

ISSN 1684–8853

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

3(94)/2018

3(94)/2018

INFORMATSIONNO- UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

REFEREED EDITION

Founder

«Information and Control Systems», Ltd.

PublisherSaint-Petersburg State University
of Aerospace Instrumentation**Editor-in-Chief**M. Sergeev
Dr. Sc., Professor, Saint-Petersburg, Russia**Deputy Editor-in-Chief**E. Krouk
Dr. Sc., Professor, Moscow, Russia**Executive secretary**

O. Muravtsova

Editorial BoardV. Anisimov
Dr. Sc., Professor, Saint-Petersburg, Russia
B. Bezruchko
Dr. Sc., Professor, Saratov, Russia
N. Blaunstein
Dr. Sc., Professor, Beer-Sheva, Israel
C. Christodoulou
PhD, Professor, Albuquerque, New Mexico, USA
A. Dudin
Dr. Sc., Professor, Minsk, BelarusI. Dumer
PhD., Professor, Riverside, USAM. Favorskaya
Dr. Sc., Professor, Krasnoyarsk, RussiaL. Fortuna
PhD, Professor, Catania, ItalyA. Fradkov
Dr. Sc., Professor, Saint-Petersburg, RussiaA. Hramov
Dr. Sc., Professor, Saratov, RussiaL. Jain
PhD, Professor, Canberra, AustraliaV. Khimenko
Dr. Sc., Professor, Saint-Petersburg, RussiaG. Maltsev
Dr. Sc., Professor, Saint-Petersburg, RussiaG. Matvienko
Dr. Sc., Professor, Tomsk, RussiaV. Melekhin
Dr. Sc., Professor, Saint-Petersburg, RussiaB. Meyer
Dr. Sc., Professor, Zurich, SwitzerlandA. Ovodenko
Dr. Sc., Professor, Saint-Petersburg, RussiaY. Podoplyokin
Dr. Sc., Professor, Saint-Petersburg, RussiaJ. Seberry
PhD, Professor, Wollongong, AustraliaA. Shalyto
Dr. Sc., Professor, Saint-Petersburg, RussiaA. Shelupanov
Dr. Sc., Professor, Tomsk, RussiaA. Shepeta
Dr. Sc., Professor, Saint-Petersburg, RussiaYu. Shokin
RAS Academician, Dr. Sc., Novosibirsk, RussiaA. Smirnov
Dr. Sc., Professor, Saint-Petersburg, RussiaV. Vassiliev
RAS Corr. Member, Dr. Sc., Professor, Saint-Petersburg, RussiaZ. Yuldashev
Dr. Sc., Professor, Saint-Petersburg, RussiaR. Yusupov
RAS Corr. Member, Dr. Sc., Professor, Saint-Petersburg, RussiaA. Zeifman
Dr. Sc., Professor, Vologda, Russia**Editor:** A. Larionova**Proofreader:** T. Zvertanovskaia**Design:** M. Chernenko, Y. Umnitcina**Layout and composition:** Y. Umnitcina**Contact information**The Editorial and Publishing Center, SUAI
67, B. Morskaia, 190000, St. Petersburg, RussiaWebsite: <http://i-us.ru/en>, e-mail: i.us.spb@gmail.com

Tel.: +7 - 812 494 70 02

The Journal was registered in the Ministry of Press,
Broadcasting and Mass Media of the Russian Federation.
Registration Certificate JD № 77-12412 from April, 19, 2002.
Re-registration in the Federal Service for Supervision in the Sphere of Telecom,
Information Technologies and Mass Communications (ROSKOMNADZOR)
due to change of the founder: «Information and Control Systems», Ltd.,
JD № FS77-49181 from March, 30, 2012.

© Corporate authors, 2018

THEORETICAL AND APPLIED MATHEMATICS**Balonin Yu. N., Sergeev A. M.** *Two-Circulant Hadamard Matrices, Weighing Matrices, and Ryser's Conjecture* 2**Ziatdinov S. I., Osipov L. A., Sokolova Y. V.** *Synthesis of Discrete Filters by Invariant Differential Equations* 10**INFORMATION PROCESSING AND CONTROL****Britov G. S.** *Optimal Terminal Diagnostics of Controlled Dynamic Systems* 17**Kuznetsov V. A., Liudaev M. U.** *Photometric Stereo Scanner with Flash Light Lamp* 25**INFORMATION AND CONTROL SYSTEMS****Martynova L. A.** *Effective Retention of AUV Position on Route Trajectory During Seismic Survey* 34**Motorin D. E., Popov S. G.** *Multi-Criteria Path Planning Algorithm for a Robot on a Multilayer Map* 45**SYSTEM AND PROCESS MODELING****Tatarnikova T. M., Volskiy A. V.** *Estimation of Probabilistic-Temporal Characteristics of Network Nodes with Traffic Differentiation* 54**INFORMATION SECURITY****Ageev S. A., Saenko I. B., Kotenko I. V.** *Method and Algorithms of Anomaly Detection in Multiservice Network Traffic based on Fuzzy Logical Inference* 61**Kozin I. S.** *Providing Personal Data Protection in an Information System based on User Behavior Analytics* 69**INFORMATION CODING AND TRANSMISSION****Trofimov A. N.** *Random Coding Bound for Channels With Memory – Decoding Function with Partial Overlapping. Part 1. Derivation of Main Expression* 79**Matveev N. V., Turlikov A. M.** *Slotted ALOHA with Iterative Procedure for Resolving Collisions. Stability and Non-Stability* 89**CONTROL IN MEDICAL AND BIOLOGICAL SYSTEMS****Limanova N. I., Ataev S. G.** *Computer Aided Tomography Picture Analysis on the base of Stage-By-Stage Binarization and its Software Implementation* 98**INFORMATION ABOUT THE AUTHORS**

107

Submitted for publication 27.04.18. Passed for printing 20.06.18. Format 60×84_{1/8}. Phototype SchoolBookC. Digital printing.Layout original is made at the Editorial and Publishing Center, SUAI,
67, B. Morskaia, 190000, St. Petersburg, RussiaPrinted from slides at the Editorial and Publishing Center, SUAI,
67, B. Morskaia, 190000, St. Petersburg, Russia

The journal is distributed by subscription. Subscription can be made in the Editorial and publishing center, SUAI as well as in any post office based on «Rospechat» catalogue: № 15385 — semiannual subscript.

3(94)/2018

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-
УПРАВЛЯЮЩИЕ
СИСТЕМЫ

Учредитель
ООО «Информационно-управляющие системы»

Издатель
Санкт-Петербургский государственный университет
аэрокосмического приборостроения

Главный редактор
М. Б. Сергеев,
д-р техн. наук, проф., Санкт-Петербург, РФ

Зам. главного редактора
Е. А. Крук,
д-р техн. наук, проф., Москва, РФ

Ответственный секретарь
О. В. Муравцова

Редакционная коллегия:
В. Г. Анисимов,
д-р техн. наук, проф., Санкт-Петербург, РФ
Б. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФ
Н. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, Израиль
В. Н. Васильев,
чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
Л. С. Джайн,
д-р наук, проф., Канберра, Австралия
А. Н. Дудин,
д-р физ.-мат. наук, проф., Минск, Беларусь
И. И. Думер,
д-р наук, проф., Риверсайд, США
А. И. Зейфман,
д-р физ.-мат. наук, проф., Вологда, РФ
К. Кристодолу,
д-р наук, проф., Альбукерке, Нью-Мексико, США
Г. Н. Мальцев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Г. Г. Матвиенко,
д-р физ.-мат. наук, проф., Томск, РФ
Б. Мейер,
д-р наук, проф., Цюрих, Швейцария
В. Ф. Мелехин,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. А. Оводенко,
д-р техн. наук, проф., Санкт-Петербург, РФ
Ю. Ф. Подоплёкин,
д-р техн. наук, проф., Санкт-Петербург, РФ
Д. Себерри,
д-р наук, проф., Волонгонг, Австралия
А. В. Смирнов,
д-р техн. наук, проф., Санкт-Петербург, РФ
М. Н. Фаворская,
д-р техн. наук, проф., Красноярск, РФ
Л. Фортуна,
д-р наук, проф., Катания, Италия
А. Л. Фрадков,
д-р техн. наук, проф., Санкт-Петербург, РФ
В. И. Хищенко,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. Е. Храмов,
д-р техн. наук, проф., Саратов, РФ
А. А. Шалыто,
д-р техн. наук, проф., Санкт-Петербург, РФ
А. А. Шелупанов,
д-р техн. наук, проф., Томск, РФ
А. П. Шепета,
д-р техн. наук, проф., Санкт-Петербург, РФ
Ю. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ
З. М. Юлдашев,
д-р техн. наук, проф., Санкт-Петербург, РФ
Р. М. Юсупов,
чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ

Редактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: М. Л. Черненко, Ю. В. Умницына
Компьютерная верстка: Ю. В. Умницына
Адрес редакции: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,
сайт: <http://i-us.ru>

Журнал зарегистрирован в Министерстве РФ по делам печати,
телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-49181 от 30 марта 2012 г.

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий,
в которых должны быть опубликованы основные научные результаты диссертации
на соискание ученой степени доктора и кандидата наук».
© Коллектив авторов, 2018

ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ МАТЕМАТИКА

- Balinin Yu. N., Sergeev A. M.** Two-Circulant Hadamard Matrices, Weighing Matrices, and Ryser's Conjecture 2
- Зиятдинов С. И., Осипов Л. А., Соколова Ю. В.** Синтез дискретных фильтров методом инвариантных дифференциальных уравнений 10

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

- Бритов Г. С.** Оптимальное терминальное диагностирование управляемых динамических систем 17
- Кузнецов В. А., Людаев М. Ю.** Стереофотометрическое трехмерное сканирование с использованием импульсного осветителя 25

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

- Мартынова Л. А.** Метод эффективного удержания положения АНПА на маршрутной траектории при ведении сейсмозащиты 34
- Моторин Д. Е., Попов С. Г.** Алгоритм многокритериального поиска траекторий движения робота на многослойной карте 45

МОДЕЛИРОВАНИЕ СИСТЕМ И ПРОЦЕССОВ

- Татарникова Т. М., Вольский А. В.** Оценка вероятностно-временных характеристик сетевых узлов с дифференциацией трафика 54

ЗАЩИТА ИНФОРМАЦИИ

- Агеев С. А., Саенко И. Б., Котенко И. В.** Метод и алгоритмы обнаружения аномалий в трафике мультисервисных сетей связи, основанные на нечетком логическом выводе 61
- Козин И. С.** Метод обеспечения безопасности персональных данных при их обработке в информационной системе на основе анализа поведения пользователей 69

КОДИРОВАНИЕ И ПЕРЕДАЧА ИНФОРМАЦИИ

- Trofimov A. N.** Random Coding Bound for Channels With Memory – Decoding Function with Partial Overlapping. Part 1. Derivation of Main Expression 79
- Матвеев Н. В., Тюрликов А. М.** Слотовый АЛОНА с итерационной процедурой разрешения коллизий. Стабильность и нестабильность 89

УПРАВЛЕНИЕ В МЕДИЦИНЕ И БИОЛОГИИ

- Лиманова Н. И., Атаев С. Г.** Метод анализа снимков компьютерной томографии на основе поэтапной бинаризации изображений и его программная реализация 98

СВЕДЕНИЯ ОБ АВТОРАХ

107

Сдано в набор 27.04.18. Подписано в печать 20.06.18. Формат 60×84^{1/8}.
Гарнитура SchoolBookC. Печать цифровая.
Усл. печ. л. 13,0. Уч.-изд. л. 18,0. Тираж 1000 экз (1-й завод 140 экз.). Заказ № 260.
Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.
Отпечатано с готовых диалогитивов в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.
Журнал распространяется по подписке. Подписку можно оформить
через редакцию, а также в любом отделении связи по каталогу «Роспечать»:
№ 15385 — полугодовой индекс.

UDC 004.67

doi:10.15217/issn1684-8853.2018.3.2

TWO-CIRCULANT HADAMARD MATRICES, WEIGHING MATRICES, AND RYSER'S CONJECTURE

Yu. N. Balonin^a, Research Fellow, tomaball@mail.ru

A. M. Sergeev^a, Senior Lecturer, asklab@mail.ru

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: Hadamard matrices and weighing matrices share the same family. The latter can fill up voids in the matrix space by setting some elements to zero, but this feature has not been properly studied yet. **Purpose:** To study how the orders of orthogonal matrices used in information processing can affect their structure. **Results:** For Ryser's conjecture about orders critical for cyclic Hadamard matrices, an extension has been suggested, covering Hadamard matrices and weighing matrices which consist of two cyclic blocks. We give examples of Hadamard matrices extended to the newly revealed critical order equal to 32, with symmetrical blocks or, on higher orders, with unsymmetrical blocks. We also present two-circulant weighing matrices which replace Hadamard matrices and alternate with them. There is an exceptional case related to the order 24 on which two-circulant Hadamard matrices or weighing matrices do not exist, forcing you to search for a solution among four-block constructions. A special set of Hadamard matrices of 20- and 52-fold orders is pointed out, as their blocks are asymmetric. A new assumption about the critical order 64 is discussed.

Keywords – Information Processing, Noise-Immune Coding, Masking Images, Orthogonal Matrices, Quasi-Orthogonal Hadamard Matrices, Belevitch Matrices, Weighing Matrices, Two-Circulant Matrices, Ryser's Conjecture.

Citation: Balonin Yu. N., Sergeev A. M. Two-Circulant Hadamard Matrices, Weighing Matrices, and Ryser's Conjecture. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 2–9. doi:10.15217/issn1684-8853.2018.3.2

Introduction

The practical interest to orthogonal (quasi-orthogonal) matrices is attributable to their features which make them highly popular in digital processing and data conversion systems. Orthogonal matrices, including Fourier matrices, Hadamard matrices, and their most close even-order interpretations which are Belevitch matrices and weighing matrices, are used in noise-proof coding, spectral expansion, image processing, code division of communication channels, security masking, etc.

The orthogonality of these matrices enables their congruent transformation. The possible orthogonal bases, including symmetrical, circulant, two-circulant and other matrix constructions, considerably expand the ways of optimization for certain data conversion problems [1, 2]. In coding theory, Hadamard matrix columns are used to build codes with large code distances [3, 4]. The special way to number the columns of such matrices in digital signal processing, image compression and masking is interpreted as a two-level representation of Walsh function.

The features of such matrices assume special importance when these conversions are implemented in specialized processors at hardware or firmware level. Since the form of the matrices, their orders and values of their elements significantly affect the choice of the corresponding filters, cost of hardware and speed of conversion, it is especially im-

portant to properly choose orthogonal matrices out of their vast variety when developing a processor.

Some modern practical applications of orthogonal matrices in genetics, biomechanics, medical technology, crystallography, video data conversion, etc. [5–7] require the fundamentals of the current digital methods to be reconsidered. From this point of view, integer values of matrix elements are not as important as the extremal properties of the matrices and their existence for all possible orders.

The theory of Hadamard matrices H_n with orthogonal columns of elements 1 and -1 was developed from simple manual calculation methods supplementing the initial sequence of Sylvester matrices towards more sophisticated ones with the use of nested matrices by Scarpis method or finite fields used by Paley [8–10].

As time elapsed, the researchers' interest moved from unstructured or semi-structured matrices to those with a clearly pronounced structure [11, 12].

Ryser was the first who noticed that the existence of orthogonal circulant matrices had a limitation. He formulated a conjecture that there were no circulant Hadamard matrices of orders $n > 4$. Turin proved in his work [2] that the conjecture was true for matrices of 8-fold orders. The trials to prove this statement for a more general case are still a subject of profound theoretical research in the area of high-order matrices, though far from any practical application. More practically essential were the suggestions to go beyond the accepted limitation at

the cost of some minor concessions. For example, Barker introduced so called Barker codes which are a source for circulant matrices of orders not higher than 13. When the order is higher than 4, they are not orthogonal but close to that in a strict sense pointed out by Barker himself: he noted that their autocorrelation function has a spike at the beginning but then deviates from zero by no more than 1, oscillating with the values 1 and -1.

Thus, the subject of circulant matrices became exhausted and the interest gradually moved to the area of two-circulant structures which had been poorly studied until recently. Note that a circulant structure is symmetric about the secondary diagonal, therefore Barker's conjecture actually describes a limitation for symmetric matrices.

Two-circulant Hadamard matrices are built on the base of two monocirculant matrices **A** and **B** of twice smaller order. **A** and **B** can be either circulant or backcirculant. Note that when applied to two-circulant Hadamard matrices, an extended interpretation of Ryser's conjecture becomes possible [13–15].

The goal of this work is to describe the structures of two-circulant Hadamard matrices found by the algorithm of search for local determinant maximum [16], different from all the above-listed classical methods, and to study in more details the peculiarities of the extended Ryser's conjecture about Hadamard matrices with two-circulant structure.

Alteration of Hadamard Matrices and Weighing Matrices

For the sake of convenience, we will consider two-circulant matrices built on a circulant **A** and backcirculant **B** matrices. Ryser's limitation is also valid for matrices of orders smaller than the critical order 4. A circulant Hadamard matrix of second order H_2 does not exist [2].

A two-circulant structure considerably expands the opportunities for combining. A circulant Ryser's matrix of order 4 can be treated as a two-circulant matrix built of monocirculant blocks $A = H_2$ and $B = J$, where **J** is a block of ones.

Fig. 1 shows an additional structure obtained by doubling a circulant diagonal Ryser's matrix H_4 by Silvester's rule. The shown two-circulant Hadamard matrices are symmetric and doubly symmetric by blocks, but so far they just slightly move Ryser's bound from order 4 to 8.

Paired elements 1 and -1 of Hadamard matrices are traditionally depicted on matrix portraits as white and black cells respectively.

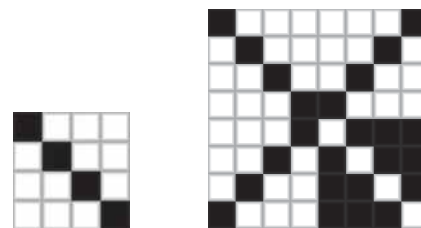
The next non-trivial generalization of Ryser's rule is that an order-12 two-circulant matrix, though still doubly symmetric, is "forced" to have zeros on both the diagonals of blocks **A** and **B**. Such matrices

were first introduced by professor J. Seberry from Wollongong Science Centre in Australia. She dubbed them "weighing matrices" and denoted as **W** [1]. Later it became common to formally denote such matrices by specifying not only the order but also the number of non-zero elements in the rows. In our case, it is $W(n, n-2)$. Elements with value 0 are usually depicted on matrix portraits in gray.

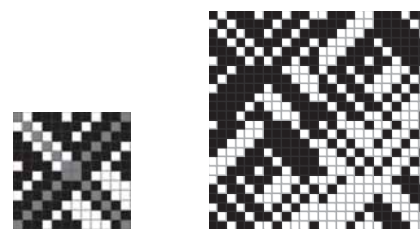
From a weighing matrix $W(n, n-2)$ you can go to a Hadamard matrix of a twice higher order H_{2n} whose blocks will coincide with those of the weighing matrix, having the same signs and values of the diagonal elements, which is a generalization of Silvester's order duplication algorithm. Fig. 2 shows both these matrices. The first of them, weighing matrix W_{12} , was found by a determinant optimization program [12, 16].

Hadamard matrices are strictly optimal by determinant. Hence, an order-12 weighing matrix which differs from them and has a simple two-circulant structure exists in the secondary maximum which is the local maximum of the determinant. This means that any sufficiently small change in the matrix elements not making the absolute value of an element higher than 1 decreases the determinant. Therefore, the determinant optimization algorithm can be used to find and analyze weighing matrices in the case when two-circulant Hadamard matrices do not exist and the absolute maximum of the determinant belongs to more complex structures.

A weighing matrix W_{12} replaces a two-circulant Hadamard matrix which does not exist for this order. Weighing matrices alternate with Hadamard matrices in two-circulant form. On order 16, we again



■ Fig. 1. Portraits of a circulant H_4 and two-circulant H_8 matrices



■ Fig. 2. Portraits of matrices W_{12} and H_{24}

meet a doubly symmetric structure (Fig. 3) which moves Ryser's bound still farther from order 4. This result argues against Ryser's conjecture.

We can presume that this rule is general for all two-circulant matrices of orders 12, 20, 28, 26, etc. On these orders, Hadamard matrices replace weighing matrices which have two zero diagonals. Hadamard matrices of orders $n = 2^k$ have structures belonging to the general sequence of Sylvester's orders; however, as we will see later, they do not keep a double-axis or ordinary symmetry. In particular, unsymmetric two-circulant Hadamard matrices produce codes of Mark Golay who used them to continue his search for Barker's monocycles, not orthogonal but close to that. On the other hand, Golay's codes do not support symmetry, so the structure we found for order 16 belongs to a different family, not yet described.

Symmetric order-16 Hadamard and Golay matrices precede the doubly symmetric weighing matrix W_{20} shown in Fig. 4 along with a Hadamard matrix of a doubled order, according to our presumptions.

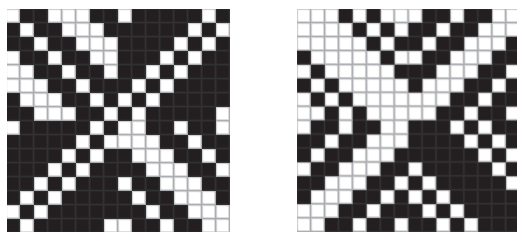
Matrix H_{40} demonstrates a *twister of Seberry* (nearly symmetric Hadamard matrices): the elements of its second subblock form a circulating structure. You can obtain it by unfolding the blocks **A** and **B** of a weighing matrix during the transition to a Hadamard matrix with circulating elements. The twister is a model of standing waves in a square pool with four oscillating areas described by the matrices of nested blocks. Hence, Hadamard matrices are a mathematical interpretation of resonances in a closed cavity (standing waves). Pioneering works in this area belong, among others, to professor J. Seberry. According to Golay's studies [2], on or-

der 20 a main maximum of the determinant exists, producing an unsymmetric matrix H_{20} (Fig. 5).

Thus, unsymmetric Golay's sequences for a side branch of orthogonal codes starting on order 20 complement symmetric sequences of weighing matrices. This fact has never been pointed out in scientific literature. Three Golay's side branches are known; the next one starts on order 52. So, order 20 is an order on which symmetry becomes the cost of an attainable absolute maximum. If you focus on symmetric structures, it makes more sense to use open weighing matrices.

Note that the period on which you meet a two-circulant Hadamard matrix grows as the problem size increases. This feature is also shared by weighing matrices. For example, on order 24 we failed to find a weighing two-circulant matrix, which happened for the first time. Instead of it, we found a more complex structure which consisted of not two but four blocks and inherited a pair of blocks from W_{12} . Accordingly, it produces a matrix H_{24} without the need to double the order, which is demonstrated in Fig. 6.

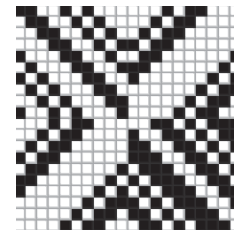
The advantage of complicating is that a four-block weighing matrix, during the transition to a



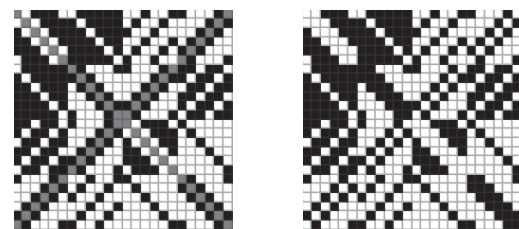
■ Fig. 3. Portraits of two matrices H_{16} with symmetric blocks



■ Fig. 4. Portraits of matrices W_{20} and H_{40}



■ Fig. 5. Unsymmetric matrix H_{20}



■ Fig. 6. Portraits of matrices W_{24} and H_{24}



■ Fig. 7. Portraits of matrices W_{28} and H_{56}

Hadamard matrix, does not require that the order is doubled. Actually, it is a general rule, because when the order is doubled, we go exactly to a four-block structure, found here in a slightly different form with respect to W_{12} . It is followed by a weighing matrix W_{28} which does not contain anything new or unexpected; by doubling the order, it produces a Hadamard matrix H_{56} . These two matrices are shown in Fig. 7.

Extended Ryser’s Bound and its Generalizations

The episode with critical order 24 shows that matrix structures can become more complex. In what follows, we confirm the result obtained in [13] as on order 32 we again meet a doubly symmetric Hadamard matrix of an order of Sylvester’s sequence. For reference, a structure which is not doubly symmetric is shown next to it in Fig. 8. Remember that the known two-circulant Golay’s matrices are not symmetric.

Analysis has shown that order 32 describes the generalization of Ryser’s conjecture for two-circulant structures. Beyond this order, it is impossible to obtain a Hadamard matrix in two-circulant form. Nevertheless, our experience with searching for symmetric structures by a determinant optimization program shows that symmetric codes still have some safety margin: a doubly symmetric weighing matrix W_{36} . By doubling its order, you can turn it into a symmetric Hadamard matrix H_{72} shown in Fig. 9.

This additional experience supplements the estimation of Ryser’s bound for two-circulant orthogo-

nal matrices performed in [13] by new details about weighing matrices.

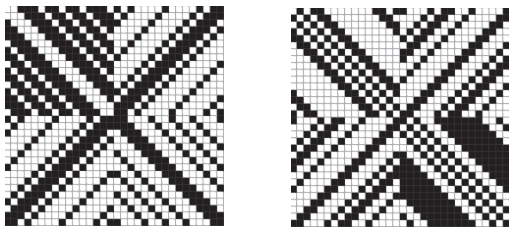
Two-circulant matrices H_{40} and H_{52} are not matrices of Sylvester’s type. They belong to sequences of unsymmetric matrices of 20-fold and 52-fold orders (Fig. 10).

The new doubly symmetric weighing matrix W_{52} we have found can produce, through doubling its order, a nearly symmetric Hadamard matrix H_{104} shown in Fig. 11. In other words, from the viewpoint of searching for symmetric matrices, this family is preferable, being considerably different from Mark Golay’s two-circulant matrices. Golay’s sequences of orders 2, 10 and 26 producing Hadamard matrices of orders 4, 20 and 52 yield unsymmetric codes, so their practical alternative can be symmetric codes of weighing matrices of respective orders.

The symmetry of orthogonal matrices on the specified orders and the very fact of their existence is a subject of modern studies. According to a new assumption which resulted from considering a chain of two-circulant Hadamard matrices, order 64 is a key for checking the extended conjecture.



■ Fig. 10. Portraits of matrices H_{40} and H_{52}



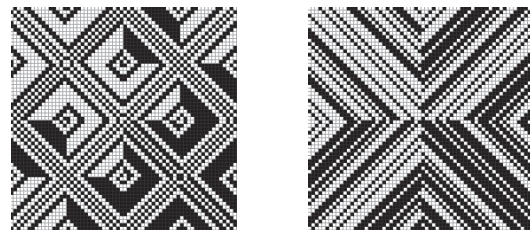
■ Fig. 8. Two matrices H_{32}



■ Fig. 11. Portraits of matrices W_{52} and H_{104}



■ Fig. 9. Portraits of matrices W_{36} и H_{72}



■ Fig. 12. Symmetric and unsymmetric matrices H_{64}

Order 64 is the simplest one for search and symmetry check of Hadamard matrices from Sylvester's sequence. For this order, we have identified the symmetric four-circulant matrix shown in Fig. 12.

If a symmetric two-circulant matrix H_{64} does not exist (which is confirmed by our experiment on searching, with a formal algorithm, for matrices optimal by determinant), then on higher orders there are no symmetric two-circulant matrices.

Chains of Golay's matrices $H_{20} - H_{40} - \dots$ and $H_{52} - H_{104} - \dots$ of 20-fold and 52-fold orders are apparently unsymmetric.

Algorithm Scheme

Fig. 13 shows a simplified scheme of the algorithm developed by the authors. It consists of three sequentially performed blocks (from left to right):

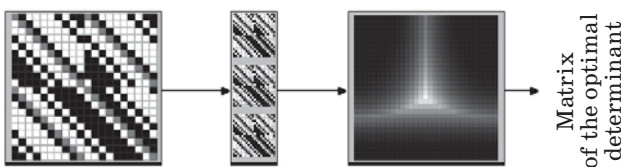
- two-circulant matrix generator;
- cross-accumulator which provides the possibility for the matrices to swap their parts **A** and **B**;
- two-circulant matrix determinant optimizer.

With software implementing the proposed algorithm, you can control the optimized determinant and visually control the structure of the resulting matrix.

The determinant optimizer is fully considered in [16] and will not be discussed here. We will only note that it is a novel scheme, not like any other one widely known from the numerical analysis literature. It optimizes not the determinant calculation algorithm, but a matrix with n^2 elements.

A matrix with a two-circulant structure has a stable local determinant, because to let the iterations go to the absolute maximum, you would have to rebuild the rigidly specified initial structure. It is much easier to convert the matrix elements 1 to 0 or -1, providing that the result is fixed in the form of a Hadamard matrix or weighing matrix if they exist. This scheme can easily be transformed for tetracycle search.

The cross-accumulator is a block which significantly reduces the time cost of generating the necessary initial matrices. At this stage, the blocks **A** and **B** can be relatively random, and the combinational circuit which allows the initial matrices to swap their already generated parts saves the computing resources. Besides, the matrices accumulat-



■ Fig. 13. Algorithm scheme

ed in the database can be cached, being spread over buffer zones according to the value of a certain cache function. Blocks which considerably differ from each other are filtered and never come to the cross-accumulator input together.

These preliminary measures enrich the output of the cross-accumulator, giving the optimizer the material it needs. Such an algorithm can be implemented in several languages. We have developed software versions in C++, Pascal and Javascript. Each of these versions has its advantages.

The implementation in C++ is the fastest one but needs a heavy compiler and is not suitable for all computers. The Pascal implementation is slower but more universal; it can be run under any version of Windows starting from XP. The Javascript implementation within mathscinet.ru international mathematical network is available for wide circles of scientists and researchers from different countries, as well as for students studying orthogonal transformations.

Javascript software has another implicit advantage: it provides the opportunity for a wide circle of users to modify the underlying algorithm. This is very valuable in a research when you need to modify both the algorithm and the construction of the calculated matrix.

Apart from the cyclic array form, there is its negacyclic form, when during a shift of a block row its last pushed-out element is carried to the beginning with its sign inverted. Besides, the size of a two-circulant matrix can be reduced by several elements using the matrix border distinctive in its wide variety of implementation. Most commonly, Hadamard and weighing matrices have structures with unary or binary border, leading to a multi-block implementation of the matrix core. Studying the cores can explain on which structure the absolute determinant maximum is attainable and where it goes to from a two-circulant structure with buffer zeros. However, the search for universal cores of orthogonal matrices is beyond the scope of this research.

The Novelty of the Results

Hadamard and weighing matrices are an object of intensive studies because of their application in information processing; in particular, image processing. New facts are regularly discovered: for example, an exhaustive search for rows in weighing matrices with a single zero (Belevitch matrices) showed recently that these matrices do not exist in a two-circulant form on order 66. The same, according to some experimental results, is true for order 86. These orders follow each other with a step of 20.

On order 46, Maton's result is widely known. He found a solution in the form of a five-block matrix:

in two of these five, not elements are shifted but entire sequences of elements. These results are very valuable because forms of matrices with an absolute determinant maximum are not theoretically known, and the discovery of Belevitch matrices of orders 66 and 86 stimulates the research in this field. So far, the question about what Maton's structure turns into when the order increases by 20 has no answer, either.

Studies of weighing matrices have been focused mostly on forms with many zeros. Considering them a natural supplement to two-circulant Hadamard matrices when the latter do not exist is a new approach. Especially new is the idea that they expand symmetric structures, because the extended Ryser's bound as such, up to order 32, was discovered and studied quite recently, with the use of supercomputers in a scientific centre in Canada. Hence, the results of our experiment introduce a significant correction to this knowledge of two-circulant matrices, showing what a symmetric structure turns into when it cannot be embodied in the form of matrices with elements 1 and -1.

The discovery of symmetric two-circulant weighing matrices $W(n, n-2)$ on orders of Hadamard matrices $n = 4t$ makes them available for practical applications and poses new theoretical problems: in particular, what Ryser's bound is for weighing matrices. Symmetric weighing matrices with two zeros supplement symmetric Belevitch matrices in the sense that in both cases they are a source for Hadamard matrices of a doubled order. Finite field theory receives a new application domain here, as we have to find out how to calculate these new matrices; sometimes, for very high orders.

For video information coding systems, the two-circulant scheme is good because it is relatively simple to implement. Omissions of two-circulant Hadamard matrices caused unnecessary problems for the applications in passing to tetracyclic constructions. In other words, the results of the research in this area seriously affects the efficiency of the applications.

Some Applications for Visual Information Transmission Systems

The transmission of visual information, i.e. images or frame-by-frame video, is an integral feature of territory monitoring systems, multifunctional registration systems, distributed industrial systems, security surveillance systems, and other systems which use open networks to build their infrastructure. The information transmitted in such distributed video systems, even not top-secret information, needs to be protected from unauthor-

ized access, distortion in the communication channels or substitution.

Generally, the process of protecting visual information goes as follows. At the transmitting side of a distributed system, the protected information is shaped. Then it is directed to a communication channel where it may become a target for the above-mentioned threats. At the receiving side, the protected information is recast into the initial form, along with finding out whether it has been distorted by noise or deliberately changed by a third party. Depending on the implementation of the method and circuitry, the visual information can come to the receiving side either with or without losses.

An effective way to protect visual information from unauthorized usage is the method of bilateral matrix masking [17]. According to it, an image (frame) P at the transmitting side is masked by an orthogonal Hadamard matrix as $Y = H^T P H$. Such a transformation visually destroys the image down to a level similar to noise, with computing cost much smaller compared to coding methods. This allows you to mask images (video frames) in real time, as fast as they come from the video camera matrix.

The shaped and masked image Y is passed through the communication channel to the receiving side where it undergoes a reverse bilateral transformation in order to obtain the initial image according to the expression $P = (H^T)^{-1} Y H^{-1}$.

The use of orthogonal Hadamard matrices H ($H^{-1} = H^T$) simplifies the computing down to $P = H Y H^T$. In this case, the reverse transformation repeats the direct one with a precision of transposing the masking matrix; also, there is no need to separately store or calculate the reverse matrix: this saves memory when the method is implemented within the system. The next advantage can be provided by switching to Hadamard matrices with symmetric structures [18] for which the amount of the stored data to produce an order- n Hadamard matrix of circulant or two-circulant structure can be no more than n elements, as many as one row of a matrix.

In a similar way, we can implement the method of noise-proof image coding used in data transmission systems when the ratio signal/noise is low, known as strip transformation [19]. However, for this transformation two different orthogonal matrices are used (those of premultiplication and postmultiplication), and the multiplication itself yields Kronecker product, adding small extra computing cost with software implementation of the method, as compared to masking.

In order to prevent image substitution or changing, the procedure of introducing digital watermarks can be applied, with the use of Hadamard transformation [20].

Conclusion

The principal question of Ryser's theory about two-circulant Hadamard matrices is determining the maximum achievable order of symmetry and the form of symmetric matrices which are exceptions. For example, the conference matrices of Maton's construction found for order 46 already have non-circulant blocks in their structure.

If there are no symmetric two-circulant structures, and Golay's codes produce only unsymmetric blocks, then the source for symmetric Hadamard matrices of orders higher than 32 will be two-circu-

lant and four-circulant weighing Seberry matrices with two zeros on their diagonals.

The extension of Ryser's conjecture can explain the peculiarities of the order alteration followed by Hadamard matrices and weighing matrices. It can also explain the difficulties experienced when searching for matrices of Hadamard family and symmetric conference matrices.

The work has been carried out with the support of Ministry of Education and Science of the Russian Federation for research within the development part of the scientific governmental task #2.2200.2017/4.6

References

1. Seberry J. *Orthogonal Designs, Hadamard Matrices, Quadratic Forms and Algebras*. Springer, International Publishing AG, 2017. 459 p. doi:10.1007/978-3-319-59032-5_1 <http://www.springer.com/us/book/9783319590318>
2. Craigen R., Kharaghani H. Hadamard Matrices and Hadamard Designs. In: *Handbook of Combinatorial Designs*. 2nd ed. C. J. Colbourn, J. H. Dinitz (eds). Boca Raton, FL, Chapman & Hall/CRC, 2007. Pp. 273–280.
3. Tarannikov Iu. V. *Kombinatornye svoistva diskretnykh struktur i prilozheniia k kriptografii* [Combinatorial Properties of Discrete Structures and Applications to Cryptography]. Moscow, MTsNMO Publ., 201. 152 p. (In Russian).
4. Jongkil Kim, Willy Susilo, Man Ho Au and Jennifer Seberry. Efficient Semi-Static Secure Broadcast Encryption Scheme. *LNCS*, Berlin, Springer Verlag, 2014, vol. 8365, pp. 62–76.
5. Petukhov S. V. *Matrichnaia genetika, algebrы geneticheskogo koda, pomekhoustoichivost'* [Matrix Genetics, Algebras of The Genetic Code, Noise Immunity]. Moscow, RKhD Publ., 2008. 316 p. (In Russian).
6. Petoukhov S. V. The Genetic Coding, United-Hypercomplex Numbers and Artificial Intelligence. *Advances in Artificial Systems for Medicine and Education*, 2017, pp. 2–13.
7. Moon Ho Lee, Han Hai, Sung Kook Lee, Petoukhov S. V. A Mathematical Proof of Double Helix DNA to Reverse Transcription RNA for Bioinformatics. *Advances in Artificial Systems for Medicine and Education*, 2017, pp. 23–38.
8. Dragomir Ž. Doković. Generalization of Scarpis' Theorem on Hadamard Matrices. *Linear and Multilinear Algebra*, 2017, vol. 65, iss. 10, pp. 1–3.
9. Dragomir Ž. Doković. Williamson Matrices of Order $4n$ for $n = 33; 35; 39$. *Discrete Math*, 1993, vol. 115, pp. 267–271.
10. Holzmann W. H., Kharaghani H., Tayfeh-Rezaie B. Williamson Matrices up to Order 59. *Designs, Codes and Cryptography*, 2008, no. 46, pp. 343–352.
11. Sergeev A. M. Generalized Mersenne Matrices and Balonin's Conjecture. *Automatic Control and Computer Sciences*, 2014, vol. 48, no. 4, pp. 214–220.
12. Balonin Ju. N., Vostrikov A. A., Sergeev A. M., Egorova I. S. On Relationships Among Quasi-Orthogonal Matrices Constructed on the Known Sequences of Prime Numbers. *Trudy SPIIRAN* [SPIIRAS Proceedings], 2017, iss. (1)50, pp. 209–223. doi:<http://dx.doi.org/10.15622/sp.50.9> (In Russian).
13. Balonin N. A., Djokovic D. Symmetry of Two Circulant Hadamard Matrices and Periodic Goley Pairs. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 3, pp. 2–16 (In Russian). doi:10.15217/issn1684-8853.2015.3.2
14. Gene Awyzio and Jennifer Seberry. On Good Matrices and Skew Hadamard Matrices. *Springer Proc. in Mathematics and Statistics: Algebraic Design Theory and Hadamard Matrices*, 2015, pp. 13–28.
15. Olivia Di Matteo, Dragomir Z. Djokovic, Ilias S. Kotsireas. Symmetric Hadamard Matrices of Order 116 and 172 Exist. *Special Matrices*, 2015, no. 3, pp. 227–234.
16. Balonin N. A., Sergeev M. B., Suzdal V. S. Dynamic Generators of the Quasiorthogonal Hadamard Matrix Family. *Trudy SPIIRAN* [SPIIRAS Proceedings], 2017, iss. (5)54, pp. 224–243 (In Russian). doi:<http://dx.doi.org/10.15622/sp.54.10>
17. Vostrikov A., Sergeev M. Expansion of the Quasi-Orthogonal Basis to Mask Images. *Smart Innovation, Systems and Technologies*, 2015, vol. 40, pp. 161–168. doi: 10.1007/978-3-319-19830-9_15
18. Sergeev A. M., Blaunstein N. S. Orthogonal Matrices with Symmetrical Structures for Image Processing. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 6, pp. 2–8 (In Russian). doi:10.15217/issn1684-8853.2017.6.2
19. Mironovskii L. A., Slaev V. A. The Strip Method of Noise-Immune Image Transformation. *Measurement Techniques*, 2006, vol. 49, no. 8, pp. 745–754.
20. Anthony T. S. Ho, Jun Shen, Soon Hie Tan. A Robust Digital Image-in-Image Watermarking Algorithm Using the Fast Hadamard Transform. *School of Electrical and Electronic Engineering*, 2003, vol. 4793, pp. 76–85.

УДК 004.67

doi:10.15217/issn1684-8853.2018.3.2

Двуматричные матрицы Адамара, взвешенные матрицы и гипотеза РайзераБалонин Ю. Н.^а, инженер, tomaball@mail.ruСергеев А. М.^а, старший преподаватель, asklab@mail.ru^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: матрицы Адамара и взвешенные матрицы образуют единое семейство, причем свойство последних заполнять пустоты матричного пространства посредством обнуления части элементов изучено недостаточно полно. **Цель:** исследование влияния порядков ортогональных матриц, используемых для обработки информации, на их структуру. **Результаты:** рассмотрено расширение гипотезы Райзера, трактующей критические для циклических матриц Адамара порядки, на матрицы Адамара и взвешенные матрицы, состоящие из двух циклических блоков. Приведены примеры матриц Адамара, расширенных до выявленных на новом критическом порядке, равном 32, с симметричными блоками, и более высоких порядках — с несимметричными блоками. Представлены чередующиеся с матрицами Адамара и заменяющие их двуматричные взвешенные симметричные и несимметричные матрицы. Приведен случай-исключение — порядок 24, на котором нет двуматричных матриц Адамара и взвешенных матриц, что вынужденно переводит решение задачи к четырехблочным конструкциям. Отмечена особая линия матриц Адамара порядков, кратных 20 и 52, выделенных среди остальных матриц асимметрией своих блоков. Сформулировано новое предположение о критическом порядке 64.

Ключевые слова — обработка информации, помехоустойчивое кодирование, маскирование изображений, ортогональные матрицы, квазиортогональные матрицы Адамара, матрицы Белевича, взвешенные матрицы, двуматричные матрицы, гипотеза Райзера.

Цитирование: Balonin Yu. N., Sergeev A. M. Two-Circulant Hadamard Matrices, Weighing Matrices, and Ryser's Conjecture // Информационно-управляющие системы. 2018. № 3. С. 2–9. doi:10.15217/issn1684-8853.2018.3.2

Citation: Balonin Yu. N., Sergeev A. M. Two-Circulant Hadamard Matrices, Weighing Matrices, and Ryser's Conjecture. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 2–9. doi:10.15217/issn1684-8853.2018.3.2

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая SCOPUS и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой — различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, снижая рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12-ти языках, включая русский (чтобы выбрать язык, кликните на зеленое поле вверху справа на стартовой странице): <https://orcid.org>

УДК 621.391

doi:10.15217/issn1684-8853.2018.3.10

СИНТЕЗ ДИСКРЕТНЫХ ФИЛЬТРОВ МЕТОДОМ ИНВАРИАНТНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

С. И. Зиятдинов^а, доктор техн. наук, профессор

Л. А. Осипов^а, доктор техн. наук, профессор

Ю. В. Соколова^а, аспирант

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: при создании таких систем обработки информации, как доплеровские измерители скорости и ускорения, системы селекции движущихся целей, радиовысотомеры с частотной модуляцией несущего сигнала, согласованные устройства обнаружения и оценки параметров объектов, используются разнообразные фильтры, синтезируемые методами билинейного z -преобразования, инвариантных импульсной и переходной характеристик. **Цель:** создание методики синтеза дискретных фильтров с использованием дифференциальных уравнений непрерывных фильтров-аналогов. **Результаты:** в общем виде для получения разностных уравнений, описывающих работу синтезируемых дискретных фильтров, производные дифференциальных уравнений заменяются аналогами в виде разностных уравнений. Приведена таблица с весовыми коэффициентами фильтров различных порядков. Расчеты показали, что при правильном выборе периода дискретности частотные свойства синтезированных фильтров практически совпадают с частотными свойствами соответствующих непрерывных фильтров. В качестве примера получены разностные уравнения для фильтров нижних и верхних частот, колебательных звеньев, режекторных и селективных фильтров на основе известных дифференциальных уравнений непрерывных фильтров-аналогов, которые могут быть использованы в измерителях скорости, дальности, углового положения объекта. **Практическая значимость:** предложенная методика позволяет создать разнообразные линейные системы, такие как фильтры нижних и верхних частот, колебательные звенья, режекторные и селективные фильтры, используемые во всевозможных системах обработки информации.

Ключевые слова — фильтр, дифференциальное уравнение, разностное уравнение, частотная передаточная функция.

Цитирование: Зиятдинов С. И., Осипов Л. А., Соколова Ю. В. Синтез дискретных фильтров методом инвариантных дифференциальных уравнений // Информационно-управляющие системы. 2018. № 3. С. 10–16. doi:10.15217/issn1684-8853.2018.3.10

Citation: Ziatdinov S. I., Osipov L. A., Sokolova Y. V. Synthesis of Discrete Filters by Invariant Differential Equations. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 10–16 (In Russian). doi:10.15217/issn1684-8853.2018.3.10

Введение

В настоящее время в связи с интенсивным развитием вычислительной техники и появлением высокоэффективных цифровых устройств, таких как микропроцессоры, микроконтроллеры, сигнальные процессоры, наибольшее распространение получили цифровые методы обработки видеосигналов. Для повышения качества изображений, подавления помех и шумов широко используются разнообразные фильтры нижних и верхних частот, полосовые и режекторные фильтры. При этом к временным и частотным свойствам фильтров в зависимости от решаемых задач предъявляются в каждом конкретном случае специфические требования.

Вопросы построения дискретных фильтров с заданными временными и частотными характеристиками в литературе достаточно хорошо исследованы [1–7]. Для синтеза дискретных фильтров по их непрерывным аналогам используются как частотные, так и временные методы [8]. В основе частотных методов лежит применение

билинейного z -преобразования, позволяющего по заданной частотной передаточной функции непрерывного фильтра разработать дискретный фильтр, частотные свойства которого практически совпадают с частотными свойствами непрерывного фильтра-аналога.

В случае синтеза дискретных фильтров во временной области используется либо метод инвариантной импульсной характеристики, либо метод инвариантной переходной характеристики [9–14]. Оба метода основаны на применении для создания дискретных фильтров отсчетов импульсной или переходной характеристик непрерывного фильтра-аналога.

Вместе с тем возможен метод синтеза дискретных фильтров на основе дифференциальных уравнений, описывающих работу непрерывных фильтров.

В литературе данному методу практически не уделено должного внимания. Однако использование дифференциальных уравнений позволяет сравнительно легко получать алгоритмы работы практически любых дискретных фильтров.

Синтез дискретных фильтров

В самом общем виде линейная система (фильтр) n -порядка описывается следующим линейным дифференциальным уравнением [15, 16]:

$$\begin{aligned}
 & a_0 \frac{d^n y(t)}{dt^n} + a_1 \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_n y(t) = \\
 & = b_0 \frac{d^m x(t)}{dt^m} + b_1 \frac{d^{m-1} x(t)}{dt^{m-1}} + \dots + b_m x(t), \\
 & m \leq n, \tag{1}
 \end{aligned}$$

где $x(t)$, $y(t)$ — входной и выходной сигналы системы; a_i , b_i — весовые коэффициенты.

Производные в соотношении (1) можно представить в виде

$$\begin{aligned}
 \frac{dy(t)}{dt} &= \lim_{T \rightarrow 0} \frac{y(t) - y(t-T)}{T}; \\
 \frac{dx(t)}{dt} &= \lim_{T \rightarrow 0} \frac{x(t) - x(t-T)}{T}; \\
 \frac{d^2 y(t)}{dt^2} &= \lim_{T \rightarrow 0} \frac{\frac{dy(t)}{dt} - \frac{dy(t-T)}{dt}}{T}; \\
 \frac{d^2 x(t)}{dt^2} &= \lim_{T \rightarrow 0} \frac{\frac{dx(t)}{dt} - \frac{dx(t-T)}{dt}}{T}; \\
 &\dots\dots \\
 \frac{d^n y(t)}{dt^n} &= \lim_{T \rightarrow 0} \frac{\frac{d^{n-1} y(t)}{dt^{n-1}} - \frac{d^{n-1} y(t-T)}{dt^{n-1}}}{T}; \\
 \frac{d^m x(t)}{dt^m} &= \lim_{T \rightarrow 0} \frac{\frac{d^{m-1} x(t)}{dt^{m-1}} - \frac{d^{m-1} x(t-T)}{dt^{m-1}}}{T}. \tag{2}
 \end{aligned}$$

В дальнейшем будем считать, что за время T не происходит заметных изменений ни функций $x(t)$, $y(t)$, ни их производных.

Тогда соотношения (2) можно заменить приближенными разностными уравнениями

$$\begin{aligned}
 \frac{dy(t)}{dt} &\approx \frac{y(t) - y(t-T)}{T}; \quad \frac{dx(t)}{dt} \approx \frac{x(t) - x(t-T)}{T}; \\
 \frac{d^2 y(t)}{dt^2} &\approx \frac{y(t) - 2y(t-T) + y(t-2T)}{T^2}; \\
 \frac{d^2 x(t)}{dt^2} &\approx \frac{x(t) - 2x(t-T) + x(t-2T)}{T^2}; \\
 &\dots\dots \\
 \frac{d^n y(t)}{dt^n} &\approx T^{-n} \sum_{i=0}^n c_i y(t-iT);
 \end{aligned}$$

$$\frac{d^m x(t)}{dt^m} \approx T^{-m} \sum_{i=0}^m c_i x(t-iT). \tag{3}$$

Значения весовых коэффициентов разностных уравнений, соответствующих производным различных порядков, представлены в таблице.

С учетом соотношений (3) дифференциальное уравнение линейной системы (1) записывается в виде конечных разностей

$$\begin{aligned}
 & a_0 T^{-n} \sum_{i=0}^n c_i y(t-iT) + \\
 & + a_1 T^{-(n-1)} \sum_{i=0}^{n-1} p_i y(t-iT) + \dots + a_n y(t) = \\
 & = b_0 T^{-m} \sum_{i=0}^m k_i x(t-iT) + \\
 & + b_1 T^{-(m-1)} \sum_{i=0}^{m-1} d_i x(t-iT) + \dots + b_m x(t). \tag{4}
 \end{aligned}$$

В данном выражении весовые коэффициенты c_i соответствуют n -й производной; p_i соответствуют $(n-1)$ -й производной; k_i соответствуют m -й производной; d_i соответствуют $(m-1)$ -й производной и т. д.

Из соотношения (4) находим выходной сигнал фильтра

$$\begin{aligned}
 y(t) &= a_n^{-1} \left[b_0 T^{-m} \sum_{i=0}^m k_i x(t-iT) + \right. \\
 & + b_1 T^{-(m-1)} \sum_{i=0}^{m-1} d_i x(t-iT) + \dots + b_m x(t) - \\
 & - a_0 T^{-n} \sum_{i=0}^n c_i y(t-iT) - a_1 T^{-(n-1)} \sum_{i=0}^{n-1} p_i y(t-iT) \times \\
 & \left. \times (t-iT) - \dots - a_{n-1} T^{-1} \sum_{i=0}^1 l_i y(t-iT) \right]. \tag{5}
 \end{aligned}$$

- Значения весовых коэффициентов
- Values of weighting factors

n	c _i									
	c ₀	c ₁	c ₂	c ₃	c ₄	c ₅	c ₆	c ₇	c ₈	c ₉
1	1	-1								
2	1	-2	1							
3	1	-3	3	-1						
4	1	-4	6	-4	1					
5	1	-5	10	-10	5	-1				
6	1	-6	15	-20	15	-6	1			
7	1	-7	21	-35	35	-21	7	-1		
8	1	-8	28	-56	70	-56	28	-8	1	
9	1	-9	36	-84	126	-126	84	-36	9	-1

При дискретном представлении входного и выходного сигналов фильтра, когда $t = iT$, работа дискретного фильтра на основании (5) определяется выражением

$$y[n] = a_n^{-1} \left\{ b_0 T^{-m} \sum_{i=0}^m k_i x[n-i] + b_1 T^{-(m-1)} \sum_{i=0}^{m-1} d_i x[n-i] + \dots + b_m x[n] - a_0 T^{-n} \sum_{i=0}^n c_i y[n-i] - a_1 T^{-(n-1)} \sum_{i=0}^{n-1} p_i y[n-i] - \dots - a_{n-1} T^{-1} \sum_{i=0}^1 l_i y[n-i] \right\}. \quad (6)$$

Данное соотношение определяет алгоритм работы дискретного фильтра n -порядка.

Примеры синтеза дискретных фильтров

Рассмотрим ряд конкретных примеров.

1. Фильтр нижних частот третьего порядка.

Частотная передаточная функция непрерывного фильтра и его дифференциальное уравнение имеют вид [7]

$$W(j\omega) = 1 / [1 + (j\omega\tau)^3]; \quad \tau^3 \frac{d^3 y(t)}{dt^3} + y(t) = x(t),$$

где τ — постоянная времени фильтра.

После замены в дифференциальном уравнении третьей производной ее разностным эквивалентом (третья строчка в таблице) получим следующее разностное уравнение дискретного фильтра нижних частот третьего порядка:

$$y[n] = \frac{1}{1+a} x[n] + \frac{a}{1+a} \{3y[n-1] - 3y[n-2] + y[n-3]\}, \quad (7)$$

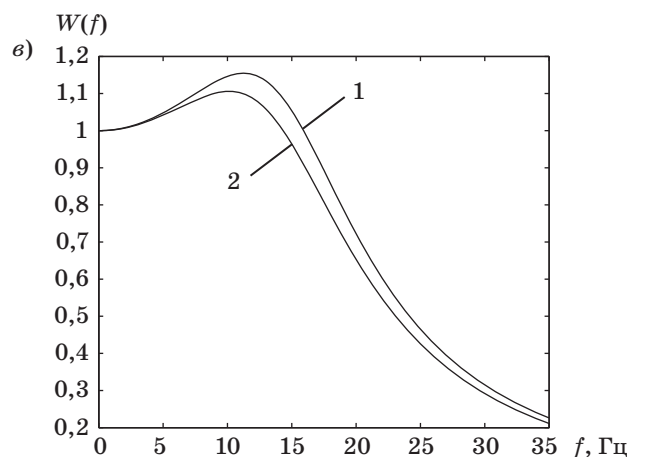
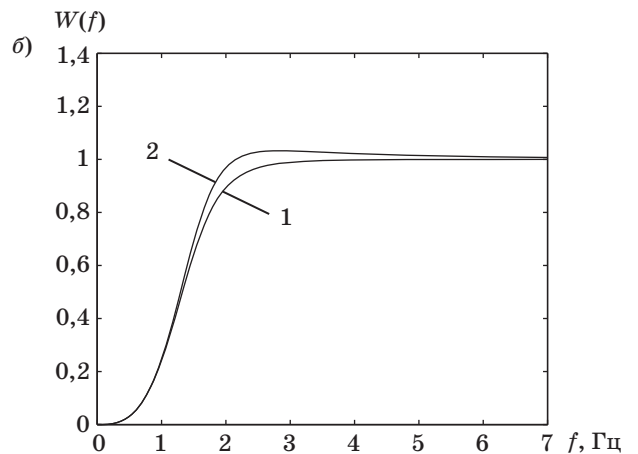
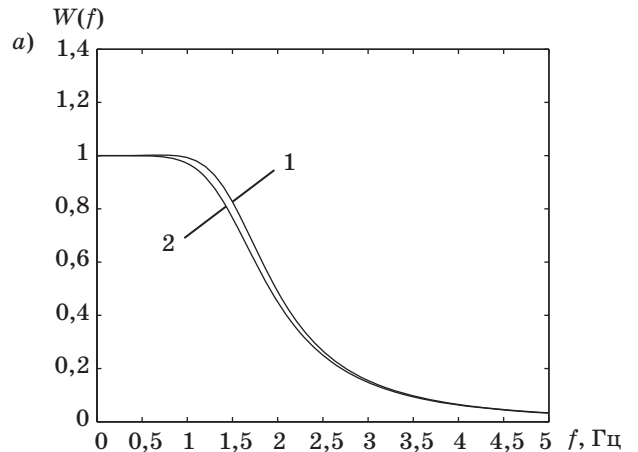
где $a = (\tau/T)^3$.

Выражение (7) позволяет записать амплитудно-частотную характеристику (АЧХ) рассматриваемого фильтра в виде

$$W(\omega) = \frac{1}{1+a} \sqrt{\left[1 - \frac{a}{1+a} (3\cos\omega T - 3\cos 2\omega T + \cos 3\omega T) \right]^2 + \left[\frac{a}{1+a} (3\sin\omega T - 3\sin 2\omega T + \sin 3\omega T) \right]^2}$$

Результаты расчета АЧХ рассматриваемого непрерывного и дискретного фильтров для случая $a = (\tau/T)^3 = 10^3$ приведены на рис. 1, а.

Из представленных АЧХ фильтров видно, что отклонение АЧХ фильтров не превышает 0,5 дБ.



■ **Рис. 1.** Амплитудно-частотные характеристики непрерывного 1 и дискретного 2 фильтра нижних частот (а); фильтра верхних частот (б); колебательного контура (в)

■ **Fig. 1.** AFC of a continuous 1 and discrete 2 LPF (a); HPF (б); oscillatory circuit (в)

2. *Фильтр верхних частот третьего порядка.*

В данном случае частотная передаточная функция и дифференциальное уравнение описываются соотношениями [7]

$$W(j\omega) = (j\omega\tau)^3 / [1 + (j\omega\tau)^3];$$

$$\tau^3 \frac{d^3 y(t)}{dt^3} + y(t) = \tau^3 \frac{d^3 x(t)}{dt^3}.$$

Применяя ранее представленную методику, получим для дискретного фильтра разностное уравнение вида

$$y[n] = \frac{a}{1+a} \{x[n] - 3x[n-1] + 3x[n-2] - x[n-3] - 3y[n-1] + 3y[n-2] - y[n-3]\}.$$

Тогда АЧХ рассматриваемого фильтра запишется следующим образом:

$$W(\omega) = \frac{a}{1+a} \sqrt{\frac{(1 - 3\cos\omega T + 3\cos 2\omega T - \cos 3\omega T)^2 + (3\sin\omega T - 3\sin 2\omega T + \sin 3\omega T)^2}{\left[1 - \frac{a}{1+a}(3\cos\omega T - 3\cos 2\omega T + \cos 3\omega T)\right]^2 + \left[\frac{a}{1+a}(3\sin\omega T - 3\sin 2\omega T + \sin 3\omega T)\right]^2}}.$$

Амплитудно-частотные характеристики непрерывного и дискретного исследуемых фильтров для случая $a = (\tau/T)^3 = 10^3$ приведены на рис. 1, б. При этом отклонение АЧХ данных фильтров не превышает 0,7 дБ.

Рассмотрим синтез более сложных линейных систем методом инвариантных дифференциальных уравнений.

3. *Последовательный колебательный контур.*

Для данной линейной системы частотная передаточная функция и дифференциальное уравнение определяются соотношениями [7]

$$W(j\omega) = 1 / [(j\omega)^2 LC + j\omega rC + 1];$$

$$LC \frac{d^2 y(t)}{dt^2} + rC \frac{dy(t)}{dt} + y(t) = x(t),$$

где L, C, r — индуктивность, емкость и сопротивление потерь колебательного контура соответственно.

Заменив в дифференциальном уравнении вторую и первую производные их разностными выражениями (вторая и первая строчки таблицы), получим следующее разностное уравнение, опре-

деляющее работу дискретного колебательного контура:

$$y[n] = a_0 x[n] + b_1 y[n-1] - b_2 y[n-2],$$

где

$$a_0 = 1 / \left(1 + \frac{LC}{T^2} + \frac{rC}{T}\right);$$

$$b_1 = \left(\frac{2LC}{T^2} + \frac{rC}{T}\right) / \left(1 + \frac{LC}{T^2} + \frac{rC}{T}\right);$$

$$b_2 = LCT^{-2} / \left(1 + \frac{LC}{T^2} + \frac{rC}{T}\right).$$

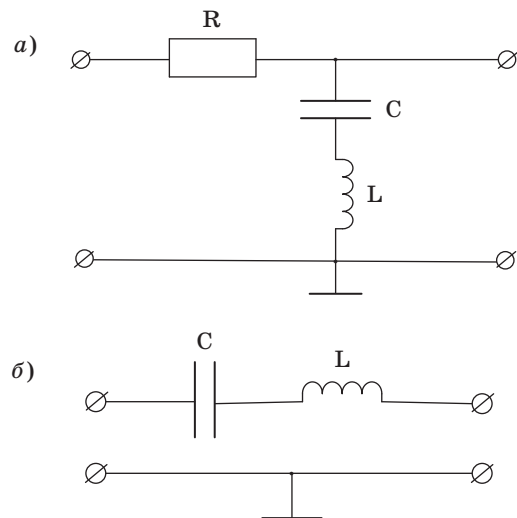
Амплитудно-частотная характеристика рассматриваемого колебательного контура имеет вид

$$W(\omega) = \frac{a_0}{\sqrt{\left((1 - b_1 \cos\omega T + b_2 \cos 2\omega T)^2 + (b_1 \sin\omega T - b_2 \sin 2\omega T)^2\right)}}.$$

Амплитудно-частотные характеристики непрерывного и дискретного последовательного колебательного контура для случая $a_0 = 0,009$, $b_1 = 1,8919$ и $b_2 = 0,9009$ показаны на рис. 1, в. Из представленных данных следует, что отличие указанных АЧХ не превышает 0,7 дБ.

4. *Режекторный фильтр.*

Существует большое количество схемотехнических решений режекторных фильтров. Пусть режекторный фильтр строится на базе последовательного колебательного контура. Схема такого фильтра приведена на рис. 2, а.



■ **Рис. 2.** Схема режекторного (а) и полосового (б) фильтров
 ■ **Fig. 2.** Scheme of the notch (а) and band (б) filters

Опуская промежуточные выкладки, запишем частотную передаточную функцию и дифференциальное уравнение данного фильтра в виде

$$W(j\omega) = [(j\omega)^2 LC + j\omega rC + 1] / [(j\omega)^2 LC + j\omega C(r + R) + 1];$$

$$\begin{aligned} LC \frac{d^2 y(t)}{dt^2} + C(R+r) \frac{dy(t)}{dt} + y(t) = \\ = LC \frac{d^2 x(t)}{dt^2} + Cr \frac{dx(t)}{dt} + x(t), \end{aligned}$$

где R — сопротивление входного резистора.

Разностное уравнение и АЧХ режекторного фильтра записываются следующим образом:

$$\begin{aligned} y[n] = a_0 x[n] + a_1 x[n-1] + a_2 x[n-2] + \\ + b_1 y[n-1] - b_2 y[n-2]; \end{aligned}$$

$$W(\omega) = \sqrt{\frac{(a_0 - a_1 \cos \omega T + a_2 \cos 2\omega T)^2 + (a_1 \sin \omega T - a_2 \sin 2\omega T)^2}{(1 - b_1 \cos \omega T + b_2 \cos 2\omega T)^2 + (b_1 \sin \omega T - b_2 \sin 2\omega T)^2}},$$

где

$$a_0 = \left(\frac{LC}{T^2} + \frac{rC}{T} + 1 \right) / \left(\frac{LC}{T^2} + \frac{C(r+R)}{T} + 1 \right);$$

$$a_1 = \left(\frac{2LC}{T^2} + \frac{rC}{T} \right) / \left(\frac{LC}{T^2} + \frac{C(r+R)}{T} + 1 \right);$$

$$a_2 = \frac{LC}{T^2} / \left(\frac{LC}{T^2} + \frac{C(r+R)}{T} + 1 \right);$$

$$b_1 = \left(\frac{2LC}{T^2} + \frac{C(R+r)}{T} \right) / \left(\frac{LC}{T^2} + \frac{C(R+r)}{T} + 1 \right);$$

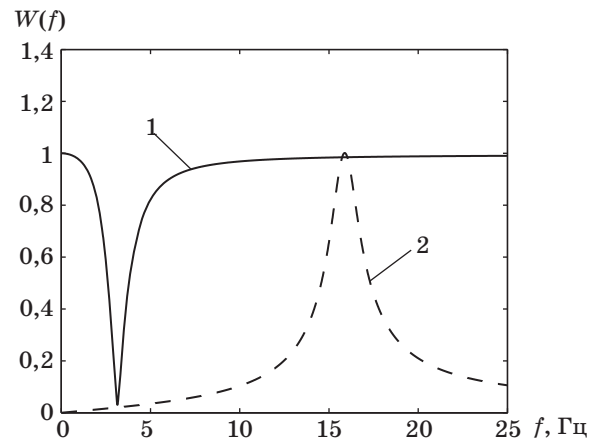
$$b_2 = \frac{LC^{-1}}{T^2} / \left(\frac{LC}{T^2} + \frac{C(R+r)}{T} + 1 \right).$$

Амплитудно-частотная характеристика режекторного фильтра для весовых коэффициентов $a_0 = 0,9877$, $a_1 = 1,9745$, $a_2 = 0,9873$, $b_1 = 1,9869$ и $b_2 = 0,9873$ показана на рис. 3 сплошной линией. Из характеристики видно, что на частоте режекции коэффициент передачи практически равен -30 дБ.

