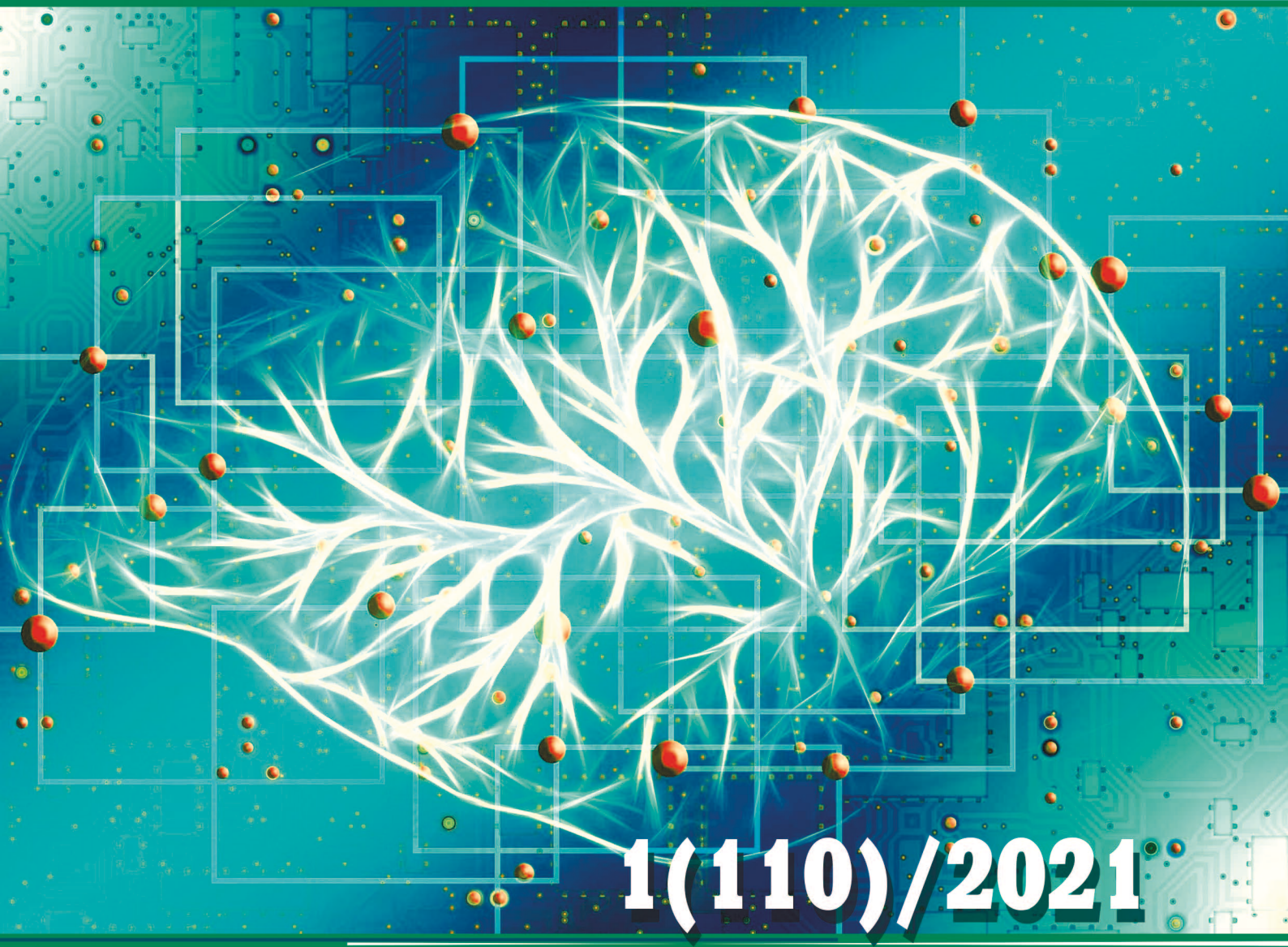
ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ



1(110)/2021

doi:10.31799/1684-8853-2021-1 ISSN 1684-8853 (print); ISSN 2541-8610 (online)

SCIENTIFIC JOURNAL

1(110)/2021

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 11.01.21. Passed for printing 24.02.21. 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:10.31799/1684-8853-2021-1 ISSN 1684-8853 (print); ISSN 2541-8610 (online)

НАУЧНЫЙ ЖУРНАЛ

1(110)/2021

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Учредитель
ООО «Информационно-управляющие системы»
Издатель
Санкт-Петербургский государственный университет аэрокосмического приборостроения

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

Сдано в набор 11.01.21. Подписано в печать 24.02.21. Формат 60×84 1/8.
Гарнитура SchoolBookC. Печать цифровая.
Усл. печ. л. 8,0. Уч.-изд. л. 11,2. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 46.
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:10.31799/1684-8853-2021-1-2-7

Articles

Three new lengths for cyclic Legendre pairs

N. A. Balonin[a], *Dr. Sc., Tech., Professor, orcid.org/0000-0001-7338-4920, korbendfs@mail.ru*
D. Ž. Đoković[b], *Dr. Sc., Tech., 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: *It is conjectured that the cyclic Legendre pairs of odd lengths > 1 always exist. Such a pair consists of two functions a, b: G → Z, whose values are +1 or −1, and whose periodic autocorrelation functions add up to the constant value −2 (except at the origin). Here G is a finite cyclic group and Z is the ring of integers. These conditions are fundamental and the closely related structure of Hadamard matrices having a two circulant core and double border is incompletely described in literature, which makes its study especially relevant.* ***Purpose:*** *To describe the two-border two-circulant-core construction for Legendre pairs having three new lengths.* ***Results:*** *To construct new Legendre pairs we use the subsets X = {x ∈ G: a(x) = −1} and Y = {x ∈ G: b(x) = −1} of G. There are 20 odd integers v less than 200 for which the existence of Legendre pairs of length v is undecided. The smallest among them is v = 77. We have constructed Legendre pairs of lengths 91, 93 and 123 reducing thereby the number of undecided cases to 17. Some new examples of cyclic Legendre pairs for lengths v ≤ 123 are given.* ***Practical relevance:*** *Hadamard matrices are used extensively in the problems of error-free coding, and compression and masking of video information. Programs for search of Hadamard matrices and a library of constructed matrices are used in the mathematical network "mathscinet.ru" together with executable on-line algorithms.*

Keywords *— Hadamard matrices, periodic autocorrelation functions, Legendre pairs, cyclic matrices, double-border two-circulant-core construction.*

For citation: Balonin N. A., Đoković D. Ž. Three new lengths for cyclic Legendre pairs. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 2–7. doi:10.31799/1684-8853-2021-1-2-7

Introduction

It is conjectured that the cyclic Legendre pairs of odd lengths > 1 always exist. See the next section for the definition of the Legendre pairs and Legendre difference families (DF). It is known that they exist for odd lengths v in the range $2 < v < 76$. The smallest unresolved case is $v = 77$. According to [1], there are four series of known cyclic Legendre pairs of odd length $v > 1$ (the first three are infinite):

(i) v is a prime number;

(ii) $2v + 1$ is a power of a prime number;

(iii) $v + 1$ is a power of 2;

(iv) $v = pq$, where p and q are prime numbers and $q - p = 2$.

We refer to (i) as the *classical series* because the construction is based on the sequence of the classical Legendre symbols [1]. The case (ii) is the *Szekeres series* provided by the well known series of so called Szekeres difference sets (in fact they are difference families) [2, 3]. The series (iii) is known as the *Galois series* [4] and (iv) is the *twin-prime series*, see e. g. [5, Theorem 9.4].

The series (iii) and (iv) as well as (i) for $v \equiv 3 \pmod 4$ are obtained from the three well known series of difference sets having the parameters $(v; (v-1)/2; (v-3)/4)$. We refer to such Legendre pairs as type 1 (see next sections).

If we start with the list of odd integers v in the range $76 < v < 200$ and remove those which satisfy at least one of the conditions (i)–(iv) above we obtain the list of 20 integers:

77, 85, 87, 91, 93, 115, 117, 123, 129, 133, 145, 147, 159, 161, 169, 175, 177, 185, 187, 195.

This is in fact the list of all cases with $v < 200$ for which the question of existence of cyclic Legendre pairs is unresolved.

In the paper [1, p. 80] the authors list 22 odd lengths < 200 for which they assert that the existence question of cyclic Legendre pairs is unresolved. However, the lengths 121 and 171 should not have been included in that list since $2 \cdot 121 + 1 = 243 = 3^5$ and $2 \cdot 171 + 1 = 343 = 7^3$ are prime powers. (On the other hand according to [6, sec. 4] the number 57 should have been included.)

There are two other series of Legendre DFs in elementary abelian groups which include some cyclic cases. One of them appears in [7] and the other in [8, Theorem 3.1]. However, while they provide new cyclic Legendre DFs they do not give new lengths v in the cyclic case.

Our main result is in section with new pairs, where we give the first examples of cyclic Legendre pairs of lengths 91, 93 and 123. Thereby we reduce to 17 the number of the undecided cases listed above.

According to [1], exhaustive computer searches for cyclic Legendre pairs where carried out for all odd $v < 48$. We consider the odd integers v in the range $48 < v < 76$ and we list the new cyclic Legendre DFs of type 2 that we constructed. Only for $v = 69$ and $v = 75$ we failed to find any new pairs.

Notation and definitions

Let G be a finite abelian group (written additively) and let v denote its order. For any function f: $G \to \mathbf{R}$ its periodic auto-correlation function, PAF_f: $G \to \mathbf{R}$, is defined by the formula $\mathrm{PAF}_f(s) = \sum_{x \in G} f(x)f(x + s)$. We refer to s as the *shift variable*.

Definition 1. We say that an ordered pair of functions (f, g) mapping $G \to \{+1, -1\}$ is a Legendre pair on G if $\mathrm{PAF}_f(s) + \mathrm{PAF}_g(s) = -2$ for all nonzero shifts s. (For $s = 0$ we have $\mathrm{PAF}_f(0) = \mathrm{PAF}_g(0) = v$.)

For any function $f: G \to \{+1, -1\}$ we set $G_f = \{x \in G: f(x) = -1\}$.

Proposition. An ordered pair of functions (f, g): $G \to \{+1, -1\}$ is a Legendre pair on G if and only if (G_f, G_g) is a difference family in G with parameters (v, k_1, k_2, λ) where $k_1 = |G_f|$, $k_2 = |G_g|$ and $\lambda = k_1 + k_2 - (v + 1)/2$. In particular the existence of Legendre pairs on G implies that v must be odd.

Proof: This follows immediately from Theorems 3 and 4 of [9].

Remark. It is customary to require that the length v of a Legendre pair is > 1. However, according to the above definition, if G is a trivial group then any pair of functions $G \to \{+1, -1\}$ is a Legendre pair. The condition in the definition holds by default (there are no nonzero shifts).

