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

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

4(107)/2020

doi:10.31799/1684-8853-2020-4 ISSN 1684-8853 (print); ISSN 2541-8610 (online)

SCIENTIFIC JOURNAL

4(107)/2020

PEER REVIEWED JOURNAL

INFORMATSIONNO-UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Contact information
The Editorial and Publishing Center, SUAI
67, B. Morskaia, 190000, St. Petersburg, Russia
Website: http://i-us.ru/en, e-mail: ius.spb@gmail.com
Tel.: +7 - 812 494 70 02

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

Submitted for publication 06.07.20. Passed for printing 21.08.20. Format 60×84$_{1/8}$.
Phototype SchoolBookC. Digital printing.

Layout original is made at the Editorial and Publishing Center, SUAI.
67, B. Morskaia, 190000, St. Petersburg, Russia
Printed from slides at the Editorial and Publishing Center, SUAI.
67, B. Morskaia, 190000, St. Petersburg, Russia

The journal is indexed in Scopus.

doi:10.31799/1684-8853-2020-4 ISSN 1684-8853 (print); ISSN 2541-8610 (online)
НАУЧНЫЙ ЖУРНАЛ

4(107)/2020

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

Адрес редакции: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,
сайт: http://i-us.ru

Журнал входит в БД SCOPUS, в RSCI на платформе Web of Science 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 06.07.20. Подписано в печать 21.08.20. Формат 60×84 1/8.
Гарнитура SchoolBookC. Печать цифровая.
Усл. печ. л. 10,6. Уч.-изд. л. 14,5. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 222.
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.

UDC 004.438
doi:10.31799/1684-8853-2020-4-2-10

Articles

Conference matrices from Legendre C-pairs

N. A. Balonin[a], *Dr. Sc., Tech., Professor, orcid.org/0000-0001-7338-4920, korbendfs@mail.ru*
D. Ž. Đoković[b], *Dr. Sc., Distinguished Professor Emeritus, orcid.org/0000-0002-0176-2395, djokovic@uwaterloo.ca*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*
[b]*University of Waterloo, Department of Pure Mathematics and Institute for Quantum Computing, Waterloo, Ontario, N2L 3G1, Canada*

Introduction: *There are just a few known methods for the construction of symmetric C-matrices, due to the lack of a universal structure for them. This obstruction is fundamental, in addition, the structure of C-matrices with a double border is incompletely described in literature, which makes its study especially relevant.* ***Purpose:*** *To describe the two-border two-circulant construction in detail we introduce the concept of the Legendre C-pairs.* ***Results:*** *The paper deals with C-matrices of order $n = 2v + 2$ with two borders and extends the so called generalized Legendre pairs, v odd, to a wider class of Legendre C-pairs with even and odd v, defined on a finite abelian group G of order v. Such a pair consists of two functions a, b: $G \to Z$, whose values are +1 or −1 except that $a(e) = 0$, where e is the identity element of G and Z is the ring of integers. To characterize the Legendre C-pairs we use the subsets $X = \{x \in G: a(x) = -1\}$ and $Y = \{x \in G: b(x) = -1\}$ of G. We show that $a(x^{-1}) = (-1)^v a(x)$ for all x. For odd v we show that X and Y form a difference family, which is not true for even v. These difference families are precisely the so called Szekeres difference sets, used originally for the construction of skew-Hadamard matrices. We introduce the subclass of the special Legendre C-pairs and prove that they exist whenever $2v + 1$ is a prime power. In the last two sections of the paper we list examples of special cyclic Legendre C-pairs for lengths $v < 70$.* ***Practical relevance:*** *C-matrices are used extensively in the problems of error-free coding, compression and masking of video information. Programs for search of conference matrices and a library of constructed matrices are used in the mathematical network "mathscinet.ru" together with executable on-line algorithms.*

Keywords *— conference matrices, skew-Hadamard matrices, periodic autocorrelation functions, Szekeres difference families, generalized Legendre pairs, constructions, telephony.*

For citation: 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

Introduction

We introduce the notion of Legendre C-pairs on a finite abelian group G of order v and use it to construct many C-matrices of order $n = 2v + 2$ with two borders and a core made up from two multi-circulants. Such a pair consists of two functions a, b: $G \to Z$, whose values are +1 or −1 except that $a(e) = 0$, where e is the identity element of G. (By Z we denote the ring of integers.) Moreover the sum of the periodic autocorrelation functions (PAF) of a and b must be −2, except at shift 0. These pairs are similar to the so called generalized Legendre (GL) pairs. While for the GL-pairs v must be odd, there is no such restriction for the Legendre C-pairs.

In Proposition 1 we show that $a(x^{-1}) = (-1)^v a(x)$ for all x and determine the cardinalities of the sets $X = \{x \in G: a(x) = -1\}$ and $Y = \{x \in G: b(x) = -1\}$. We introduce special Legendre C-pairs and in Proposition 2 and Corollary 2 we show that they exist whenever $2v + 1$ is a prime power.

In the last two sections of the paper we list examples of special cyclic Legendre C-pairs for lengths $v < 70$. In Proposition 3 we characterize the Legendre C-pairs in terms of the subsets X and Y defined above. For odd v we show that X and Y form a difference family, which is not true for even v. These difference families are precisely the so called Szekeres difference sets (see Corollary 3). Originally they were used for the construction of skew-Hadamard matrices, see [1–4]. A wider class of two-block difference families with parameters $(v; (v - 1)/2, (v - 1)/2; (v - 3)/2)$, v odd, has been investigated recently in [5]. In Fig. 3 we summarize diagramatically the main facts about such families. The existence question remains open in many cases.

Let us now recall some definitions and facts about Hadamard and conference matrices. A *Hadamard matrix* is a matrix $\mathbf{H}$ of order n with entries +1 or −1 and such that $\mathbf{H}\mathbf{H}^{\mathrm{T}} = n\mathbf{I}$ (T is the transposition operator and $\mathbf{I}$ is the identity matrix of some order, here of order n). If such a matrix exists and $n > 2$ then n must be divisible by 4. A Hadamard matrix $\mathbf{H}$ is a *skew-Hadamard matrix* if $\mathbf{H} + \mathbf{H}^{\mathrm{T}} = 2\mathbf{I}$.

A *conference matrix* (*C-matrix*) is a matrix $\mathbf{C}$ of order n whose diagonal entries are zeros, the other entries are +1 or −1, and $\mathbf{C}\mathbf{C}^{\mathrm{T}} = (n - 1)\mathbf{I}$. If such a matrix exists and $n > 1$ then n must be even. Two C-matrices of the same order are *equivalent* if one can be transformed to the other by permuting the rows and columns so that the diagonal zeros are preserved and by multiplying by −1 some

rows and some columns. Every equivalence class of C-matrices of order n contains a symmetric matrix if $n \equiv 2 \pmod 4$ and a skew-symmetric matrix if $n \equiv 0 \pmod 4$. If $\mathbf{C}$ is a skew-symmetric C-matrix of order n then $\mathbf{H} = \mathbf{C} + \mathbf{I}$ is a skew-Hadamard matrix of order n, and the converse holds.

It is well-known (see e. g. [6]) that if a C-matrix of order $n \equiv 2 \pmod 4$ exists then $n - 1$ must be a sum of two squares. Let us list such integers $n < 200$:

$$2, 6, 10, 14, 18, 26, 30, 38, 42, 46, 50, 54, 62, 66, 74, 82, 86, 90, 98, 102, 110, 114, 118, 122, 126, 138, 146, 150, 154, 158, 170, 174, 182, 186, 194, 198. \quad (1)$$

We say that a symmetric or skew-symmetric C-matrix is *normalized* if all entries of its first row are +1, except the first entry which must be 0. In that case, its *core* is the submatrix obtained by dropping the first row and column.

The basic examples of C-matrices are so called *Paley C-matrices* $\mathbf{C}$ of order $n = q + 1$ where q is a power of an odd prime (see e. g. [7] or [8, Chapter 18]). The matrix $\mathbf{C}$ is normalized and its core $\mathbf{Q}$ is a matrix of order q. The rows and columns of $\mathbf{Q}$ are labeled by the elements of the finite field GF(q) of order q. The entries of $\mathbf{Q}$ are given by the formula $\mathbf{Q}_{x,y} = \chi(x - y)$, where χ is the quadratic character of GF(q). We recall that $\chi(x) = 1$ if $x \neq 0$ is a square in GF(q), $\chi(x) = -1$ if x is not a square, and $\chi(0) = 0$. Moreover, χ satisfies the multiplicative property $\chi(xy) = \chi(x)\chi(y)$ for all x, y in GF(q). Both $\mathbf{C}$ and $\mathbf{Q}$ are symmetric if $q \equiv 1 \pmod 4$ and skew-symmetric otherwise.

There are just a few methods of construction of symmetric C-matrices. They are listed in the recent survey paper [9]. We investigate here only one of these methods, namely the *two-border two-circulant* (2b2c) *construction*. In the next section we describe this construction in detail and introduce the concept of Legendre C-pairs.

Legendre C-pairs

For the sake of simplicity, we consider first the case of the cyclic group $Z_v = \{0, 1, ..., v - 1\}$ under addition modulo v. In that case we treat the functions on Z_v as sequences of length v. Let a and b be two sequences of length v

$$a = (a_0, a_1, ..., a_{v-1}); b = (b_0, b_1, ..., b_{v-1}), \quad (2)$$

where $a_0 = 0$ while all other a_i and all the b_i are equal to +1 or −1. Recall that the value of the PAF of a sequence $x = (x_0, x_1, ..., x_{v-1})$ at shift s is

$$\mathrm{PAF}_x(s) = \sum_{i=0}^{v-1} x_i x_{i+s}.$$

Definition 1 (*cyclic case*). We say that the pair (a, b) given by (2) is a (*cyclic*) *Legendre C-pair* if the sum of the PAFs of the sequences a and b has the constant value −2, except at the shift 0 where the sum attains its peak value $2v - 1$.

The first examples of Legendre C-pairs originate from Number Theory. Indeed let $v = p$ be a prime number $\equiv 3 \pmod 4$. Then the sequence a of Legendre symbols

$$a_i = \left(\frac{i}{p}\right),\ i = 0, 1, ..., p - 1$$

has constant PAF values, $\mathrm{PAF}_a(s) = -1$ for all nonzero s. The same is true for the sequence b which is obtained from a by replacing the first term 0 by +1. Hence (a, b) is a Legendre C-pair of length p. This construction does not work when p is a prime number $\equiv 1 \pmod 4$.

The above definition differs from that of "generalised Legendre pairs" given in [10, p. 76] (see also [6]) which requires that all elements of the sequences a and b be +1 or −1. We shall refer to them simply as "Legendre pairs". For the Legendre pairs, the sum of the PAFs of a and b is required to be the constant function −2 except for the value $2v$ at shift 0. Our definition of Legendre C-pairs is designed for the construction of C-matrices and so the condition that $a_0 = 0$ is mandatory. It is mentioned in [10, p. 80] that the length of Legendre pairs must be odd. On the other hand, there exist Legendre C-pairs of even and odd lengths. Those of even length give symmetric C-matrices while the ones of odd length give skew-symmetric C-matrices (and skew-Hadamard matrices).

Now let p be a prime number $\equiv 1 \pmod 4$. Let a and b be the sequences obtained from the sequence of Legendre symbols by replacing the 0 term by +1 and −1 respectively. Then (a, b) is a Legendre pair of length p which cannot be used to produce a Legendre C-pair of the same length.

Next we show how to use a cyclic Legendre C-pair (a, b) of length v to construct a C-matrix $\mathbf{C}$ of order $n = 2v + 2$. Let $\mathbf{A}$ and $\mathbf{B}$ denote the circulant matrices whose first rows are given by a and b, respectively.

In the case when v is even, the matrix $\mathbf{C}$ is obtained by plugging the blocks $\mathbf{A}$ and $\mathbf{B}$ into the next array; the first two rows and columns of $\mathbf{C}$ form its *border* and its *core* is made up from the circulants $\mathbf{A}$ and $\mathbf{B}$ (and their transposes):

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & \mathbf{e}^{\mathrm{T}} & \mathbf{e}^{\mathrm{T}} \\ 1 & 0 & \mathbf{e}^{\mathrm{T}} & -\mathbf{e}^{\mathrm{T}} \\ \mathbf{e} & \mathbf{e} & \mathbf{A} & \mathbf{B} \\ \mathbf{e} & -\mathbf{e} & \mathbf{B}^{\mathrm{T}} & -\mathbf{A}^{\mathrm{T}} \end{pmatrix}. \quad (3)$$

(By $\mathbf{e}$ we denote a column vector of 1's.)

It is easy to verify (see Problem 18B on p. 174 and the hint on p. 490 in [8]) that if the matrix **C** defined by (3) is a C-matrix then **C** must be symmetric. Consequently, if **C** is a C-matrix then **A** must be symmetric, i. e. $a_i = a_{v-i}$ for $i = 1, \ldots, v-1$. For an example of such C-matrix see Fig. 1, *a*.

In the case when v is odd, we use the modified array

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & \mathbf{e}^{\mathrm{T}} & \mathbf{e}^{\mathrm{T}} \\ -1 & 0 & \mathbf{e}^{\mathrm{T}} & -\mathbf{e}^{\mathrm{T}} \\ -\mathbf{e} & -\mathbf{e} & \mathbf{A} & \mathbf{B} \\ -\mathbf{e} & \mathbf{e} & -\mathbf{B}^{\mathrm{T}} & \mathbf{A}^{\mathrm{T}} \end{pmatrix}. \qquad (4)$$

As in the previous case one can show that if **C** is a C-matrix then the block **A** must be skew-symmetric matrix, i. e. $a_i + a_{v-i} = 0$ for $i = 1, \ldots, v-1$. For an example see Fig. 1, *b*.

We shall now extend the definition of Legendre C-pairs to any abelian group G of order v with identity element e. Denote by * the involution of the integral group ring $Z[G]$ sending any element x to its inverse x^{-1}. For any $z \in Z[G]$ we define its *norm* $N(z)$ to be the product zz^*. For a subset X of G we say that it is *symmetric* if $X^* = X$ and that it is *skew* if G is a disjoint union of X, X^* and $\{e\}$. For any function a: $G \to Z$ we define its periodic autocorrelation function, PAF_a: $G \to Z$, by the formula

$$\mathrm{PAF}_a(s) = \sum_{x \in G} a(x)a(x+s).$$

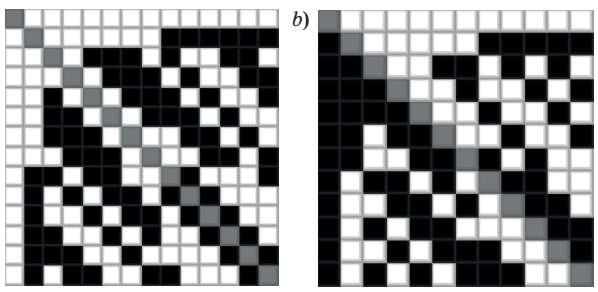
Definition 1 (*general case*). Let a and b be functions $G \to Z$ such that $a(e) = 0$ while all other values of a and all values of b belong to the set $\{+1, -1\}$. We say that the pair (a, b) is a *Legendre C-pair* if the sum of the PAFs of a and b has the constant value -2, except at the shift 0 where the sum attains its peak value $2v - 1$.

To a function a: $G \to Z$ we associate the matrix **A** of order v whose rows and columns are labeled by the elements of G and are given by the formula $\mathbf{A}_{x,y} = a(x^{-1}y)$. Such matrices are known as *G-invariant matrices* because they have the property that $\mathbf{A}_{xz,yz} = \mathbf{A}_{x,y}$ for all x, y, z in G. (By suitably arranging the indices, such matrices can be written as multi-circulants, i. e., circulants of circulants of ...). If (a, b) is a Legendre C-pair of length v and **A** and **B** are their associated matrices, it is easy to show that the matrix **C** given by (3) or (4) is a C-matrix. In that case we say that (a, b) is the *Legendre C-pair of* **C**. It is well known, see [7, Theorem 2.2], that each Paley C-matrix is equivalent to one of the form (3) or (4) with circulant blocks **A** and **B**. Hence, cyclic Legendre C-pairs of length v exist whenever $2v + 1$ is a prime power. The converse is false. For instance there exist a Legendre C-pair of length $v = 7$ while $2v + 1 = 15$ is not a prime power. For a concrete example with multi-circulant blocks **A** and **B** of order 25 see [5] or [6, section 10.3]. Another example is given below.

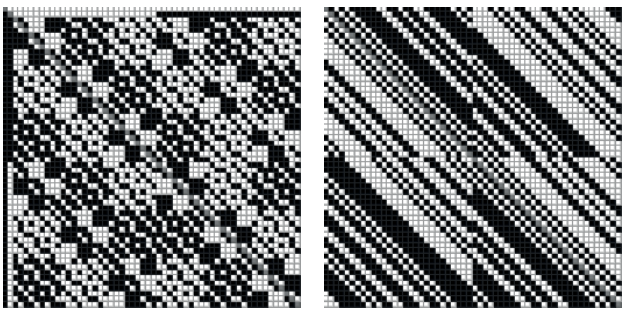
On Fig. 2 we show two skew-symmetric C-matrices. The first one has the form (4) with multi-circulant blocks **A** and **B** of order 27. It is constructed from the difference set X consisting of the nonzero squares in GF(27). To construct this field we used the primitive polynomial $x^3 - x^2 + 1$ over Z_3. The elements of GF(27) are the 27 polynomials $a + bx + cx^2$, with $a, b, c \in \{0, 1, 2\}$. We encode this polynomial by the symbol *abc*, and arrange the symbols in the lexicographic order 000, 001, 002, 010, 011, 012, ..., 222. Explicitly, we have $X = \{1, x^2, 2 + 2x + x^2, 2x + 2x^2, 2 + x + x^2, 1 + 2x + x^2, 2x + x^2, 2x, 1 + 2x^2, 1 + x, 2 + 2x^2, 1 + x + x^2, 1 + 2x\}$.

The corresponding 13 symbols are 100, 001, 221, 022, 211, 121, 021, 020, 102, 110, 202, 111, 120. The matrix **B** associated to X has -1 entries exactly at these 13 positions. It has the block-circulant structure with the first row **[U V W]**. The matrices **U**, **V**, **W** are also multi-circulants but of order 9. Their first block-rows are **[P, J, −J]**, **[−P, −Q, −Q]**, **[Q, P, P]** where **P**, **Q**, **J** are the circulants with the first rows [1, −1, 1], [1, 1, −1], [1, 1, 1], respectively. Further, $\mathbf{A} = \mathbf{B} - \mathbf{I}$.

The second matrix has the form $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^{\mathrm{T}} & \mathbf{A}^{\mathrm{T}} \end{pmatrix}$ where **A** and **B** are negacyclic blocks of size $v = 28$. We recall that a square matrix of order n is *nega-*



■ ***Fig. 1***. 2b2c conference matrices of orders 14 (*a*) and 12 (*b*), matrix portraits have white and black colors for entries 1, −1 and gray for 0



■ ***Fig. 2***. Two skew-symmetric C-matrices of order 56

cyclic if each row but the first is obtained from the previous one by the negacyclic shift

$$(x_1, x_2, x_3, ..., x_{n-1}, x_n) \to$$
$$\to (-x_n, x_1, x_2, ..., x_{n-2}, x_{n-1}).$$

The first rows of **A** and **B**, respectively, are

$a = [0, -1, 1, -1, 1, -1, -1, -1, -1, -1, 1, 1, -1, -1, 1, -1, -1, 1, 1, -1, -1, -1, -1, -1, 1, -1, 1, -1];$
$b = [1, 1, 1, -1, 1, -1, 1, 1, -1, -1, 1, 1, 1, -1, -1, 1, 1, -1, 1, -1, 1, 1, 1, 1, 1, -1, -1, -1].$

Since A is negacyclic and $a_i = a_{v-i}$ for $i = 1, ..., v - 1$ the block **A** is skew-symmetric. Hence the second matrix is also skew-symmetric.

The list (1) gives the feasible sizes $n < 200$ of symmetric C-matrices. It is known that such matrices exist when $n - 1$ is a prime power. By removing such sizes we are left with only seven cases $n = 46$, 66, 86, 118, 146, 154, 186. C-matrices of size 46 have been constructed long time ago, while it is still unknown whether they exist in the remaining six sizes, see [9]. By using a computer search, we have shown that there are no cyclic Legendre C-pairs of length 22 or 32.

The first assertion of the following proposition follows from the properties of the block **A** mentioned earlier in this section.

Proposition 1. Let (a, b) be a Legendre C-pair on an abelian group G of order v and let k_1 and k_2 be the cardinalities of the sets $\{x \in G: a(x) = -1\}$ and $\{x \in G: b(x) = -1\}$, respectively. Then $a(x^{-1}) = (-1)^v a(x)$ for all x. If v is even then $k_1 = k_2 = v/2$. If v is odd then $k_1 = k_2 = (v - 1)/2$.

Proof: Let us prove the second assertion. By the hypothesis, v is even. Let **A** and **B** be the multi-circulant matrices associated to the functions a and b, respectively. Since (a, b) is a Legendre C-pair, the matrix **C** given by (3) is a C-matrix. The first and the second row of **C** are orthogonal to the third row. This gives the two equations

$$1 + (v - 1 - 2k_1) + (v - 2k_2) = 0;$$
$$1 + (v - 1 - 2k_1) - (v - 2k_2) = 0.$$

It follows that $k_1 = k_2 = v/2$. The proof of the third assertion is similar.

Corollary 1. If (a, b) is a Legendre C-pair on an abelian group G of odd order, then after replacing $a(e) = 0$ by $a(e) = 1$ (or $a(e) = -1$) we obtain a Legendre pair. (The converse is not valid.)

Special Legendre C-pairs

We now introduce a subclass of the class of Legendre C-pairs.

Definition 2. Let (a, b) be a Legendre C-pair on G, a group of order v. If v is odd we say that (a, b) is *special* if the subset $\{x \in G: b(x) = -1\}$ is symmetric. If v is even and G is cyclic, say $G = Z_v$, we say that (a, b) is *special* if $b_{v-1-i} = -b_i$, $i = 0, 1, ..., v - 1$.

We shall now prove that cyclic special Legendre C-pairs of length v exist whenever $2v + 1$ is a prime power.

Proposition 2. The equivalence class of any Paley C-matrix **C** contains a C-matrix of the form (3) or (4) whose Legendre C-pair is cyclic and special.

Proof: We follow the proof of [7, Theorem 2.2] and modify it in order to prove our assertion. All Paley C-matrices of the same order are mutually equivalent, see [7]. Hence it suffices to construct for each odd prime power $q = 2v + 1$ a Paley C-matrix **C** of order $q + 1$, having the 2b2c form (3) or (4), whose Legendre C-pair is cyclic and special.

Let V be the 2-dimensional vector space over GF(q) with basis vectors $x = (1, 0)$ and $y = (0, 1)$. Choose a primitive element η of GF(q). Define the vectors $z_k = y - \eta^k x = (-\eta^k, 1)$ for $k = 0, 1, ..., q - 2$. We arrange the vectors x, y and the z_k as follows: $x, y, z_0, z_2, ..., z_{2v-2}, z_1, z_3, ..., z_{2v-1}$. Note that this arrangement is slightly different from the one used in [7]. We use these $q + 1$ vectors to define a Paley C-matrix **C** as usual and let (a, b) be its Legendre C-pair. It follows from [7, Theorem 2.2] that **C** has the 2b2c form.

In more details, the entries of **C** are computed as follows. Let i and j be any indices in the range 1, 2, ..., $q + 1$. Let u and w be the i-th and j-th vectors in the above list, respectively. Then the (i, j)-th entry of **C** is equal to $\chi(\det(u, w))$ where (u, w) denotes the matrix of order two made up from u and w. Moreover, $a_k = \mathbf{C}_{3,3+k}$ and $b_k = \mathbf{C}_{3,v+3+k}$ for $k = 0, 1, ..., v - 1$.

We claim that $\mathbf{C}_{3,q+1-k} = -\chi(-1)\mathbf{C}_{3,v+3+k}$ for $k = 0, 1, ..., v - 1$. Indeed for $i = 3$ and $j = q + 1 - k$ we have $u = z_0 = (-1, 1)$ and $w = z_{q-2-2k} = (-\eta^{-2k-1}, 1)$. Since $\det(u, w) = \eta^{-2k-1} - 1$, we have $\mathbf{C}_{3,q+1-k} = \chi(\eta^{-2k-1} - 1 = -\chi(-1)\chi(\eta^{2k+1} - 1)$. Similarly we have $\mathbf{C}_{3,v+3+k} = \chi(\eta^{2k+1} - 1)$. We conclude that our claim holds.

Hence the sequence b satisfies the equalities $b_{v-k} = -\chi(-1)b_k$ for $k = 0, 1, ..., v - 1$. If $q \equiv 1 \pmod 4$ this means that the Legendre C-pair (a, b) is special. The same is true in the case $q \equiv 3 \pmod 4$ after a suitable cyclic shift of b. This completes the proof.

Corollary 2. Cyclic special Legendre C-pairs of length v exist whenever $2v + 1$ is a prime power.

Many special Legendre C-pairs of odd length v can be constructed from the so called Szekeres difference sets. We recall that the *Szekeres difference sets* are in fact a difference family (DF) in an abelian group G of odd order v consisting of two blocks, X and Y, such that X is skew and $|X| = |Y| = (v - 1)/2$. Hence the parameters of such DF are

$$(v;\ (v-1)/2,\ (v-1)/2;\ (v-3)/2). \qquad (5)$$

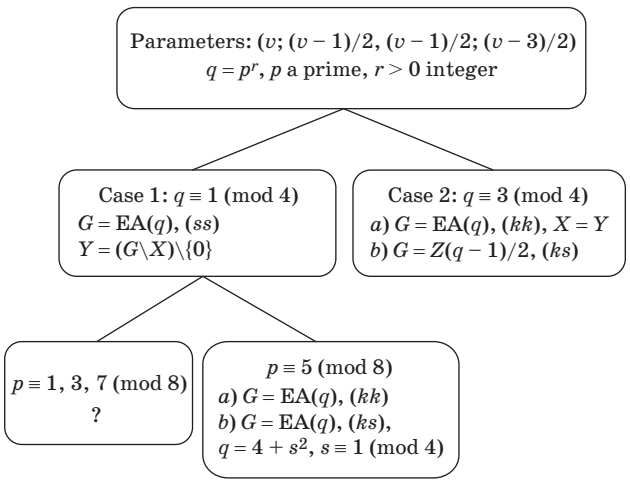
We say that a Szekeres DF (X, Y) is *special* if Y is symmetric. It is well known, see [1, Theorem 5.18] or [2], that the special Szekeres DFs in Z_v exist whenever $2v + 1$ is a prime power $\equiv 3 \pmod 4$. Let us give a simple example.

Example 1. Let $v = 5$ and note that $2v + 1 = 11$ is a prime $\equiv 3 \pmod 4$. The subsets $X = \{g, g^3\} = g + g^3$ and $Y = \{g^2, g^3\} = g^2 + g^3$ of the cyclic group $G = < g >$ of order 5 form a special Szekeres DF. The * operator sends g^i to g^{-i}. Thus $X^* = \{g^2,\ g^4\} = g^2 + g^4$, $Y^* = Y$ and we have $G = X + X^* + e$. To construct the special Legendre C-pair, we form the binary sequence $a = (1, -1, 1, -1, 1)$ by setting $a_i = -1$ if $g^i \in X$ and $a_i = 1$ otherwise, and form $b = (1, 1, -1, -1, 1)$ similarly by using Y. The sum of the PAFs of a and b is the constant function -2 apart from the origin. Thus (a, b) is a Legendre pair. By replacing the first term of a by 0 we obtain a special Legendre C-pair.

Difference families with parameters (5)

In the following diagram (Fig. 3) we summarize the main results on the existence of DFs (X, Y) with parameters (5) and X and Y symmetric or skew. By EA(q) we denote an elementary abelian group of order $q = p^r$, p a prime, r a positive integer. [The additive group of GF(q) is an EA(q).] The symmetry types of X and Y are indicated by the letter s if the block is symmetric, k if it is skew, and the symbol $*$ if no symmetry is required. For instance (ks) means that X is skew and Y is symmetric. While arbitrary Szekeres DFs have the symmetry type ($k*$), the special ones have type (ks).

In Case 1 the block X is the set of nonzero squares in GF(q). The same is true for part a) of Case 2, see [2, Theorem 4]. For part b) of Case 2 see [2, Theorem 3] and [1, Theorem 5.18]. Note that in the first subcase ($p \equiv 1, 3, 7 \pmod 8$) of Case 1 we also require that $q \equiv 1 \pmod 8$. No general result seems to be known about this subcase. In the second subcase ($p \equiv 5 \pmod 8$), part a) was proved in [2, Theorem 5] when $q \equiv 5 \pmod 8$ and in [3, 4] when $q \equiv 1 \pmod 8$. For part b) of the same subcase see [11].



■ ***Fig. 3.*** Existence of DFs (X, Y) with parameters (5)

Let us give two examples of Legendre C-pairs (X, Y) of type (kk) and length v with $2v + 1$ not a prime power. Both examples are obtained from a theorem of Szekeres, see [2, Theorem 6] and [5, Theorem 2]. First example: we start with the Szekeres DF (X, Y) in Z_{37}:

$X = \{3, 4, 5, 6, 8, 11, 13, 17, 19, 21, 22, 23, 25, 27, 28, 30, 35, 36\}$;
$Y = \{2, 3, 4, 11, 14, 15, 18, 20, 21, 24, 25, 27, 28, 29, 30, 31, 32, 36\}$.

The corresponding binary sequences a and b are

$a = (1, 1, 1, -1, -1, -1, -1, 1, -1, 1, 1, -1, 1, -1, 1, 1, 1, -1, 1, -1, 1, -1, -1, -1, 1, -1, 1, -1, -1, 1, -1, 1, 1, 1, 1, -1, -1)$;
$b = (1, 1, -1, -1, -1, 1, 1, 1, 1, 1, 1, -1, 1, 1, -1, -1, 1, 1, -1, 1, -1, -1, 1, 1, -1, -1, 1, -1, -1, -1, -1, -1, -1, 1, 1, 1, -1)$.

Since X is skew, by replacing the first term of a by 0, we obtain a Legendre C-pair of length 37. Second example: (X, Y) in Z_{61}:

$X = \{2, 3, 4, 5, 7, 14, 18, 19, 23, 24, 26, 27, 30, 32, 33, 36, 39, 40, 41, 44, 45, 46, 48, 49, 50, 51, 52, 53, 55, 60\}$;
$Y = \{3, 4, 5, 6, 8, 10, 11, 14, 17, 19, 21, 27, 28, 29, 31, 35, 36, 37, 38, 39, 41, 43, 45, 46, 48, 49, 52, 54, 59, 60\}$.

It gives the Legendre C-pair

$a = (0, 1, -1, -1, -1, -1, 1, -1, 1, 1, 1, 1, 1, 1, -1, 1, 1, 1, -1, -1, 1, 1, 1, -1, -1, 1, -1, -1, 1, 1, -1, 1, -1, -1, 1, 1, -1, -1, -1, 1, -1, -1, -1, -1, 1, -1, 1, 1, 1, 1, -1, -1, -1, -1, 1, -1, 1, 1, 1, 1, -1, -1, -1, 1, -1, 1, 1, 1, -1, -1, 1, -1, 1, -1, -1, 1, -1, -1, 1, 1, -1, 1, -1, 1, 1, 1, 1, 1, 1, -1, -1, -1, 1, -1, 1, 1, 1, -1, -1, -1, -1, -1, 1, -1, 1, -1, 1, -1, -1, 1, -1, -1, 1, 1, -1, 1, -1, 1, 1, 1, 1, 1, -1, -1)$
$b = (1, 1, 1, -1, -1, -1, -1, 1, -1, 1, -1, -1, 1, 1, -1, 1, 1, -1, 1, -1, 1, -1, 1, 1, 1, 1, 1, 1, -1, -1, -1, 1, -1, 1, 1, 1, 1, -1, -1, -1, -1, -1, 1, -1, 1, -1, 1, -1, -1, 1, -1, -1, 1, 1, -1, 1, 1, 1, -1, -1, -1, -1, 1, -1, 1, 1, 1, -1, 1, 1, -1, 1, -1, -1, 1, -1, -1, 1, 1, 1, -1, 1, -1, 1, 1, 1, 1, 1, -1, -1)$

of length 61. In general, such pairs of length q exist whenever q is a power of a prime $p \equiv 5 \pmod 8$, see [2–4].

Regarding the Legendre C-pairs of symmetry type (ks) one can ask whether they exist for lengths v when $2v + 1$ is not a prime power? It turns out that they do. The smallest such pair known to us has

length $v = 1373$ in which case $2v + 1 = 2747 = 41 \times 67$. For this see [11, Theorem 3.1].

Characterization of Legendre C-pairs

Let us give an algebraic characterization of Legendre C-pairs over an abelian group G of order v with the identity element e. Let $a, b: G \to Z$ be functions such that $a(e) = 0$ and all other values of a and all the values of b are +1 or −1. Define the subsets $X, Y \subseteq G$ by

$$X = \{x \in G: a(x) = -1\} \text{ and } Y = \{x \in G: b(x) = -1\}. \quad (6)$$

Note that $e \notin X$ because $a(e) = 0$. Also note that the pair (X, Y) determines uniquely the pair (a, b). If (a, b) is a Legendre C-pair, by using Definition 1 and [6, equation (7)], it is straightforward to verify that

$$N(G - e - 2X) + N(G - 2Y) =$$
$$= (2v - 1)e - 2(G - e) = (2v + 1)e - 2G. \quad (7)$$

Proposition 3. Let a, b, X, Y be as above. (We shall view the subsets $X, Y \subseteq G$ also as elements of the group ring $Z[G]$.) If v is even then (a, b) is a Legendre C-pair if and only if

$$XX^* + YY^* = \frac{v}{2}(e + G) - X. \quad (8)$$

If v is odd then (a, b) is a Legendre C-pair if and only if

$$XX^* + YY^* = \frac{v+1}{2}e + \frac{v-3}{2}G. \quad (9)$$

Proof: First, let v be even and assume that (a, b) is a Legendre C-pair. By Proposition 1 we have $X^* = X$ and $k_1 = k_2 = v/2$. Since $GG = vG$, $GX = k_1G$ and $GY = k_2G$, it is easy to verify that $N(G - e - 2X) = 4N(X) + 4X + e - (v + 2)G$ and $N(G - 2Y) = 4N(Y) - vG$. The sum of these two norms is equal to $4(N(X) + N(Y)) + 4X + e - 2(v + 1)G$. By comparing this expression with the right hand side of (7) we obtain (8). For the converse note that (8) implies that $X^* = X$ and so we can reverse the above arguments.

Second, let v be odd and assume that (a, b) is a Legendre C-pair. By Proposition 1 we have $X^* + X = G - e$ and $k_1 = k_2 = (v - 1)/2$. Since $GG = vG$, $GX = k_1G$ and $GY = k_2G$, it is easy to verify that $N(G - e - 2X) = 4N(X) - (v - 2)G - e$ and $N(G - 2Y) = 4N(Y) - (v - 2)G$. The sum of these two norms is equal to $4(N(X) + N(Y)) - (v - 2)G - e - (v - 2)G$. By comparing this expression with the right hand side of (7) we obtain (9).

The following corollary follows immediately from the second claim of Proposition 3.

Corollary 3. In the case when v is odd, the functions (a, b) form a Legendre C-pair if and only if the corresponding subsets X, Y form a Szekeres DF in G.

Note that if (a, b) in this corollary is a special Legendre C-pair then the corresponding Szekeres DF (X, Y) is also special.

Example 2. Let us verify the equation (9) for the sequences $a = (0 - + - +)$, $b = (- + + + -)$. In these sequences, + and − stand for +1 and −1 respectively. They form a special Legendre C-pair of length $v = 5$. By using (6) we find that $X = \{g^1, g^3\} = g + g^3$ and $Y = \{g^0, g^4\} = e + g^4$. Thus $Y^* = e + g$ and so Y is not symmetric. However its translate $Yg^{-2} = g^2 + g^3$ is symmetric. Further, $XX^* = 2e + g^2 + g^3$ and $YY^* = 2e + g + g^4$. Thus $XX^* + YY^* = 4e + g + g^2 + g^3 + g^4 = 3e + G$.

Example 3. Let us verify the equation (8) for the sequences $a = (0 + - - - +)$, $b = (- + + - - +)$. They form a special Legendre C-pair of length $v = 6$. By using (6) we find that $X = \{g^2, g^3, g^4\} = g^2 + g^3 + g^4$ and $Y = \{g^0, g^3, g^4\} = e + g^3 + g^4$. Further, we have $XX^* = 3e + 2g + g^2 + g^4 + 2g^5$ and $YY^* = 3e + g + g^2 + 2g^3 + g^4 + g^5$. Thus $XX^* + YY^* = 6e + 3g + 2(g^2 + g^3 + g^4) + 3g^5 = 3(e + G) - X$.

Algorithm for constructing negacyclic C-matrices

As stated earlier, all Paley C-matrices of the same order $n = 1 + q$ are equivalent to each other. In view of Proposition 2, the equivalence class of any Paley C-matrix contains a C-matrix of 2b2c-type. It is also well known that the same equivalence class also contains a negacyclic C-matrix, see [12, Corollary 7.2].

Let q be any odd prime power. We describe a simple algorithm which for any given q outputs a negacyclic C-matrix $\mathbf{C}$ of order $n = 1 + q$ equivalent to the Paley C-matrix of the same order. This algorithm is based on the proof of [12, Corollary 7.2]. Since $\mathbf{C}$ is negacyclic it suffices to find its first row $[c_0 = 0, c_1, c_2, ..., c_q]$.

By a theorem of Belevitch, see [12, Theorem 4.1], we have $c_{n/2+j} = (-1)^j c_{n/2-j}$ for $j = 1, 2, ..., n/2 - 1$. Thus it suffices to compute only the values of c_i for $i = 1, 2, ..., n/2$.

We assume that a suitable software for computations in finite fields is available. (One of the authors used Maple and its GF package.)

Step 1. Construct the finite field GF(q^2) and select any primitive element ε of that field.

Step 2. Construct the matrix $\mathbf{A}$ of order 2 with first row $[0, -\omega]$ and second row $[1, \tau]$ where $\omega = \varepsilon^{1+q}$ and $\tau = \varepsilon + \varepsilon^q$. Note that ω and τ belong to the subfield GF(q) while ε does not. In fact ω is a primitive element of GF(q).

Step 3. Set $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and define the vectors $x_i = \begin{pmatrix} x_i[1] \\ x_i[2] \end{pmatrix}$, $i = 1, 2, ..., n/2$ recursively by the formula $x_{i+1} = \mathbf{A}x_i$.

Step 4. Then we have $c_i = \chi(x_i[2])$ for $i = 1, 2, ..., n/2$ where χ is the quadratic character of GF(q).

This completes the description of the algorithm.

We remark that in the case $q \equiv 1 \pmod 4$ the Paley C-matrix of order $1 + q$ is also equivalent to a C-matrix of 2c-type, i. e., a C-matrix made up from two circulants as in (2) or (3) but without any border.

Example 4. We choose $q = 9$ and for a primitive polynomial $f(x)$ of degree 4 over the field GF(3) = $= Z_3$ we choose $f(x) = x^4 - x - 1$. Then GF(81) = $= Z_3[x]/(x^4 - x - 1)$, a quotient ring of $Z_3[x]$ mod the ideal (f). Denote by ε the image of x in GF(81). Then we have $\varepsilon^4 = 1 + \varepsilon$. Thus $\varepsilon^8 = 1 - \varepsilon + \varepsilon^2$, and we obtain that $\omega = \varepsilon^{10} = 1 + \varepsilon + \varepsilon^2 - \varepsilon^3$. A further computation shows that $\omega^2 = 1 - \omega$ and $\omega^3 = -1 - \omega$. As $\omega^4 = -1$ we have $\omega^5 = -\omega$, $\omega^6 = -\omega^2$ and $\omega^7 = -\omega^3$. The subfield GF(9) is given by GF(9) = $\{0, \pm 1, \pm\omega, \pm 1 \pm \omega\}$. The matrix $\mathbf{A}$ has rows $[0, -\omega]$ and $[1, \omega^2]$. The first row c of $\mathbf{C}$ has the form $c = [0, c_1, c_2, c_3, c_4, c_5, -c_4, c_3, -c_2, c_1]$. The vectors x_i are

$$x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad x_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} -\omega \\ \omega^2 \end{pmatrix}, \quad x_3 = \begin{pmatrix} -\omega^3 \\ -1-\omega \end{pmatrix},$$

$$x_4 = \begin{pmatrix} \omega + \omega^2 \\ \omega^3 - \omega^2 \end{pmatrix}, \quad x_5 = \begin{pmatrix} 1 - \omega^3 \\ 1 + \omega^2 \end{pmatrix}.$$

Finally, we compute the c_i: $c_1 = \chi(1) = 1$, $c_2 = \chi(\omega^2) = = 1$, $c_3 = \chi(-1 - \omega) = \chi(\omega^3) = -1$, $c_4 = \chi(\omega^3 - \omega^2) = \chi(1) = 1$ and $c_5 = \chi(1 + \omega^2) = \chi(\omega^3) = -1$. Thus $c = [0, 1, 1, -1, 1, -1, -1, -1, -1, 1]$.

Special Legendre C-pairs of even length

We list below the special Legendre C-pairs of even length $v < 70$, which give 2b2c-type symmetric C-matrices of order $2v + 2$. The pairs are specified by the subsets X and Y (see Proposition 3). For $v = 2, 4, 6$ we show the sequences a and b as well. In all cases $2v + 1$ is a power of a prime.

$v = 2$
[1], [0] $a = (0\ -)$, $b = (-\ +)$
$v = 4$
[1, 3], [0, 1] $a = (0\ -\ +\ -)$, $b = (-\ -\ +\ +)$
$v = 6$
[2, 3, 4], [0, 3, 4] $a = (0\ +\ -\ -\ -\ +)$, $b = (-\ +\ +\ -\ -\ +)$
$v = 8$
[2, 3, 5, 6], [0, 4, 5, 6]
$v = 12$
[1, 4, 5, 7, 8, 11], [0, 2, 3, 4, 5, 10]
$v = 14$
[3, 4, 5, 7, 9, 10, 11], [0, 1, 2, 4, 5, 7, 10]
$v = 18$
[1, 2, 4, 5, 9, 13, 14, 16, 17], [0, 1, 2, 3, 6, 8, 10, 12, 13]
$v = 20$
[2, 6, 7, 8, 9, 11, 12, 13, 14, 18], [0, 4, 5, 7, 8, 10, 13, 16, 17, 18]
$v = 24$
[2, 3, 5, 9, 10, 11, 13, 14, 15, 19, 21, 22], [0, 2, 3, 9, 12, 13, 15, 16, 17, 18, 19, 22]
$v = 26$
[1, 6, 8, 9, 10, 12, 13, 14, 16, 17, 18, 20, 25], [0, 1, 6, 7, 9, 10, 12, 14, 17, 20, 21, 22, 23]
$v = 30$
[1, 3, 8, 9, 11, 12, 14, 15, 16, 18, 19, 21, 22, 27, 29], [0, 6, 7, 11, 12, 14, 16, 19, 20, 21, 24, 25, 26, 27, 28]
$v = 36$
[2, 7, 9, 10, 11, 12, 13, 16, 17, 19, 20, 23, 24, 25, 26, 27, 29, 34], [0, 3, 5, 6, 10, 11, 12, 14, 15, 16, 18, 22, 26, 27, 28, 31, 33, 34]
$v = 40$
[1, 2, 3, 6, 8, 9, 13, 14, 16, 18, 22, 24, 26, 27, 31, 32, 34, 37, 38, 39], [0, 1, 3, 4, 6, 10, 11, 12, 13, 14, 15, 17, 18, 19, 23, 30, 31, 32, 34, 37]
$v = 44$
[1, 2, 3, 8, 11, 12, 13, 16, 18, 19, 20, 24, 25, 26, 28, 31, 32, 33, 36, 41, 42, 43], [0, 3, 5, 6, 9, 10, 12, 13, 14, 15, 16, 17, 19, 21, 23, 25, 32, 35, 36, 39, 41, 42]
$v = 48$
[3, 4, 9, 11, 12, 14, 16, 17, 18, 20, 21, 22, 26, 27, 28, 30, 31, 32, 34, 36, 37, 39, 44, 45], [0, 3, 4, 5, 8, 12, 14, 15, 16, 17, 18, 19, 21, 24, 25, 27, 34, 36, 37, 38, 40, 41, 45, 46]
$v = 50$
[2, 3, 8, 10, 11, 13, 17, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 33, 37, 39, 40, 42, 47, 48], [0, 2, 5, 9, 10, 11, 13, 14, 17, 19, 20, 21, 25, 26, 27, 31, 33, 34, 37, 41, 42, 43, 45, 46, 48]
$v = 54$
[2, 3, 4, 5, 13, 15, 16, 18, 19, 22, 23, 24, 26, 27, 28, 30, 31, 32, 35, 36, 38, 39, 41, 49, 50, 51, 52], [0, 2, 3, 5, 9, 12, 16, 17, 18, 22, 24, 27, 28, 30, 32, 33, 34, 38, 39, 40, 42, 43, 45, 46, 47, 49, 52]
$v = 56$
[3, 4, 5, 7, 9, 15, 16, 18, 20, 21, 23, 24, 25, 26, 30, 31, 32, 33, 35, 36, 38, 40, 41, 47, 49, 51, 52, 53], [0, 1, 4, 5, 6, 7, 9, 13, 19, 20, 23, 26, 28, 30, 31, 33, 34, 37, 38, 39, 40, 41, 43, 44, 45, 47, 52, 53]
$v = 60$
[1, 2, 4, 5, 7, 9, 11, 14, 15, 16, 17, 21, 22, 25, 26, 34, 35, 38, 39, 43, 44, 45, 46, 49, 51, 53, 55, 56, 58, 59], [0, 3, 4, 5, 6, 7, 9, 11, 12, 13, 14, 17, 22, 24, 25, 26, 28, 29, 32, 36, 38, 39, 40, 41, 43, 44, 49, 51, 57, 58]

$v = 62$
[3, 4, 8, 13, 14, 15, 16, 18, 19, 20, 21, 22, 24, 28, 29, 31, 33, 34, 38, 40, 41, 42, 43, 44, 46, 47, 48, 49, 54, 58, 59], [0, 2, 4, 5, 6, 9, 12, 13, 19, 21, 22, 24, 27, 31, 32, 33, 35, 36, 38, 41, 43, 44, 45, 46, 47, 50, 51, 53, 54, 58, 60]
$v = 68$
[1, 3, 4, 5, 7, 9, 10, 11, 14, 15, 18, 20, 23, 30, 31, 32, 33, 35, 36, 37, 38, 45, 48, 50, 53, 54, 57, 58, 59, 61, 63, 64, 65, 67], [0, 4, 5, 6, 7, 9, 11, 12, 13, 14, 16, 17, 22, 23, 25, 26, 28, 29, 33, 35, 36, 37, 40, 43, 46, 47, 48, 49, 52, 57, 59, 64, 65, 66]

Special Legendre C-pairs of odd length

For the sake of completeness we list here the special Szekeres DFs in cyclic groups Z_v for odd lengths $v < 70$ whenever $2v + 1$ is a prime power. In the first three cases we also give the corresponding special Legendre C-pairs. For odd v in this range, when $2v + 1$ is not a prime power, we were not able to find any special Szekeres DFs. We point out that the diagram in Fig. 2 shows (the case $q = 25 \equiv 1 \pmod 4$ and $p = 5$) that there exist Szekeres DFs in EA(25) of symmetry type (*kk*).

$v = 1$
[], [] $a = (0)$, $b = (+)$
$v = 3$
[1], [0] $a = (0 - +)$, $b = (- + +)$
$v = 5$
[3, 4], [1, 4] $a = (0 + + - -)$, $b = (+ \ - + + -)$
$v = 9$
[1, 2, 3, 5], [1, 4, 5, 8];
$v = 11$
[2, 3, 4, 6, 10], [0, 2, 3, 8, 9]
$v = 13$
[4, 7, 8, 10, 11, 12], [1, 3, 4, 9, 10, 12]
$v = 15$
[3, 5, 8, 9, 11, 13, 14], [0, 1, 2, 6, 9, 13, 14]
$v = 21$
[2, 4, 8, 11, 12, 14, 15, 16, 18, 20], [2, 3, 4, 5, 8, 13, 16, 17, 18, 19]
$v = 23$
[1, 2, 4, 5, 6, 7, 9, 12, 13, 15, 20], [0, 1, 2, 6, 8, 11, 12, 15, 17, 21, 22]
$v = 29$
[1, 7, 8, 10, 12, 15, 16, 18, 20, 23, 24, 25, 26, 27], [2, 3, 8, 10, 11, 12, 14, 15, 17, 18, 19, 21, 26, 27]
$v = 33$
[2, 3, 4, 6, 7, 8, 9, 10, 12, 17, 18, 19, 20, 22, 28, 32], [2, 3, 4, 6, 9, 10, 13, 15, 18, 20, 23, 24, 27, 29, 30, 31]
$v = 35$
[1, 3, 7, 8, 9, 10, 11, 13, 14, 15, 18, 19, 23, 29, 30, 31, 33],
[0, 2, 6, 7, 9, 12, 15, 16, 17, 18, 19, 20, 23, 26, 28, 29, 33]
$v = 39$
[1, 3, 4, 5, 6, 10, 13, 14, 16, 20, 21, 22, 24, 27, 28, 30, 31, 32, 37],
[0, 3, 5, 9, 10, 14, 16, 17, 18, 19, 20, 21, 22, 23, 25, 29, 30, 34, 36]
$v = 41$
[2, 3, 4, 5, 6, 8, 9, 11, 12, 16, 17, 19, 21, 23, 26, 27, 28, 31, 34, 40],
[2, 3, 4, 6, 7, 8, 10, 15, 17, 18, 23, 24, 26, 31, 33, 34, 35, 37, 38, 39]
$v = 51$
[5, 9, 11, 12, 13, 16, 17, 19, 21, 22, 26, 27, 28, 31, 33, 36, 37, 41, 43, 44, 45, 47, 48, 49, 50], [0, 1, 2, 3, 4, 7, 10, 11, 13, 15, 20, 22, 23, 28, 29, 31, 36, 38, 40, 41, 44, 47, 48, 49, 50]
$v = 53$
[1, 5, 8, 10, 12, 13, 14, 15, 16, 18, 21, 22, 23, 24, 27, 28, 33, 34, 36, 42, 44, 46, 47, 49, 50, 51], [2, 3, 5, 8, 12, 14, 18, 19, 20, 21, 22, 23, 26, 27, 30, 31, 32, 33, 34, 35, 39, 41, 45, 48, 50, 51]
$v = 63$
[1, 2, 5, 6, 12, 17, 18, 19, 21, 25, 27, 29, 30, 32, 35, 37, 39, 40, 41, 43, 47, 48, 49, 50, 52, 53, 54, 55, 56, 59, 60], [0, 1, 4, 6, 7, 8, 10, 11, 13, 14, 22, 23, 25, 27, 30, 31, 32, 33, 36, 38, 40, 41, 49, 50, 52, 53, 55, 56, 57, 59, 62]
$v = 65$
[2, 5, 6, 7, 9, 10, 13, 14, 15, 16, 17, 19, 20, 22, 28, 29, 31, 33, 35, 38, 39, 40, 41, 42, 44, 47, 53, 54, 57, 61, 62, 64], [4, 5, 7, 10, 12, 14, 18, 19, 20, 22, 23, 24, 25, 26, 30, 31, 34, 35, 39, 40, 41, 42, 43, 45, 46, 47, 51, 53, 55, 58, 60, 61]
$v = 69$
[1, 5, 8, 10, 11, 12, 13, 16, 17, 18, 19, 22, 24, 25, 26, 27, 28, 29, 31, 33, 35, 37, 39, 46, 48, 49, 54, 55, 60, 62, 63, 65, 66, 67], [2, 3, 5, 7, 8, 11, 12, 13, 15, 20, 21, 24, 27, 31, 32, 33, 34, 35, 36, 37, 38, 42, 45, 48, 49, 54, 56, 57, 58, 61, 62, 64, 66, 67]

Acknowledgements

The research of the first author leading to these results has received funding from the Ministry of Education and Science of the Russian Federation according to the project part of the state funding assignment No 2.2200.2017/4.6. The research of the second author was enabled in part by support provided by SHARCNET (http://www.sharcnet.ca) and Compute Canada (http://www.computecanada.ca).

References

1. Seberry J. *Orthogonal Designs: Hadamard Matrices, Quadratic Forms and Algebras*. Chapter 4. Springer, 2017. doi:10.1007/978-3-319-59032-5

2. Szekeres G. Tournaments and Hadamard matrices. *L'Enseignement Math*, 1969, vol. 15, pp. 269–278.
3. Szekeres G. Cyclotomy and complementary difference sets. *Acta Arithmetica*, 1971, vol. 18, pp. 349–353. doi:10.4064/aa-18-1-349-353
4. Whiteman A. L. An infinite family of skew Hadamard matrices. *Pacific Journal of Mathematics*, 1971, vol. 38, no. 3, pp. 817–822.
5. Blat D. and Szekeres G. A skew Hadamard matrix of order 52. *Canad. J. Math*., 1969, vol. 21, pp. 1319–1322.
6. Đoković D. Ž., Kotsireas I. S. Computational methods for difference families in finite abelian groups. *Spec. Matrices*, 2019, vol. 7, pp. 127–141. doi:10.1515/spma-2019-0012
7. Goethals J.-M., Seidel J. J. Orthogonal matrices with zero diagonal. *Canad. J. Math.,* 1967, vol. 19, pp. 1001–1010. doi:10.4153/CJM-1967-091-8
8. van Lint J. H., Wilson R. M. *A Course in Combinatorics*. Cambridge University Press, 1992. 530 p.
9. Balonin N. A., Seberry J. A review and new symmetric conference matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2014, no. 4, pp. 2–7.
10. Fletcher R. J., Gysin M., Seberry J. Application of the discrete Fourier transform to the search for generalised Legendre pairs and Hadamard matrices. *Australas. J. Combin*, 2001, vol. 23, pp. 75–86.
11. Ding C. Two constructions of $(v, (v-1)/2, (v-3)/2)$ difference families. *J. Combin. Designs*, 2008, vol. 16, no. 2, pp. 164–171. doi:10.1002/jcd.20159
12. Delsarte P., Goethals J.-M., Seidel J. J. Orthogonal matrices with zero diagonal II. *Canad. J. Math*., 1971, vol. 23, no. 5, pp. 816–832.