5. Селективный фильтр.

В основу схемотехнического решения селективного фильтра положен последовательный колебательный фильтр (рис. 2, б).

В рассматриваемом случае частотная передаточная функция, дифференциальное уравнение, разностное уравнение и АЧХ фильтра имеют вид



■ Рис. 3. Амплитудно-частотные характеристики режекторного 1 и селективного 2 фильтра

■ Fig. 3. AFC of the notch 1 and selective 2 filter

$$W(j\omega) = j\omega C / [(j\omega)^2 LC + j\omega rC + 1];$$

$$LC \frac{d^2 y(t)}{dt^2} + Cr \frac{dy(t)}{dt} + y(t) = C \frac{dx(t)}{dt};$$

$$y[n] = a(x[n] + x[n-1]) + b_1 y[n-1] - b_2 y[n-2];$$

$$W(\omega) = \sqrt{\frac{(a - a \cos \omega T)^2 + (a \sin \omega T)^2}{(1 - b_1 \cos \omega T + b_2 \cos 2\omega T)^2 + (b_1 \sin \omega T - b_2 \sin 2\omega T)^2}},$$

где $a = \frac{C}{T}$; $b_1 = \left(\frac{2LC}{T^2} + \frac{Cr}{T} \right) / \left(\frac{LC}{T^2} + \frac{Cr}{T} + 1 \right)$; $b_2 = \frac{LC}{T^2}$.

Амплитудно-частотная характеристика селективного фильтра для весовых коэффициентов $a = 0,0099$, $b_1 = 1,9802$ и $b_2 = 0,9901$ показана на рис. 3 пунктирной линией. Можно отметить достаточно хорошее совпадение АЧХ непрерывного и дискретного фильтров.

Заключение

Разработанная методика инвариантных дифференциальных уравнений позволяет синтезировать разнообразные дискретные линейные системы: фильтры нижних и верхних частот, колебательные звенья, режекторные и селективные фильтры — на основе известных дифференциальных уравнений непрерывных фильтров-аналогов с использованием разностных уравнений. Частотные свойства синтезированных дискретных фильтров при правильном выборе периода дискретности практически совпадают с частотными свойствами соответствующих непрерывных фильтров.

Литература

1. Воробьев С. Н. Цифровая обработка сигналов. — М.: Академия, 2013. — 318 с.
2. Голд Б., Рейдер Ч. Цифровая обработка сигналов. — М.: Сов. радио, 1973. — 367 с.
3. Сергиенко А. Б. Цифровая обработка сигналов. — СПб.: БХВ-Петербург, 2015. — 756 с.
4. Зиятдинов С. И. Анализ линейных систем на основе переходных характеристик // Информационно-управляющие системы. 2016. № 2. С. 104–106. doi:10.15217/issn1684-8853.2016.2.104
5. Зиятдинов С. И. Синтез рекурсивных дискретных фильтров во временной области // Изв. вузов. Радиоэлектроника. 2016. № 3. С. 3–6.
6. Зиятдинов С. И., Аграновский А. В., Осипов Л. А. Синтез комплексного фильтра с заданной передаточной функцией // Изв. вузов. Приборостроение. 2016. Т. 59. № 7. С. 542–546.
7. Зиятдинов С. И. Синтез нерекурсивных дискретных фильтров во временной области // Информационно-управляющие системы. 2016. № 5. С. 98–101. doi:10.15217/issn1684-8853.2016.5.98
8. Оппенгейм А., Шафер Р. Цифровая обработка сигналов. — М.: Техносфера, 2006. — 855 с.
9. Зиятдинов С. И., Соколова Ю. В. Анализ комплексных фильтров на основе переходных характеристик // Изв. вузов. Приборостроение. 2017. Т. 60. № 7. С. 641–646.
10. Зиятдинов С. И., Соколова Ю. В. Импульсные характеристики комплексных фильтров // Информационно-управляющие системы. 2017. № 4. С. 111–114. doi:10.15217/issn1684-8853.2017.4.111
11. Зиятдинов С. И., Соколова Ю. В. Синтез комплексных дискретных фильтров // Изв. вузов России. Радиоэлектроника. 2017. № 4. С. 12–19.
12. Oppenheim A., Schaffer R. Discrete-Time Signal Processing. — Englewood Cliffs, NJ: Prentice Hall, 1989. — 572 p.
13. Bellanger M. Digital Processing of Signals. Theory and Practice. — N. Y.: John Wiley and Sons, 1984. — 421 p.
14. Jackson L. Digital Filters and Signal Processing. Second ed. — Kluwer Academic Publishers, 1989. — 462 p.
15. Гоноровский И. С. Радиотехнические цепи и сигналы. — М.: Радио и связь, 1986. — 512 с.
16. Крук А. Е., Осипов Л. А. Синтез нелинейных импульсных систем управления при случайных воздействиях // Информационно-управляющие системы. 2012. № 3. С. 33–36.

UDC 621.391

doi:10.15217/issn1684-8853.2018.3.10

Synthesis of Discrete Filters by Invariant Differential Equations

Ziatdinov S. I.^a, Dr. Sc., Tech., ProfessorOsipov L. A.^a, Dr. Sc., Tech., ProfessorSokolova Y. V.^a, Master Student^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: When creating such information processing systems as Doppler speed and acceleration meters, moving target selection systems, radio altimeters with frequency modulation of the carrier signal, or coordinated devices for detecting and evaluating the parameters of objects, the developers use a variety of filters synthesized by the methods of bilinear z transformation and invariant impulse or transient characteristics. **Purpose:** Creating a technique for synthesizing discrete filters using differential equations of their counterpart continuous filters. **Results:** In the general form, in order to derive difference equations describing the operation of the discrete filters you synthesize, the differential equation derivatives are replaced by their counterparts in the form of difference equations. The weight coefficients for filters of various orders are listed in a table. The calculation has shown that when the discreteness period is chosen properly, the frequency characteristics of the synthesized filters practically coincide with those of the corresponding continuous filters. As an example, difference equations for low and high frequency filters, oscillatory links, notch filters and selective filters were obtained on the basis of known differential equations of the counterpart continuous filters which can be used in measuring the speed, range or angular position of an object. **Practical relevance:** The proposed technique allows you to create a variety of linear systems, such as low or high frequency filters, oscillatory links, notch or selective filters used in various information processing systems.

Keywords — Filter, Differential Equation, Difference Equation, Frequency Transfer Function.

Citation: Ziatdinov S. I., Osipov L. A., Sokolova Y. V. Synthesis of Discrete Filters by Invariant Differential Equations. *Информационно-управляющие системы* [Information and Control Systems], 2018, no. 3, pp. 10–16 (In Russian). doi:10.15217/issn1684-8853.2018.3.10

Reference

1. Vorobyov S. N. *Tsifrovaia obrabotka signalov* [Digital Signal Processing]. Moscow, Akademiia Publ., 2013. 318 p. (In Russian).
2. Gold B., Reider C. *Tsifrovaya obrabotka signalov* [Digital Signal Processing]. Moscow, Sovetskoe radio Publ., 1973. 367 p. (In Russian).
3. Sergienko A. B. *Tsifrovaya obrabotka signalov* [Digital Signal Processing]. Saint-Petersburg, BKhV-Peterburg Publ., 2015. 756 p. (In Russian).
4. Ziatdinov S. I. Analysis of Linear Systems on the base of Transient Characteristics. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 2, pp. 104–106 (In Russian). doi:10.15217/issn1684-8853.2016.2.104
5. Ziatdinov S. I. Synthesis of Recursive Discrete Filters in the Time Domain. *Izvestiia vysshikh uchebnykh zavedenii. Radioelektronika*, 2016, no. 3, pp. 3–6 (In Russian).
6. Ziatdinov S. I., Agranovsky A. V., Osipov L. A. Synthesis of Complex Filter with Specified Transfer Function. *Izvestiia vysshikh uchebnykh zavedenii. Priborostroenie* [Journal of Instrument Engineering], 2016, vol. 59, no. 7, pp. 542–546 (In Russian).
7. Ziatdinov S. I. Synthesis of Non-Recursive Discrete Filters in Temporal Range. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 5, pp. 98–101 (In Russian). doi:10.15217/issn1684-8853.2016.5.98
8. Oppenheim A., Shafer R. *Tsifrovaya obrabotka signalov* [Digital Signal Processing]. Moscow, Tekhnosfera Publ., 2006. 855 p. (In Russian).
9. Ziatdinov S. I., Sokolova Yu. V. Analysis of Complex Filters with the Use of Transition Characteristics. *Izvestiia vysshikh uchebnykh zavedenii. Priborostroenie* [Journal of Instrument Engineering], 2017, vol. 60, no. 7, pp. 641–646 (In Russian).
10. Ziatdinov S. I., Sokolova Yu. V. Impulse Characteristics of Complex Filters. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 4, pp. 111–114 (In Russian). doi:10.15217/issn1684-8853.2017.4.111
11. Ziatdinov S. I., Sokolova Yu. V. Synthesis of Complex Discrete Filters. *Izvestiia vysshikh uchebnykh zavedenii. Radioelektronika*, 2017, no. 4, pp. 12–19 (In Russian).
12. Oppenheim A., Schafer R. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ, Prentice Hall, 1989. 572 p.
13. Bellanger M. *Digital Processing of Signals. Theory and Practice*. New York, NJ, John Wiley and Sons, 1984. 421 p.
14. Jackson L. *Digital Filters and Signal Processing*. Second ed. Kluwer Academic Publishers, 1989. 462 p.
15. Honorovsky I. S. *Radiotekhnicheskie tsepi i signaly* [Radio Engineering Circuits and Signals]. Moscow, Radio i svyaz' Publ., 1986. 512 p. (In Russian).
16. Kruk A. E., Osipov L. A. Synthesis of Continuous Nonlinear Control Systems with Random Disturbances. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 3, pp. 33–36 (In Russian).

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию. Рукописи не возвращаются.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

УДК 621.38

doi:10.15217/issn1684-8853.2018.3.17

ОПТИМАЛЬНОЕ ТЕРМИНАЛЬНОЕ ДИАГНОСТИРОВАНИЕ УПРАВЛЯЕМЫХ ДИНАМИЧЕСКИХ СИСТЕМ

Г. С. Бритов^а, канд. техн. наук, доцент, britovgs@gmail.com

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Постановка проблемы: опубликованные работы по терминальному диагностированию динамических систем посвящены описанию методов расчета специальных, тестовых, сигналов, которые обеспечивают естественное движение системы, например в заданных границах. Такое движение системы осуществляется с помощью рассчитанного терминального управления. Известно, что существует множество вариантов терминального управления, которое минимизирует заданный критерий и приводит к движению системы по оптимальной траектории. **Цель:** разработка методов расчета оптимального терминального управления линейными дискретными динамическими системами с использованием в качестве критерия минимума энергетических затрат. **Методы:** теория оптимального управления дискретными системами, позволившая построить формулы получения оптимального управления для движения системы в заданных границах с минимальными энергетическими затратами. **Результаты:** разработаны методы расчета оптимального терминального управления при заданных и свободных начальных условиях системы, основанные на положениях теории оптимального управления для задач Лагранжа и Больца. Описан технологический процесс оптимального терминального диагностирования дискретных динамических систем. Полученные результаты могут позволить осуществить тестовое диагностирование дискретных динамических систем, отличающееся от известного диагностирования тем, что тестовое движение оказывается естественным, оптимальным движением системы в заданных границах, происходящее с минимальными энергетическими затратами. Теоретические предложения подтверждены компьютерным моделированием.

Ключевые слова — модель движения системы, линейная дискретная система, границы движения, оптимальное управление, оптимальное терминальное диагностирование.

Цитирование: Бритов Г. С. Оптимальное терминальное диагностирование управляемых динамических систем // Информационно-управляющие системы. 2018. № 3. С. 17–24. doi:10.15217/issn1684-8853.2018.3.17

Citation: Britov G. S. Optimal Terminal Diagnostics of Controlled Dynamic Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 17–24 (In Russian). doi:10.15217/issn1684-8853.2018.3.17

Введение

Тестовому диагностированию динамических систем, обеспечивающему контроль их технического состояния, посвящены, например, работы [1–8]. Терминальное диагностирование управляемых динамических систем, описанное в работах [9–11], можно отнести к группе методов контроля, в которых формируется сложный тестовый сигнал, но проверка диагностических признаков оказывается достаточно простой. В указанную группу входит, например, метод [12], в котором предлагается рассчитывать специальный, комплементарный, сигнал, обеспечивающий переход системы за заданное время из нулевых начальных условий опять в нулевые, конечные, условия. А в работе [13] предлагается подавать на вход системы аннулирующий сигнал, на который отсутствует выходная реакция системы. Примером такого сигнала является гармонический сигнал, частота которого равна передаточному нулю системы.

Следует обратить внимание также на структурное диагностирование. Здесь метод тестирования управляемых динамических систем основан

на использовании их структурных схем, которые описываются передаточными функциями. Этот метод приводит к новому понятию структурного диагностирования [14].

При терминальном диагностировании динамических систем осуществляют расчет терминального управления для обеспечения требуемого движения системы в заданных границах и проверки диагностических признаков. Последние определяются отклонением конечного состояния от заданной конечной границы движения системы. Известно, что существует множество вариантов расчета терминального управления. Ряд из них описан в работе [9]. Поэтому необходимо решать задачу выбора конкретного управления. Лучше всего искать оптимальное управление, которое минимизирует определенный критерий.

Целью статьи является постановка оптимизационной задачи получения управления линейной дискретной динамической системы, которое осуществляет переход системы из заданных начальных условий в требуемые конечные условия с минимальными энергетическими затратами.

Задачи оптимального управления

Рассмотрим те задачи оптимального управления линейными дискретными системами, которые целесообразно использовать при организации терминального диагностирования. Оно предполагает движение динамической системы в заданных границах. Поэтому необходимо определить математическую модель системы, которая должна быть положена в основу расчета оптимального управления. Для линейной дискретной системы используем модель движения в пространстве состояний согласно следующим рекуррентным уравнениям:

$$\mathbf{x}(t+1) = \mathbf{A} \cdot \mathbf{x}(t) + \mathbf{B} \cdot \mathbf{u}(t); \mathbf{x}(0) = \mathbf{x}_0;$$

$$\mathbf{x}(k) = \mathbf{x}_k, \mathbf{x}(t) \in R^n, \mathbf{u}(t) \in R^m.$$

Начальный вектор состояния \mathbf{x}_0 может быть установлен при тестировании системы. Конечный вектор состояния \mathbf{x}_k при числе шагов тестирования k не всегда принимается равным нулю для того, чтобы упростить получение диагностического признака. Проверка достижения системой в конце движения нулевого состояния не представляется слишком сложной задачей.

Известно, что существует множество управлений $\mathbf{u}(t)$, которые обеспечивают переход системы из известного начального состояния в требуемое конечное состояние. Поэтому вводится критерий, минимизация которого позволяет рассчитать единственное, оптимальное управление. При тестировании системы целесообразно рассчитать такой тестовый сигнал, который обладает минимальной энергией. Предлагается следующий критерий:

$$J(\mathbf{u}) = \sum_{t=0}^{k-1} \mathbf{u}(t)^T \cdot \mathbf{u}(t) \rightarrow \min.$$

Следовательно, возникает оптимизационная задача Лагранжа: найти решение рекуррентных уравнений, которое переходит из начального значения в конечное значение и минимизирует аддитивный, квадратичный критерий. Решением этой задачи будет оптимальное управление, имеющее минимальную энергию при переходе системы из начального состояния в конечное состояние по оптимальной траектории.

При тестировании может возникать ситуация, когда начальное условие неизвестно. Тогда следует рассмотреть другую оптимизационную задачу минимизации

$$J(\mathbf{x}, \mathbf{u}) = \mathbf{x}(0)^T \cdot \mathbf{x}(0) + \sum_{t=0}^{k-1} \mathbf{u}(t)^T \cdot \mathbf{u}(t) \rightarrow \min.$$

Теперь начальное условие будет рассчитано в результате решения так называемой задачи

Больца: найти решение рекуррентных уравнений, которое переходит из неизвестного начального значения в известное конечное значение при минимизации аддитивного, квадратичного критерия. Решением этой задачи будет оптимальное управление, имеющее минимальную энергию при переходе системы из рассчитанного начального состояния в известное конечное состояние по оптимальной траектории.

Методы решения поставленных задач как условно-экстремальных задач дифференциально-исчисления хорошо известны в математике. Рассмотрим применение этих методов для расчета оптимального терминального диагностирования.

Расчет терминального управления

Решение поставленной задачи Лагранжа основано на применении теоремы Ферма для функции Лагранжа. Для того чтобы с ней было удобнее работать, используется функция Гамильтона

$$H(\boldsymbol{\mu}, \mathbf{x}, \mathbf{u}) =$$

$$= \boldsymbol{\mu}^T(t+1)(\mathbf{A} \cdot \mathbf{x}(t) + \mathbf{B} \cdot \mathbf{u}(t)) - \mathbf{u}^T(t) \cdot \mathbf{u}(t).$$

Здесь вектор неопределенных множителей Лагранжа $\boldsymbol{\mu}(t)$ играет роль вектора переменных состояния сопряженной системы. Тогда вид оптимального уравнения следует из решения уравнения

$$\frac{\partial H}{\partial \mathbf{u}} = \boldsymbol{\mu}^T(t+1) \cdot \mathbf{B} - 2\mathbf{u}^T(t) = 0$$

и имеет вид

$$\mathbf{u}^0(\boldsymbol{\mu}) = 0,5\mathbf{B}^T \cdot \boldsymbol{\mu}(t+1).$$

Модифицируем функцию Гамильтона, подставляя полученную формулу оптимального управления:

$$H^0(\boldsymbol{\mu}, \mathbf{x}, \mathbf{u}) = \boldsymbol{\mu}^T(t+1) \cdot \mathbf{A} \cdot \mathbf{x}(t) + 0,25\mathbf{B} \cdot \mathbf{B}^T \boldsymbol{\mu}(t+1).$$

Тогда рекуррентные уравнения Эйлера, задающие необходимые условия экстремума, будут иметь вид

$$\mathbf{x}(t+1) = \left(\frac{\partial H}{\partial \boldsymbol{\mu}} \right)^T = \mathbf{A} \cdot \mathbf{x}(t) + 0,5\mathbf{B} \cdot \mathbf{B}^T \cdot \boldsymbol{\mu}(t+1);$$

$$\boldsymbol{\mu}(t) = \left(\frac{\partial H}{\partial \mathbf{x}} \right)^T = \mathbf{A}^T \cdot \boldsymbol{\mu}(t+1).$$

Эти уравнения должны решаться при следующих граничных условиях:

$$\mathbf{x}(0) = \mathbf{x}_0, \mathbf{x}(k) = 0.$$

Получена двухточечная граничная задача, которую можно решить аналитически. Для этого введем новую переменную, объединяющую векторы состояний исходной и сопряженной систем, и получим формулировку двухточечной граничной задачи

$$\mathbf{z}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\mu}(t) \end{bmatrix}; \quad \mathbf{z}(t+1) = \mathbf{A}_z \cdot \mathbf{z}(t);$$

$$\mathbf{B}_z \cdot \mathbf{z}(0) + \mathbf{C}_z \cdot \mathbf{z}(k) = \mathbf{d}_z.$$

Здесь использованы следующие матрицы:

$$\mathbf{A}_z = \begin{bmatrix} \mathbf{A} & 0,5\mathbf{B} \cdot \mathbf{B}^T \cdot (\mathbf{A}^T)^{-1} \\ \mathbf{0} & (\mathbf{A}^T)^{-1} \end{bmatrix};$$

$$\mathbf{B}_z = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}; \quad \mathbf{C}_z = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{E} & \mathbf{0} \end{bmatrix}; \quad \mathbf{d}_z = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{0} \end{bmatrix}.$$

Матрица \mathbf{E} является единичной матрицей. Расчет неизвестного начального условия $\mathbf{z}(0)$ можно произвести следующим образом:

$$\mathbf{z}(0) = (\mathbf{B}_z + \mathbf{C}_z \cdot \mathbf{A}_z^k)^{-1} \cdot \mathbf{d}_z.$$

Таким образом, двухточечная граничная задача приведена к задаче Коши, которую можно решить непосредственным образом. Пусть ее решение имеет вид пары

$$(\mathbf{x}^0(t), \boldsymbol{\mu}^0(t)).$$

Тогда оптимальное управление, имеющее минимальную энергию, будет функцией времени:

$$\mathbf{u}^0(t) = 0,5\mathbf{B}^T \cdot (\mathbf{A}^T)^{-1} \cdot \boldsymbol{\mu}^0(t).$$

Аналогичный подход используем для решения задачи Больца. Граничные условия для уравнений Эйлера имеют теперь вид

$$2\mathbf{x}(0) + (\mathbf{A}^T)^{-1} \cdot \boldsymbol{\mu}(0) = \mathbf{0}; \quad \mathbf{x}(k) = \mathbf{x}_k.$$

Конечные условия \mathbf{x}_k здесь не могут быть нулевыми, так как вычисление начальных условий $\mathbf{z}(0)$ по приведенной выше формуле, вообще говоря, будет невозможным. Матрицы для этой формулы

$$\mathbf{B}_z = \begin{bmatrix} 2\mathbf{E} & (\mathbf{A}^T)^{-1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}; \quad \mathbf{C}_z = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{E} & \mathbf{0} \end{bmatrix}; \quad \mathbf{d}_z = \begin{bmatrix} \mathbf{0} \\ \mathbf{x}_k \end{bmatrix}.$$

Рассчитав начальные условия $\mathbf{z}(0) = [\mathbf{x}(0); \boldsymbol{\mu}(0)]$, можно решить двухточечную задачу, так как она приведена к задаче Коши. Получим ре-

шение $\mathbf{z}^0(t) = [\mathbf{x}^0(t); \boldsymbol{\mu}^0(t)]$ и вычислим по приведенной выше формуле оптимальное управление $\mathbf{u}^0(t)$, которое осуществляет переход из рассчитанных начальных условий в заданные, ненулевые конечные условия. При этом переходе затрачивается минимальная энергия.

При назначении длительности диагностирования, которое определяется числом шагов k , необходимо учитывать следующее:

— за выбранное число шагов система должна проявить все необходимые характеристики свободного движения;

— при решении двухточечной граничной задачи возникает необходимость обращать матрицу, зависящую от числа шагов; возможно ее вырождение;

— чересчур большое число шагов увеличивает риск ложного срабатывания системы диагностирования.

Обратимся к процедуре оптимального терминального диагностирования.

Организация оптимального терминального диагностирования

Получив необходимые расчетные формулы, можно построить наглядное описание процесса терминального диагностирования. Для этого целесообразно использовать методологию стандарта IDEF3. С ее помощью создается сценарий исследуемого процесса. Он состоит из диаграмм вида PFDD (Process Flow Description Diagrams), в которых показано взаимодействие операций процесса.

В представленной на рис. 1 контекстной PFDD-диаграмме диагностирования управляемой динамической системы отмечены внешние сущности процесса:

— управляемая динамическая система, которая подвергается тестовому диагностированию с помощью подачи на ее вход рассчитанного терминального управления;

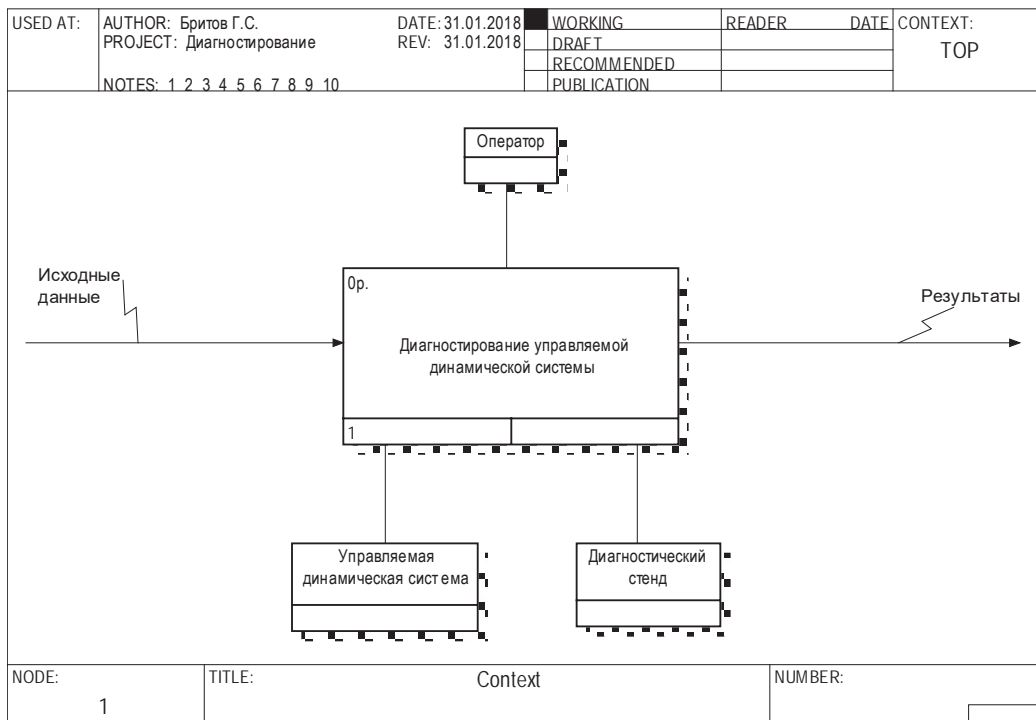
— диагностический стенд, на котором реализуется рассчитанная оптимальная траектория движения системы в пространстве состояний;

— оператор, принимающий решение по результатам тестового движения системы.

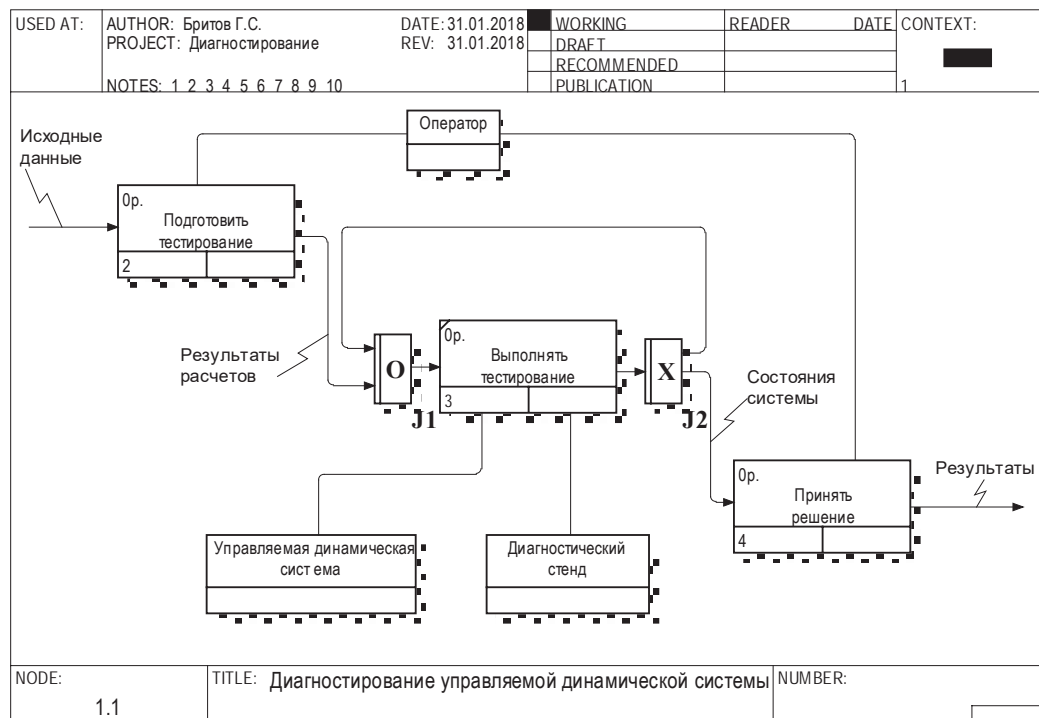
Входом в технологический процесс являются исходные данные, которые необходимы для получения модели системы с системными матрицами \mathbf{A} , \mathbf{B} , \mathbf{C} . Выходом процесса служат результаты, получаемые после выполнения тестирования системы.

Декомпозиция контекстной диаграммы приводит к PFDD-диаграмме (рис. 2), которая включает три достаточно крупные операции:

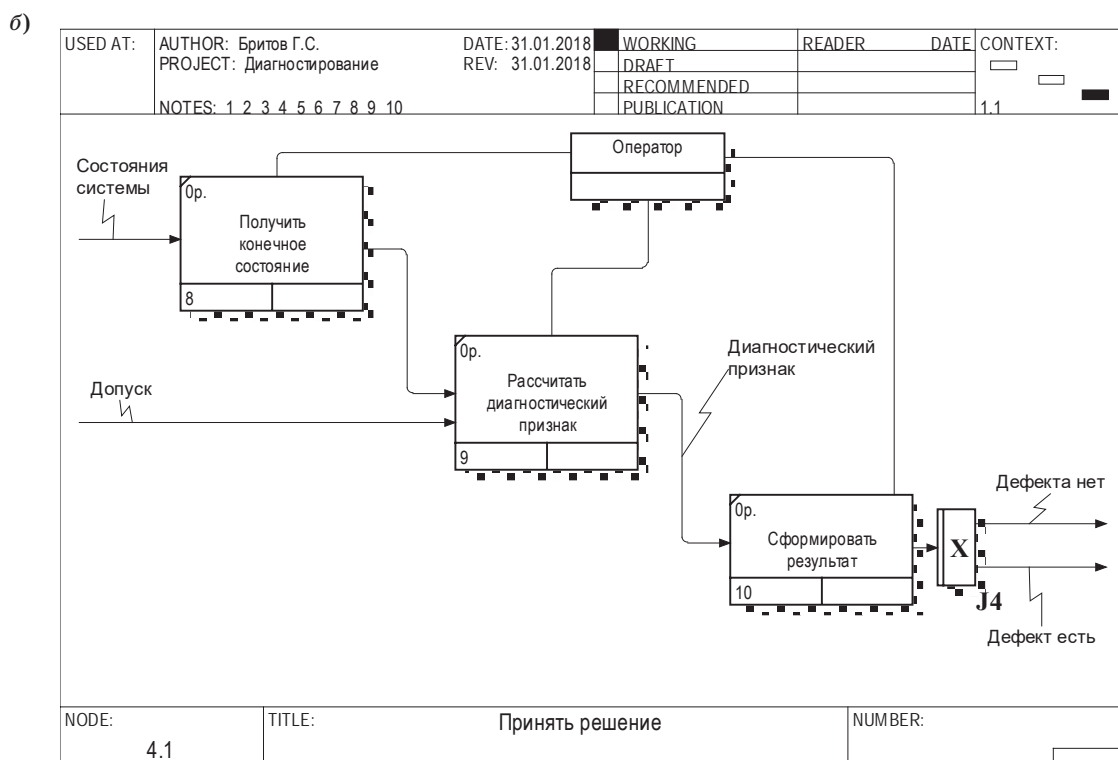
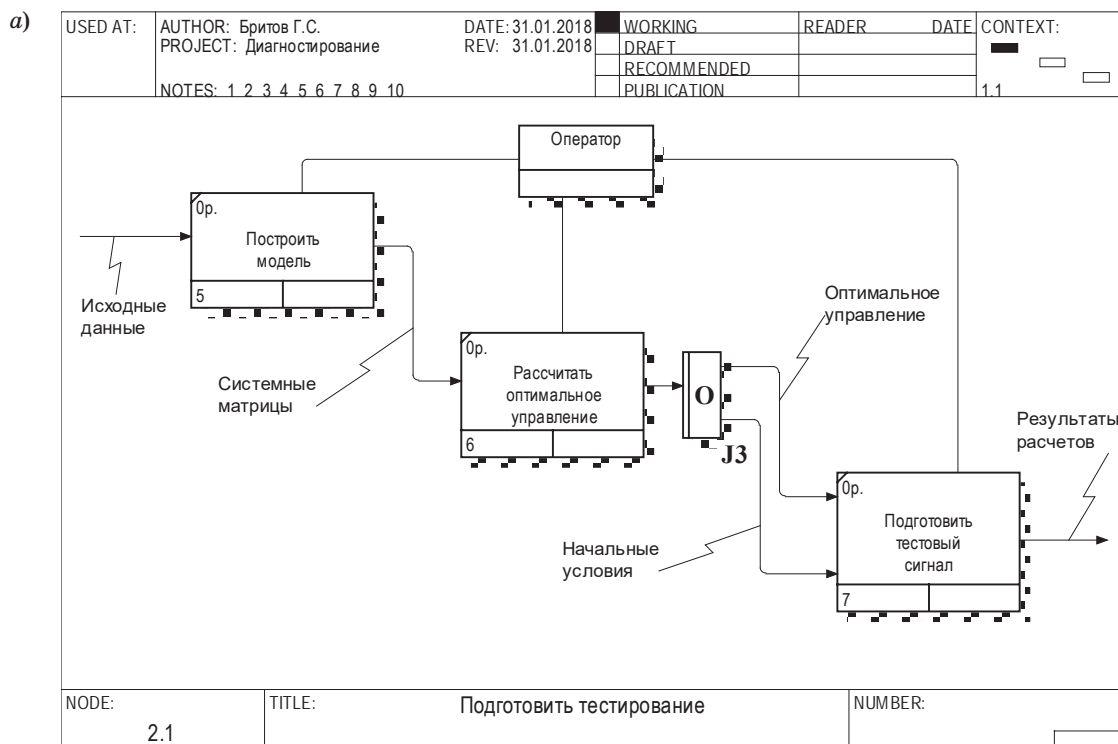
— подготовку оператором необходимых данных для построения модели тестируемой динамической системы;



■ **Рис. 1.** Контекстная PFDD-диаграмма процесса диагностирования управляемой динамической системы
 ■ **Fig. 1.** Context PFDD-diagram of the process of diagnosing of controlled dynamic systems



■ **Рис. 2.** PFDD-диаграмма второго уровня процесса диагностирования управляемой динамической системы
 ■ **Fig. 2.** PFDD-diagram of the second level of the process of diagnosing of controlled dynamic systems



■ **Рис. 3.** PFDD-диаграмма третьего уровня процесса диагностирования управляемой динамической системы: *a* — для подготовки тестирования; *b* — для принятия решения

■ **Fig. 3.** PFDD-third-level diagram of the process of diagnosing a managed dynamic system: *a* — for preparation of testing; *b* — for decision-making

— выполнение тестового движения динамической системы под действием подаваемого на ее вход рассчитанного оптимального терминального управления;

— принятие оператором решения о наличии или отсутствии дефекта после анализа по выходам системы результатов тестового движения.

Операция «Подготовить тестирование» требует декомпозиции, так как данная диаграмма является диаграммой только второго уровня дерева узлов. В PFDD-диаграмме третьего уровня декомпозиции указанной операции (рис. 3, а) участвуют три операции:

— построение модели системы с системными матрицами **A, B**;

— расчет оптимального управления системой;

— подготовка тестового сигнала, под действием которого система будет двигаться по оптимальной траектории.

Продолжая декомпозицию, получим PFDD-диаграмму третьего уровня декомпозиции операции «Принять решение» (рис. 3, б). Она включает операции, необходимые для получения результатов тестирования «Дефекта нет» или «Дефект есть».

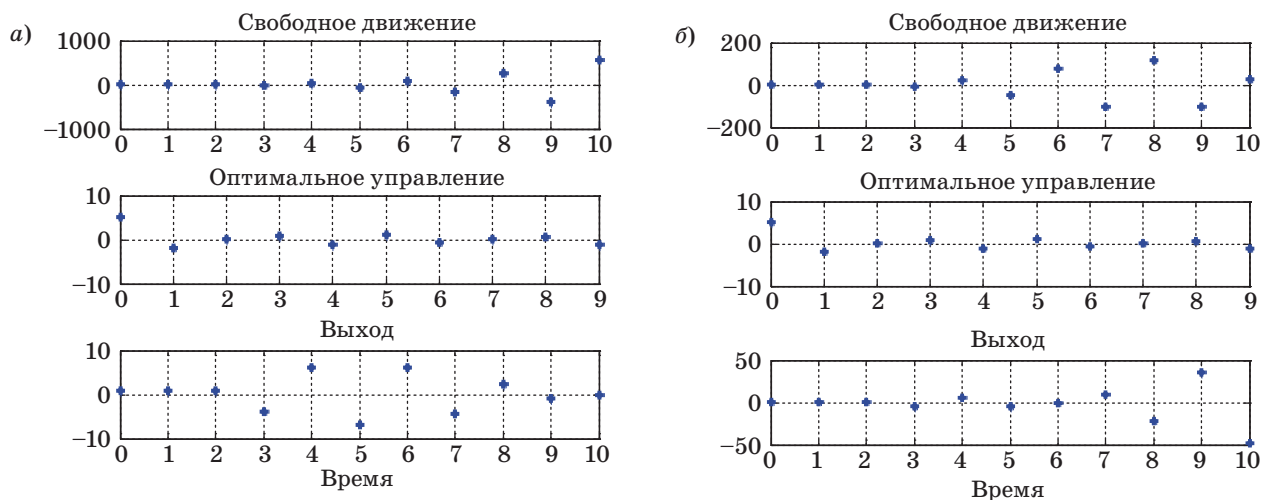
Декомпозицию технологической модели процесса оптимального терминального диагностирования управляемых динамических систем можно было бы продолжать. Однако эта работа выходит за рамки одной статьи.

Моделирование процесса диагностирования

В математическом пакете MatLab была написана программа-сценарий для моделирования процесса оптимального терминального диагно-

стирования управляемых динамических систем. Они могут быть описаны уравнениями состояний, рекуррентным уравнением заданного порядка или структурной схемой. Результатом моделирования являются графики свободного движения системы, оптимального управления, выхода, а также значение диагностического признака. Ниже эти результаты приведены для структурной схемы, включающей три неустойчивых звена первого порядка. Движение системы наблюдается с помощью выхода, которым служит выход последнего звена. Моделирование проведено для исправной системы (рис. 4, а) и системы с дефектом (рис. 4, б):

```
>> sc18
Ввод
N = 3
Введите вектор коэффициентов числителя 1-й ПФ = 1
Введите вектор коэффициентов знаменателя 1-й ПФ = [1 1.1]
W = 1
-----
s + 1.1
Введите вектор коэффициентов числителя 2-й ПФ = 1
Введите вектор коэффициентов знаменателя 2-й ПФ = [1 1.15]
W = 1
-----
s + 1.15
Введите вектор коэффициентов числителя 3-й ПФ = 1
Введите вектор коэффициентов знаменателя 3-й ПФ = [1 1.2]
W = 1
-----
s + 1.2
Wc = 1
-----
s^3 + 3.45 s^2 + 3.965 s + 1.518
Ввод столбца длиной 3 x 0 = [1;1;1]
k = 10
Ошибка в свободном коэффициенте = 0
Диагностический признак :
-1.1282e-11
Ошибка в свободном коэффициенте = 0.1
```



■ Рис. 4. Графики движения системы при отсутствии (а) и наличии (б) дефектов
 ■ Fig. 4. Graphs of system motion in the absence (а) and in the presence (б) of defects

Диагностический признак :
 -47.8226
 Конец
 >>

При отсутствии дефектов оптимальное терминальное управление приводит систему за 10 шагов к практически нулевому выходу. Об этом говорит диагностический признак, имеющий порядок -11 . При введении ошибки со значением $0,1$ оптимальное терминальное управление приводит систему за 10 шагов к выходу -50 на графике. Диагностический признак достигает значения -48 .

Таким образом, оптимальное терминальное диагностирование, обеспечивающее тестирование динамических систем, обнаруживает дефекты за счет появления ошибок в известных моделях движения системы.

Разработка системы терминального диагностирования потребует оценки качества процесса тестирования. В частности, следует определить чувствительность контроля, введя необходимую характеристику. Подробно эта задача рассматривалась в работах по функциональному диагностированию. Построение соответствующих формул для терминального диагностирования и анализ его чувствительности потребуют отдельной статьи.

Заключение

Оптимальное терминальное диагностирование продолжает специальное направление тестирования динамических систем, обеспечивающее контроль системы в естественном режиме движения ее в заданных границах. Тестовым сигналом

в данном случае является рассчитанное заранее терминальное управление. При выборе границ движения системы равными нулю диагностические признаки оказываются достаточно простыми. Приведенные в опубликованных статьях расчеты диагностического терминального управления относятся к дискретным системам. Поэтому расчетные формулы оптимального терминального управления обеспечивают самое лучшее управление по минимуму энергетических затрат. Этим получаемое тестовое движение лучше других, так как известно, чем характеризуется оптимальная траектория.

Рассмотренное оптимальное терминальное диагностирование динамических систем ограничивалось контролем, с помощью которого обнаруживалось появление дефекта. Следующей является задача локализации дефекта. Поскольку в основу терминальной диагностики положены математические модели проверяемых систем, то и локализацию дефектов целесообразно проводить с точностью до элементов этих моделей. Так, при использовании уравнений состояния многомерной системы можно выполнить локализацию дефекта с точностью до одной из матриц этих уравнений. При использовании передаточной функции скалярной системы дефект локализуется либо в числителе, либо в знаменателе этой функции.

В работах по функциональному диагностированию описан другой метод локализации дефекта. Он связан с использованием принципа диагностирования по годографам дефектов. Этот принцип требует большой подготовительной работы по созданию годографов.

Работа поддержана грантом РФФИ № 17-08-00244.

Литература

1. Мальцев Г. Н., Якимов В. Л. Достоверность многоэтапного контроля технического состояния объектов испытаний // Информационно-управляющие системы. 2018. № 1. С. 49–57. doi:10.15217/issn1684-8853.2018.1.49
2. Копкин Е. В., Бородько Д. Н., Шапова К. Е. Алгоритм построения квазиоптимальной гибкой программы анализа технического состояния объекта // Информационно-управляющие системы. 2017. № 1. С. 31–39. doi:10.15217/issn1684-8853.2017.1.31
3. Abdeerrahman A., Cerny E. Worst Tolerance Analysis and CLP-based Multifrequency Test Generation for Analogue Circuits // IEEE Transactions on Computer-Aided Design. 1999. Vol. 18. P. 332–345.
4. Lindermeir W., Graeb H., Antreich K. Analogue Testing by Characteristic Observation Inference // IEEE Transactions on Computer-Aided Design. 1999. Vol. 18. P. 1353–1368.
5. Kadim H. J. Minimal Transient Modes for Faults Detection in Analogue VLSI Circuits // Radioelectronics & Informatics. 2003. N 3. P. 82–86.
6. Мироновский Л. А., Слаев В. А. Синтез оптимальных тестовых сигналов как решение обобщенной задачи Булгакова // Автоматика и телемеханика. 2002. № 4. С. 55–66.
7. Мироновский Л. А. Тестовый контроль линейных систем управления // Информационно-управляющие системы на железнодорожном транспорте. 2005. № 5. С. 3–8.
8. Воронин В. В., Шалобанов С. С. Диагностирование непрерывных динамических систем методом пробных отклонений параметров модели // Информатика и системы управления. 2010. № 1(23). С. 121–127.
9. Бритов Г. С. Терминальное диагностирование дискретных динамических систем // Информационно-

управляющие системы. 2017. № 4. С. 18–24.
doi:10.15217/issn1684-8853.2017.4.18

10. Бритов Г. С. Диагностирование динамических систем на основе терминального управления // Системный анализ и логистика. 2016. Вып. 2(13). С. 3–12.
11. Бритов Г. С. Системный анализ терминального диагностирования динамических объектов // Системный анализ и логистика. 2017. Вып. 2(15). С. 3–11.
12. Мироновский Л. А. Диагностирование линейных систем методом комплементарного сигнала // При-

боры и системы. Управление, контроль, диагностика. 2002. № 5. С. 52–57.

13. Мироновский Л. А. Диагностирование систем управления методом аннулирующего сигнала // Информационно-управляющие системы на железнодорожном транспорте. 2001. № 5. С. 3–7.
14. Мироновский Л. А., Соловьева Т. Н. Структурное диагностирование управляемых динамических систем // Тр. XII Всерос. совещания по проблемам управления. 2014. С. 7640–7647.

UDC 621.38

doi:10.15217/issn1684-8853.2018.3.17

Optimal Terminal Diagnostics of Controlled Dynamic Systems

Britov G. S.^a, PhD, Tech., Associate Professor, britovgs@gmail.com

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: The published works on terminal diagnostics of dynamic systems discuss the methods of calculating special test signals which provide natural motion of the system, for example, within specified bounds. Such a motion is accomplished with a specially calculated terminal control. There are many options for terminal control. Therefore, it is necessary to pose and solve the problem of optimal control which would minimize the given criterion and provide that the system moves along the optimal trajectory. **Purpose:** Developing methods for calculating the optimal terminal control by linear discrete dynamic systems, using the lowest energy cost as a criterion. **Methods:** The theory of optimal control of discrete systems is used, which allows you to build formulas for obtaining optimal control over the system motion within specified boundaries with the lowest energy cost. **Results:** Methods have been developed for calculating optimal terminal control under given or free initial conditions of the system, based on the theory of optimal control for Lagrange and Bolz problems. A technological process has been described for optimal terminal diagnostics of discrete dynamic systems. The obtained results make it possible to perform test diagnostics of discrete dynamic systems, which differs from the known diagnostics in that the test motion of the system turns out to be natural and optimal within the given boundaries and have the lowest energy cost. Computer simulation confirmed the theory.

Keywords — System Motion Model, Linear Discrete System, Motion Boundaries, Optimal Control, Optimal Terminal Diagnostics.

Citation: Britov G. S. Optimal Terminal Diagnostics of Controlled Dynamic Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 17–24 (In Russian). doi:10.15217/issn1684-8853.2018.3.17

References

1. Mal'tsev G. N., Iakimov V. L. Reliability of Multi-Stage Control over Technical Condition of Tested Objects. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 49–57 (In Russian). doi:10.15217/issn1684-8853.2018.1.49
2. Kopkin E. V., Borod'ko D. N., Shapova K. E. Algorithm for Constructing a Quasi-Optimal Flexible Program for Analysis of Technical State of an Object. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 1, pp. 31–39 (In Russian). doi:10.15217/issn1684-8853.2017.1.31
3. Abdeerrahman A., Cerny E. Worst Tolerance Analysis and CLP-based Multifrequency Test Generation for Analogue Circuits. *IEEE Transactions on Computer-Aided Design*, 1999, vol. 18, pp. 332–345.
4. Lindermeir W., Graeb H., Antreich K. Analogue Testing by Characteristic Observation Inference. *IEEE Transactions on Computer-Aided Design*, 1999, vol. 18, pp. 1353–1368.
5. Kadim H. J. Minimal Transient Modes for Faults Detection in Analogue VLSI Circuits. *Radioelectronics & Informatics*, 2003, no. 3, pp. 82–86.
6. Mironovskii L. A., Slaev V. A. Synthesis Optimal Test Signals is Solution of Bulgakov Task. *Avtomatika i telemekhanika* [Automation and Remote Control], 2002, no. 4, pp. 82–86 (In Russian).
7. Mironovskii L. A. Test Checking of Linear Control Systems. *Informatsionno-upravliaiushchie sistemy na zheleznodorozhnom transporte* [Information and Management Systems of Railway Transport], 2005, no. 5, pp. 3–8 (In Russian).
8. Voronin V. V., Shalobanov S. S. Diagnosing Continuous Dynamic Systems by the Method of Test Deviations of Parameters of the Model. *Informatika i sistemy upravleniia* [Information Science and Control Systems], 2010, no. 1(23), pp. 121–127 (In Russian).
9. Britov G. S. Terminal Diagnostics of Discrete Dynamic Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 4, pp. 18–24 (In Russian). doi:10.15217/issn1684-8853.2017.4.18
10. Britov G. S. Diagnostics of Dynamic Systems based on Terminal Management. *Sistemnyi analiz i logistika* [The System Analysis and Logistics], 2016, iss. 2(13), pp. 3–12 (In Russian).
11. Britov G. S. System Analysis of Terminal Diagnostics of Dynamic Objects. *Sistemnyi analiz i logistika* [The System Analysis and Logistics], 2017, iss. 2(15), pp. 3–11 (In Russian).
12. Mironovskii L. A. Diagnosis of Linear Systems by Complementary Signal Method. *Pribory i sistemy. Upravlenie, kontrol', diagnostika*, 2002, no. 5, pp. 52–57 (In Russian).
13. Mironovskii L. A. Diagnosis Control Systems by Invalidating Signal Method. *Informatsionno-upravliaiushchie sistemy na zheleznodorozhnom transporte* [Information and Management Systems of Railway Transport], 2001, no. 5, pp. 3–7 (In Russian).
14. Mironovskii L. A., Solov'eva T. N. Structural Diagnosis of Controlled Dynamic Systems. *Trudy XII Vserossiiskogo soveshchaniia po problemam upravleniia* [Proc. of the 12th All-Russian Meeting on Governance], Moscow, IPU RAN Publ., 2014, pp. 7640–7647 (In Russian).

СТЕРЕОФОТОМЕТРИЧЕСКОЕ ТРЕХМЕРНОЕ СКАНИРОВАНИЕ С ИСПОЛЬЗОВАНИЕМ ИМПУЛЬСНОГО ОСВЕТИТЕЛЯ

В. А. Кузнецов^а, ассистент, k.avk-c@mail.ru

М. Ю. Людаев^б, инженер-программист, mike_ld@mail.ru

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения,

Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

^бРоссийский институт мощного радиостроения, 11-я линия В. О., 66, Санкт-Петербург, 199178, РФ

Постановка проблемы: метод трехмерного сканирования на основе модели отражения света поверхностью известен, но на данный момент технологий, осуществляющих по нему трехмерное сканирование, практически нет. Это обусловлено сложностью вычислений координат точек поверхности, так как в первую очередь вычисляется ориентация поверхности, неравномерное распределение погрешности которой усложняет процесс получения трехмерной модели. Помимо этого, существует ряд требований к условиям, в которых можно проводить сканирование объекта, значительно ограничивающих применение метода и увеличивающих цену сканирования. **Цель:** разработка опытного образца трехмерного сканера с использованием доступного фотооборудования; определение возможностей стереофотометрического метода для случая, когда в качестве источника света применяется импульсный осветитель. **Результаты:** получена и опробована опытная установка, позволяющая осуществлять трехмерное сканирование фрагментов объекта стереофотометрическим методом; приведены настройки аппаратуры, условия проведения процесса сканирования и характеристики установки. Для опытной установки плотность точек по осям x , y равна 0,103 мм, размер области сканирования $0,7 \times 0,7 \text{ м}^2$; для тестовой фигуры определена погрешность, обусловленная шумом матрицы фотокамеры, наличием постороннего освещения и неравномерностью освещения импульсными осветителями и в среднем составляющая 0,062 мм; среднеквадратическое отклонение по всей поверхности составляет 0,0275 мм. **Практическая значимость:** полученные результаты демонстрируют ожидаемую высокую точность сканирования стереофотометрическим методом, при этом для реализации установки использовалось доступное студийное оборудование. Специальная обработка помещения также не проводилась, что делает стереофотометрический метод более доступным для использования.

Ключевые слова — трехмерное сканирование, стереофотометрия, обработка изображений, фотографирование.

Цитирование: Кузнецов В. А., Людаев М. Ю. Стереофотометрическое трехмерное сканирование с использованием импульсного осветителя // Информационно-управляющие системы. 2018. № 3. С. 25–33. doi:10.15217/issn1684-8853.2018.3.25

Citation: Kuznetsov V. A., Liudaev M. U. Photometric Stereo Scanner with Flash Light Lamp. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 25–33 (In Russian). doi:10.15217/issn1684-8853.2018.3.25

Введение

В работах [1–9] подробно описаны методы, являющиеся базовыми для получения трехмерных моделей поверхности по набору изображений этой поверхности при различном освещении.

В настоящей статье представлена конфигурация опытной установки, которая в сочетании с разработанными алгоритмами [10–12] позволяет осуществлять сканирование практически в любом помещении.

Достоинством предлагаемой конфигурации является широкая распространенность используемого оборудования, что обеспечивает низкую стоимость и доступность трехмерного сканирования. Для фотографирования объекта человеку, владеющему навыками студийной фотосъемки, не требуется дополнительная подготовка или обучение.

Основой для вычисления ориентации поверхности является закон диффузного отражения

$$i_k = K(Ln), \quad (1)$$

где i_k — яркость пикселя на изображении; K — коэффициент отражения; L — вектор освещения; n — вектор нормали к поверхности. Если вектор освещения известен, то остаются три неизвестных параметра: коэффициент отражения K и два значения градиента p и q , определяющих положение вектора нормали к поверхности. Следовательно, для того чтобы вычислить неизвестные величины, необходимо не меньше трех измерений яркости i_k . Незвестные параметры в этом случае будут решены системой трех уравнений (1).

Использование трех изображений сильно ограничивает возможности сканирования, так как реконструируется только часть объекта, освещенная тремя линейно независимыми источниками света. Получение дополнительных снимков при другом расположении освещения позволяет увеличить видимую, доступную для сканирования область объекта. Четыре снимка [4, 5] позволяют исключить из рассмотрения один снимок или же использовать наличие затенения в нем в качестве ограничения величины градиента, избегая ошибок вычисления.

Конфигурация опытной установки

Предлагаемая конфигурация освещения (рис. 1) состоит из двух пар смежных источников света L_1, L_1^1 и L_2, L_2^1 в основных плоскостях и одного дополнительного источника (направление света L_0 от дополнительного источника совпадает с направлением оптической оси камеры). За ось z принимается направление наблюдения. Все источники света находятся под одним углом к оси z и в вычислениях принимаются за условно точечные в силу большого расстояния до объекта в сравнении с размером объекта и источника света.

При такой конфигурации объект в каждой точке освещен по крайней мере тремя источниками света. Наличие дополнительного источника света, особенно из-за отсутствия затенений на снимке I_0 , позволяет упростить и более точно осуществить сегментацию затенений на изображениях [5, 10, 13–15].

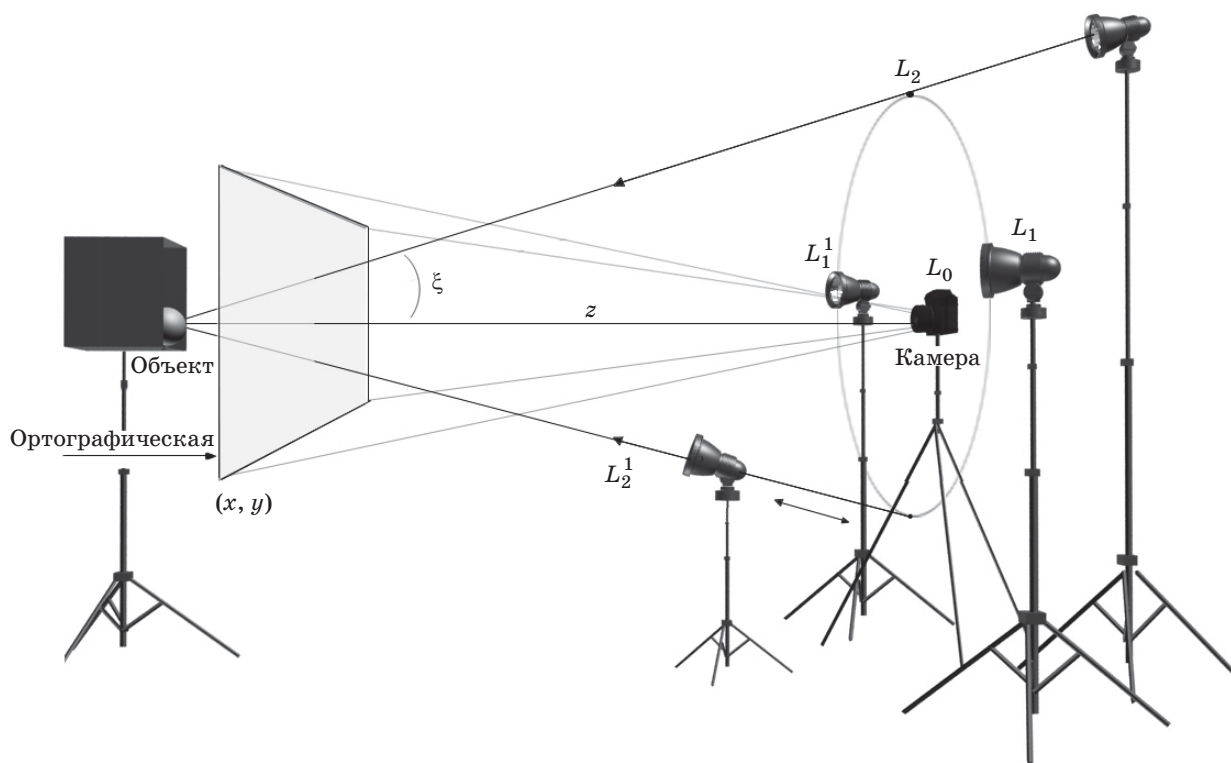
Конфигурация определяется углом ξ между направлением освещения и осью z , а также расстоянием l от камеры до объекта.

Оптимальное значение угла ξ лежит в интервале $[25^\circ; 35^\circ]$, что подтверждается результатами вычисления [16] градиента сферы и овала Кассини, полученными при компьютерном моделировании (табл. 1, ошибка представлена в ус-

ловной величине $[px]$, равной расстоянию между соседними пикселями). Рост ошибки для овала Кассини обусловлен ростом области пересечения затенений смежных снимков, что делает точное вычисление всего поля градиента невозможным, в то время как поле градиента сферы может быть вычислено полностью при любых значениях.

Из-за ограниченных размеров помещения не всегда могут быть выбраны оптимальные параметры, необходимые размеры помещения зависят от расстояния до объекта l и угла ξ : минимальная высота и ширина помещения равны $1,44l \cdot \text{tg } \xi$, т. е. при $\xi = 30^\circ$ выбор расстояния l более 4 м требует помещение высотой более 3,2 м.

Объект необходимо расположить на значительном удалении от источников света и камеры. Это позволит считать лучи, исходящие от источника, параллельными и упростить вычисления. При сканировании объекта до 10 см в любом измерении на расстоянии 4 м от источника света максимальная ошибка при вычислении ориентации поверхности составит 0,05 %, на расстоянии 2 м — 0,09 %. При получении снимков фотокамерой происходит перспективная проекция объектов на плоскость изображения. Если же размеры объекта много меньше по сравнению с расстоянием от камеры до объекта, то в процессе вычисления вместо перспективного проецирования может быть принято ортографическое [6], что оз-



■ Рис. 1. Схема опытной установки и проецирования
 ■ Fig. 1. The pilot scanner and projection model

■ **Таблица 1.** Средняя абсолютная погрешность вычисления приращений для сферы и овала Кассини, полученная при компьютерном моделировании

■ **Table 1.** Average gradient absolute calculation error for sphere and surface based on the oval of Cassini, obtained during computer modeling

ξ , град	Сфера, рх	Овал, рх
5	3,51	3,81
10	1,82	2,02
15	1,65	1,71
20	1,35	1,45
25	1,11	1,21
30	1,0	1,12
35	0,85	0,93
40	0,80	0,80
45	0,74	1,01
50	0,67	1,72
55	0,66	3,5
60	0,64	6,67

начает отсутствие на изображениях искажений формы, связанных с удалением от камеры; такая аппроксимация позволяет существенно упростить вычисления и калибровку аппаратуры.

Настройка аппаратуры

Принимая во внимание необходимость мощного источника света, а также короткого времени импульса вспышки, что уменьшает влияние движений объекта или камеры на качество снимков, источником освещения была выбрана генераторная установка Calumet Elite 2400 с четырьмя импульсными вспышками.

Необходимость малой продолжительности импульса вспышки вызвана наличием в матрицах цифровых фотоаппаратов так называемого темнового тока. Данный ток возникает по причинам теплового движения молекул в матрице и не зависит от падающего освещения. Величина сигнала, формируемого этим током, пропорциональна квадрату температуры матрицы и времени выдержки. Сигнал, наведенный темновым током, имеет шумоподобную структуру и существенно снижает точность измерений. Для уменьшения влияния темнового тока на результирующее изображение необходимо либо охлаждать матрицу (что производится в некоторых дорогих фотоаппаратах), либо уменьшать время экспонирования. При выдержке

порядка 1/200 с влияние теплового тока оказывается меньше структурного шума КМОП-матрицы, а значит, им можно пренебречь. Также необходимо учесть, что при выдержке 1/200 с часть этого времени затрачивается на движение ламелей, открывающих и закрывающих сенсор. В связи с этим время, которое сенсор остается полностью открытым, составляет не более 1/400 с. Соответственно, время импульса не должно превышать 1/400 с.

При вычислениях другие источники света не учитываются, т. е. в процессе сканирования не должно быть никаких посторонних источников света. Если используется импульсный осветитель с достаточно большим световым потоком, то специальная обработка помещения не требуется, получение снимков может производиться при тусклом постороннем освещении, паразитная засветка не окажет влияние на вычисления. Импульсный осветитель 1200 Дж с учетом потерь при создании точечного источника света 85 % испускает световой поток 270 000 лм за 1 мс, что обеспечит освещенность 1343 лк. Доля постороннего освещения в помещении с освещенностью 3 лк составляет 0,22 %, что позволяет пренебречь этим освещением при вычислениях. Для того чтобы избежать переотражений света, исходящего от осветителя (особенно от потолка при L_2), необходимо уменьшить телесный угол, в котором распространяется световой поток, при этом освещенность объекта уменьшится незначительно. Для уменьшения угла могут быть использованы шторы или сотовая насадка. В целях уменьшения доли переотраженного света на объекте используется фотофон, внутренняя поверхность которого обработана черным бархатом.

Поскольку необходимости располагать источники света на одном расстоянии нет, а устройства фиксации и размеры помещения ограничивают возможности расположения ламп, была выбрана конфигурация освещения (табл. 2), позволяющая считать источники света равномошными при угле освещения $\xi = 20^\circ$; без использования допол-

■ **Таблица 2.** Настройка освещения

■ **Table 2.** Light sources configuration

Источник света	Мощность заряда лампы, Дж	Расстояние l , м
L_1	1904	2,66
L_1^1	1904	2,66
L_2	1512	2,26
L_2^1	1512	2,26

нительного снимка. Так как была применена рассеивающая насадка, съемка производилась при высокой мощности заряда лампы.

Для получения снимков объекта использована камера Canon EOS 5D с выдержкой фотокамеры 1/200 с и диафрагмой 8.0.

При выборе выдержки желательно устанавливать минимально возможное значение, поскольку это минимизирует искажения темнового тока, но необходимо учитывать, что многие конструкции затвора позволяют получить малое время выдержки за счет частичного открытия затвора, что недопустимо для съемки с импульсными источниками света. Для камеры Canon EOS 5D время синхронизации (минимальная выдержка, при которой полностью открывается затвор) составляет 1/200 с. При выборе диафрагмы необходимо учитывать следующие факторы:

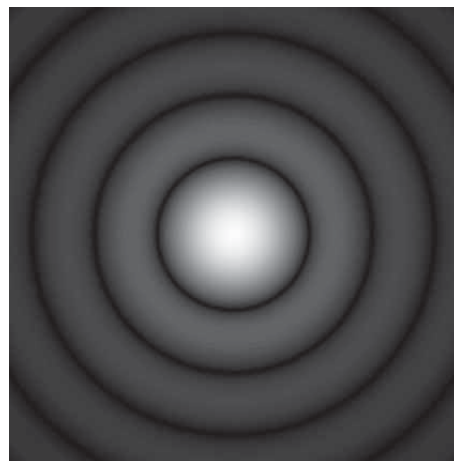
1. Глубину резкости — расстояние вдоль оптической оси объектива между двумя плоскостями в пространстве предметов, в пределах которого объекты отображаются в сопряженной фокальной плоскости субъективно резко. При проведении эксперимента необходимо, чтобы все пространственные точки области сканирования находились в зоне глубины резкости. Глубина резкости зависит от:

— фокусного расстояния объектива — по мере увеличения фокусного расстояния объектива глубина резкости увеличивается;

— используемой диафрагмы — по мере уменьшения размера отверстия диафрагмы глубина резкости увеличивается (меньшему размеру диафрагмы соответствует большее значение).

2. Аберрации объектива — современным объективам свойственен ряд специфических искажений, приводящих к снижению разрешающей способности объектива. Поскольку подобные искажения преобладают в периферической зоне линз, используемых в объективе, уменьшение диафрагмы приводит к снижению количества лучей, проходящих через периферическую зону линз и, соответственно, к увеличению разрешающей способности объектива.