If we introduce the additional parameter $n = k_1 + k_2 - \lambda$ then we have $v = 2n - 1$. By using the well known equation $k_1(k_1 - 1) + k_2(k_2 - 1) = \lambda(v - 1)$ one can easily show that

$$\{k_1, k_2\} \subseteq \left\{\frac{v-1}{2}, \frac{v+1}{2}\right\}.$$

In view of the above proposition, we shall refer to the difference families (X, Y) in G having the parameters $(v; k_1, k_2; \lambda)$ with $v = 2n - 1$ as *Legendre DFs*.

We say that a Legendre pair on a group G is *cyclic* if the group G is cyclic. In this note we deal only with the cyclic Legendre pairs and we may assume that $G = \mathbf{Z}_v$, the additive group of integers modulo v.

We give a simple example to introduce the notation that we will use in the rest of this note.

Example: $v = 39$. In this example $v = 13 \cdot 3$ and so $\mathbf{Z}_v^*$ (the group of units of the ring $\mathbf{Z}_v$) is isomorphic to $\mathbf{Z}_{12} \times \mathbf{Z}_2$. Then $H = \{1, 16, 22\}$ is the unique subgroup of $\mathbf{Z}_v^*$ of order 3. There exists an H-invariant Legendre DF with parameter set (39; 19, 19; 18), namely

$X = H\{0, 1, 2, 3, 4, 12, 14\}$;
$Y = H\{0, 2, 3, 4, 8, 14, 19\}$.

In general, if H is a subgroup of $\mathbf{Z}_v^*$ and S a subset of $\mathbf{Z}_v$ then the product HS is defined to be $HS = \{hs(\mathrm{mod}\ v)$: $h \in H$, $s \in S\}$. Note that $H\{0\} = \{0\}$.

Cyclic Legendre pairs of new lengths $v = 91, 93, 123$

We have constructed four pairwise nonequivalent Legendre DFs (X_i, Y_i) of length 91. For the definition of equivalence see the latest section of the paper. Only one DF is constructed for each of the lengths 93 and 123. Instead of Legendre pairs, we list the corresponding difference families. In each case, each block is a union of orbits of a fixed subgroup (H, H_1 or H_2) of order 3 or 5 of $\mathbf{Z}_v^*$.

$v = 91$

Four pairwise nonequivalent Legendre DFs:

(91; 45, 45; 44) $H_1 = \{1, 16, 74\}$, $H_2 = \{1, 9, 81\}$

$X_1 = H_1\{1, 2, 7, 14, 15, 17, 19, 22, 25, 28, 38, 43, 44, 50, 55\}$

$Y_1 = H_1\{2, 3, 10, 11, 14, 17, 20, 22, 28, 43, 44, 45, 49, 50, 55\}$

$X_2 = H_1\{1, 4, 5, 8, 9, 11, 15, 22, 27, 28, 34, 38, 43, 49, 50\}$

$Y_2 = H_1\{8, 9, 10, 11, 14, 17, 22, 25, 28, 33, 34, 38, 44, 50, 55\}$

$X_3 = H_1\{2, 3, 5, 9, 10, 14, 15, 20, 27, 28, 33, 34, 38, 50, 55\}$

$Y_3 = H_1\{3, 4, 11, 14, 19, 25, 27, 28, 33, 34, 43, 44, 45, 50, 55\}$

$X_4 = H_2\{2, 5, 14, 16, 19, 20, 23, 24, 29, 30, 37, 40, 46, 48, 49\}$

$Y_4 = H_2\{2, 4, 6, 8, 13, 14, 16, 23, 30, 37, 38, 39, 40, 46, 49\}$

$v = 93$

Only one Legendre DF:

(93; 46, 46; 45) $H = \{1, 25, 67\}$

$X = H\{0, 1, 2, 3, 5, 8, 10, 12, 13, 16, 22, 24, 43, 44, 47, 48\}$

$Y = H\{0, 1, 3, 4, 5, 9, 11, 12, 18, 20, 22, 37, 40, 43, 44, 51\}$

$v = 123$

Only one Legendre DF:

(123; 61, 61; 60) $H = \{1, 10, 16, 37, 100\}$

$X = H\{0, 1, 3, 6, 11, 13, 28, 29, 33, 35, 43, 45, 59\}$

$Y = H\{4, 5, 6, 11, 14, 15, 18, 19, 22, 28, 33, 41, 45\}$

Type 1 and type 2

Let (X, Y) be a Legendre DF in $\mathbf{Z}_v$. It is easy to see that if X is a difference set then Y must be a difference set too. In that case we say that (X, Y) (and its corresponding Legendre pair) is of type 1, and otherwise that it is of type 2. The Legendre pairs in the Galois and the twin-prime series as well as those in the classical series with v a prime number $\equiv 3 \pmod 4$ are of type 1. Note that two equivalent Legendre DFs must have the same type. Hence the study of type 1 Legendre DFs essentially reduces to the study of difference sets. For that reason we shall consider only the Legendre DFs of type 2.

As mentioned earlier, it is conjectured that cyclic Legendre DFs exist for all odd lengths $v > 2$. We propose a bit stronger version.

Conjecture. Legendre DFs of type 2 exist for all odd lengths $v > 8$.

One of the objectives of this note is to verify this conjecture for $v < 76$. It follows from [1] that the conjecture is true for $v < 48$. If v is a prime number $\equiv 1 \pmod 4$ then the classical Legendre pair of length v is of type 2. One can verify that the Legendre pairs in the Szekeres series having length v in the interval $4 < v < 76$ are of type 2. Thus, in order to verify the above conjecture for $v < 76$ it suffices to verify it in the cases $v = 49$, 55, 57, 59, 67, 71. This will be done in the next section.

We do not know whether all Legendre pairs of length $v > 4$ in the Szekeres series are of type 2.

New Legendre DFs of type 2

We list the cyclic Legendre DFs of type 2 and length $v > 48$ that we have constructed here. We imposed the restriction $v > 48$ because for $v < 48$ exhaustive searches have been carried out [1].

$v = 49$

The Legendre DF below is not equivalent to the one in [1, p. 85]:

(49; 24, 24; 23) $H = \{1, 18, 30\}$

$X = H\{1, 2, 8, 9, 13, 24, 26, 29\}$

$Y = H\{2, 3, 4, 6, 7, 8, 12, 37\}$

$v = 51$

The two Legendre DFs below together with the one in [1, p. 85] and another one from the Szekeres series are pairwise nonequivalent:

(51; 25, 25; 24) $H = \{1, 16\}$

$X_1 = H\{0, 1, 2, 4, 6, 8, 19, 24, 25, 28, 35, 38, 41\}$

$Y_1 = H\{0, 2, 4, 5, 9, 14, 15, 18, 21, 22, 25, 31, 35\}$

$X_2 = H\{1, 2, 9, 11, 17, 18, 19, 21, 24, 25, 28, 38, 41\}$

$Y_2 = H\{1, 3, 4, 5, 8, 9, 15, 17, 18, 19, 21, 22, 31\}$

$v = 53$

All ten Legendre DFs listed below are pairwise nonequivalent. The first five are known: the first belongs to the classical series, the second is from [7], the third from [8], the fourth from the Szekeres series, and the fifth from [1, p. 85]. We have constructed many Legendre DFs for $v = 53$ but we recorded only five of them (the last five in the list below):

(53; 26, 26; 25) $H = \{1, 10, 13, 15, 16, 24, 28, 36, 42, 44, 46, 47, 49\}$

$X_1 = H\{1, 4\}$, $Y_1 = H\{2, 5\}$

$X_2 = H\{1, 2\}$, $Y_2 = H\{1, 5\}$

$X_3 = H\{1, 2\}$, $Y_3 = H\{2, 5\}$

$X_4 = \{1, 4, 5, 6, 8, 14, 16, 17, 19, 21, 22, 23, 26, 28, 29, 33, 35, 38, 40, 41, 42, 43, 44, 46, 50, 51\}$

$Y_4 = \{4, 8, 9, 10, 12, 14, 15, 18, 19, 21, 22, 23, 24, 29, 30, 31, 32, 34, 35, 38, 39, 41, 43, 44, 45, 49\}$

$X_5 = \{5, 7, 12, 13, 15, 18, 19, 24, 26, 28, 30, 33, 35, 36, 37, 38, 39, 42, 43, 44, 46, 47, 48, 50, 51, 52\}$

$Y_5 = \{4, 7, 8, 10, 11, 14, 15, 20, 21, 23, 24, 25, 26, 29, 30, 32, 37, 40, 42, 44, 47, 48, 49, 50, 51, 52\}$

$X_6 = \{0, 1, 2, 3, 5, 9, 10, 11, 12, 14, 17, 24, 25, 26, 28, 29, 34, 35, 40, 44, 45, 46, 47, 48, 50, 51\}$

$Y_6 = \{0, 2, 4, 6, 7, 8, 12, 14, 17, 19, 20, 21, 22, 24, 27, 28, 30, 31, 34, 35, 40, 44, 46, 48, 49, 52\}$

$X_7 = \{0, 2, 3, 4, 7, 10, 11, 12, 13, 16, 17, 18, 21, 23, 32, 33, 37, 38, 39, 40, 41, 42, 45, 49, 50, 52\}$

$Y_7 = \{0, 1, 3, 6, 7, 12, 13, 15, 20, 21, 23, 24, 25, 31, 33, 35, 37, 38, 40, 42, 44, 46, 47, 48, 50, 51\}$

$X_8 = \{0, 1, 2, 6, 7, 9, 12, 13, 14, 15, 18, 19, 20, 24, 34, 36, 37, 38, 39, 41, 43, 45, 46, 49, 51, 52\}$

$Y_8 = \{0, 1, 5, 8, 9, 11, 12, 15, 16, 20, 21, 23, 24, 25, 31, 33, 35, 38, 40, 41, 43, 44, 47, 49, 51, 52\}$

$X_9 = \{0, 6, 8, 9, 10, 12, 14, 15, 23, 26, 28, 30, 31, 32, 33, 36, 37, 38, 40, 41, 42, 43, 48, 49, 50, 52\}$

$Y_9 = \{0, 1, 2, 3, 6, 8, 10, 11, 14, 15, 16, 18, 20, 23, 29, 31, 34, 35, 38, 39, 41, 42, 45, 47, 48, 52\}$

$X_{10} = \{0, 1, 6, 7, 8, 13, 14, 15, 17, 18, 19, 21, 22, 23, 24, 26, 28, 32, 37, 38, 40, 46, 48, 49, 50, 51\}$

$Y_{10} = \{0, 2, 6, 7, 10, 13, 14, 15, 16, 18, 19, 21, 22, 25, 26, 30, 31, 32, 33, 34, 36, 40, 43, 46, 48, 50\}$

$v = 55$

The Legendre DF below is not equivalent to the one listed in [1, p. 85]:

(55; 27, 27; 26) $H = \{1, 34\}$

$X = H\{1, 2, 6, 7, 8, 9, 10, 11, 15, 16, 21, 24, 27, 37, 50\}$

$Y = H\{1, 2, 3, 8, 16, 17, 19, 20, 21, 25, 27, 29, 37, 40, 42\}$

$v = 57$

In this case only two nonequivalent Legendre DFs are known. The first one was constructed in 2007 [6] and the second one constructed very recently in [10, Section 2.4]. We have constructed the six Legendre DFs below. The first five of them,

together with the two known DFs, are pairwise nonequivalent. The sixth is equivalent to the one constructed in [10]:

(57; 28, 28; 27) $H = \{1, 7, 49\}$
$X_1 = H\{0, 2, 3, 4, 8, 16, 23, 24, 30, 31\}$
$Y_1 = H\{0, 2, 3, 4, 6, 8, 16, 23, 24, 29\}$
$X_2 = H\{0, 2, 8, 10, 12, 23, 24, 29, 30, 31\}$
$Y_2 = H\{0, 3, 4, 5, 6, 22, 23, 24, 29, 31\}$
$X_3 = H\{0, 1, 3, 5, 6, 10, 16, 23, 29, 30\}$
$Y_3 = H\{0, 1, 2, 3, 15, 16, 22, 24, 29, 31\}$
$X_4 = H\{0, 1, 2, 4, 6, 11, 15, 29, 30, 31\}$
$Y_4 = H\{0, 8, 11, 12, 22, 23, 24, 29, 30, 31\}$
$X_5 = H\{0, 2, 3, 4, 6, 10, 15, 16, 29, 31\}$
$Y_5 = H\{0, 1, 3, 8, 10, 15, 16, 23, 24, 29\}$
$X_6 = H\{0, 2, 3, 4, 5, 11, 15, 16, 22, 30\}$
$Y_6 = H\{0, 1, 2, 4, 15, 16, 22, 23, 24, 30\}$

$v = 59$

First examples of Legendre DFs of length 59 and type 2:

(59; 29, 29; 28)
$X_1 = \{0, 1, 2, 3, 4, 5, 6, 10, 12, 13, 15, 16, 19, 20, 21, 25, 27, 30, 31, 33, 37, 38, 39, 41, 43, 44, 45, 52, 56\}$
$Y_1 = \{0, 1, 3, 4, 5, 6, 7, 10, 12, 13, 14, 15, 17, 18, 20, 23, 26, 27, 28, 30, 34, 35, 36, 39, 43, 45, 48, 50, 55\}$
$X_2 = \{0, 1, 2, 3, 4, 5, 7, 8, 9, 10, 12, 15, 16, 17, 22, 24, 25, 26, 28, 29, 33, 34, 38, 39, 42, 44, 48, 50, 53\}$
$Y_2 = \{0, 2, 3, 4, 6, 7, 9, 10, 12, 14, 15, 18, 19, 21, 23, 25, 29, 30, 31, 32, 33, 36, 38, 39, 43, 46, 49, 50, 51\}$
$X_3 = \{0, 1, 2, 3, 4, 5, 7, 8, 9, 10, 14, 16, 19, 20, 21, 24, 27, 28, 30, 32, 36, 37, 38, 41, 45, 47, 48, 51, 54\}$
$Y_3 = \{0, 2, 3, 4, 5, 6, 8, 9, 10, 12, 14, 16, 17, 19, 24, 25, 26, 28, 29, 31, 32, 34, 39, 42, 43, 44, 50, 54, 55\}$
$X_4 = \{0, 1, 2, 3, 4, 6, 9, 11, 12, 13, 15, 16, 18, 19, 20, 23, 24, 29, 30, 32, 36, 37, 38, 40, 41, 46, 47, 49, 56\}$
$Y_4 = \{0, 1, 2, 4, 6, 7, 9, 10, 11, 13, 14, 15, 17, 21, 22, 25, 26, 28, 30, 31, 33, 35, 36, 41, 43, 45, 46, 47, 53\}$
$X_5 = \{0, 1, 2, 3, 4, 6, 10, 12, 13, 14, 16, 18, 19, 21, 23, 26, 27, 28, 29, 31, 32, 36, 37, 40, 42, 43, 44, 49, 51\}$
$Y_5 = \{0, 1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 17, 18, 20, 21, 22, 25, 26, 31, 32, 36, 37, 38, 40, 42, 44, 45, 47, 52\}$
$X_6 = \{0, 1, 2, 3, 5, 6, 7, 8, 9, 11, 13, 15, 17, 18, 20, 24, 27, 28, 29, 31, 32, 38, 40, 41, 44, 45, 46, 54, 55\}$
$Y_6 = \{0, 1, 2, 3, 6, 7, 9, 10, 12, 13, 15, 17, 19, 20, 21, 24, 25, 27, 31, 32, 33, 36, 40, 41, 43, 44, 46, 49, 51\}$

$v = 61$

In this case apart from the classical Legendre DF there is another one provided by a lemma of J. Seberry Wallis [7], see also [11, Lemma 2]. The Legendre DF below is not equivalent to any of them:

(61; 30, 30; 29) $H = \{1, 9, 20, 34, 58\}$
$X = H\{2, 3, 4, 5, 12, 26\}$
$Y = H\{3, 4, 5, 10, 12, 13\}$

$v = 63$

The six new Legendre DFs below and the one from the Szekeres series are all pairwise nonequivalent:

(63, 31, 31, 30) $H_1 = \{1, 4, 16\}$, $H_2 = \{1, 25, 58\}$, $H_3 = \{1, 8, 11, 23, 25, 58\}$ $X_1 = H_1\{2, 3, 6, 10, 11, 22, 23, 30, 31, 42, 47\}$
$Y_1 = H_1\{1, 2, 7, 9, 10, 11, 14, 15, 21, 31, 47\}$
$X_2 = H_1\{1, 3, 9, 13, 14, 15, 21, 22, 23, 30, 47\}$
$Y_2 = H_1\{3, 5, 7, 9, 10, 11, 15, 21, 22, 23, 30\}$
$X_3 = H_1\{1, 2, 3, 6, 9, 11, 14, 21, 22, 30, 31\}$
$Y_3 = H_1\{1, 2, 3, 6, 7, 9, 11, 15, 22, 42, 47\}$
$X_4 = H_1\{2, 3, 7, 9, 10, 14, 15, 21, 26, 30, 43\}$
$Y_4 = H_1\{3, 6, 7, 10, 11, 14, 21, 27, 30, 43, 47\}$
$X_5 = H_2\{1, 3, 6, 7, 15, 17, 20, 27, 29, 40, 42\}$
$Y_5 = H_2\{3, 5, 7, 8, 10, 15, 17, 21, 27, 30, 40\}$
$X_6 = H_3\{0, 2, 9, 10, 15, 19, 27\}$
$Y_6 = H_3\{0, 2, 5, 7, 9, 15, 27\}$

$v = 65$

The Legendre DF below is not equivalent to the one in the Szekeres series:

(65; 32, 32; 31) $H = \{1, 16, 61\}$
$X = H\{1, 5, 6, 9, 18, 20, 22, 23, 24, 26, 35, 52\}$
$Y = H\{0, 1, 3, 7, 11, 13, 19, 22, 23, 24, 36, 50\}$

$v = 67$

The following Legendre DF gives the first example of Legendre pairs of length 67 and type 2:

(67; 33, 33; 32) $H = \{1, 29, 37\}$
$X = H\{1, 3, 5, 6, 10, 16, 17, 30, 34, 41, 53\}$
$Y = H\{2, 4, 6, 9, 12, 15, 16, 18, 25, 32, 41\}$

$v = 71$

We give the first example of a Legendre DF of length 71 and of type 2:

(71; 35, 35; 34) $H = \{1, 5, 25, 54, 57\}$
$X = H\{1, 2, 3, 6, 11, 14, 27\}$
$Y = H\{1, 2, 3, 9, 14, 18, 42\}$

$v = 73$

The Legendre DF below is not equivalent to the one in the classical series:

(73; 36, 36; 35) $H = \{1, 8, 64\}$
$X = H\{2, 5, 6, 7, 9, 11, 12, 17, 18, 26, 35, 42\}$
$Y = H\{1, 2, 3, 7, 9, 13, 18, 21, 26, 33, 35, 42\}$

$v = 111$

The Legendre DF below is not equivalent to the one belonging to the Szekeres series:

(111; 55, 55; 54) $H = \{1, 10, 100\}$

$X = H\{0, 1, 2, 3, 4, 7, 8, 9, 13, 16, 21, 22, 27, 41, 42, 44, 54, 62, 63\}$

$Y = H\{0, 1, 3, 4, 5, 6, 7, 8, 11, 16, 17, 21, 26, 27, 52, 53, 55, 63, 64\}$

$$v = 121$$

Note that $2v + 1 = 243 = 3^5$ is a prime power $\equiv 3 \pmod 4$. We list below two Legendre DFs (X_i, Y_i), $i = 1, 2$. The first one is equivalent to the DF in the Szekeres series. The block X_1 is skew and Y_1 is symmetric. We have constructed the second Legendre DF (X_2, Y_2) with $X_2 = X_1$ and verified that the two DFs are nonequivalent. Although Y_2 is not symmetric, the second pair still qualifies as a Szekeres difference set according to [3, Definition 5.6]:

(121; 60, 60; 59) $H = \{1, 3, 9, 27, 81\}$

$X_1 = H\{4, 10, 11, 20, 25, 26, 34, 35, 38, 40, 67, 76\}$

$Y_1 = H\{1, 7, 8, 10, 16, 20, 26, 31, 35, 38, 61, 94\}$

$X_2 = X_1$

$Y_2 = H\{1, 4, 5, 8, 11, 13, 17, 20, 22, 26, 34, 76\}$

Equivalence of Legendre pairs

To define the equivalence, we need first to define the elementary transformations on the set of Legendre pairs on a given finite abelian group G of odd order v. (We assume that G is written additively.) If f is a function $G \to \{+1, -1\}$ and $s \in G$ then we say that the function $G \to \{+1, -1\}$ sending $x \to f(s + x)$ is the *translate* of f by s.

The *elementary transformations* of a Legendre pair (f, g) are the following:

(i) interchange f and g;

(ii) replace f by $-f$;

(iii) replace f by its translate by $s \in G$;

(iv) replace f by $f \circ \iota$, where ι is the automorphism of G sending each $x \in G$ to its inverse $-x$;

(v) replace $(f; g)$ by $(f \circ \alpha, g \circ \alpha)$, where α is an automorphism of G.