УДК 004.438
doi:10.31799/1684-8853-2020-4-2-10

Конференц-матрицы на основе С-пар Лежандра

Н. А. Балонин[a], доктор техн. наук, профессор, orcid.org/0000-0001-7338-4920, korbendfs@mail.ru
Д. Ж. Джокович[б], доктор наук, профессор, orcid.org/0000-0002-0176-2395, djokovic@uwaterloo.ca
[a]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ
[б]Университет Ватерлоо, кафедра чистой математики и Институт квантовых вычислений, Ватерлоо, Онтарио, N2L 3G1, Канада

Введение: существует несколько методов построения симметричных С-матриц ввиду отсутствия для них универсальной структуры. Это ограничение принципиально, кроме того, в литературе неполно освещена структура С-матриц с парной каймой, что делает ее изучение особенно актуальным. **Цель:** детально описать бициклическую конструкцию с парной каймой и предложить концепцию С-пар Лежандра. **Результаты:** рассмотрены С-матрицы порядка $n = 2v + 2$ с парной каймой на основе адаптации так называемых обобщенных пар Лежандра нечетной длины v к более широкому случаю четных и нечетных значений v, что позволяет построить новые С-пары Лежандра в конечных абелевых группах G порядка v. Такая пара описывается двумя функциями a, b: $G \to Z$, значения которых равны +1 или −1, за исключением $a(e) = 0$, где e – единичный элемент группы G, через Z обозначено кольцо целых чисел. Для характеристики С-пар Лежандра введены два набора $X = \{x \in G: a(x) = -1\}$ и $Y = \{x \in G: b(x) = -1\}$ группы G. Показано, что $a(x^{-1}) = (-1)^v a(x)$ для всех x. Для нечетных значений v отмечено, что X и Y образуют разностное семейство, что неприменимо к четным порядкам. Это разностное семейство и разностное семейство Секереша — один и тот же класс, первоначально используемый для построения кососимметричных матриц Адамара. Введен подкласс специальных С-пар Лежандра и доказано, что они существуют для случаев, когда $2v + 1$ — степень простого числа. В последних двух разделах статьи приведены примеры специальных циклических С-пар Лежандра для размеров $v < 70$. **Практическая значимость:** С-матрицы широко используются в задачах помехоустойчивого кодирования, сжатия и маскирования видеоинформации. Программы для поиска конференц-матриц и библиотеки построенных матриц применяются в математической сети «mathscinet.ru» вместе с исполняемыми онлайн-алгоритмами.

Ключевые слова — конференц-матрицы, кососимметричные матрицы Адамара, периодические автокорреляционные функции, разностные семейства Секереша, обобщенные пары Лежандра, конструкции, телефония.

Для цитирования: Balonin N. A., Đoković D. Ž. Conference matrices from Legendre C-pairs. *Информационно-управляющие системы*, 2020, № 4, с. 2–10. doi:10.31799/1684-8853-2020-4-2-10
For citation: 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

UDC 681.5.015
doi:10.31799/1684-8853-2020-4-11-19

Articles

Research on the effectiveness of continuous-discrete cubature Kalman filter distributional-robust modifications

V. M. Chubich[a], *Dr. Sc., Tech., Professor, orcid.org/0000-0003-2006-0046, chubich@ami.nstu.ru*
S. O. Kulabukhova[a], *Master Student, orcid.org/0000-0002-5823-5641*
[a]*Novosibirsk State Technical University, 20, K. Marksa Pr., 630073, Novosibirsk, Russian Federation*

Introduction: *Usually there are some outliers (abnormal measurements) in observed data, and they can significantly affect the quality of the data processing. Many dynamic processes are described with stochastic nonlinear equations. Modern nonlinear filters that include the cubature Kalman filter, which deserves a special attention, cannot effectively process data containing abnormal measurements. One of the possible solutions to this problem is to use so-called robust methods that have good performance when one has to analyze data containing outliers. The paper deals with the common situations, when the considered process is actually continuous, but the observed data is taken discretely.* ***Purpose:*** *Identifying the most effective advanced robust modifications of the continuous-discrete cubature Kalman filter and giving the appropriate recommendations for their appliance.* ***Results:*** *Four modifications of the continuous-discrete cubature Kalman filter have been proposed based on the variational Bayesian and correntropy robust approaches to parameter estimation for stochastic processes. All the modifications have parameters with optimal values depending on both the selected mathematical model and the considered set of observations composing the sample. These values are determined numerically by minimizing the accumulated root mean square error on some grid. The research on the effectiveness of the proposed robust modifications has been carried out for the problem of tracking a space vehicle during its reentry into the atmosphere. The stochastic and the grouped outliers have been considered. Two most effective filters that have approximately equal qualities of estimation have been derived. The correntropy filter that has one configurable parameter can be recommended for practical using.* ***Practical relevance:*** *The identified most effective robust filter can be used for solving various applied problems related to the identification of stochastic nonlinear continuous-discrete systems.*

Keywords *— nonlinear filtering, cubature Kalman filter, outliers, maximum correntropy criterion, variational Bayesian estimation, stochastic continuous-discrete system.*

For citation: Chubich V. M., Kulabukhova S. O. Research on the effectiveness of continuous-discrete cubature Kalman filter distributional-robust modifications. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 11–19. doi:10.31799/1684-8853-2020-4-11-19

Introduction

In many fields of science and engineering, various applied problems related to statistical estimation of stochastic process parameters take place (see [1–3]). The optimal solution to such problems in the case of linear stochastic models can be obtained by applying the Kalman filter and its square-root modifications [4, 5]. Due to the fact that many real processes are described with nonlinear equations, nonlinear filtering methods are of particular importance.

Until the middle of the 1990s, the main method for solving nonlinear filtering problems was the extended Kalman filter [6, 7]. This filter is based on the linearization over the time domain of the state and measurement models conducted in the neighborhood of a specially defined nominal trajectory. In addition to the estimation accuracy loss, it has to be taken into account that even if there are no significant nonlinearities, the right-hand side of the state and the measurement equations may be determined with cumbersome analytical formulas. This leads to the problem of Jacobian matrices correct computing in the case of linearized models. Nowadays, the unscented (UKF) and the cubature Kalman filter (CKF) gain significant popularity as they don't have the mentioned drawbacks.

The UKF proposed in 1995 [8] and developed in [9–11] assumes using the optimal set of determined sigma points for approximating the first and the second moments of the system state vector. In each specific case, the quality of the results obtained with the UKF depends on the optimal choice of the filter parameter values.

The CKF used in this paper overcomes this drawback. For discrete systems, it was developed in 2009 [12], and after that it was modified for continuous-discrete systems in [13, 14]. To derive the CKF, the third-degree cubature rule was used for numerical approximation of a special type of multidimensional probability integrals [15, 16]. In practical terms, it is important that (unlike the UKF) the CKF has an algebraically equivalent square-root modification that provides computational robustness.

In practice, there are often the cases when observed data contain some abnormal measurements. This can be caused by failures occurring while gathering and transmitting the measurements. In mathematical terms, this can be interpreted as a

significant deviation of the actual distribution of measurement noise from the postulated one for the points corresponding to outliers. To overcome these difficulties, various robust methods can be applied (see [17–20]) including the variational Bayesian [21–23] and the correntropy approaches [24–26] used in this paper. These methods are particularly worthy to be highlighted.

This paper provides a comparative analysis of some advanced modern robust modifications of the continuous-discrete cubature Kalman filter (CD-CKF). The used modifications have been chosen on the basis of the results of studies performed in [27, 28] for linear nonstationary systems. Also the materials of some relevant publications on nonlinear robust filtering [29–32] have been taken into account. It should be emphasized that all the considered filters have been obtained by applying the corresponding discrete expressions to the continuous-discrete case. In addition, some robust modifications of the CD-CKF have been purposely derived from already known UKF modifications.

This paper is the first in the author's series of papers devoted to the construction of resistant to abnormal measurements and machine rounding errors nonlinear continuous-discrete filter. Only the first part of the specified task is considered.

Structural-probabilistic description of the model

Following [33], consider the state space model of a controlled and observed stochastic nonlinear continuous-discrete system

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{f}\left[\mathbf{x}(t),\, \mathbf{u}(t),\, t\right] + \mathbf{\Gamma}(t)\mathbf{w}(t),\ t \in [t_0,\, t_N];$$

$$\mathbf{y}(t_{k+1}) = \mathbf{h}\left[\mathbf{x}(t_{k+1}),\, \mathbf{u}(t_{k+1}),\, t_{k+1}\right] + \mathbf{v}(t_{k+1}),$$

$$k = \overline{0,\, N-1},$$

where $\mathbf{x}(\cdot)$ is the n-dimensional state vector; $\mathbf{u}(\cdot)$ is the r-dimensional predefined control vector; $\mathbf{w}(\cdot)$ is the p-dimensional process noise vector; $\mathbf{y}(\cdot)$ is the m-dimensional measurement vector; $\mathbf{v}(\cdot)$ is the m-dimensional measurement noise vector.

Suppose that

— the stochastic processes $\{\mathbf{w}(t),\ t \in [t_0,\, t_N]\}$ and $\{\mathbf{v}(t_{k+1}),\ k = 0,\ 1,\ \ldots,\ N-1\}$ are the white Gaussian noises, and

$$\mathrm{E}\left[\mathbf{w}(t)\right] = 0,\ \ \mathrm{E}\left[\mathbf{w}(t)\mathbf{w}^{\mathrm{T}}(\tau)\right] = \mathbf{Q}(t)\delta(t-\tau);$$

$$\mathrm{E}\left[\mathbf{v}(t_{k+1})\right] = 0,\ \ \mathrm{E}\left[\mathbf{v}(t_{k+1})\mathbf{v}^{\mathrm{T}}(t_{i+1})\right] = \mathbf{R}(t_{k+1})\delta_{ki},$$

$$\mathrm{E}\left[\mathbf{v}(t_{k+1})\mathbf{w}^{\mathrm{T}}(\tau)\right] = 0,\ k,\ i = \overline{0,\, N-1},\ \tau \in [t_0,\, t_N]$$

(here and further $\mathrm{E}[\cdot]$ is the mathematical mean value operator; $\delta(t-\tau)$ is the Dirac delta function; δ_{ki} is the Kronecker symbol);

— the initial state $\mathbf{x}(t_0)$ has the normal distribution with the parameters

$$\mathrm{E}\left[\mathbf{x}(t_0)\right] = \overline{\mathbf{x}}(t_0),$$

$$\mathrm{E}\left\{\left[\mathbf{x}(t_0) - \overline{\mathbf{x}}(t_0)\right]\left[\mathbf{x}(t_0) - \overline{\mathbf{x}}(t_0)\right]^{\mathrm{T}}\right\} = \mathbf{P}(t_0),$$

and it is uncorrelated with $\mathbf{w}(t)$ and $\mathbf{v}(t_{k+1})$;

— the covariance matrices of the process noise, the measurement noise and the initial state are known, $\mathbf{R}(t_{k+1})$ and $\mathbf{P}(t_0)$ are positive-definite matrixes;

— the observed data $\{\mathbf{y}(t_{k+1}),\ k = 0,\ 1,\ \ldots,\ N-1\}$ contain outliers.

CD-CKF

To construct the CKF, the third-degree Gaussian cubature rule [15, 16] is used to compute probability integrals of the following type

$$\mathrm{E}\left[\mathbf{g}(\mathbf{x})\right] = \int\limits_{\mathrm{R}^n} \mathbf{g}(\mathbf{x}) N\left(\mathbf{x} \,|\, \boldsymbol{\mu},\, \boldsymbol{\Sigma}\right)\mathrm{d}\mathbf{x} \approx$$

$$\approx \frac{1}{2n}\sum_{i=1}^{2n} \mathbf{g}\left(\boldsymbol{\mu} + \mathbf{L}_{\boldsymbol{\Sigma}}\boldsymbol{\xi}_i\right),$$

where $\mathbf{g}(\mathbf{x})$ is some vector function; $N(\mathbf{x} \,|\, \boldsymbol{\mu},\, \boldsymbol{\Sigma})$ is the probability density function of the n-dimensional normal distribution with the mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$; $\mathbf{L}_{\boldsymbol{\Sigma}}$ is the lower triangular Cholesky factor $\left(\mathbf{L}_{\boldsymbol{\Sigma}}\mathbf{L}_{\boldsymbol{\Sigma}}^{\mathrm{T}} = \boldsymbol{\Sigma}\right)$, the nodes of a cubature formula $\boldsymbol{\xi}_i$ are defined with the expression

$$\boldsymbol{\xi}_i = \begin{cases} \sqrt{n}\,\mathbf{e}_i,\ i = \overline{1,\, n}; \\ -\sqrt{n}\,\mathbf{e}_{i-n},\ i = \overline{n+1,\, 2n}, \end{cases}$$

where $\mathbf{e}_i = \left(0,\ \ldots,\ \underset{i-1}{0},\ \underset{i}{1},\ \underset{i+1}{0},\ \ldots,\ 0\right)^{\mathrm{T}}$.

Nowadays, two versions of the CD-CKF exist. The first one (see [13]) is based on the Itô-Taylor expansion of the 1.5 order used to discretize the corresponding stochastic differential equation and on applying the discrete formulas of the CKF from [12] to the obtained system. It should be noted, that preliminary discretization requires manual adjusting the optimal number of points of division for each specific problem. This leads to the lack of flexibility of the first CD-CKF version. In this work, the other version of the CD-CKF from [14] is used. This version is more accurate and actual in practical terms (see [34]), and it also doesn't have the mentioned drawback.

The CD-CKF algorithm

Step 1. Initialize the initial state and the covariance:

$$\hat{\mathbf{x}}(t_0 \mid t_0) = \overline{\mathbf{x}}(t_0),\ \mathbf{P}(t_0 \mid t_0) = \mathbf{P}(t_0).$$

Execute in a loop for $k = \overline{0,\ N-1}$:

Step 2. Obtain $\hat{\mathbf{x}}(t_{k+1} \mid t_k)$ and $\mathbf{P}(t_{k+1} \mid t_k)$ by solving the system of differential equations for $t \in [t_k, t_{k+1}]$:

$$\frac{\mathrm{d}\hat{\mathbf{x}}(t \mid t_k)}{\mathrm{d}t} = \frac{1}{2n}\sum_{i=1}^{2n} \mathbf{f}\left[\hat{\mathbf{x}}(t \mid t_k) + \mathbf{S}(t \mid t_k)\xi_i,\ \mathbf{u}(t),\ t\right];$$

$$\frac{\mathrm{d}\mathbf{P}(t \mid t_k)}{\mathrm{d}t} = \frac{1}{2n}\sum_{i=1}^{2n} \mathbf{f}\left[\hat{\mathbf{x}}(t \mid t_k) + \mathbf{S}(t \mid t_k)\xi_i,\ \mathbf{u}(t),\ t\right] \times$$
$$\times \xi_i^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}(t \mid t_k) + \frac{1}{2n}\sum_{i=1}^{2n}\mathbf{S}(t \mid t_k)\xi_i \times$$
$$\times \mathbf{f}^{\mathrm{T}}\left[\hat{\mathbf{x}}(t \mid t_k) + \mathbf{S}(t \mid t_k)\xi_i,\ \mathbf{u}(t),\ t\right] +$$
$$+ \boldsymbol{\Gamma}(t)\mathbf{Q}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(t),$$

where $\mathbf{S}(t \mid t_k) = \mathrm{Chol}[\mathbf{P}(t \mid t_k)]$ is the Cholesky factor for the matrix $\mathbf{P}(t \mid t_k)$.

Step 3. Calculate the cubature points and the propagated cubature points:

$$\mathbf{S}(t_{k+1} \mid t_k) = \mathrm{Chol}\left[\mathbf{P}(t_{k+1} \mid t_k)\right];$$
$$\boldsymbol{\chi}_i(t_{k+1} \mid t_k) = \hat{\mathbf{x}}(t_{k+1} \mid t_k) + \mathbf{S}(t_{k+1} \mid t_k)\xi_i,\ i = \overline{1,\ 2n};$$
$$\boldsymbol{\gamma}_i(t_{k+1} \mid t_k) = \mathbf{h}\left[\boldsymbol{\chi}_i(t_{k+1} \mid t_k),\ \mathbf{u}(t_{k+1}),\ t_{k+1}\right],\ i = \overline{1,\ 2n}.$$

Step 4. Find the extrapolated measurement estimate and the update vector:

$$\hat{\mathbf{y}}(t_{k+1} \mid t_k) = \frac{1}{2n}\sum_{i=1}^{2n}\boldsymbol{\gamma}_i(t_{k+1} \mid t_k);$$
$$\boldsymbol{\varepsilon}(t_{k+1}) = \mathbf{y}(t_{k+1}) - \hat{\mathbf{y}}(t_{k+1} \mid t_k).$$

Step 5. Obtain the cross-covariance matrix:

$$\mathbf{P}_{xy}(t_{k+1} \mid t_k) = \frac{1}{2n}\sum_{i=1}^{2n}\left[\boldsymbol{\chi}_i(t_{k+1} \mid t_k) - \hat{\mathbf{x}}(t_{k+1} \mid t_k)\right] \times$$
$$\times \left[\boldsymbol{\gamma}_i(t_{k+1} \mid t_k) - \hat{\mathbf{y}}(t_{k+1} \mid t_k)\right]^{\mathrm{T}}.$$

Step 6. Evaluate the covariance matrix of the prediction error:

$$\mathbf{P}_{yy}(t_{k+1} \mid t_k) = \frac{1}{2n}\sum_{i=1}^{2n}\left[\boldsymbol{\gamma}_i(t_{k+1} \mid t_k) - \hat{\mathbf{y}}(t_{k+1} \mid t_k)\right] \times$$
$$\times \left[\boldsymbol{\gamma}_i(t_{k+1} \mid t_k) - \hat{\mathbf{y}}(t_{k+1} \mid t_k)\right]^{\mathrm{T}} + \mathbf{R}(t_{k+1}). \quad (1)$$

Step 7. Calculate the Kalman gain factor:

$$\mathbf{K}(t_{k+1}) = \mathbf{P}_{xy}(t_{k+1} \mid t_k)\mathbf{P}_{yy}^{-1}(t_{k+1} \mid t_k). \quad (2)$$

Step 8. Obtain the filtered state estimate:

$$\hat{\mathbf{x}}(t_{k+1} \mid t_{k+1}) = \hat{\mathbf{x}}(t_{k+1} \mid t_k) + \mathbf{K}(t_{k+1})\boldsymbol{\varepsilon}(t_{k+1}). \quad (3)$$

Step 9. Find the corresponding error covariance matrix:

$$\mathbf{P}(t_{k+1} \mid t_{k+1}) = \mathbf{P}(t_{k+1} \mid t_k) - \mathbf{K}(t_{k+1}) \times$$
$$\times \mathbf{P}_{yy}(t_{k+1} \mid t_k)\mathbf{K}^{\mathrm{T}}(t_{k+1}). \quad (4)$$

End of the loop for k.

In order to make the further discussion brief, when describing the robust modifications of the CD-CKF we will consider this algorithm to be a basis, but will add and edit certain steps.

Robust modifications of the CD-CKF

First of all, consider the continuous-discrete variational Bayesian-based cubature Kalman filter (CD-VBCKF) proposed in [29] and derived based on [22]. This modification suggests using the following parameters: ν_0 (scalar), $\mathbf{V}_0$ (m-dimensional square matrix), L (number of iterations for the filtration step), ρ (scaling factor selected on the interval (0, 1]). The optimal values of the parameters are determined by selection.

The CD-VBCKF algorithm

Step 1. At the step 1 of the CD-CKF algorithm, the initialization of the following parameters is included:

$$v(t_0 \mid t_0) = v_0,\ \mathbf{V}(t_0 \mid t_0) = \mathbf{V}_0.$$

Execute in a loop for $k = \overline{0,\ N-1}$:

Step 2. At the step 2 of the CD-CKF algorithm, the parameter calculating is included

$$v(t_{k+1} \mid t_k) = \rho\left(v(t_k \mid t_k) - n - 1\right) + n + 1;$$
$$\mathbf{V}(t_{k+1} \mid t_k) = \rho\mathbf{V}(t_k \mid t_k).$$

Steps 3–5 are equivalent to the steps 3–5 of the CD-CKF algorithm.

Step 6. Obtain

$$\mathbf{V}^0(t_{k+1} \mid t_{k+1}) = \mathbf{V}(t_{k+1} \mid t_k);$$
$$v(t_{k+1} \mid t_{k+1}) = 1 + v(t_{k+1} \mid t_k).$$

Execute in a loop for $j = \overline{1,\ L}$:

Step 7. Obtain

$$\mathbf{R}^j(t_{k+1}) = \left(v(t_{k+1}\,|\,t_{k+1}) - n - 1\right)^{-1}\mathbf{V}^{j-1}(t_{k+1}\,|\,t_{k+1}).$$

Step 8. Calculate $\mathbf{P}^j_{yy}(t_{k+1}\,|\,t_k)$ and $\mathbf{K}^j(t_{k+1})$ using the formulas (1), (2), where the matrices $\mathbf{R}(t_{k+1})$ and $\mathbf{P}_{yy}(t_{k+1}\,|\,t_k)$ are replaced with $\mathbf{R}^j(t_{k+1})$ and $\mathbf{P}^j_{yy}(t_{k+1}\,|\,t_k)$ respectively.

Step 9. Calculate $\hat{\mathbf{x}}^j(t_{k+1}\,|\,t_{k+1})$ and $\mathbf{P}^j(t_{k+1}\,|\,t_{k+1})$ by replacing $\mathbf{K}(t_{k+1})$ and $\mathbf{P}_{yy}(t_{k+1}\,|\,t_k)$ in the formulas (3) and (4) with the matrices $\mathbf{K}^j(t_{k+1})$ and $\mathbf{P}^j_{yy}(t_{k+1}\,|\,t_k)$ respectively.

Step 10. Calculate

$$\mathbf{S}^j(t_{k+1}\,|\,t_{k+1}) = \mathrm{Chol}\left[\mathbf{P}^j(t_{k+1}\,|\,t_{k+1})\right].$$

$$\boldsymbol{\chi}^j_i(t_{k+1}\,|\,t_{k+1}) = \hat{\mathbf{x}}^j(t_{k+1}\,|\,t_{k+1}) + \mathbf{S}^j(t_{k+1}\,|\,t_{k+1})\boldsymbol{\xi}_i;$$

$$\boldsymbol{\gamma}^j_i(t_{k+1}\,|\,t_{k+1}) = \mathbf{h}\left[\boldsymbol{\chi}^j_i(t_{k+1}\,|\,t_{k+1}),\ \mathbf{u}(t_{k+1}),\ t_{k+1}\right],$$

$$i = \overline{1,\ 2n}.$$

Step 11. Obtain the matrix

$$\mathbf{V}^j(t_{k+1}\,|\,t_{k+1}) = \frac{1}{2n}\sum_{i=1}^{2n}\left[\mathbf{y}(t_{k+1}) - \boldsymbol{\gamma}^j_i(t_{k+1}\,|\,t_{k+1})\right]\times$$

$$\times\left[\mathbf{y}(t_{k+1}) - \boldsymbol{\gamma}^j_i(t_{k+1}\,|\,t_{k+1})\right]^{\mathrm{T}} + \mathbf{V}(t_{k+1}\,|\,t_k).$$

End of the loop for j.

Step 12. Obtain

$$\mathbf{V}(t_{k+1}\,|\,t_{k+1}) = \mathbf{V}^L(t_{k+1}\,|\,t_{k+1}).$$

Step 13. Obtain the filtering estimate and the corresponding covariance matrix:

$$\hat{\mathbf{x}}(t_{k+1}\,|\,t_{k+1}) = \hat{\mathbf{x}}^L(t_{k+1}\,|\,t_{k+1});$$

$$\mathbf{P}(t_{k+1}\,|\,t_{k+1}) = \mathbf{P}^L(t_{k+1}\,|\,t_{k+1}).$$

End of the loop for k.

Now consider the correntropy modifications of the CD-CKF, which have been intensively developed in the recent years. Initially, the correntropy filters were obtained for linear dynamic models (see [35–37]). Later, they were successfully adapted for solving nonlinear problems [30–32, 38–40].

The correntropy filters are constructed on the basis of the maximum correntropy criterion, and correntropy is considered to be a statistical measure of the similarity between two random variables. This measure takes into account the second and the higher-order moments. Technically, the correntropy between X, Y is determined with the formula [26]

$$C(X,\,Y) = \mathbf{E}_{XY}\left[\kappa(X,\,Y)\right] = \iint \kappa(x,\,y)f_{XY}(x,\,y)\mathrm{d}x\mathrm{d}y,$$

where $\kappa(\cdot,\cdot)$ is some continuous positive defined function (a kernel); $f_{XY}(\cdot,\cdot)$ is the joint density function of the random variables X and Y. Most commonly, the Gaussian kernel of size $\sigma > 0$ is used:

$$\kappa(x,\,y) = G_\sigma(x-y) = \exp\left(\frac{-(x-y)^2}{2\sigma^2}\right).$$

In practice, the distribution of f_{XY} is usually unknown, thus the correntropy estimate $\hat{C}(X,\,Y)$ is used instead of $C(X,\,Y)$

$$\hat{C}(X,\,Y) = \frac{1}{N}\sum_{i=1}^{N} G_\sigma(x_i - y_i).$$

Note that the kernel size significantly affects the quality of correntropy filters. There are no actual general recommendations for the optimal selection of the parameter σ value that depends on the considered sample. This is the bottleneck of all the correntropy filters. Some adaptive techniques for determining the kernel size based on the update vector are given in [26, 32, 41], but the problem has not been solved yet, because generally the results are better when the kernel size is optimal.

Consider three correntropy modifications of the CD-CKF based on different maximum correntropy criteria.

The first modification named the CD-MCCKF-1 (Continuous-discrete maximum correntropy cubature Kalman filter) contains scalar parameters σ and δ. This modification has been obtained by replacing the equations of the correntropy discrete UKF from [30] with the relevant CD-CKF formulas. The following algorithm corresponds to this modification.

The CD-MCCKF-1 algorithm

Step 1. Repeat the step 1 of the CD-CKF algorithm.

Execute in a loop for $k = \overline{0,\ N-1}$**:**

Steps 2–5 are equivalent to the steps 2–5 of the CD-CKF algorithm.

Step 6. Obtain the matrix $\mathbf{H}(t_{k+1})$**:**

$$\mathbf{H}(t_{k+1}) = \mathbf{P}^{\mathrm{T}}_{xy}(t_{k+1}\,|\,t_k)\mathbf{P}^{-1}(t_{k+1}\,|\,t_k). \qquad (5)$$

Step 7. Find $\overline{\mathbf{S}}(t_{k+1})$**:**

$$\mathbf{S}_R(t_{k+1}) = \mathrm{Chol}\left[\mathbf{R}(t_{k+1})\right];$$

$$\overline{\mathbf{S}}(t_{k+1}) = \mathrm{diag}\left[\mathbf{S}(t_{k+1}\,|\,t_k),\ \mathbf{S}_R(t_{k+1})\right].$$

Step 8. Obtain $\mathbf{D}(t_{k+1})$ and $\mathbf{W}(t_{k+1})$:

$$\mathbf{D}(t_{k+1}) = \overline{\mathbf{S}}^{-1}(t_{k+1})\begin{bmatrix} \hat{\mathbf{x}}(t_{k+1}\,|\,t_k) \\ \boldsymbol{\varepsilon}(t_{k+1}) + \mathbf{H}(t_{k+1})\hat{\mathbf{x}}(t_{k+1}\,|\,t_k) \end{bmatrix};$$

$$\mathbf{W}(t_{k+1}) = \overline{\mathbf{S}}^{-1}(t_{k+1})\begin{bmatrix} \mathbf{I}_n \\ \mathbf{H}(t_{k+1}) \end{bmatrix}.$$

Step 9. Set $i = 1$ and evaluate

$$\hat{\mathbf{x}}^0(t_{k+1}\,|\,t_{k+1}) = \left[\mathbf{W}^{\mathrm{T}}(t_{k+1})\mathbf{W}(t_{k+1})\right]^{-1} \times$$
$$\times\, \mathbf{W}^{\mathrm{T}}(t_{k+1})\mathbf{D}(t_{k+1}).$$

Execute in a loop for i:

Step 10. Calculate

$$\mathbf{e}^i(t_{k+1}) = \mathbf{D}(t_{k+1}) - \mathbf{W}(t_{k+1})\hat{\mathbf{x}}^{i-1}(t_{k+1}\,|\,t_{k+1});$$

$$\mathbf{C}_x(t_{k+1}) = \mathrm{diag}\left[G_\sigma\left(e_1^i(t_{k+1})\right), \ldots, G_\sigma\left(e_n^i(t_{k+1})\right)\right];$$

$$\mathbf{C}_y(t_{k+1}) = \mathrm{diag}\left[G_\sigma\left(e_{n+1}^i(t_{k+1})\right), \ldots, G_\sigma\left(e_{n+m}^i(t_{k+1})\right)\right];$$

$$\mathbf{P}^i(t_{k+1}\,|\,t_k) = \mathbf{S}(t_{k+1}\,|\,t_k)\mathbf{C}_x^{-1}(t_{k+1})\mathbf{S}^{\mathrm{T}}(t_{k+1}\,|\,t_k);$$

$$\mathbf{R}^i(t_{k+1}) = \mathbf{S}_R(t_{k+1})\mathbf{C}_y^{-1}(t_{k+1})\mathbf{S}_R^{\mathrm{T}}(t_{k+1});$$

$$\mathbf{B}^i(t_{k+1}) = \mathbf{H}(t_{k+1})\mathbf{P}^i(t_{k+1}\,|\,t_k)\mathbf{H}^{\mathrm{T}}(t_{k+1}) + \mathbf{R}^i(t_{k+1});$$

$$\mathbf{K}^i(t_{k+1}) = \mathbf{P}^i(t_{k+1}\,|\,t_k)\mathbf{H}^{\mathrm{T}}(t_{k+1})\left[\mathbf{B}^i(t_{k+1})\right]^{-1};$$

$$\hat{\mathbf{x}}^i(t_{k+1}\,|\,t_{k+1}) = \hat{\mathbf{x}}(t_{k+1}\,|\,t_k) + \mathbf{K}^i(t_{k+1})\boldsymbol{\varepsilon}(t_{k+1}).$$

If $\dfrac{\left\|\hat{\mathbf{x}}^i(t_{k+1}\,|\,t_{k+1}) - \hat{\mathbf{x}}^{i-1}(t_{k+1}\,|\,t_{k+1})\right\|}{\left\|\hat{\mathbf{x}}^{i-1}(t_{k+1}\,|\,t_{k+1})\right\|} \le \delta$, then the end of the loop for i, else set $i = i + 1$.

Step 11. Define the filtering estimate and the corresponding error covariance matrix:

$$\hat{\mathbf{x}}(t_{k+1}\,|\,t_{k+1}) = \hat{\mathbf{x}}^i(t_{k+1}\,|\,t_{k+1});$$

$$\mathbf{P}(t_{k+1}\,|\,t_{k+1}) = \left[\mathbf{I}_n - \mathbf{K}^i(t_{k+1})\mathbf{H}(t_{k+1})\right]\mathbf{P}^i(t_{k+1}\,|\,t_k).$$

End of the loop for k.

The next algorithm corresponds to the second modification named the CD-MCCKF-2 and obtained by applying the CKF equations from [31] to the continuous-discrete case. This modification contains a scalar parameter σ.

The CD-MCCKF-2 algorithm

Step 1. Repeat step 1 of the CD-CKF algorithm.

Execute in a loop for $k = \overline{0,\ N-1}$:

Steps 2–5 are equivalent to the steps 2–5 of the CD-CKF algorithm.

Step 6. Calculate the measurement noise covariance matrix estimate:

$$\mathbf{S}_R(t_{k+1}) = \mathrm{Chol}\left[\mathbf{R}(t_{k+1})\right];$$

$$\mathbf{e}(t_{k+1}) = \mathbf{S}_R^{-1}(t_{k+1}) \times$$
$$\times\left[\mathbf{y}(t_{k+1}) - \mathbf{h}\left[\hat{\mathbf{x}}(t_{k+1}\,|\,t_k),\ \mathbf{u}(t_{k+1}),\ t_{k+1}\right]\right];$$

$$\mathbf{C}_y(t_{k+1}) = \mathrm{diag}\left[G_\sigma(e_1(t_{k+1})),\ \ldots,\ G_\sigma(e_m(t_{k+1}))\right];$$

$$\hat{\mathbf{R}}(t_{k+1}) = \mathbf{S}_R(t_{k+1})\mathbf{C}_y^{-1}(t_{k+1})\mathbf{S}_R^T(t_{k+1}).$$

Step 7. Define $\mathbf{P}_{yy}(t_{k+1}\,|\,t_k)$ with the expression (1) replacing $\mathbf{R}(t_{k+1})$ with $\hat{\mathbf{R}}(t_{k+1})$.

Steps 8–10. Carry out the steps 7–9 of the CD-CKF algorithm.

End of the loop for k.

The last algorithm corresponds to the third modification named the CD-MCCKF-3 and obtained by replacing the correntropy discrete UKF formulas from [32] with the relevant CD-CKF equations. This modification contains a scalar parameter σ.

The CD-MCCKF-3 algorithm

Step 1. Repeat the step 1 of the CD-CKF algorithm.

Execute in a loop for $k = \overline{0,\ N-1}$:

Steps 2–6 are equivalent to the steps 2–6 of the CD-CKF algorithm.

Step 7. Calculate the matrix $\mathbf{H}(t_{k+1})$ using the expression (5).

Step 8. Calculate the measurement noise covariance matrix estimate:

$$\hat{\mathbf{R}}(t_{k+1}) = \mathbf{P}_{yy}(t_{k+1}\,|\,t_k) - \mathbf{H}(t_{k+1})\mathbf{P}(t_{k+1}\,|\,t_k)\mathbf{H}^T(t_{k+1}).$$

Step 9. Find the scalar value $L(t_{k+1})$:

$$L(t_{k+1}) = G_\sigma\left(\left[\boldsymbol{\varepsilon}^T(t_{k+1})\hat{\mathbf{R}}^{-1}(t_{k+1})\boldsymbol{\varepsilon}(t_{k+1})\right]^{1/2}\right).$$

Step 10. Calculate the Kalman gain factor $\mathbf{K}(t_{k+1})$:

$$\mathbf{B}(t_{k+1}) = \hat{\mathbf{R}}(t_{k+1}) +$$
$$+L(t_{k+1})\mathbf{H}(t_{k+1})\mathbf{P}(t_{k+1}\,|\,t_k)\mathbf{H}^{\mathrm{T}}(t_{k+1});$$

$$\mathbf{K}(t_{k+1}) = L(t_{k+1})\mathbf{P}(t_{k+1}\,|\,t_k)\mathbf{H}^{\mathrm{T}}(t_{k+1})\mathbf{B}^{-1}(t_{k+1}).$$

Step 11. Obtain the filtering estimate by repeating the step 8 of the CD-CKF algorithm.

Step 12. Calculate the error covariance matrix:

$$\mathbf{P}(t_{k+1}\,|\,t_{k+1}) = \left[\mathbf{I}_n - \mathbf{K}(t_{k+1})\mathbf{H}(t_{k+1})\right]\mathbf{P}(t_{k+1}\,|\,t_k).$$

End of the loop for k.

Comparison of the CD-CKF robust modifications

This section presents the comparison of the considered CD-CKF modifications effectiveness. The analysis has been made for the problem of tracking a space vehicle, entering the atmosphere, given in [10, 42]. In this case the state and the measurement equations in the planar Earth-centered Cartesian coordinate system are defined as follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \\ x_5(t) \end{bmatrix} = \begin{bmatrix} x_3(t) \\ x_4(t) \\ D(t)x_3(t)+G(t)x_1(t)+w_1(t) \\ D(t)x_4(t)+G(t)x_2(t)+w_2(t) \\ w_3(t) \end{bmatrix},$$

$$t \in [t_0,\ t_N],$$

$$\begin{bmatrix} y_1(t_{k+1}) \\ y_2(t_{k+1}) \end{bmatrix} = \begin{bmatrix} r(t_{k+1})+v_1(t_{k+1}) \\ \theta(t_{k+1})+v_2(t_{k+1}) \end{bmatrix},\ k=\overline{0,\ N-1}.$$

Here $x_1(t)$, $x_2(t)$ are the space vehicle coordinates; $x_3(t)$, $x_4(t)$ are the corresponding coordinate velocities; $x_5(t)$ is the parameter of the vehicle aerodynamic properties; $y_1(t_{k+1})$ is the distance to the radar; $y_2(t_{k+1})$ is the angle between the space vehicle and the horizontal axis;

$$D(t) = b(t)\exp\left(\frac{6374-R(t)}{13.406}\right)V(t);$$

$$b(t) = -0.59783\exp(x_5(t));$$

$$R(t) = \sqrt{x_1^2(t)+x_2^2(t)};$$

$$V(t) = \sqrt{x_3^2(t)+x_4^2(t)};$$

$$G(t) = -\frac{398600}{R^3(t)};$$

$$r(t_{k+1}) = \sqrt{(x_1(t_{k+1})-6374)^2+x_2^2(t_{k+1})};$$

$$\theta(t_{k+1}) = \mathrm{arctg}\left(\frac{x_2(t_{k+1})}{x_1(t_{k+1})-6374}\right).$$

Let $t_0 = 0$, $N = 150$, $t_{k+1} = 0.1(k+1)$ and assume that all priori assumptions made in the "Structural-probabilistic description of the model" section are valid, and the statistical characteristics of the noise and the initial state are defined as follows

$$\mathbf{Q} = \mathrm{diag}\left[2.4064\times10^{-4},\ 2.4064\times10^{-4},\ 0\right],$$

$$\mathbf{R} = \mathrm{diag}\left[1,\ (0.017)^2\right],$$

$$\bar{\mathbf{x}}(t_0) = \begin{bmatrix} 6500.4 \\ 349.14 \\ -1.8093 \\ -6.7967 \\ 0.6932 \end{bmatrix},$$

$$\mathbf{P}(t_0) = \mathrm{diag}\left[10^{-6},\ 10^{-6},\ 10^{-6},\ 10^{-6},\ 1\right].$$

It is worth to be mentioned that all the considered robust modifications of the CD-CKF involve using certain parameters. The optimal values of these parameters (except for $\delta_0 = 10^{-8}$ in the CD-MCCKF-1) should be found individually for each run by minimizing the accumulated root mean square error (ARMSE) on some grid.

Using the presented CD-CKF modifications, we have processed data with the *stochastic* ordering of outliers. The data have been modeled in such a way they have 20 percent noise and the outliers noise variance equal to $\mathbf{R}_A = 10000\mathbf{R}$. Data modeling was performed using software developed in the Matlab system under the assumption that outliers are uniformly distributed over the entire modeling interval at the same time points for both observation components.

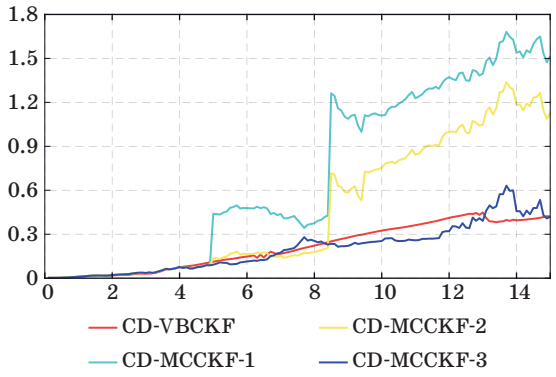
Figure 1 shows the dependence of the root mean square error (RMSE) on time.

Root mean square error value for each time point can be calculated using the equation

$$\mathrm{RMSE}(t_{k+1}) = \sqrt{\sum_{i=1}^{n}\left(x_i(t_{k+1})-\hat{x}_i(t_{k+1}\,|\,t_{k+1})\right)^2},$$

$$k=\overline{0,\ N-1}.$$

In order to reduce the impact of the observed data on the estimation results, we have modeled $M = 100$ different samples. The filtering quality has been estimated based on the ARMSE value defined in accordance with the formula [33]



■ *Fig. 1.* The RMSE values for the stochastic outliers

■ ***Table 1.*** The values of the accumulated root mean square errors for the stochastic outliers

Filters	$ARMSE_1$	$ARMSE_2$	$ARMSE_3$	$ARMSE_4$	$ARMSE_5$	ARMSE
CD-VBCKF	0.191	0.161	0.039	0.027	1.114	1.143
CD-MCCKF-1	0.674	0.795	0.094	0.109	1.245	1.630
CD-MCCKF-2	0.579	0.579	0.082	0.082	1.211	1.466
CD-MCCKF-3	0.214	0.150	0.041	0.029	1.112	1.144

$$\mathrm{ARMSE} = \sqrt{\sum_{i=1}^{n} \mathrm{ARMSE}_i^2},$$

where

$$\mathrm{ARMSE}_i = \sqrt{\frac{1}{MN} \sum_{j=1}^{M} \sum_{k=0}^{N-1} \left(x_i^j(t_{k+1}) - \hat{x}_i^j(t_{k+1} \mid t_{k+1}) \right)^2};$$

$x_i^j(t_{k+1})$ and $\hat{x}_i^j(t_{k+1} \mid t_{k+1})$ are the i-th components of the state vector and its filtering estimate for the j-th run.

The values of the accumulated root mean square errors for the various robust CD-CKF modifications are shown in the Table 1.

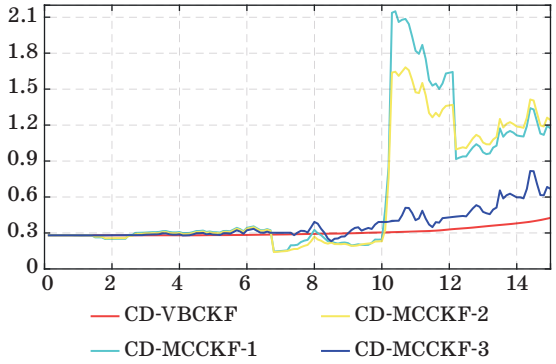
It should be emphasized that the presented data on a qualitative level repeat the results of the research carried out by the authors in [33] for a model of an underdamped oscillatory circuit.

We have also applied the presented CD-CKF modifications to data with *grouped* outliers. The outliers were organized in five groups containing six observations each. The variance of the outliers' noise was considered to be the same. The location of outlier's groups in the modeling interval is uniform and random.

Figure 2 illustrates the dependence of the RMSE on time.

One hundred different runs of the system have been made again. The aggregated values of the accumulated root mean square error are shown in the Table 2.

Hence, the CD-VBCKF and the CD-MCCKF-3 were also the most resistant modifications to the presence of grouped outliers.



■ ***Fig. 2.*** The RMSE values for the grouped outliers

Conclusion

In the paper, four distributional-robust modifications of the continuous-discrete cubature Kalman filter have been proposed. The study of the effectiveness of these modifications has been made for the problem of tracking a space vehicle during its reentry into the atmosphere. Two types of the outliers' ordering have been considered. The first one is the stochastic ordering, and the other one is the grouped ordering.

It has been found that the CD-VBCKF and the CD-MCCKF-3 provide the best results that have approximately equal qualities of estimation. Since the first filter requires finding the optimal values of four parameters (one of them is a matrix) to obtain the proper results, the second one requires estimating the only one parameter, so it seems to be appropriate to recommend the CD-MCCKF-3 for practical using.

It is further planned to modify CD-MCCKF-3 to provide computational robustness by developing a corresponding square-root modification.

■ ***Table 2.*** The values of the accumulated root mean square errors for the grouped outliers

Filters	$ARMSE_1$	$ARMSE_2$	$ARMSE_3$	$ARMSE_4$	$ARMSE_5$	ARMSE
CD-VBCKF	0.225	0.159	0.045	0.028	1.121	1.156
CD-MCCKF-1	0.806	1.049	0.122	0.133	1.451	1.972
CD-MCCKF-2	0.657	0.744	0.099	0.097	1.329	1.665
CD-MCCKF-3	0.224	0.162	0.046	0.028	1.125	1.160

References

1. Bar-Shalom Y., Li X.-R., Kirubarajan T. *Estimation with applications to tracking and navigation*. New York, John Wiley & Sons, 2001. 558 p.
2. Grewal M. S., Weill L. R., Andrews A. P. *Global positioning systems, inertial navigation, and integration*. New York, John Wiley & Sons, 2007. 392 p.
3. Keeling M., Rohani P. *Modeling infectious diseases in humans and animals*. New Jersey, Princeton University Press, 2008. 368 p.
4. Gibbs B. P. *Advanced Kalman filtering, least-squares and modeling: a practical handbook*. New Jersey, John Wiley & Sons, 2011. 605 p.
5. Bierman G. J. *Factorization methods for discrete sequential estimation*. New York, Academic Press, 1977. 241 p.
6. Jazwinski A. H. *Stochastic processes and filtering theory*. New York, Academic Press, 1970. 376 p.
7. Bhaumik S., Date P. *Nonlinear Estimation. Methods and applications with deterministic sample points*. Taylor & Francis Group, 2020. 253 p.
8. Julier S. J., Uhlmann J. K., Durrant-Whyte H. F. A new approach for filtering nonlinear systems. *Proceedings of the American Control Conference*, 1995, pp. 1628–1632.
9. Julier S. J., Uhlmann J. K., Durrant-Whyte H. F. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 2000, vol. 45, no. 3, pp. 477–482.
10. Särkkä S. On unscented Kalman filter for state estimation of continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*, 2007, vol. 52, no. 9, pp. 1631–1641.
11. Kulikova M. V., Kulikov G. Y. Numerical methods for nonlinear filtering of signals and measurements. *Computational Technologies*, 2016, vol. 21, no. 4, pp. 64–98 (In Russian).
12. Arasaratnam I., Haykin S. Cubature Kalman filters. *IEEE Transactions on Automatic Control*, 2009, vol. 54, no. 6, pp. 1254–1269.
13. Arasaratnam I., Haykin S., Hurd T. R. Cubature Kalman filtering for continuous-discrete systems: theory and simulations. *IEEE Transactions on Signal Processing*, 2010, vol. 58, no. 10, pp. 4977–4993.
14. Särkkä S., Solin A. On continuous-discrete cubature Kalman filtering. *Proceedings of the 16th IFAC Symposium on System Identification*, 2012, pp. 1221–1226.
15. Särkkä S. *Bayesian filtering and smoothing*. Cambridge University Press, 2013. 232 p.
16. Chandra K. P. B., Gu D.-W. *Nonlinear filtering. Methods and applications*. Cham, Springer, 2019. 184 p.
17. Durgaprasad G., Thakur S. S. Robust dynamic state estimation of power systems based on M-estimation and realistic modeling of system dynamics. *IEEE Transactions on Power Systems*, 1998, vol. 13, no. 4, pp. 1331–1336.
18. Zhang C., Zhi R., Li T., Corchado J. Adaptive M-estimation for robust cubature Kalman filtering. *Proceedings of the Sensor Signal Processing for Defence (SSPD)*, 2016, pp. 1–5.
19. Leong P. H., Arulampalam S., Lamahewa T. A., Abhayapala T. D. A Gaussian-sum based cubature Kalman filter for bearings-only tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 2013, vol. 49, no. 2, pp. 1161–1176.
20. Raol J. R., Gopalratnam G., Twala B. *Nonlinear filtering: concepts and engineering applications*. New York, CRC Press, 2017. 551 p.
21. Šmídl V., Quinn A. *The variational Bayes method in signal processing*. Berlin, Springer, 2006. 227 p.
22. Särkkä S., Hartikainen J. Non-linear noise adaptive Kalman filtering via variational Bayes. *2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, 2013, pp. 1–6.
23. Särkkä S., Nummenmaa A. Recursive noise adaptive Kalman filtering by variational Bayesian approximations. *IEEE Transactions on Automatic Control*, 2009, vol. 54, no. 3, pp. 596–600.
24. Principe J. C. *Information theoretic learning: Renyi's entropy and kernel perspectives*. New York, Springer-Verlag, 2010. 515 p.
25. Liu W., Pokharel P. P., Principe J. C. Correntropy: Properties and applications in non-Gaussian signal processing. *IEEE Transactions on Signal Processing*, 2007, vol. 55, no. 11, pp. 5286–5298.
26. Cinar G. T., Príncipe J. C. Hidden state estimation using the correntropy filter with fixed point update and adaptive kernel size. *IEEE World Congress on Computational Intelligence*, Brisbane, Australia, 2012, pp. 1–6.
27. Chubich V. M., Prokofieva A. E. Comparative analysis of some robust filters for non-stationary linear discrete systems. *Proceedings of Irkutsk State Technical University*, 2017, vol. 21, no. 12, pp. 123–137 (In Russian).
28. Chubich V. M., Filippova E. V. Research of the efficiency of some robust filters for non-stationary linear continuous-discrete systems. *Sovremennye naukoemkie tekhnologii*, 2018, no. 12, pp. 153–161 (In Russian).
29. Hou J., He H., Yang Y. Gao T., Zhang Y. A Variational Bayesian and Huber-based robust square root cubature Kalman filter for lithium-ion battery state of charge estimation. *Energies*, 2019, vol. 12, no. 9, pp. 1717–1739.
30. Liu X., Chen B., Xu B., Wu Z., Honeine P. Maximum correntropy unscented filter. *International Journal of Systems Science*, 2017, vol. 48, no. 8, pp. 1607–1615.
31. Liu X., Hua Q., Zhao J. Yue P. Maximum correntropy square-root cubature Kalman filter with application to SINS/GPS integrated systems. *ISA Transactions*, 2018, vol. 80, pp. 195–202.
32. Wang G., Li N., Zhang Y. Maximum correntropy unscented Kalman and information filters for non-Gaussian measurement noise. *Journal of the*

Franklin Institute, 2017, vol. 354, no. 18, pp. 8659–8677.
33. Kulabukhova S. O., Chubich V. M. Quality analysis of the continuous-discrete cubature Kalman filter robust modifications. *Proceedings of the Conference of Young Scientists "Science. Technology. Innovation"*, Novosibirsk, 2019, Part 2, pp. 38–42 (In Russian).
34. Kulikov G. Y., Kulikova M. V. Accurate cubature and extended Kalman filtering methods for estimating continuous-time nonlinear stochastic systems with discrete measurements. *Applied Numerical Mathematics*, 2017, vol. 111, pp. 260–275.
35. Chen B., Liu X., Zhao H., Principe J. C. Maximum correntropy Kalman filter. *Automatica*, 2017, vol. 76, pp. 70–77.
36. Izanloo R., Fakoorian S. A., Yazdi H. S., Simon D. Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise. *2016 Annual Conference on Information Science and Systems (CISS)*, 2016, pp. 500–505.
37. Fakoorian S., Mohammadi A., Azimi V., Simon D. Robust Kalman-type filter for non-Gaussian noise: Performance analysis with unknown noise covariances. *Journal of Dynamic Systems, Measurement, and Control*, 2019, vol. 141, no. 9, pp. 091011-1–8.
38. Liu X., Qu H., Zhao J., Chen B. Extended Kalman filter under maximum correntropy criterion. *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 1733–1737.
39. Liu X., Qu H., Zhao J., Yue P., Wang M. Maximum correntropy unscented Kalman filter for spacecraft relative state estimation. *Sensors*, 2016, vol. 16, no. 9, pp. 1530–1546.
40. Wang G., Zhang Y., Wang X. Iterated maximum correntropy unscented Kalman filters for non-Gaussian systems. *Signal Processing*, 2019, no. 163, pp. 87–94.
41. Huang F., Zhang J., Zhang S. Adaptive filtering under a variable kernel width maximum correntropy criterion. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2017, vol. 64, no. 10, pp. 1247–1251.
42. Julier S. J., Uhlmann J. K. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 2004, vol. 92, no. 3, pp. 401–422.

УДК 681.5.015
doi:10.31799/1684-8853-2020-4-11-19

Исследование эффективности робастных к аномальным наблюдениям модификаций непрерывно-дискретного кубатурного фильтра Калмана

В. М. Чубич[a], доктор техн. наук, профессор, orcid.org/0000-0003-2006-0046, chubich@ami.nstu.ru
С. О. Кулабухова[a], магистрант, orcid.org/0000-0002-5823-5641
[a]Новосибирский государственный технический университет, К. Маркса пр., 20, Новосибирск, 630073, РФ

Введение: характерное для практики присутствие в экспериментальных данных выбросов — аномальных наблюдений — способно существенно повлиять на качество обработки указанных данных. Многие динамические процессы описываются стохастическими нелинейными уравнениями. Современные нелинейные фильтры, среди которых кубатурный фильтр Калмана заслуживает особого внимания, не способны эффективно обрабатывать данные с аномальными наблюдениями. Одним из возможных путей решения этой проблемы является применение так называемых робастных методов, устойчивых к наличию выбросов в измерительных данных. **Цель исследования:** выявить наиболее эффективные из современных перспективных робастных модификаций непрерывно-дискретного кубатурного фильтра Калмана и дать соответствующие рекомендации по их применению. **Результаты:** рассмотрены часто возникающие на практике ситуации, когда процесс протекает непрерывно, а данные наблюдений снимаются дискретно. На основе вариационного байесовского и коррентропийного робастных подходов к оцениванию параметров случайных процессов предложены четыре модификации непрерывно-дискретного кубатурного фильтра Калмана. Во всех модификациях присутствуют параметры, оптимальные значения которых зависят как от используемой математической модели, так и от конкретной реализации выборки. Эти значения определяются численно путем минимизации на некоторой сетке накопленной средней квадратичной ошибки. Проведено исследование эффективности предложенных робастных модификаций на примере задачи слежения за космическим аппаратом при его входе в атмосферу в условиях случайного и группированного характерах расположения аномальных наблюдений. Выявлены два наилучших фильтра с близким качеством оценивания. К практическому применению рекомендован коррентропийный фильтр, имеющий один настраиваемый параметр. **Практическая значимость:** выявленный наиболее эффективный робастный фильтр можно использовать при решении различных прикладных задач, связанных с идентификацией стохастических нелинейных непрерывно-дискретных систем.

Ключевые слова — нелинейная фильтрация, кубатурный фильтр Калмана, выбросы, критерий максимальной коррентропии, вариационное байесовское оценивание, стохастическая непрерывно-дискретная система.