3. Дифракционный предел оптики — даже в случае полной компенсации аберраций объектива при проецировании точки на матрицу видна не точка, а размытый кружок рассеяния с максимумом в центре и интерференционными кольцами, расположенными вокруг него. Согласно критерию Рэлея, два дифракционных круга рассеяния воспринимаются зрением как два в том случае, если между их максимумами расположен минимум с интенсивностью на 20 % меньше. В противном случае они воспринимаются глазом как один. Так как объектив и диафрагма имеют круглую форму, то точка приобретает вид нерезкого круга, который называется



■ Рис. 2. Диск Эйри
■ Fig. 2. The Airy disk

дискон Эйри (рис. 2). Чтобы две точки, проецируемые на матрицу, воспринимались как две, они должны удовлетворять критерию Рэлея. Размер диска Эйри для объектива с условием соблюдения критерия Рэлея можно вычислить по формуле

$$D = 1,22\lambda K_d, \quad (2)$$

где λ — длина световой волны (для цифровой фотографии берется по зеленому цвету 500 нм); K_d — диафрагменное число.

Для различимости двух пикселей должно выполняться условие $D < r$, где r — расстояние между пикселями на матрице. Поскольку в камере Canon EOS 5D расстояние между пикселями составляет 8,2 мкм, диафрагменное число не должно превышать 13,4. Для повышения контрастности в высокочастотной области стоит ограничить число диафрагмы значением не более 10.

4. Уменьшение размера отверстия диафрагмы приводит к уменьшению светового потока, попадающего на матрицу. Компенсировать уменьшение светового потока можно либо за счет увеличения выдержки (что приведет к увеличению темнового тока), либо за счет увеличения мощности светового импульса, используемого для освещения объекта съемки.

Сканирование плоской поверхности и калибровка

Целями сканирования плоской поверхности являются: определение погрешности вычисления ориентации поверхности, определение характеристик сканера и коррекция значений яркости на исходных снимках.

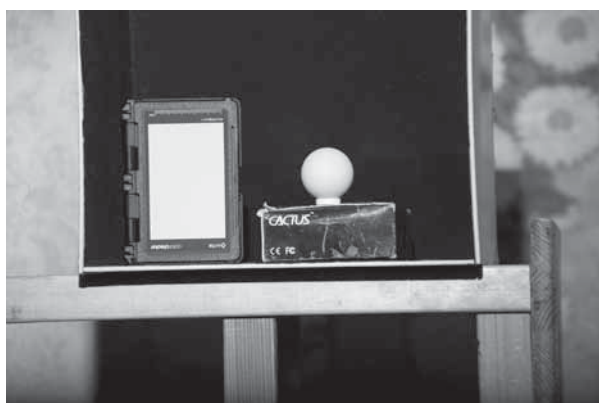
Для калибровки и определения погрешности яркости на исходных снимках может быть использована любая однородная матовая плоская поверхность с неизменным значением яркостей для всех снимков:

$$i_1 = i_1^1 = i_2 = i_2^1 = |L_1|K\cos\xi.$$

Следовательно, для калибровки необходимо получить набор снимков этой поверхности, удовлетворяющий двум условиям.

1. Объект на каждом снимке освещен равномерно, т. е. $i_k(x, y) = \text{const}$. В противном случае источник света не удовлетворяет условиям сканирования или же необходимо увеличить расстояние от камеры до объекта l .

2. Каждая точка объекта имеет одну яркость на всех снимках $i_1 = i_1^1 = i_2 = i_2^1$. В противном случае, и если при этом условие 1 выполнено, необ-



■ **Рис. 3.** Снимок I_1 объекта и серой карты
 ■ **Fig. 3.** Image I_1 of the object and gray card

ходимо отрегулировать мощность источников света и расстояние каждого источника света. Возможна корректировка яркости уже полученных снимков.

В большинстве случаев поверхность не освещена равномерно. В работе проведена корректировка значений яркости точек изображения в зависимости от их расположения относительно участка поверхности, освещенного равномерно.

Для коррекции значений яркости используется серая карта, расположенная слева от сканируемого объекта (рис. 3). Набор кадрированных участков серой карты на снимках представлен на рис. 4, а–г.

Если источники света были установлены не точно или размеры объекта больше области, которая может быть освещена равномерно, то необходимо проанализировать изменение яркости на плоской поверхности.

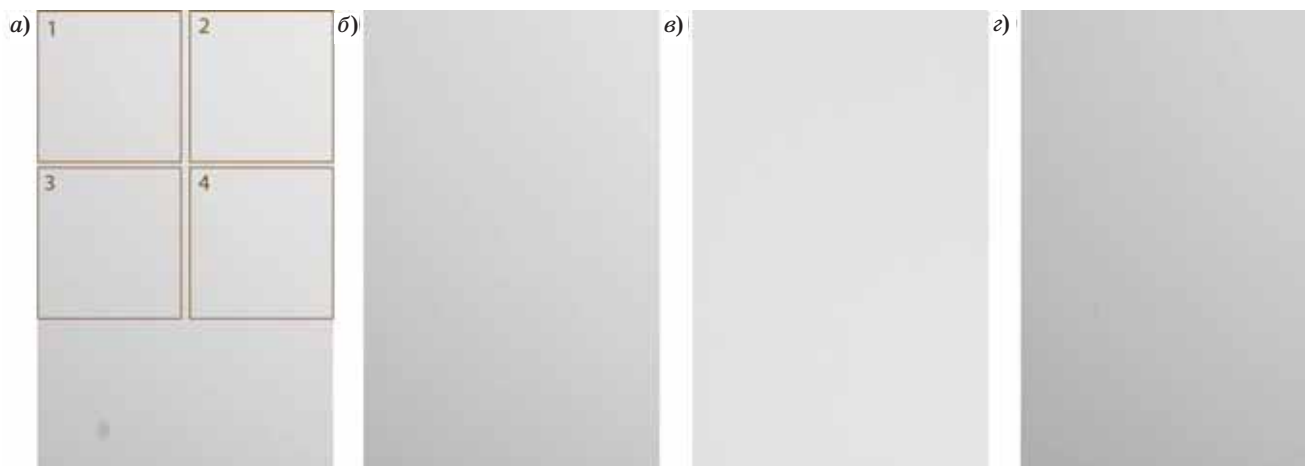
1. Разбить всю плоскость на четыре участка (см. рис. 4, а).

2. Для каждой части определить средние значения градиента яркости $\Delta i_{x \text{ ср}}$, $\Delta i_{y \text{ ср}}$, если оба значения меньше некоторого порогового значения, то коррекция значений яркости не производится.

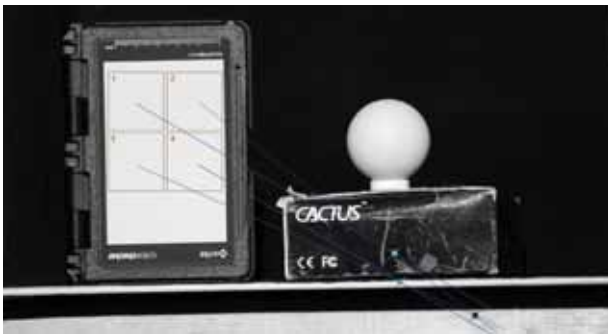
3. Из каждой области провести линию $y = ax + b$, где $a = \Delta i_{y \text{ ср}} / \Delta i_{x \text{ ср}}$, само значение b корректируется на шаге 4.

4. Определяется центр области максимальной освещенности. Предполагается, что центр области не может находиться слева от серой карты. Если точек пересечения несколько (рис. 5), то значения b линий корректируются до минимального значения величины $\sum_j \sum_i (x_i - x_j)^2 + (y_i - y_j)^2$, где

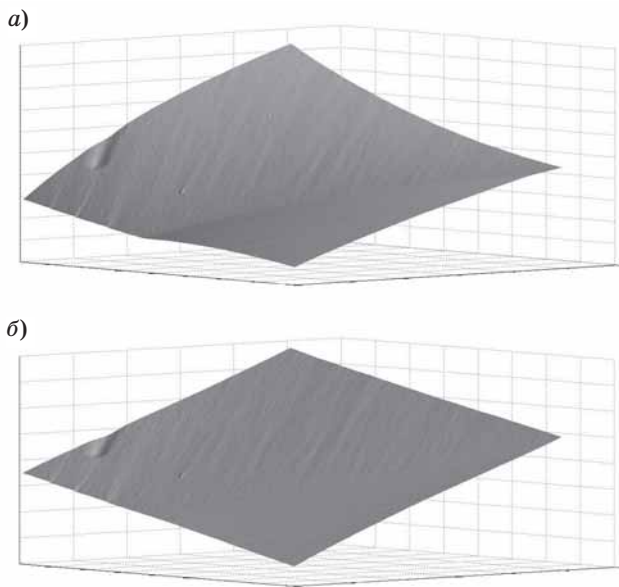
i, j — номера точек пересечения; x_i, y_i — координаты точки пересечения.



■ **Рис. 4.** Участок серой карты на снимках I_1 (а), I_1^1 (б), I_2 (в), I_2^1 (г)
 ■ **Fig. 4.** Gray card patch of images I_1 (а), I_1^1 (б), I_2 (в), I_2^1 (г)



■ **Рис. 5.** Определение центра светового пятна
 ■ **Fig. 5.** Determination of the light spot center



■ **Рис. 6.** Реконструированная плоская поверхность: *a* — без коррекции значений яркости снимков I_1, I_1^1, I_2, I_2^1 ; *б* — после коррекции значений яркости снимков I_1, I_1^1, I_2, I_2^1
 ■ **Fig. 6.** Result of the flat surface scanning: *a* — without brightness correction of images I_1, I_1^1, I_2, I_2^1 ; *б* — with brightness correction of images I_1, I_1^1, I_2, I_2^1

5. В зависимости от удаления d от центра области максимальной освещенности значение яркости умножается на коэффициент усиления яркости f , определяемый по снимку источника света или матовой плоской освещенной этим источником света поверхности. Предполагается, что коэффициент f не зависит от направления от центра.

Полученные результаты построения плоской поверхности представлены на рис. 6, *a* и *б*. Область искажения формы у левого края поверхности вызвана пылинкой на матрице фотоаппарата, что привело к повышению погрешности. Для интегрирования градиента используется локальный алгоритм интегрирования на основе криволинейного интеграла [7, 9, 17].

Среднее значение градиента поверхности, представленной на рис. 6, *б*, составило 0,006 мм для приращения p и 0,017 мм для приращения q , среднеквадратическое отклонение (СКО) значений приращений составило 0,0009 и 0,0013 мм соответственно.

Определение погрешности сканирования

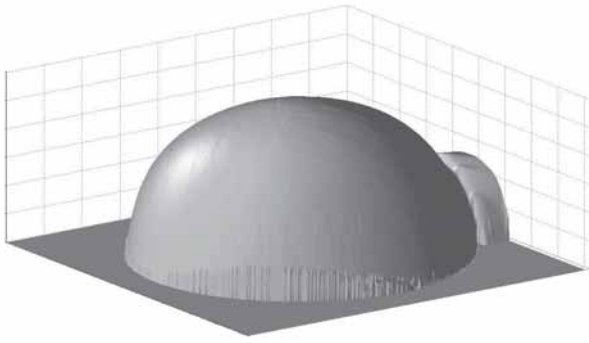
Вычисление ориентации поверхности осуществляется для каждой видимой точки поверхности. Для интегрирования градиента используется локальный алгоритм интегрирования на основе криволинейного интеграла [8]. Пример области, занимаемой сканируемым объектом на исходных снимках, представлен на рис. 7, *a–д*.

Результат сканирования объекта, близкого к сферической форме (теннисный мячик), представлен на рис. 8. Сфера использовалась в качестве калибровочной поверхности [1], так как охватывает весь диапазон возможных значений приращений, а также часто выступает в качестве тестового или опорного объекта [4, 5].

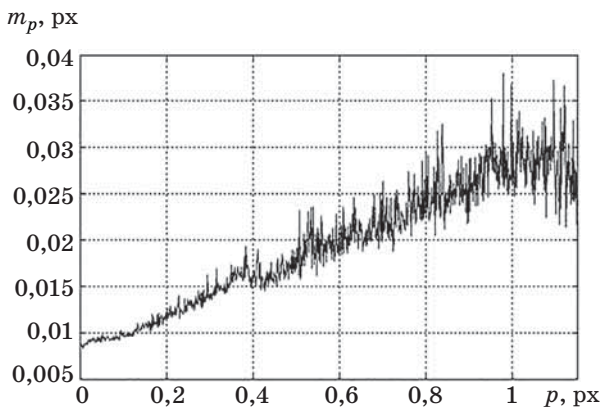
Расстояние между пикселями вычислено на основе ширины теннисного мяча 40 мм на изображении 388 px и составляет 0,103 мм.



■ **Рис. 7.** Исходные снимки I_1 (*a*), I_1^1 (*б*), I_2 (*в*), I_2^1 (*г*) и видимая область (*д*)
 ■ **Fig. 7.** Source images I_1 (*a*), I_1^1 (*б*), I_2 (*в*), I_2^1 (*г*) and visible area (*д*)



■ **Рис. 8.** Результат сканирования теннисного мяча
 ■ **Fig. 8.** Result of the table tennis ball scanning



■ **Рис. 9.** Абсолютная погрешность вычисления приращения поверхности теннисного мяча в зависимости от величины приращения
 ■ **Fig. 9.** Gradient absolute calculation error for tennis ball in dependence of gradient value

Для определения погрешности эксперимента было произведено 10 повторений сканирования теннисного мяча. Исходя из выводов, представленных в работах [8, 10], погрешность вычисления приращений зависит от величины приращений, что подтверждается результатами тестового сканирования (рис. 9).

Относительная и абсолютная погрешности вычисления градиента были вычислены для каждой точки поверхности на основе СКО и среднего значения при 10 итерациях и уровне доверия 95 %. Относительная погрешность вычисления градиента принимает значения от 4 % при малых p до 14 % на границе сканируемого объекта, среднее значение относительной погрешности 6,6 %. Полученная средняя абсолютная погрешность вычисления градиента составляет 0,02 px (0,002 мм).

Абсолютная погрешность вычисления третьей координаты была вычислена для каждой точки поверхности на основе СКО и среднего значения при 10 итерациях и уровне доверия 95 %. Абсолютная погрешность вычисления третьей координаты принимает значения до 1,5 px (0,154 мм). Полученная средняя абсолютная погрешность вычисления третьей координаты составляет 0,604 px (0,062 мм), СКО по всей поверхности составляет 0,26 px (0,028 мм).

Заключение

Характеристики трехмерного сканера, независимые от сканируемой поверхности и конфигурации устройства сканирования, определены в процессе компьютерного моделирования и анализа погрешностей.

Плотность точек по осям x, y определяется разрешением камеры и всегда равна 1 px — условной величине, равной расстоянию между соседними пикселями.

Минимальное расстояние между точками по оси z равно 0. Максимальное расстояние между точками по оси z теоретически может достигать бесконечности. Так как слабоосвещенных областей на изображениях значительно меньше, в целом облако точек достаточно плотное.

Под временем сканирования подразумевается время получения изображений, все необходимые вычисления наличия объекта не требуют. Если рассматривать сканирование в автоматическом режиме с дополнением к представленной аппаратуре устройств управления, то обработка одного вида объекта может быть произведена за 0,1 с; с учетом перезарядки источников света 2,6 с и установленном количестве ракурсов 36 общее время составит 97,2 с. При ручном переключении источников света время обработки одного вида объекта составило 12 с.

Возможность получения цвета объекта без дополнительной обработки является одной из отличительных особенностей фотометрического метода сканирования. Для большинства существующих методов такой возможности нет.

Для опытной установки полученная средняя абсолютная погрешность вычисления третьей координаты составляет 0,062 мм, СКО по всей поверхности составляет 0,0275 мм.

Размеры предельной области сканирования опытной установки определены исходя из размеров светового пятна и составляют 0,7 м. Плотность точек по осям x, y равна 0,103 мм.

Работа выполнена при поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере.

Литература

1. Woodham R. J. Gradient and Curvature from the Photometric-Stereo Method, Including Local Confidence Estimation // *Journal of Optical Society of America*. Nov. 1994. Vol. 11. N 11. P. 3050–3068. doi:10.1364/JOSAA.11.003050
2. Basri R., Jacobs D., Kemelmacher I. Photometric Stereo with General, Unknown Lighting // *International Journal of Computer Vision*. 2007. Vol. 72(3). P. 239–257.
3. Hertzmann A., Seitz S. M. Example-based Photometric Stereo: Shape Reconstruction with General Varying BDRFs // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Aug. 2005. Vol. 27. N 8. P. 1254–1264. doi:10.1109/TPAMI.2009.102
4. Solomon F., Ikeuchi K. Extracting the Shape and Roughness of Specular Lobe Objects using Four Light Photometric Stereo // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Apr. 1996. Vol. 18. N 4. P. 449–454.
5. Barsky S., Petrou M. The 4-source Photometric Stereo Technique for Three-Dimensional Surfaces in the Presence of Highlights and Shadows // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Oct. 2003. Vol. 25. N 10. P. 1239–1252. doi:10.1109/TPAMI.2003.1233898
6. Woodham R. J. Photometric Method for Determining Surface Orientation from Multiple Images // *Optical Engineering*. 1980. Vol. 19. N 1. P. 139–144.
7. Красильников Н. Н. Метод получения 3D-изображений, основанный на диффузном отражении света сканируемыми объектами // *Информационно-управляющие системы*. 2009. № 6. С. 7–11.
8. Красильников Н. Н., Красильникова О. И. Исследование погрешностей определения координаты глубины при 3D-сканировании методом, основанным на диффузном отражении света // *Информационно-управляющие системы*. 2012. № 3. С. 2–8.
9. Klette R., Schlüns K. Height Data from Gradient Fields // *Proc. Machine Vision Applications, Architectures, and Systems Integration V. SPIE 2908*. Boston, Massachusetts, Nov. 18–19, 1996. P. 204–215.
10. Кузнецов В. А. Алгоритм сегментации исходных снимков для фотометрического метода трехмерного сканирования // *Информационно-управляющие системы*. 2015. № 3. С. 29–34. doi:10.15217/issn1684-8853.2015.3.29
11. Кузнецов В. А. Использование неравномерной шкалы квантования интенсивности света для вычисления градиента поверхности фотометрическим методом трехмерного сканирования // *Известия высших учебных заведений. Приборостроение*. 2015. № 10(58). С. 816–822. doi:10.17586/0021-3454-2015-58-10-816-822
12. Кузнецов В. А. Фотометрический метод вычисления ориентации поверхности с высокой точностью // *Научная сессия ГУАП: сб. докл. Ч. 1. Технические науки*. СПб.: ГУАП, 2015. С. 260–263.
13. Chandraker M., Agarwal S., Kriegman D. Shadowcuts: Photometric Stereo with Shadows // *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*. 2007. doi:10.1109/CVPR.2007.383288
14. Hernandez C., Vogiatzis G., Cipolla R. Overcoming Shadows in 3-Source Photometric Stereo // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Feb. 2011. Vol. 33(2). P. 419–426.
15. Sunkavalli K., Zickler T., Pfister H. Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows // *Inc: ECCV 2*. Vol. 6312. of *Lecture Notes in Computer Science*. 2010. P. 251–264.
16. Кузнецов В. А. Определение максимальной ошибки вычисления ориентации поверхности для метода трехмерного сканирования, основанного на модели отражения света поверхностью // *Научная сессия ГУАП: сб. докл. Ч. 1. Технические науки*. СПб.: ГУАП, 2014. С. 177–179.
17. Kuznetsov V. A. The Treatment of the Surface Gap for Photometric Stereo // *Bulletin of the UNESCO Department “Distance Education in Engineering” of the SUAI: Collection of the Papers*. 2016. Iss. 1. P. 50–52.

UDC 612.8

doi:10.15217/issn1684-8853.2018.3.25

Photometric Stereo Scanner with Flash Light Lamp

Kuznetsov V. A.^a, Assistant Professor, k.avk-c@mail.ru
 Liudaev M. U.^b, Programmer Engineer, mike_ld@mail.ru

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

^bRussian Institute for Power Radiobuilding, 11 Line V. O., 66, 199178, Saint-Petersburg, Russian Federation

Introduction: The photometric stereo technique is well known, but currently there are no good technologies which use it to perform 3D scanning. This is due to the complexity of calculating the coordinates of the surface points, as you calculate only the surface orientation whose uneven error distribution makes it difficult to obtain a 3D model. In addition, there are a number of demands about the conditions under which it is possible to scan an object. This considerably limits the application of photometric stereo and raises the price of the scanning. **Purpose:** Developing a pilot model of a 3D scanner using the available photo equipment, and determining

the possibilities of the photometric stereo technique for the case of using a flash light. **Results:** A pilot model of a scanner based on the photometric stereo method was obtained and tested. The paper discusses the characteristics of this model, configuration of the equipment and conditions of the scanning process. For the pilot model, the resolution along the x and y axes is 0.103 mm, the size of the scan area is 0.7×0.7 m². The average coordinate calculation error caused by the camera matrix noise, ambient light effect and uneven illumination from the flash light lamps has been estimated for a test object and found equal to 0.062 mm. The standard deviation over the area is 0.0275 mm. **Practical relevance:** The obtained results demonstrate the expected high accuracy of the photometric stereo technique, with inexpensive equipment in use. The experiment was carried out without a special treatment of the room. These results make the photometric stereo technique more accessible for use.

Keywords — Photometric Stereo, 3D Scanner, Image Processing, Photography.

Citation: Kuznetcov V. A., Liudaev M. U. Photometric Stereo Scanner with Flash Light Lamp. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 25–33 (In Russian). doi:10.15217/issn1684-8853.2018.3.25

References

- Woodham R. J. Gradient and Curvature from the Photometric-Stereo Method, Including Local Confidence Estimation. *Journal of Optical Society of America*, Nov. 1994, vol. 11, no. 11, pp. 3050–3068. doi:10.1364/JOSAA.11.003050
- Basri R., Jacobs D., Kemelmacher I. Photometric Stereo with General, Unknown Lighting. *International Journal of Computer Vision*, 2007, vol. 72, no. 3, pp. 239–257.
- Hertzmann A. and Seitz S. M. Example-based Photometric Stereo: Shape Reconstruction with General Varying BDRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, August 2005, vol. 27, no. 8, pp. 1254–1264. doi:10.1109/TPAMI.2009.102
- Solomon F. and Ikeuchi K. Extracting the Shape and Roughness of Specular Lobe Objects using four Light Photometric Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, April 1996, vol. 18, no. 4, pp. 449–454.
- Barsky S., Petrou M. The 4-source Photometric Stereo Technique for Three-dimensional Surfaces in the Presence of Highlights and Shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, October 2003, vol. 25, no. 10, pp. 1239–1252. doi:10.1109/TPAMI.2003.1233898
- Woodham R. J. Photometric Method for Determining Surface Orientation from Multiple Images. *Optical Engineering*, 1980, vol. 19, no. 1, pp. 139–144.
- Krasilnikov N. N. A Method of 3D Scanning based on Scanned Objects' Light Radiosity. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2009, no. 6, pp. 7–11 (In Russian).
- Krasilnikov N. N., Krasilnikova O. I. The Investigation of Measurement Errors of Depth Coordinate during 3D Scanning by the Method based on Diffuse Reflection of Light. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 3, pp. 2–8 (In Russian).
- Klette R. and Schlüns K. Height Data from Gradient Fields. *Proc. Machine Vision Applications, Architectures, and Systems Integration V*, SPIE 2908, Boston, Massachusetts, Nov. 18–19, 1996, pp. 204–215.
- Kuznetcov V. A. Image Segmentation Technique for Photometric Stereo. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 3, pp. 29–34 (In Russian). doi:10.15217/issn1684-8853.2015.3.29
- Kuznetcov V. A. Non-uniform Quantization Scale for Surface Gradient Calculation with Photometry Method of Three-dimensional Scanning. *Izvestiia vysshikh uchebnykh zavedenii. Priborostroenie* [Journal of Instrument Engineering], 2015, vol. 58, no. 10, pp. 816–822 (In Russian). doi: 10.17586/0021-3454-2015-58-10-816-822
- Kuznetcov V. A. The Photometric Stereo Technique with High Accuracy. *Nauchnaia sessiia GUAP: sbornik dokladov. Ch. 1. Tekhnicheskie nauki* [Collection of Presentations at the Scientific Session of the SUAI], Saint-Petersburg, SUAI, 2015, pp. 260–263 (In Russian).
- Chandraker M., Agarwal S., Kriegman D. Shadowcuts: Photometric Stereo with Shadows. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2007. doi:10.1109/CVPR.2007.383288
- Hernandez C., Vogiatzis G., Cipolla R. Overcoming Shadows in 3-Source Photometric Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Feb. 2011, vol. 33(2), pp. 419–426.
- Sunkavalli K., Zickler T., Pfister H. Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows. *Inc: ECCV 2, Lecture Notes in Computer Science*, 2010, vol. 6312, pp. 251–264.
- Kuznetcov V. A. Determination of the Maximum Surface Orientation Error for the Photometric Stereo Technique. *Nauchnaia sessiia GUAP: sbornik dokladov. Ch. 1. Tekhnicheskie nauki* [Collection of Presentations at the Scientific Session of the SUAI], 2014, pp. 177–179 (In Russian).
- Kuznetcov V. A. The Treatment of the Surface Gap for Photometric Stereo. *Bulletin of the UNESCO Department "Distance Education in Engineering" of the SUAI: Collection of the Papers*, 2016, iss. 1, pp. 50–52.

МЕТОД ЭФФЕКТИВНОГО УДЕРЖАНИЯ ПОЛОЖЕНИЯ АНПА НА МАРШРУТНОЙ ТРАЕКТОРИИ ПРИ ВЕДЕНИИ СЕЙСМОРАЗВЕДКИ

Л. А. Мартынова^а, доктор техн. наук, старший научный сотрудник, martynowa999@bk.ru
^аАО «Концерн «ЦНИИ «Электроприбор», Малая Посадская ул., 30, Санкт-Петербург, 197046, РФ

Постановка проблемы: в процессе ведения сейсморазведки в подледных условиях автономные необитаемые подводные аппараты могут отклоняться от маршрутной траектории вследствие влияния течения и других случайных нестационарных факторов. В связи с этим для ведения эффективной сейсморазведки необходимо постоянно удерживать аппарат на маршрутной траектории. **Цель:** разработка метода эффективного удержания автономного необитаемого подводного аппарата на маршрутной траектории. **Метод:** после выявления факта отклонения аппарата от маршрутной траектории для последующего его восстановления на маршрутной траектории использован один из методов наведения на цель — метод управления по отклонению угла линии визирования. **Результаты:** получены математические зависимости оценки отклонения аппарата от маршрутной траектории; реализован метод восстановления аппарата на маршрутной траектории, доработанный с учетом ограничений на маневренные характеристики аппарата; разработана математическая модель и проведена оптимизация параметров предложенного метода на основе оценки эффективности ведения сейсморазведки. По результатам моделирования определены значения параметров метода удержания автономных необитаемых подводных аппаратов на маршрутной траектории, обеспечивающих эффективное ведение поиска, с учетом расстояния между маршрутными точками, ограничений на скорость поворота автономных необитаемых подводных аппаратов и допустимой дистанции их отклонения от маршрутной траектории. **Практическая значимость:** разработанный метод позволяет эффективно восстанавливать положение аппарата на маршрутной траектории с учетом ограничений и особенностей ведения сейсморазведки в заданном регионе.

Ключевые слова — автономный необитаемый подводный аппарат, маршрутная траектория, течение, оценка эффективности.

Цитирование: Мартынова Л. А. Метод эффективного удержания положения АНПА на маршрутной траектории при ведении сейсморазведки // Информационно-управляющие системы. 2018. № 3. С. 34–44. doi:10.15217/issn1684-8853.2018.3.34

Citation: Martynova L. A. Effective Retention of AUV Position on Route Trajectory during Seismic Survey. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 34–44 (In Russian). doi:10.15217/issn1684-8853.2018.3.34

Введение

При ведении сейсморазведки по поиску углеводорода в морском дне Арктического шельфа [1–4] в подледных условиях в настоящее время используются сейсморазведочные станции, которые после их распределения на водной поверхности опускаются на дно для прослушивания отраженного сейсмосигнала, создаваемого излучателем. Стремительное совершенствование возможностей автономных необитаемых подводных аппаратов (АНПА) создает все предпосылки для их использования в разных отраслях народного хозяйства, в том числе и в сейсморазведке. При использовании группы АНПА на всех аппаратах размещают по одному геофону [5–8]. Каждый АНПА группы с геофоном самостоятельно перемещается в заданную маршрутным заданием точку и опускается на дно. Геофон ведет прием сигналов, создаваемых излучателем и отраженных от толщи морского дна, затем АНПА всплывает и перемещается в новую маршрутную точку и т. д. [9–11].

Процесс перемещения АНПА сопровождается отклонением их положения от заданной маршрутной траектории. Отклонения могут быть вы-

званы внешними возмущающими факторами: течением, изменением параметров морской среды, погрешностями собственного определения местоположения [12, 13] и т. д. Отклонение положения АНПА от заданной маршрутной траектории приводит к нарушению взаимного положения АНПА в группе и снижению возможности обнаружения местоположения залежей углеводородов. Поэтому для эффективного ведения сейсморазведки с помощью АНПА необходимо в случае отклонения АНПА от маршрутной траектории восстанавливать его положение на маршрутной траектории.

Имеющиеся работы в области уточнения положения аппарата и прогнозирования параметров его движения [14–16] не позволяют использовать полученные результаты для подводных аппаратов, поскольку подводная среда отличается чрезмерным затуханием звукового сигнала, ограниченностью связи, влиянием параметров морской среды [17].

Для достижения поставленной в работе цели — разработки метода эффективного удержания АНПА на маршрутной траектории — были решены следующие задачи:

— проанализированы факторы, оказывающие влияние на положение АНПА;

— проанализированы имеющиеся подходы к восстановлению положения АНПА на маршрутной траектории в случае выявления отклонения его положения от заданного;

— разработан алгоритм восстановления положения АНПА на маршрутной траектории с учетом ограничений на маневренные возможности АНПА;

— определены показатели эффективности удержания АНПА на маршрутной траектории;

— разработана математическая модель для проведения исследований по выбору оптимальных параметров метода удержания положения АНПА на заданной траектории;

— получены оптимальные значения параметров метода удержания АНПА на маршрутной траектории, обеспечивающие эффективное ведение сейсморазведки;

— по результатам численного эксперимента выданы рекомендации.

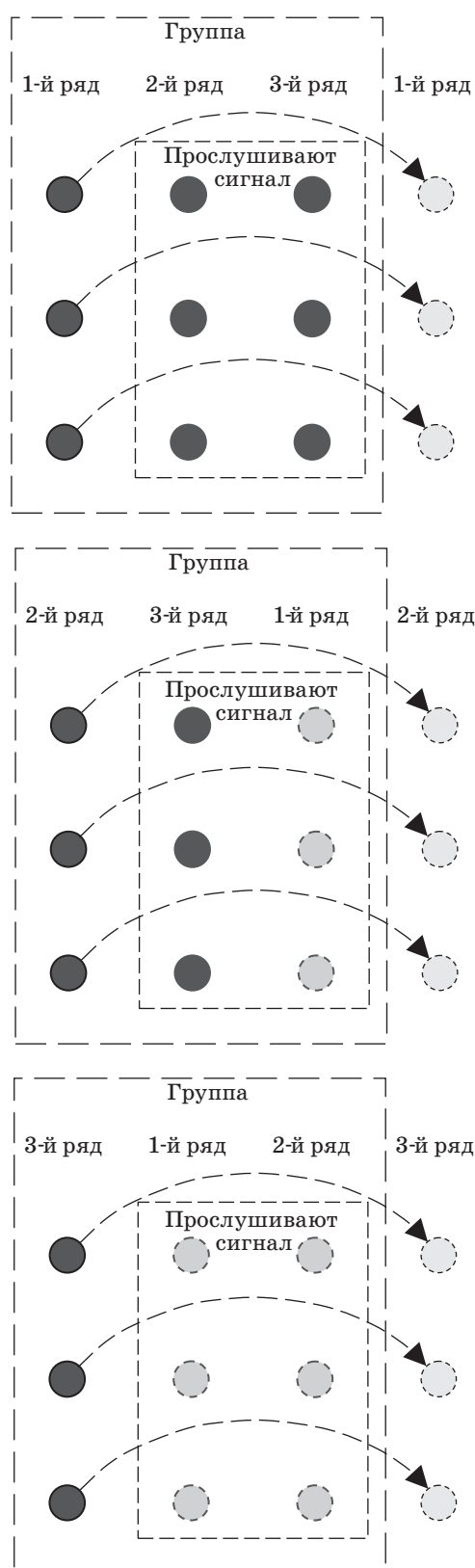
Описание поведения группы АНПА

Перед началом выполнения сейсморазведки в каждый аппарат вводится маршрутное задание с указанием положения маршрутных точек. Линия, соединяющая маршрутные точки, является маршрутной траекторией. После того как аппараты доставлены в район выполнения сейсморазведки, они выстраиваются группой в исходное положение.

В ходе ведения сейсморазведки излучатель зондирует донную поверхность импульсными сигналами. Большая часть АНПА из группы с помощью геофонов принимает отраженный сигнал, создаваемый излучателем. В это время другая часть АНПА перемещается для занятия новой позиции, не участвуя при этом в приеме отраженных сигналов. После занятия новой позиции прием отраженных сигналов осуществляет группа АНПА в обновленном составе, при этом часть АНПА, не задействованная в приеме отраженного сигнала, перемещается в места новой своей дислокации в соответствии с маршрутным заданием. Схема поэтапного перемещения АНПА группы, в результате которого происходит поступательное движение группы АНПА, приведена на рис. 1. Общая схема взаимного положения излучателя и группы аппаратов представлена на рис. 2, а и б.

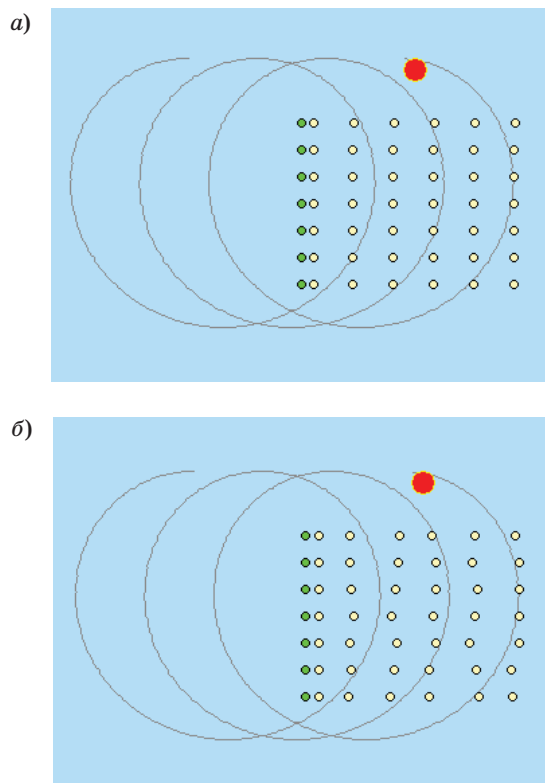
По результатам постобработки зарегистрированных сигналов определяется местоположение залежей углеводородов.

Из приведенного описания тактической ситуации при ведении сейсморазведки с использованием АНПА следует, что для снижения погрешности определения местоположения залежей



■ **Рис. 1.** Схема последовательного перемещения ряда АНПА, обеспечивающего согласованное перемещение группы АНПА с излучателем

■ **Fig. 1.** Scheme of sequential displacement of a series of AUV, ensuring a coordinated movement of the AUV group with the radiator



■ **Рис. 2.** Положение АНПА без влияния (а) и с влиянием (б) возмущающих факторов
 ■ **Fig. 2.** The position of the AUV without influence (a) and with the influence of (b) perturbing factors

углеводородов аппарату необходимо с высокой точностью оказаться в окрестности точки для последующего опускания на дно и прослушивания в этой точке отраженного сигнала. Однако на положение АНПА при перемещении влияние оказывают внешние факторы (например, течение), которые вызывают отклонение АНПА от заданной маршрутной траектории.

Задача состоит в удержании АНПА на маршрутной траектории так, чтобы обеспечить прибытие АНПА в окрестность заданной точки с погрешностью в пределах заданной. Разработан метод для решения указанной задачи.

Метод удержания положения АНПА на маршрутной траектории

Удержание АНПА на маршрутной траектории определяется:

- собственным позиционированием аппарата;
- результатами анализа влияния внешних факторов, которые вызывают отклонение АНПА от заданной маршрутной траектории;
- алгоритмом возврата на маршрутную траекторию.

Собственное позиционирование АНПА в момент движения происходит с использованием бортовой инерциальной навигационной системы и (или) вспомогательных средств — донных или поверхностных маяков-ответчиков. Частота обновления данных менее 1 с.

Анализ положения АНПА на маршрутной траектории заключается в определении отклонения положения АНПА от маршрутной траектории и сравнении полученного отклонения с допустимой дистанцией. Выбор допустимой дистанции определяется возможностью возвращения АНПА на маршрутную траекторию. На маршрутную траекторию аппарат должен вернуться до местоположения ближайшей маршрутной точки.

Расчет отклонения АНПА от маршрутной траектории между i -й и $(i + 1)$ -й маршрутными точками будем проводить следующим образом. Для одной из координат положения АНПА, например X , определяется соответствующая ей вторая координата Y , как точка, принадлежащая маршрутной траектории:

$$Y_{\text{м.т}} = Y_i + (X_{\text{АНПА}} - X_i) \frac{(Y_{i+1} - Y_i)}{(X_{i+1} - X_i)},$$

где $Y_{\text{м.т}}$ — координата АНПА при нахождении его на маршрутной траектории; $X_{\text{АНПА}}$ — текущая координата АНПА; X_i, Y_i — координаты i -й точки; X_{i+1}, Y_{i+1} — координаты $(i + 1)$ -й точки.

Если текущее отклонение АНПА от маршрутной траектории превышает значение допустимой дистанции $D_{\text{доп}}$, т. е. если выполняется условие $|Y_{\text{АНПА}} - Y_{\text{м.т}}| > D_{\text{доп}}$, то АНПА производит корректировку своего положения путем маневрирования по курсу.

Выбор алгоритма корректировки определяется обеспечением прибытия АНПА в маршрутную точку с погрешностью не выше заданной.

Возможность возвращения АНПА на маршрутную траекторию определяется:

- оставшимся расстоянием до маршрутной точки;
- ограничениями, накладываемыми на маневренные характеристики АНПА;
- алгоритмом, выбранным для формирования курса АНПА, обеспечивающего возвращение на маршрутную траекторию.

При выборе алгоритма учитывалась целесообразность перехода от возвращения АНПА на маршрутную траекторию к прибытию АНПА в маршрутную точку. Поэтому были рассмотрены алгоритмы, ориентированные на прибытие АНПА в окрестность маршрутной точки:

- метод погони, условием выполнения которого является непрерывное направление продольной оси АНПА на ближайшую маршрутную точку;

— метод наведения с упреждением, вектор скорости АНПА направлен в упрежденную точку «встречи» с ближайшей маршрутной точкой;

— метод «накрытия» ближайшей маршрутной точки, при котором должно быть обеспечено равенство углов между «АНПА — i -я маршрутная точка» и «АНПА — $(i + 1)$ -я маршрутная точка».

В настоящей работе с учетом специфики ведения сейсморазведки выбран алгоритм «наведения на цель» [18–21], обеспечиваемый алгоритмом управления по отклонению угла линии визирования. Этот алгоритм отличается тем, что не требует каких-либо дополнительных специальных измерений, например расстояния до маршрутной траектории. В качестве цели рассматривается заданная маршрутная точка траектории. В результате определяется курс, на который должен лечь АНПА. Для исключения ситуации «промаха» АНПА относительно заданной маршрутной точки, когда АНПА входит в область вокруг заданной маршрутной точки траектории, система управления АНПА запускает наведение на следующую маршрутную точку траектории. Описанный алгоритм «наведения на цель» минимизирует время восстановления положения АНПА на маршрутной траектории, в том числе и в условиях течения как стационарного, так и нестационарного.

Задача АНПА заключается в совмещении линии визирования «АНПА — $(i + 1)$ -я маршрутная точка» с идеальной траекторией. Условием выполнения совмещения линии является равенство нулю угла β между линией визирования и линией маршрутной траектории. Заданный угол курса АНПА определяется выражением [7]

$$\psi = \alpha + K_1 + K_1 K_u \int \beta dt,$$

где α — угол наклона траектории АНПА (рис. 3); K_1, K_u — коэффициенты системы управления.

Приведенные выше выражения позволяют определить текущий курс АНПА так, чтобы ему оказаться в окрестности маршрутной точки.

Схема взаимного расположения АНПА и предыдущей (i -й) и ближайшей ($(i + 1)$ -й) маршрутных точек представлена на рис. 3.

Достоинством алгоритма является то, что отсутствует колебательный характер траектории движения АНПА при переходе на новый участок галса [7]. При этом АНПА следует траекторией, близкой к идеальной, как при отсутствии, так и при наличии течения.

Однако в приведенном алгоритме восстановления положения АНПА на маршрутной траектории отсутствуют ограничения на маневренные характеристики. Для их учета модифицируем приведенный алгоритм, добавив ограничение на корректировку текущего курса АНПА:

$$\psi = \alpha + K_1 + K_1 K_u \int \beta dt$$

при $K_1 K_u \int \beta dt \leq \psi_{огр},$

где $\psi_{огр}$ — ограничение на угол поворота.

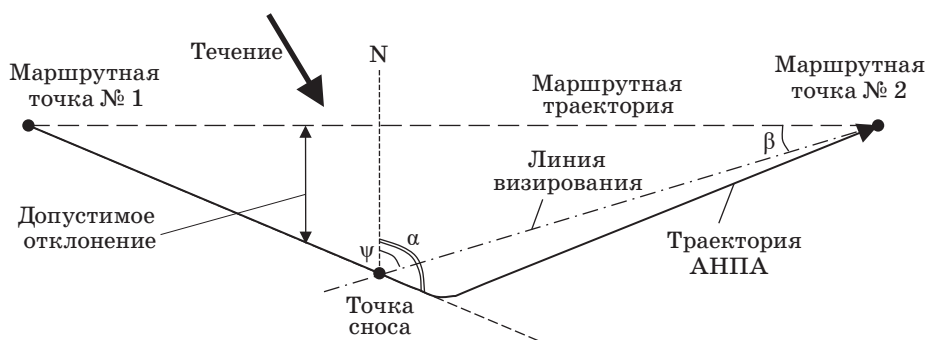
При этом указанный ранее параметр — допустимое отклонение — зависит от ограничения на скорость изменения угла поворота АНПА, вызванного его инерционностью.

Поэтому, для того чтобы можно было воспользоваться предложенным в настоящей работе методом удержания АНПА на маршрутной траектории, необходимо определить параметр метода — допустимое отклонение $D_{доп}$ с учетом ограничения на скорость изменения курса:

$$D_{доп i} = f(i, V_{теч}, R_{i,i+1}, Q_{огр}, D_{м.т}),$$

где $V_{теч}$ — скорость течения, измеряемая лагом, которым оборудован АНПА; $R_{i,i+1}$ — расстояние между заданными маршрутными точками; $Q_{огр}$ — ограничение на скорость изменения курса АНПА; $D_{м.т}$ — допустимое отклонение от маршрутной точки для прибывающего в нее АНПА.

С точки зрения периодичности корректировки положения АНПА, определяющей допустимое отклонение, необходимо отметить следующее.



■ **Рис. 3.** Схема изменения курса АНПА при восстановлении своего положения на маршрутной траектории

■ **Fig. 3.** Scheme change course of the AUV when restoring its position on the route trajectory

Автономный необитаемый подводный аппарат в ходе своего перемещения подвержен не только постоянным возмущающим факторам, таким как течение, но и случайным факторам, связанным, например, с вихревыми образованиями, турбулентностью, изменением ее параметров, а также случайными погрешностями определения собственного местоположения и т. д. В такой ситуации вести постоянную корректировку собственного положения аппарату нецелесообразно.

С другой стороны, слишком большая дистанция отклонения может привести к нехватке времени для возврата к маршрутной траектории из-за ограничений на скорость изменения угла поворота АНПА. Поэтому необходимо было определить оптимальное значение допустимого отклонения АНПА от маршрутной траектории, при котором восстановление АНПА на маршрутной траектории обеспечивает эффективное ведение сейсморазведки.

Для определения оптимального значения допустимого отклонения АНПА от маршрутной траектории необходимо было, прежде всего, выработать показатель эффективности ведения сейсморазведки. Использование показателя эффективности позволяет по результатам численного эксперимента определить оптимальное значение допустимой дистанции отклонения АНПА от маршрутной траектории.

Эффективность E ведения сейсморазведки в формализованном виде может быть представлена следующим образом:

$$E = f(i, D_{\text{доп } i}, V_{\text{теч}}, R_{i,i+1}, Q_{\text{огр}}, D_{\text{м.т}}) \rightarrow \max.$$

Показатель эффективности характеризует успешность сейсморазведки. Успешность сейсморазведки определяется погрешностью определения положения обнаруженного месторождения залежей углеводородов. Погрешность определения местоположения залежей углеводородов зависит, в свою очередь, от равномерности распределения азимутов в ходе проведения приема отраженных сигналов. Поэтому в данной работе в качестве показателя эффективности ведения сейсморазведки, по аналогии с работами [9, 10], принято среднеквадратическое отклонение (СКО) азимутов направлений прихода отраженных сигналов.

Для определения параметров эффективного удержания АНПА на маршрутной траектории было использовано математическое моделирование.

Описание математической модели

Для выбора параметра метода разработаны две математические модели: одна, так называемая макромодель, предназначена для оценки эф-

фективности ведения сейсморазведки. Модель основана на математической модели [10], в которой воспроизводились процессы перемещения излучателя, периодически посылающего импульсные сигналы в толщу морского дна, и группы АНПА. Излучатель и АНПА при моделировании представлены в виде материальных точек, движение которых определяется курсом, скоростью, дифферентом, а положение характеризуется глубиной и координатами.

Макромодель использована для численных экспериментов, направленных на определение отклонения и восстановления положения АНПА на траектории.

Математическая модель в полном объеме воспроизводит поведение излучателя и группы АНПА при ведении сейсморазведки в районе заданной площади.

В математической модели в каждом такте имитации осуществляется:

- воспроизведение движения излучателя;
- воспроизведение движения каждого АНПА группы;
- решение навигационной задачи каждым АНПА;
- постоянная оценка дистанций между АНПА группы;
- обмен информацией между АНПА;
- излучение излучателем зондирующего импульсного сигнала;
- прием геофонами отраженного от месторождений углеводородов сигнала.

Месторождение углеводородов моделировалось в виде эллипсоида, размеры которого являлись случайными и разыгрывались равномерно из заданного диапазона. Положение месторождения в заданном районе также было случайным и разыгрывалось равномерно.

В результате математическая модель и ее программная реализация позволяют оценить эффективность ведения сейсморазведки при различных отклонениях положения АНПА в момент прибытия в маршрутную точку.

Это в свою очередь дает возможность сформулировать определенные требования к параметрам метода удержания АНПА на заданной траектории, которые обеспечивали бы прибытие АНПА в маршрутную точку с погрешностями не выше тех, которые определены в макромодели.

Поэтому для оптимизации параметров метода удержания АНПА на траектории специально была разработана вторая математическая модель. Она рассматривает перемещение отдельно взятого АНПА, моделирует отклонение АНПА с маршрутной траектории, вызванное течением, которое характеризуется направлением и скоростью. В модели реализован алгоритм возвращения АНПА на маршрутную траекторию в случае

отклонения от нее на дистанцию, превышающую допустимую.

Дополнительно разработанная математическая модель позволяет провести:

- оценку отклонения каждым АНПА своего положения от заданного на маршрутной траектории;

- анализ превышения отклонения положения АНПА от маршрутной траектории, по результатам которого определяется необходимость восстановления положения АНПА на маршрутной траектории;

- корректировку собственного положения каждым аппаратом в случае отклонения от заданной траектории с использованием предложенного в настоящей работе метода.

Благодаря модели оценивалась погрешность отклонения положения АНПА от маршрутной точки в момент его прибытия в ближайшую маршрутную точку.

Результатом использования этой математической модели является получение погрешности прибытия АНПА в окрестность маршрутной точки при различных параметрах метода удержания аппарата на траектории.

С использованием описанных двух математических моделей и их программных реализаций проведены исследования по выбору параметров метода удержания АНПА на маршрутной траектории.

Результаты исследований по оптимизации параметров метода удержания АНПА на маршрутной траектории

Исследования по оптимизации параметров метода удержания АНПА на маршрутной траектории проводились в два этапа.

На первом этапе определялись допустимые отклонения положения АНПА в момент прибытия

в маршрутную точку, при которых ведение сейсморазведки являлось эффективным.

На втором этапе полученные допустимые отклонения использовались для определения оптимального значения допустимой дистанции. При этом учитывались следующие параметры:

- ограничения на угловую скорость поворота АНПА при маневрировании;

- расстояние между маршрутными точками.

Рассматривались различные варианты течения, вызывающего отклонение АНПА от маршрутной траектории.

Результаты моделирования при различных параметрах тактической ситуации позволили оптимизировать параметры метода удержания положения АНПА на маршрутной траектории, обеспечивающего минимальную погрешность отклонения положения аппарата от маршрутной точки.

По полученным результатам определялись варианты сочетания параметров метода удержания АНПА на маршрутной траектории.

В ходе проведения численного эксперимента в качестве исходных данных приняты:

- размер района ведения сейсморазведки 1000×1000 м;

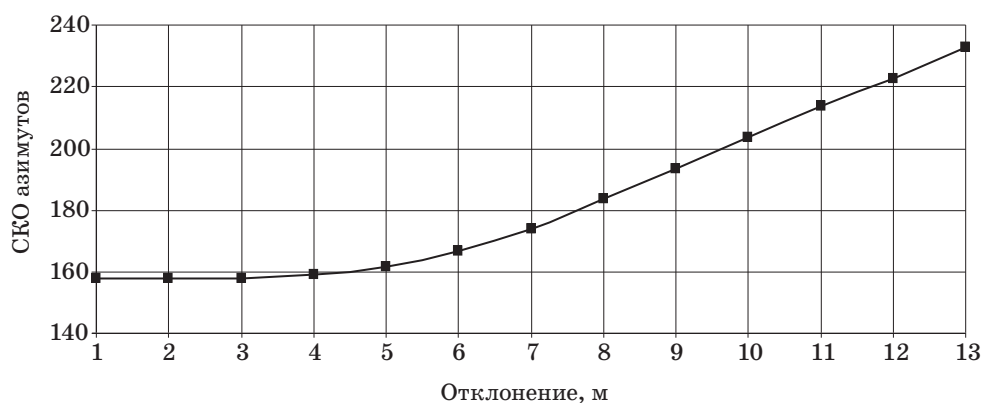
- скорость движения АНПА 1 уз;

- количество аппаратов в группе 49: 7×7 ед.

Рассматривались варианты оценки эффективности сейсморазведки при различном отклонении АНПА от маршрутной траектории в момент его прибытия в маршрутную точку.

Размер отклонения варьировался в диапазоне от 1 до 13 м, направление отклонения разыгрывалось равновероятно относительно маршрутной точки. Результаты исследований в виде зависимости СКО азимутов от погрешности прибытия АНПА в маршрутную точку представлены на рис. 4.

Зависимость носит монотонный характер, и с увеличением погрешности прибытия АНПА в маршрутную точку происходит увеличение СКО



■ Рис. 4. Зависимость СКО азимутов от погрешности прибытия АНПА в маршрутную точку

■ Fig. 4. Dependence of standard deviation azimuths on the arrival error of the AUV on the waypoint

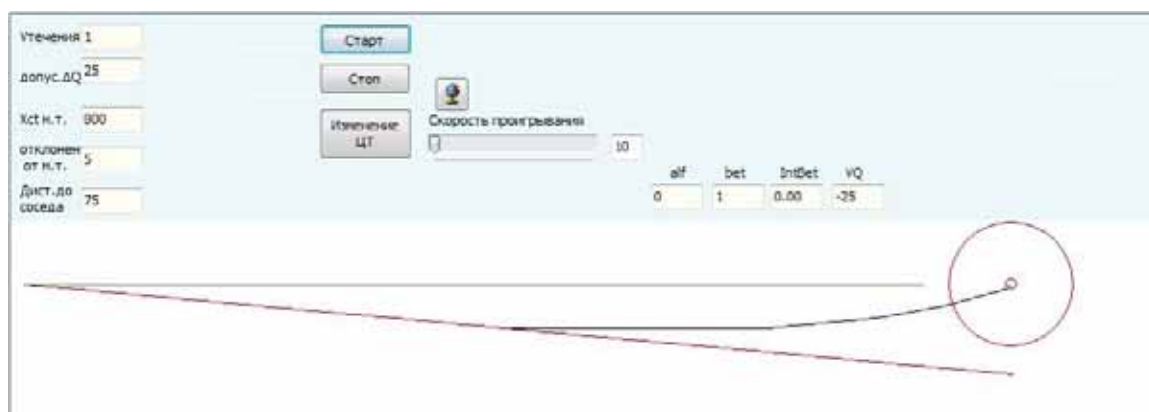
азимуты от 160 до 235. Представленный график позволяет определить допустимые значения погрешности прибытия АНПА в окрестность маршрутной точки. Так, при принятии решения о том, что СКО азимуты в пределах до 180 позволяет считать ведение сейсмозазведки эффективным, означает, что допустимое отклонение АНПА от маршрутной точки не должно превышать 7,7 м.

После того как с использованием макромодели были получены ограничения по допустимому отклонению АНПА от маршрутной точки, на следующем этапе исследований они были использованы для выбора параметров корректировки положения АНПА.

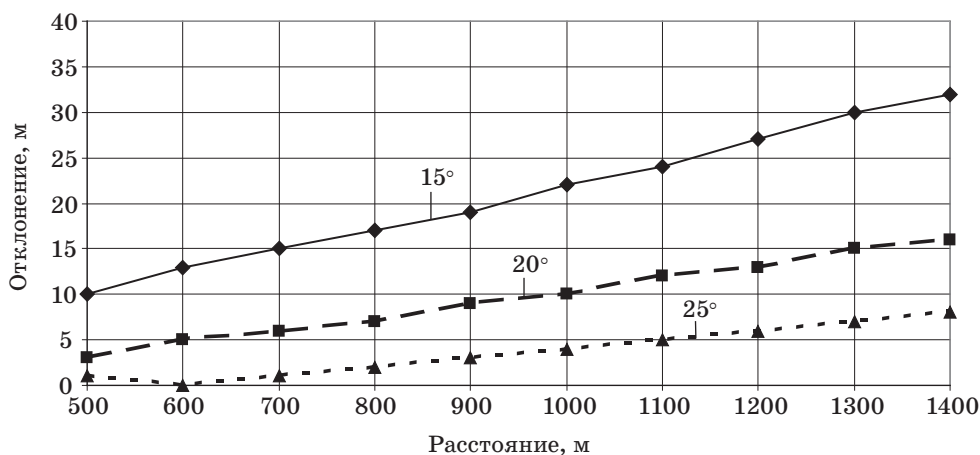
На экране программы, с использованием которой происходило исследование (рис. 5), линия зеленого цвета отражает маршрутную траекторию, линия красного цвета — траекторию АНПА в отсутствие корректировки, линия синего цвета —

траекторию АНПА с учетом корректировки курса в момент выявления отклонения. Положение маршрутной точки показано кружком малого радиуса, окрестность маршрутной точки ограничена кругом большого радиуса.

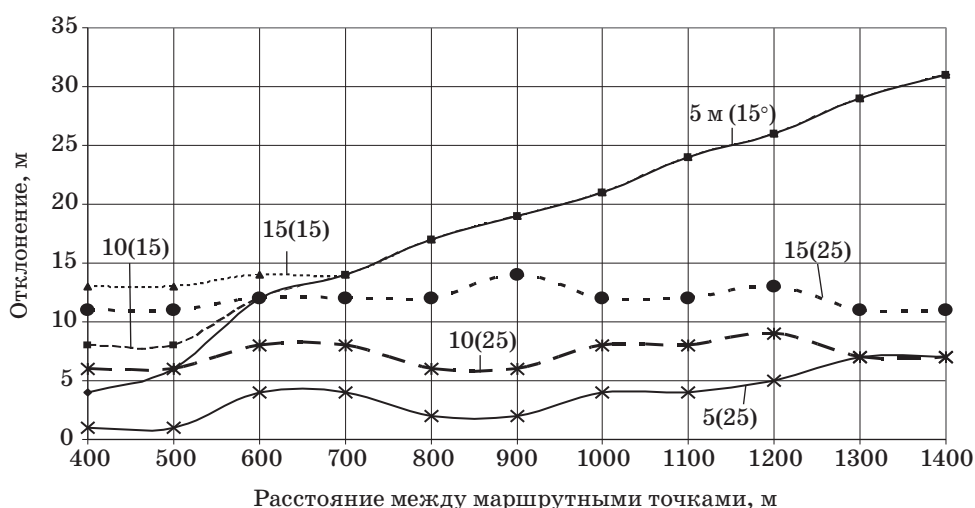
При проведении исследований принималось, что допустимая дистанция — фиксированная (составляла 25 м), и целью эксперимента было определение влияния расстояния между маршрутными точками на погрешность местоположения АНПА при различных ограничениях на маневрирование по курсу. Рассматривались варианты расположения маршрутных точек в диапазоне от 500 до 1400 м с шагом 100 м. Ограничения на маневр по курсу рассматривались в трех вариантах: 15, 20 и 25° за время, допустимое для возвращения на траекторию. Значение допустимой дистанции принималось равным 25 м. Результаты исследований представлены на рис. 6.



■ Рис. 5. Окна программы для определения оптимальных параметров метода
 ■ Fig. 5. Windows program to determine the optimal parameters of the method



■ Рис. 6. Зависимость отклонения положения АНПА от расстояния между маршрутными точками при различных ограничениях на маневр АНПА
 ■ Fig. 6. Dependence of the deviation of the position of the AUV on the distance between the waypoints under various restrictions on the maneuver of the AUV



■ **Рис. 7.** Зависимость отклонения положения АНПА от допустимой дистанции, расстояния между маршрутными точками и различными ограничениями на маневр по курсу

■ **Fig. 7.** Dependence of the deviation of the position of the AUV from the permissible distance, the distance between the waypoints and various restrictions on maneuvering at the rate

Из рисунка видно, что дистанция между маршрутными точками и ограничения на изменение курса оказывают влияние на отклонение положения АНПА от маршрутной точки. Так, при изменении дистанции между маршрутными точками от 400 до 1400 м погрешность увеличилась от 3 до 7 м при ограничении на изменение курса, равное 15° , от 5 до 17 м — при ограничении на изменение курса 20° и от 10 до 34 м — при ограничении на изменение курса 25° .

На следующем этапе исследований с использованием этой же математической модели оценивалось влияние допустимой дистанции на погрешность попадания АНПА в окрестность маршрутной точки. Рассматривались ограничения на курс 15° и 25° . Допустимая дистанция при проведении исследований принималась 15, 20 и 25 м.

При проведении исследований рассматривались следующие варианты:

— ограничения на изменение курса 15° и 25° ;

— допустимая дистанция, на которой АНПА принимает решение относительно корректировки курса для восстановления своего положения на маршрутной траектории, 5, 10, 15 м;

— расстояние между маршрутными точками в диапазоне от 400 до 1400 м с шагом 100 м.

Результаты исследований (рис. 7) позволяют определить оптимальные значения допустимой дистанции в зависимости от погрешности положения АНПА в окрестности маршрутной точки. Так, например, в том случае, если по результатам первого этапа выявлено, что допустимым значением погрешности является дистанция 8 м, то это значение обеспечивают только варианты:

— при ограничении на скорость изменения курса 15° допустимые дистанции 5 и 10 м при расстоянии до маршрутной точки 400–500 м;

— при ограничении 25° допустимые дистанции 5 и 10 м для всех дистанций между маршрутными точками в диапазоне 400 ÷ 1400 м.

В том случае, если по результатам моделирования с использованием макромодели принято значение погрешности 12 м, что соответствует СКО = 220, то допустимыми можно считать варианты:

— при ограничении на изменение курса 15° , в дополнение к указанным ранее, и варианты расстояния между маршрутными точками 600 м;

— при ограничении на изменение курса 25° — практически все рассмотренные дистанции между маршрутными точками, т. е. от 400 до 1400 м, и при допустимых отклонениях от маршрутной траектории 5, 10 и 15 м.

Таким образом, полученные результаты моделирования позволяют определить допустимые дистанции отклонения АНПА от маршрутной траектории, на которых принимается решение относительно корректировки положения АНПА, обеспечивающее эффективное ведение сейсморазведки.

Заключение

Разработанный метод удержания АНПА на маршрутной траектории в случае отклонения его на дистанцию, превышающую допустимую, позволяет прибыть АНПА в маршрутную точку с погрешностью не выше заданной — с учетом

ограничений, налагаемых на маневрирование АНПА по курсу.

Разработанные математические модели: оценки эффективности сейсморазведки и оценки отклонения АНПА от маршрутной точки с реализацией предложенного метода удержания АНПА на маршрутной траектории — были использованы для определения оптимальных параметров метода, обеспечивающих эффективное ведение сейсморазведки.

Результаты моделирования позволили определить значения параметров метода удержания АНПА на маршрутной траектории с учетом рас-

стояния между маршрутными точками, ограничений на скорость поворота АНПА, допустимой дистанции отклонения АНПА от маршрутной траектории, обеспечивающих эффективное ведение поиска.

Результаты исследований позволяют учитывать особенности влияния внешних факторов на движение АНПА и его маневренные характеристики при подготовке маршрутных заданий каждому АНПА.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 17-08-00666).

Литература

1. **Фасхутдинов Р. В.** Подводные камни морской сейсморазведки // *www.Korabel.ru*. 2016. Вып. 2(32). С. 98–106.
2. **Лаверов Н. П., Дмитриевский А. Н., Боговяленский В. И.** Фундаментальные аспекты освоения нефтегазовых ресурсов арктического шельфа России// *Арктика: экология и экономика*. 2011. № 1. С. 25–37.
3. **Лаверов Н. П., Лобковский Л. И., Тулупов А. В., Воронов М. А., Ганжа О. Ю.** Перспективы донной сейсморазведки в Российской Федерации // *Арктика: экология и экономика*. 2011. № 4. С. 4–13.
4. **Крылов А. А.** Некоторые аспекты донных сейсмологических исследований и мониторинга на объектах нефтегазовой добычи и транспортировки на шельфе// *Защита окружающей среды в нефтегазовом комплексе*. 2015. № 9. С. 20–25.
5. Пат. 2381530 РФ. Морская автономная донная станция для выполнения геофизических и геологоразведочных работ/ А. И. Машошин, С. В. Жуменов, В. Б. Зиннатов, С. Ю. Ермаков, А. С. Смирнов (РФ). — № 2008124873/28; заявл. 10.06.2008; опубл. 10.02.2010, Бюл. № 4. — 8 с.
6. Пат. 2438149 РФ. Автономная донная станция для сейсмических наблюдений/ И. А. Ильин, П. Г. Бродский, В. С. Аносов, Д. Г. Левченко, Е. Р. Павлюкова, В. П. Ленюков, В. В. Чернявец, А. А. Зайцев, А. Л. Гвоздецкий (РФ). — № 2010103093/28; заявл. 29.01.2010; опубл. 27.12.2011, Бюл. № 36. — 14 с.
7. Пат. 2435180 РФ. Подводная геофизическая станция/ С. Я. Суконкин, С. В. Белов, А. В. Кошурников, П. Ю. Пушкарев, Н. П. Рыбаков, С. Ю. Чернявец, В. В. Червинчук (РФ). — № 2010113601/28, заявл. 07.04.2010; опубл. 27.11.2011, Бюл. № 33. — 13 с.
8. **Максимов С. В.** О применении автономных необитаемых подводных аппаратов в технологии тотальной донной сейсморазведки// *Технические проблемы освоения Мирового океана: тр. 4-й Всерос. науч.-техн. конф., Владивосток, 2011. Секция 2. 2011. С. 220–224.*
9. **Мартынова Л. А.** Метод согласованного поведения излучателя и автономных необитаемых подводных аппаратов для эффективного ведения сейсморазведки // *Информационно-управляющие системы*. 2017. № 1. С. 83–92. doi:10.15217/issn1684-8853.2017.1.83
10. **Мартынова Л. А.** Инструментарий для исследований эффективности ведения сейсморазведки с использованием автономных необитаемых подводных аппаратов // *Информационно-управляющие системы*. 2017. № 2. С. 77–87. doi: 10.15217/issn1684-8853.2017.2.77
11. **Мартынова Л. А., Машошин А. И.** Построение системы управления автономных необитаемых подводных аппаратов на базе мультиагентной технологии // *Изв. ЮФУ. Технические науки*. 2016. № 2. С. 38–48.
12. **Степанов О. А., Торопов А. Б.** Методы нелинейной фильтрации в задаче навигации по геофизическим полям. Ч. 1: Обзор алгоритмов // *Гироскопия и навигация*. 2015. Т. 6. № 3. С. 102–125. doi:10.17285/0869-7035.2015.23.3.102-125
13. **Степанов О. А., Торопов А. Б.** Методы нелинейной фильтрации в задаче навигации по геофизическим полям. Ч. 2: Современные тенденции развития // *Гироскопия и навигация*. 2015. № 4. С. 147–159. doi:10.17285/0869-7035.2015.23.4.147-159
14. **Шпаков В. М.** О ситуационно-событийном подходе к управлению взаимодействующими дискретно-непрерывными процессами // *Информационно-управляющие системы*. 2017. № 2. С. 26–33. doi:10.15217/issn1684-8853.2017.2.26
15. **Бритов Г. С.** Терминальное диагностирование дискретных динамических систем // *Информационно-управляющие системы*. 2017. № 4. С. 18–24. doi:10.15217/issn1684-8853.2017.4.18
16. **Фуртат И. Б.** Алгоритм компенсации помех измерения и возмущений // *Информационно-управляющие системы*. 2017. № 5. С. 21–29. doi:10.15217/issn1684-8853.2017.5.21

17. Малышкин Г. С. Сравнительная эффективность классических и быстрых проекционных алгоритмов при разрешении слабых гидроакустических сигналов // Акустический журнал. 2017. Т. 63. № 2. С. 196–208.
18. Edwards D. B. A Leader-Follower Algorithm for Multiple AUV Formations // Proc. of IEEE/OES Autonomous Underwater Vehicles. 2004. 10 p.
19. Edwards D. B. Decentralized Control of Multiple Autonomous Underwater. https://www.researchgate.net/publication/235092321_Decentralized_Control_of_Multiple_Autonomous_Underwater_Vehicles (дата обращения: 15.11.2017).
20. Лямина Е. А. Алгоритмы управления движением группы АНПА по поисковым траекториям. <http://sntbul.bmstu.ru/doc/637281.html> (дата обращения: 15.11.2017).
21. Rastgoftar H. Continuum Deformation of Multi-Agent Systems. — Springer International Publishing, 2016. — 217 p.

