyj tehnologichnyj universytet. – № a201505937; zajavl. 16.06.2015; opubl. 24.06.2016, Bjul. № 12.
4. Rudnyc'kyj V. M., Faure E'. V., Shvydkyj V. V., and Shherba A. I. Pat. 107657 Ukrai'na, MPK H03M 13/09 (2006.01), H04K 1/06 (2006.01), G09C 1/06 (2006.01). Sposib kombinovanogo koduvannja informacii'; zajavnyk ta patentovlasnyk Cherkas'kyj derzhavnyj tehnologichnyj universytet. № a201508148 ; zajavl. 17.08.2015; opubl. 24.06.2016, Bjul. № 12.
5. Faure E'. V. Faktorial'noe kodirovanie s vosstanovleniem dannykh, Visnik Cherkas'kogo derzhavnogo tekhnologichnogo universitetu, 2016, No. 2, pp. 33–39. DOI: 10.24025/bulletinchstu.v1i2.82932.
6. Faure E'. V. Metod povysheniya effektivnosti faktorial'nogo kodirovaniya s vosstanovleniem dannykh, Visnik Cherkas'kogo derzhavnogo tekhnologichnogo universitetu, 2016, No. 4, pp. 57–61.
7. Faure E'. V. Faktorial'noe kodirovanie s neskol'kimi kontrol'nymi summami, Visnik Zhitomirs'kogo derzhavnogo tekhnologichnogo universitetu, 2016, No. 3, pp. 104–113. DOI: 10.26642/tn-2016-3(78)-104-113.
8. Faure E'. V., Harin O. O., Shvydkyj V. V., and Shherba A. I. Pat. 117004 Ukrai'na, MPK H03M 13/09 (2006.01), H04L 1/16 (2006.01), G04C 1/06 (2006.01). Sposib faktorial'nogo koduvannja z vidnovlennjam danyh; zajavnyk ta patentovlasnyk Cherkas'kyj derzhavnyj tehnologichnyj universytet. № u201613641 ; zajavl. 30.12.2016; opubl. 12.06.2017, Bjul. № 11.
9. Faure E'. V. Faktorial'noe kodirovanie s ispravleniem oshibok, *Radio Electronics, Computer Science, Control,* 2017, No. 3, pp. 130–138. DOI: 10.15588/1607-3274-2017-3-15.
10. Graham R. L., Knuth D. E., and Patashnik O. Concrete mathematics: a foundation for computer science. 2nd ed. Reading, Massachusetts, Addison-Wesley, 1994, 657 p. ISBN13: 978-0201558029, ISBN10: 0201558025.
11. MacMahon P. A. The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects, American Journal of Mathematics, 1913, No. 35, Vol. 3, pp. 281–322. DOI: 10.2307/2370312.
12. Comtet L. Advanced Combinatorics, Dordrecht, D. Reidel Publishing Company, 1974, 343 p. DOI: 10.1007/978-94-010-2196-8.
13. Moritz R. H. and Williams R. C. A Coin-Tossing Problem and Some Related Combinatorics, Mathematics Magazine, 1988, No. 1, Vol. 61, pp. 24-29. DOI: 10.2307/2690326.
14. Mendes A. A Note on Alternating Permutations, The American Mathematical Monthly, 2007, No. 5, Vol. 114, pp. 437-440.
15. Stanley R. P. Enumerative Combinatorics, V.1., 2nd ed., New York, Cambridge University Press, 2011, 725 p. ISBN13: 978-1107602625, ISBN10: 1107602629.
16. The on-line encyclopedia of integer sequences. A008302, Rezhim dostupa: http://oeis.org/A008302.
17. Knuth Donald E. [Per. s angl. pod red. Tertyshnogo V. T. i Krasikova I.V.]. Iskusstvo programmirovaniya. V 7 t. Vol. 3. Sortirovka i poisk, 2-e izd., M, OOO «I.D. Vil'yams», 2007, 832 p. ISBN: 978-5-8459-0082-1.
18. Fink L. M. Teoriya peredachi diskretnykh soobshchenii. [Izd. 2-e.]. M., Sovetskoe radio, 1970, 728 p.

УДК 004.422.83

Федасюк Д. В.[1], Чопей Р.С.[2]

[1]Д-р техн. наук, професор, проректор Національного університету «Львівська політехніка», Львів, Україна
[2]Аспірант кафедри програмного забезпечення Національного університету «Львівська політехніка», Львів, Україна

АЛГОРИТМ ПОБУДОВИ ГРАФУ ПОТОКУ КЕРУВАННЯ ЗА ТЕКСТОМ ПРОГРАМИ МОВОЮ С

Актуальність. Робота присвячена проблемі автоматизованої побудови графу потоку керування за текстом програми мовою С, що є важливим етапом структурного тестування вбудованих систем.

Мета роботи – створення алгоритму побудови графу потоку керування за текстом програми, що має високу швидкість роботи та дозволяє подальше застосування існуючими засобами автоматизованого тестування.

Метод. Запропоновано алгоритм побудови графу потоку керування за текстом програми мовою С, який здійснює попередню обробку вхідного тексту програми шляхом видалення коментарів та порожніх рядків, визначає кількість вершин та ребер графу потоку керування на основі синтаксичного аналізу тексту програм, а також формує, заповнює та зберігає матрицю інцидентності в окремому текстовому файлі, що дозволяє подальше використання засобами тестування, які у якості вхідних даних приймають граф потоку керування, окрім того, текстовий файл може використовуватись засобами для графічного представлення графу потоку керування.

Результати. Розроблено програмний модуль, що реалізує запропонований алгоритм, який використано при проведенні експериментів з дослідження залежності часу побудови графу потоку керування від кількості рядків тексту програми.

Висновки. Проведені експерименти підтвердили працездатність запропонованого алгоритму та розробленого на його основі програмного модуля, довели придатність результатів роботи алгоритму, що дозволяє рекомендувати його для подальшого застосування засобами автоматизованого тестування для зменшення часових витрат на тестування вбудованого програмного забезпечення.

Ключові слова: граф потоку керування, автоматизоване тестування, Keil uVision, ARM, вбудовані системи**.**

НОМЕНКЛАТУРА

ПЗ – програмне забезпечення;

ARM – покращена RISC машина;

N – кількість відповідних ідентифікаторів case та default;

edgeC – лічильник вершин;

CSV – значення, розділені комою;

$I(|V|\text{r}|E|)$ – матриця інцидентності;

V – кількість рядків у матриці інцидентності, і, відповідно, вершин у графі керування;

E – кількість стовпців у матриці інцидентності, і, відповідно, ребер у графі керування;

M – мова програмування С;

$f_i(\cdot)$ – функція програми;

n – кількість функцій;

G – граф потоку керування;

O_{En} – обчислювальна складність n-го етапу роботи алгоритму.

$O_{\text{ВК}}$ – обчислювальна складність процедури видалення коментарів;

$O_{\text{МП}}$ – обчислювальна складність процедури модифікації тексту програми;

$O_{\text{ФЛ}}$ – обчислювальна складність процедури формування списків ознак;

$O_{\text{ЗаМ}}$ – обчислювальна складність процедури заповнення матриці;

$O_{\text{ЗбМ}}$ – обчислювальна складність процедури збереження матриці.

ВСТУП

Графи потоку керування (Control flow graph, CFG) є загальноприйнятим засобом візуального представлення множини всіх можливих шляхів виконання програми. Кожен оператор представлений вузлом в графі, ребра відображають потік управлення між ними. Графи потоку керування застосовуються при аналізі потоку керування та оптимізації коду компілятором [1–3], для визначення та оцінки метрик програмного коду, зокрема складності програми [4–7], а також для структурного тестування програмного забезпечення (ПЗ), що використовує граф потоку керування для генерування тестових випадків [8–11], та визначення тестового покриття [12].

Об’єктом дослідження є процес побудови графу потоку керування.

Наявність графу потоку керування, сформованого у вигляді, зручному для опрацювання комп’ютером (зокрема, матрицею інцидентності), дозволить використати існуючі засоби [8–12] для обчислення метрик розроблюваного вбудованого ПЗ, а програмна реалізація алгоритму побудови графу потоку керування може бути оформлена у окремий модуль засобів автоматизованого тестування вбудованого ПЗ, що дозволить істотно зменшити витрати часу та зусиль команди тестувальників, а також уникнути помилок через неуважність, що зазвичай супроводжує монотонну рутинну роботу.

Огляд інтегрованих середовищ розроблення коду для мікроконтролерів ARM показав, що всі вони будують граф потоку керування автоматично, однак, жоден з них не дає можливості видобути граф потоку керування та зберегти його у вигляді, придатному для застосування в якості вхідних даних засобів автоматизованого тестування ПЗ.

Предмет дослідження становлять алгоритми побудови графу потоку керування за текстом програми.

Метою роботи є розроблення алгоритму автоматизованої побудови графу потоку керування за текстом програми, написаним мовою С, а також дослідження ефективності його роботи та можливості застосування його існуючими засобами тестування ПЗ.

© Федасюк Д. В., Чопей Р. С., 2018

DOI 10.15588/1607-3274-2018-2-17

1 ПОСТАНОВКА ЗАДАЧІ

Нехай задано текст програми мовою програмування С, представлений набором функцій у вигляді $T =< C, \widetilde{F} >$, де $\widetilde{F} = \{f_i(\cdot), i = 1,2,...n\}$.

Тоді задача побудови графу потоку керування, представленого у вигляді метриці інцидентності $G = I(|V| \times |E|)$, полягатиме в синтаксичному аналізі тексту програми для визначення кількості вершин і ребер графу та взаємозв'язків між ними.

Для оцінювання якості одержаних результатів може бути використано порівняння графу потоку керування, побудованого за одержаною матрицею інцидентності, та графом потоку керування, вручну створеного фахівцем з забезпечення якості ПЗ згідно з існуючими правилами.

2 ОГЛЯД ЛІТЕРАТУРИ

У [13] запропоновано алгоритм перетворення тексту програми у граф-схему, що може застосовуватися для структурного проектування складних інформаційно-вимірювальних системи. Наведено приклад застосування алгоритму перетворення, реалізованого у вигляді окремого програмного модуля для тексту програми мовою Object Pascal.

У [14] розглянуто проблему побудови графу потоку керування класичними алгоритмами для коду планувальника операційної системи. Представлено алгоритм перетворення асемблерного коду планувальника у граф потоку керування, що не виключає гілки з затримками. Проведені експерименти для систем на базі архітектури TMS320C6200 від компанії Texas Instruments показали, що побудовані за представленим алгоритмом графи є більш достовірними у порівнянні з графами, побудованими за класичними алгоритмами. А відтак, оптимізація обсягу вихідного коду та часу виконання програми стає більш ефективною.

У [15] представлено архітектуру засобу для тестування програмного забезпечення вбудованих систем. Запропонований засіб використовує машинний код, отриманий у процесі компілювання ПЗ, написаного мовою Java, та формальну специфікацію. Окрім архітектури засобу, представлено алгоритм побудови графу потоку керування з машинного коду та алгоритм заміни виклику функції, представленого вершиною графу, його графом потоку керування.

Головним недоліком наявних алгоритмів побудови графу потоку керування [13–15] є відсутність можливості їх застосування для ПЗ, написаного мовою С, яка на сьогодні утримує лідерство з-поміж мов для розроблення вбудованого ПЗ.

3 МАТЕРІАЛИ ТА МЕТОДИ

Для формалізації алгоритму використаємо наступні поняття:

Обчислювальна операція – найменша автономна часина мови програмування. Для будь-якої мови програмування виділяють три різновиди обчислювальних операцій – слідування, галуження, та цикли:

– слідування – виклики функцій, та прості операції (інкремент/декремент та інші);

– галуження – оператори галуження для мови С (if, else if, else, switch);

– цикли – оператори циклів для мови С ("for", "while", "do").

Граф потоку керування – орієнтований граф, що представлений у формі матиці, та містить один вхідний та один вихідний вузол. Вершинами графу представлені символьні оператори чи функції, а ребрами – зв'язок між ними.

Матриця інцидентності (Incidence matrix) – одна з форм представлення графа, в якій вказується зв'язок між інцидентними елементами графа (ребро і вершина). Стовпці матриці відповідають ребрам, рядки – вершинам.

Заповнення матриці відбувається згідно з наступними правилами:

– комірка матриці набуває значення «1», якщо ребро виходить з вершини;

– комірка матриці набуває значення «–1», якщо ребро входить у вершину;

– комірка матриці набуває значення «0», якщо ребро не пов'язане з вершиною.

Алгоритм побудови графу потоку керування базується на об'єднанні декількох складових: попередня обробка тексту програми, визначення кількості вершин графу, визначення кількості ребер графу, заповнення матриці на основі визначених взаємозв'язків між вершинами графу. Процесу побудови графу потоку керування зображено на рис. 1.

Побудова графу потоку керування передбачає чотири етапи роботи:

Етап 1. Попередня обробка тексту програми

На цьому етапі роботи алгоритму з тексту програми видаляються коментарі, порожні рядки, а також проводиться модифікація тексту програми, якщо виклик функції або умову галуження записано у декілька рядків.

Етап 2. Визначення кількості вершин графу

На цьому етапі роботи алгоритму виконується автоматична нумерація рядків тексту програми за винятком наступних випадків:

– рядок тексту програми, що містить лише спеціальний символ "{" або "}", не збільшує лічильник рядків;

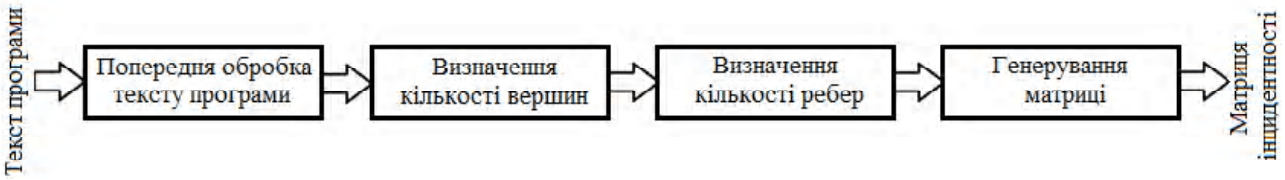


Рисунок 1 – Схема процесу побудови графу потоку керування за текстом програми, мовою С

– рядок тексту програми, що містить ідентифікатор: case, default, break, return, else, else if, не збільшує лічильник рядків;

– рядок тексту програми, що містить ініціалізацію змінної чи екземпляра структури, не збільшує лічильник рядків;

– лічильник рядків програмного коду збільшується на одиницю після останнього рядка тексту програми;

Етап 3. Визначення кількості ребер графу

На цьому етапі роботи алгоритму виконується нумерація ребер, шляхом збільшення лічильника ребер, на основі синтаксичного аналізу тексту програми, а саме пошуку у ньому ідентифікаторів галуження (if, switch, else, else if), ідентифікаторів циклу (for, do, while) та ідентифікаторів слідування (рис. 2).

Визначення кількості ребер відбувається згідно з правилами:

– лічильник ребер збільшується на одиницю, якщо рядок тексту програми відповідає обчислювальному процесу – слідування;

– лічильник ребер збільшується на одиницю, якщо рядок тексту програми містить ідентифікатори if, для яких є відповідний ідентифікатор else;

– лічильник ребер збільшується на одиницю, якщо рядок тексту програми містить ідентифікатор else або else if;

– лічильник ребер збільшується на два, якщо рядок тексту програми містить ідентифікатори if, для яких відсутні відповідні ідентифікатори else та else if;

– лічильник ребер збільшується на два, якщо рядок тексту програми містить ідентифікатор циклу – for або while;

– лічильник ребер збільшується на *N*, якщо рядок тексту програми містить ідентифікатор галуження – switch.

Етап 4. Формування матриці інцидентності

На цьому етапі роботи алгоритму створюється матриця з розмірами, що були визначені на попередніх етапах, та виконується її заповнення. Цей етап роботи є найбільш трудомістким, адже передбачає визначення взаємозв'язків у програмному коді. Для цього створюється чотири списки:

List1 – містить номер вершини, якій відповідає рядок тексту програми.

List2 – містить номер вершини, яка є наступною після ідентифікатора галуження або ідентифікатора циклу.

List3 – містить маску, що відповідає типу рядка, та може набувати таких значень:

– functionName – рядок, що містить ім'я функції;

– returnLine – рядок, що містить ідентифікатор return;

– followLine – рядок, що містить виклик іншої функції або присвоєння значення у змінну;

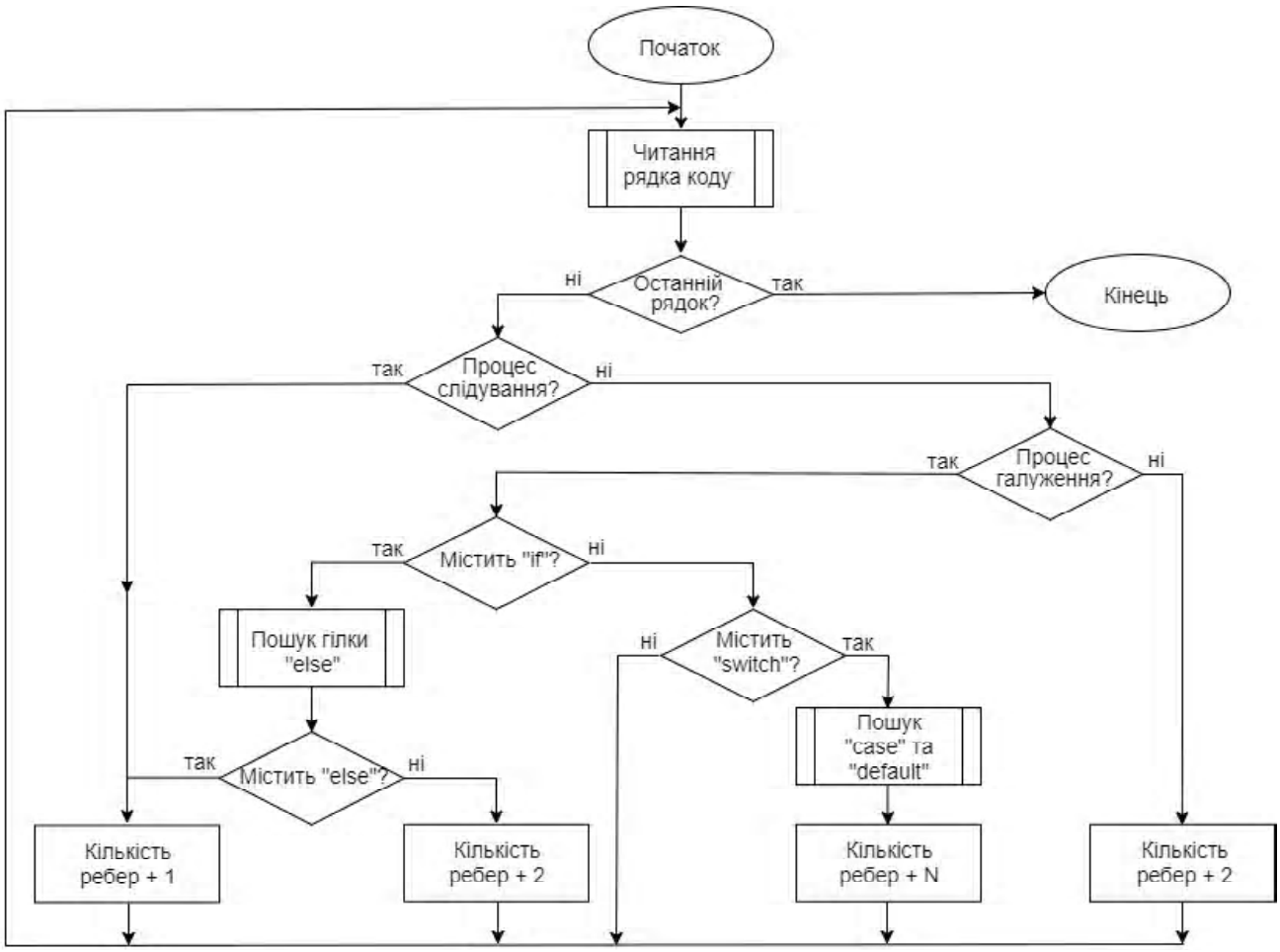


Рисунок 2 – Блок-схема етапу визначення кількості ребер у графі

– lastLineInsideIF – останній рядок вихідного коду під конструкцією if;

– lastLineInsideCycle – останній рядок вихідного коду під конструкцією for або while;

– lastLineInsideCase – останній рядок вихідного коду під конструкцією case;

List4 – містить номер вершини ідентифікатора if, switch, for або while, під якою знаходиться рядок тексту програми.

Наступним кроком після створення списків є заповнення цих списків та створення змінної – *edgeC*, що буде використано при оновленні інформації у матриці. Оновлення матриці відбувається згідно з правилами:

– якщо елемент *List1* містить 0 і елемент *List3* містить маску functionName, встановлюємо у комірці [0,0] значення «1», що відповідає виходу ребра з вершини 0.

