

ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

3(112)/2021

3(112)/2021

PEER REVIEWED JOURNAL

INFORMATSIONNO- UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

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 Board

S. Andreev

Dr. Sc., Tampere, Finland

V. Anisimov

Dr. Sc., Professor, Saint-Petersburg, Russia

B. Bezruchko

Dr. Sc., Professor, Saratov, Russia

N. Blaunstein

Dr. Sc., Professor, Beer-Sheva, Israel

M. Buzdalov,

PhD, Researcher, Saint-Petersburg, Russia

C. Christodoulou

PhD, Professor, Albuquerque, New Mexico, USA

A. Dudin

Dr. Sc., Professor, Minsk, Belarus

I. Dumer

PhD., Professor, Riverside, USA

M. Favorskaya

Dr. Sc., Professor, Krasnoyarsk, Russia

L. Fortuna

PhD, Professor, Catania, Italy

A. Fradkov

Dr. Sc., Professor, Saint-Petersburg, Russia

A. Hramov

Dr. Sc., Professor, Indianapolis, Russia

L. Jain

PhD, Professor, Canberra, Australia

V. Khimenko

Dr. Sc., Professor, Saint-Petersburg, Russia

G. Matvienko

Dr. Sc., Professor, Tomsk, Russia

A. Myllari

PhD, Professor, Grenada, West Indies

Y. Podoplyokin

Dr. Sc., Professor, Saint-Petersburg, Russia

K. Samouylov

Dr. Sc., Professor, Moscow, Russia

J. Seberry

PhD, Professor, Wollongong, Australia

A. Shalyto

Dr. Sc., Professor, Saint-Petersburg, Russia

A. Shepeta

Dr. Sc., Professor, Saint-Petersburg, Russia

Yu. Shokin

RAS Academician, Dr. Sc., Novosibirsk, Russia

A. Smirnov

Dr. Sc., Professor, Saint-Petersburg, Russia

T. Sutikno

PhD, Associate Professor, Yogyakarta, Indonesia

Z. Yuldashev

Dr. Sc., Professor, Saint-Petersburg, Russia

R. Yusupov

RAS Corr. Member, Dr. Sc., Professor, Saint-Petersburg, Russia

A. Zeifman

Dr. Sc., Professor, Vologda, Russia

Editor: A. Larionova**Proofreader:** T. Zvertanovskaia**Design:** M. Chernenko, Y. Umnitsina**Layout and composition:** Y. Umnitsina**Contact information**

The Editorial and Publishing Center, SUAI

67, B. Morskaia, 190000, St. Petersburg, Russia

Website: <http://i-us.ru/en>, e-mail: i-us.spb@gmail.com

Tel.: +7 - 812 494 70 02

THEORETICAL AND APPLIED MATHEMATICS**Balonin Yu. N., Sergeev A. M., Vostrikov A. A.** *The rise in maximum determinant matrix complexity*

2

INFORMATION PROCESSING AND CONTROL**Osipov V. Yu., Nikiforov V. V.** *Coding and robustness of signal processing in streaming recurrent neural networks*

9

Parfiryev A. V., Parfiryeva O. V., Dushkin A. V. *Method and algorithm for determining the geographic coordinates of ground objects from an unmanned aerial vehicle*

19

SYSTEM AND PROCESS MODELING**Lebedev I. S.** *Dataset segmentation considering the information about impact factors*

29

INFORMATION CODING AND TRANSMISSION**Kruglika S. A., Kucherova G. A., Nazirkhanova K. N., Filitova M. E.** *Information-theoretic problems of DNA-based storage systems*

39

INFORMATION CHANNELS AND MEDIUM**Ziatdinova S. I., Osipova L. A.** *Suppression of correlated interference by adaptive notch filters under pulse repetition period modulation*

53

CHRONICLES AND INFORMATION*International Conference «Big Data Days 2021»*

61

INFORMATION ABOUT THE AUTHORS

62

3(112)/2021

ИНФОРМАЦИОННО-
УПРАВЛЯЮЩИЕ
СИСТЕМЫ

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

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:

С. Д. Андреев,
д-р техн. наук, Тампере, ФинляндияВ. Г. Анисимов,
д-р техн. наук, проф., Санкт-Петербург, РФБ. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФН. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, ИзраильМ. В. Буздалов,
канд. техн. наук, научный сотрудник, Санкт-Петербург, РФЛ. С. Джайн,
д-р наук, проф., Канберра, АвстралияА. Н. Дудин,
д-р физ.-мат. наук, проф., Минск, БеларусьИ. И. Думер,
д-р наук, проф., Риверсайд, СШАА. И. Зейфман,
д-р физ.-мат. наук, проф., Вологда, РФК. Кристофолу,
д-р наук, проф., Альбукерке, Нью-Мексико, СШАГ. Г. Матвиенко,
д-р физ.-мат. наук, проф., Томск, РФА. А. Мюллер,
д-р наук, проф., Гренада, Вест-ИндияЮ. Ф. Подоплёкин,
д-р техн. наук, проф., Санкт-Петербург, РФК. Е. Самуилов,
д-р техн. наук, проф., Москва, РФД. Себерри,
д-р наук, проф., Волонгонг, АвстралияА. В. Смирнов,
д-р техн. наук, проф., Санкт-Петербург, РФТ. Сутикнуо,
д-р наук, доцент, Джокьякарта, ИндонезияМ. Н. Фаворская,
д-р техн. наук, проф., Красноярск, РФЛ. Фортуна,
д-р наук, проф., Катания, ИталияА. Л. Фрадков,
д-р техн. наук, проф., Санкт-Петербург, РФВ. И. Хименко,
д-р техн. наук, проф., Санкт-Петербург, РФА. Е. Храмов,
д-р физ.-мат. наук, Иннополис, РФА. А. Шалыто,
д-р техн. наук, проф., Санкт-Петербург, РФА. П. Шепета,
д-р техн. наук, проф., Санкт-Петербург, РФЮ. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФЗ. М. Юлдашев,
д-р техн. наук, проф., Санкт-Петербург, РФР. 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

ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ МАТЕМАТИКА

Balotin Yu. N., Sergeev A. M., Vostrikov A. A. The rise in maximum determinant matrix complexity

2

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

Осипов В. Ю., Никифоров В. В. Кодирование и устойчивость обработки сигналов в потоковых рекуррентных нейронных сетях

9

Parfiriyev A. V., Parfiriyeva O. V., Dushkin A. V. Method and algorithm for determining the geographic coordinates of ground objects from an unmanned aerial vehicle

19

МОДЕЛИРОВАНИЕ СИСТЕМ И ПРОЦЕССОВ

Лебедев И. С. Сегментирование множества данных с учетом информации воздействующих факторов

29

КОДИРОВАНИЕ И ПЕРЕДАЧА ИНФОРМАЦИИ

Круглик С. А., Кучеров Г. А., Назирханова К. Н., Филитов М. Е. Теоретико-информационные проблемы ДНК-памяти

39

ИНФОРМАЦИОННЫЕ КАНАЛЫ И СРЕДЫ

Зиятдинов С. И., Осипов Л. А. Подавление коррелированных помех адаптивными режекторными фильтрами при модуляции периода повторения импульсов

53

ХРОНИКА И ИНФОРМАЦИЯ

Международная конференция «Big Data Days 2021»

61

СВЕДЕНИЯ ОБ АВТОРАХ

62

Журнал входит в БД SCOPUS и в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук.

Сдано в набор 11.05.21. Подписано в печать 25.06.21. Формат 60×84/8.

Гарнитура SchoolBookC. Печать цифровая.

Усл. печ. л. 7,6. Уч.-изд. л. 10,5. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 244.

Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.

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

Отпечатано с готовых диапозитивов в редакционно-издательском центре ГУАП.

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

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.

© Коллектив авторов, 2021

Распространяется бесплатно.

UDC 519.614

doi:10.31799/1684-8853-2021-3-2-8

The rise in maximum determinant matrix complexity

Yu. N. Balonin^a, Research Fellow, orcid.org/0000-0002-5102-4139

A. M. Sergeev^a, PhD, Tech., Associate Professor, orcid.org/0000-0002-4788-9869

A. A. Vostrikov^a, PhD, Tech., Associate Professor, orcid.org/0000-0002-8513-368, vostricov@mail.ru

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

Introduction: In spite of the apparent relation between maximum determinant matrices of even (Hadamard matrices) and odd orders, the latter have particularly complex patterns of repetitive elements. This is what makes them unique and attractive for various applications in visual data processing, coding and masking. **Purpose:** Developing the theory of maximum determinant matrices, with the focus on using computer-aided analysis, and calculating unique matrices with unique pattern structures in their portraits. **Results:** We have found some peculiarities of maximum determinant matrices, outlined their families related to Fermat numbers, demonstrated the complication of patterns in other matrices as their orders grow. The presumption about the increasing complexity of structures as the matrix orders grow is confirmed by a chain of matrix portraits we demonstrate. As applied to orthogonal Belevitch matrices, it follows that they cannot be found even in small orders such as 66 or 86.

Keywords – determinant, determinant maximum, quasi-orthogonal matrices, Hadamard matrices, Mersenne matrices, Fermat matrices, cyclic matrices.

For citation: Balonin Yu. N., Sergeev A. M., Vostrikov A. A. The rise in maximum determinant matrix complexity. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 2–8. doi:10.31799/1684-8853-2021-3-2-8

Introduction

Mathematical objects, such as matrices with small number of their element values (levels), are corresponded by determinant-extreme solutions which find practical use in image processing and noise-immune coding problems [1, 2], such as impulse-noise-resistant coding of images transferred via telecommunication channels [3], formation of nested code structures based on Barker codes with improved auto-correlation functions [4], image transformation for better protection from unauthorized use [5], etc.

In spite of the successes in its practical use, the theory of building few-level matrices is far from being completed. The material it has accumulated has not been fully systematized yet. For example, it is well known that prime-number sizes of matrix blocks can be matched by simple symmetric or skew-symmetric patterns formed by a cyclic shift of the upper row of the block, moving the pushed-out elements into the beginning. Non-prime-number sizes can strike you with a great diversity of possible pattern structures, still poorly studied.

In some cases, the structure of matrices or their blocks stays cyclic regardless of whether their sizes are expressed by prime or composite numbers. Although the problem of searching for few-level matrices has been in discussion for about 150 years, starting from Sylvester's works [6], the attention to the existence of universal structures (symmetric, cyclic, two circulant, negacyclic, edgeless or edged) for Hadamard matrices [7] was paid only recently

[8, 9]. Non-universal structures are studied even worse.

In other cases, splitting a matrix into one or two cyclic blocks is closely associated with the primeness of the number which expresses the matrix order. This seemingly natural rule has some exceptions [10] which ruin the integrity of the concept that a matrix order should be followed both numerical and pattern (pattern is ornament, structure) characteristics. The idea about a possible solution becomes rather ambiguous.

Maximum determinant matrices [11–13] are particularly unlucky, exactly because their properties, including pattern properties, are strongly pronounced. Researchers prefer to study isolated odd orders [14, 15], missing the overall picture, especially because they do not collate these orders with few-level orthogonal Hadamard and Belevitch matrices [7, 16, 17]. For example, Bush-type Hadamard matrices [10] are constrained by large square arrays of 1s on their diagonal. Like Belevitch matrices with 0s on their diagonal, they have even orders, but their pattern characteristics make them closer to the most complicated case, odd-order matrices.

In this work, we focus not on particular determinant-extremal matrices as mathematical objects difficult to calculate, and not even on their calculation methods, but on what unites them with and separates them from all the matrices from the vast Hadamard family, being figured out during the study of determinant properties. The purpose of this work is to present some common patterns in maximum determinant matrices, showing how uni-

versal their structures are and how their complexity grows at higher orders.

Simple and complex patterns of even-order matrices

In order to study few-level matrix patterns, let us discuss two different cases which have influenced the matrix pattern theory.

In Fig. 1, you can see an Hadamard matrix with elements 1 and -1 shown on its portrait as white and dark cells, respectively. The matrix is of order 16, consisting of a cyclic block and an edge. This is a classical combination, as it was in the 1960s when it was discovered that there are no edgeless cyclic matrices in orders higher than 4.

The first row of the nested cyclic block of order $p = n - 1$ can be easily found by an algorithm working in finite field arithmetics $GF(p)$. When M. Hall discovered a cyclic block of a composite size 15 [10, 18], it raised some perplexity because even when there is a field $GF(p^m)$, the internal block can be split into cyclic subblocks of size p . Later, two points were realized about such cases.

The first, almost apparent point is that 15 is a product of two close prime numbers 3 and 5. The second, less known one can be checked by Fig. 1. In all “normal” matrices obtained by finite field algorithms, the elements placed symmetrically about the diagonal always differ in sign. Therefore, the edge is often inverted by multiplying the first row (or column) by -1 . But the above-mentioned matrix violates this rule, which means that a composite construction can always be recognized by its pattern. Mathematics leaves no room for arbitrariness and wanton deviation from its fundamental principles.

On evenly odd orders, instead of Hadamard matrices we search for Belevitch matrices which dif-

fer from Hadamard ones by a chain of 0s on their diagonal. In Fig. 1, they are highlighted in red. If the field $GF(p)$ exists, it exists for all prime p , and a Belevitch matrix is easy to find. Nevertheless, the 0s on the diagonal are not 1s, and Belevitch matrices do not share the property of Hadamard matrices about having a certain universal structure formed at most by four cyclic blocks, a couple of negacyclic blocks or a couple of cyclic blocks and a double edge [9, 10]. Apparently, there is a reason for Belevitch matrix patterns to become more complex, and this reason manifests itself in a very unexpected fashion.

In a chain of matrices of orders 6, 26, 46, 66, 86, 106, 126..., with an interval 20, the first matrix of order 6 has the above-mentioned simple structure which corresponds to the prime number $p = 5$. In fields $GF(25)$, $GF(125)$, the internal block consists of cyclic subblocks, therefore the matrix orders do not cause any size problems.

The order $46 = 5 \times 9 + 1$ is a different matter. It is easy to see that the edge in this decomposition corresponds to 1, while the numbers 3 and 5 form a close couple (9 is a prime number power). Nevertheless, two blocks out of 5 in this decomposition partially lose their cyclic structure: they consist of 1, -1 , -1 sequences cyclically shifted by three. The solution is very uncommon for Hadamard matrices.

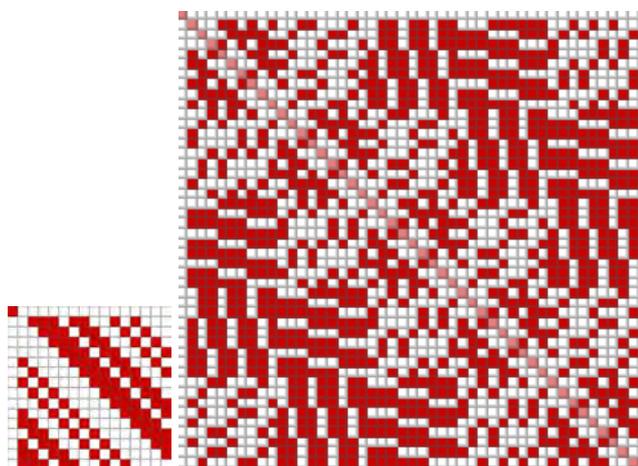
But really interesting is that this trick does not work for matrices of adjacent orders 66 and 86. However, the precondition for their existence is met, as the values $p = n - 1$ in both cases can be split into two squares: $65 = 16 + 49$, $85 = 36 + 49$.

Note that the simplest one- or two-blocked edgeless or double-edged structures have been verified in a computer experiment (in [8] and other works to be published). There are two assumptions left: either Belevitch matrices of order 66 and 86 differ from each other, having, in respect to the matrix of order 46, a new unpredictable pattern element, or, which is less probable, the demand for such an element contradicts the symmetry, and these matrices do not exist. Both these cases are probable, having some strict logic behind them, which is difficult to check at the current level of computing power.

We do not exhaustively understand the nature of these maximum determinant matrices on the set of orthogonal matrices of the same order. But we do know something about their pattern invariants.

Simple and complex patterns in odd-order matrices

Maximum determinant matrices are square real-number matrices of order n with an extremal (largest in magnitude) determinant among the matrices with elements not exceeding 1 in magnitude.



■ Fig. 1. Portraits of Hadamard matrix of order 16 (left) and Belevitch matrix of order 46 (right)

As a matter of fact, at the moment there is no theory of such matrices. We will try to explain why.

Among them, there are few odd-order matrices with a simple structure. However, there is a presumption for orders equal to prime Fermat numbers, which is crucial for this topic. It is systemically important, so we will discuss it first.

Prime Fermat numbers are classical objects in the theory of numbers. The first five of them are $F_k = 3, 5, 17, 257, 65537$. Back in 1796, C. F. Gauss, in his student days, discovered an unexpected relation between them and geometrical figures. He inscribed a regular 17-gon in a circle and proved a more general proposition that if a regular polygon has a number of sides equal to a prime Fermat number, it can be constructed by means of compasses and ruler.

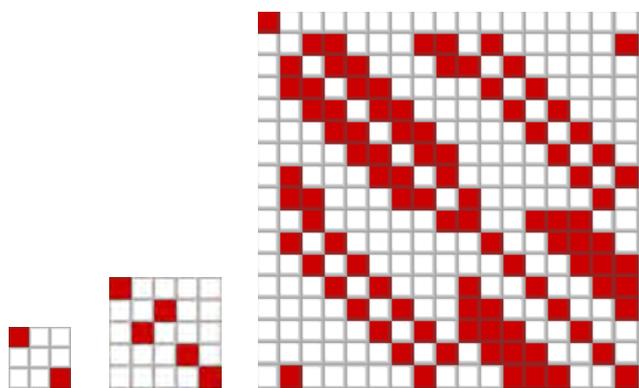
Conjecture. Fermat matrices of the form

$$F = \begin{pmatrix} -1 & \mathbf{e}^T & \mathbf{e}^T \\ \mathbf{e} & \mathbf{A} & \mathbf{B} \\ \mathbf{e} & \mathbf{B}^T & -\mathbf{A}^T \end{pmatrix}$$

built on a pair of cyclic blocks \mathbf{A} and \mathbf{B} with an edge vector \mathbf{e} , of orders equal to prime numbers $n = F_k = 3, 5, 17, 257, 65537, \dots$ are maximum determinant matrices. $F_{k-1}/(2F_k - 1)^{1/2} \times \mathbf{B}$, $\mathbf{B} = (n - 1)^{(n-1)/2}(2n - 1)^{1/2}$ is Barba's upper-bound estimation of the odd-order matrix determinant [19].

For matrices of the first three orders whose portraits are shown in Fig. 2, these patterns are easy to find with the use of a computer.

Gauss's theorem suggests that due to deep interdisciplinary links in mathematics, these matrices are of special importance, not less than the 17-gon had for Gauss. His discovery is highly valued by geometers and other specialists who want to see the integrity of a knowledge area artificially divided for the convenience of study. Here we add a new abstract example to it.



■ Fig. 2. Portraits of two circulant Fermat matrices of orders 3, 5, and 17

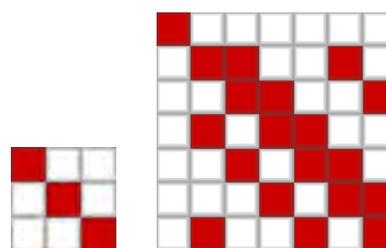
Note that if we discard the edge, the core of such matrices will be an Hadamard matrix. Regular Hadamard matrices of orders 8, 32, 64 similar to the above-mentioned matrices and their other versions easily found in a two circulant form by doubling Golay sequences [10] given in Table do not have this feature. We suppose that any other Hadamard matrices with any patterns in their portraits do not have it either. Therefore, the position taken by the Fermat matrices discussed above is absolutely unique. They are some kind of ideal "correct" matrices. They did not receive much attention so far because nobody thought about patterns. The competition was for uniqueness, not for residual universality.

It is easy to guess that the next to come are Mersenne numbers in the form $n = 2^k - 1$. Among them, there are many prime numbers, including 15 discussed above. However, they give us only two maximum determinant matrices with a prime structure: of orders 3 and 7. Their portraits are shown in Fig. 3. As we can see, they can be considered two circulant (splitting the core into 4 squares) or edged monocyclic ones. They have a maximum determinant, and here is the end of the Mersennian chain. It lacks the quality of the Fermat matrix chain.

Maximum determinant matrices are not orthogonal. Unlike Hadamard matrices, they are not constricted by the orthogonality condition, but they coincide with Hadamard matrices on evenly even orders. Therefore, Ryser's circulant structure border [20] which is set for Hadamard matrices at a

■ Golay sequences

Length n	Golay sequences
2	$\mathbf{a} = [1, 1], \mathbf{b} = [1, -1]$
10	$\mathbf{a} = [1, 1, -1, 1, -1, 1, -1, 1, 1]$ $\mathbf{b} = [1, 1, -1, 1, 1, 1, 1, 1, -1, -1]$
26	$\mathbf{a} = [1, 1, 1, 1, -1, 1, 1, -1, -1, 1, -1, 1, -1, -1, 1, -1, 1, -1, 1, 1]$ $\mathbf{b} = [1, 1, 1, 1, -1, 1, 1, -1, -1, 1, -1, 1, 1, 1, 1, -1, 1, -1, -1, 1, 1, 1, -1, -1]$



■ Fig. 3. Portraits of Mersenne matrices of orders 3 and 7

very low level (order 4), in this case is expanded up to three matrices of orders 3, 5, and 13. You can see their portraits in Fig. 4. Note that 5 and 13 are related to the key lengths of Golay sequences.

Now, it would be logical to presume that, by analogy with Gauss's theorem interpreted in an abstract way, there can be no simple pattern maximum determinant matrices.

There are certain starting orders in which you can meet some simple pattern matrices without violating the general principle. Besides, Barba [21] in his upper-bound estimation of maximum determinant noticed two chains of rapidly growing orders, so called "Barba's whiskers" in which integer matrices follow these integer border points (Gauss's points) by determinant.

Note that Fermat matrices do not go as far as this border. They were overlooked by researchers who focused on optimistic estimation to the detriment of common sense which suggested focusing on resolvable orders.

The orders found by Barba are not resolvable. The most well-known solution is the one of order 25 proposed by Raghavarao in [22]. It stimulated Andre Brouwer [23] to prove a theorem that there are no explicit solutions for them (we also think so). Brouwer proved that for a half of all orders a solu-

tion would consist of blocks with Hadamard matrix properties and an edge. Since the structure of these Hadamard matrices is not specified, this path leads to nowhere.

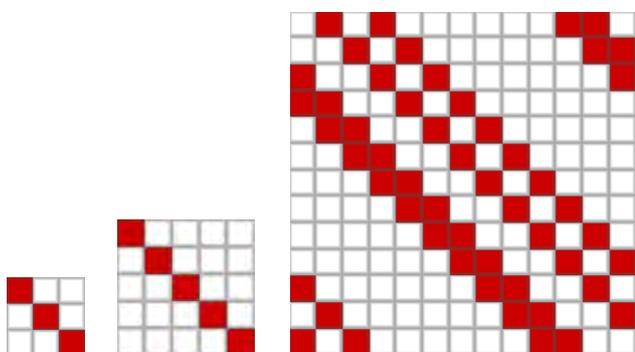
Still it is interesting to compare this path with Fermat matrices which exceed all such matrices. In order to complete our review of odd-order maximum determinant matrices, let us present the matrices of orders 9, 11, and 15 which were omitted so far. You can see their patterns in Fig. 5.

What we need to do now is to find the fundamental difference between these three matrix patterns. These matrices are unique, although all the three are constructed by the use of the same building block: Mersenne matrix of order 3 with a usual or inverted (in the third solution) diagonal when it forms a new block out of 1 or -1. Another difference is the block edge of different width, with lengthwise elements found in the Belevitch matrix of order 46.

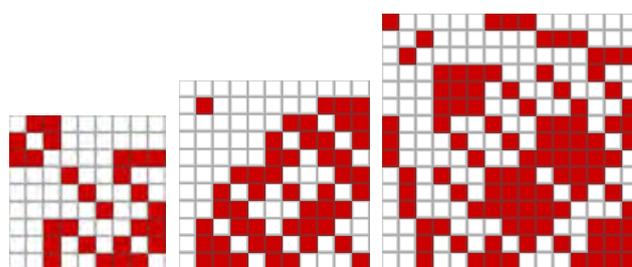
It is difficult to find solutions, but software with the self-describing name "Kaleidoscope" can synthesize them, dealing with a Mersenne matrix as if it were an Hadamard one. You still should remember how few good solutions can be found among Mersenne matrices.

Now let us consider matrices of the next odd orders, namely 19, 21, and 23. Their patterns are shown in Fig. 6.

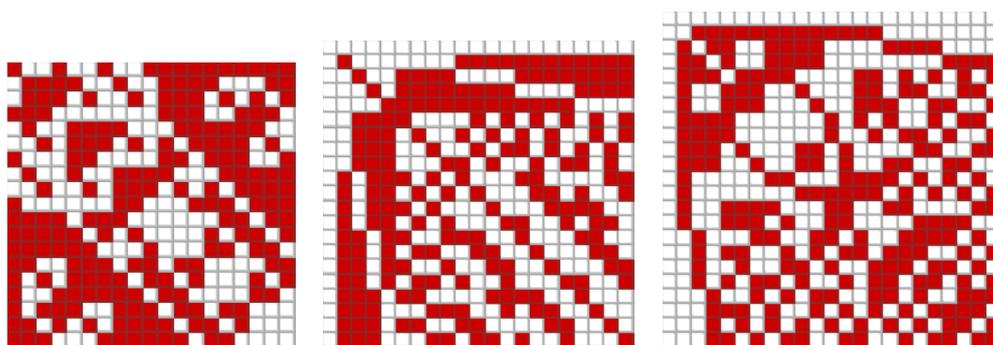
These matrices could be expected to be simpler because, unlike the previous ones, they all are



■ Fig. 4. Portraits of three cyclic versions of maximum determinant matrices



■ Fig. 5. Patterns of matrices of orders 9, 11, and 15



■ Fig. 6. Patterns of matrices of orders 19, 21, and 23

of prime orders. In their structure, we still find similar low-order blocks. However, in order 19, the starting block of size 10×10 containing 9 Mersenne matrices with a usual or inverted edge also contains an extra inversion of the first and then of every third element, both vertically and horizontally. Definitely, it is a new element, never met before. Possibly, this is the last symmetric matrix out of the considered ones. The matrix of order 21 is constructed from blocks of a higher order 4, and it is not generally symmetric. The matrix of order 23 is more complicated. We cannot be definite about the order of its blocks, or where its block edge is, or what can be considered as its core.

Complex patterns of matrices of Barba's orders

The basic inequality for determinants of matrices **A** of odd orders was suggested by Guido Barba [21] in the following form:

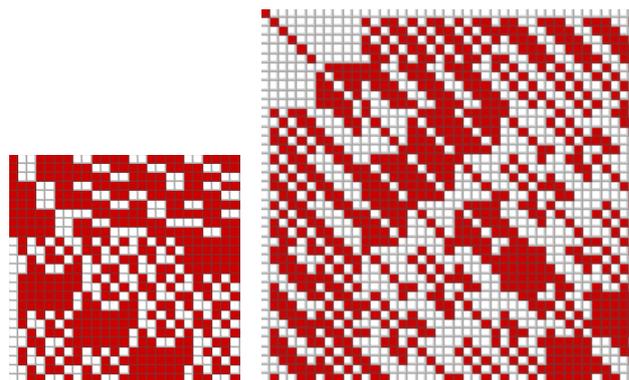
$$|\det(\mathbf{A})|^2 \leq \det((n - 1)\mathbf{I} + \mathbf{J}) = (n - 1)^{n-1}(2n - 1),$$

where **I** is an identity matrix, and **J** is a matrix of 1s.

Since the optimal matrix and its determinant are integer, the maximum is achievable in orders for which $2n - 1$ is a square of an integer. Otherwise, square rooting the expression on the right will give us an irrational number a fortiori greater than the determinant. Hence, the border is achievable in a sequence of orders described by a sum of squares of two close numbers $n = a^2 + (a + 1)^2$:

- 5, 13, 25 (Raghavarao),
- 41 (Bridges, Hall and Hayden),
- 61 (Brouwer),
- 85 (approximation by Solomon and Orrick),
- 113 (Brouwer), etc.

The names in brackets belong to the authors who were first to find matrices of the respective orders. These are matrices from two families, and their portraits are shown in Fig. 7, because the pattern complexity alternates every other order and matrices of orders 41, 85, etc. are much more difficult to find. Originally, these matrices were described by Raghavarao [22], and Andre Brouwer noticed that they were vitrages, i.e. matrices obtained by inserting blocks into Hadamard matrices. The principle of growing pattern complexity is true for them, too. The algorithm for constructing Barba's border matrices was described by Andre Brouwer and his followers [24] only roughly; to apply it, you need an Hadamard matrix, but the way to construct Hadamard matrices is not fully known. We also do not know Barba matrices of orders 145, 181, 221, 265, 313, 365, 421, 481, 545, 613, 685, etc.



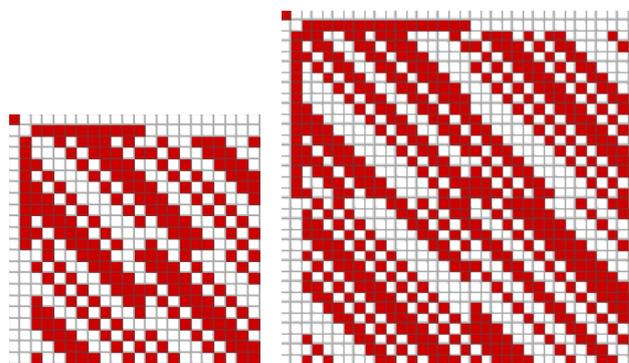
■ Fig. 7. Portraits of Barba matrices of orders 25 and 41

Thus, in each considered order, we can see an original element as a pattern part, and all these elements are unique. Our presumption (conjecture) is that this is a key feature of odd-order maximum determinant matrices everywhere outside the orders equal to Fermat numbers. This point dramatically reduces the chances to calculate them, lending weight to the alternative proposals about searching for matrices which achieve the maximum determinant on a preset structure borrowed from the optimal solutions for Fermat matrices [25, 26].

Pattern symmetry border

The presence of a symmetry border can be traced on two matrices of low even orders 22 and 34. The best pattern solution for orders 22, 34 and similar to them is based on the fact that $n - 1$ is not resolvable into a sum of two squares (except order 58), which gives us double-edged two circulant matrices. You can see the portraits of maximum determinant matrices of orders 22 and 34 in Fig. 8.

This structure is universal for Hadamard matrices. The symmetry border for edgeless two circulant



■ Fig. 8. Portraits of double-edged two circulant matrices of orders 22 and 34

matrices in Hadamard matrix orders was found in the critical order 32 [8]. In maximum determinant matrices, this border is shifted much farther for Barba's double orders, but is also limited (most likely, by order 82). The matrix of order 22 is unique in being the only doubly symmetric edged two circulant matrix whose both blocks **A** и **B** are symmetric.

Conclusion

The efficiency of an algorithm depends on the representation complexity of the mathematical object in the problem to be solved. The same object, for example a number, can be written either as a Roman number or in a positional notation. The latter provides simpler rules which, in old days, fostered the ideas of al-Khwarizmi whose name and whose book gave life to the words “algorithm” and “algebra”.

Such an obvious example of advantageous object representation stimulated the research of computational complexity conducted by some famous logicians of the 20th century: Kurt Gödel, Andrey

Markov, Pyotr Novikov, Alan Turing, and Alonzo Church.

Our current research is devoted to representation of maximum determinant matrices: this is the pattern whose properties are discussed in this article. We have specified several primary families, including matrices of orders equal to Mersenne or Fermat numbers, and made a presumption about the leading role of Fermat matrices based on their comparison with objects from a different field of mathematics studied by Gauss. Fermat matrices of high and super high orders can become another example of a discovery “with the point of your pen” due to finding a relation between objects of very different nature from different mathematical fields.

Financial support

The paper was prepared with the financial support of the Ministry of Science and Higher Education of the Russian Federation, grant agreement No. FSRF-2020-0004.

References

- Ahmed N., Rao R. *Orthogonal Transforms for Digital Signal Processing*. Berlin-Heidelberg-New York, Springer-Verlag, 1975. 263 p.
- Wang R. *Introduction to Orthogonal Transforms with Applications in Data Processing and Analysis*. Cambridge University Press, 2010. 504 p.
- Mironovskii L. A., Slaev V. A. The strip method of noise-immune image transformation. *Measurement Techniques*, 2006, no. 8, pp. 6–12.
- Sergeev M. B., Nenashv V. A., Sergeev A. M. Nested code sequences of Barker — Mersenne — Raghavarao. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 3, pp. 71–81 (In Russian). doi:10.31799/1684-8853-2019-3-71-81
- Vostrikov A. A., Sergeev M. B., Balonin N. A., Sergeev A. M. *Use of symmetric Hadamard and Mersenne matrices in digital image processing*. In: *Procedia Computer Science*, 2018. Pp. 1054–1061. doi:10.1016/j.procs.2018.08.042
- Sylvester J. J. Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. *London Edinburgh and Dublin Philos. Mag. and J. Sci.*, 1867, vol. 34, pp. 461–475.
- Hadamard J. Résolution d'une question relative aux déterminants. *Bulletin des Sciences Mathématiques*, 1893, vol. 17, pp. 240–246 (In French).
- Seberry J., Yamada M. *Hadamard Matrices: Constructions using Number Theory and Linear Algebra*. John Wiley & Sons, 2020. 352 p.
- Djokovic D. Z. *Some new symmetric Hadamard matrices*. arXiv:2101.05429 [math.CO]. Available at: <https://arxiv.org/abs/2101.05429>, free access (accessed 14 January 2021).
- Handbook of Combinatorial Designs*. Second Edition (Discrete Mathematics and its Applications). Charles J. Colbourn and Jeffrey H. Dinitz (Eds.). Chapman and Hall/CRC, 2006. 1000 p.
- Osborn J. H. *The Hadamard Maximal Determinant Problem*. Honours thesis. University of Melbourne, 2002. 144 p. [Electronic resource]. Available at: <http://maths-people.anu.edu.au/~osborn/publications/pubsall.html>, free access (accessed 22 November 2020).
- Neubauer M. G., and Radcliffe A. J. The maximum determinant of +1 matrices. *Linear Algebra and its Applications*, 1997, vol. 257, pp. 289–306.
- Seberry J., Xia T., Koukouvinos C., Mitrouli M. The maximal determinant and subdeterminants of ± 1 matrices. *Linear Algebra and its Applications*, 2003, vol. 373, pp. 297–310. *Combinatorial Matrix Theory Conference* (Pohang, 2002). doi:10.1016/S0024-3795(03)00584-6
- Orrick W. P. The maximal $\{-1,1\}$ -determinant of order 15. *Metrika*, 2005, no. 62, pp. 195–219.
- Brent R. P., Orrick W. P., Osborn J. H., Zimmermann P. *Maximal determinants and saturated D-optimal designs of orders 19 and 37* (preprint) [Electronic resource]. Available at: <http://arxiv.org/abs/1112.4160> (accessed 05 March 2020).
- Seberry J., Yamada M. *More on Maximal Determinant Matrices*. In: *Hadamard Matrices: Constructions using Number Theory and Linear Algebra*. John Wiley & Sons, 2020. Pp. 245–249.

17. Balonin N. A., Đoković D. Ž. Conference matrices from Legendre C-pairs. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 2–10. doi:10.31799/1684-8853-2020-4-2-10
18. Hall M. A survey of difference sets. *Proc Amer. Math. Soc.*, 1956, vol. 7, pp. 975–986.
19. Balonin N. A., Sergeev M. B., Vostrikov A. A. Prime Fermat numbers and maximum determinant matrix conjecture. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 2, pp. 2–9. doi:10.31799/1684-8853-2020-2-2-9
20. 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
21. Barba G. Intorno al teorema di Hadamard sui determinanti a valore Massimo. *Giorn. Mat. Battaglini*, 1933, vol. 71, pp. 70–86 (In Italian).
22. Raghavarao D. Some optimum weighing designs. *Ann. Math. Statist.*, 1959, no. 30, pp. 295–303.
23. Brouwer A. E. An infinite series of symmetric designs. *Math. Centrum Amsterdam Report ZW 202/83*, 1983, pp. 1–5.
24. Popescu O., Rose C., Popescu D. Maximizing the determinant for a special class of block-partitioned matrices. *Mathematical Problems in Engineering*, 2004, no. 1, pp. 49–61.
25. Balonin N. A., Sergeev M. B., Seberry J., Sinitsyna O. I. Circles on lattices and Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 3, pp. 2–9 (In Russian). doi:10.31799/1684-8853-2019-3-2-9
26. Balonin N. A., Sergeev M. B., Seberry J., Sinitsyna O. I. Circles on lattices, and maximum determinant matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 6, pp. 2–11 (In Russian). doi:10.31799/1684-8853-2020-6-2-11

УДК 519.614

doi:10.31799/1684-8853-2021-3-2-8

Рост сложности матриц максимума детерминанта

Ю. Н. Балонин^а, научный сотрудник, orcid.org/0000-0002-5102-4139

А. М. Сергеев^а, канд. техн. наук, доцент, orcid.org/0000-0002-4788-9869

А. А. Востриков^а, канд. техн. наук, доцент, orcid.org/0000-0002-8513-368, vostricov@mail.ru

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

Введение: несмотря на очевидную связь матриц максимума детерминанта четных (матриц Адамара) и нечетных порядков, последние имеют особенно сложные структуры, что делает их уникальными и привлекательными для различных применений в обработке визуальной информации, ее кодировании и маскировании. **Цель:** развитие теории матриц максимального детерминанта с опорой на использование компьютерного анализа, а также вычисление новых матриц с уникальными структурами орнамента их портретов. **Результаты:** выявлены особенности матриц максимального детерминанта, выделены их семейства на порядках, связанных с числами Ферма, показано усложнение структурных закономерностей в матрицах других порядков по мере их роста. Предположение об усложнении структур по мере роста порядков матриц подтверждается цепочкой демонстрируемых матричных портретов. Применительно к ортогональным матрицам Белевича следует, что они не могут быть найдены даже на таких небольших порядках, как 66 или 86.