Definition 2. We say that two Legendre pairs on G are equivalent if one can be transformed to the other by performing a finite sequence of elementary transformations.

The effect on (G_f, G_g) of the above elementary transformations is as follows:

(i)′ interchange G_f and G_g;

(ii)′ replace G_f by $G \backslash G_f$;

(iii)′ replace G_f by the translate $G_f - s$;

(iv)′ replace G_f by $-G_f$;

(v)′ replace (G_f, G_g) by $(\alpha^{-1}(G_f), \alpha^{-1}(G_g))$.

We define the equivalence of Legendre DFs on G by using the (i)′–(v)′ as elementary transformations of pairs (G_f, G_g). Then two Legendre pairs are equivalent if and only if their Legendre DFs are equivalent. We remark that because of (ii)′, two equivalent Legendre DFs may have different parameter sets.

Financial support and acknowledgements

The research of the first author leading to these results has received financial support of the Ministry of Science and Higher Education of the Russian Federation, agreement No FSRF-2020-0004. 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).

The authors wish to sincerely thank Tamara Balonina for converting this note into printing format.

References

1. 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, no. 23, pp. 75–86.
2. Szekeres G. Tournaments and Hadamard Matrices. *Enseignement Math.*, 1969, no. 15, pp. 269–278.
3. Seberry J. *Orthogonal Designs, Hadamard matrices, Quadratic Forms and Algebras*. Springer, 2017. 476 p.
4. Schroeder M. R. (W. D.) *Number Theory in Science and Communication*. Berlin; New York, Springer-Verlag, 1984. Vol. 7. 350 p.
5. Moore Emily H., Pollatsek Harriet S. *Difference Sets: Connecting Algebra, Combinatorics, and Geometry*. Ser. Student Mathematical Library. AMS, 2013. Vol. 67. 314 p.
6. Chiarandini M., Kotsireas I. S., Koukouvinos C., Paquete L. Heuristic algorithms for Hadamard matrices with two circulant cores. *Theoretical Computer Science*, 2008, vol. 407, pp. 274–277.
7. Seberry Wallis J. Some remarks on supplementary difference sets. *Colloquia Mathematica Societatis Janos Bolyai*, 1973, no. 10, pp. 1503–1526.
8. Cunsheng Ding. Two constructions of $(v, (v-1)/2, (v-3)/2)$ difference families. *J. Combin. Designs*, 2008, no. 16, pp. 164–171.
9. Đokovic Dragomir Z., Kotsireas Ilias S. Computational methods for difference families in finite abelian groups. *Spec. Matrices*, 2019, no. 7, pp. 127–141.
10. Arasu K. T., Bulutoglu D. A., Hollon J. R. Legendre G-pairs and the theoretical unication of several G-array families. *Combinatorics*, arXiv:2004.05608v1 [math.CO], 12 Apr 2020.
11. Đokovic Dragomir Z. Survey of cyclic $(v; r; s; \lambda)$ difference families with $v \leq 50$. *Facta Universitatis (Nis), Ser. Math. Inform.*, 1997, no. 12, pp. 1–13.

УДК 004.438
doi:10.31799/1684-8853-2021-1-2-7

Три новые длины циклических пар Лежандра

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

Введение: согласно гипотезе, циклические пары Лежандра нечетной длины > 1 всегда существуют. Такая пара состоит из двух функций a, b: $G \to Z$, принимающих значения +1 или –1, с периодическими автокорреляционными функциями, принимающими в сумме постоянное значение –2 (исключая начальную точку). Здесь G — конечная циклическая группа, Z — кольцо целых чисел. Эти условия являются фундаментальными и тесно связаны со структурой бициклических матриц Адамара с двойной каймой, недостаточно полно описанной в литературе, что делает ее исследование особенно актуальным. **Цель:** дополнить описание бициклической конструкции с двойной каймой тремя новыми решениями пар Лежандра. **Результаты:** для характеристики пар Лежандра использованы подмножества $X = \{x \in G: a(x) = -1\}$ и $Y = \{x \in G: b(x) = -1\}$ из G. Есть 20 нечетных целых чисел v, меньших 200, для которых существование пар Лежандра длины v не доказано. Наименьшее из них — $v = 77$. Построены пары Лежандра длиной 91, 93 и 123, в результате количество нерешенных случаев сократилось до 17. Приводятся примеры циклических пар Лежандра для длин $v \le 123$. **Практическая значимость:** матрицы Адамара широко используются в задачах помехоустойчивого кодирования, сжатия и маскирования видеоинформации. Программы поиска матриц Адамара и библиотека построенных матриц используются в математической сети mathscinet.ru вместе с исполняемыми онлайн алгоритмами.

Ключевые слова — матрицы Адамара, периодические автокорреляционные функции, пары Лежандра, циклические матрицы, бициклические конструкции.

Для цитирования: Balonin N. A., Đoković D. Ž. Three new lengths for cyclic Legendre pairs. *Информационно-управляющие системы*, 2021, № 1, с. 2–7. doi:10.31799/1684-8853-2021-1-2-7
For citation: Balonin N. A., Đoković D. Ž. Three new lengths for cyclic Legendre pairs. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 2–7. doi:10.31799/1684-8853-2021-1-2-7

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

UDC 004.057.4
doi:10.31799/1684-8853-2021-1-8-16

Articles

Analysis of requirements for modern spacecraft onboard network protocols

V. L. Olenev[a], *PhD, Tech., Associate Professor, orcid.org/0000-0002-1817-2754, Valentin.Olenev@guap.ru*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*

Introduction: *New technologies are replacing the onboard space networks based on bus topologies. One of these technologies is SpaceWire. New communication protocols are being developed, expanding SpaceWire functionality. The protocol developers should provide all the required technical characteristics for data transmission and processing.* ***Purpose:*** *New technologies are replacing the onboard space networks based on bus topologies. One of these technologies is SpaceWire. New communication protocols are being developed, expanding SpaceWire functionality. The protocol developers should provide all the required technical characteristics for data transmission and processing.* ***Results:*** *The analysis of the existing demands on communication protocols resulted in a set of consolidated requirements for the physical-network layers' protocols and the transport layer protocols. The requirements cover the speed, latencies, transmission distance, transmitted information amount, fault detection functionality, time synchronization between the devices, quality of service, main user data types, and data transfer modes at the transport level. The existing SpaceWire protocols are defined as a special class of protocols, possessing unique characteristics.* ***Practical relevance:*** *The performed analysis can simplify the implementation of new onboard communication protocols and provide a required level of technique for new generation spacecraft.*

Keywords *— onboard networks, communication protocols, technical requirements, SpaceWire.*

For citation: Olenev V. L. Analysis of requirements for modern spacecraft onboard network protocols. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 8–16. doi:10.31799/1684-8853-2021-1-8-16

Introduction

Communication technologies for onboard communication networks are rapidly developing. New standards and protocols, new principles and mechanisms of data transmission, new equipment that implement these mechanisms and protocols appear [1, 2]. The MIL-STD 1553 [3] bus has been used for data exchange in onboard systems since the 1970s, but the rapidly growing requirements for the functionality of spacecraft make its further use impossible. As a result, network topologies are replacing the buses according to the demands of the leading industrial companies. One of such technologies is SpaceWire.

Open standard ECSS-E-50-12C (SpaceWire protocol) was specifically developed for space applications, so it has a low implementation cost and complexity, high performance, and flexible architecture [4]. SpaceWire met all the requirements for aerospace applications [5, 6] and has become the dominant technology used for small sized spacecraft, landing modules, etc. [7, 8]. Later, the SpaceWire protocol was supplemented with the GigaSpaceWire protocol [9], which provides gigabit speeds, and in 2019, the next-generation standard ECSS-E-ST-50-11C (SpaceFibre protocol) was released [10]. However, currently SpaceWire remains the main protocol used in real missions.

The SpaceWire, GigaSpaceWire, and SpaceFibre protocol specifications cover the OSI model layers from physical to network, and a number of transport protocols with different functionality and complexity have been developed. These transport protocols greatly extend the functionality of SpaceWire family protocols. Transport protocols were developed for SpaceWire, but due to compatibility at the network level, they can be used for GigaSpaceWire and SpaceFibre. An analysis of the existing transport protocols is given in the article [11]; [12] provides a detailed overview and comparison of the standards ECSS-E-ST-50-52C (RMAP), ECSS-S-ST-50-53C (CPTP), SMCS-ASTD-PS-001 (STUP), and the protocols STP [13] and JRDDP [14]. Overview shows that these transport layer protocols are not sufficient to provide different types of quality of service, reliable data delivery, and configuration flexibility. Therefore, the transport layer protocols continue to be improved within the missions of various space agencies. One such development by JAXA is the SpaceWire-R protocol [15], which introduced guaranteed data delivery and transport connections. ESA introduced the SpaceWire-D protocol [16], which for the first time introduced deterministic data delivery in the SpaceWire network [17]. For Russian spacecraft using SpaceWire networks, the STP-ISS protocol was created [18]. By that time STP-ISS provided the necessary transport-level mechanisms for the Russian industry. However, the requirements for on-board networks are changing as the technical capabilities for implementing different protocols.

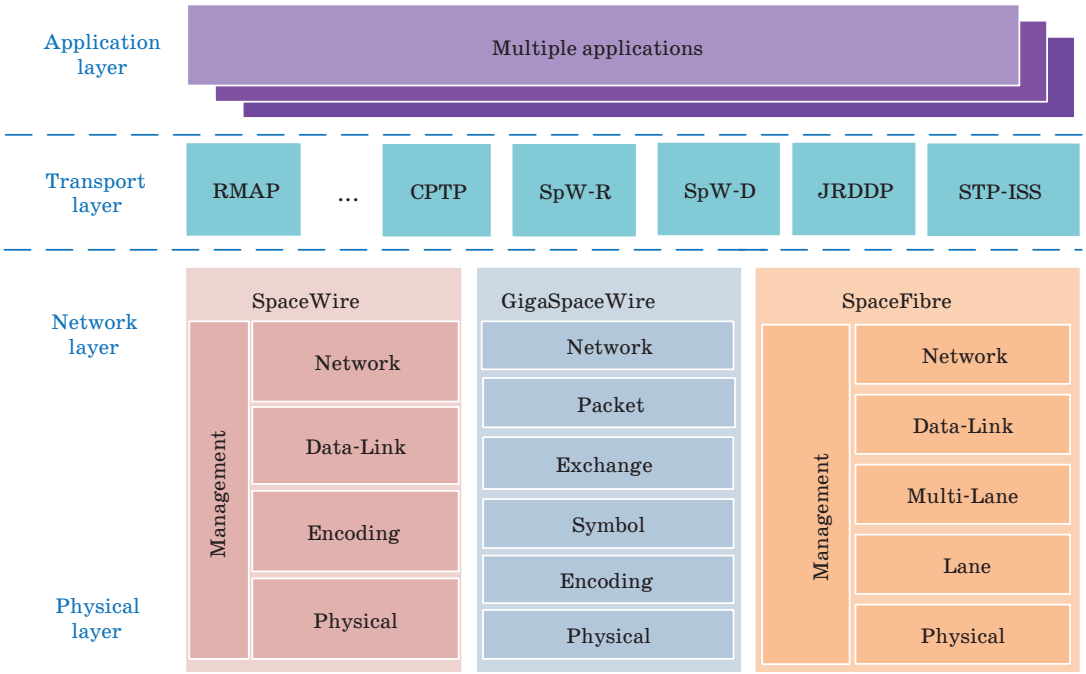
The urgent remaining task is to form a consolidated set of requirements for communication protocols for onboard space networks it will further allow analyzing existing technologies for compliance and setting tasks for improving existing data transmission standards. Thus, current paper will consider the existing requirements for communication protocols. Based on this analysis, a set of consolidated requirements for the onboard space protocols will be derived. In addition, paper will show that the SpaceWire family protocols satisfies these requirements. The requirements for protocols presented in the article available in open sources, are collected from leading companies of space industry, space industry experts, real developers of onboard space equipment.