UDC 519.87

doi:10.15217/issn1684-8853.2018.3.34

Effective Retention of AUV Position on Route Trajectory during Seismic SurveyMartynova L. A.^a, Dr. Sc., Tech., Associate Professor, martynowa999@bk.ru^aConcern CSRI Elektropribor, JSC State Research Center of Russia, 30, Malaya Posadskaya St., 197046, Saint-Petersburg, Russian Federation

Introduction: during subglacial seismic survey, autonomous underwater vehicles can deviate from the route trajectory due to currents and other random non-stationary factors. To maintain an effective survey, it is necessary to permanently keep the device on the route trajectory. **Purpose:** Developing a method for effective retention of an AUV on a route trajectory. **Method:** An AUV which has deviated from the route trajectory is sent back to it by a method of aiming at the target, which uses the sight line deflecting angle. **Results:** Mathematical dependencies of the deviation from the route trajectory were obtained. A method of restoring the device on the route trajectory was developed and modified in order to take into account the limitations on the maneuvering characteristics of the device. A mathematical model was built, and parameterization of the proposed method was conducted on the basis of seismic survey efficiency evaluation. The simulation results made it possible to determine the parameters of the method used to hold an AUV on the route trajectory, taking into account the distance between the waypoints, the restrictions on the device rotation speed and the permissible deviation from the route trajectory. **Practical relevance:** The developed method allows you to efficiently restore the position of an AUV on the route trajectory, taking into account the limitations and peculiarities of conducting a seismic survey in a given area.

Keywords — Autonomous Underwater Vehicle, Route Trajectory, Water Currents, Efficiency Evaluation.

Citation: Martynova L. A. Effective Retention of AUV Position on Route Trajectory during Seismic Survey. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 34–44 (In Russian). doi:10.15217/issn1684-8853.2018.3.34

References

- Faskhutdinov R. V. Reefs of Naval Seismic Exploration. *www.Korabel.ru*, 2016, iss. 2(32), pp. 98–106 (In Russian).
- Laverov N. P., Dmitriyevsky A. N., Bogoyavlensky V. I. Fundamental Aspects of Development of Oil and Gas Resources of the Arctic Shelf of Russia. *Arktika: ekologiya i ekonomika* [Arctic: Ecology and Economy], 2011, no. 1, pp. 25–37 (In Russian).
- Laverov N. P., Lobkovsky L. I., Tulupov A. V., Ravens M. A., Ganzha O. Y. The Prospects of Ground Seismic Exploration in the Russian Federation. *Arktika: ekologiya i ekonomika* [Arctic: Ecology and Economy], 2011, no. 4, pp. 4–13 (In Russian).
- Krylov A. A. Some Aspects of Ground Seismological Researches and Monitoring on Objects of Oil and Gas Production and Transportation on the Shelf. *Zashchita okruzhaiushchei sredy v neftegazovom komplekse* [Environment Protection in an Oil and Gas Complex], 2015, no. 9, pp. 20–25 (In Russian).
- Mashoshin A. I., Zhumenkov S. V., Zinnatov V. B., Ermakov S. Y., Smirnov A. S. *Morskaya avtonomnaya donnaia stantsiya dlia vypolneniya geofizicheskikh i geologorazvedochnykh rabot* [The Sea Autonomous Ground Station for Accomplishment of Geophysical and Exploration Works]. Patent RU, no. 2381530, 2008.
- Il'in I. A., Brodskij P. G., Anosov V. S., Levchenko D. G., Pavljukova E. R., Len'kov V. P., Chernjavets V. V., Zajtsev A. A., Gvozdet'skij A. L. *Avtonomnaya donnaia stantsiya dlia seismicheskikh nabludений* [Independent Bottom Station for Seismic Observations]. Patent RU, no. 2438149, 2010.
- Sukonkin S. J., Rybakov N. P., Belov S. V., Chervinchuk S. J., Koshurnikov A. V., Pushkarev P. J., Chernjavets V. V. *Podvodnaya geofizicheskaya stantsiya* [Underwater Geophysical Station]. Patent RU, no. 2435180, 2010.
- Maximov S. V. About use of Autonomous Uninhabited Submersibles in Technology of Total Ground Seismic Exploration. *Trudy 4-i Vserossiiskoi nauchno-tekhnicheckoi konferentsii "Tekhnicheskie problemy osvoeniia Mirovogo okeana"* [Proc. of the 4th All-Russian Scientific and Technical Conf. "Engineering Problems of Development of the World Ocean"], Vladivostok, 2011, Section 2, pp. 220–224 (In Russian).
- Martynova L. A. Concerted Action of a Radiator and Autonomous Uninhabited Submersibles for Effective Seismic Exploration. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 1, pp. 83–92 (In Russian). doi: 10.15217/issn1684-8853.2017.1.83
- Martynova L. A. Tools for Studying the Efficiency of Seismic Exploration with Standalone Unmanned Submersibles. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 2, pp. 77–87 (In Russian). doi: 10.15217/issn1684-8853.2017.2.77
- Martynova L. A., Mashoshin A. I. Formation of AUV Control System based on Multi-Agent Technology. *Izvestiia IuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, no. 2, pp. 38–48 (In Russian).
- Stepanov O. A., Toropov A. B. Nonlinear Filtering for Map-Aided Navigation. Part 1. An Overview of Algorithms. *Giroskopiia i navigatsiia* [Gyroscopy and Navigation], 2015, vol. 6, no. 3, pp. 324–337 (In Russian).
- Stepanov O. A., Toropov A. B. Nonlinear Filtering for Map-Aided Navigation. Part 2. Current Trends of Develop-

- ment. *Giroskopiia i navigatsiia* [Gyroscopy and Navigation], 2015, vol. 6, no. 4, pp. 147–159 (In Russian).
14. Shpakov V. M. About Situation-Event Approach to Control of Interacting Discrete-Continuous Processes. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 2, pp. 26–33 (In Russian). doi:10.15217/issn1684-8853.2017.2.26
 15. Britov G. S. Terminal Diagnostics of Discrete Dynamic Systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 4, pp. 18–24 (In Russian). doi:10.15217/issn1684-8853.2017.4.18
 16. Furtat I. B. Algorithm for Compensation of Measurement Noises and Disturbances. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2017, no. 5, pp. 21–29 (In Russian). doi:10.15217/issn1684-8853.2017.5.21
 17. Malyskhin G. S. The Comparative Efficiency of Classical and Fast Projection Algorithms in the Resolution of Weak Hydroacoustic Signals. *Akusticheskiy zhurnal* [Acoustical Physics], 2017, vol. 63, no. 2, pp. 216–228 (In Russian).
 18. Edwards D. B. A Leader-Follower Algorithm for Multiple AUV Formations. *Proc. of IEEE/OES Autonomous Underwater Vehicles*, 2004, 10 p.
 19. Edwards D. B. *Decentralized Control of Multiple Autonomous Underwater*. Available at: https://www.researchgate.net/publication/235092321_Decentralized_Control_of_Multiple_Autonomous_Underwater_Vehicles (accessed 15 November 2017).
 20. Lyamina E. A. *Algoritmy upravleniia dvizheniem gruppy ANPA po poiskovym traektoriiam* [Algorithms of Motion Control of the AUV Group on Search Trajectories]. Available at: <http://sntbul.bmstu.ru/doc/637281.html> (accessed 15 November 2017).
 21. Rastgoftar H. *Continuum Deformation of Multi-Agent Systems*. Springer International Publishing, 2016. 217 p.

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, электронные адреса авторов, которые по требованию ВАК должны быть опубликованы на страницах журнала. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисуночные подписи и названия таблиц на русском и английском языках.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени — эта информация будет опубликована в ссылке на первой странице.

Формулы набирайте в Word, не используя формульный редактор (Mathtype или Equation), при необходимости можно использовать формульный редактор; для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта никогда не пользуйтесь вкладкой Other..., используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; в формулах не отделяйте пробелами знаки: + = -.

Для набора формул в Word никогда не используйте Конструктор (на верхней панели: «Работа с формулами» — «Конструктор»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Иллюстрации предоставляются отдельными исходными файлами, поддающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Coreldraw (*.cdr); Excel (*.xls); Word (*.docx); Adobe Illustrator (*.ai); AutoCad (*.dxf); Matlab (*.ps, *.pdf или экспорт в формат *.ai);

— если редактор, в котором Вы изготавливаете рисунок, не позволяет сохранить в векторном формате, используйте функцию экспорта (только по отношению к исходному рисунку), например, в формат *.ai, *.esp, *.wmf, *.emf, *.svg;

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисуночных подписей обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (<http://i-us.ru/paperrules>) по разным стандартам: Литература — СИБИБД РФ, References — один из мировых стандартов.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Оформление статей».

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Кому: Редакция журнала «Информационно-управляющие системы»
Тел.: (812) 494-70-02
Эл. почта: i-us.spb@gmail.com
Сайт: www.i-us.ru

УДК 519.876.5, 519.873, 004.023:896
doi:10.15217/issn1684-8853.2018.3.45

АЛГОРИТМ МНОГОКРИТЕРИАЛЬНОГО ПОИСКА ТРАЕКТОРИЙ ДВИЖЕНИЯ РОБОТА НА МНОГОСЛОЙНОЙ КАРТЕ

Д. Е. Моторин^а, аспирант, d.e.motorin@gmail.com

С. Г. Попов^а, канд. техн. наук, доцент, popovserge@spbstu.ru

^аСанкт-Петербургский политехнический университет Петра Великого, Политехническая ул., 29, Санкт-Петербург, 195251, РФ

Постановка проблемы: сложная среда характеризуется возможностью декомпозиции воздействующих на робота факторов на независимые слои. В процессе движения робота в сложной среде на него воздействуют негативные факторы, влияющие на возможность достижения цели. В связи с этим возникает проблема выбора траектории движения при одновременной минимизации нежелательных воздействий на робота и длины пройденного пути. **Цель:** разработка и анализ результатов работы алгоритма двухкритериальной оптимизации траектории движения робота с учетом желаемых критериев взаимодействия со средой и длины траектории. **Результаты:** разработан и реализован алгоритм поиска кратчайшего пути на карте, каждый слой которой отображает свойство пространства и позволяет учитывать взаимодействие робота и среды с учетом длины пройденной траектории. Реализация алгоритма встроена в модель управления группой роботов. Для анализа алгоритма рассмотрены тестовые многослойные карты с добавлением гауссова шума. В результате симуляции построено множество траекторий движения с учетом коэффициентов влияния свойств пространства на объект управления при заданных начальном и конечном положении на карте. Построено пространство состояний движения робота, представленное в виде зависимостей влияния свойств среды на робота от длины траектории и риска отказов на протяжении всего пути. **Практическая значимость:** разработанный алгоритм может быть применен в системах планирования индивидуального или группового движения роботов. Полученное пространство состояний отражает диапазоны эффективных характеристик робота при выполнении поставленных задач в заданной среде. Продолжением работы станет применение разработанного алгоритма при поиске пути на разномасштабных картах и построение сетей траекторий в пространстве состояний группы роботов.

Ключевые слова — поиск траекторий, многослойные карты, управление, робот, пространство состояний, эвристический алгоритм, реалистичная среда.

Цитирование: Моторин Д. Е., Попов С. Г. Алгоритм многокритериального поиска траекторий движения робота на многослойной карте // Информационно-управляющие системы. 2018. № 3. С. 45–53. doi:10.15217/issn1684-8853.2018.3.45

Citation: Motorin D. E., Popov S. G. Multi-Criteria Path Planning Algorithm for a Robot on a Multilayer Map. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 45–53 (In Russian). doi:10.15217/issn1684-8853.2018.3.45

Введение

Управление роботами и их группами включает в себя ряд сложных задач, начиная с механических принципов движения и заканчивая интеллектуальными алгоритмами принятия решений о выборе глобальных целей для выполнения. В настоящий момент центральной проблемой управления и планирования движения является поиск траекторий. Построение пути на плоской карте с дискретными препятствиями является хорошо изученной задачей как в теоретическом, так и в практическом плане. Для ее решения, как правило, используются классические алгоритмы поиска на графах, такие как, например, алгоритмы Дейкстры, A^* , волновой, Флойда — Уоршелла [1, 2]. Они широко и успешно применяются при наличии одного слоя карты. При этом сложно распространить их на многослойные карты, учитывающие несколько свойств среды.

Сравнение стандартных алгоритмов поиска траекторий рассматривается в работах [2–4]. Авторы оптимизируют алгоритмы A^* , Дейкстры

и собственный алгоритм НСТН для систем с низкими вычислительными ресурсами. Анализ показал, что алгоритм A^* наиболее эффективен по критерию минимизации длины траектории. Из стандартных поисковых алгоритмов наиболее распространенный для исследования и модификации также алгоритм A^* . Например, его модификация [5] позволяет минимизировать число перспективных вершин в процессе поиска, а эвристика, оценивающая перспективность перехода на основе переменной длины шага поиска [6], сокращает время получения результата более чем в два раза. Критерий поискового алгоритма [7] учитывает не только время получения результата, но и динамику среды, состоящую в перемещении препятствий, что позволяет избегать коллизий в процессе движения. Модифицированные способы применения поисковых алгоритмов дают возможность решать сложные задачи планирования гибридных траекторий в двух приближениях [8] или разрешать коллизии траекторий [9] с использованием алгоритмов поиска промежуточных целевых точек.

Помимо классических алгоритмических подходов используются эволюционные алгоритмы [10], в частности генетические [11]. В работе [12] рассматривается модификация алгоритма с переменной длиной хромосом с апробацией в натурном эксперименте по движению робота. Сокращение вычислительных затрат [13] обеспечивается реализацией иерархического генетического алгоритма с редукцией области поиска траекторий. Генетические алгоритмы используются при построении траектории для задачи максимального покрытия, в работе [14] предложен критерий минимизации затрачиваемой роботом энергии.

Сравнительно новым подходом при поиске траекторий является муравьиный алгоритм [15–17]. Модификация этого метода [18], состоящая в редукции сложности задачи коммивояжера, указанием обязательного посещения желаемых узлов применена для решения задачи построения индивидуальных туристических маршрутов. Оптимизация муравьиного алгоритма для статических карт различного размера с типовым и случайным распределением препятствий представлена в работе [19], где исследована зависимость длины пути от размера популяции. Алгоритм [20] предоставляет решение задачи поиска траектории движения транспортного средства в реальном времени в городских условиях при имеющемся прогнозе дорожной обстановки.

Указанные выше методы и алгоритмы рассматривают двумерные плоские карты с дискретными препятствиями, что не всегда объективно отражает ситуацию в окружающей среде. Для более точного описания обстановки используются объемные трехмерные карты. Например, в статье [21] рассматривается оптимизация траектории движения самолетов относительно совместной минимизации критериев расстояния и времени полета с заданием начальных параметров оптимизации. Для поиска траектории летающих роботов с вращающимся крылом [22] применяется комбинация алгоритма A^* и эвристической функции оценки перспективности следующего узла, в которой используется значение потенциального поля, что позволяет строить плавные траектории.

В рассмотренных ранее алгоритмах и методах применяются однослойные карты, что существенно снижает возможности управления, эффективность построенных путей и увеличивает неопределенность и вероятность отказа роботов во время движения. Поиск траекторий в сложных средах, как в работе [23], предполагает вычисление траектории движения надводного робота объединением векторного поля и маски запрещенных к движению областей. Предложенный подход [24] позволяет строить траектории в условиях динамической карты, состоящей из объем-

ного слоя поверхности ландшафта и слоя с областями вероятных отказов.

Если рассматривать окружающий мир как набор отдельных слоев, содержащих измеримые значения свойств среды, то задача может быть решена методами многокритериальной оптимизации. В данной статье авторами предложен эвристический алгоритм поиска траекторий на многослойных картах.

Постановка задачи многокритериального поиска на многослойной карте

Рассматривается карта E , размеченная регулярной сеткой, состоящая из m слоев. Каждый слой отображает свойство пространства, например высоту ландшафта местности над уровнем моря, радиационный фон, температуру. Каждый узел сетки принимает значения в диапазоне определенным выбранным свойством слоя от минимального до максимального значения $[\min_i \dots \max_i]$.

В начальный момент времени робот R расположен в начальной точке сетки A_R и должен переместиться в целевую точку Z_R . Требуется построить траекторию $L = \{A_R, l_1, \dots, l_w, \dots, Z_R\} \in L : l_w \in E$, представляющую собой набор точек сетки, в которые последовательно проходит робот во время движения.

При этом робот строит траекторию движения с учетом вектора $G = \{g_1, \dots, g_m\}$, отражающего эмпирические коэффициенты подверженности робота свойствам среды.

Планирование траектории движения осуществляется на множестве E — окружающей робота среде:

$$E = \{E_1, \dots, E_i \leftarrow [e_{xyi} \in \zeta_i], \dots\}, i = 1, \dots, m,$$

где E_i — отдельный слой, характеризующий свойство, влияющее на робота, каждый элемент e_{xyi} которого отражает значение свойства из возможных значений $\zeta_i \in [\min_i, \dots, \max_i]$ в координатах сетки (x, y) . Поиск траектории осуществляется посредством эвристического правила

$$F(e_{xyi}) = h(f(g_i \times e_{xyi}, g'_i \times e'_{xyi}), \rho) \rightarrow \min, \quad (1)$$

где e_{xyi} — значение свойства в i -м слое пространства на траектории движения робота; g_i — коэффициент влияния слоя на вероятность отказа объекта управления; ρ — расстояние от текущего положения до целевого; h — функция преобразования значений слоев в вероятность отказа; f — функция, задающая отношение слоев:

$$f = g_1 e_{xy1} + g_2 e_{xy2} + g_3 e_{xy3}. \quad (2)$$

Представленное правило используется при формировании критерия

$$\sum F(L) \rightarrow \min, \quad (3)$$

где L — траектория движения робота.

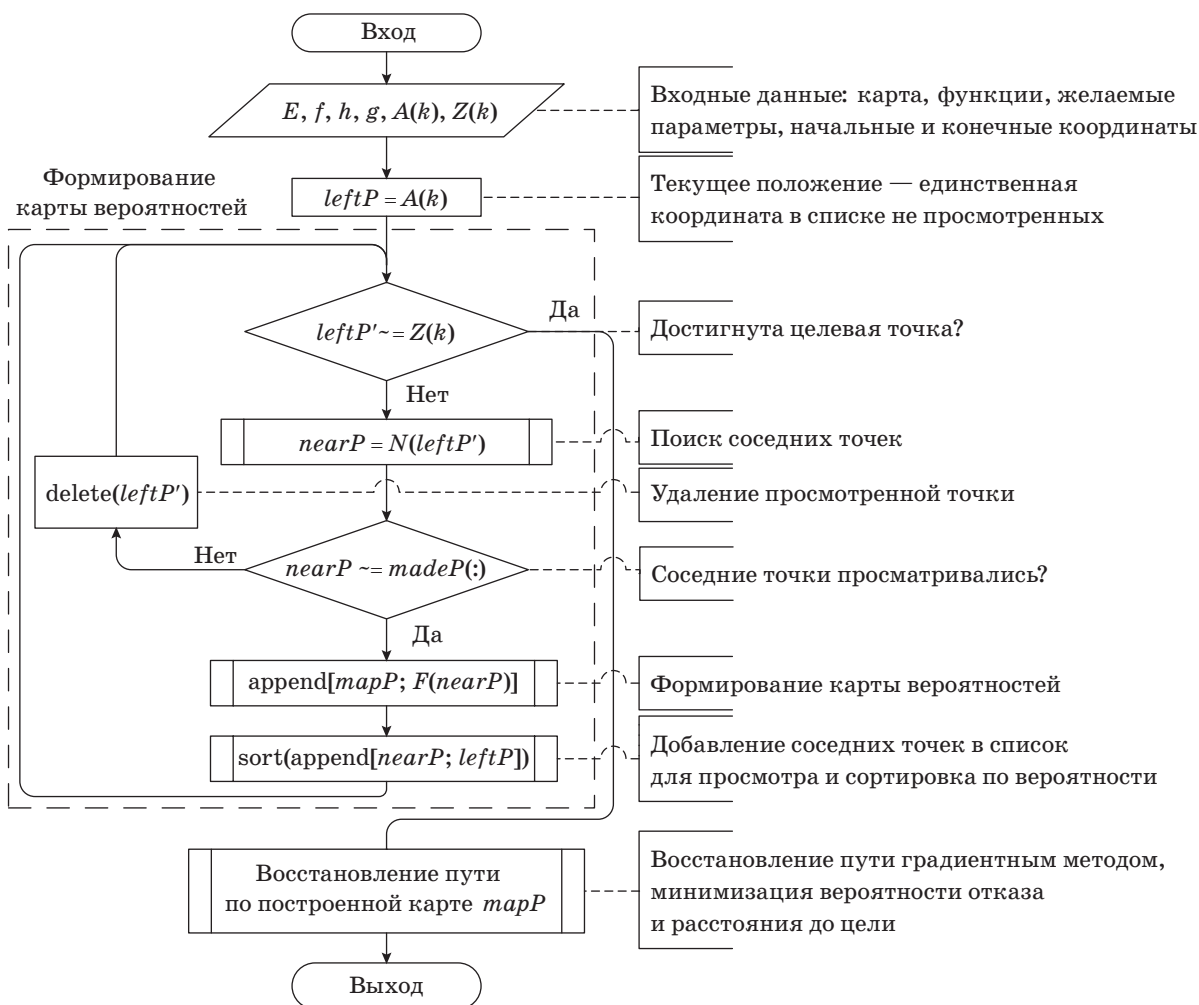
В основе алгоритма лежит минимизация вероятности отказа робота при переходах из одной дискретной точки в другую и минимизация геометрического расстояния, пройденного роботом.

Алгоритм поиска траектории на многослойной карте

Для решения поставленной задачи разработан алгоритм поиска траекторий на многослойной карте (рис. 1). Представленный алгоритм выполняется в два этапа: на первом формируется карта вероятностей, на втором строится траектория движения.

Формирование карты начинается с текущего положения объекта управления $A(k)$ в пространстве карты E , размеченной регулярной сеткой на каждом слое, которое записывается в вектор просматриваемых значений $leftP$. Далее анализируются соседние ячейки сетки $nearP$, если они еще не были просмотрены (вектор $madeP$). В найденных точках $nearP$ рассчитываются вероятности отказа оборудования, после чего точки добавляются в общую карту вероятностей $mapP$. Точки $nearP$ и $leftP$ объединяются в один список и сортируются по величине суммы вероятности отказа. Для следующих итераций берется точка с минимальной суммой F . Итерации заканчиваются в момент достижения конечной точки $Z(k)$, так как сумма вероятностей отказов в ней будет минимальной, а переход из другого положения увеличит вероятность отказа.

Формируемая карта вероятности $mapP$ на начальном этапе состоит из запрещенных для про-



■ **Рис. 1.** Алгоритм формирования карты вероятностей отказа и поиска траекторий
 ■ **Fig. 1.** The algorithm for generating failure probability maps and path planning

хода значений, заменяемых в процессе выполнения первого этапа поисков на вычисленные с помощью выражения (1) значения, при этом итоговая карта $mapP$ не является многослойной.

На втором шаге используется стандартный метод градиентного спуска для поиска траекторий по минимальным значениям вероятности отказов, начиная с конечной точки.

Реализация алгоритма

Алгоритм реализован в пакете MatLab и встроено в модель системы планирования движения группы роботов как программная функция поиска траекторий движения робота из текущего положения в целевую точку `robot_search_multilayer_path`.

Входными данными являются координаты начальной и конечной точек движения, желаемые параметры оптимизации в виде вектор-столбца и пространство действия.

Выходными данными является траектория как последовательность точек перемещения робота.

Основной цикл функции работает до тех пор, пока не найдена конечная точка либо пока не кончатся не просмотренные точки на карте. В ходе его реализации из упорядоченной очереди $leftP$ берется координата с минимальной суммарной вероятностью отказа и расстоянием до конечной точки $leftP'$. Она имеет группу точек-соседей четырехсвязной области $nearP$. Каждая точка из $nearP$ проверяется на выход за пределы физических границ карты и принадлежность к стеку просмотренных точек $madeP$. Если точка проходит проверку, то вызывается подфункция расчета суммарного риска по слоям карты `robot_risk_fun`.

Риск рассчитывается в соответствии с формулой (1). Коэффициенты g_i задаются пользователем и являются входными данными для функции расчета риска. Реализация функции расстояния до цели $\rho = \sqrt{(e_{xy_T} - e_{xy_C})^2 + (e_{xy_T} - e_{xy_C})^2}$ использует оценку по прямой, где (e_{xy_T}, e_{xy_C}) — координаты текущего положения; (e_{xy_C}, e_{xy_C}) — координаты цели; $\sqrt{\quad}$ — операция взятия квадратного корня. Функция f реализована суммой произведения значений свойств слоев и коэффициентов g влияния на робота; h определяет функцию перевода физических характеристик слоя в функцию риска.

Сумма вычисленной функции `robot_risk_fun` для координат $nearP$ и значения $leftP'$ текущей точки добавляются в список $leftP$ в соответствии с сортировкой по возрастанию. Проверенная точка $leftP'$ добавляется в список $madeP$.

При попадании в целевую точку поиск завершается сформированной картой $mapP$, так как последнее выбранное значение имеет минимальную суммарную вероятность отказа для всей траектории. Для построения этой траектории на последнем этапе алгоритм возвращается по минимальным значениям вероятности от конечной точки к начальной, записывая последовательность точек перемещения.

Модель движения робота

Для тестирования алгоритма используется модель в системе MatLab, позволяющая симулировать движение группы роботов в пространстве с несколькими слоями. Модель содержит модули: отображения пространства движения объектов управления; генерации случайных и загрузки пользовательских карт; выбора постановок выполняемых задач; определения параметров и реализации взаимодействия среды и роботов; планирования операций роботов; разрешения коллизий траекторий; реализации динамических изменений среды.

На изображении интерфейса модели (рис. 2, а) указаны: слои свойств пространства движения группы роботов, положение роботов, грузов, целей и траектории движения из начальной точки в конечную при транспортировке груза для однородных и гетерогенных групп роботов. Дополнительно могут быть визуализированы слои карты, траектории, пространство состояний.

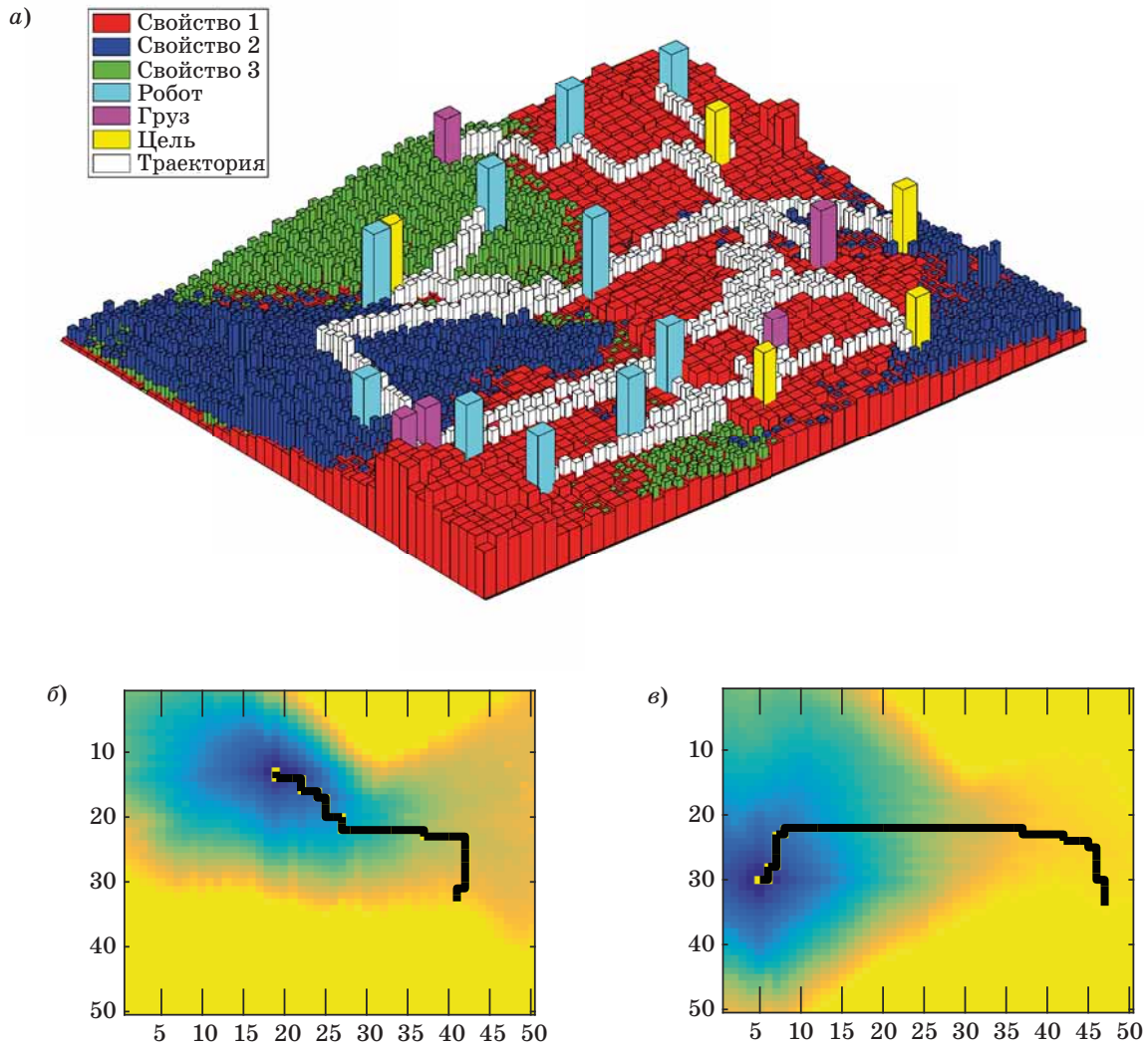
Построение траекторий осуществляется в соответствии с алгоритмом (см. рис. 1).

На рис. 2, а представлено решение транспортной задачи доставки грузов в целевые точки группой гетерогенных роботов, транспортеров и погрузчиков; на карте показано пересечение трех слоев пространства. На рис. 2, б и в продемонстрированы индивидуальные траектории роботов на картах вероятностей отказов, генерируемых в процессе поиска траектории.

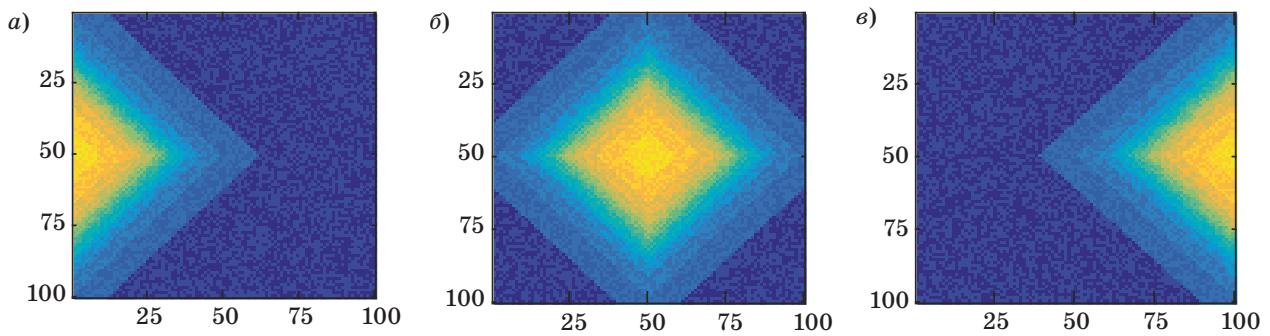
Моделирование

Для проверки работоспособности алгоритма реализована трехслойная карта (рис. 3, а–в). Каждый слой хранит характеристики среды, используемые для расчета коэффициентов риска по формуле (1).

Каждый слой карты представляет собой функцию с единственным максимумом, суммированную с функцией гауссова шума и отражающую значения свойства пространства движения. Слои смещают траекторию объекта управления с учетом коэффициентов влияния G .



■ **Рис. 2.** Представление среды в модели (а) и примеры отражения слоя вероятности отказа и построенной на нем траектории (б, в)
 ■ **Fig. 2.** Representation of the environment in the model (a) and examples of the failure probability layer and planned trajectory on it (б, в)



■ **Рис. 3.** Значения уровня свойств слоев. При построении траекторий слой смещает траекторию вправо (а); отталкивает от центра карты (б); смещает влево (в)
 ■ **Fig. 3.** Layer property level values. When constructing trajectories, the layer shifts the trajectory to the right (a); pushes away from the center of the map (б); shifts to the left (в)

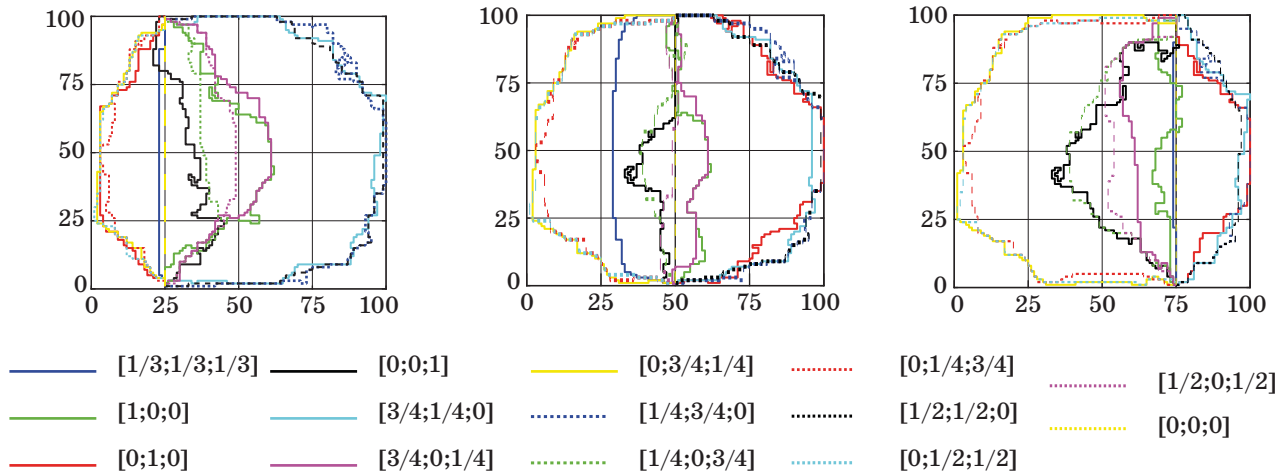
Результаты работы алгоритма построения траекторий для различных значений коэффициентов влияния слоев на роботов представлены на рис. 4. Различные траектории движения демонстрируют отклонения от кратчайшего расстояния, что подтверждает работоспособность предложенного алгоритма. Отклонение путей от идеальных прямых обусловлено добавлением аддитивного шума и коэффициентами влияния слоев, что полностью предсказывает предложенный алгоритм.

Коэффициенты G , помимо влияния слоев на отклонение траектории, также оказывают влияние на приоритетность критериев; например, на рис. 4 при случае $[0;0;0]$ траектория превращается в прямую. Если указать для модели дискретный слой, вероятность отказа на котором будет либо нулевой, либо бесконечной, то ал-

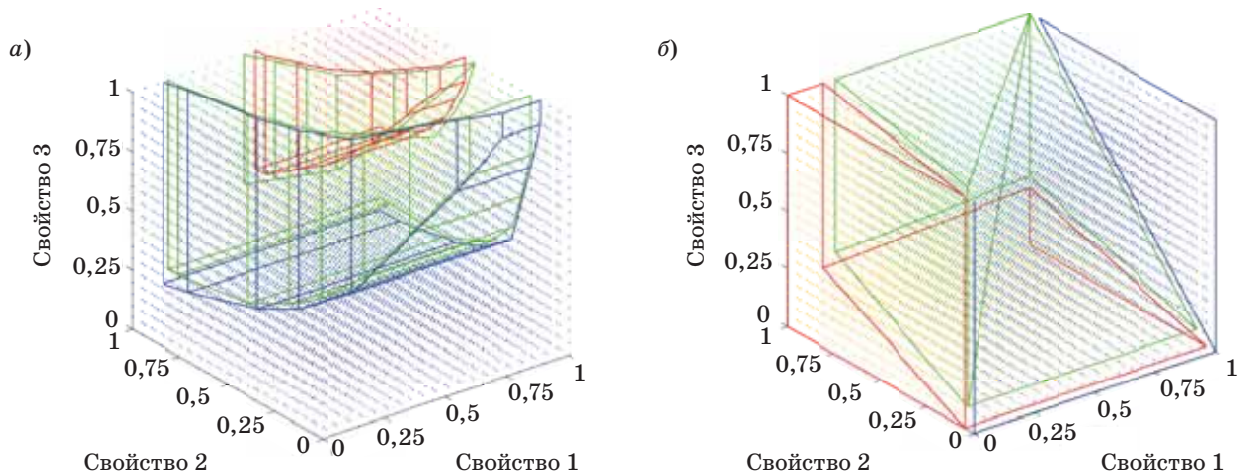
горитм будет работать по аналогии с алгоритмом A^* .

Множество построенных траекторий может быть отображено в виде пространства состояний. Пространство состояний для робота, траектории которого показаны на рис. 4 (в центре), представлены на рис. 5, где приняты следующие цветовые обозначения: красный цвет — высокие значения параметров (вероятности или длины), зеленый — средние, синий — малые. Стоит отметить, что изменения величин происходят плавно, а не скачкообразно, и границы областей являются условными.

На рис. 5, *a* отражена зависимость длины траектории от коэффициентов влияния свойства слоя на путь движения робота; большое влияние всех слоев пространства очевидно увеличивает вероятность отказа (область, ограниченная крас-



■ **Рис. 4.** Траектории движения робота при различных значениях G
 ■ **Fig. 4.** Robot' motion trajectories for different values of G



■ **Рис. 5.** Пространство состояний: *a* — вероятность; *б* — длина траектории
 ■ **Fig. 5.** State space: *a* — probability; *б* — trajectory length

ным цветом), но при снижении влияния хотя бы одного из слоев менее чем на 50 % (при 100%-м влиянии остальных) создаются условия, позволяющие строить траектории со средними и низкими вероятностями отказов. На рис. 5, б представлена вероятность достижения цели роботом. Можно заметить, что даже при максимальном влиянии свойств 1 и 3 траектории минимальны, как и вероятности; более того, длина траектории зависит в основном от свойства 2, что логично следует из рис. 3, б.

Заключение

В статье разработан и исследован алгоритм планирования траектории движения робота на многослойных картах. Встроенный в модель алгоритм может быть использован для планирования группового движения роботов.

Построены семейства траекторий при трех различных значениях начальных и целевых условий, демонстрирующие корректную работу алгоритмов и зависимость от значений свойств окружающего пространства и желаемых коэффициентов.

Построенное пространство состояний позволяет планировать движение робота при заданных условиях среды. Даны рекомендации по модификации параметров объекта управления для оптимизации времени выполнения задачи или вероятности выполнения поставленных перед системой задач.

В продолжение данной работы планируется модификация алгоритма для работы с пространством состояний и построения в нем траекторий в целях оптимизации параметров вероятности отказов и длин траекторий в динамически изменяемых средах.

Работа выполнена при финансовой поддержке РФФИ, проект № 16-29-04319.

Литература

1. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein. Introduction to Algorithms. 3rd. — MIT Press, 2009. — 1313 p.
2. Eraghi N. O., López-Colino F., de Castro A., Garrido J. Path Length Comparison in Grid Maps of Planning Algorithms: HCTNav, A* and Dijkstra // Design of Circuits and Integrated Systems, Madrid, 2014. P. 1–6. doi:10.1109/DCIS.2014.7035557
3. Столяров А. А., Санников Е. В. Выбор эффективного алгоритма планирования для формирования информационной подсистемы движения мобильного робота // Universum: Технические науки: электрон. науч. журн. 2015. № 8-9(20). <http://7universum.com/ru/tech/archive/item/2586> (дата обращения: 28.04.2018).
4. Нейдорф Р. А., Полях В. В., Черногоров И. В., Ярахмедов О. Т. Исследование эвристических алгоритмов в задачах прокладки и оптимизация маршрутов в среде с препятствиями // Изв. ЮФУ. Технические науки. 2016. № 3 (176). С. 127–143.
5. Lin M., Yuan K., Shi C., Wang Y. Path Planning of Mobile Robot based on Improved A* Algorithm // 2017 29th Chinese Control and Decision Conf. (CCDC), Chongqing, 2017. P. 3570–3576. doi:10.1109/CCDC.2017.7979125
6. Da K., Xiaoyu L., Bi Z. Variable-Step-Length A* Algorithm for Path Planning of Mobile Robot // 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, 2017. P. 7129–7133. doi:10.1109/CCDC.2017.7978469
7. Cherni F., Boutereaa Y., Rekik C., Derbel N. Autonomous Mobile Robot Navigation Algorithm for Planning Collision-Free Path Designed in Dynamic Environments // 2015 9th Jordanian International Electrical and Electronics Engineering Conf. (JIEEEEC), Amman, 2015. P. 1–6. doi:10.1109/JIEEEEC.2015.7470747
8. Даринцев О. В., Мигранов А. Б. Синтез гибридных интеллектуальных алгоритмов планирования траектории // Фундаментальные исследования. 2015. № 12. Ч. 4. С. 676–681. <http://www.fundamental-research.ru/ru/article/view?id=39603> (дата обращения: 28.04.2018).
9. Моторин Д. Е., Попов С. Г., Курочкин Л. М. Алгоритм разрешения коллизий при планировании движения группы роботов в условиях пространственно-ситуационной неопределенности // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2017. Т. 10. № 2. С. 32–44. doi:10.18721/JCSTCS.10203
10. Дивеев А. И., Шмалько Е. Ю. Эволюционные методы вычислений для синтеза управления группой роботов и поиска оптимальных траекторий их движения // Cloud of Science. 2017. Т. 4. № 3. С. 395–414.
11. Федоренко К. В., Оловянный А. Л. Исследование основных параметров генетического алгоритма применительно к задаче поиска оптимального маршрута // Вестник Государственного университета морского и речного флота имени адмирала С. О. Макарова. 2017. Т. 9. № 4. С. 714–723. doi:10.21821/2309-5180-2017-9-4-714-723
12. Ni J., Wang K., Huang H., Wu L., Luo C. Robot Path Planning based on an Improved Genetic Algorithm with Variable Length Chromosome // 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Changsha, 2016. P. 145–149. doi:10.1109/FSKD.2016.7603165
13. Li J., Huang Y., Xu Z., Wang J., Chen M. Path Planning of UAV based on Hierarchical Genetic Algo-

- rithm with Optimized Search Region // 2017 13th IEEE Intern. Conf. on Control & Automation (ICCA), Ohrid, 2017. P. 1033–1038. doi:10.1109/ICCA.2017.800320323
14. Schäfle T. R., Mohamed S., Uchiyama N., Sawodny O. Coverage Path Planning for Mobile Robots using Genetic Algorithm with Energy Optimization // 2016 International Electronics Symposium (IES), Denpasar, 2016. P. 99–104. doi:10.1109/ELECSYM.2016.7860983
 15. Ming K. Solving Path Planning Problem based on Ant Colony Algorithm // 2017 29th Chinese Control and Decision Conference (CCDC), Chongqing, 2017. P. 5391–5395. doi:10.1109/CCDC.2017.7979455
 16. Ватугин Э. И., Титов В. С. Анализ результатов применения алгоритма муравьиной колонии в задаче поиска пути в графе при наличии ограничений // Изв. ЮФУ. Технические науки. 2014. № 12 (161). С. 111–120.
 17. Мартынов А. В., Курейчик В. М. Гибридный алгоритм решения задачи коммивояжера // Изв. ЮФУ. Технические науки. 2015. № 4 (165). С. 36–44.
 18. Zhang W., Gong X., Han G., Zhao Y. An Improved Ant Colony Algorithm for Path Planning in One Scenic Area With Many Spots // IEEE Access. 2017. Vol. 5. P. 13260–13269. doi:10.1109/ACCESS.2017.2723892
 19. Uriol R., Moran A. Mobile Robot Path Planning in Complex Environments using Ant Colony Optimization Algorithm // 2017 3rd Intern. Conf. on Control, Automation and Robotics (ICCAR), Nagoya, 2017. P. 15–21. doi:10.1109/ICCAR.2017.7942653
 20. Xiao S. Optimal Travel Path Planning and Real Time Forecast System based on Ant Colony Algorithm // 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conf. (IAEAC), Chongqing, 2017. P. 2223–2226. doi:10.1109/IAEAC.2017.8054413
 21. Lin T., Zhang K., Cui N., Tu Z., Zhang H. Path Planning of Aircraft based on Adaptive Multiobjective Estimation of Distribution Algorithm // 2016 IEEE Symp. Series on Computational Intelligence (SSCI), Athens, 2016. P. 1–8. doi:10.1109/SSCI.2016.7850199
 22. Tan J., Zhao L., Wang Y., Zhang Y., Li L. The 3D Path Planning based on A* Algorithm and Artificial Potential Field for the Rotary-Wing Flying Robot // 2016 8th Intern. Conf. on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, 2016. P. 551–556. doi:10.1109/IHMSC.2016.155
 23. Song R., Liu W., Liu Y., Bucknall R. A Two-Layered Fast Marching Path Planning Algorithm for an Unmanned Surface Vehicle Operating in a Dynamic Environment // OCEANS 2015, Genova, Genoa, 2015. P. 1–8. doi:10.1109/OCEANS-Genova.2015.7271405
 24. Chen H., Wang H., Jiang L. Path Planning of UAV based on Cultural Algorithm in Dynamic Environments // 2016 6th Intern. Conf. on Electronics Information and Emergency Communication (ICEIEC), Beijing, 2016. P. 130–134. doi:10.1109/ICEIEC.2016.7589704

UDC 519.876.5, 519.873, 004.023:896
doi:10.15217/issn1684-8853.2018.3.45

Multi-Criteria Path Planning Algorithm for a Robot on a Multilayer Map

Motorin D. E.^a, Post-Graduate Student, d.e.motorin@gmail.com

Popov S. G.^a, PhD, Tech., Associate Professor, popovserge@spbstu.ru

^aPeter the Great St. Petersburg Polytechnic University, 29, Politekhnicheskaya St., 195251, Saint-Petersburg, Russian Federation

Introduction: A complex environment is characterized by the possibility to decompose the factors affecting the robot into independent layers. As the robot is moving in a complex environment, it is exposed to negative factors which affect its ability to achieve the goal. The problem arises how to choose the motion trajectory while minimizing the negative effects on the robot and the distance covered. **Purpose:** Developing and analyzing an algorithm for two-criteria optimization of robot motion trajectory, taking into account the desired criteria about the interaction with the environment and the trajectory length. **Results:** We have developed and implemented an algorithm for shortest path search on a map each layer of which displays a property of the space and allows you to take into account the interaction between the robot and the environment, as well as the distance covered. The algorithm implementation is incorporated into the robot group control model. To analyze the algorithm, test multilayer maps were used, with the addition of Gaussian noise. The simulation results are a set of trajectories reflecting the coefficients with which the space properties affect the robot when the initial and final positions on the map are given. A space of the robot motion states demonstrates how the influence of the environment properties on the robot depends on the trajectory length and on the failure risk throughout the path. **Practical relevance:** The developed algorithm can be applied in planning systems of individual or group motion of robots. The resulting state space reflects the ranges of effective characteristics of the robot when performing tasks in a given environment. As the next step, the developed algorithm will be applied to plan paths on multiscale maps, and sets of trajectories will be built in the state space of a group of robots.

Keywords — Trajectory Planning, Multilayer Maps, Control, Robot, State Space, Heuristic Algorithm, Realistic Environment.

Citation: Motorin D. E., Popov S. G. Multi-Criteria Path Planning Algorithm for a Robot on a Multilayer Map. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 45–53 (In Russian). doi:10.15217/issn1684-8853.2018.3.45

References

1. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein. *Introduction to Algorithms*. 3rd. MIT Press, 2009. 1313 p.
2. Eraghi N. O., López-Colino F., de Castro A., Garrido J. Path Length Comparison in Grid Maps of Planning Algorithms: HCTNav, A* and Dijkstra. *Design of Circuits and Integrated Systems*, Madrid, 2014, pp. 1–6. doi:10.1109/DCIS.2014.7035557
3. Stolyarov A. A., Sannikov E. V. Choice of Effective Planning Algorithm for Formation of Informative Subsystem of Mobile Robot Motion. *Universum: Tekhnicheskie nauki*, 2015, no. 8-9(20). Available at: <http://7universum.com/ru/tech/archive/item/2586> (accessed 28 April 2018) (In Russian).
4. Neydorf R. A., Polyakh V. V., Chernogorov I. V., Yarakhmedov O. T. The Study of Heuristic Algorithms in the Tasks Strip, and Optimization of Routes in an Environment with Obstacles. *Izvestiya YUFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2016, no. 3 (176), pp. 127–143 (In Russian).
5. Lin M., Yuan K., Shi C., Wang Y. Path Planning of Mobile Robot based on Improved A* Algorithm. *2017 29th Chinese Control and Decision Conf. (CCDC)*, Chongqing, 2017, pp. 3570–3576. doi:10.1109/CCDC.2017.7979125
6. Da K., Xiaoyu L., Bi Z. Variable-Step-Length A* Algorithm for Path Planning of Mobile Robot. *2017 29th Chinese Control and Decision Conf. (CCDC)*, Chongqing, 2017, pp. 7129–7133. doi:10.1109/CCDC.2017.7978469
7. Cherni F., Boutereaa Y., Rekek C., Derbel N. Autonomous Mobile Robot Navigation Algorithm for Planning Collision-Free Path Designed in Dynamic Environments. *2015 9th Jordanian Intern. Electrical and Electronics Engineering Conf. (JIEEEEC)*, Amman, 2015, pp. 1–6. doi:10.1109/JIEEEEC.2015.7470747
8. Darintsev O. V., Migranov A. B. Synthesis of the Hybrid Intelligent Algorithm of Planning Path. *Fundamental'nye issledovaniya* [Fundamental Research], 2015, no. 12, part 4, pp. 676–681. Available at: <http://www.fundamental-research.ru/ru/article/view?id=39603> (accessed 28 April 2018) (In Russian).
9. Motorin D. Ye., Popov S. G., Kurochkin L. M. An Algorithm for Collision Avoidance in Path Planning for a Group of Robots in a Spatio-Situational Indeterminacy. *Nauchno-tekhnicheskie vedomosti SPbGPU. Informatika. Telekomunikatsii. Upravlenie* [St. Petersburg Polytechnic University Journal of Engineering Science and Technology], 2017, vol. 10, no. 2, pp. 32–44 (In Russian). doi:10.18721/JCSTCS.10203
10. Diveev A. I., Shmal'ko E. Yu. Evolutionary Computing Methods for Control Synthesis a Group of Robots and the Search for Optimal Trajectories of their Movement. *Cloud of Science*, 2017, vol. 4, no. 3, pp. 395–414 (In Russian).
11. Fedorenko K. V., Olovyannikov A. L. Research of the Main Parameters of the Genetic Algorithm for the Problem of Searching the Optimal Route. *Vestnik Gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S. O. Makarova*, 2017, vol. 9, no. 4, pp. 714–723 (In Russian). doi:10.21821/2309-5180-2017-9-4-714-723
12. Ni J., Wang K., Huang H., Wu L., Luo C. Robot Path Planning based on an Improved Genetic Algorithm with Variable Length Chromosome. *2016 12th Intern. Conf. on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, Changsha, 2016, pp. 145–149. doi:10.1109/FSKD.2016.7603165
13. Li J., Huang Y., Xu Z., Wang J., Chen M. Path Planning of UAV based on Hierarchical Genetic Algorithm with Optimized Search Region. *2017 13th IEEE Intern. Conf. on Control & Automation (ICCA)*, Ohrid, 2017, pp. 1033–1038. doi:10.1109/ICCA.2017.800320323
14. Schäfle T. R., Mohamed S., Sawodny O. Coverage Path Planning for Mobile Robots using Genetic Algorithm with Energy Optimization. *2016 International Electronics Symposium (IES)*, Denpasar, 2016, pp. 99–104. doi:10.1109/ELECSYM.2016.7860983
15. Ming K. Solving Path Planning Problem based on Ant Colony Algorithm. *2017 29th Chinese Control And Decision Conf. (CCDC)*, Chongqing, 2017, pp. 5391–5395. doi:10.1109/CCDC.2017.7979455
16. Vatutin Je. I., Titov V. S. Analysis of the Results of Applying the Ant Colony Optimization Algorithm to the Shortest Path Problem in the Graphs with Constraints. *Izvestiya YUFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, no. 12 (161), pp. 111–120 (In Russian).
17. Martinov A. V., Kureichik V. M. Hybrid Approach for Travelling Salesman Problem. *Izvestiya YUFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, no. 4 (165), pp. 36–44 (In Russian).
18. Zhang W., Gong X., Han G., Zhao Y. An Improved Ant Colony Algorithm for Path Planning in One Scenic Area with Many Spots. *IEEE Access*, 2017, vol. 5, pp. 13260–13269. doi:10.1109/ACCESS.2017.2723892
19. Uriol R., Moran A. Mobile Robot Path Planning in Complex Environments using Ant Colony Optimization Algorithm. *2017 3rd Intern. Conf. on Control, Automation and Robotics (ICCAR)*, Nagoya, 2017, pp. 15–21. doi:10.1109/ICCAR.2017.7942653
20. Xiao S. Optimal Travel Path Planning and Real Time Forecast System based on Ant Colony Algorithm. *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conf. (IAEAC)*, Chongqing, 2017, pp. 2223–2226. doi:10.1109/IAEAC.2017.8054413
21. Lin T., Zhang K., Cui N., Tu Z., Zhang H. Path Planning of Aircraft based on Adaptive Multiobjective Estimation of Distribution Algorithm. *2016 IEEE Symp. Series on Computational Intelligence (SSCI)*, Athens, 2016, pp. 1–8. doi:10.1109/SSCI.2016.7850199
22. Tan J., Zhao L., Wang Y., Zhang Y., Li L. The 3D Path Planning based on A* Algorithm and Artificial Potential Field for the Rotary-Wing Flying Robot. *2016 8th Intern. Conf. on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, Hangzhou, 2016, pp. 551–556. doi:10.1109/IHMSC.2016.155
23. Song R., Liu W., Liu Y., Bucknall R. A Two-Layered Fast Marching Path Planning Algorithm for an Unmanned Surface Vehicle Operating in a Dynamic Environment. *OCEANS 2015*, Genova, Genoa, 2015, pp. 1–8. doi:10.1109/OCEANS-Genova.2015.7271405
24. Chen H., Wang H., Jiang L. Path Planning of UAV based on Cultural Algorithm in Dynamic Environments. *2016 6th Intern. Conf. on Electronics Information and Emergency Communication (ICEIEC)*, Beijing, 2016, pp. 130–134. doi:10.1109/ICEIEC.2016.7589704

ОЦЕНКА ВЕРОЯТНОСТНО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК СЕТЕВЫХ УЗЛОВ С ДИФФЕРЕНЦИАЦИЕЙ ТРАФИКА

Т. М. Татарникова^а, доктор техн. наук, профессор, *tm-tatarn@yandex.ru*

А. В. Вольский^а, магистрант, *volskiyav62@gmail.com*

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Постановка проблемы: появление новых приложений реального времени, требовательных к качеству обслуживания, привело к необходимости классификации IP-трафика и соответствующих моделей его дифференцированной обработки на сетевых узлах. **Цель:** оценка влияния механизма дифференцированных услуг DiffServ на качество обслуживания сетевого трафика. **Результаты:** выполнено моделирование обслуживания трафика на сетевом узле, реализующем механизм DiffServ, с целью оценки характеристик, обеспечивающих качество обслуживания. Разработаны модели бесприоритетного, приоритетного и взвешенного справедливого обслуживания. Результаты моделирования показали, что применяемый механизм DiffServ может способствовать эффективной реализации качества обслуживания. **Практическая значимость:** разработанные модели являются инструментом управления очередями и планирования обработки пакетов на сетевом узле, реализующем механизм DiffServ.

Ключевые слова — сетевой трафик, механизм дифференциации трафика, качество обслуживания, приоритетное обслуживание, взвешенное справедливое обслуживание, настраиваемая очередь, маршрутизатор, IP-приложение.

Цитирование: Татарникова Т. М., Вольский А. В. Оценка вероятностно-временных характеристик сетевых узлов с дифференциацией трафика // Информационно-управляющие системы. 2018. № 3. С. 54–60. doi:10.15217/issn1684-8853.2018.3.54

Citation: Tatarnikova T. M., Volskiy A. V. Estimation of Probabilistic-Temporal Characteristics of Network Nodes with Traffic Differentiation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 54–60 (In Russian). doi:10.15217/issn1684-8853.2018.3.54

Введение

Растущий спрос на многоадресную рассылку в IP-сетях и появление новых приложений реального времени привели к расширению архитектуры и протоколов сети Интернет, поддерживающих интегрированные услуги [1, 2].

В настоящее время разработано два стандарта для обеспечения качества обслуживания (Quality of Services — QoS) трафика реального времени: так называемые интегрированные услуги (Integrated Services — IntServ) и дифференцированные услуги (Differentiated Services — DiffServ) [3].

Согласно концепции IntServ, приложения могут выбирать для своих потоков данных любой из контролируемых уровней QoS, применяя для этих целей механизм резервирования. Данный процесс предусматривает резервирование ресурсов для каждого потока во всех промежуточных маршрутизаторах на пути от отправителя к получателю с использованием сигнализации «из конца в конец» [4]. Реализация IntServ требует хранения состояний всех активных соединений, число которых может достигать нескольких тысяч в каждый момент времени, что, естественно, увеличивает нагрузку на маршрутизаторы. Более того, каждый раз при изменении топологии все зарезервированные маршруты необходимо устанавливать заново [5].

Механизм DiffServ преодолевает ограничения, присущие IntServ. DiffServ разделяет трафик на классы, вводя несколько уровней QoS. Таким образом механизм DiffServ реализует приоритизацию трафика, однако точные значения показателей качества или гарантии их обеспечения не входят в его функции. В действительности немногие приложения нуждаются в жестких гарантиях QoS. Классификация трафика в сочетании с адаптивной природой многих приложений оказываются достаточными условиями для обеспечения нормальной работы IP-сетей с дифференциацией услуг. К тому же отправитель и получатель не обмениваются информацией о требованиях к качеству обслуживания, что исключает временные затраты на прокладку маршрута, присущие IntServ [6].

Пока некоторые приложения, функционирующие на базе Интернета, не всегда правильно классифицируются. Так, например, с помощью программы-анализатора трафика WireShark у двух различных VoIP-утилит — skype и discord — были зафиксированы разные приоритеты. В skype трафик маркируется действительно как priority (приоритетный), в то время как в discord маркируется как routine (без приоритета), что не гарантирует обеспечение QoS.

Появление новых приложений, требовательных к качеству обслуживания, актуализирует

моделирование механизма DiffServ как инструмента управления очередями и планирования обработки пакетов на сетевых узлах.

Принципы работа механизма DiffServ

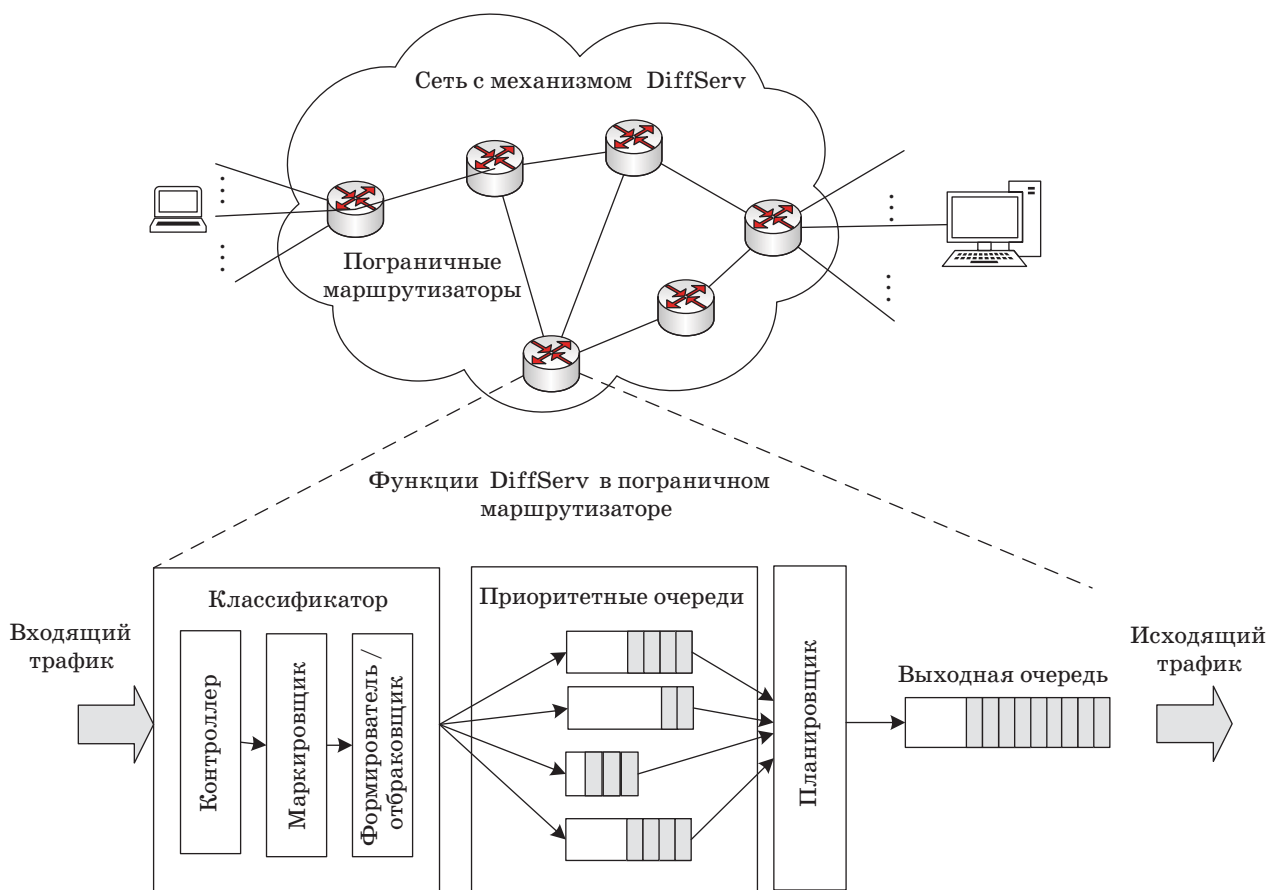
Архитектура DiffServ предполагает наличие классификаторов и формирователей трафика на границе сети с функциями дифференциации трафика (рис. 1) [7, 8].

Пакеты классифицируются и маркируются так, чтобы они могли получить определенный режим обработки в каждом маршрутизаторе на протяжении всего пути следования в соответствии с политикой пошаговой обработки (Per-Hop Behavior — PHB) [9].

Политика пошаговой обработки — это способ резервирования ресурсов маршрутизатора, об-

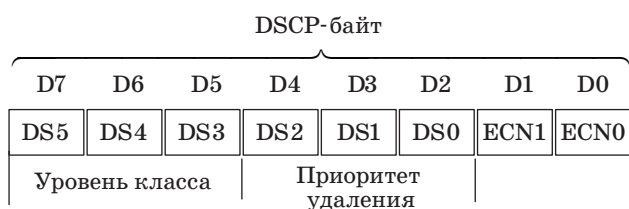
служивающего потоки трафика. Правила PHB реализуются с помощью нескольких механизмов управления буфером и планирования обработки пакетов [10].