Ключевые слова — детерминант, максимум детерминанта, квазиортогональные матрицы, матрицы Адамара, матрицы Мерсенна, матрицы Ферма, циклические матрицы.

Для цитирования: Balonin Yu. N., Sergeev A. M., Vostrikov A. A. The rise in maximum determinant matrix complexity. *Информационно-управляющие системы*, 2021, № 3, с. 2–8. doi:10.31799/1684-8853-2021-3-2-8

For citation: Balonin Yu. N., Sergeev A. M., Vostrikov A. A. The rise in maximum determinant matrix complexity. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 2–8. doi:10.31799/1684-8853-2021-3-2-8

Дополнение к статье

Sivak S. A., Royak M. E., Stupakov I. M., Voznuk E. S., Aleksashin A. S.

The implementation of the boundary element method to the Helmholtz equation of acoustics, опубликованной в № 2 за 2021 год:

Financial support

The reported study was funded by RFBR, project no. 19-31-90052.

Финансовая поддержка

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-31-90052.

Кодирование и устойчивость обработки сигналов в потоковых рекуррентных нейронных сетях

В. Ю. Осипов^а, доктор техн. наук, профессор, orcid.org/0000-0001-5905-4415, osipov_vasiliy@mail.ru

В. В. Никифоров^а, доктор техн. наук, профессор, orcid.org/0000-0002-6836-6303

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

Введение: при обосновании перспективных архитектур потоковых рекуррентных нейронных сетей возникает необходимость оценивания их устойчивости обработки различных входных сигналов. Для этого строятся диаграммы устойчивости, содержащие результаты имитационного моделирования для каждого из узлов этих диаграмм. На такое оценивание могут затрачиваться значительные временные и вычислительные ресурсы, особенно при анализе больших нейронных сетей. **Цель:** поиск методов оперативного построения таких диаграмм и оценивания устойчивости работы потоковых рекуррентных нейронных сетей. **Результаты:** анализ особенностей исследуемых диаграмм устойчивости показал, что узлы диаграмм группируются в сплошные зоны с одинаковыми характеристиками соотношения дефектов обработки входных сигналов. Предложен метод построения этих диаграмм на основе принципа обхода границ их зон. Таким путем можно избежать выполнения имитационного моделирования для внутренних узлов каждой зоны. Моделирование выполняется только для узлов, примыкающих к границам зон. Число узлов, для которых необходимо выполнять сеансы имитационного моделирования, снижается на порядок. Исследовано влияние видов кодирования входных сигналов на устойчивость потоковых рекуррентных нейронных сетей. Показано, что представление входных сигналов в виде последовательностей единичных образов с пересекающимися элементами может обеспечить большую устойчивость, чем без их пересечения.

Ключевые слова – рекуррентная нейронная сеть, устойчивость, моделирование, кодирование и обработка сигналов.

Для цитирования: Осипов В. Ю., Никифоров В. В. Кодирование и устойчивость обработки сигналов в потоковых рекуррентных нейронных сетях. *Информационно-управляющие системы*, 2021, № 3, с. 9–18. doi:10.31799/1684-8853-2021-3-9-18

For citation: Osipov V. Yu., Nikiforov V. V. Coding and robustness of signal processing in streaming recurrent neural networks. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 9–18 (In Russian). doi:10.31799/1684-8853-2021-3-9-18

Введение

Потоковые рекуррентные нейронные сети (РНС) ориентированы на непрерывную ассоциативную обработку различных сигналов. В качестве таких сигналов могут выступать предварительно преобразованные к требуемому виду видео- и звуковые потоки, информация от датчиков скорости, температуры, давления, расхода энергии и др. Эти РНС применимы для распознавания и фильтрации динамических образов, прогнозирования и восстановления прошедших событий, управления различными процессами. Такие сети могут наделяться различными пространственно-временными структурами. Они способны самообучаться, формировать модели воспринимаемых сигналов и использовать их для выработки творческих решений. В них реализованы возможности управления ассоциативным запоминанием и вызовом сигналов из памяти, а также переключения от решения одних задач к другим [1–5].

Несмотря на достигнутые результаты в исследовании потоковых РНС, их потенциал во многом не изучен, не снят ряд проблемных вопросов, свя-

занных с кодированием входных и внутренних сигналов, а также с обеспечением устойчивости функционирования этих нейронных сетей.

В общем случае перед подачей на вход потоковых РНС сигналов они должны быть преобразованы к входному языку сети. Используется принцип аналогии с биологическими системами. Примерами таких аналогий выступают преобразования оптических сигналов сетчаткой глаза, звуковых волн ворсинками слухового аппарата и др. В искусственных РНС, как и в естественных нейронных сетях, реализуется также внутрисетевое кодирование сигналов.

От реализуемых видов внешнего и внутрисетевого кодирования сигналов в потоковых РНС во многом зависит их устойчивость, а также возможности пространственно-временного связывания сигналов на элементах сети, запоминания и извлечения из нейросетевой памяти.

Известны некоторые принципы кодирования информации в биологических системах [6–9]. По аналогии с биологическими принципами предложены варианты предварительного кодирования сигналов перед подачей их в искусственные нейронные сети [1, 10–14].

В общем случае «сигналы раскладывают на пространственно-частотные составляющие. Каждую из этих составляющих преобразуют в последовательность единичных образов. Частоту и фазу следования этих импульсов задают как функции от амплитуды и фазы соответствующей пространственно-частотной составляющей. При таком преобразовании сигналы в дискретном времени представляются в виде последовательных совокупностей единичных импульсов (образов)» [4].

Реализуются и другие варианты [15–18]. Во всех случаях возможно задание различных параметров предварительного кодирования сигналов. Остается открытым вопрос: какой из вариантов предварительного кодирования сигналов является наилучшим?

Кроме этого, при обработке сигналов в потоковых РНС большое влияние на ее результаты оказывают также значения порогов возбуждения нейронов и уровень ассоциативного взаимодействия [19]. При несоблюдении ряда условий потоковые, как и другие РНС, могут перевозбуждаться или быть слишком заторможенными. Желательно за счет обоснования параметров потоковых РНС уже на этапе их проектирования обеспечивать последующую устойчивую работу при непрерывной обработке различных сигналов.

Для этого могут применяться известные методы анализа устойчивости РНС [1, 4, 19–21] при варьировании различными параметрами входных сигналов и самой сети. Среди этих методов применительно к потоковым РНС наибольший интерес представляет метод [4], позволяющий находить границы параметров устойчивого функционирования нейронных сетей. Однако в исходном варианте этот метод характеризуется большой вычислительной сложностью, что затрудняет применение его на практике. Необходим поиск новых решений, позволяющих снизить вычислительную сложность анализа устойчивости потоковых РНС и определиться с целесообразными подходами к кодированию обрабатываемых в них сигналов.

Особенности потоковых РНС и постановка задачи

Основу исследуемых потоковых РНС составляет двухслойная нейронная сеть с блоком управления [1]. Принципиальные отличия потоковых РНС от традиционных нейросетевых решений состоят в том, что первые наделяются прозрачными логическими структурами за счет реализации пространственных сдвигов сигналов (в виде последовательных совокупностей единичных образов (СЕО)) при передаче от слоя к слою [1]. За счет

этих сдвигов РНС могут быть наделены линейными, спиральными, петлевыми и другими логическими структурами [1–4]. В этих РНС используются импульсные нейроны с тремя состояниями: ожидание, возбуждение и невосприимчивость. Время невосприимчивости каждого нейрона после возбуждения больше, чем время задержки импульсов в двухслойных контурах сети [5]. Для каждого i -го нейрона на выходе $y_i(t + 1)$ формируется единичный образ по правилу

$$y_i(t + 1) = \begin{cases} 1, & \text{если } \sum_{j=1}^J w_{ij}(t) \cdot x_j(t) \geq Thr \text{ и } \Delta t_i > \Delta T; \\ 0 & \text{в других случаях,} \end{cases} \quad (1)$$

где J — число нейронов, связанных с i -м нейроном; $w_{ij}(t)$ — вес ij -го синапса, связывающего i -й нейрон с j -м нейроном; $x_j(t)$ — значение потенциала на выходе j -го нейрона на момент времени t ; Thr — порог возбуждения нейронов, находящихся в режиме ожидания; Δt_i — время, прошедшее после предыдущего возбуждения i -го нейрона; ΔT — время невосприимчивости нейрона после возбуждения.

Веса $w_{ij}(t)$ синапсов в такой РНС определяют через произведения весовых коэффициентов $k_{ij}(t)$, функций ослабления $\beta_{ij}(t)$, $\eta_{ij}(t)$ расходящихся и сходящихся единичных образов [4, 5]:

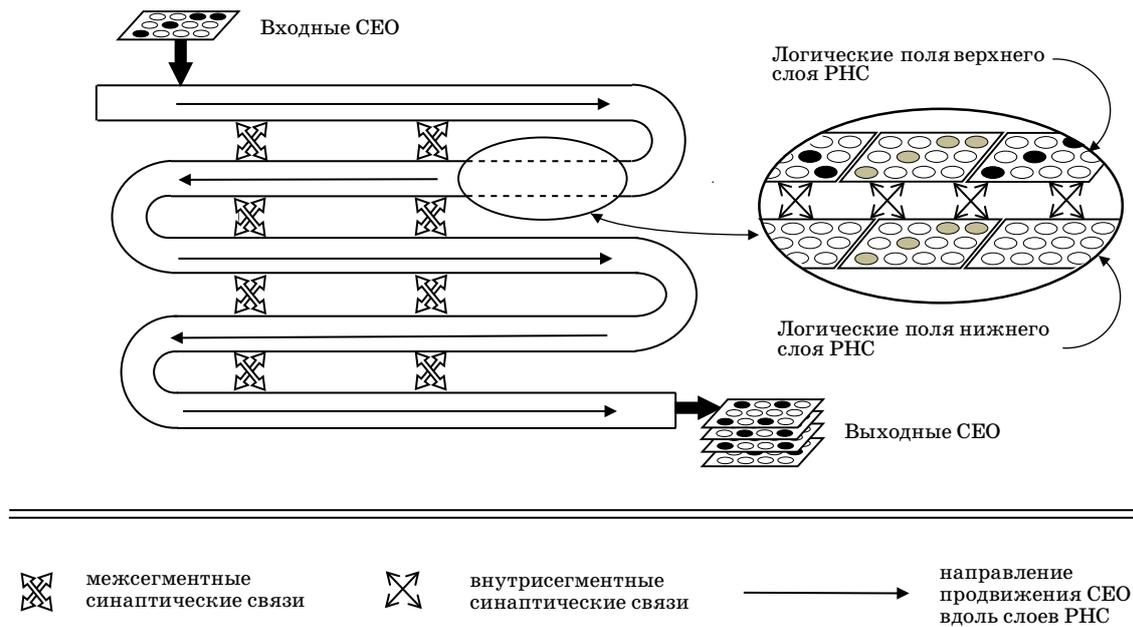
$$w_{ij}(t) = k_{ij}(t) \cdot \beta_{ij}(t) \cdot \eta_{ij}(t). \quad (2)$$

Весовые коэффициенты являются функциями прошедших через синапсы единичных импульсов [1, 18]. Функции ослабления зависят от параметров расходимости и сходимости этих импульсов, а также удаленности взаимодействующих нейронов с учетом принятого масштаба α . Изменяя $\beta_{ij}(t)$ или $\eta_{ij}(t)$ в зависимости от текущих состояний слоев, можно управлять ассоциативным взаимодействием обрабатываемых в РНС сигналов. Управлять также можно порогами возбуждения нейронов.

Заметим, что за счет приоритетности коротких связей между нейронами в потоковых РНС может обеспечиваться однозначное соответствие между входным и выходным пространством сети [1].

Пример спиральной структуры потоковой РНС, ориентированной на устранение искажений в последовательности входных СЕО, показан на рис. 1.

Внутренняя организация каждого из сегментов соответствует схеме в правой части рис. 1. Сегмент содержит два слоя нейронов — верхний и примыкающий к нему нижний слой. Каждый слой строится в виде последовательности прямоугольных полей нейронов. Отдельное поле имеет



■ **Рис. 1.** Пример структуры спиральной РНС
 ■ **Fig. 1.** An example of the structure of a spiral RNN

формат, идентичный формату кодов входного потока. В примере рис. 1 каждое поле содержит 12 нейронов, изображенных кружками (черный — нейрон возбужден, светлый — нейрон в состоянии ожидания, серый — нейрон в состоянии невосприимчивости). Конкретное поле верхнего слоя с непосредственно примыкающим к нему полем нижнего слоя составляет пару противоположащих полей; каждый сегмент представляет собой последовательность таких пар. Нейроны верхнего и нижнего слоя связываются внутрисегментными синапсами, обеспечивающими потактовое движение введенных СЕО от входного поля к выходу сети. Такое движение достигается должной настройкой параметров внутрисегментных синапсов.

При функционировании РНС могут наблюдаться дефекты типа 0 (пропуск единицы) и типа 1 (ложная единица).

Для конкретных условий применения потоковых РНС необходимо определить набор параметров сети, который обеспечивал бы требуемую безошибочность функционирования сети.

В работе [4] предложен метод, «позволяющий определять число $Q(l)$ успешных вариантов восстановления искаженной последовательности СЕО при l циклах обучения:

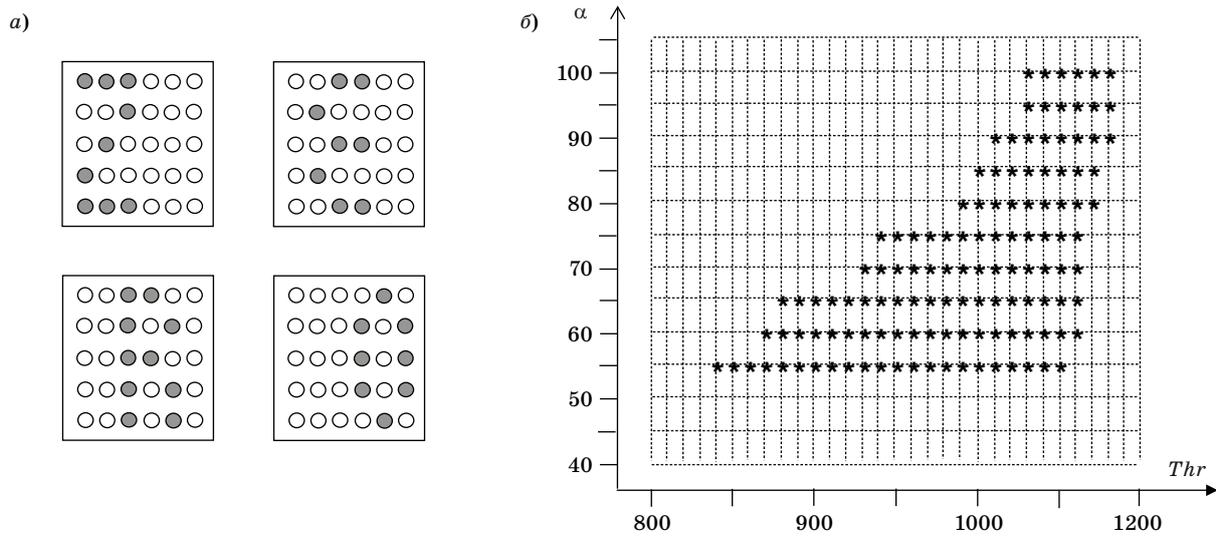
$$Q(l) = \sum_{z=1}^Z \sum_{k=1}^K \delta_{zk}(Thr_z, \alpha_k, l),$$

где $\delta_{zk}(Thr_z, \alpha_k, l)$ — булева функция успешного восстановления последовательности искажен-

ных СЕО, своеобразных кадров; α_k — величина масштаба расстояний между нейронами; Z — число градаций порога возбуждения нейронов; K — число градаций масштаба расстояний между нейронами. Заметим, что $\delta_{zk}(Thr_z, \alpha_k, l)$ определяется по результатам испытаний на конкретной модели РНС. В результате таких испытаний для каждого значения l формируется своя диаграмма $\delta_{zk}(Thr_z, \alpha_k, l)$ [4]. Эту диаграмму можно называть диаграммой относительной устойчивости РНС. Вариант такой диаграммы с набором СЕО для составления входной последовательности показан на рис. 2, а и б. Успешные восстановления последовательностей СЕО обозначены на диаграмме в виде символов *. Число таких восстановлений зависит от масштаба диаграммы и заданных дискретностей $\Delta\alpha, \Delta Thr$ параметров α, Thr .

В работе [4] для оценки величины $Q(l)$ реализуются сеансы моделирования для каждого из узлов диаграммы. Для построения этой диаграммы число сеансов моделирования исчисляется в сотнях единиц. Если учесть, что пространство комбинаций значений параметров настройки сети существенно многомерно, то для поиска комбинации, подходящей для конкретного набора входных СЕО, может потребоваться построение десятков и сотен тысяч рассматриваемых диаграмм. Это миллион сеансов моделирования.

Требуется разработать метод ускоренного построения таких диаграмм и анализа устойчивости потоковых РНС.



■ **Рис. 2.** Пример построения диаграммы $\delta_{zk}(Thr_z, \alpha_k, l)$: а — набор СЕО для составления входной последовательности; б — диаграмма

■ **Fig. 2.** An example of constructing a diagram $\delta_{zk}(Thr_z, \alpha_k, l)$: а — the collection of SSP for compiling an input sequence; б — the diagram

Метод построения диаграмм устойчивости потоковых РНС

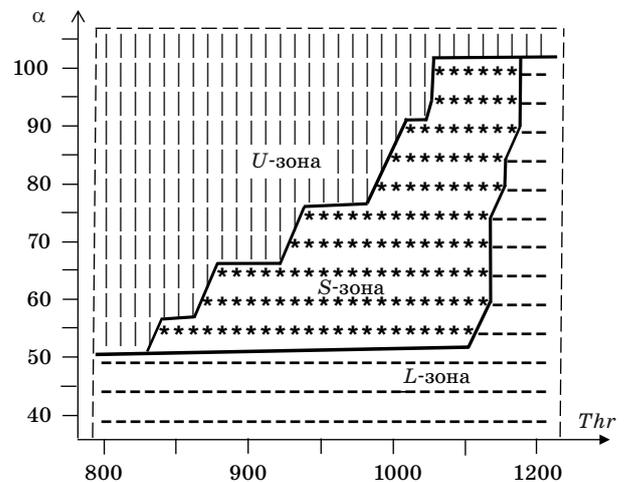
Для пояснения предлагаемого метода уточним особенности диаграммы $\delta_{zk}(Thr_z, \alpha_k, l)$, которая при дискретных значениях α, Thr может рассматриваться как решетка с конкретными узлами. При обработке заданной входной последовательности СЕО рекуррентной нейронной сетью с определенной структурой с различными вариантами значений α, Thr множество узлов, получаемых на диаграмме, можно разбить на три зоны. К ним относятся: *S*-зона, отражающая отсутствие возможных дефектов; *L*-зона, в которой число дефектов типа «0» больше числа дефектов типа 1; *U*-зона с числом дефектов типа 1, равным числу дефектов типа 0 либо превышающим его. Пример такой диаграммы приведен на рис. 3.

При моделировании потоковых РНС установлено, что если узел (α, Thr) принадлежит *U*-зоне, то при $\alpha' > \alpha$ узел (α', Thr) также входит в эту зону. Если узел (α, Thr) принадлежит *L*-зоне, то при $\alpha' < \alpha$ узел (α', Thr) относится к *L*-зоне. С учетом этого обстоятельства предлагается метод ускоренного оценивания значения относительной устойчивости РНС. Метод основан на построении нижней границы *U*-зоны и верхней границы *L*-зоны.

Построение нижней границы *U*-зоны предлагается выполнять в два этапа.

Этап 1. Поиск крайней правой верхней точки (α^*, Thr^*) , принадлежащей нижней границе *U*-зоны.

Если *U*-зона не пустая, то в силу принятых правил правый верхний узел находится либо на горизонтали от $(\alpha_{high}, Thr_{left})$ до $(\alpha_{high}, Thr_{right})$, либо на вертикали от $(\alpha_{low}, Thr_{right})$ до $(\alpha_{high}, Thr_{right})$. Этот этап требует не более чем $\max\{\alpha_{high}, Thr_{right}\}$ шагов моделирования. На каждом шаге этапа 1 моделирование выполняется для нового узла $(\alpha_{cur}, Thr_{cur})$. Устанавливаются значения $\alpha_{cur} = \alpha_{high}, Thr_{cur} = Thr_{right}$. Выполняется сеанс моделирования для узла $(\alpha_{high}, Thr_{right})$. Если этот узел принадлежит *U*-зоне, то искомый узел (α^*, Thr^*) находится на вертикали от $(\alpha_{low}, Thr_{right})$ до $(\alpha_{high}, Thr_{right})$. Иначе (α^*, Thr^*) находится на горизонтали от $(\alpha_{high}, Thr_{left})$ до $(\alpha_{high}, Thr_{right-1})$.



■ **Рис. 3.** Зоны *U, S* и *L* диаграммы

■ **Fig. 3.** Zones *U, S* and *L* of the diagram

При поиске по горизонтали в начале каждого очередного шага значение Thr_{cur} уменьшается на величину ΔThr шага решетки по координате Thr , $Thr_{cur} = Thr_{cur} - \Delta Thr$. Если новое значение Thr_{cur} меньше, чем Thr_{left} , то делается вывод, что в диаграмме U -зона отсутствует. Отсутствует и ее нижняя граница, а также искомый узел (α^* , Thr^*). В этом случае этап 1 завершен, переход к этапу 2 не требуется.

При $Thr_{cur} \geq Thr_{left}$ выполняется сеанс моделирования с новым значением Thr_{cur} . Если $(\alpha_{high}, Thr_{cur})$ принадлежит U -зоне, то $Thr^* = Thr_{cur}$, $\alpha^* = \alpha_{high}$ и полагается, что правый верхний узел ее границы найден. Это означает, что этап 1 завершен, выполняется переход к этапу 2. В противном случае, если узел $(\alpha_{high}, Thr_{cur})$ не принадлежит U -зоне, выполняется переход к следующему шагу этапа 1.

Поиск по вертикали выполняется аналогично. На каждом шаге Thr_{cur} остается равным Thr_{right} , значение α_{cur} уменьшается на $\Delta\alpha$. При $\alpha_{cur} < \alpha_{low}$ делается вывод, что U -зона занимает всю диаграмму. Переход к этапу 2 не требуется. Если $(\alpha_{cur}, Thr_{right})$ принадлежит U -зоне, то правый верхний узел ее границы найден, $Thr^* = Thr_{right}$, $\alpha^* = \alpha_{cur} - \Delta\alpha$, выполняется переход к этапу 2.

Этап 2. Поиск остальных узлов нижней границы U -зоны.

К первому шагу этапа 2 значения α_{cur}, Thr_{cur} представляют известные координаты α^*, Thr^* крайнего правого верхнего узла (α^*, Thr^*) нижней границы U -зоны, $\alpha_{cur} = \alpha^*, Thr_{cur} = Thr^*$. К каждому очередному шагу значения α_{cur}, Thr_{cur} — это найденные координаты следующего узла этой границы. Каждый очередной шаг этапа 2 обеспечивает определение координат следующего узла нижней границы U -зоны. К началу очередного шага уже известно, что узел $(\alpha_{cur}, Thr_{cur})$ является узлом границы U -зоны. Если $\alpha_{cur} \leq \alpha_{low}$ или $Thr_{cur} < Thr_{left}$, то построение границы U -зоны завершено.

Для определения координат следующего узла границы U -зоны в общем случае возможен один из трех вариантов: $(\alpha_{cur}, Thr_{cur} - \Delta Thr)$, $(\alpha_{cur} - \Delta\alpha, Thr_{cur} - \Delta Thr)$, $(\alpha_{cur} - \Delta\alpha, Thr_{cur})$. Если два предыдущих узла искомой границы лежат на одной горизонтали, то узел $(\alpha_{cur} - \Delta\alpha, Thr_{cur})$ не может принадлежать границе U -зоны. В этом случае выполняется сеанс моделирования для узла $(\alpha_{cur} - \Delta\alpha, Thr_{cur} - \Delta Thr)$. Если этот узел принадлежит U -зоне, то это очередной узел границы этой зоны. Координатам α_{cur}, Thr_{cur} присваиваются значения $\alpha_{cur} - \Delta\alpha, Thr_{cur} - \Delta Thr$, выполняется переход к следующему шагу этапа 2.

При условии, что два предыдущих узла искомой границы не лежат на одной горизонтали, выполняется сеанс моделирования для узла $(\alpha_{cur} - \Delta\alpha, Thr_{cur})$. Если этот узел принадлежит

U -зоне, то он принимается как очередной узел границы этой зоны. Координате α_{cur} присваивается значение $\alpha_{cur} - \Delta\alpha$, выполняется переход к следующему шагу этапа 2. В противном случае при $Thr_{cur} = Thr_{left}$ построение границы U -зоны завершено. При $Thr_{cur} > Thr_{left}$ выполняется моделирование для узла $(\alpha_{cur} - \Delta\alpha, Thr_{cur} - \Delta Thr)$. Если этот узел принадлежит U -зоне, то он принимается очередным узлом границы этой зоны и координатам α_{cur}, Thr_{cur} присваиваются, соответственно, значения $(\alpha_{cur} - \Delta\alpha, Thr_{cur} - \Delta Thr)$. В противном случае очередным узлом U -зоны принимается узел $(\alpha_{cur}, Thr_{cur} - \Delta Thr)$, координатам α_{cur}, Thr_{cur} присваиваются, соответственно, значения $\alpha_{cur}, Thr_{cur} - \Delta Thr$. В обоих случаях выполняется переход к следующему шагу этапа 2.

Аналогичным образом строится верхняя граница L -зоны. На этапе 1 поиск крайнего левого нижнего узла (α^*, Thr^*) этой зоны начинается с узла $(\alpha_{low}, Thr_{left})$. Если L -зона не пустая, то ее крайний левый нижний узел находится либо на горизонтали от $(\alpha_{low}, Thr_{left})$ до $(\alpha_{low}, Thr_{right})$, либо на вертикали от $(\alpha_{low}, Thr_{left})$ до $(\alpha_{high}, Thr_{left})$. Если $(\alpha_{high}, Thr_{left})$ принадлежит L -зоне, делается вывод, что эта зона занимает всю диаграмму и переход к этапу 2 не требуется. На этапе 2 последовательность узлов, принадлежащих границе L -зоны, строится снизу вверх (в противоположность границе U -зоны, которая строится сверху вниз).

Когда границы U -зоны и L -зоны построены, значение относительной устойчивости РНС определяется как число узлов диаграммы, лежащих между этими границами.

Предложенный метод вместо $N_\alpha \cdot N_{Thr}$ требует не более чем $2(N_\alpha + N_{Thr})$ выполнений сеанса моделирования, где N_α и N_{Thr} — соответственно число строк и столбцов формируемой диаграммы относительной устойчивости РНС.

Исходные данные и результаты моделирования

С использованием предложенного метода осуществлялось оценивание относительной устойчивости для двухслойной РНС с числом нейронов в каждом слое 360 единиц. Каждый слой разбивался на $6 \times 2 = 12$ логических полей размером 6×5 нейронов. Продвижение обрабатываемых совокупностей единичных образов вдоль слоев выполнялось по спирали. В отличие от рис. 1 в РНС формировалось не пять, а два сегмента. При определении весов синапсов (2) весовые коэффициенты $k_{ij}(t)$ определялись как

$$k_{ij}(t) = b \cdot \text{th}(g_{ij}(t)), g_{ij}(t) = g_{ij}(t-1) \pm \Delta g_{ij}(t-1),$$

где $\text{th}(\cdot)$ — гиперболический тангенс; $g_{ij}(t)$ — параметр, зависящий от числа прошедших через ij -й синапс единичных импульсов (образов) на момент времени t ; b — положительный коэффициент.

Функция ослабления $\beta_{ij}(t)$ расходящихся единичных образов в (2) задавалась в виде $\beta_{ij}(t) = 1 - r_{ij}(t)/(\alpha + r_{ij}(t))$, где $r_{ij}(t)$ — расстояние между i -м и j -м нейроном; α — масштаб расстояний между нейронами, являющийся одной из координат рассматриваемых диаграмм. Значение функции $\eta_{ij}(t)$ в (2) считалось равным единице.

В РНС подавались наборы непересекающихся входных кодов, которые приведены в трех следующих вариантах:

$$\mathbf{A}_1 = \begin{pmatrix} 110000 \\ 110000 \\ 000000 \\ 000000 \\ 000000 \end{pmatrix}, \mathbf{A}_2 = \begin{pmatrix} 000011 \\ 000011 \\ 000000 \\ 000000 \\ 000000 \end{pmatrix},$$

$$\mathbf{A}_3 = \begin{pmatrix} 000000 \\ 000000 \\ 000000 \\ 110000 \\ 110000 \end{pmatrix}, \mathbf{A}_4 = \begin{pmatrix} 000000 \\ 000000 \\ 000000 \\ 000011 \\ 000011 \end{pmatrix}; \quad (3)$$

$$\mathbf{B}_1 = \begin{pmatrix} 101000 \\ 000000 \\ 101000 \\ 000000 \\ 000000 \end{pmatrix}, \mathbf{B}_2 = \begin{pmatrix} 000000 \\ 010100 \\ 000000 \\ 010100 \\ 000000 \end{pmatrix},$$

$$\mathbf{B}_3 = \begin{pmatrix} 000000 \\ 001010 \\ 000000 \\ 001010 \\ 000000 \end{pmatrix}, \mathbf{B}_4 = \begin{pmatrix} 000000 \\ 000000 \\ 000101 \\ 000000 \\ 000101 \end{pmatrix}; \quad (4)$$

$$\mathbf{C}_1 = \begin{pmatrix} 100100 \\ 000000 \\ 000000 \\ 100100 \\ 000000 \end{pmatrix}, \mathbf{C}_2 = \begin{pmatrix} 000000 \\ 010010 \\ 000000 \\ 000000 \\ 010010 \end{pmatrix},$$

$$\mathbf{C}_3 = \begin{pmatrix} 000000 \\ 001001 \\ 000000 \\ 000000 \\ 001001 \end{pmatrix}, \mathbf{C}_4 = \begin{pmatrix} 010010 \\ 000000 \\ 000000 \\ 010010 \\ 000000 \end{pmatrix}. \quad (5)$$

Вариант (3) относится к плотным кодам, (4) — к раздвинутым кодам, (5) — к широким кодам. В (3)–(5) нули соответствуют отсутствию единичных образов. Нежирными единицами отмечены единичные образы, которые перед вводом в сеть искажаются на нули. Жирными единицами обозначены активные единичные образы.

Согласно (3)–(5) выполняются условия: $\mathbf{A}_i \circ \mathbf{A}_j = \mathbf{Z}$, $\mathbf{B}_i \circ \mathbf{B}_j = \mathbf{Z}$, $\mathbf{C}_i \circ \mathbf{C}_j = \mathbf{Z}$, где \circ — поэлементное умножение матриц при i, j , не равном j ; \mathbf{Z} — нулевая матрица. Заметим, что в примерах (3)–(5) в каждом из входных кодов имеется 13,3 % активных элементов, из которых искажается половина.

Результаты построения искомых диаграмм для вариантов входных кодов (3)–(5) приведены в табл. 1, из которой следует, что для более плотных кодов значения относительной устойчивости Q работы потоковой РНС наибольшие. Эта тенденция может объясняться тем, что в режиме обучения при обработке более плотных кодов формируются более насыщенные синаптические связи. Отметим, что в диаграммах (см. табл. 1) нет U -зон. Подобные диаграммы могут возникать в случае отсутствия входных пересечений входных кодов. Для построения подобных диаграмм, в которых заведомо отсутствуют U -зоны, допустимо использовать упрощенный алгоритм ускоренной оценки значений относительной устойчивости. Достаточно ограничиться построением верхней границы L -зоны. В общем случае для выполнения ускоренного алгоритма построения диаграммы требуется число сеансов моделирования равно числу узлов, лежащих на ее периметре. При упрощенном варианте этого алгоритма число сеансов моделирования уменьшается вдвое.

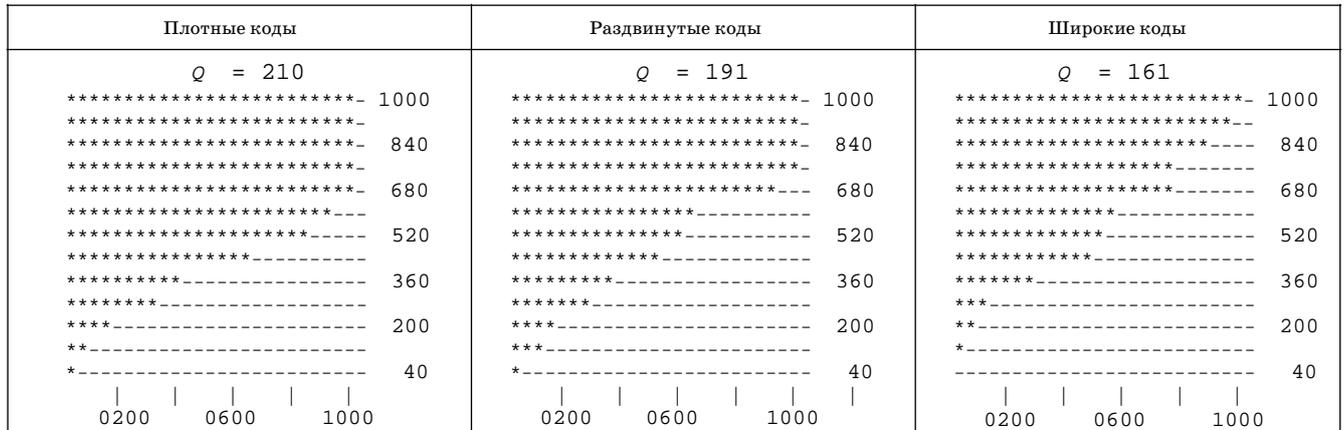
Выражения

$$\mathbf{A}_1^* = \begin{pmatrix} 110000 \\ 110000 \\ 000000 \\ 000000 \\ 000000 \end{pmatrix}, \mathbf{A}_2^* = \begin{pmatrix} 000000 \\ 011000 \\ 011000 \\ 000000 \\ 000000 \end{pmatrix},$$

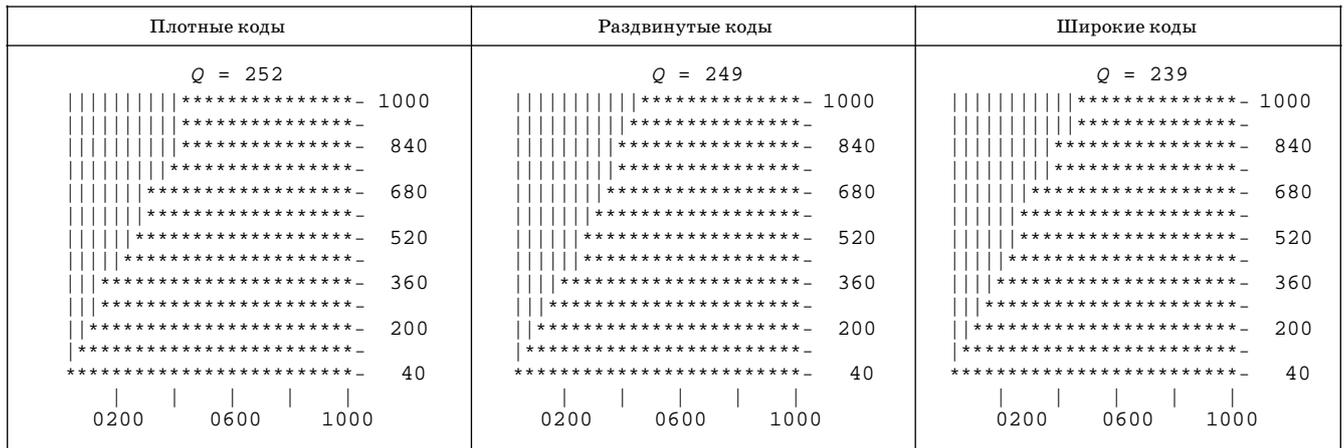
$$\mathbf{A}_3^* = \begin{pmatrix} 000000 \\ 000000 \\ 001100 \\ 001100 \\ 000000 \end{pmatrix}, \mathbf{A}_4^* = \begin{pmatrix} 000000 \\ 000000 \\ 000000 \\ 000110 \\ 000110 \end{pmatrix}; \quad (6)$$

$$\mathbf{B}_1^* = \begin{pmatrix} 001000 \\ 100000 \\ 001000 \\ 100000 \\ 000000 \end{pmatrix}, \mathbf{B}_2^* = \begin{pmatrix} 000000 \\ 000010 \\ 001000 \\ 000010 \\ 001000 \end{pmatrix},$$

■ Таблица 1. Диаграммы относительной устойчивости РНС для вариантов (3)–(5) кодирования входных сигналов
 ■ Table 1. Diagrams of the relative RNN stability for (3)–(5) coding variants of input signals



■ Таблица 2. Диаграммы относительной устойчивости РНС для вариантов (6)–(8)
 ■ Table 2. Diagrams of the relative RNN stability for (6)–(8) coding variants



$$B_3^* = \begin{pmatrix} 000000 \\ 000000 \\ 000101 \\ 000000 \\ 001010 \end{pmatrix}, B_4^* = \begin{pmatrix} 000000 \\ 000100 \\ 000001 \\ 000100 \\ 000001 \end{pmatrix}; \quad (7)$$

$$C_1^* = \begin{pmatrix} 000100 \\ 100000 \\ 000000 \\ 000100 \\ 100000 \end{pmatrix}, C_2^* = \begin{pmatrix} 000000 \\ 100100 \\ 000000 \\ 000000 \\ 010010 \end{pmatrix},$$

$$C_3^* = \begin{pmatrix} 010000 \\ 000010 \\ 000000 \\ 010000 \\ 000010 \end{pmatrix}, C_4^* = \begin{pmatrix} 000000 \\ 010010 \\ 000000 \\ 000000 \\ 001001 \end{pmatrix} \quad (8)$$

соответствуют другим трем вариантам кодов, но уже с пересекающимися единичными образами: $A_1^* \circ A_2^* \neq Z$, $A_2^* \circ A_3^* \neq Z$, $A_3^* \circ A_4^* \neq Z$, $B_1^* \circ B_2^* \neq Z$, $B_2^* \circ B_3^* \neq Z$, $B_3^* \circ B_4^* \neq Z$, $C_1^* \circ C_2^* \neq Z$, $C_2^* \circ C_3^* \neq Z$, $C_3^* \circ C_4^* \neq Z$.