Overview of communication protocol requirements

The SpaceWire family of protocols can be divided into two main categories — these are protocols that describe the OSI (Open Systems Interconnection) layers from physical to network: SpaceWire, GigaSpaceWire and SpaceFibre, and many transport layer protocols that provide various end-to-end functionality (Fig. 1). Requirements for protocols at different layers of the OSI model will also differ and will be grouped into network-physical layer protocols and transport layer protocols.

A survey of experts from the Russian space industry conducted on the technical requirements for communication protocols of the network-physical layers and the streaming service of the transport layer showed the following. Data packets should not be delivered to the wrong destination or filtered in the receiver. The network should be free of deadlocks, so packets that cannot reach the destination within a given fixed time should be dropped. The required Quality of Service (QoS) is guaranteed delivery, guaranteed bandwidth, multiple levels of priority (from fore to six). The network should provide reliable time synchronization — incorrect time stamps and interrupt codes should not be delivered to recipients. It should be possible to transfer high-speed data streams (for example, video streams) over separate virtual connections, while other traffic should continue to be transmitted over other connections.

In terms of the streaming data transmission (for example, video data), separate requirements were presented. It takes into account the features of the industry standards ARINC-818-2 and CCSDS 766.1-B-1, the characteristics of streaming video. A relatively constant rate of data entry into the network should be ensured. It could be achieved by limiting the maximum size and intensity of packet transmission to the network throughout the entire information exchange session. An important characteristic of streaming traffic is also the low latency for transmitting real-time streaming data. Therefore, protocols with the transport connection establishment should be used, which will reduce the length of the header of packets containing useful data. It is also preferable to use a simple data delivery mechanism implemented in hardware. It should be done to



■ ***Fig. 1.*** SpaceWire/SpaceFibre standards family

have a mode without buffering on the transmitter and receiver, with packet delivery without acknowledgements and resending. It is important to control the delivery of data to the receiver: checking the correctness of the packet header and the payload field; detection of the erroneous packets, lost packets and packets reordering. The new protocols should be compatible with SpaceWire\SpaceFibre networks [19, 20].

Another source of requirements is the comparison of communication architectures given in [21]. It contains the following NASA requirements for a rigid real-time distributed control system for mission-critical security systems on a manned spacecraft. The first parameter is high reliability and high accessibility. The use of modular components at all levels to ensure high reusability, flexibility and scalability. These components should support Plug & Play technology and, if possible, perform hot-swapping. The Plug & Play technology for SpaceWire has several implementations adapted to the requirements of various space agencies, and is successfully developing [22, 23]. Comprehensive functionality for fault detection, isolation and recovery (FDIR) and health monitoring should be provided. It is also noted that the ability to transmit large amounts of data at extremely high speeds is not mandatory. Most control circuits operate at a frequency of 100 Hz or less. For example, the space shuttle main engine controller operates at 50 Hz, and the flight control loop in the space shuttle computers operates at 25 Hz.

G. Kopets in his book [24] provides a set of requirements for the communication infrastructure of distributed real-time systems. The first group of requirements relates to temporary properties. The message transfer delay should be as low as possible to minimize the idle time of the control commands. It is necessary to ensure a minimum jitter, that is, the difference between the worst-case message latency and the best-case message latency. In such systems, it is necessary to have a global time value for all network nodes with proper accuracy (time synchronization). A reliable time synchronization algorithm should set the internal time of the network nodes so close to each other that the amount of time discrepancy during the offline operation does not exceed the specified accuracy interval. A fault-tolerant time synchronization algorithm should allow the specified number of errors in the network. Such a requirement in SpaceWire networks is represented by a mechanism for sending of high-priority timestamps, but time synchronization mechanisms are not described. They are represented at the transport layer, and only in the STP-ISS protocol [25].

The second group of requirements relates to error detection and recovery mechanisms. Reliability of communication should be ensured through the use of reliable channel coding or algorithms based on broadcast. In systems that do not operate in real time, reliability can be achieved through retransmission. Mechanisms for control of the malfunction of components in time are needed. For example, the communication system should contain information on the permitted behavior of the component in time and can disable the component that violates the rules of operation (babbling idiots avoiding). Each network element should report about all component failures. It is necessary to use end-to-end confirmation of the success or failure of any action for any scenario. It is important to use mechanisms that ensure determinism. These requirements are fully taken into account in the second edition of the STP-ISS protocol, which provides determinism and mechanisms for detecting of duplicated control commands.

The last group of requirements covered in the book, which indirectly affects data transmission protocols, relates to the physical structure of a real-time communication system. This requirement is low cost and low weight of equipment.

Let us also consider the requirements of the Russian and European industry for communication protocols for on-board systems. Industrial companies within the framework of the FP7 Program SpaceWire-RT project [26] jointly elaborated this list of requirements.

To ensure timely data delivery, support of transmission rates of up to 20 Gbit/s for remote sensing missions, and speeds of up to 400 Mbit/s for low latency routing is required. The operation of the devices should be possible at a distance of up to 100 m for spacecraft applications, where equipment can be installed at a long distance. Additionally operation of the devices should be possible at a distance of 1 to 10 m for operation at high data rates between closely located equipment.

It is necessary to support the transmission of application messages with a size of at least 32 MB. Such packets are used to transmit raw data. Message sizes from 8 B to 64 KB should be supported to transmit commands and telemetry from application processes. In this case, the maximum latency for the transit of command packets over the network in real-time applications should be less than 100 ms. and for time synchronization packets up to 100 ns.

Protocols should provide capabilities for reliable delivery of important data that should be delivered without corruption. Determinism and configurable automatic confirmation for controlling non-intelligent devices and sensors should be provided. The protocols should provide mechanisms for automatic FDIR. At the same time, recovery could be implemented in most applications in soft-

ware. In some applications where short response times are hardly achievable, automatic recovery could be implemented at the network layer.

Time-critical commands require support for multi-path data transmission and support for multicast data transmission, for example, to deliver data to devices in a redundant system. The requirements states the need to support the transmission of timestamps.

The quality characteristics are given in a generalized table describing the support for communication requirements. Table 1 shows the main characteristics for each of the required traffic types.

Next, consider the requirements for the functionality of the onboard network from "Academician M. F. Reshetnev "Information Satellite System" within the framework of a joint project aimed at creating a modern transport protocol STP-ISS [15].

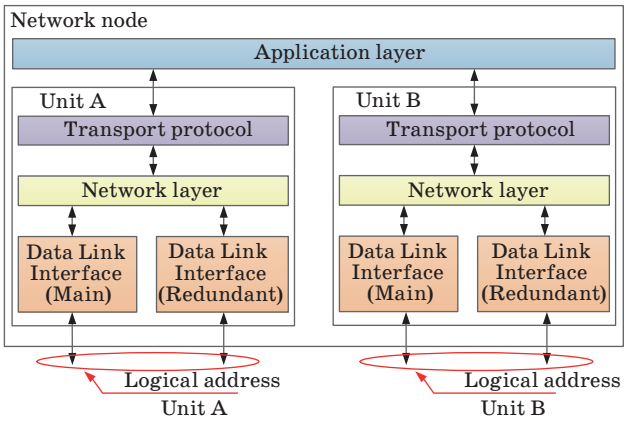
The maximum number of logical addresses specified in the corresponding protocol should determine the number of nodes in the network. From one to three logical addresses could be set for one network node. In accordance with the SpaceWire protocol, it is necessary to preserve the possibility of dividing the network into regions. The maximum number of logical addresses of a particular protocol should also determine the number of nodes in each region. When using path or regional-logical addressing, the number of transit switches in the network (or subnet) should be no more than 15.

From one to three units could represent each node of the onboard network. Each unit should have a separate logical address. Fig. 2 illustrates this requirement for a network node.

The requirements of "Academician M. F. Reshetnev "Information Satellite System" are mostly focused on transport layer protocols, since the lower layer protocol (SpaceWire) has already been defined as a main protocol for future missions. The transport protocol is needed to provide transport services for onboard networks and should describe the data processing and exchange mechanisms, packet formats. It should transmit the following user data types from the transmitter application layer to the receiver application layer: control commands, application messages, time-codes, interrupt signals and interrupt acknowledgements. Urgent packets and regular packets could represent application messages. The protocol itself should provide data transmission in two modes: connection oriented and connectionless.

Connection establishment is performed separately for each pair of receiver-transmitter remote network nodes (Fig. 3). The connection establishment initiator could be either an active or a passive device. Only one type of data should be transmitted over the transport connection. Control commands should not be transmitted in connection-oriented mode. The maximum number of transport connections for each node shall not exceed 8 connections in one direction. Within each transport connection, a flow control mechanism should be provided. Flow control means sending the information on the remaining free space in the receiving buffer to the transmitter of the transport connection.