– якщо елемент *List3* містить маску followLine, встановлюємо значення «-1» у комірці [*List1[n], edgeC*], збільшуємо лічильник вершин на одиницю та встановлюємо значення «1» у комірці [*List1[n], edgeC*];

– якщо елемент *List3* містить маску lastLineInsideIF встановлюємо значення «–1», у комірці [*List1[n], edgeC*], збільшуємо лічильник вершин на 1 та встановлюємо значення «1», у комірці [*List1[n], edgeC*]. У комірці [*List2[n], edgeC*] встановлюємо значення «–1» та встановлюємо значення «1», у комірці [*List4[n], edgeC*];

– якщо елемент *List3* містить маску lastLineInsideCycle встановлюємо значення «–1», у комірці [*List1[n], edgeC*], збільшуємо лічильник вершин на 1 та встановлюємо значення «1» у комірці [*List1[n], edgeC*], встановлюємо значення «–1» у комірці [*List4[n], edgeC*] і збільшуємо лічильник вершин на 1, встановлюємо значення «1» у комірці [*List4[n], edgeC*];

– якщо елемент *List3* містить маску lastLineInsideCase встановлюємо значення «–1» у комірці [*List1[n], edgeC*], збільшуємо лічильник вершин на 1 та встановлюємо значення «1» у комірці [*List1[n], edgeC*]. У комірці [*List2[n], edgeC*] встановлюємо значення «–1», а також встановлюємо значення «1» у комірці [*List4[n], edgeC*];

– якщо елемент *List3* містить маску returnLine встановлюємо значення «–1» у комірці [*List1[n], edgeC*];

Останнім кроком роботи цього алгоритму є збереження матриці у окремому текстовому файлі з форматом даних CSV.

4 ЕКСПЕРИМЕНТИ

Розглянемо застосування розробленого алгоритму для функції читання значення температури, що містить основні ідентифікатори мови С (дод. А).

У відповідності до представленого алгоритму виконуємо процедуру попереднього опрацювання тексту програми, а також визначаємо кількість вершин та ребер. Текст програми після процедури попереднього опрацювання та визначення кількості вершин наведений у дод. Б.

На основі аналізу тексту програми виявлено:

– 14 рядків, що відповідають обчислювальному процесу слідування – 14 ребер;

– 2 рядки, що містять ідентифікатор циклу for та while, для кожного з яких необхідно 2 ребра – 4 ребра для обох;

– 3 рядки, що містять ідентифікатор галуження if, else if та else, – 3 ребра;

– 1 рядок, що містить ідентифікатор галуження switch, для якого є два ідентифікатори case та один ідентифікатор default – 3 ребра;

– загальна кількість ребер – 24.

– загальна кількість вершин – 19.

Створюємо матрицю інцидентності $I(|19|\times|24|)$ та чотири списки для 19 вершин, які будуть заповнені згідно з наведеними вище правилами. Результати аналізу взаємозв'язків у тексті програми наведені у табл. 1.

Таблиця 1 – Списки ознак для функції ReadTemperature

List1	List2	List3	List4
0	–	functionName	–
1	6	followLine	–
2	6	followLine	1
3	6	followLine	1
4	6	followLine	1
5	6	lastLineInsideCycle	1
6	9	followLine	–
7	9	followLine	6
8	9	lastLineInsideCycle	6
9	–	followLine	–
10	14	followLine	–
11	14	lastLineInsideIF	10
12	14	lastLineInsideIF	10
13	14	lastLineInsideIF	10
14	18	followLine	–
15	18	lastLineInsideCase	14
16	18	lastLineInsideCase	14
17	18	lastLineInsideCase	14
18	18	returnLine	–

Матриця, заповнена в результаті роботи алгоритму, має вигляд:

$$I = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & -1 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & -1
\end{bmatrix} \quad (1)$$

Отримана матриця буде збереженою для подальшого використання. Також граф потоку керування, представлений у вигляді матриці інцидентності, можемо представити графічно з допомогою Інтернет-сервісу [16]. Побудований за матрицею (1) граф наведений на рис. 3.

5 РЕЗУЛЬТАТИ

Враховуючи модульну структуру розробленого алгоритму обчислювальна складність алгоритму є сумою складності для кожного етапу роботи. Перший етап передбачає видалення коментарів та модифікацію вихідного коду при необхідності, складність якого визначається згідно (2):

$$O_{E1} = O_{\text{ВК}}(1) + O_{\text{МП}}(N). \quad (2)$$

Другий етап роботи алгоритму передбачає визначення кількості вершин у графі. Обчислювальна складність цього етапу роботи залежить лише від кількості рядків тексту програми, що рівна $O_{E2}(N)$.

Третій етап роботи алгоритму передбачає визначення кількості ребер у графі. Складність цього етапу роботи збільшується за умови наявності у програмному коді операторів галуження “if” та “switch”. Тому вважаємо, що складність цього етапу $O_{E3}(N\hat{}2)$.

Етап генерування матриці є найбільш трудомістким. Складність цього етапу рівна (3):

$$O_{E4} = O_{\text{ФЛ}}(N\hat{}2) + O_{\text{ЗаМ}}(N) + O_{\text{ЗбМ}}(1). \quad (3)$$

Загальна складність алгоритму побудови графу потоку керування визначається:

$$O_{E1} + O_{E2} + O_{E3} + O_{E4} = O_{\text{ВК}}(1) + O_{МП}(N) + O_{E2}(N) + \\ + O_{E3}(N\hat{}2) + O_{\text{ФЛ}}(N\hat{}2) + O_{\text{ЗаМ}}(N) + O_{\text{ЗбМ}}(1) = O(N\hat{}2). \quad (4)$$

Згідно з проведеним розрахунком складності алгоритму визначаємо час побудови графу потоку керування, для тексту програми обсягом 50, 100 та 120 тис. рядків. Розрахунок буде виконаний за умови, що комп’ютер виконує 1 млрд. операцій в секунду. Результати розрахунку наведені у табл. 2.

Таблиця 2 – Час побудови графу потоку керування для різного обсягу вхідних даних

Загальна к-ть рядків тексту програми, тис.	К-ть рядків коду після попередньої обробки, тис.	Час побудови, с
50	49	2,5
100	98	10
120	118	14,4

Отримані теоретичні результати розрахунку часу роботи алгоритму, свідчать про придатність результатів та доцільність реалізації алгоритму для проведення експериментальних досліджень.

Реалізація алгоритму здійснювалась у вигляді окремого програмного модуля, з використанням мови програмування C#. Окрім того, для проведення експериментів з великим обсягом вхідних даних реалізовано програмний модуль, який почергово передає текст функцій, з проектів розроблених у середовищі програмування Keil uVision. Експериментальні дослідження проведені на персональному комп’ютері Lenovo W520, з процесором Intel Core i5, що працює з частотою 2,2 ГГц та обсягом оперативної пам’яті 8 ГБайт. У якості вхідних даних використано три комерційні продукти, що містять 56, 90, та 150 тисяч рядків тексту програми. Результати експерименту наведені у табл. 3.

На основі отриманих теоретичних та експериментальних даних побудовано графік залежності часу побудови від кількості рядків тексту програми (рис. 4).

Як видно з рис. 4, теоретичні результати розрахунку та експериментальні дані дещо відрізняються, що зумовлено витратами часу на послідовне виконання алгоритму для кожної функції комерційного проекту.

Відзначимо, що для зменшення часу побудови графу потоку керування можливе паралельне виконання другого та третього етапів.

6 ОБГОВОРЕННЯ

Запропонований алгоритм побудови графу потоку керування порівняно з алгоритмом побудови на основі асемблерного коду [14] забезпечує більшу наочність перетворення, адже граф потоку керування будується напряму з тексту програми мовою С, в той час як граф потоку керування, побудований за асемблерним кодом, може дещо відрізнятись, що зумовлено процесом перетворення тексту програми у асемблерний код. Також запропонований алгоритм потребує більших часових витрат на побудову графу, ніж алгоритм побудови графу потоку керування за асемблерним кодом [14].

Таблиця 3 – Час побудови графів потоку керування для комерційних продуктів

Загальна к-ть рядків тексту програми, тис.	К-ть рядків коду після попередньої обробки	Час побудови, с
56	54 390	23
90	87 690	30
150	147 300	50,2

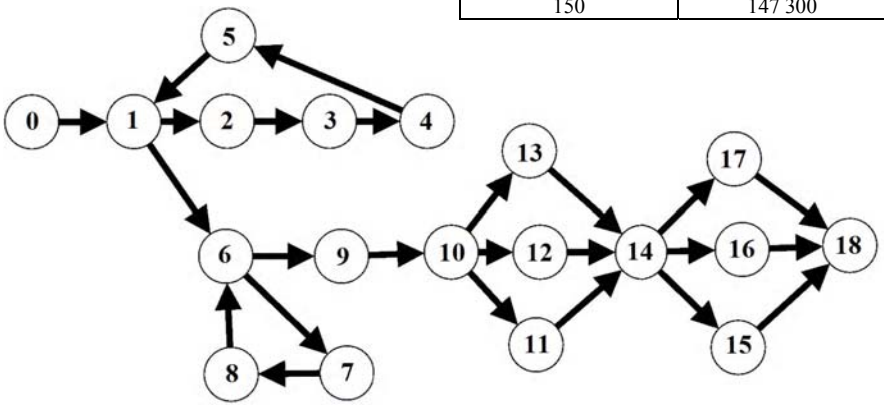


Рисунок 3 – Граф потоку керування функції “ReadTemperature”

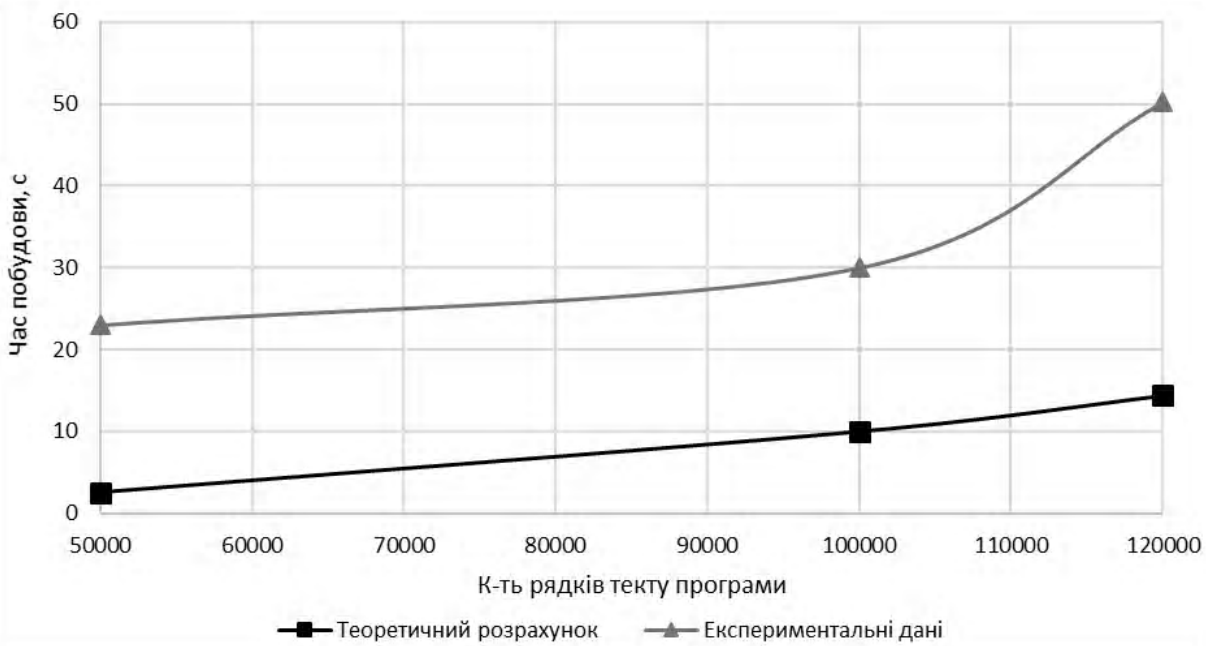


Рисунок 4 – Залежність часу побудови від кількості рядків тексту програми

Недоліком запропонованого алгоритму є його залежність від обчислювальних ресурсів, що збільшуються з кількістю рядків тексту програми.

Зменшення часу побудови графу потоку керування за рахунок паралельного виконання другого та третього етапу виконання, зростатиме з збільшенням кількості рядків тексту програми.

ВИСНОВКИ

У роботі вирішено актуальну задачу автоматизованої побудови графу потоку керування за текстом програми мовою С.

Наукова новизна роботи полягає в тому, що запропоновано алгоритм, який автоматично будує граф потоку керування з тексту програми мовою С, характеризується невисокою обчислювальною складністю, що, у свою чергу дозволяє знизити вимоги до обчислювальних ресурсів і пам'яті персонального комп'ютера, на якому він виконується.

Практична цінність отриманих результатів полягає в тому, що розроблено програмний модуль, який реалізує запропонований алгоритм та дозволяє виконувати побудову графу потоку керування, представленого у вигляді матриці інцидентності.

Перспективи подальших досліджень полягають у застосуванні запропонованого алгоритму до формування вхідних даних для алгоритму генерування тестових випадків, що інтегрований у засобі автоматизованого тестування найбільш тривалого часу виконання програми.

ПОДЯКИ

Роботу виконано в рамках госпдоговірної науково-дослідної роботи кафедри програмного забезпечення Національного університету «Львівська політехніка» на тему «Автоматизоване генерування тестових сценаріїв для забезпечення якості апаратного та мікропрограмного забезпечення, для Італійської компанії Dinamica Generale S.p.A., що є лідером в галузі розробки вбудованих систем для агропромисловості.

СПИСОК ЛІТЕРАТУРИ

1. Compilers: Principles, techniques, and tools / [eds.: A. V. Aho, M. S. Lam, R. Sethi, J. D. Ullman]. – Amsterdam: Addison Wesley, 2007. – 1038 p. ISBN: 0321486811
2. Graph-theoretic constructs for program control flow analysis : Research Report : RC-3923 / F. Allen, J. Cocke. – IBM Thomas J. Watson Research Center – IBM, 1972. – 130 p.
3. Ferrante J. The program dependence graph and its use in optimization / J. Ferrante, K. J. Ottenstein, J. D. Warren // ACM Transactions on Programming Languages and Systems. – 1987. – № 9. – P. 319–349. DOI: 10.1145/24039.24041
4. Gold R. On cyclomatic complexity and decision graphs / Robert Gold // Proceedings of the 10th International Conference of Numerical Analysis and Applied Mathematics (ICNAAM '12), 2007. – Vol. 1479. – P. 2170–2173. DOI: 10.1063/1.4756622
5. Henderson-Sellers B. The theoretical extension of two versions of cyclomatic complexity to multiple entry/exit modules / B. Henderson-Sellers, D. Tegarden // Software Quality Journal. – 1994. – Vol. 3. – P. 253–269. DOI: 10.1007/BF00403560
6. Jalote P. An integrated approach to software engineering / P. Jalote. – New York : Springer, 2005. – 566 p. ISBN: 978-0-387-28132-2
7. Sommerville I. Software Engineering / I. Sommerville. – Pearson : Addison Wesley, 2004. – 750 p. ISBN: 0321210263
8. Frankl P. G. Provable improvements on branch testing / P. G. Frankl, E. J. Weyuker. // IEEE Transactions on Software Engineering. – 1993. – Vol. 19. – P. 962–975. DOI: 10.1109/32.245738
9. Gold R. Control flow graphs and code coverage / R. Gold. // International Journal of Applied Mathematics & Computer Science. – 2010. – Vol. 20. – P. 739–749. DOI: 10.2478/v10006-010-0056-9
10. Zhu H. Software unit test coverage and adequacy / H. Zhu, P. A. Hall, J. H. May. // ACM Computing Surveys. – 1997. – Vol. 29. – P. 366–427. DOI: 10.1145/267580.267590
11. Fedasyuk D. Architecture of a tool for automated testing the worst-case execution time of real-time embedded systems' firmware / D. Fedasyuk, R. Chopey, B. Knysh. // 14th International conference, The experience of designing and application of cad systems in microelectronics. – 2017. – P. 278–282. DOI: 10.1109/CADSM.2017.7916134
12. Hossain M. I. Integration testing approach using usage patterns of global variables [Electronic resource] / M. I. Hossain, W. J. Lee, S. Youngsul. – Researchgate, 2012. Access mode: https://www.researchgate.net/publication/255173431_Integration_Testing_Approach_using_Usage_Patterns_of_Global_Variables.
13. Муха Ю. П. Алгоритм для преобразования программного кода на языке высокого уровня в граф-схему / Ю. П. Муха, В. А. Секачев // Известия волгоградского государственного технического университета. – 2009. – № 3. – С. 58–63.
14. Cooper K. D. Building a control-flow graph from scheduled assembly code [Electronic resource] / K. D. Cooper, T. J. Harvey, T. Waterman. – Researchgate, 2013. Access mode: https://www.researchgate.net/publication/2572750.

15. Amine A. Generating control flow graph from Java card byte code / A. Amine, B. Mohammed, L. Jean-Louis // Information Science and Technology (CIST). – 2014. – P. 206–212. DOI: 10.1109/CIST.2014.7016620

16. Graph Online [Electronic resource]: Creating graph from incidence matrix. Access mode: http://graphonline.ru/en/

Стаття надійшла до редакції 14.08.2017.
Після доробки 25.09.2017.

Федасюк Д. В.[1], Чопей Р. С.[2]

[1]Д-р. техн. наук, профессор, проректор Национального университета «Львовская политехника», Львов, Украина

[2]Аспирант кафедры программного обеспечения Национального университета «Львовская политехника», Львов, Украина

АЛГОРИТМ ПОСТРОЕНИЯ ГРАФА ПОТОКА УПРАВЛЕНИЯ ПО ТЕКСТУ ПРОГРАММЫ НА ЯЗЫКЕ С

Актуальность. Работа посвящена проблеме автоматизированного построения графа потока управления по тексту программы на языке С, что является важным этапом структурного тестирования встроенных систем.

Цель работы – создание алгоритма построения графа потока управления по тексту программы, что имеет высокую скорость работы и позволяет дальнейшее применение существующими средствами автоматизированного тестирования.

Метод. Предложен алгоритм построения графа потока управления по тексту программы на языке С, который осуществляет предварительную обработку исходного текста программы путем удаления комментариев и пустых строк, определяет количество вершин и ребер графа потока управления на основе синтаксического анализа текста программ, а также формирует, заполняет и сохраняет матрицу инцидентности в отдельном текстовом файле, что позволяет дальнейшее использование средствами тестирования, которые в качестве входных данных принимают граф потока управления, кроме того, текстовый файл может использоваться средствами для графического представления графа потока управления.

Результаты. Разработан программный модуль, реализующий предложенный алгоритм, который используется при проведении экспериментов по исследованию зависимости времени построения графа потока управления от количества строк текста программы.

Выводы. Проведенные эксперименты подтвердили работоспособность предложенного алгоритма и разработанного на его основе программного модуля и доказали пригодность результатов работы алгоритма, что позволяет рекомендовать его для дальнейшего применения средствами автоматизированного тестирования для уменьшения временных затрат на тестирование встроенного программного обеспечения.

Ключевые слова: граф потока управления, автоматизированное тестирование, Keil uVision, ARM, встраиваемые системы.

Fedasyuk D. V.[1], Chopey R. S.[2]

[1]Dr.Sc., Professor, Vice-Rector of Lviv Polytechnic National University, Lviv, Ukraine

[2]Post-graduate student of the Department of Software, Lviv Polytechnic National University, Lviv, Ukraine

THE ALGORITHM OF CONSTRUCTING CONTROL FLOW GRAPH BASED ON A PROGRAM WRITTEN IN C

Actuality. The work considers the problem of automated construction of the control flow graph using the text of a program written in C. The graph construction is an important step in the structural testing of embedded systems.

Objective. The work is aimed at creation of a fast algorithm for constructing the control flow graph from a source code. Such an automatically generated graph can be used by existing tools for automated testing.

Method. We propose an algorithm for constructing the control flow graph from the program written in C language, which preprocesses the source code of the program by removing comments and blank lines, determines the number of vertices and edges of the control flow graph by means of program text syntactic analysis, and then formats, fills and stores the incidence matrix in a separate text file. This file can be passed as input data for some existing automated testing tools. Moreover, the content of the text file can be visualized by existing tools for graphic representation of graphs.

Results. We've developed a software module that implements the proposed algorithm. The module can be used for experiments aimed at studying the dependence of the control flow graph's time construction on the amount of lines in a source code.

Conclusions. The conducted experiments confirmed the operability of the proposed algorithm and the software module developed on its basis. The results have proved the applicability of the algorithm, thus it can be recommended for further use together with automated testing tools to reduce the time required for testing of embedded software.

Keywords: control flow graph, automated testing, Keil uVision, ARM, embedded systems.