В диаграммах табл. 2, в отличие от табл. 1, присутствуют U -зоны. Однако L -зона практически отсутствует. Такое обстоятельство может объясняться тем, что в режиме обучения в ходе обработки пересекающихся кодов для небольших значений α формируются сравнительно более насыщенные синаптические связи. В результате в левом верхнем углу диаграммы появляется U -зона, а L -зона в правом нижнем углу исчезает. При этом значения относительной устойчивости РНС для (6)–(8) выше, чем для (3)–(5).

Для данных, приведенных в табл. 1 и 2, нет необходимости выполнять полный алгоритм ускоренного построения диаграмм относительной устойчивости. Достаточно реализовать

упрощенный алгоритм. В случае табл. 1 достаточно найти границу U -зоны, для табл. 2 — границу L -зоны. Для столбцов «Плотные коды», «Раздвинутые коды» и «Широкие коды» табл. 1 потребовалось выполнить соответственно 34, 35 и 38 сеансов моделирования. Для аналогичных столбцов табл. 2 потребовалось 23, 24 и 24 сеанса моделирования. Полный алгоритм ускоренного построения диаграмм на этих же данных предусматривает выполнение 76 сеансов моделирования вместо 325 сеансов для известного метода [4].

Заключение

Полученные новые результаты по свойствам потоковых РНС состоят в следующем.

Предложен метод ускоренного оценивания значения относительной устойчивости РНС, который позволяет в $N_\alpha \cdot N_{Thr} / 2(N_\alpha + N_{Thr})$ раз уменьшить число реализуемых сеансов модели-

рования. Метод основан на построении границ интересующей зоны, содержащей все положительные решения. Полученные результаты моделирования подтверждают преимущества предложенного метода по сравнению с известными решениями.

Показано, что на относительную устойчивость РНС существенное влияние оказывает принятый вариант кодирования входных сигналов. Установлено, что эта устойчивость повышается, если входные последовательные сигналы будут иметь частичные пересечения элементов. Она также растет с повышением плотности используемых кодов. С учетом этого при поиске целесообразных архитектур РНС необходимо их согласовывать с возможными вариантами кодирования входных сигналов.

Предложенный метод рекомендуется применять при исследовании перспективных моделей РНС, которые в дальнейшем могут реализовываться аппаратным путем.

Литература

- Osipov V., Osipova M. Space-time signal binding in recurrent neural networks with controlled elements. *Neurocomputing*, 2018, no. 308, pp. 194–204. <https://doi.org/10.1016/j.neucom.2018.05.009>
- Osipov V., Nikiforov V., Zukova N., Miloserdov D. Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers. *Neural Computing and Applications*, 2020, no. 32, pp. 14885–14897. <https://doi.org/10.1007/s00521-020-04843-5>
- Osipov V., Kuleshov S., Zaytseva A., Levonevskiy D., Miloserdov D. Neural network forecasting of news feeds. *Expert Systems with Applications*, 2021, no. 169, Article 114521. <https://doi.org/10.1016/j.eswa.2020.114521>
- Осипов В. Ю., Никифоров В. В. Возможности рекуррентных нейронных сетей с управляемыми элементами по восстановлению потоков кадров. *Информационно-управляющие системы*, 2019, № 5, с. 10–17. doi:10.31799/1684-8853-2019-5-10-17
- Осипов В. Ю., Милосердов Д. И. Нейросетевое прогнозирование событий для роботов с непрерывным обучением. *Информационно-управляющие системы*, 2020, № 5, с. 33–42. doi:10.31799/1684-8853-2020-5-33-42
- Rolls E., Treves A. The neural encoding of information in the brain. *Progress in Neurobiology*, 2011, no. 95, pp. 448–490. doi:10.1016/j.pneurobio.2011.08.002
- Rozzi S., Fogassi L. Neural coding for action execution and action observation in the prefrontal cortex and its role in the organization of socially driven behavior. *Frontiers in Neuroscience*, 2017, no. 11, Article 492. <https://doi.org/10.3389/fnins.2017.00492>
- Kong Q., Han J., Zeng Yi., Xu B. Efficient coding matters in the organization of the early visual system. *Neural Networks*, 2018, no. 108, pp. 218–226. <https://doi.org/10.1016/j.neunet.2018.04.019>
- Kim T., Bair W., and Pasupathy A. Neural coding for shape and texture in macaque area V4. *The Journal of Neuroscience*, 2019, no. 39, pp. 4760–4774. <https://doi.org/10.1523/JNEUROSCI.3073-18.2019>
- Gerstner W., Kistler W. M. *Spiking neuron models*. Cambridge University Press, 2002. <https://doi.org/10.1017/CBO9780511815706>
- Kreiman G. Neural coding: computational and biophysical perspectives. *Physics of Life Reviews*, 2004, no. 1, pp. 71–102. <https://doi.org/10.1016/j.plrev.2004.06.001>
- Cessac B., Paugam-Moisy H., Vieville T. Overview of facts and issues about neural coding by spikes. *Journal of Physiology — Paris*, 2010, no. 104, pp. 5–18. <https://doi.org/10.1016/j.jphysparis.2009.11.002>
- Palm G. Neural associative memories and sparse coding. *Neural Networks*, 2013, no. 37, pp. 165–171. <https://doi.org/10.1016/j.neunet.2012.08.013>
- Ruys de Perez A., Matusевич L. F., Shiu A. Neural codes and the factor complex. *Advances in Applied Mathematics*, 2020, no. 114, Article 101977. <https://doi.org/10.1016/j.aam.2019.101977>
- Zhengzhong L., Schwartz D., Ditzler G., Koyluoglu O. The impact of encoding-decoding schemes and weight normalization in spiking neural networks. *Neural Networks*, 2018, no. 108, pp. 365–378. <https://doi.org/10.1016/j.neunet.2018.08.024>
- Azarfar A., Calcini N., Huang C., Zeldenrust F., Celikel T. Neural coding: A single neuron’s perspective. *Neuroscience & Biobehavioral Reviews*, 2018, no. 94,

- pp. 238–247. <https://doi.org/10.1016/j.neubiorev.2018.09.007>
17. Guo W., Fouda M. E., Eltawil A. M., and Salama K. N. Neural coding in spiking neural networks: A comparative study for robust neuromorphic systems. *Frontiers in Neuroscience*, 2021, no. 15, Article 638472. <https://doi.org/10.3389/fnins.2021.638472>
18. Zeldenrust F., Wadman W., Englitz B. Neural coding with bursts—current state and future perspectives. *Frontiers in Computational Neuroscience*, 2018, vol. 12, Article 48. <https://doi.org/10.3389/fncom.2018.00048>
19. Haykin S. *Neural networks and learning machines*. Third Ed. New-York, Prentice Hall, 2008. 938 p.
20. Aouiti C., M’hamdi M. S., Chérif F., Alimi A. M. Impulsive generalized high-order recurrent neural networks with mixed delays: Stability and periodicity. *Neurocomputing*, 2018, no. 321, pp. 296–307. <https://doi.org/10.1016/j.neucom.2017.11.037>
21. Ozcan N. New conditions for global stability of neutral-type delayed Cohen – Grossberg neural networks. *Neural Networks*, 2018, no. 106, pp. 1–7. <https://doi.org/10.1016/j.neunet.2018.06.009>

UDC 004.8

doi:10.31799/1684-8853-2021-3-9-18

Coding and robustness of signal processing in streaming recurrent neural networks

V. Yu. Osipov^a, Dr. Sc., Tech., Professor, orcid.org/0000-0001-5905-4415, osipov_vasily@mail.ru

V. V. Nikiforov^a, Dr. Sc., Tech., Professor, orcid.org/0000-0002-6836-6303

^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: When substantiating promising architectures of streaming recurrent neural networks, it becomes necessary to assess their stability in processing various input signals. For this, stability diagrams are constructed containing the results of simulation for each of the nodes of these diagrams. Such an estimation can be time-consuming and computationally intensive, especially when analyzing large neural networks. **Purpose:** Search for methods of quick construction of such diagrams and assessing the stability of streaming recurrent neural networks. **Results:** Analysis of the features of the stability diagrams under study showed that the nodes of the diagrams are grouped into continuous zones with the same ratio characteristics of the input signal processing defects. With this in mind, the article proposes a method for constructing these diagrams based on bypassing the boundaries of their zones. With this approach, you do not have to perform simulation for the interior nodes of each zone. The simulation should be performed only for the nodes adjacent to zone boundaries. Due to this, the number of nodes for which you need to perform simulation sessions is reduced by an order of magnitude. The influence of the input signal coding types on the streaming recurrent neural network stability has been investigated. It is shown that the representation of input signals in the form of sequences of single pulses with intersecting elements can provide greater stability as compared to pulses without any intersection.

Keywords — recurrent neural network, stability, simulation, coding and signal processing.

For citation: Osipov V. Yu., Nikiforov V. V. Coding and robustness of signal processing in streaming recurrent neural networks. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 9–18 (In Russian). doi:10.31799/1684-8853-2021-3-9-18

References

- Osipov V., Osipova M. Space-time signal binding in recurrent neural networks with controlled elements. *Neurocomputing*, 2018, no. 308, pp. 194–204. <https://doi.org/10.1016/j.neucom.2018.05.009>
- Osipov V., Nikiforov V., Zukova N., Miloserdov D. Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers. *Neural Computing and Applications*, 2020, no. 32, pp. 14885–14897. <https://doi.org/10.1007/s00521-020-04843-5>
- Osipov V., Kuleshov S., Zaytseva A., Levonevskiy D., Miloserdov D. Neural network forecasting of news feeds. *Expert Systems with Applications*, 2021, no. 169, Article 114521. <https://doi.org/10.1016/j.eswa.2020.114521>
- Osipov V. Yu., Nikiforov V. V. Recurrent neural networks with controlled elements in restoring frame flows. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 5, pp. 10–17 (In Russian). doi:10.31799/1684-8853-2019-5-10-17
- Osipov V. Yu., Miloserdov D. I. Neural network event forecasting for robots with continuous training. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 5, pp. 33–42 (In Russian). doi:10.31799/1684-8853-2020-5-33-42
- Rolls E., Treves A. The neural encoding of information in the brain. *Progress in Neurobiology*, 2011, no. 95, pp. 448–490. doi:10.1016/j.pneurobio.2011.08.002
- Rozzi S., Fogassi L. Neural coding for action execution and action observation in the prefrontal cortex and its role in the organization of socially driven behavior. *Frontiers in Neuroscience*, 2017, no. 11, Article 492. <https://doi.org/10.3389/fnins.2017.00492>
- Kong Q., Han J., Zeng Yi., Xu B. Efficient coding matters in the organization of the early visual system. *Neural Networks*, 2018, no. 108, pp. 218–226. <https://doi.org/10.1016/j.neunet.2018.04.019>
- Kim T., Bair W., and Pasupathy A. Neural coding for shape and texture in macaque area V4. *The Journal of Neuroscience*, 2019, no. 39, pp. 4760–4774. <https://doi.org/10.1523/JNEUROSCI.3073-18.2019>
- Gerstner W., Kistler W. M. *Spiking neuron models*. Cambridge University Press, 2002. <https://doi.org/10.1017/CBO9780511815706>
- Kreiman G. Neural coding: computational and biophysical perspectives. *Physics of Life Reviews*, 2004, no. 1, pp. 71–102. <https://doi.org/10.1016/j.plrev.2004.06.001>
- Cessac B., Paugam-Moisy H., Vieville T. Overview of facts and issues about neural coding by spikes. *Journal of Physiology — Paris*, 2010, no. 104, pp. 5–18. <https://doi.org/10.1016/j.jphysparis.2009.11.002>
- Palm G. Neural associative memories and sparse coding. *Neural Networks*, 2013, no. 37, pp. 165–171. <https://doi.org/10.1016/j.neunet.2012.08.013>

14. Ruys de Perez A., Matusevich L. F., Shiu A. Neural codes and the factor complex. *Advances in Applied Mathematics*, 2020, no. 114, Article 101977. <https://doi.org/10.1016/j.aam.2019.101977>
15. Zhengzhong L., Schwartz D., Ditzler G., Koyluoglu O. The impact of encoding-decoding schemes and weight normalization in spiking neural networks. *Neural Networks*, 2018, no. 108, pp. 365–378. <https://doi.org/10.1016/j.neunet.2018.08.024>
16. Azarfar A., Calcini N., Huang C., Zeldenrust F., Celikel T. Neural coding: A single neuron's perspective. *Neuroscience & Biobehavioral Reviews*, 2018, no. 94, pp. 238–247. <https://doi.org/10.1016/j.neubiorev.2018.09.007>
17. Guo W., Fouda M. E., Eltawil A. M., and Salama K. N. Neural coding in spiking neural networks: A comparative study for robust neuromorphic systems. *Frontiers in Neuroscience*, 2021, no. 15, Article 638472. <https://doi.org/10.3389/fnins.2021.638474>
18. Zeldenrust F., Wadman W., Englitz B. Neural coding with bursts—current state and future perspectives. *Frontiers in Computational Neuroscience*, 2018, vol. 12, Article 48. <https://doi.org/10.3389/fncom.2018.00048>
19. Haykin S. *Neural networks and learning machines*. Third Ed. New-York, Prentice Hall, 2008. 938 p.
20. Aouiti C., M'hamdi M. S., Chérif F., Alimi A. M. Impulsive generalized high-order recurrent neural networks with mixed delays: Stability and periodicity. *Neurocomputing*, 2018, no. 321, pp. 296–307. <https://doi.org/10.1016/j.neucom.2017.11.037>
21. Ozcan N. New conditions for global stability of neutral-type delayed Cohen – Grossberg neural networks. *Neural Networks*, 2018, no. 106, pp. 1–7. <https://doi.org/10.1016/j.neunet.2018.06.009>

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, что снижает 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 сверху справа на стартовой странице): <https://orcid.org>

Method and algorithm for determining the geographic coordinates of ground objects from an unmanned aerial vehicle

A. V. Parfiryev^a, PhD, Tech., Associate Professor, orcid.org/0000-0003-4112-896X

O. V. Parfiryeva^a, Head of the Training Laboratory Department, orcid.org/0000-0001-7097-8897

A. V. Dushkin^{a,b}, Dr. Sc., Tech., Associate Professor, orcid.org/0000-0002-8078-8971, a_dushkin@mail.ru

^aMilitary Educational and Scientific Center of the Air Forces – Zhukovskii and Gagarin Air Forces Academy, 54a, Staryh Bolshevikov St., 394064, Voronezh, Russian Federation

^bNational Research University of Electronic Technology, Shokin Sq., 1, 124498, Moscow, Zelenograd, Russian Federation

Introduction: As practice shows, the accuracy of determining the coordinates of objects is influenced by many factors associated with the presence of errors in measuring the angular coordinates of the optical system, the distance to the object and the presence of an inhomogeneous terrain. **Purpose:** Improving the accuracy of determining the geographic coordinates of ground objects from an unmanned aerial vehicle. **Results:** A method and an algorithm for determining geographic coordinates based on the use of a digital terrain model and optimization methods have been developed. The accuracy of calculating the coordinates of the object is increased by minimizing the error in measuring the declination angle, azimuth to the target and slant range. To confirm the analytical calculations, a field experiment was carried out with a car on the ground. At a considerable distance, at which the slant range was 900 m, several data freeze frames were taken. As a result of calculations, the geographical coordinates of the car were obtained in two ways (traditional and developed). Ultimately, the accuracy of calculating coordinates using the developed method is 4.8 times higher. **Practical relevance:** The method and algorithms for information processing proposed in the work will make it possible to create a number of hardware and software solutions for guidance and target designation systems.

Keywords – digital image processing, correlation filtering, PID controller, optimization algorithm, optoelectronic system.

For citation: Parfiryev A. V., Parfiryeva O. V., Dushkin A. V. Method and algorithm for determining the geographic coordinates of ground objects from an unmanned aerial vehicle. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 19–28. doi:10.31799/1684-8853-2021-3-19-28

Introduction

Robotics and unmanned aircraft are currently some of the most promising areas of research. Unmanned aerial vehicles (UAVs) have become widespread in both the civilian and military segments and perform an increasing range of tasks, which include the following: monitoring the environment and ground objects, extinguishing fires, video filming, delivery of goods, organizing sports, competitions, etc. [1, 2]. One of the determining factors in the development of unmanned aviation is the growing role of UAVs in organizing aerial reconnaissance and striking targets in conditions of local conflicts and potential conflicts of a global scale in the future [3, 4].

Analysis of the experience of the combat use of UAV groups in military conflicts of recent years, in particular, in Syria, Libya and Nagorno-Karabakh in the period 2018–2020 revealed some of the shortcomings of optoelectronic systems (OES). One of the disadvantages of such systems, operating on the basis of inertial sensors, is a large error in determining the coordinates of the location of ground objects for issuing target indications for their destruction,

due to the error in determining the angles of declination and azimuth. This disadvantage is associated with the presence of instrumental errors of inertial sensors, the accuracy of their installation in the block of sensing elements and the quality of setting the initial coordinates of the strapdown inertial navigation system (SINS). The main types of error manifestation, which are an additive component of measurement errors: change in the position of zero signals; emission of random components; conversion errors of scale factors [5, 6]. The multiplicative component is the scale factor error and manifests itself in the nonlinearity of the transformation of physical measurements into output parameters. In this case, the presence of random noise is always observed in the output signals of the sensors [5]. Thus, a constant increase in the total error is achieved during a prolonged UAV flight, which directly affects the calculation of the horizontal distance to a ground object, especially at large distances (over 5 km) and at large elevation angles. Another drawback that affects the accuracy of determining geographic coordinates is the error in measuring the slant range to a ground object by a laser rangefinder at large distances.

There are a sufficient number of methods and methods for determining the coordinates of ground objects from UAVs, but not all of them meet the safety requirements for their use in the conditions of enemy air defense.

1. With the help of digital image processing and a number of auxiliary systems, including satellite radio navigation and inertial [7]. The disadvantage of this method is the need for shooting in nadir and the impossibility of using it at night.

2. Using an OES with a digital image processing unit installed on board an unmanned aerial vehicle [8, 9]. The author of these works uses ideal azimuth and elevation angles as initial data without taking into account the instrumental error in their measurements, which leads to inaccuracies in calculating the geographic coordinates of ground objects, especially at large distances relative to the UAV. Also, the disadvantage of this approach is the impossibility of taking into account the errors in determining the satellite coordinates of the UAV associated with the shift in the coordinates of the satellite or the clock of the receiver and the satellite [10, 11].

3. With the use of active and passive radars, the disadvantages of which are the cumbersomeness of the equipment, high cost, limitation in terms of targets and visibility [12–15].

4. Based on the use of laser radiation energy in conjunction with SINS and satellite radio navigation system. The disadvantage of this method is the low degree of sampling in the construction of the terrain and, accordingly, the drift of inertial sensors.

Improving the accuracy of determining geographic coordinates can be achieved in the following ways.

1. The use of expensive high-precision SINS, satellite navigation systems with differential modes, entailing the development of new high-grade optoelectronic systems. In turn, the development of new systems intended for installation on UAVs involves taking into account a number of restrictions on the mass of the payload, dimensions, etc.

2. Using additional information about the terrain (digital elevation model) to refine the coordinates issued by the module for calculating geographic coordinates and motion parameters of the captured object.

Formulation of the problem

It is required to determine the geographic coordinates of a moving ground object using an OES installed on a UAV (helicopter or aircraft type). The UAV must be in the air at an altitude and distance from the object, which ensures the correct

operation of the OES measuring instruments in accordance with their technical characteristics. The ground object should be located in the line of sight of a day or night camera and not be overlapped by other objects. Meteorological conditions in the OES coverage area contribute to the maximum visibility range of a ground object.

A UAV with an OES must include the following necessary equipment: a gyro-stabilized platform; day camera; infrared night camera (wavelength from 7 to 14 μm); laser rangefinder; GLONASS or GPS receiver; onboard computer.

Requirements for OES: viewing angle: 360 degrees; the presence of digital image stabilization; a gyro-stabilized system based on a gyroscope and an accelerometer of medium accuracy based on microelectromechanical systems (MEMS) with active stabilization to provide a clear image from cameras in motion and during vibrations; the presence of a three-axis magnetic declination sensor; laser radiation wavelength: 0.9 μm ; the wavelength of the received electromagnetic radiation by the infrared camera: 7–14 μm ; the wavelength of the received electromagnetic radiation by the day camera: 0.5–0.9 μm ; an adjusted system of day and night cameras and a laser rangefinder.

Determining the coordinates of moving ground objects from photographs obtained by a day or night camera with a UAV involves the solution of three particular problems:

1) tracking a ground object by video sequence to determine its center of coordinates in the image coordinate system;

2) development of an automatic tracking system based on a gyro-stabilized platform;

3) development of a method for determining geographic coordinates based on a digital elevation model (DEM).

Let's consider options for solving these problems.

Algorithm for tracking video sequence

Before tracking an object, it must be detected. Such localization of a ground object can be automatic or manual. Machine learning algorithms based on Convolutional Neural Networks (CNN) can be used for automatic localization. The process of automatic localization in this case can be conditionally divided into several parts: selection of the most effective neural network model; creation of a special data structure; preparation of a training sample; training the model. To work with regional convolutional neural networks, you can use the open source TensorFlow machine learning computer library from Google, where the OpenCV library takes over the main work of capturing images from a camera

and digital processing. To solve the problem of localizing ground objects in the image with simultaneous classification, the following CNN topologies can be used: R-CNN, Fast R-CNN, Faster R-CNN, YOLO, SSD, FPN, and RetinaNet [16]. To increase the number of classified objects, the neural network is retrained on the basis of one of the above models. Variants of learning algorithms are described in [17, 18]. The result of applying the Faster R-CNN model after the training process on a series of aerial photographs of aircraft and helicopters is shown in Fig. 1.

The bottom line is pretty good. However, in addition to expensive and energy-consuming equipment, neural network approaches have the following serious disadvantages regarding the UAV power supply system.

1. Even on modern video cards it is difficult to achieve high performance from networks. For comparison, the YoloV3 classifier network with a relatively simple architecture can produce an image in 50 ms [19].

2. The object of interest may not be included in the nomenclature of the neural network classifier. It must be retrained on new datasets (dataset), which leads to a decrease in efficiency in decision making.

Due to the fact that the result of the operation of the neural network is not always stable, the object of interest, including when it has a complex shape, may be missed in some frames at different shooting angles. This situation can lead to failures in the system of automatic guidance of the optical axis of the OES camera to a ground object to determine its coordinates. For more stable tracking of a ground object, it is necessary to use other algorithms called

optical trackers — ASEF (Average of Synthetic Exact Filters) and MOSSE (Minimum Output Sum of Squared Error), and to search and classify objects in the image, a neural network.

These two algorithms have a lot in common. They differ only in the method of filter averaging and in that MOSSE creates better filters than ASEF [19].

The general view of the MOSSE filter is represented by the expression:

$$\hat{W}_i^* = \frac{A_i}{B_i}, \quad (1)$$

where

$$A_i = \eta \sum_{p=1}^P \hat{G}_p \odot \hat{F}^{P*} + (1-\eta)A_{i-1};$$

$$B_i = \eta \sum_{p=1}^P \hat{F}^p \odot \hat{F}^{P*} + (1-\eta)B_{i-1}.$$

Here A_i and B_i are filter components; $\hat{F}^p \odot \hat{F}^{P*}$ — the energy spectrum of the rectangle from the p -th image; \hat{G}_p — frequency matrix of the desired ideal response; η — learning rate during tracking; P — number of transformations.

Filter (1) is elementwise multiplied by the frequency matrix \hat{F} , obtained from the next frame and preprocessed. Preprocessing includes non-linear transformation of image pixels (logarithm), which will bring too light and too dark areas to the average. View normalization will work best here:

$$I = \text{sign}(I - 127) \sqrt{|I - 127|}. \quad (2)$$



■ Fig. 1. The result of the detection and classification of objects in the photograph

Next, you need to apply the expression:

$$A = \exp \frac{(x-i)^2 + (y-j)^2}{\sigma^2}. \quad (3)$$

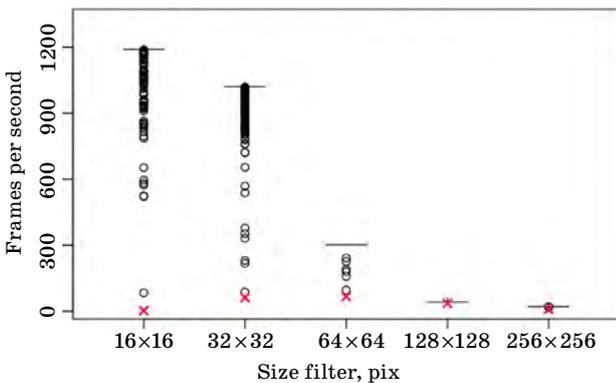
You can also use a 2D Henning window to smoothly zero out the pixel values at the edges of the image. After receiving $\hat{\mathbf{W}}_i^*$ the response is calculated in the frequency domain:

$$\hat{\mathbf{G}}_{\text{conv}} = \hat{\mathbf{F}} \odot \hat{\mathbf{W}}^*. \quad (4)$$

Then the inverse Fourier transform is applied to it to transform it into the spatial domain and the coordinates of the maximum in the resulting \mathbf{G}_{conv} matrix are found. The coordinates found indicate the new position of the object. The object being tracked may disappear from the frame or be covered by another object. To identify such situations, it is necessary to analyze in more detail the response received from the last frame. To do this, find the maximum g_{max} of the matrix \mathbf{G}_{conv} , then the 11×11 square is excluded from the center of this matrix, and the mathematical expectation μ_G and the standard deviation σ_G are calculated from the remaining elements. Calculate the Peak to Sidelobe Ratio (PSR) using the formula

$$PSR = \frac{g_{\text{max}} - \mu_G}{\sigma_G} \quad (5)$$

and, if this value is above a certain threshold, then the object is considered detected. The value of this threshold for stable detection is 20 or more. In Fig. 2 shows the data on the frame refresh rate (frames per second) for various filter sizes, due to the increase in the computational load on the on-board computer processor. The software implementation of the presented optical tracker MOSSE is presented in [20]. Application of this approach allows obtaining initial data that can be interpreted into control signals of an automatic control system (ACS) by a



■ Fig. 2. Frame refresh rate (fps) for different filter sizes

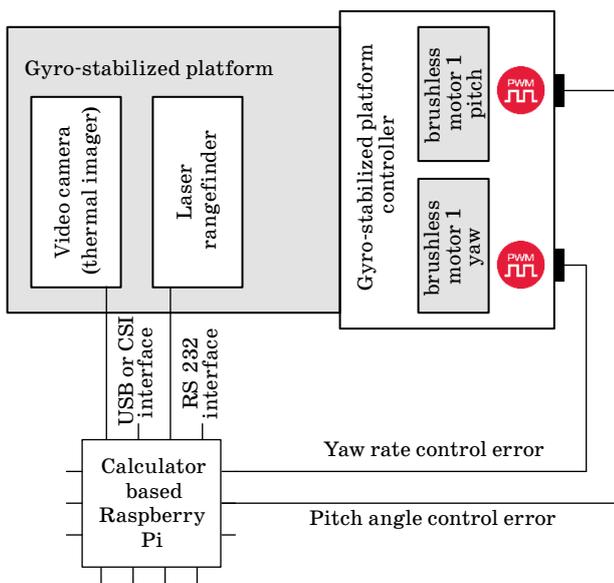
gyro-stabilized UAV platform for aligning the line of sight of the optical system with the direction to the tracked ground object.

Automatic tracking system based on a gyro-stabilized platform

The alignment of the line of sight of the OES camera is carried out by changing two angular coordinates — pitch and yaw. On board the UAV, as a rule, there is already a gyro-stabilized platform with a controller. The controller is a computer with many digital inputs and outputs for external control, reading signals from inertial sensors, issuing control signals for brushless motors, current state and operating mode. We will consider the implementation of the control method using the example of the BaseCam SimpleBGC 32-bit controller with an open programming interface for building external control modules. For data exchange, the controller provides control and serial interfaces — UART and I2C, as well as digital outputs, working on receiving a pulse-width modulation (PWM) signal to control the angular coordinates of the platform. To organize platform control in order to align the line of sight with the direction to the captured ground object, it is necessary to send a PWM signal to the corresponding inputs of the RC_PITCH and RC_YAW controller. To generate control signals, a Raspberry Pi-based calculator is required. The general diagram of an automatic object tracking device, which includes a three-axis gyro-stabilized platform with a control controller, a video camera (thermal imager), a laser rangefinder, and a Raspberry Pi platform for image processing and generation of control signals, is shown in Fig. 3.

The signals from the calculator are directly proportional to the deviation of the center of the object from the center of the frame and represent two signals of misalignment in the x and y coordinates. The frame center coordinates are the reference signal. To compensate for mismatch signals, a device and a tracking-type automatic control system have been developed, the functional diagram of which is shown in Fig. 4 and 5.

In Fig. 4 introduced the following designations: $x_o(t)$, $y_o(t)$ — reference signal, respectively, in the coordinates x and y ; $x_m(t)$, $y_m(t)$ — the current signal, respectively, along the x and y coordinates; CorDev1, CorDev2 — sequential correcting device, respectively, along the coordinates x (diagram Fig. 5, a) and y (diagram Fig. 5, b); CtrlDev — control device for converting a PWM signal into voltage; Amp — amplifying device; ExDev1, ExDev2 — actuators (platform motors); CtrlObj — control object (platform with a video camera or thermal imager);



■ Fig. 3. Diagram of the automatic tracking system of the object and the controller of the gyro-stabilized platform

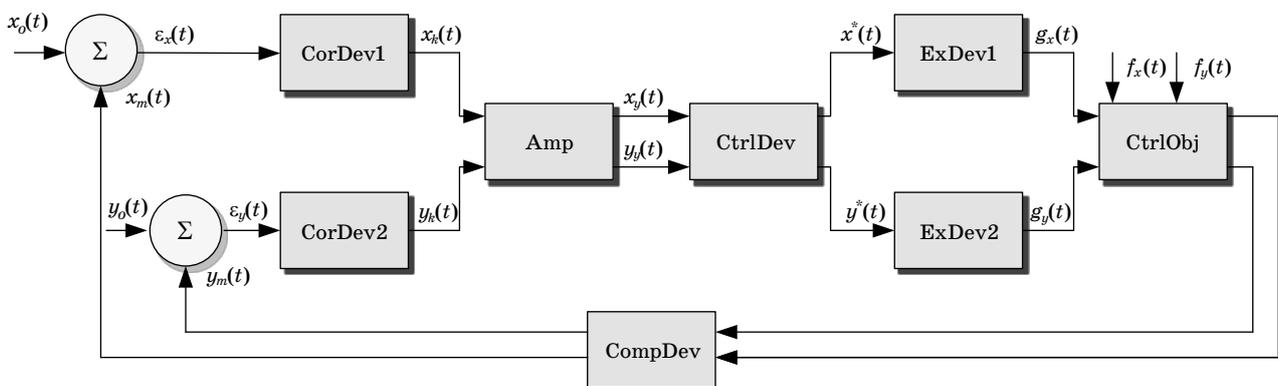
CompDev is a computing device that issues control signals to the controller of the gyro platform [21].

The purpose of tracking is the most accurate and fast alignment of the line of sight of the camera and the beam of the laser rangefinder with the direc-

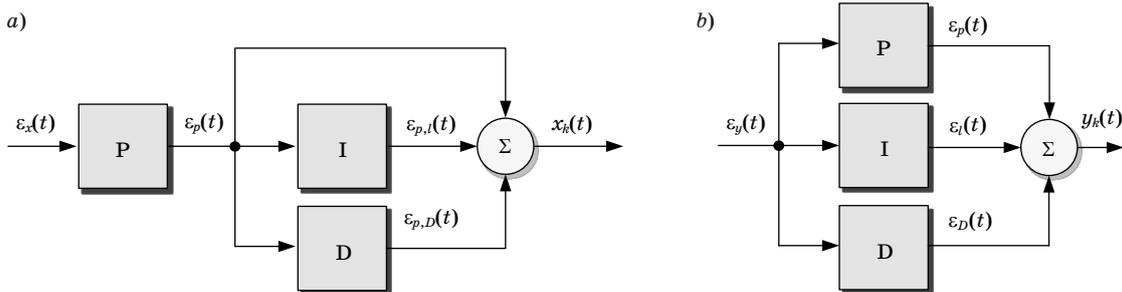
tion to the object. This is how the correct reading of the initial data is achieved for the implementation of the algorithm for determining the coordinates of ground objects from the UAV using the video stream.

Method for determining geographic coordinates based on digital elevation model

The solution to the third particular problem is based on the use of DEMs obtained using the methods of classical digital photogrammetry in conjunction with computer vision algorithms. For this, the advanced Agisoft Metashape software was used. Its advantage is that machine learning technologies are used in the analysis and post-processing, which leads to an increase in the accuracy of the results obtained. With the help of Metashape, during image processing, images are converted into dense point clouds, and a digital elevation model is built only from those points of a dense cloud that are classified as points of the earth's relief. A terrain model can be created in the form of a triangular irregular network, in the form of a digital elevation model (DEM), and also as a set of contour lines (isohypsum) records with the required elevation of the relief section. The terrain model in the form of a triangular irregular network has higher resolu-



■ Fig. 4. Functional diagram of the developed tracking type ACS



■ Fig. 5. Functional diagrams of correcting devices based on the PID controller: a — coordinate x; b — coordinate y

tion in areas where the surface is extremely irregular, and lower resolution in areas with a uniform surface. DEM accuracy can reach 16 cm at a scale of 1:500. The DEM interval (grid step) can reach 0.003 arc seconds (0.1 m), depending on the technical characteristics of the camera, shooting altitude, weather conditions, etc. A terrain model in the form of a DEM is presented in the following formats: GeoTIFF, Arc / Info ASCII Grid (ASC), Band interleaved (BIL), XYZ or as tiles in the Sputnik KMZ format. The DEM format is accepted by all modern programming languages, including Python. On the Internet, you can find services that allow you to get DEM in one click (for example, *Open Aerial Map*. <https://map.openaerialmap.org/>). Fig. 6 is an illustration of this method.