Для цитирования: Chubich V. M., Kulabukhova S. O. Research on the effectiveness of continuous-discrete cubature Kalman filter distributional-robust modifications. *Информационно-управляющие системы*, 2020, № 4, с. 11–19. doi:10.31799/1684-8853-2020-4-11-19

For citation: Chubich V. M., Kulabukhova S. O. Research on the effectiveness of continuous-discrete cubature Kalman filter distributional-robust modifications. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 11–19. doi:10.31799/1684-8853-2020-4-11-19

УДК 004.934.2
doi:10.31799/1684-8853-2020-4-20-30

Обзоры
Surveys

Аналитический обзор подходов к распознаванию тональности русскоязычных текстовых данных

А. А. Двойникова[a], *программист, orcid.org/0000-0001-8047-6639*
А. А. Карпов[a], *доктор техн. наук, доцент, orcid.org/0000-0003-3424-652X, karpov@iias.spb.su*
[a]*Санкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ*

Введение: *в последние годы анализ тональности, или сентимент-анализ, высказываний пользователей находит практическое применение во многих областях: оценка качества товаров и услуг по отзывам покупателей в Интернете, анализ негативных эмоций в сообщениях, прогноз фондовых рынков, политических ситуаций на основе новостных лент и многих других. В связи с этим разрабатываются разнообразные системы и методы для сентимент-анализа русскоязычных текстовых данных.* ***Цель:*** *выполнение подробного обзора подходов и сравнительного анализа существующих баз данных в области сентимент-анализа текстов на русском языке.* ***Результаты:*** *аналитический обзор подходов к анализу тональности русскоязычных текстовых данных показал, что для сентимент-анализа текстов сейчас имеется множество разнообразных методов предобработки текстовых данных, их векторизации и машинной классификации. Из сравнительного анализа существующих баз данных по данной тематике можно сделать вывод, что автоматический сентимент-анализ русскоязычных текстов развит значительно меньше, чем для других основных мировых языков. Исследование программных систем для анализа текстов на русском языке демонстрирует, что пока русскоязычный анализ тональности показывает относительно низкую точность по сравнению с англоязычным, одной из причин этого может являться сложная структура русского языка. В статье описываются основные нерешенные проблемы анализа тональности русскоязычных текстов.* ***Обсуждение:*** *в дальнейших исследованиях планируется реализовать сентимент-анализ разговорной речи дикторов с использованием аудиоданных, для чего необходимо сначала получить орфографическую транскрипцию речи для каждого диктора.*

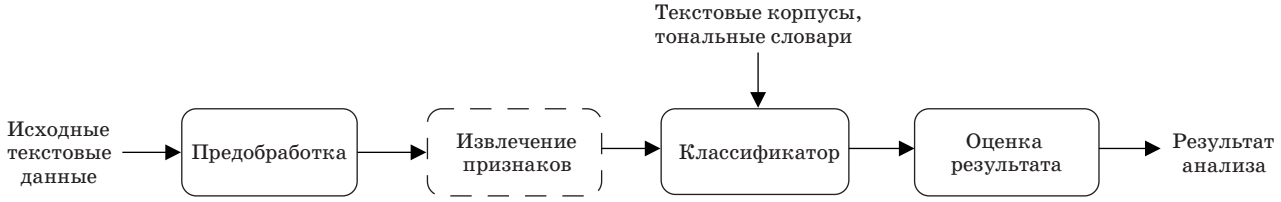
Ключевые слова *— тональность текстовых данных, векторизация текста, сентимент-анализ, компьютерная паралингвистика.*

Для цитирования: Двойникова А. А., Карпов А. А. Аналитический обзор подходов к распознаванию тональности русскоязычных текстовых данных. *Информационно-управляющие системы*, 2020, № 4, с. 20–30. doi:10.31799/1684-8853-2020-4-20-30
For citation: Dvoynikova A. A., Karpov A. A. Analytical review of approaches to Russian text sentiment recognition. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 20–30 (In Russian). doi:10.31799/1684-8853-2020-4-20-30

Введение

Анализ тональности текста, или сентимент-анализ (sentiment analysis), — область компьютерной лингвистики и интеллектуального анализа текста, ориентированная на извлечение из него субъективных мнений и эмоций человека. Анализ тональности находит практическое применение во многих областях: оценка качества товаров и услуг по отзывам покупателей в Интернете, анализ негативных эмоций в сообщениях, прогноз фондовых рынков, политических ситуаций на основе новостных лент [1]. Также сентимент-анализ необходим в автоматизированных системах, в которых человек общается с машиной на естественном языке. Чтобы проанализировать такой объем информации, в последние годы были предложены многочисленные методы автоматического сентимент-анализа, которые рассмотрены в данной статье.

Анализ тональности текстов происходит в несколько этапов (рис. 1). На первом этапе выполняется предобработка исходного текста, далее извлекаются информативные признаки (векто-



■ ***Рис. 1.*** Этапы анализа тональности текста
■ ***Fig. 1.*** The stages of sentiment analysis of text

ризация текста), на их основе строится классификатор (распознаватель) тональности, и последним этапом является оценка результата работы. Этап векторизации текста для лингвистических методов классификации не является обязательным, так как такие классификаторы работают непосредственно с текстами, а не с их векторами.

Предобработка текста

Предобработка текста является первым этапом в его анализе. Она необходима для того, чтобы выделить из «зашумленного» текста релевантную информацию. Предобработка текста включает в себя приведение всех слов к единому регистру, удаление знаков пунктуации, удаление стоп-слов, токенизацию, нормализацию слов и при необходимости иные операции.

При приведении всех слов к единому регистру, как правило, все прописные символы преобразуются в их строчные формы, поскольку предполагается, что прописные или строчные формы слов не имеют различий. Во всех текстах присутствуют знаки пунктуации, выполняющие чаще всего синтаксическую функцию, поэтому при анализе эмоций в тексте нет необходимости сохранять их. Также при обработке текста удаляются стоп-слова — слова, не содержащие смысловой нагрузки, например предлоги, союзы, частицы и т. п. Необходимым этапом предобработки для последующего компьютерного анализа текста является токенизация слов — разбиение текста на отдельные значимые единицы (токены) [2]. Самый простой способ токенизировать русскоязычный текст — разделить его на слова по пробелам. Парадигмы слов в русском языке имеют большое количество словоформ, передающих одинаковый смысл. Форма слова не всегда несет в себе полезную информацию, поэтому при анализе текста рекомендуется производить нормализацию всех слов, т. е. представление слова в его начальной форме. Нормализация может осуществляться двумя способами: лемматизацией и стеммингом. *Лемматизация* — преобразование слова к его начальной форме (лемме). Лемматизация основана на морфологическом словаре. Если слово не присутствует в словаре, то строится гипотеза о способах изменения слова и получения для него леммы. *Стемминг* — получение основы слова, при этом у слов отбрасываются окончания, суффиксы, приставки. Тем самым все слова в тексте приводятся к единой форме. Стемминг основан на морфологических правилах и не требует наличия словаря.

Каждый из этапов предобработки текста позволяет снизить размерность признакового пространства. В зависимости от исходного текста предобработка может включать в себя только несколько операций, а каждая операция может дорабатываться вручную с учетом всех исключений.

Извлечение признаков из текста

Перед тем как использовать машинный классификатор, необходимо представить текст в числовом виде (признаковое описание), т. е. векторизовать текст. Рассмотрим несколько современных способов векторизации текста.

BoW (Bag of Words — «мешок слов») — метод, представляющий текст в виде неупорядоченного набора слов [3]. Каждому слову присваивается свой вес, часто используются веса TF-IDF, отражающие отношение частоты слова в документе к частоте слова во всех документах.

One-hot encoding (прямое кодирование) — метод, преобразующий слова в бинарные векторы [4]. Размер каждого вектора равен объему всех слов в тексте. Перед кодированием все слова, присутствующие в тексте, располагаются по алфавиту.

SVD (Singular Value Decomposition) — метод, преобразующий текст в разреженную матрицу $\mathbf{A}_{n\times m} = \{\mathbf{a}_{i1}, \mathbf{a}_{i2}, \mathbf{a}_{i3}, ..., \mathbf{a}_{im}\}$, где $\mathbf{a}_{ij}$ — взвешенный вектор-столбец частоты терминов предложения i в рассматриваемом документе [5]. Если в документе содержится всего n терминов и m предложений, то на выходе будет матрица размерностью $n \times m$.

Word2Vec (инструментарий, разработанный компанией Google) — нейронная сеть, которая генерирует векторы слов [6]. Она обучена на двух алгоритмах: BoW (предсказывает слово с учетом контекста) и Skip-gram (предсказывает контекст с учетом слова). Word2Vec сначала строит словарь из обучающего текстового корпуса и анализирует векторные представления каждого слова. Кроме того, Word2Vec имеет возможность рассчитать косинусное расстояние между словами.

Glove — метод, разработанный в Стэнфордском университете (США) [7]. В его основе лежит способ подсчета частоты появления слов в текстовом корпусе. Фактически он состоит из двух основных этапов: на первом происходит построение матрицы смежности из обучающего корпуса, а на втором — факторизация матрицы для получения векторов.

FastText — метод, преобразующий в векторы не только слова, но и символьные n-граммы, из которых составлены слова [8].

BERT (Bidirectional Encoder Representations from Transformers) — нейронная сеть, разработанная компанией Google [9]. BERT обучали на

корпусе текстов из Wikipedia и сборнике книг BookCorpus. Идея векторизации в BERT заключается в том, что каждому слову из текста присваивается число, обозначающее порядковый номер слова в словаре, далее это число преобразуется в вектор из 512 символов. Словарь, который использует данная нейросеть, построен таким образом, что слова, близкие по смыслу, располагаются рядом. Тем самым нейронная сеть BERT векторизует текст, учитывая близость слов. Существуют также модификации BERT, например DistilBERT [10]. Это более легкая и быстрая версия BERT, которая примерно соответствует его производительности. Авторы работы [11] показали, что перевод обучения с многоязычной модели BERT на одноязычную модель для русского языка приводит к значительному росту производительности при выполнении анализа эмоций в тексте.

ELMo (Embeddings from Language Models) — нейронная сеть, которая генерирует контекстное представление слов [12]. Модели ELMo обучены на корпусе объемом 1 млрд слов, собранных из новостных лент сети Walmart. Для обучения ELMo применяется минимальная предобработка текстовых данных, выполняется только токенизация и лемматизация. ELMo позволяет векторизовать слова, учитывая контекст до и после этого слова. Идея модели состоит в том, чтобы сначала построить для каждого слова в тексте посимвольный эмбеддинг (embedding) слова, а потом для них применить нейросеть LSTM (Long Short-Term Memory) таким образом, что получатся эмбеддинги, учитывающие контекст, в котором встретилось слово.

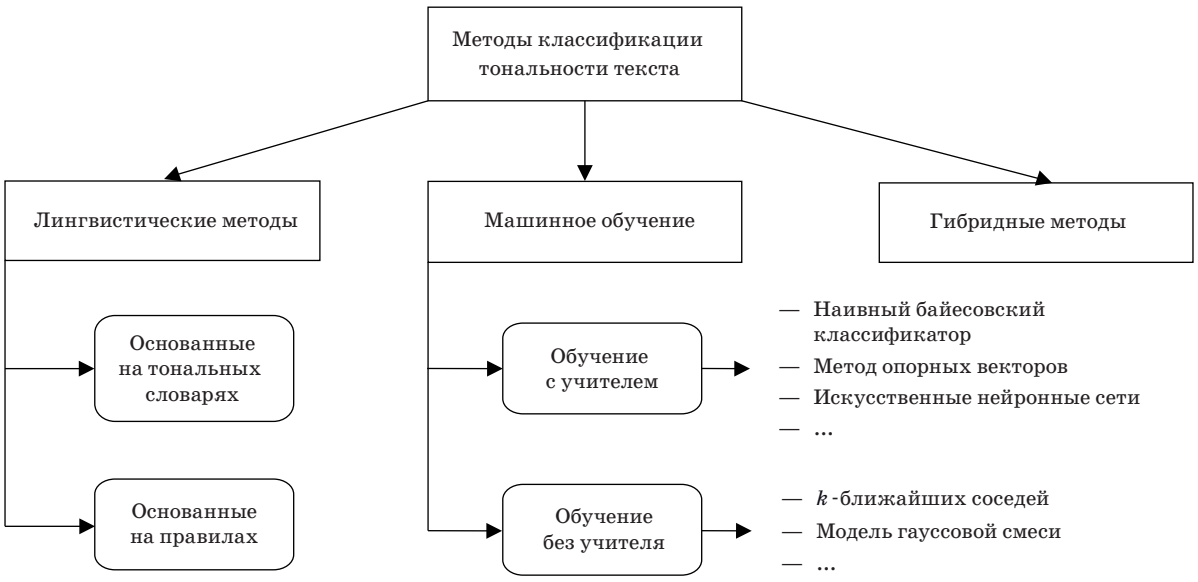
Классификация тональности текстовых данных

В настоящее время существует несколько основных методов для определения (классификации) тональности текста [13]. Все их можно разделить на несколько основных типов (рис. 2), в том числе лингвистические методы, методы на основе машинного обучения и гибридные методы. Рассмотрим все эти методы более подробно.

Первый лингвистический метод основан на *тональных словарях*. Тональный словарь представляет собой набор слов или биграмм, которым задается определенный вес принадлежности к позитивному или негативному классу. При анализе текста каждое слово ищется в этом словаре, и его вес записывается. Если слова нет в словаре, то его класс считается нейтральным, и вес равняется нулю. После того как все веса получены, высчитывается принадлежность данного текста к определенному классу тональности. Данный метод был использован для сентимент-анализа в работе [14].

Второй лингвистический метод основан на *правилах*. Для работы этого метода необходим большой набор продукционных правил конструкции «если –> то». Этот метод также подразумевает использование тональных словарей, в которых слова принадлежат определенному классу. Задача сентимент-анализа решается при помощи метода, основанного на правилах, например в работе [15].

Методы на основе машинного обучения подразделяются на обучение *с учителем (*supervised learning) и *без учителя* (unsupervised learning). Метод обучения с учителем основан на том,



■ ***Рис. 2.*** Систематизация методов классификации тональности текста
■ ***Fig. 2.*** Systematization of methods of classifying the tonal of text

чтобы обучить классификатор на заранее размеченных обучающих текстовых данных [16]. Наиболее распространенные методы в области тонального анализа: наивный байесовский классификатор, метод опорных векторов, логистическая регрессия и искусственные нейронные сети такие, как сверточные (Convolutional Neural Network — CNN), рекуррентные нейронные сети (Recurrent Neural Network — RNN), нейросетевые модели с длинной краткосрочной памятью (Long Short-Term Memory — LSTM) и управляемым рекуррентным блоком (Gated Recurrent Unit — GRU). Авторы статьи [17] сравнивают работу различных традиционных методов обучения с учителем при анализе тональности текстов в социальных сетях. В статье [18] использовались различные нейронные сети для анализа негативных текстовых сообщений. В отличие от метода обучения с учителем, метод обучения без учителя определяет взаимосвязь и закономерности между объектами без размеченных обучающих данных [19]. К таким методам можно отнести модель гауссовой смеси и k-ближайших соседей.

Существуют также гибридные методы, объединяющие в себе несколько различных методов. В работе [20] для задачи классификации текста использовался гибридный метод, объединяющий тональные словари и метод опорных векторов. В статье [21] авторы для решения задачи сентимент-анализа объединяли CNN и k-ближайших соседей.

После этапа классификации сентимент-анализа текстов следует количественная оценка результатов, которая может быть проведена с использованием набора следующих статистических показателей: точности (accuracy или precision), полноты (recall) и F-меры (F-score) [13].

Корпусы для анализа тональности текстов

Несмотря на актуальность анализа тональности русскоязычных текстов, количество аннотированных корпусов для русского языка невелико [22]. На начало 2020 г. в свободном доступе нам удалось найти четыре тональных словаря и семь текстовых корпусов, предназначенных для задачи сентимент-анализа русскоязычных текстов.

Русскоязычные тональные словари в свободном доступе

При использовании метода, основанного на тональных словарях, для автоматической классификации текстов необходимо опираться на словарь, в котором содержатся слова с разметкой принадлежности их к определенному сентименту. Разметка может быть бинарная (2 класса), тернарная (3 класса) и многоклассовая (больше трех классов). Известны несколько тональных словарей для русского языка. Базовые сведения об этих словарях, включая количество содержащихся в них слов, а также количество рассматриваемых классов, представлены в табл. 1.

Тональный словарь RuSentiLex [23] может содержать как отдельные слова, так и словосочетания, для которых указаны их характеристики, обозначающие часть речи или синтаксический тип группы, их лемматизированную форму, тональность, источник информации. В зависимости от контекста одно и то же слово может принимать разное значение тональности. Поэтому авторы словаря ввели отдельный класс тональности, обозначающий смешанную оценку слова. Также авторы отчасти решили проблему со словами, имеющими несколько значений. Они перечисляют все значения слова по тезаурусу РуТез [24] и дают отсылку на соответствующее понятие,

■ ***Таблица 1***. Тональные словари для русскоязычного анализа тональности текста
■ ***Table 1***. Tonal dictionaries for Russian-language text tonality analysis

Название словаря (ссылка для доступа)	Число слов	Число классов	Классы с количеством слов
RuSentiLex (https://www.labinform.ru/pub/rusentilex/index.htm)	16 057	4	Положительные (3785), отрицательные (10 234), нейтральные (1747), смешанная оценка (291)
LinisCrowd (http://linis-crowd.org/)	7545	5	Сильно отрицательные (228), отрицательные (1598), нейтральные (4864), положительные (806), сильно положительные (49)
WordNetAffect (http://lilu.fcim.utm.md/resourcesRoRuWNA_ru.html)	2401	6	Радость (749), страх (617), гнев (398), печаль (445), отвращение (74), удивление (118)
Словарь Белякова [27]	690	2	Положительные (300), отрицательные (390)

имя понятия прописывают в кавычках. В таких случаях каждому значению слова присваивается свое значение тональности.

LinisCrowd [25] — тональный словарь на основе пользовательского интернет-контента социально-политической тематики. Изначально словарь составлялся по размеченным текстам, полученным из социальной сети Facebook. Впоследствии словарь расширился за счет добавления к нему других словоформ, а также слов из других словарей.

WordNetAffect [26] является лексическим ресурсом, который содержит слова, описывающие эмоции. Он был создан на базе онтологии WordNet — семантического лексикона английского языка — путем выбора и разметки наборов синонимов (синсетов) эмоциональными концепциями. Наборы синонимов были вручную размечены эмоциональными метками, далее они были дополнительно переразмечены на шесть эмоциональных категорий. Для русского языка авторы словаря вручную перевели синсеты WordNetAffect с английского языка.

Тональный словарь из работы Белякова [27] содержит 690 основ эмоциональных слов. Словарь разбит на два класса: основы русскоязычных слов с положительной и отрицательной эмоциональной окраской.

Русскоязычные эмоционально окрашенные текстовые корпусы в свободном доступе

Существует несколько эмоционально окрашенных текстовых корпусов для русского языка, их основные характеристики представлены в табл. 2.

Крупнейшая российская конференция по компьютерной лингвистике «Диалог» ежегодно проводит соревнования по компьютерному анализу русского языка (http://www.dialog-21.ru/evaluation/), одним из основных направлений соревнований является анализ тональности текстов. Так, в 2015 и 2016 гг. организаторы предоставили текстовые корпусы SentiRuEval. SentiRuEval-2015 [28] содержит в себе отзывы, собранные из сети Twitter, о ресторанах и автомобилях. Помимо общей тональности отзыва, SentiRuEval-2015 содержит различные целевые аспекты оцениваемого объекта. Каждый из этих аспектов также может иметь тональную оценку. SentiRuEval-2016 [29] включает отзывы о банках и мобильных операторах, собранные из Twitter. Разметка отзывов показывает объект отзыва и отношение субъекта к этому объекту.

LinisCrowd [25] — коллекция документов, посвященных социально-политической тематике. В качестве источника данных использовались записи блог-платформы «Живой Журнал». RuSentiment [22] — текстовый корпус, имеет в своем составе посты, собранные из социальной сети «ВКонтакте», на разные тематики. Некоторые посты могут не иметь разметку по тональности, но они могут относиться к определенному классу высказывания (шаблонные приветствия, благодарственные и поздравительные сообщения). RuTweetCorp [30] — корпус русскоязычных twitter-постов, автоматически размеченных на два класса. Корпусы РОМИП 2012 [31] и Auto_reviews [32] также находятся в свободном доступе.

■ ***Таблица 2***. Текстовые корпусы для исследований русскоязычного сентимент-анализа
■ ***Table 2***. Text corpora for research of Russian-language sentimental analysis

Название корпуса (ссылка для доступа)	Тематика текстов	Число фраз	Число классов
RuTweetCorp (https://study.mokoron.com/#download)	Широкая	226 914	2
РОМИП 2012 (http://romip.ru/ru/2012/tracks.html)	Книги, фильмы и фотокамеры	50 247	2, 3, 5
RuSentiment (https://github.com/strawberrypie/rusentiment)	Широкая	31 185	3
LinisCrowd (http://linis-crowd.org/)	Социально-политическая	26 873	5
SentiRuEval-2016 (http://www.dialog-21.ru/evaluation/2016/sentiment/)	Банки и мобильные операторы	23 595	3
SentiRuEval-2015 (http://www.dialog-21.ru/evaluation/2015/sentiment/)	Рестораны и автомобили	17 628	4
Auto_reviews (https://github.com/oldaandozerskaya/auto_reviews)	Автомобили	6152	5

Сентимент-анализ может также применяться при анализе разговорной речи дикторов. Для решения такой задачи можно использовать мультимодальный корпус RAMAS [33]. Он содержит около семи часов аудио- и видеозаписей интерактивных диалогов, разыгранных несколькими актерами. Перед анализом текстовой составляющей высказываний дикторов необходимо для начала получить орфографическую транскрипцию аудиофайлов, которая не предоставляется разработчиками.

Программные системы для сентимент-анализа

Экспериментальные системы для русскоязычного сентимент-анализа

На соревнованиях в рамках конференции «Диалог» в 2012 г. был предоставлен корпус РОМИП [31]. Авторы работы [34] показали наилучший результат классификации на 5 классов, применив *n*-граммный метод опорных векторов, использовав двоичные веса вместо традиционного TF-IDF, а также обучив модель на комбинированных корпусах. С применением данного подхода на корпусе РОМИП получено среднее значение F-меры = 30,63 %.

На соревнованиях в 2013 г. использовался тот же текстовый корпус, что и в 2012-м, дополнительно к нему организаторы включили корпус из новостных лент, который содержит прямую и косвенную речь с оценкой тональности высказывания (http://romip.ru/ru/collections/sentiment-news-collection-2012.html). Метка тональности текста может принимать одно из четырех значений: положительная, отрицательная, смешанная или нет оценки. В этом корпусе содержится около 5 тыс. новостных фрагментов. Авторы работы [35] достигли наилучшего значения F-меры = 65,9 % для бинарной классификации и 35,36 % для 5-классовой задачи, использовав метод максимальной энтропии и опорных векторов соответственно.

В 2015 г. на «Диалоге» была поставлена более широкая задача. Участникам был предоставлен корпус SentiRuEval-2015, и им необходимо было выделить аспектные термины, определить их тональность и тональность отзыва в целом. Лучший результат решения данной задачи описан в работе [36]. Автор работы применял рекуррентные нейронные сети и получил результат F-меры, равный 61,9 и 64,7 % для отзывов о ресторанах и автомобилях соответственно.

В 2016 г. соревнования проходили на текстовом корпусе SentiRuEval-2016. Задача участников состояла в том, чтобы определить репутационное отношение твита по отношению к конкретной компании. Авторы работы [37] использовали двухслойную нейронную сеть GRU (управляемый рекуррентный блок) и подавали входной вектор в обратной последовательности. При помощи этого метода достигнута F-мера = 55,17 и 55,94 % для отзывов о банках и мобильных операторах соответственно.

Программные продукты для анализа тональности русскоязычных текстов

Задача определения тональности текста является коммерчески востребованной, в связи с этим разрабатываются различные ориентированные компьютерные системы, анализирующие тональность текстов. На начало 2020 г. нам удалось найти пять программных систем в свободном доступе, предназначенных для сентимент-анализа русскоязычных текстов.

SentiFinder [38] — программный модуль высокоскоростной системы лингвистического анализа текстов Eureka Engine. Он определяет тональность текстов на русском, английском и армянском языках. Особенностью данного модуля является то, что он позволяет оценить степень эмоциональности высказывания. Он предназначен для определения тональности отзывов о различных продуктах, а также новостных лент и блогов.

Semantria [39] — программный модуль сентимент-анализа на базе платформы Lexalytics. Система позволяет классифицировать тональность сообщений на нескольких европейских языках, в том числе и на русском. Semantria предназначена для анализа текстов в области маркетинга.

SentiScan — технология распознавания тональности текста на базе платформы YouScan [40]. Классификатор SentiScan обучался на данных, которые содержали в себе отзывы о товарах из различных отраслей. YouScan является коммерческим продуктом, но у него есть бесплатный пробный период, который предоставляется по запросу.

SentiStrength — программный продукт для анализа настроений пользователя [41]. Он предназначен для анализа коротких социальных интернет-текстов. Результатом анализа текста являются две оценки, которые принимают значения от −5 (крайне отрицательно) до 1 (не отрицательно) и от 1 (не положительно) до 5 (крайне положительно). Изначально SentiStrength разрабатывался для анализа английского языка, но впоследствии адаптирован для других языков, в том числе для русского.

Texterra — приложение для анализа тональности новостных сообщений [42]. Анализируемые тексты могут быть из определенных областей: политики, финансов, Интернета, здоровья и постов Twitter. Демоверсия Texterra находится в сво-

■ ***Таблица 3***. Программные продукты для анализа тональности текстов
■ ***Table 3***. Software for analyzing the tonality of texts

Название системы (ссылка для доступа)	Число классов	Используемые методы	Ограничения демоверсии
SentiFinder (http://eurekaengine.ru/ru/demo/)	3	Случайный лес и градиентный бустинг	Анализ текстов объемом не более 10 тыс. символов
Semantria (https://www.lexalytics.com/demo)	3	Нет данных	Анализ текстов объемом до 16 384 символов
SentiScan (https://youscan.io/ru/demo)	3	Метод, основанный на правилах и машинном обучении	Не известны
SentiStrength (http://sentistrength.wlv.ac.uk/)	2	Метод, основанный на тональных словарях и правилах	Анализ сообщений до 100 символов
Texterra (https://texterra.ispras.ru/demo)	3	Метод опорных векторов	Не известны

бодном доступе, ее разработчики предоставляют возможность анализировать фактические новости, собранные с платформы Яндекс.Новости и Twitter, а также пользовательские тексты, введенные вручную.

В свободном доступе можно найти только демоверсии упомянутых систем. Ссылка на демоверсии, их ограничения, а также основные сведения о самих программных продуктах для анализа тональности русскоязычных текстов представлены в табл. 3.

В основе программных продуктов для сентимент-анализа текстов на русском языке лежат, как правило, традиционные методы обучения и не используются нейронные сети. Такой подход может быть обоснован тем, что нейронные сети требуют большого объема обучающих данных, а также большого количества вычислительных и временных ресурсов для их обучения.

Заключение

В статье представлен обзор подходов к анализу тональности русскоязычных текстовых данных. Наличие многочисленных работ на тему анализа тональности текста говорит о том, что данная задача является актуальной и коммерчески востребована во многих сферах, включая рекламу, политику, маркетинг и т. п. Это подтверждается увеличением с каждым годом количества конференций в области анализа текста, а также количества публикаций по анализу как русскоязычных данных, так и текстов на других языках. Однако системы сентимент-анализа русскоязычных текстов развиты меньше, чем для основных мировых языков. По данным академии Google за 2019 г., было опубликовано всего около 28 000 работ по сентимент-анализу русскоязычных текстов, тогда как по англоязычным текстам вышло около 43 000 публикаций. Также русскоязычный сентимент-анализ показывает довольно низкую точность по сравнению с англоязычным, что связано со сложной структурой русского языка. Чтобы подтвердить это утверждение, можно рассмотреть работы по сентимент-анализу чешского языка, так как грамматики русского и чешского языков схожи. В работах [43–45] проводится анализ тональности текстов на английском и чешском языках, по результатам исследования видно, что точность распознавания сентимента в английском языке выше, чем в чешском.

В дальнейших исследованиях мы планируем реализовать сентимент-анализ разговорной речи дикторов с использованием корпуса RAMAS [33]. Для этого необходимо будет получить орфографическую транскрипцию речи для каждого диктора. На основе полученных данных планируется построить классификатор, используя для начала метод на основе тональных словарей, а в последующем и другие методы классификации, описанные в статье.

Финансовая поддержка

Исследование проведено при поддержке РФФИ (проект № 18-07-01407), РНФ (проект № 18-11-00145, раздел 3) и бюджетной темы № 0073-2019-0005.

Литература

1. **Ениколопов С. Н., Кузнецова Ю. М., Смирнов И. В., Станкевич М. А., Чудова Н. В.** Создание инструмента автоматического анализа текста в интересах социогуманитарных исследований. Часть 1. Методические и методологические аспекты. *Искусственный интеллект и принятие решений*, 2019, № 2, с. 28–38. doi:10.14357/20718594190203
2. **Поляков Е. В., Восков Л. С., Абрамов П. С., Поляков С. В.** Исследование обобщенного подхода к решению задач анализа настроений коротких текстовых сообщений в задачах обработки естественного языка. *Информационно-управляющие системы*, 2020, № 1, с. 2–14. doi:10.31799/1684-8853- 2020-1-2-14
3. **Soumya G. K., Joseph S.** Text classification by augmenting bag of words (BOW) representation with co-occurrence feature. *IOSR Journal of Computer Engineering*, 2014, vol. 16(1), pp. 34–38.
4. **Potdar K., Pardawala T. S., Pai C. D.** A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications*, 2017, vol. 175, no. 4, pp. 7–9.
5. **Steinberger J., Jezek K.** Text summarization and singular value decomposition *Proceedings of International Conference on Advances in Information Systems*, Springer, Berlin, Heidelberg, 2004, pp. 245–254.
6. **Mikolov T., Chen K., Corrado G., Dean J.** Efficient estimation of word representations in vector space. *Proceedings of the International Conference on Learning Representations (ICLR 2013)*, 2013. https://openreview.net/forum?id=idpCdOWtqXd60#7b076554-87ba-4e1e-b7cc-2ac107ce8e4d (дата обращения: 02.05.2020).
7. **Pennington J., Socher R., Manning C. D.** Glove: Global vectors for word representation. *Proceedings of International Conference on Empirical Methods in Natural Language Processing (EMNLP-2014)*, 2014, pp. 1532–1543.
8. **Pylieva H., Chernodub A., Grabar N., Hamon T.** Improving automatic categorization of technical vs. Laymen medical words using fasttext word embeddings. *Proceedings of the 1st International Workshop on Informatics and Data-Driven Medicine*, *IDDM 2018*, 2018, pp. 93–102.
9. **Devlin J., Chang M., Lee K., Toutanova K.** Bert: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 2019, vol. 1 (Long and Short Papers), pp. 4171–4186. doi:10.18653/v1/N19-1423
10. **Sanh V., Debut L., Chaumond J., Wolf T.** DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108* (дата обращения: 05.04.2020).
11. **Куратов Ю., Архипов М.** Адаптация глубоких двунаправленных многоязычных моделей на основе архитектуры transformer для русского языка. *Компьютерная лингвистика и интеллектуальные технологии*, 2019, вып. 18, с. 333–339.
12. **Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L.** Deep contextualized word representations. *NAACL-HLT*, 2018, vol. 1 (Long Papers), pp. 2227–2237. doi:10.18653/v1/N18-1202
13. **Dvoynikova A., Verkholyak O., Karpov A.** Analytical review of methods for identifying emotions in text data. *CEUR-WS*, 2020, vol. 2552, pp. 8–21.
14. **Тутубалина Е. В., Иванов В. В., Загулова М., Мингазов Н., Алимова И., Малых В.** Тестирование методов анализа тональности текста, основанных на словарях. *Электронные библиотеки*, 2015, т. 18, № 3-4, с. 138–162.
15. **Паничева П. В.** Система сентиментного анализа ATEX, основанная на правилах, при обработке текстов различных тематик. *Компьютерная лингвистика и интеллектуальные технологии*, 2013, вып. 12, т. 2, с. 101–113.
16. **Котельников Е. В., Клековкина М. В.** Автоматический анализ тональности текстов на основе методов машинного обучения. *Компьютерная лингвистика и интеллектуальные технологии*, 2012, вып. 11, т. 2, с. 27–36.
17. **Maltseva A. V., Makhnytkina O. V., Shilkina N. E., Lizunova I. A.** Social media sentiment analysis with context space model. *Communications in Computer and Information Science*, 2020, vol. 1135, pp. 399–412. doi:10.1007/978-3-030-39296-3_29
18. **Aken B., Risch J., Krestel R., Loser A.** Challenges for toxic comment classification: An in-depth error analysis. *EMNLP*, 2018, pp. 33–42.
19. **Воронина И. Е., Гончаров В. А.** Анализ эмоциональной окраски сообщений в социальных сетях (на примере сети «ВКонтакте»). *Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии*, 2015, № 4, с. 151–158.
20. **Konig A. C., Brill E.** Reducing the human overhead in text categorization. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006, pp. 598–603.
21. **Lakshmi B. S., Raj P. S., Vikram R. R.** Sentiment analysis using deep learning technique CNN with KMeans. *International Journal of Pure and Applied Mathematics*, 2017, vol. 114, no. 11, pp. 47–57.
22. **Rogers A., Romanov A., Rumshisky A., Volkova S., Gronas M., Gribov A.** Rusentiment: An enriched sentiment analysis dataset for social media in Russian. *Proceedings of the 27th International Conference on Computational Linguistics*, 2018, pp. 755–763.
23. **Loukachevitch N., Levchik A.** Creating a general Russian sentiment lexicon. *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16)*, 2016, pp. 1171–1176.

24. **Loukachevitch N., Dobrov B.** RuThes linguistic ontology vs. Russian wordnets. *Proceedings of the 7th Global Wordnet Conference*, 2014, pp. 154–162.
25. **Алексеева С. В., Кольцов С. Н., Кольцова О. Ю.** Linis-crowd. org: лексический ресурс для анализа тональности социально-политических текстов на русском языке. *Труды XVIII объединенной конференции «Интернет и современное общество» (IMS-2015)*, 2015, с. 25–34.
26. **Sokolova M., Bobicev V.** Classification of emotion words in Russian and Romanian languages. *Proceedings of the International Conference RANLP-2009*, 2009, pp. 416–420.
27. **Беляков М. В.** Анализ новостных сообщений сайта МИД РФ методом сентимент-анализа (статья 2). *Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика*, 2016, № 4, с. 115–124.
28. **Loukachevitch N., Blinov P., Kotelnikov E., Rubtsova Y., Ivanov V., Tutubalina E.** SentiRuEval: testing object-oriented sentiment analysis systems in Russian. *Компьютерная лингвистика и интеллектуальные технологии*, 2015, вып. 14, т. 2, с. 3–13.
29. **Lukashevich N. V., Rubtsova Y. V.** SentiRuEval-2016: overcoming time gap and data sparsity in tweet sentiment analysis. *Компьютерная лингвистика и интеллектуальные технологии*, 2016, вып. 15, с. 416–426.
30. **Рубцова Ю. В.** Построение корпуса текстов для настройки тонового классификатора. *Программные продукты и системы*, 2015, № 1(109), с. 72–78.
31. **Chetviorkin I., Braslavskiy P., Loukachevich N.** Sentiment analysis track at ROMIP 2011. *Компьютерная лингвистика и интеллектуальные технологии*, 2012, вып. 11, т. 2, с. 1–14.
32. **Глазкова А. В.** Оценка степени близости категорий текстов при решении задач классификации электронных документов. *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*, 2015, № 2 (31), с. 18–25. doi:10.17223/19988605/31/2
33. **Perepelkina O., Kazimirova E., Konstantinova M.** RAMAS: Russian multimodal corpus of dyadic interaction for affective computing. *Proceedings of 20th International Conference on Speech and Computer SPECOM-2018*, Springer, Cham, 2018, pp. 501–510.
34. **Pak A., Paroubek P.** Language independent approach to sentiment analysis (LIMSI participation in ROMIP'11). *Компьютерная лингвистика и интеллектуальные технологии*, 2012, вып. 11, т. 2, с. 37–50.
35. **Blinov P. D., Klelovkina M. V., Ktelnikov E. V., Pestov O. A.** Research of lexical approach and machine learning methods for sentiment analysis. *Компьютерная лингвистика и интеллектуальные технологии*, 2013, вып. 12, т. 2, с. 51–61.
36. **Тарасов Д.** Глубокие рекуррентные нейронные сети для аспектно-ориентированного анализа тональности отзывов пользователей на различных языках. *Компьютерная лингвистика и интеллектуальные технологии*, 2015, вып. 14, т. 2, с. 53–64.
37. **Trofimovich J.** Comparison of neural network architectures for sentiment analysis of Russian tweets. *Компьютерная лингвистика и интеллектуальные технологии*, 2016, вып. 15, с. 50–59.
38. **Zafar L., Afzal M. T., Ahmed U.** Exploiting polarity features for developing sentiment analysis tool. *CEUR-WS*, 2017, vol. 1874, no. 4. http://ceur-ws.org/Vol-1874/paper_4.pdf (дата обращения: 02.05.2020).
39. **Зверева П. П.** Сентимент-анализ текста (на материале печатных текстов газеты "The New York Times" о России и россиянах). *Вестник Московского государственного областного университета. Серия: Лингвистика*, 2014, № 5, с. 32–37.
40. **Кривоногова С. А.** Психоэмоциональная окрашенность текста: теория и методы исследования. *Материалы 68-й научной конференции «Наука ЮУРГУ»*, 2016, т. 100, с. 368–375.
41. **Thelwall M.** The heart and soul of the web? Sentiment strength detection in the social web with SentiStrength. *Cyberemotions, Springer, Cham*, 2017, pp. 119–134.
42. **Mayorov V., Andrianov I.** MayAnd at SemEval-2016 Task 5: Syntactic and word2vec-based approach to aspect-based polarity detection in Russian. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 2016, pp. 325–329.
43. **Hercig T., Brychcin T., Svoboda L., Konkol M., Steinberger J.** Unsupervised methods to improve aspect-based sentiment analysis in Czech. *Computacion y Sistemas*, 2016, vol. 20 (3), pp. 365–375. doi:10.13053/cys-20-3-2469
44. **Hercig T., Brychcin T., Svoboda L., Konkol M.** Uwb at semeval-2016 task 5: Aspect based sentiment analysis. *SemEval-2016*, 2016, pp. 342–349.
45. **Prikrylova K., Kubon V., Veselovska K.** The role of conjunctions in adjective polarity analysis in Czech. *Computacion y Sistemas*, 2016, vol. 20 (3), pp. 377–386. doi:10.13053/cys-20-3-2460

UDC 004.934.2
doi:10.31799/1684-8853-2020-4-20-30

Analytical review of approaches to Russian text sentiment recognition

A. A. Dvoynikova[a], Programmer, orcid.org/0000-0001-8047-6639
A. A. Karpov[a], Dr. Sc., Tech., Associate Professor, orcid.org/0000-0003-3424-652X, karpov@iias.spb.su
[a]Saint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: In recent years, sentiment analysis has found practical application in many areas, such as evaluating the quality of products and services based on customers' online reviews, analyzing negative emotions in messages, forecasting stock markets or political situations based on news data. In this regard, a large number of systems and methods for Russian text sentiment analysis are being developed. **Purpose:** A detailed review of approaches, and comparative analysis of available databases in the field of Russian text sentiment analysis. **Results:** Our analytical review of the approaches to Russian text data sentiment analysis has shown that there are a large number of ways for preprocessing, vectorization and machine classification of the text data. Studying the available databases shows that the Russian text sentiment analysis is less developed than that for other major world languages. Studying the existing software systems for Russian text analysis reveals their low accuracy compared to English, which can be caused by the sophisticated structure of Russian. **Discussion:** In our further research, we plan to implement sentiment analysis of spoken speech using audio data. To do this, we will need to obtain a spelling transcription of speech for each speaker.
Keywords — text tonality, text vectorization, sentiment analysis, computational paralinguistic.

For citation: Dvoynikova A. A., Karpov A. A. Analytical review of approaches to Russian text sentiment recognition. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 20–30 (In Russian). doi:10.31799/1684-8853-2020-4-20-30

References

1. Enikolopov S. N., Kuznetsova Y. M., Smirnov I. V., Stankevich M. A., Chudova N. V. Creating a text analysis tool for socio-humanitarian research. Part 1. Methodical and methodological aspects. *Artificial Intelligence and Decision Making*, 2019, no. 2, pp. 28–38 (In Russian). doi:10.14357/20718594190203
2. Polyakov E. V., Voskov L. S., Abramov P. S., Polyakov S. V. Generalized approach to sentiment analysis of short text messages in natural language processing. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 1, pp. 2–14 (In Russian). doi:10.31799/1684-8853- 2020-1-2-14
3. Soumya G. K., Joseph S. Text classification by augmenting bag of words (BOW) representation with co-occurrence feature. *IOSR Journal of Computer Engineering*, 2014, vol. 16(1), pp. 34–38.
4. Potdar K., Pardawala T. S., Pai C. D. A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications*, 2017, vol. 175, no. 4, pp. 7–9.
5. Steinberger J., Jezek K. Text summarization and singular value decomposition *Proceedings of International Conference on Advances in Information Systems*, Springer, Berlin, Heidelberg, 2004, pp. 245–254.
6. Mikolov T., Chen K., Corrado G., Dean J. Efficient estimation of word representations in vector space. *Proceedings of the International Conference on Learning Representations (ICLR 2013)*, 2013. Available at: https://openreview.net/forum?id=idpCdOWtqXd60#7b076554-87ba-4e1e-b7cc-2ac107ce8e4d (accessed 2 May 2020).
7. Pennington J., Socher R., Manning C. D. Glove: Global vectors for word representation. *Proceedings of International Conference on Empirical Methods in Natural Language Processing (EMNLP-2014)*, 2014, pp. 1532–1543.
8. Pylieva H., Chernodub A., Grabar N., Hamon T. Improving automatic categorization of technical vs. Laymen medical words using fasttext word embeddings. *Proceedings of the 1st International Workshop on Informatics and Data-Driven Medicine*, *IDDM 2018*, 2018, pp. 93–102.
9. Devlin J., Chang M., Lee K., Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 2019, vol. 1 (Long and Short Papers), pp. 4171–4186. doi:10.18653/v1/N19-1423
10. Sanh V., Debut L., Chaumond J., Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108* (accessed 05 April 2020).
11. Kuratov Yu., Arkhipov M. Adaptation of deep bidirectional multilingual transformers for russian language. *Computational Linguistics and Intellectual Technologies*, 2019, iss. 18, pp. 333–339 (In Russian).
12. Peters M., Neumann M., Iyyer M., Gardner M., Clark C., Lee K., Zettlemoyer L. Deep contextualized word representations. *NAACL-HLT*, 2018, vol. 1 (Long Papers), pp. 2227–2237. doi:10.18653/v1/N18-1202
13. Dvoynikova A., Verkholyak O., Karpov A. Analytical review of methods for identifying emotions in text data. *CEUR-WS*, 2020, vol. 2552, pp. 8–21.
14. Tutubalina E. V., Ivanov V. V., Zagulova M. A., Mingazov N. R., Alimova I. S., Malykh V. A. Sentiment classification of reviews and twitter posts based on dictionaries. *Russian Digital Libraries Journal*, 2015, vol. 18, no. 3-4, pp. 138–162 (In Russian).
15. Panicheva P. V. ATEX: a rule-based sentiment analysis system processing texts in various topics. *Computational Linguistics and Intellectual Technologies*, 2013, iss. 12, vol. 2, pp. 101–113 (In Russian).
16. Kotelnikov E. V., Klekovkina M. V. Automatic text tonality analysis based on machine learning methods. *Computational Linguistics and Intellectual Technologies*, 2012, iss. 11, vol. 2, pp. 27–36 (In Russian).
17. Maltseva A. V., Makhnytkina O. V., Shilkina N. E., Lizunova I. A. Social media sentiment analysis with context space model. *Communications in Computer and Information Science*, 2020, vol. 1135, pp. 399–412. doi:10.1007/978-3-030-39296-3_29
18. Aken B., Risch J., Krestel R., Loser A. Challenges for toxic comment classification: An in-depth error analysis. *EMNLP*, 2018, pp. 33–42.
19. Voronina I. E., Goncharov V. A. Analysis of the emotional color of messages in social networks (for example, the "Vkontakte network"). *Bulletin of the Voronezh State University. Series: System Analysis and Information Technologies*, 2015, no. 4, pp. 151–158 (In Russian).
20. Konig A. C., Brill E. Reducing the human overhead in text categorization. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006, pp. 598–603.
21. Lakshmi B. S., Raj P. S., Vikram R. R. Sentiment analysis using deep learning technique CNN with KMeans. *International Journal of Pure and Applied Mathematics*, 2017, vol. 114, no. 11, pp. 47–57.
22. Rogers A., Romanov A., Rumshisky A., Volkova S., Gronas M., Gribov A. Rusentiment: An enriched sentiment analysis dataset for social media in Russian. *Proceedings of the 27th International Conference on Computational Linguistics*, 2018, pp. 755–763.
23. Loukachevitch N., Levchik A. Creating a general Russian sentiment lexicon. *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16)*, 2016, pp. 1171–1176.
24. Loukachevitch N., Dobrov B. RuThes linguistic ontology vs. Russian wordnets. *Proceedings of the 7th Global Wordnet Conference*, 2014, pp. 154–162.

25. Alexeeva S., Kolcov S., Koltsova O. Linis-crowd.org: A lexical resource for Russian sentiment analysis of social media. *Trudy XVIII ob"edinennoj konferencii «Internet i sovremennoe obshchestvo» (IMS-2015)* [Proceedings of the XVIII Joint Conference "Internet and Modern Society" (IMS-2015)], 2015, pp. 25–34 (In Russian).
26. Sokolova M., Bobicev V. Classification of emotion words in Russian and Romanian languages. *Proceedings of the International Conference RANLP-2009*, 2009, pp. 416–420.
27. Belyakov M. V. The analysis of news messages on the RF ministry of foreign affairs affairs website by the sentimental analysis (article 2). *Bulletin of the Peoples' Friendship University of Russia. Series: Theory of Language. Semiotics. Semantics*, 2016, no. 4, pp. 115–124 (In Russian).
28. Loukachevitch N., Blinov P., Kotelnikov E., Rubtsova Y., Ivanov V., Tutubalina E. SentiRuEval: testing object-oriented sentiment analysis systems in Russian. *Computational Linguistics and Intellectual Technologies*, 2015, iss. 14, vol. 2, pp. 3–13.
29. Lukashevich N. V., Rubtsova Y. V. SentiRuEval-2016: overcoming time gap and data sparsity in tweet sentiment analysis. *Computational Linguistics and Intellectual Technologies*, 2016, iss. 15, pp. 416–426.
30. Rubcova U. V. Building a text corpus for setting up a tone classifier. *Software & Systems*, 2015, no. 1(109), pp. 72–78 (In Russian).
31. Chetviorkin I., Braslavskiy P., Loukachevich N. Sentiment analysis track at ROMIP 2011. *Computational Linguistics and Intellectual Technologies*, 2012, iss. 11, vol. 2, pp. 1–14.
32. Glazkova A. V. The evaluation of the proximity of text categories for solving electronic documents classification tasks. *Bulletin of Tomsk State University. Management, Computer Engineering and Informatics*, 2015, no. 2(31), pp. 18–25 (In Russian). doi:10.17223/19988605/31/2
33. Perepelkina O., Kazimirova E., Konstantinova M. RAMAS: Russian multimodal corpus of dyadic interaction for affective computing. *Proceedings of 20th International Conference on Speech and Computer SPECOM-2018*, Springer, Cham, 2018, pp. 501–510.
34. Pak A., Paroubek P. Language independent approach to sentiment analysis (LIMSI participation in ROMIP'11). *Computational Linguistics and Intellectual Technologies*, 2012, iss. 11, vol. 2, pp. 37–50.
35. Blinov P. D., Klelovkina M. V., Ktelnikov E. V., Pestov O. A. Research of lexical approach and machine learning methods for sentiment analysis. *Computational Linguistics and Intellectual Technologies*, 2013, iss. 12, vol. 2, pp. 51–61.
36. Tarasov D. S. Deep recurrent neural networks for multiple language aspect-based sentiment analysis of user reviews. *Computational Linguistics and Intellectual Technologies*, 2015, iss. 14, vol. 2, pp. 53–64 (In Russian).
37. Trofimovich J. Comparison of neural network architectures for sentiment analysis of russian tweets. *Computational Linguistics and Intellectual Technologies*, 2016, iss. 15, pp. 50–59.
38. Zafar L., Afzal M. T., Ahmed U. Exploiting polarity features for developing sentiment analysis tool. *CEUR-WS*, 2017, vol. 1874, no. 4. Available at: http://ceur-ws.org/Vol-1874/paper_4.pdf (accessed 2 May 2020).
39. Zvereva P. Sentiment-analysis of text (texts about Russia and the Russians from The New York Times). *Bulletin of the Moscow State Regional University. Series: Linguistics*, 2014, no. 5, pp. 32–37 (In Russian).
40. Krivonogova S. A. Psychoemotional color of the text: theory and research methods. *Materialy 68-j nauchnoj konferencii «Nauka YUURGU»* [Materials of the 68th Scientific Conference "Science of the South Ural State University"], 2016, vol. 100, pp. 368–375 (In Russian).
41. Thelwall M. The heart and soul of the web? Sentiment strength detection in the social web with SentiStrength. *Cyberemotions, Springer, Cham*, 2017, pp. 119–134.
42. Mayorov V., Andrianov I. MayAnd at SemEval-2016 Task 5: Syntactic and word2vec-based approach to aspect-based polarity detection in Russian. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 2016, pp. 325–329.
43. Hercig T., Brychcin T., Svoboda L., Konkol M., Steinberger J. Unsupervised methods to improve aspect-based sentiment analysis in Czech. *Computacion y Sistemas*, 2016, vol. 20 (3), pp. 365–375. doi:10.13053/cys-20-3-2469
44. Hercig T., Brychcin T., Svoboda L., Konkol M. Uwb at semeval-2016 task 5: Aspect based sentiment analysis. *SemEval-2016*, 2016, pp. 342–349.
45. Prikrylova K., Kubon V., Veselovska K. The role of conjunctions in adjective polarity analysis in Czech. *Computacion y Sistemas*, 2016, vol. 20 (3), pp. 377–386. doi:10.13053/cys-20-3-2460

УДК 519.87
doi:10.31799/1684-8853-2020-4-31-41

Научные статьи
Articles

Метод выбора архитектуры мультиагентной системы управления автономного необитаемого подводного аппарата

Л. А. Мартынова[а], *доктор техн. наук, старший научный сотрудник, orcid.org/0000-0002-5613-0838, martynowa999@bk.ru*
Н. К. Киселев[б], *первый заместитель главного конструктора, orcid.org/0000-0002-5401-4470*
А. А. Мысливый[в], *канд. воен. наук, заместитель начальника отдела, orcid.org/0000-0002-6741-3139*
[а]*АО «Концерн «ЦНИИ «Электроприбор», Малая Посадская ул., 30, Санкт-Петербург, 197046, РФ*
[б]*АО «ЦКБ «Лазурит», Свободы ул., 57, Нижний Новгород, 603951, РФ*
[в]*НИИ ОСИС ВМФ, Разводная ул., 17, Санкт-Петербург, Петергоф, 198516, РФ*

Введение: *постоянное совершенствование автономных необитаемых подводных аппаратов, усложнение их систем и использование гибридной системы энергообеспечения привели к необходимости разработки системы управления с использованием мультиагентной технологии. К настоящему времени сформировалось большое количество стилей мультиагентных архитектур, преимущественно в области организации производства и разработки программного обеспечения. В связи с этим возникает задача выбора наиболее подходящего стиля архитектуры мультиагентной системы управления автономного необитаемого подводного аппарата с гибридной системой энергообеспечения с учетом его особенностей.* ***Цель:*** *разработка метода выбора наиболее подходящего стиля мультиагентной архитектуры на множестве альтернативных вариантов.* ***Метод:*** *в основу разработанного метода положена сравнительная оценка различных стилей архитектур по нефункциональным требованиям. Для этого специально разрабатывается целевой граф с учетом особенностей проектируемого аппарата. Кроме того, при формировании итогового результата использован алгоритм распространения меток как наиболее подходящий для рассматриваемой задачи.* ***Результаты:*** *предложенный метод выбора стиля архитектуры включает в себя выработку показателей, по которым целесообразно вести сравнение альтернативных вариантов; формирование различных стилей архитектур, наиболее подходящих для разрабатываемого аппарата; анализ положительных и отрицательных влияний стиля архитектуры на нефункциональные требования; формализацию этих влияний в виде качественных или количественных меток; получение итоговой оценки путем применения алгоритма распространения меток.* ***Практическая значимость:*** *предложенный метод позволяет осуществить выбор наиболее целесообразной архитектуры мультиагентной системы управления автономного необитаемого подводного аппарата. Метод может быть использован также для более широкого круга робототехнических комплексов наземного и воздушного базирования.*

Ключевые слова *— автономный необитаемый подводный аппарат, архитектура мультиагентной системы управления, оценка эффективности, нефункциональные требования, алгоритм распространения меток.*

Для цитирования: Мартынова Л. А., Киселев Н. К., Мысливый А. А. Метод выбора архитектуры мультиагентной системы управления автономного необитаемого подводного аппарата. Информационно-управляющие системы, 2020, № 4, с. 31–41. doi:10.31799/1684-8853-2020-4-31-41

For citation: Martynova L. A., Kiselev N. K., Myslivyi A. A. Choice of architecture for a multi-agent control system of an autonomous underwater vehicle. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 31–41 (In Russian). doi:10.31799/1684-8853-2020-4-31-41

Введение

Совершенствование технологий привело к возможности создания сложных автономных необитаемых подводных аппаратов (АНПА) для выполнения продолжительных миссий с переходом на дальние расстояния. Спецификой функционирования АНПА является невозможность использовать спутниковые навигационные системы, а также ограниченность дальности гидроакустической связи и объема передаваемой информации. В результате при выполнении маршрутного задания АНПА приходится преодолевать дальние расстояния практически без сторонней помощи, без дозаправки, с крайне редкой обсервацией по сигналам спутниковых навигационных систем. Создание таких АНПА сопровождается разработкой сложных систем: системы освещения обстановки; навигационной системы; энергетической системы; интегрированной системы управления [1, 2], взаимодействующей с системами управления остальных систем АНПА, — каждая из которых имеет собственную локальную систему управления. В результате сложилась необходимость формирования интегрированной системы управления на мультиагентной основе. Движение АНПА и выполнение других функций осуществляются путем взаимодействия всех систем-агентов на основе равноправного общения между собой. Все системы взаимоувязаны в виде логической архитектуры, определяемой целями и задачами, стоящими перед АНПА. Логическая

архитектура включает в себя функциональную, поведенческую и временную архитектуры. Функциональная архитектура определяет преобразования, проводимые системой при выполнении своего назначения. Поведенческая архитектура определяет последовательность выполнения действий, условия для управления системой или потоком данных, уровень производительности, необходимый для удовлетворения системных требований. Временная архитектура определяет синхронные и асинхронные аспекты функций системы.