REFERENCES

1. Aho A. V., Lam M. S., Sethi R., Ullman J. D. : eds. Compilers: Principles, techniques, and tools. Amsterdam, Addison Wesley, 2007, 1038 p. ISBN: 0321486811
2. Allen F., Cocke J. Graph-theoretic constructs for program control flow analysis : Research Report : RC-3923, *IBM Thomas J. Watson Research Center – IBM,* 1972, 130 p.
3. Ferrante J., Ottenstein K. J., Warren J. D. The program dependence graph and its use in optimization, *ACM Transactions on Programming Languages and Systems*, 1987, No. 9, pp. 319–349. DOI: 10.1145/24039.24041
4. Gold R. On cyclomatic complexity and decision graphs, *Proceedings of the 10th International Conference of Numerical Analysis and Applied Mathematics (ICNAAM '12),* 2007, Vol. 1479, pp. 2170–2173. DOI: 10.1063/1.4756622
5. Henderson-Sellers B., Tegarden D. The theoretical extension of two versions of cyclomatic complexity to multiple entry/exit modules, *Software Quality Journal*, 1994, Vol. 3, pp. 253–269. DOI: 10.1007/BF00403560
6. Jalote P. An integrated approach to software engineering. New York, Springer, 2005, 566 p. ISBN: 978-0-387-28132-2
7. Sommerville I. Software Engineering. Pearson, Addison Wesley, 2004, 750 p. ISBN: 0321210263
8. Frankl P. G., Weyuker E. J. Provable improvements on branch testing, *IEEE Transactions on Software Engineering*, 1993, Vol. 19, pp. 962–975. DOI: 10.1109/32.245738
9. Gold R. Control flow graphs and code coverage, *International Journal of Applied Mathematics & Computer Science*, 2010, Vol. 20, pp. 739–749. DOI: 10.2478/v10006-010-0056-9
10. Zhu H., Hall P. A., May J. H. Software unit test coverage and adequacy, *ACM Computing Surveys*, 1997, Vol. 29, pp. 366–427. DOI: 10.1145/267580.267590
11. Fedasyuk D., Chopey R., Knysh B. Architecture of a tool for automated testing the worst-case execution time of real-time embedded systems' firmware, *14th International conference, The experience of designing and application of cad systems in microelectronics*, 2017, pp. 278–282. DOI: 10.1109/CADSM.2017.7916134
12. Hossain M. I., Lee W. J., Youngsul S. Integration testing approach using usage patterns of global variables [Electronic resource], *Researchgate, 2012. Access* mode: https://www.researchgate.net/publication/255173431_Integration_Testing_Approach_using_Usage_Patterns_of_Global_Variables.
13. Muha Ju. P., Sekachjov V. A. Algoritm dlja preobrazovanija programmnogo koda na jazyke vysokogo urovnja v graf-shemu, *Izvestija volgogradskogo gosudarstvennogo tehnicheskogo universiteta*, 2009, No. 3, pp. 58–63.
14. Cooper K. D., Harvey T. J., Waterman T. Building a control-flow graph from scheduled assembly code [Electronic resource], *Researchgate, 2013*. Access mode: https://www.researchgate.net/publication/2572750.
15. Amine A., Mohammed B., Jean-Louis L. Generating control flow graph from Java card byte code, *Information Science and Technology (CIST),* 2014, pp. 206–212. DOI: 10.1109/CIST.2014.7016620
16. Graph Online [Electronic resource]: Creating graph from incidence matrix. Access mode: http://graphonline.ru/en/

ДОДАТОК А

```
short ReadTemperature (void)
{
unsigned Counter, elems,TemperatureType;
float TemperatureBuffer[10];
float Tmp = 0.0;
float summ = 0.0;
float average = 0.0;
/* the TemperatureBuffer used
for average temperature value */
for (Counter = 0; Counter < 10; Counter++)
{
StartConversionOfInternalADC_1();
osDelay(10);
TemperatureBuffer[Counter]          =
ReadTemperatureADC_1();
elems++;
}
/* calculate average temperature */
while(elems > 0)
{
summ+= TemperatureBuffer[elems];
elems—;
}
average = summ/10;
/* defina th temperature type */
if (average > 40)
{
TemperatureType = 1;
}
else if ((average < 40) &&
(average > 10))
{
TemperatureType = 2;
}
else
{
TemperatureType = 3;
}
switch(TemperatureType) /* show message on
display */
{
case 1:
{
printf(“Temperature type - Hot, Temperature =
%f\n”, average);
}
break;
case 2:
{
printf(“Temperature type - Normal, Temperature
= %f\n”, average);
}
break;
default:
{
printf(“Temperature type - Cold, Temperature
= %f\n”, average);
}
break;
}
return ((short)average);/*return temperature
value*/
}
```

ДОДАТОК Б

```
(0)  short ReadTemperature (void)
{
unsigned Counter, elems,TemperatureType;
float TemperatureBuffer[10];
float Tmp = 0.0;
float summ = 0.0;
float average = 0.0;
 (1)      for (Counter = 0; Counter < 10;
Counter++)
{
(2)
StartConversionOfInternalADC_1();
(3)osDelay(10);
(4)TemperatureBuffer[Counter]      =
ReadTemperatureADC_1();
(5)elems++;
}
(6) while(elems > 0)
{
(7)summ+= TemperatureBuffer[elems];
(8)elems—;
}
(9)average = summ/10;
(10) if (average > 40)
{
(11)TemperatureType = 1;
}
else if ((average < 40) && (average > 10))
{
(12)TemperatureType = 2;
}
else
{
(13)TemperatureType = 3;
}
(14)switch(TemperatureType)
{
case 1:
{
(15) printf(“Temperature type - Hot,
Temperature = %f\n”, average);
}
break;
case 2:
{
(16)printf(“Temperature type - Normal,
Temperature = %f\n”, average);
}
break;
default:
{
(17) printf(“Temperature type - Cold,
Temperature = %f\n”, average);
}
break;
}
(18) return ((short)average);
}
```

УДК 004.93 : 621.311.238

Шитикова Е. В.[1], Табунщик С. С.[2], Табунщик Г. В.[3]

[1]Младший научный сотрудник кафедры программных средств Запорожского национального технического университета, Запорожье, Украина

[2]Ведущий инженер кафедры программных средств Запорожского национального технического университета, Запорожье, Украина

[3]Канд. техн. наук, доцент, профессор кафедры программных средств Запорожского национального технического университета, Запорожье, Украина

МЕТОД ФОРМИРОВАНИЯ ОБЪЕМА РАБОТ ДЛЯ ПРОГРАММ ИСПЫТАНИЙ НА ОСНОВЕ НЕЧЕТКОГО ВЫВОДА

Актуальность. Снижение затрат процесса испытаний может быть достигнуто за счет рационального планирования всех работ и оптимального распределения ресурсов. Одним из основных документов, регламентирующих цель и задачи испытаний, объем, порядок и методы проведения работ, является программа испытаний, процесс создания которой представляется как последовательность взаимосвязанных задач. В виду высоких затрат материальных и нематериальных ресурсов повышение качества проведения испытаний возможно за счет обоснованного выбора объема работ и контролируемых переменных. Объектом исследования является процесс испытаний сложных технических систем. Предметом исследования – методы сбора и анализа параметров работы, контролируемых в процессе испытаний сложной технической системы.

Цель работы – разработка метода формирования объема работ для программ испытаний сложных технических систем.

Метод. На основе разработанной ранее теоретико-множественной модели программы испытаний получен метод формирования объема работ для программы испытаний сложных технических систем, основанный на комбинации прямого логического вывода для формирования множества контролируемых параметров и нечетком логическом выводе для композиции работ, проведение которых целесообразно за один запуск сложной технической системы. Комбинация различных видов работ позволяет снизить прямые затраты и уменьшить стоимость процесса испытаний.

Результаты. Разработанный метод апробирован для испытаний газотурбинных установок наземного применения. В соответствии с классификацией затрат на процесс испытаний, использование метода формирования работ для программы испытаний позволило снизить затраты на ресурсы труда, капитала и альтернативные ресурсы, за счет сокращения количества экспериментов для подтверждения работоспособности газотурбинной электростанции.

Выводы. Разработанный метод формирования объема работы для программы испытаний сложных технических систем позволяет комбинировать работы, которые возможно выполнить за один запуск сложной технической системы. Применение разработанного метода позволило снизить затраты процесса испытаний за счет уменьшения времени и трудозатрат на создание программы испытаний, а также снижения количества запусков газотурбинной установки наземного применения при увеличении объема проведенных исследований за один запуск.

Ключевые слова: газотурбинные установки наземного применения, программа испытаний, контролируемые параметры, объем работ, нечеткий логический вывод, затраты.

НОМЕНКЛАТУРА

ГТП – газотурбинный привод;

ГТУ НП газотурбинная установка наземного применения;

САУ – система автоматизированного управления;

СТС сложная техническая система;

A^{CV*} – матрица отношений, отражающая попарные сравнения количества совпадающих элементов в подмножествах CV_i^* и CV_j^*;

a_{ij}^{CV*} – элемент матрицы A^{CV*};

Alt – множество альтернативных ресурсы для всех работ из подмножества *Work*;

Alt_i – альтернативные ресурсы для работы $work_i$;

$Behavior = \{behavior_1, ..., behavior_n\}$ – подмножество, содержащее диапазоны режимов работы СТС, на которых должны быть проведены соответствующие проверки, где n число работ;

$Controlled\ variable = \{cv_1, ..., cv_m\}$ – подмножество информативных (значимых) параметров работы СТС, где m количество параметров;

$Cost = \{cost_1, ..., cost_n\}$ – подмножество, содержащее затраты на проведение соответствующих работ, где n число работ;

$cost_i'$ – минимизированные затраты на проведение работы $work_i$;

CV^* – обобщающее множество параметров работы, которые необходимо контролировать при выполнении всех работ из подмножества *Work*;

CV_i^*, CV_j^* подмножества множества CV^*, содержащие достаточное количество контролируемых параметров, характеризующих эффективность работ $work_i$ и $work_j$ соответственно;

G – база предикативных правил;

Goal – цель испытаний;

K – множество ресурсов капитала для всех работ из подмножества *Work*;

K_i – ресурсы капитала для работы $work_i$;

L – множество ресурсов труда для всех работ из подмножества *Work*;

L_i – ресурсы труда для работы $work_i$;

M^{μ_A} – матрица функций принадлежности;

Obj – подмножество, описывающее объект испытаний;

r_i – цены ресурсов капитала K_i;

$R_{work \to CV}$ – отношение, которое для каждой работы $work_i$ ставит в соответствие подмножество CV_i^*;

© Шитикова Е. В., Табунщик С. С., Табунщик Г. В., 2018

DOI 10.15588/1607-3274-2018-2-18

Sim – матрица, содержащая отдельные степени принадлежности sim_{ij};

Span – подмножество, содержащее диапазоны допустимых значений контролируемых параметров;

Time = {$time_1$, ..., $time_n$} – подмножество, содержащее время на проведение соответствующих работ, где n – число работ;

TP – теоретико-множественная модель программы испытаний;

$T(CV^*)$ – нечеткий терм;

u_i – цены альтернативных ресурсов Alt_i;

Var – множество, содержащее все измеряемые и контролируемые САУ параметры работы СТС;

Verification – кортеж, описывающий проверки (работы), которые необходимо провести в процессе испытаний;

w_i – цены ресурсов труда L_i;

Work = {$work_1$, ..., $work_n$} – подмножество, содержащее перечень (наименование) проверок (работ), где n число работ;

$work^*_k$ – подмножество множества $Work^*$, содержащего перечень скомпонованных работ, где k – число запусков СТС;

ε – количество совпадающих элементов в паре сравниваемых подмножеств CV_i^* и CV_j^*;

$\mu_A\left(a_{ij}^{CV*}\right)$ – функция принадлежности a_{ij}^{CV*} нечетким термам $T(CV^*)$;

μ_{Rnm} ($work_n$, cv_m) – характеристическая функция, принимающая значения 0 или 1 в зависимости от соответствия параметра cv_m рассматриваемой работе $work_n$.

ВВЕДЕНИЕ

Процесс испытаний позволяет объективно установить соответствие показателей качества продукции требованиям нормативной документации, выявить фактические значения этих показателей, проверить, удовлетворяет ли продукция требованиям стандартов, сравнить качество изделий с качеством аналогов, подтвердить их надежность и безопасность [1].

Эффективное управление всем процессом испытаний сложных технических систем (СТС) в первую очередь зависит от рационального планирования всех работ и оптимального распределения ресурсов [2].

Проведение испытаний СТС и их узлов сопряжено с очень большими затратами. Затраты производства могут быть представлены как ценность израсходованных ресурсов в фактических ценах их приобретения [3].

Поскольку избежать этих затрат невозможно, надо особое внимание уделить этапу планирования испытаний при этом повышая их информативность, т.е. увеличивать объем, точность, достоверность результатов и сведений, получаемых в результате проведения каждой экспериментальной работы. Увеличивая объем измерений, можно получить дополнительную информацию и в одном и том же испытании выполнить несколько различных исследований [4].

Одним из основных документов, регламентирующих цель и задачи испытаний, объем, порядок и методы проведения работ, является программа испытаний [5].

Процесс создания программы испытаний может быть представлен как последовательность взаимосвязанных задач:

– идентификация объекта испытаний, определение цели и задач испытаний;

– определение порядка проведения испытаний и их обеспечения (распределение и использование материальных, временных и людских ресурсов);

– уточнение объема работ и их последовательности, определение оцениваемых характеристик и расчетных соотношений, методик испытаний;

– описание порядка обработки полученных результатов;

– определение формы и порядка отчетности.

Решение вышеперечисленных задач должно быть отражено в программе испытаний.

Объектом исследования является процесс испытаний сложных технических систем.

Предметом исследования – методы сбора и анализа параметров работы, контролируемых в процессе испытаний СТС.

Целью данной работы является разработка метода формирования объема работ для программ испытаний СТС.

1 ПОСТАНОВКА ЗАДАЧИ

При определении объема работ с формированием перечня необходимых проверок и контролируемых характеристик (параметров работы) необходимо учитывать следующее:

– процессы, протекающие в СТС, разнообразны и в тоже время взаимозависимы и случайны, что существенно усложняет их исследование и определение корреляционных связей между параметрами рабочего процесса;

– задача еще более усложняется при исследовательских испытаниях, во время которых ведется отработка конструктивных нововведений, взаимная работа которых до этого не исследовалась, что вносит дополнительную неопределенность контролируемых переменных;

– в виду высоких затрат материальных и нематериальных ресурсов повышение качества проведения испытаний возможно за счет обоснованного выбора объема работ и контролируемых переменных.

Теоретико-множественная модель программы испытаний может быть представлена в следующем виде:

$$TP = \langle Obj, Goal, Verification, Controlled\ vriable, Span \rangle, \quad (1)$$

где подмножество *Verification* может быть представлено кортежем:

$$Verification = <Work, Behavior, Cost, Time>. \quad (2)$$

Подмножество *Cost* представляет собой затраты на проведение соответствующих работ. Различают затраты в длительном периоде и в коротком периоде. В рамках данной работы рассмотрим затраты для процесса испытаний СТС в течение короткого периода, т.е. краткосрочные затраты.

В течение короткого периода неизменными считаются производственные мощности предприятия, а их использование может изменяться за счет изменения объема переменных факторов (числа обрабатываемых человека-часов, расхода сырья и материалов и т.д.) [3].

Наипростейший способ формирования множества *Controlled variable* включить в него все измеряемые и контролируемые системой автоматизированного управления (САУ) параметры (аналоговые и дискретные сигналы от контрольно-измерительной аппаратуры), которые используются для управления работой СТС. Однако, полученное в результате множество будет с одной стороны избыточным за счет наличия неинформативных параметров работы, а с другой стороны недостаточным, т.к. будут отсутствовать расчетные параметры работы, в том числе временные показатели (например, для газотурбинной электростанции такие параметры как «время восстановления частоты и напряжения сети при мгновенном сбросе/набросе нагрузки», которые характеризуют качество вырабатываемой электрической энергии).

Поэтому были поставлены следующие задачи:

– определить множество контролируемых характеристик (параметров работы) СТС в процессе испытаний;

– разработать метод формирования объема работ для программ испытаний СТС.

2 ОБЗОР ЛИТЕРАТУРЫ

Структуры и содержание программ испытаний изделий различных видов и назначения изложены в нормативных материалах и соответствующих государственных стандартах на программы и методики испытаний конкретных изделий. В целом для СТС применимы стандарты ГОСТ 15.309, ДСТУ ГОСТ 15.001 [6–7], регламентирующие разработку, испытания, приемку и постановку на производство продукции производственно-технического назначения. Также для конкретных видов СТС правила приемки представлены стандартами на соответствующие технические условия, например ГОСТ 28775, ГОСТ 29328 [8-9]. Данные стандарты описывают виды испытаний, в общих чертах требования к объему необходимых проверок, что впоследствии берется за основу при написании программ испытаний.

При определении множества контролируемых параметров необходимо провести рациональный и критический анализ данных. Для этого могут быть использованы различные методы анализа данных. Знания, которыми располагает человек, в какой-то степени всегда неполны, приближенны, ненадежны. Тем не менее, на основе таких знаний удается делать достаточно обоснованные выводы и принимать решения. Для решения такой проблемы применимы различные подходы: нейронные сети, data mining, нечеткая логика и др. [10–12].

Для формирования объема работ была выбрана нечеткая логика, основным отличием которой является введение лингвистических переменных (субъективных категорий), в качестве значений (термов) которых выступают нечеткие множества. Математическая теория нечетких множеств позволяет описывать нечеткие понятия и знания, оперировать ими и делать нечеткие заключения. Нечеткое управление оказывается особенно полезным, когда исследуемые процессы слишком сложны для анализа с помощью общепринятых методов или доступные источники информации интерпретируются лишь качественно, неточно или неопределенно [13–14].

3 МАТЕРИАЛЫ И МЕТОДЫ

Для уменьшения стоимости процесса испытаний необходимо снижение прямых затрат за счет комбинации различных видов работ. Рассмотрим метод определения оптимального объема работ и формирования множества контролируемых параметров [15].

На вход системы поступает множество *Var*, содержащее все измеряемые и контролируемые САУ параметры работы СТС.

На первом этапе формируется множество информативных (значимых) параметров работы СТС *Controlled variable*, которое будет включать как измеряемые, так и расчетные параметры работы СТС.

Далее на втором этапе определяется перечень работ, которые необходимо выполнить в процессе испытаний СТС, для чего формируется множество *Verification*, представленное формулой (2).

При формировании подмножества *Cost* примем, что для каждой *i*-ой работы процесса испытаний используются ресурсы труда L_i, капитала K_i и альтернативные ресурсы Alt_i, цены которых w_i, r_i и u_i заданы, тогда общие затраты предприятия для каждой из работ процесса испытаний могут быть представлены простым тождеством:

$$cost_i \equiv w_i L_i + r_i K_i + u_i Alt_i. \quad (3)$$

При этом цены альтернативных ресурсов u_i будут рассчитываться как ценность других благ, которые можно было бы получить при наиболее выгодном из всех возможных альтернативных направлений их использования [3].

На третьем этапе для каждой из работ $work_i$ подмножества *Work* необходимо поставить в соответствие подмножество $CV_i^* \subseteq$ *Controlled variable*, которое бы включало достаточное количество контролируемых параметров, характеризующих эффективность работы исследуемого объекта (системы, агрегата, процесса и т.д.) во время испытаний [16]. Например, для определения эффективности работы системы вентиляции такой СТС как газотурбинная установка наземного применения (ГТУ НП) такие параметры как «Давление масла на входе в ГТП» или «Расход топливного газа» не являются информативными в отличие от параметров «Температура воздуха в контейнере отсека ГТП» и «Температура наружного воздуха».

Для этого была сформирована база предикативных правил *G*, которая для каждой из возможных работ», проводимых при испытаниях СТС, содержит правила вывода типа «ЕСЛИ – ТО».

С помощью системы продукций, используя правило вывода modus ponens, было сформировано отношение

$$R_{work \to CV} = \begin{array}{c|ccc} & cv_1 & \cdots & cv_m \\ work_1 & \mu_{R_{11}}(work_1, cv_1) & \cdots & \mu_{R_{1m}}(work_1, cv_m) \\ \vdots & \vdots & \vdots & \vdots \\ work_n & \mu_{R_{n1}}(work_n, cv_1) & \cdots & \mu_{R_{nm}}(work_n, cv_m) \end{array}, \quad (4)$$

Тогда обобщающее множество параметров работы установки CV^*, которые необходимо контролировать при выполнении всех работ из подмножества *Work* можно представить в следующем виде:

$$CV^* = \begin{pmatrix} CV_1^* \\ \vdots \\ CV_n^* \end{pmatrix}. \quad (5)$$

Композиция работ из подмножества *Work* позволит сократить количество запусков СТС в процессе испытаний.

Для этого необходимо определить элементы каких из подмножеств CV_i^* приблизительно равны друг другу, что может свидетельствовать о целесообразности совместного проведения данных работ за один запуск. Именно субъективность лингвистического критерия «приблизительно равны» и является той неопределенностью, которая послужила предпосылкой к использованию нечеткого логического вывода.