Geographic coordinates are calculated by solving a multiparameter optimization problem based on an optimization algorithm that minimizes the functional of the residual, which is the square of the difference between the height of the object, calculated by the goniometric-rangefinder method h_c , and the height of the object from the DEM h^* :

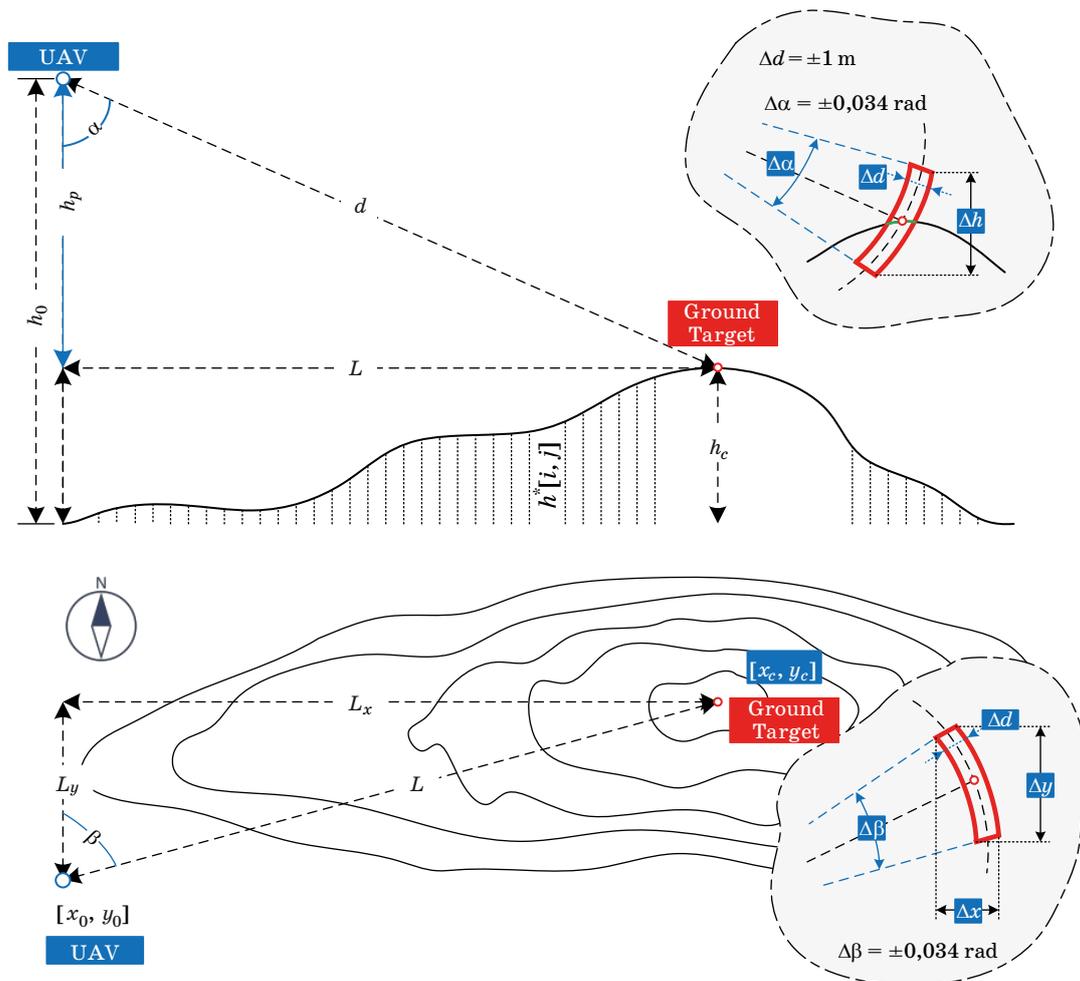
$$\mathfrak{Z}(\alpha, \beta, d) : \min_{\mathfrak{Z}} \|h_c(\alpha, d) - h^*(x_0 + L_x, y_0 + L_y)\|_2; \quad (6)$$

$$\begin{cases} L_x = d \cdot \sin(\alpha) \cdot \sin(\beta); \\ L_y = d \cdot \sin(\alpha) \cdot \cos(\beta); \end{cases} \quad (7)$$

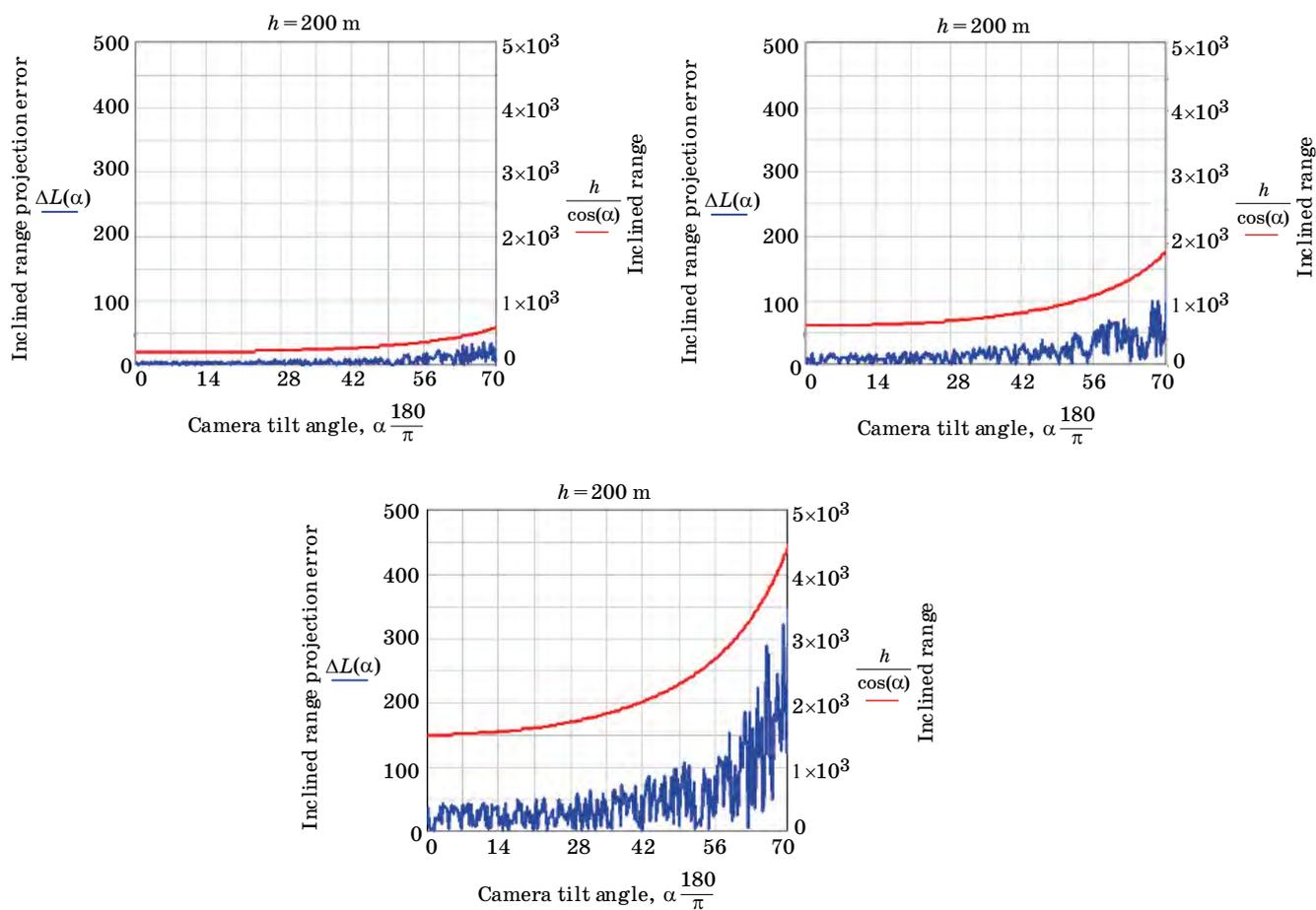
$$h_c = h_0 - d \cdot \cos(\alpha). \quad (8)$$

In expression (6): α, β — current angles of sight and azimuth to the ground object, obtained from the controller of the UAV's gyro-stabilized platform; d — slant range to a ground object, measured by a laser rangefinder; L is the horizontal range to a ground object; x_0, y_0, h_0 — geographic coordinates of the UAV location (latitude, longitude, altitude) obtained using a GLONASS or GPS receiver.

The accuracy of calculating the coordinates of the object is increased by minimizing the error in measuring the angle α , azimuth to the ground object β and slant range d . In accordance with analytical calculations, the error in determining the coordinates of a ground object by the traditional method is within the red zone. The error in deter-



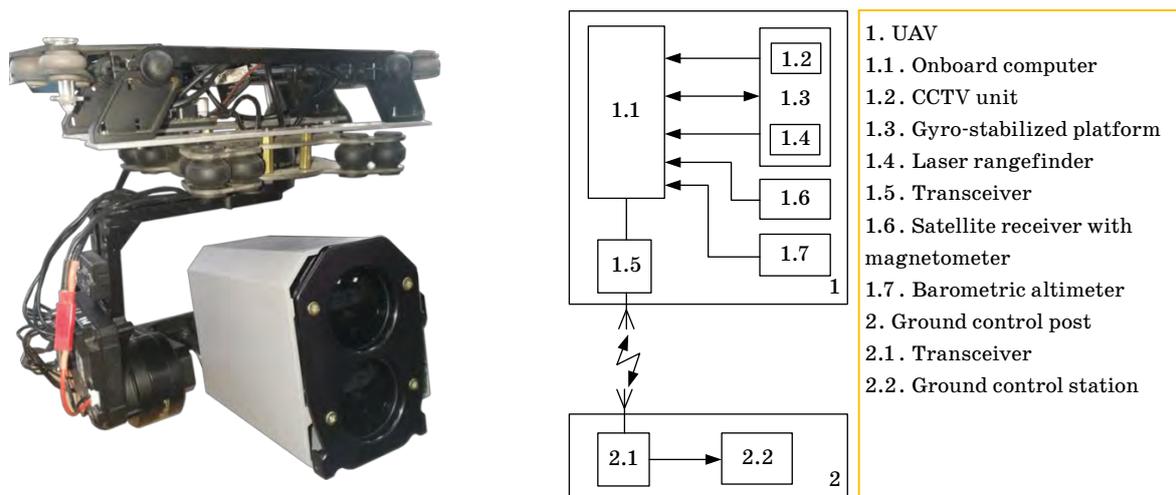
■ Fig. 6. Illustration of a method for determining the coordinates of ground objects



■ Fig. 7. Graphs of horizontal range error distribution



■ Fig. 8. Evaluation of the calculation of geographic coordinates in two ways



■ Fig. 9. General view of the ECO and its block diagram

mining the coordinates by the developed method lies at the intersection of the red zone with the horizontal surface. The dynamic values of the analytically calculated error in measuring the horizontal range, depending on the angle of sight of the OES and the height of the UAV, are shown in Fig. 7.

To confirm the analytical calculations, an experiment was carried out with a car on the ground. At a considerable distance, at which the slant range was 900 m, several data freeze frames were taken. As a result of calculations, the geographical coordinates of the car were obtained in two ways (traditional — red circles and developed — blue circles, Fig. 8).

Ultimately, the accuracy of calculating coordinates by the developed method was increased by 4.8 times. A general view of the OES installed on a drone-type UAV and its block diagram implementing this method is shown in Fig. 9.

Conclusion

Thus, an algorithm and a method for determining the geographic coordinates of ground objects based on the OES installed on board the UAV have been developed. This makes it possible to more efficiently and accurately calculate the coordinates of any ground object visible in the visible and infrared ranges of the electromagnetic spectrum from the UAV in a time mode close to real. The method is based on an innovative approach to using the DEM obtained using the methods of classical digital photogrammetry in conjunction with modern computer vision algorithms. The degree of discreteness of the DEM, which is achieved by the quality of photographs, the resolution of the camera, the height of the survey, and computational costs, will have a direct impact on the estimation of the calculation

of the coordinates of a ground object. The results of the experiment show that the accuracy of calculating the coordinates by the developed method can be increased by 4.8 times in comparison with the traditional goniometric-rangefinder method.

Reference

1. Kartenichev A. Yu., Panfilova E. V. Fire extinguishing technologies using unmanned aerial vehicles. *Pozharnaya bezopasnost': problemy i perspektivy*, 2019, no. 10, pp. 149–151 (In Russian).
2. Makarov K. S. Block diagram of a complex with an unmanned aerial vehicle intended for monitoring the territory and objects in conditions of low availability of communication channels. *Auditorium*, 2019, no. 4(24), pp. 48–53. Available at: <https://auditorium.kursksu.ru/#new-number?id=156> (accessed 20 January 2021) (In Russian).
3. Nartov M. V., Polyaniin K. S. Weapon system designed to counter unmanned aerial vehicles. *Nauka bez granits*, 2018, no. 1(18), pp. 48–51 (In Russian).
4. Afonin I. E., Makarenko S. I., Petrov S. V., Privalov A. A. Analysis of the experience of the combat use of groups of unmanned aerial vehicles to defeat anti-aircraft missile systems of the air defense system in military conflicts in Syria, Libya and Nagorno-Karabakh. *Systems of Control, Communication and Security*, 2020, no. 4, pp. 163–191 (In Russian). doi:10.24411/2410-9916-2020-10406
5. Andreyev S. V., Ilinykh V. V., Ilinykh O. A., Chertkov M. S., Klyuchnikov A. V. Estimation of inertial sensors error impact on platformless inertial navigation system accuracy. *Journal of "Almaz — Antey" Air and Space Defence Corporation*, 2018, no. 2, pp. 29–34. doi:10.38013/2542-0542-2018-2-29-34

6. Egoshkin N. A. *Metody vysokotochnoi geometricheskoi obrabotki informatsii ot sovremennykh sistem kosmicheskogo zondirovaniya Zemli*. Dis. dokt. tekh. nauk [Methods for high-precision geometric processing of information from modern space sensing systems of the Earth. Dr. tech. sci. diss.]. Ryazan', RGRTU Publ., 2019. 35 p. (In Russian).
7. Kaplin A. Yu. *Sposob opredeleniya koordinat nazemnykh ob»ektov pri fotos»emke s bespilotnogo letatel'nogo apparata* [Method for Determining the Coordinates of Ground Objects when Photographing from an Unmanned Aerial Vehicle]. Patent RU, no. RU 2726902, 2020.
8. Shipko V. V., Rubinov V. I., Sharonov I. E., Khanov A. S. Algorithm for determining the coordinates and parameters of the movement of ground objects. *Vestnik VGTU*, 2018, no. 1, pp. 43–48 (In Russian).
9. Shipko V. V. On some features of determining the coordinates of moving surface targets by the opto-electronic system of an unmanned aerial vehicle of the "multicopter" type. *Vozdushno-kosmicheskie sily. Teoriya i praktika*, 2019, no. 12, pp. 252–261 (In Russian).
10. Makarenkov A. Modular transfer function compensation for hyperspectral data from Resurs-P satellite system. *Image and Signal Processing for Remote Sensing XXIV*, 2018, vol. 10789, pp. 301–307.
11. Egoshkin N. Standard image products of Russian highly elliptical remote sensing system "Arktika-M". *Image and Signal Processing for Remote Sensing XXIV*, 2018, vol. 10789, pp. 430–436.
12. Nasenkov I. G. *Sposob opredeleniya uglovykh koordinat tseli s pomoshch'yu lineinoi antennoi reshetki* [Method for determining the angular coordinates of the target using a linear antenna array]. Patent RU, no. RU 2638174, 2017.
13. Smirnova D. M. *Obnaruzhenie i izmerenie koordinat dvizhushchikhsya nazemnykh ob»ektov v mnogopozitsionnoi prosvetnoi radiolokatsionnoi sisteme*. Dis. kand. tekh. nauk [Detection and measurement of coordinates of moving ground objects in a multi-position transmissive radar system. PhD diss.]. Nizhnii Novgorod, NGTU Publ., 2012. 16 p. (In Russian).
14. Chernodarov A. V. Flight testing of a distributed inertial-satellite micronavigation system for a synthetic aperture radar. *Nauchnyi vestnik MGTU GA*, 2017, no. 1, pp. 222–231 (In Russian).
15. Bazhenov A. V. The use of unmanned aerial vehicles for radar search of aircraft and ships that have suffered an accident. *Sistemy upravleniya, svyazi i bezopasnosti*, 2020, no. 2, pp. 212–230 (In Russian).
16. *Raspoznavanie ob»ektov v rezhime real'nogo vremeni na iOS s pomoshch'yu YOLOv3* [Real-time Object Recognition on iOS with YOLOv3]. Available at: <https://habr.com/ru/post/460869> (accessed 12 February 2021).
17. Dushkin A. V., Porfiriev A. V., Sumin V. I. Algorithm of measurement information processing for hardware and software complex capture and automatic tracking of unmanned aerial vehicle. *Proceedings. 2017 2nd International Ural Conference on Measurements (UralCon)*. South Ural State University, Chelyabinsk, Russia, 2017, 16–19 Oct, pp. 199–204. doi:10.1109/URALCON.2017.8120710
18. Parfiryev A. V., Dushkin A. V., Dubrovin A. S., Stepanov L. V. Control of unmanned aerial vehicles based on the detection algorithm. *Journal of Physics: Conf. Series* 1202 (2019) 012014. doi:10.1088/1742-6596/1202/1/012014
19. *Optical trackers: ASEF and MOSSE*. Available at: <https://sudonull.com/post/12040-Optical-trackers-ASEF-and-MOSSE> (accessed 20 January 2021).
20. Parfiryev A. V., Ischuk I. N., Dushkin A. V. Algorithm for controlling the trajectory of an unmanned aerial vehicle with the possibility of flying around obstacles. *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus 2020)*, Moscow, Russia, 2020, 27–30 Jan, pp. 2395–2400. doi:10.1109/EIConRus49466.2020.9039467
21. Parfiryev A. V., Dushkin A. V. The software implementation of the system of automatic observation of ground objects based on correlation analysis. *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus 2019)*, Moscow, Russia, 2019, 28–31 Jan, pp. 1749–1753. doi:10.1109/EIConRus.2019.8656636

УДК 623.1/.7

doi:10.31799/1684-8853-2021-3-19-28

Способ и алгоритм определения географических координат наземных объектов с беспилотного летательного аппаратаА. В. Парфирьев^а, канд. техн. наук, доцент, orcid.org/0000-0003-4112-896XО. В. Парфирьева^а, начальник отделения учебной лаборатории, orcid.org/0000-0001-7097-8897А. В. Душкин^{а,б}, доктор техн. наук, доцент, orcid.org/0000-0002-8078-8971, a_dushkin@mail.ru^аВоенный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия им. профессора Н. Е. Жуковского и Ю. А. Гагарина», Старых Большевиков ул., 54а, Воронеж, 394064, РФ^бНациональный исследовательский университет «Московский институт электронной техники», Шокина пл., 1, Москва, Зеленоград, 124498, РФ

Введение: как показывает практика, на точность определения координат наземных объектов летательными аппаратами с помощью оптических приборов влияет множество факторов, связанных с погрешностями в измерении угловых координат оптической системы, дальности до объекта и наличием неоднородного рельефа местности. **Цель:** разработка алгоритма повышения эффективности и точности определения географических координат наземных объектов с беспилотного летательного аппарата. **Результаты:** разработаны способ и алгоритм определения географических координат, основанный на применении цифровой модели рельефа местности и методов оптимизации. Точность расчета координат объекта повышается за счет минимизации ошибки измерения угла склонения, азимута на наземный объект и наклонной дальности. Для подтверждения аналитических расчетов проведен натурный эксперимент с автомобилем на местности. На значительном расстоянии, при котором наклонная дальность составила 900 м, произведено несколько стоп-кадров данных. В результате расчетов получены географические координаты автомобиля двумя способами — традиционным и разработанным. В конечном итоге точность расчета координат разработанным способом в 4,8 раза выше. **Практическая значимость:** предложенный в работе способ и алгоритмы обработки информации позволяют создать ряд аппаратных и программных решений для систем наведения и целеуказания.

Ключевые слова — цифровая обработка изображений, корреляционная фильтрация, ПИД-регулятор, оптимизационный алгоритм, оптико-электронная система.

Для цитирования: Parfiriyev A. V., Parfiriyeva O. V., Dushkin A. V. Method and algorithm for determining the geographic coordinates of ground objects from an unmanned aerial vehicle. *Информационно-управляющие системы*, 2021, № 3, с. 19–28. doi:10.31799/1684-8853-2021-3-19-28

For citation: Parfiriyev A. V., Parfiriyeva O. V., Dushkin A. V. Method and algorithm for determining the geographic coordinates of ground objects from an unmanned aerial vehicle. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 19–28. doi:10.31799/1684-8853-2021-3-19-28

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.

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

Сегментирование множества данных с учетом информации воздействующих факторов

И. С. Лебедев^а, доктор техн. наук, профессор, orcid.org/0000-0001-6753-2181, isl_box@mail.ru

^аСанкт-Петербургский федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

Введение: применение методов машинного обучения предполагает сбор и обработку в автономном режиме значений, поступающих от регистрирующих элементов. Большинство моделей обучается на исторических данных, а затем применяется в задачах прогнозирования, классификации, поиска влияющих факторов и воздействий, анализа состояния. В течение времени могут меняться диапазоны регистрируемых значений, что сказывается на качестве результатов классификационных алгоритмов и приводит к тому, что модели должны обучаться постоянно либо перенастраиваться с учетом поступающих значений параметров. **Цель:** разработка методики, повышающей показатели качества алгоритмов машинного обучения в условиях динамически изменяющихся и нестационарных сред, где распределение данных может изменяться с течением времени. **Методы:** разбиение (сегментирование) множества данных на основе информации о факторах, влияющих на диапазоны значений целевых переменных. **Результаты:** предложена методика сегментирования множества данных, основанная на учете факторов, которые влияют на изменение диапазонов значений целевых переменных. Выявление воздействий дает возможность сформировать выборки исходя из текущих и предполагаемых ситуаций. На примере датасета PowerSupply реализовано разбиение множества данных на подмножества, учитывающее влияние факторов на диапазоны значений. Приведена формализация внешних воздействий на основе продукционных правил. Показана обработка факторов с помощью функции принадлежности (индикаторной функции). С ее использованием произведено разбиение выборки данных на конечное число непересекающихся измеримых подмножеств. Приведены экспериментальные значения функции потерь MSE нейросети для предлагаемой методики на выбранном датасете. Показаны результаты качественных показателей классификации (Accuracy, AUC, F-мера) для различных классификаторов. **Практическая значимость:** результаты могут быть использованы при разработке классификационных моделей методов машинного обучения. Предложенная методика позволяет повысить показатели качества классификации в изменяющихся условиях функционирования.

Ключевые слова – машинное обучение, сегментирование множества данных, воздействующие факторы, изменяющиеся условия.

Для цитирования: Лебедев И. С. Сегментирование множества данных с учетом информации воздействующих факторов. *Информационно-управляющие системы*, 2021, № 3, с. 29–38. doi:10.31799/1684-8853-2021-3-29-38

For citation: Lebedev I. S. Dataset segmentation considering the information about impact factors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 29–38 (In Russian). doi:10.31799/1684-8853-2021-3-29-38

Введение

Развитие технологий вызывает лавинообразный рост информации. В связи с этим необходимо разрабатывать эффективные методы анализа и обработки постоянно увеличивающихся объемов данных в различных информационных системах.

Традиционное применение методов машинного обучения предполагает сбор и обработку в автономном режиме значений, поступающих от регистрирующих элементов. Большинство моделей обучается на исторических данных, а затем применяется в задачах прогнозирования, классификации, поиска влияющих факторов и воздействий, анализа состояния.

Во многих информационных системах наблюдения производятся одновременно множеством регистрирующих элементов, информация представляется временными рядами. В разных ситуациях могут меняться диапазоны регистри-

руемых параметров; это влияет на качество результатов классификационных алгоритмов и приводит к тому, что модели должны обучаться постоянно либо перенастраиваться с учетом непрерывно поступающих значений.

В динамически меняющихся и нестационарных средах распределение данных может становиться другим с течением времени, что приводит к «дрейфу концепций» [1, 2], когда возникают изменения условного распределения выходных данных от значений входных признаков, в то время как распределение входных данных может оставаться неизменным.

Рост объемов разнородной информации о поведении информационных, физических процессов, протекающих в технологических системах, и требования повышения качества анализа состояния элементов и узлов обуславливают необходимость адаптировать методы машинного обучения к возникающим изменениям диапазонов значений целевых переменных.

Обзор существующих методов

Большая часть исследований, направленных на повышение качественных показателей идентификации состояния информационных систем в условиях динамически протекающих процессов, фокусируется на проблемах классификации и адаптации дрейфа концепций. Основные виды этого явления представлены и описаны в ряде работ [3–5].

Методы обнаружения и обработки дрейфа концепций делятся на контролируемые, требующие заранее заданной модели или значений параметров, и неконтролируемые подходы.

Трансформация свойств анализируемого явления, связанного с изменением диапазонов значений переменных его кортежей, приводит к тому, что модель становится неактуальной. В связи с этим возникает ряд фундаментальных аспектов применения методов машинного обучения. Текущий этап исследований сфокусирован на задачах обнаружения и реакции на дрейф, идентификации ложных «выбросов» данных, устойчивости к ошибкам первого и второго рода, быстрого обнаружения аномальных событий на небольшом количестве наблюдений.

Используется несколько направлений к решению обозначенных проблемных вопросов.

Первое связано с ансамблями классифицирующих алгоритмов, обученных на подмножествах данных [6–9]. Суть методов состоит в объединении прогнозов моделей. Дрейф концепций определяется анализом статистического расхождения результатов, выдаваемых классификаторами. Если установленная доля ответов находится выше порога, то рассматривается гипотеза о смещении значений целевых переменных. Однако эти методы не являются универсальными, имеют сложности, связанные с формированием модели производящих оценку достоверности классификаторов.

Второе направление базируется на контроле распределения вероятностей. Такие методы направлены на обнаружение возможных изменений диапазонов обрабатываемых данных. Они требуют большого количества ресурсов и в определенных ситуациях характеризуются высокой частотой ложных тревог [10–12].

Третье направление — разработка моделей проявления дрейфа концепта. Такие модели не являются универсальными [10, 13–15], требуют больших вычислительных затрат и адаптации классифицирующих алгоритмов. Они основаны на предварительных знаниях о свойствах концептов, присутствующих в данных. В случае большого количества анализируемых целевых переменных формируется множество классифицирующих моделей и разрабатываются сложные решающие правила.

В большинстве случаев применяемые на сегодня методы являются узкоспециализированными и требуют существенных затрат на реализацию [14, 16–18].

В реальной среде данные всегда имеют несовершенную форму, являются примерами несбалансированных выборок. В то же время существует ряд факторов, влияющих на значения показателей. Они могут быть известны заранее, действовать с определенной периодичностью и изменять регистрируемые результаты в пределах некоторого заранее оцениваемого диапазона.

Использование информации о влияющих на диапазоны значений факторах дает возможность сформировать выборки, позволяющие повысить качество алгоритмов классификации, вследствие чего предлагается методика разбиения (сегментации) множества данных на основе выбранных факторов. В результате ее применения получается несколько подмножеств, каждое из которых определено исходя из влияния выбранного фактора. В дальнейшем для повышения качественных показателей на каждую выборку может назначаться свой классифицирующий алгоритм либо его параметры, влияющие на результат классификации, которые могут меняться с учетом влияния фактора.

Описание предлагаемой методики

Одним из основополагающих факторов, определяющих результат алгоритмов машинного обучения, является формирование обучающего подмножества. Наличие качественных выборок данных во многих случаях гораздо важнее качества алгоритмов [19–21]. Ошибки формирования множеств примеров, на которых обучаются и тестируются классификаторы, предопределяют эффективность модели. Однако при этом необходимо учитывать, что могут проявляться изменения в распределении данных во времени в различных формах.

Формализованное описание постановки задачи можно представить следующим образом.

X — множество описаний объектов, $x \in X$ — это d -мерный кортеж признаков в предопределенном векторном пространстве $X = R^d$.

Множество классов, поставленных в соответствие описаниям x , отмеченных метками $\{c_1, c_2, \dots, c_l\} \in C$, разбивается на бинарное подмножество состояний, объединенных классами $\{C_1, C_2\} \in C$.

Имеется множество факторов V , влияющих на значения признаков. Множество X характеризуется множеством значений признаков $f: X \xrightarrow{V} D_f$.

С учетом действующего в текущий момент времени фактора $v \in V$ определяется признаковое описание объекта наблюдения $X_i \in X$ в виде $X_i = (f_1(v, x), \dots, f_n(v, x))$ при заданных f_1, \dots, f_n .

Временная метка	Значения признаков			
	Признак 1 f_1	Признак 2 f_2	...	Признак n f_n
t_0	x_{10}	x_{20}	...	x_{n0}
t_1	x_{11}	x_{21}	...	x_{n1}
t_2	x_{12}	x_{22}	...	x_{n2}
...
t_m	x_{1m}	x_{2m}	...	x_{nm}

Влияющий фактор	Метка класса (принадлежность событию)	Временная метка	Значения признаков			
			Признак 1 f_1	Признак 2 f_2	...	Признак n f_n
v_1	c_0	t_0	x_{10}	x_{20}	...	x_{n0}
	c_1	t_1	x_{11}	x_{21}	...	x_{n1}
	c_0	t_2	x_{12}	x_{22}	...	x_{n2}
	c_0	t_3	x_{10}	x_{20}	...	x_{n0}
v_2	c_0	t_4	x_{11}	x_{21}	...	x_{n1}
	c_1	t_5	x_{12}	x_{22}	...	x_{n2}
	c_1	t_6	x_{10}	x_{20}	...	x_{n0}
	c_0	t_7	x_{11}	x_{21}	...	x_{n1}
...
v_k	c_1	t_m	x_{1m}	x_{2m}	...	x_{nm}

■ **Рис. 1.** Преобразование датасета
 ■ **Fig. 1.** Dataset transformation

Обучающее множество приобретает вид размеченной выборки $\{v_j, \{(x_i, c_i)\}_{i=1}^N\}_{j=1}^M$, где i — количество кортежей в момент, когда оказывал влияние рассматриваемый j -й фактор v_j .

Необходимо построить классифицирующий алгоритм, учитывающий влияние фактора v , $\alpha: X \xrightarrow{V} C$. Алгоритм определяет по входному признаковому описанию соответствие классу.

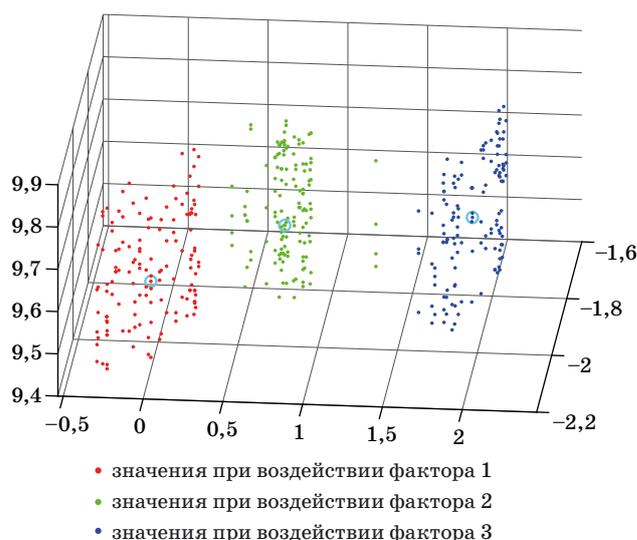
Обучение модели усложняется не только большой размерностью признакового пространства, но и наличием влияющих на значения признаков переменных факторов.

Основным ограничением методов машинного обучения является то, что алгоритмы классификации не всегда могут быть эффективны в условиях постоянно функционирующей под влиянием различных внешних и внутренних воздействий системы. Система находится в динамике, осуществляются постоянные переходы из одного состояния в другое. Внешние и внутренние факторы изменяют значения характеристик. Например, подключение нового устройства сетевого сегмента может в течение определенного момента времени вызывать изменения анализируемых значений параметров, таких как рост количества служебных сообщений, конфигурационных команд, увеличение времени задержки и т. д.

Накопленная статистическая информация о функционировании системы представляется в виде матрицы, где временной отметке соответствуют значения признаков, полученных от регистрирующих устройств и элементов. Различные факторы влияют на свойства объектов. Изменения значений признаков являются результатом воздействий факторов. Часть из них, например сезонность, являются известными в текущий момент времени.

Внесение дополнительных полей в разметку обучающего подмножества дает возможность учитывать информацию о влияющих факторах. На рис. 1 представлен вид записей исходного датасета и датасета после разметки и добавления информации о влияющих факторах. В результате такого представления записи связываются с отдельными факторами.

Появляется возможность разбить пространство на подпространства, где происходит классификация отдельных признаковых сегментов, и рассматривать множество полученных значений для каждого фактора. На рис. 2 показан пример



■ **Рис. 2.** Пример значений из датасета при воздействии разных факторов
 ■ **Fig. 2.** Dataset values example exposed to various factors

разбиения множества данных на подмножества с учетом наблюдаемых факторов.

Задача отнесения состояния наблюдаемого объекта решается с учетом текущего информационного воздействия.

Анализ и учет влияющих на данные факторов позволяют реализовать разбиение множества на подмножества. В дальнейшем, определяя свойства полученных выборок, можно осуществить решение задачи о применении наиболее эффективных алгоритмов обработки.

Применение методики

Рассмотрим классификатор $\varphi(x, \mathbf{W})$. На вход поступает кортеж значений x . Для принятия решения используется весовая матрица \mathbf{W} . Возможны два направления разделения множества данных: использование продукционных правил и функций принадлежности.

Применение продукций предполагает, что влияющие на значения данных факторы поддаются эвристике. Она дает возможность определить правила, учитывающие воздействия на значения выборки. Например, периодичность процессов в экономике, энергетике, других областях позволяет сформировать ряд продукционных правил. В общем виде такая модель представляется в предикативном виде:

$$M = \langle \Phi, V, \mathbf{W}, \mathbf{X} \rangle,$$

где Φ — классифицирующие алгоритмы, использующие весовые матрицы для сравнения поступающего кортежа данных; V — множество влияющих факторов на целевые переменные в выборке данных; \mathbf{W} — множество весовых матриц классификаторов, значения матриц зависят от фактора, влияющего на данные в системе; \mathbf{X} — множество описаний объектов, состоящее из подмножества выборок данных, каждому подмножеству соответствуют свои весовые матрицы классификаторов.

Выбор значений $w_j \in \mathbf{W}$ может быть осуществлен на основе продукционной модели. Определяется подмножество данных \mathbf{X}_j с учетом влияния фактора v_j . Каждому подмножеству может назначаться классифицирующий алгоритм $\varphi \in \Phi$. Сгруппированным переменным «влияющий фактор — подмножество» v_j, \mathbf{X}_j определяется матрица w_i с учетом свойств классифицирующих алгоритмов. Продукция, реализующая правило, примет вид

$$(v_j, \mathbf{X}_j) \xrightarrow{\varphi_k} w_j. \quad (1)$$

Выражение (1) позволяет после определения текущего воздействия v_j использовать матрицу w_j .

Поступающий на вход новый кортеж x идентифицируется классификатором $\varphi(x, w_j)$.

В дальнейшем возможно реализовать правила не только для разбиения на сегменты множества данных, но и для выбора соответствующего классификатора в момент действия фактора. Это позволит учитывать изменения в данных и повышать показатели качества классифицирующей модели в целом.

Второе направление базируется на использовании функции принадлежности. Оно может применяться, когда существуют воздействия, поддающиеся аналитическому описанию (например, длина светового дня в зависимости от времени года, широта места в подсистеме подачи электроэнергии на объекты городского хозяйства, часы пиковой нагрузки в информационной системе). Такие факторы v_j могут быть обработаны с помощью функции принадлежности (индикаторной функции). На ее основе производится разбиение выборки данных \mathbf{X} на конечное число не пересекающихся измеримых подмножеств $\mathbf{X}_1, \dots, \mathbf{X}_n$. В простейшем случае функция принадлежности μ подмножества $\mathbf{X}_j \in \mathbf{X}$, где x — кортеж обучающей выборки, может быть представлена в виде

$$\mu_{\mathbf{X}_j}(x, v_j) = \begin{cases} 1, & x \in \mathbf{X}_j \\ 0, & x \notin \mathbf{X}_j \end{cases}. \quad (2)$$

Выражение (2) позволяет определить принадлежность элемента выборки данных $x \in \mathbf{X}$ подмножеству \mathbf{X}_j в момент действия фактора v_j .

В общем случае выборка состоит из n подмножеств. Принадлежность подмножеству определяется функциями $\mu_{\mathbf{X}_1}(x), \mu_{\mathbf{X}_2}(x), \dots, \mu_{\mathbf{X}_n}(x)$.

Получается разбиение $\mathbf{X}_1 \cup \mathbf{X}_2 \cup \dots \cup \mathbf{X}_n = \mathbf{X}$ при условии $\mathbf{X}_j \cap \mathbf{X}_i = \emptyset \forall i \neq j$. Объединение подмножеств совпадает с множеством \mathbf{X} , подмножества не пересекаются.

Классификацию становится возможным осуществить на каждом из подмножеств $\mathbf{X}_1, \dots, \mathbf{X}_n$. Формируется тестовая и обучающая выборка с учетом действующих факторов v_j . В зависимости от обрабатываемого подмножества классификатор $\varphi(x, w)$ может быть дополнен функцией $\psi(v_j)$. Функция $\psi(v_j)$ учитывает фактор v_j , влияющий на подмножество \mathbf{X}_j , и определяет по его значению весовую матрицу $w_j = \psi(v_j)$. Классификатор примет вид $\varphi(x, \psi(v_j))$.

В качестве одной из мер оценки модели может быть применена функция потерь.

Функция потерь $L(v_j)$ для подмножества \mathbf{X}_j определяется выражением

$$L(v_j) = \frac{1}{N} \sum_i L_i(\varphi(x_i, \psi(v_j)), c_i) + \lambda R(\psi(v_j)), \quad (3)$$

где $R(\psi(v_j))$ — функция регуляризации; λ — коэффициент регуляризации. Они влияют на диспер-

сию и смещение ответов классификатора. Регуляризация предназначена для добавления дополнительных ограничений, предотвращающих переобучение.

Средняя сумма потерь для данных множества X

$$L = \frac{1}{M} \sum_{j=1}^M L(v_j). \quad (4)$$

Применяя (3) и (4) и минимизируя среднюю сумму потерь, можно найти оптимальные параметры на основе выражения

$$L = \frac{1}{MN} \sum_{j=1}^M \sum_{i=1}^N (L_i(\varphi(x_i, \psi(v_j)), c_i) + \lambda R(\psi(v_j))) \rightarrow \min. \quad (5)$$

В предлагаемой методике с использованием выражения (5) появляется возможность определить разбиение выборки на подмножества с учетом факторов V .

Эксперимент

Эксперимент проводился на множестве датасета PowerSupply [22], содержащего информацию о почасовой подаче электроэнергии итальянской энергетической компании. В нем представлены данные о потребляемой мощности из электрических сетей каждый час с 1995 по 1998 год. Различные факторы (рабочее и нерабочее время дня, будни и выходные, сезонность, перемена погоды) вызывают явление дрейфа концепта.

В эксперименте рассматривалось предсказание рабочего и нерабочего времени по входным данным потребляемой мощности. В качестве воздействующего фактора, на основе которого производилась сегментация выборки данных, была выбрана сезонность. В первом случае разбиение производилось по датам перехода на летнее и зимнее время, во втором — по календарным временам года.

Общий вид датасета представлен на рис. 3. В горизонтальной плоскости на осях показаны дни наблюдений и часы, по вертикальной оси отложена потребляемая мощность.

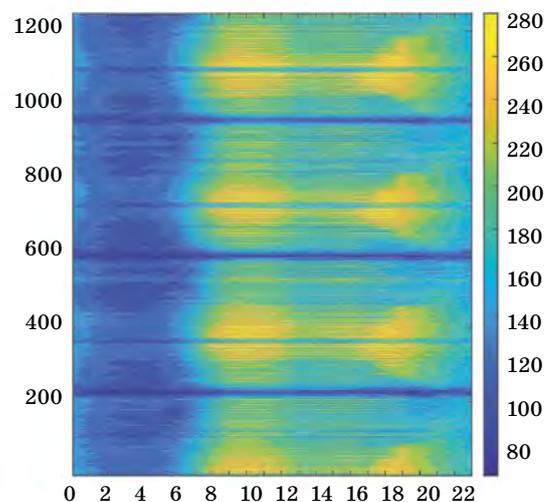
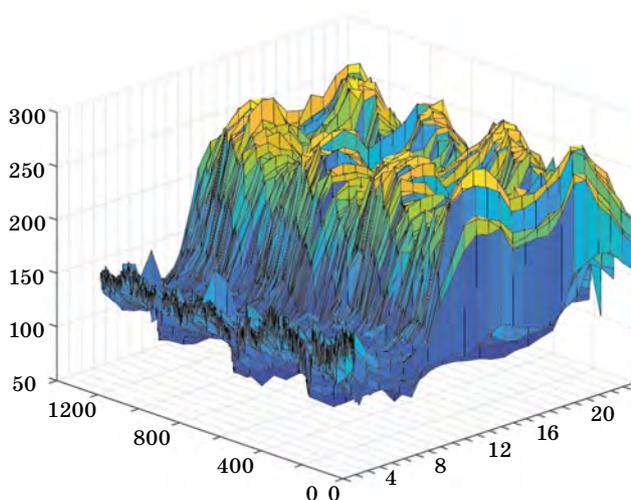
Даты перехода на летнее и зимнее время дают возможность реализовать правило или функцию принадлежности, с помощью которых выборка делится на два непересекающихся подмножества (рис. 4, а и б).