Однако логическая архитектура не способна обеспечить слаженную работу всех агентов, основанную на обработке информации в реальном масштабе времени, безотказности системы, предсказуемости поведения агентов, наиболее критичных для морской робототехники. Поэтому при проектировании системы управления АНПА необходимо учитывать не только функциональные требования, но и так называемые «нефункциональные» требования, определяющие свойства, которые система должна демонстрировать, или ограничения, которые она должна соблюдать, не относящиеся к поведению системы [3]. От того, насколько удачно выбрана архитектура системы управления, зависит успешность и эффективность функционирования АНПА в целом.

К сожалению, в настоящее время, в отличие от оценки функциональности АНПА [4–8], нефункциональным требованиям в области морской робототехники не уделено должного внимания, что может привести при проектировании сложного АНПА к многократному повторному перепроектированию вплоть до полного отказа от проекта.

В отечественной литературе нефункциональные требования к АНПА не рассматривались: в середине 2000-х применительно к АНПА описаны различные виды архитектур [9–17], однако для сложных АНПА они не могут быть использованы без усовершенствования.

В зарубежных источниках формированию современных сложных АНПА уделено достаточно внимания [18–30], однако без сравнительного анализа альтернативных вариантов построения. В то же время работы по формированию различных стилей архитектур мультиагентных систем в других областях, например, при организации крупных предприятий [31] или при разработке сложного программного обеспечения [32] несколькими коллективами, позволили создать ряд альтернативных стилей архитектур [33], часть из которых может быть использована также и в робототехнике.

В качестве наиболее целесообразных архитектур для применения в робототехнике в работе [34] предложено рассматривать архитектуры «Структура-5» и «Совместное предприятие». Однако указанные архитектуры можно применить к простым роботам с несложными системами управления, и, кроме того, в них отсутствует энергетическая составляющая, наиболее критичная для АНПА, поскольку специальных дозаправок по ходу выполнения маршрутного задания в морской среде пока не предусмотрено. Последнее обстоятельство вызвало необходимость использовать гибридную систему энергообеспечения (СЭО).

В связи с этим возникла задача формирования стиля мультиагентной архитектуры, наиболее подходящего под особенности функционирования сложного АНПА с гибридной СЭО.

Для того чтобы сформировать стиль мультиагентной архитектуры АНПА, необходимо прежде всего методически определить последовательность выполняемых действий. Методическая последовательность действий легла в основу специально разработанного метода, описание которого приведено в настоящей работе.

Постановка задачи и последовательность ее решения

Пусть разрабатываемый АНПА включает в себя следующие основные системы: навигации (СН); освещения обстановки (СОО); гидроакустической и радиосвязи (СРС); энергообеспечения (СЭО); аварийную, а также движительно-рулевой комплекс и полезную нагрузку.

Пусть на этапе разработки АНПА уже определена логическая архитектура, не учитывающая нефункциональные требования.

Пусть имеется множество стилей мультиагентных архитектур $\{A_1, ..., A_m\}$ размерности m.

Пусть имеется множество критериев $\{K_1, ..., K_n\}$ размерности n, сформированных по нефункциональным требованиям.

Пусть получены оценки S_{ij}, $i = 1, ..., m$; $j = 1, ..., n$ каждого нефункционального показателя $K_j \in \{K_1, ..., K_n\}$ каждого стиля архитектуры A_i.

Пусть по совокупности критериев $K_1, ..., K_n$ получены итоговые оценки S_i, $i = 1, ..., m$ каждого стиля архитектуры A_i, образующие множество $\{S_1, ..., S_m\}$ по правилам, описанным ниже.

Необходимо выбрать такой стиль архитектуры $A \in \{A_1, ..., A_m\}$, который доставлял бы максимум итоговой оценке $S_i(K_1, ..., K_n) \forall i = 1, ..., m$.

Таким образом, получив результат $S(K_1, ..., K_n)$, мы получим тем самым наиболее подходящий стиль архитектуры.

Для решения поставленной задачи необходимо:

— сформировать множество показателей $\{K_1, ..., K_n\}$, основанных на нефункциональных требованиях;

— сформировать множество альтернативных стилей архитектур $\{A_1, ..., A_m\}$ мультиагентной системы управления АНПА;

— сформировать оценки S_i по показателям $\{K_1, ..., K_n\}$ каждого альтернативного стиля архитектуры из множества $\{A_1, ..., A_m\}$;

— определить итоговые оценки S_i по совокупности параметров $\{K_1, ..., K_n\}$ каждого альтернативного стиля архитектуры из множества $\{A_1, ..., A_m\}$;

— осуществить выбор наиболее подходящего стиля архитектуры $A \in \{A_1, ..., A_m\}$.

Последовательное решение перечисленных задач составляет основу предлагаемого метода выбора стиля мультиагентной архитектуры АНПА.

Формирование показателей для выбора стиля мультиагентной архитектуры

Показатели, по которым предлагается осуществить выбор наиболее подходящего стиля архитектуры мультиагентной системы управления АНПА, представляют собой нефункциональные требования к мультиагентной системе.

Применительно к робототехнике наиболее подходящими требованиями определены [34]:

— работа в режиме реального времени;

— координация работы агентов;

— предсказуемость поведения агентов и глобального поведения системы;

— адаптация;

— безопасность;

— отказоустойчивость;

— масштабируемость.

Под *режимом реального времени* понимается способность агентов обрабатывать данные в таком темпе, при котором обеспечивается взаимодействие вычислительной системы с внешними по отношению к ней процессами в темпе, соизмеримом со скоростью протекания этих процессов.

Под *координацией* понимается способность агентов мультиагентных систем координировать свои действия с другими агентами для достижения общей цели или своих локальных целей.

Под *предсказуемостью* понимается прогнозирование поведения агентов, которое может быть затруднено из-за их адаптивности и отзывчивости к неожиданным ситуациям.

Под *адаптивностью* понимается способность агентов адаптироваться к изменениям в их окружении.

Под *безопасностью* функционирования мультиагентной системы понимается проверка подлинности данных, полученных от источников, например, путем их идентификации со своими собственными данными.

Под *отказоустойчивостью* понимается обнаружение неисправности или восстановление системы после отказа элементов, агентов или систем; защита услуг, предоставляемых другим агентам, от прерываний; исключение отказа всей системы в целом в случае отказа только одного из агентов.

Под *масштабируемостью* понимается возможность добавлять новые программные и аппаратные модули и эффективная коммуникация потока данных.

Для решения поставленной задачи на данном этапе исследований выберем основные показатели, наиболее критичные для АНПА с гибридной СЭО, и проведем их объективизацию.

При рассмотрении любой мультиагентной системы основным вопросом является координация действий агентов, так как обработка данных может вестись в различном темпе, объем данных может быть различным, темп поступления также может различаться. В то же время в обработке должны использоваться данные, соответствующие одному и тому же моменту времени вне зависимости от момента их поступления. Поэтому координация опирается на работу в режиме реального времени.

Кроме того, при использовании в АНПА гибридной СЭО, включающей в себя разнородные источники электропитания, на передний план выходят вопросы безопасности, так как при неудачном подключении источников питания или потребителей, ошибочной передаче команд на включение или переключение может возникнуть предаварийная или аварийная ситуация. Поэтому вопросы безопасности для АНПА с гибридной СЭО являются основополагающими.

Таким образом, наиболее критичными показателями для АНПА с гибридной СЭО являются координация, работа в режиме реального времени и безопасность.

Такие показатели, как адаптируемость и масштабируемость для данного АНПА не критичны, так как разрабатываемый АНПА — уникальный, все особенности конструктива можно предусмотреть заранее, и что-то добавлять или исключать не планируется. Формировать системы, что называется, с запасом, на всякий случай — нецелесообразно, поскольку те негативные явления, которые могут произойти в ходе модернизации, прогнозируются и учитываются заранее при проектировании. При разработке уникального аппарата рациональнее по максимуму задействовать имеющиеся ресурсы без учета возможной перспективы.

Рассмотрим подробнее, чем определяются выбранные показатели.

Координация, выражаемая способностью агентов системы координировать свои действия, определяется синхронизацией и постоянным общением агентов между собой, направленным на обеспечение синхронизации. Для этого, с одной стороны, должна быть предварительная дого-

воренность, каким образом происходит синхронизация, а с другой стороны, обработку информации следует производить так, чтобы успевать выполнять договоренности о правилах синхронизации. Сказанное означает следующее. Пусть правилами функционирования мультиагентной системы определено, что синхронизация происходит путем обработки данных в определенные моменты времени с постоянным интервалом. Для выдерживания этих интервалов должны использоваться или точно выверенные внутренние часы каждого агента, или единое время для всех агентов, которое постоянно рассылается системой единого времени, входящей в состав оборудования АНПА. В любом случае агенты должны функционировать так, чтобы успевать к заданному моменту времени справиться с внутренней обработкой данных и переслать их заинтересованным агентам. Поэтому, кроме синхронизации, важно иметь факторы, способствующие возможности выполнять обработку данных в режиме реального времени.

На работу в режиме реального времени влияние оказывает скорость обработки данных и доставки сообщений между агентами. Скорость обработки данных зависит от объема обрабатываемой информации и использования быстрых алгоритмов обработки (возможно, в ущерб точности). На скорость доставки влияние оказывают прямое общение между взаимодействующими агентами и объем информационных потоков: чем больше поток, тем дольше он доставляется и обрабатывается на предмет целостности переданных данных, их безопасности, корректности и т. д.

Безопасность обеспечивается постоянным контролем истинности поступающих данных и результатов обработки в целях непротиворечивости прогнозируемым результатам. Кроме того, в процессе функционирования в максимально сжатые сроки должны быть выявлены и оперативно устранены нестыковки в данных. В этом случае безопасности способствуют также ограничение информационных потоков, прямое общение между агентами и быстрые алгоритмы обработки поступающих данных.

После того как для проведения сравнительного анализа стилей архитектур мультиагентных систем выбраны показатели, сформируем с их учетом альтернативные стили мультиагентных систем, наиболее подходящие под особенности АНПА.

Формирование стилей архитектур системы управления АНПА

При функционировании мультиагентной системы управления главным является обеспечение координации взаимодействия между агентами, поэтому стили мультиагентных архитектур определяются той системой-агентом, которая будет являться координатором в АНПА с гибридной СЭО.

Поскольку задачей разрабатываемого АНПА является прибытие в заданную конечную точку маршрута, то системой, координирующей работу остальных систем мультиагентной системы управления, можно назначить навигационную систему. Это положение согласуется с предложенным в работе [33] стилем архитектуры, использующим в качестве координатора навигационную систему (рис. 1). Однако для более адекватного отражения работы АНПА с гибридной СЭО в схему [33] была добавлена одна из важнейших ключевых систем — гибридная СЭО, выраженная на схеме агентом «Энергетика».

В соответствии с этой схемой агент «Управление механическими средствами» управляет устройствами, клапанами, приводами и т. д. Агент «Глобальный планировщик» осуществляет стратегическое планирование маршрута движения АНПА. Агент «Сенсоры» является опорным агентом получения реальной информации от датчиков, который интегрирует ее в последовательную интерпретацию в реальном масштабе времени для агента «Навигатор». Агент «Энергетика» получает данные по удельному расходу энергоресурса от механических средств, сопоставляет данные с собственной информацией о запасах энергоресурса и сообщает эти данные агенту «Навигация». Тот сопоставляет полученные данные с текущим положением АНПА на маршрутной траектории и по результатам оценивает оставшийся путь и достаточность энергоресурса для завершения маршрутного задания и прибытия АНПА в конечную точку маршрута.

Другим подходом к контролю возможности достижения АНПА конечной точки маршрута яв-

■ ***Рис. 1.*** Архитектура мультиагентной системы «Структура-5», дополненная агентом «Энергетика»

■ ***Fig. 1.*** The architecture of the multi-agent system “Structure-5”, supplemented by the agent “Energy”

ляется оценка достаточности энергоресурса, проводимая в гибридной СЭО (агент «Энергетика»). В функцию агента «Энергетика» входят контроль расхода энергоресурса при движении АНПА вдоль маршрутной траектории и оценка соответствия текущего расхода плану расхода энергоресурса. По результатам сопоставления данных о положении АНПА на маршрутной траектории, выдаваемых агентом «Навигация», оценивается достаточность энергоресурса для преодоления всего запланированного маршрута. В связи с этим координатором может также являться агент «Энергетика». Сама по себе СЭО — гибридная, постоянно внутри себя принимающая решение относительно того, какой именно источник электроэнергии задействовать так, чтобы энергии хватило на завершение маршрутного задания. Поэтому для оперативного управления энергосистемой все исполнительные органы должны быть рядом с координатором, и в этом случае координатором целесообразно назначить агента «Энергетика».

Вместе с тем функционирование и навигационной системы, и энергетической тесно связано с функционированием остальных систем АНПА. Так, например, при отклонении АНПА от маршрута он должен произвести маневр для восстановления своего положения на маршрутной траектории, а также необходимы данные системы освещения обстановки о том, не попал ли АНПА в результате отклонения в неблагоприятную зону. В связи с этим еще одним альтернативным вариантом является формирование специального агента «Диспетчер» и назначение его координатором. «Диспетчер» собирает все данные, благодаря чему имеет полную информацию обо всех системах АНПА.

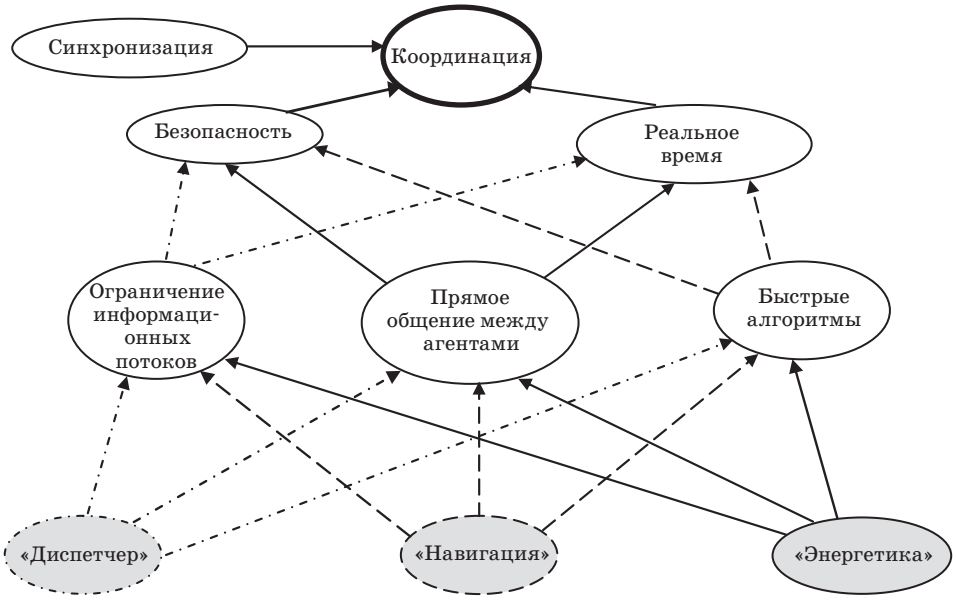
Таким образом, по результатам анализа работы АНПА наиболее перспективными для дальнейшего рассмотрения выбраны три альтернативных стиля архитектуры: «Навигация», «Энергетика», «Диспетчер», — отличающиеся агентом-координатором, именем которого они названы.

Представим сформированные стили архитектуры и выбранные ранее показатели в виде целевого графа (рис. 2). У построенного графа ребра характеризуют отношения между нефункциональными требованиями, а узлы являются целями и подцелями.

На следующем этапе исследований из сформированных стилей мультиагентных архитектур необходимо определить наиболее подходящий. Проанализируем достоинства и недостатки каждого из рассматриваемых стилей архитектур по выбранным показателям.

Ограничение информационных потоков. Агент «Диспетчер» аккумулирует, по сути, всю информацию у себя, и это приводит к перегрузке его вычислительных ресурсов и сети обмена данными. Агент «Навигация» и агент «Энергетика» способствуют ограничению информационных потоков, общаясь только с теми агентами, с которыми это необходимо.

Прямое общение между агентами. Агент «Диспетчер» исключает возможность прямого общения агентов друг с другом, агент «Навигация» и агент «Энергетика» такую возможность обеспечивают.



■ ***Рис. 2.*** Граф целей и подцелей АНПА с гибридной системой энергообеспечения
■ ***Fig. 2.*** Graph of goals and sub-goals of the AUV with a hybrid energy supply system

Быстрые алгоритмы. При работе в режиме реального времени агент «Диспетчер» проигрывает, поскольку вынужден обрабатывать большой объем данных по широкому спектру задач и раздавать результаты обработки всем потребителям-подсистемам. Агент «Навигация» в части навигации самостоятельно принимает решение и тут же корректирует положение АНПА на маршрутной траектории. Агент «Энергетика» также проводит оперативно оценку по собственным алгоритмам.

С учетом результатов анализа выберем далее подход к оценке этих показателей.

Получение оценки показателей для выбора стиля архитектуры

После того как сформирован целевой граф, задачу получения оценки можно формализовать как обнаружение сообществ в графе [35]. В настоящее время существуют три основных класса алгоритмов, которые обеспечивают хорошую точность в выявлении структуры сильно пересекающихся сообществ. Эти классы включают алгоритмы, основанные на базе графических моделей, методы локальной оптимизации и методы распространения меток.

Для рассматриваемой задачи получения оценки показателей в работе [31] применительно к робототехнике наиболее подходящим указан метод распространения меток (Label Propagation — LP), поскольку он позволяет практически без изменений уложить всю цепочку зависящих факторов в стройный связный граф. Парадигма LP может рассматриваться как популярное направление современных методов обнаружения сообществ. Простые и интуитивно понятные методы из данного класса обеспечивают идеальное сочетание свойств приемлемой точности обнаружения сообществ, низкой вычислительной сложности, простоты реализации с точки зрения современных вычислительных парадигм распределенной обработки графов [36, 37].

Общие и отличительные характеристики методов в этом семействе состоят в процессе обмена метками сообществ между узлами графа, которые накапливают пришедшие метки и отправляют сообщения об обновленной коллекции меток соседним узлам. Первый раз модель LP была введена авторами работы [38].

Преимуществами использования данного алгоритма являются:

— возможность получения качественного и (или) количественного значения оценки;

— доказанность сходимости алгоритма к конечному результату.

Для получения оценки каждого альтернативного стиля архитектуры используем разложение цели на подцели (см. рис. 2). Для этого воспользуемся подходом [31], в котором разложение цели на подцели подчинено правилам AND или OR. Пусть переходы из одного узла графа в другой характеризуются метками: S (Satisfied — повышение) и D (Denied — понижение). Тогда выполнение одной из задач может привести к понижению уровня, а выполнение другой — наоборот, к его повышению. Затем каждому узлу графа G ставятся в соответствие две переменные Sat(G) и Den(G), принимающие одно из значений из множества $\{F; P; N\}$, где (F — полное, P — частичное, N — никакое), при этом $F > P > N$.

Использование алгоритма LP основано на понятиях Initial (стартовое), Current (текущее) и Old (предыдущее) значений переменных. Пара $\{\text{Sat}(G_i); \text{Den}(G_i)\}$ является меткой для G_i. Алгоритм LP [14] заключается в следующем. Сначала происходит инициализация массива Current стартовым значением Initial. Затем для каждого узла графа G_i и для каждого перехода из этого узла в смежный узел происходит обновление пары $\{\text{Sat}(G_i); \text{Den}(G_i)\}$ с учетом повышения (Set) или понижения (Den) текущего его значения. Полученный результат сравнивается с предыдущим значением Old, и по результатам сравнения возвращается максимум как новое текущее значение G_i Current. Эта процедура выполняется до тех пор, пока не окажется, что дальнейшее обновление невозможно, т. е. выполнено условие Current = Old. В результате будут получены качественные оценки рассматриваемых типов архитектур, и по максимуму полученного результата будет осуществлен выбор наиболее целесообразного стиля архитектуры.

В том случае, если есть все возможности хотя бы частично наделить цели и подцели количественными характеристиками, то можно использовать метод LP по количественным характеристикам.

При количественном анализе используются две действительные константы inf и sup такие, что $0 < \text{inf} < \text{sup}$, для каждого узла графа G введены две действительные переменные Sat(G); Den(G) в диапазоне интервала [inf; sup]. Для обработки целевых отношений используются два оператора $\otimes$ и $\oplus$, означающие, соответственно, конъюнкцию и дизъюнкцию; кроме того, могут быть использованы отрицание дизъюнкции и конъюнкции. Также приписываем каждому целевому отношению $+S$, $-S$, $+D$, $-D$ вес $w \in]\text{inf}; \text{sup}[$. По аксиомам для инвариантов и правилам, приведенным в работе [31], вычисляем текущее значение узла графа, основываясь на прежнем значении. В данном случае принимается вероятностная модель, в которой увеличение Sat(G) представляется как вероятность того, что G увеличивается, а снижение Den(G) представляется

как вероятность того, что G снижается. В алгоритме LP полагается inf = 0, sup = 1, и операции ⊗, ⊕, inv() определены как

$$p_1 \otimes p_2 = def\, p_1 \cdot p_2;$$
$$p_1 \oplus p_2 = def\, p_1 + p_2 - p_1 \cdot p_2;$$
$$inv(p_1) = 1 - p_1.$$

Приведенные выражения означают вероятности конъюнкции и дизъюнкции двух независимых событий с вероятностями p_1 и p_2, а также событие отрицания исходного события. В этом смысле правила вычислений в узлах графа соответствуют правилам Байеса. Отметим, что качественный подход можно трактовать как частный случай количественного подхода с

$$D = \{F; P; N\}, \oplus = \min() \text{ и } \otimes = \max().$$

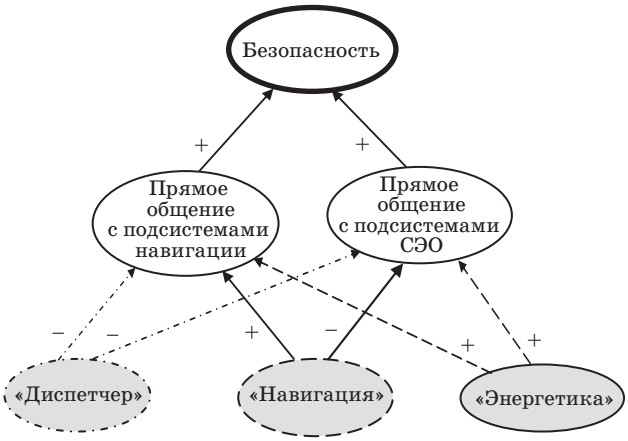
Отличительной особенностью количественного подхода к анализу по сравнению с качественным является то, что элементы Initial, Current и Old теперь выбираются численно из диапазона [0; 1]; целевой граф содержит также веса $+S$, $-S$, $+D$, $-D$.

Таким образом, в ходе выбора наиболее целесообразного стиля мультиагентной архитектуры были выработаны показатели, сформированы наиболее подходящие для АНПА с гибридной СЭО стили архитектуры, представлен целевой граф, определен метод оценки выработанных показателей для каждого рассматриваемого стиля архитектуры. По совокупности сформировался метод выбора архитектуры мультиагентной системы управления АНПА.

Пример применения разработанного метода выбора

Рассмотрим АНПА с гибридной СЭО, который в целях своей безопасности должен двигаться в различных скоростных режимах. Различные скоростные режимы движения АНПА связаны с необходимостью проходить узкости, в которых для повышения устойчивости движения и удержания курса следует повышать скорость. Это автоматически означает переключение с одного источника электроэнергии на другой. При этом должно быть понимание точного положения АНПА в узкости, поскольку высокоскоростной режим для АНПА является запредельным, и как только необходимость в нем пропадет, АНПА должен перейти на обычный скоростной режим.

Для АНПА с гибридной СЭО осуществим выбор стиля архитектуры, используя в качестве цели «Безопасность». Рассмотрим подграф (рис. 3) целевого графа (см. рис. 2), разделив показатель «Прямое общение между агентами» на два: «Прямое общение с подсистемами навигации» и «Прямое общение с подсистемами СЭО».



■ ***Рис. 3.*** Целевой граф «Безопасность»
■ ***Fig. 3.*** Security Target Graph

Основным источником опасности при использовании гибридной СЭО является повышение внутриотсечной температуры из-за нерационального подключения потребителей к токопроводам, запитываемым разнородными источниками электроэнергии. Вторым неблагоприятным фактором, оказывающим влияние на безопасность, является нехватка энергоресурса для прибытия АНПА в заданную конечную точку маршрута.

Проанализируем по показателю «Безопасность» каждый из трех рассматриваемых стилей архитектуры, отличающихся координатором: «Диспетчер», «Навигация», «Энергетика».

Координатор «Диспетчер» способен обеспечить наибольшую безопасность, так как владеет всей информацией, но поступающей с некоторым запаздыванием, так как информацию поставляют все агенты, ввиду чего возможно скопление, переполнение, задержки в обработке. Это может оказаться критичным в плане оперативности функционирования системы для предотвращения аварий. «Диспетчер» общается с агентами «Навигация» и «Энергетика» и не имеет прямого общения с подсистемами этих агентов.

Координатор «Навигация» обеспечивает меньшую безопасность по сравнению с координатором «Диспетчер», так как не владеет всей ситуацией в целом и не может сформировать интегральную оценку признаков опасности. Однако, с другой стороны, по напрямую поступающим данным способен прямо или косвенно оценить правильность или ошибочность поступающей информации и оперативно принять решение относительно вероятного возникновения опасности. В то же время отсутствие прямого доступа к подсистемам

гибридной СЭО не позволяет оперативно оценить основной источник опасности.

Координатор «Энергетика» не способен объективно оценить сложившуюся ситуацию из-за недостатка поступающей информации. Однако, учитывая то, что сама по себе гибридная СЭО является наиболее вероятным источником опасности, с этой точки зрения наиболее безопасно именно в ней разместить анализ предаварийной ситуации, поскольку обеспечен оперативный прием данных и их обработка, прямое общение с подсистемами гибридной СЭО, непосредственными устройствами, формирующими эти данные. Плюсы координатора «Энергетика» заключаются в том, что все ключевые данные, необходимые для безопасного движения АНПА вдоль маршрутной траектории, сосредоточены в агенте «Энергетика». Прямое общение агента «Энергетика» с подсистемами СЭО, входящими в него, исключают нагревание внутри корпуса АНПА, контролируют расход ресурса, оценивают его достаточность для прибытия в заданную точку и принимают решение о выборе скоростного режима движения АНПА, напрямую связанного с управлением маршевым двигателем. Поэтому и данные о потребляемой мощности маршевого двигателя также должны поступать координатору — агенту «Энергетика».

В связи со сказанным представляется наиболее предпочтительным назначить координатором агента «Энергетика». При этом за навигацией остается контроль соответствия положения АНПА заданному маршруту, который заранее проложен с учетом карты глубин, береговой черты, положения опасных зон.

Результаты анализа выразим на рис. 3 указанием рядом с каждым ребром графа положительного «+» или отрицательного «–» влияния особенностей стиля архитектуры на целевой показатель.

Результаты оценки стиля по показателям занесем в таблицу. Итоговый результат оценки выведен по правилам алгоритма LP.

■ Качественные оценки стилей мультиагентной системы управления

■ Qualitative estimate of multi-agent management system styles

Цель «Безопасность»	«Диспетчер»	«Навигация»	«Энергетика»
Прямое обращение к навигации	FD	FS	PS
Прямое обращение к СЭО	FD	FD	FS
Итого	FD	N	PS

Из приведенных результатов видно, что назначение координатором агента «Энергетика» для АНПА с гибридной СЭО позволило получить максимальный результат PS, поскольку FD < N < PS.

Однако если СЭО не гибридная, а однородная, то тогда основная опасность возникновения аварий связана с отклонением АНПА от маршрутной траектории и возможным его попаданием в неблагоприятные районы препятствий, течения, загрязнения и т. д. В этом случае отсутствует необходимость «Прямого общения с подсистемами СЭО», а вот «Прямое общение к подсистемам навигации» для принятия оперативного решения требуется. Поэтому предпочтение целесообразно отдать стилю архитектуры, в котором координатором будет являться агент «Навигация».

Заключение

Рассмотрены особенности функционирования АНПА с гибридной системой энергообеспечения и мультиагентной системой управления. В результате анализа выявлена необходимость выбора стиля архитектуры мультиагентной системы с использованием нефункциональных требований.

Для выбора стиля архитектуры мультиагентной системы управления разработан метод, включающий в себя формирование множества наиболее подходящих для АНПА показателей; множество стилей мультиагентных систем управления, наиболее подходящих для АНПА; получение оценки рассматриваемых стилей по показателям и итоговой оценки по каждому альтернативному стилю. Получение такой оценки позволяет осуществить выбор наиболее подходящего стиля мультиагентной архитектуры.

На примере анализа фрагмента целевого графа показано использование разработанного метода для выбора наилучшего решения.

Финансовая поддержка

Работа выполнена при финансовой поддержке РФФИ (проект 20-08-00130 а).

Литература

1. **Мартынова Л. А., Машошин А. И., Пашкевич И. В., Соколов А. И.** Система управления — наиболее сложная часть автономных необитаемых подводных аппаратов. *Морская радиоэлектроника*, 2015, № 4(54), с. 27–33.
2. **Мартынова Л. А., Машошин А. И., Пашкевич И. В.** Система поддержки разработки алгоритмов систе-

мы управления АНПА. *Известия ЮФУ. Технические науки*, 2015, № 10(171), с. 178–190.
3. **Chung L. K., Nixon B., Yu E., Mylopoulos J.** *Non-Functional Requirements in Software Engineering*. Kluwer Publishing, 2000. 441 p.
4. **Мартынова Л. А.** Решение задачи подводного наблюдения в условиях применения интеллектуальных помех. *Информационно-управляющие системы*, 2018, № 1, с. 31–41. doi:10.15217/issnl684-8853.2018.1.31
5. **Мартынова Л. А.** Метод эффективного удержания положения АНПА на маршрутной траектории при ведении сейсморазведки. *Информационно-управляющие системы*, 2018, № 3, с. 34–44. doi:10.15217/issn1684-8853.2018.3.34
6. **Мартынова Л. А., Карсаев О. В.** Метод координации поведения группы автономных необитаемых подводных аппаратов на мультиагентной основе при ведении сейсморазведки. *Известия ЮФУ. Технические науки*, 2018. № 1 (195), с. 52–67. doi:10.23683/2311-3103-2018-1-52-67
7. **Безрук Г. Г., Мартынова Л. А., Саенко И. Б.** Динамический способ поиска антропогенных объектов в морском дне с использованием автономных необитаемых подводных аппаратов. *Труды СПИИРАН*, 2018, вып. 58, с. 203–226. doi:10.15622/sp.58.9
8. **Martynova L. A., Bezruk G. G., Myslivyi A. A.** Application of differential mode for AUV location. *Информационно-управляющие системы*, 2018, № 4, с. 15–23. doi:10.31799/1684-8853-2018-4-15-23
9. **Пшихопов В. Х., Чернухин Ю. В., Федотов А. А., Гузик В. Ф., Медведев М. Ю., Гуренко Б. В., Пьявченко А. О., Сапрыкин Р. В., Переверзев В. А., Приемко А. А.** Разработка интеллектуальной системы управления автономного подводного аппарата. *Известия ЮФУ. Технические науки*, 2014, № 3, с. 87–101.
10. **Pshikhopov V. Kh., Medvedev M. Yu., Gaiduk A. R., Gurenko B. V.** Control system design for autonomous underwater vehicle. *Latin American Robotics Symposium and Competition*, Arequipa, Peru, 2013. doi:10.1109/LARS.2013.61
11. **Pshikhopov V., Chernukhin Y., Guzik V., Medvedev M., Gurenko B., Piavchenko A., Saprikin R., Pereversev V., Krukhmalev V.** Implementation of intelligent control system for autonomous underwater vehicle. *Applied Mechanics and Materials*, 2015, vol. 701–702, pp. 704–710. doi:10.4028/www.scientific.net/AMM. 701-702.704
12. **Gurenko B. V., Fedorenko R., Beresnev M., Saprykin R.** Development of simulator for intelligent autonomous underwater vehicle. *Applied Mechanics and Materials*, 2015, vol. 799–800, pp. 1001–1005. doi:http://dx.doi.org/10.4028/www.scientific.net/AMM.799-800.1001
13. **Kostukov V. A., Kulchenko A. E., Gurenko B. V.** Model parameters research procedure for underwater vehicle. *Proc. of XXXVI–XXXVII International Conference*, Novosibirsk, 2015, no. 11–12 (35), pp. 75–79.
14. **Инзарцев А. В., Львов О. Ю., Сидоренко А. В., Хмельнов Д. Б.** Архитектурные конфигурации систем управления АНПА. *Подводные исследования и робототехника*, 2006, № 1, с. 18–30.
15. **Киселев Л. В., Инзарцев А. В., Матвиенко Ю. В.** Создание интеллектуальных АНПА и проблемы интеграции научных исследований. *Подводные исследования и робототехника*, 2006, № 1, с. 6–17.
16. **Инзарцев А. В., Киселев Л. В., Костенко В. В., Матвиенко Ю. В., Павин А. М., Щербатюк А. Ф.** *Подводные робототехнические комплексы: системы, технологии, применение*. Владивосток, Ин-т проблем морских технологий ДВО РАН, 2018. 368 с.
17. **Боровик А. И., Наумов Л. А.** Компонентно-ориентированная система управления АНПА ММТ-2012. *Известия ЮФУ. Технические науки*, 2014, № 3, с. 102–112.
18. **Zhang L., Jiang D., Zhao J.** The basic control system of an ocean exploration AUV. *Applied Mechanics and Materials*, 2013, vol. 411–414, pp. 1757–1761. https://doi.org/10.4028/www.scientific.net/amm.411-414.1757
19. **Freire L. O., Oliveira L. M., Vale R. T. S., Medeiros M., Diana R. E. Y., Lopes R. M., Pellini E. L., de Barros E. A.** Development of an AUV control architecture based on systems engineering concepts. *Ocean Engineering*, 2018, vol. 151, pp. 157–169. https://doi.org/10.1016/j.oceaneng.2018.01.016.
20. **Aili A., Ekelund E.** *Model-based Design, Development and Control of an Underwater Vehicle*. MSc Thesis. Linköping University, 2016. 102 p.
21. **Blanke M., Lindegaard K.-P., Fossen T. I.** Dynamic model for thrust generation of marine propellers. *IFAC Proceedings Volumes*, 2000, no. 33(21), pp. 353–358.
22. **Deutsch C., Moratelli L., Thuné S., Kuttenkeuler J., Söderling F.** Design of an AUV research platform for demonstration of novel technologies. *2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)*, Nov 2018, pp. 1–8.
23. **Perez T.** *Ship Motion Control: Course Keeping and Roll Stabilisation Using Rudder and Fins*. Springer Science & Business Media, 2006. 109 p. doi:10.1007/1-84628-157-1
24. **Tanakitkorn K., Wilson Ph. A., Turnock S. R., Phillips A. B.** Depth control for an over-actuated, hover-capable autonomous underwater vehicle with experimental verification. *Mechatronics*, 2017, no. 41, pp. 67–81.
25. **Vervoort J. H. A. M.** *Modeling and Control of an Unmanned Underwater Vehicle*. MSc Thesis. University of Canterbury, 2008. 119 p.
26. **Wehbe B., Fabisch A., and Krell M. M.** Online model identification for underwater vehicles through incremental support vector regression. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, Vancouver, BC, Canada, pp. 4173–4180. doi:10.1109/IROS.2017.8206278
27. **Weiss J. D., Du Toit N. E.** Real-time dynamic model learning and adaptation for underwater vehicles. *2013 OCEANS*, San Diego, 2013, pp. 1–10.

28. **Yoerger D. R., Cooke J. G., Slotine J. J. E.** The influence of thruster dynamics on underwater vehicle behavior and their incorporation into control system design. *IEEE Journal of Oceanic Engineering*, 1990, no. 15(3), pp. 167–178.
29. **Yuh J., Marani G., Blidberg D. R.** Applications of marine robotic vehicles. *Intelligent Service Robotics*, 2011, no. 4(4), pp. 221–231.
30. **Yuh J.** Design and control of autonomous underwater robots. *A Survey. Autonomous Robots*, 2000, no. 8(1), pp. 7–24.
31. **Giorgini P., Mylopoulos J., Nicchiarelli E., Sebastiani R.** Reasoning with goal models. *Proceedings of the 21st International Conference on Conceptual Modeling (ER 2002)*, Tampere, Finland, October 2002. doi:10.1007/3-540-45816-6_22. https://www.researchgate.net/publication/226665392_Reasoning_with_Goal_Models. (дата обращения: 21.04.2019).
32. **Чеглаков А. Л., Нехотина В. С.** Оценка нефункциональных требований ИТ-архитектуры с использованием онтологий. *Национальная ассоциация ученых (НАУ)*, 2015, # IV (9), с. 44–46. https://national-science.ru/ (дата обращения: 21.04.2019).
33. **Giorgini P., Kolp M., Mylopoulos J.** Multi-agent architectures as organizational structures. *Autonomous Agent and Multi-Agent Systems*, 2006, no. 13, pp. 1–2. https://www.academia.edu/2731942/Multi-agent_architectures_as_organizational_structures (дата обращения: 21.04.2019).
34. **Innocenti Badano B. M.** A multi-agent architecture with distribution for an autonomous robot. *2009 Universitat de Girona*. https://www.tdx.cat/bitstream/handle/10803/7749/Tbi1de1.pdf;sequence = 1 (дата обращения: 13.08.2019).
35. **Buzun N., Korshunov A.** Innovative methods and measures in overlapping community detection. *Proceedings of International Workshop on Experimental Economics in Machine Learning 2012*, KU-Leuven, 2012, pp. 20–31.
36. **Malewicz G., Austern M. H., Bik A. J. C., Dehnert J. C., Horn I., Leiser N., Czajkowski G.** Pregel: a system for large-scale graph processing. *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data (SIGMOD '10)*, ACM, New York, NY, USA, 2010, pp. 135–145.
37. **Xin R., Gonzalez J., Franklin M., Stoica I.** GraphX: A resilient distributed graph system on spark. *GRADES (SIGMOD Workshop)*, 2013, pp. 1–6. https://doi.org/10.1145/2484425.2484427
38. **Raghavan U. N., Albert R., Kumara S.** Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 2007, vol. 76, no. 3, p. 036106.

UDC 519.87
doi:10.31799/1684-8853-2020-4-31-41

Choice of architecture for a multi-agent control system of an autonomous underwater vehicle

L. A. Martynova[a], Dr. Sc., Tech., Senior Researcher, orcid.org/0000-0002-5613-0838, martynowa999@bk.ru
N. K. Kiselev[b], First Deputy Chief Designer, orcid.org/0000-0002-5401-4470
A. A. Myslivyi[c], PhD, Militarhy, Deputy Head of Division, orcid.org/0000-0002-6741-3139
[a]Concern CSRI Elektropribor, JSC State Research Center of Russia, 30, Malaya Posadskaya St., 197046, Saint-Petersburg, Russian Federation
[b]JSC Central Design Bureau Lazurit, 57, Svobody St., 603951, Nizhnij Novgorod, Russian Federation
[c]Research Institute of OSIS Navy, 17, Razvodnaya St., 198516, Petergof, Sankt-Peterburg, Russian Federation

Introduction: The continuous improvement of autonomous underwater vehicles, the complexity of their systems and the use of a hybrid energy supply system have led to the need of developing a control system using multi-agent technology. To date, a large number of styles of multi-agent architectures have been formed, mainly in the field of organizing the manufacture and developing software. It is important to choose the most suitable architecture style for a multi-agent control system of an autonomous underwater vehicle with a hybrid energy supply system, taking into account its features. **Purpose:** The development of a method for choosing the most suitable style of a multi-agent architecture among a variety of alternative options. Method: The developed method is based on comparative assessment of various architecture styles according to non-functional requirements. For this purpose, a target graph is specially developed, taking into account the features of the device to be designed. In addition, when generating the final result, the label distribution algorithm was used as the most suitable one for this problem. **Results:** The proposed method of choosing the architecture style includes the following components: developing indicators by which it is advisable to compare the alternative options; forming various styles of architectures most suitable for the device under construction; analyzing the positive and negative effects of the architecture style according to non-functional requirements; formalizing these influences in the form of qualitative or quantitative labels; obtaining the final grade by applying the label distribution algorithm. **Practical relevance:** The proposed method allows you to select the most appropriate architecture for a multi-agent control system of an autonomous underwater vehicle. The method can also be used for a wider range of ground-based and air-based robotic systems.

Keywords — autonomous underwater vehicle, multi-agent control system architecture, performance evaluation, non-functional requirements, label propagation algorithm.

For citation: Martynova L. A., Kiselev N. K., Myslivyi A. A. Choice of architecture for a multi-agent control system of an autonomous underwater vehicle. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 31–41 (In Russian). doi:10.31799/1684-8853-2020-4-31-41

References

1. Martynova L. A., Mashoshin A. I., Pashkevich I. V., Sokolov A. I. Control system is the most complicated part of autonomous underwater vehicles. *Marine Radio electronics*, 2015, no. 4 (54), pp. 27–33 (In Russian).
2. Martynova L. A., Mashoshin A. I., Pashkevich I. V. The support system for design of AUV integrated control system algorithms. *Izvestiya SFedU. Engineering Sciences*, 2015, no. 10(171), pp. 178–190 (In Russian).
3. Chung L. K., Nixon B., Yu E., Mylopoulos J. *Non-Functional Requirements in Software Engineering*. Kluwer Publishing, 2000. 441 p.
4. Martynova L. A. The solution of the problem of underwater observation in the conditions of application of intellectual interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 1, pp. 31–41 (In Russian). doi:10.15217/issnl684-8853.2018.1.31
5. Martynova L. A. The method of effectively maintaining the position of the AUV on the route trajectory during seismic surveys. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 34–44 (In Russian). doi:10.15217/issn1684-8853.2018.3.34
6. Martynova L. A., Karsaev O. V. A method for coordinating the behavior of a group of autonomous underwater vehicles on a multi-agent basis during seismic surveys. *Izvestiya SFedU. Engineering Sciences*, 2018, no. 1(195), pp. 52–67 (In Russian). doi:10.23683/2311-3103-2018-1-52-67
7. Bezruk G. G., Martynova L. A., Saenko I. B. Dynamic method of searching anthropogenic objects in use of seabed with autonomous underwater vehicles. *SPIIRAS Proceedings*, 2018, no. 3(58), pp. 203–226 (In Russian). doi:10.15622/sp.58.9
8. Martynova L. A., Bezruk G. G., Myslivyi A. A. Application of differential mode for AUV location. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 4, pp. 15–23 (In Russian). doi:10.31799/1684-8853-2018-4-15-23
9. Pshikhopov V. Kh., Chernukhin Yu. V., Fedotov A. A., Guzik V. F., Medvedev M. Yu., Gurenko B. V., Piavchenko A. O., Saprikin R. V., Pereversev V. A., Priemko A. A. Development of intelligent control system for autonomous underwater vehicle. *Izvestiya SFedU. Engineering Sciences*, 2014, no. 3(152), pp. 87–101.
10. Pshikhopov V. Kh., Medvedev M. Yu., Gaiduk A. R., Gurenko B. V. Control system design for autonomous underwater vehicle. *Latin American Robotics Symposium and Competition*, Arequipa, Peru, 2013. doi:10.1109/LARS.2013.61
11. Pshikhopov V., Chernukhin Y., Guzik V., Medvedev M., Gurenko B., Piavchenko A., Saprikin R., Pereversev V., Krukhmalev V. Implementation of intelligent control system for autonomous underwater vehicle. *Applied Mechanics and Materials*, 2015, vol. 701–702, pp. 704–710. doi:10.4028/www.scientific.net/AMM.701-702.704
12. Gurenko B. V., Fedorenko R., Beresnev M., Saprykin R. Development of simulator for intelligent autonomous underwater vehicle. *Applied Mechanics and Materials*, 2015, vol. 799–800, pp. 1001–1005. doi:http://dx.doi.org/10.4028/www.scientific.net/AMM.799-800.1001
13. Kostukov V. A., Kulchenko A. E., Gurenko B. V. Model parameters research procedure for underwater vehicle. *Proc. of XXXVI–XXXVII International Conference*, Novosibirsk, 2015, no. 11–12 (35), pp. 75–79.
14. Inzartsev A. V., Lvov O. Yu., Sidorenko A. V., Khmelnov D. B. Architectural configurations of AUV control systems. *Underwater Investigations and Robotics*, 2006, no. 1, pp. 18–30 (In Russian).
15. Kiselev L. V., Inzartsev A. V., Matvienko Yu. V. The Creation of intelligent AUVs and the problems of the integration of scientific research. *Underwater Investigations and Robotics*, 2006, no. 1, pp. 6–17 (In Russian).
16. Inzartsev A. V., Kiselev L. V., Kostenko V. V., Matvienko Yu. V., Pavin A. M., Sherbatyuk A. F. *Underwater Robotics: System, Technologies, Application*. Vladivostok, IPMT FEB RAS Publ., 2018. 368 p. (In Russian).
17. Borovik A. I., Naumov L.A. Component-oriented management system AUV MMT-2012. *Izvestiya SFedU. Engineering Sciences*, 2014, no. 3, pp. 102–112 (In Russian).
18. Zhang L., Jiang D., Zhao J. The basic control system of an ocean exploration AUV. *Applied Mechanics and Materials*, 2013, no. 411–414, pp. 1757–1761. https://doi.org/10.4028/www.scientific.net/amm.411-414.1757
19. Freire L. O., Oliveira L. M., Vale R. T. S., Medeiros M., Diana R. E. Y., Lopes R. M., Pellini E. L., de Barros E. A. Development of an AUV control architecture based on systems engineering concepts. *Ocean Engineering*, 2018, vol. 151, pp. 157–169. https://doi.org/10.1016/j.oceaneng.2018.01.016.
20. Aili A., Ekelund E. *Model-based Design, Development and Control of an Underwater Vehicle*. MSc Thesis. Linköping University, 2016. 102 p.
21. Blanke M., Lindegaard K.-P., Fossen T. I. Dynamic model for thrust generation of marine propellers. *IFAC Proceedings Volumes*, 2000, no. 33(21), pp. 353–358.
22. Deutsch C., Moratelli L., Thuné S., Kuttenkeuler J., Söderling F. Design of an AUV research platform for demonstration of novel technologies. *2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)*, Nov 2018, pp. 1–8.
23. Perez T. *Ship Motion Control: Course Keeping and Roll Stabilisation Using Rudder and Fins*. Springer Science & Business Media, 2006. 109 p. doi:10.1007/1-84628-157-1
24. Tanakitkorn K., Wilson Ph. A., Turnock S. R., Phillips A. B. Depth control for an over-actuated, hover-capable autonomous underwater vehicle with experimental verification. *Mechatronics*, 2017, no. 41, pp. 67–81.
25. Vervoort J. H. A. M. *Modeling and Control of an Unmanned Underwater Vehicle*. MSc Thesis. University of Canterbury, 2008. 119 p.
26. Wehbe B., Fabisch A., Krell M. M. Online model identification for underwater vehicles through incremental support vector regression. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, Vancouver, BC, Canada, pp. 4173–4180. doi:10.1109/IROS.2017.8206278
27. Weiss J. D., Du Toit N. E. Real-time dynamic model learning and adaptation for underwater vehicles. *2013 OCEANS*, San Diego, 2013, pp. 1–10.
28. Yoerger D. R., Cooke J. G., Slotine J. J. E. The influence of thruster dynamics on underwater vehicle behavior and their incorporation into control system design. *IEEE Journal of Oceanic Engineering*, 1990, no. 15(3), pp. 167–178.
29. Yuh J., Marani G., Blidberg D. R. Applications of marine robotic vehicles. *Intelligent Service Robotics*, 2011, no. 4(4), pp. 221–231.
30. Yuh J. Design and control of autonomous underwater robots. *A Survey. Autonomous Robots*, 2000, no. 8(1), pp. 7–24.
31. Giorgini P., Mylopoulos J., Nicchiarelli E., Sebastiani R. Reasoning with goal models. *Proceedings of the 21st International Conference on Conceptual Modeling (ER 2002)*, Tampere, Finland, October 2002. doi:10.1007/3-540-45816-6_22. Available at: https://www.researchgate.net/publication/226665392_Reasoning_with_Goal_Models. (accessed 21 April 2019).
32. Cheglakov A. L., Nekhotina V. S. Assessment of non-functional requirements of IT architecture using ontologies. *National Science Journal*, 2015, # IV (9), pp. 44–46. Available at: https://national-science.ru/ (accessed 21 April 2019).
33. Giorgini P., Kolp M., Mylopoulos J. Multi-agent architectures as organizational structures. *Autonomous Agent and Multi-Agent Systems*, 2006, no. 13, pp. 1–2. Available at: https://www.academia.edu/2731942/Multi-agent_architectures_as_organizational_structures (accessed 21 April 2019).
34. Innocenti Badano B. M. A multi-agent architecture with distribution for an autonomous robot. *2009 Universitat de Girona*. Available at: https://www.tdx.cat/bitstream/handle/10803/7749/Tbi1de1.pdf;sequence = 1 (accessed 13 August 2019).
35. Buzun N., Korshunov A. Innovative methods and measures in overlapping community detection. *Proceedings of International Workshop on Experimental Economics in Machine Learning 2012*, KU-Leuven, 2012, pp. 20–31.
36. Malewicz G., Austern M. H., Bik A. J. C, Dehnert J. C., Horn I., Leiser N., Czajkowski G. Pregel: a system for large-scale graph processing. *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data (SIGMOD '10)*, ACM, New York, NY, USA, 2010, pp. 135–145.
37. Xin R., Gonzalez J., Franklin M., Stoica I. GraphX: A resilient distributed graph system on spark. *GRADES (SIGMOD Workshop)*, 2013, pp. 1–6. https://doi.org/10.1145/2484425.2484427
38. Raghavan U. N., Albert R., Kumara S. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 2007, vol. 76, no. 3, p. 036106.

УДК 004.942
doi:10.31799/1684-8853-2020-4-42-49

Научные статьи
Articles

Количественный анализ программы для управления бортовой вычислительной сетью

К. Н. Рождественская[a], *ассистент, orcid.org/0000-0003-4930-6898, rogdkn@yandex.ru*
[a]*Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ*

Постановка проблемы: *программа для управления бортовой вычислительной сетью принимает, отправляет и обрабатывает большой объем данных. При этом не должны возникать перегрузки или замедление процесса обработки данных, так как программа ответственна за своевременную обработку возможных критических ситуаций. Следовательно, необходимо провести количественный анализ программы для исследования ее динамических характеристик.* ***Цель:*** *анализ динамического поведения программы для управления бортовой вычислительной сетью при различных начальных условиях и внешних воздействиях.* ***Результаты:*** *построена линейная динамическая система, описывающая часть конкретной программы для управления бортовой вычислительной сетью и являющаяся известным менеджером Plug-and-Play. Она представлена в виде графа переходов, отражающего относительные доли данных, поступающие от одного состояния к другому. Выбраны три ситуации, при которых программа для управления бортовой вычислительной сетью может испытывать перегрузки или замедления. Определены системные матрицы коэффициентов, входа и выхода для каждого из рассматриваемых случаев и способы избегания критических ситуаций. Проведено компьютерное моделирование с помощью программы-сценария, созданной в математическом пакете MatLab. Результатами моделирования являются графики загрузки и управления программой во времени. Загрузка и управление меняются в зависимости от целевой загрузки и вычислительных способностей линейной динамической системы. Количественный анализ, представленный в работе, выполняет расчет поведения менеджера Plug-and-Play во времени при его конкретных характеристиках.* ***Практическая значимость:*** *представленные формулы и созданная программа-сценарий позволяют промоделировать работу программы управления бортовой вычислительной сетью с разными характеристиками нагрузки, зависящими от аппаратной и программной реализации в конкретном проекте, для предотвращения сбоев и отказов работы программы и оборудования в ходе эксплуатации.*

Ключевые слова *— администрирование бортовой сети, мониторинг бортовой сети, Plug-and-Play, менеджер, линейная динамическая система, количественный анализ, MatLab.*

Для цитирования: Рождественская К. Н. Количественный анализ программы для управления бортовой вычислительной сетью. *Информационно-управляющие системы*, 2020, № 4, с. 42–49. doi:10.31799/1684-8853-2020-4-42-49
For citation: Rozhdestvenskaya K. N. Quantitative analysis of an onboard computer network administration program. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 42–49 (In Russian). doi:10.31799/1684-8853-2020-4-42-49

Введение

Программа для управления бортовой вычислительной сетью выполняет сбор данных о сетевых компонентах, анализ полученных данных и корректировку работы сетевых компонентов [1]. В работах [1, 2] дано краткое описание, обоснование и обзор выбранного направления исследования управления сетями SpaceWire [3–7]. В настоящей статье основное внимание уделяется количественному анализу менеджера Plug-and-Play (PnP) для сети SpaceWire [8–12]. Взаимодействие бортовой вычислительной сети и программы для управления ею в виде линейной динамической системы (ЛДС) подробно анализируется в настоящей статье. Предлагается количественный анализ представленной части программы для управления вычислительной сетью по шагам с помощью изучения динамических характеристик, определяемых поступлением определенного количества данных. Для подтверждения результатов используется программа-сценарий, написанная в математическом пакете MatLab.