На четвертом этапе попарные сравнения количества совпадающих элементов в подмножествах CV_i^* и CV_j^* представляем матрицей отношений $A^{CV*} = \left\{a_{ij}^{CV*}\right\}$ при

$$a_{ij}^{CV*} = \frac{\varepsilon}{\min\left(\left|CV_i^*\right|, \left|CV_j^*\right|\right)}. \quad (6)$$

При этом матрица A^{CV*} будет симметрична относительно главной диагонали, поэтому достаточно представления только половины матрицы.

На пятом этапе реализован нечеткий логический вывод. Для введения нечеткости определяем соотношение количества совпадающих элементов для всех возможных сочетаний пар подмножеств CV_i^* и $CV_j^* \subseteq CV^*$ с помощью нечетких термов $T(CV^*)$ = (*низкое совпадение, среднее совпадение, высокое совпадение)*.

Далее для каждого определенного элемента матрицы a_{ij}^{CV*} из верхнего треугольника экспертными методами задаются $\mu_A\left(a_{ij}^{CV*}\right) \in [0,1]$ функции принадлежности нечетким термам $T(CV^*)$. Таким образом, будет сформирована матрица функций принадлежности $M^{\mu_A} = \left\{\mu_A(a_{ij}^{CV*})\right\}$.

Логический вывод осуществляется с использованием базы правил нечеткого вывода G'.

Поскольку допускается применение нескольких нечетких правил, в системе предусмотрен блок агрегирования, формирующий на выходе системы матрицу $Sim=\{sim_{ij}\}$, где sim_{ij} – отдельные степени принадлежности [17].

Исходя из лингвистических значений sim_{ij} все работы, для которых контролируемые параметры, содержащиеся в подмножествах CV_i^* и CV_j^*, имеют «высокое совпадение» и равны режимы работы СТС, на которых должны быть проведены соответствующие проверки, рекомендуются к проведению за один запуск. Тогда будет справедлива запись:

IF sim_{ij}=“ВЫСОКОЕ совпадение” AND $behavior_i$ = = $behavior_j$ THEN $work_i \in work^*_k$ AND $work_j \in work^*_k$ [17].

Полагая цены ресурсов в коротком периоде неизменными, можно утверждать, что снижение затрат возможно за счет снижения количества используемых ресурсов [3]. Тогда композиция работ с условием минимизации количества запусков в конечном счете приведет к уменьшению затрат на проведение испытаний за счет уменьшения количества всех используемых ресурсов *L*, *K* и *Alt*. Поэтому с учетом формулы (3) будет справедлива запись:

$$cost_i' = \arg\min(cost_i) \equiv w_i \arg\min(L_i) + r_i \arg\min(K_i) + u_i \arg\min(Alt_i). \quad (7)$$

4 ЭКСПЕРИМЕНТЫ

Данный метод был опробован на этапе планирования испытаний вновь создаваемой модификации газотурбинной электростанции ЭГ-6000, производства АО «МОТОР СИЧ». На основе разработанной модели программы испытаний (1) и метода формирования объема работ была сформирована программа исследовательских испытаний для данной электростанции.

Количество контролируемых САУ входных и выходных аналоговых и дискретных сигналов составляет порядка 600 наименований (это сигналы, используемые для управления работой установки). Было выделено 95 информативных (значимых) параметров работы электростанции с формированием множества *Controlled variable*. Определено 19 работ, необходимых к проведению в процессе исследовательских испытаний, информация о которых была внесена во множество *Verification*.

Было построено отношение $R_{work \to CV}$, на основании которого сформировано итоговое множество CV^*, в котором каждое из 19 подмножеств CV_i^* содержит достаточное количество контролируемых параметров, характеризующих эффективность каждой из 19 работ $work_i$.

На следующем этапе работы метода были проведены попарные сравнения количества совпадающих элементов во всех подмножествах CV_i^* и сформирована матрица отношений A^{CV*} размерностью 19x19, элементы которой рассчитаны по формуле (5).

Далее была получена матрица $M^{\mu A} = \{\mu_A(a_{ij}^{CV^*})\}$, элементы которой определены как треугольные функции принадлежности элементов матрицы A^{CV*} нечетким термам $T(CV^*)$, при $\mu_A(a_{ij}^{CV^*}) \in [0,1]$.

С использованием базы правил нечеткого вывода сформирована матрица частных степеней сравнения $Sim=\{sim_{ij}\}$, на основании которой были выделены работы, рекомендованные к совместного проведению за один запуск ГТУ НП.

5 РЕЗУЛЬТАТЫ

Применение разработанного метода формирования объема работ для программы испытаний позволило сократить количество экспериментов (проверок) для подтверждения работоспособности функциональных систем ГТУ НП и правильности принятых конструктивных решений до 17 при уменьшении количества запусков с взятием нагрузки электростанцией до 9.

Для процесса испытаний ГТУ НП совокупные затраты можно подразделить как показано на рис. 1.

В виду того, что в рамках данной работы были рассмотрены затраты в коротком периоде, то Косвенные затраты в виде Капитальных расходов (рис.1) процесса испытаний ГТУ НП во внимание не принимались.

Таким образом, были снижены затраты на проведение процесса испытаний ГТУ НП за счет снижения ресурсов труда *L* в качестве уменьшения трудозатрат на проведение испытаний, за счет снижения количества запусков снижено использование ресурсов капитала *K* (материалов) и альтернативных ресурсов *Alt* в качестве уменьшения выработки собственного ресурса работы ГТУ НП (снижено количество часов наработки за весь цикл испытаний).

6 ОБСУЖДЕНИЕ

В табл. 1 приведена оценка результатов применения метода формирования объема работ для программы испытаний.

Применение данного метода формирования объема работ позволило сократить время и трудозатраты на создание программы испытаний на 37,5%, а также скомпоновать выделенные работы и сократить количество планируемых экспериментов на 11% при уменьшении пусков на 19%, что привело к снижению всех используемых ресурсов.

Дальнейшее направление работы будет также посвящено этапу планирования испытаний СТС, в частности автоматизации процесса создания графиков проведения испытаний, подбора квалифицированного персонала и т. д.

ВЫВОДЫ

Работа посвящена решению задачи определения объема работ в процессе испытаний с формированием перечня контролируемых параметров.

На основе разработанной ранее теоретико-множественной модели программы испытаний, содержащей множество проверок (работ) *Verification*, которые необходимо провести в процессе испытаний и множество контролируемых параметров *Controlled variable*, был разработан метод формирования объема работ для программы испытаний СТС.

Научной новизной работы является разработанный метод формирования объема работ для программы испытаний сложной технической системы, который в отличие от существующих основан на комбинации пря-

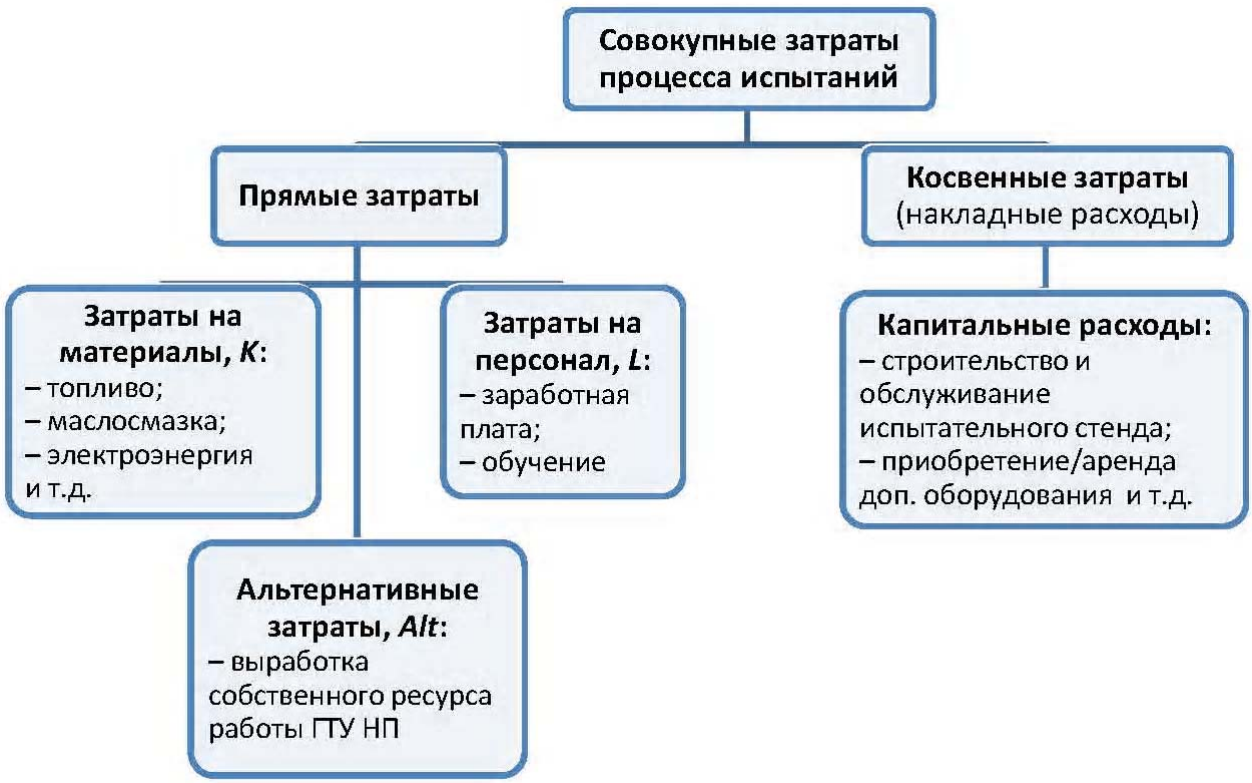


Рисунок 1 – Совокупные затраты процесса испытаний ГТУ НП

Таблица 1 – Сравнительная оценка результатов применения метода формирования объема работ для программы испытаний

Критерии сравнения	До применения метода	После применения метода
Время создания программы испытаний, дней	20	12,5
Трудозатраты на создание программы испытаний, чел./ч	160	100
Общее количество экспериментов (проверок) проводимых в процессе исследовательских испытаний	19	17
Количество экспериментов (проверок) с запуском электростанции и взятием нагрузки	11	9

мого логического вывода и нечеткого логического вывода, что позволило сформировать множество контролируемых параметров и осуществить композицию работ, которые целесообразно провести за один запуск СТС.

Практическая ценность данной работы заключается в том, что применение разработанного метода позволило снизить затраты процесса испытаний за счет уменьшения времени и трудозатрат на создание программы испытаний, а также снижения количества запусков ГТУ НП при увеличении объема проведенных исследований за один запуск.

Данный метод был апробирован для планирования испытаний газотурбинной электростанции.

БЛАГОДАРНОСТИ

Работа выполнена в рамках госбюджетной научно-исследовательской темы Запорожского национального технического университета «Інформаційна система діагностування розподілених мінікомп'ютерних систем в багатокомпонентному зовнішньому середовищі» (номер государственной регистрации 0117U000615) и при поддержке международного проекта Erasmus + Internet of Things: Emerging Curriculum for Industry and Human Applications ALIOT 573818-EPP-1-2016-1-UK-EPPKA2-CBHE-JP.

СПИСОК ЛИТЕРАТУРЫ

1. Handbook of Performability Engineering / [ed.: Krishna B. Misra]. – London : Springer-Verlag London Limitedr, 2008. – 1295 p. DOI 10.1007/978-1-84800-131-2.
2. Шитикова Е. В. Метод управления неопределенностью в условиях ограниченных ресурсов / Е. В. Шитикова, Г. В. Табунщик // Радіоелектроніка, інформатика, управління. – 2015. – № 2(33). – С. 87–95. DOI: 10.15588/1607-3274-2015-2-11.
3. Гальперин В. М. Микроэкономика / В. М. Гальперин, С. М. Игнатьев, В. И. Моргунов. – Т.1. – СПб. : Экономическая школа, 2004. – 349 с.
4. Теоретические основы испытаний и экспериментальная отработка сложных технических систем : учебное пособие / [Л. Н. Александровская, В. И. Круглов, А. Г. Кузнецов и др.]. – М. : Логос, 2003. – 736 с.
5. Випробування і контроль якості продукції. Терміни та визначення : ДСТУ 3021–95. – [Чинний від 1996-01-01]. – К. : Держстандарт України, 1995. – 74 с. – (Національний стандарт України).
6. Система разработки и постановки продукции на производство. Испытания и приемка выпускаемой продукции. Основные положения : ГОСТ 15.309-98. – [Дата введения 01.01.2000]. – М. : Стандартинформ, 2008. – 16 с. – (Межгосударственный стандарт).
7. Система розроблення та поставлення продукції на виробництво : ДСТУ ГОСТ 15.001:2009. – [Чинний від 2009-02-01]. – К. : Держстандарт України, 2009. – 9 с. – (Національний стандарт України).
8. Агрегаты газоперекачивающие с газотурбинным приводом. Общие технические условия : ГОСТ 28775-90. – [Дата введения 01.01.1992]. – М. : Стандартинформ, 2005. – 11 с. – (Межгосударственный стандарт).
9. Установки газотурбинные для привода турбогенераторов. Общие технические условия : ГОСТ 29328-92. – [Дата введения 01.01.1993]. – М. : ИПК Издательство стандартов, 2004. – 12 с. – (Межгосударственный стандарт).
10. Parallel computing system resources planning for neuro-fuzzy models synthesis and big data processing / [A. Oliinyk, S. Skrupsky, S. Subbotin et al.] // Radio Electronics, Computer Science, Control. – 2016. – Vol. 4. – P. 61–69. DOI: 10.15588/1607-3274-2016-4-8.
11. The model for estimation of computer system used resources while extracting production rules based on parallel computations / [A. A. Oliinyk, S. Yu. Skrupsky, V. V. Shkarupylo et al.] // Radio Electronics, Computer Science, Control. – 2017. – Vol. 1. – P. 142–152. DOI: 10.15588/1607-3274-2017-1-16.
12. Performance Measurement with Fuzzy Data Envelopment Analysis / [eds.: Ali Emrouznejad, Madjid Tavana]. – Heidelberg : Springer, 2014. – 288 p. DOI 10.1007/978-3-642-41372-8.
13. Вахитов А. Р. Использование нечеткого логического вывода для интеллектуального анализа данных / А. Р. Вахитов, В. А. Силич // Известия Томского политехнического университета. – 2010. – Т. 317, № 5. – С. 171–174.
14. Исследование нечетких нейронных сетей в задачах макроэкономического прогнозирования / [Ю. П. Зайченко, Севаее Фатма, К. М. Титаренко и др.] // Системні дослідженні та інформаційні технології. – 2004. – № 2. – С. 70–86.
15. Шитикова Е. В. Модельно-ориентированное проектирование испытаний слжных технических объектов : монография / Е. В. Шитикова, С. Б. Беликов, Г. В. Табунщик. – Запорожье : Дике поле, 2017. – 170 с.
16. Метод синтеза множества контролируемых переменных для диагностирования отказов турбореактивных двигателей в условиях неопределенности входных данных / [В. Е. Стрелец, М. Л. Угрюмов, Е. М. Угрюмова и др.] // Вісник двигунобудування. – 2015. – № 2. – С. 120–124.
17. Шитікова О. В. Інформаційна технологія підтримки процесу випробувань газотурбінних установок наземного використання з урахуванням невизначеності : автореф. дис. … к. техн. наук : 05.13.06 «Інформаційні технології» / Шитікова О.В ; М-во освіти і науки України, Херсонський Національний Технічний Університет. – Херсон, 2017. – 24 с.

Статья поступила в редакцию 31.08.2017.
После доработки 25.09.2017.

Шитікова О. В.[1], Табунщик С. С.[2], Табунщик Г. В.[3]

[1]Молодший науковий співробітник кафедри програмних засобів Запорізького національного технічного університету, Запоріжжя, Україна

[2]Провідний фахівець кафедри програмних засобів Запорізького національного технічного університету, Запоріжжя, Україна

[3]Канд. техн. наук, доцент, професор кафедри програмних засобів Запорізького національного технічного університету, Запоріжжя, Україна

МЕТОД ФОРМУВАННЯ ОБСЯГУ РОБІТ ДЛЯ ПРОГРАМ ВИПРОБУВАНЬ НА ОСНОВІ НЕЧІТКОГО ВИВЕДЕННЯ

Актуальність. Зниження витрат процесу випробувань може бути досягнуто за рахунок раціонального планування всіх робіт і оптимального розподілу ресурсів. Одним з основних документів, що регламентує мету і завдання випробувань, обсяг, порядок і методи проведення робіт, є програма випробувань, процес створення якої представляється як послідовність взаємопов'язаних завдань. У зв'язку зі значними матеріальними і нематеріальними витратами ресурсів, підвищення якості проведення випробувань можливо за рахунок обгрунтованого вибору обсягу робіт і контрольованих змінних. Об'єктом дослідження є процес випробувань складних технічних систем. Предметом дослідження – методи збору та аналізу параметрів роботи, що контролюються під час випробувань складних технічних систем.

Мета роботи – розробка методу формування обсягу робіт для програм випробувань складних технічних систем.

Метод. Розроблено метод формування обсягу робіт для програми випробувань складних технічних систем, заснований на комбінації прямого логічного виведення для формування множени параметрів, що контролюються і нечіткому логічному виведення для композиції робіт, проведення яких доцільно за один запуск складної технічної системи. Комбінація різноманітних видів робіт дозволяє знизити прямі витрати та зменшити вартість процесу випробувань.

Результати. Розроблений метод апробований для планування випробувань газотурбінних установок наземного використання. Відповідно до класифікації витрат на процес випробувань, використання метода формування робіт для програми випробувань дозволило знизити витрати на ресурси труда, капіталу та альтернативні ресурси, за рахунок зниження кількості експериментів для підтвердження працездатності газотурбінної електростанції.

Висновки. Розроблений метод формування об'єму робіт для програми випробувань складних технічних систем дозволяє комбінувати роботи, що можливо виконати за один запуск складної технічної системи. Застосування розробленого методу дозволило знизити витрати процесу випробувань за рахунок зменшення часу та трудовитрат на створення програми випробувань, а також зниження кількості запусків газотурбінної установки наземного використання при збільшенні обсягу проведених досліджень за один запуск.

Ключові слова: газотурбінні установки наземного використання, програма випробувань, контрольовані параметри, обсяг робіт, нечіткий логічний вивід, витрати.

Shytikova O[1]., Tabunshchyk S.[2], Tabunshchyk G.[3]

[1]Junior Researcher, Software Tools Department, Zaporizhzhia National Technical University, Zaporozhia, Ukraine

[2]Leading Engineer, Software Tools Department, Zaporizhzhia National Technical University, Zaporozhia, Ukraine

[3]PhD, Professor, Software Tools Department, Zaporizhzhia National Technical University, Zaporozhia, Ukraine

METHOD FOR FORMING TEST WORKFLOW BASED ON FUZZY INFERENCE

Actuality. Test process costs can be reduced by rational planning of all works and the optimal allocation of the resources. One of the main documents which regulate the purpose and tasks of the test process, its' scope, procedure and methods of carrying out the work, is a test program. The process of it creation is represented as a sequence of interrelated tasks. In the view of the high costs of material and non-material resources, the improvement of the testing is possible due to a reasonable choice of scope of work and controlled variables. The object of the research – the test process of the compound technical system. The subject of the research – the method for collection and analysis of controlled parameters within test process of the compound technical system.

Goal of the research is to develop a method for forming the scope of testing workflows which will be included in the compound technical system test program.

Method. A method for forming of the testing workflows scope for a test program for compound technical systems is developed by the authors. It is based on a combination of direct logical inference for the formation of a set of monitored parameters and a fuzzy logical conclusion for the composition of works that are expedient for one launch of the compound technical systems. Combination of different work types reduces direct costs and reduces the test process cost.

Results. Developed method was used for the tests of the gas turbine unit for terrestrial usage. According to the test costs classification, usage of the suggested approach for forming test process workflows reduced labor costs, capital (materials) resources and alternative resources, due to the reducing of experiments number for a gas turbine power station.

Conclusions. Method for workflows forming for test program of compound technical systems allows combining working processes, which are possible to run within one launch of the technical system. The application of the method has reduced test process costs by reducing time and man-hour for creation of test program, as well as reducing number of launches of the gas turbine unit, with the increase of work scope for one launch.

Keywords: gas turbine unit for terrestrial usage, test program, controlled parameters, scope of work, fuzzy inference, costs.