Второе разбиение было осуществлено для анализа изменений показателей качества классификаторов. Общая выборка была разделена на четыре части и содержала значения потребляемой энергии в весенние, летние, осенние, зимние месяцы. В дальнейшем два способа разбиения использовались для сравнения результатов классифицирующих алгоритмов.

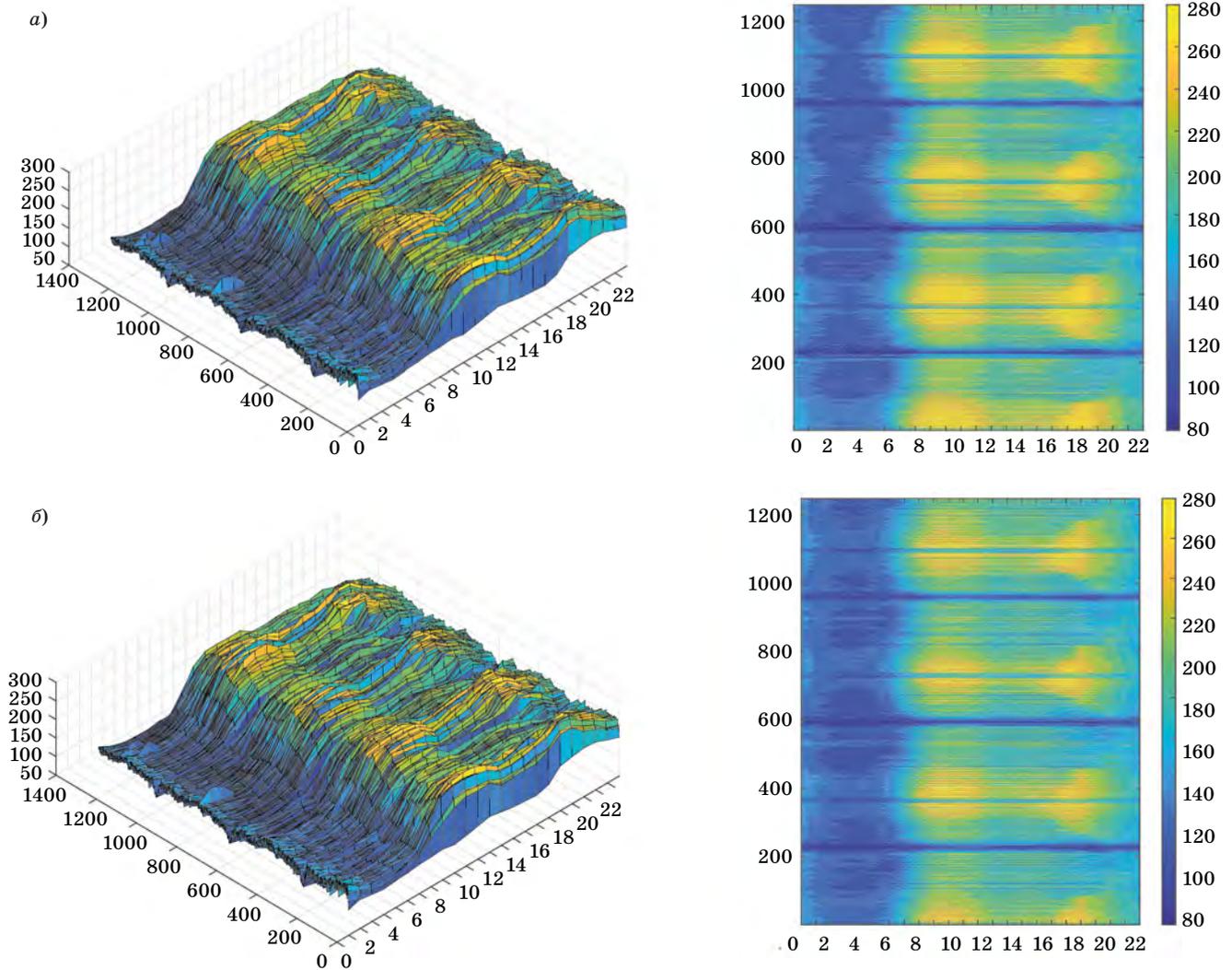
В качестве обучающих данных использовался переменный объем каждого подмножества, составляющий от 5 до 35 %.

Статистические свойства предсказываемой целевой переменной изменяются с течением времени. В рассматриваемых случаях на значения входящих в кортеж данных влияет заранее определенный сезонный фактор.

В первой части эксперимента для оценки влияния на классифицирующий алгоритм разбиения данных на подмножества была реализована



■ **Рис. 3.** Поверхность датасета потребления электроэнергии
 ■ **Fig. 3.** The dataset surface of electricity consumption



■ **Рис. 4.** Поверхность подмножества датасета потребления электроэнергии летнего (а) и зимнего (б) времени
 ■ **Fig. 4.** The dataset surface of electricity consumption “Summer Time” (a) and “Winter Time” (б)

двухслойная нейросеть. Значение точности сети (Ассигасу) составляло около 0,75. Задача сети состояла в том, чтобы по входному кортежу признаков потребления электроэнергии пользователями спрогнозировать, происходит данная ситуация в рабочее или нерабочее время.

В качестве меры была выбрана среднеквадратичная функция потерь (MSE). Она является одним из основных показателей в задачах регрессии и чувствительна к выбросам данных.

Среднеквадратичная функция потерь для всего множества, представленного на рис. 3:

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2. \quad (6)$$

Для разделенного датасета, состоящего из двух или четырех частей ($M = 2, M = 4$), выражение (6) будет выглядеть следующим образом:

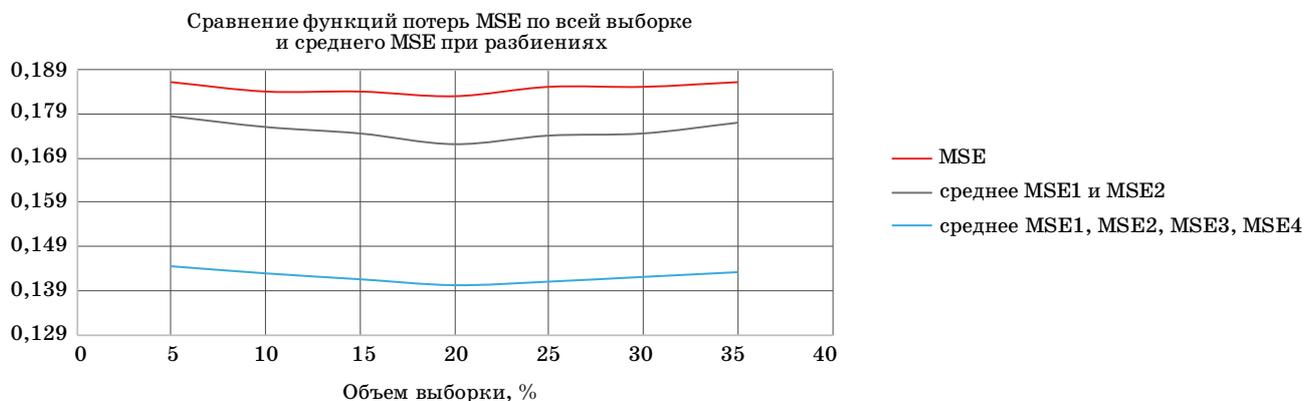
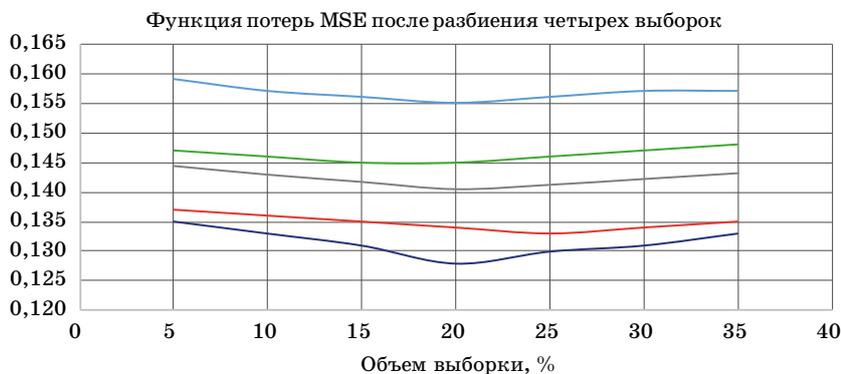
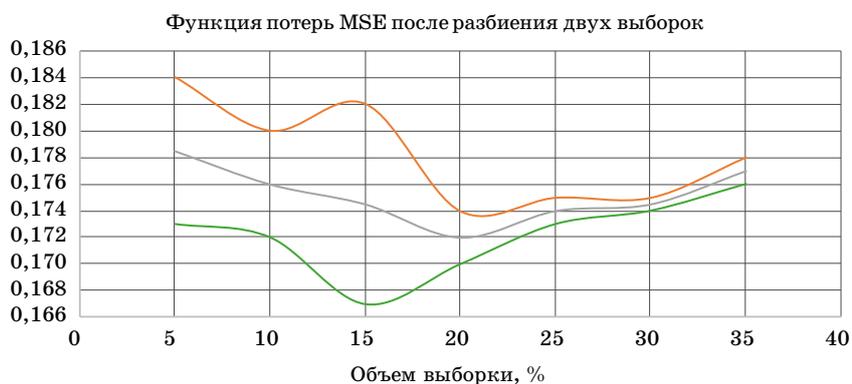
$$L_{MSE_{x_1, x_2}} = \frac{1}{MN} \sum_{j=1}^M \sum_{i=1}^N (x_{ji} - \hat{x}_{ji})^2. \quad (7)$$

В результате проведения эксперимента по обучению и оценки модели были получены следующие графики функции потерь (рис. 5), определяемой выражениями (6) и (7).

Графики показывают более низкие значения функции потерь для модели, когда пространство сегментировано, что позволяет говорить о целесообразности разделения на подмножества на основе информации о влияющем факторе.

Кроме того, на графиках виден момент переобучения нейросети для выбранных подмножеств, что позволяет оценить свойства данных для дальнейших стратегий обучения на подмножествах.

Во второй части эксперимента для общей оценки влияния применения подмножеств на ка-



■ Рис. 5. Значения функций потерь
 ■ Fig. 5. Loss functions values

чество результатов моделей машинного обучения были выбраны различные алгоритмы: линейного дискриминантного анализа (LD), квадратичного дискриминантного анализа (QD), наивного байесовского классификатора (NB), ближайших соседей (KNN).

На каждый классифицирующий алгоритм подавались наборы данных из двух подмножеств зимнего и летнего времени (см. рис. 3, 4) и четырех подмножеств весенних, летних, зимних, осенних месяцев.

Результаты тестирования классификаторов (AUC — площадь под ROC-кривой, точность Accuracy, F-мера) для целого множества X и усредненные значения для множеств X_{11} и X_{12} и множеств X_{21} , X_{22} , X_{23} , X_{24} приводятся в таблице.

Результаты тестирования демонстрируют, что разбиение общей выборки на отдельные подмножества в основном позволяет повысить ряд показателей качества классификации для выбранных алгоритмов.

Разделение на два подмножества позволяет улучшить результаты на 1 %. При разбиении на четыре подмножества наблюдается повышение значений AUC, Accuracy, F-меры для каждого отдельного классификатора уже на 5–8 %, что сопоставимо с результатами ансамблевых моделей.

Можно осуществить дальнейшую операцию разбиения на подмножества, используя информацию о выходных и праздничных днях, погодных условиях и т. д. Предложенное решение можно использовать как дополнение к различным классификационным моделям. Допустима и более сложная сегментация, учитывающая дополнительные параметры, которая за счет уменьшения явления «выбросов» данных позволит повысить качественные показатели.

Заключение

Основные проблемные вопросы методов машинного обучения лежат в области формирования выборок данных, определяющих достижение заданных показателей. Качество обучающих подмножеств повышается путем устранения шумов, удаления дисбаланса классов, обнаружения дрейфа концепта.

Предложена методика сегментации выборок данных, основанная на учете факторов, которые влияют на изменение диапазонов целевых переменных. Выявление воздействий дает возможность сформировать сегментированные выборки данных исходя из текущих предполагаемых ситуаций. Для каждого полученного множества возможен поиск лучшей модели, реализующей классификационную задачу.

Наборы данных, тестовые множества имеют свои свойства. Применение предложенного решения позволяет уменьшить влияние шумовых данных, избежать введения дополнительных затрат на борьбу с явлением дрейфа концепта, повысить показатели полноты и точности за счет уменьшения разброса параметров.

Однако имеется несколько моментов, которые необходимо учитывать для реализации предлагаемой методики. Внутри сформированных сегментов выборок могут происходить изменения целевых переменных под влиянием других факторов. В обрабатываемых множествах могут быть другие концепции, влияющие на конечный результат. Для эффективного применения классифицирующей модели необходимо исследовать объемы, свойства обучаемых выборок, ограничения, связанные с длинами сегментов, пороговыми значениями диапазонов целевых переменных.

■ Результаты классифицирующих алгоритмов

■ Results of classifying algorithms

Модель	Объем выборки, %	AUC			Accuracy			F-мера		
		X	Среднее $X_{11}+X_{12}$	Среднее $X_{21}+X_{22}+X_{23}+X_{24}$	X	Среднее $X_{11}+X_{12}$	Среднее $X_{21}+X_{22}+X_{23}+X_{24}$	X	Среднее $X_{11}+X_{12}$	Среднее $X_{21}+X_{22}+X_{23}+X_{24}$
LD	5	0,76	0,78	0,81	0,727905	0,730246	0,783536	0,790824	0,775388	0,821392
	35	0,77	0,78	0,82	0,717489	0,717681	0,784194	0,757563	0,758894	0,825022
QD	5	0,95	0,95	0,97	0,834698	0,872516	0,924455	0,872141	0,895602	0,943755
	35	0,96	0,96	0,99	0,873597	0,861341	0,925849	0,896387	0,886896	0,947586
NB	5	0,89	0,91	0,94	0,803321	0,811307	0,866226	0,844574	0,84626	0,892745
	35	0,91	0,92	0,96	0,812483	0,812493	0,864015	0,844612	0,84462	0,898299
KNN	5	0,88	0,95	0,98	0,827748	0,878408	0,925379	0,85605	0,898913	0,953179
	35	0,95	0,96	0,99	0,881582	0,895829	0,942752	0,901138	0,913712	0,968623

Литература

1. Schlimmer J. C., Granger R. H. Incremental learning from noisy data. *Machine Learning*, 1986, no. 1, pp. 317–354. doi:10.1023/A:1022810614389
2. Widmer G., Kubat M. Learning in the presence of concept drift and hidden contexts. *Machine Learning*, 1996, no. 23(1), pp. 69–101. doi:10.1007/BF00116900
3. Gama J., Žliobait I., Bifet A., Pechenizkiy M., Bouchachia A. A survey on concept drift adaptation. *ACM Computing Surveys*, 2014, no. 46(4), pp. 1–37.
4. Sung-Yu T., Jen-Yuan C. Parametric study and design of deep learning on leveling system for smart manufacturing. *IEEE International Conference on Smart Manufacturing, Industrial & Logistics Engineering (SMILE)*, February 8–9, 2018, pp. 48–52. doi:10.1109/SMILE.2018.8353980
5. Lu J., Liu A., Dong F., Gu F., Gama J., Zhang G. Learning under concept drift: a review. *IEEE Transactions on Knowledge and Data Engineering*, 2019, no. 31(12), pp. 2346–2363.
6. Wang L. Y., Park C., Yeon K., Choi H. Tracking concept drift using a constrained penalized regression combiner. *Computational Statistics & Data Analysis*, 2017, no. 108, pp. 52–69.
7. Khan S., Yairi T. A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, 2018, no. 107, pp. 241–265. doi:10.1016/j.ymsp.2017.11.024
8. Salehi H., Burgueño R. Emerging artificial intelligence methods in structural engineering. *Engineering Structures*, 2018, no. 171, pp. 170–189. doi:10.1016/j.engstruct.2018.05.084
9. Zhou Z.-H., Feng J. Deep forest. *National Science Review*, 2019, vol. 6, no. 1, pp. 74–86. doi:10.1093/nsr/nwy108
10. Sethi T., Kantardzic M. Handling adversarial concept drift in streaming data. *Expert Systems with Applications*, 2018, vol. 97, pp. 18–40. doi:10.1016/j.eswa.2017.12.022
11. Takacs A., Toledano-Ayala M., Dominguez-Gonzalez A., Pastrana-Palma A., Velazquez D. T., Ramos J. M., Rivas-Araiza A. E. Descriptor generation and optimization for a specific outdoor environment. *IEEE Access*, 2020, vol. 8, pp. 2169–3536. doi:10.1109/ACCESS.2020.2975474
12. Saadallah A., Moreira-Matias L., Sousa R., Khiairi J., Jenelius E., Gama J. Bright-drift-aware demand predictions for taxi networks. *IEEE Transactions on Knowledge and Data Engineering*, 2020, vol. 32, pp. 234–245.
13. Шелухин О. И., Симомян А. Г., Ванюшина А. В. Влияние структуры обучающей выборки на эффективность классификации приложений трафика методами машинного обучения. *Т-Сотт: Телекоммуникации и транспорт*, 2017, т. 11, № 2, с. 25–31.
14. Lughofer E., Weigl E., Heidl W., Eitzinger C., Radauer T. Recognizing input space and target concept drifts in data streams with scarcely labeled and unlabeled instances. *Information Sciences*, 2016, vol. 35, pp. 127–151.
15. Sethi T., Kantardzic M. Handling adversarial concept drift in streaming data. *Expert Systems with Applications*, 2018, vol. 97, pp. 18–40.
16. Рзаев Б. Т., Лебедев И. С. Применение бэггинга при поиске аномалий сетевого трафика. *Научно-технический вестник информационных технологий, механики и оптики*, 2021, т. 21, № 2, с. 50–56. doi:10.17586/2226-1494-2021-21-2-50-56
17. Maletzke A., dos Reis D., Cherman E., Batista G. (2019). DyS: A frame work for mixture models in quantification. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, no. 1, pp. 4552–4560. doi:10.1609/aaai.v33i01.33014552
18. Gomes H. M., Bifet A., Read J., Barddal J. P., Enembreck F., Pfharinger B., Holmes G. Adaptive random forests for evolving data stream classification. *Machine Learning*, 2017, vol. 106, iss. 9–10, pp. 1469–1495. doi:10.1007/s10994-017-5642-8
19. Бажаев Н. А., Лебедев И. С., Кривцова И. Е. Анализ статистических данных мониторинга сетевой инфраструктуры для выявления аномального поведения локального сегмента системы. *Научно-технический вестник информационных технологий, механики и оптики*, 2017, т. 17, № 1, с. 92–99. doi:10.17586/2226-1494-2017-92-99
20. Oikarinen E., Tiittanen H., Henelius A. Detecting virtual concept drift of regressors without ground truth values. *Data Mining and Knowledge Discovery*, 2021, vol. 35, iss. 3, pp. 821–859. doi:10.1007/s10618-021-00739-7
21. Maletzke A., dos Reis D., Batista G. Combining instance selection and self-training to improve data stream quantification. *Journal of the Brazilian Computer Society*, 2018, vol. 24, no. 12, pp. 123–141. doi:10.1186/s13173-018-0076-0
22. PowerSupply dataset. <http://www.cse.fau.edu/~xqzhu/stream.html> (дата обращения: 27.03.2021).

UDC 621.396

doi:10.31799/1684-8853-2021-3-29-38

Dataset segmentation considering the information about impact factorsI. S. Lebedev^a, Dr. Sc., Tech., Professor, orcid.org/0000-0001-6753-2181, isl_box@mail.ru^aSaint-Petersburg Federal Research Center of the RAS, 39, 14 Line V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: The application of machine learning methods involves the collection and processing of data which comes from the recording elements in the offline mode. Most models are trained on historical data and then used in forecasting, classification, search for influencing factors or impacts, and state analysis. In the long run, the data value ranges can change, affecting the quality of the classification algorithms and leading to the situation when the models should be constantly trained or readjusted taking into account the input data. **Purpose:** Development of a technique to improve the quality of machine learning algorithms in a dynamically changing and non-stationary environment where the data distribution can change over time. **Methods:** Splitting (segmentation) of multiple data based on the information about factors affecting the ranges of target variables. **Results:** A data segmentation technique has been proposed, based on taking into account the factors which affect the change in the data value ranges. Impact detection makes it possible to form samples based on the current and alleged situations. Using PowerSupply dataset as an example, the mass of data is split into subsets considering the effects of factors on the value ranges. The external factors and impacts are formalized based on production rules. The processing of the factors using the membership function (indicator function) is shown. The data sample is divided into a finite number of non-intersecting measurable subsets. Experimental values of the neural network loss function are shown for the proposed technique on the selected dataset. Qualitative indicators (Accuracy, AUC, F-measure) of the classification for various classifiers are presented. **Practical relevance:** The results can be used in the development of classification models of machine learning methods. The proposed technique can improve the classification quality in dynamically changing conditions of the functioning.

Keywords — machine learning, segmentation of multiple data, affecting factors, changing conditions.

For citation: Lebedev I. S. Dataset segmentation considering the information about impact factors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 29–38 (In Russian). doi:10.31799/1684-8853-2021-3-29-38

References

- Schlimmer J. C., Granger R. H. Incremental learning from noisy data. *Machine Learning*, 1986, no. 1, pp. 317–354. doi:10.1023/A:1022810614389
- Widmer G., Kubat M. Learning in the presence of concept drift and hidden contexts. *Machine Learning*, 1996, no. 23(1), pp. 69–101. doi:10.1007/BF00116900
- Gama J., Žliobait I., Bifet A., Pechenizkiy M., Bouchachia A. A survey on concept drift adaptation. *ACM Computing Surveys*, 2014, no. 46(4), pp. 1–37.
- Sung-Yu T., Jen-Yuan C. Parametric study and design of deep learning on leveling system for smart manufacturing. *IEEE International Conference on Smart Manufacturing, Industrial & Logistics Engineering (SMILE)*, February 8–9, 2018, pp. 48–52. doi:10.1109/SMILE.2018.8353980
- Lu J., Liu A., Dong F., Gu F., Gama J., Zhang G. Learning under concept drift: a review. *IEEE Transactions on Knowledge and Data Engineering*, 2019, no. 31(12), pp. 2346–2363.
- Wang L. Y., Park C., Yeon K., Choi H. Tracking concept drift using a constrained penalized regression combiner. *Computational Statistics & Data Analysis*, 2017, no. 108, pp. 52–69.
- Khan S., Yairi T. A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, 2018, no. 107, pp. 241–265. doi:10.1016/j.ymssp.2017.11.024
- Salehi H., Burgueño R. Emerging artificial intelligence methods in structural engineering. *Engineering Structures*, 2018, no. 171, pp. 170–189. doi:10.1016/j.engstruct.2018.05.084
- Zhou Z.-H., Feng J. Deep forest. *National Science Review*, 2019, vol. 6, no. 1, pp. 74–86. doi:10.1093/nsr/nwyl08
- Sethi T., Kantardzic M. Handling adversarial concept drift in streaming data. *Expert Systems with Applications*, 2018, vol. 97, pp. 18–40. doi:10.1016/j.eswa.2017.12.022
- Takacs A., Toledano-Ayala M., Dominguez-Gonzalez A., Pastrana-Palma A., Velazquez D. T., Ramos J. M., Rivas-Araiza A. E. Descriptor generation and optimization for a specific outdoor environment. *IEEE Access*, 2020, vol. 8, pp. 2169–3536. doi:10.1109/ACCESS.2020.2975474
- Saadallah A., Moreira-Matias L., Sousa R., Khiari J., Jenelius E., Gama J. Bright-drift-aware demand predictions for taxi networks. *IEEE Transactions on Knowledge and Data Engineering*, 2020, vol. 32, pp. 234–245.
- Sheluhin O. I., Simonyan A. G., Vanyushina A. V. Influence of training sample structure on traffic application efficiency classification using machine-learning methods. *T-Comm*, 2017, vol. 11, no. 2, pp. 25–31 (In Russian).
- Lughofer E., Weigl E., Heidl W., Eitzinger C., Radauer T. Recognizing input space and target concept drifts in data streams with scarcely labeled and unlabelled instances. *Information Sciences*, 2016, vol. 35, pp. 127–151.
- Sethi T., Kantardzic M. Handling adversarial concept drift in streaming data. *Expert Systems with Applications*, 2018, vol. 97, pp. 18–40.
- Rzayev B. T., Lebedev I. S. Applying bagging in finding network traffic anomalies. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 2, pp. 50–56 (In Russian). doi:10.17586/2226-1494-2021-21-2-50-56
- Maletzke A., dos Reis D., Cherman E., Batista G. (2019). DyS: A frame work for mixture models in quantification. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, no. 1, pp. 4552–4560. doi:10.1609/aaai.v33i01.33014552
- Gomes H. M., Bifet A., Read J., Barddal J. P., Enembreck F., Pfahringer B., Holmes G. Adaptive random forests for evolving data stream classification. *Machine Learning*, 2017, vol. 106, iss. 9–10, pp. 1469–1495. doi:10.1007/s10994-017-5642-8
- Bazhayev N. A., Lebedev I. S., Krivtsova I. E. Analysis of statistical data from network infrastructure monitoring to detect abnormal behavior of system local segments. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2017, vol. 17, no. 1, pp. 92–99 (In Russian). doi:10.17586/2226-1494-2017-17-1-92-99
- Oikarinen E., Tiittanen H., Henelius A. Detecting virtual concept drift of regressors without ground truth values. *Data Mining and Knowledge Discovery*, 2021, vol. 35, iss. 3, pp. 821–859. doi:10.1007/s10618-021-00739-7
- Maletzke A., dos Reis D., Batista G. Combining instance selection and self-training to improve data stream quantification. *Journal of the Brazilian Computer Society*, 2018, vol. 24, no. 12, pp. 123–141. doi:10.1186/s13173-018-0076-0
- PowerSupply dataset. Available at: <http://www.cse.fau.edu/~xqzhu/stream.html> (accessed 27 March 2021).

Теоретико-информационные проблемы ДНК-памяти

С. А. Круглик^{а,б}, младший научный сотрудник, orcid.org/0000-0001-9557-5197 stanislav.kruglik@skoltech.ru

Г. А. Кучеров^{а,в}, канд. физ.-мат. наук, ведущий научный сотрудник, orcid.org/0000-0001-5899-5424

К. Н. Назирханова^г, аспирант, orcid.org/0000-0002-7447-9857

М. Е. Филитов^а, магистрант, orcid.org/0000-0003-2421-0777

^аСколковский институт науки и технологий, Большой б-р, 30, стр. 1, Москва, 121205, РФ

^бМосковский физико-технический институт, Институтский пер., 9, Долгопрудный, Московская обл., 141701, РФ

^вНациональный центр научных исследований, Университет Густава Эйфеля, 77454 Марн-ля-Валле, Франция

^гСтэнфордский университет, 94305 Стэнфорд, Калифорния, США

Введение: взрывной рост объемов производимой человечеством информации ставит новые фундаментальные задачи, связанные с ее эффективным хранением и доступом к ней. Широко используемые при этом магнитные, оптические и полупроводниковые устройства хранения имеют ряд существенных недостатков, связанных, прежде всего, с ограничениями на объем и долговечность хранения. Одной из возможных альтернатив, активно исследуемой в последние годы, является хранение данных с помощью молекул ДНК. **Цель:** обзор текущего состояния методов хранения информации с помощью молекул ДНК и связанных теоретико-информационных проблем. **Результаты:** сделан обзор современного состояния дел в разработке систем ДНК-памяти. Проведен анализ типов ошибок, возникающих в таких системах, и корректирующих кодов для выявления и исправления этих ошибок. Показаны недостатки предложенных на сегодня кодов и указаны возможные направления их улучшения. Приведен анализ существующих теоретико-информационных моделей каналов для систем ДНК-памяти и присущих им ограничений. В заключении обзора сформулированы основные проблемы на пути создания практических систем ДНК-памяти, решению которых послужит дальнейшее развитие теоретико-информационных методов, рассмотренных в настоящем обзоре.

Ключевые слова – системы хранения информации, ДНК-память, каналы передачи информации, пропускная способность канала, ошибки замены, ошибки вставки, ошибки выпадения.

Для цитирования: Круглик С. А., Кучеров Г. А., Назирханова К. Н., Филитов М. Е. Теоретико-информационные проблемы ДНК-памяти. *Информационно-управляющие системы*, 2021, № 3, с. 39–52. doi:10.31799/1684-8853-2021-3-39-52

For citation: Kruglik S. A., Kucherov G. A., Nazirkhanova K. N., Filitov M. E. Information-theoretic problems of DNA-based storage systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 39–52 (In Russian). doi:10.31799/1684-8853-2021-3-39-52

Введение

Мы живем в эпоху цифровых технологий, в которой множество различных устройств ежедневно генерируют огромное количество данных. При этом как число таких устройств, так и объемы генерируемой ими информации растут с экспоненциальной скоростью [1]. По имеющимся оценкам, суммарный объем производимой человечеством информации достигает миллиардов терабайт в год [2, 3]. Столь быстрый рост поднимает множество вопросов, связанных, прежде всего, с хранением данных и управлением ими. В частности, остро стоит проблема увеличения емкости существующих хранилищ [4]. В настоящее время большая часть данных хранится на магнитных и оптических устройствах, таких как компакт-диски, жесткие диски и магнитные ленты. Еще недавно магнитные устройства являлись наиболее популярным и доступным решением, но затем их заменили оптические устройства. Теперь же и

они вытесняются более функциональными и дешевыми устройствами флэш-памяти. Однако все вышеперечисленные носители имеют ряд общих недостатков. Во-первых, их емкости ограничены. Например, максимальный объем, который может хранить магнитная лента, это эксабайт данных, но такое хранилище может стоить весьма дорого в обслуживании и занимает значительное пространство [5]. Кроме того, все эти устройства имеют низкую плотность хранения, как правило, не превышающую тысячи гигабайт на квадратный миллиметр [6]. Еще одной проблемой, связанной с хранением данных на существующих носителях, является возможность потери данных с течением времени. Все это требует разработки принципиально новых способов хранения информации [7, 8].

Одним из таких методов, активно изучаемых в последнее время, является хранение информации с использованием молекул ДНК. Отметим, что вскоре после открытия структуры ДНК

в 1953 г. [9] некоторые известные ученые высказали предположения о возможном использовании ДНК для хранения произвольной цифровой информации. Об этих перспективах говорил выдающийся физик Ричард Фейнман в своей лекции «There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics» в 1959 г. В 60-х годах подобные идеи высказывал Норберт Винер [10], а в Советском Союзе — физик и радиоинженер М. С. Нейман [11, 12]. Последний в своих работах изложил соображения о возможных способах реализации данных систем и некоторые предварительные расчеты, которые, однако, в то время были далеки от практической реализации. Появлению и раннему развитию идей ДНК-памяти посвящен исторический обзор [13].

Первые успешные попытки сохранить информацию с использованием ДНК относятся к 1988 г., когда коллективу ученых удалось вставить в плазмидную ДНК бактерии *Escherichia coli* искусственный фрагмент из 28 нуклеотидов, из которых 18 кодировали простой символ-пиктограмму, а оставшиеся 10 содержали метаинформацию для декодирования. Впоследствии этот фрагмент был успешно извлечен из бактериальной ДНК с помощью секвенирования [14]. Интересно, что целью этого эксперимента было создание нового типа художественного объекта. Подобные эксперименты проводились и позже, в конце 90-х: так, например, в работе [15] описан эксперимент по передаче секретных сообщений, закодированных в растворе ДНК, а в [16] — эксперимент по кодированию в ДНК коротких предложений на естественном языке. Однако эти эксперименты имели целью кодирование в ДНК лишь очень небольшого объема информации, измеряемого десятками байтов, и не допускали масштабирования.

Ситуация изменилась в 2012-м, когда с использованием новых технологий синтеза и секвенирования ДНК был закодирован набор из 643 Кбайт данных, состоящий из книги, 11 изображений JPG и одной программы JavaScript [17]. Годом позже другим коллективом ученых была представлена схема хранения 739 Кбайт произвольной цифровой информации с использованием ДНК [18]. Эти работы открыли новый этап в развитии систем ДНК-памяти, послужив началом серии экспериментов по кодированию в молекулах ДНК все больших объемов информации. Так, в 2018 г. авторам работы [19] удалось сохранить 200 Мбайт пользовательских данных, а уже в 2019-м авторы работы [20] смогли сохранить 16 Гбайт англоязычной Википедии.

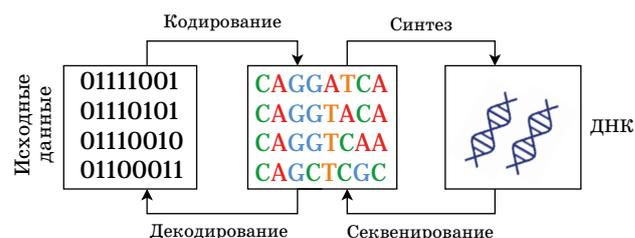
В последние годы к тематике ДНК-памяти прикован большой интерес исследовательского сообщества [21, 22], и многие ученые рассматривают ее как перспективный носитель для долговременного хранения гигантских объемов

информации, с которыми сегодня имеет дело человечество. Скорость размещения информации в ДНК-памяти и доступа к ней вряд ли в обозримом будущем сможет конкурировать с ныне используемыми носителями, однако по плотности информации, стоимости хранения и долговечности ДНК-память может их превзойти качественным образом.

Главным препятствием для практической реализации ДНК-памяти являются ошибки, т. е. модификации в последовательностях ДНК, неизбежно возникающие в процессе манипуляции с ДНК. В связи с этим в последние несколько лет произошел всплеск работ по применению идей теории кодирования для надежной реализации ДНК-памяти. Анализу типов ошибок и методам их коррекции уделено главное внимание в настоящем обзоре. В качестве основных источников информации использованы труды ведущих конференций по теории информации и теоретической информатике последних лет, а также специализированная научная периодика.

В упрощенном виде процедура хранения информации с использованием ДНК может быть представлена следующим образом: первоначально исходная информация (двоичный код) преобразуется в четверичный алфавит, соответствующий четырем нуклеотидам, составляющим ДНК. Затем в целях борьбы с возможными ошибками к полученной таким образом последовательности применяется помехоустойчивое кодирование. После чего сгенерированная синтетическая ДНК помещается в специальное хранилище или же встраивается в существующую ДНК живого организма. Для извлечения информации применяются процедуры секвенирования и декодирования имеющейся нуклеотидной последовательности [23–26]. Эти шаги схематически представлены на рис. 1 и будут рассмотрены более подробно в разделе «Технологии систем ДНК-памяти» данного обзора.

Ошибки, возникающие в системах ДНК-памяти, не ограничиваются простой заменой символов, а включают также выпадения и вставки, когда некоторые из передаваемых символов «те-



■ Рис. 1. Общая архитектура системы ДНК-памяти
 ■ Fig. 1. General architecture of DNA-based storage systems

ряются» или же, наоборот, в последовательность вставляются символы, ей не принадлежащие. Встречаются также определенные комбинации таких ошибок, в частности дубликации и пакетные выпадения или вставки [27–29]. Обзору наиболее часто встречающихся в системах ДНК-памяти ошибок, а также кодов, их исправляющих, посвящен раздел «Ошибки в системах ДНК-памяти и коды, их исправляющие» настоящего исследования.

Пользовательские данные представляются в системах ДНК-памяти в виде большого числа коротких последовательностей нуклеотидов, также называемых олигонуклеотидами. При этом в процессе манипуляций с ними, помимо вышеописанных ошибок на уровне символов, могут также происходить ошибки на уровне самих последовательностей, такие как изменение числа или удаление некоторых из них. Это, в свою очередь, приводит к изменению в пуле последовательностей на выходе системы ДНК-памяти по отношению ко входу. Рассмотрению возникающих математических моделей каналов передачи информации, а также построению кодов для них посвящен раздел «Модели каналов для систем ДНК-памяти».

Технологии систем ДНК-памяти

Рассмотрим более подробно этапы хранения информации на основе ДНК (см. рис. 1). Перед началом процедуры кодирования данные необходимо привести в формат, соответствующий четырем нуклеотидам (А, С, G, Т), образующим молекулы ДНК. В теории информации такая процедура называется кодированием источника [30, 31]. Так, например, в своей ранней работе [15] авторы использовали простое отображение букв английского алфавита, знаков препинания и цифр в последовательности из нуклеотидов по заранее определенному правилу, по которому буква D английского алфавита преобразуется в последовательность TTG, буква N — в TCT, буква A — в CGA. В результате слово DNA преобразуется в последовательность нуклеотидов TTGTCTCGA. Позднее другими учеными [17] была предложена схема кодирования с помощью ДНК относительно больших объемов разнородной информации, изначально представленной в html-файле. При этом символ 0 представлялся в виде нуклеотида А или С, выбираемого случайным образом. Символ же 1 представлялся в виде нуклеотида Т или G, также выбираемого случайно. В данном случае двоичная последовательность 0100 представлялась в виде AGAC. В работе [18] к данным перед их преобразованием в последовательность нуклеотидов авторы применили троичный ал-

горитм Хаффмана для кодирования источника. При этом кодирование в алфавит нуклеотидов было определено таким образом, чтобы избежать повторения подряд одного нуклеотида в целевой последовательности ДНК. Это достигалось с помощью специальной таблицы преобразования, определяющей правило отображения текущего троичного символа в нуклеотид в зависимости от значения предыдущего нуклеотида. В частности, при предыдущем нуклеотиде А символ 0 отображался в С, а при С — уже в G. При этом отображение первого символа определялось по правилу, при котором предыдущим нуклеотидом является А. Например, последовательность 0020 представлялась в виде CGCG. Данное требование вызвано повышенной вероятностью возникновения ошибки, свойственной современным технологиям секвенирования, при секвенировании гомополимерных (состоящих из одного нуклеотида) участков ДНК [32]. Другим важным для кодирования обстоятельством, влияющим на уровень ошибок и надежность хранения, является доля GC нуклеотидов [33]. Этот фактор учитывался, например, в работе [23], где кодирование было организовано таким образом, чтобы избежать гомополимерных участков длины больше 3, а также олигонуклеотидов с GC-содержанием больше 55 % либо меньше 45 %. Это достигалось путем предварительного применения преобразования Луби с различными псевдослучайными параметрами и отбрасывания неподходящих последовательностей. При этом преобразование бит в последовательность нуклеотидов осуществлялось по заранее определенному правилу $00 \rightarrow A$, $01 \rightarrow C$, $10 \rightarrow G$, $11 \rightarrow T$. В таком случае строка 0100 представлялась в виде CA.