Линейная динамическая система взаимодействия менеджера PnP с сетью

Взаимодействие программы с управляемой сетью происходит посредством отправки и приема пакетов данных [7]. Граф переходов с состояниями, отражающими описанное взаимодействие, представлен на рис. 1 [1].

Состояния представленного графа следующие: C_1 — формирование команды; C_2 — прием пакета; C_3 — обработка пакета; C_4 — обновление служебных структур [1].

На ребрах графа переходов указаны коэффициенты, которые отражают следующие доли данных, поступающих на соответствующее состояние:

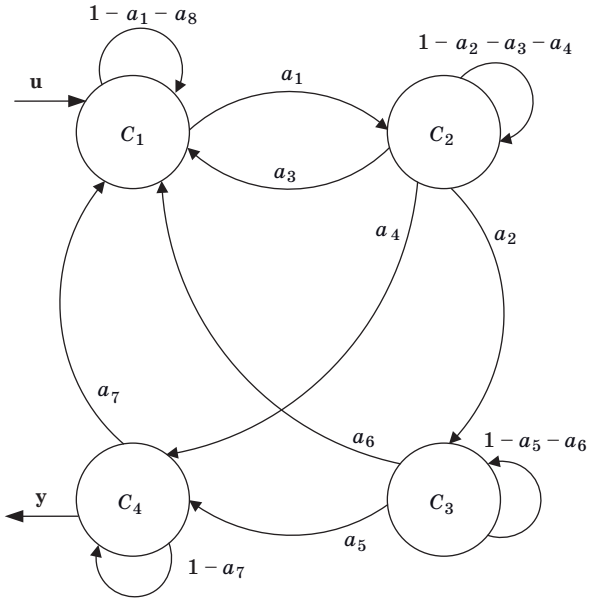
a_1 — доля данных, принимаемых из системы;

a_2 — доля принятых данных, требующих обработки;

a_3 — доля данных, которые требуют повторного запроса;

a_4 — доля данных, которые не были приняты;

a_5 — доля данных, которые были успешно обработаны;



■ ***Рис. 1.*** Линейная динамическая система взаимодействия программы с управляемой сетью

■ ***Fig. 1.*** The linear dynamical system describing interact program with controlled network

a_6 — доля данных, которые не обработаны;

a_7 — доля данных, которые требуют запроса;

a_8 — доля данных, возвращаемых в сеть.

Определим переменные, которые используются в математических моделях ЛДС. Вектор **u** представляет управление в ЛДС как входные данные системы. Вектор **y** — обработанное количество данных, определяемое как выходные данные системы. Вектор $\mathbf{x}(t) = [x_1(t); x_2(t); x_3(t); x_4(t)]$ определяет состояния ЛДС как количество данных в системе. Количественный анализ предполагает, что все представленные переменные имеют относительный характер, поэтому принадлежат интервалу [0; 1]. То есть, если в каком-либо состоянии количество данных определено как 1, то это состояние загружено полностью. Можно сказать, что совокупность конкретных числовых значений всех переменных полностью определяет состояние системы в заданный момент времени. Для исследования динамического поведения системы при различных условиях и воздействиях необходимо определить системные матрицы коэффициентов, входа и выхода [13].

Матрица коэффициентов строится на основе графа переходов:

$$\mathbf{A} = \begin{bmatrix} 1-a_1-a_8 & a_3 & a_6 & a_7 \\ a_1 & 1-a_2-a_3-a_4 & 0 & 0 \\ 0 & a_2 & 1-a_5-a_6 & 0 \\ 0 & a_4 & a_5 & 1-a_7 \end{bmatrix}.$$

Матрица входов имеет такой вид:

$$\mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

так как именно в состоянии системы C_1 происходит поставка пакетов из бортовой вычислительной сети.

Матрица выхода

$$\mathbf{C} = [0 \quad 0 \quad 0 \quad 1].$$

Матрица выхода имеет указанный вид, так как в состоянии системы C_4 происходит обновление служебных структур и принимается решение о дальнейшем ходе работы менеджера PnP. В состоянии C_4 определяется доля обработанных данных.

Задачей исследования ЛДС менеджера PnP является получение динамических характеристик. К ним относятся устойчивость, управляемость и наблюдаемость [13–17]. Для выполнения поставленной задачи определим способы вычисления каждой из характеристик.

Устойчивость определяет стабильную работу системы не только в нормальном режиме, но и при отклонении от нормы, при влиянии внешних воздействий. Она является также способностью системы возвращаться к уравновешенному состоянию. Для вычисления этой динамической характеристики необходимо найти спектральный радиус матрицы коэффициентов [15, 18]. Условие устойчивости не будет выполняться, если спектральный радиус больше единицы, в этом случае происходит перегрузка менеджера PnP, и он не справится с обработкой пакетов данных из сети. То есть положительный спектральный радиус, меньший единицы, означает устойчивую работу менеджера PnP, но чем ближе значение к единице, тем выше нагрузка.

Управляемость ЛДС показывает, можно ли ее перевести из любого начального состояния в любое конечное состояние за заданное число шагов [15, 16]. Другими словами, можно ли начать работу с любой начальной загрузки и перейти к любой конечной загрузке. Для управляемой ЛДС можно определить такую поставку данных в систему, при которой будет выполнен описанный переход. Условие управляемости требует построить матрицу управляемости [19, 20]

$$\mathbf{G} = \left[\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \mathbf{A}^3\mathbf{B}\right].$$

Ранг матрицы **G** должен быть равен размерности пространства состояний, то есть четырем.

Тогда ЛДС является управляемой. Для обеспечения перехода из начального состояния в любое заданное конечное состояние за k шагов построим матрицу достижимости [14, 17, 19]

$$\mathbf{G}_k = \left[\mathbf{B},\ \mathbf{AB},\ \ldots,\ \mathbf{A}^{k-1}\mathbf{B}\right].$$

Программу поставки данных можно представить в виде вектора управления

$$\mathbf{u}(k-1,0) = \left[\mathbf{u}(k-1);\ \mathbf{u}(k-2),\ \ldots,\ \mathbf{u}(0)\right].$$

Можно получить линейное алгебраическое уравнение

$$\mathbf{G}_k\mathbf{u}(k-1,0) = \mathbf{x}(k) - \mathbf{A}^k\mathbf{x}(0).$$

Его решение будет иметь вид

$$\mathbf{u}(k-1,0) = \mathbf{G}_k^+(\mathbf{x}(k) - \mathbf{A}^k\mathbf{x}(0),$$

где $\mathbf{G}_k^+$ — псевдообратная матрица матрицы $\mathbf{G}_k$.

Расчет поставки данных позволит определить такое управление менеджером PnP, которое не приведет к простою (значение меньше нуля) и не перегрузит (значение превышает единицу) ни одного из его состояний. Простой или перегрузка будут означать потерю производительности и могут привести к выходу из строя менеджера PnP. Расчет поставки данных в систему по полученным формулам произведен в программе-сценарии, написанной с помощью математического пакета MatLab.

Наблюдаемость ЛДС определяет, можно ли, наблюдая только за изменениями выхода, восстановить начальное состояние системы [14, 17]. В исследуемом случае ЛДС наблюдаема тогда и только тогда, когда выполняется условие

$$rank([\mathbf{C}u^{\mathrm{T}},\ \mathbf{A}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}},\ \left(\mathbf{A}^{\mathrm{T}}\right)^2\mathbf{C}^{\mathrm{T}},\ \left(\mathbf{A}^{\mathrm{T}}\right)^3\mathbf{C}^{\mathrm{T}})] = 4.$$

Оценка начального состояния может быть выполнена по четырем сообщениям выхода $\mathbf{y}(t)$ следующим образом:

$$\hat{\mathbf{x}}(0) = \mathbf{H}^{-1}\begin{bmatrix}\mathbf{y}(0)\\ \mathbf{y}(1)\\ \mathbf{y}(2)\\ \mathbf{y}(3)\end{bmatrix}.$$

Данная динамическая характеристика позволит сделать вывод о процессах, происходящих внутри менеджера PnP, для проверки в ходе его эксплуатации. Чтобы это было возможно сделать, необходимо, чтобы ранг матрицы наблюдаемости был равен рангу матрицы коэффициентов. Пример оценки начального состояния системы также выполнен в программе-сценарии, написанной с помощью математического пакета MatLab.

Для исследования динамических характеристик определим следующие три ситуации.

1. Линейная динамическая система принимает и отправляет относительно большой объем данных, обрабатывает его, но отдает в бортовую вычислительную сеть только половину обработанных данных. Такая ситуация может быть связана с невысокой производительностью выходного порта устройства или комплексным механизмом передачи информации [21]. Пакеты данных, поступающие в ЛДС, являются корректными, т. е. повторные команды в сеть не отправляются. В такой ситуации коэффициенты соответствующей матрицы имеют вид: $a_1 = 0{,}9$, $a_2 = 0{,}7$, $a_3 = 0{,}1$, $a_4 = 0{,}1$, $a_5 = 0{,}5$, $a_6 = 0{,}1$, $a_7 = 0{,}5$, $a_8 = 0{,}45$.

2. Изменим относительный объем доли данных, которые передаются между состояниями в ЛДС, предположим, что данных поступает меньше, чем менеджер PnP может обработать. Такая ситуация может возникнуть, когда менеджер обладает высокой вычислительной способностью или устройств в сети немного. Коэффициенты приобретают вид: $a_1 = 0{,}5$, $a_2 = 0{,}3$, $a_3 = 0{,}1$, $a_4 = 0{,}1$, $a_5 = 0{,}8$, $a_6 = 0{,}1$, $a_7 = 0{,}9$, $a_8 = 0{,}5$.

3. Изменим соотношение относительных объемов получаемой и обрабатываемой информации следующим образом: ЛДС принимает данных больше, чем может обработать. Такая ситуация возможна, когда устройств в сети слишком много или в сети постоянно случаются критические ситуации из-за некачественного оборудования. Коэффициенты приобретают вид: $a_1 = 0{,}5$, $a_2 = 0{,}9$, $a_3 = 0{,}1$, $a_4 = 0{,}1$, $a_5 = 0{,}3$, $a_6 = 0{,}1$, $a_7 = 0{,}2$, $a_8 = 0{,}5$.

Компьютерное моделирование взаимодействия менеджера PnP с бортовой вычислительной сетью

Выполним компьютерное моделирование первой указанной выше ситуации, когда ЛДС принимает и отправляет относительно большой объем данных, обрабатывает его, но отдает в бортовую вычислительную сеть только половину обработанных данных. Матрица переходов приобретает вид

$$\mathbf{A} = \begin{bmatrix}0{,}1 & 0{,}1 & 0{,}1 & 0{,}5\\ 0{,}9 & 0{,}1 & 0 & 0\\ 0 & 0{,}7 & 0{,}4 & 0\\ 0 & 0{,}1 & 0{,}5 & 0{,}05\end{bmatrix}.$$

Спектральный радиус такой матрицы равен 0,919657. ЛДС является устойчивой к внешним воздействиям, но степень устойчивости очень мала. Значит, менеджер PnP работает почти на пределе своих вычислительных возможностей.

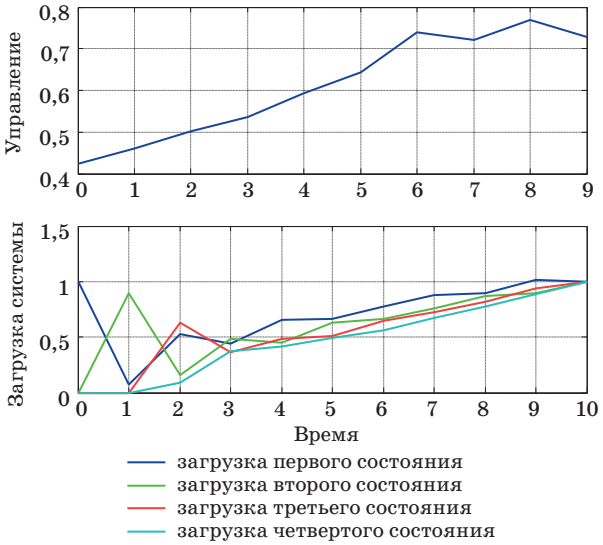
Матрица управляемости для первой ситуации имеет вид

$$\mathbf{G} = \begin{bmatrix} 1 & -0{,}35 & 0{,}21 & 0{,}01 \\ 0 & 0{,}9 & -0{,}22 & 0{,}17 \\ 0 & 0 & 0{,}63 & 0{,}94 \\ 0 & 0 & 0{,}09 & 0{,}34 \end{bmatrix}.$$

Ее ранг равен четырем. Значит, ЛДС управляема, и возможен переход от любого начального состояния в любое конечное состояние. На рис. 2 показано управление $\mathbf{u}(t)$ при $\mathbf{x}(k) = [1; 1; 1; 1]$. На рисунке видно, как управление плавно переводит систему к требуемой загрузке $\mathbf{x}(k)$, при этом не перегружая систему, оставаясь в допустимом пределе [0,4; 0,8].

Наблюдаемость ЛДС в представленной ситуации будет присутствовать, так как $rank([\mathbf{C}^T, \mathbf{A}^T\mathbf{C}^T, (\mathbf{A}^T)^2\mathbf{C}^T, (\mathbf{A}^T)^3\mathbf{C}^T)] = 4$. Следовательно, по измерениям данных на выходе системы можно будет восстановить ее состояние в начальный момент времени.

Выполним моделирование перехода системы из полностью загруженного первого состояния и пустых остальных состояний в полностью загруженные все состояния. На рис. 2 видно, что система под действием рассчитанного управления начала работу из первого состояния и закончила работу полной загрузкой всей системы.



■ ***Рис. 2.*** Управление и загрузка в системе для первой выбранной ситуации

■ ***Fig. 2.*** Control and loading graph in system for first situation

Для второй выбранной ситуации, когда данных поступает меньше, чем менеджер PnP может обработать, матрица коэффициентов становится следующей:

$$\mathbf{A} = \begin{bmatrix} 0{,}5 & 0{,}1 & 0{,}1 & 0{,}9 \\ 0{,}5 & 0{,}5 & 0 & 0 \\ 0 & 0{,}3 & 0{,}1 & 0 \\ 0 & 0{,}1 & 0{,}8 & -0{,}4 \end{bmatrix}.$$

Спектральный радиус изменится и станет равным 0,863485. Значит, ЛДС станет более устойчивой, и перегрузки менеджера PnP в данном случае не будет. Ранг матрицы управляемости по-прежнему равен четырем:

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0{,}05 & 0{,}08 \\ 0 & 0{,}5 & 0{,}25 & 0{,}15 \\ 0 & 0 & 0{,}15 & 0{,}09 \\ 0 & 0 & 0{,}05 & 0{,}15 \end{bmatrix}.$$

Следовательно, существует такая поставка данных, которая приведет от любого начального состояния в любое конечное, как произошло и в первом случае.

Наблюдаемость ЛДС во втором случае сохраняется, поскольку $rank([\mathbf{C}^T, \mathbf{A}^T\mathbf{C}^T, (\mathbf{A}^T)^2\mathbf{C}^T), (\mathbf{A}^T)^3\mathbf{C}^T)] = 4$.

Третья рассматриваемая ситуация возникает, когда ЛДС принимает данных больше, чем может обработать. Матрица коэффициентов для такого случая выглядит следующим образом:

$$\mathbf{A} = \begin{bmatrix} 0 & 0{,}1 & 0{,}1 & 0{,}2 \\ 0{,}5 & -0{,}1 & 0 & 0 \\ 0 & 0{,}9 & 0{,}6 & 0 \\ 0 & 0{,}1 & 0{,}3 & 0{,}8 \end{bmatrix}.$$

Здесь спектральный радиус равен 0,923564, что означает, что менеджер PnP близок к перегрузке.

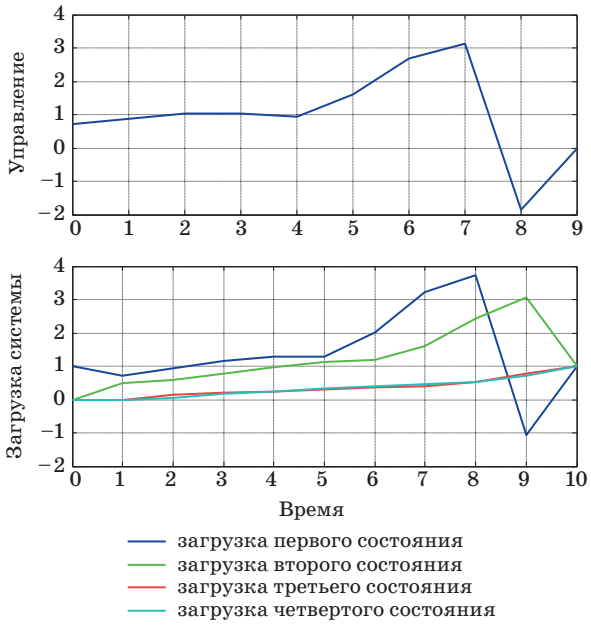
Линейная динамическая система остается управляемой, так как ранг матрицы $\mathbf{G}$ равен четырем, и наблюдаемой, так как $rank([\mathbf{C}^T, \mathbf{A}^T\mathbf{C}^T, (\mathbf{A}^T)^2\mathbf{C}^T), (\mathbf{A}^T)^3\mathbf{C}^T)] = 4$.

При расчете управления необходимо учитывать корректность целевой загрузки. Пример ошибочной работы ЛДС представлен на рис. 3, когда коэффициенты следующие: $a_1 = 0{,}5$, $a_2 = 0{,}3$, $a_3 = 0{,}1$, $a_4 = 0{,}1$, $a_5 = 0{,}8$, $a_6 = 0{,}1$, $a_7 = 0{,}9$, $a_8 = 0{,}5$, — и требуемая целевая загрузка $\mathbf{x}(10) = [1; 1; 1; 1]$.

За счет управления ЛДС пытается загрузить первое и второе состояния, чтобы объем поступающих данных на третье и четвертое состояния достиг требуемой загрузки, но после 4-го шага

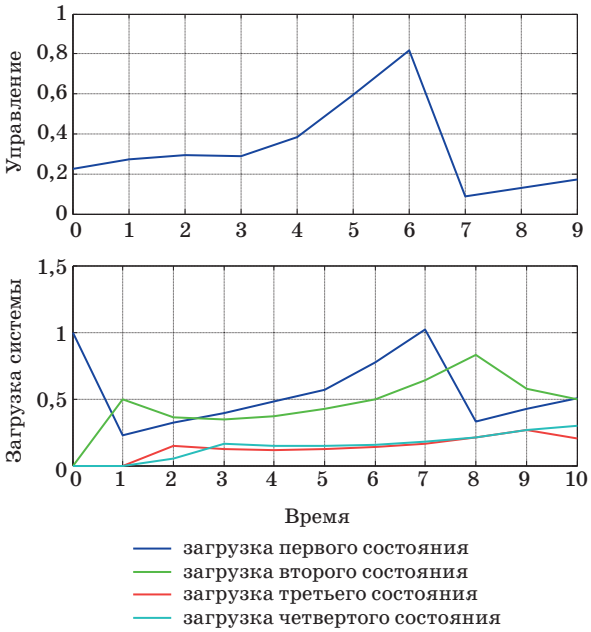
управление становится отрицательным. Это недопустимый вариант управления.

При уменьшении целевой загрузки до $\mathbf{x}(t) = [0{,}5;\ 0{,}5;\ 0{,}2;\ 0{,}3]$ управление изменится до приемлемого вида (рис. 4).



■ ***Рис. 3.*** Ошибочное управление и загрузка системы

■ ***Fig. 3.*** Erroneous incorrect control and loading system



■ ***Рис. 4.*** Приемлемое управление и загрузка системы

■ ***Fig. 4.*** Acceptable control and loading system

Таким образом, необходимо согласовывать уровни загрузки для тех состояний системы, на которые поступает больший объем относительных данных. С другой стороны, возможно перераспределение относительного объема данных, передающегося между состояниями, при сохранении целевой загрузки $\mathbf{x}(10) = [1;\ 1;\ 1;\ 1]$.

Заключение

Исследование менеджера PnP как программы для управления бортовой вычислительной сетью начато с временного анализа, основанного на конечных автоматах [1, 22]. Данная работа является продолжением исследования, предлагается количественный анализ на основе современной теории управления. Исследование продолжится в следующей работе, которая будет представлять вероятностный анализ программы для управления бортовой вычислительной сетью на основе марковских процессов.

Для проведения количественного анализа менеджера PnP в части его взаимодействия с бортовой вычислительной сетью была построена линейная динамическая система, определены и представлены исследуемые динамические характеристики, определены три ситуации, для которых производится компьютерное моделирование. В трех исследованных ситуациях отсутствует накопление данных в работе менеджера PnP, взаимодействие программы для управления с сетью управляемо. Для него может быть рассчитана поставка данных, обеспечивающая любую загрузку, и, следя только за состоянием ЛДС при обработке данных, можно рассчитать нагрузку всех четырех состояний программы. Показано, что некорректная целевая загрузка может привести к некорректному управлению, которое перегрузит отдельные состояния и может привести к критичным ситуациям в работе менеджера PnP.

Выполнено компьютерное моделирование с помощью математического пакета MatLab. Результаты моделирования представлены в виде графиков, отражающих во времени загрузку системы и ее управление. Результаты моделирования подтвердили и дополнили полученные расчеты.

Количественный анализ и созданная программа-сценарий позволят выполнить расчет загрузки конкретной бортовой вычислительной сети, на которую оказывают влияние программная и аппаратная составляющие, а также произвести моделирование ее поведения во времени, чтобы предугадать сбои и отказы программы управления при несбалансированной нагрузке или недостаточных аппаратных или программных ресурсов.

Литература

1. **Рождественская К. Н.** Временной анализ системы управления в сети обработки данных. *Информационно-управляющие системы*, 2019, № 1, с. 32–39. doi:10.31799/1684-8853-2019-1-32-39
2. **Шейнин Ю. Е., Рождественская К. Н., Евдокимов А. С., Дымов Д. В., Кочура С. Г.** SpaceWire-Plug-and-Play для перспективных бортовых сетей КА АО «ИСС». *Современные проблемы радиоэлектроники*, 2018, с. 196–200. http://efir.sfu-kras.ru/downloads/sbornik-spr-2018.pdf (дата обращения: 03.02.2020).
3. Space engineering — SpaceWire — Links, nodes, routers and networks. ECSS-E-ST-50-12C Rev.1 DIR3, November 23, 2015. ESA Requirements and Standards Division, 2015. 124 p.
4. **Clancy S. C., Chase M. D., Yarlagadda A., Starch M. D., Lux J. P.** SpaceWire as a Cube-Sat instrument interface. *Proc. 8th Intern. SpaceWire Conf.*, 2018, Los Angeles, USA, 2018, pp. 26–30.
5. **Windsor J., Gasti W., Clerigo I.** BepiColombo — building a robust data management subsystem utilizing SpaceWire networks. *7th Intern. SpaceWire Conf.*, 2016, Yokohama, Japan, 2016, pp. 112–119.
6. **Tomitaka M., Igarashi Y., Ichikawa S., Inaba N., Tomiki A., Matsuzaki K., Kobayashi R., Kumakiri M., Fujishiro I., Nomachi M.** Feasibility study of wireless communication system operating on SpaceWire network. *Proc. 8th Intern. SpaceWire Conf.*, 2018, Los Angeles, USA, 2018, pp. 118–122.
7. **Siegle F., Leoni A.** Standardization efforts for a network management and discovery protocol for SpaceFibre. *Proc. 8th Intern. SpaceWire Conf.*, 2018, Los Angeles, USA, 2018, pp. 133–137.
8. **Sheynin Y. E., Suvorova E. A., Rozhdestvenskaya K. N.** Management in perspective distributed onboard computing systems based on SpaceWire standard. *Wave Electronics and its Application in Information and Telecommunication Systems*, 2019, pp. 1–5. doi:10.1109/WECONF.2019.8840122
9. **Рождественская К. Н., Евдокимов А. С.** Архитектура и организация сети SpaceWire при применении протокола SpaceWire-Plug-and-Play. *Научная сессия ГУАП*, 2017, с. 198–203.
10. **Новиков В. М., Платошин Г. А., Шейнин Ю. Е.** Особенности применения интерфейса SpaceWire в комплексах бортового оборудования. *Труды ГОСНИИАС. Серия: Вопросы Авионики*, 2018, № 7(40), с. 41–69.
11. **Шейнин Ю. Е., Оленев В. Л., Лавровская И. Я., Дымов Д. В., Кочура С. Г.** Протоколы для бортовых сетей перспективных космических аппаратов на основе технологий SpaceWire и SpaceWire-Plug-and-Play. *Решетнёвские чтения*, 2016, № 1, с. 651–652.
12. **Romanowski K., Tyczka P., Holubowicz W., Renk R., Kollias V. D., Pogkas N., Jameux D.** SpaceWire network management using network discovery and configuration protocol. *Proc. 7th Intern. SpaceWire Conf.*, 2016, Yokohama, Japan, 2016, pp. 45–50.
13. **Бураков В. В.** *Модели оценивания и алгоритмы управления качеством программных средств*: автореф. дис. ... доктора техн. наук/ СПб., ГУАП, Санкт-Петербург, 2010. 42 с.
14. **Цюцзе Ю., Гунбо Л.** Реализация метода пространства состояний системы в среде MathCad. *Молодежь и современные информационные технологии: сборник трудов XIII Международной научно-практической конференции студентов, аспирантов и молодых ученых*, 2016, с. 302–303. http://earchive.tpu.ru/handle/11683/16943 (дата обращения: 03.02.2020).
15. **Зубов Н. Е., Микрин Е. А., Рябченко В. Н.** *Матричные методы в теории и практике систем автоматического управления летательных аппаратов*. М., Московский государственный технический университет им. Н. Э. Баумана, 2016. 672 с.
16. **Шумафов М. М.** Стабилизация линейных систем управления. Проблема назначения полюсов. Обзор. *Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия*, 2019, т. 6 (64), вып. 4, с. 564–591. https://doi.org/10.21638/11701/spbu01.2019.404 (дата обращения: 03.02.2020).
17. **Фокеева Л. Х., Богданов Х. У., Аблулкина Н. В.** Определение управляемости и наблюдаемости при исследовании систем. *Материалы научной сессии ученых Альметьевского государственного нефтяного института*, 2016, № 2, с. 39–42.
18. **Борисов Д. В.** Динамический анализ состояния системы управления. *World Science: Problems and Innovations: сборник статей XXVII Международной научно-практической конференции*, 2018, с. 86–90.
19. **Бойков И. В.** Достаточные условия устойчивости систем обыкновенных дифференциальных уравнений с запаздываниями, зависящими от времени. *Известия вузов. Поволжский регион. Физико-математические науки. Часть I. Линейные уравнения*, 2018, № 4 (48), с. 3–19.
20. **Гукасян А. А.** О математическом моделировании процесса обслуживания и условия ее управляемости. *Известия Национальной академии наук РА. Механика*, 2017, № 3(70), с. 26–38.
21. **Olenev V. L., Podgornova E., Lavrovskaya I. I., Sheynin Yu. E.** Deterministic services for SpaceWire networks. *Proc. 7th Intern. SpaceWire Conf.*, 2016, Yokohama, Japan, 2016, pp. 159–166.
22. **Рождественская К. Н.** Анализ поведения менеджера Plug-and-Play для сетей SpaceWire. *Научная сессия ГУАП*, 2018, с. 225–232.

UDC 004.942
doi:10.31799/1684-8853-2020-4-42-49

Quantitative analysis of an onboard computer network administration program

K. N. Rozhdestvenskaya[a], Assistant Professor, orcid.org/0000-0003-4930-6898, rogdkn@yandex.ru
[a]Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation

Introduction: An onboard computer network administration program receives, sends and processes a huge amount of data. The data processing should not be overloaded or slowed down, as the program is supposed to deal with possible critical situations on time. Therefore, the program should be subject to quantitative analysis in order to study its dynamic characteristics. **Purpose:** Analyzing the dynamic behavior of an onboard computer network administration program under various initial conditions and external influences. **Results:** A linear dynamic system was constructed, specifying a part of a particular onboard computer network administration program, representing a well-known plug-and-play manager. It is presented as a transition graph, showing relative parts of data coming from one state to another. Three situations were selected in which the program may be overloaded or slowed down. The system matrix were determined for the coefficients, inputs and outputs in each of these situation, along with the ways to avoid them. Computer simulation was carried out, using a script program created in MatLab mathematical package. The results of the simulation are charts showing how the program is loaded and controlled in time. The loading and control can change depending on the targeted load and the computational power of the linear dynamic system. The quantitative analysis presented in the article calculates the behavior of the plug-and-play manager in time, determined by its particular characteristics. **Practical relevance:** The presented formulas and developed program script allow us to simulate the functioning of an onboard computer network administration program under various load characteristics depending on the hardware and software implementation in a particular project, in order to avoid failures and errors during its operation.

Keywords — onboard network administration, onboard network monitoring, Plug-and-Play manager, linear dynamic system, quantitative analysis, MatLab.

For citation: Rozhdestvenskaya K. N. Quantitative analysis of an onboard computer network administration program. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 42–49 (In Russian). doi:10.31799/1684-8853-2020-4-42-49

References

1. Rozhdestvenskaya K. N. Temporal analysis of a control system in a data processing network. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 1, pp. 32–39 (In Russian). doi:10.31799/1684-8853-2019-1-32-39
2. Sheinin Iu. E., Rozhdestvenskaya K. N., Evdokimov A. S., Dymov D. V., Kochura S. G. SpaceWire-Plug-and-Play for future onboard JSC spacecraft networks. *Sovremennye problemy radioelektroniki* [Modern problems of radio electronics], 2018, pp. 196–200. Available at: http://efir.sfukras.ru/downloads/sbornik-spr-2018.pdf (accessed 03 February 2020) (In Russian).
3. Space engineering — SpaceWire — Links, nodes, routers and networks. ECSS-E-ST-50-12C Rev. 1 DIR3, November 23, 2015. ESA Requirements and Standards Division, 2015. 124 p.
4. Clancy S. C., Chase M. D., Yarlagadda A., Starch M. D., Lux J. P. SpaceWire as a Cube-Sat instrument interface. *Proc. 8th Intern. SpaceWire Conf.*, 2018, Los Angeles, USA, 2018, pp. 26–30.
5. Windsor J., Gasti W., Clerigo I. BepiColombo — building a robust data management subsystem utilizing SpaceWire networks. *7th Intern. SpaceWire Conf.*, 2016, Yokohama, Japan, 2016, pp. 112–119.
6. Tomitaka M., Igarashi Y., Ichikawa S., Inaba N., Tomiki A., Matsuzaki K., Kobayashi R., Kumakiri M., Fujishiro I., Nomachi M. Feasibility study of wireless communication system operating on SpaceWire network. *Proc. 8th Intern. SpaceWire Conf.*, 2018, Los Angeles, USA, 2018, pp. 118–122.
7. Siegle F., Leoni A. Standardization efforts for a network management and discovery protocol for SpaceFibre. *Proc. 8th Intern. SpaceWire Conf.*, 2018, Los Angeles, USA, 2018, pp. 133–137.
8. Sheynin Yu. E., Suvorova E. A., Rozhdestvenskaya K. N. Management in perspective distributed onboard computing systems based on SpaceWire standard. *Wave Electronics and its Application in Information and Telecommunication Systems*, 2019, pp. 1–5. doi:10.1109/WECONF.2019.8840122
9. Rozhdestvenskaya K. N., Evdokimov A. S. Architecture and organization of the SpaceWire network using the SpaceWire-Plug-and-Play protocol. *Nauchnaya sessiya GUAP*, 2017, pp. 198–203 (In Russian).
10. Novikov V. M., Platoshin G. A., Sheynin Yu. E. SpaceWire interface application features in avionics suites. *GosNIIAS Transactions. "Issues of Avionics" series*, 2018, no. 7(40), pp. 41–69 (In Russian).
11. Sheynin Yu. E., Olenev V. L., Lavrovskaya I. I., Dymov D. V., Kochura S. G. Protocols for prospective spacecraft on-board networks with SpaceWire and SpaceWire-Plug-and-Play technologies. *Reshetnevskie chteniya*, 2016, no. 1, pp. 651–652 (In Russian).
12. Romanowski K., Tyczka P., Holubowicz W., Renk R., Kollias V. D., Pogkas N., Jameux D. SpaceWire network management using network discovery and configuration protocol. *Proc. 7th Intern. SpaceWire Conf.*, 2016, Yokohama, Japan, 2016, pp. 45–50.
13. Burakov V. V. *Modeli ocenivaniya i algoritmy upravleniya kachestvom programmnyh sredstv.* Dis. doktor tehn. nauk [Evaluation models and algorithms for software quality management. Dr. tech. sci. diss.]. Saint-Petersburg, GUAP Publ., 2010, 42 p. (In Russian).
14. Cyucze Yu., Gunbo L. Implementation of the system state space method in MathCad. *Sbornik trudov XIII Mezhdunarodnoj nauchno-prakticheskoj konferencii studentov, aspirantov i molodyh uchenyh "Molodezh' i sovremennye informacionnye tekhnologii"* [Proc. XIII Int. Conf. "Youth and modern information technologies"]. Tomsk, 2016, pp. 302–303. Available at: http://earchive.tpu.ru/handle/11683/16943 (accessed 03 February 2020) (In Russian).
15. Zubov N. E., Mikrin E. A., Ryabchenko V. N. *Matrichnye metody v teorii i praktike sistem avtomaticheskogo upravleniya letatel'nyh apparatov* [Matrix methods in the theory and practice of automatic control systems for aircraft]. Moscow, Moskovskij gosudarstvennyj tekhnicheskij universitet im. N. E. Baumana Publ., 2016. 672 p. (In Russian).
16. Shumafov M. M. Stabilization of linear control systems. Pole assignment problem. A survey. *Vestnik of Saint Petersburg University. Mathematics. Mechanics. Astronomy*, 2019, vol. 6 (64), no. 4, pp. 564–591. Available at: https://doi.org/10.21638/11701/spbu01.2019.404 (accessed 03 February 2020) (In Russian).
17. Fokeeva L. K., Bogdanov K. U., Abdulkina F. V. Controllability and observability identification in the analysis of the system. *Materialy nauchnoj sessii uchenyh Al'met'evskogo gosudarstvennogo neftyanogo institute* [Materials of the academic session of scientists of the Almetyevsk state oil institute], 2016, no. 2, pp. 39–42 (In Russian).
18. Borisov D. V. Dynamic analysis of the state of the control system. *Sbornik statej XXVII Mezhdunarodnoj nauchno-prakticheskoj konferencii "World Science: Problems and*

Innovations" [Collection of Articles of the XXVII International Scientific and Practical Conference "World Science: Problems and Innovations"], 2018, pp. 86–90 (In Russian).
19. Bojkov I. V. Sufficient conditions for the stability of systems of ordinary differential time-dependent delay equations. *Izvestiya vuzov. Povolzhskij region. Fiziko-matematicheskie nauki. CHast' I. Linejnye uravneniya*, 2018, no. 4 (48), pp. 3–19 (In Russian).
20. Ghukasyan A. A. On the mathematical modeling of maintenance process and the condition of its controllability. *Proceedings of National Academy of Sciences of Armenia*, 2017, no. 3(70), pp. 26–38. Available at: http://mechanics.sci.am/docs/v70i3a3 (accessed 03 February 2020) (In Russian).
21. Olenev V. L., Podgornova E., Lavrovskaya I. I., Sheynin Yu. E. Deterministic services for SpaceWire networks. *Proc. 7th Intern. SpaceWire Conf.*, 2016, Yokohama, Japan, 2016, pp. 159–166.
22. Rozhdestvenskaya K. N. Analysis of Plug-and-Play manager behavior for SpaceWire networks. *Nauchnaya sessiya GUAP*, 2018, pp. 225–232 (In Russian).

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 вверху справа на стартовой странице): https://orcid.org

УДК 004.056.5
doi:10.31799/1684-8853-2020-4-50-60

Научные статьи
Articles

Метод обнаружения DoS-атак на прикладном уровне в сетях «издатель-подписчик»

Д. И. Дикий[a], *аспирант, orcid.org/0000-0002-8819-8423, dimandikiy@mail.ru*
[a]*Университет ИТМО, Кронверкский пр., 49, Санкт-Петербург, 197101, РФ*

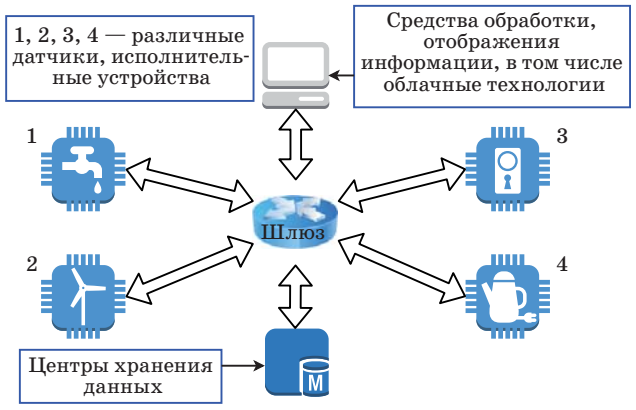
Введение: *для развития киберфизических систем разрабатываются новые технологии и протоколы передачи данных, которые призваны сократить энергетические затраты устройств на коммуникацию. Одним из современных подходов передачи данных для киберфизических систем является модель «издатель-подписчик», которая подвержена угрозе реализации атаки типа «отказ в обслуживании».* ***Цель:*** *разработка модели детектирования атаки типа «отказ в обслуживании», реализуемой на прикладном уровне сетей вида «издатель-подписчик», на основе анализа трафика методами машинного обучения.* ***Результаты:*** *разработана модель средства обнаружения атаки типа «отказ в обслуживании», учитывающая три вида сообщений: подключение, подписку, публикацию. Такой подход позволяет точнее идентифицировать источник атаки, которым может выступать узел сети, конкретное устройство или учетная запись пользователя. В качестве классификаторов были рассмотрены многослойный перцептрон, алгоритм «случайный лес» и метод опорных векторов различных конфигураций. Сгенерированы обучающие и тестовые наборы данных по предложенному вектору признаков. Оценка качества классификации производилась путем расчета F1-меры, коэффициента корреляции Метьюса и точности. Лучшие показатели по всем метрикам принадлежат модели многослойного перцептрона и методу опорных векторов с полиномиальным ядром и методом оптимизации Sequential Minimal Optimization. Однако для последнего метода характерно незначительное снижение качества классификации при ширине окна анализа трафика, близкой к максимальному периоду отправки легальных сообщений обучающего набора данных.* ***Практическая значимость:*** *результаты исследования могут быть использованы для проектирования средств обнаружения вторжений киберфизических систем, использующих модель «издатель-подписчик», а также иных систем, построенных на этом подходе.*

Ключевые слова *— DoS-атаки, «издатель-подписчик», машинное обучение, метод опорных векторов, «случайный лес», искусственная нейронная сеть.*

Для цитирования: Дикий Д. И. Метод обнаружения DoS-атак на прикладном уровне в сетях «издатель-подписчик». *Информационно-управляющие системы*, 2020, № 4, с. 50–60. doi:10.31799/1684-8853-2020-4-50-60
For citation: Dikii D. I. DoS attack detection at application level in publish-subscribe networks. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 50–60 (In Russian). doi:10.31799/1684-8853-2020-4-50-60

Введение

В настоящее время огромное внимание уделяется киберфизическим системам (КФС) [1] и их частным реализациям в виде умного города, умного дома, интернета вещей, которые позволяют автоматизировать производственные, бытовые и другие процессы. Одним из подходов, который используется для организации межмашинной коммуникации в КФС, является модель передачи данных «издатель-подписчик» [2]. Данная модель позволяет организовать передачу показателей датчиков и данных исполнительных устройств одновременно целой группе получателей. Модель «издатель-подписчик» реализована в таких протоколах, как AMQP, MQTT, XMPP и др. Существует два варианта организации сетевого взаимодействия по этой модели: распределенный и основанный на использовании шлюза. Второй вариант (рис. 1) наиболее распространен в небольших вычислительных сетях и строится по топологии «звезда». Он обладает простотой в развертывании и масштабируемости, легок в администрировании. При этом весь трафик проходит через шлюз, который отвечает за адресацию и логистику сообщений. Наряду со множеством преимуществ, которые предоставляет модель «издатель-подписчик», появляются новые угрозы информационной безопасности, свойственные именно этой модели, например, несанкционированный доступ к информации из-за



■ ***Рис. 1.*** Структура сети КФС по модели «издатель-подписчик», использующей шлюз
■ ***Fig. 1.*** CPS network structure according to the publish-subscribe model using a gateway

отсутствия разграничения доступа к ней на шлюзе, атаки типа «отказ в обслуживании», сканирование сети на наличие открытых портов и хостов, угрозы неавторизованного доступа к шлюзу.

Защита информации в КФС является актуальной задачей. Эти системы способствовали развитию новых технологий, которые должны обеспечивать надежную коммуникацию между устройствами на больших расстояниях, например, технологии LoRa, XNB и др. Кроме разработок на физическом уровне, создаются протоколы передачи данных поверх уже существующих сетей, например, протоколы прикладного уровня CoAP, MQTT. Одним из главных требований к этим протоколам является уменьшение размера служебных заголовков. Это требование основано на том, что многие устройства КФС обладают автономным ограниченным источником питания. Сокращение объемов передаваемой информации позволяет увеличить продолжительность использования этих устройств.

Одной из наиболее актуальных угроз КФС является возможность реализации атаки типа «отказ в обслуживании». Этот вид атаки может быть реализован на физическом, а также на сетевом и прикладном уровнях. Если рассматривать беспроводные сенсорные сети как элемент КФС, то для них характерны атаки типа «отказ в обслуживании» в виде зашумления радиосигнала — jamming attack [3, 4]. Помещая источник сильного шума вблизи приемников и передатчиков, можно добиться нарушения передачи данных [5]. Другой вид атак, также вызывающий отказ в обслуживании, характерен для КФС в виде одноранговых сетей — атака Сивиллы [6]. Также для КФС типичны атаки, заключающиеся в злонамеренном истощении элементов автономного питания устройства [7]. Реализация этой атаки приведет к временной неработоспособности устройства и затратам на замену элемента питания. Атаки типа «отказ в обслуживании» используются как инструмент деструктивного воздействия в ботнет-сетях [8].

В отличие от стандартных методов реализации атаки типа «отказ в обслуживании» на сетевом уровне по протоколам TCP/IP, сеть, построенная на модели «издатель-подписчик», может быть выведена из строя путем атаки на прикладном уровне [9–11]. Так как шлюз является узким местом сети, то чаще всего именно он становится целью атаки. Суть атаки сводится к генерации большого числа запросов таким образом, чтобы шлюз не справился с нагрузкой. Следовательно, разработка методов детектирования этого вида атак на прикладном уровне является актуальной задачей.

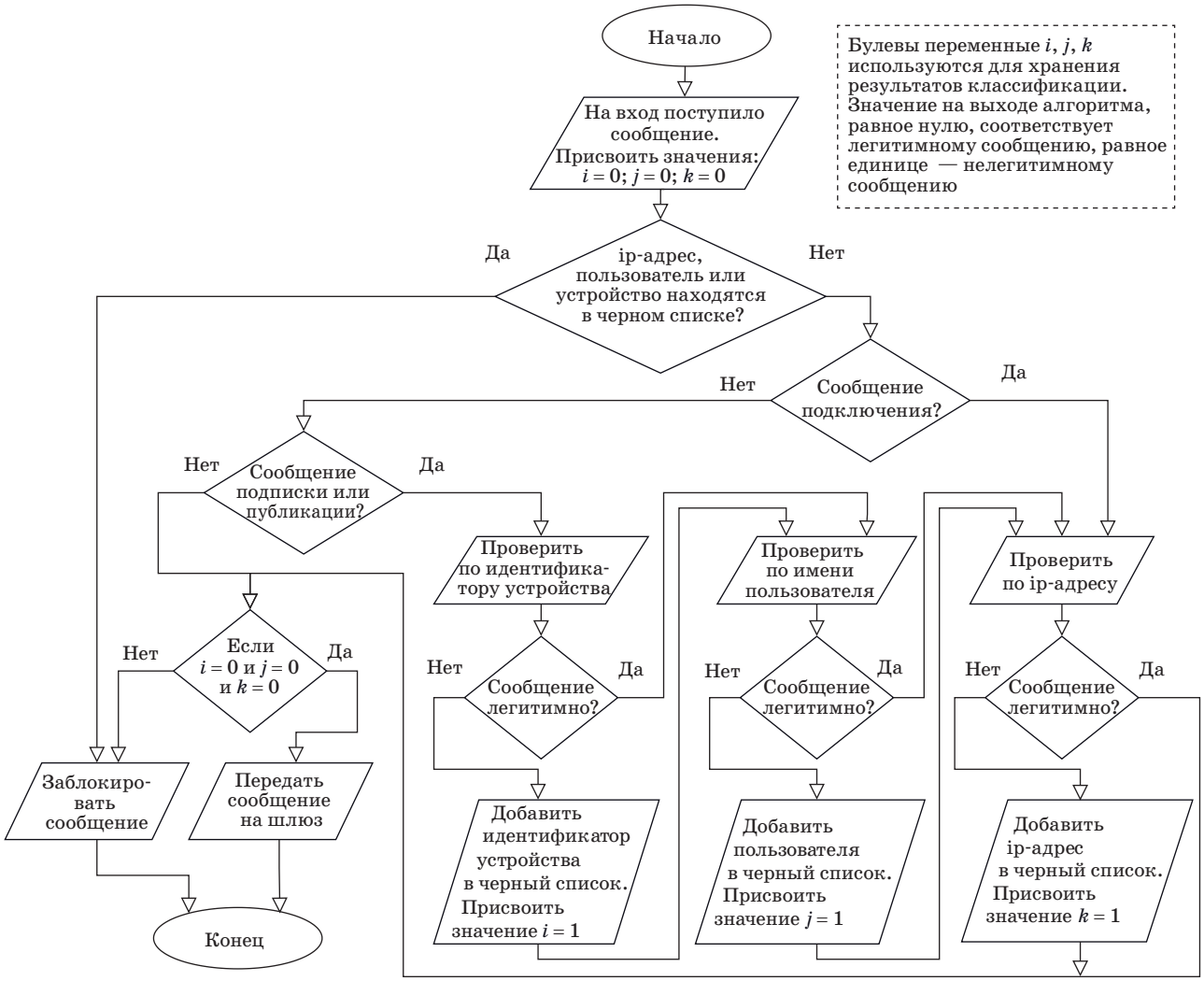
Цель исследования состоит в разработке модели средства детектирования атаки типа «отказ в обслуживании», реализуемой на прикладном уровне сетей вида «издатель-подписчик», на основе анализа трафика методами машинного обучения. В качестве объекта исследования в данной работе рассмотрен протокол MQTT, который реализует исследуемую модель межмашинной коммуникации. В отличие от большинства других протоколов, модель «издатель-подписчик» в протоколе MQTT является основной, а не расширяет уже имеющийся функционал в рамках протокола.

Возможности протокола MQTT как инструмента генерации большого числа запросов в рамках атаки типа «отказ в обслуживании» были рассмотрены во многих работах. Например, в работе [12] предложен метод, основанный на нечеткой логике, который детектирует атаки, производимые с помощью сообщений подключения к шлюзу. В этом исследовании удалось добиться точности классификации 0,909. Протокол MQTT поддерживает три уровня качества обслуживания (Quality of Service — QoS). В зависимости от выбранного уровня QoS изменяется количество служебных сообщений, участвующих в передаче информации. Следовательно, появляется возможность оптимизировать нагрузку на сеть, изменяя уровень QoS [13]. Аналогичный подход к защите от атак сводится к оптимизации загрузки шлюза на основе показаний использования вычислительных ресурсов его центрального процессора [14]. Ограничение частоты сообщений может значительно снизить риск реализации угрозы [15]. Для обнаружения сетевой аномалии недостаточно учитывать только частоту сообщений и уровень QoS. Другими важными параметрами являются криптографические преобразования, значительно влияющие на время обработки сообщения и, следовательно, на загруженность шлюза. Большое значение имеет размер полезной нагрузки. Чем больше полезная нагрузка, тем больше временных и вычислительных ресурсов потребуется на шлюзе [16].

Алгоритм детектирования атаки

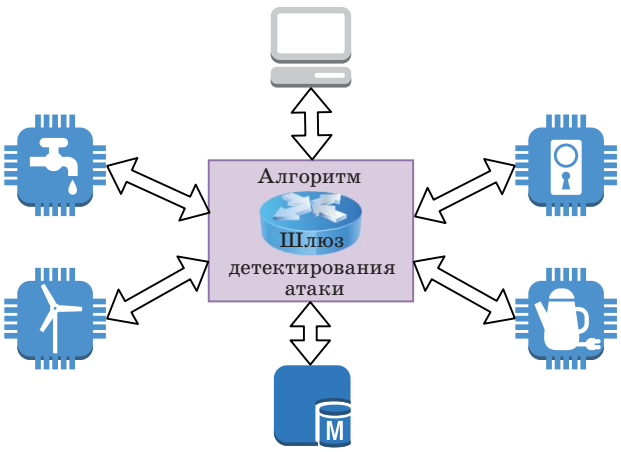
Модель «издатель-подписчик» предполагает три ключевые фазы коммуникации: подключение к шлюзу, подписку на тему, публикацию сообщения на тему. В протоколе MQTT эти три фазы реализованы посредством трех типов сообщений: connect, subscribe, publish [17]. Аналогичные механизмы предусмотрены и в других протоколах «издатель-подписчик». Эти три вида сообщений будут использованы для анализа сетевого трафика. Кроме того, как показано в работе [9], злоупотребление любым из этих трех видов сообщений способно вызвать нестабильную работу шлюза.

Отправителя сообщения можно идентифицировать несколькими способами. Сообщения под-



■ ***Рис. 2.*** Алгоритм анализа трафика в сети «издатель-подписчик»
■ ***Fig. 2.*** Algorithm of traffic analysis in the publish-subscribe network

ключения к шлюзу не требуют авторизации на шлюзе и могут быть отправлены любым устройством, которое знает адрес и порт шлюза и имеет доступ к сети. В таком случае отправителя можно идентифицировать только по его сетевому адресу. Имя пользователя и идентификатор устройства, указываемые в теле сообщения подключения, могут быть любыми, в том числе несуществующими. Сообщения подписки и публикации обрабатываются только от авторизованных на шлюзе устройств. Отправителя этих сообщений можно идентифицировать по сетевому адресу, имени пользователя и уникальному идентификатору устройства. Таким образом, в разработанном алгоритме детектирования атаки (рис. 2) применяется либо один, либо три классификатора в зависимости от вида анализируемого сообщения. Булевы переменные i, j, k используются для хранения результатов классификации по



■ ***Рис. 3.*** Схема расположения средства защиты от атаки типа «отказ в обслуживании», реализующего предлагаемый алгоритм
■ ***Fig. 3.*** Scheme of the denial-of-service attack protection that implements the proposed algorithm

идентификатору устройства, имени пользователя и ip-адресу соответственно. Значение этих переменных на выходе алгоритма, равное нулю, соответствует легитимному сообщению, равное единице — нелегитимному сообщению.

Средство обнаружения атаки типа «отказ в обслуживании», реализующее предлагаемый алгоритм, следует располагать на входе шлюза таким образом, чтобы все входящие сообщения сперва обрабатывались классификаторами и уже затем шлюзом (рис. 3).

Методы и оценка качества классификации

Классификация сетевого трафика на аномальный и легитимный часто осуществляется методами машинного обучения. Большинство работ по детектированию атаки типа «отказ в обслуживании» посвящены анализу трафика по TCP-протоколу. К самым распространенным методам относятся алгоритм k-ближайших соседей, наивный байесовский классификатор, метод опорных векторов, искусственные нейронные сети (ИНС), деревья решений и некоторые другие [18–21]. В данном исследовании в качестве классификаторов рассмотрены метод опорных векторов, ИНС, алгоритм «случайный лес». Выбор этих методов определяется тем, что они показали свою высокую эффективность при решении аналогичных задач на сетевом уровне.

Метод опорных векторов основан на построении оптимальной гиперплоскости в многомерном пространстве, разделяющей объекты различных классов. Этот подход был использован для детектирования атак в работах [22, 23] и показал отличные результаты. На форму гиперплоскости огромное влияние оказывает функция ядра. Наиболее распространенными функциями ядра являются линейная, полиномиальная, радиально-базисная, представленные следующими уравнениями соответственно:

$$K(x_i, x_j) = x_i^{\mathrm{T}} x_j; \quad (1)$$

$$K(x_i, x_j) = (x_i^{\mathrm{T}} x_j + c)^k; \quad (2)$$

$$K(x_i, x_j) = \exp(\gamma \|x_i - x_j\|^2), \quad (3)$$

где x_i и x_j — элементы классифицируемого множества; c, γ — константы; k — степень полинома.

При использовании метода опорных векторов применяются различные методы оптимизации, например метод SMO (Sequential Minimal Optimization) [24].

Искусственные нейронные сети нашли широкое применение в решении задач распознавания образов, в том числе и детектирования атаки типа «отказ в обслуживании». Например, в работе [25] точность ИНС составила порядка 0,99. Моделей ИНС существует довольно большое количество: это и обычный многослойный перцептрон Ф. Розенблатта, и рекуррентные ИНС, и ИНС с краткосрочной памятью (LSTM), и др. Основой ИНС является нейрон с соответствующей функцией активации. Наиболее часто в задачах распознавания образов используют сигмоидальную функцию активации нейрона

$$F(x) = 1/(1 + e^{(-x)}), \quad (4)$$

где x — это сумма произведений выходных сигналов нейронов предыдущего слоя на соответствующий весовой коэффициент.

Обучение ИНС производится, как правило, алгоритмом обратного распространения ошибки либо генетическим алгоритмом.

Класс алгоритмов, основанный на деревьях решений, также применяется для детектирования атаки типа «отказ в обслуживании» [26]. Идея алгоритма дерева решений сводится к построению направленного графа от корня к листьям таким образом, чтобы распределение вероятностей в листьях было равномерным. Классификатор на основе композиции из нескольких деревьев решений, генерируемых методом бутсрепа, получил большое распространение под названием «случайный лес». Оптимизация деревьев решений в рамках алгоритма «случайный лес» производится по оценкам энтропии, индекса Джинни или частоты ошибочных классификаций, представленными следующими уравнениями соответственно:

$$I = -\sum P(w_j)\log_2 P(w_j); \quad (5)$$

$$I = 1 - \sum P^2(w_j); \quad (6)$$

$$I = 1 - \max P(w_j), \quad (7)$$

где $P(w_j)$ — вероятность отнесения объекта w к классу j.