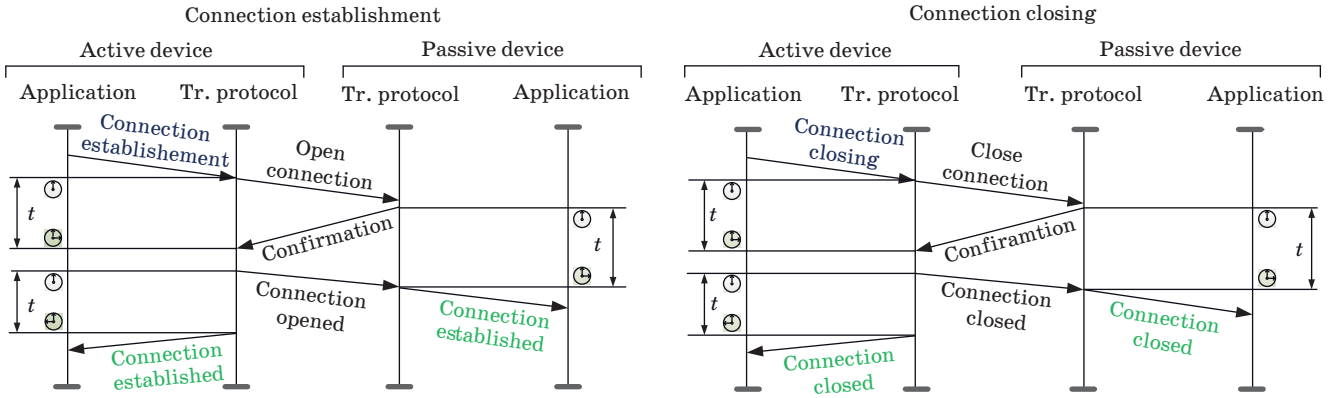
In the connectionless mode, the segment of data transmitted to the transport protocol shall not be larger than 2048 B. Limit for the connection-oriented mode is 64 KB. If the size of the application message exceeds the maximum allowed size of the data segment, the application should perform segmentation of this message.



■ ***Fig. 2.*** Logical addresses distribution for the network node unit

■ ***Table 1.*** Main characteristics for required classes of data

Class of data	Distance	Speed	Latency	Packet size	Quality of service
Data	Short and long	From low to very high	Not important	Short to long	Reserved bandwidth
Control commands	Short and long	Low	Low	Short to long	Deterministic delivery
Telemetry	Short and long	Low	Low	Short	Reserved bandwidth
Time stamps	Short and long	Low	Very low	Short	High priority



■ ***Fig. 3.*** Example of transport connection establishment and closing mechanisms

The transport protocol should implement the following mechanisms for transmission errors detection: CRC check, validation of the packet data field length, confirmation of successful data reception, timeouts for detecting lost data packets. Data should be prioritized for different information flows (at least 3 priority levels for data packets and control commands). The transport protocol should contain a separate logical buffer for each priority data coming from the application layer.

Requirements for the provided quality of service for information flows should be provided in accordance with Table 2. Requirements for the data transmission latency are given for a data channel with transmission through 8 transit switches and transmission rates of 20 and 50 Mbit/s.

The column "Quality of service" states the following QoS types:

a) priority — higher priority data transmitted first;

b) scheduling — a single schedule is created for the whole network; transmitting nodes are allowed to send data according to this schedule. It is not managed in switches. Schedule is available in the local and remote nodes exchanging data, in the source and in the receiver (Fig. 4);

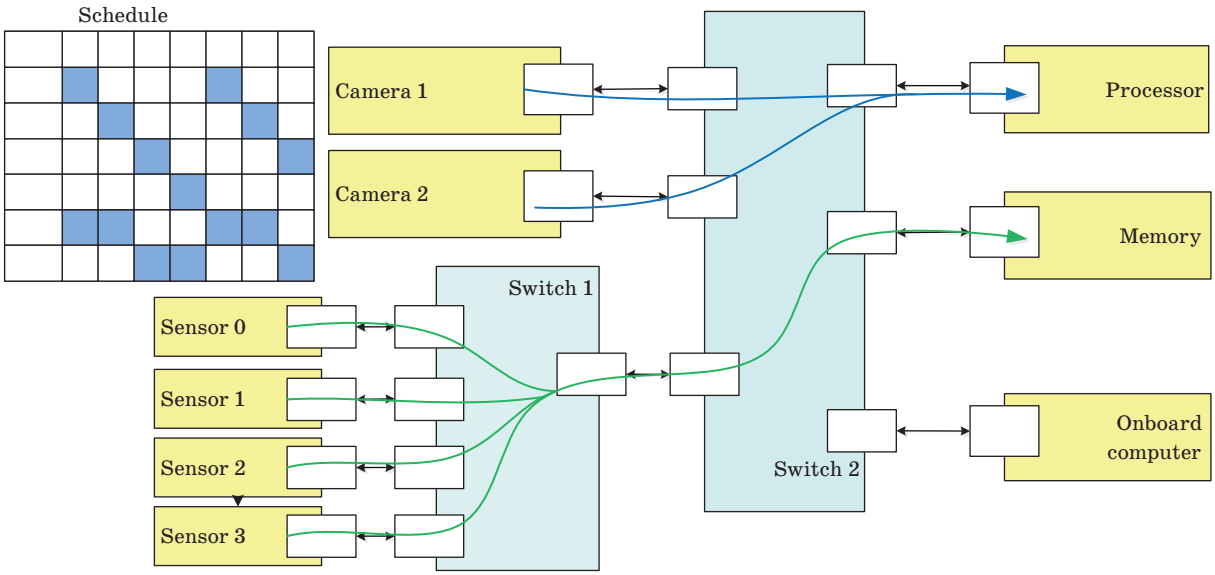
c) guaranteed delivery — confirmation of correct data delivery, re-sending by the source if there is no confirmation during the timeout (Fig. 5);

d) not-guaranteed delivery.

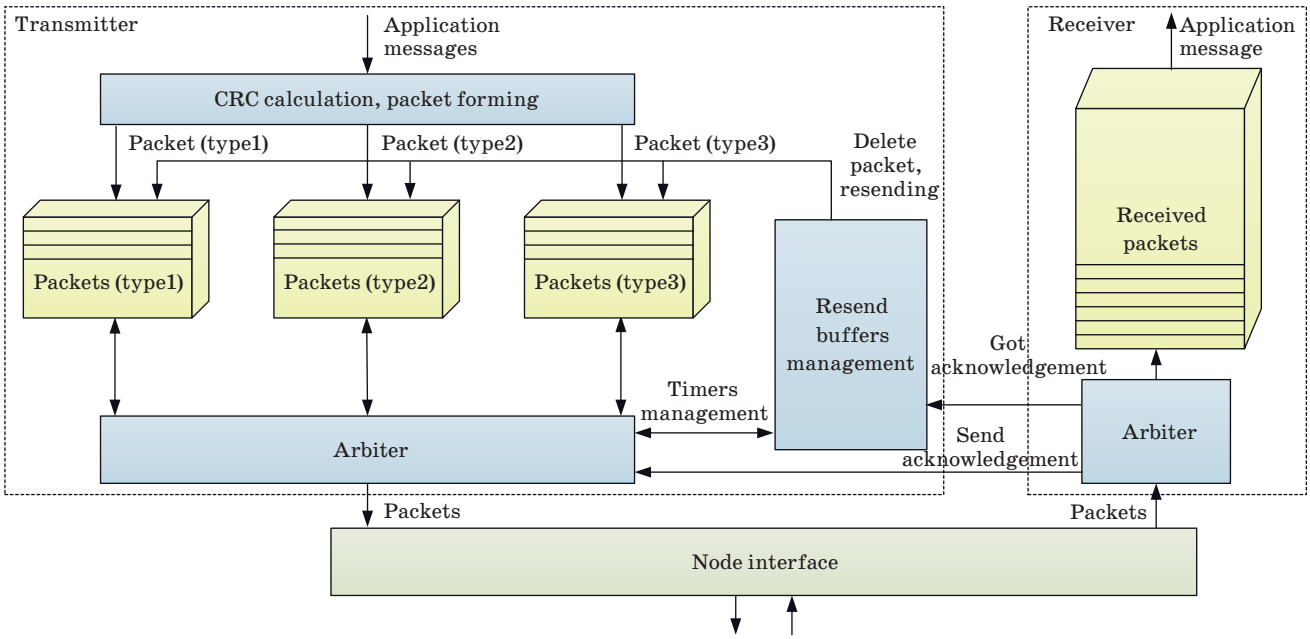
The scheduling quality of service should be carried out in accordance with the specified schedule. It is formed at the stage of the transport protocol configuration. At the same time, the protocol itself should be able to operate without the scheduling quality of service.

■ ***Table 2.*** Data streams' quality of service

Data type	Length	Generation frequency	Delay, 20 (50) Mbps	Quality of service	Priority	Confirmation
Control command	4 B	≥ 1 ms	≤ 1 (0.5) ms	a, b, c	1	Yes
Urgent message	64 B	≥ 0.5 ms	≤ 1 (0.8) ms	a, b, c	2	Yes
	2 KB	≥ 5 ms	≤ 8 (5) ms			
	64 KB	≥ 100 ms	≤ 230 (150) ms			
Regular message	64 B	≥ 0.5 ms	≤ 1.5 (1) ms	a, b, c, d	3	Yes /No
	2 KB	≥ 5 ms	≤ 10 (7) ms			
	64 KB	≥ 100 ms	≤ 300 (220) ms			
Time-code	6 bit	≥ 60 s	≤ 0.1 ms	a	0	No
Interrupt. interrupt acknowledge	5+1 bit	≥ 5 ms	≤ 0.1 ms	a	0	Yes /No



■ ***Fig. 4.*** Example of data transmission scheduling for onboard networks



■ ***Fig. 5.*** Guaranteed quality of service example

The transport protocol should provide the further distribution for the time-codes and system interrupts from the local node application to the network. Similarly, the transport protocol should accept time-codes and system interrupts from the network and transmit them to the applications. The protocol could use the system time information from the time-codes to implement the Scheduling quality of service.

The transport protocol should detect duplicate control commands on the receiver and discard the duplicates.

It is necessary to provide the possibility of simplified configuration of the protocol for small-size networks (no more than 2 hops).

Consolidated requirements for the communication protocols at the physical-network level

Based on the provided industry requirements, it is possible to form consolidated requirements for communication protocols. Let us start with

the requirements at the physical-network levels:

1) support data transfer rates up to 20 Gbps, including intermediate speeds;

2) ensuring operation at distances up to 100 m;

3) reliable communication through the use of reliable channel coding;

4) providing comprehensive functionality for FDIR and efficiency monitoring at the data link layer;

5) point-to-point determinism;

6) ensuring time synchronization between devices. It is provided by supporting the transmission of time-codes. At the same time, synchronization should be reliable — erroneous time-codes and interrupt codes should not appear on the network and should be discarded;

7) support for transmitting of information with a size of 32 MB or more. For transmitting commands and telemetry — from 8 B to 64 KB;

8) support maximum data transfer latency:

a) for transmitting control commands less than 100 ms;

b) for time synchronization up to 100 ns;

c) for communication between two processor modules up to 100 ns in one link;

9) support multipath data transmission;

10) support multicast data transmission.

Consolidated requirements for communication protocols at the transport layer

The requirements at the transport layer relate to end-to-end transmission between the information transmitter and the receiver.