Для борьбы с ошибками, возникающими в процессе хранения данных с использованием ДНК, в информационную последовательность вносят дополнительную избыточность. Различают физическую и логическую избыточность. Физическая избыточность предполагает увеличение «покрытия», или, иными словами, числа копий молекул ДНК, хранящих информацию об одном и том же участке исходной последовательности. Например, в работе [18] использовалось четырехкратное покрытие, тогда как в работе [34] для кодирования одной и той же информации использовалось несколько молекул ДНК, полученных путем «сдвига фазы» в процессе преобразования исходной двоичной последовательности в последовательность нуклеотидов. К сожалению, подобные методы не позволяют полностью застраховать информацию от возникающих ошибок. Другой независимый способ исправления возникающих ошибок — использование логической избыточности, задаваемой с помощью помехоустойчивых кодов. Этот метод

требует значительно меньшего объема дополнительной информации, что, в свою очередь, увеличивает итоговую плотность хранения. Коды, используемые для обнаружения и коррекции возникающих ошибок, в частности ошибок типа вставки и выпадения, будут рассмотрены в следующем разделе.

После приведения исходной последовательности к виду, соответствующему четырем нуклеотидам, ее необходимо преобразовать в нуклеотидную последовательность и поместить соответствующую молекулу ДНК в некоторое хранилище. Большинство существующих экспериментальных систем оперируют с искусственно синтезированными олигонуклеотидами, хранящимися в виде раствора, однако ДНК простейших живых организмов (как правило, бактерий или других микроорганизмов) также потенциально может быть использована для хранения синтезированной ДНК. Такой подход применен в ранней работе [14]. В дальнейшем было установлено [19], что бактерия способна нести в себе около одного мегабайта пользовательской информации, что сопоставимо с информацией, хранящейся в ее собственном геноме [35]. Несмотря на это кодирование информации *in vivo*, помимо очевидных технологических ограничений, вряд ли может быть масштабировано на большие объемы информации. В настоящем обзоре мы сосредоточим наше внимание на кодировании *in vitro*.

Современные технологии синтеза ДНК позволяют синтезировать одновременно на одном микрочипе множество коротких одноцепочечных олигонуклеотидов, представленных во многих экземплярах. Количество различных синтезируемых олигонуклеотидов может достигать нескольких миллионов, а каждый олигонуклеотид может быть представлен десятками или сотнями тысяч копий, однако из-за технологических погрешностей количество копий для разных олигонуклеотидов может существенно различаться [36–38]. С другой стороны, длины этих фрагментов очень малы, порядка 200 нуклеотидов. При этом сам процесс синтеза подвержен ошибкам типа замены, вставки и выпадения нуклеотидов, в результате которых экземпляры одного и того же олигонуклеотида могут слегка различаться. Отдельно отметим появление новых перспективных технологий синтеза, которые могут существенно сократить его стоимость и увеличить скорость, а также уменьшить число возникающих ошибок в ближайшем будущем [39].

Для представления исходной последовательности в виде набора олигонуклеотидов каждый олигонуклеотид должен содержать информацию, обычно называемую индексом, о позиции соответствующего фрагмента во входной последовательности. Кроме того, по краям олигонуклео-

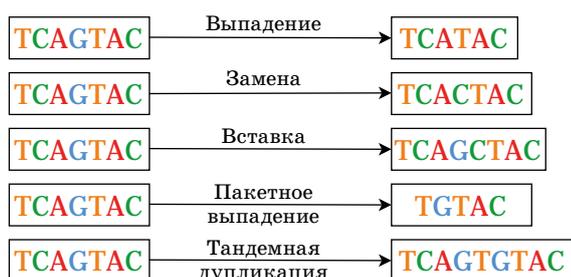
тида добавляются специальные последовательности — праймеры для полимеразной цепной реакции (ПЦР). С помощью ПЦР синтезированные олигонуклеотиды затем амплифицируются, т. е. «размножаются» до большего числа копий. Этот этап также позволяет отфильтровать «неправильно синтезированные» олигонуклеотиды, не содержащие праймеров ПЦР, однако вносит дополнительную неравномерность в число копий олигонуклеотидов из-за неравномерной амплификации.

Для извлечения информации из ДНК-памяти применяются стандартные технологии секвенирования ДНК [40]. С появлением так называемых секвенаторов нового поколения в середине 2000-х годов эти технологии вышли на качественно новый уровень, позволяя считывать фрагменты ДНК общим объемом миллиарды баз (нуклеотидов) за один цикл работы. Одними из наиболее распространенных на сегодня являются секвенаторы компании Illumina. По сравнению с другими высокопроизводительными технологиями, секвенаторы Illumina характеризуются относительно низким уровнем ошибок, порядка 0,1 %. Подавляющее большинство этих ошибок являются заменами, причем частота вставок и выпадений примерно на два порядка ниже частоты замен. При этом, как отмечалось, уровень ошибок зависит от последовательности: он может быть существенно выше для последовательностей с очень низким или, наоборот, очень высоким GC-содержанием, а также для последовательностей, содержащих длинные (шесть нуклеотидов и более) гомополимерные фрагменты [37].

Ошибки в системах ДНК-памяти и коды, их исправляющие

Типичными ошибками в процессе манипуляции с молекулами ДНК являются замены, вставки и выпадения нуклеотидов. Кроме того, случаются пакетные вставки и выпадения, т. е. одновременные вставки или выпадения коротких фрагментов, а также дубликации, когда в последовательность вставляется копия некоторого его фрагмента. Примеры ошибок представлены на рис. 2.

Модель, когда ошибки ограничиваются только заменами символов, хорошо изучена в теории кодирования, и для нее известно много конструкций кодов с хорошими корректирующими свойствами, в том числе конструкции оптимальных кодов. Модель ошибок типа вставки и выпадения, исследование которой началось в работах [27–29], оказывается гораздо более сложной. В этом разделе мы сосредоточимся на кодах, исправляющих именно такие ошибки, а также их



■ **Рис. 2.** Примеры ошибок в молекулах ДНК
 ■ **Fig. 2.** Examples of errors in DNA-molecules

комбинации. Отметим, что некоторые существующие модели систем ДНК-памяти ограничиваются рассмотрением кодов, исправляющих лишь ошибки замены [41], хотя эта модель ошибок плохо соответствует практическим системам.

В некоторых работах для исправления вставок/выпадения применялись различные эвристики, например, в работе [42] в центр каждой последовательности ДНК вставлялся специальный маркер, позволяющий локализовать данную ошибку и тем самым привести ее к ошибке замены. В качестве корректирующих кодов в работе [42] рассматривались коды с малой плотностью проверок на четность длины 256. В работах [43–46] применялись коды Рида — Соломона, а в [23] применены фонтанные коды. Основным недостатком данных конструкций является то, что результирующие коды гарантируют исправление лишь небольшого числа ошибок вставки/выпадения, что делает затруднительным применение этих кодов в системах ДНК-памяти.

Рассмотрим теперь ошибки типа вставки и выпадения. К первой конструкции кодов для их исправления можно отнести коды Варшавова — Тененгольца [28, 47]. Такие коды состоят из всех двоичных векторов длины n , поэлементная сумма которых принимает некоторое фиксированное значение по модулю n . Отметим, что данные коды исправляют одну ошибку вставки или выпадения и являются асимптотически оптимальными. Систематическая версия данных кодов представлена в работе [48], а на их основе были построены асимптотически оптимальные коды, исправляющие две последовательные вставки или выпадения [49].

Обобщение кодов Варшавова — Тененгольца на случай большего числа вставок и выпадений является сложной научной задачей. Большинство предложенных явных кодовых конструкций, в частности конструкция из работы [50], обобщающая их на случай исправления до пяти вставок и выпадений, являются неоптимальными с точки зрения кодовой скорости, а также не обладают эффективными алгоритмами кодирования и де-

кодирования. Коды, исправляющие до двух вставок и выпадений и улучшающие конструкцию [50], были получены с помощью построения кодовой книги путем перебора [51]. Конструкции, не использующие перебор, получены в работе [52].

Первая конструкция кодов для исправления фиксированного числа вставок и выпадений, обладающая малой избыточностью и эффективными алгоритмами кодирования и декодирования, была представлена в работе [53]. Эта конструкция обобщает идею разделения кодового слова на блоки фиксированного размера и последующего отделения их друг от друга путем добавления длинных последовательностей из нулей [54]. Процедура декодирования исходного кодового слова при этом осуществляется по мажоритарному правилу. Впоследствии полученный результат был улучшен [55, 56].

Все вышеописанные конструкции подразумевают исправление фиксированного числа вставок и выпадений, в то время как наиболее интересным с практической точки зрения является случай фиксированной доли вставок и выпадений. Первые коды для этого случая были построены в работе [57]. Данная конструкция основывается на каскадной схеме, подразумевающей использование внутренних кодов, полученных с помощью перебора и внешних кодов Рида — Соломона, и обладает полиномиальными алгоритмами кодирования и декодирования. Верхняя граница на скорость кодов, обладающих данным свойством, получена в работе [58]. Там же показано, что коды из работы [57] являются асимптотически оптимальными.

Конструкции кодов, исправляющих вставки и выпадения для двух наиболее интересных случаев: высокой кодовой скорости и большой доли вставок/выпадений, — были предложены в работе [54]. Эти конструкции также основываются на каскадной схеме, состоящей из кода с хорошими корректирующими свойствами с точки зрения исправления ошибок замены (в частности, кода Рида — Соломона и кодов Парвареша — Варди) и кодов для исправления вставок/выпадений со значительно меньшей длиной блока. Последнее свойство позволяет декодировать данный код методом простого перебора. При этом в случае высокой доли вставок/выпадений необходимо специальным образом учитывать влияние неправильного декодирования внутреннего кода на внешний код. В случае высокой кодовой скорости, в свою очередь, необходимо разделять блоки внутренних кодовых слов с помощью блоков из последовательных нулей. Ключевой особенностью предложенного подхода является определение параметров декодирования внутреннего кода, позволяющее внешнему коду восстановить исходное сообщение с помощью списочного де-

кодера. Улучшение данной конструкции для исправления большей доли выпадений предложено в работе [59]. Обобщение на случай фиксированной доли как вставок, так и выпадений представлено в работе [60].

Еще одним методом, позволяющим исправлять вставки и выпадения, является преобразование корректирующих кодов для ошибок замены с помощью строк синхронизации [61]. Его основная идея состоит в добавлении к каждому символу исходной кодовой последовательности специальным образом построенной синхронизирующей строки над малым алфавитом и исправлении произошедших ошибок типа вставки/выпадения с помощью декодера исходного кода.

Детерминированный способ построения синхронизирующих строк предложен в работе [62]. В случае если после передачи по каналу со вставками/выпадениями получена синхронизирующая строка S' , то алгоритм приведения ошибок типа вставки/выпадения к ошибкам замены последовательно находит наибольшие соответствия между S и S' . Если какой-либо элемент исходной строки находится во всех полученных ранее соответствиях, то он поступает на вход декодера исходного кода. В работе [61] приведены оценки на число исправляемых таким образом вставок и выпадений, связанные с корректирующей способностью применяемого кода и свойствами синхронизирующей строки.

Последние результаты для кодов, исправляющих фиксированную долю вставок и выпадений, получены в работах [63, 64] в рамках исследования связанной задачи об обмене документами, популярной в области теоретической информатики. В этом контексте кодирование происходит с помощью добавления к исходной строке дополнительного сообщения — скетча (sketch), с помощью которого подверженная ошибкам исходная строка может быть впоследствии декодирована.

Отметим, что приведенные выше коды способны эффективно исправлять фиксированную долю ошибок лишь определенного типа, тогда как в системах ДНК-памяти могут одновременно происходить как вставки и выпадения, так и замены символов. При этом различные типы ошибок имеют различные вероятности, определяемые структурой конкретной системы.

Важным обобщением ошибок вставки и выпадения индивидуальных символов, возникающих в системах ДНК-памяти, являются вставки или выпадения последовательных символов, также называемых пакетными. Впервые такая постановка задачи введена в работе [49], в которой рассматривались группы, состоящие не более чем из двух последовательных выпадений, для которых были предложены корректирующие коды с оптимальной избыточностью. Отметим различие

между кодами, исправляющими до k последовательных вставок или выпадений, и кодами, исправляющими ровно k последовательных вставок или выпадений. Применение последних, в свою очередь, не гарантирует исправление меньшего числа ошибок в общем случае.

В работе [65] исследованы коды, исправляющие k последовательных вставок или выпадений (но не комбинацию их), обобщенные также на случай не более k последовательных вставок или выпадений. Результат для случая ровно k последовательных вставок или выпадений был в дальнейшем улучшен в работе [66], а для случая не более k таких ошибок — в работах [67, 68]. В частности, в работе [68] получены коды с асимптотически оптимальным значением избыточности. При этом все отмеченные выше конструкции основываются на различных модификациях кодов Варшавова — Тененгольца. Отметим, что возможный случай нескольких пакетных выпадений ограниченной длины является менее изученной проблемой. Для решения последней в работе [69] предложен метод синдромной компрессии (syndrome compression) в применении к конструкции кодов из работы [53]. При этом полученные таким образом коды не являются асимптотически оптимальными.

Другим распространенным типом вставок являются дубликации, когда некоторый фрагмент последовательности копируется в позицию справа от него. Ошибки дубликации были впервые рассмотрены в работе [70], в которой исследован случай дубликации одного символа в двоичном алфавите и получена асимптотическая верхняя граница на мощность такого кода. Асимптотически оптимальные коды для исправления фиксированного числа ошибок дубликации одного символа, а также коды с эффективными алгоритмами кодирования и декодирования предложены в работах [71, 72]. В [73] выведена оптимальная конструкция для исправления неограниченного числа дубликаций фиксированной длины. Кроме того, в этой работе проведено обобщение на случай тандемных дубликаций, при которых фрагмент преобразуется в две последовательные копии. Для этого случая была представлена конструкция, исправляющая тандемные дубликации длины до 1, 2 или 3. Оптимальность этих кодов показана в работе [74]. В работе [75] рассмотрена задача исправления одной тандемной дубликации, но уже ограниченной длины. Авторами предложены границы на мощность таких кодов, а также явные конструкции на основе кодов Варшавова — Тененгольца, не являющиеся оптимальными.

В заключение раздела отметим, что в случае систем ДНК-памяти конфигурации и доля возможных ошибок определяются их характери-

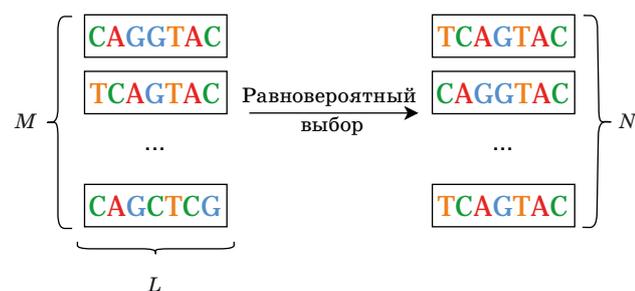
стиками. Это ставит задачу построения для них оптимальных кодов, обладающих вычислительно-эффективными алгоритмами кодирования и декодирования. Помимо этого, важной является задача дальнейшего улучшения существующих кодовых конструкций, в частности кодов, исправляющих несколько пакетных выпадений фиксированной длины, и кодов, исправляющих тандемные дубликации ограниченной длины.

Модели каналов для систем ДНК-памяти

Как отмечалось ранее, при хранении информации с использованием ДНК пользовательские данные преобразуются в большое число олигонуклеотидов. В ходе манипуляций с ними, а именно их синтеза и чтения (секвенирования), возникают ошибки, что приводит к появлению новых олигонуклеотидов. При этом даже в процессе хранения олигонуклеотиды могут быть подвержены химическим преобразованиям, которые приводят к «потере» некоторых из них на этапе чтения [37]. Наконец, в силу специфических свойств ПЦР распределение числа копий олигонуклеотидов весьма неравномерно: некоторые могут быть представлены намного большим числом копий (возможно, содержащих ошибки), чем другие. Это приводит к задаче моделирования всего процесса хранения информации с помощью ДНК-памяти, т. е. построению соответствующих ей теоретико-информационных моделей каналов, а также изучению их характеристик и кодов для них.

В одной из первых работ в этом направлении была рассмотрена модель, в которой на вход канала поступает набор из M последовательностей длины L . На выход, в свою очередь, поступает набор из N последовательностей, равновероятно выбранных из входных [76]. Такой канал проиллюстрирован на рис. 3.

Авторами работы получена оценка на пропускную способность данного канала. Отметим, что такая постановка задачи исключает возможные ошибки в последовательностях, а пред-



■ *Рис. 3.* Пример модели канала ДНК-памяти
 ■ *Fig. 3.* Example of DNA-storage channel model

положение о равновероятном выборе не вполне адекватно отражает реальность. В дальнейшем эта постановка была расширена в работе [77], в которой в последовательностях могли происходить ошибки замены. Схема кодов, достигающих пропускной способности, основанная на случайном кодировании и декодере, группирующем последовательности в соответствии с расстоянием Хэмминга между ними, представлена в работе [78].

Упрощенная постановка вышеописанных задач рассмотрена в работах [79, 80], где последовательности лишь перемешивались без проведения процедуры случайного выбора. При этом были рассмотрены варианты данных каналов с ошибками замены и стираниями, получены значения пропускной способности для них, а также приведены схемы, с помощью которых они достигаются. В последних применялись кодирования порядка с помощью индексов, а также коды, достигающие пропускной способности для двоично-симметричного канала или же двоично-стирающего канала. Отметим, что в этих работах не рассматривались ошибки типа вставки/выпадения, для которых до сих пор неизвестны точные значения пропускной способности для всех областей возможных параметров [81], а также способы получения выходных последовательностей, отличные от равновероятного выбора.

Еще одной возможной моделью каналов для ДНК-памяти является модель, представленная в работе [82]. В ней на вход канала поступает набор из M последовательностей длины L . На выход канала поступает набор из $M - s$ последовательностей, из которых в t последовательностях происходит до e ошибок типа вставки/выпадения, а также замены. Параметры данной модели являются фиксированными величинами. Интересным фактом является то, что в рассматриваемых каналах код, исправляющий выпадения, не способен исправлять и вставки, и выпадения. Это отличает рассматриваемую задачу от задачи исправления ошибок в символах последовательностей, рассмотренной в предыдущем разделе. Авторами работы [82] выведены границы на избыточность данных кодов, а также представлены конструкции на основе индексации и кодирования всех информационных подпоследовательностей кодом с максимально достижимым кодовым расстоянием. Также предложен способ дальнейшего уменьшения избыточности с помощью использования наиболее значимых бит в индексах. Отметим, что данная модель не учитывает неравномерность распределения числа копий различных последовательностей, а также различия в вероятностях разных типов ошибок.

Интересной теоретико-информационной моделью каналов для ДНК-памяти также является

ся модель, представленная в работе [83]. В ней, как и в предыдущем случае, на вход поступает набор из M последовательностей длины L . На выход же канала поступает набор из T последовательностей, в которых произошло K замен, где $M - K \leq T \leq M$. При этом считается, что возможная потеря последовательностей является следствием произошедших в них ошибок замены. С помощью границы для упаковки сфер авторами получена граница для существования кодов, исправляющих ошибки в таком канале. Кроме того, была представлена явная конструкция, исправляющая несколько произошедших замен, основанная на идее конкатенации последовательностей в специальном порядке и применении кода Рида — Соломона. Существенным недостатком данной модели является отсутствие учета возможных ошибок типа вставки/выпадения.

Кроме описанных, еще одной недавно предложенной моделью канала для ДНК-памяти [84] является модель, в которой на вход канала поступает набор из M последовательностей длины L , а на выходе получается набор из уже M последовательностей, часть из которых добавлена дополнительно. При этом часть исходных последовательностей может отсутствовать на выходе канала, а в части могут происходить ошибки замены. Отметим, что, в отличие от предыдущих постановок, M не обязательно является фиксированной величиной. Для работы в данном канале авторами была введена новая метрика, называемая последовательно-множественным расстоянием, и предложены способы исправления ошибок в ней на основе кодов постоянного веса и кодирования индексов. Подчеркнем, что, как и в предыдущем случае, авторы не рассматривают возможные ошибки типа вставки/выпадения.

В заключение добавим, что построение новых теоретико-информационных моделей каналов ДНК-памяти, более точно учитывающих нерав-

номерность числа копий различных последовательностей и возможные типы ошибок в них, по-прежнему является важной теоретической задачей.

Заключение

Системы ДНК-памяти являются перспективным способом хранения пользовательской информации, но на сегодня они еще далеки от практического внедрения. Этапы «записи» и «чтения» информации пока остаются более медленными и более дорогими в сравнении с традиционными носителями. Кроме того, проблеме борьбы с различными ошибками, возникающими на этапах синтеза, хранения и секвенирования ДНК, нельзя считать полностью решенной. Однако на этом пути в последние годы достигнут осязаемый прогресс, что и явилось главной темой настоящего обзора. Как часто случается, возникновение новой области приложений дает импульс для появления и развития новых моделей, алгоритмов и теоретических результатов, мотивированных новыми постановками задач. В данном случае это относится к теории информации и теории кодирования. Зачастую эти результаты представляют независимый интерес, что мы также попытались отразить в настоящем обзоре. Скажем в заключение, что обзор посвящен теоретико-информационным аспектам и не претендует на полное освещение этой новой области на стыке теории информации и биоинформатики.

Финансовая поддержка

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов 19-01-00364, 20-07-00652 и 20-17-50170.

Литература

1. Aftab U., Siddiqui G. F. Big data augmentation with data warehouse: a survey. *Proceedings of 2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018, pp. 2775–2784. doi:10.1109/BigData.2018.8622182
2. Lee S. Y. DNA data storage is closer than you think. 2019. <https://www.scientificamerican.com/article/dna-data-storage-is-closer-than-you-think/> (дата обращения: 12.03.2021).
3. Reinsel D., Gantz J., Rydning J. The digitization of the world from edge to core. *An IDC White Paper US4413318*, 2018. <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-data-age-whitepaper.pdf> (дата обращения: 12.03.2021).
4. Bohannon J. DNA: The ultimate hard drive. *Science*, 2012. <https://www.sciencemag.org/news/2012/08/dna-ultimate-hard-drive> (дата обращения: 12.03.2021).
5. Colen C. DNA data storage — setting the data density record with DNA fountain. 2017. <https://www.twistbioscience.com/blog/perspectives/dna-data-storage-setting-data-density-record-dna-fountain> (дата обращения: 12.03.2021).
6. Ceze L., Nivala J., Strauss K. Molecular digital data storage using DNA. *Nature Reviews Genetics*, 2019, vol. 20, pp. 456–466. doi:10.1038/s41576-019-0125-3
7. Extance A. How DNA could store all the world's data. *Nature*, 2016, vol. 537, pp. 22–24. doi:10.1038/537022a
8. Zhirnov V., Zadegan R. M., Sandhu G. S., Church G. M., Hughes W. L. Nucleic acid memory.

- Nature Material*, 2016, vol. 15, pp. 366–370. doi:10.1038/nmat4594
9. **Watson J. D., Crick F. H.** Molecular structure of nucleic acids. *Nature*, 1953, vol. 171, pp. 737–738. doi:10.1038/171737a0
 10. **Wiener N.** Interview: machines smarter than men? *US News World Rep*, 1964, vol. 56, pp. 84–86.
 11. **Нейман М. С.** Некоторые фундаментальные вопросы микроминиатюризации. *Радиотехника*, 1964, т. 12, № 1, с. 3–12. https://2a008ed5-a-62cb3a1a-s-sites.googlegroups.com/site/msneiman1905/Neiman-1964_Micromini.pdf (дата обращения: 12.03.2021).
 12. **Нейман М. С.** О молекулярных системах памяти и направлениях мутаций. *Радиотехника*, 1965, т. 20, № 6, с. 1–8. https://2a008ed5-a-62cb3a1a-s-sites.googlegroups.com/site/msneiman1905/Neiman-1965_Molecul.pdf (дата обращения: 12.03.2021).
 13. **Реброва И. М., Реброва О. Ю.** Запоминающие устройства на основе искусственной ДНК: рождение идеи и первые публикации. *Вопросы истории естествознания и техники*, 2019, т. 41, № 4, с. 666–676. doi:10.31857/S020596060013006-8
 14. **Davis J.** Microvenus. *Art Journal*, 1996, vol. 55, no. 1, pp. 70–74. doi:10.2307/777811
 15. **Clelland C. T., Risca V., Bancroft C.** Hiding messages in DNA microdots. *Nature*, 1999, vol. 399, pp. 533–534. doi:10.1038/21092
 16. **Bancroft C., Bowler T., Bloom B.** Long-term storage of information in DNA. *Science*, 2001, vol. 293, iss. 5536, pp. 1763–1765. doi:10.1126/science.293.5536.1763c
 17. **Church G. M., Gao Y., Kosuri S.** Next-generation digital information storage in DNA. *Science*, 2012, vol. 337, pp. 1628. doi:10.1126/science.1226355
 18. **Goldman N., Bertone P., Chen S., Dessimoz C., Leproust E. M., Sipos B., Birney E.** Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, 2013, vol. 494, pp. 77–79. doi:10.1038/nature11875.
 19. **Organick L., Ang S. D., Chen Y., et al.** Random access in large-scale DNA data storage. *Nature Biotechnology*, 2018, vol. 36, pp. 242–248. doi:10.1038/nbt.4079
 20. **Shankland S.** Startup packs all 16GB of Wikipedia onto DNA strands to demonstrate new storage tech. *Cnet*, 2019. <https://www.cnet.com/news/startup-packs-all-16gb-wikipedia-onto-dna-strands-demonstrate-new-storage-tech/> (дата обращения: 12.03.2021).
 21. **Extance A.** How DNA could store all the world's data. *Nature*, 2016, vol. 537, pp. 22–24. doi:10.1038/537022a
 22. **Service R.** DNA could store all of the world's data in one room. *Science*, 2017. doi:10.1126/science.aal0852. <https://www.sciencemag.org/news/2017/03/dna-could-store-all-worlds-data-one-room> (дата обращения: 12.03.2021).
 23. **Erlich Y., Zielinski D.** DNA fountain enables a robust and efficient storage architecture. *Science*, 2017, vol. 355, iss. 6328, pp. 950–954. doi:10.1126/science.aaj2038
 24. **Yazdi H. T., Kiah H. M., Garcia-Ruiz E., Ma J., Zhao H., Milenkovic O.** DNA-based storage: trends and methods. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2015, vol. 1, no. 3, pp. 230–248. doi:10.1109/TMBMC.2016.2537305
 25. **Yiming D., Fajia S., Zhi P., Qi O., Long Q.** DNA storage: research landscape and future prospects. *National Science Review*, 2020, vol. 7, iss. 6, pp. 1092–1104. doi:10.1093/nsr/nwaa007
 26. **Yim S. S., McBee R. M., Song A. M., Huang Y., Sheth R. U., Wang H. H.** Robust direct digital-to-biological data storage in living cells. *Nature Chemical Biology*, 2021, vol. 17, pp. 246–253. doi:10.1038/s41589-020-00711-4
 27. **Gilbert E.** Synchronization of binary messages. *IRE Transactions on Information Theory*, 1960, vol. 6, no. 4, pp. 470–477. doi:10.1109/TIT.1960.1057587
 28. **Левенштейн В. И.** Двоичные коды с исправлением выпадений, вставок и замещений символов. *Доклады АН СССР*, 1965, т. 163, с. 845–848.
 29. **Sellers F.** Bit loss and gain correction code. *IRE Transactions on Information Theory*, 1962, vol. 8, no. 1, pp. 35–38. doi:10.1109/TIT.1962.1057684
 30. **Xiong Z.** Multiterminal source coding: theory, code design and applications. *Proceedings of the Fifth International Workshop on Signal Design and its Applications in Communications*, IEEE, 2011, pp. 3–3. doi:10.1109/IWSDA.2011.6159430
 31. **Luby M.** LT codes. *The 43rd Annual IEEE Symposium on Foundations of Computer Science*, IEEE, 2002, pp. 271–280. doi:10.1109/SFCS.2002.1181950
 32. **Niedringhaus T. P., Milanova D., Kerby M. B., Snyder M. P., Barron A. E.** Landscape of next-generation sequencing technologies. *Analytical Chemistry*, 2011, vol. 83, no. 12, pp. 4327–4341. doi:10.1021/ac2010857
 33. **Schwartz J. J., Lee C., Shendure J.** Accurate gene synthesis with tag-directed retrieval of sequence-verified DNA molecules. *Nature Methods*, 2012, vol. 9, pp. 913–915. doi:10.1038/nmeth.2137
 34. **Yachie N., Sekiyama K., Sugahara J., Ohashi Y., Tomita M.** Alignment-based approach for durable data storage into living organisms. *Biotechnol Progress*, 2007, vol. 23, no. 2, pp. 501–505. doi:10.1021/bp060261y
 35. **Itaya M., Tsuge K., Koizumi M., Fujita K.** Combining two genomes in one cell: stable cloning of the *synechocystis* PCC6803 genome in the *bacillus subtilis* 168 genome. *Proceedings of the National Academy of Sciences of the USA*, 2005, vol. 102, no. 44, pp. 15971–15976. doi:10.1073/pnas.0503868102
 36. **Kosuri S., Church G. M.** Large-scale de novo DNA synthesis: technologies and applications. *Nature Methods*, 2014, vol. 11, no. 5, pp. 499–507. doi:10.1038/nmeth.2918
 37. **Heckel R., Mikutis G., Grass R. N.** A characterization of the DNA data storage channel. *Science Reports*, 2019, vol. 9, pp. 1–10. doi:10.1038/s41598-019-45832-6

38. Schmidt T., Beliveau B., Uca Y., Theilmann M., Cruz F. D., Wu C.-T., Shih W. M. Scalable amplification of strand subsets from chip-synthesized oligonucleotide libraries. *Nature Communications*, 2015, vol. 6, pp. 1–11. doi:10.1038/ncomms963439.
39. Lee H. H., Kalhor R., Goela N., Bolot J., Church G. M. Terminator-free template-independent enzymatic DNA synthesis for digital information storage. *Nature Communications*, 2019, vol. 10, pp. 1–12. doi:10.1038/s41467-019-10258-1
40. Kulski J. K. Next-generation Sequencing — an overview of the history, tools, and “omic” applications. Next generation sequencing — advances, applications and challenges. *IntechOpen*, 2016, pp. 1–15. doi:10.5772/61964
41. Organick L., Chen Y.-J., Ang S. D., Lopez R., Liu X., Strauss K., Ceze L. Probing the physical limits of reliable DNA data retrieval. *Nature Communications*, 2020, vol. 11, pp. 1–12. doi:10.1038/s41467-020-14319-8
42. Chandak S., Tatwawadi K., Lau B., Mardia J., Kubit M., Neu J., Griffin P., Wootters M., Weissman T., Ji H. Improved read/write cost tradeoff in DNA-based data storage using LDPC codes. *Proceedings of 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2019, pp. 147–156. doi:10.1109/ALLERTON.2019.8919890
43. Grass R. N., Heckel R., Puddu M., Paunescu D., Stark W. J. Robust chemical preservation of digital information on DNA in silica with error-correcting codes. *Angewandte Chemie*, 2015, vol. 54, pp. 2552–2555. doi:10.1002/anie.201411378
44. Anavy L., Vaknin I., Atar O., Amit R., Yakhini Z. Data storage in DNA with fewer synthesis cycles using composite DNA letters. *Nature Biotechnology*, 2019, vol. 37, pp. 1229–1236. doi:10.1038/s41587-019-0240-x
45. Meiser L. C., Antkowiak P. L., Koch J., Chen W. D., Kohll A. X., Stark W. J., Heckel R., Grass R. N. Reading and writing digital data in DNA. *Nature Protocols*, 2020, vol. 15, pp. 86–101. doi:10.1038/s41596-019-0244-5
46. Lopez R., Chen Y.-J., Ang D. S., Yekhanin S., Makarychev K., Racz M. Z., Seelig G., Strauss K., Ceze L. DNA assembly for nanopore data storage readout. *Nature Communications*, 2019, vol. 10, pp. 1–9. doi:10.1038/s41467-019-10978-4
47. Варшамов Р. Р., Тененгольц Г. М. Код, исправляющий одиночные несимметричные ошибки. *Автоматика и телемеханика*, 1965, т. 26, № 2, с. 286–290.
48. Saowapa K., Kaneko H., Fujiwara E. Systematic binary deletion/insertion error-correcting codes capable of correcting random bit errors. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 2000, vol. E83-A, no. 12, pp. 2699–2705.
49. Левенштейн В. И. Асимптотически оптимальный двоичный код с исправлением выпадений одного или двух соседних символов. *Проблемы кибернетики*, 1967, т. 19, с. 298–304.
50. Helberg A. S. J., Ferreira H. C. On multiple insertion/deletion correcting codes. *IEEE Transactions on Information Theory*, 2002, vol. 48, no. 1, pp. 305–308. doi:10.1109/18.971760
51. Swart T. G., Ferreira H. C. A note on double insertion/deletion correcting codes. *IEEE Transactions on Information Theory*, 2003, vol. 49, no. 1, pp. 269–273. doi:10.1109/TIT.2002.806155
52. Guruswami V., Hastad J. Explicit two-deletion codes with redundancy matching the existential bound. *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, SIAM, 2021, pp. 1–8. doi:10.1137/1.9781611976465.2
53. Brakensiek J., Guruswami V., Zbarsky S. Efficient low-redundancy codes for correcting multiple deletions. *IEEE Transactions on Information Theory*, 2018, vol. 64, no. 5, pp. 3403–3410. doi:10.1109/TIT.2017.2746566
54. Guruswami V., Wang C. Deletion codes in the high-noise and high-rate regimes. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 4, pp. 1961–1970. doi:10.1109/TIT.2017.2659765
55. Sima J., and Bruck J. On optimal k-deletion correcting codes. *IEEE Transactions on Information Theory*, 2021. doi:10.1109/TIT.2020.3028702
56. Sima J., Gabrys R., and Bruck J. Optimal systematic t-deletion correcting codes. *Proceedings of 2020 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2020, pp. 769–774. doi:10.1109/ISIT44484.2020.9173986
57. Schulman L. J., Zuckerman D. Asymptotically good codes correcting insertions, deletions, and transpositions. *IEEE Transactions on Information Theory*, 1999, vol. 45, no. 7, pp. 2552–2557. doi:10.1109/18.796406
58. Kulkarni A. A., Kiyavash N. Nonasymptotic upper bounds for deletion correcting codes. *IEEE Transactions on Information Theory*, 2013, vol. 59, no. 8, pp. 5115–5130. doi:10.1109/TIT.2013.2257917
59. Bukh B., Guruswami V., and Hastad J. An improved bound on the fraction of correctable deletions. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 1, pp. 93–103. doi:10.1109/TIT.2016.2621044
60. Guruswami V., and Li R. Efficiently decodable insertion/deletion codes for high-noise and high-rate regimes. *Proceedings of 2016 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2016, pp. 620–624. doi:10.1109/ISIT.2016.7541373
61. Haeupler B., Shahrasbi A. Synchronization strings: codes for insertions and deletions approaching the Singleton bound. *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2017)*, ACM, 2017, pp. 33–46. doi:10.1145/3055399.3055498
62. Cheng K., Haeupler B., Li X., Shahrasbi A., Wu K. Synchronization strings: highly efficient determinis-

- tic constructions over small alphabets. *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM, 2019, pp. 2185–2204. doi:10.1137/1.9781611975482.132
63. **Cheng K., Jin Z., Li X., and Wu K.** Deterministic document exchange protocols, and almost optimal binary codes for edit errors. *Proceedings of 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 2018, pp. 200–211. doi:10.1109/FOCS.2018.00028
64. **Haeupler B.** Optimal document exchange and new codes for insertions and deletions. *Proceedings 2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 2019, pp. 334–347. doi:10.1109/FOCS.2019.00029
65. **Schoeny C., Wachter-Zeh A., Gabrys R., Yaakobi E.** Codes correcting a burst of deletions or insertions. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 4, pp. 1971–1985. doi:10.1109/TIT.2017.2661747
66. **Saeki T., Nozaki T.** An improvement of non-binary code correcting single b-burst of insertions or deletions. *Proceedings of 2018 International Symposium on Information Theory and its Applications (ISITA)*, IEICE, 2018, pp. 6–10. doi:10.23919/ISITA.2018.8664217
67. **Gabrys R., Yaakobi E., Milenkovic O.** Codes in the damerau distance for deletion and adjacent transposition correction. *IEEE Transactions on Information Theory*, 2018, vol. 64, no. 4, pp. 2550–2570. doi:10.1109/TIT.2017.2778143
68. **Lenz A., Polyanskii N.** Optimal codes correcting a burst of deletions of variable length. *2020 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2020, pp. 757–762. doi:10.1109/ISIT44484.2020.9174288
69. **Sima J., Gabrys R., and Bruck J.** Syndrome compression for optimal redundancy codes. *Proceedings of 2020 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2020, pp. 751–756. doi:10.1109/ISIT44484.2020.9174009
70. **Левенштейн В. И.** Двоичные коды с исправлением вставок и выпадений символа 1. *Проблемы передачи информации*, 1965, т. 1, № 1, с. 8–17.
71. **Dolecek L., Anantharam V.** Repetition error correcting sets: explicit constructions and prefixing methods. *SIAM Journal on Discrete Mathematics*, 2010, vol. 23, no. 4, pp. 2120–2146. doi:10.5555/1958171.1958195
72. **Mahdavifar H., Vardy A.** Asymptotically optimal sticky-insertion correcting codes with efficient encoding and decoding. *Proceedings of 2017 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 2688–2692. doi:10.1109/ISIT.2017.8007016
73. **Jain S., Farnoud F., Schwartz M., Bruck J.** Duplication-correcting codes for data storage in the DNA of living organisms. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 8, pp. 4996–5010. doi:10.1109/ISIT.2016.7541455
74. **Kovacevic M.** Codes correcting all patterns of tandem-duplication errors of maximum length 3. <https://arxiv.org/pdf/1911.06561.pdf> (дата обращения: 12.03.2021).
75. **Nazirkhanova K., Medova L., Kruglik S., and Frolov A.** Codes correcting bounded length tandem duplication. *Proceedings of 2020 International Symposium on Information Theory and its Applications (ISITA)*, IEICE, 2020, pp. 299–303. doi:10.34385/proc.65.B06-6
76. **Heckel R., Shomorony I., Ramchandran K., Tse D. N. C.** Fundamental limits of DNA storage systems. *Proceedings of 2017 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 3130–3134. doi:10.1109/ISIT.2017.8007106
77. **Lenz A., Siegel P. H., Wachter-Zeh A., Yaakobi E.** An upper bound on the capacity of the DNA storage channel. *2019 IEEE Information Theory Workshop (ITW)*, IEEE, 2019, pp. 1–5. doi:10.1109/ITW44776.2019.8989388
78. **Lenz A., Siegel P. H., Wachter-Zeh A., Yaakobi E.** Achieving the capacity of the DNA storage channel. *Proceedings of ICASSP 2020 — 2020 IEEE International Conference on Acoustics, Speech and Signal*, IEEE, 2020, pp. 8846–8850. doi:10.1109/ICASSP40776.2020.9053049
79. **Shin S., Heckel R., Shomorony I.** Capacity of the erasure shuffling channel. *Proceedings of ICASSP 2020 — 2020 IEEE International Conference on Acoustics, Speech and Signal*, IEEE, 2020, pp. 8841–8845. doi:10.1109/ICASSP40776.2020.9053486
80. **Shomorony I., Heckel R.** Capacity results for the noisy shuffling channel. *Proceedings of 2019 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2019, pp. 762–766. doi:10.1109/ISIT.2019.8849789
81. **Mitzenmacher M.** A survey of results for deletion channels and related synchronization channels. *Probability Surveys*, 2009, vol. 6, pp. 1–33. doi:10.1214/08-PS141
82. **Lenz A., Siegel P. H., Wachter-Zeh A., Yaakobi E.** Coding over sets for DNA storage. *IEEE Transactions on Information Theory*, 2020, vol. 66, no. 4, pp. 2331–2351. doi:10.1109/TIT.2019.2961265
83. **Sima J., Raviv N., and Bruck J.** On coding over sliced information. *IEEE Transactions on Information Theory*, 2021. doi:10.1109/TIT.2021.3063709
84. **Song W., Cai K., Schouhamer Immink K. A.** Sequence-subset distance and coding for error control in DNA-based data storage. *IEEE Transactions on Information Theory*, 2020, vol. 66, no. 10, pp. 6048–6065. doi:10.1109/TIT.2020.3002611

UDC 004.074

doi:10.31799/1684-8853-2021-3-39-52

Information-theoretic problems of DNA-based storage systemsS. A. Kruglik^{a,b}, Junior Researcher, orcid.org/0000-0001-9557-5197,

stanislav.kruglik@skoltech.ru

G. A. Kucherov^{a,c}, PhD, Phys.-Math., Leading Researcher, orcid.org/0000-0001-5899-5424K. N. Nazirkhanova^d, Post-Graduate Student, orcid.org/0000-0002-7447-9857M. E. Filitov^a, Master Student, orcid.org/0000-0003-2421-0777^aSkolkovo Institute of Science and Technology, bld. 1, 30, Bolshoy Boulevard, 121205, Moscow, Russian Federation^bMoscow Institute of Physics and Technology, 9, Institutskiy Per., 141701, Dolgoprudny, Moscow region, Russian Federation^cCentre National de Recherche Scientifique, Université Gustave Eiffel, 77454 Marne-la-Vallée, France^dStanford University, 94305 Stanford, CA, USA

Introduction: Currently, we witness an explosive growth in the amount of information produced by humanity. This raises new fundamental problems of its efficient storage and processing. Commonly used magnetic, optical, and semiconductor information storage devices have several drawbacks related to small information density and limited durability. One of the promising novel approaches to solving these problems is DNA-based data storage. **Purpose:** An overview of modern DNA-based storage systems and related information-theoretic problems. **Results:** The current state of the art of DNA-based storage systems is reviewed. Types of errors occurring in them as well as corresponding error-correcting codes are analyzed. The disadvantages of these codes are shown, and possible pathways for improvement are mentioned. Proposed information-theoretic models of DNA-based storage systems are analyzed, and their limitation highlighted. In conclusion, main obstacles to practical implementation of DNA-based storage systems are formulated, which can be potentially overcome using information-theoretic methods considered in this overview.