В данной работе в качестве классификаторов будут рассмотрены следующие алгоритмы: ИНС в виде многослойного перцептрона, «случайный лес», метод опорных векторов с линейной и радиально-базисной функциями ядра, метод опорных векторов с методом оптимизации SMO полиномиальной и радиально-базисной функциями ядра.

Для оценки качества классификации рассчитывались точность, F1-мера и коэффициент корреляции Мэтьюса. Значения этих параметров вычисляются исходя из количества правильно и ошибочно классифицированных объектов тестовой выборки:

— точность

$$A = (TP + TN)/(TP + TN + FP + FN); \quad (8)$$

— F1-мера

$$P = TP/(TP + FP); \quad (9)$$

$$R = TP/(TP + FN); \quad (10)$$

$$F = (2 \times P \times R)/(P + R); \quad (11)$$

— коэффициент корреляции Мэтьюса

$$M = (TP \times TN - FP \times FN)/((TP + FN) \times \times (TP + FP) \times (TN + FP) \times (TN + FN))^{1/2}. \quad (12)$$

В формулах (8)–(12) TP соответствует истинно положительным классификациям, TN — истинно отрицательным, FP — ложноположительным, а FN — ложноотрицательным.

Вектор признаков

Для классификации трафика должен быть сформирован вектор признаков с учетом модели «издатель-подписчик». Так как исследуемая модель передачи информации состоит из трех основных фаз: подключения, подписки и публикации, — то для каждой из них должен быть сформирован характерный ей вектор признаков, как представлено в табл. 1.

Основой атаки типа «отказ в обслуживании» является повышенная частота сообщений. Чтобы своевременно обнаружить атаку на начальном ее этапе, необходимо определить оптимальный интервал времени (далее — ширина окна), в течение которого рассчитываются характеристики трафика для дальнейшего анализа. Малые значения ширины окна приведут к росту ложноотрицательных результатов. С другой стороны, большие значения ширины окна отрицательно влияют на быстродействие детектирования. Вторым параметром, также влияющим на загруженность шлюза при обработке сообщений подключения, является среднее значение времени между двумя последовательными сообщениями. Для сообщений подключения важное значение имеет факт использования криптографических преобразований, в том числе процесс генерации общего сессионного ключа между шлюзом и устройством. Как правило, в КФС используется протокол TLS, который требует большого количества вычислительных и временных затрат во время установления защищенного соединения. В качестве идентификатора сообщения может выступать только сетевой адрес устройства.

Вектор признаков сообщения подписки на тему состоит из тех же параметров, что и вектор признаков сообщения подключения к шлюзу. Наличие криптографических преобразований, а это, как правило, симметричное шифрование, незначительно влияет на производительность шлюза. Источник атаки по сообщению подписки на тему можно идентифицировать не только по его сетевому адресу, но и по имени пользователя и идентификатору устройства, так как такие запросы возможно отправлять только с авторизованного устройства.

Для сообщения публикации помимо вышеперечисленных параметров необходимо учитывать размер полезной нагрузки. Чем больше размер нагрузки, тем больше вычислительных операций производится на шлюзе. К тому же на производительность шлюза как ретранслятора сообщений будет оказывать влияние количество подписчиков на тему. В большинстве протоколов модели «издатель-подписчик» для обеспечения гарантированности доставки сообщений используются уровни QoS, что также влияет на нагрузку сети. Этот механизм предусмотрен в таких протоколах, как DDS, MQTT, AMQP.

■ ***Таблица 1***. Вектор признаков трафика модели «издатель-подписчик» в зависимости от вида сообщений
■ ***Table 1***. Traffic feature vector of publish-subscribe model depending on message type

Параметр	Вид сообщения		
	Подключение	Подписка	Публикация
IP-адрес	+	+	+
Имя пользователя	–	+	+
Идентификатор устройства	–	+	+
Частота сообщений	+	+	+
Среднее значение интервала времени между сообщениями	+	+	+
Наличие криптографических преобразований трафика	+	+	+
Уровни качества QoS	–	–	+
Размер полезной нагрузки	–	–	+
Количество подписчиков	–	–	+

Наборы данных

На настоящий момент в открытом доступе отсутствуют наборы данных сетевого трафика по модели «издатель-подписчик», содержащие примеры атаки типа «отказ в обслуживании» на прикладном уровне со всей необходимой для данного исследования информацией. В связи с этим

для обучения и тестирования классификаторов были сгенерированы соответствующие наборы данных. Моделирование трафика производилось с нескольких ЭВМ, на которых были запущены MQTT-клиенты, созданные с помощью библиотеки paho-mqtt [27]. Сбор и анализ данных производился на модифицированном шлюзе Moquette с открытым исходным кодом [28]. В качестве шлюза использовался одноплатный микрокомпьютер Raspberry Pi 3 model B. Алгоритмы классификации использовались из программного пакета WEKA [29]. Основные настраиваемые параметры моделируемых наборов данных представлены в табл. 2.

Для легитимного трафика программным путем задавался максимальный период между сообщениями. Сообщения публикации дополнительно характеризовались количеством подписчиков и размером полезной нагрузки. В рамках исследования было смоделировано по два обучающих и тестовых набора для легитимного трафика. Главным отличием этих наборов данных является различный максимальный период времени между отправкой легальных сообщений, а для сообщений публикации также изменялся размер полезной нагрузки и количество подписчиков. Использование нескольких обучающих наборов данных позволит оценить влияние отличий в этих наборах на результаты классификации.

Набор данных, содержащий аномальный трафик, включал в себя примеры не только частой отправки сообщений, но и примеры со значительно большими размерами нагрузки и количеством получателей для сообщений публикации. Например, трафик сообщений большого размера, но с обычной частотой и количеством получателей.

Тестовый набор данных генерировался аналогичным образом, только с расширенными границами изменяемых параметров: периодичности, размера нагрузки, количества подписчиков. Расширение границ тестовых наборов производилось для того, чтобы оценить способность

■ ***Таблица 2***. Основные задаваемые параметры при генерации наборов данных
■ ***Table 2***. Main changed parameters of dataset generation

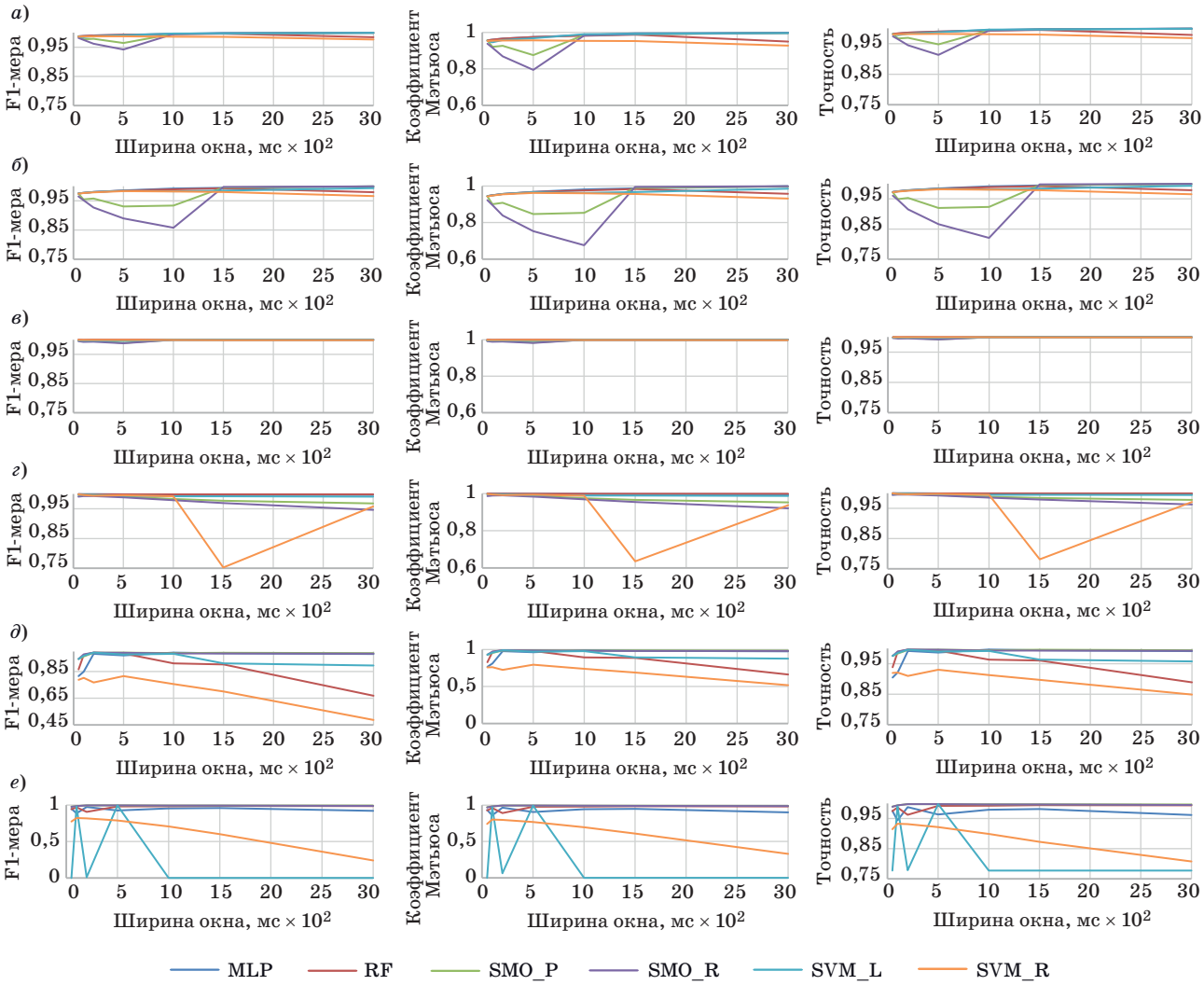
Набор данных	Легитимный трафик, вариант 1	Легитимный трафик, вариант 2	Аномальный трафик
Подключение			
Частота сообщений			
Обучение	Не реже одного сообщения в 500 мс	Не реже одного сообщения в 1000 мс	Максимальная частота
Тестирование	Не реже одного сообщения в 250 мс	Не реже одного сообщения в 500 мс	То же
Подписка			
Частота сообщений			
Обучение	Не реже одного сообщения в 500 мс	Не реже одного сообщения в 5000 мс	Максимальная частота
Тестирование	Не реже одного сообщения в 250 мс	Не реже одного сообщения в 2500 мс	То же
Публикация			
Частота сообщений			
Обучение	Не реже одного сообщения в 500 мс	Не реже одного сообщения в 5000 мс	Максимальная частота
Тестирование	Не реже одного сообщения в 250 мс	Не реже одного сообщения в 2500 мс	То же
Размер полезной нагрузки, байт			
Обучение	1–80	1–800	60 000–80 000
Тестирование	1–120	1–1200	40 000–80 000
Количество подписчиков, ед.			
Обучение	1–5	1–10	50–100
Тестирование	1–8	1–15	35–100

методов машинного обучения классифицировать трафик, явно не относящийся к одному из двух классов согласно обучающим наборам.

Результаты исследования

Для решения задачи поиска оптимального классификатора необходимо определить минимальную ширину окна, при которой результаты классификации достаточно высоки. В данной работе были рассмотрены следующие значения ширины окна: 50, 100, 200, 500, 1000, 1500, 3000 мс. Результаты классификации в виде значений F1-меры, коэффициента корреляции Мэтьюса и точности представлены на рис. 4, где MLP соответствует модели многослойного перцептрона, RF — алгоритму «случайный лес», SMO_P — методу опорных векторов с полиномиальным ядром и оптимизацией SMO, SMO_R — методу опорных векторов с радиально-базисным ядром и оптимизацией SMO, SVM_L — методу опорных векторов с линейным ядром, SVM_R — методу опорных векторов с радиально-базисным ядром.

Все три метрики (точность, F1-мера и коэффициент корреляции Мэтьюса) показывают одинаковую динамику в рамках одного и того же классификатора. Для сообщений подключения результаты всех классификаторов, кроме методов SMO_P и SMO_R, находятся в пределах небольшого диапазона для всех значений ширины окон. При этом стоит отметить, что при использовании



■ ***Рис. 4.*** Значения F1-меры, коэффициента корреляции Мэтьюса и точности классификаций при использовании легитимного трафика первого варианта для сообщений подключения (*а*), подписки (*в*), публикации (*д*); при использовании легитимного трафика второго варианта для сообщений подключения (*б*), подписки (*г*), публикации (*е*)

■ ***Fig. 4.*** The values of F1-score, Matthews correlation coefficient and accuracy of the classifications using the legitimate traffic of the first variant for message type connect (*а*), subscribe (*в*), publish (*д*); using the legitimate traffic samples of the second variant for message type connect (*б*), subscribe (*г*), publish (*е*)

алгоритмов с оптимизацией SMO худшие результаты находятся в области ширины окна, равной максимальному периоду отправки легальных сообщений в обучающем наборе. Таким образом, наибольшее влияние обучающего набора на итоговый результат оказывается при использовании SMO_P и SMO_R.

В случае сообщений подписки все классификаторы показали отличные результаты. Исключение составляет метод SVM_R, который показал один результат, резко контрастирующий на общем фоне. Как и для сообщений подключения, методы SMO_P и SMO_R показывают ухудшение качества классификации в области значений ширины окна, близкой к максимальному периоду отправки легальных сообщений, но в меньших масштабах. Высокие показатели всех классификаторов связаны с тем, что время обработки этого вида сообщений намного меньше, чем у сообщений подключения и публикации.

Наиболее интересная ситуация наблюдается для сообщений публикации. Метод SVM_L не справился с поставленной задачей при нескольких значениях ширины окна на одном наборе данных и показал не самые лучшие результаты на другом. Качество распознавания с помощью метода SVM_R уступает другим алгоритмам. Аналогичная динамика, но с чуть большими значениями метрик, наблюдается у алгоритма «случайный лес».

Таким образом, среди всех рассмотренных алгоритмов можно выделить модель многослойного перцептрона, значение F1-меры которого не опустилось ниже уровня 0,9 при ширине окна более 100 мс на всех исследуемых наборах данных. Динамика алгоритма «случайный лес» зависит от максимального периода отправки легитимных сообщений обучающего набора. При ширине окна меньшей, чем этот период, алгоритм показывает хорошие результаты, но при последующем увеличении ширины окна качество классификации начинает ухудшаться. Алгоритмы SVM_R и SVM_L показали нестабильную работу. Напротив, метод опорных векторов в виде SMO_P и SMO_R показал высокое качество распознавания — значения F1-меры не опускались ниже уровня 0,85. При этом результаты SMO_P лучше, чем у SMO_R.

Заключение

Одной из самых легко реализуемых угроз в исследуемых сетях является атака типа «отказ в обслуживании» на прикладном уровне. Предложен алгоритм детектирования атаки, позволяющий определить ее источник: узел сети, отдельное устройство или учетную запись. В качестве анализаторов трафика были рассмотрены классификаторы на основе методов машинного обучения, из которых наиболее подходящими для детектирования атаки по предлагаемому вектору признаков являются модель многослойного перцептрона при ширине окна более 100 мс и метод опорных векторов с полиномиальным ядром и методом оптимизации SMO. Однако при использовании последнего стоит учитывать локальную особенность снижения качества распознавания при ширине окна, приблизительно равной периоду отправки легальных сообщений в обучающей выборке. Результаты данного исследования будут полезны при проектировании программных и аппаратных средств защиты сетей КФС. В отличие от исследований других научных групп, чьи работы посвящены отдельному виду сообщений модели «издатель-подписчик», данная работа позволяет рассмотреть предлагаемый алгоритм детектирования атаки типа «отказ в обслуживании» как комплексное решение для анализа трафика всех трех видов сообщений. Произведено сравнение применения нескольких методов машинного обучения для решения поставленной задачи. Дальнейшие исследования будут продолжены в области анализа и проектирования систем защиты сетей, реализующих модель «издатель-подписчик» на базе конкретных протоколов с учетом их особенностей.

Финансовая поддержка

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-37-90051.

Financial support

The reported study was funded by RFBR, project number 19-37-90051.

Литература

1. **Lee J., Bagheri B., Kao H.** A Cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 2015, vol. 3, pp. 18–23. doi:10.1016/j.mfglet.2014.12.001
2. **Henneke D., Elattar M., Jasperneite J.** Communication patterns for cyber-physical systems. *2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA)*, Luxembourg, 2015, pp. 1–4. doi:10.1109/ETFA.2015.7301623
3. **Vadlamani S., Eksioglu B., Medal H., Nandi A.** Jamming attacks on wireless networks: A taxonomic survey. *International Journal of Production Economics*,

2016, vol. 172, pp. 76–94. doi:10.1016/j.ijpe.2015.11.008
4. **Li Y., Shi L., Cheng P., Chen J., Quevedo D. E.** Jamming attacks on remote state estimation in cyber-physical systems: A game-theoretic approach. *IEEE Transactions on Automatic Control*, 2015, vol. 60, no. 10, pp. 2831–2836. doi:10.1109/TAC.2015.2461851
5. **Zhang H., Qi Y., Wu J., Fu L., He L.** DoS attack energy management against remote state estimation. *IEEE Transactions on Control of Network Systems*, 2018, vol. 5, no. 1, pp. 383–394. doi:10.1109/TCNS.2016.2614099
6. **Polyzos G. C., Fotiou N.** Building a reliable Internet of things using information-centric networking. *Journal of Reliable Intelligent Environments*, 2015, vol. 1, pp. 47–58. doi:10.1007/s40860-015-0003-5
7. **Desnitsky V. A., Kotenko I. V., Rudavin N. N.** Protection mechanisms against energy depletion attacks in cyber-physical systems. *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, Saint-Petersburg and Moscow, Russia, 2019, pp. 214–219. doi:10.1109/EIConRus.2019.8656795
8. **Perrone G., Vecchio M., Pecori R., Giaffreda R.** The day after mirai: A survey on MQTT security solutions after the largest cyber-attack carried out through an army of IoT devices. *The 2nd International Conference on Internet of Things, Big Data and Security (IoTBDS)*, 2017, Porto, Portugal, pp. 246–253. doi:0.5220/0006287302460253
9. **Дикий Д. И.** Анализ протокола MQTT на атаки «отказ в обслуживании». *Научно-технический вестник информационных технологий, механики и оптики*, 2020, т. 20, № 2, с. 185–194. doi:10.17586/2226-1494-2020-20-2-185-194
10. **Chifor B., Patriciu V.** Mitigating DoS attacks in publish-subscribe IoT networks. *The 9th International Conference: Electronics, Computers and Artificial Intelligence (ECAI 2017)*, 2017, pp. 1–6. doi:10.1109/ECAI.2017.8166463
11. **Nebbione G., Calzarossa M.** Security of IoT application layer protocols: challenges and findings. *Future Internet*, 2020, vol. 12, no. 55, pp. 1–20. doi:10.3390/fi12030055
12. **Haripriya A. P., Kulothungan K.** Secure-MQTT: an efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for internet of things. *Eurasip Journal on Wireless Communications and Networking*, 2019, vol. 90. doi:10.1186/s13638-019-1402-8
13. **Potrino G., De Rango F., Fazio P.** A Distributed mitigation strategy against DoS attacks in edge computing. *2019 Wireless Telecommunications Symposium (WTS)*, New York City, NY, USA, 2019, pp. 1–7. doi:10.1109/WTS.2019.8715543
14. **Jo H., Jin H.** Adaptive periodic communication over MQTT for large-scale cyber-physical systems. *IEEE 3rd International Conference on Cyber-Physical Systems, Networks, and Applications*, 2015, Hong Kong, pp. 66–69. doi:10.1109/CPSNA.2015.21
15. **Potrino G., De Rango F., Santamaria A. F.** Modeling and evaluation of a new IoT security system for mitigating DoS attacks to the MQTT broker. *IEEE Wireless Communications and Networking Conference (WCNC)*, 2019, Marrakesh, Morocco, 2019, pp. 1–6. doi:10.1109/WCNC.2019.8885553
16. **Firdous S. N., Baig Z., Valli C., Ibrahim A.** Modelling and evaluation of malicious attacks against the IoT MQTT protocol. *IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Exeter, 2017, pp. 748–755. doi:10.1109/iThings-GreenCom-CPSCom-SmartData.2017.115
17. OASIS Standart MQTT Version 3.1.1. OASIS. 2014. 81 p. http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.pdf (дата обращения: 04.04.2020).
18. **Gharibian F., Ghorbani A. A.** Comparative study of supervised machine learning techniques for intrusion detection. *The Fifth Annual Conference on Communication Networks and Services Research (CNSR '07)*, 2007, Fredericton, NB, pp. 350–358. doi:10.1109/CNSR.2007.22
19. **Anbar M., Abdullah R., Hasbullah I. H., Chong Y., Elejla O. E.** Comparative performance analysis of classification algorithms for intrusion detection system. *The 14th Annual Conference on Privacy, Security and Trust (PST)*, 2016, Auckland, pp. 282–288. doi:10.1109/PST.2016.7906975
20. **Barati M., Abdullah A., Udzir N. I., Mahmod R., Mustapha N.** Distributed denial of service detection using hybrid machine learning technique. *The International Symposium on Biometrics and Security Technologies (ISBAST)*, 2014, Kuala Lumpur, pp. 268–273. doi:10.1109/ISBAST.2014.7013133
21. **Xiao L., Wan X., Lu X., Zhang Y., Wu D.** IoT security techniques based on machine learning: how do IoT devices use AI to enhance security? *IEEE Signal Processing Magazine*, 2018, vol. 35, no. 5, pp. 41–49. doi:10.1109/MSP.2018.2825478
22. **Jianjian D., Yang T., Feiyue Y.** A novel intrusion detection system based on IABRBFSVM for wireless sensor networks. *Procedia Computer Science*, 2018, vol. 131, pp. 1113–1121. doi:10.1016/j.procs.2018.04.275
23. **Abusitta A., Bellaiche M., Dagenais M.** An SVM-based framework for detecting DoS attacks in virtualized clouds under changing environment. *Journal of Cloud Computing*, 2018, vol. 7, no. 9. doi:10.1186/s13677-018-0109-4
24. **Platt J. C.** Sequential minimal optimization: A fast algorithm for training support vector machines. *Microsoft Research, Technical Report MSR-TR-98-14*, 1998, pp. 1–21.
25. **Hodo E., Bellekens X., Hamilton A., Dubouilh P.-L., Iorkyase E., Tachtatzis C., Atkinson R.** Threat analy-

sis of IoT networks using artificial neural network intrusion detection system. *International Symposium on Networks, Computers and Communications (ISNCC)*, Yasmine Hammamet, 2016, pp. 1–6. doi:10.1109/ISNCC.2016.7746067
26. **Sangkatsanee P., Wattanapongsakorn N., Charnsripinyo C.** Practical real-time intrusion detection using machine learning approaches. *Computer Communications*, 2011, vol. 34, pp. 2227–2235. doi:10.1016/j.comcom.2011.07.001
27. Paho-MQTT library for Python. https://pypi.org/project/paho-mqtt/ (дата обращения: 04.04.2020).
28. Moquette project open source code. https://github.com/moquette-io/moquette (дата обращения: 04.04.2020).
29. WEKA project. https://www.cs.waikato.ac.nz/ml/weka/ (дата обращения: 04.04.2020).

UDC 004.056.5
doi:10.31799/1684-8853-2020-4-50-60

DoS attack detection at application level in publish-subscribe networks

D. I. Dikii[a], Post-Graduate Student, orcid.org/0000-0002-8819-8423, dimandikiy@mail.ru
[a]ITMO University, 49, Kronverkskii Pr., 197101, Saint-Petersburg, Russian Federation

Introduction: For the development of cyberphysical systems, new technologies and data transfer protocols are being developed, in order to reduce the energy costs of communication devices. One of the modern approaches to data transmission in cyberphysical systems is the publish-subscribe model, which is subject to a denial-of-service attack. **Purpose:** Development of a model for detecting a DoS attack implemented at the application level of publish-subscribe networks based on the analysis of their traffic using machine learning methods. **Results:** A model is developed for detecting a DoS attack, operating with three classifiers depending on the message type: connection, subscription, and publication. This approach makes it possible to identify the source of an attack. That can be a network node, a particular device, or a user account. A multi-layer perceptron, the random forest algorithm, and a support vector machine of various configurations were considered as classifiers. Training and test data sets were generated for the proposed feature vector. The classification quality was evaluated by calculating the F1 score, the Matthews correlation coefficient, and accuracy. The multilayer perceptron model and the support vector machine with a polynomial kernel and SMO optimization method showed the best values of all metrics. However, in the case of the support vector machine, a slight decrease in the prediction quality was detected when the width of the traffic analysis window was close to the longest period of sending legitimate messages from the training data set. **Practical relevance:** The results of the research can be used in the development of intrusion detection features for cyberphysical systems using the publish-subscribe model, or other systems based on the same approach.
Keywords — DoS attack, publish-subscribe, machine learning, SVM, random forest, ANN.

For citation: Dikii D. I. DoS attack detection at application level in publish-subscribe networks. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 50–60 (In Russian). doi:10.31799/1684-8853-2020-4-50–60

References

1. Lee J., Bagheri B., Kao H. A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 2015, vol. 3, pp. 18–23. doi:10.1016/j.mfglet.2014.12.001
2. Henneke D., Elattar M., Jasperneite J. Communication patterns for cyber-physical systems. *2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA)*, Luxembourg, 2015, pp. 1–4. doi:10.1109/ETFA.2015.7301623
3. Vadlamani S., Eksioglu B., Medal H., Nandi A. Jamming attacks on wireless networks: A taxonomic survey. *International Journal of Production Economics*, 2016, vol. 172, pp. 76–94. doi:10.1016/j.ijpe.2015.11.008
4. Li Y., Shi L., Cheng P., Chen J., Quevedo D. E. Jamming attacks on remote state estimation in cyber-physical systems: A game-theoretic approach. *IEEE Transactions on Automatic Control*, 2015, vol. 60, no. 10, pp. 2831–2836. doi:10.1109/TAC.2015.2461851
5. Zhang H., Qi Y., Wu J., Fu L., He L. DoS attack energy management against remote state estimation. *IEEE Transactions on Control of Network Systems*, 2018, vol. 5, no. 1, pp. 383–394. doi:10.1109/TCNS.2016.2614099
6. Polyzos G. C., Fotiou N. Building a reliable Internet of things using information-centric networking. *Journal of Reliable Intelligent Environments*, 2015, vol. 1, pp. 47–58. doi:10.1007/s40860-015-0003-5
7. Desnitsky V. A., Kotenko I. V., Rudavin N. N. Protection mechanisms against energy depletion attacks in cyber-physical systems. *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, Saint-Petersburg and Moscow, Russia, 2019, pp. 214–219. doi:10.1109/EIConRus.2019.8656795
8. Perrone G., Vecchio M., Pecori R., Giaffreda R. The day after mirai: A survey on MQTT security solutions after the largest cyber-attack carried out through an army of IoT devices. *The 2nd International Conference on Internet of Things, Big Data and Security (IoTBDS)*, 2017, Porto, Portugal, pp. 246–253. doi:10.5220/0006287302460253
9. Dikii D. I. Denial-of-service attack analysis by MQTT protocol. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 2, pp. 185–194 (In Russian). doi:10.17586/2226-1494-2020-20-2-185-194
10. Chifor B., Patriciu V. Mitigating DoS attacks in publish-subscribe IoT networks. *The 9th International Conference: Electronics, Computers and Artificial Intelligence (ECAI 2017)*, 2017, pp. 1–6. doi:10.1109/ECAI.2017.8166463
11. Nebbione G., Calzarossa M. Security of IoT application layer protocols: Challenges and findings. *Future Internet*, 2020, vol. 12, no. 55, pp. 1–20. doi:10.3390/fi12030055
12. Haripriya A. P., Kulothungan K. Secure-MQTT: an efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for internet of things. *Eurasip Journal on Wireless Communications and Networking*, 2019, vol. 90. doi:10.1186/s13638-019-1402-8
13. Potrino G., De Rango F., Fazio P. A distributed mitigation strategy against DoS attacks in edge computing. *2019 Wireless Telecommunications Symposium (WTS)*, New York City, NY, USA, 2019, pp. 1–7. doi:10.1109/WTS.2019. 8715543
14. Jo H., Jin H. Adaptive periodic communication over MQTT for large-scale cyber-physical systems. *IEEE 3rd International Conference on Cyber-Physical Systems, Networks, and Applications*, 2015, Hong Kong, pp. 66–69. doi:10.1109/CPSNA.2015.21

15. Potrino G., De Rango F., Santamaria A. F. Modeling and evaluation of a new IoT security system for mitigating DoS attacks to the MQTT broker. *IEEE Wireless Communications and Networking Conference (WCNC)*, 2019, Marrakesh, Morocco, 2019, pp. 1–6. doi:10.1109/WCNC.2019.8885553
16. Firdous S. N., Baig Z., Valli C., Ibrahim A. Modelling and evaluation of malicious attacks against the IoT MQTT protocol. *IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Exeter, 2017, pp. 748–755. doi:10.1109/iThings-GreenCom-CPSCom-SmartData.2017.115
17. *OASIS Standart MQTT Version 3.1.1. OASIS*. 2014. 81 p. Available at: http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.pdf (accessed 04 April 2020).
18. Gharibian F., Ghorbani A. A. Comparative study of supervised machine learning techniques for intrusion detection. *The Fifth Annual Conference on Communication Networks and Services Research (CNSR '07)*, 2007, Fredericton, NB, pp. 350–358. doi:10.1109/CNSR.2007.22
19. Anbar M., Abdullah R., Hasbullah I. H., Chong Y., Elejla O. E. Comparative performance analysis of classification algorithms for intrusion detection system. *The 14th Annual Conference on Privacy, Security and Trust (PST)*, 2016, Auckland, pp. 282–288. doi:10.1109/PST.2016.7906975
20. Barati M., Abdullah A., Udzir N. I., Mahmod R., Mustapha N. Distributed denial of service detection using hybrid machine learning technique. *The International Symposium on Biometrics and Security Technologies (ISBAST)*, 2014, Kuala Lumpur, pp. 268–273. doi:10.1109/ISBAST.2014.7013133
21. Xiao L., Wan X., Lu X., Zhang Y., Wu D. IoT security techniques based on machine learning: How do IoT devices use AI to enhance security? *IEEE Signal Processing Magazine*, 2018, vol. 35, no. 5, pp. 41–49. doi:10.1109/MSP.2018. 2825478
22. Jianjian D., Yang T., Feiyue Y. A novel intrusion detection system based on IABRBFSVM for wireless sensor networks. *Procedia Computer Science*, 2018, vol. 131, pp. 1113–1121. doi:10.1016/j.procs.2018.04.275
23. Abusitta A., Bellaiche M., Dagenais M. An SVM-based framework for detecting DoS attacks in virtualized clouds under changing environment. *Journal of Cloud Computing*, 2018, vol. 7, no. 9. doi:10.1186/s13677-018-0109-4
24. Platt J. C. Sequential minimal optimization: A fast algorithm for training support vector machines. *Microsoft Research, Technical Report MSR-TR-98-14*, 1998, pp. 1–21.
25. Hodo E., Bellekens X., Hamilton A., Dubouilh P.-L., Iorkyase E., Tachtatzis C., Atkinson R. Threat analysis of IoT networks using artificial neural network intrusion detection system. *International Symposium on Networks, Computers and Communications (ISNCC)*, Yasmine Hammamet, 2016, pp. 1–6. doi:10.1109/ISNCC.2016.7746067
26. Sangkatsanee P., Wattanapongsakorn N., Charnsripinyo C. Practical real-time intrusion detection using machine learning approaches. *Computer Communications*, 2011, vol. 34, pp. 2227–2235. doi:10.1016/j.comcom.2011.07.001
27. *Paho-MQTT library for Python*. Available at: https://pypi.org/project/paho-mqtt/(accessed 04 April 2020).
28. *Moquette project open source code*. Available at: https://github.com/moquette-io/moquette (accessed 04 April 2020).
29. *WEKA project*. Available at: https://www.cs.waikato.ac.nz/ml/weka/(accessed 04 April 2020).

УДК 004.93'1
doi:10.31799/1684-8853-2020-4-61-77

Научные статьи
Articles

Высоконадежная аутентификация по рукописным паролям на основе гибридных нейронных сетей с обеспечением защиты биометрических эталонов от компрометации

А. Е. Сулавко[a], *канд. техн. наук, доцент, orcid.org/0000-0002-9029-8028, sulavich@mail.ru*
[a]*Омский государственный технический университет, Мира пр., 11, Омск, 644050, РФ*

Введение: *нейросетевые преобразователи «биометрия-код» являются идеологической основой для серии стандартов ГОСТ Р 52633 (не имеющих к настоящему моменту мировых аналогов), которые могут быть востребованы при разработке средств высоконадежной биометрической аутентификации и электронной подписи с биометрической активацией.* ***Цель:*** *разработать модель преобразователя «биометрия-код» для высоконадежной биометрической аутентификации по рукописным паролям с высокой устойчивостью к атакам на извлечение знаний.* ***Результаты:*** *продемонстрирована уязвимость нейросетевых преобразователей «биометрия-код», позволяющая совершать быстрый направленный перебор конкурирующих примеров для компрометации биометрического образа и личного ключа его владельца. Описан метод эффективной защиты от данной атаки. Предложена гибридная модель нейросетевого преобразователя «биометрия-код» (основанная на новом типе гибридных нейронных сетей), не компрометирующего биометрический эталон и ключ (пароль) пользователя и устойчивого к подобным атакам. Экспериментально подтверждена высокая надежность и эффективность предложенной модели в задачах верификации рукописных паролей. Показатели надежности генерации ключа из рукописного пароля составили: FRR = 11,5 %, FAR = 0,0009 % при длине ключа 1024 бит (с учетом предъявления подделок рукописного образа).* ***Практическая значимость:*** *результаты будут востребованы в приложениях информационной безопасности и при реализации электронного документооборота.*

Ключевые слова *— распознавание образов, разностные функционалы Байеса, обработка коррелированных биометрических параметров, защита информации, автоматическая настройка нейронных сетей, плотности вероятности, «широкие» нейронные сети, преобразователи «биометрия-код», рукописный почерк.*

Для цитирования: Сулавко А. Е. Высоконадежная аутентификация по рукописным паролям на основе гибридных нейронных сетей с обеспечением защиты биометрических эталонов от компрометации. *Информационно-управляющие системы*, 2020, № 4, с. 61–77. doi:10.31799/1684-8853-2020-4-61-77

For citation: Sulavko A. E. Highly reliable authentication based on handwritten passwords using hybrid neural networks with protection of biometric templates from being compromised. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 61–77 (In Russian). doi:10.31799/1684-8853-2020-4-61-77

Введение

Наступила эпоха Big Data. Растет количество интеллектуальных технологий и онлайн-сервисов, которые несложно освоить массовому потребителю. При этом обеспечить безопасность виртуального образа пользователя становится все сложней — вместе с количеством личных кабинетов возрастает число паролей, которые нужно хранить. Сегодня пользователь нуждается не только в надежной аутентификации, но и в защите аутентификационных данных от компрометации.

Пароли и криптографические ключи являются отчуждаемыми от владельца и поэтому подвержены «человеческому фактору». Длинный ключ (пароль) является надежным, только если были соблюдены все правила при его генерации — информационная энтропия ключа (пароля) должна быть сопоставима с его длиной. Но случайный длинный пароль почти невозможно запомнить. Выходом из ситуации является «привязка» всех ключей и паролей субъекта к его биометрическим параметрам с помощью преобразователя «биометрия-код» (ПБК) [1]. ПБК можно сравнить с интеллектуальным «черным ящиком», который «знает» своего владельца и безопасно хранит его пароль или криптографический ключ. ПБК обучается формировать и отдавать пользователю его пароль (ключ) при предъявлении биометрического образа. При предъявлении образа любого другого субъекта ПБК должен формировать случайный бинарный код, близкий по информационной энтропии к «белому шуму». Предполагается, что сами пароли и ключи генерируются перед обучением ПБК в соответствии с принятыми нормами. Данные обученного ПБК (ключ и биометрический эталон) должны быть защищены от компрометации при хранении и передаче по каналам связи без применения сторонних средств шифрования [2]. Хакеры не должны иметь возможность извлечь знания из обученного ПБК. Концепция ПБК может использоваться как основа для средств высоконадежной биометрической

аутентификации, а также электронной подписи с биометрической активацией.

Большинство методов биометрической аутентификации базируется на статических биометрических образах (отпечатках пальцев, радужке и т. п.). Статические образы наиболее уязвимы перед атаками представления, так как их невозможно держать в секрете. Открытый образ может быть изучен злоумышленником и фальсифицирован. Поэтому для аутентификации желательно использовать тайный образ, характеризующий особенности воспроизведения пароля его владельцем.

Настоящее исследование посвящено разработке модели ПБК для высоконадежной аутентификации на основе рукописных паролей.

Описание проблемы

Биометрический образ (пример образа) — это биометрические данные человека, подвергающиеся в дальнейшем масштабированию, удалению шумов и другой обработке в целях вычисления вектора биометрических параметров (признаков). При обучении ПБК создается биометрический эталон пользователя, который связывается с криптографическим ключом или паролем. Высвобождение ключа (пароля) происходит на этапе аутентификации, шифрования/дешифрования контента или создания электронной подписи. Предполагается, что обучение должно выполняться в доверенной среде, но обученный ПБК размещается в потенциально враждебной среде. При этом во многих приложениях, где требуется обеспечить анонимность пользователей (например, в медицине), ПБК каждого пользователя должен быть обезличен (не связан с его персональными данными).

Далее под *ключом* будет подразумеваться как непосредственно ключ шифрования или электронной подписи, так и пароль пользователя. Принципиального отличия в реализации ПБК при связывании биометрии с паролем или ключом нет. В зависимости от применения ПБК на практике могут предъявляться требования к длине и информационной энтропии ключа, а также соответствующим свойствам ПБК.

Наконец, следует дополнительно пояснить, что понимается под высокой надежностью работы ПБК. Надежность определяется показателями *FRR* и *FAR* — вероятностями ошибок «ложного отказа» (ответ ПБК не совпадает с ключом субъекта при предъявлении образа легитимным пользователем («Своим»)) и «ложного допуска» (ответ ПБК совпадает с ключом пользователя при предъявлении образа нелегитимным субъектом («Чужим»)).

Чтобы балансировать FRR и FAR, требуется изменять *порог принятия*. На практике для установки ненулевого порога требуется дополнительно корректировать ответ ПБК, например, с помощью кодов, исправляющих ошибки, и хранить синдромы ошибок, что снижает защищенность биометрического эталона и ключа от компрометации. *Нулевой порог принятия* означает, что система принимает ответ как правильный, только если он строго равен ключу пользователя (расстояние между ответом ПБК и ключом равно нулю).

В настоящей работе рассматривается атака *«извлечения знаний» из ПБК*, направленная на фальсификацию биометрического образа и получение злоумышленником личного ключа субъекта быстрее, чем при реализации атаки полного перебора ключей. Как и в случае атак на парольные системы, злоумышленник пытается сократить количество вариантов для перебора биометрических образов.

Сценарии атак на биометрические системы классифицируют по уровню знаний нарушителя об атакуемой стороне [3]. Будем исходить из того, что алгоритмы извлечения признаков из рукописного образа и функционирования ПБК не являются секретными. Кроме того, злоумышленник обладает параметрами обученного ПБК и выборкой образов «Чужие» неограниченного объема. Он может построить обученный ПБК, чтобы реализовать перебор образов «Чужие» с целью получить на выходах ПБК личный ключ пользователя, правильность которого он может проверить (рис. 1), не имея представления о том, какие именно выходные биты ПБК не совпали с соответствующими битами ключа пользователя (какие разряды оказались ошибочными). Битовую последовательность на выходе из ПБК будем называть *ответом*.

■ ***Рис. 1.*** Попытка подбора рукописного пароля для дешифрования контента

■ ***Fig. 1.*** Trying to matching handwriting password for decrypting the content

Тем не менее злоумышленник не обладает цифровой копией образа пользователя-жертвы (т. е. полными данными о написании пароля, включая динамику изменения скорости и давления пера), считается, что в этом случае ПБК будет скомпрометирован. Однако разглашение текстового содержания и даже компрометация внешнего вида рукописного пароля не должна приводить к компрометации ПБК. В этом случае пользователю требуется время, чтобы повторно обучить ПБК с использованием другого тайного биометрического образа. Именно такой сценарий атаки рассматривается в работе: злоумышленнику известны все алгоритмы, у него имеется база данных с параметрами обученных ПБК, имеется изображение рукописного образа пользователя-жертвы.

Стандарты ПБК и достигнутые ранее результаты

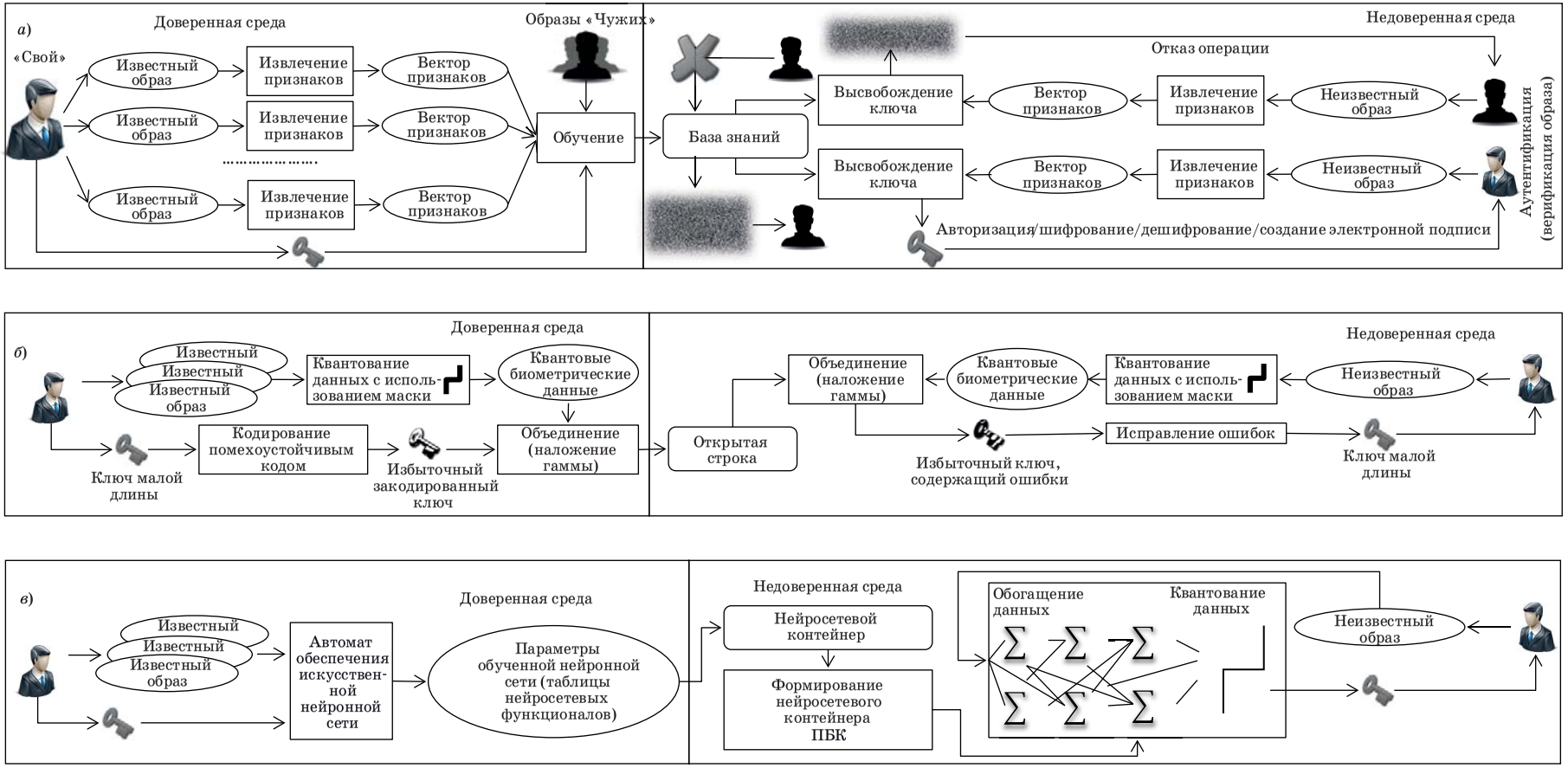
Общая схема работы ПБК представлена на рис. 2, *а*. На текущий момент сложилось два основных подхода к построению ПБК [4]: на основе нечеткого экстрактора (рис. 2, *б*), в рамках которого для дополнительной защиты применяется криптография (стандарты ISO/IEC 19792:2009 [5], 24761:2009 [6] и 24745:2011 [7]), и нейросетевой подход, поддерживаемый серией ГОСТ Р 52633 [8–14] (рис. 2, *в*), не имеющих к настоящему времени мировых аналогов.

В основе классического нечеткого экстрактора [15] применяются методы помехоустойчивого кодирования для исправления ошибок, возникающих из-за невозможности точного повторного воспроизведения биометрического образа. Этот подход имеет принципиальные недостатки [4]. В работе [16] приводятся уязвимости нечетких экстракторов, позволяющие выполнять направленный перебор биометрических данных для получения ключа. Об утечке конфиденциальности в некоторых схемах нечетких экстракторов идет речь и в других работах [17]. Также нечеткие экстракторы не способны к полноценному обучению. Они квантуют «сырые» биометрические данные, при этом подавляются шумы оборудования, но не учитывается характер распределения значений признаков пользователей. По этой причине их можно применять, только если образы высокоинформативные (как отпечаток пальца или радужка). Для рукописных образов число ошибок оказывается значительным [18].

Искусственные нейронные сети (ИНС) кодируют данные об особенностях признаков пользователей весовыми коэффициентами, что не дает прямого наблюдения за биометрическими эталонами [19]. Нейросетевой ПБК строится персонально для каждого субъекта, при этом формируется ИНС, количество входов которой равно числу признаков, а количество выходов — длине его личного ключа (см. рис. 2, *в*). Каждый нейрон последнего слоя генерирует один бит. Нейронная сеть обучается на биометрических образах пользователя и образах «Чужих», чтобы вырабатывать ключ субъекта при поступлении на вход его биометрического образа. Хорошо обученная нейронная сеть не нуждается в дополнительной корректировке выходов (ответа). Обучение нейросетевого ПБК должно быть абсолютно устойчивым, при этом объем обучающей выборки «Чужие» может быть сколь угодно большим. Разработчик биометрической системы может заготовить репрезентативную выборку «Чужие» заранее и использовать ее для обучения каждого ПБК. Однако число примеров образа «Свой» должно быть малым (по ГОСТ Р 52633.5-2011 достаточно 11 примеров [13]), нельзя заставлять пользователя сотни раз вводить биометрический образ. Это обстоятельство накладывает существенные ограничения на архитектуру ИНС, используемую в основе ПБК.

На сегодня построить ПБК на основе многослойных нейронных сетей затруднительно [4]. Исследования показывают, что на практике для обучения «глубокой» сети требуется более сотни примеров рукописного образа «Свой», чтобы надежность решений была приемлемой [1]. Практически все итерационные алгоритмы обучения теряют устойчивость при изменении параметров ИНС или объема обучающей выборки. Чем ниже качество биометрического образа, тем больший объем выборки требуется. Поэтому построение ПБК с большим количеством бинарных выходов на базе многослойных ИНС видится затруднительным (уже при длине ключа/пароля 32 бита количество классов нейронной сети должно составлять $2^{32} = 4\ 294\ 967\ 296$, что больше половины населения планеты). Кроме того, сверточные сети имеют ряд уязвимостей, обусловленных их изначальной ориентированностью на обработку графических образов [3] (наложение различного вида шума на изображение подписи существенно увеличивает FAR).

Активные исследования ведутся в области неглубоких сетей (shallow networks), способных к универсальной аппроксимации (для этого требуется потенциально неограниченное число скрытых нейронов, которое играет роль сложности модели ИНС). На данный момент проведена оценка ограничений малых ИНС, сформулирован ряд теорем [20], получены нижние оценки сложности ИНС в зависимости от соотношения между областью значений аппроксимируемой функции и размерностью входа [21].



■ ***Рис. 2.*** Иллюстрация принципов работы ПБК (слева — обучение ПБК, справа — высвобождение ключа): *а* — общая схема; *б* — нечеткий экстрактор; *в* — нейросетевой ПБК

■ ***Fig. 2.*** Illustration of the principles of operation of converters “biometrics to code” (on the left — training, on the right — the release of the key): *а* — general scheme; *б* — fuzzy extractor; *в* — neural network converter

■ ***Таблица 1***. Достигнутый уровень надежности при аутентификации по рукописным образам

■ ***Table 1***. Achieved reliability level for handwritten authentication

Метод/подход, особенности	FRR, %	FAR, %	С учетом подделок
Многослойный персептрон (MLP) + метод главных компонент PCA [23]	6,4	7,4	+
Авторский метод реализации ПБК [24]. Ключ 256 бит. Защита повышает число ошибок	38,75	13,45	–
CNN + метод опорных векторов [25]. Обучающая выборка «Свой», 30 примеров	2,17	13	+
CNN [26]. Обучающая выборка «Свой», 30 примеров	1,48 2,63	1,48 2,63	+ +
CNN [27]	2,42	2,42	+
Рекуррентные ИНС [27]	2,37	2,37	+
Нечеткий экстрактор [28]	9	9	–
«Широкие» ИНС [29]	10	10^{-7}	–

Большие нейронные сети из малого числа слоев (одного или двух) легли в основу стандартов ГОСТ Р 52633. Эти сети принято называть «широкими» [4]. Важным отличием «широких» сетей является процедура автоматического и абсолютно устойчивого послойного обучения (без использования алгоритма градиентного спуска). Идеологом и основателем научного направления, связанного с «широкими» ИНС и нейросетевыми ПБК на их основе, является А. И. Иванов [22]. За последние 20 лет под его авторством вышло множество работ, посвященных данному направлению. Сегодня развитием нейросетевых ПБК занимаются преимущественно ученые из России и Казахстана [1, 19, 22].

Приведем сопоставительные данные о надежности биометрической аутентификации на основе рукописных образов (подписей и паролей) при помощи методов, которые позволяют защитить эталоны и личные ключи субъектов от компрометации (табл. 1).

Преимущество «широкой» сети — в устойчивом обучении, увеличение числа нейронов влияет на время обучения линейно, в отличие от «глубокой» сети.

База рукописных образов

Для проведения исследований собрана база естественных рукописных образов. В соответствии с ISO/IEC 19795-3 [30] при формировании базы были учтены следующие факторы:

— «старение эталона»: база включает рукописные образы испытуемых, полученные в разные дни (с интервалом до нескольких недель);

— пол и возраст распределены равномерно на интервале от 18 до 35 лет;

— усилия злоумышленника: образ пароля каждого испытуемого пытались повторить еще пять субъектов по 10 раз, предварительно изучив его внешний вид на экране монитора и имея представление о темпе почерка его владельца. Согласно ISO/IEC 19795-3 [30], такой вид подделки рукописного образа соответствует 4-й степени фальсификации.

Для ввода рукописных образов использовались планшеты фирмы Wacom с частотой опроса 200 точек в секунду и 1024 уровнями давления пера на планшет. Ввод осуществлялся при помощи программного модуля, разработанного на языке C++ для семейства ОС Windows, все примеры сохранены как тестовые файлы. На момент эксперимента база насчитывала более 27 000 примеров рукописных паролей 260 испытуемых, включая примеры их подделок.

Предлагаемая процедура извлечения признаков

В компьютерном представлении рукописный образ состоит из функций положения пера $x(t)$, $y(t)$ и давления пера на планшет $\boldsymbol{p(t)}$ (аналогом является сила нажатия, которую способны регистрировать некоторые модели), где t — это время в дискретной форме. Образ преобразуется в вектор признаков $\bar{\mathbf{a}} = \{a_1, \ldots, a_N\}$ фиксированной длины ($N = 782$). «Широкие» ИНС позволяют использовать как можно больше признаков и не рассматривать разные методы их извлечения как альтернативные. Поэтому различные подходы к извлечению признаков могут и должны применяться совместно. В настоящей работе комбинировались следующие способы извлечения признаков [4, 18]:

— образ делится на 16 равных отрезков, строится матрица расстояний между их краями в двух- и трехмерном пространстве ($p(t)$ — третье измерение);

вычисляются:

— коэффициенты корреляции между $x(t)$, $y(t)$, $p(t)$, $x'(t)$, $y'(t)$, $p'(t)$ и функцией скорости пера $v_{xy}(t)$, производной от $x(t)$, $y(t)$;

— параметры внешнего вида образа: угол наклона, отношение длины к ширине, центр в трехмерном пространстве с координатами x, y, p;

— средние значения фрагментов функций $p(t)$, $x'(t)$, $y'(t)$, $v_{xy}(t)$ (образ делится на пять равных по числу точек отрезков);

— детализирующие коэффициенты быстрого вейвлет-преобразования Хаара (алгоритм Малла), полученные на четырех уровнях разложения (низкие частоты) для $x(t)$, $y(t)$, $p(t)$, $v_{xy}(t)$;

— усредненный амплитудный спектр, полученный с помощью Short-Time Fourier Transform (размер окна — 128 отчетов, шаг — 16 отчетов) для $x(t)$, $y(t)$, $p(t)$, $v_{xy}(t)$.

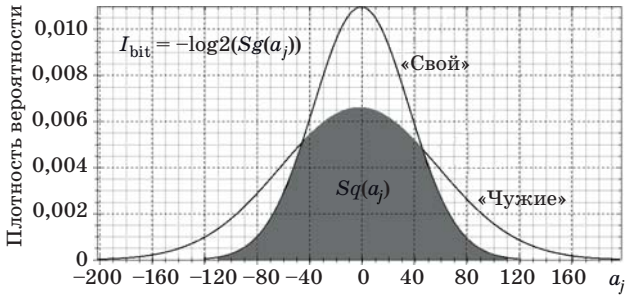
Сочетание нескольких методов получения признаков не снижает вероятность случайного совпадения и даже затрудняет попытки намеренной подделки.