Информация о классе трафика передается в IP-пакетах посредством маркировки поля DiffServ размером 1 байт, получившего название точки кода дифференцированных услуг (Differentiated Services Code Point — DSCP). Формат DSCP-байта приведен на рис. 2. Биты DS5...DS3 кодируют уровень класса обслуживания от 0 — минимального приоритета до 7 — максимального приоритета, биты DS2...DS0 кодируют приоритет удаления от 0, когда приоритет удаления максимальный, до 7, когда приоритет удаления минимальный. В итоге получается код приоритета — число от 0 до 63, где чем больше число, тем трафик важнее. Например, для VoIP-трафика применяется класс сервиса 5 (DSCP-байт равен 0xA0 или 10100000b),



Контроллер определяет соответствие потока трафика определенному классу.
 Маркировщик относит пакет к одному из потоков DiffServ.
 Формирователь задерживает пакеты для выравнивания потока в соответствии с классом.
 Отбраковщик удаляет пакеты для обеспечения соответствия потока классу.

■ **Рис. 1.** Архитектура метода DiffServ
 ■ **Fig. 1.** The architecture of the DiffServ method



■ **Рис. 2.** Формат DSCP-байта
 ■ **Fig. 2.** DSCP byte format

а для обычного трафика класс сервиса 0 (DSCP-байт равен 0x00 или 00000000b). Биты ECN1, ECN0 не определены [11, 12].

Алгоритмы обработки очередей

Известны несколько алгоритмов управления очередями, которые нашли практическое применение при обработке трафика в сетевых узлах. Их классификация приведена на рис. 3.

Принцип работы алгоритма FIFO («первым пришел — первым ушел», First In — First Out) состоит в следующем: в случае перегрузки пакеты помещаются в очередь, а если перегрузка устраняется, пакеты передаются на выход в том порядке, в котором поступили. Несмотря на простую реализацию и отсутствие потребности в конфигурировании, очереди FIFO не справляются с поддержкой дифференцированного качества обслуживания [13].

Приоритетное обслуживание обеспечивает безусловный приоритет одних пакетов над другими. Всего выделено четыре очереди с приоритетами: высоким, средним, нормальным и низким. Обработка ведется последовательно от высокого к низкому приоритету, начиная с высокоприоритетной очереди, и до ее полной очистки, затем переходит к менее приоритетным очередям. Очевидно, возможна монополизация канала высоко-



■ **Рис. 3.** Классификация алгоритмов обработки очередей маршрутизатора
 ■ **Fig. 3.** Classification of router queue processing algorithms

коприоритетными очередями, что приводит к потере низкоприоритетного трафика при высокой интенсивности поступления высокоприоритетных данных. По умолчанию всем приоритетным очередям отводятся буферы одинакового размера, но возможно административное выделение размера буфера. Пакет, поступивший в то время, когда буфер заполнен, просто отбрасывается [14].

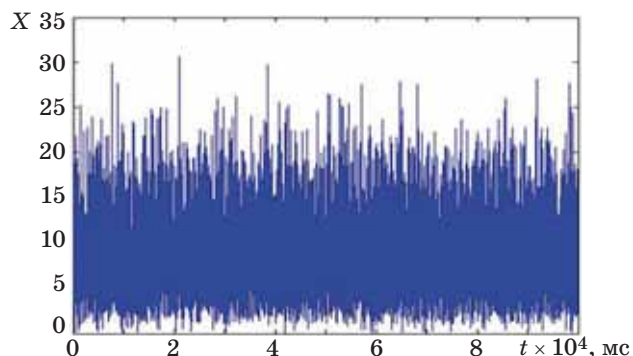
Алгоритм настраиваемых очередей предусматривает управление долей полосы пропускания канала для каждой очереди. Всего поддерживается 17 очередей, где нулевая очередь зарезервирована для управляющих высокоприоритетных пакетов и пользователю недоступна. Каждая очередь содержит счетчик байтов, который в начале обхода содержит заданное значение, уменьшающееся на размер пакета при его обслуживании. Очереди обходятся последовательно, начиная с первой. Очереди настраиваются администратором сети [15].

Алгоритм взвешенного справедливого обслуживания разработан для того, чтобы для всех классов трафика можно было предоставить определенный минимум пропускной способности или удовлетворить требования к задержкам. Под весом какого-либо класса понимается доля выделяемой данному виду трафика пропускной способности выходного интерфейса [16]. Вес класса трафика назначается автоматически в зависимости от:

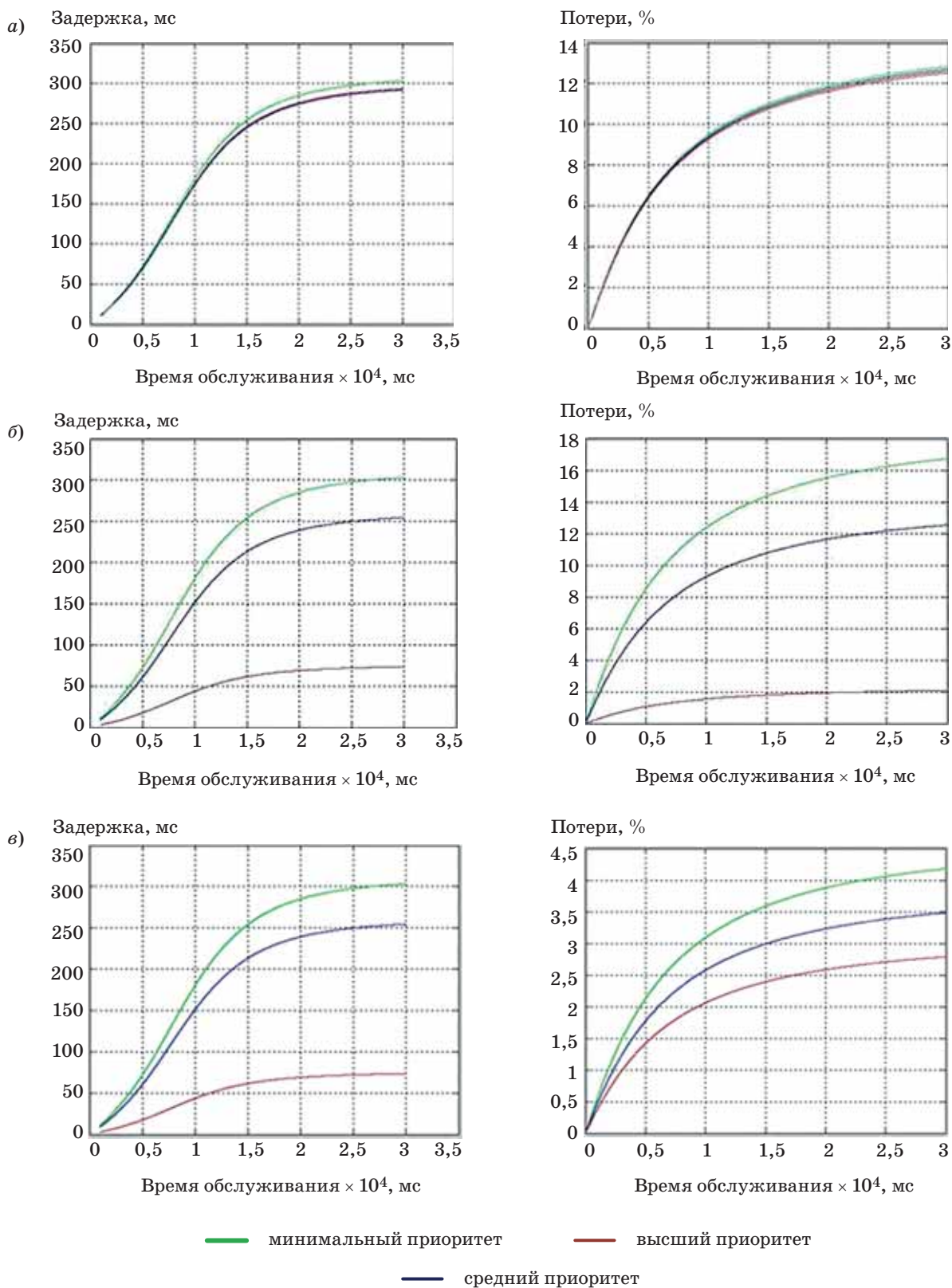
- обслуживания очередей на основе классов: пропускная способность распределяется по классам в абсолютном значении или в процентах относительно установленного значения на планировщике;
- вероятности отбрасывания пакетов.

Модель маршрутизатора, реализующего механизм DiffServ

Маршрутизатор представлен системой массового обслуживания. На его вход поступает сетевой трафик, требующий дифференцированно-



■ **Рис. 4.** Диаграмма сетевого трафика, AR-модель
 ■ **Fig. 4.** Network traffic diagram, AR-model



■ **Рис. 5.** Характеристики обслуживания трафика по алгоритму FIFO (а); по алгоритму приоритетного обслуживания (б); по алгоритму справедливой очереди (в)
 ■ **Fig. 5.** Characteristics of traffic servicing by FIFO algorithm (a); priority service algorithm (б); weighted fair queuing algorithm (в)

го обслуживания [17]. Модель трафика должна соответствовать современному представлению о трафике реально функционирующих сетей с коммутацией пакетов и агрегацией данных различных приложений.

Трафик проходит все элементы архитектуры DiffServ (см. рис. 1). Модели маршрутизатора отличаются механизмами управления буфером и планирования обработки пакетов. Реализованы модели обслуживания FIFO, приоритетное обслуживание, справедливая очередь.

Результаты статистического анализа измерений сетевого трафика, представленные в многочисленных публикациях, доказали, что модель сетевого трафика должна отображать фрактальность его характеристик [18].

К фрактальным свойствам характеристик трафика относят такие понятия, как медленно затухающие дисперсии, долговременная зависимость, самоподобность. Эти свойства, полученные опытным путем, существенно отличаются от свойств, наблюдаемых у трафика, сгенерированного традиционными моделями. Основной причиной этого различия является структура лежащей в основе зависимости. Особенностью самоподобного трафика является устойчивость кластеризации, в то время как традиционные модели пакетного трафика являются кратковременно зависимыми, т. е. имеют экспоненциально затухающие корреляции, а данные измеренного пакетного трафика проявляют долговременную зависимость, т. е. гиперболически затухающие корреляции [19].

На сегодня разработано множество моделей, предназначенных для имитации фрактального трафика. Анализ доступных публикаций по моделированию сетевого трафика позволил выделить авторегрессионные модели (Autoregressive Models — AR), получившие наибольшее распространение благодаря свойству длительной памяти ти самоподобных процессов [20].

AR-модель — это модель временных рядов, в которой значения временного ряда в данный момент линейно зависят от предыдущих значений этого же ряда. Как разновидности таких моделей используются модели «скользящего среднего», смысл которых заключается в том, что учитывается только ближайшее прошлое на за-

данное количество отсчетов по времени в глубину и следующая пачка трафика строится на основе только этих данных:

$$X_t = c + \sum_{i=1}^p a_i X_{t-i} + \varepsilon_t,$$

где c — постоянная; p — размер пачки трафика; a — коэффициенты авторегрессии; ε_t — белый шум.

Диаграмма сетевого трафика, смоделированного по методу «скользящего среднего», показана на рис. 4.

Результаты моделирования с разными механизмами управления буфером приведены на рис. 5, а–в. Представленные зависимости демонстрируют, что введение механизмов дифференциации позволяет значительно улучшить характеристики трафика, чувствительного к задержкам и потерям. Так, например, введение приоритетов позволило снизить задержку почти в 5 раз и потери в 4–6 раз. Эта тенденция сохраняется и при других параметрах входного потока и времени обслуживания маршрутизатора.

Заключение

Появление новых приложений реального времени, требовательных к качеству обслуживания, привело к необходимости дифференциации IP-трафика.

В целях оценки характеристик, обеспечивающих качество обслуживания, выполнено моделирование механизма DiffServ как инструмента управления очередями и планирования обработки пакетов на сетевых узлах. Разработаны модели беспriorитетного, приоритетного и взвешенного справедливого обслуживания.

Как показывают результаты моделирования, применяемый механизм DiffServ может способствовать эффективной реализации QoS.

Настройка политики пошаговой обработки в сети с дифференциацией услуг позволит использовать Интернет как универсальную платформу глобальных коммуникаций, которая не только представляет инструмент для классификации трафика, но и гарантирует обеспечение показателей QoS.

Литература

1. Tanenbaum A., Wetherall D. Computer Networks. 5th ed. — Prentice Hall, 2010. — 960 p.
2. Kilkki K. Differentiated Services for the Internet. — Indianapolis: Macmillan Technical Published, 1999. — 400 p.

3. ITU-T Recommendation Y.1291. An Architectural Framework for Support of Quality of Service in Packet Networks. May 2004. https://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKewiqw6fQ49PaAhXFFywKHUp4ASwQFggrMAA&url=https%3A%2F%2Fwww.itu.int%2Frec%2Fdologin_pub.asp%3Flang%3

- De%26id%3DТ-REC-Y.1291-200405-I!!PDF-E%-26type%3Ditems&usg=AOvVaw03tHDrq48RVxkVnz0wpuLe (дата обращения: 12.03.2018).
4. Богатырев В. А., Кармановский Н. С., Попцова Н. А., Паршутин С. А., Воронина Д. А., Богатырев С. В. Имитационная модель поддержки проектирования инфокоммуникационных резервированных систем // Научно-технический вестник информационных технологий, механики и оптики. 2016. Т. 16. № 5(105). С. 831–838. doi:10.17586/2226-1494-2016-16-5-831-838
 5. Evans J. W., Filsfils C. Deploying IP and MPLS QoS for Multiservice Networks: Theory & Practice. — Morgan Kaufmann Published, 2010. — 456 p.
 6. Vegesna S. R. IP Quality of Service. — Cisco Press Published, 2003. — 368 p.
 7. Blake S., Black D., Carlson M., Davies E., Wang Z., Weiss W. RFC 2475. An Architecture for Differentiated Services. December 1998. <http://www.rfc-base.org/rfc-2475.html> (дата обращения: 12.03.2018).
 8. Tonenbaum A., Austin T. Structured Computer Organization. 6th ed. — Prentice Hall, 2012. — 800 p.
 9. Jacobson V., Nichols K., Poduri K. RFC 2598. An Expedited Forwarding PHB. June 1999. <https://tools.ietf.org/html/rfc2598> (дата обращения: 12.03.2018).
 10. Heinanen J., Baker F., Weiss W., Wroclawski J. RFC 2597 Assured Forwarding PHB Group. June 1999. <https://tools.ietf.org/html/rfc2597> (дата обращения: 12.03.2018).
 11. Nichols K., Blake S., Baker F., Black D. RFC 2474. Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers. December 1998. <https://tools.ietf.org/html/rfc2474> (дата обращения: 12.03.2018).
 12. Гольдштейн Б. С., Пинчук А. В., Суховицкий А. Л. IP-телефония. — М.: Радио и связь, 2006. — 336 с.
 13. Богатырев В. А., Богатырев А. В., Богатырев С. В. Оценка надежности выполнения кластерами запросов реального времени // Изв. вузов. Приборостроение. 2014. Т. 57. № 4. С. 46–48.
 14. Олифер В., Олифер Н. Компьютерные сети. Принципы, технологии, протоколы. — СПб.: Питер, 2016. — 992 с.
 15. Кутузов О. И., Татарникова Т. М. Моделирование систем и сетей телекоммуникаций. — СПб.: РГГМУ, 2012. — 134 с.
 16. Кутузов О. И., Сергеев В. Г., Татарникова Т. М. Коммутаторы в корпоративных сетях. Моделирование и расчет. — СПб.: Судостроение, 2003. — 170 с.
 17. Tatarnikova T., Kolbanov M. Statement of a Task Corporate Information Networks Interface Centers Structural Synthesis // IEEE EUROCON 2009, Saint-Petersburg, 2009. P. 1883–1887.
 18. Zwart A. P. Queueing Systems with Heavy Tails. — Eindhoven University of Technology, 2001. — 227 p.
 19. Шелухин О. И., Осин А. В., Смольский С. М. Самоподобие и фракталы. Телекоммуникационные приложения. — М.: Физматлит, 2008. — 368 с.
 20. Костромицкий А. И., Волотка В. С. Подходы к моделированию самоподобного трафика // Восточно-Европейский журнал передовых технологий. 2010. № 46. С. 46–49.

UDC 004.7

doi:10.15217/issn1684-8853.2018.3.54

Estimation of Probabilistic-Temporal Characteristics of Network Nodes with Traffic DifferentiationTatarnikova T. M.^a, Dr. Sc., Tech., Professor, tm-tatarn@yandex.ruVolskiy A. V.^a, Graduate Student, volskiyav62@gmail.com^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: The release of new real-time applications demanding of service quality led to the need to classify IP traffic and the corresponding models of its differentiated processing on network nodes. **Purpose:** Assessing the impact of DiffServ differentiated service on the quality of network traffic service. **Results:** Traffic on a DiffServ network node has been simulated in order to evaluate the performance characteristics which provide the quality of services. Models have been developed for routine, priority and weighted fair service. The simulation results showed that DiffServ can contribute to effective maintenance of QoS. **Practical relevance:** The developed models are a tool for managing queues and scheduling packet processing on a DiffServ network node.

Keywords — Network Traffic, Differentiated Services, Quality of Services, Priority Queuing, Weighted Fair Queuing, Custom Queuing, Router, IP Application.

Citation: Tatarnikova T. M., Volskiy A. V. Estimation of Probabilistic-Temporal Characteristics of Network Nodes with Traffic Differentiation. *Informatsionno-upravlyaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 54–60 (In Russian). doi:10.15217/issn1684-8853.2018.3.54

References

1. Tanenbaum A., Wetherall D. *Computer Networks*. 5th ed. Prentice Hall, 2010. 960 p.
2. Kilkki K. *Differentiated Services for the Internet*. Indianapolis, Macmillan Technical Published, 1999. 400 p.
3. *ITU-T Recommendation Y.1291. An Architectural Framework for Support of Quality of Service in Packet Networks*. May 2004. Available at: https://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiqW-6fQ49PaAhXFFyWkHUUp4ASwQFggrMAA&url=https%3A%2F%2Fwww.itu.int%2Frec%2Fdologin_pub.asp%3Flang%3De%26id%3DT-REC-Y.1291-200405-I!!PDF-E%-26type%3Ditem&usg=AOvVaw03tHDrq48RVxkVnz0wpuLe (accessed 12 March 2018).
4. Bogatyrev V. A., Karmanovsky N. S., Poptcova N. A., Parshutina S. A., Voronina D. A., Bogatyrev S. V. Simulation Model for Design Support of Infocomm Redundant Systems. *Nauchno-tehnicheskij vestnik informacionnykh tehnologij, mehaniki i optiki* [Scientific and Technical Journal of Information Technologies, Mechanics and Optics], 2016, vol. 16, no. 5, pp. 831–838 (In Russian) doi:10.17586/2226-1494-2016-16-5-831-838
5. Evans J. W., Filsfils C. *Deploying IP and MPLS QoS for Multiservice Networks: Theory & Practice*. Morgan Kaufmann Published, 2010. 456 p.
6. Vegesna S. R. *IP Quality of Service*. Cisco Press Published, 2003. 368 p.
7. Blake S., Black D., Carlson M., Davies E., Wang Z., Weiss W. *RFC 2475. An Architecture for Differentiated Services*. December 1998. Available at: <http://www.rfc-base.org/rfc-2475.html> (accessed 12 March 2018).
8. Tanenbaum A., Austin T. *Structured Computer Organization*. 6th ed. Prentice Hall, 2012. 800 p.
9. Jacobson V., Nichols K., Poduri K. *RFC 2598. An Expedited Forwarding PHB*. June 1999. Available at: <https://tools.ietf.org/html/rfc2598> (accessed 12 March 2018).
10. Heinanen J., Baker F., Weiss W., Wroclawski J. *RFC 2597. Assured Forwarding PHB Group*. June 1999. Available at: <https://tools.ietf.org/html/rfc2597> (accessed 12 March 2018).
11. Nichols K., Blake S., Baker F., Black D. *RFC 2474. Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*. December 1998. Available at: <https://tools.ietf.org/html/rfc2474> (accessed 12 March 2018).
12. Goldstein B. S., Pinchuk A. V., Suhovitsky A. L. *IP-telefoniia* [IP Telephony]. Moscow, Radio i sviaz' Publ., 2006. 336 p. (In Russian).
13. Bogatyrev V. A., Bogatyrev A. V., Bogatyrev S. V. Evaluation of Reliability Performance Cluster Real-Time Query. *Izvestiia vysshikh uchebnykh zavedenii. Priborostroenie* [Journal of Instrument Engineering], 2014, vol. 57, no. 4, pp. 46–48 (In Russian).
14. Olifer V., Olifer N. *Kompiuternye seti. Printsipy, tekhnologii, protokoly* [Computer Networks. Principles, Technologies, Protocols]. Saint-Petersburg, Piter Publ., 2016. 992 p. (In Russian).
15. Kutuzov O. I., Tatarnikova T. M. *Modelirovanie sistem i setei telekommunikatsii* [Simulation of Systems and Telecommunications Networks]. Saint-Petersburg, RGGMU Publ., 2012. 134 p. (In Russian).
16. Kutuzov O. I., Sergeev V. G., Tatarnikova T. M. *Kommutatory v korporativnykh setiakh. Modelirovanie i raschet* [Switches in the Corporate Networks. Simulation and Calculation]. Saint-Petersburg, Sudostroenie Publ., 2003, 170 p. (In Russian).
17. Tatarnikova T., Kolbanev M. Statement of a Task Corporate Information Networks Interface Centers Structural Synthesis. *IEEE EUROCON 2009*, Saint-Petersburg, 2009, pp. 1883–1887.
18. Zwart A. P. *Queueing Systems with Heavy Tails*. Eindhoven University of Technology, 2001. 227 p.
19. Shelukhin O. I., Osin A. V., Smol'skij S. M. *Samopodobie i fraktaly. Telekommunikatsionnye prilozheniia* [Self-similarity and Fractals. Telecommunication Applications]. Moscow, Fizmatlit Publ., 2008. 368 p. (In Russian).
20. Kostromickij A. I., Volotka V. S. Approaches to Modeling Self-similar Traffic. *Vostochno-Evropeiskii zhurnal perezodnykh tekhnologii* [Eastern-European Journal of Enterprise Technologies], 2010, no. 46, pp. 46–49 (In Russian).

УВАЖАЕМЫЕ АВТОРЫ!

Научная электронная библиотека (НЭБ) продолжает работу по реализации проекта SCIENCE INDEX. После того как Вы регистрируетесь на сайте НЭБ (<http://elibrary.ru/defaultx.asp>), будет создана Ваша личная страничка, содержание которой составят не только Ваши персональные данные, но и перечень всех Ваших печатных трудов, имеющих в базе данных НЭБ, включая диссертации, патенты и тезисы к конференциям, а также сравнительные индексы цитирования: РИНЦ (Российский индекс научного цитирования), h (индекс Хирша) от Web of Science и h от Scopus. После создания базового варианта Вашей персональной страницы Вы получите код доступа, который позволит Вам редактировать информацию, помогая создавать максимально объективную картину Вашей научной активности и цитирования Ваших трудов.

МЕТОД И АЛГОРИТМЫ ОБНАРУЖЕНИЯ АНОМАЛИЙ В ТРАФИКЕ МУЛЬТИСЕРВИСНЫХ СЕТЕЙ СВЯЗИ, ОСНОВАННЫЕ НА НЕЧЕТКОМ ЛОГИЧЕСКОМ ВЫВОДЕ

С. А. Агеев^а, канд. техн. наук, доцент, ведущий научный сотрудник, serg123_61@mail.ru
И. Б. Саенко^а, доктор техн. наук, профессор, ведущий научный сотрудник, ibsaen@comsec.spb.ru
И. В. Котенко^а, доктор техн. наук, профессор, заведующий лабораторией проблем компьютерной безопасности, ivkote@comsec.spb.ru

^аСанкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

Введение: важнейшим требованием, предъявляемым к системам обнаружения и предупреждения вредоносных вторжений в современные телекоммуникационные инфраструктуры, является способность обнаруживать аномалии и, соответственно, угрозы вторжения в реальном времени. Сложность этой задачи во многом обусловлена нестационарностью, неполнотой, неточностью априорных знаний о законах распределения, которым подчиняются потоки в трафике мультисервисных сетей связи, их многообразием, а также изменяющимся характером злонамеренных действий со стороны атакующего, которые приводят компьютерные системы в небезопасное состояние. **Цель:** повышение оперативности и достоверности процесса обнаружения аномалий в сетевом трафике в условиях неполноты и высокой разнородности анализируемой информации. **Результаты:** предложены гибридные адаптивные метод и алгоритмы обнаружения аномалий в трафике мультисервисных сетей связи, работающие в режиме реального времени. Гибридный метод объединяет механизм безыдентификационной адаптации к изменяющимся параметрам трафика и нечеткий логический вывод, используемый для регулирования параметров алгоритмов и анализа выходных данных. Адаптивные алгоритмы ориентированы на комбинированную реализацию процедур модифицированной стохастической аппроксимации и псевдоградиентного поиска. Проведенная экспериментальная оценка показала, что алгоритмы имеют функциональные характеристики, максимально близкие к потенциально достижимым. **Практическая значимость:** разработанные метод и алгоритмы могут быть реализованы на существующих аппаратно-программных платформах на основе технологии интеллектуальных агентов. Их совместное использование с уже существующими методами и алгоритмами обнаружения вторжений может существенно повысить эффективность систем защиты информации в мультисервисных сетях связи.

Ключевые слова — мультисервисная сеть связи, обнаружение аномалий, трафик, стохастическая аппроксимация, нечеткий логический вывод.

Цитирование: Агеев С. А., Саенко И. Б., Котенко И. В. Метод и алгоритмы обнаружения аномалий в трафике мультисервисных сетей связи, основанные на нечетком логическом выводе // Информационно-управляющие системы. 2018. № 3. С. 61–68. doi:10.15217/issn1684-8853.2018.3.61

Citation: Ageev S. A., Saenko I. B., Kotenko I. V. Method and Algorithms of Anomaly Detection in Multiservice Network Traffic based on Fuzzy Logical Inference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 61–68 (In Russian). doi:10.15217/issn1684-8853.2018.3.61

Введение

В настоящее время в различных информационных инфраструктурах (системах управления, мобильных системах, на транспорте, в энергетике, экономике и т. д.) успешно внедряются технологии высокоскоростных телекоммуникаций и сетей нового поколения [1, 2]. Достижения в развитии этих технологий привели к концепции мультисервисной сети связи (МСС), ядром которой являются опорные IP-сети, интегрирующие услуги передачи речи, данных и мультимедиа и реализующие принцип конвергенции услуг электросвязи [3, 4].

Множество основных сервисов, предоставляемых пользователям с помощью МСС, хорошо известно. Однако появление большого количества

дополнительных сервисов у МСС делает достаточно острой проблему обеспечения ее информационной безопасности. Серьезность этой проблемы возрастает по следующим причинам: 1) из-за реализации процедур динамического изменения топологии МСС; 2) из-за возможного добавления или исключения из сети различного, априори неопределенного, числа абонентов; 3) в связи с динамическим изменением пространственного расположения абонентов; 4) из-за взаимодействия и сопряжения МСС друг с другом и т. д. Оперативность реагирования системы управления МСС на внешние и внутренние деструктивные воздействия приобретает особое значение для безопасности сети. Поэтому отсутствие аномалий в трафике сети является одним из критериев безопасного функционирования МСС.

При этом считается, что нормальный трафик соответствует политике безопасности в сети.

Трафик в МСС является достаточно разнообразным [2]. Он состоит, в том числе, из следующих компонентов: 1) мультимедийного трафика, который очень чувствителен к задержкам; 2) трафика передачи данных; 3) трафика передачи сигнальной информации; 4) трафика электронной почты. При этом заданные требования к качеству сервисов должны выполняться полностью. Однако существуют объективные трудности в построении системы управления МСС и в защите сетевой и абонентской информации. Эти трудности вызваны сложностью структуры МСС, разнородностью сети, необходимостью анализа большого количества различных сетевых и информационных параметров. Поэтому оперативное обнаружение аномалий сетевого трафика является одной из ключевых задач управления МСС и представляет собой актуальную научную проблему.

В настоящей статье рассматривается новый подход к обнаружению аномалий трафика в МСС, основанный на применении нечеткого логического вывода. При этом правила такого вывода используются совместно с модифицированными безыдентификационными алгоритмами.

Основной теоретический вклад работы заключается в следующем. Во-первых, обоснована модель мультисервисного трафика. Во-вторых, предложены методы и алгоритмы обнаружения аномалий трафика в МСС, основанные на применении нечеткого логического вывода. Наконец, экспериментально подтверждено, что предложенные алгоритмы обладают практически максимально возможным быстродействием.

Состояние исследований

Основные архитектурные решения, связанные с построением МСС, отражены во многих работах, например [1–4]. В них подчеркивается, что важная особенность построения МСС заключается в сильной сегментации топологии сети и в наличии нескольких точек сопряжения МСС с другими сетями. В результате общий трафик МСС уже нельзя контролировать из одной точки сети.

Для контроля трафика сетей, подобных МСС, в работах [4, 5] предлагается применять технологию интеллектуальных агентов. Интеллектуальные агенты выполняют сбор данных о трафике сети, предварительную обработку этих данных и передачу данных в центральные устройства системы управления. Таким образом, интеллектуальные агенты самостоятельно вырабатывают и реализуют часть управляющих решений.

Отмечается высокая важность задачи обеспечения функционирования систем обнаружения

аномалий трафика в режиме времени, близком к реальному [6]. Автор предлагает использовать смешанную централизованно-децентрализованную структуру для построения системы управления МСС. Такая структура позволяет значительно повысить оперативность принятия решений по противодействию деструктивным воздействиям на сеть, а также снизить служебный трафик управления.

Идея использования интеллектуальных агентов для контроля трафика МСС получила дальнейшее развитие в работе [7]. Там утверждается, что для реализации в МСС технологии интеллектуальных агентов необходимо разрабатывать и применять алгоритмы обнаружения аномалий мультисервисного трафика, которые являются простыми в реализации и устойчивыми к изменению параметров трафика. При этом данные алгоритмы должны функционировать в режиме времени, близком к реальному. Однако алгоритмы, предложенные в этой работе, не учитывали нечеткие факторы.

Таким образом, решения по алгоритмам обнаружения аномалий сетевого трафика, представленные в известных работах, не отвечают требованиям МСС. Это в основном связано с необходимостью обработки в этих алгоритмах нечеткой информации. В то же время известен ряд работ [8–17], в которых сделаны отдельные попытки применения методов обработки нечетких знаний для моделирования и оценки сетевого трафика. Однако непосредственное применение результатов, полученных в этих работах, для МСС является невозможным. Это объясняется тем, что статистические свойства трафика в МСС сильно отличаются как для различных условий эксплуатации, так и для различных приложений в МСС.

Известны другие подходы к обнаружению аномального трафика в сети, например, использующие непараметрические кумулятивные суммы [18], максимальную энтропийную оценку [19], историю изменений сетевого трафика [20], временные ряды от базы данных управления [21]. Однако эти подходы не могут быть отнесены к методам и алгоритмам, которые обеспечивают функционирование сети с данным качеством в режиме реального времени и могут быть достаточно просто реализованы. В нашей статье мы намереваемся устранить этот недостаток.

Модель мультисервисного трафика

Для разработки методов обнаружения аномалий в трафике МСС необходимо прежде всего сформировать модель мультисервисного трафика.

Модель мультисервисного трафика в общем виде представляет собой объединение множе-

ства различных стохастических процессов (СП). Поэтому предлагаемый подход к формированию модели мультисервисного трафика основан на учете следующих факторов: законов распределения СП; стационарности СП; самоподобия СП; характеристик, выбранных для анализа СП.

Отмечается [8, 9], что трафик для различных приложений в МСС может быть аппроксимирован с помощью многих вероятностных распределений, основными из которых являются распределения Пуассона, Парето, Вейбулла, лог-нормальное и экспоненциальное. Для моделирования различных типов трафика применяются различные законы распределения. Например, если моделируемый трафик является «Аудио» или «Видео», то тогда он наделяется эффектом самоподобия, и для его моделирования применяется распределение Парето. Если моделируемый трафик сформирован протоколами SMTP/TCP, то тогда применяются распределение Пуассона или экспоненциальное распределение. Полный перечень законов распределения трафика в МСС и их распределение по уровням модели ISO/OSI можно найти в статье [22].

Стационарность или нестационарность СП также является важным фактором. Проще всего решать задачу обнаружения аномалий в трафике, если он является стационарным. Однако во многих работах, например в [8, 9], отмечается, что трафик в МСС является нестационарным по своей природе. Это существенно осложняет обнаружение аномалий в трафике, так как аномалии могут восприниматься как нормальное поведение трафика.

Анализ трафика может вестись с помощью различных характеристик случайного процесса. Основными такими характеристиками являются максимальное, минимальное и среднее значение интенсивности процесса, среднее квадратическое отклонение и другие. Ниже в работе будет рассматриваться среднее значение интенсивности процесса, которое рассчитывается по формуле

$$S_{mid} = \frac{1}{N} \sum_{i=1}^N S_i, \quad (1)$$

где N — размер выборки; S_i — элемент i в выборке.

Выражения для расчета остальных характеристик стохастических процессов приведены в работе [22].

Метод и алгоритмы обнаружения аномалий трафика

К разработанным алгоритмам обнаружения аномалий в трафике МСС предъявляются следующие требования: функционирование в режиме реального или близкого к реальному времени; под-

держание заданного качества сервиса; простота реализации. Алгоритмы относятся к классу гибридных адаптивных алгоритмов идентификации параметров трафика. Они применяются как для стационарных, так и для нестационарных трафиков. Трафики моделируются посредством СП. Каждый СП относится к соответствующему классу, определяемому законом распределения СП.

Сущность гибридного метода обнаружения аномалий в трафике МСС состоит в том, что, с одной стороны, используются алгоритмы безыдентификационной адаптации к изменяющимся параметрам СП. С другой стороны, для настройки параметров алгоритмов и принятия решений используется нечеткий логический вывод. Применение такого гибридного подхода вызвано необходимостью оценки как текущих точечных, так и интегральных параметров трафика. Оценка интегральных параметров трафика производится в скользящем окне. Оценка текущих точечных параметров выполняется от точки к точке одновременно с процедурой интегральной оценки. При этом оценки, полученные в скользящем окне, используются как начальные значения для безыдентификационной процедуры. Для проведения этой процедуры используются два алгоритма, отличающиеся своими возможностями по аппроксимации СП. Комбинация таких подходов позволяет как оценивать интегральные свойства СП, так и отслеживать динамику его поведения. Применение нечеткого логического вывода позволяет строить параметрические оценки параметров алгоритма по малому числу наблюдений. В этом случае становится возможным делать нечеткие выводы относительно аномалий трафика.

Пусть СП задан в дискретных значениях времени $t_i = i, i = 1, 2, \dots$. Первый алгоритм является алгоритмом *модифицированной стохастической аппроксимации* (МСА) [11]

$$S_{i+1} = S_{i-1} + \mu_i (S_i - S_{i-1}), \quad (2)$$

где S_i — среднее значение интенсивности трафика в момент времени i ; μ_i — параметр алгоритма на шаге i .

Второй алгоритм используется в случае, когда плотность вероятности значений СП является симметричной. Он основан на применении *псевдоградиентных процедур* (ПГП) вида

$$S_{i+1} = S_{i-1} + \mu_i \text{sign} (S_i - S_{i-1}). \quad (3)$$

Такой выбор первого и второго алгоритмов обеспечивает оценку динамических свойств СП. Из-за того, что алгоритмы МСА и ПГП относятся к классу безыдентификационных алгоритмов, время анализа трафика и обнаружения аномалий существенно сокращается.

В качестве целевой функции предлагается квадратичная функция вида

$$M\{(S_i - S_{i-1})^2\}, \quad (4)$$

где $M\{\cdot\}$ — функция вычисления математического ожидания.

Параметр μ_i для МСА и ППП должен удовлетворять следующим условиям:

$$0 < \mu_i < 1, \mu_i = \text{const}. \quad (5)$$

Особенностью алгоритмов МСА и ППП является тот факт, что необходимо подстраивать значение параметра μ_i для различных оцениваемых СП и их статистических свойств. Предлагается процедуру подстройки этих алгоритмов осуществлять на основе метода нечеткого вывода Мамдани [5, 11], согласно которому идентификация параметров алгоритмов МСА и ППП производится с помощью правил, имеющих в общем случае следующий вид:

$$\text{IF } \langle S_i = A \rangle \text{ AND } \langle \sigma = B \rangle \text{ AND} \\ \langle \rho = D \rangle \text{ THEN } \mu = R, \quad (6)$$

где A , B и D — нечеткие пороговые значения, определяемые в ходе обучения системы нечеткого логического вывода.

Идентификация аномалий в трафике МСС осуществляется на основе нечеткого логического вывода следующего вида:

$$\text{IF } \langle S_i = A \rangle \text{ AND} \\ \langle A \text{ соответствует политике безопасности} \rangle \\ \text{THEN } R, \quad (7)$$

где R — значение лингвистической переменной, оценивающей наличие аномалий.

Для выполнения процедуры идентификации аномалий необходимо предварительно проводить обучение системы нечеткого логического вывода по экспериментальным данным. Экспериментальные данные формируются заранее.

При выборе размера скользящего окна необходимо найти разумный компромисс между скоростью изменения значений СП, размером окна и репрезентативностью выборки значений СП. Этот компромисс необходим для устранения эффекта излишнего сглаживания значений СП.

Вывод о характере поведения трафика в МСС также осуществляется на основе метода Мамдани.

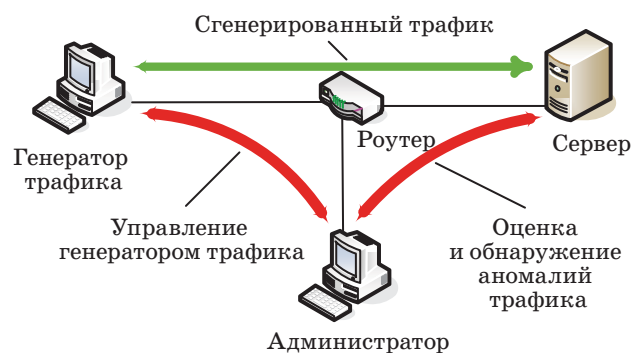
Анализ экспериментальных данных

Для экспериментальной проверки предложенных методов и алгоритмов был использован инструментальный стенд (рис. 1). Стенд состоял

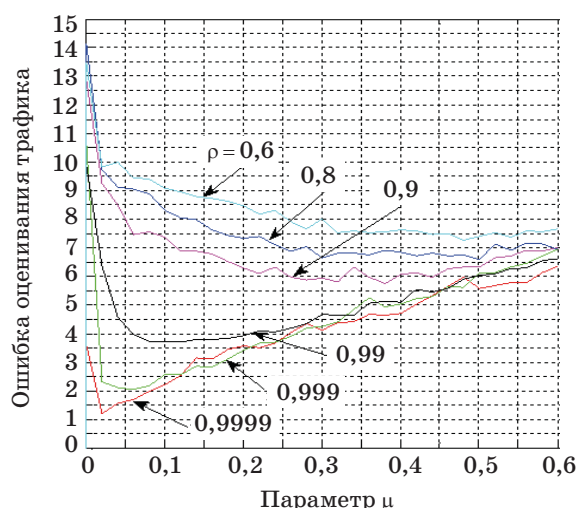
из сервера и двух рабочих станций, объединенных в сеть с помощью маршрутизатора. Одна рабочая станция являлась генератором трафика. На второй станции работал администратор сети. Администратор сети управлял генератором. Генератор формировал трафик с заданным законом распределения. При этом аномалии внедрялись в трафик случайным образом в соответствии с выбранным типом компьютерной атаки. Трафик проходил через маршрутизатор на сервер. Задача администратора заключалась в оценке и определении аномалий трафика.

Генератор формировал стационарные и нестационарные трафики с распределениями Пуассона и Парето. Эти распределения являются наиболее характерными для МСС. Для генерации аномалий использовались атаки типа DoS, приводящие к перегрузке канала связи. Нестационарные процессы моделировались как мультипликативные СП с детерминированными и случайными модулирующими функциями. Случайные модулирующие функции представляли собой случайные процессы авторегрессии первого порядка с различными коэффициентами корреляции. Адаптация к параметрам пуассоновского СП осуществлялась с помощью алгоритма МСА. Для СП с законом Парето применялся алгоритм ППП.

Настройка параметров алгоритмов адаптации осуществлялась с помощью правил, имеющих вид (6). Нечеткие константы A , B и D для этих правил определялись предварительно по результатам обучения. Для этой цели использовались экспериментальные результаты, представленные на рис. 2, на которых показаны зависимости среднеквадратического значения ошибки оценивания интенсивности трафика нестационарного СП от значения коэффициента шага алгоритма МСА при различных значениях коэффициента корреляции, модулирующего СП для пуассоновских трафиков. Случайный процесс имел среднее значение $a_0 = 100$ и среднеквадратическое отклонение $\sigma = 10$. Коэффициент корреляции тренда ρ



■ Рис. 1. Структура инструментального стенда
■ Fig. 1. Testbed structure



■ **Рис. 2.** Зависимость ошибки оценивания трафика от величины параметра μ
 ■ **Fig. 2.** Dependence of the traffic assessment error on the parameter μ

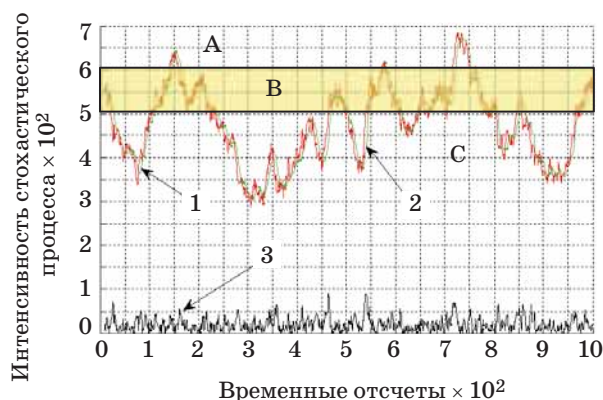
для различных зависимостей изменялся в диапазоне от 0,6 до 0,9999. Видно, что при малых значениях μ погрешности оценки трафика сильно отличаются друг от друга в зависимости от значения ρ .

При больших значениях μ эти погрешности принимают приблизительно равные значения.

По данным, показанным на рис. 2, формируются правила настройки параметров МСА. Например, при $S_i \approx 100$, $\sigma \approx 10$ и $\rho \approx 0,9999$ параметр алгоритма μ приблизительно равен 0,025. Для $\rho \approx 0,99$ параметр μ будет иметь значение, приблизительно равное 0,12. Параметры σ и ρ оцениваются в скользящем окне.

Пример численного моделирования нестационарного пуассоновского СП представлен на рис. 3. Коэффициент корреляции процесса изменялся в экспериментах от 0,9 до 0,99999. СП имел следующие параметры: начальное значение среднего значения тренда $m_0 = 500$, коэффициент корреляции тренда $\rho = 0,99$, размер скользящего окна $r = 10$, параметр $\mu = 0,05$, среднеквадратическое отклонение тренда $\sigma = 100$. Как видно из рисунка, предложенный алгоритм работает практически без задержек и имеет достаточно высокую точность оценивания.

С определением аномалий связаны три зоны (см. рис. 3). Зона А определяет, что риск наличия аномалий есть, и уровень риска является неприемлемым. Зона В показывает, что риск наличия аномалий есть, и уровень риска является приемлемым. В трафике, расположенном в зоне С, аномалий нет. Границы зон А, В и С определяются на этапе предварительной настройки базы знаний машины нечеткого вывода в соответствии

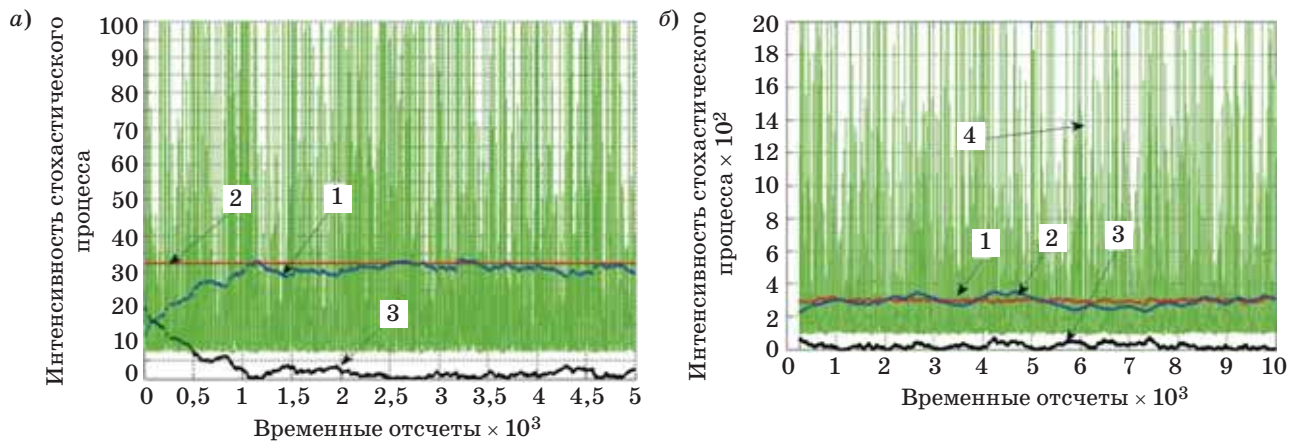


■ **Рис. 3.** Оценка нестационарного тренда с распределением Пуассона: 1 — истинное значение тренда; 2 — оценка тренда; 3 — ошибка оценивания тренда
 ■ **Fig. 3.** Assessment of a non-stationary trend with the Poisson distribution: 1 — the true value of a trend; 2 — trend assessment; 3 — estimation error of a trend

■ **Таблица 1.** Параметры стационарного и нестационарного процесса
 ■ **Table 1.** Parameters of stationary and nonstationary process

Параметр	Стационарный процесс	Нестационарный процесс
Параметр Херста Н	0,85	0,75
Среднее значение СП	32,5	250
Параметр распределения СП	1,3	—
Параметр алгоритма МСА μ	0,004	0,13
Относительная погрешность оценивания Δ , %	$\leq 6,2$	$\leq 9,1$

с принятыми правилами политики безопасности. Значения оценок тренда на границах зон А, В и С используются для построения правил Мамдани по формуле (7). Например, правило для вывода об отсутствии аномалий имеет следующий вид: **IF** $\langle S_i \in \text{Зона C} \rangle$ **THEN** *Аномалий трафика нет.*
 Результаты оценки параметров (табл. 1) самоподобных трафиков с распределением Парето представлены на рис. 4, а и б.



■ **Рис. 4.** Оценка стационарного (а) и нестационарного (б) тренда самоподобного трафика с распределением Парето: 1 — оценка тренда; 2 — истинное значение тренда; 3 — ошибка оценивания тренда; 4 — значения СП

■ **Fig. 4.** Assessment of a stationary (а) and non-stationary (б) trend of a self-similar traffic with the Pareto distribution: 1 — trend assessment; 2 — the true value of a trend; 3 — estimation error of a trend; 4 — stochastic process values

■ **Таблица 2.** Экспериментальные результаты оценки алгоритмов

■ **Table 2.** Experimental results of the algorithms assessment

Распределение	Коэффициент корреляции ρ	Средняя относительная погрешность оценки, %	Время обнаружения, временные отсчеты
Парето (H = 0,75)	0,99	12,0	80
	0,999	10,0	65
	0,9999	8,0	50
Пуассона	0,9	8,6	34
	0,99	7,6	27
	0,999	6,4	20
	0,9999	5,3	15

Итоговые результаты оценки разработанных алгоритмов определения аномалий в трафике МСС для рассмотренных выше распределений приведены в табл. 2.

Из этой таблицы видно, что величина относительной погрешности алгоритмов не превышает 12 % от истинного значения тренда. Этот результат следует считать достаточно хорошим, учитывая, что алгоритмы работают в реальном времени, так как полученные для времени обнаружения значения лежат в диапазоне от 15 до 80

временных отсчетов. Таким образом, разработанные алгоритмы обнаружения аномалий в трафике МСС полностью удовлетворяют требованиям оперативности и точности.

Заключение

В настоящей работе предложен гибридный метод обнаружения аномалий в трафике МСС, использующий алгоритмы безыдентификационной адаптации и нечеткого вывода Мамдани.

Ключевой особенностью мультисервисного трафика как объекта оценивания на существование аномалий является наличие в нем стохастических процессов, подчиненных различным законам распределения. Для экспериментальной оценки предложенных метода и алгоритмов были выбраны законы распределения Пуассона и Парето, определяющие предельные случаи регулярности трафика.

Экспериментальная оценка предложенных метода и алгоритмов показала, что они позволяют оценивать тренды указанных стохастических процессов в реальном времени, с высокой точностью и с сохранением качества обслуживания.

Дальнейшие направления исследований связаны с разработкой методов обучения предложенных алгоритмов с использованием тестовых трафиков.

Работа выполнена при частичной финансовой поддержке РФФИ (проекты 16-29-09482, 18-07-01369 и 18-07-01488), бюджетной темы № АААА-А16-116033110102-5, а также при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-У01).

Литература

1. Гольдштейн А. Б., Гольдштейн Б. С. SOFT-SWITCH. — СПб.: БХВ, 2006. — 368 с.
2. Kanáliková A. Services in NGN — Next Generation Networks // Journal of Information, Control and Management Systems. 2005. Vol. 3. N 2. P. 97–102.
3. Агеев С. А., Шерстюк Ю. М., Саенко И. Б., Полубелова О. В. Концептуальные основы автоматизации управления защищенными мультисервисными сетями // Проблемы информационной безопасности. Компьютерные системы. 2011. № 3. С. 30–39.
4. Агеев С. А., Бушуев А. С., Егоров Ю. П., Саенко И. Б. Концепция автоматизации управления информационной безопасностью в защищенных мультисервисных сетях специального назначения // Автоматизация процессов управления. 2011. № 1. С. 50–57.
5. Gorodetski V., Kotenko I., Karsaev O. Multi-Agent Technologies for Computer Network Security: Attack Simulation, Intrusion Detection and Intrusion Detection Learning // International Journal of Computer Systems Science & Engineering. 2003. N 4. P. 191–200.
6. Paxson V. A System for Detecting Network Intruders in Real-Time // Proc. of the 7th USENIX Security Symp. 1998. P. 2435–2463.
7. Paxson V. A System for Detecting Network Intruders in Real-Time // Computers Networks. 1999. N 31. P. 2435–2463.
8. Laskin N., Lambadaries I., Harmatzis F. C., Devetsikiotis M. Fractional Levy Motion and its Application to Network Traffic Modeling // Elsevier Comp. Network. 2002. Vol. 40. P. 363–375.
9. Шелухин О. И., Осин А. В., Смольский С. М. Самоподобие и фракталы. Телекоммуникационные приложения. — М.: Физматлит, 2008. — 368 с.
10. Ageev S., Vasil'ev K. Adaptive Algorithms for Decorrelation to Image Processing // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Application. 2001. Vol. 11. N 1. P. 131–134.
11. Takagi T., Sugeno M. Fuzzy Identification of Systems and its Applications to Modeling and Control // IEEE Trans. on System, Man and Cybernetics. 1985. Vol. 15. N 1. P. 11–132.
12. Derrac J., García S., Herrera F. Fuzzy Nearest Neighbor: Taxonomy, Experimental Analysis and Prospects // Inf. Sci. 2014. Vol. 260. P. 98–119.
13. Li X., Lara-Rosano F. Adaptive Fuzzy Petri Nets for Dynamic Knowledge Representation and Inference // Expert System Applications. 2000. Vol. 19. N 3. P. 235–241.
14. Xing-zhu W. Network Information Security Situation Assessment based on Bayesian Network // International Journal of Security and its Applications. 2016. Vol. 10. N 5. P. 129–138.
15. Papageorgiou E. I. Fuzzy Cognitive Maps for Applied Sciences and Engineering: from Fundamentals to Extensions and Learning Algorithms. — Springer-Verlag Berlin Heidelberg, 2014. — 395 p. doi:10.1007/978-3-642-39739-4
16. Astanin S. V., Zhukovskaya N. K. Business Processes Control via Modeling by Fuzzy Situational Networks // Automation and Remote Control. 2014. Vol. 75. N 3. P. 570–579.
17. Nikolaev A. B., Sapego Yu. S., Jakubovich A. N., Berner L. I., Stroganov V. Yu. Fuzzy Algorithm for the Detection of Incidents in the Transport System // International Journal of Environmental & Science Education. 2016. Vol. 11. N 16. P. 9039–9059.
18. Wang H., Zhang D., Shin K. G. Detecting SYN Flooding Attacks // Proc. of IEEE INFOCOM. 2002. 10 p.
19. Staniford S., Hoagland J., McAlerney J. M. Practical Automated Detection of Stealthy Portscans // J. Comput. Secur. 2002. Vol. 10. N 1–2. P. 105–136.
20. Brutlag J. D. Aberrant Behavior Detection in Time Series for Network Service Monitoring // Proc. of the 14th Systems Administration Conf. 2000. P. 139–146.
21. Thottan M., Ji C. Anomaly Detection in IP Networks // IEEE Trans. Signal Processing. 2003. Vol. 51. N 8. P. 2191–2204.
22. Котенко И. В., Саенко И. Б., Агеев С. А., Копчак Я. М. Обнаружение аномального трафика в сетях Интернета вещей на основе нечеткого логического вывода // XVIII Междунар. конф. по мягким вычислениям и измерениям (SCM/2015): сб. докл. 2015. Т. 1. С. 9–14.

UDC 681.142.33:681.14

doi:10.15217/issn1684-8853.2018.3.61

Method and Algorithms of Anomaly Detection in Multiservice Network Traffic based on Fuzzy Logical Inference

Ageev S. A.^a, PhD, Tech., Associate Professor, Leading Researcher, serg123_61@mail.ruSaenko I. B.^a, Dr. Sc., Tech., Professor, Leading Researcher, ibsaen@comsec.spb.ruKotenko I. V.^a, Dr. Sc., Tech., Professor, Head of Laboratory of Computer Security Problems, ivkote@comsec.spb.ru^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: The major requirement imposed to the systems of detection and prevention of malicious invasions into modern telecommunication infrastructures is the ability to find anomalies and invasion threats in real time. The complexity of this problem is in many respects caused by the incompleteness, discrepancy and diversity of the a priori knowledge about the distribution laws peculiar

to the traffic of multiservice communication networks, as well as by the changing nature of malicious actions which make computer systems unsafe. **Purpose:** Increasing the speed and reliability of network traffic anomaly detection when the analyzed information is incomplete and highly heterogeneous. **Results:** A hybrid method and adaptive algorithms have been proposed for real-time anomaly detection in multiservice communication network traffic. The hybrid method unites the mechanism of non-identificational adaptation to the changing traffic parameters with the fuzzy logical inference used for regulating the algorithm parameters and for analyzing the output data. The adaptive algorithms are focused on combined implementation of modified stochastic approximation and pseudo-gradient search procedures. An experimental assessment has shown that the functional characteristics of the algorithms are close to the potentially achievable ones. **Practical relevance:** The developed method and algorithms can be implemented on available hardware-software platforms on the basis of intellectual agent technology. Sharing them with already existing methods and algorithms of invasion detection can considerably increase the efficiency of information security systems in multiservice communication networks.

Keywords — Multiservice Communication Network, Anomaly Detection, Traffic, Stochastic Approximation, Fuzzy Logical Inference.

Citation: Ageev S. A., Saenko I. B., Kotenko I. V. Method and Algorithms of Anomaly Detection in Multiservice Network Traffic based on Fuzzy Logical Inference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 61–68 (In Russian). doi:10.15217/issn1684-8853.2018.3.61

References

- Gol'dshtejn A. B., Gol'dshtejn B. S. *SOFTSWITCH* [SOFTSWITCH]. Saint-Petersburg, BkHV Publ., 2006. 368 p. (In Russian).
- Kanálíková A. Services in NGN — Next Generation Networks. *Journal of Information, Control and Management Systems*, 2005, vol. 3, no. 2, pp. 97–102.
- Ageev S. A., Sherstjuk Y. M., Saenko I. B., Polubelova O. V. Conceptual Basics of Protected Multiservice Networks' Control Automation. *Problemy informacionnoj bezopasnosti. Komp'yuternye sistemy* [Information Security Problems. Computer Systems], 2011, no. 3, pp. 30–39 (In Russian).
- Ageev S. A., Bushuev A. S., Egorov Y. P., Saenko I. B. The Concept of Information Security Control Automation in the Protected Multiservice Networks of a Special Purpose. *Avtomatizacija processov upravlenija*, 2011, no. 1, pp. 50–57 (In Russian).
- Gorodetski V., Kotenko I., Karsaev O. Multi-agent Technologies for Computer Network Security: Attack Simulation, Intrusion Detection and Intrusion Detection Learning. *International Journal of Computer Systems Science & Engineering*, 2003, no. 4, pp. 191–200.
- Paxson V. A System for Detecting Network Intruders in Real-Time. *Proc. of the 7th USENIX Security Symp.*, 1998, pp. 2435–2463.
- Paxson V. A System for Detecting Network Intruders in Real-Time. *Computers Networks*, 1999, no. 31, pp. 2435–2463.
- Laskin N., Lambadaries I., Harmatzis F. C., Devetsikiotis M. Fractional Levy Motion and its Application to Network Traffic Modeling. *Elsevier Comp. Network*, 2002, vol. 40, pp. 363–375.
- Sheluhin O. I., Osin A. V., Smol'skij S. M. *Samopodobie i fraktaly. Telekomunikacionnye prilozheniya* [Self-Similarity and Fractals. Telecommunication Applications]. Moscow, Fizmatlit Publ., 2008. 368 p. (In Russian).
- Ageev S., Vasil'ev K. Adaptive Algorithms for Decorrelation to Image Processing. *Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Application*, 2001, vol. 11, no. 1, pp. 131–134.
- Takagi T., Sugeno M. Fuzzy Identification of Systems and its Applications to Modeling and Control. *IEEE Trans. on System, Man and Cybernetics*, 1985, vol. 15, no. 1, pp. 11–132.
- Derrac J., García S., Herrera F. Fuzzy Nearest Neighbor: Taxonomy, Experimental Analysis and Prospects. *Inf. Sci.*, 2014, vol. 260, pp. 98–119.
- Li X., Lara-Rosano F. Adaptive Fuzzy Petri Nets for Dynamic Knowledge Representation and Inference. *Expert System Applications*, 2000, vol. 19, no. 3, pp. 235–241.
- Xing-zhu W. Network Information Security Situation Assessment based on Bayesian Network. *International Journal of Security and its Applications*, 2016, vol. 10, no. 5, pp. 129–138.
- Papageorgiou E. I. Fuzzy Cognitive Maps for Applied Sciences and Engineering: from Fundamentals to Extensions and Learning Algorithms. Springer-Verlag Berlin Heidelberg, 2014. 395 p. doi:10.1007/978-3-642-39739-4
- Astanin S. V., Zhukovskaya N. K. Business Processes Control via Modeling by Fuzzy Situational Networks. *Automation and Remote Control*, 2014, vol. 75, no. 3, pp. 570–579.
- Nikolaev A. B., Sapego Yu. S., Jakubovich A. N., Berner L. I., Stroganov V. Yu. Fuzzy Algorithm for the Detection of Incidents in the Transport System. *International Journal of Environmental & Science Education*, 2016, vol. 11, no. 16, pp. 9039–9059.
- Wang H., Zhang D., Shin K. G. Detecting SYN Flooding Attacks. *Proc. of IEEE INFOCOM*, 2002, 10 p.
- Staniford S., Hoagland J., McAlerney J. M. Practical Automated Detection of Stealthy Portscans. *J. Comput. Secur.*, 2002, vol. 10, no. 1–2, pp. 105–136.
- Brutlag J. D. Aberrant Behavior Detection in Time Series for Network Service Monitoring. *Proc. of the 14th Systems Administration Conf.*, 2000, pp. 139–146.
- Thottan M., Ji C. Anomaly Detection in IP Networks. *IEEE Trans. Signal Processing*, 2003, vol. 51, no. 8, pp. 2191–2204.
- Kotenko I. V., Saenko I. B., Ageev S. A., Kopchak Y. M. Abnormal Traffic Detection in Networks of the Internet of Things based on Fuzzy Logical Inference. *Sbornik dokladov XVIII Mezhdunarodnoj konferentsii po myagkim vychisleniyam i izmereniyam (SCM'2015)* [Proc. of the XVIII Int. Conf. on Soft Computing and Measurements (SCM'2015)], 2015, vol. 1, pp. 9–14 (In Russian).

МЕТОД ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ПЕРСОНАЛЬНЫХ ДАННЫХ ПРИ ИХ ОБРАБОТКЕ В ИНФОРМАЦИОННОЙ СИСТЕМЕ НА ОСНОВЕ АНАЛИЗА ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ

И. С. Козин^а, ведущий специалист, van@trioptimum.com

^аАО «Кронштадт Технологии», Малый пр. В. О., 54, корп. 5, лит. П, Санкт-Петербург, 199178, РФ

Постановка проблемы: с развитием информационных технологий появляются новые классы средств защиты персональных данных при их обработке в информационных системах. Одним из классов средств защиты информации являются системы анализа поведения пользователей. При разработке средств анализа поведения все большее распространение получают методы машинного обучения, в том числе с применением математического аппарата теории искусственных нейронных сетей. Однако подходы к разработке средств защиты информации, основанные на машинном обучении, на сегодня изучены недостаточно. **Цель:** разработка метода создания искусственной нейронной сети, обеспечивающей проведение анализа санкционированного поведения пользователей информационной системы и выявление аномалий в поведении, сигнализирующих о совершении противоправных действий. **Результаты:** обзор подходов к обеспечению безопасности информации с применением искусственных нейронных сетей показал активное их развитие по разным направлениям, в том числе в направлении выявления аномалий. Разработан метод создания искусственной нейронной сети, включающий предложения по определению типа нейронной сети, области числовых значений входных и выходного сигналов, количества слоев и нейронов в слоях, метода обучения, а также типа активационных функций. В качестве входных значений предложено использовать характеристики поведения пользователя: набор данных, с которыми работает пользователь; место доступа к информационной системе; набор действий, которые совершает пользователь; время, в которое осуществляются доступ или определенные действия; общая продолжительность проводимых в течение определенного времени работ. На примере времени выполнения доступа пользователя предложен подход к присвоению характеристике пользователя числовых значений, основанный на применении математического аппарата теории нечетких множеств. **Практическая значимость:** обученная нейронная сеть обеспечивает более оперативное выявление аномалий в поведении пользователя, чем анализ специалиста по обеспечению безопасности информации без использования специальных средств автоматизации.

Ключевые слова – информационная безопасность, персональные данные, анализ поведения, искусственная нейронная сеть, теория нечетких множеств.

Цитирование: Козин И. С. Метод обеспечения безопасности персональных данных при их обработке в информационной системе на основе анализа поведения пользователей// Информационно-управляющие системы. 2018. № 3. С. 69–78. doi:10.15217/issn1684-8853.2018.3.69

Citation: Kozin I. S. Providing Personal Data Protection in an Information System based on User Behavior Analytics. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 69–78 (In Russian). doi:10.15217/issn1684-8853.2018.3.69

Введение

Неблагоприятная геополитическая обстановка и активизация действий террористов обуславливают актуальность проблем обеспечения безопасности, одним из важнейших аспектов которой является информационная безопасность. С учетом положений Указа Президента [1] в качестве одного из основных направлений обеспечения безопасности информации можно выделить защиту персональных данных. Под персональными данными понимается любая информация, относящаяся к прямо или косвенно определенному или определяемому физическому лицу (субъекту персональных данных) [2]. Примерами персональных данных являются паспортные данные, сведения о состоянии здоровья, политических взглядах, религиозных и философских убежде-

ниях. Нарушение безопасности персональных данных может привести к материальному и моральному ущербу как для субъекта персональных данных, так и для организации-оператора персональных данных.

В настоящее время существует множество технических решений по обеспечению безопасности информации при ее обработке в информационной системе. Одним из таких решений являются системы управления информационной безопасностью (СУИБ). СУИБ начали получать активное распространение примерно с 2012 г., во время динамичного развития технологий аналитической работы с большими объемами данных и машинного обучения.

Одним из направлений развития СУИБ стала разработка систем анализа поведения пользователей (САПП). В задачи САПП входит анализ

действий пользователей (состав обрабатываемых данных, контроль используемых устройств и приложений, учет взаимодействий с другими пользователями и т. п.) и выявление аномалий в их поведении. В общем случае в ходе работы САПП каждому пользователю присваивается определенный уровень надежности, отражающий общую адекватность его поведения в удобном для восприятия администратором безопасности виде. Примерами практического применения САПП являются выявление аномальных действий, совершаемых от имени:

— служебных учетных записей (например, использование учетных записей, предназначенных для обеспечения определенных сервисов, в иных целях);

— привилегированных учетных записей (администратор домена осуществляет массовый сбор рабочих материалов пользователей и т. п.);

— учетных записей обычных пользователей (активный анализ доступных сетевых ресурсов, доступ в несвойственное для пользователя время или из несвойственного места, параллельный доступ из нескольких мест, резко возросшие объемы исходящего в Интернет трафика и т. п.).