Keywords — data storage systems, DNA-based memory, communication channel, channel capacity, substitution errors, insertion errors, deletion errors.

For citation: Kruglik S. A., Kucherov G. A., Nazirkhanova K. N., Filitov M. E. Information-theoretic problems of DNA-based storage systems. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 39–52 (In Russian). doi:10.31799/1684-8853-2021-3-39-52

Funding

The reported study was funded by RFBR, projects 19-01-00364, 20-07-00652 and 20-17-50170.

References

- Aftab U., Siddiqui G. F. Big data augmentation with data warehouse: a survey. *Proceedings of 2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018, pp. 2775–2784. doi:10.1109/BigData.2018.8622182
- Lee S. Y. *DNA data storage is closer than you think*. 2019. Available at: <https://www.scientificamerican.com/article/dna-data-storage-is-closer-than-you-think/> (accessed 12 March 2021).
- Reinsel D., Gantz J., Rydning J. The digitization of the world from edge to core. *An IDC White Paper US4413318*, 2018. Available at: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-data-age-whitepaper.pdf> (accessed 12 March 2021).
- Bohannon J. DNA: The ultimate hard drive. *Science*, 2012. Available at: <https://www.sciencemag.org/news/2012/08/dna-ultimate-hard-drive> (accessed 12 March 2021).
- Colen C. *DNA data storage — setting the data density record with DNA fountain*. 2017. Available at: <https://www.twistbioscience.com/blog/perspectives/dna-data-storage-setting-data-density-record-dna-fountain> (accessed 12 March 2021).
- Ceze L., Nivala J., Strauss K. Molecular digital data storage using DNA. *Nature Reviews Genetics*, 2019, vol. 20, pp. 456–466. doi:10.1038/s41576-019-0125-3
- Extance A. How DNA could store all the world's data. *Nature*, 2016, vol. 537, pp. 22–24. doi:10.1038/537022a
- Zhirnov V., Zadegan R. M., Sandhu G. S., Church G. M., Hughes W. L. Nucleic acid memory. *Nature Material*, 2016, vol. 15, pp. 366–370. doi:10.1038/nmat4594
- Watson J. D., Crick F. H. Molecular structure of nucleic acids. *Nature*, 1953, vol. 171, pp. 737–738. doi:10.1038/171737a0
- Wiener N. Interview: machines smarter than men? *US News World Rep*, 1964, vol. 56, pp. 84–86.
- Neiman M. S. Some fundamental issues of microminiaturization. *Radiotekhnika*, 1964, vol. 12, no. 1, pp. 3–12 (In Russian). Available at: https://2a008ed5-a62cb3a1a-sites.googlegroups.com/site/msneiman1905/Neiman-1964_Micromini.pdf (accessed 12 March 2021).
- Neiman M. S. On the molecular memory systems and the directed mutations. *Radiotekhnika*, 1965, vol. 20, no. 6, pp. 1–8 (In Russian). Available at: https://2a008ed5-a62cb3a1a-sites.googlegroups.com/site/msneiman1905/Neiman-1965_Molecul.pdf (accessed 12 March 2021).
- Rebrova I. M., Rebrova O. Y. Synthesized DNA-based data storage devices: the birth of the idea and the first publications. *Studies in the History of Science and Technology*, 2019, vol. 41, no. 4, pp. 666–676 (In Russian). doi:10.31857/S020596060013006-8
- Davis J. Microvenus. *Art Journal*, 1996, vol. 55, no. 1, pp. 70–74. doi:10.2307/777811
- Clelland C. T., Risco V., Bancroft C. Hiding messages in DNA microdots. *Nature*, 1999, vol. 399, pp. 533–534. doi:10.1038/21092
- Bancroft C., Bowler T., Bloom B. Long-term storage of information in DNA. *Science*, 2001, vol. 293, iss. 5536, pp. 1763–1765. doi:10.1126/science.293.5536.1763c
- Church G. M., Gao Y., Kosuri S. Next-generation digital information storage in DNA. *Science*, 2012, vol. 337, pp. 1628. doi:10.1126/science.1226355
- Goldman N., Bertone P., Chen S., Dessimoz C., LeProust E. M., Sipos B., Birney E. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, 2013, vol. 494, pp. 77–79. doi:10.1038/nature11875
- Organick L., Ang S. D., Chen Y., et al. Random access in large-scale DNA data storage. *Nature Biotechnology*, 2018, vol. 36, pp. 242–248. doi:10.1038/nbt.4079
- Shankland S. Startup packs all 16GB of Wikipedia onto DNA strands to demonstrate new storage tech. *Cnet*, 2019. Available at: <https://www.cnet.com/news/startup-packs-all-16gb-wikipedia-onto-dna-strands-demonstrate-new-storage-tech/> (accessed 12 March 2021).
- Extance A. How DNA could store all the world's data. *Nature*, 2016, vol. 537, pp. 22–24. doi:10.1038/537022a
- Service R. DNA could store all of the world's data in one room. *Science*, 2017. doi:10.1126/science.aal0852 Available at: <https://www.sciencemag.org/news/2017/03/dna>

- could-store-all-worlds-data-one-room (accessed 12 March 2021).
23. Erlich Y., Zielinski D. DNA fountain enables a robust and efficient storage architecture. *Science*, 2017, vol. 355, iss. 6328, pp. 950–954. doi:10.1126/science.aa.j2038
 24. Yazdi H. T., Kiah H. M., Garcia-Ruiz E., Ma J., Zhao H., Milenkovic O. DNA-based storage: trends and methods. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2015, vol. 1, no. 3, pp. 230–248. doi:10.1109/TMBMC.2016.2537305
 25. Yiming D., Fajia S., Zhi P., Qi O., Long Q. DNA storage: research landscape and future prospects. *National Science Review*, 2020, vol. 7, iss. 6, pp. 1092–1104. doi:10.1093/nsr/nwaa007
 26. Yim S. S., McBees R. M., Song A. M., Huang Y., Sheth R. U., Wang H. H. Robust direct digital-to-biological data storage in living cells. *Nature Chemical Biology*, 2021, vol. 17, pp. 246–253. doi:10.1038/s41589-020-00711-4
 27. Gilbert E. Synchronization of binary messages. *IRE Transactions on Information Theory*, 1960, vol. 6, no. 4, pp. 470–477. doi:10.1109/TIT.1960.1057587
 28. Levinstein V. I. Binary codes for the correction of deletions insertions, and changes of symbols. *Reports of the USSR Academy of Sciences*, 1965, vol. 163, pp. 845–848.
 29. Sellers F. Bit loss and gain correction code. *IRE Transactions on Information Theory*, 1962, vol. 8, no. 1, pp. 35–38. doi:10.1109/TIT.1962.1057684
 30. Xiong Z. Multiterminal source coding: theory, code design and applications. *Proceedings of the Fifth International Workshop on Signal Design and its Applications in Communications*, IEEE, 2011, pp. 3–3. doi:10.1109/IWSDA.2011.6159430
 31. Luby M. LT codes. *The 43rd Annual IEEE Symposium on Foundations of Computer Science*, IEEE, 2002, pp. 271–280. doi:10.1109/SFCS.2002.1181950
 32. Niedringhaus T. P., Milanova D., Kerby M. B., Snyder M. P., Barron A. E. Landscape of next-generation sequencing technologies. *Analytical Chemistry*, 2011, vol. 83, no. 12, pp. 4327–4341. doi:10.1021/ac2010857
 33. Schwartz J. J., Lee C., Shendure J. Accurate gene synthesis with tag-directed retrieval of sequence-verified DNA molecules. *Nature Methods*, 2012, vol. 9, pp. 913–915. doi:10.1038/nmeth.2137
 34. Yachie N., Sekiyama K., Sugahara J., Ohashi Y., Tomita M. Alignment-based approach for durable data storage into living organisms. *Biotechnol Progress*, 2007, vol. 23, no. 2, pp. 501–505. doi:10.1021/bp060261y
 35. Itaya M., Tsuge K., Koizumi M., Fujita K. Combining two genomes in one cell: stable cloning of the synechocystis PCC6803 genome in the bacillus subtilis 168 genome. *Proceedings of the National Academy of Sciences of the USA*, 2005, vol. 102, no. 44, pp. 15971–15976. doi:10.1073/pnas.0503868102
 36. Kosuri S., Church G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nature Methods*, 2014, vol. 11, no. 5, pp. 499–507. doi:10.1038/nmeth.2918
 37. Heckel R., Mikutis G., Grass R. N. A characterization of the DNA data storage channel. *Science Reports*, 2019, vol. 9, pp. 1–10. doi:10.1038/s41598-019-45832-6
 38. Schmidt T., Beliveau B., Uca Y., Theilmann M., Cruz F. D., Wu C.-T., Shih W. M. Scalable amplification of strand subsets from chip-synthesized oligonucleotide libraries. *Nature Communications*, 2015, vol. 6, pp. 1–11. doi:10.1038/ncomms9634
 39. Lee H. H., Kalthor R., Goela N., Bolot J., Church G. M. Terminator-free template-independent enzymatic DNA synthesis for digital information storage. *Nature Communications*, 2019, vol. 10, pp. 1–12. doi:10.1038/s41467-019-10258-1
 40. Kulski J. K. Next-generation Sequencing — an overview of the history, tools, and “omic” applications. Next generation sequencing — advances, applications and challenges. *IntechOpen*, 2016, pp. 1–15. doi:10.5772/61964
 41. Organick L., Chen Y.-J., Ang S. D., Lopez R., Liu X., Strauss K., Ceze L. Probing the physical limits of reliable DNA data retrieval. *Nature Communications*, 2020, vol. 11, pp. 1–12. doi:10.1038/s41467-020-14319-8
 42. Chandak S., Tatwawadi K., Lau B., Mardia J., Kubit M., Neu J., Griffin P., Wooters M., Weissman T., Ji H. Improved read/write cost tradeoff in DNA-based data storage using LDPC codes. *Proceedings of 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2019, pp. 147–156. doi:10.1109/ALLERTON.2019.8919890
 43. Grass R. N., Heckel R., Puddu M., Paunesco D., Stark W. J. Robust chemical preservation of digital information on DNA in silica with error-correcting codes. *Angewandte Chemie*, 2015, vol. 54, pp. 2552–2555. doi:10.1002/anie.201411378
 44. Anavy L., Vaknin I., Atar O., Amit R., Yakhini Z. Data storage in DNA with fewer synthesis cycles using composite DNA letters. *Nature Biotechnology*, 2019, vol. 37, pp. 1229–1236. doi:10.1038/s41587-019-0240-x
 45. Meiser L. C., Antkowiak P. L., Koch J., Chen W. D., Kohll A. X., Stark W. J., Heckel R., Grass R. N. Reading and writing digital data in DNA. *Nature Protocols*, 2020, vol. 15, pp. 86–101. doi:10.1038/s41596-019-0244-5
 46. Lopez R., Chen Y.-J., Ang S. G., Yekhanin S., Makarychev K., Racz M. Z., Seelig G., Strauss K., Ceze L. DNA assembly for nanopore data storage readout. *Nature Communications*, 2019, vol. 10, pp. 1–9. doi:10.1038/s41467-019-10978-4
 47. Varshamov R. R., Tenengolts G. M. Codes which correct single asymmetric errors. *Automation and Remote Control*, 1965, vol. 26, no. 2, pp. 286–290.
 48. Saowapa K., Kaneko H., Fujiwara E. Systematic binary deletion/insertion error-correcting codes capable of correcting random bit errors. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 2000, vol. E83-A, no. 12, pp. 2699–2705.
 49. Levenshtein V. I. Asymptotically optimum binary codes with correction for losses of one or two adjacent bits. *Problems of Cybernetics*, 1967, vol. 19, pp. 298–304.
 50. Helberg A. S. J., Ferreira H. C. On multiple insertion/deletion correcting codes. *IEEE Transactions on Information Theory*, 2002, vol. 48, no. 1, pp. 305–308. doi:10.1109/18.971760
 51. Swart T. G., Ferreira H. C. A note on double insertion/deletion correcting codes. *IEEE Transactions on Information Theory*, 2003, vol. 49, no. 1, pp. 269–273. doi:10.1109/TIT.2002.806155
 52. Guruswami V., Hastad J. Explicit two-deletion codes with redundancy matching the existential bound. *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, SIAM, 2021, pp. 1–8. doi:10.1137/1.9781611976465.2
 53. Brakensiek J., Guruswami V., Zbarsky S. Efficient low-redundancy codes for correcting multiple deletions. *IEEE Transactions on Information Theory*, 2018, vol. 64, no. 5, pp. 3403–3410. doi:10.1109/TIT.2017.2746566
 54. Guruswami V., Wang C. Deletion codes in the high-noise and high-rate regimes. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 4, pp. 1961–1970. doi:10.1109/TIT.2017.2659765
 55. Sima J., and Bruck J. On optimal k-deletion correcting codes. *IEEE Transactions on Information Theory*, 2021. doi:10.1109/TIT.2020.3028702
 56. Sima J., Gabrys R., and Bruck J. Optimal systematic t-deletion correcting codes. *Proceedings of 2020 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2020, pp. 769–774. doi:10.1109/ISIT44484.2020.9173986
 57. Schulman L. J., Zuckerman D. Asymptotically good codes correcting insertions, deletions, and transpositions. *IEEE Transactions on Information Theory*, 1999, vol. 45, no. 7, pp. 2552–2557. doi:10.1109/18.796406
 58. Kulkarni A. A., Kiyavash N. Nonasymptotic upper bounds for deletion correcting codes. *IEEE Transactions on Information Theory*, 2013, vol. 59, no. 8, pp. 5115–5130. doi:10.1109/TIT.2013.2257917
 59. Bukh B., Guruswami V., and Hastad J. An improved bound on the fraction of correctable deletions. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 1, pp. 93–103. doi:10.1109/TIT.2016.2621044
 60. Guruswami V., and Li R. Efficiently decodable insertion/deletion codes for high-noise and high-rate regimes. *Proceedings of 2016 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2016, pp. 620–624. doi:10.1109/ISIT.2016.7541373
 61. Haeupler B., Shahrabi A. Synchronization strings: codes for insertions and deletions approaching the Singleton bound. *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2017)*, ACM, 2017, pp. 33–46. doi:10.1145/3055399.3055498
 62. Cheng K., Haeupler B., Li X., Shahrabi A., Wu K. Synchronization strings: highly efficient deterministic constructions over small alphabets. *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM, 2019, pp. 2185–2204. doi:10.1137/1.9781611975482.132

63. Cheng K., Jin Z., Li X., and Wu K. Deterministic document exchange protocols, and almost optimal binary codes for edit errors. *Proceedings of 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 2018, pp. 200–211. doi:10.1109/FOCS.2018.00028
64. Haeupler B. Optimal document exchange and new codes for insertions and deletions. *Proceedings 2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 2019, pp. 334–347. doi:10.1109/FOCS.2019.00029
65. Schoeny C., Wachter-Zeh A., Gabrys R., Yaakobi E. Codes correcting a burst of deletions or insertions. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 4, pp. 1971–1985. doi:10.1109/TIT.2017.2661747
66. Saeki T., Nozaki T. An improvement of non-binary code correcting single b-burst of insertions or deletions. *Proceedings of 2018 International Symposium on Information Theory and its Applications (ISITA)*, IEICE, 2018, pp. 6–10. doi:10.23919/ISITA.2018.8664217
67. Gabrys R., Yaakobi E., Milenkovic O. Codes in the damerau distance for deletion and adjacent transposition correction. *IEEE Transactions on Information Theory*, 2018, vol. 64, no. 4, pp. 2550–2570. doi:10.1109/TIT.2017.2778143
68. Lenz A., Polyanskii N. Optimal codes correcting a burst of deletions of variable length. *2020 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2020, pp. 757–762. doi:10.1109/ISIT44484.2020.9174288
69. Sima J., Gabrys R., and Bruck J. Syndrome compression for optimal redundancy codes. *Proceedings of 2020 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2020, pp. 751–756. doi:10.1109/ISIT44484.2020.9174009
70. Levenshtein V. I. Binary codes capable of correcting spurious insertions and deletions of ones. *Problems on Information Transmission*, 1965, vol. 1, no. 1, pp. 8–17.
71. Dolecek L., Anantharam V. Repetition error correcting sets: explicit constructions and prefixing methods. *SIAM Journal on Discrete Mathematics*, 2010, vol. 23, no. 4, pp. 2120–2146. doi:10.5555/1958171.1958195
72. MahdaviFar H., Vardy A. Asymptotically optimal sticky-insertion correcting codes with efficient encoding and decoding. *Proceedings of 2017 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 2688–2692. doi:10.1109/ISIT.2017.8007016
73. Jain S., Farnoud F., Schwartz M., Bruck J. Duplication-correcting codes for data storage in the DNA of living organisms. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 8, pp. 4996–5010. doi:10.1109/ISIT.2016.7541455
74. Kovacevic M. Codes correcting all patterns of tandem-duplication errors of maximum length 3. Available at: <https://arxiv.org/pdf/1911.06561.pdf> (accessed 12 March 2021).
75. Nazirkhanova K., Medova L., Kruglik S., and Frolov A. Codes correcting bounded length tandem duplication. *Proceedings of 2020 International Symposium on Information Theory and its Applications (ISITA)*, IEICE, 2020, pp. 299–303. doi:10.34385/proc.65.B06-6
76. Heckel R., Shomorony I., Ramchandran K., Tse D. N. C. Fundamental limits of DNA storage systems. *2017 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 3130–3134. doi:10.1109/ISIT.2017.8007106
77. Lenz A., Siegel P. H., Wachter-Zeh A., Yaakobi E. An upper bound on the capacity of the DNA storage channel. *2019 IEEE Information Theory Workshop (ITW)*, IEEE, 2019, pp. 1–5. doi:10.1109/ITW44776.2019.8989388
78. Lenz A., Siegel P. H., Wachter-Zeh A., Yaakobi E. Achieving the capacity of the DNA storage channel. *Proceedings of ICASSP 2020 — 2020 IEEE International Conference on Acoustics, Speech and Signal*, IEEE, 2020, pp. 8846–8850. doi:10.1109/ICASSP40776.2020.9053049
79. Shin S., Heckel R., Shomorony I. Capacity of the erasure shuffling channel. *Proceedings of ICASSP 2020 — 2020 IEEE International Conference on Acoustics, Speech and Signal*, IEEE, 2020, pp. 8841–8845. doi:10.1109/ICASSP40776.2020.9053486
80. Shomorony I., Heckel R. Capacity results for the noisy shuffling channel. *Proceedings of 2019 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2019, pp. 762–766. doi:10.1109/ISIT.2019.8849789
81. Mitzenmacher M. A survey of results for deletion channels and related synchronization channels. *Probability Surveys*, 2009, vol. 6, pp. 1–33. doi:10.1214/08-PS141
82. Lenz A., Siegel P. H., Wachter-Zeh A., Yaakobi E. Coding over sets for DNA storage. *IEEE Transactions on Information Theory*, 2020, vol. 66, no. 4, pp. 2331–2351. doi:10.1109/TIT.2019.2961265
83. Sima J., Raviv N., and Bruck J. On coding over sliced information. *IEEE Transactions on Information Theory*, 2021. doi:10.1109/TIT.2021.3063709
84. Song W., Cai K., Schouhamer Immink K. A. Sequence-subset distance and coding for error control in DNA-based data storage. *IEEE Transactions on Information Theory*, 2020, vol. 66, no. 10, pp. 6048–6065. doi:10.1109/TIT.2020.3002611

Подавление коррелированных помех адаптивными режекторными фильтрами при модуляции периода повторения импульсов

С. И. Зиятдинов^а, доктор техн. наук, профессор, orcid.org/0000-0002-8109-5456, ziat.53@mail.ru

Л. А. Осипов^а, доктор техн. наук, профессор, orcid.org/0000-0002-8009-5706

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

Введение: рассматривается вопрос подавления коррелированных помех адаптивными комплексными режекторными фильтрами различных порядков. Для устранения зависимости коэффициента передачи полезного сигнала от его частоты использована модуляция периода повторения. **Цель:** исследование влияния модуляции периода повторения импульсов на коэффициент подавления коррелированных помех. **Методы:** оптимизация параметров режекторных фильтров осуществлялась на основе критерия минимума средней дисперсии коррелированных помех на выходе фильтров за время модуляции периода повторения. **Результаты:** получены выражения для дисперсии коррелированных помех на выходе комплексных адаптивных фильтров различных порядков при модуляции периода повторения. Приведены соотношения для нахождения оптимальных значений частоты настройки и коэффициентов режекторных фильтров, минимизирующих уровень коррелированных помех на их выходе. Получены выражения для коэффициентов подавления коррелированных помех режекторными фильтрами в условиях модуляции периода повторения импульсов. Представлены графики зависимостей коэффициента подавления коррелированных помех от относительной величины девиации периода повторения зондирующих сигналов для различных значений ширины спектральной плотности коррелированных помех при оптимальных и неоптимальных значениях частоты настройки и коэффициентов режекторных фильтров. Показано, что использование модуляции периода повторения зондирующих импульсов приводит к снижению коэффициента подавления коррелированных помех. При этом адаптация весовых коэффициентов для принятых моделей режекторных фильтров и коррелированных помех обеспечивает увеличение коэффициента подавления. **Практическая значимость:** полученные результаты позволяют при разработке и исследовании систем подавления коррелированных помех с учетом допустимых потерь коэффициента подавления обоснованно выбрать величину девиации периода повторения входных импульсов для решения задачи устранения эффекта «слепых» частот.

Ключевые слова – коррелированные помехи, режекторные фильтры, адаптация, модуляция периода повторения.

Для цитирования: Зиятдинов С. И., Осипов Л. А. Подавление коррелированных помех адаптивными режекторными фильтрами при модуляции периода повторения импульсов. *Информационно-управляющие системы*, 2021, № 3, с. 53–60. doi:10.31799/1684-8853-2021-3-53-60

For citation: Ziatdinov S. I., Osipov L. A. Suppression of correlated interference by adaptive notch filters under pulse repetition period modulation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 53–60 (In Russian). doi:10.31799/1684-8853-2021-3-53-60

Введение

При разработке систем оценки параметров и обнаружения сигналов, систем автоматического измерения угловых координат, дальности и скорости движения объектов для подавления коррелированных помех, таких как отражения от земной поверхности либо метеообразований, используются режекторные фильтры различных порядков [1–5]. В случаях неизвестных характеристик помех применяются адаптивные режекторные фильтры, характеристики которых изменяются в соответствии с частотными свойствами коррелированных помех [6–10].

На практике в зависимости от тактических задач устройства подавления коррелированных помех выполняются либо в виде режекторных фильтров различных порядков, либо в виде полосовых фильтров [6, 7, 11, 12]. В обоих случаях су-

ществленным недостатком используемых средств подавления коррелированных помех и выделения полезного сигнала является наличие эффекта «слепых» частот, когда при доплеровских частотах обрабатываемых сигналов, кратных частоте повторения импульсов, наблюдается уменьшение коэффициента передачи полезного сигнала [2].

Для устранения эффекта «слепых» частот используют, как правило, изменение периода следования (частоты повторения) импульсов [1, 2]. В системах подавления коррелированных помех, выполненных на базе полосовых фильтров, осуществляется модуляция частоты повторения импульсов. При этом, например, на отрезке времени T импульсы следуют с частотой повторения $F_{п1}$, а на следующем отрезке времени T — с частотой $F_{п2}$ и т. д.

В случае использования для подавления коррелированных помех режекторных фильтров на базе устройств череспериодного вычитания при-

меняется модуляция периода повторения в виде, например, двух чередующихся периодов $T_{п1}$ и $T_{п2}$.

В работах [7, 13–15] показано, что коэффициент подавления коррелированных помех в значительной степени зависит от величины периода повторения зондирующих импульсов. При уменьшении периода повторения происходит увеличение коэффициента подавления коррелированных помех. С увеличением периода повторения резко проявляется эффект «слепых» частот. Значение периода повторения в целом определяет тактические характеристики всего комплекса обработки сигналов. Однако в вышеперечисленных работах не рассматривается работа режекторных фильтров в условиях модуляции периода повторения импульсов.

В статье рассматривается адаптивная система подавления коррелированных помех на базе адаптивных комплексных режекторных фильтров различных порядков. Для устранения эффекта «слепых» частот использована модуляция периода повторения импульсов. Исследование влияния модуляции периода повторения импульсов на коэффициент подавления коррелированных помех адаптивными режекторными фильтрами различных порядков составляет основное содержание статьи.

Комплексный адаптивный режекторный фильтр первого порядка

Алгоритм работы рассматриваемого режекторного фильтра определяется следующим соотношением [16–18]:

$$z_{\text{ВЫХ}}(t) = z_{\text{ВХ}}(t) + az_{\text{ВХ}}(t - T_{\text{П}})e^{j\omega_0 T_{\text{П}}}, \quad (1)$$

где $z_{\text{ВХ}}(t)$ и $z_{\text{ВЫХ}}(t)$ — комплексные значения входной и выходной коррелированной помехи; $t = iT_{\text{П}}$, $T_{\text{П}}$ — период следования отсчетов значений входной и выходной коррелированной помехи, $i = 0, 1, 2, \dots$; a , $\omega_0 T_{\text{П}}$ — модуль и аргумент комплексного весового коэффициента, ω_0 — частота настройки зоны режекции фильтра.

Если в формуле (1) принять $a = -1$ и $\omega_0 = 0$, то получим известный алгоритм работы устройства однократного череспериодного вычитания [2, 14].

При заданном периоде повторения $T_{\text{П}}$ определим оптимальные значения модуля весового коэффициента a и частоты настройки ω_0 режекторного фильтра, минимизирующие уровень помехи на его выходе. Для этого найдем корреляционную функцию помехи на выходе режекторного фильтра

$$B_{\text{ВЫХ}}(\tau) = [z_{\text{ВХ}}(t) + az_{\text{ВХ}}(t - T_{\text{П}})e^{j\omega_0 T_{\text{П}}}] \times [z_{\text{ВХ}}^*(t - \tau) + az_{\text{ВХ}}^*(t - T_{\text{П}} - \tau)e^{-j\omega_0 T_{\text{П}}}], \quad (2)$$

В данном выражении черта сверху означает статистическое усреднение; значок * — комплексную сопряженность.

С учетом выражения (1) соотношение для корреляционной функции (2) принимает вид

$$\begin{aligned} B_{\text{ВЫХ}}(\tau) &= \overline{z_{\text{ВХ}}(t)z_{\text{ВХ}}^*(t - \tau)} + \\ &+ \overline{az_{\text{ВХ}}(t - T_{\text{П}})z_{\text{ВХ}}^*(t - \tau)e^{j\omega_0 T_{\text{П}}}} + \\ &+ \overline{az_{\text{ВХ}}(t)z_{\text{ВХ}}^*(t - T_{\text{П}} - \tau)e^{-j\omega_0 T_{\text{П}}}} + \\ &+ \overline{a^2 z_{\text{ВХ}}(t - T_{\text{П}})z_{\text{ВХ}}^*(t - T_{\text{П}} - \tau)} = \\ &= (1 + a^2)B_{\text{ВХ}}(\tau) - aB_{\text{ВХ}}(\tau - T_{\text{П}})e^{j\omega_0 T_{\text{П}}} - \\ &- aB_{\text{ВХ}}(\tau + T_{\text{П}})e^{-j\omega_0 T_{\text{П}}}, \end{aligned} \quad (3)$$

где $B_{\text{ВХ}}(\tau)$ — корреляционная функция комплексной входной помехи $z_{\text{ВХ}}(t)$.

Положив в (3) $\tau = 0$, находим дисперсию помехи на выходе рассматриваемого режекторного фильтра

$$\begin{aligned} \sigma_{\text{ВЫХ}}^2 &= (1 + a^2)B_{\text{ВХ}}(0) + \\ &+ a[B_{\text{ВХ}}(-T_{\text{П}})e^{j\omega_0 T_{\text{П}}} + B_{\text{ВХ}}(T_{\text{П}})e^{-j\omega_0 T_{\text{П}}}], \end{aligned} \quad (4)$$

Согласно работе [9], корреляционная функция входной стационарной комплексной помехи $z_{\text{ВХ}}(t)$ имеет вид

$$B_{\text{ВХ}}(\tau) = \sigma_{\text{П}}^2 \rho_{\text{П}}(\tau)(\cos \omega_{\text{П}} \tau + j \sin \omega_{\text{П}} \tau), \quad (5)$$

где $\sigma_{\text{П}}^2$, $\rho_{\text{П}}(\tau)$ и $\omega_{\text{П}}$ — дисперсия, модуль коэффициента корреляции и средняя частота спектральной плотности входной комплексной помехи, которая при симметричной спектральной плотности соответствует положению ее максимума на оси частот.

После подстановки (5) в (4) и проведения несложных преобразований получим

$$\begin{aligned} \sigma_{\text{ВЫХ}}^2 &= \sigma_{\text{П}}^2 \{ (1 + a^2) + a\rho_{\text{П}}(T_{\text{П}}) \times \\ &\times [e^{j(\omega_0 - \omega_{\text{П}})T_{\text{П}}} + e^{-j(\omega_0 - \omega_{\text{П}})T_{\text{П}}}] \} = \\ &= \sigma_{\text{П}}^2 [1 + a^2 + 2a\rho_{\text{П}}(T_{\text{П}})\cos \Delta\omega T_{\text{П}}]. \end{aligned} \quad (6)$$

В данном соотношении $\Delta\omega_0 = \omega_0 - \omega_{\text{П}}$ — расстройка по частоте режекторного фильтра и средней частоты спектральной плотности помехи на входе.

В работе [13] дан сравнительный анализ двух алгоритмов для адаптации режекторного фильтра — точного алгоритма, когда учитываются корреляционные свойства сигнала и ищется максимум отношения сигнал/помеха на выходе режекторного фильтра, и приближенного алгоритма, при котором сигнал не учитывается, а ищется

ся минимум дисперсии коррелированной помехи на выходе режекторного фильтра.

В результате исследований показано, что с учетом инвариантности параметров режекторного фильтра к корреляционным характеристикам сигнала и в предположении равномерного распределения частоты сигнала приближенный алгоритм оптимизации по критерию минимума дисперсии помехи на выходе режекторного фильтра в случае относительной ширины спектральной плотности помехи $\Delta f_{\Pi} T_{\Pi} < 0,2 \dots 0,3$ практически не уступает точному алгоритму. При этом упрощенный алгоритм оптимизации значительно проще в реализации.

С учетом этого обстоятельства в качестве критерия оптимизации параметров рассматриваемого режекторного фильтра примем минимум дисперсии помехи на его выходе, т. е. $\sigma_{\text{ВЫХ}}^2 \rightarrow \min$.

Найдем значения его частоты настройки ω_0 и весового коэффициента a , минимизирующие дисперсию помехи на его выходе. Для этого необходимо решить следующую систему уравнений:

$$\begin{cases} \frac{\partial \sigma_{\text{ВЫХ}}^2}{\partial \Delta \omega} = 0; \\ \frac{\partial \sigma_{\text{ВЫХ}}^2}{\partial a} = 0. \end{cases} \quad (7)$$

Дифференцируя соотношение (6) по параметру $\Delta \omega$, получим

$$\frac{\partial \sigma_{\text{ВЫХ}}^2}{\partial \Delta \omega} = 2a T_{\Pi} \rho_{\Pi}(T_{\Pi}) \sin \Delta \omega T_{\Pi}. \quad (8)$$

Будем считать, что в составе радиотехнического комплекса имеются средства для измерения частоты коррелированной помехи, значение которой используется для адаптации режекторного фильтра.

Приравняем производную (8) к нулю. В результате найдем оптимальное значение частоты настройки режекторного фильтра $\omega_{0\text{opt}} = \omega_{\Pi}$ ($\Delta \omega = 0$).

Второе уравнение системы (7) имеет вид

$$\frac{\partial \sigma_{\text{ВЫХ}}^2}{\partial a} = 2a + 2\rho_{\Pi}(T_{\Pi}) \cos \Delta \omega T_{\Pi} = 0. \quad (9)$$

При $\Delta \omega = 0$ из формулы (9) получим оптимальное значение модуля весового коэффициента рассматриваемого режекторного фильтра $a_{\text{opt}} = -\rho_{\Pi}(T_{\Pi})$.