В работе [31] предложена универсальная шкала для оценки информационной емкости биометрического образа. Информативность j-го признака оценивается через построение функций плотности вероятности для классов образов «Свой» и «Чужой» (рис. 3). Закон распределения большинства представленных признаков близок к нормальному [18] (реже встречается логнормальное и двойное экспоненциальное распределения).

После построения функций плотности вероятности вычисляется площадь их пересечения $Sq(a_j)$, которая равна вероятности ошибки при классификации образа по признаку a_j (интеграл от функции плотности вероятности равен вероятности). Вероятность переводится в собственную информацию по формуле

$$I_{\text{bit}} \approx -\log_2 Sq(a_j).$$

Для каждого подписанта следует выбирать наиболее информативные признаки индивидуально — при обучении нейросетевого ПБК. Информативность отдельно взятого признака I_{bit} — это показатель его уникальности. Но на общее количество информации (о субъекте) в биометрическом образе влияют также корреляционные связи между признаками. Часть информации всегда «переходит» в матрицу парных коэффициентов корреляции признаков, которая почти уникальна у каждого рукописного образа. Эту дополнительную информацию (как будет показано далее) можно использовать при распознавании.



■ ***Рис. 3.*** Информативность j-го признака для одного из испытуемых

■ ***Fig. 3.*** Information capacity of the j-th feature for some subject

Нейросетевые ПБК на базе классических нейронов

Первый слой классической «широкой» нейронной сети, обучаемой по алгоритму ГОСТ 52633.5, обогащает входные данные, второй — играет роль кодов, исправляющих ошибки [4]. Однако по сравнению с нечеткими экстракторами нейросетевая коррекция ошибочных разрядов ключа обладает гораздо меньшей избыточностью [19].

Рассмотрим однослойные «широкие» сети. Классический нейрон базируется на функционале

$$y = \sum_{j=1}^{n} \mu_j a_j \qquad (1)$$

и пороговой функции активации

$$f(y) = \begin{cases} 0, \text{ если } y < \mu_0 \\ 1, \text{ если } y > \mu_0 \end{cases}; \qquad (2)$$

модули весов нейронов первого слоя вычисляются по формуле [4]

$$\mu_j = |m_s(a_j) - m_o(a_j)| \,/\, \sigma_s(a_j) \cdot \sigma_o(a_j), \qquad (3)$$

где y — отклик нейрона на образ «Свой» или «Чужой»; n — количество входов нейрона; a_j — значение j-го признака (входа нейрона); $f(y)$ — ответ нейрона; μ_0 — порог активации нейрона; $m_o(a_j)$ и $\sigma_o(a_j)$ — математическое ожидание и среднеквадратичное отклонение значений j-го признака для образа «Свой»; $m_s(a_j)$ и $\sigma_s(a_j)$ — аналогичные показатели образов для «Чужих». Таким образом, сеть из нейронов (1) с функцией активации (2) после обучения представляет собой нейросетевой ПБК. Ответ нейросетевого ПБК складывается из битовых значений на выходах нейронов (путем их конкатенации).

Если нейрон настроен на выход «1» при поступлении образа «Свой», то знак весового коэффициента выбирается исходя из правила: «+» при $m_s(a_j) < m_o(a_j)$, иначе «−». Если нейрон настраивается на «нулевой» бит, то знаки инвертируются. Параметры $m_o(a_j)$, $\sigma_o(a_j)$, $m_s(a_j)$ и $\sigma_s(a_j)$ после обучения удаляются, чтобы не компрометировать эталон. Остаются таблицы связей и весов μ, из которых нельзя непосредственно вычислить $m_o(a_j)$. Параметры обученных нейронов (связи и веса) называют *нейросетевым контейнером*.

Порог μ_0 нейрона обычно настраивается исходя из откликов нейрона на обучающие примеры «Чужой» по правилу

$$\mu_0 = m_s(y)\alpha, \quad (4)$$

где α — единый для всех нейронов эмпирически подбираемый коэффициент, влияющий на баланс между FRR и FAR. В теории такой способ должен дать вероятность ошибки «ложного принятия» нейроном образа «Чужой», приближенно равную 0,5. Если допустить, что выходы нейронов независимы, каждый добавляемый нейрон, который настраивается по формуле (4), должен снижать FAR примерно в 2 раза. Информационная энтропия (далее просто энтропия) ответов нейросетевого ПБК на образы «Чужих» в этом случае должна быть близка к длине ключа. Но в действительности чем больше входов у нейронов, тем выше корреляция между их выходами и FAR, но чем больше информации поступает на вход каждому нейрону, тем ниже FRR. Действует и обратная логика: чем больше нейронов, тем ниже FAR, но выше FRR.

Выполненные в настоящей работе оценки показали высокий уровень FRR при настройке порогов нейронов по формуле (4), поэтому предложен альтернативный способ настройки порогов нейронов исходя из откликов y на примеры «Свой», не использовавшиеся при вычислении весов:

$$\mu_0 = m_o(y) - \sigma_o(y)\alpha.$$

Число входов каждого нейрона предлагается определять так, чтобы сумма информативности связанных с ними признаков была $\Sigma I_{\text{bit}} > 1$. Номера связанных с нейроном признаков определяются случайно, но среди тех, коэффициент корреляции r между которыми принимает высокие значения ($r > 0{,}5$).

Предлагаемая модель гибридного ПБК на базе разностных нейронов Байеса и классических нейронов

Функционал (1) теряет мощность при усилении корреляционных связей между признаками. По этой причине количество информации на входе классического нейрона всегда меньше, чем сумма собственной информации (см. рис. 3) всех признаков. Абсолютно иначе дело обстоит, если вместо функционала (1) использовать разностный байесовский функционал

$$d_t = \sum_{j=1}^{n}\left|\frac{|m_o(a_t)-a_t|}{\sigma_o(a_t)} - \frac{|m_o(a_j)-a_j|}{\sigma_o(a_j)}\right|, \; j \neq t. \quad (5)$$

Многомерный разностный функционал Байеса (5) дает тем меньшее значение, чем выше коэффициент корреляции признака под номером t с признаками под номерами j [4]. На его основе возможна нейросетевая обработка коррелированных сочетаний признаков. Недостатком является то, что он полностью компрометирует биометрический эталон (параметры $m_o(a_j)$ не должны использоваться при расчете близости после обучения ПБК).

Предлагается функционал, обладающий аналогичными свойствами, но компрометирующий биометрический эталон лишь частично:

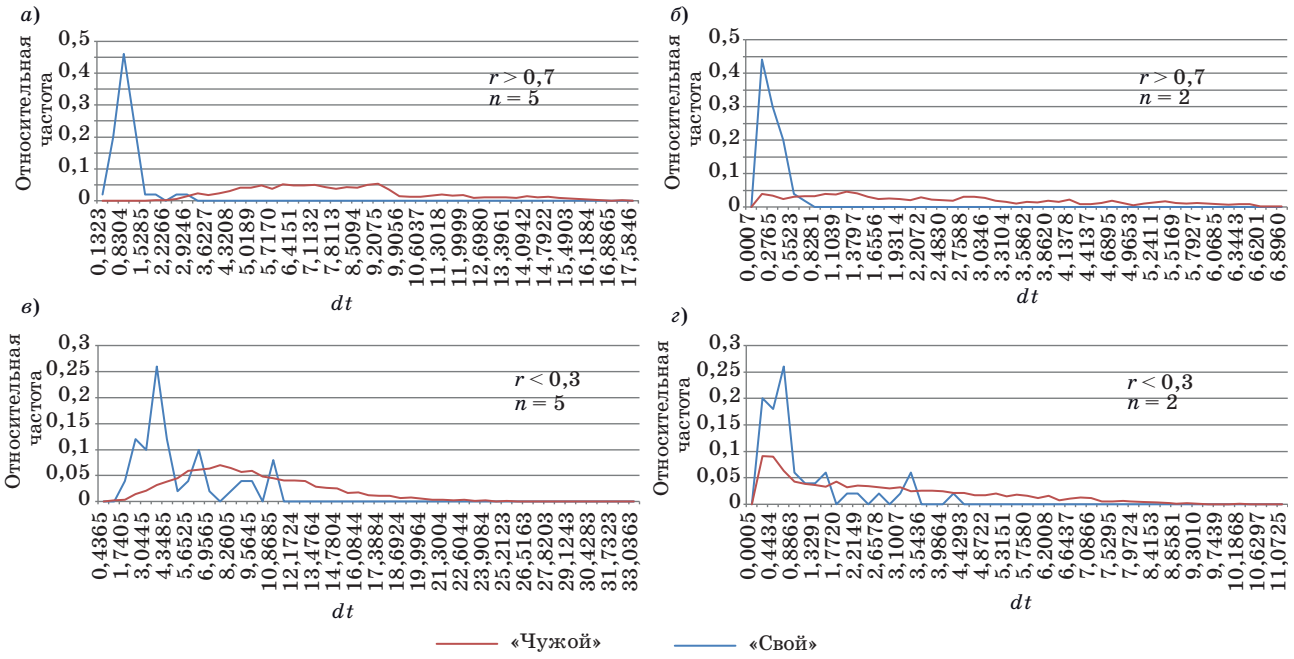
$$d_t = \sum_{j=1}^{n}\left|\left|\frac{a_t}{\sigma_o(a_t)} - \frac{a_j}{\sigma_o(a_j)}\right| - \Delta m_{tj}\right|,$$
$$\Delta m_{tj} = \left|\frac{m_o(a_t)}{\sigma_o(a_t)} - \frac{m_o(a_j)}{\sigma_o(a_j)}\right|, \; j \neq t. \quad (6)$$

Свойства функционала (6) демонстрируются на рис. 4, *а–г*: чем больше n и выше корреляция между признаками r, тем меньше вероятность ошибок распознавания образа. Это справедливо для любых признаков, независимо от их физического смысла. Убедиться в этом несложно, достаточно сгенерировать описания абстрактных классов образов в двух пространствах признаков с нормальным законом распределения: независимых и зависимых. Сгенерировать значения независимых признаков под заданные параметры распределения можно методом Монте-Карло (классы должны отличаться параметрами $m_o(a_j)$ и $\sigma_o(a)$). Воссоздать положительную зависимость между признаками можно, отдельно сортируя значения каждого признака по возрастанию [31]. Далее следует построить нейрон Байеса на основе функционала (6), сформировать образы из сгенерированных данных (в двух вариантах — на основе независимых признаков и зависимых) и обработать эти образы при помощи нейрона Байеса, после чего построить эмпирические плотности вероятности откликов (6) этого нейрона (см. рис. 4).

Разностные нейроны Байеса на базе функционала (6) с функцией активации (2) обучаются путем вычисления параметров Δm и $\sigma_o(a)$. Это позволяет не хранить параметры $m_o(a)$. Введем также функционал

$$d_t = \sum_{j=1}^{n}\left|\frac{a_t}{\sigma(a_t)} - \frac{a_j}{\sigma(a_j)}\right|, \; j \neq t \quad (7)$$

обладающий идентичными свойствами в плане обработки сильно коррелированных признаков. Преимущество метрики (7) заключается в том,



■ ***Рис. 4.*** Эмпирические плотности вероятности откликов разностных нейронов Байеса (6) при распознавании абстрактных образов при $I_{\text{bit}} \approx 0{,}5$ в случае зависимых ($r > 0{,}7$) и независимых ($r < 0{,}3$) признаков: *а*, *в* — $n = 5$; *б*, *г* — $n = 2$

■ ***Fig. 4.*** Empirical probability densities of responses of Bayes difference neurons (6) when recognizing abstract images at $I_{\text{bit}} \approx 0.5$: in the case of dependent ($r > 0.7$) and independent ($r < 0.3$) features: *а*, *в* — $n = 5$; *б*, *г* — $n = 2$

что для ее корректной работы требуется хранить только параметры $\sigma(a)$. Это могут быть среднеквадратичные отклонения признаков как для класса образов «Свой» $\sigma_o(a)$, так и для класса «Чужие» $\sigma_s(a)$, что также допустимо. При этом $\sigma_s(a)$ вообще не компрометирует биометрический эталон пользователя.

Порог срабатывания функции активации (2) для разностного нейрона Байеса вычисляется по правилу

$$\mu_0 = m_o(d_t) + \sigma_o(d_t)\beta,$$

где β — коэффициент баланса FRR и FAR для нейронов Байеса. Несмотря на идентичные свойства метрик (5)–(7), стопроцентной корреляции между их откликами d_t не наблюдается, что позволяет строить разностные нейроны Байеса на основе всех указанных метрик и объединять их в единую сеть. Каждый нейрон должен быть связан с признаками, коэффициент корреляции между которыми принимает достаточно высокие значения ($r_{t,j} > 0{,}5$). Количество входов нейрона не лимитируется: чем больше признаков с примерно равным (и высоким) уровнем взаимной корреляции входит в нейрон, тем стабильнее он работает.

Классические нейроны показывают хороший результат, если признаки достаточно информативны и слабо коррелированы ($r_{t,j} < 0{,}5$). Разностные нейроны Байеса обладают почти противоположными свойствами: они ориентированы на обработку сильно зависимых признаков ($r_{t,j} > 0{,}5$), так как неявно извлекают дополнительную информацию из корреляционной матрицы признаков, что было продемонстрировано на рис. 4. Таким образом, если объединить классические нейроны и разностные нейроны Байеса в гибридную сеть, можно снизить показатели FRR и FAR, а также повысить энтропию ответов ПБК. При формировании гибридной сети следует отдельно настраивать сегмент из классических нейронов и сегмент из байесовских нейронов, объединяя их в один слой. Формирование связей и настройка классических нейронов может производиться в соответствии с ГОСТ Р 52633.5 [13], с единственным обязательным отличием — при создании связей для каждого нейрона выбираются признаки с уровнем взаимной корреляции $r_{t,j} < 0{,}5$. При создании связей для нейрона Байеса признаки выбираются случайно [4], но среди тех, что имеют взаимную корреляционную зависимость $r_{t,j} > 0{,}5$.

При построении гибридного ПБК необходимо учитывать тот факт, что разностные нейроны Байеса компрометируют биты ключа пользователя. Каждый нейрон должен давать на выходе один бит ключа (верный или неверный, зависит от преодоления порога μ_0 нейроном), который необходимо хранить (решение этой проблемы будет показано далее).

Об оценке стойкости ПБК к атакам

Энтропия ответов ПБК на образы «Чужие» является важным показателем, так как она связана с FAR: чем ниже FAR, тем выше энтропия [19]. Приблизительную оценку энтропии можно получить, вычислив собственную информацию события «ложного допуска»:

$$E(\mathrm{FAR}) \approx -\log_2 \mathrm{FAR}.$$

Точный расчет многомерной энтропии длинных бинарных последовательностей прямым численным экспериментом является технически нерешаемой задачей, как и расчет сверхнизких показателей FAR. Чтобы получить $E(\mathrm{FAR}) = 256$ бит, требуется, чтобы $\mathrm{FAR} < 10^{-77}$. Для проверки такой экстремально низкой вероятности прямым численным экспериментом не хватит населения планеты (даже если каждый человек придумает тысячи независимых рукописных паролей). По причине сложности сбора больших выборок к биометрическим системам предъявляются не столь жесткие требования, как к паролям.

Рассмотрим подходы к оценке FAR в биометрических системах.

Первый подход заключается в проведении прямого численного эксперимента и вычислении FAR как отношения числа ошибок к числу опытов по распознаванию «Чужих» с определением доверительных вероятности и интервала. Такая идеология лежит в основе ISO/IEC 19795-1 [5]. Если следовать только этому подходу (например, правилам «трех» или «тридцати»), то оценить надежность ПБК, обладающего действительно высокой энтропией, невозможно.

Второй подход основан на построении двух функций плотности вероятности для расстояний Хэмминга между ключом (паролем) пользователя и ответом ПБК (нейронной сети или нечеткого экстрактора) на образы «Свой» и образы «Чужие» соответственно. Далее вычисляется площадь их пересечения (по аналогии с процедурой оценки информативности признака, см. рис. 3). Если задать *порог принятия*, то можно вычислить FRR и FAR [19]. Однако данный способ дает приблизительную оценку. Для большей точности вводятся поправки, учитывающие показатели стабильности, информативности и коррелированности признаков. Плотности вероятности для ПБК можно описать нормальным законом распределения лишь условно, так как для иных архитектур ИНС (например, гибридных, состоящих не только из классических нейронов) закон распределения расстояний Хэмминга может быть другой.

Третий подход изложен в ГОСТ Р 52633.3-2011 [11]. Согласно стандарту, в эксперименте необходимо использовать не только естественные образы «Чужой», но и синтетические, генерируемые на основе скрещивания естественных (по методике ГОСТ Р 52633.2-2010 [10]). Если для естественных образов «Чужих» не наблюдается ошибок 2-го рода («ложного совпадения ключа»), то естественные образы скрещиваются. При скрещивании следует выбирать пары из «Чужих», которые дают ответы ПБК, наиболее близкие в метрике Хэмминга к ключу пользователя. Далее по аналогичному принципу могут скрещиваться синтетические образы. Каждая новая популяция синтетических образов «Чужих» все ближе к образу «Свой». Процесс тестирования ПБК прекращается, когда для очередной популяции будут зафиксированы ошибки 2-го рода или синтезировано определенное число популяций «Чужих». В процессе тестирования сокращается количество попыток предъявления конкурирующих примеров, а точность оценки FAR увеличивается на порядки при сохранении статистической значимости.

Результаты по оценке вероятности ошибочных решений нейросетевых ПБК

При разработке алгоритма тестирования ПБК решено опираться на третий подход (ГОСТ 52633.3-2011 [11]) и метод перекрестного сравнения (кроссвалидации) [4, 18].

По требованиям ГОСТ 52633.3 для тестирования средств высоконадежной биометрической аутентификации при условии высокого уровня взаимного доверия между донором биометрии и владельцем средства биометрической аутентификации (тестирования) необходимо не менее 128 примеров естественных образов «Чужих», которые ранее не были использованы при обучении тестируемого ПБК. Используемые при тестировании биометрические образы должны быть независимыми и формироваться по ГОСТ Р 52633.1 [9]. В соответствии с ГОСТ 52633.5 для обучения нейросетевого ПБК требуется не менее 64 независимых образов «Чужих» [13]. Под независимыми образами подразумеваются примеры различных рукописных паролей, воспроизведенных разными субъектами.

Для каждого испытуемого генерировался случайный ключ и формировался ПБК. Для обучения ПБК пользователя использовалось 15 примеров его образа и 64 примера образов других случайно выбранных испытуемых («Чужих»).

Далее проводились испытания надежности работы ПБК. Для оценки FRR использовалось по 30 образов от каждого испытуемого, не вошедших в обучающую выборку. FRR определялась как отношение числа зарегистрированных ошибок «ложного отказа» к общему количеству опытов ($30 \times 260 = 7800$).

Чтобы оценить вероятность ошибки «ложного допуска», относительно каждого испытуемого формировалась тестовая выборка образов «Чужих», которая состояла из примеров рукописного пароля оставшихся 195 подписантов, не вошедших в выборку обучения (бралось по 10 образов на подписанта). Кроме того, в тестовую выборку испытуемого вошли 50 подделок его образа. Так для каждого испытуемого сформирована тестовая выборка из 2000 естественных образов «Чужих» (всего $2000 \times 260 = 520\,000$ примеров).

Тестирование FAR выполнялось независимо для каждого испытуемого на основании отобранных 2000 примеров «Чужих». Если ошибки не фиксируются, производится скрещивание 10 % «Чужих» (20 образов из 2000, воспроизведенных различными подписантами), которые дают ответы ПБК, наиболее близкие в метрике Хэмминга к ключу пользователя, по следующей формуле:

$$a_{j,k} = \frac{c+1-k}{c+1} a_{j,A} + \frac{k}{c+1} a_{j,B},$$

где c — количество синтетических примеров, порождаемых парой «сильных Чужих» A и B предыдущего поколения; k — номер синтетического примера; j — номер признака. Этот способ синтеза «Чужих» сохраняет естественные корреляционные связи образов A и B.

Формируется новая популяция из 2000 синтетических образов, для которых также вычисляются ответы ПБК. Далее по аналогичному принципу скрещиваются синтетические образы. Тестирование повторяется, пока не будут зафиксированы ошибки или расстояние Хэмминга между ответами ПБК и ключом испытуемого не перестанет уменьшаться. FAR определялась по формуле

$$FAR = \frac{\sum_{i=1}^{260} \sum_{p=0}^{p\_\max_i} \frac{er_{i,p}}{2 \cdot 10^{3+p}}}{260},$$

где $p\_\max_i$ — количество популяций «Чужих» для i-го испытуемого; $er_{i,p}$ — количество ошибок «ложного допуска» для i-го испытуемого из блока при тестировании на примерах из p-й популяции.

Описанная методика дает точную и достоверную ненулевую оценку FAR. Выборочные результаты тестирования нескольких вариаций ПБК на базе однослойных «широких» сетей представлены в табл. 2.

При обучении MLP на больших объемах данных «глубину» персептрона обычно стараются повысить, а размерность пространства признаков снизить (например, с помощью PCA). Для нейросетевых автоматов биометрической аутентификации увеличение количества слоев не актуально. Гораздо эффективнее соизмеримо повышать число входов и выходов сети. Это приводит к медленному росту качества решений — снижению вероятностей ошибок и повышению энтропии ответов ПБК при предъявлении образов «Чужих» (см. табл. 2). Однако рост постепенно останавливается. После этого наращивать число нейронов не имеет смысла. По этой причине энтропия ответов нейросетевых ПБК на образы «Чужих» не соответствует их длине. Такой подход к повышению стойкости ПБК имеет общие черты с повышением стойкости парольной защиты путем увеличения длины их контрольных сумм (хотя реальная стойкость также зависит от степени случайности самих паролей).

■ ***Таблица 2***. Результаты тестирования ПБК на базе классических нейронов

■ ***Table 2***. Results of testing classical neural net "biometrics to code" converters

Число нейронов L и входов n	FRR, % (порог = 0)	FAR, % (порог = 0)	EER, % (порог > 0)	α
$L = 1024$, $n \geq 5(\Sigma I > 1)$	14	0,0016	1,9	4,75
$L = 512$, $n \geq 5(\Sigma I > 1{,}5)$	17,5	0,0072	2,2	4
$L = 512$, $n \geq 5(\Sigma I > 1)$	14,5	0,0019	1,9	4,5
$L = 512$, $n = 10$	16	0,0058	2	4
$L = 512$, $n = 20$	23	0,0207	2,7	3
$L = 512$, $n = 150$	50	0,233	8,3	1,5
$L = 256$, $n = 10$	25	0,0094	2,3	3
$L = 128$, $n = 10$	19	0,0333	2,8	3
$L = 64$, $n \geq 5(\Sigma I > 1)$	32	0,0348	3,9	3
$L = 64$, $n = 10$	26,5	0,0462	4,1	2,5
$L = 64$, $n = 10$	53,5	0,0125	4,1	2
$L = 64$, $n = 50$	12	0,4671	6	2

Установлено, что свыше 50 % «наиболее близких "Чужих"» попали в намеренные подделки. Обычно это происходит, когда пароль испытуемого является достаточно простым (из четырехпяти символов). Самые сильные синтетические образы часто получались из двух подделок, воспроизведенных разными людьми. Таким образом, динамика рукописного пароля не является абсолютно устойчивой к подделкам, выполняемым путем копирования внешнего вида.

Результаты по тестированию нейросетевых ПБК на предмет возможности извлечения знаний

Процедура обучения «широкой» ИНС является однонаправленной и не подразумевает обратной разработки. Однако восстановление биометрического образа и личного ключа пользователя из таблиц нейросетевых функционалов все же возможно.

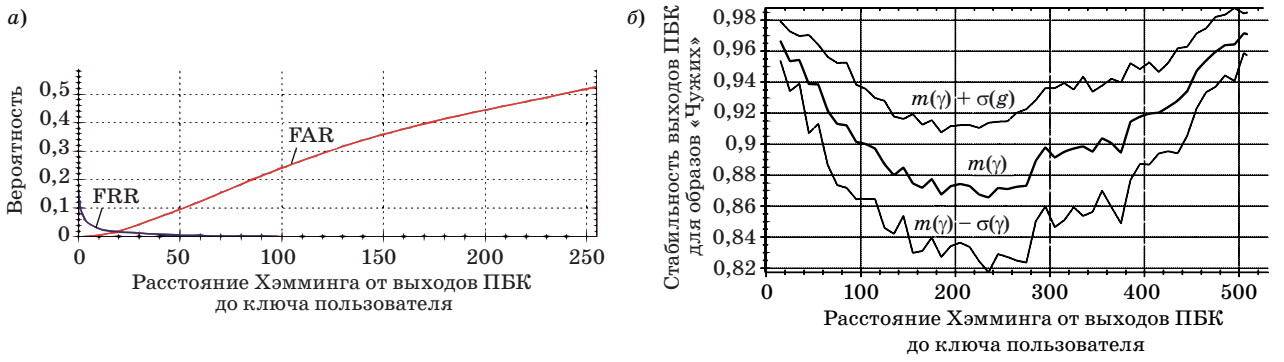
Контролируя допустимое число ошибочных бит в ответе «широкой» сети, можно балансировать FRR и FAR (например, применяя корректирующие коды или второй слой нейронов для исправления нескольких неверных бит) (рис. 5, *а*). Однако эта возможность одновременно является уязвимостью. Хакер может собрать большую базу примеров произвольных паролей, воспроизведенных различными подписантами («Чужими»), и оценить среднюю стабильность ответов ПБК для каждого подписанта («Чужого») по формуле

$$\gamma_k = \sum_{l=1}^{L} 2\left|P_l(1) - 0{,}5\right|, \qquad (8)$$

где k — номер подписанта («Чужого»); L — количество нейронов; l — номер нейрона; $P_l(1)$ — вероятность (или относительная частота) появления «единицы» (можно заменить на $P_l(0)$) в l-м разряде ответа ПБК (в выходе l-го нейрона) на примеры образа k-го «Чужого». Оценка относительных частот $P_l(1)$ и $P_l(0)$ может проводиться на основании нескольких образов подписанта (в настоящей работе использовалось по 10 примеров от каждого «Чужого»).

Вычислены показатели стабильности ответов ПБК (8) при предъявлении примеров каждого «Чужого». По результатам эксперимента показатель стабильности (8) для каждого «Чужого» оказался зависим от среднего количества ошибочных бит соответствующих ответов ПБК (рис. 5, *б*). Таким образом, даже в отсутствие явной индикации близости ответов ПБК и ключа пользователя хакер может осуществить направленный перебор синтетических образов, скрещивая примеры рукописных паролей разных «Чужих», которые дают наиболее стабильный ответ. Через несколько поколений скрещивания, «двигаясь» в направлении повышения стабильности ответов ПБК, удается подобрать «Чужого», почти или полностью идентичного образу «Свой» (в зависимости от объема исходной базы «Чужих», наличия в базе злоумышленника подделок образа «Свой» и того, насколько качественно они выполнены). Данная атака снижает количество вариантов перебора на несколько порядков. Даже если злоумышленник не обладает примерами подделок и какой-либо информацией о пароле пользователя-жертвы, данная атака вполне осуществима (в этом случае нарушителю потребуется гораздо больше времени).

Также видно (см. рис. 5, *б*), что образы «Чужие» обладают так называемой *«симметрией» стабильности ответов относительно образа «Свой»* (свойством «симметрии»). Это означает, что стабильность ответов ПБК при предъявлении образов «Чужих» возрастает, но не только если ответы близки (в метрике Хэмминга) к ключу пользователя, но и если они близки к инверсии ключа (инверсный код возникает, если все биты ответа ПБК являются ошибочными). Инверсный код можно обратить и получить ключ пользователя. Из этого следует, что у каждого образа «Свой» в нейросетевом логическом базисе существует его инверсия. Если на вход классического нейросетевого ПБК (обученного по ГОСТ 52633.5 [13]) по-



■ ***Рис. 5.*** Результаты тестирования классических нейросетевых ПБК с параметрами $L = 1024$, $n \geq 5$: *а* — вероятности ошибок в зависимости от порога принятия; *б* — стабильность ответов ПБК на образы «Чужих» в зависимости от количества ошибочных бит

■ ***Fig. 5.*** Test results of classical neural net “biometrics to code” converter with parameters $L = 1024$, $n \geq 5$: *а* — probability of errors depending on the threshold of acceptance; *б* — stability of converter responses to “Strangers” depending on the number of error bits

дать инверсию образа «Свой», то на выходе ПБК появится инверсный ключ, который можно обратить. Данное свойство позволяет ускорить процедуру направленного перебора биометрических образов в 2 раза (осуществляя одновременно поиск наиболее близкого и наиболее дальнего образа «Чужого» относительно образа «Свой»).

Предлагаемый способ защиты гибридных нейросетевых контейнеров

В работе [19] предложено защищать нейросетевые контейнеры путем применения обратимых и необратимых преобразований. Усовершенствуем данный подход, чтобы защитить гибридный ПБК.

Каждый нейрон имеет таблицы связей и весов. Для защиты таблиц нейросетевых функционалов нужно применять механизм защищенного нейросетевого контейнера (ЗНК) (рис. 6, *а* и *б*). Нейроны можно выстроить в цепочку (см. рис. 6, *а*). После обучения ПБК таблицы каждого нейрона шифруются наложением гаммы, представляющей собой контрольную сумму выходов всех предыдущих нейронов в цепочке:

$$tables'_l = \\ = XOR(tables_l, hash(pass, bit_1, \ldots, bit_{l-1})), \quad (9)$$

где $tables_l$ — таблицы параметров соответствующего нейрона; $hash()$ — криптографическая хеш-функция (например, md5); $pass$ — пароль, который является опциональным и служит для дополнительной (двухфакторной) защиты; bit_l — выход, на который настраивается l-й нейрон в цепочке. В настоящем исследовании пароль не использовался.

При обработке биометрического образа нейросетевым ПБК в режиме ЗНК происходит процесс «распаковки» нейронов — параметры каждого следующего нейрона в цепочке дешифруются по той же формуле (9). Для получения на выходе ПБК верного ключа пользователя требуется, чтобы все нейроны «проголосовали» правильно. Если хотя бы один нейрон в цепочке выдаст ошибочный бит, это повлечет неверную дешифровку параметров всех последующих нейронов. В свою очередь последующие нейроны будут давать случайные выходы, и возникнет эффект хеширования биометрического образа «Чужого». В итоге ответы нейросетевого ПБК становятся случайными, их энтропия возрастает. При этом важен тот факт, что FRR и FAR в режиме ЗНК не меняются при пороге принятия, равном нулю (рис. 7, *а*). Стабильность ответов ПБК при поступлении на вход образов «Чужих» становится низкой и перестает возрастать, если образ «Чужого» близок к образу «Своего» (график на рис. 5, *б* в режиме ЗНК становится почти прямым, рис. 7, *б*). Однако режим ЗНК все же накладывает ограничения: балансировать FRR и FAR, корректируя несколько ошибочных ответов ПБК, становится невозможно.

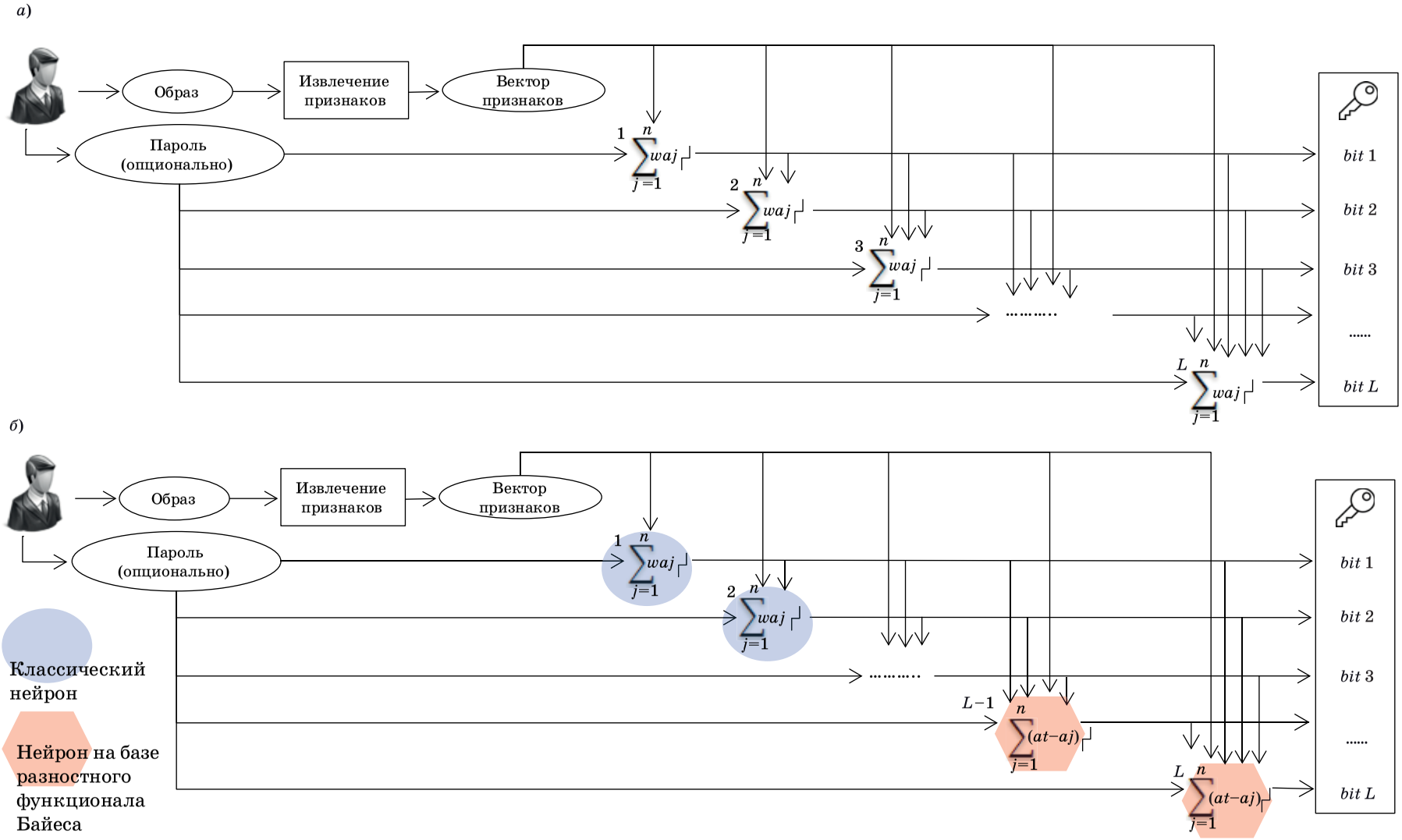
Результаты по оценке вероятностей ошибочных решений гибридных ПБК с применением предложенной схемы защиты

Разностные нейроны Байеса нужно использовать совместно с классическими нейронами, применяя механизм ЗНК (см. рис. 6, *б*). Предлагается размещать классические нейроны в начале цепочки, а байесовские нейроны — в конце (в силу того, что последние в незащищенном виде компрометируют часть ключа). Нейроны Байеса будут надежно защищены, если классических нейронов будет много (достаточно 256).

В режиме ЗНК порядок нейронов в цепочке (см. рис. 6, *б*) играет важную роль. Предлагается расположить классические нейроны в порядке уменьшения суммарной информативности их входов (ΣI_{bit}), а разностные нейроны Байеса расположить в порядке снижения размерности и корреляционной зависимости входов. В этом случае удается повысить энтропию ответов ПБК при идентичных показателях FRR и FAR. Этот эффект имеет простое объяснение. Если сначала располагаются нейроны с более стабильной статистикой выходов, то при поступлении образа «Чужой» эти нейроны среагируют первыми, и процесс «хеширования» запустится раньше (больше нейронов будет дешифровано неверно), при поступлении образа «Свой» ничего не изменится.

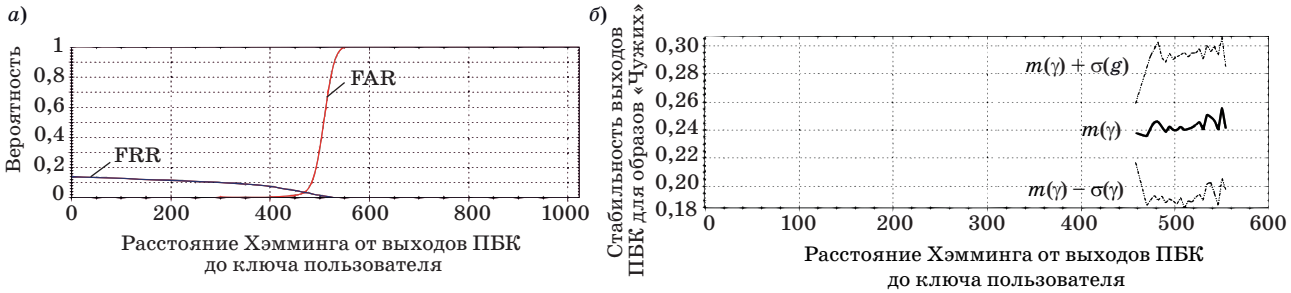
По результатам эксперимента (рис. 8, *а–в*) установлено, что средние показатели стабильности ответов гибридного ПБК при поступлении образов «Чужих» в режиме ЗНК гораздо ниже ($m(\gamma) = 0{,}24$), чем без защиты ($m(\gamma) = 0{,}92$), и еще ниже ($m(\gamma) = 0{,}23$) при ранжировании нейронов в соответствии с информативностью и коррелированностью входов. Также можно видеть, что у гибридного ПБК отсутствует уязвимость, связанная со свойством «симметрии» (см. рис. 8, *а*).

Установлено, что комплексирование двух видов нейронов существенно снижает вероятности ошибок: FRR = 11,5 %, FAR = 0,0009 %, EER ≈ ≈ 1,6 %, $L = 1024$ (512 классических и 512 нейронов Байеса), $\alpha = 4{,}25$, $\beta = 25{,}5$ (рис. 9, *а–в*). Результат [29] превосходит полученный в настоящей работе потому, что при тестировании в работе [29] не учитывались подделки и для вычисления FAR применялся менее точный метод



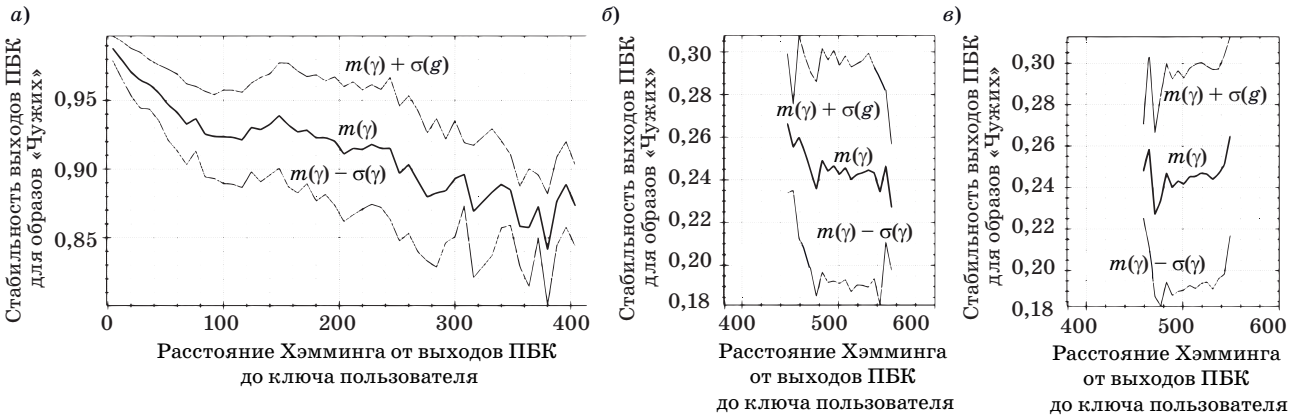
■ ***Рис. 6***. Механизм ЗНК применительно к нейросетевому ПБК на базе классических нейронов (*а*) и гибридному ПБК (*б*)

■ ***Fig. 6***. The mechanism of a protected neural network container in relation to converter “biometrics to code” based on classical neurons (*a*) and hybrid converter (*б*)



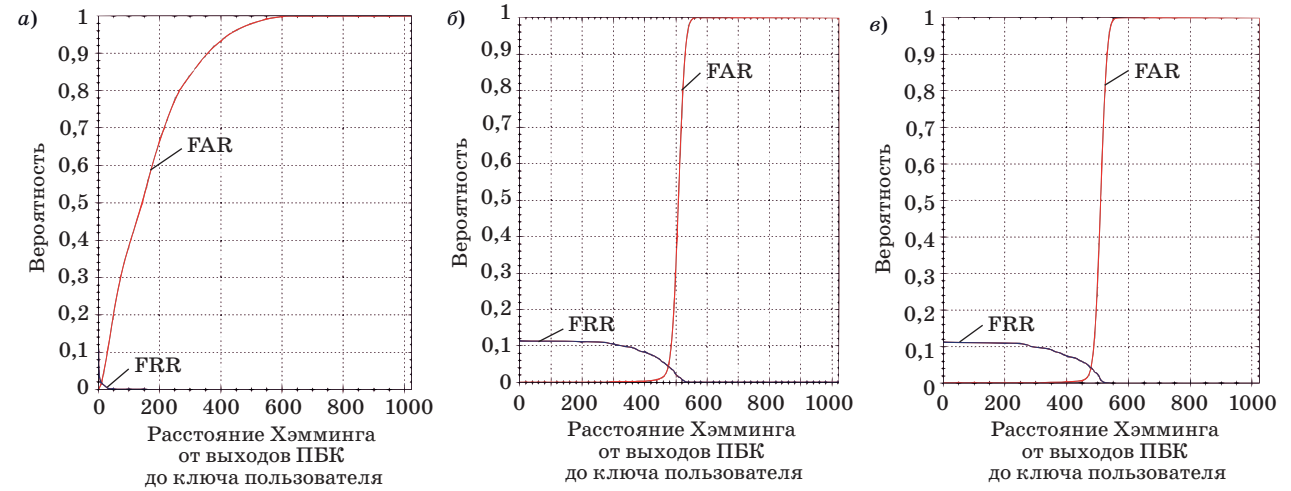
■ ***Рис. 7.*** Результаты тестирования классических нейросетевых ПБК в режиме ЗНК с параметрами $L = 1024$, $n \geq 5$: *а* — вероятности ошибок в зависимости от порога принятия; *б* — стабильность ответов ПБК на образы «Чужих» в зависимости от количества ошибочных бит

■ ***Fig. 7.*** Testing results of classical neural net "biometrics to code" converters in protection mode with parameters $L = 1024$, $n \geq 5$: *а* — probability of errors depending on the threshold; *б* — stability of converters responses to "Strangers" images depending on the number of erroneous bit



■ ***Рис. 8.*** Стабильность ответов гибридных ПБК при поступлении образов «Чужих» в зависимости от числа ошибочных бит: *а* — обычный режим; *б* — режим ЗНК; *в* — режим ЗНК с ранжированием нейронов

■ ***Fig. 8.*** Stability of hybrid "biometrics to code" responses to "Strangers" images depending on the number of erroneous bits: *а* — normal mode; *б* — protection mode; *в* — protection mode with ranking of neurons



■ ***Рис. 9.*** Вероятности ошибок в зависимости от порога принятия: *а* — обычный режим; *б* — режим ЗНК; *в* — режим ЗНК с ранжированием нейронов

■ ***Fig. 9.*** Probabilities of errors depending on the threshold: *а* — normal mode; *б* — protection mode; *в* — protection mode with ranking of neurons

(на базе второго подхода — оценки пересечения функций плотности вероятности для расстояний Хэмминга между ответами ПБК и ключами).

Снижение вероятностей ошибок указывает на то, что ошибочные решения нейронов Байеса слабо коррелированы с ошибками классических нейронов.

Заключение

Установлено, что механизм защищенного нейросетевого контейнера можно успешно применять в гибридных нейронных сетях, состоящих из классических нейронов и разностных нейронов Байеса. Предположены новые варианты построения разностных нейронов Байеса, не компрометирующих и частично компрометирующих биометрический эталон пользователя (даже без применения метода защиты нейросетевых контейнеров). Продемонстрирована их эффективность при распознавании образов в пространстве сильно коррелированных признаков.

Экспериментально подтверждена высокая надежность верификации рукописных образов на базе предложенной модели гибридной нейробайесовской сети (с учетом предъявления подделок рукописных паролей испытуемых). Показатели ошибок аутентификации (высвобождения ключа пользователя) составили: FRR = 11,5 %, FAR = 0,0009 % (EER ≈ 1,6 %) при длине ключа 1024 бита. Достигнутые показатели не являются предельными.

Направления будущих исследований могут быть связаны с применением механизмов защиты нейросетевых контейнеров в отношении других архитектур гибридных нейронных сетей, способных к быстрому обучению.

Финансовая поддержка

Работа выполнена при поддержке Российского научного фонда по гранту № 17-71-10094.

Financial support

This work was supported by the Russian Science Foundation, No. 17-71-10094.

Литература

1. **Akhmetov B. S., Ivanov A. I., Alimseitova Z. K.** Training of neural network biometry-code converters. *2018 News of the National Academy of Sciences of the Republic of Kazakhstan. Series of Geology and Technical Sciences*, 2018, vol. 1, no. 427, pp. 61–68.
2. **Jain A. K., Nandakumar K., Nagar A.** Biometric template security. *EURASIP J. Adv. Signal Process*, 2008, pp. 113:1–113:17.
3. **Hafemann L. G., Sabourin R., Oliveira L. S.** Characterizing and evaluating adversarial examples for offline handwritten signature verification. *IEEE Transactions on Information Forensics and Security*, 2019, vol. 14, iss. 8, pp. 2153–2166. doi:10.1109/TIFS.2019.2894031
4. **Ivanov A. I., Lozhnikov P. S., Sulavko A. E.** Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms. *Computer Optics*, 2017, no. 5, pp. 765–774. doi:10.18287/2412-6179-2017-41-5-765-774
5. ISO/IEC 19792:2009. Information technology — Security techniques — Security evaluation of biometrics. International Organization for Standardization, 2009. 37 p.
6. ISO/IEC 24761:2009. Information technology — Security techniques — Authentication context for biometrics. International Organization for Standardization, 2011. 50 p.
7. ISO/IEC 24745:2011. Information technology — Security techniques — Biometric information protection. International Organization for Standardization, 2011. 50 p.
8. ГОСТ Р 52633.0-2006. Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации. М., Стандартинформ, 2007. 25 с.
9. ГОСТ Р 52633.1-2009. Защита информации. Техника защиты информации. Требования к формированию баз естественных биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации. М., Стандартинформ, 2010. 24 с.
10. ГОСТ Р 52633.2-2010. Защита информации. Техника защиты информации. Требования к формированию синтетических биометрических образов, предназначенных для тестирования средств высоконадежной биометрической аутентификации. М., Стандартинформ, 2011. 22 с.
11. ГОСТ Р 52633.3-2011. Защита информации. Техника защиты информации. Тестирование стойкости средств высоконадежной биометрической защиты к атакам подбора. М., Стандартинформ, 2012. 16 с.
12. ГОСТ Р 52633.4-2011. Защита информации. Техника защиты информации. Интерфейсы взаимодействия с нейросетевыми преобразователями биометрия-код. М., Стандартинформ, 2012. 46 с.
13. ГОСТ Р 52633.5-2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа. М., Стандартинформ, 2012. 20 с.

14. ГОСТ Р 52633.6-2012. Защита информации. Техника защиты информации. Требования к индикации близости предъявленных биометрических данных образу «Свой». М., Стандартинформ, 2012. 24 с.
15. **Dodis Y., Reyzin L., Smith A.** Fuzzy extractors: How to generate strong keys from biometrics and other noisy. *EuroCrypt*, 2004, pp. 523–540.
16. **Иванов А. И., Сомкин С. А., Андреев Д. Ю., Малыгина Е. А.** О многообразии метрик, позволяющих наблюдать реальные статистики распределения биометрических данных «нечетких экстракторов» при их защите наложением гаммы. *Вестник УрФО. Безопасность в информационной сфере*, 2014, № 2(12), с. 16–23.
17. **Ignatenko T., Frans M. J.** Willems. Information leakage in fuzzy commitment schemes. *IEEE Transactions on Information Forensics and Security*, 2010, vol. 5, no. 2, pp. 337–348. doi:10.1109/TIFS.2010.2046984
18. **Ложников П. С., Сулавко А. Е., Еременко А. В., Волков Д. А.** Экспериментальная оценка надежности верификации подписи сетями квадратичных форм, нечеткими экстракторами и персептронами. *Информационно-управляющие системы*, 2016, № 5, с. 73–85. doi:10.15217/issn1684-8853.2016.5.73
19. **Ахметов Б. С., Иванов А. И., Фунтиков В. А., Безяев А. В., Малыгина Е. А.** *Технология использования больших нейронных сетей для преобразования нечетких биометрических данных в код ключа доступа*. Алматы, LEM, 2014. 144 с.
20. **Kůrková V., Sanguineti M.** Probabilistic lower bounds for approximation by shallow perceptron networks. *Neural Networks*, 2017, vol. 91, pp. 34–41.
21. **Kůrková V., Sanguineti M.** Model complexities of shallow networks representing highly varying functions. *Neurocomputing*, 2016, vol. 171, pp. 598–604.
22. **Иванов А. И.** Нейросетевая защита конфиденциальных биометрических образов гражданина и его личных криптографических ключей. Пенза, ПНИЭИ, 2014. 57 с.
23. **Iranmanesh V., Ahmad S. M. S., Adnan W. A. W., Yussof S., Arigbabu O. A., Malallah F. L.** Online handwritten signature verification using neural network classifier based on principal component analysis. *Scientific World Journal*, 2014, vol. 2014, pp. 1–8.
24. **Iranmanesh V.** Online signature template protection by shuffling and one time pad schemes with neural network verification. *Proceedings of the International Conference on Computer Science and Computational Mathematics (ICCSCM '13)*, 2013, pp. 53–59.
25. **Hafemann L. G., Sabourin R., Oliveira L. S.** Writer-independent feature learning for offline signature verification using deep convolutional neural networks. *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016. doi:10.1109/IJCNN.2016. 7727521
26. **Souza V. L. F., Oliveira A. L. I., Sabourin R.** A writer-independent approach for offline signature verification using deep convolutional neural networks features. *2018 7th Brazilian Conference on Intelligent Systems (BRACIS)*, 2018. doi:10.1109/BRACIS.2018. 00044
27. **Díaz M., Fischer A., Ferrer M. A., Plamondon R.** A perspective analysis of handwritten signature technology. *ACM Computing Surveys*, 2019, vol. 51, iss. 6, article 117, pp. 1–37.
28. **Maiorana E., Campisi P.** Fuzzy commitment for function based signature template protection. *IEEE Signal Processing Letters*, 2010, vol. 17, pp. 249–252.
29. **Malygin A., Seilova N., Boskebeev K., Alimseitova Zh.** Application of artificial neural networks forhandwritten biometric images recognition. *Computer Modelling and New Technologies*, 2017, vol. 21(1), pp. 31–38.
30. ГОСТ Р ИСО/МЭК ТО 19795-3-2009. Автоматическая идентификация. Идентификация биометрическая. Эксплуатационные испытания и протоколы испытаний в биометрии. Часть 3. Особенности проведения испытаний при различных биометрических модальностях. М., Стандартинформ, 2010. 28 с.
31. **Sulavko A. E., Zhumazhanova S. S., Fofanov G. A.** Perspective neural network algorithms for dynamic biometric pattern recognition in the space of interdependent features. *Proceedings of 2018th Conference "Dynamics of Systems, Mechanisms and Machines"*, Omsk, 2018, pp. 1–12.

UDC 004.93'1
doi:10.31799/1684-8853-2020-4-61-77

Highly reliable authentication based on handwritten passwords using hybrid neural networks with protection of biometric templates from being compromised

A. E. Sulavko[a], PhD, Tech., Associate Professor, orcid.org/0000-0002-9029-8028, sulavich@mail.ru
[a]Omsk State Technical University, 11, Mira Pr., 644050, Omsk, Russian Federation

Introduction: Biometrics-to-code converters based on neural networks are the ideological basis for a series of GOST R 52633 standards (unparalleled anywhere in the world) and can be used in the development of highly reliable biometric authentication and electronic signature with biometric activation. **Purpose:** Developing a model of a biometrics-to-code converter for highly reliable biometric authentication by handwritten passwords with high resistance to attacks on knowledge extraction. **Results:** We demonstrated the vulnerability of neural networks which makes it possible to perform quick directed enumeration of competing examples in order to compromise a biometric pattern and the personal key of its owner. We described a method of effective protection against this attack, and proposed a hybrid model for a biometrics-to-code converter based on a new type of hybrid neural networks, which does not compromise

the biometric standard and the user's key (password), being resistant to such attacks. The high reliability and effectiveness of the proposed model has been experimentally confirmed in handwritten password verification. The reliability indicators for generating a key from a handwritten password were: FRR = 11.5%, FAR = 0.0009% with a key length of 1024 bits (taking into account the presented fakes of a handwritten pattern). **Practical relevance:** The results can be used in information security applications or electronic document management.

Keywords — pattern recognition, Bayesian differential measures, correlated biometric features, information protection, quick tuning of neural networks, probability density, "wide" neural networks, biometrics-to-code converters, handwritten patterns.