Выявление аномалий, распределенных во времени или среди нескольких пользователей, как правило, затруднительно, если оно основано на применении экспертных систем, и требует значительных временных и вычислительных ресурсов. Из-за большого разнообразия действий пользователей даже регулярные обновления базы данных правил экспертной системы не способны гарантировать точной идентификации всего диапазона аномалий. Одним из подходов к устранению указанных затруднений может быть использование в составе САПП интеллектуальных подсистем, разработанных с помощью методик машинного обучения.

Машинное обучение является одним из направлений развития искусственного интеллекта и за счет применения различных математических аппаратов (таких как математическая статистика, теория вероятностей, численные методы оптимизации и т. п.) позволяет решать задачи классификации, кластеризации, систематизации, предсказания и регрессии. В направлении машинного обучения можно выделить искусственные нейронные сети (ИНС).

Развитие ИНС вдохновляется биологией. При работе с ИНС, рассматривая различные сетевые конфигурации и алгоритмы, исследователи применяют термины, заимствованные из принципов организации мозговой деятельности. В силу ограниченности знаний о работе мозга разработчикам ИНС приходится выходить за пределы современных биологических знаний в поисках структур, способных выполнять полезные функ-

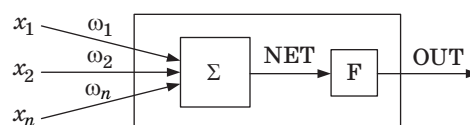
ции [3]. Элементарной структурной единицей ИНС является искусственный нейрон [4–6], схематично представленный на рис. 1.

В общем случае основными элементами искусственного нейрона являются: входные сигналы x (соответствующие сигналам, приходящим в синапсы биологического нейрона), веса ω_n (на которые умножаются входные сигналы, соответствуют «силе» биологических синаптических связей), суммирующий блок Σ (принимающий и суммирующий входные сигналы, соответствует телу биологического нейрона), выходной сигнал NET (создается суммирующим блоком, является алгебраической суммой взвешенных входов), активационная функция F (преобразующая выходной сигнал NET) и непосредственно выходной нейронный сигнал OUT. Математически нейрон может быть представлен формулой

$$OUT = f \left(\sum_{n=1}^m (x_n \omega_n) \right).$$

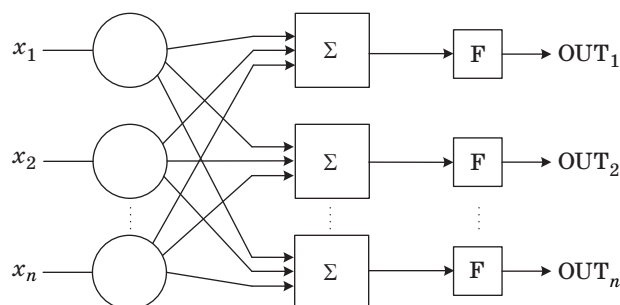
Один нейрон способен выполнять простейшие процедуры распознавания, однако для серьезных нейронных вычислений необходимо соединять нейроны в сети (рис. 2).

Способность ИНС обучаться, адаптироваться под новые типы поведения и распознавать их, даже если ранее она с ними не сталкивалась, придает системе защиты информации, разработанной с их применением, определенную гибкость. Обучают ИНС на определенной выборке приме-



■ Рис. 1. Обобщенная схема нейрона

■ Fig. 1. General scheme of the neuron



■ Рис. 2. Обобщенная схема однослойной ИНС

■ Fig. 2. General scheme of the one-layer neural network

ров, после чего ее реакция анализируется, и ИНС настраивается таким образом, чтобы достичь удовлетворительных результатов.

Обзор подходов к обеспечению безопасности с применением ИНС

Универсальность, которая изначально была заложена в ИНС, обуславливает активное развитие этого направления как в целом [4, 5, 7–9], так и в области обеспечения информационной безопасности: при защите от сетевых атак и вторжений [10–14], антивирусной защите [15], фильтрации спама [16, 17], анализе безопасности [18–20], анализе угроз [21, 22], разработке адаптивных средств защиты [23–25], выявлении аномалий [26] и пр. [27, 28]. В настоящее время существует большое количество разных конфигураций нейронных сетей с различными принципами функционирования [10]. Однако практически все они связаны с выбором и анализом некоторых частных видов структур с известными свойствами (сети Хопфилда, Гроссберга, Кохонена) [4, 29]. Как отмечается в работе [11], наиболее популярными и изученными являются следующие: многослойный перцептрон, сети Кохонена, нейронные сети встречного распространения, сети Хопфилда и Хэмминга, сеть с радиальными базисными элементами (RBF), вероятностная нейронная сеть (PNN), обобщенно-регрессионная нейронная сеть (GRNN) и линейные нейронные сети. Использование существующих ИНС открывает широкие возможности в области обеспечения информационной безопасности, но вместе с тем оно связано и с рядом проблемных вопросов [4, 7, 30, 31].

Постановка задачи определения типа и основных характеристик ИНС

В настоящей статье предлагается метод обеспечения безопасности персональных данных при их обработке в информационной системе, основанный на проведении анализа санкционированного поведения пользователей информационной системы персональных данных с применением ИНС.

Предполагается, что каждый пользователь информационной системы персональных данных обладает набором характеристик, совокупность которых выражает его уникальное типовое поведение. К таким характеристикам предлагается отнести:

— набор данных, с которыми работает пользователь (файлы, папки, сетевые объекты, интернет-сайты и т. п.);

— место осуществления доступа к информационной системе персональных данных (конкретный компьютер, № помещения, здание, город, страна и т. п.);

— набор действий, которые выполняет пользователь (чтение, запись, копирование, модификация и т. п.);

— время, в которое осуществляется доступ или выполняются определенные действия (время суток, день недели, определенные числа и т. п.);

— общую продолжительность выполняемых в течение определенного времени действий.

Предложенный набор характеристик не является исчерпывающим, но позволяет построить уникальную модель поведения пользователя. Отступление от модели поведения (выявление аномалий в поведении) может свидетельствовать о совершении противоправных действий. Примерами таких действий являются:

— массовое удаление материалов, к которым имеет доступ пользователь (практикуется многими недовольными работниками при увольнении);

— использование чужой учетной записи (практикуется пользователями, несерьезно относящимися к правилам разграничения доступа);

— беспорядочное ознакомление или копирование корпоративной информации (практикуется любопытными пользователями и инсайдерами).

Каждая из представленных характеристик пользователя может быть рассмотрена применительно к группе пользователей. Таким образом, появляется возможность выявлять аномалии в поведении не только пользователя, но и групп. Такой подход может найти применение при сговоре среди пользователей и совершении санкционированных неправомерных действий, распространенных среди нескольких человек и потому особо затруднительных в выявлении.

Каждую характеристику пользователя (или группы пользователей) можно описать в виде коэффициентов x_n , $n \in \{1; 5\}$, где n выражает порядковый номер характеристики:

x_1 — набор данных, с которыми работает пользователь;

x_2 — точка доступа пользователя к ИСПДн;

x_3 — набор совершаемых пользователем действий;

x_4 — время осуществления доступа;

x_5 — общая продолжительность проводимых работ.

Близкий по своему составу набор характеристик поведения использовался ранее [32]. Совокупность характеристик поведения x обозначим вектором поведения x .

В качестве математического аппарата выявления аномалий в поведении пользователя предлагается использовать теорию ИНС. Таким образом, на входе нейросети должно быть пять вход-

ных сигналов x_1-x_5 , и построение ИНС сводится к решению следующих задач:

- определение типа необходимой ИНС;
- определение подхода к присвоению числовых значений входным сигналам ИНС, отражающим поведение пользователя или группы пользователей (x_1-x_5);
- определение необходимого количества слоев ИНС и количества нейронов в слоях ИНС;
- выбор метода обучения ИНС;
- выбор активационных функций;
- выбор области значений выходного сигнала NET, сигнализирующего о наличии аномалий в поведении пользователя или группы пользователей.

Построение ИНС

1. Тип искусственной нейронной сети.

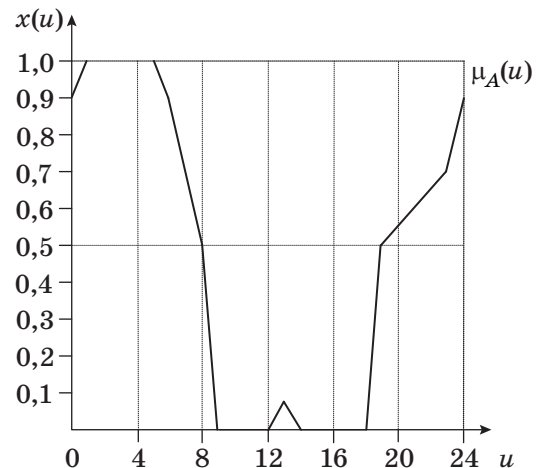
Поскольку на настоящий момент не существует строгой теории по выбору ИНС [27], за основу разрабатываемой ИНС предлагается взять хорошо изученный многослойный полносвязный перцептрон без обратных связей.

2. Вход.

Для присвоения числовых значений входным сигналам ИНС предлагается использовать математический аппарат теории нечетких множеств, позволяющий присваивать вербальным характеристикам (более свойственно, менее свойственно и т. п.) числовые значения. Основной трудностью, мешающей применению теории нечетких множеств при решении практических задач, является то, что функция принадлежности должна быть задана вне самой теории и, следовательно, ее адекватность не может быть проверена непосредственно средствами теории. В каждом известном методе построения функции принадлежности формулируются свои требования и обоснования к выбору именно такого построения [33]. В рамках решаемой задачи предлагается рассматривать характеристики поведения пользователя x_n как характеристические функции принадлежности $\mu_A(u)$ множеству значений аномального поведения A , заданные на универсальном множестве U и принимающие значения, равные единице, на тех элементах множества U , которые принадлежат множеству A , и значения, равные нулю, на тех элементах, которые не принадлежат множеству A :

$$\mu_A(u) = \begin{cases} 1, & \text{если } u \in A \\ 0, & \text{если } u \notin A \end{cases}$$

При этом для каждой функции принадлежности должны рассматриваться свои множества. В качестве примера на рис. 3 представлена диа-



■ Рис. 3. Диаграмма Заде

■ Fig. 3. Zade diagram

грамма Заде, демонстрирующая возможную зависимость значения характеристической функции принадлежности $\mu_A(u)$ множеству значений аномального поведения A в зависимости от времени доступа пользователя к ресурсам сети, на которой:

U — множество значений времени суток, в которое может быть осуществлен доступ к ресурсам сети, $U = \{u, u \in R : 0 \leq u \leq 24\}$;

A — множество значений времени суток, доступ в которое аномален для конкретного пользователя;

$\mu_A(u)$ — характеристическая функция принадлежности множеству значений времени суток, доступ в которое аномален для конкретного пользователя.

В предложенном примере предполагается следующее:

— продолжительность рабочего дня пользователя составляет девять часов (с девяти утра до шести вечера), включая перерыв на обед (в районе часа дня);

— пользователю не свойственно приходить на работу раньше начала рабочего дня, задерживаться по окончании рабочего дня и работать в околообеденное время.

Несущее множество аномального поведения можно записать следующим образом:

$$A = 0,9/0 + 1,0/1 + 1,0/5 + 0,9/6 + 0,5/8 + 0/9 + 0/12 + 0,1/13 + 0/14 + 0/18 + 0,5/19 + 0,7/23 + 0,9/24,$$

где запись, например, вида $0,9/0$ выражает не деление на ноль, а значение функции принадлежности (в данном примере $0,9$) при осуществлении доступа в полночь.

Общая форма записи нечетких подмножеств будет иметь следующий вид:

$$\begin{aligned}
 A = & \sum_{u=0}^{24} \mu_A(u) / u = \sum_{u=0}^0 0,9 / u + \sum_{u=1}^1 1,0 / u + \\
 & + \sum_{u=5}^5 1,0 / u + \sum_{u=6}^6 0,9 / u + \sum_{u=8}^8 0,5 / u + \sum_{u=9}^9 0 / u + \\
 & + \sum_{u=12}^{12} 0 / u + \sum_{u=13}^{13} 0,1 / u + \sum_{u=14}^{14} 0 / u + \sum_{u=18}^{18} 0 / u + \\
 & + \sum_{u=19}^{19} 0,5 / u + \sum_{u=23}^{23} 0,7 / u + \sum_{u=24}^{24} 0,9 / u,
 \end{aligned}$$

где запись вида $\sum_{u=0}^{24} \mu_A(u) / u$ не предполагает сум-

му, но предполагает объединение по всем элементам конечного несущего множества значений u . Точками перехода ($x(u) = 0,5$) для функции принадлежности являются $u = 8$ и $u = 19$.

3. Количество слоев и нейронов в слоях.

Как показал А. Н. Колмогоров [34], любую непрерывную функцию n переменных на единичном отрезке $[0; 1]$ можно представить в виде суммы конечного числа одномерных функций:

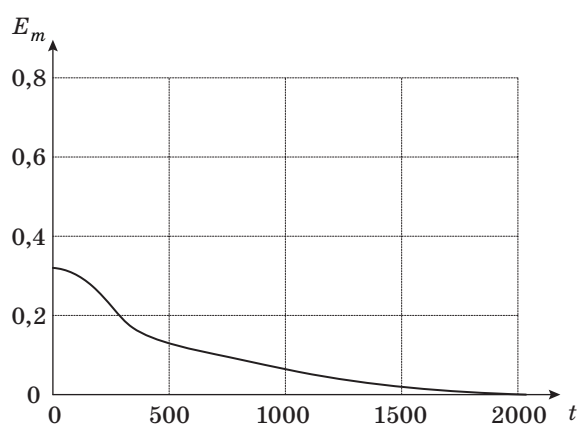
$$f(x_1, x_2, \dots, x_n) = \sum_{p=1}^{2n+1} g \left(\sum_{i=1}^n \lambda_i \varphi_p(x_i) \right),$$

где функции g и φ_p являются одномерными и непрерывными; $\lambda_i = \text{const}$ для всех i . Из этого следует, что любую непрерывную функцию $f: [0; 1]^n \rightarrow [0; 1]$ можно аппроксимировать при помощи трехслойной нейронной сети, имеющей n входных, $2n + 1$ скрытых и один выходной нейрон. Данный результат обобщен на многослойную сеть с алгоритмом обратного распространения ошибки [35–37]. Таким образом, при построении ИНС предлагается использовать три слоя: первый слой будет включать в себя три нейрона, второй слой — семь нейронов, последний слой (выходной) — один нейрон.

4. Метод обучения.

В качестве метода обучения предлагается использовать алгоритм обратного распространения ошибки [38]. Данный алгоритм позволяет минимизировать среднеквадратичную ошибку ИНС. На рис. 4 продемонстрирована зависимость среднеквадратичной ошибки E_m от номера итерации t на примере прогнозирования функции $y = 0,1 \sin(3x) + 0,5$ с 20 точками обучающей выборки и восемью предсказанными точками [39].

Согласно методу градиентного спуска изменение весовых коэффициентов и порогов нейронной сети происходит по следующему правилу:



■ **Рис. 4.** График убывания ошибки на обучающей выборке
 ■ **Fig. 4.** Graph of falling of error on the learning sample

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \alpha \frac{\partial E}{\partial \omega_{ij}(t)}; \quad (1)$$

$$T_j(t+1) = T_j(t) - \alpha \frac{\partial E}{\partial T_j(t)}; \quad (2)$$

где $E = \frac{1}{2} \sum_j (y_j - t_j)^2$ — среднеквадратичная ошибка ИНС для одного образа.

В работе [40] показано, что

$$\frac{\partial E}{\partial \omega_{ki}} = \gamma_i F'(S_i) \gamma_k; \quad (3)$$

$$\frac{\partial E}{\partial T_i} = -\gamma_i F'(S_i), \quad (4)$$

где γ_i — выходное значение i -го нейрона.

В результате подстановки (3) и (4) в (1) и (2) получаются выражения, устанавливающие порядок изменения весовых коэффициентов и порогов нейронов, которого необходимо придерживаться для минимизации среднеквадратичной ошибки ИНС:

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \alpha \gamma_j F'(S_j) y_i; \quad (5)$$

$$T_j(t+1) = T_j(t) + \alpha \gamma_j F'(S_j). \quad (6)$$

Выражения (5) и (6), называемые обобщенным дельта-правилом, определяют правило обучения многослойных ИНС в общем виде.

5. Активационные функции.

Для обеспечения сходимости алгоритма обратного распространения ошибки в качестве активационной функции предполагается использовать гиперболический тангенс. Таким образом, ИНС будет являться гомогенной, а выходное значе-

ние j -го нейрона определяться следующим образом:

$$y = \text{tg}(S_j) = \frac{e^{S_j} - e^{-S_j}}{e^{S_j} + e^{-S_j}},$$

где S_j — взвешенная сумма j -го нейрона. Поскольку производная этой функции имеет вид $F'(S_j) = 1 - y_j^2$, правило обучения можно представить в виде

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \alpha \gamma_j (1 - y_j^2) y_i;$$

$$T_j(t+1) = T_j(t) + \alpha \gamma_j (1 - y_j^2),$$

где t — номер итерации; α — значение шага обучения; γ_i — значение ошибки для i -го нейрона; T_j — значение порога j -го нейрона.

Ошибка для j -го нейрона выходного и скрытого слоев определяется следующим образом:

$$\gamma_j = y_j - t_j;$$

$$\gamma_j = \sum_i \gamma_i (1 - y_i^2) \omega_{ij}.$$

Для выполнения алгоритма обратного распространения ошибки необходимо выполнить следующие действия.

1. Задать значения шага обучения α ($0 < \alpha < 1$) и желаемой среднеквадратичной ошибки E_m .

2. Придать случайные числовые значения весовым коэффициентам и пороговым значениям ИНС. В соответствии с рекомендацией, представленной в работе [41], весовым коэффициентам ω_{ij} предлагается придавать значения, примерно равные $\frac{1}{\sqrt{n(i)}}$, где $n(i)$ — число элементов в слое i .

3. Последовательно подать образы из обучающей выборки на вход ИНС. При этом для каждого входного образа необходимо выполнить следующие действия:

а) произвести фазу прямого распространения образа по ИНС, при этом вычисляется выходная активность всех нейронов ИНС

$$y_j = F\left(\sum_i \omega_{ij} y_i - T_j\right);$$

б) осуществить фазу обратного распространения сигнала, в результате которой определится ошибка нейронов γ_j для всех слоев ИНС. При этом, соответственно, для выходного и скрытого слоев

$$\gamma_j = y_j - t_j;$$

$$\gamma_j = \sum_i \gamma_i F'(S_j) \omega_{ji}, \quad (7)$$

где i характеризует нейронные элементы следующего слоя по отношению к слою j ;

в) изменить весовые коэффициенты и пороги нейронных элементов для каждого слоя ИНС в соответствии с (5) и (6):

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \alpha \gamma_j F'(S_j) y_i;$$

$$T(t+1) = T_j(t) + \alpha \gamma_j F'(S_j).$$

4. Вычислить суммарную среднеквадратичную ошибку E :

$$E = \frac{1}{2} \sum_{k=1}^L \sum_j (y_j^k - t_j^k)^2,$$

где L — размерность обучающей выборки.

5. Если $E > E_m$, происходит переход к шагу 3, в противном случае алгоритм заканчивается. Таким образом, алгоритм функционирует до тех пор, пока суммарная среднеквадратичная ошибка ИНС не станет меньше заданной.

Для нейтрализации застраивания метода градиентного спуска в нежелательных минимумах предлагается применять метод тяжелого шарика [41]. В этом случае модификация синаптических связей ИНС будет осуществляться в соответствии с выражением

$$\Delta \omega_{ij}(t+1) = -\alpha \gamma_j F'(S_j) y_i + \xi \Delta \omega_{ij}(t),$$

где ξ — моментный параметр, выбираемый из диапазона $[0; 1]$. В соответствии с рекомендациями, представленными в работе [41], предлагается использовать значение $\xi = 0,9$.

6. Выход.

Для удобства интерпретации результата работы ИНС уместным является использование диапазона выходных значений от нуля (для обозначения отсутствия аномалии) до единицы (для обозначения наличия аномалии).

Однако с учетом выбора, сделанного в пользу использования в качестве метода обучения алгоритма обратного распространения ошибки (для гарантированной минимизации среднеквадратичной ошибки) и гиперболических тангенсов в качестве активационных функций (для обеспечения сходимости алгоритма обратного распространения ошибки), выходной сигнал NET будет принимать значения из диапазона от минус единицы до единицы. Таким образом, наличие аномалии в поведении пользователя предлагается интерпретировать выходом NET, равным единице, а отсутствие аномалии — выходом NET, равным минус единице.

Результаты определения типа и основных характеристик ИНС

В качестве итоговой конфигурации был выбран трехслойный полносвязный гомогенный перцептрон без обратных связей с пятью входными сигналами, одиннадцатью нейронами и гиперболическими тангенсами в качестве функций активации. В первом слое содержится три нейрона, в скрытом слое — семь и в выходном — один. Итоговая конфигурация подготовленной ИНС представлена на рис. 5.

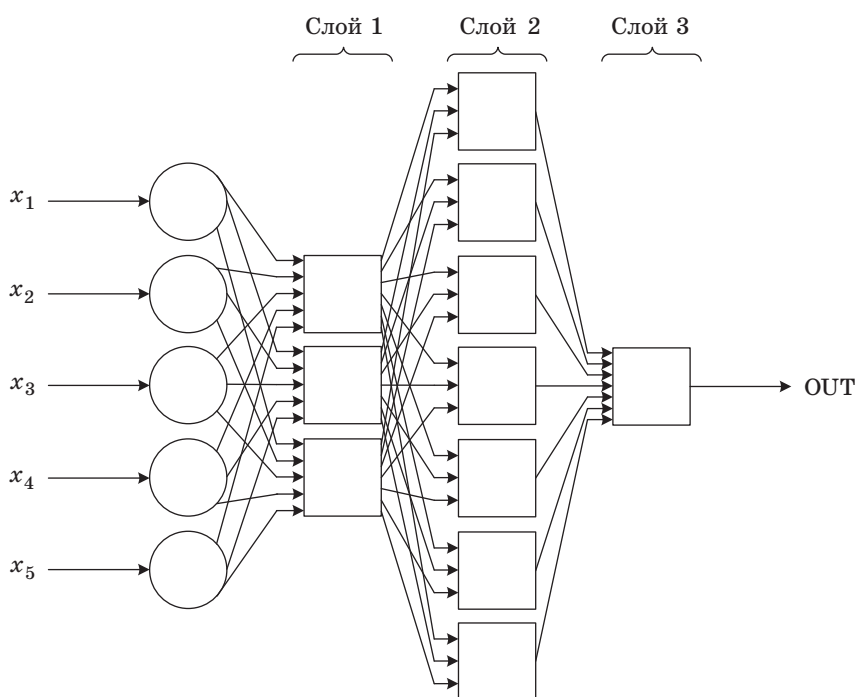
При возрастании числа нейронов в скрытых слоях, с одной стороны, растет точность ИНС, но с другой — при слишком большой размерности скрытых слоев возникает явление перетренировки сети, ухудшающее обобщающие способности ИНС. Таким образом, число тренировочных образцов должно быть больше числа нейронов в скрытом слое. В качестве обучающей выборки были взяты 30 комбинаций вектора x , выражающих наборы поведенческих характеристик пользователя и формирующих образы, подаваемые на вход ИНС. В связи с ограниченной областью значений гиперболического тангенса обучающая выборка была предварительно масштабирована к соответствующему диапазону значений. ИНС обучалась на обучающей выборке до достижения заданной среднеквадратичной ошибки. По окончании обучения для определения точности работы обученной сети использовалась тестовая

выборка, состоящая из 10 комбинаций входных и выходных значений. При использовании тестовой выборки на входы сети подавались значения входов из тестовой выборки, затем значения полученных выходов сравнивались со значениями выходов тестовой выборки. В случае отличия значений полученных выходов от значений выходов тестовой выборки сеть проходила дообучение. Для итогового тестирования использовалась тестовая выборка, состоящая из пяти комбинаций.

Заключение

Рассмотрен актуальный подход к защите информации, основанный на применении технологии машинного обучения (теории искусственных нейронных сетей), отличающейся от известных уникальным составом характеристик поведения пользователя (составом входных характеристик ИНС) и подходом к выбору ИНС. Подготовлена выборка характеристик санкционированных действий пользователей. Данная выборка использовалась для обучения трехслойного перцептрона. Обучение проводилось в программном обеспечении SPSS Statistics (разработка компании IBM), предназначенном для статистической обработки данных.

После обучения нейронной сети была проведена проверка эффективности ее работы с помощью контрольной выборки. Относительная погреш-



■ Рис. 5. Итоговая конфигурация ИНС
 ■ Fig. 5. Total configuration of neural network

ность классификации данных составила примерно 10 %, что является достаточно хорошим результатом. При обучении ИНС использовались небольшие выборки, предназначенные для де-

монстрации общих принципов работы ИНС. Для применения представленного подхода в решении реальных прикладных задач защиты информации выборки должны быть как можно больше.

Литература

1. Указ Президента РФ от 06.03.1997 № 188 «Об утверждении перечня сведений конфиденциального характера». <http://www.kremlin.ru/acts/bank/10638> (дата обращения: 20.05.2018).
2. Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных». <https://rg.ru/2006/07/29/personalnye-dannye-dok.html> (дата обращения: 20.05.2018).
3. Основы теории нейронных сетей. <http://www.intuit.ru/studies/courses/88/88/info> (дата обращения: 06.12.2017).
4. **Круглов В. В., Дли М. И., Голунов Р. Ю.** Нечеткая логика и искусственные нейронные сети. — М.: Физматлит, 2001. — 224 с.
5. **Воронцов К. В.** Лекции по искусственным нейронным сетям. <http://www.machinelearning.ru/wiki/images/c/cc/Voron-ML-NeuralNets.pdf> (дата обращения: 02.04.2018).
6. **Лю Б.** Теория и практика неопределенного программирования. — М.: БИНОМ, 2005. — 416 с.
7. **Барский А. Б.** Нейронные сети: распознавание, управление, принятие решений. — М.: Финансы и статистика, 2004. — 176 с.
8. **Беркинблит М. Б.** Нейронные сети. — М.: МИРОС, 1993. — 96 с.
9. **Callan R.** The Essence of Neural Networks. — London: Prentice Hall Europe, 1999. — 248 p.
10. **Лукашик Е. П., Кочетов Д. А.** Применение нейронных сетей для обнаружения сетевых атак // Традиционная и инновационная наука: история, современное состояние, перспективы: сб. ст. по итогам Междунар. науч.-практ. конф./под ред. А. А. Сукиясяна. — Стерлитамак: Агентство международных исследований, 2017. С. 24–27.
11. **Зубков Е. В., Белов В. М.** Методы интеллектуального анализа данных и обнаружения вторжений // Вестник СибГУТИ. 2016. № 1. С. 118–133.
12. **Частикова В. А., Картамышев Д. А.** К вопросу защиты информации от сетевых атак на основе нейронных сетей // Науч. тр. Кубанского государственного технологического университета. 2014. № 6. С. 101–104.
13. **Кондратьев А. А., Талалаев А. А., Тищенко И. П., Фраленко В. П., Хачумов В. М.** Методологическое обеспечение интеллектуальных систем защиты от сетевых атак // Современные проблемы науки и образования. 2014. № 2. С. 119.
14. **Васильев В. И., Шарабыров И. В.** Обнаружение атак в локальных беспроводных сетях на основе интеллектуального анализа данных // Изв. ЮФУ. Технические науки. 2014. № 2(151). С. 57–67.
15. **Zhi-Peng Pan, Chao Feng, Chao-Jing Tang.** Malware Classification based on the Behavior Analysis and Back Propagation Neural Network. ITM Web of Conferences 7, 02001. 2016. P. 1–5. https://www.itm-conferences.org/articles/itmconf/pdf/2016/02/itmconf_ita2016_02001.pdf (дата обращения: 02.03.2018). doi:10.1051/itmconf/20160702001
16. **Ларионова А. В., Хорев П. Б.** Метод фильтрации спама на основе искусственной нейронной сети // Наукоедение. 2016. Т. 8. № 3(34). <https://naukovedenie.ru/PDF/04TVN316.pdf> (дата обращения: 03.04.2018).
17. **Ларионова А. В., Хорев П. Б.** Оценка эффективности фильтрации спама на основе искусственной нейронной сети // Наукоедение. 2016. Т. 8. № 2(33). <https://naukovedenie.ru/PDF/134TVN216.pdf> (дата обращения: 03.04.2018).
18. **Трапезников Е. В., Данилова О. Т.** Модель анализа защиты информации на основе нейронной сети // Динамика систем, механизмов и машин. 2016. Т. 2. № 1. С. 302–308.
19. **Данилова О. Т., Трапезников Е. В.** Разработка модели, анализирующей функцию безопасности в системе информационной защиты, на основе нейронной сети // Информационное противодействие угрозам терроризма. 2015. № 24. С. 24–29.
20. **Гильмуллин Т. М., Гильмуллин М. Ф.** Подходы к автоматизации процесса валидации уязвимостей, найденных автоматическими сканерами безопасности, при помощи нечетких множеств и нейронных сетей // Фундаментальные исследования. 2014. № 11. Ч. 2. С. 266–279.
21. **Цырульник В. Ф., Кадочникова Н. А.** Оценка актуальности угроз информационной безопасности с помощью программной реализации обученной нейронной сети // Научный альманах. 2016. № 4–3(18). С. 211–215.
22. **Соловьев С. В., Мамута В. В.** Применение аппарата нейросетевых технологий для определения актуальных угроз безопасности информации информационных систем // Наукоемкие технологии в космических исследованиях Земли. 2016. Т. 8. № 5. С. 78–82.
23. **Мухин В. Е., Корнага Я. И., Шешин В. В.** Адаптивные средства защиты компьютерных систем на основе модифицированных нейронных сетей Кохонена // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. 2014. № 2(193). С. 31–38.
24. **Штеренберг С. И., Виткова Л. А., Просихин В. П.** Методика применения концепции адаптивной са-

- моразвивающейся системы // Информационные технологии и телекоммуникации. 2014. № 4(8). С. 126–133.
25. Котенко И. В., Нестерук Ф. Г., Шоров А. В. Гибридная адаптивная система защиты информации на основе биометафор «нервных» и нейронных сетей // Инновации в науке. 2013. № 16-1. С. 79–83.
26. Вишняков В. А., Коваль О. И., Моздурани Шираз М. Г. Использование нейронных сетей для обнаружения и распознавания аномалий в корпоративной информационной системе предприятия // Докл. Белорусского государственного университета информатики и радиоэлектроники. 2016. № 4(98). С. 86–92.
27. Улезло Д. С., Кадан А. М. Методы машинного обучения в решении задач информационной безопасности // Intelligent Technologies for Information Processing and Management (ITIPM'2015): Proc. of the 3rd Intern. Conf. 2015. С. 41–44.
28. Цветкова О. Л., Крепер А. И. О применении теории искусственных нейронных сетей в решении задач обеспечения информационной безопасности // Символ науки. 2017. № 04-2. С. 105–107.
29. Галушкин А. И. Нейрокомпьютеры и их применение. — М.: ИПРЖР, 2000. — 416 с.
30. Щавелев Л. В. Способы аналитической обработки данных для поддержки принятия решений. <http://infovisor.ivanovo.ru/press/paper04.html> (дата обращения: 03.04.2018).
31. Хайкин С. Нейронные сети: полный курс. — М.: Вильямс, 2006. — 1104 с.
32. Shelestov A., Skakun S., Kissul O. Complex Neural Network Model of User Behavior in Distributed System// Knowledge-Dialogue-Solutions: Intern. Conf. 2007. P. 1–8. http://inform.ikd.kiev.ua/content/ua/publications/articles/content/KDS07-Shelestov_Skakun_Kussul.pdf (дата обращения: 02.03.2018).
33. Ronald R. Yager. Fuzzy Set and Possibility Theory: Recent Developments. — N. Y.: Pergamon, 1982. — 408 p.
34. Колмогоров А. Н. О представлении непрерывных функций нескольких переменных в виде суперпозиций непрерывных функций одного переменного и сложения // Докл. АН СССР. 1957. Т. 114. С. 953–956.
35. Hornik K., Stinchcombe M., White H. Multilayer Feedforward Networks are Universal Approximators// Neural Network. 1989. N 2(5). P. 359–366.
36. Rojas R. Theorie der Neuronalen Netze: Eine Systematische Einführung. — Berlin: Springer-Verlag, 2013. — 446 p.
37. Maxwell T., Giles C. L., Lee Y. C., Chen H. H. Nonlinear Dynamics of Artificial Neural Systems // AIP Conf. Proc. 1986. Vol. 151. P. 299–304. doi:10.1063/1.36227
38. Rumelhart D., Hinton G., Williams R. Learning Representations by Back-Propagating Errors// Nature. 1986. Vol. 323. P. 533–536. doi:10.1038/323533a0
39. Рудой Г. И. Выбор функции активации при прогнозировании нейронными сетями // Машинное обучение и анализ данных. 2001. Т. 1. № 1. С. 16–39.
40. Головкин В. А. Нейронные сети. Обучение, организация и применение. — М.: ИПРЖР, 2001. — 256 с.
41. Hertz J., Krogh A., Palmer R. Introduction to the Theory of Neural Computation. — Redwood City: Addison Wesley, 1991. — 327 p.

UDC 004.056

doi:10.15217/issn1684-8853.2018.3.69

Providing Personal Data Protection in an Information System based on User Behavior AnalyticsKozin I. S.^a, Lead Specialist, van@trioptimum.com^aJSC «Kronstadt Technologies», 54, Bldg 5, Litas P, Malyy Pr. V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: As information technologies develop, new classes of personal data protection software come out. One of them is User Behavior Analytics. In the development of such systems, machine learning methods are widely used, including the applications of artificial neural network theory. However, the approaches to protection software development based on machine learning have not been studied well enough. **Purpose:** Developing a method of creating an artificial neural network which would provide the analysis of authorized behavior of information system users and the detection of abnormalities in their behavior signaling about criminal activity. **Results:** A review of the approaches to provide information security with artificial neural networks demonstrates their active development in several directions, including the detection of abnormalities. A new method to create an artificial neural network has been developed, including the proposals about determining the network type, the range of numeric values for the input and output signals, the number of the layers and neurons in a layer, the learning method, and the type of the activation functions. As the input values, user behavior characteristics can be used, namely: a set of the user's data, the point of access to the information system, the set of user's actions, the time of the access or time of certain actions, the general duration of runtime operations. With user's access time as an example, an approach has been proposed to assign numeric values to the characteristic of a user, based on fuzzy set theory application. **Practical relevance:** A trained neural network provides a more efficient detection of abnormalities in user behavior than an information security specialist without special automation tools.

Keywords — Information Security, Personal Data, Behavior Analytics, Artificial Neural Networks, Fuzzy Set Theory.

Citation: Kozin I. S. Providing Personal Data Protection in an Information System based on User Behavior Analytics. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 69–78 (In Russian). doi:10.15217/issn1684-8853.2018.3.69

References

1. Ukaz Prezidenta RF ot 06.03.1997 N 188 "Ob utverzhdenii perechnia svedenii konfidentsial'nogo kharaktera". Available at: <http://www.kremlin.ru/acts/bank/10638> (accessed 20 May 2018).
2. Federal'nyi zakon ot 27.07.2006 N 152-FZ "O personal'nykh dannykh". Available at: <https://rg.ru/2006/07/29/personal-jnye-dannye-dok.html> (accessed 20 May 2018).
3. *Osnovy teorii neironnykh setei* [The Fundamentals of the Theory of Artificial Neural Networks]. Available at: <http://www.intuit.ru/studies/courses/88/88/info> (accessed 6 December 2017).
4. Kruglov V. V., Dli M. I., Golunov R. Iu. *Nechetkaia logika i iskusstvennyye neironnye seti* [Fuzzy Logic and Artificial Neural Networks]. Moscow, Fizmatlit Publ., 2001. 224 p. (In Russian).
5. Vorontsov K. V. *Lektsii po iskusstvennym neironnym setiam* [Lectures on Artificial Neural Networks]. Available at: <http://www.machinelearning.ru/wiki/images/c/cc/Voron-ML-NeuralNets.pdf> (accessed 2 April 2018).
6. Liu B. *Teoriia i praktika neopredelennogo programmirovaniia* [Theory and Practice of Indefinite Programming]. Moscow, BINOM Publ., 2005. 416 p. (In Russian).
7. Barskii A. B. *Neironnye seti: raspoznavanie, upravlenie, priniatie reshenii* [Neural Networks: Recognition, Management, Decision Making]. Moscow, Finansy i statistika Publ., 2004. 176 p. (In Russian).
8. Berkinblit M. B. *Neironnye seti* [Neural Networks]. Moscow, MIOSCO Publ., 1993. 96 p. (In Russian).
9. Callan R. *The Essence of Neural Networks*. London, Prentice Hall Europe, 1999. 248 p.
10. Lukashchik E. P., Kochetov D. A. The use of Neural Networks to Detect Network Attacks. *Scientific Collection "Traditsionnaia i innovatsionnaia nauka: istoriia, sovremennoe sostoiianie, perspektivy"* [Traditional and Innovative Science: History, Current State, Prospects], eds. A. A. Sukiasian, Sterlitamak, Agentstvo mezhdunarodnykh issledovaniia Publ., 2017, pp. 24–27 (In Russian).
11. Zubkov E. V., Belov V. M. Methods of Data Mining and Intrusion Detection. *Vestnik SibGUTI*, 2016, no. 1, pp. 118–133 (In Russian).
12. Chastikova V. A., Kartamyshev D. A. To the Issue of Protection of Information from Network Attacks based on Neural Networks. *Nauchnye trudy Kubanskogo gosudarstvennogo tekhnologicheskogo universiteta* [Scientific Collection of Kuban State Technical University], 2014, no. 6, pp. 101–104 (In Russian).
13. Kondrat'ev A. A., Talalaev A. A., Tishchenko I. P., Fralenko V. P., Khachumov V. M. Methodological Provision of Intelligent Systems to Protect against Network Attacks. *Sovremennye problemy nauki i obrazovaniia*, 2014, no. 2, p. 119 (In Russian).
14. Vasil'ev V. I., Sharabyrov I. V. Detection of Attacks in Local Wireless Networks based on Data Mining. *Izvestiia IuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2014, no. 2(151), pp. 57–67 (In Russian).
15. Zhi-Peng Pan, Chao Feng, Chao-Jing Tang. Malware Classification based on the Behavior Analysis and Back Propagation Neural Network. *ITM Web of Conferences 7, 02001*, 2016, pp. 1–5. Available at: https://www.itm-conferences.org/articles/itmconf/pdf/2016/02/itmconf_ita2016_02001.pdf (accessed 2 March 2018).
16. Larionova A. V., Khorev P. B. The Method of Spam Filtering based on an Artificial Neural Network. *Naukovedenie*, 2016, vol. 8, no. 3(34). Available at: <https://naukovedenie.ru/PDF/04TVN316.pdf> (accessed 3 April 2018) (In Russian).
17. Larionova A. V., Khorev P. B. Evaluating the Effectiveness of Spam Filtering based on an Artificial Neural Network. *Naukovedenie*, 2016, vol. 8, no. 2(33). Available at: <https://naukovedenie.ru/PDF/134TVN216.pdf> (accessed 3 April 2018) (In Russian).
18. Trapeznikov E. V., Danilova O. T. The Model of the Analysis of Protection of the Information on the basis of a Neural Network. *Dinamika sistem, mekhanizmov i mashin*, 2016, vol. 2, no. 1, pp. 302–308 (In Russian).
19. Danilova O. T., Trapeznikov E. V. Development of a Model Analyzing the Security Function in the Information Protection System, on the Neural Network. *Informatsionnoe protivodeistvie ugrozam terrorizma*, 2015, no. 24, pp. 24–29 (In Russian).
20. Gil'mullin T. M., Gil'mullin M. F. Approaches to Automating the Validation Process of Vulnerabilities Found by Automatic Security Scanners, using Fuzzy Sets and Neural Networks. *Fundamental'nye issledovaniia*, 2014, no. 11, part 2, pp. 266–279 (In Russian).
21. Tsyurul'nik V. F., Kadochnikova N. A. Assessment of the Urgency of Information Security Threats through the Implementation of a Trained Neural Network. *Nauchnyi al'manakh*, 2016, no. 4-3(18), pp. 211–215 (In Russian).
22. Solov'ev S. V., Mamuta V. V. Application of the Apparatus of Neural Technologies for Determining the Urgency of Information Security Threats to Information Systems. *Naukoemkie tekhnologii v kosmicheskikh issledovaniakh Zemli*, 2016, vol. 8, no. 5, pp. 78–82 (In Russian).
23. Mukhin V. E., Kornaga Ia. I., Steshin V. V. Adaptive Means of Protection of Computer Systems based on Modified Neural Networks of Kohonen. *Nauchno-tekhnicheskie vedomosti Sankt-Peterburgskogo gosudarstvennogo politekhnicheskogo universiteta. Informatika. Telekomunikatsii. Upravlenie* [St. Petersburg Polytechnic University Journal of Engineering Science and Technology], 2014, no. 2(193), pp. 31–38 (In Russian).
24. Shterenberg S. I., Vitkova L. A., Prosikhin V. P. The Method of Application of the Concept of Adaptive Self-Developing System. *Informatsionnye tekhnologii i telekommunikatsii*, 2014, no. 4(8), pp. 126–133 (In Russian).
25. Kotenko I. V., Nesteruk F. G., Shorov A. V. Hybrid Adaptive Information Protection System based on the Biometaphors of "Nervous" and Neural Networks. *Innovatsii v nauke*, 2013, no. 16-1, pp. 79–83 (In Russian).
26. Vishniakov V. A., Koval' O. I., Mozdurani Shiraz M. G. The Use of Neural Networks to Detect and Recognize Anomalies in the Corporate Enterprise Information System. *Doklady Belorusskogo gosudarstvennogo universiteta informatiki i radioelektroniki*, 2016, no. 4(98), pp. 86–92 (In Russian).
27. Ulezlo D. S., Kadan A. M. Methods of Machine Learning in Solving Information Security Problems. *Proc. of the 3rd Intern. Conf. "Intelligent Technologies for Information Processing and Management"* (ITIPM'2015), 2015, pp. 41–44 (In Russian).
28. Tsvetkova O. L., Kreper A. I. On the Application of the Theory of Artificial Neural Networks in Solving Problems of Providing Information Security. *Simvol nauki*, 2017, no. 04-2, pp. 105–107 (In Russian).
29. Galushkin A. I. *Neirokomp'yutery i ikh primenenie* [Neurocomputers and their Application]. Moscow, IPRZhR Publ., 2000. 416 p. (In Russian).
30. Shchavalev L. V. *Sposoby analiticheskoi obrabotki dannykh dlia podderzhki priniatiia reshenii* [Methods of Analytical Data Processing to Support Decision Making]. Available at: <http://infovisor.ivanovo.ru/press/paper04.html> (accessed: 3 April 2018).
31. Khaikin S. *Neironnye seti: polnyi kurs* [Neural Networks: Full Course]. Moscow, Vil'iams Publ., 2006. 1104 p. (In Russian).
32. Andrii Shelestov, Serhiy Skakun, Olga Kissul. Complex Neural Network Model of User Behavior in Distributed System. *Intern. Conf. "Knowlegge-Dialogue-Solutions"*, 2007, pp. 1–8. Available at: http://inform.ikd.kiev.ua/content/ua/publications/articles/content/KDS07-Shelestov_Skakun_Kussul.pdf (accessed 2 March 2018).
33. Ronald R. Yager. *Fuzzy Set and Possibility Theory: Recent Developments*. Pergamon, New York, 1982. 408 p.
34. Kolmogorov A. N. On the Representation of Neural Functions of Several Variables in the Form of Superpositions of Continuous Functions of One Variable and Addition. *Doklady AN SSSR* [Reports of the USSR Academy of Sciences], 1957, vol. 114, pp. 953–956 (In Russian).
35. Hornik K., Stinchcombe M., White H. Multilayer Feedforward Networks are Universal Approximators. *Neural Networks*, 1989, no. 2(5), pp. 359–366.
36. Rojas R. *Theorie der Neuronalen Netze: Eine Systematische Einfuehrung*. Springer-Verlag, Berlin, 2013. 446 p. (In German).
37. Maxwell T., Giles C. L., Lee Y. C., Chen H. H. Nonlinear Dynamics of Artificial Neural Systems. *AIP Conf. Proc.*, 1986, vol. 151, pp. 299–304. doi:10.1063/1.36227
38. Rumelhart D., Hinton G., Williams R. Learning Representations by Back-Propagating Errors. *Nature*, 1986, vol. 323, pp. 533–536. doi:10.1038/323533a0
39. Rudoi G. I. Selecting the Activation Function when Predicting Neural Networks. *Mashinnoe obuchenie i analiz dannykh*, 2001, vol. 1, no. 1, pp. 16–39 (In Russian).
40. Golovko V. A. *Neironnye seti. Obuchenie, organizatsiia i primenenie* [Neural Networks. Training, Organization and Application]. Moscow, IPRZhR Publ., 2001. 256 p. (In Russian).
41. Hertz J., Krogh A., Palmer R. *Introduction to the Theory of Neural Computation*. Addison Wesley, Redwood City, 1991. 327 p.

UDC 621.391

doi:10.15217/issn1684-8853.2018.3.79

RANDOM CODING BOUND FOR CHANNELS WITH MEMORY — DECODING FUNCTION WITH PARTIAL OVERLAPPING

Part 1. Derivation of Main Expression

A. N. Trofimov^A, PhD, Tech., Associate Professor, andrei.trofimov@vu.spb.ru

^ASaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: The problem of calculating the random coding exponent in the full range of code rates for finite-state channels is not completely solved and remains relevant. It seems that a good approximation to the optimal random coding exponent can be found by using a mismatched decoding function. **Purpose:** Deriving a random coding exponent close to the optimal one. **Results:** A new random coding bound is presented for a wide class of channels, including those for which the complete random coding exponent was not previously derived. The derivation of this bound is based on the use of a mismatched decoding function which depends on two parameters: the length W of the segment of the channel output sequence and the length B of the segment of the channel input sequence. The values of W and B greatly influence the values of the random coding exponent and the complexity of its calculation.

Keywords — Random Coding Bound, Finite-State Channel, Mismatched Decoding, Perron — Frobenius Theorem.

Citation: Trofimov A. N. Random Coding Bound for Channels with Memory — Decoding Function with Partial Overlapping. Part 1. Derivation of Main Expression. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 79–88. doi:10.15217/issn1684-8853.2018.3.79

Introduction

According to the random coding theorem [1] there exists a block code of length N and rate R , $R = N^{-1} \log_2 M$, M is the number of code words, such that the error probability of the maximum likelihood (ML) decoding is upper bounded as $P_e < \exp_2(-N(E_r(R) - o(N)))$, where $\exp_2 = 2^x$, and $E_r(R)$ the exponent of the random coding bound. The function $E_r(R) > 0$, if $R < C$, where C is the channel capacity. Another quantity characterizing the channel is the computational cut-off rate R_0 . The value R_0 can be found from the function $E_r(R)$ as $R_0 = E_r(0)$. The values of $E_r(R)$ and R_0 are calculated by averaging over a code ensemble and optimizing over a distribution on this ensemble. Generally, the exact computation of the function $E_r(R)$ is a difficult problem which can only be solved for some channel models. However, sometimes it is possible to obtain a bound for the function $E_r(R)$ for the decoding algorithm used a decoding function ψ which differs from the ML decoding function (mismatched decoding; see, for example, [2–6] and references therein). In this case it is possible to obtain near-exact characteristics of the transmission reliability. This approach is used in this paper.

For simplicity, we assume that the distribution on the code ensemble corresponds to the *independent uniformly distributed* (i.u.d.) code symbols. In this case we have the function $E_r^*(R; \psi) \leq E_r(R)$ and quantities $R_0^*(\psi) \leq R_0$, $C^*(\psi) \leq C$ (the asterisk in the

superscript hereafter means that the code symbols are chosen as i. u. d. random variables).

We assume that the channel model is given by the well known finite-state model [1]. For this model we present the derivation of the random coding exponent for the decoding function which is in the form of the product of the *a posteriori* probabilities (APP) of segments of the channel input sequence relative to the overlapped segments of the channel output sequence. One of the main challenges of this study was the choice of a suitable *decoding function* enabling a good final result. With the usage of such a decoding function the problem can be reduced to the evaluation of the logarithm of a bilinear form defined by a power of a nonnegative matrix. Then we used a known technique based on the usage of the Perron — Frobenius theorem to obtain the final result. Using this approach a new suboptimal random coding bound which is applicable to a wide class of the channel with memory has been obtained. The discrete-time model with intersymbol interference, additive noise and fixed inputs is one of the important examples of such channels [7].

Under these or similar conditions the *maximum achievable information rate* for the finite-state channel model has been estimated in [8–10] using a simulation-based algorithm. Bounds on the computational cut-off rate R_0^* for the finite-state channel models were studied, particularly, in [11–16]. Some of the very first random coding bounds for chan-

nels with memory were published in [17, 18]. In [18] random coding bounds for the discrete additive finite-state channel were obtained for mismatched decoding different from the ML decoding function, and that publication gave a basic idea for this paper. The main known results for the random coding bounds for discrete-time channels with time-invariant intersymbol interference (ISI) consist in evaluation the value R_0^* for *ML decoding* and *i. u. d. code symbols*. In [11], this task was solved for a particular case of the ISI channel with the length of interference equal to 1. More general cases of ISI were studied in [12–16]. Therefore, one may say that for channels with ISI it is known how to calculate the value of R_0 (the technique based on the Perron — Frobenius theorem) and the value of *maximum achievable information rate* C^* (simulation-based algorithm [9]). Taking these values into account let us introduce function $\tilde{E}_r^*(R)$, as

$$\tilde{E}_r^*(R) = \begin{cases} R_0^* - R, & 0 \leq R \leq R_0^*, \\ 0, & R \geq C^*. \end{cases} \quad (1)$$

Note that the function $\tilde{E}_r^*(R)$ in is not defined for rates in the range $R_0^* < R < C^*$. Let us denote by $E_r^*(R)$ the random coding exponent obtained for the discrete-time ISI channel, ML decoding, and a code ensemble with i. u. d. code symbols. Then, for the values of $R \in [0, R_0^*] \cup \{C^*\}$ the inequality $E_r^*(R) \geq \tilde{E}_r^*(R)$ is valid. Clearly, $E_r^*(R) = \tilde{E}_r^*(R)$ in the interval $0 \leq R \leq R_{cr}$, where R_{cr} is the *critical rate*, and in the point $R = C^*$. In this paper we introduce a decoding function ψ with partial overlapping which depends on two integer parameters W and B , $W \geq B \geq 0$. With the use of this function we obtain the random coding exponent $E_r^*(R; \psi)$ which can be a good approximation of the function $E_r^*(R)$ if $W, B \rightarrow \infty$ even for the those values of R for which the function $\tilde{E}_r^*(R)$ is not defined.

This paper is the first part of a general study, consisting of two parts, published separately. In the second part of the work we intend to present a number of examples and their discussion.

Notation and Basic Equations

Let $p_{y|x}(y|x)$ be the transition probability of the discrete-time channel; for the continuous-output channel it is instead a probability density function (p.d.f.); $\mathbf{x} \in X^N$, where X be a discrete input channel alphabet and $q_x = |X| < \infty$; $\mathbf{y} \in Y^N$, where Y is the channel output alphabet and N is the length of a block code. For the continuous channel output $Y = \mathbb{R}$. The set Y may also represent a quantized version of the continuous channel outputs, i. e. $|Y| = q$. In this study consider this case.

The notation $p_{y|x}(y|x)$ will mainly be used when vectors \mathbf{x} and \mathbf{y} have the equal lengths, e. g., N . For subvectors, or segments of vectors, \mathbf{x} and \mathbf{y} the notation \mathbf{x} and \mathbf{y} is used. The difference between them is noted due to the use of ordinary and *sans serif* font. This notation is context-dependent; in particular, the length of \mathbf{x} and/or \mathbf{y} can differ in the various contexts.

To indicate a segment of an arbitrary vector \mathbf{z} we use the notation $\mathbf{z}_a^b = (z^{(\max(1,a))}, z^{(\max(1,a)+1)}, \dots, z^{(\min(b,L))})$, where L is length of the vector \mathbf{z} .

Let $p_{\mathbf{x}}(\mathbf{x})$ be a distribution on the code ensemble, where $\mathbf{x} = (x^{(1)}, x^{(2)}, \dots, x^{(N)})$, and $p_x(x^{(n)})$ be a one-dimensional distribution giving the distribution of a single code symbol, $n = 1, 2, \dots, N$.

We assume the decoding rule is given by $\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} \psi(\mathbf{y}; \mathbf{x})$, where $\psi(\mathbf{y}; \mathbf{x})$ is a real-valued positive decoding function, and the maximization is performed over all code words.

Using standard techniques [1] one can obtain the random coding bound $P_e < \exp_2(-NE_r(N, R, \psi))$, where P_e is the block error probability; $E_r(N, R, \psi)$ is the generalized exponent of the random coding bound defined as

$$E_r(N, R, \psi) = \max_{p_{\mathbf{x}}} \max_{1 \geq \rho \geq 0} \max_{\lambda > 0} (E_0(N, \psi, \rho, \lambda, p_{\mathbf{x}}) - \rho R),$$

where λ, ρ are the optimization parameters; R is the code rate, and (see also [5])

$$\begin{aligned} E_0(N, \psi, \rho, \lambda, p_{\mathbf{x}}) &= \\ &= -\frac{1}{N} \log \sum_{\mathbf{y}} \sum_{\mathbf{x}} p_{\mathbf{x}}(\mathbf{x}) p_{y|x}(\mathbf{y}|\mathbf{x}) \psi(\mathbf{y}; \mathbf{x})^{-\lambda \rho} \times \\ &\quad \times \left(\sum_{\mathbf{x}'} p_{\mathbf{x}}(\mathbf{x}') \psi(\mathbf{y}; \mathbf{x}')^{\lambda} \right)^{\rho}. \end{aligned} \quad (2)$$

Hereafter $\log(\cdot)$ denotes the binary logarithm. The expression is an approximation to the exponent of random coding for a channel with a continuous output and the accuracy of this approximation increases with increasing number of quantization levels q .

Assigning $\psi(\mathbf{y}; \mathbf{x}) = p_{y|x}(y|x)$ corresponds to ML decoding; in this case the optimal value of the parameter λ is equal to $\lambda = 1/(1 + \rho)$, and we get classical expression for the random coding exponent [1]

$$\begin{aligned} E_r(N, R, p_{y|x}) &= E_r(N, R) = \\ &= \max_{p_{\mathbf{x}}} \max_{1 \geq \rho \geq 0} (E_0(N, \rho, p_{\mathbf{x}}) - \rho R), \end{aligned}$$

where

$$E_0(N, \rho, p_{\mathbf{x}}) = -\frac{1}{N} \log \sum_{\mathbf{y}} \left(\sum_{\mathbf{x}} p_{\mathbf{x}}(\mathbf{x}) p_{y|x}(\mathbf{y}|\mathbf{x})^{1+\rho} \right)^{1+\rho}.$$

As $N \rightarrow \infty$, there is an *asymptotic* generalized random coding bound $P_e < \exp_2(-N(E_r(R; \psi) - o(N)))$, where

$$E_r(R; \psi) = \max_{p_x} \max_{1 \geq \rho \geq 0} \left(\max_{\lambda > 0} (E_0(\psi, \rho, \lambda, p_x) - \rho R) \right),$$

and

$$E_0(\psi, \rho, \lambda, p_x) = \lim_{N \rightarrow \infty} E_0(N, \psi, \rho, \lambda, p_x).$$

If $\psi(\mathbf{y}; \mathbf{x}) = p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$, then again the optimal value of the parameter λ is $\lambda = 1/(1 + \rho)$ and we get asymptotic expression for the asymptotic random coding exponent for ML decoding [1]

$$E_r(R, p_{\mathbf{y}|\mathbf{x}}) = E_r(R) = \max_{p_x} \max_{1 \geq \rho \geq 0} (E_0(\rho, p_x) - \rho R),$$

where $E_0(\rho, p_x) = \lim_{N \rightarrow \infty} E_0(N, \rho, p_x)$. It is known [1] that $E_r(R) > 0$, if $R < C$, where C is the channel capacity, which can be found from the function $E_0(\rho, p_x)$ as

$$C = \max_{p_x} \left. \frac{\partial E_0(\rho, p_x)}{\partial \rho} \right|_{\rho=0}.$$

We assume that the code ensemble distribution is given as $p_x(\mathbf{x}) = \prod_{n=1}^N p_x(x^{(n)})$, $\mathbf{x} \in X^N$, and $p_x(x) = 1/q_x$, $x \in X$, i. e. it corresponds to the i. u. d. channel input. This assumption leads to *loss of optimality* but simplifies further consideration. Under these assumptions one can derive the suboptimal exponent of the random coding bound in *asymptotic form*

$$E_r^*(R; \psi) = \max_{1 \geq \rho \geq 0} \left(\max_{\lambda > 0} E_0^*(\psi, \rho, \lambda) - \rho R \right), \quad (3)$$

where

$$E_0^*(\psi, \rho, \lambda) = (1 + \rho) \log q_x - \lim_{N \rightarrow \infty} \frac{1}{N} \log \sum_{\mathbf{y}} \sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) \psi(\mathbf{y}; \mathbf{x})^{-\lambda \rho} \times \left(\sum_{\mathbf{x}'} \psi(\mathbf{y}; \mathbf{x}')^\lambda \right)^\rho. \quad (4)$$

By analogy with the channel capacity C let us define the lower bound on maximum achievable code rate $C^*(\psi)$ as

$$C^*(\psi) = \left. \frac{\partial \max_{\lambda > 0} E_0^*(\psi, \rho, \lambda)}{\partial \rho} \right|_{\rho=0} \leq C. \quad (5)$$

The value $R_0^*(\psi) = \max_{\lambda > 0} E_0^*(\psi, 1, \lambda)$ gives a bound on the cut-off rate $R_0 = \max_{p_x} E_0(1, p_x)$; evidently the inequalities $R_0^*(\psi) \leq R_0^* \leq R_0$ are valid.

Similarly we can write the equations for function $E_r^*(N, R)$ for the ML decoding as

$$E_r^*(N, R) = \max_{1 \geq \rho \geq 0} (E_0^*(N, \rho) - \rho R), \quad (6)$$

where

$$E_0^*(N, \rho) = (1 + \rho) \log q_x - \frac{1}{N} \log \sum_{\mathbf{y}} \left(\sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}} \right)^{1+\rho}, \quad (7)$$

and the asymptotic random coding exponent as $E_r^*(R) = \lim_{N \rightarrow \infty} E_r^*(N, R)$ with bound on the maximum achievable information rate

$$R_{\max}^*(N) = \left. \partial E_0^*(N, \rho) / \partial \rho \right|_{\rho=0}, \quad R_{\max}^*(N) \leq C.$$

Channel Model and Decoding Function

Let the channel transition probabilities be given as $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) = \left(\sum_{\mathbf{s}} p_{\mathbf{y}\mathbf{x}\mathbf{s}}(\mathbf{y}, \mathbf{x}, \mathbf{s}) \right) / p_x(\mathbf{x})$, where $\mathbf{s} = (s^{(0)}, s^{(1)}, \dots, s^{(n)}, \dots)$ is the sequence of the channel states, $s^{(n)} \in S$, S is a set of the channel states and $|S| < \infty$, $p_{\mathbf{y}\mathbf{x}\mathbf{s}}(\mathbf{y}, \mathbf{x}, \mathbf{s})$ is the simultaneous probability of the vectors \mathbf{y} , \mathbf{x} and \mathbf{s} . Note that

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) = \frac{\sum_{\mathbf{s}} p_{\mathbf{y}\mathbf{x}\mathbf{s}}(\mathbf{y}, \mathbf{x}, \mathbf{s})}{p_x(\mathbf{x})} = \frac{\sum_{\mathbf{s}} p_{\mathbf{y}|\mathbf{x}\mathbf{s}}(\mathbf{y}|\mathbf{x}, \mathbf{s}) p_{\mathbf{x}\mathbf{s}}(\mathbf{x}, \mathbf{s})}{p_x(\mathbf{x})} = \sum_{\mathbf{s}} p_{\mathbf{y}|\mathbf{x}\mathbf{s}}(\mathbf{y}|\mathbf{x}, \mathbf{s}) p_{\mathbf{s}|\mathbf{x}}(\mathbf{s}|\mathbf{x}),$$

where $p_{\mathbf{y}|\mathbf{x}\mathbf{s}}(\mathbf{y}|\mathbf{x}, \mathbf{s})$ is the conditional probability of the channel output for the fixed vectors \mathbf{x} and \mathbf{s} , and $p_{\mathbf{s}|\mathbf{x}}(\mathbf{s}|\mathbf{x})$ is a conditional probability of the channel states for the given input vector \mathbf{x} .

Let us assume that the channel is a *probabilistic finite-state machine*, i. e.

$$p_{\mathbf{y}|\mathbf{x}\mathbf{s}}(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{n=1}^N p_{y|x s}(y^{(n)} | x^{(n)}, s^{(n-1)});$$

$$p_{\mathbf{s}|\mathbf{x}}(\mathbf{s}|\mathbf{x}) = p_s(s^{(0)}) \prod_{n=1}^N p_{s|x s}(s^{(n)} | x^{(n)}, s^{(n-1)}),$$

where $p_s(\cdot)$ is an unconditional (stationary) distribution on the set of channel states; $p_{s|x s}(s^{(n)} | x^{(n)}, s^{(n-1)})$ is the conditional channel state transition probabil-

ity. In addition, we assume that the input channel symbol $x^{(n)}$ and the current channel state $s^{(n-1)}$ are independent. Such a model is quite general enough and has been widely used (see, for example, [8, 9] and references therein). An explanatory illustration is given in Fig. 1 [9].

Some particular cases of this model are:

1. State transitions not depending on the input symbol (*channel with freely evolving states*) [9], i. e. $p_{s|xs}(s^{(n)} | x^{(n)}, s^{(n-1)}) = p_{s|s}(s^{(n)} | s^{(n-1)})$. The Gilbert — Elliot [19, 20] channel is a specific example of such a model.

2. Deterministic state transitions (*deterministic finite state machine*). In this case the state transition is given as

$$p_{s|xs}(s^{(n)} | x^{(n)}, s^{(n-1)}) = \begin{cases} 1, & \text{if } s^{(n-1)} \xrightarrow{x^{(n)}} s^{(n)}; \\ 0, & \text{otherwise.} \end{cases}$$

The ISI channel is an example of a channel in this class.

The probabilities $p_{y|x}(y | x)$ can be represented as

$$p_{y|x}(y | x) = \sum_{s^{(0)}} \dots \sum_{s^{(N-1)}} \sum_{s^{(N)}} p_s(s^{(0)}) \times \left(\prod_{n=1}^N p_{y|xs}(y^{(n)} | x^{(n)}, s^{(n-1)}) p_{s|xs}(s^{(n)} | x^{(n)}, s^{(n-1)}) \right),$$

or in the following form of *matrix product* proved similarly to the derivation of [1, eq. (5.9.39)]:

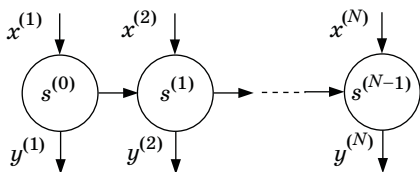
$$p_{y|x}(y | x) = \mathbf{p}_s \left(\prod_{n=1}^N \mathbf{P}(y^{(n)} | x^{(n)}) \right) \mathbf{1}^T, \quad (8)$$

where

$$\mathbf{P}(y | x) = [p_{y|xs}(y | x, s) p_{s|xs}(s' | x, s)] \quad (9)$$

is a matrix of size $|S| \times |S|$; $\mathbf{p}_s = [p_s(1), \dots, p_s(|S|)]$ is the vector of the unconditional state probabilities at $n = 0$, and $\mathbf{1} = (1, \dots, 1)$ is vector of 1's of dimensions $1 \times |S|$.

Let us choose the decoding function as the product of APP of input segments of length $2B + 1$ for fixed output segments of length $2W + 1$, $W \geq B \geq 0$ are integer parameters,



■ Fig. 1. Finite-state channel transitions

$$\psi(y; \mathbf{x}) = \prod_{n=0}^{N(B)-1} \Pr \left[\mathbf{x}_{n(2B+1)-B}^{n(2B+1)+B} | \mathbf{y}_{n(2B+1)-W}^{n(2B+1)+W} \right],$$

and

$$N(B) = \lceil (N - B) / (2B + 1) \rceil. \quad (10)$$

The value $N(B)$, defined in (10), gives the number of the code block segments, or subblocks, of the length $2B + 1$. As will be seen in the following, for such a decoding function, it is possible to obtain a good final result for the suboptimal random coding exponent.

Let us denote

$$k(n) = n(2B + 1) + 1. \quad (11)$$

This value gives the position of the central element of the n^{th} code block segment of length $2B + 1$.