С учетом полученных результатов минимальное значение дисперсии помехи на выходе режекторного фильтра записывается следующим образом:

$$\sigma_{\text{ВЫХ}}^2 \min = \sigma_{\Pi}^2 [1 - \rho_{\Pi}^2(T_{\Pi})].$$

В случае использования для устранения эффекта «слепых» частот двух чередующихся периодов повторения $T_{\Pi 1}$ и $T_{\Pi 2}$ выражение (6) для дисперсии помехи на выходе принимает вид

$$\begin{aligned} \sigma_{\text{ВЫХ}1}^2 &= \sigma_{\Pi}^2 [1 + a_1^2 + 2a_1 \rho_{\Pi}(T_{\Pi 1}) \cos \Delta \omega_1 T_{\Pi 1}]; \\ \sigma_{\text{ВЫХ}2}^2 &= \sigma_{\Pi}^2 [1 + a_2^2 + 2a_2 \rho_{\Pi}(T_{\Pi 2}) \cos \Delta \omega_2 T_{\Pi 2}]. \end{aligned}$$

Среднее значение дисперсии помехи на выходе режекторного фильтра при использовании двух чередующихся периодов повторения импульсов составит

$$\sigma_{\text{ср}}^2 = (\sigma_{\text{ВЫХ}1}^2 + \sigma_{\text{ВЫХ}2}^2) / 2.$$

В качестве критерия оптимизации параметров режекторного фильтра примем минимум среднего значения дисперсии помехи на его выходе за период модуляции, т. е. $\sigma_{\text{ср}}^2 \rightarrow \min$.

При оптимальных значениях параметров $\Delta \omega_1 = \Delta \omega_2 = 0$, $a_{1\text{opt}} = -\rho_{\Pi}(T_{\Pi 1})$ и $a_{2\text{opt}} = -\rho_{\Pi}(T_{\Pi 2})$ минимальное значение среднего значения дисперсии помехи на выходе режекторного фильтра имеет вид

$$\sigma_{\text{ср min}}^2 = \sigma_{\Pi}^2 \{1 - 0,5[\rho_{\Pi}^2(T_{\Pi 1}) + \rho_{\Pi}^2(T_{\Pi 2})]\}.$$

Коэффициент подавления коррелированной помехи, приведенный ко входу рассматриваемого режекторного фильтра, будем оценивать следующим соотношением:

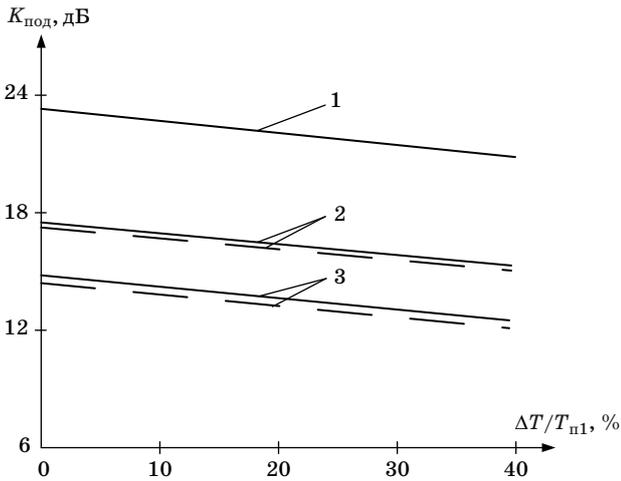
$$K_{\text{под}} = \frac{4\sigma_{\Pi}^2}{\sigma_{\text{ср min}}^2} = 4\{1 - 0,5[\rho_{\Pi}^2(T_{\Pi 1}) + \rho_{\Pi}^2(T_{\Pi 2})]\}^{-1}.$$

Для конкретизации полученных результатов примем гауссову аппроксимацию модуля коэффициента корреляции коррелированной помехи на входе [19, 20]

$$\rho_{\Pi}(\tau) = \exp(-\Delta \omega_{\Pi}^2 \tau^2 / 4),$$

где $\Delta \omega_{\Pi}$ — ширина спектральной плотности помехи.

Графики зависимости коэффициента подавления коррелированной помехи от относительной величины девиации периода повторения $\Delta T / T_{\Pi 1} = (T_{\Pi 2} - T_{\Pi 1}) / T_{\Pi 1}$ для различных значений ширины спектральной плотности коррелированной помехи $\Delta f_{\Pi} = \Delta \omega_{\Pi} / 2\pi$ при оптимальных значениях параметров $\Delta \omega_1 = \Delta \omega_2 = 0$, $a_{1\text{opt}} = -\rho_{\Pi}(T_{\Pi 1})$ и $a_{2\text{opt}} = -\rho_{\Pi}(T_{\Pi 2})$ показаны на рис. 1 сплошной линией. Здесь же пунктирной линией представлены аналогичные зависимости для случая отсутствия адаптации весового коэффициента, который принимался равным $a_1 = a_2 = -1$. Расчеты проводились для $T_{\Pi 1} = 1$ мс.



■ **Рис. 1.** Коэффициент подавления коррелированной помехи режекторным фильтром первого порядка: 1 — $\Delta f_{\text{п}} = 30$ Гц; 2 — $\Delta f_{\text{п}} = 60$ Гц; 3 — $\Delta f_{\text{п}} = 90$ Гц

■ **Fig. 1.** The coefficient of suppression of correlated interference by a first-order filter: 1 — $\Delta f_{\text{п}} = 30$ Hz; 2 — $\Delta f_{\text{п}} = 60$ Hz; 3 — $\Delta f_{\text{п}} = 90$ Hz

Из полученных результатов следует, что использование для устранения эффекта «слепых» частот модуляции периода повторения зондирующих импульсов приводит к снижению коэффициента подавления коррелированных помех. Так, при реализации системы подавления коррелированных помех в виде комплексного адаптивного режекторного фильтра первого порядка изменение относительной девиации периода повторения в пределах от 0 до 40% в случае точной настройки фильтра на среднюю частоту спектральной плотности коррелированной помехи приводит к уменьшению коэффициента подавления на 1,8 дБ. Использование адаптации весового коэффициента для принятой модели режекторного фильтра при изменении относительной ширины спектральной плотности коррелированной помехи в пределах $\Delta\omega_{\text{п}}T_{\text{п1}} = 0,19 \div 0,57$ ($\Delta f_{\text{п}} = 30 \div 90$ Гц, $T_{\text{п1}} = 10^{-3}$ с) дает увеличение коэффициента подавления на 0,29 дБ.

В тех случаях, когда режекторный фильтр первого порядка не обеспечивает требуемого подавления коррелированной помехи, используются режекторные фильтры более высоких порядков. Рассмотрим работу адаптивного режекторного фильтра второго порядка в условиях модуляции периода повторения импульсов.

Комплексный адаптивный режекторный фильтр второго порядка

Пусть для устранения эффекта «слепых» частот аналогично предыдущему случаю использу-

ются два чередующихся периода повторения импульсов $T_{\text{п1}}$ и $T_{\text{п2}}$, порядок следования которых показан на рис. 2.

Помеху на выходе адаптивного режекторного фильтра второго порядка можно записать в виде

$$z_{\text{ВЫХ}}(t_3) = z_{\text{ВХ}}(t_3) + a_1 z_{\text{ВХ}}(t_2) e^{j\omega_1 T_{\text{п2}}} + a_2 z_{\text{ВХ}}(t_1) e^{j\omega_2 (T_{\text{п1}} + T_{\text{п2}})};$$

$$z_{\text{ВЫХ}}(t_4) = z_{\text{ВХ}}(t_4) + a_3 z_{\text{ВХ}}(t_3) e^{j\omega_1 T_{\text{п1}}} + a_4 z_{\text{ВХ}}(t_2) e^{j\omega_2 (T_{\text{п1}} + T_{\text{п2}})};$$

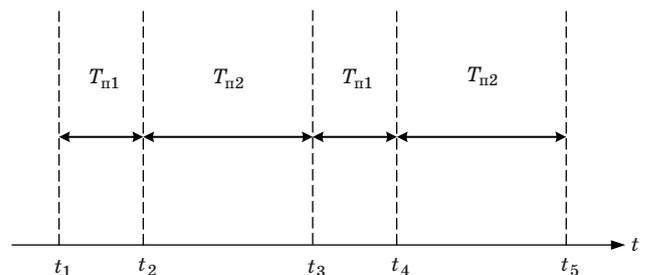
и т. д. В данных выражениях a_1, a_2, a_3, a_4 — модули весовых коэффициентов, а $\omega_1 T_{\text{п1}}, \omega_1 T_{\text{п2}}, \omega_2 (T_{\text{п1}} + T_{\text{п2}})$ — их аргументы. При $a_1 = a_3 = -2, a_2 = a_4 = 1$ и $\omega_1 = \omega_2 = 0$ данное соотношение определяет алгоритм работы известного устройства двукратного череспериодного вычитания [2, 14].

Дисперсия помехи на выходе комплексного адаптивного режекторного фильтра второго порядка в моменты времени t_3 и t_4 записывается следующим образом:

$$\sigma_{\text{ВЫХ1}}^2 = \sigma_{\text{п}}^2 \{1 + a_1^2 + a_2^2 + 2a_1 \rho_{\text{п}}(T_{\text{п2}}) \times \cos[(\omega_1 - \omega_{\text{п}})T_{\text{п2}}] + 2a_2 \rho_{\text{п}}(T_{\text{п1}} + T_{\text{п2}}) \times \cos[(\omega_2 - \omega_{\text{п}})(T_{\text{п1}} + T_{\text{п2}})] + 2a_1 a_2 \rho_{\text{п}}(T_{\text{п1}}) \times \cos[\omega_2 (T_{\text{п1}} + T_{\text{п2}}) - \omega_1 T_{\text{п2}} - \omega_{\text{п}} T_{\text{п1}}]\};$$

$$\sigma_{\text{ВЫХ2}}^2 = \sigma_{\text{п}}^2 \{1 + a_3^2 + a_4^2 + 2a_3 \rho_{\text{п}}(T_{\text{п1}}) \times \cos[(\omega_1 - \omega_{\text{п}})T_{\text{п1}}] + 2a_4 \rho_{\text{п}}(T_{\text{п1}} + T_{\text{п2}}) \times \cos[(\omega_2 - \omega_{\text{п}})(T_{\text{п1}} + T_{\text{п2}})] + 2a_3 a_4 \rho_{\text{п}}(T_{\text{п2}}) \times \cos[\omega_2 (T_{\text{п1}} + T_{\text{п2}}) - \omega_1 T_{\text{п2}} - \omega_{\text{п}} T_{\text{п1}}]\}.$$

Оптимальные значения параметров ω_1, ω_2 и a_1, a_2 , минимизирующих дисперсии помехи на выходе режекторного фильтра $\sigma_{\text{ВЫХ1}}^2$ и $\sigma_{\text{ВЫХ2}}^2$, находятся из решения системы уравнений



■ **Рис. 2.** Модуляция периода повторения импульсов
 ■ **Fig. 2.** Modulation of the pulse repetition period

$$\begin{cases}
 \frac{\partial \sigma_1^2}{\partial \omega_1} = -a_1 T_{п2} \rho_{п}(T_{п2}) \sin[(\omega_1 - \omega_{п})T_{п2}] + \\
 + a_1 a_2 T_{п2} \rho_{п}(T_{п1}) \sin[\omega_2(T_{п1} + T_{п2}) - \\
 - \omega_1 T_{п2} - \omega_{п} T_{п1}] = 0; \\
 \frac{\partial \sigma_2^2}{\partial \omega_2} = -a_2 (T_{п1} + T_{п2}) \rho_{п}(T_{п1} + T_{п2}) \times \\
 \times \sin[(\omega_2 - \omega_{п})(T_{п1} + T_{п2})] - \\
 - a_1 a_2 (T_{п1} + T_{п2}) \rho_{п}(T_{п1}) \times \\
 \times \sin[\omega_2(T_{п1} + T_{п2}) - \omega_1 T_{п2} - \omega_{п} T_{п1}] = 0; \\
 \frac{\partial \sigma_1^2}{\partial a_1} = a_1 + \rho_{п}(T_{п2}) \cos[(\omega_1 - \omega_{п})T_{п2}] + \\
 + a_2 \rho_{п}(T_{п1}) \cos[\omega_2(T_{п1} + T_{п2}) - \omega_1 T_{п2} - \omega_{п} T_{п1}] = 0; \\
 \frac{\partial \sigma_2^2}{\partial a_2} = a_2 + \rho_{п}(T_{п1} + T_{п2}) \times \\
 \times \cos[(\omega_2 - \omega_{п})(T_{п1} + T_{п2})] + a_1 \rho_{п}(T_{п1}) \times \\
 \times \cos[\omega_2(T_{п1} + T_{п2}) - \omega_1 T_{п2} - \omega_{п} T_{п1}] = 0.
 \end{cases} \quad (10)$$

Решению первого и второго уравнений системы (10) соответствуют оптимальные значения параметров $\omega_1 = \omega_2 = \omega_{п}$. Оптимальные значения коэффициентов a_1 и a_2 , удовлетворяющие третьему и четвертому уравнению системы (10) при $\omega_1 = \omega_2 = \omega_{п}$, находятся из решения системы уравнений

$$\begin{cases}
 a_1 + a_2 \rho_{п}(T_{п1}) + \rho_{п}(T_{п2}) = 0; \\
 a_2 + \rho_{п}(T_{п1} + T_{п2}) + a_1 \rho_{п}(T_{п1}) = 0.
 \end{cases}$$

Из данной системы уравнений следует, что

$$\begin{aligned}
 a_{1 \text{ opt}} &= \frac{\rho_{п}(T_{п1} + T_{п2}) \rho_{п}(T_{п1}) - \rho_{п}(T_{п2})}{1 - \rho_{п}(T_{п1})^2}; \\
 a_{2 \text{ opt}} &= -\rho_{п}(T_{п1} + T_{п2}) - a_1 \rho_{п}(T_{п1}).
 \end{aligned}$$

Аналогичным образом нетрудно получить следующие выражения для оптимальных значений коэффициентов a_3 и a_4 :

$$\begin{aligned}
 a_{3 \text{ opt}} &= \frac{\rho_{п}(T_{п1} + T_{п2}) \rho_{п}(T_{п2}) - \rho_{п}(T_{п1})}{1 - \rho_{п}(T_{п2})^2}; \\
 a_{4 \text{ opt}} &= -\rho_{п}(T_{п1} + T_{п2}) - a_3 \rho_{п}(T_{п2}).
 \end{aligned}$$

Как и ранее, в качестве критерия оптимизации параметров режекторного фильтра примем минимум среднего значения дисперсии помехи на его выходе за период модуляции, т. е. $\sigma_{ср}^2 \rightarrow \min$.

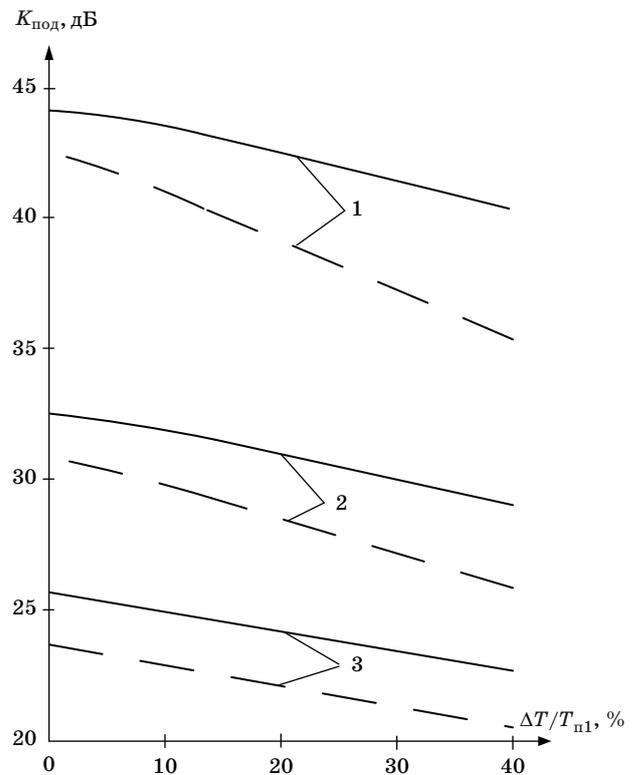
В результате минимальное среднее значение дисперсии помехи на выходе режекторного фильтра при использовании двух чередующихся периодов повторения зондирующих импульсов составит

$$\sigma_{ср \text{ min}}^2 = (\sigma_{\text{вых}1 \text{ min}}^2 + \sigma_{\text{вых}2 \text{ min}}^2) / 2.$$

Оптимальные значения частоты настройки режекторного фильтра и коэффициентов обеспечивают максимальный коэффициент подавления коррелированной помехи, который с учетом приведения ко входу режекторного фильтра рассчитывается с помощью соотношения

$$K_{\text{под}} = 16 \sigma_{п}^2 / \sigma_{ср \text{ min}}^2.$$

Графики зависимости коэффициента подавления коррелированной помехи от относительной величины девиации периода повторения $\Delta T / T_{п1} = (T_{п2} - T_{п1}) / T_{п1}$ для различных значений ширины спектральной плотности помехи $\Delta f_{п} = \Delta \omega_{п} / 2\pi$ при оптимальных значениях частоты настройки и коэффициентов показаны на рис. 3 сплошной линией. Здесь же пунктирной линией представлены аналогичные зависимости для случая отсутствия адаптации коэффициентов, которые принимались равными $a_1 = a_3 = -2$, $a_2 = a_4 = 1$ при $\omega_1 = \omega_2 = \omega_{п}$. Расчеты проводились для $T_{п1} = 1$ мс, $\rho_{п}(\tau) = \exp(-\Delta \omega_{п}^2 \tau^2 / 4)$.



■ **Рис. 3.** Коэффициент подавления коррелированной помехи режекторным фильтром второго порядка: 1 — $\Delta f_{п} = 30$ Гц; 2 — $\Delta f_{п} = 60$ Гц; 3 — $\Delta f_{п} = 90$ Гц
 ■ **Fig. 3.** The coefficient of suppression of correlated interference by a second-order filter: 1 — $\Delta f_{п} = 30$ Hz; 2 — $\Delta f_{п} = 60$ Hz; 3 — $\Delta f_{п} = 90$ Hz

Исследования показали, что для системы подавления коррелированных помех в виде комплексного адаптивного режекторного фильтра второго порядка использование для устранения эффекта «слепых» частот модуляции периода повторения зондирующих импульсов приводит к уменьшению коэффициента подавления. В случае изменения относительной девиации периода повторения в пределах от 0 до 40% наблюдается уменьшение коэффициента подавления при относительной ширине спектральной плотности коррелированной помехи $\Delta\omega_{\text{п}}T_{\text{п1}} = 0,19$ ($\Delta f_{\text{п}} = 30$ Гц, $T_{\text{п1}} = 10^{-3}$ с) на 3,2 дБ; при $\Delta\omega_{\text{п}}T_{\text{п1}} = 0,38$ — на 3 дБ; при $\Delta\omega_{\text{п}}T_{\text{п1}} = 0,57$ — на 2,8 дБ.

Адаптация весовых коэффициентов режекторного фильтра при точной настройке фильтра на среднюю частоту спектральной плотности коррелированной помехи в случае изменения относительной ширины спектральной плотности помехи в пределах $\Delta\omega_{\text{п}}T_{\text{п1}} = 0,57 \div 0,19$ при относительной девиации периода повторения 40% позволяет увеличить коэффициент подавления коррелированной помехи на $2,8 \div 5,8$ дБ.

Заключение

Использование модуляции периода повторения импульсов приводит к уменьшению значения коэффициента подавления коррелированных помех. Для системы подавления коррелированных помех в виде комплексного режекторного адаптивного фильтра первого порядка изменение относительной девиации периода повторения в пределах от 0 до 40% дает уменьшение коэффициента подавления на 1,8 дБ. В то же время для комплексного режекторного адаптивного фильтра второго порядка данные изменения составляют 4,4 дБ. Адаптация весовых коэффициентов режекторных фильтров к частотным свойствам коррелированных помех является практически сложной задачей. Для режекторного фильтра первого порядка при соотношении $\Delta\omega_{\text{п}}T_{\text{п1}} = 0,57$ адаптация весовых коэффициентов приводит к увеличению коэффициента подавления на 0,29 дБ, для фильтра второго порядка увеличение коэффициента подавления равняется 2,8 дБ.

Литература

1. *Radar Handboor*. Editor by M. I. Skolnik. 3rd ed. New York, McGraw-Hill, 2008. 1352 p.
2. Бакулев П. А., Стенин В. М. *Методы и устройства селекции движущихся целей*. М., Радио и связь, 1986. 286 с.
3. Попов Д. И. Автокомпенсация доплеровской фазы пассивных помех. *Цифровая обработка сигналов*, 2009, № 2, с. 30–33.
4. Ястребов А. В. Адаптивный фильтр на основе методов степенных рядов. *Изв. вузов. Радиоэлектроника*, 2017, № 1, с. 5–11.
5. Журавский В. Н., Силин С. И. Синтез адаптивного обнаружителя в спектральной области. *Радиотехника и электроника*, 2018, т. 63, № 12, с. 1277–1284.
6. Попов Д. И. Адаптивное подавление пассивных помех. *Цифровая обработка сигналов*, 2014, № 4, с. 32–37.
7. Попов Д. И. Адаптация нерекурсивных режекторных фильтров. *Изв. вузов. Радиоэлектроника*, 2009, т. 52, № 4, с. 46–55.
8. Зиятдинов С. И. Влияние рассогласования параметров квадратурных каналов на работу адаптивной системы селекции движущихся целей. *Изв. вузов. Радиоэлектроника*, 2020, т. 23, № 3, с. 6–9.
9. Мичурин С. В. Влияние ошибок настройки квадратурных каналов когерентной системы на оценку параметров комплексного сигнала. *Изв. вузов. Приборостроение*, 2020, т. 63, № 6, с. 495–500.
10. *Микропроцессорные системы автоматического управления*. Под общ. ред. В. А. Бесекерского. Л., Машиностроение, 1988. 355 с.
11. Kholodnyak D., Turgaliev A., Baskakova A. A. Method to design lumped-element tunable bandpass filters with constant absolute bandwidth. *Proc. of 44th Eur. Microwave Conf.*, Roma, 6–9 Oct, 2014, Nijkerk, EuMA, 2014, pp. 335–338.
12. Баскакова А. Э., Тургалиев В. М., Холодняк Д. В. Перестраиваемый полосно-пропускной фильтр на элементах с сосредоточенными параметрами с независимым непрерывным управлением. *Изв. вузов. Радиоэлектроника*, 2016, № 3, с. 25–32.
13. Попов Д. И. Адаптивные режекторные фильтры с комплексными весовыми коэффициентами. *Вестник Концерна ПВО «Алмаз — Антей»*, 2015, № 2, с. 21–26.
14. Бакулев П. А. *Радиолокационные системы*. М., Радиотехника, 2004. 319 с.
15. Кокоскин А. В., Коротков В. А., Коротков К. В., Новичихин Е. П. Методы улучшения различимости объектов при наличии гидрометеоров. *Журнал радиоэлектроники*, 2015, № 10, с. 1–14. <http://jre.cplire.ru/jre/contents.htm>. (дата обращения: 03.03.2021).
16. Зиятдинов С. И., Соколова Ю. В. Синтез комплексных дискретных фильтров. *Изв. вузов. Радиоэлектроника*, 2017, № 4, с. 12–19.
17. Зиятдинов С. И., Соколова Ю. В. Синтез комплексных дискретных фильтров на основе переходных характеристик. *Изв. вузов. Приборостроение*, 2017, т. 60, № 7, с. 641–647.

18. Котоусов А. С., Морозов А. К. *Оптимальная фильтрация и компенсация помех*. М., Горячая линия-Телеком, 2008. 166 с.
19. Малинкин В. Б. *Адаптивные фильтры в телекоммуникационных системах*. Новосибирск, СибГУТИ, 2005. 223 с.

20. Пинчук А. Н. Модель пассивной помехи при радиолокации низковисотных целей над морской поверхностью. *Радиоэлектроника*, 2015, № 3, с. 15–19.

UDC 621.396: 681.323

doi:10.31799/1684-8853-2021-3-53-60

Suppression of correlated interference by adaptive notch filters under pulse repetition period modulation

S. I. Ziatdinov^a, Dr. Sc., Tech., Professor, orcid.org / 0000-0002-8109-5456, ziat.53@mail.ru

L. A. Osipov^a, Dr. Sc., Tech., Professor, orcid.org / 0000-0002-8009-5706

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

Introduction: We discuss the problem of correlated noise suppression by adaptive complex notch filters of various orders. In order to eliminate the dependence of the transmission coefficient of the useful signal on its frequency, the pulse repetition period is modulated. **Purpose:** Studying the influence of pulse repetition period modulation on the correlated noise suppression coefficient. **Methods:** The notch filter parameters were optimized with the criterion of minimum average dispersion of correlated noise at the output of the filters during the repetition period modulation. **Results:** Expressions are obtained for the variance of correlated noise at the output of complex adaptive filters of various orders when the repetition period is modulated. Relationships are given for finding the optimal values of the tuning frequency and coefficients of the notch filters which minimize the correlated noise level at their output. Expressions are obtained for the coefficients of correlated noise suppression by notch filters in the context of pulse repetition period modulation. The graphs are presented showing how the correlated noise suppression coefficient depends on the relative value of the probing signal repetition period deviation for various values of the correlated noise spectral density width at optimal or non-optimal values of the tuning frequency and coefficients of the notch filters. It is shown that the use of probing pulse repetition period modulation leads to a decrease in the correlated noise suppression coefficient. On the other hand, the adaptation of the weighting coefficients for the adopted models of notch filters and correlated interference provides an increase in the suppression coefficient. **Practical relevance:** When developing or studying correlated noise suppression systems, the obtained results make it possible, taking into account the permissible losses of the suppression coefficient, to reasonably choose the input pulse repetition period deviation value in order to eliminate the effect of “blind” frequencies.

Keywords — correlated interference, notch filters, adaptation, repetition period modulation.

For citation: Ziatdinov S. I., Osipov L. A. Suppression of correlated interference by adaptive notch filters under pulse repetition period modulation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 3, pp. 53–60 (In Russian). doi:10.31799/1684-8853-2021-3-53-60

References

1. *Radar Handbook*. Ed. by M. I. Skolnik. 3rd ed. New York, McGraw-Hill, 2008. 1352 p.
2. Bakulev P. A., Stenin V. M. *Metody i ustroystva selektsii dvizhushchihся celej* [Methods and devices for selection of moving targets]. Moscow, Radio i sviaz' Publ., 1986. 286 p. (In Russian).
3. Popov D. I. Autocompensation of the Doppler phase of passive interference. *Digital Signal Processing*, 2009, no. 2, pp. 30–33 (In Russian).
4. Yastrebov A. V. Adaptive filter based on power series methods. *Izvestiia vysshikh uchebnykh zavedenii. Radioelektronika*, 2017, no. 1, pp. 5–11 (In Russian).
5. Zhuravsky V. N., Silin S. I. Synthesis of an adaptive detector in the spectral region. *Journal of Communications Technology and Electronics*, 2018, vol. 63, no. 12, pp. 1277–1284 (In Russian).
6. Popov D. I. Adaptive suppression of passive interference. *Digital Signal Processing*, 2014, no. 4, pp. 32–37 (In Russian).
7. Popov D. I. Adaptation of non-recursive notch filters. *Izvestiia vysshikh uchebnykh zavedenii. Radioelektronika*, 2009, vol. 52, no. 4, pp. 46–55 (In Russian).
8. Ziatdinov S. I. Influence of mismatch of parameters of quadrature channels on the adaptive system of selection of moving targets. *Izvestiia vysshikh uchebnykh zavedenii. Radioelektronika*, 2020, vol. 23, no. 3, pp. 6–9 (In Russian).
9. Michurin S. V. Influence of errors in tuning quadrature channels of a coherent system on the estimation of parameters of a complex signal. *Journal of Instrument Engineering*, 2020, vol. 63, no. 6, pp. 495–500 (In Russian).
10. *Mikroprocessornye sistemy avtomaticheskogo upravleniya* [Microprocessor automatic control systems]. V. A. Besekersky ed. Leningrad, Mashinostroenie Publ., 1988. 355 p. (In Russian).
11. Kholodnyak D., Turgaliev A., Baskakova A. A. Method to design lumped-element tunable bandpass filters with constant absolute bandwidth. *Proc. of 44th Eur. Microwave Conf.*, Roma, 6–9 Oct. 2014, Nijkerk, EuMA, 2014, pp. 335–338.
12. Baskakova A. E., Turgaliev V. M., Kholodnyak D. V. Tunable band-pass filter on elements with concentrated parameters with independent continuous control. *Izvestiia vysshikh uchebnykh zavedenii. Radioelektronika*, 2016, no. 3, pp. 25–32 (In Russian).
13. Popov D. I. Adaptive notch filters with complex weight coefficients. *Journal of “Almaz — Antey” Air and Space Defence Corporation*, 2015, no. 2, pp. 21–26 (In Russian).
14. Bakulev P. A. *Radiolokatsionnye sistemy* [Radar systems]. Moscow, Radiotekhnika Publ., 2004. 319 p. (In Russian).
15. Kokoshkin A. V., Korotkov V. A., Korotkov K. V., Novichikhin E. P. Methods for improving the distinguishability of objects in the presence of hydrometeors. *Journal of Radio Electronics*, 2015, no. 10, pp. 1–14 (In Russian). Available at: <http://jre.cplire.ru/jre/contents.htr>. (accessed 3 March 2021).

16. Ziatdinov S. I., Sokolova Yu. V. Synthesis of complex discrete filters. *Izvestia vysshikh uchebnykh zavedenii. Radioelektronika*, 2017, no. 4, pp. 12–19 (In Russian).
17. Ziatdinov S. I., Sokolova Yu. V. Synthesis of complex discrete filters based on transient characteristics. *Journal of Instrument Engineering*, 2017, vol. 60, no. 7, pp. 641–647 (In Russian).
18. Kotousov A. S., Morozov A. K. *Optimal'naya fil'traciya i kompensaciya pomekh* [Optimal filtering and noise compensation]. Moscow, Goriachaia liniia-Telekom Publ., 2008. 166 p. (In Russian).
19. Malinkin V. B. *Adaptivnye fil'try v telekommunikacionnyh sistemah* [Adaptive filters in telecommunication systems]. Novosibirsk, SibGUTI Publ., 2005. 223 p. (In Russian).
20. Pinchuk A. N. Model of passive interference in low-altitude radar targets over the sea surface. *Journal of Radio Electronics*, 2015, no. 3, pp. 15–19 (In Russian).

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.).

Текст рукописи должен быть оригинальным, а цитирование и самоцитирование корректно оформлено.

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, ORCID и электронный адрес одного из авторов. При написании аннотации не используйте аббревиатур и не делайте ссылку 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.

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, подающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, подающихся 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 и названий таблиц на русском и английском языках 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, doi;

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

— 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/>): Литература и References.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Руководство для авторов».

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



Международная конференция Big Data Days 2021

28–30 сентября 2021

Конференция Big Data Days 2021 посвящена техническим докладам в областях больших данных, высоких нагрузок, обработки и анализа данных и машинного обучения.

Конференция объединяет разработчиков, специалистов в области ИТ и пользователей, позволяя им поделиться опытом, обсудить передовые методы и достигнутые успехи в отношении сценариев использования и коммерческого применения.

Мероприятие предназначено для обучения, информирования и вдохновения — организовано людьми, увлеченными большими данными и исследованием данных.

Big Data Days 2021 — трехдневное мероприятие, состоящее из одного дня мастер-классов и двух дней конференции.

Организатор

ООО «ДАТА МИНЕР ЛАБС»

Формат проведения

В этом году Big Data Days Moscow пройдет в гибридном формате: вы сможете посещать мастер-классы и слушать доклады офлайн или онлайн — на ваше усмотрение. Часть спикеров поделится своими знаниями очно, а некоторые

сделают это онлайн. Однако вы сможете взаимодействовать с ними в любом выбранном вами формате! Конференция на месте будет проводиться с соблюдением всех необходимых требований гигиены и дистанции, поэтому количество билетов на очную часть конференции может быть ограничено в соответствии с требованиями.

Направления работы

- Big Data
- High Load
- Data Science
- Озеро данных
- Машинное обучение
- Искусственный интеллект
- Прогнозная аналитика
- Облачные технологии
- Другие связанные темы

Дополнительная информация и справки

Почтовый адрес для доставки документов:
105203, г. Москва, ул. Первомайская, д. 116, 41
Тел.: +7 (906) 084 33 24
Эл. адреса:
info@bigdatadays.ru,
tickets@bigdatadays.ru
Сайт: <https://bigdatadays.ru/>

**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

**ВОСТРИКОВ
Антон
Александрович**



Доцент кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2000 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети». В 2004 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 45 научных публикаций и двух свидетельств о регистрации программного продукта. Область научных интересов — распределенные и встраиваемые информационно-управляющие системы, обработка визуальной информации, оптико-информационные системы.
Эл. адрес: vostricov@mail.ru

**ДУШКИН
Александр
Викторович**



Профессор кафедры информационной безопасности Московского института электронной техники, старший научный сотрудник НИЦ Военно-воздушной академии им. профессора Н. Е. Жуковского и Ю. А. Гагарина, Воронеж. В 1989 году окончил Воронежское высшее военное инженерное училище радиоэлектроники по специальности «Радиоэлектронные средства». В 2012 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 450 научных публикаций и 50 свидетельств о государственной регистрации программ для ЭВМ. Область научных интересов — радиоэлектронные и информационные системы, специальные аппаратно-программные комплексы.
Эл. адрес: a_dushkin@mail.ru

**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

**КРУГЛИК
Станислав
Александрович**



Младший научный сотрудник, аспирант Сколковского института науки и технологий, старший преподаватель кафедры радиотехники и систем управления Московского физико-технического института. В 2017 году окончил Московский физико-технический институт по специальности «Прикладные математика и физика». Является автором 26 научных публикаций. Область научных интересов — теория информации, теория кодирования, защита информации.
Эл. адрес: stanislav.kruglik@skoltech.ru

**КУЧЕРОВ
Григорий
Аронович**



Ведущий научный сотрудник Сколковского института науки и технологий, ведущий научный сотрудник Французского национального центра научных исследований (CNRS). В 1982 году окончил Новосибирский электротехнический институт по специальности «Прикладная математика». В 1988 году защитил диссертацию на соискание ученой степени кандидата физ.-мат. наук, в 2000 году — хабилитацию в Университете Анри Пуанкаре, Нанси, Франция. Является автором около 100 научных публикаций. Область научных интересов — дискретные алгоритмы, структуры данных, вычислительная биология и биоинформатика, большие данные, программная инженерия.
Эл. адрес: G.Kucherov@skoltech.ru

ЛЕБЕДЕВ
Илья
Сергеевич



Профессор, главный научный сотрудник Санкт-Петербургского федерального исследовательского центра РАН.

В 1998 году окончил Санкт-Петербургское высшее военное училище ПВО по специальности «Инженер-математик».

В 2012 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 80 научных публикаций.

Область научных интересов — информационные технологии, машинное обучение, компьютерная лингвистика.

Эл. адрес: isl_box@mail.ru

НАЗИРХАНОВА
Камилла
Назирова



Аспирант Стэнфордского университета, США.

В 2020 году окончила магистратуру МФТИ и Сколковского института науки и технологий по специальности «Прикладные математика и физика».

Является автором шести научных публикаций.

Область научных интересов — теория кодирования, распределенные системы хранения данных.

Эл. адрес: nazirk@stanford.edu

НИКИФОРОВ
Виктор
Викентьевич



Профессор, главный научный сотрудник лаборатории информационно-вычислительных систем и технологий программирования Санкт-Петербургского института информатики и автоматизации РАН.

В 1965 году окончил Ленинградский политехнический институт по специальности «Радиотехника».

В 1991 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором 120 научных публикаций.

Область научных интересов — системы реального времени, встроенные системы, операционные системы.

Эл. адрес: nik@iias.spb.su

ОСИПОВ
Василий
Юрьевич



Профессор, главный научный сотрудник, руководитель лаборатории информационно-вычислительных систем и технологий программирования Санкт-Петербургского института информатики и автоматизации РАН.

В 1981 году окончил Высшее военно-морское училище радиоэлектроники им. А. С. Попова по специальности «Радиотехнические средства».

В 2000 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 100 научных публикаций и семи патентов на изобретения.

Область научных интересов — интеллектуальные системы, нейронные сети, моделирование, информационная безопасность. Эл. адрес: osipov_vasily@mail.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

ПАРФИРЬЕВ
Андрей
Владимирович



Доцент кафедры робототехнических комплексов и систем воздушного базирования Военно-воздушной академии им. профессора Н. Е. Жуковского и Ю. А. Гагарина, Воронеж.

В 2009 году окончил Тамбовское высшее военное авиационное инженерное училище радиоэлектроники по специальности «Автоматизированные системы управления и обработки данных».

В 2015 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 82 научных публикаций и пяти патентов на изобретения.

Область научных интересов — цифровая обработка изображений, системы автоматического управления, системы навигации беспилотных летательных аппаратов.

Эл. адрес: keeperate@mail.ru

ПАРФИРЬЕВА
Оксана
Владимировна



Начальник отделения учебной лаборатории кафедры робототехнических комплексов и систем воздушного базирования Военно-воздушной академии им. профессора Н. Е. Жуковского и Ю. А. Гагарина, Воронеж.

В 2011 году окончила Борисоглебский государственный педагогический институт по специальности «Физико-математическое образование», в 2018 году — Воронежский государственный университет по специальности «Бухгалтерский учет, статистика».

Является автором одной научной публикации.

Область научных интересов — цифровая обработка изображений, системы автоматического управления, системы навигации беспилотных летательных аппаратов.

Эл. адрес: aregard@inbox.ru

ФИЛИТОВ
Михаил
Егорович



Магистрант Сколковского института науки и технологий образовательной программы «Информационные науки и технологии».

В 2020 году окончил бакалавриат Высшей школы экономики по специальности «Прикладная математика и информатика».

Область научных интересов — теория информации, биоинформатика, распределенные системы.

Эл. адрес: filitovme@gmail.com

СЕРГЕЕВ
Александр
Михайлович



Доцент 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».

В 2020 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 35 научных публикаций.

Область научных интересов — численные методы, теория вычислительных процессов, проектирование специализированных процессоров.

Эл. адрес: asklab@mail.ru