For citation: Sulavko A. E. Highly reliable authentication based on handwritten passwords using hybrid neural networks with protection of biometric templates from being compromised. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 61–77 (In Russian). doi:10.31799/1684-8853-2020-4-61-77

References

1. Akhmetov B. S., Ivanov A. I., Alimseitova Z. K. Training of neural network biometry-code converters. *2018 News of the National Academy of Sciences of the Republic of Kazakhstan. Series of Geology and Technical Sciences*, 2018, vol. 1, no. 427, pp. 61–68.
2. Jain A. K., Nandakumar K., Nagar A. Biometric template security. *EURASIP J. Adv. Signal Process*, 2008, pp. 113:1–113:17.
3. Hafemann L. G., Sabourin R., Oliveira L. S. Characterizing and evaluating adversarial examples for offline handwritten signature verification. *IEEE Transactions on Information Forensics and Security*, 2019, vol. 14, iss. 8, pp. 2153–2166. doi:10.1109/TIFS.2019.2894031
4. Ivanov A. I., Lozhnikov P. S., Sulavko A. E. Evaluation of signature verification reliability based on artificial neural networks, Bayesian multivariate functional and quadratic forms. *Computer Optics*, 2017, no. 5, pp. 765–774. doi:10.18287/2412-6179-2017-41-5-765-774
5. ISO/IEC 19792:2009. Information technology — Security techniques — Security evaluation of biometrics. International Organization for Standardization, 2009. 37 p.
6. ISO/IEC 24761:2009. Information technology — Security techniques — Authentication context for biometrics. International Organization for Standardization, 2011. 50 p.
7. ISO/IEC 24745:2011. Information technology — Security techniques — Biometric information protection. International Organization for Standardization, 2011. 50 p.
8. State Standard 52633.0-2006. Data protection. Information security technique. High Reliability Biometric Authentication Requirements. Moscow, Standardinform Publ., 2007. 25 p. (In Russian).
9. State Standard 52633.1-2009. Data protection. Information security technique. Requirements for the formation of databases of natural biometric images intended for testing highly reliable biometric authentication. Moscow, Standardinform Publ., 2010. 24 p. (In Russian).
10. State Standard 52633.2-2010. Data protection. Information security technique. Requirements for the formation of synthetic biometric images intended for testing highly reliable biometric authentication tools. Moscow, Standardinform Publ., 2011. 22 p. (In Russian).
11. State Standard 52633.3-2011. Data protection. Information security technique. Testing the resistance of highly reliable biometric protection to selection attacks. Moscow, Standardinform Publ., 2012. 16 p. (In Russian).
12. State Standard 52633.4-2011. Data protection. Information security technique. Interfaces for interaction with neural network biometrics to code converters. Moscow, Standardinform Publ., 2012. 46 p. (In Russian).
13. State Standard 52633.5-2011. Data protection. Information security technique. Automatic training of neural network biometrics to code converters. Moscow, Standardinform Publ., 2012. 20 p. (In Russian).
14. State Standard 52633.6-2012. Data protection. Information security technique. Requirements for indicating the proximity of biometric data presented to the image of "Own". Moscow, Standardinform Publ., 2012. 24 p. (In Russian).
15. Dodis Y., Reyzin L., Smith A. Fuzzy extractors: How to generate strong keys from biometrics and other noisy. *EuroCrypt*, 2004, pp. 523–540.
16. Ivanov A., Somkin S., Andreev D., Malygina E. Diversity metrics to watch actual biometric data distribution statistics "fuzzy extractors" in their protection of a range. *UrFR Newsletter. Information Security*, 2014, no. 2(12), pp. 16–23 (In Russian).
17. Ignatenko T., Frans M. J. Willems. Information leakage in fuzzy commitment schemes. *IEEE Transactions on Information Forensics and Security*, 2010, vol. 5, no. 2, pp. 337–348. doi:10.1109/TIFS.2010.2046984
18. Lozhnikov P. S., Sulavko A. E., Eremenko A. V., Volkov D. A. Experimental evaluation of reliability of signature verification by quadratic form networks, fuzzy extractors and perceptrons. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 5, pp. 73–85 (In Russian). doi:10.15217/issn1684-8853.2016.5.73
19. Ahmetov B. S., Ivanov A. I., Funtikov V. A., Bezjaev A. V., Malygina E. A. *Tekhnologiia ispol'zovaniia bol'shikh neironnykh setei dlia preobrazovaniia nechetkikh biometricheskikh dannykh v kod kliucha dostupa* [Technology of using large neural networks for fuzzy transformation of biometric data in the access code key]. Almaty, LEM Publ., 2014. 144 p. (In Russian).
20. Kůrková V., Sanguineti M. Probabilistic lower bounds for approximation by shallow perceptron networks. *Neural Networks*, 2017, vol. 91, pp. 34–41.
21. Kůrková V., Sanguineti M. Model complexities of shallow networks representing highly varying functions. *Neurocomputing*, 2016, vol. 171, pp. 598–604.
22. Ivanov A. I. *Neirosetevaia zashchita konfidentsial'nykh biometricheskikh obrazov grazhdanina i ego lichnykh kriptograficheskikh kliuchei* [Neural protection of sensitive biometric images of the citizen and his personal cryptographic keys]. Penza, PNIEI Publ., 2014. 57 p. (In Russian).
23. Iranmanesh V., Ahmad S. M. S., Adnan W. A. W., Yussof S., Arigbabu O. A., Malallah F. L. Online handwritten signature verification using neural network classifier based on principal component analysis. *Scientific World Journal*, 2014, vol. 2014, pp. 1–8.
24. Iranmanesh V. Online signature template protection by shuffling and one time pad schemes with neural network verification. *Proceedings of the International Conference on Computer Science and Computational Mathematics (ICCSCM '13)*, 2013, pp. 53–59.
25. Hafemann L. G., Sabourin R., Oliveira L. S. Writer-independent feature learning for offline signature verification using deep convolutional neural networks. *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016. doi:10.1109/IJCNN.2016.7727521
26. Souza V. L. F., Oliveira A. L. I., Sabourin R. A writer-independent approach for offline signature verification using deep convolutional neural networks features. *2018 7th Brazilian Conference on Intelligent Systems (BRACIS)*, 2018. doi:10.1109/BRACIS.2018.00044
27. Díaz M., Fischer A., Ferrer M. A., Plamondon R. A perspective analysis of handwritten signature technology. *ACM Computing Surveys*, 2019, vol. 51, iss. 6, article 117, pp. 1–37.
28. Maiorana E., Campisi P. Fuzzy commitment for function based signature template protection. *IEEE Signal Processing Letters*, 2010, vol. 17, pp. 249–252.
29. Malygin A., Seilova N., Boskebeev K., Alimseitova Zh. Application of artificial neural networks forhandwritten biometric images recognition. *Computer Modelling and New Technologies*, 2017, vol. 21(1), pp. 31–38.
30. ISO/IEC TR 19795-3:2007. Information technology — Biometric performance testing and reporting — Part 3: Modality-specific testing. International Organization for Standardization Publ., 2007. 19 p. (In Russian).
31. Sulavko A. E., Zhumazhanova S. S., Fofanov G. A. Perspective neural network algorithms for dynamic biometric pattern recognition in the space of interdependent features. *Proceedings of 2018th Conference "Dynamics of Systems, Mechanisms and Machines"*, Omsk, 2018, pp. 1–12.

УДК 621.391
doi:10.31799/1684-8853-2020-4-78-86

Научные статьи
Articles

Аддитивная граница вероятности ошибки в дискретном канале передачи информации с помехоустойчивым кодированием и группированием ошибок

Г. Н. Мальцев[a], *доктор техн. наук, профессор, orcid.org/0000-0002-6755-5700*
В. В. Джумков[a], *канд. техн. наук, доцент, orcid.org/0000-0002-6385-7285, valentin32k@mail.ru*
[a]*Военно-космическая академия им. А. Ф. Можайского, Ждановская наб., 13, Санкт-Петербург, 197198, РФ*

Введение: *анализ показателей достоверности передачи информации при использовании помехоустойчивого кодирования в каналах с группированием ошибок, в частности в радиоканалах с помехами и замираниями принимаемых сигналов, затрудняется необходимостью использовать модели дискретных каналов передачи информации, учитывающие группирование ошибок и отличающиеся от традиционной биномиальной модели. Сложность аналитического описания таких моделей приводит к тому, что в практике анализа показателей качества передачи информации по каналам с группированием ошибок широкое распространение получает имитационное моделирование, а разработка аналитических моделей дискретных каналов передачи информации с группированием ошибок является одним из современных направлений развития теории помехоустойчивого кодирования.* ***Цель исследования:*** *определение аддитивной границы вероятности ошибки на бит информации для дискретного канала передачи информации с группированием символьных ошибок при его описании моделью Эллиота — Гильберта.* ***Результаты:*** *для случая передачи информации с использованием группового помехоустойчивого кода получены аналитические выражения для аддитивной границы вероятности ошибки на бит информации в дискретном канале передачи информации с группированием символьных ошибок. Полученные выражения учитывают особенности передачи информации по каналу с группирующимися символьными ошибками, в частности, отличие вероятностей различных сочетаний одинакового количества ошибок. Представлены примеры расчета вероятности ошибки на бит информации для случая использования помехоустойчивых кодов, исправляющих ошибки. Показано, что при любой длине кода использование модели Эллиота — Гильберта позволяет существенно уточнить результаты расчетов вероятности ошибочного приема сообщения в каналах с группированием символьных ошибок по сравнению с исходной биномиальной моделью. Полученные результаты сравниваются с результатами имитационного моделирования.* ***Практическая значимость:*** *полученные результаты могут быть использованы при проектировании и анализе характеристик систем передачи информации различного назначения, функционирующих в условиях группирования ошибок. Использование аналитических выражений для расчета вероятностных показателей достоверности передачи информации позволяет отказаться от сложного имитационного моделирования процесса передачи информации в каналах с группированием ошибок на этапе выбора помехоустойчивого кода и его параметров.*

Ключевые слова *— канал передачи информации, группирование символьных ошибок, модель Эллиота — Гильберта, вероятность ошибки на бит информации.*

Для цитирования: Мальцев Г. Н., Джумков В. В. Аддитивная граница вероятности ошибки в дискретном канале передачи информации с помехоустойчивым кодированием и группированием ошибок. *Информационно-управляющие системы*, 2020, № 4, с. 78–86. doi:10.31799/1684-8853-2020-4-78-86

For citation: Maltsev G. N., Dzhumkov V. V. Additive boundary of error probability in a discrete data transmission channel with noise-immune coding and grouping of errors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 78–86 (In Russian). doi:10.31799/1684-8853-2020-4-78-86

Введение

Помехоустойчивое кодирование широко используется в канальных протоколах передачи информации современных информационно-управляющих систем [1]. При проектировании систем передачи информации выбор помехоустойчивого кода является одним из наиболее важных вопросов, поскольку его корректирующими способностями в значительной степени определяются достижимые показатели качества передачи информации в условиях помех, прежде всего вероятностные показатели достоверности: вероятность ошибочного приема сообщения (кодового слова), вероятность ошибочного приема символа, вероятность ошибки на бит информации [2–5]. При этом на правильность выбора параметров помехоустойчивого кода, обеспечивающих достижение требуемого качества передачи информации, существенным образом влияет точность используемого метода расчета вероятностных показателей достоверности, который должен учитывать статистические свойства потока ошибок в ожидаемых условиях передачи информации.

Метод расчета вероятностных показателей достоверности передачи информации в каналах с помехами зависит от используемой модели канала передачи информации. Наиболее общим является метод расчета вероятностей ошибочного приема сообщения и ошибки на бит информации для биномиальной модели дискретного канала передачи информации, предполагающей незави-

симость символьных ошибок при посимвольном приеме кодовых слов. Этот стандартный метод может быть применен к любому блоковому или сверточному коду с посимвольным приемом сообщений [6], однако не учитывает группирования (пакетирования) символьных ошибок в дискретном канале передачи информации, что характерно, например, для радиоканалов передачи информации, функционирующих в условиях помех и замираний принимаемых сигналов [7].

Сложность аналитического описания моделей дискретных каналов передачи информации с группированием символьных ошибок приводит к тому, что в настоящее время в качестве основного способа анализа показателей качества передачи информации в таких каналах рассматривается имитационное моделирование [8, 9]. В настоящей работе на примере группового помехоустойчивого кода, используемого для исправления ошибок, получено аналитическое выражение для аддитивной границы вероятности ошибки на бит информации для каналов с группирующимися символьными ошибками при описании их статистических свойств моделью Эллиота — Гильберта [8, 10], которое в известных работах по теории помехоустойчивого кодирования не представлено. Аддитивный характер найденной границы обусловлен ее определением как суммы вероятностей возникновения различных векторов (сочетаний) символьных ошибок, приводящих к ошибке при декодировании сообщения.

Вывод выражения для аддитивной границы вероятности ошибки в дискретном канале передачи информации с группированием ошибок для группового помехоустойчивого кода и модели Эллиота — Гильберта

В общем случае выбор помехоустойчивого кода системы передачи информации осуществляется исходя из зависимостей, связывающих вероятностные показатели достоверности передачи информации с параметрами канала передачи информации и параметрами помехоустойчивого кода. Будем рассматривать влияние параметров канала передачи информации с группированием ошибок на достоверность передачи информации при использовании группового (линейного) помехоустойчивого кода. К этому классу помехоустойчивых кодов относятся такие блочные помехоустойчивые коды, как коды Хэмминга, коды Рида — Соломона, коды Рида — Маллера, коды Боуза — Чоудхури — Хоквингема (БЧХ-коды) и др. При этом полученные результаты могут быть распространены и на случаи использования других классов помехоустойчивых кодов.

С учетом известных свойств групповых кодов [6], в частности обязательного наличия в множестве кодовых слов нулевого кодового слова для определения вероятности ошибочного приема сообщения при использовании групповых кодов без потери общности рассуждений, может быть рассмотрен эффект передачи нулевого кодового слова. В этом случае число символов, в которых принятое кодовое слово отличается от переданного (нулевого), может быть заменено его весом — числом единиц в принятом кодовом слове. Тогда для биномиальной модели дискретного канала передачи информации с независимыми символьными ошибками вероятность ошибочного приема кодового слова для группового помехоустойчивого (n, k)-кода с исправлением ошибок может быть ограничена сверху неравенством

$$P_{\text{ош}} \le \sum_{j=\mu+1}^{n} n_j \Pr(B_j \mid A_0), \qquad (1)$$

где n — разрядность кодового слова; μ — кратность ошибок в кодовом слове, исправляемых рассматриваемым помехоустойчивым (n, k)-кодом; n_j — число кодовых слов веса j; $\Pr(B_j|A_0)$ — вероятность того, что решение принимается в пользу кодового слова B_j веса j, в то время как передавалось нулевое кодовое слово A_0, $j = 1, ..., n$.

Вероятности $\Pr(B_j|A_0)$, $j = 1, ..., n$, входящие в неравенство (1), рассчитываются для заданной вероятности ошибочного приема символа кодового слова p_0. Величина p_0 в большинстве случаев определяется для модели канала передачи информации с аддитивным белым гауссовым шумом, что позволяет использовать аналитические выражения, связывающие величину p_0 с отношением сигнал/шум в канале передачи информации [9, 10].

Кратность исправляемых ошибок μ связана с общим числом символов n и числом информационных символов в кодовом слове k известной границей Хэмминга [6, 10]

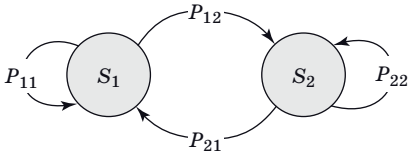
$$2^{n-k} \ge 1 + \sum_{i=1}^{\mu} C_n^k. \qquad (2)$$

Для безызбыточного кода, не обладающего корректирующими способностями, $n = k$ и $\mu = 0$. Достоинством использования в прикладных задачах анализа дискретных каналов передачи информации с помехоустойчивым кодированием границы Хэмминга является то, что она задает минимальное число $n - k$ поверочных символов, при котором существует корректирующий код, гарантированно исправляющий ошибки с кратностью μ. Поэтому если определяемый границей Хэмминга (2) помехоустойчивый (n, k)-код существует, то он гарантированно будет обе-

спечивать заданную кратность исправляемых или обнаруживаемых ошибок. В сравнении с границей Хэмминга другая известная граница Варшамова — Гильберта [6] показывает, при каком числе $n - k$ проверочных символов определенно существует код, исправляющий ошибки кратности μ. Поэтому при фиксированных значениях n и μ граница Хэмминга всегда дает большее число $n - k$ проверочных символов, чем граница Варшамова — Гильберта, и лишь в предельном случае эти граничные значения совпадают, но граница Варшамова — Гильберта при этом определяет лишь возможность существования кода, исправляющего ошибки кратностью μ.

Выражение (1) определяет исходную верхнюю границу вероятности ошибочного приема кодового слова в дискретном канале передачи информации при посимвольном приеме кодовых слов для случая независимых символьных ошибок. В то же время в случае группирования символьных ошибок значения вероятностных показателей достоверности передачи информации, полученные с использованием биномиальной модели дискретного канала передачи информации с независимыми ошибками, могут существенно отличаться от истинных и требуют уточнения. Учет группирования символьных ошибок в дискретном канале передачи информации осуществляется в рамках модели Эллиота — Гильберта [8, 11]. Данная модель описывает изменение состояния дискретного канала передачи информации между двумя состояниями — «хорошим» и «плохим». Ошибки происходят в этих состояниях с различными вероятностями, группируясь, когда дискретный канал находится в «плохом» состоянии, например, на интервалах замираний принимаемого сигнала. Данная модель широко используется при анализе помехоустойчивости систем передачи информации с группированием символьных ошибок и может быть развита на случай дифференцированного представления дискретного канала более чем двумя состояниями [12, 13].

Схема модели Эллиота — Гильберта для канала передачи информации с группированием ошибок представлена на рис. 1. Дискретный канал передачи информации в произвольный момент времени может находиться в одном из двух состояний — S_1 и S_2, в которых возможно возникновение ошибок при приеме символа кодового слова с вероятностями p_{01} и p_{02} соответственно. Переходы между состояниями происходят с вероятностями P_{xy}, где x, $y = \{1, 2\}$. Вероятности P_{xy} образуют матрицу переходных вероятностей размерностью 2×2. Нахождение дискретного канала передачи информации в состояниях S_1 и S_2 характеризуется средними длинами состояний — D_1 и D_2 соответственно. Средняя длина состояния равна среднему числу символов (бит) кодового слова, при последовательной передаче которых дискретный канал остается в данном состоянии. Вероятности переходов между состояниями P_{xy} выражаются через средние длины состояний: $P_{12} = 1/D_1$, $P_{11} = 1 - P_{12}$, $P_{21} = 1/D_2$, $P_{22} = 1 - P_{22}$.



■ ***Рис. 1***. Схема модели Эллиота — Гильберта для канала передачи информации с группированием ошибок

■ ***Fig. 1***. Elliot — Guilbert model diagram for a burst-noise channel

Условия передачи сообщений в модели Эллиота — Гильберта характеризует совокупность параметров p_{01}, p_{02}, D_1, D_2. Вероятности ошибочного приема символа кодового слова p_{01} и p_{02}, как и вероятность ошибочного приема символа кодового слова p_0 для биномиальной модели дискретного канала передачи информации с независимыми ошибками, в большинстве случаев определяются для модели канала передачи информации с аддитивным белым гауссовым шумом, но с различными уровнями шума в состояниях канала S_1 и S_2. Средние длины состояний D_1 и D_2 задаются с учетом временных характеристик изменения условий передачи информации, например замираний принимаемого сигнала, определяющих длину «пакета» ошибок.

Особенностью передачи информации в условиях группирования символьных ошибок является то обстоятельство, что вероятности различных сочетаний одинакового количества ошибок при приеме символов кодового слова в пределах кодового слова различны. В этом случае не может быть использовано весовое суммирование вероятностей различного числа символьных ошибок в кодовом слове, аналогичное выражению (1), и необходимо выполнять суммирование по всем возможным ошибкам при приеме кодового слова. Тогда вероятность ошибочного приема кодового слова для группового помехоустойчивого (n, k)-кода с исправлением ошибок определяется суммой вероятностей трансформации нулевого кодового слова в другие $2^k - 1$ разрешенных кодовых слова и ограничена сверху неравенством

$$P_{\text{ош}} \le \sum_{i=1}^{2^k-1} \Pr(X_i \mid A_0), \qquad (3)$$

где k — число разрядов кодового слова, в которых передаются информационные символы; $\Pr(X_i|A_0)$ — вероятность того, что решение примется в пользу кодового слова X_i, в то время как передавалось нулевое кодовое слово A_0, $i = 1$, ...,

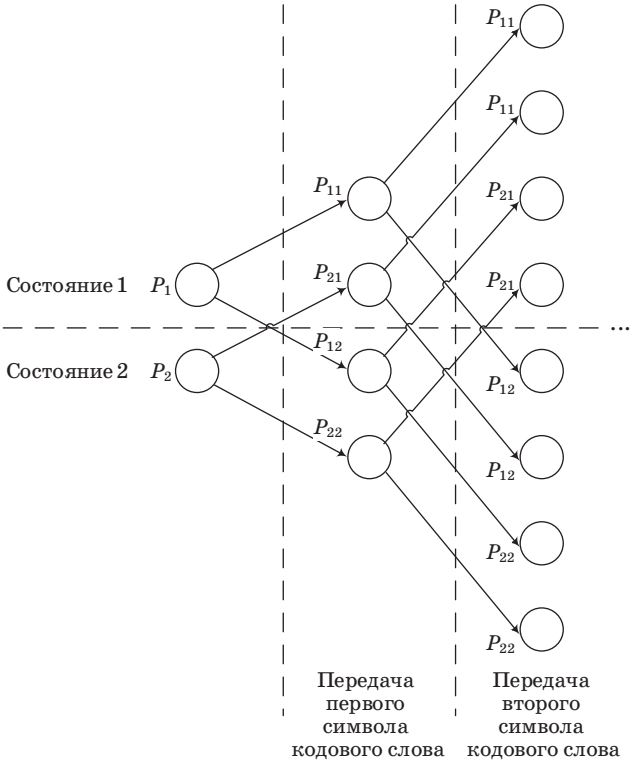
$2^k - 1$. Как и для неравенства (1), в случае безызбыточного кода, не обладающего корректирующими способностями, $n = k$ и $\mu = 0$.

Суммирование в выражении (3) осуществляется по $2^k - 1$ кодовым словам, которые совместно с нулевым кодовым словом являются разрешенными. Остальные $2^n - 2^k$ кодовых слов являются запрещенными и исправляются рассматриваемым корректирующим (n, k)-кодом с кратностью исправляемых ошибок μ. Практически допущение о том, что все ошибки кратности, меньшей и равной μ, исправляются, а все ошибки большей кратности приводят к ошибочному приему кодового слова, соответствует использованию к случаю так называемых совершенных (плотноупакованных) кодов, для которых неравенство для границы Хэмминга (2) превращается в равенство [6].

Вероятности $\Pr(X_i | A_0)$, $i = 1, ..., 2^k - 1$, входящие в неравенство (3), могут быть найдены как суммы вероятностей возникновения векторов ошибок $\mathbf{Z}_i$, которые приведут к трансформации нулевого кодового слова A_0 в разрешенное кодовое слово X_i. Под вектором ошибок n-разрядного кодового слова $\mathbf{Z}_i$ в общем случае понимается вектор размерности n, характеризующий определенное i-е сочетание символьных ошибок при приеме кодового слова. Каждый элемент вектора ошибок соответствует одному из разрядов кодового слова и принимает единичное значение в случае ошибки и нулевое значение при отсутствии ошибки в соответствующем разряде. Для описания вероятности возникновения определенного вектора ошибок $\Pr(\mathbf{Z}_i)$ в канале передачи информации с группированием ошибок необходимо рассмотреть передачу кодового слова по дискретному каналу, описываемому моделью Эллиота — Гильберта.

Граф возможных переходов между состояниями канала передачи информации с группированием ошибок при передаче следующих друг за другом двух символов кодового слова для модели Эллиота — Гильберта представлен на рис. 2. В момент начала передачи кодового слова дискретный канал может находиться в одном из двух состояний с вероятностями P_1 и P_2. В процессе передачи первого символа кодового слова дискретный канал может остаться в исходном состоянии либо перейти в другое состояние с вероятностями перехода P_{xy}, где $x, y = \{1, 2\}$. Это состояние становится исходным состоянием дискретного канала к моменту начала передачи второго символа кодового слова. В процессе передачи второго символа кодового слова дискретный канал также может остаться в исходном состоянии либо перейти в другое состояние с вероятностями перехода P_{xy} и т. д.

При принятых допущениях вероятность ошибочного приема первого символа кодового слова определяется выражением



■ ***Рис. 2.*** Граф переходов между состояниями дискретного канала передачи информации с группированием ошибок при передаче следующих друг за другом символов кодового слова

■ ***Fig. 2.*** A transition graph between burst-noise channel states when transmitting codeword symbols following each other

$$\Pr(\mathbf{Z}[1]) = p_{01}(P_1P_{11} + P_2P_{21}) + p_{02}(P_1P_{12} + P_2P_{22}) = \\ = [P_1 \;\; P_2] \times \mathbf{M} \times \begin{bmatrix} p_{01} \\ p_{02} \end{bmatrix}, \quad (4)$$

где P_1, P_2 — вероятности первого и второго состояний дискретного канала передачи информации соответственно; p_{01}, p_{02} — вероятности ошибочного приема символа кодового слова в первом и втором состояниях дискретного канала передачи информации соответственно; $\mathbf{M} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$ — матрица переходных вероятностей модели Эллиота — Гильберта. Вероятность правильного приема первого символа кодового слова определяется выражением

$$\Pr(\overline{\mathbf{Z}[1]}) = [P_1 \;\; P_2] \times \mathbf{M} \times \begin{bmatrix} q_{01} \\ q_{02} \end{bmatrix}, \quad (5)$$

где q_{01}, q_{02} — вероятности правильного приема символа кода в первом и втором состояниях дискретного канала соответственно ($q_{01} = 1 - p_{01}$, $q_{02} = 1 - p_{02}$).

В результате последовательного выполнения аналогичных преобразований могут быть получены выражения для вероятности ошибочного приема второго и последующих символов кодового слова. Для j-го символа кодового слова вероятность ошибочного приема определяется выражением

$$\Pr(\mathbf{Z}[j]) = \begin{bmatrix} P_1 & P_2 \end{bmatrix} \times \mathbf{M}^j \times \begin{bmatrix} p_{01} \\ p_{02} \end{bmatrix}, \quad (6)$$

а вероятность правильного приема определяется выражением

$$\Pr(\overline{\mathbf{Z}[j]}) = \begin{bmatrix} P_1 & P_2 \end{bmatrix} \times \mathbf{M}^j \times \begin{bmatrix} q_{01} \\ q_{02} \end{bmatrix}. \quad (7)$$

В результате обобщения выражений (4)–(7) может быть составлено выражение для расчета вероятности возникновения определенного вектора ошибок $\mathbf{Z}_i$

$$\Pr(\mathbf{Z}_i) = \begin{bmatrix} P_1 & P_2 \end{bmatrix} \times \left(\prod_{j=1}^{n} \mathbf{M} \times \right.$$
$$\left. \times \left(\begin{bmatrix} p_{01} & 0 \\ 0 & p_{02} \end{bmatrix} \times \mathbf{Z}[j] + \begin{bmatrix} q_{01} & 0 \\ 0 & q_{02} \end{bmatrix} \times \overline{\mathbf{Z}[j]} \right) \right) \times \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad (8)$$

где $\mathbf{Z}[j] = 1$, $\overline{\mathbf{Z}[j]} = 0$ в случае наличия ошибки в j-м символе кодового слова и $\mathbf{Z}[j] = 0$, $\overline{\mathbf{Z}[j]} = 1$ в случае отсутствия ошибки в j-м символе кодового слова. Сочетание значений $\mathbf{Z}[j]$, $j = 1, \ldots, n$ определяет вектор ошибок $\mathbf{Z}_i$ размерности n, характеризующий определенное i-е сочетание ошибок при приеме кодового слова.

Используя выражение (8), можно определить верхнюю границу (3) для вероятности ошибочного приема сообщения $P_{\text{ош}}$ как сумму вероятностей возникновения векторов ошибок $\Pr(\mathbf{Z}_i)$, при которых выделенное при декодировании кодовое слово X_i будет отличаться от нулевого кодового слова A_0:

$$P_{\text{ош}} \le \sum_{i=1}^{2^k-1} \Pr(\mathbf{Z}_i). \quad (9)$$

Вероятность ошибочного приема сообщения $P_{\text{ош}}$ является исходной для определения вероятности ошибки на бит информации P_b. В общем случае группового помехоустойчивого (n, k)-кода с исправлением ошибок переход от верхней границы вероятности ошибочного приема сообщения $P_{\text{ош}}$, определяемой выражением (9), к вероятности ошибки на бит информации выполняется в соответствии с выражением

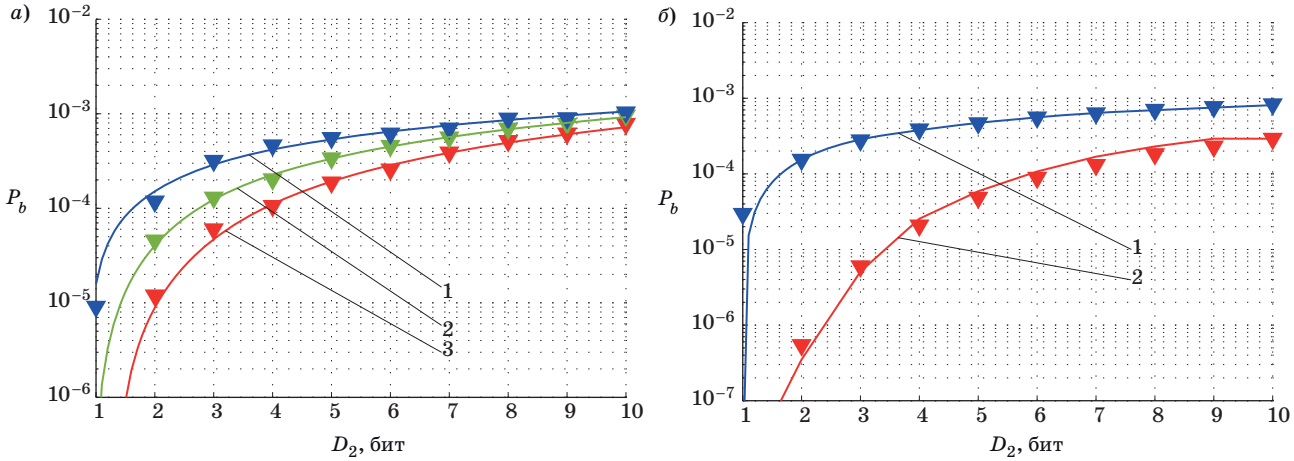
$$P_b \approx P_{\text{ош}}/k. \quad (10)$$

Использование в качестве показателя достоверности вероятности ошибки на бит информации P_b обусловлено необходимостью сравнивать коды различной длины n и k. Входящая в выражение (10) вероятность ошибочного приема сообщения $P_{\text{ош}}$ учитывает результат исправления ошибок помехоустойчивым (n, k)-кодом в пределах его корректирующих возможностей, а деление на число информационных символов в кодовом слове k позволяет отнести показатель достоверности к единице передаваемой информации.

Результаты расчетов аддитивной границы вероятности ошибки в дискретном канале передачи информации с группированием ошибок при использовании помехоустойчивых кодов

В соответствии с выражениями (9) и (10) были проведены расчеты аддитивной границы вероятности ошибки на бит информации P_b в канале с группированием ошибок для групповых помехоустойчивых кодов. Кроме того, для рассмотренных условий передачи информации и групповых помехоустойчивых кодов было проведено имитационное моделирование процесса передачи информации в каналах с группированием символьных ошибок. Результаты расчетов и имитационного моделирования представлены на рис. 3, *а* и *б* в виде зависимостей вероятности ошибки на бит информации P_b от средней длины «плохого» состояния канала D_2 при фиксированных остальных параметрах модели Эллиота — Гильберта: средняя длина «хорошего» состояния канала $D_1 = 1000$ бит, вероятности ошибочного приема символа кодового слова в «хорошем» и «плохом» состояниях канала $p_{01} = 10^{-3}$ и $p_{02} = 0{,}3$ соответственно.

На рисунке 3, *а* приведены зависимости вероятности ошибки на бит информации P_b от средней длины «плохого» состояния канала D_2 для простейших коротких помехоустойчивых кодов (15, 5), (15, 7), (15, 11). На рис. 3, *б* приведены зависимости вероятности ошибки на бит информации P_b от средней длины «плохого» состояния канала D_2 для более длинных помехоустойчивых кодов (31, 26) и (31, 11). Выбранные для анализа (15, 11) и (31, 26) являются кодами Хэмминга, а коды (15, 5), (15, 7) и (31, 11) являются БЧХ-кодами, их корректирующие способности удовлетворяют границе Хэмминга (2). На обоих рисунках зависимости, полученные аналитически, обозначены сплошными линиями, а треугольными маркерами показаны результаты имитационного моделирования. Имитационное моделирование проводилось по методике, изложенной в работах [11, 14]. Сходимость результатов имитационного моделирования и ана-



■ ***Рис. 3.*** Зависимости вероятности ошибки на бит информации P_b от средней длины «плохого» состояния канала D_2 для помехоустойчивых кодов: *а —1* — код Хэмминга (15, 11); *2* — БЧХ-код (15, 7); *3* — БЧХ-код (15, 5); *б — 1* — код Хэмминга (31, 26); *2* — БЧХ-код (31, 11)

■ ***Fig. 3.*** Bit error probability P_b dependence on channel "bad" state average length D_2 for different noise-immune code: *а —1* — Hamming code (15, 11); *2* — BCH-code (15, 7); *3* — BCH-code (15, 5); *б — 1* — Hamming code (31, 26); *2* — BCH-code (31, 11)

литических расчетов позволяет сделать вывод об адекватности полученных выражений.

Приведенные зависимости показывают, что даже небольшие «пакеты» символьных ошибок со средней длиной «плохого» состояния канала в единицы бит при средней длине «хорошего» состояния канала 1000 бит приводят к заметному увеличению вероятности ошибки на бит передаваемой информации. При этом повышение корректирующей способности кода для канала с группирующимися ошибками дает существенно меньший выигрыш в достоверности передачи информации, чем для канала с независимыми ошибками. Так, для дискретного канала передачи информации со средней длиной «плохого» состояния канала $D_2 = 2$ и средней длиной «хорошего» состояния канала $D_1 = 1000$ использование БЧХ-кода (15, 5), исправляющего три ошибки, позволяет снизить вероятность ошибки на бит информации P_b на порядок по сравнению с кодом Хэмминга (15, 11), исправляющим одну ошибку. В то же время для дискретного канала передачи информации со средней длиной «плохого» состояния канала $D_2 = 10$ и тех же кодов вероятность ошибки на бит информации P_b снижается всего в 1,25 раза.

Если бы для расчета вероятности ошибочного приема сообщения $P_{ош}$ и связанной с ней вероятности ошибки на бит информации P_b использовалась биномиальная модель дискретного канала передачи информации и верхняя граница (1), то полученные результаты соответствовали бы $D_2 = 1$ и значения рассматриваемых вероятностных показателей достоверности передачи информации для каналов с группированием символьных ошибок были бы завышены. Представленные выражения определяют методику расчета верхней границы (3), позволяющую получить на основе модели Эллиота — Гильберта более точные оценки вероятности ошибочного приема сообщения и вероятности ошибки на бит информации в каналах с группирующимися ошибками в сравнении с известными приближенными методиками [15–17]. Из представленных результатов видно, что при любой длине кода использование модели Эллиота — Гильберта позволяет существенно уточнить результаты расчетов вероятностных показателей достоверности передачи информации в каналах с группированием символьных ошибок по сравнению с исходной биномиальной моделью.

К недостаткам рассмотренного способа следует отнести высокую вычислительную сложность при его применении, которая существенно увеличивается при увеличении длины кода. Этим объясняется то обстоятельство, что представленные результаты расчетов и имитационного моделирования относятся к сравнительно коротким помехоустойчивым кодам. В то же время анализ показателей достоверности передачи информации в каналах с группированием ошибок методом имитационного моделирования имеет еще большую вычислительную сложность, что приводит к задаче оптимизации точности вычислений при ограниченных вычислительных ресурсах [8].

Заключение

Полученные результаты показывают важность адекватного описания условий передачи информации в каналах с группирующимися ошибками.

Полученные аналитические выражения позволяют выполнять расчеты показателей достоверности передачи информации в каналах передачи информации с помехоустойчивым кодированием и группированием символьных ошибок. В настоящее время в качестве основного способа анализа показателей качества передачи информации в таких каналах рассматривается имитационное моделирование. Рассмотренная методика описания характеристик дискретного канала передачи информации с группированием символьных ошибок и расчета аддитивной границы вероятности ошибки на бит информации имеет достаточно высокую вычислительную сложность, однако хорошее совпадение результатов аналитических расчетов и имитационного моделирования позволяет отказаться от имитационного моделирования на этапе проектирования систем передачи информации, функционирующих в условиях группирования ошибок. При этом обеспечивается более точное оценивание вероятностных показателей достоверности передачи информации в сравнении с более простыми методиками расчета, дающими, как показано в работе, существенную погрешность оценивания вероятности ошибки на бит информации.

Представленные результаты могут быть использованы при проектировании и анализе характеристик радиотехнических систем передачи информации различного назначения, функционирующих в условиях группирования ошибок, которое может быть связано как с условиями распространения сигналов в радиоканале передачи информации, так и с воздействием различного вида помех.

Литература

1. **Семенов Ю. А.** *Алгоритмы телекоммуникационных сетей. Часть 1. Алгоритмы и протоколы каналов и сетей передачи данных*. М., Бином. Лаборатория знаний, 2016. 637 с.
2. **Зубарев А. Е., Позов А. В., Приходько А. И.** Анализ методов расчета битовой вероятности ошибки при когерентном приеме сигналов с *M*-ичной фазовой манипуляцией. *Международный научно-исследовательский журнал*, 2019, № 1(79), с. 53–59. doi:10.23670/IRJ.2019.79.1.009
3. **Владимиров С. С.** Сравнение вероятностных характеристик 8-разрядных кодов с прямой коррекцией ошибок. *Информационные технологии и телекоммуникации*, 2019, т. 7, № 1, с. 21–30. doi:10.31854/2307-1303-2019-7-1-21-30. http://itt.sut.ru (дата обращения: 29.03.2020).
4. **Струков А. П.** Метод аналитического расчета вероятности символьной и битовой ошибок сигнала с амплитудно-фазовой манипуляцией в нелинейном канале. *Ракетно-космическое приборостроение и информационные системы*, 2017, т. 4, вып. 4, с. 83–88. doi:10.17238/issn2409-0239.2017.4.83
5. **Чиров Д. С., Лобов Е. М.** Выбор сигнально-кодовой конструкции для командно-телеметрической линии радиосвязи с беспилотным летательным аппаратом средней и большой дальности. *T-Comm: Телекоммуникации и транспорт*, 2017, т. 11, № 10, с. 21–28.
6. **Clark G. C. Jr., Cain J. B.** *Error-Correction Coding for Digital Communications*. Springer Science & Business Media, 2013. 422 p.
7. **Шевченко В. А., Снедков Д. М.** Критическая скорость кодирования для некогерентных каналов связи с группированием ошибок, вызванным замираниями и воздействием импульсной помехи. *Известия института инженерной физики*, 2018, № 1(47), с. 39–46.
8. **Мелентьев О. Г.** *Теоретические аспекты передачи данных по каналам с группирующимися ошибками*. М., Горячая линия–Телеком, 2007. 232 с.
9. **Шевченко В. А., Пашинцев В. П.** Метод расчета вероятности ошибки на бит в каналах связи с блочными замираниями для приемника с линейным сложением мягких решений некогерентного демодулятора. *Инфокоммуникационные технологии*, 2019, т. 17, № 4, с. 372–382. doi:10.18469/ikt.2019.17.4.03
10. **Нурматов А. Т., Селихов Ю. Р., Нурматова Е. В.** Проектирование систем оценки состояния нестационарного дискретного канала связи. *Вестник компьютерных и информационных технологий*, 2017, № 4(154), с. 29–38. doi:10.14489/VKIT.2017.04.PP.029-038
11. **Афанасьев В. Б., Давыдов А. А., Зигангиров Д. К.** Оценка доли стираний, исправляемых линейными кодами. *Информационные процессы*, 2016, т. 16, № 4, с. 382–404.
12. **Trofimov A. N.** Random coding bound for channels with memory — decoding function with partial overlapping. Part 1. Derivation of main expression. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 79–88. doi:10.15217/issn1684-8853.2018.3.79
13. **Мальцев Г. Н., Джумков В. В.** Обобщенная модель дискретного канала передачи информации с группированием ошибок. *Информационно-управляющие системы*, 2013, № 1, с. 27–33.
14. **Кузнецов В. С., Волков А. С. Быков А. В.** Разработка и моделирование метода декодирования помехоустойчивого блокового кода с применением второго алгоритма Чейза. *Наукоемкие технологии в космических исследованиях Земли*, 2018, т. 10, № 5, с. 46–55. doi:10.24411/2409-5419-2018-10165
15. **Bildea A., Alphand O., Rousseau F., Duda A.** Link quality estimation with the Gilbert—Elliot model for wireless sensor networks. *IEEE 26th Annual Interna-*

tional Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Hong Kong, 2015, pp. 1–6.
16. **Nielsen J. J., Leyva-Mayorga I., Popovski P.** Reliability and error burst length analysis of wireless multi-connectivity. *2019 IEEE 16th International Symposium on Wireless Communication Systems (ISWCS)*, Oulu, 2019, pp. 107–111.
17. **da Silva C. A. G., Pedroso C. M.** MAC-layer packet loss models for Wi-Fi networks: Survey. *IEEE Access*, 2019, vol. 7, pp. 180512–180531. doi:10.1109/ACCESS.2019.2958260

UDC 621.391
doi:10.31799/1684-8853-2020-4-78-86

Additive boundary of error probability in a discrete data transmission channel with noise-immune coding and grouping of errors

G. N. Maltsev[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-6755-5700
V. V. Dzhumkov[a], PhD, Tech., Associate Professor, orcid.org/0000-0002-6385-7285, valentin32k@mail.ru
[a]A. F. Mozhaiskiy Military Space Academy, 13, Zhdanovskaia Emb., 197198, Saint-Petersburg, Russian Federation

Introduction: Data transmission reliability analysis when using noise-immune coding in channels with grouping of errors (in particular, in radio channels with interference and fading of the received signals) is complicated by the need to use discrete data transmission channel models which take into account the error grouping, differing from the traditional binomial model. The complexity of the analytical description of such models leads to the fact that the quality indicators of data transmission over channels with error grouping are usually analyzed by simulation methods, and the development of analytical models of data transmission discrete channels with grouping of errors is one of the modern direction in the noise-immune coding theory development. **Purpose:** Finding the additive boundary of a bit error probability for data transmission discrete channel with grouping of symbol errors, described by Elliot — Hilbert model. **Results:** For the case of data transmission using a group noise-immune code, analytical expressions are obtained for calculating the additive boundary of a bit error probability in a discrete data transmission channel with grouping of symbol errors. The obtained expressions take into account the features of data transmission over a channel with error grouping, in particular, the fact that the probabilities of various combinations of the same number of errors are not equal to each other. Examples are presented of calculating a bit error probability for the case of using noise-immune codes which correct errors. It is shown that for any code length, the use of the Elliot — Hilbert model allows you to substantially refine the results of calculating the probabilistic indicators of the reliability of data transmission in channels with error grouping, as compared to the original binomial model. The obtained results are compared to the results of the simulation. **Practical relevance:** The results can be used in the design and analysis of the characteristics of data transmission systems for various purposes, operating under conditions of error grouping. Using analytical expressions to calculate the probability indicators of the reliability of data transfer allows you to abandon complex simulation modeling of transmitting data in channels with error grouping at the stage of choosing a noise-immune code and its parameters.
Keywords — data transmission channel, grouping of symbol errors, Elliot — Hilbert model, bit error probability.

For citation: Maltsev G. N., Dzhumkov V. V. Additive boundary of error probability in a discrete data transmission channel with noise-immune coding and grouping of errors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 78–86 (In Russian). doi:10.31799/1684-8853-2020-4-78-86

References

1. Semenov Yu. A. *Algoritmy telekommunikatsionnykh setey. Chast' 1. Algoritmy i protokoly kanalov i setey peredachi dannykh* [Telecommunication network algorithms. Part 1. Algorithms and protocols of channels and data transmission networks]. Moscow, Binom. Laboratoriya znaniy Publ., 2016. 637 p. (In Russian).
2. Zubarev A. E., Pozov A. V., Prikhodko A. I. Calculating method analysis of bit probability of error at cogerent reception of signals with *M*-ary phase manipulation. *Research Journal of International Studies*, 2019, no. 1(79), pp. 53–59 (In Russian). doi:10.23670/IRJ.2019.79.1.009
3. Vladimirov S. Comparison of the probabilistic characteristics of 8-bit codes with forward error correction. *Informacionnye tekhnologii i telekommunikacii*, 2019, vol. 7, no. 1, pp. 21–30. doi:10.31854/2307-1303-2019-7-1-21-30. Available at: http://itt.sut.ru (accessed 29 March 2020) (In Russian).
4. Strukov A. P. A method of analytical calculation of SER and BER for APSK modulation in the nonlinear channel with AWGN. *Rocket-space device engineering and information systems*, 2017, vol. 4, iss. 4, pp. 83–88 (In Russian). doi:10.17238/issn2409-0239.2017.4.83
5. Chirov D. S., Lobov E. M. Choice of signal-code construcure for the command-telemetry radio communication line with medium and long range unmanned aerial vehicles. *T-Comm*, 2017, vol. 11, no. 10, pp. 21–28 (In Russian).
6. Clark G. C. Jr., Cain J. B. *Error-Correction Coding for Digital Communications*. Springer Science & Business Media, 2013. 422 p.
7. Shevchenko V. A., Snedkov D. M. Critical coding rate for incoherent communication channels with error grouping caused by fading and impulse noise interference. *Izvestiya instituta inzhenernoy fiziki*, 2018, no. 1(47), pp. 39–46 (In Russian).
8. Melent'yev O. G. *Teoreticheskiye aspekty peredachi dannykh po kanalam s gruppiruyushchimisya oshibkami* [Theoretical aspects of data transmission on channels with grouping errors]. Moscow, Goryachaya liniya–Telekom Publ., 2007, 232 p. (In Russian).
9. Shevchenko V. A., Pashintsev V. P. The method of calculating the error probability per bit in the communication channels with block freeze-ups for the receiver with linear addition of soft solutions of the incoherent demodulator. *Infokommunikacionnye tehnologii*, 2019, vol. 17, no. 4, pp. 372–382. doi:10.18469/ikt.2019.17.4.03 (In Russian).
10. Nurmatov A. T., Selikhov Yu. R., Nurmatova E. V. Non-stationary discrete communication channel state assessment system design. *Herald Computer and Information Technologies*, 2017, no. 4(154), pp. 29–38 (In Russian). doi:10.14489/VKIT.2017.04.PP.029-038
11. Afanas'yev V. B., Davydov A. A., Zigangirov D. K. Estimation of the proportion of erasures corrected by linear codes.

Informatsionnyye protsessy, 2016, vol. 16, no. 4, pp. 382–404 (In Russian).
12. Trofimov A. N. Random coding bound for channels with memory — decoding function with partial overlapping. Part 1. Derivation of main expression. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 79–88. doi:10.15217/issn1684-8853.2018. 3.79
13. Maltsev G. N., Dzhumkov V. V. Generalized model of a discrete communication channel in conditions of burst errors. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2013, no. 1, pp. 27–33 (In Russian).
14. Kuznetsov V. S., Volkov A. S., Bykov A. V. Development and modeling of decoding method of error correction block code using the second Chase algorithm. *High Technologies in Earth Space Research*, 2018, vol. 10, no. 5, pp. 46–55 (In Russian). doi:10.24411/2409-5419-2018-10165
15. Bildea A., Alphand O., Rousseau F., Duda A. Link quality estimation with the Gilbert—Elliot model for wireless sensor networks. *IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, 2015, pp. 1–6.
16. Nielsen J. J., Leyva-Mayorga I., Popovski P. Reliability and error burst length analysis of wireless multi-connectivity. *2019 IEEE 16th International Symposium on Wireless Communication Systems (ISWCS)*, Oulu, 2019, pp. 107–111.
17. da Silva C. A. G., Pedroso C. M. MAC-layer packet loss models for Wi-Fi networks: Survey. *IEEE Access*, 2019, vol. 7, pp. 180512–180531. doi:10.1109/ACCESS.2019.2958260

БАЛОНИН Николай Алексеевич

Профессор кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1982 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Автоматика и телемеханика».
В 2008 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 100 научных публикаций, в том числе трех монографий.
Область научных интересов — теория динамических систем, теория идентификации, теория операторов, теория матриц, вычислительные методы, интернет-робототехника, интернет-книги с исполняемыми алгоритмами, научные социальные сети.
Эл. адрес: korbendfs@mail.ru

ДВОЙНИКОВА Анастасия Александровна

Программист лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН .
В 2018 году окончила Университет ИТМО по специальности «Информационная безопасность».
Является автором восьми научных публикаций.
Область научных интересов — сентимент-анализ, компьютерная паралингвистика.
Эл. адрес:
dvoynikova.a@iias.spb.su

ДИКИЙ Дмитрий Игоревич

Аспирант факультета безопасных информационных технологий Университета ИТМО, Санкт-Петербург.
В 2015 году окончил бакалавриат Балтийского федерального университета, в 2017 году — магистратуру Университета ИТМО по специальности «Информационная безопасность».
Является автором более 20 научных публикаций и пяти свидетельств о регистрации программ для ЭВМ.
Область научных интересов — информационная безопасность, интернет вещей, машинное обучение, криптографические системы с открытым ключом.
Эл. адрес: dimandikiy@mail.ru

ДЖОКОВИЧ Драгомир

Почетный профессор кафедры теоретической математики Университета Ватерлоо, Ватерлоо, Онтарио, Канада.
В 1960 году окончил Белградский университет по специальности «Электротехника», Белград, Югославия.
В 1963 году защитил диссертацию на соискание ученой степени доктора наук в Белградском университете.
Является автором более 200 научных публикаций.
Область научных интересов — линейная и полилинейная алгебра, теория групп, алгебра Ли и групп Ли, квантовая запутанность, комбинаторика.
Эл. адрес: djokovic@uwaterloo.ca

ДЖУМКОВ Валентин Валентинович

Доцент кафедры космических радиотехнических систем Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.
В 2008 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Радиоэлектронные системы».
В 2014 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 20 научных публикаций.
Область научных интересов — помехоустойчивость радиоканалов управления и информационного обмена с космическими аппаратами, совершенствование методов и алгоритмов информационного обмена с космическими аппаратами, выбор параметров канальных протоколов.
Эл. адрес:
valentin32k@mail.ru

КАРПОВ Алексей Анатольевич

Главный научный сотрудник, руководитель лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН.
В 2002 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети».
В 2013 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 300 научных публикаций, включая три монографии и три патента.
Область научных интересов — многомодальные интерфейсы и системы, речевые технологии, автоматическое распознавание и синтез речи, компьютерная паралингвистика.
Эл. адрес: karpov@iias.spb.su

КИСЕЛЕВ Николай Константинович

Первый заместитель главного конструктора АО ЦКБ «Лазурит», Нижний Новгород.
В 2000 году окончил Нижегородский государственный университет им. Н. И. Лобачевского по специальности «Радиофизика и электроника».
Область научных интересов — моделеориентированная системная инженерия, системы управления автономными подводными средствами.
Эл. адрес: kiselevu@gmail.com

МАЛЬЦЕВ Георгий Николаевич

Профессор кафедры космических радиотехнических систем Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург, заслуженный деятель науки РФ, действительный член Академии космонавтики им. К. Э. Циолковского.
В 1980 году окончил Военный инженерный Краснознаменный институт им. А. Ф. Можайского по специальности «Радиотехнические системы комплексов».
В 1994 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 300 научных публикаций и 27 патентов на изобретения.
Область научных интересов — обработка сигналов в радиотехнических и оптико-электронных информационных системах и др.
Эл. адрес:
georgy_maltsev@mail.ru

МЫСЛИВЫЙ Александр Александрович

Научный сотрудник Научно-исследовательского института оперативно-стратегических исследований строительства ВМФ, Санкт-Петербург.
В 2008 году окончил Санкт-Петербургский военно-морской институт по специальности «Инженер-гидрограф».
В 2016 году защитил диссертацию на соискание ученой степени кандидата военных наук.
Является автором 20 научных публикаций.
Область научных интересов — морская робототехника, модели применения, навигация, системный анализ, исследование операций.
Эл. адрес: aam-07@mail.ru

КУЛАБУХОВА Светлана Олеговна

Магистрант кафедры теоретической и прикладной информатики Новосибирского государственного технического университета.
В 2018 году окончила бакалавриат Новосибирского государственного технического университета по специальности «Прикладная математика и информатика».
Является автором восьми научных публикаций.
Область научных интересов — параметрическая идентификация динамических систем.
Эл. адрес:
kulabuhova.s@gmail.com

МАРТЫНОВА Любовь Александровна

Старший научный сотрудник, ведущий научный сотрудник научно-исследовательского центра «Системы освещения обстановки» АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург.
В 1985 году окончила Ленинградский кораблестроительный институт по специальности «Прикладная математика».
В 2013 году защитила диссертацию на соискание ученой степени доктора технических наук.
Является автором 90 научных публикаций.
Область научных интересов — математическое моделирование, системный анализ, обработка информации, управление сложными системами.
Эл. адрес:
martynowa999@bk.ru

РОЖДЕСТВЕНСКАЯ Ксения Николаевна

Ассистент кафедры аэрокосмических компьютерных и программных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2012 году окончила магистратуру Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Встроенные системы обработки информации и управления».
Является автором 24 научных публикаций и пяти свидетельств о государственной регистрации программ на ЭВМ.
Область научных интересов — вычислительные сети и системы, администрирование, SpaceWire, SpaceWire-Plug-and-Play.
Эл. адрес: rogdkn@yandex.ru

СУЛАВКО
Алексей
Евгеньевич

Доцент кафедры комплексной защиты информации Омского государственного технического университета.
В 2009 году окончил Сибирскую государственную автомобильно-дорожную академию по специальности «Комплексное обеспечение информационной безопасности автоматизированных систем».
В 2014 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 130 научных публикаций и одного патента на изобретение.
Область научных интересов — распознавание образов, машинное обучение, биометрия, искусственный интеллект, защита информации, искусственные нейронные сети.
Эл. адрес: sulavich@mail.ru

ЧУБИЧ
Владимир
Михайлович

Профессор, заведующий кафедрой теоретической и прикладной информатики Новосибирского государственного технического университета.
В 1984 году окончил с отличием Новосибирский электротехнический институт по специальности «Прикладная математика».
В 2014 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 100 научных и учебно-методических публикаций, включая две монографии.
Область научных интересов — активная параметрическая идентификация стохастических динамических систем на основе планирования эксперимента.
Эл. адрес: chubich@ami.nstu.ru

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.

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

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

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

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, ORCID и электронный адрес одного из авторов. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисуночные подписи и названия таблиц на русском и английском языках.

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

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

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

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

Иллюстрации предоставляются отдельными исходными файлами, поддающимися редактированию:

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

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

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

Наличие подрисуночных подписей 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 «Правила 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: ius.spb@gmail.com
Сайт: www.i-us.ru