For equiprobable segments $\mathbf{x}_{k(n)-B}^{k(n)+B}$ we can write the expression for the decoding function $\psi(y; \mathbf{x})$ in another equivalent form

$$\psi(y; \mathbf{x}) = \prod_{n=0}^{N(B)-1} p_{y|x} \left(\mathbf{y}_{k(n)-W}^{k(n)+W} | \mathbf{x}_{k(n)-B}^{k(n)+B} \right), \quad (12)$$

where $p_{y|x}(\cdot)$ is the conditional probability for segments of *different*, in general, lengths $2W + 1$ and $2B + 1$. Note, that in general $p_{y|x}(\cdot) \neq p_{x|y}(\cdot)$. The difference in these values is emphasized by their index notation typed in Roman bold font and *sans serif* bold font respectively.

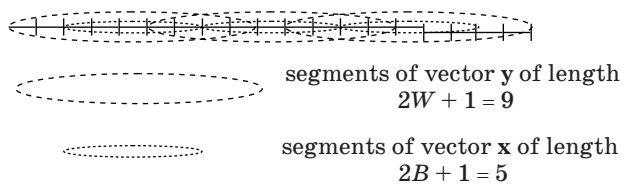
The segments $\mathbf{x}_{k(n)-B}^{k(n)+B}$ of length $2B + 1$ of the input vector \mathbf{x} do not overlap, but the segments $\mathbf{y}_{k(n)-W}^{k(n)+W}$ of length $2W + 1$ overlap on a segment of length $2(W - B)$. The illustration for $B = 2$, $W = 4$, $2B + 1 = 5$, $2W + 1 = 9$ is shown in Fig. 2.

Let us note some specific cases of the decoding function (12).

Case 1. If $W = B = 0$, then $N(B) = N$, and the decoding function is

$$\psi(y; \mathbf{x}) = \prod_{n=0}^{N-1} p_{y|x} \left(\mathbf{y}_{n+1}^{n+1} | \mathbf{x}_{n+1}^{n+1} \right) = \prod_{n=0}^{N-1} p_{y|x} \left(y^{(n)} | x^{(n)} \right),$$

i. e. it is matched with the memoryless channel.



■ Fig. 2. Positions of the subblocks in product (12)

Case 2. If $B = 0$ and $W \geq 1$, then $N(B) = N$, and the decoding function is

$$\begin{aligned} \psi(\mathbf{y}; \mathbf{x}) &= \prod_{n=0}^{N-1} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{n+1}^{n+W} | \mathbf{x}^{(n)}) = \\ &= \prod_{n=0}^{N-1} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{n-W}^{n+W} | \mathbf{x}^{(n)}), \end{aligned}$$

and such decoding is equivalent to APP symbol decoding in a window of the length $2W + 1$, and it is similar to windowed version of the Bahl — Cocke — Jelinek — Raviv (BCJR) algorithm [21], (see e. g. [22]).

Case 3. If $W = B = N$, then $N(B) = 1$, and the decoding function is

$$\begin{aligned} \psi(\mathbf{y}; \mathbf{x}) &= p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{1-N}^{1+N} | \mathbf{x}_{1-N}^{1+N}) = \\ &= p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_1^N | \mathbf{x}_1^N) = p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x}), \end{aligned}$$

and corresponds to the ML decoding.

It can be shown (see Appendix A) that the probabilities $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ on the right-hand side of can be found for any $\mathbf{y} \in Y^{2W+1}$ and $\mathbf{x} \in X^{2B+1}$ as

$$\begin{aligned} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x}) &= q_x^{-2(W-B)} \mathbf{p}_s \left(\prod_{l=1}^{W-B} \mathbf{P}(\mathbf{y}^{(l)}) \right) \times \\ &\times \left(\prod_{l=W-B+1}^{W+B+1} \mathbf{P}(\mathbf{y}^{(l)} | \mathbf{x}^{(l)}) \right) \left(\prod_{l=W+B+2}^{2W+1} \mathbf{P}(\mathbf{y}^{(l)}) \right) \mathbf{1}^T, \end{aligned} \quad (13)$$

where $\mathbf{y}^{(l)}$, $\mathbf{x}^{(l)}$ are components of the vectors \mathbf{y} and \mathbf{x} respectively; $\mathbf{P}(\mathbf{y}|\mathbf{x})$ is $|S| \times |S|$ matrix defined by equation (9), and $\mathbf{P}(\mathbf{y})$ is matrix defined as the sum

$$\mathbf{P}(\mathbf{y}) = \sum_{\mathbf{x} \in X} \mathbf{P}(\mathbf{y} | \mathbf{x}). \quad (14)$$

Suboptimal Random Coding Exponent

Consider the sums in the right-hand part of the equation (4). It can be shown (see Appendix B) that

$$\begin{aligned} \sum_{\mathbf{x}} \psi(\mathbf{y}; \mathbf{x})^\lambda &= a(\mathbf{y}_1^{k(n_1-1)+W}) \times \\ &\times \left(\prod_{n=n_1}^{n_2} D_1(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda) \right) b(\mathbf{y}_{k(n_2+1)-W}^N), \end{aligned}$$

where

$$D_1(\mathbf{y}; \lambda) = \sum_{\mathbf{x} \in X^{2B+1}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x})^\lambda, \quad \mathbf{y} \in Y^{2W+1},$$

n_1 is the least positive integer such that $k(n_1) - W \geq 1$, and n_2 is the greatest integer such that, $k(n_2) + W \leq N$, or according to the definition (11)

$$n_1 = \lceil W / (2B + 1) \rceil, \quad n_2 = \lfloor (N - W - 1) / (2B + 1) \rfloor. \quad (15)$$

We denote by $a(\cdot)$ and $b(\cdot)$ the non-essential (not affecting the exponent of random coding) positive factors. It is also shown in Appendix B that

$$\begin{aligned} \sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x}) \psi(\mathbf{y}; \mathbf{x})^{-\lambda\rho} &= \mathbf{p}_s \mathbf{A}(\mathbf{y}_1^{k(n_1-1)+W}) \times \\ &\times \left(\prod_{n=n_1}^{n_2} \mathbf{D}_2(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda\rho) \right) \mathbf{B}(\mathbf{y}_{k(n_2+1)-W}^N) \mathbf{1}^T, \end{aligned}$$

where $\mathbf{D}_2(\mathbf{y}; \lambda\rho)$ is $|S| \times |S|$ matrix defined as

$$\begin{aligned} \mathbf{D}_2(\mathbf{y}; \lambda\rho) &= \\ &= \sum_{\mathbf{x} \in X^{2B+1}} \mathbf{P}_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{W-B+1}^{W+B+1} | \mathbf{x}) p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x})^{-\lambda\rho}, \end{aligned} \quad (16)$$

$\mathbf{y} \in Y^{2W+1}$; $\mathbf{A}(\cdot)$ и $\mathbf{B}(\cdot)$ are inessential nonnegative matrix multipliers of dimension $|S| \times |S|$, and

$$\mathbf{P}_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x}) = \prod_{l=1}^{2B+1} \mathbf{P}(\mathbf{y}^{(l)} | \mathbf{x}^{(l)}), \quad \mathbf{y} \in Y^{2B+1}, \quad \mathbf{x} \in X^{2B+1}, \quad (17)$$

where the matrices $\mathbf{P}(\cdot)$ are defined by the equation (9). Note, that $\mathbf{y}_{W-B+1}^{W+B+1}$ in the right-hand part of (16) is the middle part of the vector \mathbf{y} having length $2B + 1$.

Using these notations we get from (4) that

$$\begin{aligned} E_0^*(\psi, \rho, \lambda) &= (1 + \rho) \log q_x - \\ &- \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbf{p}_s \sum_{\mathbf{y}} \left(\mathbf{U}(\mathbf{y}_1^{k(n_1-1)+W}) \times \right. \\ &\times \left. \left(\prod_{n=n_1}^{n_2} \mathbf{D}(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda, \rho) \right) \mathbf{V}(\mathbf{y}_{k(n_2+1)-W}^N) \right) \mathbf{1}^T, \end{aligned} \quad (18)$$

where

$$\begin{aligned} \mathbf{U}(\mathbf{y}) &= a(\mathbf{y})^\rho \mathbf{A}(\mathbf{y}), \quad \mathbf{y} \in Y^{k(n_1-1)+W}; \\ \mathbf{V}(\mathbf{y}) &= \mathbf{B}(\mathbf{y}) b(\mathbf{y})^\rho, \quad \mathbf{y} \in Y^{N-k(n_2+1)+W+1}, \end{aligned} \quad (19)$$

$$\mathbf{D}(\mathbf{y}; \lambda, \rho) = D_1(\mathbf{y}; \lambda)^\rho \mathbf{D}_2(\mathbf{y}; \lambda\rho), \quad \mathbf{y} \in Y^{2W+1}. \quad (20)$$

In general (see Appendix C), the sum over \mathbf{y} on the right-hand side of (16) can be written in the form

$$\sum_{\mathbf{y}} \prod_{n=n_1}^{n_2} \mathbf{D}(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda, \rho) = \mathbf{F} \mathbf{K}(\lambda, \rho)^{n_2-n_1+1} \mathbf{G}, \quad (21)$$

where

$$\mathbf{K}(\lambda, \rho) = [\mathbf{K}_{ij}(\lambda, \rho)] = \begin{bmatrix} \mathbf{K}_{11}(\lambda, \rho) & \dots & \mathbf{K}_{1q^{2(W-B)}}(\lambda, \rho) \\ \dots & \dots & \dots \\ \mathbf{K}_{q^{2(W-B)}1}(\lambda, \rho) & \dots & \mathbf{K}_{q^{2(W-B)}q^{2(W-B)}}(\lambda, \rho) \end{bmatrix} \quad (22)$$

is a square block matrix of order $|S|q^{2(W-B)}$ built of the blocks $\mathbf{K}_{ij}(\lambda, \rho)$ of dimension $|S| \times |S|$ defined as

$$\mathbf{K}_{ij}(\lambda, \rho) = \begin{cases} \mathbf{D}(\mathbf{y}; \lambda, \rho), & W \geq 2B+1; \\ \sum_{\mathbf{y}_{2^{B+1}}^{2(W-B)+1}} \mathbf{D}(\mathbf{y}; \lambda, \rho), & W < 2B+1 \end{cases} \quad (23)$$

for $i, j = 1, \dots, q^{2(W-B)}$. The matrix $\mathbf{D}(\mathbf{y}; \lambda, \rho)$ in equations (21) and (23) is defined in (20). The correspondence of the indices i, j and the vector \mathbf{y} in the expression (23) is given as $i \leftrightarrow \mathbf{y}_1^{2(W-B)}$ and $j \leftrightarrow \mathbf{y}_{2^{B+1}}^{2W+1}$. In other words, the components of the vectors $\mathbf{y}_1^{2(W-B)}$ and $j \leftrightarrow \mathbf{y}_{2^{B+1}}^{2W+1}$ are the digits in q -ary representation of the indices i and j respectively.

The matrices \mathbf{F} and \mathbf{G} on the right-hand side of (21) are nonnegative matrix multipliers of dimension $|S| \times |S|q^{2(W-B)}$ and $|S|q^{2(W-B)} \times |S|$.

With these definitions equation (18) can be rewritten as follows

$$E_0^*(\psi, \rho, \lambda) = (1+\rho) \log q_x - \lim_{N \rightarrow \infty} \frac{1}{N} \log(\mathbf{f} \mathbf{K}(\lambda, \rho)^{n_2 - n_1 + 1} \mathbf{g}), \quad (24)$$

where n_1, n_2 are defined in equations (15) and $\mathbf{f} = \mathbf{p}_s \mathbf{G}$; $\mathbf{g} = \mathbf{G} \mathbf{1}^T$ are inessential nonnegative vectors. In what follows we use the following assertion.

Corollary from the Perron — Frobenius theorem [1, 23]. Let \mathbf{A} be a nonnegative irreducible square matrix, \mathbf{a} and \mathbf{b} be nonnegative vectors of the corresponding dimensions, then

$$\lim_{N \rightarrow \infty} N^{-1} \log \mathbf{a} \mathbf{A}^N \mathbf{b} = \log r(\mathbf{A}),$$

where $r(\mathbf{A})$ is maximum eigenvalue (spectral radius) of the matrix \mathbf{A} .

Using this corollary and the definitions in (15) we get from the equation (24) that, if the matrix $\mathbf{K}(\lambda, \rho)$ is irreducible, then

$$E_0^*(\psi, \rho, \lambda) = (1+\rho) \log q_x - (2B+1)^{-1} \log r(\mathbf{K}(\lambda, \rho)), \quad (25)$$

where $r(\mathbf{K}(\lambda, \rho))$ is maximum eigenvalue (spectral radius) of the matrix $\mathbf{K}(\lambda, \rho)$. The similar approach has been used many times in early publications such as [1, 17, 18], and later, for example, in [12–16].

Let us consider the conditions for matrix $\mathbf{K}(\lambda, \rho)$, defined in (22), to be irreducible. Obviously for the matrix $\mathbf{K}(\lambda, \rho)$ to be irreducible it is sufficient that each of its blocks $\mathbf{K}_{ij}(\lambda, \rho)$ is irreducible. Matrices $\mathbf{K}_{ij}(\lambda, \rho)$ defined by the equation (23) are linear combinations of the matrices $\mathbf{P}_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ [see (16) and (17)]. Hence, for irreducibility of the matrix $\mathbf{K}_{ij}(\lambda, \rho)$ it is sufficient that the matrices in (17) are irreducible. This condition is satisfied if the matrices (9) are irreducible. Irreducibility of the matrices (9) means that any channel state is reachable from any other state over a finite number of steps when receiving independent, equally distributed symbols to the channel input. Below we assume that the matrices (9) are irreducible for any x and y .

After substitution (25) into (3) we have the following theorem.

Theorem. Let channel be specified by conditional probabilities (8), where the matrices (9) are irreducible, and let the decoding function ψ be given by equation (12) with integer parameters W and B , where $W \geq B \geq 0$. Then the achievable random coding exponent $E_r^*(R; \psi)$ for the code ensemble with i. u. d. code symbols is

$$E_r^*(R; \psi) = \max_{0 \leq \rho \leq 1} (E_0^*(\psi, \rho) - \rho R),$$

where $E_0^*(\psi, \rho) = \max_{\lambda > 0} E_0^*(\psi, \rho, \lambda) = (1+\rho) \log q_x - (2B+1)^{-1} \log \left(\min_{\lambda > 0} r(\mathbf{K}(\lambda, \rho)) \right)$, and $r(\mathbf{K}(\lambda, \rho))$ is the maximum eigenvalue (spectral radius) of matrix $\mathbf{K}(\lambda, \rho)$, given in equation (22).

Conclusion

In this paper representing the first part of the general study we built a random coding bound applicable to a wide class of channels with memory defined as probabilistic finite-state machine. This class of the models describes many transmission channels important for theory and practice. Among them, we can highlight channel models with intersymbol interference, which are widely used for description of data transmission and recording systems. To obtain the main result, an approach used mismatched decoding function is applied. The choice of decoding function was the main problem of this study. In the second part of the work we will give examples of calculating the exponent of random coding for several models of channels with memory, their discussion and comparison of these results with known ones.

Appendix A

Let us consider the computation of the probabilities $p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k-W}^{k+W} | \mathbf{x}_{k-B}^{k+B})$ for some k . The segment \mathbf{x}_{k-W}^{k+W} can be represented as a concatenation $\mathbf{x}_{k-W}^{k+W} = (\mathbf{x}_{k-W}^{k-B-1}, \mathbf{x}_{k-B}^{k+B}, \mathbf{x}_{k+B+1}^{k+W})$. Therefore,

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k-W}^{k+W} | \mathbf{x}_{k-B}^{k+B}) = \sum_{\mathbf{x}_{k-W}^{k-B-1}} \sum_{\mathbf{x}_{k+B+1}^{k+W}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k-W}^{k+W} | \mathbf{x}_{k-W}^{k+W}) \times p_{\mathbf{x}}(\mathbf{x}_{k-W}^{k-B-1}) p_{\mathbf{x}}(\mathbf{x}_{k+B+1}^{k+W}).$$

For i. u. d. components of the vectors \mathbf{x}_{k-W}^{k-B-1} and \mathbf{x}_{k+B+1}^{k+W} the following equation is valid: $p_{\mathbf{x}}(\mathbf{x}_{k-W}^{k-B-1}) = p_{\mathbf{x}}(\mathbf{x}_{k+B+1}^{k+W}) = q_x^{-(W-B)}$. Hence,

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k-W}^{k+W} | \mathbf{x}_{k-B}^{k+B}) = q_x^{-2(W-B)} \sum_{\mathbf{x}_{k-W}^{k-B-1}} \sum_{\mathbf{x}_{k+B+1}^{k+W}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k-W}^{k+W} | \mathbf{x}_{k-W}^{k+W}). \quad (\text{A1})$$

Then, using the equation (8), we have

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k-W}^{k+W} | \mathbf{x}_{k-B}^{k+B}) = \mathbf{p}_s \left(\prod_{l=k-W}^{k+W} \mathbf{P}(y^{(l)} | x^{(l)}) \right) \mathbf{1}^T = \mathbf{p}_s \left(\prod_{l=k-W}^{k-B-1} \mathbf{P}(y^{(l)} | x^{(l)}) \right) \left(\prod_{l=k-B}^{k+B} \mathbf{P}(y^{(l)} | x^{(l)}) \right) \times \left(\prod_{l=k+B+1}^{k+W} \mathbf{P}(y^{(l)} | x^{(l)}) \right) \mathbf{1}^T$$

and further

$$\sum_{\mathbf{x}_{k-W}^{k-B-1}} \sum_{\mathbf{x}_{k+B+1}^{k+W}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k-W}^{k+W} | \mathbf{x}_{k-B}^{k+B}) = \mathbf{p}_s \left(\prod_{l=k-W}^{k-B-1} \mathbf{P}(y^{(l)}) \right) \left(\prod_{l=k-B}^{k+B} \mathbf{P}(y^{(l)} | x^{(l)}) \right) \times \left(\prod_{l=k+B+1}^{k+W} \mathbf{P}(y^{(l)}) \right) \mathbf{1}^T,$$

where $\mathbf{P}(y) = \sum_{x \in X} \mathbf{P}(y | x)$. Then using equation (A1) and notation (14) we obtain the expression (13) for $k = W + 1$.

Appendix B

To calculate the sum $\sum_{\mathbf{x}} \psi(\mathbf{y}; \mathbf{x})^\lambda$, let us represent the vector \mathbf{x} as a sequence of $N(B)$ subvectors

of length $2B + 1$ except perhaps the first and last segments,

$$\mathbf{x} = \left(\mathbf{x}_{k(0)-B}^{k(0)+B}, \mathbf{x}_{k(1)-B}^{k(1)+B}, \dots, \mathbf{x}_{k(n)-B}^{k(n)+B}, \dots, \mathbf{x}_{k(N(B)-1)-B}^{k(N(B)-1)+B} \right), \quad (\text{B1})$$

where $k(n)$ is defined in equation (11). Then from the equation (12), it follows that

$$\psi(\mathbf{y}; \mathbf{x})^\lambda = a_1^\lambda \left(\prod_{n=n_1}^{n_2} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k(n)-W}^{k(n)+W} | \mathbf{x}_{k(n)-B}^{k(n)+B}) \right) b_1^\lambda, \quad (\text{B2})$$

where n_1 is the smallest integer such that $k(n_1) - W \geq 1$, n_2 is the greatest integer such that $k(n_2) + W \leq N$, and a_1, b_1 are positive multipliers depending on initial and final segments of the vectors \mathbf{x} and \mathbf{y} .

After a summation over i. u. d. and disjoint segments of the vector \mathbf{x} , we have

$$\sum_{\mathbf{x}} \psi(\mathbf{y}; \mathbf{x})^\lambda = a(\lambda) \left(\prod_{n=n_1}^{n_2} D_1(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda) \right) b(\lambda), \quad (\text{B3})$$

where

$$D_1(\mathbf{y}; \lambda) = \sum_{\mathbf{x} \in X^{2B+1}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x})^\lambda, \quad \mathbf{y} \in Y^{2W+1},$$

and $a(\lambda), b(\lambda)$ are positive multipliers obtained after summing the quantities a_1^λ and b_1^λ over the initial and final segments of the vector \mathbf{x} .

Let us now derive an expression for the sum $\sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x}) \psi(\mathbf{y}; \mathbf{x})^{-\lambda\rho}$. In this case, we also represent the vector \mathbf{x} as a sequence of subvectors of length $2B + 1$ except the first and last ones, which may have a different length [see equation (B1)]. Using equation (B1) one can rewrite the expression

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x}) = \mathbf{p}_s \mathbf{A}_1 \left(\prod_{n=n_1}^{n_2} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}_{k(n)-B}^{k(n)+B} | \mathbf{x}_{k(n)-B}^{k(n)+B}) \right) \mathbf{B}_1 \mathbf{1}^T, \quad (\text{B4})$$

where, as before, n_1 is the least integer such that $k(n_1) - W \geq 1$, and n_2 is the greatest integer such that $k(n_2) + W \leq N$; \mathbf{A}_1 и \mathbf{B}_1 nonnegative matrix multipliers of dimension $|S| \times |S|$, corresponding to the first and last segments of the vector \mathbf{x} , and

$$p_{\mathbf{y}|\mathbf{x}}(\mathbf{y} | \mathbf{x}) = \prod_{l=1}^{2B+1} \mathbf{P}(y^{(l)} | x^{(l)}), \quad \mathbf{y} \in Y^{2B+1}, \quad \mathbf{x} \in X^{2B+1}, \quad (\text{B5})$$

where the matrices $\mathbf{P}(\mathbf{y}^{(l)}|\mathbf{x}^{(l)})$ in the right-hand part of (B5) is defined by equation (9).

Using expressions (B4) and (B2), we can write

$$\begin{aligned} & p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})\psi(\mathbf{y}; \mathbf{x})^{-\lambda\rho} = \\ & = \mathbf{p}_s \mathbf{a}_1^{-\lambda\rho} \mathbf{A}_1 \left(\prod_{n=n_1}^{n_2} \mathbf{P}_{\mathbf{y}|\mathbf{x}} \left(\mathbf{y}_{k(n)-B}^{k(n)+B} | \mathbf{x}_{k(n)-B}^{k(n)+B} \right) \times \right. \\ & \left. \times \mathbf{P}_{\mathbf{y}|\mathbf{x}} \left(\mathbf{y}_{k(n)-W}^{k(n)+W} | \mathbf{x}_{k(n)-B}^{k(n)+B} \right)^{-\lambda\rho} \right) b_1^{-\lambda\rho} \mathbf{B}_1 \mathbf{1}^T. \end{aligned}$$

Summing over i. u. d. and disjoint components of the vector \mathbf{x} , we get

$$\begin{aligned} & \sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})\psi(\mathbf{y}; \mathbf{x})^{-\lambda\rho} = \\ & = \mathbf{p}_s \mathbf{A} \left(\prod_{n=n_1}^{n_2} \sum_{\mathbf{x} \in X^{2B+1}} \mathbf{P}_{\mathbf{y}|\mathbf{x}} \left(\mathbf{y}_{k(n)-B}^{k(n)+B} | \mathbf{x} \right) \times \right. \\ & \left. \times \mathbf{P}_{\mathbf{y}|\mathbf{x}} \left(\mathbf{y}_{k(n)-W}^{k(n)+W} | \mathbf{x} \right)^{-\lambda\rho} \right) \mathbf{B}_1^T, \end{aligned} \quad (\text{B6})$$

where \mathbf{A} и \mathbf{B} are inessential matrix multipliers obtained after summation $\mathbf{a}_1^{-\lambda\rho} \mathbf{A}_1$ and $b_1^{-\lambda\rho} \mathbf{B}_1$ over the initial and final segments of the vector \mathbf{x} respectively. And finally it follows from (B6) that

$$\begin{aligned} & \sum_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})\psi(\mathbf{y}; \mathbf{x})^{-\lambda\rho} = \\ & = \mathbf{p}_s \mathbf{A} \left(\prod_{n=n_1}^{n_2} \mathbf{D}_2 \left(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda\rho \right) \right) \mathbf{B}_1^T, \end{aligned} \quad (\text{B7})$$

where $\mathbf{D}_2(\mathbf{y}; \lambda\rho)$ is matrix $|\mathcal{S}| \times |\mathcal{S}|$, defined as

$$\begin{aligned} \mathbf{D}_2(\mathbf{y}; \lambda\rho) &= \sum_{\mathbf{x} \in X^{2B+1}} \mathbf{P}_{\mathbf{y}|\mathbf{x}} \left(\mathbf{y}_{W-B+1}^{W+B+1} | \mathbf{x} \right) \times \\ & \times p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})^{-\lambda\rho}, \quad \mathbf{y} \in Y^{2W+1}, \end{aligned} \quad (\text{B8})$$

and matrix $\mathbf{P}_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x})$ is given by equation (B5).

Appendix C

Consider two adjacent terms in the product on the left-hand side of (21)

$$\begin{aligned} & \prod_{n=n_1}^{n_2} \mathbf{D} \left(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda, \rho \right) = \mathbf{D} \left(\mathbf{y}_{k(n_1)-W}^{k(n_1)+W}; \lambda, \rho \right) \times \\ & \times \mathbf{D} \left(\mathbf{y}_{k(n_1+1)-W}^{k(n_1+1)+W}; \lambda, \rho \right) \dots \mathbf{D} \left(\mathbf{y}_{k(n_2)-W}^{k(n_2)+W}; \lambda, \rho \right). \end{aligned}$$

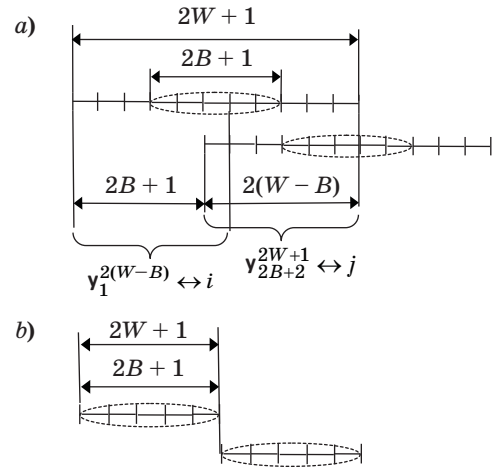
They depend on two adjacent blocks $\mathbf{y}_{k(n)-W}^{k(n)+W}$ and $\mathbf{y}_{k(n+1)-W}^{k(n+1)+W}$. The indices of the first block are

$n(2B+1)+1-W, \dots, n(2B+1)+1+W$, and indices of the second block are $(n+1)(2B+1)+1-W, \dots, (n+1)(2B+1)+1+W$, i. e. the second set of indices is shifted right to $2B+1$ positions. An intersection of the position numbers exists on an interval of length $(2W+1) - (2B+1) = 2(W-B)$. If $W > B$, then this intersection is not empty; an illustration is shown in Fig. C1, a. Since $W \geq B$, the case $W = B$ is also possible. In this case there is no intersection and the summing in over vector \mathbf{y} is reduced to summing separate factors over segments $\mathbf{y}_{k(n)-W}^{k(n)+W}$.

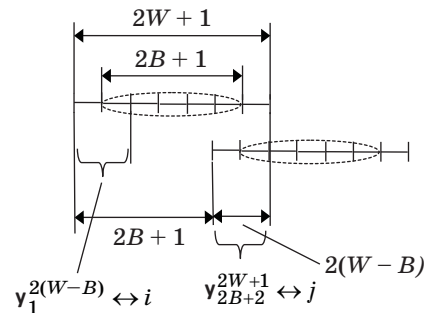
This case is simpler to analyze and hence omitted. The illustration is given in Fig. C1, b.

Consider the case that $W > B$. Let us introduce indices $i, j = 1, 2, \dots, q^{2(W-B)}$ and establish a one-to-one correspondence between indices i, j and vectors $\mathbf{y} \in Y^{2W+1}$ as $i \leftrightarrow \mathbf{y}_1^{2(W-B)}$ и $j \leftrightarrow \mathbf{y}_{2B+2}^{2W+1}$. Then we consider two cases: 1) the vectors $\mathbf{y}_1^{2(W-B)}$ and \mathbf{y}_{2B+2}^{2W+1} have common elements (see Fig. C1, a), and 2) the vectors $\mathbf{y}_1^{2(W-B)}$ and \mathbf{y}_{2B+2}^{2W+1} have no common elements (see Fig. C2).

The first case takes place, if $2W+1 < 2 \times (2B+1)$, or if $W \geq 2B+1$, because W is integer.



■ Fig. C1. Layout of adjacent segments: a — intersection of segments, $W > B$; b — disjoint segments, $W = B$



■ Fig. C2. Intersection of adjacent segments, $2W < 4B+1$

The second case takes place if $W < 2B + 1$. Since in the second case two blocks $\mathbf{y}_1^{2(W-B)}$ and \mathbf{y}_{2B+2}^{2W+1} have no common elements, it is possible to sum independently over elements with indices from $2(W - B) + 1$ to $2B + 1$. Therefore, we can define the $|S| \times |S|$ matrices $\mathbf{K}_{ij}(\lambda, \rho)$ as

$$\mathbf{K}_{ij}(\lambda, \rho) = \begin{cases} \mathbf{D}(\mathbf{y}; \lambda, \rho), W \geq 2B + 1; \\ \sum_{\mathbf{y}_{2(W-B)+1}^{2B+1}} \mathbf{D}(\mathbf{y}; \lambda, \rho), W < 2B + 1, \end{cases} \quad (\text{C1})$$

where $\mathbf{y} \in Y^{2W+1}$, $i \leftrightarrow \mathbf{y}_1^{2(W-B)}$ and $j \leftrightarrow \mathbf{y}_{2B+2}^{2W+1}$.

Also define a block matrix $\mathbf{K}(\lambda, \rho) = [\mathbf{K}_{ij}(\lambda, \rho)]$ of order $|S|q^{2(W-B)}$. Then the sum over \mathbf{y} on right-hand side of (21) can be written as

$$\begin{aligned} & \sum_{\mathbf{y}} \left(\mathbf{U} \left(\mathbf{y}_1^{k(n_1-1)+W} \right) \prod_{n=n_1}^{n_2} \mathbf{D} \left(\mathbf{y}_{k(n)-W}^{k(n)+W}; \lambda, \rho \right) \times \right. \\ & \quad \left. \times \mathbf{V} \left(\mathbf{y}_{k(n_2+1)-W}^N \right) \right) = \\ & = \mathbf{F} \prod_{n=n_1}^n \mathbf{K}(\lambda, \rho) \mathbf{G} = \mathbf{F} \mathbf{K}(\lambda, \rho)^{n_2-n_1+1} \mathbf{G}, \quad (\text{C2}) \end{aligned}$$

where matrices $\mathbf{U}(\cdot)$ and $\mathbf{V}(\cdot)$ are defined by equations (19), \mathbf{F} is a block matrix of size $|S| \times |S|q^{2(W-B)}$, and \mathbf{G} is a block matrix of size $|S|q^{2(W-B)} \times |S|$.

The matrices \mathbf{F} and \mathbf{G} on the right-hand side of (C2) do not affect the asymptotic expression for the random coding exponent, so the description of their structure is not given due to space savings.

References

1. **Gallager R.** *Information Theory and Reliable Communication*. New York, John Wiley & Sons, 1968. 588 p.
2. **Ganti A., Lapidoth A., Telatar I. E.** Mismatched Decoding Revisited: General Alphabets, Channels with Memory, and the Wide-Band Limit. *IEEE Transactions on Information Theory*, 2000, vol. 46, no. 7, pp. 2315–2328.
3. **Lapidoth A., Narayan P.** Reliable Communication under Channel Uncertainty. *IEEE Transactions on Information Theory*, 1998, vol. 44, no. 7, pp. 2148–2177.
4. **Merhav N., Kaplan G., Lapidoth A., Shamai (Shitz) S.** On Information Rate for Mismatched Decoders. *IEEE Transactions on Information Theory*, 1994, vol. 40, no. 10, pp. 1953–1967.
5. **Kaplan G., Shamai (Shitz) S.** Information Rates and Error Exponents of Compound Channels with Application to Antipodal Signaling in a Fading Environment. *AEU International Journal of Electronics and Communications*, 1993, vol. 47, no. 4, pp. 228–230.
6. **Scarlett J., Martinez A., Guillén I Fàbregas A.** Mismatched Decoding: Error Exponents, Second-Order Rates and Saddlepoint Approximations. *IEEE Transactions on Information Theory*, 2014, vol. 60, no. 5, pp. 2647–2666.
7. **Forney Jr. G. D.** Maximum Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference. *IEEE Transactions on Information Theory*, 1972, vol. IT-18, no. 3, pp. 363–378.
8. **Kavcic A., Xiao Ma, Mitzenmacher M.** Binary Intersymbol Interference Channels: Gallager Codes, Density Evolution, and Code Performance Bounds. *IEEE Transactions on Information Theory*, 2003, vol. 47, no. 7, pp. 1636–1652.
9. **Arnold D. M., Loeliger H.-A., Vontobel P. O., Kavcic A., Zeng W.** Simulation-based Computation of Information Rates for Channels with Memory. *IEEE Transactions on Information Theory*, 2006, vol. 52, no. 8, pp. 3498–3508.
10. **Rusek F., Fertonani D.** Bounds on the Information Rate of Intersymbol Interference Channels based on Mismatched Receivers. *IEEE Transactions on Information Theory*, 2012, vol. 58, no. 3, pp. 1470–1482.
11. **Sasano H., Kasahara M., Namekawa T.** Evaluation of the Exponent Function $E(R)$ for Channels with Intersymbol Interference. *Electronics and Communications in Japan*, 1982, vol. 65-A, no. 8, pp. 28–37.
12. **Biglieri E.** The Computational Cutoff Rate of Channel Having Memory. *IEEE Transactions on Information Theory*, 1981, vol. 27, pp. 352–357.
13. **Raghavan S. A., Wolf J. K., Milstein L. B.** On the Cutoff Rate of a Discrete Memoryless Channel with (d, k) -constrained Input Sequences. *IEEE Journal on Selected Areas in Communications*, 1992, vol. 10, no. 1, pp. 233–241.
14. **Raghavan S., Kaplan G.** Optimum Soft Decision Demodulation for ISI Channels. *IEEE Transactions on Communications*, 1993, vol. 41, no. 1, pp. 83–89.
15. **Shamai (Shitz) S., Raghavan S.** On the Generalized Symmetric Cutoff Rate for Finite-State Channels. *IEEE Transactions on Information Theory*, 1995, vol. 41, no. 9, pp. 1333–1346.
16. **Trofimov A., Chan Keong Sann.** Complexity-Performance Trade-off for Intersymbol Interference Channels — Random Coding Analysis. *IEEE Transactions on Magnetics*, 2010, vol. 46, no. 4, pp. 1077–1091.
17. **Egarmin V. K.** Lower and Upper Bounds on Decoding Error Probability for Discrete Channels. *Problemy peredachi informatsii*, 1969, vol. 5, no. 1, pp. 23–39 (In Russian).
18. **Poltyrev G. Sh., Shekhunova N. A.** Decoding in Discrete Channels with Memory. *Voprosy kibernetiki*, 1977, vol. 34, pp. 130–158 (In Russian).
19. **Gilbert E. N.** Capacity of a Burst-Noise Channel. *Bell System Technical Journal*, 1960, vol. 39, no. 5, pp. 1253–1265.

20. Elliott E. O. Estimates of Error Rates for Codes on Burst-Noise Channels. *Bell System Technical Journal*, 1963, vol. 42, no. 5, pp. 1977–1997.
21. Bahl L. R., Cocke J., Jelinek F., Raviv J. Optimal Decoding of Linear Codes for Minimum Symbol Error Rate. *IEEE Transactions on Information Theory*, 1974, vol. 20, no. 2, pp. 284–287.
22. Benedetto S., Divsalar D., Montorsi G., Pollara F. Soft-output Decoding Algorithms for Continuous Decoding of Parallel Concatenated Convolutional Codes. *Proc. IEEE Intern. Conf. on Communication (ICC 96)*, 1996, vol. 1, pp. 112–117.
23. Gantmacher F. R. *Teoriia matrits* [Theory of Matrices]. Moscow, Nauka Publ., 1988. 552 p. (In Russian).

УДК 621.391

doi:10.15217/issn1684-8853.2018.3.79

Граница случайного кодирования для каналов с памятью — декодирующая функция с частичным перекрытием

Часть 1: Вывод основного выражения

Трофимов А. Н.^а, канд. техн. наук, доцент, andrei.trofimov@vu.spb.ru

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: задача вычисления экспоненты случайного кодирования во всем диапазоне скоростей кода для каналов с конечным числом состояний не решена полностью и остается актуальной. Представляется, что хорошее приближение к оптимальной экспоненте случайного кодирования может быть найдено при использовании несогласованной декодирующей функции. **Цель:** построить экспоненту случайного кодирования, близкую к оптимальной. **Результаты:** представлена новая граница случайного кодирования, применимая для широкого класса каналов, в том числе для тех, для которых полная экспонента случайного кодирования ранее не была построена. Вывод этой границы основан на использовании несогласованной декодирующей функции, которая зависит от двух параметров: длины сегмента выходной последовательности канала W и длины сегмента последовательности на входе канала B . Величины W и B в существенной степени влияют на значения экспоненты случайного кодирования и на сложность ее вычисления.

Ключевые слова — граница случайного кодирования, канал с конечным числом состояний, несогласованное декодирование, теорема Перрона — Фробениуса.

Цитирование: Trofimov A. N. Random Coding Bound for Channels With Memory – Decoding Function with Partial Overlapping. Part 1: Derivation of Main Expression // Информационно-управляющие системы. 2018. № 3. С. 79–88. doi:10.15217/issn1684-8853.2018.3.79

Citation: Trofimov A. N. Random Coding Bound for Channels With Memory – Decoding Function with Partial Overlapping. Part 1. Derivation of Main Expression. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 79–88. doi:10.15217/issn1684-8853.2018.3.79

СЛОТОВЫЙ ALOHA С ИТЕРАЦИОННОЙ ПРОЦЕДУРОЙ РАЗРЕШЕНИЯ КОЛЛИЗИЙ. СТАБИЛЬНОСТЬ И НЕСТАБИЛЬНОСТЬ

Н. В. Матвеев^а, ассистент, *n.matveev@vu.spb.ru*

А. М. Тюрликов^а, доктор техн. наук, профессор, *turlikov@vu.spb.ru*

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Постановка проблемы: в настоящее время в сотовых сетях нового поколения рассматриваются сценарии крупномасштабных систем межмашинной связи (*Massive Machine Type Communication*) и класс алгоритмов случайного множественного доступа слотовый ALOHA с итерационной процедурой разрешения коллизий (*Coded Random Access*). Алгоритмы этого класса позволяют поддерживать большое число устройств, но являются нестабильными. Нестабильность приводит к увеличению времени доставки сообщения от абонента до базовой станции в процессе работы крупномасштабных систем межмашинной связи. **Цель:** обоснование нестабильности класса алгоритмов слотового ALOHA с итерационной процедурой разрешения коллизий при любой интенсивности входного потока; предложение метода его стабилизации; определение интенсивности, до которой система будет стабильна. **Результаты:** введена модель системы случайного множественного доступа для класса алгоритмов *Coded Random Access* и пуассоновского входного потока. Функционирование модели описано с помощью марковской цепи со счетным числом состояний. Доказано, что марковская цепь является невозвратной при любой отличной от нуля интенсивности входного потока. Таким образом доказана нестабильность системы множественного доступа для любого алгоритма из класса *Coded Random Access* и предложена модификация этих алгоритмов. Функционирование модели для предложенной модификации описано с помощью двумерной марковской цепи со счетным числом состояний. Если интенсивность входного потока не превышает некоторое предельное значение, то двумерная марковская цепь является эргодической. Это предполагает, что предложенная модификация алгоритмов обеспечивает стабильную работу системы. Для любого алгоритма из рассматриваемого класса предложен способ определения численного значения предельной интенсивности входного потока, до которого система будет стабильна. **Практическая значимость:** предложенная модификация алгоритмов может быть использована при разработке протоколов, ориентированных на сценарий с большим числом устройств, низкой интенсивностью появления сообщения в расчете на одно устройство и большой входной суммарной интенсивностью в расчете на систему в целом.

Ключевые слова — уровень доступа к среде, слотовый ALOHA, крупномасштабные системы межмашинной связи, последовательное погашение интерференции, случайный кодовый доступ, 5G, *Irregular Repetition Slotted ALOHA*.

Цитирование: Матвеев Н. В., Тюрликов А. М. Слотовый ALOHA с итерационной процедурой разрешения коллизий. Стабильность и нестабильность // Информационно-управляющие системы. 2018. № 3. С. 89–97. doi:10.15217/issn1684-8853.2018.3.89

Citation: Matveev N. V., Turlikov A. M. Slotted ALOHA with Iterative Procedure for Resolving Collisions. Stability and Non-Stability. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 89–97 (In Russian). doi:10.15217/issn1684-8853.2018.3.89

Введение

В настоящее время разрабатывается новый стандарт мобильной связи 5G. Идея стандарта пятого поколения неразрывно связана с парадигмой *massive Machine Type Communication* (mMTC), которая характеризуется полностью автоматизированными генерацией, изменением, передачей и обработкой данных электронными устройствами с минимальным вмешательством человека. Предполагается, что количество mMTC-устройств к 2020 г. составит порядка одного миллиона на квадратный километр, и каждое устройство нерегулярно будет передавать небольшой объем данных [1]. Текущая организация сетей не позволит справиться с таким количеством устройств. Очевидно, что необходима разработка алгоритмов повышения скорости передачи данных. Так, в рамках проекта

METIS показано, что для сценария множественного доступа некоординированный доступ является эффективным в случае возникновения большого количества абонентов, что характерно для mMTC [2]. А одной из возможных технологий множественного доступа, позволяющих увеличить скорость передачи данных, является *Coded Random Access* (CRA), или слотовый ALOHA (SA) с итерационной процедурой разрешения коллизий. CRA — это технология случайного множественного доступа с использованием алгоритма последовательного погашения интерференции (SIC), в которой теория и инструменты исправления стираний в системах с LDPC применены для увеличения производительности SA [2–4]. Алгоритмы, основанные на данном подходе, описываются в работах [2–4]. Самым эффективным по средней скорости передачи данных алгоритмом CRA является *Irregular Repetition*

Slotted ALOHA (IRSA) [3]. Несмотря на эффективность, прямое использование технологии CRA невозможно из-за ее нестабильности при случайном входном потоке, что приводит к увеличению задержки в системе [5–8].

Основной целью данной работы является доказательство нестабильности алгоритмов, предложенных в работах [2–4] в сценарии со случайным входным потоком, разработка метода стабилизации и доказательство его работоспособности.

Модель системы

Будем рассматривать модель системы, базирующуюся на следующих допущениях.

Допущение 1: в системе имеется множество абонентов и базовая станция (БС). Каждый абонент генерирует сообщение и по общему каналу множественного доступа передает на БС. БС по обратному широкополосному каналу передает управляющую информацию абонентам, БС точно знает число активных абонентов в системе.

Допущение 2: все время разделено на равные отрезки — фреймы. Каждый фрейм состоит из n слотов. Количество слотов в каждом фрейме — постоянная величина. Длительность слота равна времени передачи сообщения. Границы фреймов и слотов точно известны всем абонентам. Абонент может передавать сообщение только в соответствии с некоторым алгоритмом A из класса алгоритмов случайного множественного CRA (класс алгоритмов будет определен далее). Алгоритм A известен каждому абоненту.

Допущение 3: абонент в соответствии с алгоритмом A передает одну или несколько копий сообщения, каждое сообщение содержит номера слотов, в которых передавались копии. На БС работает процедура погашения интерференции.

Допущение 4: в каждом слоте может произойти одно из трех событий:

- успех — когда в слоте передавал только один абонент;
- пусто — когда в слоте ни один абонент не передавал;
- конфликт — когда в слоте передавали два и более абонентов.

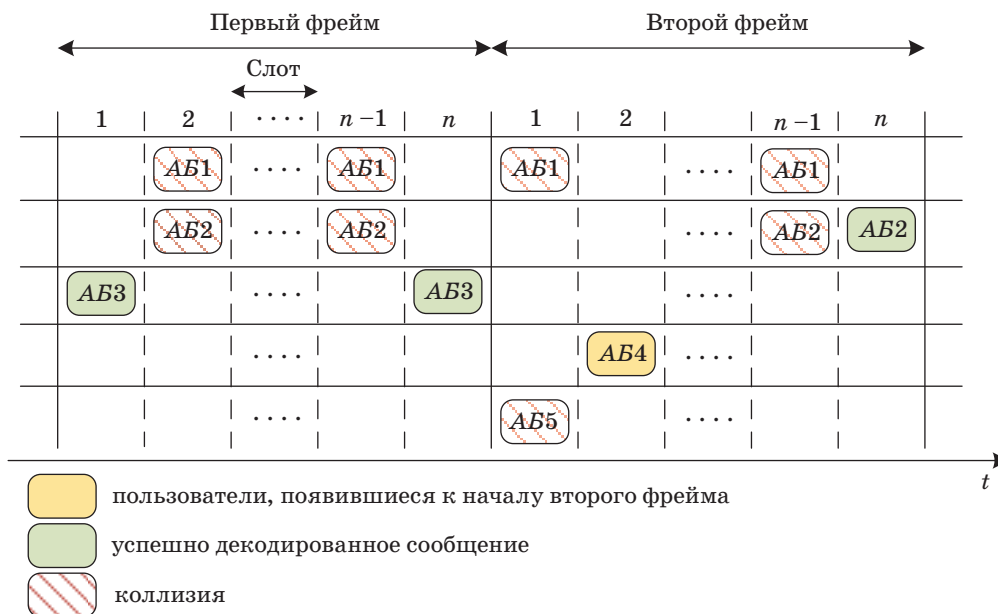
В системе возможно разрешение конфликта. Для этого используется процедура погашения интерференции, которая также называется итерационной процедурой погашения интерференции.

Допущение 5: в каждом фрейме количество пришедших абонентов фиксировано. Абоненты пытаются передать сообщение и затем уходят из системы.

Допущения 1–5 взяты из статей [3, 4], в работах рассматривалось повышение эффективности алгоритмов SA.

Алгоритм последовательного погашения интерференции

Для разрешения некоторых конфликтов используется процедура погашения интерференции. Покажем пример работы процедуры SIC на рис. 1.



■ *Рис. 1.* Пример работы алгоритма CRA: AB — абонент
 ■ *Fig. 1.* Example of operation of the CRA algorithm: AB — subscriber

Два сообщения от АБЗ были приняты успешно в первом фрейме. Однако АБ1 и АБ2 выбрали для передачи одни и те же слоты. В связи с этим БС не удалось декодировать сообщения, эти абоненты будут передавать в следующем фрейме. Во втором фрейме вторая копия сообщения от АБ2 была принята успешно. Зная номер слота, в котором АБ2 передавал вторую копию, можно восстановить сообщения от АБ1. Повторяя данную процедуру, можно восстановить сообщение от АБ5.

Все алгоритмы CRA можно разделить по количеству копий на два класса:

— *регулярные* — когда количество копий одного сообщения в фрейме от каждого абонента фиксировано, например:

- всегда передается одна копия — алгоритм SA;

- всегда передаются две копии — алгоритм Contention Resolution Diversity Slotted ALOHA (CDRSA);

— *нерегулярные* — когда количество копий сообщения в фрейме выбирается случайно из фиксированного набора копий, например алгоритм IRSA.

В общем виде любой алгоритм А можно записать с помощью полинома

$$\Pi_A = \sum_{i=0}^k \pi_i x^i,$$

где π_i — вероятность передачи i копий; k — максимальное число копий.

Соответственно, представим перечисленные алгоритмы:

- алгоритм SA $\Pi_{SA} = 1x^1$;

- алгоритм CDRSA $\Pi_{CDRSA} = 1x^2$;

- алгоритм IRSA $\Pi_{IRSA} = 0,5x^2 + 0,28x^3 + 0,22x^8$.

В работах [3, 4] повышалась такая характеристика алгоритма SA, как максимизация числа успешных переданных сообщений. В рамках модели для алгоритма случайного множественного доступа А и для фиксированного количества слотов в фрейме, равного n , для фиксированного числа абонентов в начале фрейма m можно ввести функцию

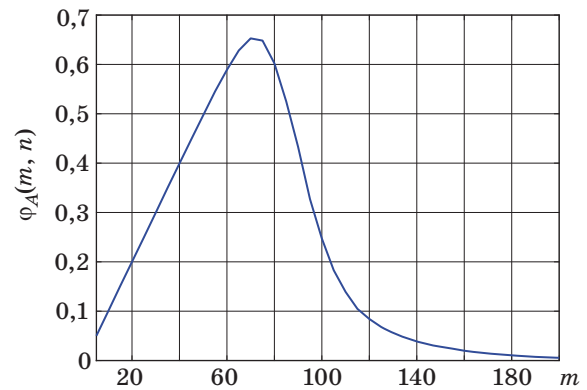
$$\varphi_A(m, n) = \frac{E[M_t | m]}{n}, \quad (1)$$

где M_t — число успешно переданных сообщений.

Для выражения (1) можно ввести следующую лемму.

Лемма 1. Функция $\varphi_A(m, n)$ является унимодальной и $\lim_{m \rightarrow \infty} \varphi_A(m, n) = 0$.

Справедливость леммы следует непосредственно из определения функции $\varphi_A(m, n)$.



■ **Рис. 2.** Зависимость функции $\varphi_A(m, n)$ от m при $n = 100$

■ **Fig. 2.** The function $\varphi_A(m, n)$ versus m for $n = 100$

Наглядный вид функции $\varphi_A(m, n)$ из леммы 1 представлен на рис. 2. При этом в качестве алгоритма А выбран алгоритм IRSA [3] из класса CRA. Далее в статье все представленные графики и численные расчеты приведены для этого алгоритма и числа слотов во фрейме $n = 100$.

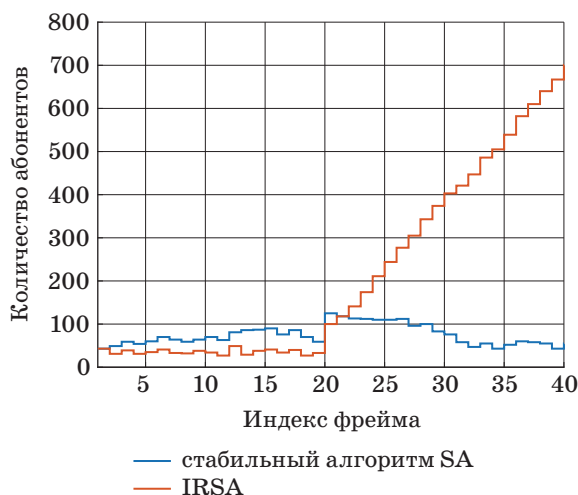
Нестабильность класса алгоритмов CRA

Нужно отметить, что изначальная модель системы не подходит для моделирования систем mMTC, вследствие допущения 5. Изменим его, чтобы приблизить модель к реальности.

Допущение 5: в каждом слоте количество пришедших абонентов распределено по закону Пуассона с параметром λ . Новые абоненты, которые появились в фрейме t , пытаются передать сообщение в фрейме $t + 1$ вместе с абонентами, которым не удалось передать свое сообщение в предыдущих фреймах.

Недостатком технологии CRA является увеличение задержки в системе при случайном входном потоке, что доказано в работе [6]. Так как число абонентов, поступающих в систему в каждом фрейме, случайно, в один момент времени в систему может поступить очень большое количество сообщений. Отличием стабильной системы от нестабильной является то, что в этом случае абоненты все равно будут покидать стабильную систему, и если некоторое время выходной поток будет превышать входной, число абонентов в системе будет стабильно уменьшаться, в то же время в нестабильной системе абоненты будут накапливаться, и в итоге задержка станет бесконечно большой.

Для иллюстрации нестабильности версии алгоритма IRSA проведено моделирование системы при низкой интенсивности входного потока и искусственно введено в систему дополнительно не-



■ **Рис. 3.** Пример нестабильной работы алгоритма IRSA и стабильной работы SA с интенсивностью входного потока $\lambda = 0,34$ и $n = 100$

■ **Fig. 3.** Example of unstable operation of IRSA and stable operation of ALOHA with average arrival rate $\lambda = 0.34$ and $n = 100$

которое число абонентов, которое можно оценить как достаточное для перегрузки системы, этот эффект продемонстрирован на рис. 3. Такое событие может возникнуть, даже если не вмешиваться в работу системы, однако, так как вероятность такого события достаточно маленькая, то для проявления такого эффекта нужно достаточно длительное время моделировать систему.

Описание функционирования модели с помощью случайных процессов

В рамках описанной модели процесс появления и ухода сообщений из системы может быть представлен в виде следующего рекуррентного соотношения:

$$L_{t+1} = L_t - M_t + V_t, \tag{2}$$

где L_t — число абонентов в фрейме t ; M_t — количество абонентов, успешно передавших свое сообщение в фрейме t ; V_t — количество абонентов, появившихся в системе за время фрейма t .

Последовательность дискретных случайных величин L_t образует цепь Маркова со счетным числом состояний. Из допущений принятой модели следует, что данная цепь является однородной, неприводимой, апериодической. В зависимости от выбранной стратегии и интенсивности входного потока марковская цепь L_t может быть невозвратной или эргодической. Для определения условий невозвратности и эргодичности будем использовать критерий Фостера [9]. Дадим формулировку критерия Фостера в форме, представленной в работе [10].

Пусть L_t для $t = 0, 1, \dots$ — однородная неприводимая апериодическая цепь Маркова, принимающая значения из счетного множества $\{0, 1, \dots\}$. Тогда:

— если существуют такие числа $\delta > 0$ и $K > 0$, что для всех $l > K$ справедливо

$$E[L_{t+1} - L_t | L_t = l] < -\delta \tag{3}$$

и для $l \leq K$ справедливо

$$E[L_{t+1} | L_t = l] < \infty, \tag{4}$$

то L_t — эргодическая цепь;

— если существует такое число $K > 0$, что для всех $l > K$

$$E[L_{t+1} - L_t | L_t = l] < 0, \tag{5}$$

то L_t — возвратная цепь;

— если существуют такие числа $\delta > 0$ и $K > 0$, что для всех $l > K$ справедливо

$$E[L_{t+1} - L_t | L_t = l] \geq \delta, \tag{6}$$

то L_t — невозвратная цепь.

Доказательство нестабильности класса алгоритмов CRA

Основываясь на критерии Фостера, нестабильность класса алгоритмов CRA можно сформулировать в виде следующей теоремы.

Теорема 1. Цепь Маркова L_t , каждое состояние которой равно числу абонентов в системе, является невозвратной при любых значениях входного потока.

Доказательство: Для доказательства возьмем математическое ожидание от рекуррентного соотношения (2):

$$E[L_{t+1}] = E[L_t] - E[M_t] + E[V_t]. \tag{7}$$

Пусть число абонентов в фрейме t будет фиксировано: $L_t = l$, тогда

$$E[M_t | L_t = l] = \varphi_A(l, n)n \text{ и } E[V_t] = \lambda n.$$

Получим выражение

$$E[L_{t+1} - L_t | L_t = l] = -\varphi_A(l, n)n + \lambda n.$$

Поскольку $\lim_{l \rightarrow \infty} \varphi_A(l, n) = 0$ (см. лемму 1), при любых фиксированных λ и n всегда найдутся числа $\delta > 0$ и $K > 0$ такие, что для всех $l > K$

$$\lambda n - \varphi_A(l, n)n \geq \delta.$$

Согласно критерию Фостера, такая цепь является невозвратной. Из этого следует, что данная

система является нестабильной при любом алгоритме A из класса алгоритмов CRA и любой интенсивности λ .

Стабилизация класса алгоритмов CRA

Для достижения стабилизации расширим класс алгоритмов CRA следующим образом: БС устанавливает вероятность передачи сообщения p , каждый абонент в текущем фрейме передает с вероятностью p и не передает с вероятностью $1 - p$, где $p \in [0, 1]$.

Введем вспомогательную модель, оставим допущения 1–4 без изменений, изменим допущение 5 следующим образом:

— число абонентов m в начале каждого фрейма случайно и распределено по пуассоновскому закону, среднее число абонентов $E[m] = g$;

— абоненты, которым не удалось передать в фрейме t , уходят из системы.

Для этой модели введем вспомогательную функцию и лемму.

Для алгоритма случайного множественного доступа A и для фиксированного количества слотов в фрейме, равного n , можно ввести следующую функцию:

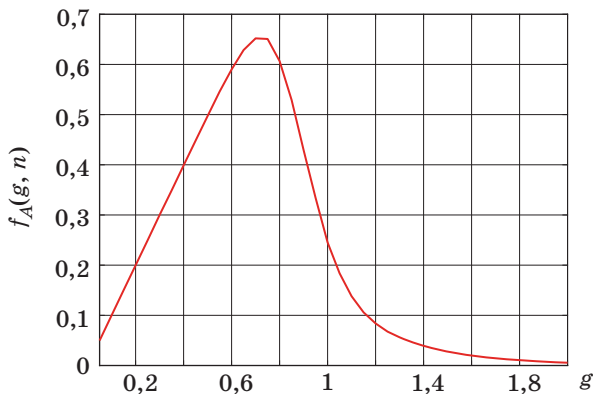
$$f_A(g, n) = \frac{E[M_t; g]}{n} \quad (8)$$

Лемма 2. Функция $f_A(g, n)$ является унимодальной и $\lim_{g \rightarrow \infty} f_A(g, n) = 0$.

Справедливость леммы следует непосредственно из определения функции $f_A(g, n)$.

Наглядный вид функции $f_A(g, n)$ из леммы 2 представлен на рис. 4.

Введем вспомогательную модель, оставим допущения 1–4 без изменений, изменим допущение 5:



■ **Рис. 4.** Зависимость функции $f_A(g, n)$ от g при $n = 100$

■ **Fig. 4.** The function $f_A(g, n)$ versus g for $n = 100$

— зафиксируем число абонентов в начале фрейма $L_t = m$;

— все абоненты передают сообщение с вероятностью p (с вероятностью $1 - p$ не передают);

— абоненты, которым не удалось передать в фрейме t , уходят из системы.

Для этой модели введем вспомогательную функцию и лемму.

Пусть к началу фрейма в системе находится m абонентов, которые используют стратегию A . Для фиксированных n , p , и A введем функцию

$$\gamma_A(p, m, n) = \frac{E[M_t]}{n} \quad (9)$$

Будем задавать вероятность передачи сообщения для фиксированных параметров g, n, m следующим образом:

$$p(g, m, n) = \begin{cases} 1, & \text{если } m/n < g; \\ g \frac{n}{m} & \text{иначе.} \end{cases} \quad (10)$$

Лемма 3. Найдется число $K > 0$ такое, что для всех $m > K$ верно следующее утверждение:

$$\gamma_A(p(g_0, m, n), m, n) \geq f_A(g_0, n),$$

где $g_0 = \max[f_A(g, n)]$ (функция $f_A(g, n)$ определена в лемме 2).

Для наглядного примера вида функции $\gamma_A(p(g_0, m, n), m, n)$ из леммы 3 использован метод Монте-Карло. Результат моделирования представлен на рис. 5 для $n = 100$.

Тогда стабильность расширенного класса алгоритмов CRA можно сформулировать в виде следующей теоремы.

Теорема 2. Цепь Маркова L_t , каждое состояние которой равно числу абонентов в системе, является эргодической при значениях входного потока $0 < \lambda < f_A(g_0, n)$, где g_0 — точка максимума для функции $f_A(g_0, n)$.

Доказательство: Для доказательства возьмем математическое ожидание (7).

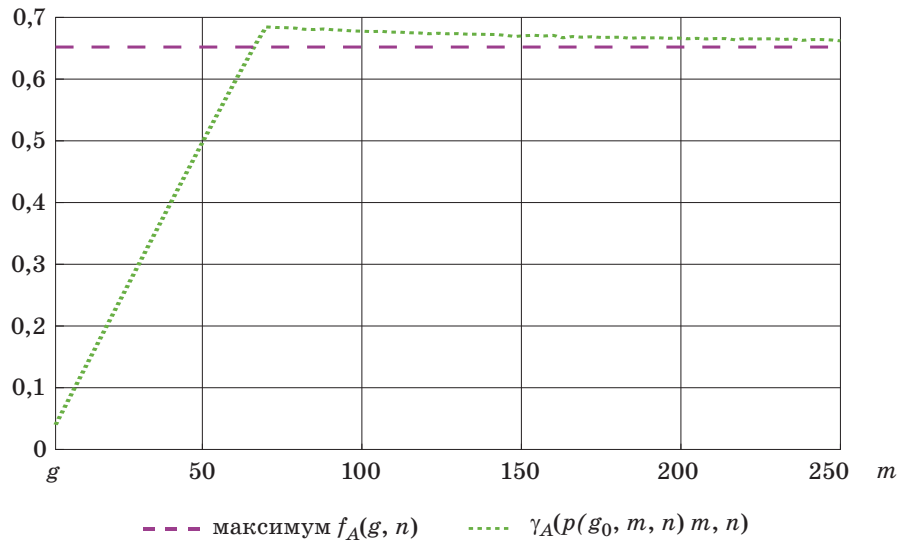
Пусть число абонентов в фрейме t будет фиксировано: $L_t = l$, тогда

$$E[M_t | L_t = l] = \gamma_A(p(g_0), l, n) \text{ и } E[V_t] = \lambda n.$$

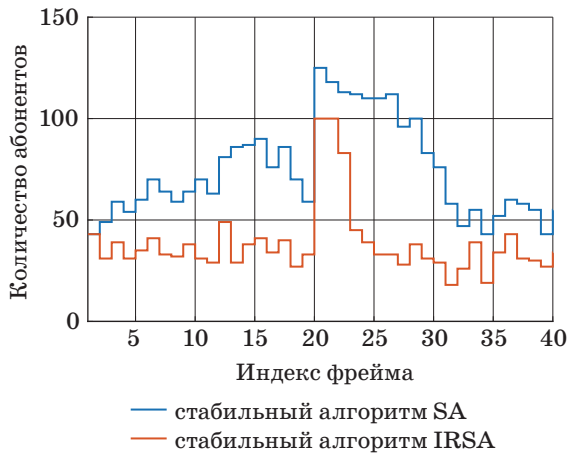
Получим выражение

$$E[L_{t+1} - L_t | L_t = l] = -\gamma_A(p(g_0), l, n)n + \lambda n.$$

Докажем первое условие (3) критерия Фостера для эргодической цепи. Согласно лемме 3, при значениях входного потока $0 < \lambda < f_A(g_0, n)$ и



■ **Рис. 5.** Зависимость функции $\gamma_A(p(g_0, m, n), m, n)$ от m при $n = 100$
 ■ **Fig. 5.** The function $\gamma_A(p(g_0, m, n), m, n)$ versus m for $n = 100$



■ **Рис. 6.** Пример стабильной работы алгоритма IRSA с предлагаемой модификацией с интенсивностью входного потока $\lambda = 0,34$ и $n = 100$
 ■ **Fig. 6.** Example stable operation of IRSA with proposed modification with average arrival rate $\lambda = 0.34$ and $n = 100$

любом конечном n всегда найдутся числа $\delta > 0$ и $K > 0$ такие, что для всех $l > K$

$$\lambda n - \gamma_A(p(g_0), l, n)n \leq -\sigma.$$

Докажем второе условие (4) критерия Фостера для эргодической цепи. Вследствие того, что порог K конечен из первого условия критерия Фостера, при значениях входного потока $0 < \lambda < f_A(g_0, n)$ и любом конечном n для всех $l \leq K$ справедливо

$$E[L_{t+1} | L_t = l] < \infty.$$

Из критерия Фостера следует, что данная система является стабильной.

Предложенный способ позволил стабилизировать алгоритм IRSA при интенсивности входного потока $\lambda \in (0; 0,65)$ и значении $n = 100$. Данный подход для любого алгоритма из класса CRA для фиксированного значения n позволяет получить g_0 и значения входного потока, при которых гарантируется стабильность.

На рис. 6 представлена производительность алгоритма IRSA с предложенной модификацией с входными параметрами, аналогичными на рис. 3.

Как видно из графика, после появления 100 абонентов в начале 20-го фрейма IRSA показывает стабильность. Увеличение количества абонентов не наблюдается.

Оценка числа активных абонентов в системе при случайном входном потоке

Примем, что БС не известно точное число активных абонентов. Поэтому БС после принятия сообщений за текущий фрейм вычисляет количество слотов, в которых не было передано сообщений, — N^{empty} , количество слотов, в которых были переданы все сообщения, — $N^{success}$, количество слотов с коллизиями после процедуры SIC — $N^{collision}$. На основании полученных значений для следующего фрейма вычисляется оценка числа активных абонентов. Для этого воспользуемся подходом, предложенным в работе [11], для алгоритма адаптивный АЛОНА и обобщенным в работе [12] на случай, когда абонент передает несколько копий одного сообщения, однако в от-

сутствие на БС итерационной процедуры разрешения коллизии.