REFERENCES

1. Krishna B. Misra : ed. Handbook of Performability Engineering. London, Springer-Verlag London Limitedr, 2008, 1295 p. DOI 10.1007/978-1-84800-131-2.
2. Shitikova E. V., Tabunshhik G. V. Metod upravleniya neopredelennost'yu v usloviyax ogranichennyx resursov, *Radio Electronics, Computer Science, Control,* 2015, No. 2(33), pp. 87–95. DOI: 10.15588/1607-3274-2015-2-11.
3. Gal'perin V. M., Ignat'ev S. M., Morgunov V. I. Mikroe'konomika, Vil.1., Sankt-PeterburgE'konomicheskaya shkola, 2004, 349 p.
4. Aleksandrovskaya L. N., Kruglov V. I., Kuznecov A. G. i dr. Teoreticheskie osnovy ispytanij i e'ksperimental'naya otrabotka slozhnyx texnicheskix sistem : uchebnoe posobie. Moscow, Logos, 2003, 736 p.
5. Vyprobuvannja i kontrol' jakosti produkcii'. Terminy ta vyznachennja : DSTU 3021–95. [Effective from 1996-01-01]. Kiev, Derzhstandart Ukrai'ny, 1995, 74 p. (National standard of Ukraine).
6. Sistema razrabotki i postanovki produkcii na proizvodstvo. Ispytaniya i priemka vypuskaemoj produkcii. Osnovnye polozheniya : GOST 15.309-98. [Effective from 01.01.2000]. Moscow, Standartinform, 2008, 16 p. (Interstate standard).
7. Systema rozroblennja ta postavlennja produkcii' na vyrobnyctvo : DSTU GOST 15.001:2009. [Effective from 2009-02-01]. Kiev, Derzhstandart Ukrai'ny, 2009, 9 p. (National standard of Ukraine).
8. Agregaty gazoperekachivayushhie s gazoturbinnym privodom. Obshhie texnicheskie usloviya : GOST 28775-90. [Effective from 01.01.1992]. Moscow, Standartinform, 2005, 11 p. (Interstate standard).
9. Ustanovki gazoturbinnye dlya privoda turbogeneratorov. Obshhie texnicheskie usloviya : GOST 29328-92. [Effective from 01.01.1993]. Moscow, IPK Izdatel'stvo standartov, 2004, 12 p. (Interstate standard).
10. Oliinyk A., Skrupsky SSubbotin S. et al Parallel computing system resources planning for neuro-fuzzy models synthesis and big data processing, *Radio Electronics Computer Science Control*, 2016, Vol. 4, pp. 61–69. DOI: 10.15588/1607-3274-2016-4-8.
11. Oliinyk A. A., Skrupsky S. Yu., Shkarupylo V. V. et al. The model for estimation of computer system used resources while extracting production rules based on parallel computations, *Radio Electronics, Computer Science, Control,* 2017, Vol.1, pp. 142–152. DOI: 10.15588/1607-3274-2017-1-16.
12. Ali Emrouznejad, Madjid Tavana: eds. Performance Measurement with Fuzzy Data Envelopment Analysis. Heidelberg, Springer, 2014, 288 p. DOI 10.1007/978-3-642-41372-8.
13. Vaxitov A. R., Silich V. A. Ispol'zovanie nechetkogo logicheskogo vyvoda dlya intellektual'nogo analiza dannyx, *Izvestiya Tomskogo politexnicheskogo universiteta*, 2010, Vol. 317, No. 5, pp. 171–174.
14. Zajchenko Yu. P., Sevaee Fatma, Titarenko K. M. i dr. Issledovanie nechetkix nejronnyx setej v zadachax makroe'konomicheskogo prognozirovaniya, *Sistemni doslidzhenni ta informacijni texnologiï*, 2004, No. 2, pp. 70–86.
15. Shitikova E. V., Belikov S. B., Tabunshhik G. V. Model'no-orientirovannoe proektirovanie ispytanij slzhnyx texnicheskix ob'ektov : monografiya. Zaporozh'e, Dyke pole, 2017, 170 p.
16. Strelec V. E., Ugryumov M. L., Ugryumova E. M. i dr. Metod sinteza mnozhestva kontroliruemyx peremennyx dlya diagnostirovaniya otkazov turboreaktivnyx dvigatelej v usloviyax neopredelennosti vxodnyx dannyx, *Visnyk dvygunobuduvannja*, 2015, No. 2, pp. 120–124.
17. Shytikova O. V. Informacijna tehnologija pidtrymky procesu vyprobuvan' gazoturbinnyh ustanovok nazemnogo vykorystannja z urahuvannjam nevyznachenosti : avtoref. dys. … kand. tekhn. nauk : 05.13.06 «Informacijni tehnologii'». M-vo osvity i nauky Ukrainy, Khersonskyi Natsionalnyi Tekhnichnyi Universytet. Kherson, 2017, 24 p.

УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

CONTROL IN TECHNICAL SYSTEMS

УДК 004.415.25

Святный В. А.[1], Иванов А. Ю.[2], Иванов Ю. А.[3]

[1]Д-р техн. наук, заведующий кафедры компьютерной инженерии Донецкого национального технического университета, Покровск, Украина

[2]Ст. преподаватель кафедры компьютерной инженерии Донецкого национального технического университета, Покровск, Украина

[3]Канд. техн. наук, инженер-программист, Киев, Украина

ИССЛЕДОВАНИЕ ЦИКЛИЧЕСКОГО РАСПИСАНИЯ ДЛЯ МОДЕЛИ СЕТЕВОГО ДИНАМИЧЕСКОГО ОБЪЕКТА РЕАЛЬНОГО ВРЕМЕНИ

Актуальность. Рассмотрено создание моделей автоматизированных систем управления проветриванием, предназначенных для предотвращения возникновения аварийных ситуаций в шахтной вентиляционной сети. Решается задача исследования и реализации программных средств моделирования, которые дают возможность выполнения экспериментов с реальными наукоемкими объектами.

Цель работы – повышение эффективности цифровой полунатурной модели шахтной вентиляционной сети за счет разработки способа организации вычислительного процесса, учитывающего разночастотные параметры моделей, и его оптимизации.

Метод. Исследован, разработанный авторами, алгоритм планирования и реализация циклического расписания для модели динамического объекта реального времени, представленного шахтной вентиляционной сетью. Анализ математической модели аэродинамических процессов позволил отнести ее к классу жестких систем реального времени. Вычисления фазовых переменных, изменяющихся с различной скоростью, выполняется путем расщепления исходного алгоритма на параллельные потоки в программной модели. Численное интегрирование системы дифференциальных уравнений математической модели объекта выполнено в соответствии с алгоритмом статических приоритетов и управлением при помощи временной шкалы теории расписаний. Определена допустимость расписаний для директивных сроков переменных объекта. Алгоритм планирования учитывает временные параметры программных средств как однопроцессорных, так и симметричных мультипроцессорных систем.

Результаты. Анализ результатов моделирования предложенной модели с использованием циклических расписаний показывает соответствие получаемых результатов моделирования эталонным.

Выводы. Основным достоинством модели является повышение производительности при выполнении обработки данных за счет уменьшения использования процессорного времени. Реализованный алгоритм циклического расписания полунатурной моделирующей платформы используется для последующей разработки системы управления технологическими и аварийными режимами проветривания в шахтах.

Ключевые слова: моделирование, циклическое расписание, многоядерная система, расширение реального времени.

НОМЕНКЛАТУРА

РВ-цикл – повторяющееся выполнение части всех потоков интегрирования переменных в реальном времени;

РРВ – расширение реального времени для операционной системы;

РИ – различная инерционность;

ШВС – шахтная вентиляционная сеть;

ОДУ – обыкновенное дифференциальное уравнение;

ЗМ – задача моделирования;

SMP – symmetric multiprocessing;

Q_j – значения расхода воздуха в ветвях сети;

P_B – давление вентилятора;

C – значение расхода метана;

τ_i – время вычисления i-ой переменной процессором;

T_i – директивный срок вычисления i-ой переменной в реальном времени;

p – накладные временные расходы операционной системы;

$\Delta_i\tau_i$ – время выполнения части потока на РВ-цикле;

L_j – период РВ-цикла;

h_i – шаг интегрирования численного метода.

ВВЕДЕНИЕ

Для полунатурных моделирующих устройств динамических объектов реального времени с реальной аппа-

© Святный В. А., Иванов А. Ю., Иванов Ю. А., 2018
DOI 10.15588/1607-3274-2018-2-19

ратурой наиболее характерными типами приложений являются hardware in the loop (HIL) и software in the loop (SIL) [2, 3]. Особенность моделируемых процессов в этих приложениях состоит в том, что их математическое описание содержит системы дифференциальных уравнений. Для моделей жестких систем уравнений [4] является существенной разница в скорости изменения фазовых переменных, описывающих объект моделирования. Частоты переменных моделей реального времени изменяются в диапазоне: сотни герц для процессов медленной динамики и сотни килогерц в сверхбыстрых переходных процессах [5]. Обеспечение таких характеристик требует специальной организации вычислений при создании моделей реального времени [6, 7].

В качестве динамического объекта для исследования циклического расписания выбрана ШВС. Фрагмент традиционного для ШВС тестового объекта представляется в виде ориентированного графа (рис. 1), состоящего из ветвей (выработок) и узлов (мест соединения выработок друг с другом). Его ветви обозначены переменными $Q_j[\text{м}^3/c]$. Аэродинамические процессы в тестовой для ШВС модели вызываются вентилятором, который установлен в Q_B, развивает давление $P_B = 4100[H/\text{м}^2]$. Проектирование систем управления проветриванием ШВС базируется на анализе изменения значений расхода воздуха Q, метана C и давления P во всех ветвях. В любом сечении выработки уравнения аэродинамического состояния описываются дифференциальными уравнениями в частных производных [8, 9] для воздуха и метана. Для упрощения математической модели с достаточной для практики точностью допустимо использовать и широко применяются ОДУ. В j-ой ветви процессы изменения расхода воздуха являются первой частотной составляющей системы.

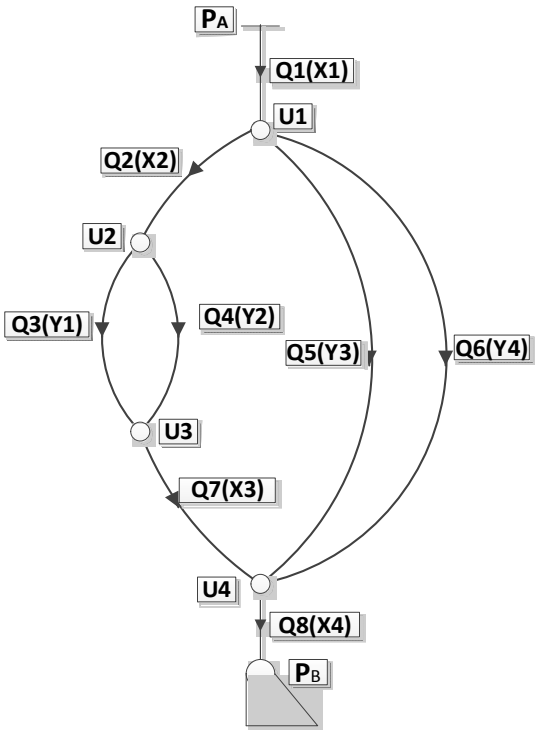


Рисунок 1 – Схема ШВС

Период времени T_1 для шага интегрирования по расходу воздуха (h_1) в модели соответствует собственной частоте изменения Q_j реальной сети и составляет 2000 [мкс]. Второй частотной составляющей системы ОДУ является изменения показателей метана в выработках. Период времени T_2, соответствующий собственной частоте изменения метана C (шага интегрирования h_2) составляет 9000 [мкс]. Технологические постоянные, которые задействованы в уравнениях аэродинамики ШВС, подробно описаны в [8, 9].

Представление динамики аэрогазодинамического объекта математической моделью с различными скоростями (РИ) изменения параметров используется при решении задач с критичными требованиями реактивности реагирования при установленных директивных сроках. Примерами таких задач являются: распространение пожаров в шахтах, режим помпажа. Эти явления сопровождаются резкими срывами воздушных потоков, сильным гидравлическим ударам.

Задачи исследования технологических и аварийных режимов управления проветриванием путем регулирования расхода воздуха для поддержания концентрации метана в нормативных пределах в вентиляционной сети как сложном РИ объекте реального времени имеют практическую значимость для угледобывающей отрасли.

1 ПОСТАНОВКА ЗАДАЧИ

Задачи оптимизации работы средств моделирования сложных динамических систем в реальном масштабе времени требуют учета особенностей моделей при реализации. В статье предлагается разработать на основании измеренных временных характеристик и физических РИ параметров ШВС равномерное циклическое расписание [13] и оптимизировать [14] программные средства организации вычислений модели динамической системы реального времени. Исследование циклического расписания провести для реализации, как для однопроцессорной, так и для многоядерной системы.

2 ОБЗОР ЛИТЕРАТУРЫ

Цифровая часть модели жесткой динамической системы рассмотрена как один их видов SIL, HIL-задач реального времени [1]. Вычисления фазовых переменных в модели выполняется путем расщепление исходного алгоритма способом параллелизма задач [10]. Каждая из задач должна быть запущена и полностью выполнить свою работу за время τ_i, (i=1–n) на каждом соответствующем этой задаче периоде времени T_i точно в срок. Временные модели вычисления фазовых переменных при интегрировании, рассматриваются как периодические задачи, задаваемые набором (τ_i, T_i). При этом величина $\sum_{i=1}^{n} \frac{\tau_i}{T_i}$ определяет загрузку процессора.

Для формального описания модели как системы реального времени используется модель synchronous data flow [10]. Описание процесса выполнения модели целесообразно формализовать методами теории расписаний [11]. Доказано [12], что допустимость и качество построения расписания зависит исключительно от директив-

ных частот вызова программ обработки или скоростей процессов, протекающих в модели.

Для разработки модели ШВС представляется перспективным применение оптимизированных циклических расписаний. Задача оптимизации рассматриваемого типа циклического расписания решена в [13] для однопроцессорной и в [14] для многоядерной SMP системы.

Анализ литературы [15] показывает, что для выполнения программ в режиме реального времени имеется широкий выбор различных сред. Расширения для операционных систем Linux и Windows, функциональность и временные характеристики, сопоставимые со специализированными операционными системами, обеспечивают экономию на программном и аппаратном обеспечении. Основу РРВ составляет специально реализованный слой аппаратной абстракции, которые позволяют работать с аппаратными ресурсами на уровне ядра операционной системы в режиме реального времени. В качестве платформы для разработки системы моделирования ШВС использована РРВ RTX 2012 для Windows фирмы Interval Zero [16] с монопольная архитектурой и разрешением таймера системы 0,01 [мс].

3 МАТЕРИАЛЫ И МЕТОДЫ

Первым шагом создания модели является анализ и преобразование ее математического описания. Для системы уравнений, описывающих расход воздуха Q_j, применяется метод прямых, что позволяет получить систему ОДУ, описывающую неустановившееся движение воздуха в ШВС [8]. Разностные уравнения всех ветвей получены разбиением на участки длиной 40 метров. Каждая j ветвь, будет разбита на $U_j = \{20,4,49,36,43,24,4,17\}$ участков, соответственно. Общее количество уравнений, описывающих аэродинамические процессы объекта, равно удвоенной сумме количества всех участков $2\times\sum_{j=1}^{8}U_j - 4 = 390$ за исключением четырех вершин, в которых ветви соединяются, и давления в них будут одинаковы. Для каждой из ветвей $(Y_1 - Y_4)$ в модель включены 4 уравнения, описывающие газовые процессы. Тогда общее количество уравнений, которые описывают аэрогазодинамическую модель объекта, составит 406.

В цифровой модели выполняется обработка задач с учетом особенностей РИ объектов в виде модели параллельных заданий. Интегрирование системы по воздуху и метану выполняется в отдельных потоках программного проекта в соответствии с алгоритмом статических приоритетов. Каждый параллельный поток модели содержит программу расчета соответствующих переменных системы ОДУ. Выбрана бесприоритетная циклическая дисциплина, причем первым выполняется поток с наибольшей частотой повторения.

Для разработки эффективного расписания проведены измерения временных характеристик данной платформы, после чего генерируется расписание и датаграмма выполнения. Она передается на модуль системы управления для запуска HIL моделирования. На вход модуля планирования подаются полученные на этапе профилирования τ_i и p (накладные временные расходы), а также n, T_i от разработчиков модели, что позволяет проверить условие существования расписания $\sum_{i=1}^{n}\frac{\tau_i}{T_i}\le 1$ [7]. В зависимости от целевой платформы выполняется расчет численных значений модели по алгоритму оптимизации для системы с однопроцессорной цифровой частью или для SMP системы. Расчет долей выполнения каждого потока $\Delta_i\tau_i = \frac{\tau_i}{T_i}L_j$, ($j$ – номер ядра), выполняется на основании периода L_j РВ-цикла. После чего полученные параметры модели преобразуются в датаграмму расписания для управления процессами реального времени на основе принципа временной шкалы. Вызовы потоков интегрирования системы по воздуху и метану выполняется согласно разработанной структуре – датаграмме управления циклическим расписанием обработки (рис. 2).

В основе реализации управления вычислением каждой фазовой переменной лежит установка таймера с заданным L_j для каждого из ядер. При создании, таймеру устанавливается максимальный приоритет в системе, чтобы обеспечить порядок следования согласно разработанному расписанию. Обработчик таймера выполняет последовательный запуск всех назначенных на ядро M_j потоков ЗМ в указанном порядке. Каждому потоку предоставляется время для выполнения его части ($\Delta_i\tau_i$), после чего его работа приостанавливается. Только один из потоков ядра может завершаться полностью на одном РВ-цикле в том случае, если его период вызова соответствует периоду цикла. Работа всех остальных потоков будет приостановлена и продолжена уже в следующем цикле. После вызова всех необходимых потоков

Рисунок 2 – Формат датаграммы расписания

работа функции обработчика таймера прекращается. Ее повторный вызов инициируется сигналом от таймера, который соответствует наступлению нового РВ-цикла.

Реализация управления циклическим расписанием реального времени на системе RTX 2012 выполнена в среде Microsoft Visual Studio (VS) на языке C++ с использованием специальных системных функций WINAPI библиотек RTX. Все выполняемые потоки созданы с установкой параметра *Affinitymask,* который задает их привязку к необходимому ядру системы [17,18]. Для организации функционирования обработчика таймера РВ-цикла используется функция RtCreateTimerEx(), которая позволяет запустить таймер с необходимой частотой и выполнить его привязку к заданному ядру.

Основными исполнительными частями модуля управления потоками являются: главная функция, функции обработки событий таймера и функции расчета значений уравнений модели. Главная функция выполняет инициализацию начальных значений модели, создание объектов таймеров, событий управления, потоков интегрирования, запуск таймеров и потоков [19]. Количество функций обработки событий от таймера, которые вызываются каждый раз равномерно с установленным периодом L_j, соответствует числу задействованных в расписании ядер. Функции расчета значений фазовых переменных выполняют интегрирование согласно численному методу Рунге-Кутты и их количество соответствует числу потоков ЗМ.

4 ЭКСПЕРИМЕНТЫ

Разработка параллельной модели для исследования циклического расписания выполнена на четырех ядерной системе с Windows 7 и установленным РРВ. Измерение временных характеристик исследуемых функций потоков выполнено с использованием стандартных инструментов профилирования среды разработки VS с учетом механизмов синхронизации и защиты данных [19]. Это позволяет выполнять разработку и исследование всех программных компонентов ЗМ в единой оболочке. С помощью утилит VS получены величины $\tau_1 = 1260$ [мкс] для функций воздуха и $\tau_2 = 100$ [мкс] для расчета значений системы по метану. Директивные величины $T_1 = 2000$ [мкс], $T_2 = 9000$ [мкс] выбраны так, чтобы они соответствовали частотам изменения физических величин в объекте. Эти периоды соответствуют выбору, соответственно $h_1 \approx 0{,}002$ и $h_2 \approx 0{,}009$. Измерение временных задержек среды реального времени, в которой выполняется ЗМ, основано на использовании программируемых таймеров. В циклическом расписании реального времени измерены и учтены временные задержки тестовой платформы: управления функцией обработки (максимальное значение отклонения от таймера системы d не превышает 0,00002 [мс]), передачи управления из функции обработки событий таймера, конкретной функции обработки одного из потоков (значение задержки не более 0,02 [мкс]), самой функции обработки (составило в худшем случае не более 0,06 [мкс]). Разработанная на подготовительном этапе сервисная программа определения задержек, основывается на измерении времени с использованием функции RtGetClockTimerPeriod (CLOCK_FASTEST). Величина p может быть рассчитана как сумма всех трех задержек. В данном случае она составляет p=0,1 [мкс].

С использованием модуля расчета значения РВ-цикла определено оптимальное [15, 16] значение L=1000[мкс]. На основании полученных исходных параметров циклического расписания ШВС (M=2 p=0,1): характеристик потока расчета по воздуху ($\tau_i = 1260$, $T_i = 2000$) и характеристик потока расчета по метану ($\tau_i = 2000$, $T_i = 9000$) выполняется проверка условия существования расписания:

$$\sum_{i=1}^{M} \frac{\tau_i}{T_i'} = 0{,}6411 < 1 \ - \text{условие выполняется.}$$

Вычисляются доли выполнения каждого потока $\Delta_i \tau_i = \frac{\tau_i}{T_i} L$ на периоде РВ-цикла, на основании полученных значений L с использованием модуля разработки расписаний:

$$\Delta_1 \tau_1 = \frac{1260}{2000} \cdot 1000 = 630 \ \text{[мкс]},$$

$$\Delta_2 \tau_2 = \frac{100}{9000} \cdot 1000 = 11{,}1 \ \text{[мкс]}.$$

Полученные временные интервалы должны быть обеспечены при обработке в функциях программных потоков. Такой вариант организации вычислительного процесса соответствует построенной датаграмме № 1:

1

1000 2 1 630 2 12

Для реализации на двуядерной системе, модулем генерации расписаний каждый из потоков был назначен на отдельное ядро с расчетными периодами циклов: $L_1 = 2000$ [мкс] и $L_2 = 9000$ [мкс] (рис. 2).