1) required quality of service: Guaranteed delivery, guaranteed bandwidth, priorities (from fore to six);

2) exchange of end-to-end acknowledgements;

3) support for the following main user data types:

a) control commands;

b) urgent packets;

c) regular packets;

d) time-codes;

e) interrupt codes and interrupt acknowledgments;

4) data transmission in connection-oriented and connectionless modes;

5) connection-oriented mode: connection for each receiver-transmitter pair; both active and passive devices could be initiators of data exchange; 8 unidirectional transport connections maximum; one transport connection for one data type; flow control mechanism; data segment length up to 64 KB;

6) connectionless mode: data segment length up to 2048 B;

7) command control length up to 4 B;

8) quality of service for Control commands and Urgent packets: priority, scheduling, guaranteed delivery;

9) quality of service for Regular Messages: priority, scheduling, guaranteed delivery, non-guaranteed delivery;

10) the transfer of time-code is carried out only in a non-guaranteed mode;

11) detection of duplicate control commands at the node receiver, discarding duplicates;

12) the possibility of a simplified protocol configuration for simple networks;

13) possibility to switch off the scheduling quality of service.

Additional requirements for the protocols are also:

1) interfaces to access the functions of the protocol (Service Access Points);

2) minimum data transfer delays;

3) the smallest possible footprint of the chip and the energy consumption. Since the mechanisms described in the protocols can be difficult to implement, require the additional memory (for example, segmentation, transport connections) and, as a result, occupy a large chip area. This may lead to exceeding the permissible weight and energy consumption characteristics for the spacecraft.

Conclusion

This article discusses the requirements for communication protocols for onboard networks from various open sources: scientific literature, project reports, and scientific articles. These requirements provide the important vision of what future data transfer protocols for spacecraft should look like. The analysis showed that many sources has the similar requirements. Therefore, on their basis, a number of main characteristics for the protocols of the physical-network layers and the transport layer are elaborated.

The article focuses on certain aspects of the operation of onboard networks and does not consider all possible parameters of the spacecraft operation. However, taking into account the authority of the analyzed sources, it can be concluded that these requirements are currently the focus of the global space industry in terms of communication protocols.

The requirements obtained during the analysis reflect the distinctive features of the SpaceWire/SpaceFibre technology. It is a simple routing mechanism without buffering, which is able to transmit data of various unlimited lengths, and also has a

relatively small hardware implementation cost. It is the flexibility and scalability of protocols, the availability of opportunities for simplified configuration and assembly of networks using Plug-n-play technology. In addition, the SpaceWire family describes specialized types of high-priority packets for transmitting time data, as well as time synchronization mechanisms. It is important that this is an open technology that provides almost unlimited opportunities to expand the protocol family by transport layer protocols while maintaining compatibility with previous versions. The combination of these characteristics, concentrated in the protocols of the SpaceWire family, distinguishes these communication protocols for on-board networks into a separate group of space protocols that have characteristics that are not presented in other protocols.

The analysis of the requirements shows that the SpaceWire/SpaceFibre protocol family meets them. So they can be considered as the main ones for further application in the space industry. At the transport layer, at the moment, only the STP-ISS protocol meets all the described requirements due to the support of most mechanisms for ensuring the quality of service, reliable data transmission and various implementation profiles. At the same time, combinations of other existing protocols that solve specific narrowly focused tasks, also could provide the necessary performance characteristics. The use of the STP-ISS protocol is limited by the desire and ability of companies to use third-party protocols, so the development of new protocols is inevitable. The proposed requirements should be taken into account when developing these protocols in order to provide the important services for the new generation of onboard equipment.

Financial support

The paper was prepared with the financial support of the Ministry of Science and Higher Education and of the Russian Federation, grant agreement No. FSRF-2020-0004 “Scientific basis for architectures and communication systems development of the onboard information and computer systems new generation in aviation, space systems and unmanned vehicles”.

References

1. Tavoularis A., Vlagkoulis V., Kostopoulos F., Le Ngoc T., Dellandrea B., Fossati L., Ilstad J., Jameux D. SpaceWire components, long paper: An IP core for the SpW family of protocols. *2016 International SpaceWire Conference (SpaceWire)*, Yokohama, 2016, pp. 1–8. doi:10.1109/SpaceWire.2016.7771642
2. Kapranova E. A., Nenashev V. A., Sergeev M. B. Compression and coding of images for satellite systems of Earth remote sensing based on quasi-orthogonal matrices. *Proc. of SPIE, Image and Signal Processing for Remote Sensing XXIV*, Berlin, Germany, 2018, vol. 10789, pp. 1078923-1–1078923-6. doi:https://doi.org/10.1117/12.2324249
3. Orly S., Elovici Y., Shabtai A., Shugol G., Tikochinski R., Kur S. Protecting military avionics platforms from attacks on MIL-STD-1553 communication bus. *Computing Research Repository*, 2017, pp. 1–15.
4. Parkes S. *SpaceWire Users Guide*. Dundee, Star-Dundee, 2020. 117 p.
5. Nepomnyashii O. V., Postnikov A. I., Goreva V. V., Varochkin S. S. Architecture of onboard management system for small satellites based on networking technologies. *Issledovaniia naukograda*, 2017, no. 1, pp. 22–29 (In Russian).
6. Parkes S., Armbruster P. SpaceWire: a spacecraft onboard network for real-time communications. *14th IEEE-NPSS Real Time Conference*, 2005, Stockholm, 2005, pp. 6–10. doi:10.1109/RTC.2005.1547397
7. Notebaert O., Montano G., Planche T., Pruvost C., Wartel F., Schüttauf A., Herpel H., Honvault C., Jameux D. Towards SpaceWire-2: Space robotics needs: SpaceWire missions and applications, long paper, *2016 International SpaceWire Conference (SpaceWire)*, Yokohama, 2016, pp. 1–9. doi:10.1109/SpaceWire.2016.7771614
8. Dello Sterpaio L., Marino A., Nannipieri P., Dinelli G., Davalle D., Fanucci L. A complete EGSE solution for the SpaceWire and SpaceFibre protocol based on the PXI industry standard. *Sensors*, 2019, vol. 19, no. 22, p. 5013. doi:10.3390/s19225013
9. Yablokov E., Sheynin Y., Suvorova E. GigaSpaceWire — gigabit links for SpaceWire networks. *Proceedings of the 5th International SpaceWire Conference*, Gothenburg, 2013. pp. 28–34.
10. Parkes S., Ferrer Florit A., Gonzalez-Villafranca A. SpaceFibre interfaces and architectures. *In 2019 IEEE Aerospace Conference*, IEEE, 2019. pp. 1–8. https://doi.org/10.1109/AERO.2019.8741961
11. Peng T., Weps B., Höflinger K., Borchers K., Lüdtke D., Gerndt A. A new SpaceWire protocol for reconfigurable distributed on-board computers: SpaceWire networks and protocols, long paper. *2016 International SpaceWire Conference (SpaceWire)*, Yokohama, 2016. pp. 1–8. doi:10.1109/SpaceWire.2016.7771624
12. Olenev V. L., Lavrovskaya I. I., Korobkov I. L., Dymov D. V. Analysis of the transport protocol requirements for the SpaceWire on-board networks of spacecrafts. *Proceedings of 15th Seminar of Finnish-Russian University Cooperation in Telecommunications (FRUCT) Program*, Saint-Petersburg, 2014, pp. 65–71. doi:10.1109/FRUCT.2014.6872424

13. Sheynin Y., Suvorova E., Schutenko F., Goussev V. Streaming transport protocols for SpaceWire networks. *International SpaceWire Conference 2010*, Saint-Petersburg, 2010, pp. 56–59.
14. Sandia National Laboratories. *Joint Architecture System Reliable Data Delivery Protocol (JRDDP)*. Albuquerque, SNL, 2011. 72 p.
15. Mich W., Romanowski K., Tyczka P., Renk R., Kollias V., Pogkas N. Implementation and validation of the SpaceWire-R protocol. *2016 International SpaceWire Conference (SpaceWire)*, Yokohama, 2016, pp. 1–4. doi:10.1109/SpaceWire.2016.7771601
16. Gibson D., Parkes S., McClements C., Mills S. SpaceWire-D prototype and demonstration system. *2016 International SpaceWire Conference (SpaceWire)*, Yokohama, 2016, pp. 1–7, doi:10.1109/SpaceWire.2016.7771645
17. Gibson D. *Deterministic SpaceWire Networks*. University of Dundee, 2017. 297 p.
18. Sheinin Iu. E., Olenev V. L., Lavrovskaya I. Ia., Dymov D. V., Kochura S. G. Development, analysis and design of STP-ISS transport protocol for onboard SpaceWire networks. *Issledovaniia naykograda*, 2016, no. 1-2, pp. 21–30 (In Russian).
19. Korobkov I., Suvorova E., Sheynin Y., Olenev V. Streaming services over SpaceFibre networks. *Proceedings of 7th International SpaceWire Conference 2016*, Yokohama, 2016, pp. 151–158. doi:10.1109/SpaceWire.2016.7771621
20. Korobkov I. Adaptive data streaming service for onboard spacecraft networks. *Proceedings of 17th Conference of Open Innovations Association Finnish-Russian University Cooperation in Telecommunications (FRUCT) Program*, Yaroslavl, 2015, pp. 291–298.
21. National Aeronaut Administration (NASA). *Comparison of Communication Architectures for Spacecraft Modular Avionics Systems*. LLC-Create Space, 2018. 36 p.
22. 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. *Modern Problems of Radio Electronics*, 2018, pp. 196–200. Available at: http://efir.sfu-kras.ru/downloads/sbornik-spr-2018.pdf (accessed 03 February 2020) (In Russian).
23. SkiSys SSL/08717/DOC/001. User Requirements — SpaceWire Plug-and-Play Protocol. *Network Discovery Protocols*, 2012. 84 p.
24. Kopetz H. *Real-Time Systems. Design Principles for Distributed Embedded Applications*. Second ed. Springer US, 2011. 378 p. doi:10.1007/978-1-4419-8237-7
25. Dymov D., Sheynin Y., Olenev V. STP-ISS transport protocol application for SpaceFibre on-board networks. *2020 7th International Conference on Control, Decision and Information Technologies (CoDIT)*, Prague, 2020, pp. 914–919. doi:10.1109/CoDIT49905.2020.9263976
26. *D1.1 Consolidated set of Requirements for SpaceWire-RT*. Available at: http://spacewire-rt.org/Data/Docs/SpWRT_D1-1_v2-00.pdf (assessed 2 December 2020).