Следуя работе [11], составим рекуррентное выражение для оценки числа активных абонентов

$$S_{t+1} = \max(1, S_t + aN_t^{empty} + bN_t^{succes} + cN_t^{collision}),$$

где S_t — оценка числа активных абонентов в фрейме t ; a, b, c — коэффициенты, которые влияют на работу алгоритма и являются константами; N_t^{empty} — количество пустых слотов в фрейме t ; N_t^{succes} — количество слотов, в которых произошел успех в фрейме t ; $N_t^{collision}$ — число слотов с неразрешенными конфликтами в фрейме t .

В рамках описанной модели процесс появления и ухода сообщений из системы может быть представлен в виде следующего рекуррентного соотношения:

$$L_{t+1} = L_t - M_t(S_t, L_t) + V_t. \quad (11)$$

Последовательность пар случайных величин (L_t, S_t) образует двумерную марковскую цепь. Найдем средний снос случайных величин (L_t, S_t)

$$E[L_{t+1} - L_t | L_t = l, S_t = s] = E[M_t(l, s)] - \lambda,$$

$$E[S_{t+1} - S_t | L_t = l, S_t = s] = aE[N^{empty}] + bE[N^{succes}] + c[N^{collision}].$$

Следуя работе [9], рассмотрим предельный переход $\sqrt{l^2 + s^2} \rightarrow \infty$, введем величину $k = l/s$ и найдем граничную вектор-функцию среднего сноса $\mu = (\mu_l(k), \mu_s(k))$

$$\mu_l(k) = E[M_t(k)] - \lambda,$$

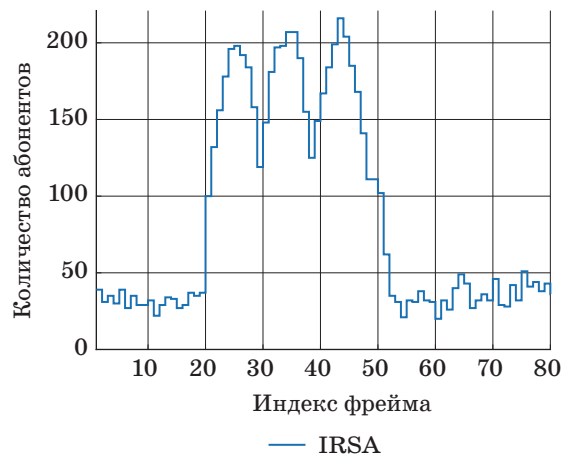
$$\mu_s(k) = aE[N^{empty}] + bE[N^{succes}] + c[N^{collision}].$$

Рассмотрим уравнение

$$\mu_l(k) = k\mu_s(k).$$

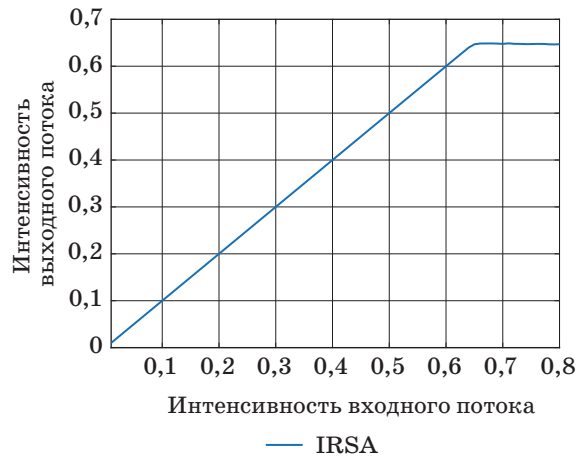
Цепь Маркова (L_t, S_t) будет эргодической, если для любого корня k уравнения имеют место неравенства $\mu_l(k) < 0$ и $\mu_s(k) < 0$ [9]. Всегда возможно подобрать коэффициенты a, b, c и корень k , при которых система будет стабильна до предельной интенсивности входного потока (см. предыдущий раздел).

Для решения уравнения зафиксировали алгоритм IRSA, интенсивность входного потока $\lambda = 0,65$, коэффициенты: $a = -1, b = -1$. Найти с помощью численного метода c и k . В итоге $c = 9, k = 9$.



■ **Рис. 7.** Пример работы алгоритма IRSA при использовании оценки числа активных абонентов в системе с интенсивностью входного потока $\lambda = 0,34$ и $n = 100$

■ **Fig. 7.** Example stable operation of IRSA using the estimation of the number of active users in the system with average arrival rate $\lambda = 0.34$ and $n = 100$



■ **Рис. 8.** Интенсивность выходного потока от входного

■ **Fig. 8.** Intensity of the input stream from the output

Для проверки стабильности алгоритма IRSA построен график (рис. 7), аналогичный графикам на рис. 3 и 6. В фрейме с номером 20 в систему искусственно добавлено 100 абонентов, что привело к бесконечному росту числа абонентов. Из графика видно, что число абонентов не стремится к бесконечности.

Из графика зависимости интенсивности выходного потока от интенсивности входного (рис. 8) видно, что алгоритм стабилен при использовании предложенного алгоритма до предельной интенсивности входного потока.

Заключение

В данной работе была доказана нестабильность класса алгоритмов случайного множественного доступа CRA, предложенного в статьях [3, 4], для сценария со случайным входным потоком при любой интенсивности. Для стабилизации исходных алгоритмов предложено расширение класса алгоритмов CRA за счет введения специальной процедуры управления передачей абонентов. Доказано, что предложенный способ

позволяет стабилизировать алгоритмы CRA до определенной интенсивности входного потока. Для конкретного алгоритма CRA предложен способ получения значения максимальной интенсивности входного потока, до которой система будет стабильна. Для модели, приближенной к реальности, когда БС не известно число активных абонентов в системе, предложена оценка числа активных абонентов в системе, которая гарантирует стабильную работу системы до максимального значения интенсивности входного потока.

Литература

1. Recommendation ITU-R M.2083-0 (09/2015). IMT Vision Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond. https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf (дата обращения: 18.04.2018).
2. ICT-317669-METIS/D2.4. Proposed Solutions for New Radio Access. https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D2.4_v1.pdf (дата обращения: 18.04.2018).
3. Liva G. Graph-based Analysis and Optimization of Contention Resolution Diversity Slotted ALOHA // IEEE Transactions on Communications. 2011. N 59. P. 477–487. doi:10.1109/TCOMM.2010.120710.100054
4. Paolini E., Stefanovic C., Liva G., Popovski P. Coded Random Access: Applying Codes on Graphs to Design Random Access Protocols // IEEE Communications Magazine. 2015. N 53 P. 144–150. doi:10.1109/MCOM.2015.7120031
5. Apanasenko N., Matveev N., Turlikov A. Stability and Delay of Algorithms of Random Access with Successive Interference Cancellation // Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer, Cham, 2017. P. 510–518. doi:10.1007/978-3-319-67380-6_47
6. Апанасенко Н. В., Матвеев Н. В., Тюрликов А. М. Метод стабилизации алгоритмов множественного доступа с последовательным погашением интерференции // Научная сессия ГУАП: сб. докл. 2017. С. 260–266.
7. Meloni A., Murrone M. CRDSA, CRDSA++ and IRSA: Stability and Performance Evaluation // 2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC), Baiona, 2012. P. 220–225. doi:10.1109/ASMS-SPSC.2012.6333080
8. Kissling C. On the Stability of Contention Resolution Diversity Slotted ALOHA (CRDSA) // Global Telecommunications Conference (GLOBECOM 2011), Kathmandu, 2011. P. 1–6. doi:10.1109/GLOCOM.2011.6134329
9. Foster F. G. On the Stochastic Matrices Associated with Certain Queuing Processes // The Annals of Mathematical Statistics. 1953. N 24. P. 355–360. doi:10.1214/aoms/1177728976
10. Цыбаков Б. С., Михайлов В. А. Свободный синхронный доступ пакетов в широкополосный канал с обратной связью // Проблемы передачи информации. 1978. Т. 14. № 4. С. 32–59.
11. Михайлов В. А. Геометрический анализ устойчивости цепей Маркова в R_+^n и его приложение к вычислению пропускной способности адаптивного алгоритма случайного множественного доступа // Проблемы передачи информации. 1988. Т. 24. Вып. 1. С. 61–73.
12. Galinina O., Turlikov A., Andreev S., Koucheryavy Y. Stabilizing Multi-Channel Slotted ALOHA for Machine-Type Communications // 2013 IEEE Intern. Symp. on Information Theory, Istanbul, 2013. P. 2119–2123. doi:10.1109/ISIT.2013.6620600

UDC 004.728.3.057.4

doi:10.15217/issn1684-8853.2018.3.89

Slotted ALOHA with Iterative Procedure for Resolving Collisions. Stability and Non-Stability

Matveev N. V.^a, Assistant Professor, n.matveev@vu.spb.ru

Turlikov A. M.^a, Dr. Sc., Tech., Professor, turlikov@vu.spb.ru

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: Cellular networks of the new generation consider Massive Machine Type Communication scenarios and a class of multiple random access algorithms Slotted ALOHA with coded random access. The algorithms of this class allow you to maintain a

large number of devices, but they are unstable. Their non-stability leads to a longer time of delivering a message from a subscriber to the base station during the operation of large-scale systems of inter-machine communication. **Purpose:** Substantiation of non-stability of Slotted ALOHA algorithms with coded random access at any intensity of the input stream; proposal of a method for its stabilization; determination of the intensity at which the system will be stable. **Results:** A model of a random multiple access system is introduced for coded random access and a Poisson input flow. The functioning of the model is described with the use of a Markov chain with a countable number of states. It is proved that the Markov chain is non-returnable for any non-zero intensity of the input stream. Thus, the non-stability of a multiple access system for any coded random access algorithm is proved, and a modification of these algorithms is proposed. The operation of the model for the proposed modification is described using a two-dimensional Markov chain with a countable number of states. If the intensity of the input stream does not exceed a certain value limit, the two-dimensional Markov chain is ergodic. This suggests that the proposed modification of the algorithms ensures stable operation of the system. For any algorithm from the considered class, a method is proposed for determining the numerical value of the limiting intensity of the input stream at which the system is stable. **Practical relevance:** The proposed modification of the algorithms can be used to develop protocols oriented to a scenario with a large number of devices, a low message rate per device and a large total input intensity in the whole system.

Keywords — Medium Access Control, Slotted ALOHA, Massive Machine Type Communication, Successive Interference Cancellation, Coded Random Access, 5G, Irregular Repetition Slotted ALOHA.

Citation: Turlikov A. M., Matveev N. V. Slotted ALOHA with Iterative Procedure for Resolving Collisions. Stability and Non-Stability. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 89–97 (In Russian). doi:10.15217/issn1684-8853.2018.3.89

References

1. *Recommendation ITU-R M.2083-0 (09/2015). IMT Vision — Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond*. Available at: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf (accessed 18 April 2018).
2. *ICT-317669-METIS/D2.4. Proposed Solutions for New Radio Access*. Available at: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D2.4_v1.pdf (accessed 18 April 2018).
3. Liva G. Graph-based Analysis and Optimization of Contention Resolution Diversity Slotted ALOHA. *IEEE Transactions on Communications*, 2011, no. 59, pp. 477–487. doi:10.1109/TCOMM.2010.120710.100054
4. Paolini E., Stefanovic C., Liva G., Popovski P. Coded Random Access: Applying Codes on Graphs to Design Random Access Protocols. *IEEE Communications Magazine*, 2015, no. 53, pp. 144–150. doi:10.1109/MCOM.2015.7120031
5. Apanasenko N., Matveev N., Turlikov A. Stability and Delay of Algorithms of Random Access with Successive Interference Cancellation. *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*, Springer, Cham, 2017, pp. 510–518. doi:10.1007/978-3-319-67380-6_47
6. Apanasenko N. V., Matveev N. V., Turlikov A. M. Method of Stabilization of Random Access Algorithms with Successive Interference Cancellation. *Sbornik dokladov Nauchnoi sessii GUAP*, 2017, pp. 260–266 (In Russian).
7. Meloni A., Murrioni M. CRDSA, CRDSA++ and IRSA: Stability and Performance Evaluation. *2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC)*, Baiona, 2012, pp. 220–225. doi:10.1109/ASMS-SPSC.2012.6333080
8. Kissling C. On the Stability of Contention Resolution Diversity Slotted ALOHA (CRDSA). *Global Telecommunications Conf. (GLOBECOM 2011)*, Kathmandu, 2011, pp. 1–6. doi:10.1109/GLOCOM.2011.6134329
9. Foster F. G. On the Stochastic Matrices Associated with Certain Queuing Processes. *The Annals of Mathematical Statistics*, 1953, no. 24, pp. 355–360. doi:10.1214/aoms/1177728976
10. Tsybakov B. S., Mikhailov V. A. Free Synchronous Packet Access in a Broadcast Channel with Feedback. *Problemy peredachi informatsii*, 1978, vol. 14, no. 4, pp. 32–59 (In Russian).
11. Mikhailov V. A. Geometrical Analysis of the Stability of Markov Chains in and its Application to Throughput Evaluation of the Adaptive Random Multiple Access Algorithm. *Problemy peredachi informatsii* [Problems of Information Transmission], 1988, vol. 24, iss. 1, pp. 47–56 (In Russian).
12. Galina O., Turlikov A., Andreev S., Koucheryavy Y. Stabilizing Multi-Channel Slotted ALOHA for Machine-Type Communications. *2013 IEEE Intern. Symp. on Information Theory*, Istanbul, 2013. pp. 2119–2123. doi:10.1109/ISIT.2013.6620600

МЕТОД АНАЛИЗА СНИМКОВ КОМПЬЮТЕРНОЙ ТОМОГРАФИИ НА ОСНОВЕ ПОЭТАПНОЙ БИНАРИЗАЦИИ ИЗОБРАЖЕНИЙ И ЕГО ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Н. И. Лиманова^а, доктор техн. наук, доцент, nataliya.i.limanova@gmail.com

С. Г. Атаев^а, ассистент, sergeyataev@gmail.com

^аПоволжский государственный университет телекоммуникации и информатики, Льва Толстого ул., 23, Самара, 443010, РФ

Введение: снимки внутренних структур тела, получаемые методами компьютерной томографии, содержат в себе большие объемы информации, требуют комплексного всестороннего анализа, занимающего длительное время и представляющего сложность для врача, проводящего диагностику. Повысить точность постановки диагноза позволяет применение контрастного вещества, но данный способ диагностического исследования дороже и имеет ряд противопоказаний. **Цель:** разработка метода автоматического определения параметров объектов на снимках компьютерной томографии, предоставляющего диагностирующим врачам требуемую информацию. **Результаты:** предложен метод параметризации объектов, расположенных на слоях снимка компьютерной томографии, выполняющий их поочередный анализ и определяющий область снимка, относящуюся к исследуемому объекту. В процессе последующего исследования объект рассматривается как совокупность вокселей, отображающих среднюю плотность ткани в соответствующей области объекта. Далее вычисляются площади объекта на каждом срезе, а затем его суммарный объем. Определение выборочной несмещенной дисперсии для значений плотности объекта, а также гистограмма этих значений дают возможность диагностирующему врачу делать выводы о свойствах данного объекта и оценивать вероятность наличия патологических образований. Фильтрация вокселей объекта по значению плотности позволяет локализовать области объекта с определенными свойствами и оценить степень симметричности новообразований, свидетельствующую о вероятности их патологической природы. **Практическая значимость:** разработанный метод и его программная реализация позволяют, во-первых, частично автоматизировать процесс анализа снимков компьютерной томографии, предоставляя специалистам значения параметров объекта, выбранного на снимке: послойных площадей, объема, распределения значений плотности внутри объекта, выборочной несмещенной дисперсии для распределения значений плотности объекта, оценки симметричности новообразований внутри объекта; во-вторых, проводить локализацию тканей определенной плотности и выявлять патологические образования без использования контрастирования.

Ключевые слова — компьютерная томография, анализ изображений, бинаризация изображений, программное обеспечение, системы компьютерной диагностики.

Цитирование: Лиманова Н. И., Атаев С. Г. Метод анализа снимков компьютерной томографии на основе поэтапной бинаризации изображений и его программная реализация // Информационно-управляющие системы. 2018. № 3. С. 98–106. doi:10.15217/issn1684-8853.2018.3.98

Citation: Limanova N. I., Ataev S. G. Computer Aided Tomography Picture Analysis on the base of Stage-By-Stage Binarization and its Software Implementation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 98–106 (In Russian). doi:10.15217/issn1684-8853.2018.3.98

Введение

Системы компьютерной диагностики, представляющие специалистам помощь в анализе и интерпретации результатов медицинской визуализации — активно развивающаяся область исследований в сфере алгоритмизации и разработки программного обеспечения [1, 2]. В настоящее время компьютерная томография (КТ) используется в медицине для решения широкого спектра диагностических задач [3]. Снимки внутренних структур тела, получаемые методами КТ, содержат в себе большие объемы информации и требуют комплексного всестороннего анализа, занимающего длительное время и представляющего для врача, проводящего диагностику, определенную сложность. Системы компьютерной диагностики чаще всего не способны заменить диагностирующих медицинских специалистов, однако их пред-

назначение заключается в представлении информации со снимка в удобном для восприятия виде и возможности извлечь из него дополнительную информацию, полезную для постановки диагноза.

Рассмотрим одну из областей применения КТ — исследование верхнечелюстных пазух. На снимке КТ эти полости находятся на нескольких соседних слоях, и для определения их параметров требуется высокая степень внимательности со стороны диагностирующего специалиста, а также применение сложных косвенных методов, не дающих гарантированно точных результатов. Повысить точность позволяет контрастное вещество. Однако данный способ диагностического исследования имеет ряд противопоказаний и более высокую стоимость. К измеряемым параметрам относятся: объем исследуемой полости, координаты ее геометрического центра, выборочная несмещенная дисперсия для значений плотности. Альтернативу контрасти-

рованию может составить поэтапная бинаризация изображений исследуемых объектов. В данной работе предлагается автоматизировать определение подобных характеристик объекта путем поиска границ исследуемой полости (а следовательно, множества пикселей, принадлежащих проекции полости на слой КТ) и последующего анализа значений яркости множества пикселей (проецирующих плотность объекта на соответствующие слои томографического снимка).

Изобретение многослойной КТ расширило диагностические возможности медицинских специалистов, позволив им получать снимки внутренней структуры исследуемых анатомических объектов. Каждый слой снимка КТ проецирует на себя среднюю плотность тканей в соответствующем пространстве. Благодаря этому набор слоев многослойного томографического снимка позволяет получить исчерпывающую информацию о внутреннем строении объекта в пространстве и провести ее автоматизированный анализ.

На сегодняшний момент существует ряд методов улучшения визуализации снимков КТ [4, 5] и реконструкции 3D-моделей изучаемого объекта [6–8]. Существующие подходы к диагностическому процессу позволяют упростить чтение снимков, но не приводят к автоматизированному подсчету параметров исследуемых объектов. Другие известные методы интерпретируют каждый слой снимка отдельно. Примером такого подхода является классификация легочных узлов на доброкачественные и злокачественные [9–11]. Это оправдано в тех случаях, когда решается задача классификации объектов, т. е. определяются их качественные признаки принадлежности к какому-либо типу объектов [12, 13].

Для решения задачи определения параметров объекта необходима комплексная информация о его строении, для получения которой требуется анализировать в совокупности все слои, содержащие изображение данного объекта. При автоматизированной обработке снимков легких и мозга [14–16] для их исследования применяются методы сегментации [17], пороговой обработки (разграничения снимка по принципу яркости пикселей) и анализа гистограмм [18–19]. Однако существует класс задач параметризации небольших объектов, которые, в отличие от легких и мозга, занимают только малую часть снимка. Для подобных объектов указанные выше способы автоматизированного анализа являются неприменимыми.

Метод параметризации объектов КТ

В работе предлагается решение задачи параметризации исследуемых малых объектов путем поиска границ выбранного на одном из слоев

снимка объекта КТ на основе поэтапной бинаризации слоев с последующим анализом множеств пикселей, проецирующих плотность объекта на слои томографического снимка и, следовательно, характеризующих строение этого объекта. Далее рассмотрим особенности хранения информации на слоях снимков КТ, предоставляющие такую возможность.

Результатом томографического исследования является набор слоев, полученных с определенным интервалом и проецирующих на себя внутреннее строение исследуемых анатомических структур. Каждый слой проекции томограммы дает точное значение плотности любого из пикселей на снимке, которые впоследствии отображаются как оттенки серого. Чем светлее оттенок серого, тем плотнее ткань в пределах воксела (воксел — аналог пиксела для трехмерного пространства: область, проецирующаяся на пиксел томографического среза) [20]. Таким образом, наличие информации о границах исследуемого объекта дает исчерпывающие сведения о свойствах объекта в каждой точке пространства и предоставляет возможность проводить его дальнейший анализ и исследование (а также позволяет частично автоматизировать этот процесс).

В связи с тем, что разрешающая способность мониторов не может отразить весь диапазон градаций плотностей («шкала Хаунсфилда»), каждый томографический снимок имеет свой собственный диапазон плотностей, отображаемых с помощью оттенков, находящихся между черным и белым цветами. Данный диапазон зависит от специфики снимаемых объектов, и точная информация о плотности тканей за пределами данного диапазона теряется [21]. Границы диапазона задаются двумя значениями: центром диапазона и его шириной. Чисто белому цвету пиксела соответствует ткань, превосходящая по плотности диапазон отображения, чисто черному цвету — ткань, менее плотная, чем диапазон отображения. Промежуточные оттенки серого цвета пиксела проецируют значения плотности из интервала отображения и отражают точную величину плотности ткани, находящейся в конкретной точке пространства. Пример слоя томографического снимка представлен на рис. 1. Значения W: 90 и C: 40 в нижнем левом углу рисунка свидетельствуют о центре шкалы Хаунсфилда 40HU (Hounsfield unit — единица Хаунсфилда) и ширине шкалы 90HU.

Рассмотрим разработанный метод программной параметризации объектов на примере исследования верхнечелюстных пазух. Параметризация какого-либо объекта, отображенного на снимке, проводится в два этапа: сначала выполняется сегментация слоев снимка КТ в целях определения границ исследуемого объекта, затем

совокупность пикселей, относящихся к исследуемому объекту, анализируется для определения его параметров.

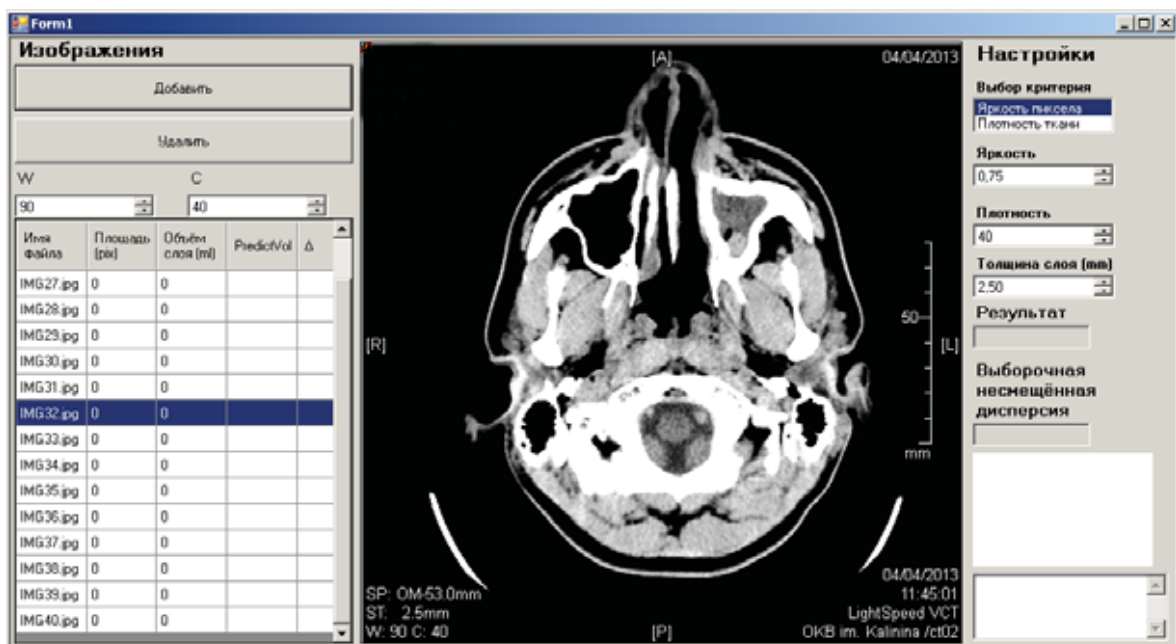
Функционирование предложенного алгоритма сегментации снимков и автоматизированного определения границ исследуемого объекта начинается с загрузки в программу пользователем

упорядоченных слоев томографического снимка и выбора объекта, например, путем указания на слое любой точки, принадлежащей исследуемому объекту. В программу необходимо ввести также значение толщины среза слоев и указать либо граничное значение яркости пикселей, либо ограничение по плотности для пикселей объекта, которое послужит критерием проверки принадлежности пикселей выбранному объекту. На рис. 2 представлен интерфейс программного обеспечения после загрузки в него слоев томографического снимка.

К множеству пикселей на данном слое, относящихся к целевому объекту, добавляется стартовый пиксел. В дальнейшем каждый из пикселей на границе целевой области сравнивается с максимальным значением яркости пикселей исследуемого объекта. Если его яркость меньше заданного значения, он добавляется к множеству пикселей, принадлежащих объекту, после чего происходит рассмотрение смежных с ним пикселей. Иначе он добавляется к множеству пикселей, относящихся к границе объекта, и не участвует в дальнейшем расширении области объекта. На слоях, где граница объекта имеет замкнутый характер, этих действий достаточно для идентификации объекта на снимке. Если граница разомкнута, необходимо дотраивать ее искусственно. Требуется выбрать стартовый слой таким образом, чтобы отображаемый на нем объект имел замкнутые границы.



■ *Рис. 1.* Пример слоя снимка КТ
 ■ *Fig. 1.* CT layer example

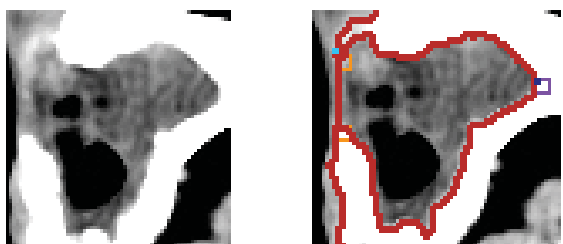


■ *Рис. 2.* Внешний вид интерфейса разработанного программного обеспечения после загрузки слоев томографического снимка
 ■ *Fig. 2.* Developed software graphic user interface with CT layers loaded

Поиск пикселей в рамках стартового слоя заканчивается в момент, когда все граничные пиксели рассмотрены и дальнейшее расширение исследуемой области невозможно. Далее следует проверка возможности расширения исследуемой области на соседние слои до тех пор, пока это возможно. Для этого необходимо повторять поиск пикселей для каждого нового соседнего слоя. На каждом из этих слоев поиск начинается с заполнения множества пикселей в соответствии с наличием на соседних слоях граничных пикселей, относящихся к исследуемому объекту, и проходит аналогично поиску на стартовом слое. В тот момент, когда на рассматриваемом слое не оказывается пикселей, удовлетворяющих критерию принадлежности, расширение области на соседние слои в данном направлении завершается. Как только завершается расширение области на соседние слои в обоих направлениях, поиск новых пикселей объекта заканчивается.

В силу анатомических особенностей исследуемого типа объектов (верхнечелюстных пазух, соединяемых соустьем с носовой полостью) замкнутая граница, очевидно выраженная разницей яркостей, встречается не на всех срезах томографического снимка. Для корректного определения области исследуемого объекта на таких срезах предусмотрен метод распознавания таких слоев и последующего замыкания границы объекта. Слева на рис. 3 показан фрагмент исходного слоя, в левой части которого есть разрыв на границе объекта, справа показан контур границы, определенный и достроенный описываемым алгоритмом. Замыкание границы происходит полностью в автоматическом режиме и не требует вмешательства врача в этот процесс. Анализ отдельных слоев также выполняется программно.

Рассмотрим подробнее предлагаемый метод. При работе алгоритма на слоях снимка, где объект обладает замкнутой границей, расширение области исследуемого объекта в какой-то момент становится невозможным. Это является крите-



■ **Рис. 3.** Пример томографического слоя с достроенной границей исследуемого объекта

■ **Fig. 3.** CT layer example with automated borderline closing

рием прекращения исследования текущего слоя и перехода к следующему. Однако в случае отсутствия у объекта четких однозначных границ его область, определяемая данным алгоритмом, будет расширяться бесконечно. Для того чтобы избежать этого, вводится ограничение на максимальную площадь исследуемого объекта на слое, определяемое перед началом анализа этого слоя. В качестве вспомогательной информации на этой стадии работы алгоритма используется экстраполяция значений площади на нескольких предыдущих слоях с целью предсказать площадь объекта на текущем слое. Максимальная площадь, после достижения которой расширение области объекта прекращается, определяется путем перемножения предсказанной площади исследуемого объекта на заранее заданную константу, которую назовем коэффициентом ограничения площади (β). Данная константа позволяет превысить предсказанное значение площади исследуемого объекта, но не позволяет расширяться бесконечно, что привело бы к зависанию программной реализации алгоритма.

Поскольку значение площади объекта, полученное при помощи экстраполяции, в дальнейшем будет уточняться, для получения этих значений было решено использовать полином Лагранжа третьей степени. Более сложные расчеты в данном методе решения задачи неуместны, потому что итоговый результат почти не зависит от конкретного значения, полученного на этой стадии.

Превышение площадью объекта данного ограничения в процессе ее расширения свидетельствует о том, что объект не имеет замкнутой границы на данном слое. После этого область объекта на слое обнуляется, и алгоритм производит поиск оптимальных точек для замыкания между ними границы. Для этого используется множество точек, граничных области исследуемого объекта, обладающих при этом яркостью большей, чем яркость пикселей в пределах объекта. Проводится поиск пары точек, для которых определяется значение весовой функции, в числителе которой — длина границы между двумя этими точками, а в знаменателе — геометрическое расстояние между проверяемыми точками:

$$w(x_1, y_2, x_1, y_2) = \frac{p^\alpha}{d},$$

где p — длина границы между точками; α — варьируемый коэффициент; d — геометрическое расстояние между проверяемыми точками.

Очевидно, что благодаря форме исследуемого объекта (верхнечелюстной пазухи) для пикселей границы, расположенных в соустье, будет наблюдаться рост показателя p и локальный минимум

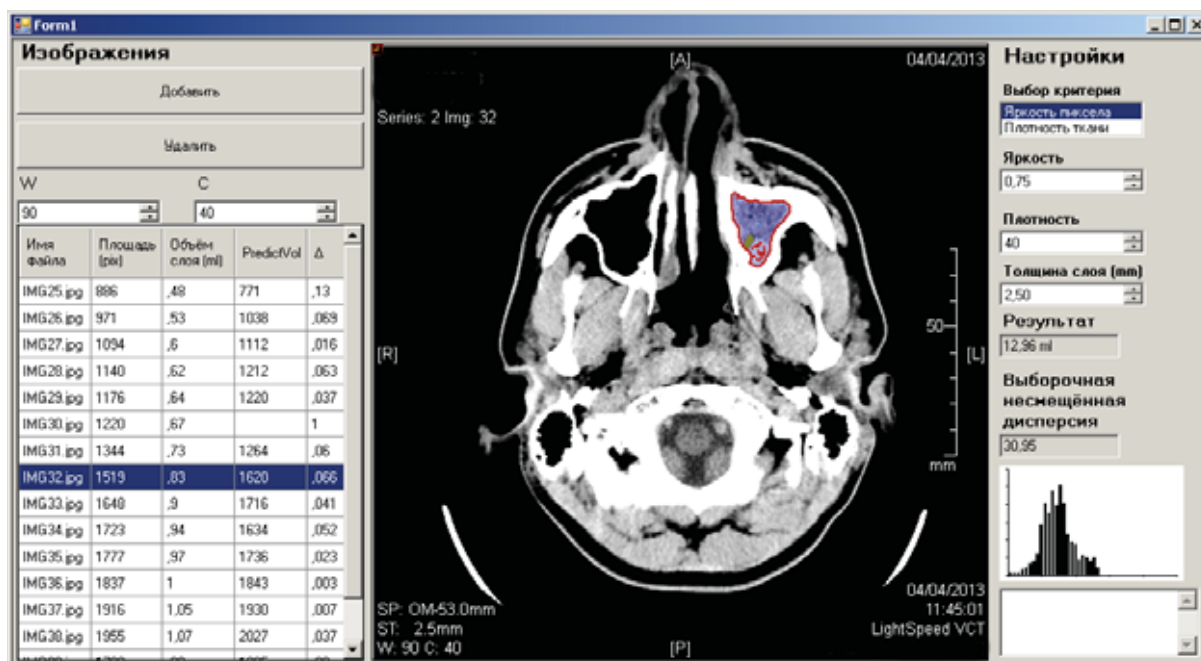
показателя d . Разработанный метод достройки границы обладает эвристическим характером и допускает колебания результатов. Для того чтобы его результаты были стабильными, числитель возводится в степень, которую назовем степенным весовым коэффициентом (α). Варьирование этим значением позволяет получить несколько потенциальных пар точек, которые могут быть использованы для замыкания границы. После этого площадь объекта определяется несколько раз для каждого из результатов. Увеличение значения степенного весового коэффициента α приводит к отдалению границы от центра объекта, позволяя таким образом выбрать из нескольких вариантов границ объекта наиболее подходящей. Окончательным вариантом конфигурации границы объекта считается тот, который приводит к наименьшему отклонению площади от предсказанного экстраполяцией значения.

На стадии первоначального поиска зоны объекта необходимо, чтобы его граница, даже если в ней отсутствует замкнутость, была непрерывна, поэтому для правильной работы метода замыкания границы требуется внутреннюю область объекта исследовать целиком. Следовательно, для того чтобы гарантировать стабильную работу метода замыкания границ, необходимо использовать достаточно большое значение коэффициента ограничения площади β . Слишком большие значения, однако, увеличивают время работы алгоритма и вызывают появление лишних значений во множестве пикселей, соответствующих гра-

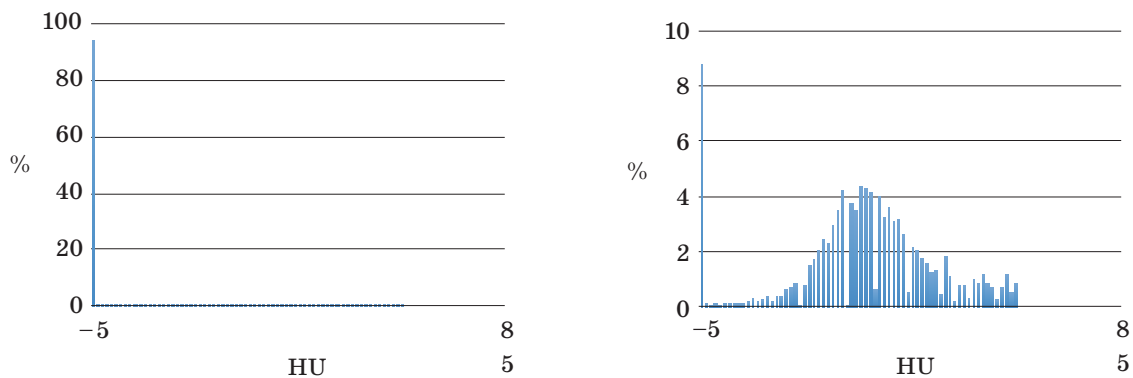
нице объекта, приводящих к возможным аномалиям в его работе. Авторами экспериментально определены оптимальные границы данного коэффициента в диапазоне $1,4 \div 1,6$.

Последующий этап бинаризации заключается в разделении выделенного объекта на области с различной плотностью. Такое разделение позволяет делать выводы о степени симметричности внутреннего строения объекта. Рассмотрим два множества пикселей. Первое множество содержит все пиксели объекта, второе включает только те из них, яркость которых равна нулю (черный цвет соответствует областям объекта, заполненным воздухом). Центр каждого множества пикселей вычисляется как среднее арифметическое координат этих пикселей по осям X, Y, Z . Чем больше расстояние между центрами этих множеств, тем более несимметрично расположены наросты внутри объекта и тем выше вероятность их патологической природы.

Результатом работы описанного алгоритма сегментации слоев снимка КТ является информация о совокупности пикселей на срезах КТ, принадлежащих исследуемому объекту, которая позволяет проводить детальный анализ объекта по снимку и определять его параметры, такие как объем, выборочная несмещенная дисперсия распределения значений его плотности, послойные площади и объемы объекта. Программная реализация разработанного алгоритма позволяет также получить гистограмму распределения плотностей исследуемого образования, по кото-



■ Рис. 4. Интерфейс программного обеспечения, визуализирующий результаты анализа
 ■ Fig. 4. Developed software user interface with object analysis result



■ **Рис. 5.** Гистограммы распределения плотности верхнечелюстных пазух (слева верхнечелюстная пазуха с нормальным строением, справа деформированная верхнечелюстная пазуха с патологическими образованиями внутри)

■ **Fig. 5.** Density distribution histograms of maxillary sinus (left — maxillary sinus with normal anatomic structure, right — deformed maxillary sinus with possible pathologic growths)

рой можно судить о наличии или отсутствии патологии внутри полости и о характере выявленной патологии.

Рассмотрим параметры исследуемого объекта, полученные при помощи разработанного алгоритма. Объем верхнечелюстной пазухи на каждом слое определялся исходя из количества принадлежащих ей пикселей. Суммарный объем пазухи составил 12,96 мл. Результаты послойного анализа пазухи с подсчетом параметров отображены на рис. 4. Разработанный алгоритм и его программная реализация позволяют проводить визуализацию патологических образований без использования контрастирования. Например, на рис. 4 голубым цветом выделена область, плотность которой существенно отличается от плотности воздуха, заполняющего верхнечелюстную пазуху здорового человека. Темно-зеленым выделена та часть пазухи, которая заполнена воздухом. Подобная локализация тканей определенной плотности, их программное определение и визуализация на снимке способны составить альтернативу применению контрастирования.

Гистограммы распределения плотностей двух верхнечелюстных пазух представлены на рис. 5. Левая пазуха заполнена преимущественно воздухом. Это наглядно видно по ее гистограмме, в которой преобладает крайний левый столбец, соответствующий меньшей плотности, чем нижняя граница шкалы Хаунсфилда, т. е. воздуху. В правой полости находятся патологические образования, ее гистограмма отражает их распределение. Показатель выборочной несмещенной дисперсии для значений плотности объекта, отражаемой его пикселями, можно рассматривать как критерий степени неоднородности внутреннего содержания исследуемого объекта. Для правой верхнечелюстной пазухи (выделена на рис. 4) это значение равно 30,95, в то время как для ле-

вой оно равно 8,01. Таким образом, значения дисперсии, наряду с гистограммой плотности, являются численными показателями степени неоднородности содержания верхнечелюстных пазух и свидетельствуют о возможном наличии в них патологических образований.

Заключение

Применение разработанного алгоритма позволяет автоматизировать процесс анализа послойных снимков КТ путем вычисления параметров объектов и их визуализации. Дополнительное разделение исследуемого объекта на зоны дает возможность локализовать области с определенными свойствами, делать выводы о строении объекта и определять потенциальные зоны патологических образований подобно тому, как это достигается применением контрастирования. Визуализация гистограммы плотности дает сведения о характере тканей исследуемого объекта, а выборочная несмещенная дисперсия плотности свидетельствует о степени неоднородности внутреннего строения объекта.

Возможные пространственные искажения изображений от центра к периферии снимка не оказывают заметного влияния на точность диагностики исследуемых объектов. Погрешность разработанного метода имеет систематический характер в силу сходности анатомического строения однотипных исследуемых объектов у различных индивидуумов. Как известно, систематические погрешности простейшим образом корректируются (вычитаются из результата измерения). Что касается программной параметризации верхнечелюстных пазух, выполненной в данной работе, то погрешность предложенного метода в десятки раз меньше погрешности косвенных вычисле-

ний объема пазух, основанных на максимальных фронтальных, вертикальных и сагиттальных размерах этих объектов на отдельных слоях, порядок проведения которых описан в работе [22].

Таким образом, автоматизированное вычисление параметров объектов позволяет дополнить объем диагностической информации, доступной медицинским специалистам. Определение центров исследуемых объектов, выполняемое на осно-

ве обработки множества пикселей, обладает большей точностью, чем косвенные методы их определения [22], применяемые в настоящее время. Разработанное программное обеспечение прошло апробацию в Самарской областной клинической больнице имени В. Д. Середавина и дает возможность проводить диагностику по томографическим снимкам, полученным без использования контрастирования.

Литература

1. **Van Ginneken B., Schaefer-Prokop CM., Prokop M.** Computer-Aided Diagnosis: how to Move from the Laboratory to the Clinic // *The British Journal of Radiology*. 2005. N 78(1). P. 3–19. doi:10.1259/bjr/82933343
2. **Doi K.** Current Status and Future Potential of Computer-Aided Diagnosis in Medical Imaging // *The British Journal of Radiology*. 2005. Vol. 78. P. 3–19. doi:10.1259/bjr/82933343
3. **Hofer M.** CT Teaching Manual: Systematic Approach to CT Reading. — Thieme, 2007. — 224 p.
4. **Seletchi E., Dului O.** Image Processing and Data Analysis in Computed Tomography // *Romanian Journal of Physics*. 2007. N 1. P. 764–774. http://www.nipne.ro/rjp/2007_52_5-6/0667_0667.pdf (дата обращения: 15.12.2017).
5. **Bousson N., Fayad H., Le Pogam A., Pradier O., Visyikis D.** Image Processing Methods in CT for Radiotherapy Applications // *Theory and Applications of CT Imaging and Analysis*. — InTech, 2011. P. 127–142. <http://cdn.intechopen.com/pdfs-wm/14772.pdf> (дата обращения: 15.12.2017).
6. **Maher M., Kalra M., Sahani D., Perumpillichira J., Rizzo S., Saini S., Mueller P.** Techniques, Clinical Applications and Limitations of 3D Reconstruction in CT of the Abdomen // *Korean Journal of Radiology*. 2004. N 5(1). P. 55–67. doi:10.3348/kjr.2004.5.1.55
7. **Kim H. C., Park S. H., Park S. I., Shin H. C., Park S. J., Kim H. H., Kim Y. T., Bae W. K., Kim I. Y.** Three-Dimensional Reconstructed Images using Multidetector Computed Tomography in Evaluation of the Biliary Tract // *Abdominal Imaging*. 2004. N 29(4). P. 472–478. doi:10.1007/s00261-003-0123-x
8. **Xu F., Mueller K.** Real-Time 3D Computed Tomographic Reconstruction using Commodity Graphics Hardware // *Physics in Medicine and Biology*. 2007. N 52(12). P. 3405–3419. doi:10.1088/0031-9155/52/12/006
9. **El-Baz A., Beache G., Gimel'farb G., Suzuki K., Okada K., Elnakib A., Soliman A., Abdollahi B.** Computer-Aided Diagnosis Systems for Lung Cancer: Challenges and Methodologies // *International Journal of Biomedical Imaging*. 2013. N 2013. P. 1–46. doi:10.1155/2013/942353
10. **Chen H., Xu Y., Ma Y., Ma B.** Neural Network Ensemble-based Computer-Aided Diagnosis for Differentiation of Lung Nodules on CT Images: Clinical Evaluation // *Academic Radiology*. 2016. N 3(82). P. 595–602. doi:10.1016/j.acra.2009.12.009
11. **Chen H., Wang X. H., Ma D. Q., Ma B. R.** Neural Network-based Computer-Aided Diagnosis in Distinguishing Malignant from Benign Solitary Pulmonary Nodules by Computed Tomography // *Chinese Medical Journal*. 2007. N 120(14). P. 1211–1215. http://124.205.33.103:81/ch/reader/create_pdf.aspx?file_no=200771851284230&year_id=2007&quarter_id=14&falg=1 (дата обращения: 15.12.2017).
12. **Montejo L., Jia J., Kim H., Netz U., Blaschke S., Müller G., Hielscher A.** Computer-Aided Diagnosis of Rheumatoid Arthritis with Optical Tomography. Part 1: Feature Extraction // *Journal of Biomedical Optics*. 2013. N 18(7). P. 123–137. doi:10.1117/1.JBO.18.7.076001
13. **Montejo L., Jia J., Kim H., Netz U., Blaschke S., Müller G., Hielscher A.** Computer-Aided Diagnosis of Rheumatoid Arthritis with Optical Tomography. Part 2: Image Classification // *Journal of Biomedical Optics*. 2013. N 18(7). P. 138–150. doi:10.1117/1.JBO.18.7.076002
14. **Hussien R., Supriyanto E., Anati N., Azmira N., Illani N., Mahfooz R.** Digital Processing for Computed Tomography Images: Brain Tumor Extraction and Histogram Analysis // *Mathematics and Computers in Contemporary Science*. — WSEAS Press, 2013. P. 119–128.
15. **Homma N.** CT Image based Computer-Aided Lung Cancer Diagnosis // *Theory and Applications of CT Imaging and Analysis*. — InTech, 2011. P. 3–14. <http://www.intechopen.com/books/theory-and-applications-of-ct-imaging-and-analysis/ct-image-based-computer-aided-lung-cancer-diagnosis> (дата обращения: 15.12.2017).
16. **Zhou C., Chan H-P., Sahiner B.** Computer-Aided Detection of Pulmonary Embolism in Computed Tomographic Pulmonary Angiography (CTPA): Performance Evaluation with Independent Data Sets // *Medical Physics*. 2009. N 36(8). P. 3385–3396. doi:10.15217/issn1684-8853.2016.3.2
17. **Gonzalez R., Woods E.** Digital Image Processing. — Prentice Hall, 2007. — 976 p.

18. Pratt W. Digital Image Processing: PIKS Scientific Inside. — Wiley-Interscience, 2007. — 812 p.
 19. Seletchi E., Sutac V. Image Analysis in X-ray Computed Tomography // ICVL. 2006. P. 187–194. http://fmi.unibuc.ro/cniv/2006/disc/icvl/documente/pdf/met/16_seletchi.pdf (дата обращения: 15.12.2017).

20. Romans L. Computed Tomography for Technologists: A Comprehensive Text. — LWW, 2004. — 400 p.
 21. Kelley L., Petersen C. Workbook for Sectional Anatomy for Imaging Professionals. — Mosby, 2012. — 256 p.
 22. Волков А. Г. Лобные пазухи. — Ростов н/Д: Феникс, 2000. — 512 с.

UDC 004.932.2

doi:10.15217/issn1684-8853.2018.3.98

Computer Aided Tomography Picture Analysis on the base of Stage-By-Stage Binarization and its Software Implementation

Limanova N. I.^a, Dr. Sc., Tech., Associate Professor, nataliya.i.limanova@gmail.com

Ataev S. G.^a, Assistant Professor, sergeyataev@gmail.com

^aPovolzhskiy State University of Telecommunications and Informatics, 23, L'va Tolstogo St., 443010, Samara, Russian Federation

Introduction: Computed tomography shots of a body contain big amounts of information and require complex comprehensive analysis which takes a long time and often is difficult for a diagnostician. The diagnostics accuracy can be improved by the use of contrast dye, but this method is more expensive and has a number of contraindications. **Purpose:** Developing an automated object detection method for computed tomography shots which would provide physicians with the necessary information. **Results:** A parameterization method is proposed for objects on the layers of a CT shot. It performs sequential analysis of the layers and determines the area related to the examined object. In the course of the further analysis, the object is considered as an array of voxels, each of them characterizing an average density of the object in a particular zone. Furthermore, the areas of each layer are calculated, resulting in the total volume of the object. The determination of the selective dispersion for the density values of the object zones along with the density values histogram allow the diagnostician to make conclusions about the object properties and to evaluate the possibility of pathologies. Filtering the object voxels by their density values makes it possible to localize areas of the object with certain properties and to estimate the degree of symmetry for new growths in order to judge about their possible pathological nature. **Practical relevance:** The developed method along with its software implementation allows you to partially automate CT shot analysis by providing the medical specialists with the values of the examined object parameters: areas of each layer, total volume, sample variance of the density values, and symmetry estimation. This method makes it possible to localize the object zones with particular density values and to detect possible pathologic growths without the use of contrast dye.

Keywords — Computed Tomography, Image Analysis, Image Binarization, Software, Computer Aided Diagnostics.

Citation: Limanova N. I., Ataev S. G. Computer Aided Tomography Picture Analysis on the base of Stage-By-Stage Binarization and its Software Implementation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 98–106 (In Russian). doi:10.15217/issn1684-8853.2018.3.98

References

1. Van Ginneken B., Schaefer-Prokop CM., Prokop M. Computer-Aided Diagnosis: how to Move from the Laboratory to the Clinic. *Radiology*, 2011, vol. 261, no. 3, pp. 719–732. doi:10.1148/radiol.11091710
2. Doi K. Current Status and Future Potential of Computer-Aided Diagnosis in Medical Imaging. *The British Journal of Radiology*, 2005, vol. 78, pp. 3–19. doi:10.1259/bjr/82933343
3. Hofer M. *CT Teaching Manual: Systematic Approach to CT Reading*. Thieme, 2007. 224 p.
4. Seletchi E., Dului O. Image Processing and Data Analysis in Computed Tomography. *Romanian Journal of Physics*, 2007, vol. 1, pp. 764–774. Available at: http://www.nipne.ro/rjp/2007_52_5-6/0667_0667.pdf (accessed 15 December 2017).
5. Boussion N., Fayad H., Le Pogam A., Pradier O., Visyikis D. Image Processing Methods in CT for Radiotherapy Applications. In: *Theory and Applications of CT Imaging and Analysis*. InTech, 2011. P. 127–142. Available at: <http://cdn.intechopen.com/pdfs-wm/14772.pdf> (accessed 15 December 2017).
6. Maher M., Kalra M., Sahani D., Perumpillichira J., Rizzo S., Saini S., Mueller P. Techniques, Clinical Applications and Limitations of 3D Reconstruction in CT of the Abdomen. *Korean Journal of Radiology*, 2004, vol. 5, no. 1, pp. 55–67. doi:10.3348/kjr.2004.5.1.55
7. Kim H. C., Park S. H., Park S. I., Shin H. C., Park S. J., Kim H. H., Kim Y. T., Bae W. K., Kim I. Y. Three-Dimensional Reconstructed Images using Multidetector Computed Tomography in Evaluation of the Biliary Tract. *Abdominal Imaging*, 2014, vol. 29, no. 4, pp. 472–478. doi:10.1007/s00261-003-0123-x
8. Xu F., Mueller K. Real-Time 3D Computed Tomographic Reconstruction using Commodity Graphics Hardware. *Physics in Medicine and Biology*, 2007, vol. 52, no. 12, pp. 3405–3419. doi:10.1088/0031-9155/52/12/006
9. El-Baz A., Beache G., Gimel'farb G., Suzuki K., Okada K., Elnakib A., Soliman A., Abdollahi B. Computer-Aided Diagnosis Systems for Lung Cancer: Challenges and Methodologies. *International Journal of Biomedical Imaging*, 2013, vol. 2013, pp. 1–46. doi:10.1155/2013/942353
10. Chen H., Xu Y., Ma Y., Ma B. Neural Network Ensemble-based Computer-Aided Diagnosis for Differentiation of Lung Nodules on CT Images: Clinical Evaluation. *Academic Radiology*, 2010, vol. 17, no. 5, pp. 595–602. doi:10.1016/j.acra.2009.12.009
11. Chen H., Wang X. H., Ma D. Q., Ma B. R. Neural Network-based Computer-Aided Diagnosis in Distinguishing Malignant from Benign Solitary Pulmonary Nodules by Computed Tomography. *Chinese Medical Journal*, 2007, vol. 120, no. 14, pp. 1211–1215. Available at: <http://124.205.33.103:81/ch/>

- reader/create_pdf.aspx? file_no=200771851284230&year_id=2007&quarter_id=14&falq=1 (accessed 15 December 2017).
12. Montejo L., Jia J., Kim H., Netz U., Blaschke S., Müller G., Hielscher A. Computer-Aided Diagnosis of Rheumatoid Arthritis with Optical Tomography. Part 1: Feature Extraction. *Journal of Biomedical Optics*, 2013, vol. 18, no. 7, pp. 123–137. doi:10.1117/1.JBO.18.7.076001
 13. Montejo L., Jia J., Kim H., Netz U., Blaschke S., Müller G., Hielscher A. Computer-Aided Diagnosis of Rheumatoid Arthritis with Optical Tomography, Part 2: Image Classification. *Journal of Biomedical Optics*, 2013, vol. 18, no. 7, pp. 138–150. doi:10.1117/1.JBO.18.7.076002
 14. Hussien R., Supriyanto E., Anati N., Azmira N., Ilani N., Mahfooz R. Digital Processing for Computed Tomography Images: Brain Tumor Extraction and Histogram Analysis. *Mathematics and Computers in Contemporary Science*, 2013, WSEAS Press, pp. 119–128.
 15. Homma N. CT Image based Computer-Aided Lung Cancer Diagnosis. In: *Theory and Applications of CT Imaging and Analysis*. InTech, 2011. P. 3–14. Available at: <http://www.intechopen.com/books/theory-and-applications-of-ct-imaging-and-analysis/ct-image-based-computer-aided-lung-cancer-diagnosis> (accessed 15 December 2017).
 16. Zhou C., Chan H-P., Sahiner B. Computer-Aided Detection of Pulmonary Embolism in Computed Tomographic Pulmonary Angiography (CTPA): Performance Evaluation with Independent Data Sets. *Medical Physics*, 2009, vol. 36, no. 8, pp. 3385–3396. doi:10.1118/1.3157102
 17. Gonzalez R., Woods E. *Digital Image Processing*. Prentice Hall, 2007. 976 p.
 18. Pratt W. *Digital Image Processing: PIKS Scientific Inside*. Wiley-Interscience, 2007. 812 p.
 19. Seletchi E., Sutac V. Image Analysis in X-ray Computed Tomography. *The 1st Intern. Conf. on Virtual Learning (ICVL)*, 2006, pp. 187–194. Available at: http://fmi.unibuc.ro/cniv/2006/disc/icvl/documente/pdf/met/16_seletchi.pdf (accessed 15 December 2017).
 20. Romans L. *Computed Tomography for Technologists: A Comprehensive Text*. LWW, 2010. 400 p.
 21. Kelley L., Petersen C. *Workbook for Sectional Anatomy for Imaging Professionals*. Mosby, 2012. 256 p.
 22. Volkov A. G. *Lobnye pazukhi* [Frontal Sinus]. Rostov-on-Don: Fenix Publ., 2000. 512 p. (In Russian).

**Научный журнал
«ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ»**

выходит каждые два месяца.

Стоимость годовой подписки (6 номеров) для подписчиков России — 6000 рублей, для подписчиков стран СНГ — 6600 рублей, включая НДС 18%, таможенные и почтовые расходы.

Подписку на печатную версию журнала можно оформить в любом отделении связи по каталогу:

«Роспечать»: № 15385 — полугодовой индекс,

а также через посредство подписных агентств:

«Северо-Западное агентство „Прессинформ“»

Санкт-Петербург, тел.: (812) 335-97-51, 337-23-05,

эл. почта: press@crp.spb.ru, zajavka@crp.spb.ru,

сайт: <http://www.pinform.spb.ru>

«МК-Периодика» (РФ + 90 стран)

Москва, тел.: (495) 681-91-37, 681-87-47,

эл. почта: export@periodicals.ru, сайт: <http://www.periodicals.ru>

«Деловая пресса»

Москва, тел.: (495) 962-11-11, эл. почта: podpiska@delpress.ru,

сайт: <http://delpress.ru/contacts.html>

«Коммерсант-Курьер»

Казань, тел.: (843) 291-09-99, 291-09-47, эл. почта: kazan@komcur.ru,

сайт: <http://www.komcur.ru/contacts/kazan/>

«Урал-Пресс» (филиалы в 40 городах РФ)

Сайт: <http://www.ural-press.ru>

«Идея» (Украина)

Сайт: <http://idea.com.ua>

«BTL» (Узбекистан)

Сайт: <http://btl.sk.uz/ru/cat17.html> и др.

На электронную версию нашего журнала (все выпуски, годовая подписка, один выпуск, одна статья) вы можете подписаться на сайтах НЭБ: <http://elibrary.ru>;

РУКОНТ: <http://www.rucont.ru>; ИВИС: <http://www.ivis.ru/>

Полнотекстовые версии журнала за 2002–2016 гг.

в свободном доступе на сайте журнала (<http://www.i-us.ru>),

НЭБ (<http://www.elibrary.ru>)

и Киберленинки (<http://cyberleninka.ru/>

journal/n/informatsionno-upravlyayuschiesistemy).

АГЕЕВ
Сергей
Александрович



Ведущий научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН. В 1983 году окончил радиотехнический факультет Ульяновского политехнического института по специальности «Радиотехника». Является автором 70 научных публикаций. Область научных интересов — информационная и сетевая безопасность, искусственный интеллект, мультисервисные сети связи. Эл. адрес: serg123_61@mail.ru

АТАЕВ
Сергей
Геннадьевич



Ассистент кафедры информационных систем и технологий Поволжского государственного университета телекоммуникаций и информатики, Самара. В 2014 году окончил Поволжский государственный университет телекоммуникаций и информатики по специальности «Программное обеспечение вычислительной техники и автоматизированных систем». Является автором 17 научных публикаций. Область научных интересов — теория графов, методы обработки и анализа изображений, системы компьютерной диагностики, системы поддержки принятия решений. Эл. адрес: sergeyataev@gmail.com

БАЛОНИН
Юрий
Николаевич



Научный сотрудник кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2010 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети». Является автором 17 научных публикаций. Область научных интересов — вычислительные методы, теория чисел. Эл. адрес: yuraball@mail.ru

БРИТОВ
Георгий
Семенович



Доцент кафедры информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1963 году окончил Ленинградский институт авиационного приборостроения по специальности «Авиационное приборостроение». В 1968 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 100 научных публикаций и 13 патентов на изобретения. Область научных интересов — теория надежности и техническая диагностика информационных систем. Эл. адрес: bgs@ibi.metrocom.ru

ВОЛЬСКИЙ
Александр
Владимирович



Магистрант кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2017 году окончил бакалавриат Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Инфокоммуникационные технологии и системы связи». Область научных интересов — моделирование систем инфокоммуникаций. Эл. адрес: volskiyav62@gmail.com

ЗИАТДИНОВ
Сергей
Ильич



Профессор кафедры информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1970 году окончил Ленинградский институт авиационного приборостроения по специальности «Электронные устройства». В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором 140 научных публикаций и 16 патентов на изобретения. Область научных интересов — аналоговая и цифровая обработка сигналов, автоматические системы управления. Эл. адрес: kaf53@guap.ru

**КОЗИН
Иван
Сергеевич**



Ведущий специалист отдела защиты информации АО «Кронштадт Технологии», Санкт-Петербург.

В 2011 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Прикладная информатика (в социальной сфере)».

Является автором пяти научных публикаций.

Область научных интересов — информационные системы персональных данных, методы защиты конфиденциальной информации, теория искусственных нейронных сетей

Эл. адрес:
van@trioptimum.com

**КУЗНЕЦОВ
Виталий
Александрович**



Ассистент кафедры информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 2012 году окончил магистратуру Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Информационные системы».

Является автором десяти научных публикаций.

Область научных интересов — машинное зрение, цифровая обработка изображений, трехмерное сканирование.

Эл. адрес: k.avk-c@mail.ru

**ЛЮДАЕВ
Михаил
Юрьевич**



Инженер-программист ПАО «Российский институт мощного радиостроения», Санкт-Петербург.

В 2005 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Инженер-схемотехник».

Является автором 12 научных публикаций.

Область научных интересов — цифровая обработка изображений, цифровая обработка сигналов, построение распределенных вычислительных систем.

Эл. адрес: mike_ld@mail.ru

**КОТЕНКО
Игорь
Витальевич**



Профессор, заведующий лабораторией проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН.

В 1983 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Математическое обеспечение автоматизированных систем управления», в 1987 году — Военную академию связи по специальности «Инженерная автоматизированных систем управления».

В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 450 научных публикаций.

Область научных интересов — безопасность компьютерных сетей, обнаружение компьютерных атак, межсетевые экраны, защита от вирусов и сетевых червей и др.

Эл. адрес: ivkote@comsec.spb.ru

**ЛИМАНОВА
Наталья
Игоревна**



Доцент, заведующая кафедрой информационных систем и технологий Поволжского государственного университета телекоммуникаций и информатики, Самара.

В 1986 году окончила Куйбышевский авиационный институт им. акад. С. П. Королева по специальности «Конструирование и производство радиоаппаратуры».

В 2007 году защитила диссертацию на соискание ученой степени доктора технических наук.

Является автором более 200 научных публикаций и 18 патентов на изобретения.

Область научных интересов — автоматизация и управление в информационных системах, теория и алгоритмы анализа и обработки изображений и др.

Эл. адрес:
nataliya.i.limanova@gmail.com

**МАРТЫНОВА
Любовь
Александровна**



Старший научный сотрудник, ведущий научный сотрудник научно-исследовательского центра «Системы освещения обстановки» АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург.

В 1985 году окончила Ленинградский кораблестроительный институт по специальности «Прикладная математика».

В 2013 году защитила диссертацию на соискание ученой степени доктора технических наук.

Является автором 90 научных публикаций.

Область научных интересов — системный анализ, математическое моделирование, оценка эффективности, обработка разнородной информации.

Эл. адрес:
martynowa999@bk.ru

МАТВЕЕВ
Николай
Валентинович



Ассистент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2015 году окончил магистратуру Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Инфокоммуникационные технологии и системы связи». Является автором пяти научных публикаций. Область научных интересов — математическая оптимизация, беспроводные сети, LTE, теория вероятностей, имитационное моделирование. Эл. адрес: n.matveev@vu.spb.ru

ОСИПОВ
Леонид
Андроникович



Профессор, заведующий кафедрой информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1968 году окончил Ленинградский институт авиационного приборостроения по специальности «Электрооборудование летательных аппаратов». В 1995 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором около 200 научных публикаций, в том числе соавтором трех монографий. Область научных интересов — компьютерное управление нелинейными объектами. Эл. адрес: kaf53@guar.ru

САЕНКО
Игорь
Борисович



Профессор, ведущий научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН. В 1981 году окончил Белорусский государственный университет по специальности «Радиофизика и электроника», в 1989 году — Военную академию связи по специальности «Инженерная автоматизированных систем управления». В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 350 научных публикаций и семи патентов на изобретения. Область научных интересов — автоматизированные информационные системы, информационная безопасность и др. Эл. адрес: ibsaen@comsec.spb.ru

МОТОРИН
Дмитрий
Евгеньевич



Аспирант кафедры телематики Санкт-Петербургского политехнического университета Петра Великого. В 2014 году окончил Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» по специальности «Автоматизация и управление производственными комплексами и подвижными объектами» и Технический университет Софии по специальности «Мехатронные системы». Является автором девяти научных публикаций и двух свидетельств регистрации программы для ЭВМ. Область научных интересов — системы управления, гетерогенные группы роботов, информационные системы. Эл. адрес: d.e.motorin@gmail.com

ПОПОВ
Сергей
Геннадьевич



Доцент кафедры телематики Санкт-Петербургского политехнического университета Петра Великого. В 1996 году окончил Санкт-Петербургский государственный технический университет по специальности «Системы автоматизации проектных работ». В 2010 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 50 научных публикаций. Область научных интересов — теория баз данных, информационные системы, автомобильная телематика, беспроводные сети. Эл. адрес: popovserge@spbstu.ru

СЕРГЕЕВ
Александр
Михайлович



Старший преподаватель кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2004 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети». Является автором 32 научных публикаций. Область научных интересов — численные методы, теория вычислительных процессов, проектирование специализированных процессоров. Эл. адрес: asklab@mail.ru

СОКОЛОВА
Юлия
Витальевна



Магистрант кафедры информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения. Является автором одной научной публикации. Область научных интересов — информационные технологии. Эл. адрес: kaf53@guar.ru

ТАТАРНИКОВА
Татьяна
Михайловна



Профессор кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1993 году окончила Восточно-Сибирский технологический институт по специальности «Электронно-вычислительные машины, комплексы, системы и сети». В 2007 году защитила диссертацию на соискание ученой степени доктора технических наук. Является автором более 100 научных публикаций. Область научных интересов — инфокоммуникации, взаимодействие неоднородных сетей. Эл. адрес: tm-tatarn@yandex.ru

ТРОФИМОВ
Андрей
Николаевич



Доцент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1976 году окончил Ленинградский институт авиационного приборостроения по специальности «Автоматизированные системы управления». В 1982 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 41 научной публикации и шести патентов на изобретения. Область научных интересов — теория передачи дискретных сообщений, теория информации, теория кодирования. Эл. адрес: andrei.trofimov@vu.spb.ru

ТЮРЛИКОВ
Андрей
Михайлович



Профессор, директор Института информационных систем и защиты информации, заведующий кафедрой инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1980 году окончил Ленинградский институт авиационного приборостроения по специальности «Информационные системы управления». В 2011 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 80 научных публикаций. Область научных интересов — многоабонентные системы связи, системы дистанционного обучения, протоколы передачи данных в реальном масштабе времени, алгоритмы сжатия видеoinформации. Эл. адрес: turlikov@vu.spb.ru