По полученным датаграммам выполняется настройка параметров таймеров и функций задержки в системы управления вычислительным процессом. Для первого расписания выполняется создание и привязка таймера (*hTimer*) к одному из ядер системы реального времени функцией *RtCreateTimerEx()*. Для запуска таймера (*hTimer*) устанавливается интервал вызова функцией *RtSetTimerRelative()*. В функции обработки таймера выполняется возобновление работы программного потока расчета системы ОДУ по воздуху *ResumeThread()* и ожидание его выполнения в течение 630 [мкс] (функция RtSleepFt(6300), после чего он останавливается функцией *SuspendThread()*. Далее происходит аналогичный вызов программных потоков по метану с ожиданием 12 [мкс] функцией *RtSleepFt()*. Для второй датаграммы создается два таймера *hTimer1* (2 [мс]) и *hTimer2* (9 [мс]), каждый из них привязывается к собственному ядру. В функциях обработки каждого из потоков выполняется ожидание необходимого временного интервала программного потока: по воздуху функцией *RtSleepFt(12600)* и по метану функцией *RtSleepFt(1000)*.

5 РЕЗУЛЬТАТЫ

Экспериментальное исследование эффективности при многопоточном [18] моделировании аэрогазодинамического объекта выполнена для трех различных режимов работы (рис.3): интегрирование модели по метану и воздуху выполняются традиционно с одинаковой частотой (рис. 3а); вычисления согласно усовершенствованному методу планирования на одном ядре с использованием РИ по сгенерированному расписанию (рис. 3б); вычисления модели с РИ по составленному расписанию на двух ядрах (рис. 3в).

Цель первого эксперимента заключается в определении эталонных расчетных значений для оценки эффективности разработанного подхода. С использованием функций замера времени RTXAPI, были получены затраты процессорного времени при моделировании: время работы потока обработки по воздуху (24,591 [с]), время работы потока обработки по метану (1,997 [с]) и суммарное процессорное время модели (27,221 [с]). Реальное время выполнения определялось от точки входа в функцию обработки, то есть после срабатывания сигнала от таймера до передачи управления во время возврата из функции.

Второй эксперимент выполнялся на том же временном интервале, но каждое из уравнений системы по Q, C просчитывается с собственной частотой, которые соответствует периодам (T_1, T_2). Порядок и временные параметры работы модели системы управления устанавливались согласно рассчитанной датаграммы №1. В результате профилирования модели получены следующие результаты: время работы потока обработки по воздуху (24,585 [с]), время работы потока обработки по метану (0,444 [с]) и суммарное процессорное время ЗМ (25,643 [с]). Полученные результаты свидетельствуют о существенном уменьшении временных затрат на интегрирование системы по метану, которое равно $\frac{2}{0{,}445} = 4{,}494$ раза. Общее сокращение используемого процессорного времени составило 5,7%.

Для третьего эксперимента выполнение потоков происходит на разных ядрах системы. Потоки, моделирующие систему по воздуху, запускаются с периодом T_1 на первом ядре, а поток, в котором интегрируются значения системы по метану с периодом T_2 выполняются на втором ядре. Анализ такой организации вычислительного процесса показал результаты: время работы потока обработки по воздуху (24,590 [с]), время работы потока обработки по метану (0,445 [с]) и суммарное процессорное время (25,590 [с]). Суммарное процессорное время в данном случае рассчитывается по длительности максимального из выполняемых потоков, в связи с тем, что они функционируют параллельно [19]. При такой организации вычислений фазовых переменных, сокращение процессорного времени составило 1,8% по отношению к одноядерной РИ датаграмме. Повышение эффективности при этом достигается путем построения оптимального расписания [13, 14] и директивного его исполнения разработанной системой управления программными средствами. Для более сложных задач, в которых имеется значительное количество разделяемых физических переменных полезность предложенного подхода будет возрастать.

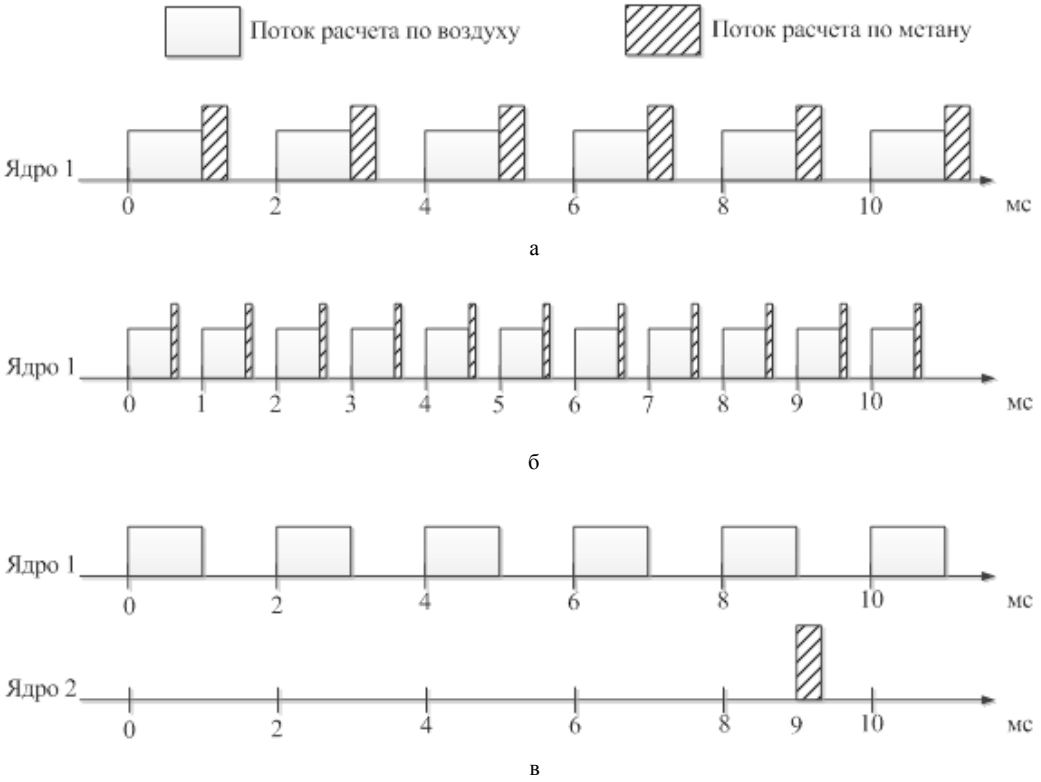


Рисунок 3 – Диаграммы выполнения экспериментов:
а – расчет с единой частотой на одном ядре; б – расчет с разными собственными частотами на одном ядре; в – расчет с разными собственными частотами на двух ядрах

6 ОБСУЖДЕНИЕ

Для верификации поведения воздушных и газовых процессов в разработанной модели, было выполнено сравнение с известными показателями в установившемся режиме. Также было выполнено сравнение графиков с результатами моделирования разработанной в среде MATLAB программой. В ней был выбран аналогичный метод решения с применением функции ode45 (метод Рунге-Кутты 4–5-го порядков). Графики изменения значений расхода воздуха (рис. 4) как в системе MATLAB, так и полученные разработанной моделью для трех экспериментов совпадают.

ВЫВОДЫ

Разработана программная аэрогазодинамическая РИ модель ШВС, которая представляется в виде графа. В объекте исследованы процессы по воздуху, протекающие во всех ветвях, а также изменение концентрации метана для выработок. Научные результаты, полученные в статье: основываясь на РИ фазовых переменных ШВС впервые предложена и исследована HIL модель на базе усовершенствованного авторами циклического расписания для решения уравнений модели. Отличительной чертой расписания является равномерное выполнение потоков вычисление фазовых переменных при постоянном базовом периоде РВ-цикла, получены условия его существования.

Практическая значимость полученных результатов заключается в том, что разработано программное обеспечение, реализующее предложенный метод планирования циклических расписаний, на основе которого решена задача построения многопоточной модели ШВС для решения задач прогнозирования в угледобывающей отрасли. В разработанной модели время расчета системы ОДУ по метану уменьшилось в 4,5 раза.

Дальнейшее повышение эффективности разработанных методов планирования и расширение класса создаваемых на их основе моделей может быть выполнено за счет учета информационной связности переменных в кэше процессоров SMP системы.

БЛАГОДАРНОСТИ

Работа выполнена в рамках госбюджетной научно-исследовательской темы кафедры «Компьютерная инженерия» Н-26-10 «Теоретический анализ и исследование процессов управления и обработки данных в компьютерных системах» Донецкого национального технического университета г. Покровск.

СПИСОК ЛИТЕРАТУРЫ

1. Bélanger J. The What, Where and Why of Real-Time Simulation / J. Bélanger, P. Venne, J. N. Paquin // PES-GM-Tutorial_04, IEEE PES general meeting. – 2010. – P. 37–49
2. Abel D. Rapid control prototyping / D. Abel, A. Bollig. – Berlin : Springer, 2006. – 400 p.
3. Demers S. A Generic Solution to Software-in-the-Loop / S. Demers, P. Gopalakrishnan, L. Kant // Military Communications Conference. – Orlando : MILCOM, 2007. – P. 1–6.
4. Hairer E. Solving ordinary differential equations I: Nonstiff problems 2nd rev. ed. / E. Hairer, S. P. Norsett, G. Wanner. – Heidelberg : Springer, 2009. – 474 p.
5. Pimentel J. C. G. Hardware Emulation for Real-Time Power System Simulation / J. C. G. Pimentel, L. Hoang // IEEE International Symposium on Industrial Electronics. – Montreal : IEEE, 2006. – № 2. – P. 1560–1565.
6. Mohammadi A. G. Scheduling Algorithms for Real-Time Systems / A. G. Mohammadi, S. G. Akl // School of Computing, Queen's University, Technical Report N 2005. – 2005. – 499 p.
7. Liu C. L. Scheduling algorithms for multiprogramming in a hard real-time environment / C. L. Liu, J. Layland // Journal of the ACM. – 1973. – № 20. – P. 46–61.
8. Forschungsgebiet: parallele Simulationstechnik / [Feldman L. P., Svjatnyj V. A., Resch M., Zeitz M.] // Наукові праці Донецького національного технічного університету, серія «Проблеми моделювання та автоматизації проектування». – 2008. – № 7(150). – С. 9–39.
9. Абрамов Ф. А. Моделирование динамических процессов рудничной аэрологии / Ф. А. Абрамов, Л. П. Фельдман, В. А. Святный. – К. : Наукова думка, 1981. – 284 с.
10. Tanenbaum A. Structured Computer Organization (6th Edition) / A. Tanenbaum, T. Austin. – Pearson, 2012. – 800 p.
11. Conway R. W. Theory of Scheduling (2nd ed.) / R. W. Conway, W. L. Maxwell, L. W. Miller – New York : DOVER PUBLICATIONS, 1967. – 294 p.

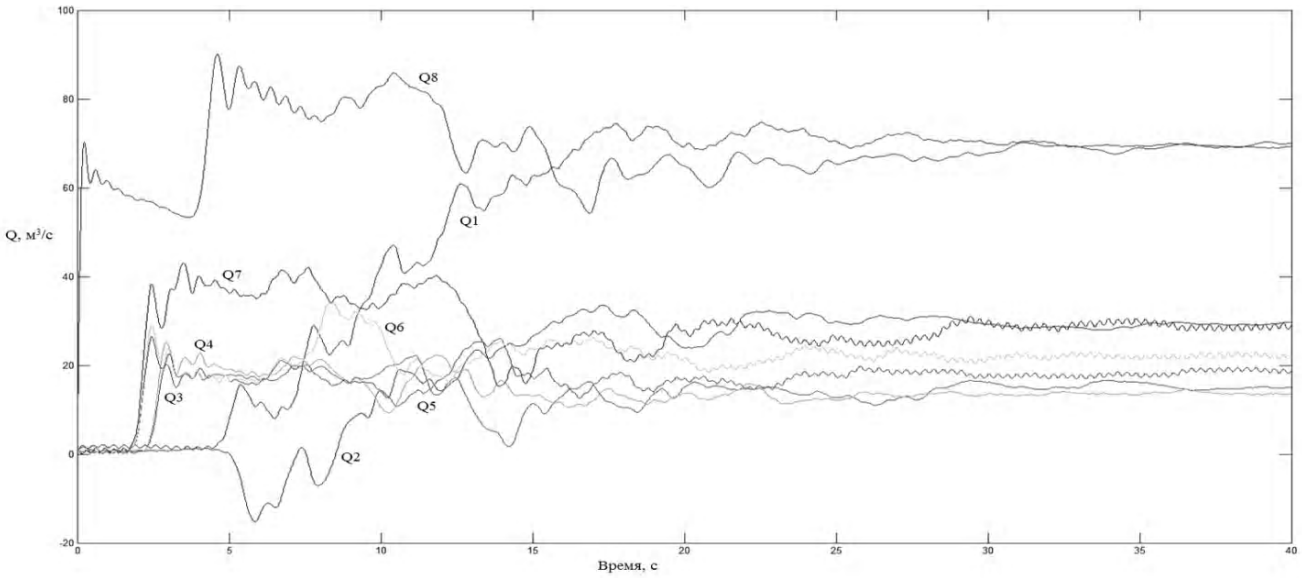


Рисунок 4 – Изменение расхода воздуха в ветвях ШВС при моделировании

12. Гильман А. Л. Расписания в задачах организации периодической обработки инфомации / А. Л. Гильман, Я. Г. Хаит // Техническая кибернетика. – Харьков : Известия академии наук СССР, 1970. – С. 125–130.
13. Иванов Ю. А. Алгоритм решения задачи оптимизации для циклических расписаний с прерываниями / Ю. А. Иванов // Радиоэлектроника, информатика, управление». – 2012. – № 2 (27). – С. 97–102. ISSN 1607-3274.
14. Иванов Ю. А. Оптимизация вычислительного процесса на многоядерной системе с использованием алгоритма искусственной пчелиной колонии / Ю. А. Иванов // Наукові праці ДонНТУ, Серія «Інформатика, кібернетика та обчислювальна техніка». – 2013. – № 2 (18). – С. 19–25. ISSN 1996-1588.
15. RTX 2012 Release Notes : [каталог]. – [IntervalZero, Inc., 2012]. – 20 p.
16. Kithara Real-Time Suite Version 9.23 : [каталог]. – [Kithara Software GmbH, 2011]. – 105 p.
17. Williams A. C++ Concurrency in Action: Practical Multithreading / A. Williams. – New York : Manning Publications, 2012. – 528 p.
18. Akhter S. Multi-core programming: increasing performance through software multi-threading / S. Akhter, J. Roberts. – Santa-Clara : Intel Press, 2015. – 360 p.
19. Wilkinson B. Parallel Programming Techniques & Applications Using Networked Workstations & Parallel Computers / B. Wilkinson, M. Allen. – Toronto : Pearson, 2004. – 496 p.

Статья поступила в редакцию 26.09.2017.
После доработки 10.11.2017.

Святний В. А.[1], Іванов О. Ю.[2], Іванов Ю. О.[3]
[1]Д-р техн. наук, професор, завідувач кафедрою комп'ютерної інженерії Донецького національного технічного університету, Покровськ, Україна
[2]Старший викладач кафедри комп'ютерної інженерії Донецького національного технічного університету, Покровськ, Україна
[3]Канд. техн. наук, доцент, інженер-програміст, Київ, Україна

ДОСЛІДЖЕННЯ ЦИКЛІЧНОГО РОЗКЛАДУ ДЛЯ МОДЕЛІ МЕРЕЖЕВОГО ДИНАМІЧНОГО ОБ'ЄКТА РЕАЛЬНОГО ЧАСУ

Актуальність. Розглянуте створення моделей автоматизованих систем управління провітрюванням призначених для запобігання виникненню аварійних ситуацій в шахтній вентиляційній мережі. Вирішується задача дослідження та реалізації програмних засобів моделювання, які дають можливість виконання експериментів з реальними наукоміcткими об'єктами.

Мета роботи – підвищення ефективності цифрової напівнатурної моделі шахтної вентиляційної мережі за рахунок розробки способу організації обчислювального процесу, що враховує різночастотні параметри моделей, та його оптимізація.

Метод. Виконано дослідження розробленого алгоритму планування та реалізація циклічного розкладу для моделі динамічного об'єкту реального часу представленого шахтною вентиляційною мережею. Аналіз математичної моделі аеродинамічних процесів дозволив віднести її до класу жорстких систем реального часу. Обчислення фазових змінних, що змінюються із різною частотою, виконуються шляхом розщеплення початкового алгоритму на паралельні потоки програмної моделі. Чисельне інтегрування системи диференційних рівнянь математичної моделі об'єкта виконано відповідно з алгоритмом статичних пріоритетів та керуванням за допомоги часової шкали теорії розкладів. Визначена допустимість розкладів для директивних строків змінних об'єкту. Алгоритм планування враховує параметри програмних засобів як однопроцесорних, так і симетричних мультипроцесорних систем.

Результати. Аналіз результатів моделювання запропонованої моделі з використанням циклічних розкладів показав відповідність отриманих результатів моделювання до еталонних.

Висновки. Основною перевагою моделі є підвищення продуктивності при виконані обробки даних за рахунок зменшення використання процесорного часу. Реалізований алгоритм циклічного розкладу напівнатурної моделюючої платформи використовується для наступної розробки систем управління технологічними та аварійними режимами провітрювання у шахтах.

Ключові слова: моделювання, циклічний розклад, багатоядерна система, розширення реального часу.

Svjatnij V. A.[1], Ivanov A. Y.[2], Ivanov Y. A.[3]
[1]Dr. Sc., Professor, Chief of the Department of Computer Engineering of the Donetsk National Technical University, Ukraine
[2]Senior lecturer of the Donetsk National Technical University, Ukraine
[3]PhD, software engineer, Kiev, Ukraine

CYCLIC SCHEDULE RESEARCH FOR THE MODEL OF REAL-TIME DYNAMIC NETWORK OBJECT

Context. Development of models of automated control systems for ventilation, designed to prevent emergencies in the mine ventilation network, is considered in the paper. The task of research and implementation of software modeling tools that allows the execution of experiments with real science-intensive objects is being solved.

Objective is to improve effectiveness of the real-time model of a mine ventilation network by developing and optimizing a method for organizing a computational process that considers the different frequency parameters of the models.

Method. We proposed and investigated the algorithm of scheduling and implementation of the cyclic schedule for the model of the dynamic real-time object, presented by the mine ventilation network. We analyzed the mathematical model of aerodynamic processes and classified it as a stiff real-time system. Calculations of phase variables with varying velocities are performed by splitting the original algorithm into parallel flows in the program model. Numerical integration of the differential equations system of the mathematical model of the object is performed in accordance with the algorithm of static priorities and controlled using the time scale. The schedules feasibility is determined for the target object limited parameters. The scheduling algorithm considers the time parameters of software tools for both uniprocessor and symmetric multiprocessor systems.

Results. Analysis of simulation results of the proposed model with using cyclic schedules shows the correspondence of the obtained modeling results to the reference one.

Conclusions. The main advantage of the model is the performance increase when performing data processing due to the reduction of the CPU usage time. Implemented cyclic schedule algorithm for real-time simulation platform can be used for development of control system for technological and emergency modes of ventilation in mines.

Keywords: simulation, cyclic schedule, multi-core system, real-time extension.

REFERENCES

1. Bélanger J., Venne P., Paquin J. N. The What, Where and Why of Real-Time Simulation, *PES-GM-Tutorial_04, IEEE PES general meeting,* 2010, pp. 37–49
2. Abel D., Bollig A. Rapid control prototyping. Berlin, Springer, 2006, 400 p.
3. Demers S., Gopalakrishnan P., Kant L. A Generic Solution to Software-in-the-Loop, *Military Communications Conference*. Orlando, MILCOM, 2007, pp. 1–6.
4. Hairer E., Norsett S. P., Wanner G. Solving ordinary differential equations I: Nonstiff problems 2nd rev. ed. Heidelberg, Springer, 2009, 474 p.
5. Pimentel J. C. G., Hoang L. Hardware Emulation for Real-Time Power System Simulation, *IEEE International Symposium on Industrial Electronics*. Montreal, IEEE, 2006, No. 2, pp. 1560–1565.
6. Mohammadi A. G., Akl S. G. Scheduling Algorithms for Real-Time Systems, *School of Computing, Queen's University, Technical Report N 2005*, 2005, 499 p.
7. Liu C. L., Layland J. Scheduling algorithms for multiprogramming in a hard real-time environment, *Journal of the ACM*, 1973, No. 20, pp. 46–61.
8. Feldman L. P., Svjatnyj V. A., Resch M., Zeitz M.Forschungsgebiet: parallele Simulationstechnik, *Naukovi pratsi Donetskogo natsionalnogo technicheskogo universitetu, seriya «Problemy modeluvannya ta avtomatizacii proektuvannya»,* 2008, No. 7(150), pp. 9–39.
9. Abramov F. A., Feldman L. P., Svyatnyy V. A. Modelirovanie dinamicheskih procesov rudnichnoy aerologii. Kiev, Naukova dumka, 1981, 284 p.
10. Tanenbaum A., Austin T. Structured Computer Organization (6th Edition). Pearson, 2012, 800 p.
11. Conway R. W., Maxwell W. L., Miller L. W. Theory of Scheduling (2nd ed.). New York, DOVER PUBLICATIONS, 1967, 294 p.
12. Gilman A. L., Hayt Y. G. Raspisaniya v zadachah organizacii periodicheskoy obrabotki informacii, *Tehnicheskaya kibernetika*. Kharkov, Izvestiya akademii nauk SSSR, 1970, pp. 125–130.
13. Ivanov Y. A. Algoritm resheniya zadachi optimizacii dlya ciklicheskih raspisaniy s preryvaniyami, *Radio Electronics, Computer Science, Control*, 2012, No. 2(27), pp. 97–102.
14. Ivanov Y. A. Optimizaciya vichislitelnogo processa na mnogoyadernoy sisteme s ispolzovaniem algoritma iskusstvenoy pchelinoy kolonii, *Invormatyka, kibenytyka ta obchysluvalna tehnika,* 2013, No. 2 (18), pp. 19–25. ISSN 1996-1588.
15. RTX 2012 Release Notes : [catalog]. [IntervalZero, Inc., 2012], 20 p.
16. Kithara Real-Time Suite Version 9.23 : [catalog]. [Kithara Software GmbH, 2011], 105 p.
17. Williams A. C++ Concurrency in Action: Practical Multithreading. New York, Manning Publications, 2012, 528 p.
18. Akhter S., Roberts J. Multi-core programming: increasing performance through software multi-threading. Santa-Clara, Intel Press, 2015, 360 p.
19. Wilkinson B., Allen M. Parallel Programming Techniques & Applications Using Networked Workstations & Parallel Computers. Toronto, Pearson, 2004, 496 p.

UDC 629.783.05

Firsov S. N.[1], Pishchukhina O. A.[2]

[1]Ph.D., Dr. Sc., Professor, Department of Electrical Engineering, Acting Dean of the Faculty of Rocket and Space Technology, National Aerospace University, Kharkiv, Ukraine
[2]PhD., Associate Professor, Department of Informatics, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

INTELLIGENT SUPPORT OF MULTILEVEL FUNCTIONAL STABILITY OF CONTROL AND NAVIGATION SYSTEMS

Relevance. Ensuring functionally steady operation of dynamic objects' movement in the presence of abnormal situations is one of the most prospective research problems in the area of technical diagnostics and restoration of system operability. Use of intelligent control methods and algorithms for solving this scientific and technical task can significantly extend the functionality and improve the performance of control systems.

The aim is the formation of an approach to determine the control effect that ensures the functional stability of the CNS in the presence of abnormal situations.

Method. The concept of ensuring functionally stable control of dynamic objects' movement has been offered. Well-known diagnostic methods and tools have been systematized and on that basis new models and methods for deep diagnosis of the functional state of the control and navigation systems up to the reason of abnormality have been developed. Models and methods for multi-level parrying of the reason of abnormal situations through control over the diagnosis have been synthesized, they use such redundant resources as signal and parametric adjustment, reconfiguration of algorithms and commutation of equipment.

Results. The results of the solution of a number of combined scientific and technical problems aimed at the multilevel ensuring of the functional stability of CNS have been presented.

Conclusion. Analyses of tendencies of theoretical investigation and practical achievements in ensuring functionally steady control of CNS at occurrence of abnormal situations has been carried out. The concept of intelligent support of the functionally steady control of CNS has been developed. The concept is based on the principles of multilevel hierarchical diagnosis of CNS to the parried reason of abnormal situation, as well as on the situational approach to eliminate the consequences of failures in accordance with the level of its application. As a result of analysis and synthesis of known approaches to diagnosing the technical state of dynamic objects, the new models and methods for deep diagnosis of the CNS' functional state have been formed, that makes possible to determine the failure up to the reason of abnormality. Models and methods for multi-level parrying of abnormality through control over the diagnosis with the use of available intelligent on-board resources have been offered. A hardware-software complex for experimental research of theoretical positions have been developed, and conducted experiments have been verified the possibility of extending the functionality of the system to counter abnormal situations. As a result of conducted experiments in real-time scale, diagnosis and restoration of CNS' functional state during a time not exceeding 42.2%–48.4% of the transient time of the system in normal operation mode have been implemented.

Keywords: control, stabilization, orientation, functionally steady movement control, abnormal situation, diagnosis, parrying the reason of abnormality.

NOMENCLATURE

DLM is a diagnostic logical model;
TLM is a tabular logical model;
CNS is a control and navigation systems;
FDM is a functional diagnostic model;
SFM is a "small" failures model;
HSC is a hardware and software complex;
PM is a power amplifier;
AM is a actuating mechanism;

$\tilde{y}(t)$ is a vector of the measured variables of the perturbed movement, $\tilde{y}(t) \in Y^n$;

$A(\lambda)$, $B(\lambda)$ is a matrixes of the appropriate dimensionalities characterizing a linearized part of the system;

λ is a μ-dimensional vector of the diagnostic parameters of the reasons of abnormal situations;

$F[\lambda, u(t)]$ is a r-dimensional vector function describing essential static nonlinearities of the system;

$u(t)$ is a r-dimensional vector of input signals, $u(t) \in U^r$;

$\Delta y(t)$ is a vector of deviations of the measured variables of the perturbed movement from the nominal ones $\Delta y(t) \in Y^n$;

R, T is a matrixes of the appropriate dimensionalities;

$\Delta u(t)$ is a input signals; $\Delta u(t) \in U^r$;

$\psi(t)$ is a m-dimensional vector of the controlling influences, providing execution of $\Delta y(t) \underset{t \to t_k}{\to} 0$ for the finite number of iterations t_k;

$\hat{y}(t)$ is a output of a reference model;

$\Delta\lambda_i$ is a direct diagnostic sign $\Delta\lambda_i = \lambda_i - \lambda_{iH}$;

λ_{iH} is a nominal value of a parameter of the abnormal situation;

G is a Luenberger matrix ;

$\frac{\partial[\cdot]}{\partial\lambda_i}$ is a partial derivative;

$\frac{D[\cdot]}{\partial\lambda_i}$ is a generalized differentiation;

$S_2(.)$ is a two-digit predicate;

© Firsov S. N., Pishchukhina O. A., 2018
DOI 10.15588/1607-3274-2018-2-20

δ_i is a threshold value depending on the accuracy of data representation and the subsequent data processing;

$V_\lambda(t)$ is a function of sensitivity on the diagnostic parameter λ_i;

σ_j is a threshold value of a ratio under the sign of a predicate $S_2[\cdot]$;

t_O is a time point of the abnormal situation affecting the system;

t_{det} is a time point of its detection;

t_{cd} is a time point of the completion of diagnosing and determination of parameters of the abnormal situation;

t_{pc} is a time point of the beginning of parrying by signal setup;

t_{pr} is a time point of the beginning of parrying by equipment reconfiguration;

t_{cp} is a time point of completion of the procedure of parrying of an abnormal situation.

INTRODUCTION

Exploration of natural phenomena like storms, tornados, earthquakes, volcanic eruptions, including effects they cause, are inseparably linked with the practice of application of autonomous dynamic objects. In most cases, the distinguished characteristic of such objects functioning is their remoteness and, very frequently, impossibility of operative elimination of abnormalities in their operation. Those circumstances define the appearance of new specific requirements to operability of both autonomous systems in general and elements ensuring their operation through the whole life cycle at the design , development and operation stages.

Ensuring functionally steady operation of dynamic objects in the presence of abnormal situations is one of the most prospective research problems in the area of technical diagnostics and restoration of system operability. Use of intelligent control methods and algorithms for solving this scientific and technical task can significantly extend the functionality and improve the performance of control systems, which defines the topical object and subject of conducted research. The object of research is the processes of the functioning of control and navigation systems in the presence of abnormal situations. The subject of research is models and methods for ensuring the functional stability of control and navigation systems. The aim of research is the formation of an approach to determine the control effect that ensures the functional stability of the CNS in the presence of abnormal situations.

1 PROBLEM STATEMENT

The emergency mode of functioning of a linearized object with essential nonlinearities in a case of the impact of disturbances is described in the Cauchy form by the vector equation of the perturbed movement:

$$\dot{\tilde{y}}(t) = A(\lambda)\tilde{y}(t) + B(\lambda)F[\lambda, u(t)],\ \tilde{y}(t_0) = \tilde{y}_0. \quad (1)$$

For the description of a deviation of CNS perturbed movement in the abnormal situation (1) from the reference one the following equation is used:

$$\Delta\dot{y}(t) = R\Delta u(t) + T\psi(t);\ \Delta y(t_0) = \tilde{y}_0; \Delta y(t) \in Y^n. \quad (2)$$

The aim is to define the control action, providing the functional stability of CNS while abnormal situations.

2 REVIEW OF THE LITERATURE

Failures and malfunction of elements and units of control systems, as a result of abnormal situations appearance, are the one of the main problem while fulfilling the missions [1–3]. According to the results of statistical studies, the most vulnerable systems, which are subjects to technical breakages, have been control and navigation systems (CNS), and also its main measuring and actuating subsystems.

Necessity of the solution of the given problem has contributed into increasing the functionality of dynamic objects in terms of their application and efficiency [4–5]. It has fostered the progress in extension of the functional capabilities of CNS for parrying the abnormal situations of different kind, which is implemented by ensuring the functional stability of the most vulnerable subsystems to different breakages, failures and other abnormalities of CNS and their elements. The functional stability is ability of the system, device or units to perform the required functions during the preset time at the presence of abnormal situations [6–7]. It allows overcoming the objective contradiction consisting in need of fault-free mission fulfilling, on the one hand, and existence of limited resources for maintenance of their operability in real time, on the other hand. Using of standard methods and technologies to support operability and stability of similar systems doesn't cover all set of the reasons of abnormal situations origins, and the unreasonable adjusting of redundancies leads to increase in their weight, overall dimensions, energy consumption and other performance indices growth. For effective mission implementation by the autonomous dynamic object, methods of restoration of its functional properties have been offered. The aim of those methods implementation is to ensure the functionally steady operating of CNS throughout the required period in the conditions of counteraction to different destabilizing factors leading to the partial and/or complete loss of operability.

The theoretical research started in the early seventies in the field of the functional stability of CNS now continues for the purpose of creation of perfect design tools for complex technical objects development. At the same time, CNS functional stability ensuring is rather new direction which is at the stage of its formation. O. Mashkov in paper [8] describes the basic reasons of this direction emergence and obvious advantages of functionally stable systems in comparison with unstable ones. A few scientists have made a contribution to the development of the functional stability theory using some kind of intelligent approaches, for example, Ju. Kravchenko offered an actual direction connected with application of hardware redundancy to restore technical objects [9]; P. Castaldi, N. Mimm and S. Simani considered fault-tolerant approach on the base of software and hardware redundancy with regard to non-linearity of control systems [10]; J. Huang improved the common approach having implemented control reconfigurability through fuzzy logic techniques for sensor placement. According to those groups of researchers and their published papers ensuring the functional stability for CNS allows to reduce the time of system debugging, its

energy consumption and mass-dimensional characteristics and leads to extension of system life cycle. The major drawback of those scientific ideas and technical solutions is absence of a complex approach to intelligent technical diagnostics before fault's elimination and further flexible system restoration with regard to available onboard resources.

One of perspective ways to support the fault-tolerant operation of CNS is its intelligence consisting in ensuring the system with ability of its self-diagnosis [11-12], and then having got the diagnosis results to recover its operability. The analysis of tendencies of small-size autonomous dynamic objects development shows, that the quality assessment is carried out not only by a criteria of functionality and payload mass, but also on indices of the onboard "intelligence" power. Until recently, CNS of the previous generations were getting improved mostly at the expense of the hardware component, that in the case of the modern tactical-economic requirements and appearance of new, more complex challenges is not sufficient.

Research of some domestic and foreign researchers are devoted to issues of CNS intelligence under conditions of uncertainty. The intelligent automatic control systems steady against failures, designed after appearance of new direction of the theory of automatic control [12–13], have become a result of such development. Results of the research, devoted to the development of CNS tolerant to failures, are provided in the articles [6–8]. In scientific works [4–6] the systems, capable to parry failures in aircraft control systems, and in [8–10] systems, compensating failures in the executive mechanisms, are described. The author of [12–13] selects the following necessary and sufficient conditions of ensuring functional stability of CNS: (1) presence of robustness property (no sensitivity while changing of parameters or structures of system); (2) reconfigurability (changing of the CNS parameters and structure for the purpose of compensations of uncontrollable occurred anomalies).

The analysis of tendencies in the field of theoretical researches in ensuring functional stability of CNS promotes identification of the new scientific direction of complex studying of a problem of functional stability and development of approaches to its intelligence on the basis of two independent directions of a research connected with diagnosing and parrying of the reasons of abnormal situations.

3 MATERIALS AND METHODS

Functional resistance of CNS to abnormal situations is carried out by fulfillment of the principle of three-level hierarchy. The first level of hierarchy – subsystem. At this level are devices, which are carrying out target functional transformations for CNS, are developed and designed.

To suppress the influence of abnormal situations for an autonomous dynamic object and subsystem communications, the second level of the self-organization hierarchy is formed – system level. At this level the diagnosis of a functional condition of the whole CNS is made, which means detecting the abnormal situation itself, defining a place of the failure and its class, i.e. providing the full diagnosis of the abnormal situation which has caused deterioration in quality of functioning of CNS. Having obtained the diagnosis, the procedure of parrying of the abnormal situation by redundant means of CNS starts.

The third level of self-organization – mega system level. At this level, diagnosing of all the mission or project is carried out in general, if the abnormal situation has not been managed as diagnosed one at the lower levels: subsystem and system. With that, the abnormal situation is localized and classified under available information, as well as its reason and type is defined. On the basis of this diagnosis and the available onboard resources of an autonomous dynamic object, as well as taking into account political, military, economic and strategic factors, the decision about abnormal situation parrying is made on the mega-system.

While using methods of redundancy control, synthesis of algorithms has to be made with use of the diagnostic models applied at the first stage. Thus, the offered concept to support the functional stability provides adaptation, structural and semantic unity of models and, respectively, the algorithms used at all stages of control.

Among a number of approaches to diagnosing of technical systems in the context of the considered concept, the most perspective one is a signal-parametric approach, as it is based on the assumption, that origin of any reason of the abnormal situation $d_i \in D_j$ represents an uncertain event. Within this approach, special functional diagnostic models (FDM) formalize description of the system, which is necessary for the solution of diagnostic tasks for connecting the direct and not measured signs with indirect and measured ones. In general form, FDM connects a deviation from reference value of object's signal characteristics ($\Delta x(t)$, $\Delta y(t)$) to deviations from rated values of the appropriate target diagnostic signs $\Delta\lambda_i$.

For the non-linear system, it is possible to develop several types of FDM. We will provide, for example, a diagnostic model for "small" failures ($\Delta\lambda_i >> \Delta\lambda_i^2$), so-called SFM-model:

$$\Delta\dot{y}(t) = G\Delta y(t) + \frac{\partial A(\lambda)}{\partial\lambda_i}\hat{y}(t)\Delta\lambda_i + \frac{\partial B(\lambda)}{\partial\lambda_i}F\left[\lambda_н, u(t)\right]\Delta\lambda_i,$$

$$\Delta y(t_0) = \tilde{y}. \qquad (3)$$

Connection between direct and indirect diagnostic signs is non-linear (3), and, therefore, for conversion to qualitative models we use different conversions of a sign $\Delta y(t)$ to a numerical sign, and further from an absolute scale of signs to a scale of denomination we apply the two-valued predicate equations:

$$Z_j = S_2\left(\eta_{ij} - \delta_i\right) = \begin{cases} 1, & \eta_{ij} \geq \delta_i, \\ 0, & \eta_{ij} < \delta_i, \ j = \overline{1,m}. \end{cases} \qquad (4)$$

Consistently applying (4), we obtain rectangular Boolean tables of $\mu \times m$ size. The number of lines in the table corresponds to number of $\Delta\lambda_i$, and the number of columns – $\Delta y(k)$. This table represents a variant of diagnostic logical model (DLM) – tabular logical model (TLM).

Initial data for development of models and tools of CNS diagnosing are results of a research of a subject to diagnosing, namely: hierarchy of diagnostic models, requirements to diagnosing depth, diagnosing time, reliability of the diagnosis, set of initial signals and requirement to the mode of input signals. The result of the solution of the major tasks is the machine realized algorithms and all process of diagnosing of CNS in the form of the tree-like knowledge base. Joints of a tree contains the corresponding signs, representing the two-digit predicate equations on a set of data array of discrete values of system signals, which are available to measurement:

$$\mu_j = S_2\left[\Delta y(t), V_\lambda(t), \sigma_j\right]. \quad (5)$$

The equations (5) requires solution on the corresponding data array of discrete values with use of a confidence coefficient, reducing the influence of the destabilizing factors on decision quality. As a result of the solution (5), we receive value 0 or 1, determining the direction of further moving on a tree. The tree finishes with leaves, which correspond to possible conditions and the set of abnormal situations of CNS.

Application of FDM has allowed to reduce a number of the control points, representing additional measurable outputs of the system.

Parrying of the reasons of abnormal situations is feasible by means of the appropriate redundant resources at the appropriate level of ensuring the functional stability of CNS. The more excess resources are available, the quality of parrying is higher. In practice a set of redundant resources – R is limited. The obvious contradiction between the diagnosing depth (which indirect characteristic is the power of a set D) and flexibility of restoration (indirect quantitative assessment of which is the power of a set R) can be resolved by finding a compromise.

For the purpose of specificity of the task of multi-level parrying, we will consider restoration of the functional status of CNS. For the solution of this task, CNS should contain redundant resources in its structure, such as: a) supplemental signal entries; b) means of parametric adjustment; c) redundant instruments, units, devices; d) additional algorithms and programs; e) other resources.

For the description of specific of control mentioning described redundancy, we use equation of CNS's movement, perturbed of an abnormal situation (2).

The control input is defined by approaches to ensure stability of dynamic systems. The most effective is the discrete analog of the second Lyapunov's method.

The second Lyapunov's method is a formation of the special extra scalar function named Lyapunov's function $V\left[\Delta y(k)\right]$, and further investigation of its properties, and also properties of its first difference, defined on necessary trajectory of equation.

The linear system described by the equation (2) is asymptotically steady, if $|\sigma_i| < 1$, $i = \overline{1,m}$. Also, if in the area Ω for system (2) there is the definitely positive function $V\left[\Delta y(k)\right]$, which first difference $\Delta V\left[k, k+1\right]$ is definitely negative, then equilibrium position is asymptotically steady, i.e.

$$\lim_{k\to\infty} \Delta y(k) = 0.$$

We select the compensating influence $\psi(k)$, coming from a following inequality $\Delta V\left[k, k+1\right] < 0$ for a finite difference of $\Delta V\left[k, k+1\right]$ in the area Ω, i.e. from execution of a condition of definite negativity of the function:

$$\psi^T(k) = -2\Delta y^T(k) R^T Q T\left[T^T Q T\right]^{-1}. \quad (6)$$

Free choice of Lyapunov's functions structure allows to obtain different options of restoration of the functional state, different in structure and quality indices of dynamics reflecting compensating of consequences of abnormal situation impact.

Self-organization of CNS is based both on obtaining the complete diagnosis about the functional state after an abnormal situation and on flexible use of the redundant resources allowing to parry the revealed abnormal situation. Thus, high-quality parrying of abnormal situations requires: full diagnosis of functional state of an object of diagnosing, reflecting time of appearance of an abnormal situation, its place, class and specific physical reason; availability of redundant resources, covering a set of compensated abnormal situations; concept of a choice of redundant resources, considering previous abnormal situations proceeding from the current situation and tasks of a mission; methods of effective parrying of an abnormal situation of an object of diagnosing; resources for determination of the technical condition of a restored object of diagnosing.

4 EXPERIMENTS

For experimental verification of the offered concept of ensuring the functional stability the appropriate pilot studies on the developed hardware and software complex (HSC) have been conducted. In particular, computer simulation of abnormal situations of functionally steady CNS, being in the modes of stabilizing and orientation of a platform, have been carried out and their influence on a movement deviation from rated one have been investigated using HSC.

On figure 1 the results of experimental research are presented: influence of power amplifier (PA) and actuating mechanism (AM) transmission ratio reduction by 30% in the mode of stabilizing of angular position of a platform. The dotted line represents graphs of CNS output signals changing in the system without functional stability, and a solid line – in the case of the permanently connected subsystem to ensure the functional stability of CNS.

5 RESULTS

Investigated failure occurred in AM and time point of its detection matches time point of parrying through reconfiguration of equipment ($t_{det} = t_{pr}$). After determination of failure class and failure mode, parrying of a failure through signal setup is carried out ($t_{ps} = t_{cd}$). The total time t_{cp1}, spent for failure detection, determination of failure and its parrying is:

$$t_{cp_1} = (t_{\det} - t_o) + (t_{cd} - t_{\det}) + (t_{cp} - t_{ps}) = t_{cp} - t_o = 1.42\ c.$$

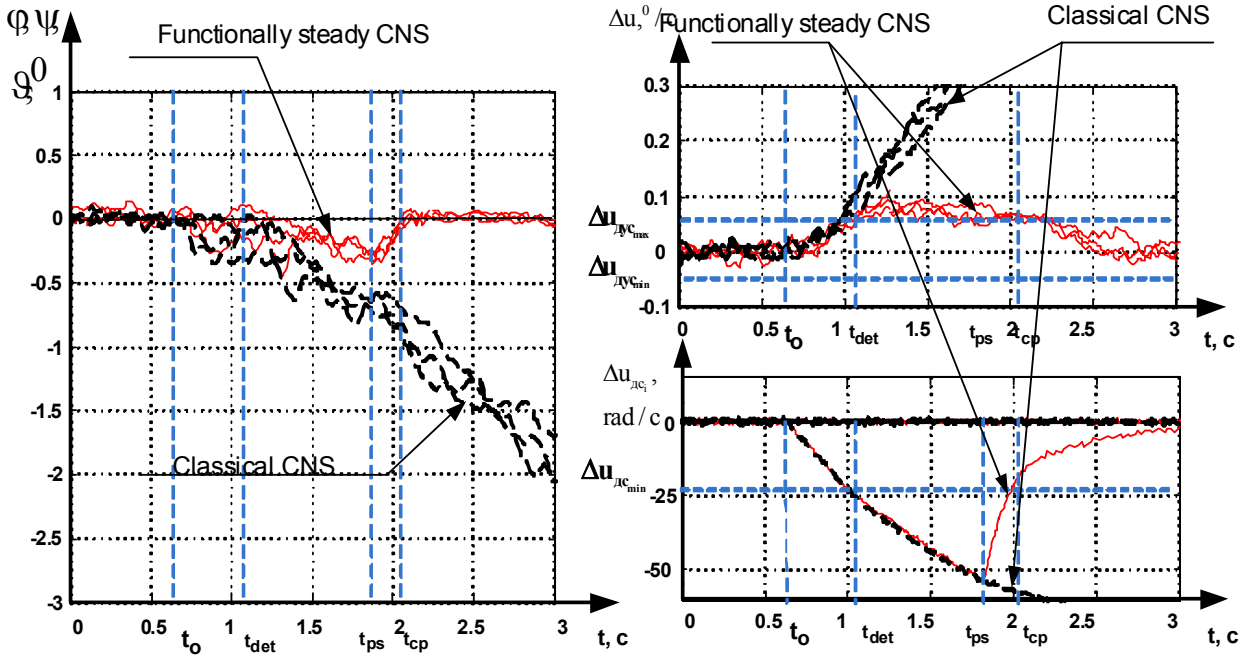


Figure 1 – Results of a research of functionally steady CNS

As can be seen from the presented graphs, the system without a property of functional stability deviates from the predetermined trajectory. In system with functional stability, the system parries failure and moves on the required trajectory after the moment the failure has been detected.

6 DISCUSSION

The world practice of CNS exploitation shows their low-efficiency at various abnormal situations arising in the process of application. A significant proportion of accidents happens due to failures in CNS itself. Traditional engineering approach to parrying of abnormal situations doesn't allow to meet a number of modern and perspective requirements. Theoretical research in the field of functional stability develops in the direction of the fragmentary solution of separate aspects of parrying of abnormal situations. For elimination of the specified contradictions, the complex of the combined scientific and technical tasks aimed at multi-level ensuring CNS with functional stability has been solved.

To compare with analogues [9–11] offered by the groups of researchers, the applied approach of situation-dependent parrying is based on the principles of adaptation to CNS specifics and possible abnormal situations and tasks of a mission. Within current approach, parrying of abnormal situations is made after detailed, deep and high-quality diagnosing not only on the base of object's technical condition, but including its functional state.

Thus, the new concept of ensuring functional stability of CNS has been formed. It is based on the principles and results of the modern control theory, applied to the solution of functional stability problem, and on the methods of signal-parametric diagnosing of systems in contrast to common fault-tolerant approaches. With that, for the implementation of the concept application, more advanced models, allowing to design an independent dynamic object with CNS both in nominal and in abnormal operation have been offered. Methods to recover abnormal situations include methods of intelligent analysis on the base tree-like knowledge base and further flexible reconfiguration of the available onboard resources, which was outside the scope of recent results presented by researchers.

CONCLUSIONS

Analyses of tendencies of theoretical investigation and practical achievements in ensuring functionally steady control of CNS at occurrence of abnormal situations has been carried out. The concept of intelligent support of the functionally steady control of CNS has been developed. The concept is based on the principles of multilevel hierarchical diagnosis of CNS to the parried reason of abnormal situation, as well as on the situational approach to eliminate the consequences of failures in accordance with the level of its application. As a result of analysis and synthesis of known approaches to diagnosing the technical state of dynamic objects, the new models and methods for deep diagnosis of the CNS' functional state have been formed, that makes possible to determine the failure up to the reason of abnormality. Models and methods for multi-level parrying of abnormality through control over the diagnosis with the use of available intelligent on-board resources have been offered. A hardware-software complex for experimental research of theoretical positions have been developed, and conducted experiments have been verified the possibility of extending the functionality of the system to counter abnormal situations. As a result of conducted experiments in real-time scale, diagnosis and restoration of CNS' functional state during a time not exceeding 42.2%–48.4% of the transient time of the system in normal operation mode have been implemented.

ACKNOWLEDGEMENTS

Research is carried out in accordance with the plan of research works within the framework of "Concept of State Policy Implementation in the Sphere of Space Activities on

the period up to 2031" (approved by the Cabinet of Ministers of Ukraine on 30.03.2011, No. 238-p) and also according to the Agreement on Cooperation with the Aerospace Company «MAVA TECHNOLOGIES», Mexico.

REFERENCES

1. Gostev V. I. Intermodular exchange of diagnosis information in Selfdiagnosis of complex systems [Text] / V. I. Gostev, O. A. Mashkov, V. A. Mashkov // Cybernetics and Computing Technology, Discrete Control Systems. – 1999. – Vol. 105, № 2. – P. 95–103.
2. Gostev V.I. Methods for providing fault tolerance of airborne computer system through their self-control [Text] / V. I. Gostev, O. A. Mashkov, V. A. Mashkov // Cybernetics and Computing Technology, Discrete Control Systems. – 1999. – Vol. 109, № 4 – P. 33–44.
3. Mashkov O. Fault-Tolerant of Computing Systems based on the Self-diagnosis by the traveling kerning Principle [Text] / O. Mashkov, V. Mashkov // Cybernetics and Computing Technology, Discrete Control Systems. – 1999. – Vol. 106, №3. – P. 89–94.
4. Gostev V. I. Self-diagnostic of modular systems in random performance of elementary tests [Text] / V. I. Gostev, O. A. Mashkov, V. A. Mashkov // Cybernetics and Computing Technology, Discrete Control Systems. – 1995. – Vol. 99, № 3 – P. 104–112.
5. Firsov S. N. Formation of Fault-Tolerant Flywheel Engine Units in Satellite Stabilization and Attitude Control Systems [Text] / S. N. Firsov // Journal of Computer and Systems Sciences International. – 2014. – Vol. 53, №4. – P. 601 – 609.
6. Firsov S. Hardware and Software Package for Search, Detection and First Aid Means Delivery in Rough Terrain on Basis of a Three Rotor Unmanned Aerial Vehicle [Text] / S. Firsov, N. Plavynska, K. Rudenko // Transport Problems. – 2014. – Vol. 9, Issue 2. – P. 69–75.
7. Машков О. А. Оцінка ефективності застосування функціонально стійких систем дистанційно пілотованих літальних апаратів при моніторингу навколишнього природного середовища в умовах турбулентної атмосфери [Текст] / О. А. Машков, О. М. Щукін // Моделювання та інформаційні технології. – 2013. – Вип. 69, №2. – С. 152–158.
8. Кравченко Ю. В. Сучасний стан та шляхи розвитку теорії функціональної стійкості [Текст] / Ю. В. Кравченко, С. А. Мюкіс // Моделювання та інформаційні технології. – 2013. – Вип. 68, № 1. – С. 60–68.
9. Castaldi, P. Fault Tolerant Control Schemes for Nonlinear Models of Aircraft and Spacecraft Systems [Text] / P. Castaldi, N. Mimm, S. Simani // 18th IFAC: World Congress, 28.08–02.09 2011. – P. 13705–13710.
10. Huang Jianzhuang. Fault-Tolerant Sensor Placement Based on Control Reconfigurability [Text] / Jianzhuang Huang, N. Eva Wu // 18th IFAC: World Congress, 28.08–02.09 2011 – P. 14814–14819.
11. Firsov Sergii Compact vertical take-off and landing aerial vehicle for monitoring tasks in dense urban areas [Text] / Sergii Firsov, Igor Kulik // Transport Problems. – 2015. – Vol. 10, Issue 8. – P. 29– 34.
12. Firsov Sergii. Assessment of functional sustainability of computer-integrated technological systems [Text] / Sergii Firsov, O. A. Pishchukhina // Open information and computer-integrated technologies. – 2016. – Vol. 72, №4. – P. 112–121.
13. Firsov S. N. Hardware-software complex for experimental tests of control processes and diagnostics of small spacecraft's faults [Text] / S. N. Firsov, O. V. Reznikova – Devices and systems. Management, monitoring, diagnostics. – 2014. – Vol. 6. – P. 60–69.

Article was submitted 03.09.2017.
After revision 12.11.2017.

Фірсов С. М.[1], Піщухіна О. О.[2]

[1]Д-р техн. наук, професор кафедри електротехніки, виконуючий обов'язки декану факультету ракетно-космічної техніки Національного аерокосмічного університету ім. М.Є. Жуковського «ХАІ», Харків, Україна

[2]Канд. техн. наук, доцент кафедри інформатики Харківського Національного університету радіоелектроніки, Харків, Україна

ІНТЕЛЕКТУАЛЬНА ПІДТРИМКА БАГАТОРІВНЕВОЇ ФУНКЦІОНАЛЬНОЇ СТІЙКОСТІ СИСТЕМ УПРАВЛІННЯ ТА НАВІГАЦІЇ

Актуальність. Забезпечення функціонально стійкого управління рухом динамічних об'єктів за наявності нештатних ситуацій є однією з найбільш перспективних дослідницьких задач в області технічного діагностування та відновлення працездатності систем. Використання методів і алгоритмів інтелектуального управління для вирішення даної науково-технічної проблеми дозволяє значно розширити функціональні можливості і поліпшити показники якості систем управління.

Мета роботи – формування підходу до визначення керуючого впливу, що забезпечує функціональну стійкість систем керування та навігації (СКН) в нештатних ситуаціях.

Метод. Запропоновано концепцію забезпечення функціонально стійкого управління рухом динамічних об'єктів. Систематизовано відомі інструментальні засоби діагностування, на підставі яких розроблені нові моделі і методи глибокого діагностування функціонального стану систем управління рухом і навігації з точністю до причини позаштатної ситуації. Синтезовано моделі і методи багаторівневого парирування причин нештатних ситуацій за допомогою управління за діагнозом такими надмірними засобами, як сигнальне і параметричне підстроювання, реконфігурація алгоритмів і комутація апаратури.

Результати. Наведено результати вирішення комплексу взаємопов'язаних науково-технічних задач, спрямованих на багаторівневе забезпечення функціональної стійкості систем керування та навігації.

Висновки. Проведено аналіз тенденцій теоретичних досліджень та практичних розробок щодо забезпечення функціонально стійкого управління систем керування та навігації за появи нештатних ситуацій. Розроблено концепцію інтелектуального забезпечення функціонально стійкого управління СКН. Концепція базується на принципах багаторівневого ієрархічного діагностування СКН до причини нештатної ситуації, а також ситуаційному підході усунення наслідків відмов відповідно до рівня його застосування. В результаті аналізу і синтезу відомих підходів до діагностування технічного стану динамічних об'єктів сформовано нові моделі і методи глибокого діагностування функціонального стану СКН, що дозволяють визначити відмову з точністю до причини позаштатної ситуації. Запропоновано моделі та методи, що забезпечують багаторівневе парирування нештатних ситуацій СКН в результаті управління за діагнозом з використанням наявних засобів інтелектуалізації бортового обладнання. Розроблено апаратно-програмний комплекс для експериментального дослідження теоретичних положень, що підтвердило можливість розширення функціональних можливостей системи для парирування нештатних ситуацій. В результаті проведення експериментів у реальному масштабі часу забезпечено діагностування та відновлення функціонального стану СКН за час, що не перевищує 42,2%-48,4% часу перехідного процесу системи в штатному режимі функціонування.

Ключові слова: керування, стабілізація, орієнтація, функціонально-стійке керування рухом, нештатна ситуація, діагностування, парирування причини нештатної ситуації.

Фирсов С. Н.[1], Пищухина О. А.[2]

[1]Д-р техн. наук, профессор кафедры электротехники, исполняющий обязанности декана факультета ракетно-космической техники Национального аэрокосмического университета им. Н.Е. Жуковского «ХАИ», Харьков, Украина

[2]Канд. техн. наук, доцент кафедры информатики Харьковского национального университета радиоэлектроники, Харьков, Украина

ИНТЕЛЛЕКТУАЛЬНАЯ ПОДДЕРЖКА МНОГОУРОВНЕВОЙ ФУНКЦИОНАЛЬНОЙ УСТОЙЧИВОСТИ СИСТЕМ УПРАВЛЕНИЯ И НАВИГАЦИИ

Актуальность. Обеспечение функционально устойчивого управления движением динамических объектов при наличии нештатных ситуаций представляет собой одну из наиболее перспективных исследовательских задач в области технического диагностирования и восстановления работоспособности систем. Использование методов и алгоритмов интеллектуального управления для решения данной научно-технической проблемы позволяет значительно расширить функциональные возможности и улучшить показатели качества систем управления.

Цель работы – формирование подхода к определению управляющего воздействия, обеспечивающего функциональную устойчивость систем управления и навигации (СУН) в нештатных ситуациях.

Метод. Предложена концепция обеспечения функционального устойчивого управления движением динамических объектов. Систематизированы известные инструментальные средства диагностирования, на основании которых разработаны новые модели и методы глубокого диагностирования функционального состояния систем управления движением и навигации с точностью до причины нештатной ситуации. Синтезированы модели и методы многоуровневого парирования причин нештатных ситуаций посредством управления по диагнозу такими избыточными средствами, как сигнальная и параметрическая подстройки, а также реконфигурация алгоритмов и коммутация аппаратуры.

Результаты. Приведены результаты решения комплекса взаимосвязанных научно-технических задач, направленных на многоуровневое обеспечение функциональной устойчивости СУН.

Выводы. Проведен анализ тенденций теоретических исследований и практических разработок по обеспечению функционально устойчивого управления систем управления и навигации при появлении нештатных ситуаций. Разработана концепция интеллектуального обеспечения функционально устойчивого управления СУН. Концепция базируется на принципах многоуровневого иерархического диагностирования СУН до парируемой причины нештатных ситуаций, а также ситуационном подходе устранения последствий отказов в соответствии с уровнем его применения. В результате анализа и синтеза известных подходов к диагностированию технического состояния динамических объектов сформированы новые модели и методы глубокого диагностирования функционального состояния СУН, позволяющие определить отказ с точностью до причины нештатной ситуации. Предложены модели и методы, обеспечивающие многоуровневое парирование нештатных ситуаций СУН в результате управления по диагнозу с использованием доступных средств интеллектуализации бортового оборудования. Разработан аппаратно-программный комплекс для экспериментального исследования теоретических положений, что подтвердило возможность расширения функциональных возможностей системы для парирования нештатных ситуаций. В результате проведения экспериментов в реальном масштабе времени обеспечено диагностирование и восстановление функционального состояния СУН за время, не превышающее 42,2%–48,4% времени переходного процесса системы в штатном режиме функционирования.

Ключевые слова: управление, стабилизация, ориентация, функционально устойчивое управление движением, нештатная ситуация, диагностирование, парирование причины нештатной ситуации.

REFERENCES

1. Gostev V. I., Mashkov O. A., Mashkov V. A. Intermodular exchange of diagnosis information in Selfdiagnosis of complex systems, *Cybernetics and Computing Technology, Discrete Control Systems*, 1999, Vol. 105, No. 2, pp. 95–103..
2. Gostev V. I., Mashkov O. A., Mashkov V. A. Methods for providing fault tolerance of airborne computer system through their self-control [Text], *Cybernetics and Computing Technology, Discrete Control Systems,*1999, Vol. 109, No. 4, pp. 33–44.
3. Mashkov O., Mashkov V. Fault-Tolerant of Computing Systems based on the Self-diagnosis by the traveling kerning Principle, *Cybernetics and Computing Technology, Discrete Control Systems*, 1999, Vol. 106, No. 3, pp. 89–94.
4. Gostev V. I., Mashkov O. A., Mashkov V. A. Self-diagnostic of modular systems in random performance of elementary tests, *Cybernetics and Computing Technology, Discrete Control Systems*, 1995, Vol. 99, No. 3, pp. 104–112.
5. Firsov, S. N. Formation of Fault-Tolerant Flywheel Engine Units in Satellite Stabilization and Attitude Control Systems, *Journal of Computer and Systems Sciences International*, 2014, Vol. 53, No. 4, pp. 601–609.
6. Firsov S., Plavynska N., Rudenko K. Hardware and Software Package for Search, Detection and First Aid Means Delivery in Rough Terrain on Basis of a Three Rotor Unmanned Aerial Vehicle, *Transport Problems*, 2014, Vol. 9, Issue 2, pp. 69–75.
7. Mashkov O. A., O. M. Shhukin Ocinka efektyvnosti zastosuvannja funkcional'no stijkyh system dystancijno pilotovanyh lital'nyh aparativ pry monitoryngu navkolyshn'ogo pryrodnogo seredovyshha v umovah turbulentnosti atmosfery, *Modeljuvannja ta informacijni tehnologii',* 2013, – Vol. 69, No. 2, pp. 152–158.
8. Kravchenko Ju.V., Mykus' S. A. Suchasnyj stan ta shljahy rozvytku teorii' funkcional'noi' stijkosti, *Modeljuvannja ta informacijni tehnologii',* 2013, Vol. 68, Issue 1, pp. 60–68.
9. Castaldi P., Mimm N., Simani S. Fault Tolerant Control Schemes for Nonlinear Models of Aircraft and Spacecraft Systems [Text], *18th IFAC: World Congress*, 28.08–02.09 2011, pp. 13705–13710.
10. Huang, Jianzhuang, Eva Wu N. Fault-Tolerant Sensor Placement Based on Control Reconfigurability, *18th IFAC: World Congress*, 28.08–02.09 2011, pp. 14814–14819.
11. Firsov Sergii, Igor Kulik Compact vertical take-off and landing aerial vehicle for monitoring tasks in dense urban areas, *Transport Problems*, 2015, Vol. 10, Issue 8, pp. 29–34.
12. Firsov Sergii, Pishchukhina O. A. Assessment of functional sustainability of computer-integrated technological systems [Text], *Open information and computer-integrated technologies*, 2016, Vol. 72, No. 4, pp. 112–121.
13. Firsov S. N., Reznikova O. V. Hardware-software complex for experimental tests of control processes and diagnostics of small spacecraft's faults, *Devices and systems. Management, monitoring, diagnostics*, 2014, Vol. 6, pp. 60–69.

Наукове видання

Радіоелектроніка,

інформатика,

управління

№ 2/2018

Науковий журнал

Головний редактор – д-р техн. наук С. О. Субботін

Заст. головного редактора – д-р техн. наук Д. М. Піза

Комп'ютерне моделювання та верстання С. В. Зуб
Редактор англійських текстів С. О. Субботін

Оригінал-макет підготовлено у редакційно-видавничому відділі ЗНТУ

Свідоцтво про державну реєстрацію
КВ № 6904 від 29.01.2003.

Підписано до друку 04.06.2018. Формат 60×84/8.
Папір офс. Різогр. друк. Ум. друк. арк. 21,39.
Тираж 300 прим. Зам. № 559.

69063, м. Запоріжжя, ЗНТУ, друкарня, вул. Жуковського, 64

Свідоцтво суб'єкта видавничої справи
ДК № 2394 від 27.12.2005.