УДК 004.057.4
doi:10.31799/1684-8853-2021-1-8-16

Анализ требований к современным протоколам для бортовых сетей космических аппаратов

В. Л. Оленев[a], канд. техн. наук, доцент, orcid.org/0000-0002-1817-2754, Valentin.Olenev@guap.ru
[a]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: на смену устаревающим бортовым космическим сетям на базе шинных топологий приходят новые технологии, одной из которых является SpaceWire. Разрабатываются новые протоколы, расширяющие возможности SpaceWire. Необходима уверенность в том, что они будут обеспечивать все технические возможности для передачи и обработки данных на борту космических аппаратов. **Цель:** анализ существующих и разработка обобщенных требований к коммуникационным протоколам для бортовых космических сетей, которые позволят учитывать современные запросы космической индустрии. **Результаты:** в результате проведенного анализа сформирован набор консолидированных требований к протоколам физического–сетевого уровней и отдельно к протоколам транспортного уровня. Описаны требования, касающиеся скорости, задержек и расстояния передачи, объема передаваемой информации, функциональности для обнаружения неисправностей, синхронизации времени между устройствами, необходимых качеств сервиса и их свойств, основных пользовательских типов данных и режимов передачи данных на транспортном уровне. Существующие протоколы семейства SpaceWire выделены в отдельный класс протоколов, обладающих характеристиками, не присущими другим космическим протоколам. **Практическая значимость:** проведенный анализ позволит в значительной степени упростить процесс создания новых коммуникационных протоколов бортовых космических сетей, а также обеспечить необходимый уровень технологического оснащения космических аппаратов нового поколения.

Ключевые слова — бортовые космические сети, коммуникационные протоколы, технические требования, SpaceWire.

Для цитирования: Olenev V. L. Analysis of requirements for modern spacecraft onboard network protocols. *Информационно-управляющие системы*, 2021, № 1, с. 8–16. doi:10.31799/1684-8853-2021-1-8-16
For citation: Olenev V. L. Analysis of requirements for modern spacecraft onboard network protocols. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 8–16. doi:10.31799/1684-8853-2021-1-8-16

УДК 004.416
doi:10.31799/1684-8853-2021-1-17-27

Научные статьи
Articles

Автоматизированный подход к семантическому поиску по программной документации на основе алгоритма Doc2Vec

А. Д. Ковалев[a], *аспирант, ассистент, orcid.org/0000-0002-4610-5524, kov3000@ya.ru*
И. В. Никифоров[a], *канд. техн. наук, доцент, orcid.org/0000-0002-2330-2197*
П. Д. Дробинцев[a], *канд. техн. наук, доцент, orcid.org/0000-0003-1116-7765*
[a]*Санкт-Петербургский политехнический университет Петра Великого, Политехническая ул., 29, Санкт-Петербург, 195251, РФ*

Введение: *одним из значимых этапов жизненного цикла разработки программного обеспечения является этап его поддержки, когда заказчики могут обращаться в службу поддержки компании-поставщика с вопросами, проблемами и предложениями. Для решения поступившего запроса инженеры пользуются соответствующей документацией. С целью снизить трудоемкость и повысить качество этапа сопровождения можно автоматизировать поиск необходимых страниц, параграфов и предложений документации.* ***Цель:*** *разработка подхода к семантическому поиску по документации с использованием алгоритма машинного обучения Doc2Vec для автоматизации решения запросов заказчиков.* ***Результаты:*** *предложен подход к семантическому поиску по текстовым файлам документации и вики-страницам с использованием алгоритма машинного обучения Doc2Vec. Страницы документации, которые имеют семантическое сходство с текстовым описанием неразрешенного запроса заказчика, помогают разработчику более эффективно обрабатывать входящий запрос. На базе предложенного подхода разработан программный инструмент, предоставляющий инженеру отчет со ссылками на семантически близкие к нерешенному запросу страницы документации. Во время испытаний инструмента установлены оптимальные параметры алгоритма Doc2Vec, которые обеспечивают необходимое качество семантического поиска. Идея эксперимента заключалась в применении инструмента к нерешенным запросам и оценке его эффективности. Предложенный подход и реализующий его инструмент успешно протестированы на проекте с открытым исходным кодом Apache Kafka. В рамках эксперимента загружено и проанализировано 100 запросов из системы отслеживания ошибок Jira. Результаты эксперимента показывают преимущество использования инструмента в процессе поддержки программного продукта. Среднее время анализа документации сократилось по сравнению с традиционным ручным подходом.* ***Практическая значимость:*** *результаты исследований использованы при решении реальных запросов заказчиков. Разработанный подход и реализованное на его основе программное средство позволяют сократить трудоемкость этапа сопровождения.*

Ключевые слова *— поддержка программного обеспечения, автоматизация, Doc2Vec, машинное обучение, семантический поиск, документация.*

Для цитирования: Ковалев А. Д., Никифоров И. В., Дробинцев П. Д. Автоматизированный подход к семантическому поиску по программной документации на основе алгоритма Doc2Vec. *Информационно-управляющие системы*, 2021, № 1, с. 17–27. doi:10.31799/1684-8853-2021-1-17-27

For citation: Kovalev A. D., Nikiforov I. V., Drobintsev P. D. Automated approach to semantic search through software documentation based on Doc2Vec algorithm. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 17–27 (In Russian). doi:10.31799/1684-8853-2021-1-17-27

Введение

Этап сопровождения программного обеспечения (ПО) является одним из наиболее значимых и трудоемких этапов жизненного цикла разработки [1, 2]. В процессе сопровождения ПО разработчики компании-поставщика исправляют ошибки и решают проблемы, возникающие в ходе эксплуатации программного продукта на стороне заказчика. Например, причина неправильной работы ПО может быть связана с некорректной конфигурацией системы или с дефектом в коде программы. На данном этапе происходит прием и обработка запросов заказчика инженером технической поддержки или программистом. Запрос, как правило, написан на естественном, неформализованном языке и содержит описание проблемы в программном продукте. Для удобного хранения и управления такими запросами обычно используются системы отслеживания ошибок [3]. В данной работе рассмотрена система отслеживания ошибок Jira [4].

В процессе решения проблемы заказчика, как правило, необходимо обратиться к соответствующей документации или базе знаний. Документация может быть представлена в виде локальных текстовых файлов разных форматов (PDF, DOC, RTF и др.), а также содержаться на удаленных сайтах. Частный случай сайта с документацией — это вики-система Confluence. Именно ее мы и рассмотрим в данной работе.

Документация может содержать очень большой объем информации, и поиск необходимых страниц может занять много времени, что делает

процесс изучения документации трудоемким. Как правило, поиск в текстах документов осуществляется по ключевым словам. Однако данный подход не эффективен, если поисковый запрос становится слишком большим и начинает содержать больше пяти слов. Главным недостатком поиска по ключевым словам является невозможность определить синонимы к словам в поисковом запросе. Также в процессе поиска можно пропустить смысловые части документов, которые совпадают по семантике, но не содержат ключевые слова.

В решении задачи семантического поиска по документации могут помочь алгоритмы машинного обучения и нейронные сети [5], а именно алгоритм Doc2Vec [6]. За счет использования программного средства, основанного на алгоритме Doc2Vec, возможно снизить трудоемкость и повысить эффективность процесса сопровождения. Повышение качества сопровождения помогает выстроить долгосрочные отношения поставщика ПО с заказчиком.

Целью данной работы является сокращение трудоемкости поиска необходимой информации по документации за счет автоматизированного подхода, основанного на применении алгоритма Doc2Vec. Для достижения поставленной цели необходимо разработать программное средство, которое реализует предложенный подход, а также показать эффективность применения предложенного подхода и его реализации в программном средстве на актуальных данных.

Актуальность исследования обусловлена тем, что в процессе развития и усложнения программных продуктов увеличивается количество написанного исходного кода. Это неизбежно приводит к повышению числа дефектов и недоработок в ПО, что является причиной обращения заказчиков в службу поддержки компании-поставщика [7].

Обзор литературы

Существует множество исследований в области повышения эффективности сопровождения программных продуктов за счет внедрения современных методов работы с программной документацией.

Aghajani E. и др. [8] провели крупномасштабное эмпирическое исследование, в котором проанализировали и классифицировали 878 артефактов, имеющих отношение к программной документации. Результатом их работы явился подробный обзор проблем, связанных с документацией, и ряд действенных предложений как для исследователей, так и для практиков. Кроме того, рассмотрены решения, которые применяются при возникновении этих проблем.

На практике документация, как правило, обладает многочисленными проблемами, такими как недостаточное содержание и устаревшая неоднозначная информация. Чтобы противостоять этому, исследователи изучают разработку систем, которые автоматически генерируют высококачественную документацию. Liu M. и др. в статье [9] представляют инструмент OpenAPIDocGen2, который генерирует документацию на основе анализа исходного кода. Данный инструмент создает комбинированную документацию, которая включает в себя описания функций, директивы, концепции предметной области, а также примеры и сценарии использования.

Современная программная документация так же сложна, как и само ПО. В течение жизненного цикла документация накапливает множество «почти повторяющихся» фрагментов, т. е. фрагментов текста, которые были скопированы из одного источника и позже изменены различными способами. Такие дубликаты снижают качество документации и затрудняют ее дальнейшее использование. Luciv D. V. и др. [10] представляют алгоритм обнаружения дубликатов в программных документах, основанный на использовании адаптированного инструмента Clone Miner.

Для того чтобы эффективно пользоваться документацией, недостаточно правильно ее написать и оформить. Также необходимо обеспечить эффективный поиск по ней. Калиниченко А. В. [11] предлагает интерактивный метод поиска похожих текстовых документов, позволяющий повысить пертинентность поиска. Диалог с пользователем, используемый в методе, позволяет уточнить информационную потребность и построить более точный поисковый запрос.

Также существует множество исследований в области технологий семантического поиска по текстовым документам. Семантический поиск является актуальной областью исследований в сфере компьютерных наук [12].

Использование онтологий — один из подходов к семантическому поиску. Например, Kassim J. M. и Rahmany M. [13] предлагают семантическую поисковую систему, которая состоит из сканера онтологий, аннотатора онтологий, веб-сканера, модуля семантического поиска и обработчика запросов. Они используют онтологию для хранения структуры слов и создания информационных структур, связанных с доменом.

Другая группа подходов к поиску семантически похожих текстов в корпусе документов включает представление документов в виде числовых векторов. Существует множество техник численного представления документов, например, bag-of-words, TF-IDF (term frequency-inverse document frequency), латентное распределение

Дирихле (Latent Dirichlet Allocation — LDA) и алгоритмы машинного обучения.

Латентное распределение Дирихле можно в основном рассматривать как модель, которая разбивает набор документов на темы, представляя документ как смесь тем с их распределениями вероятностей.

Wei Xing и Croft W. [14] применяют алгоритм LDA для Ad-hoc поиска. Они предлагают модель документа на основе LDA в рамках фреймворка для языкового моделирования и оценивают ее на нескольких тестовых наборах данных TREC.

Ai Wang, Yao Dong Li и Wei Wang [15] предлагают метод межъязыкового поиска на основе LDA, который не полагается на пословный перевод запроса или документа.

Существуют исследовательские работы, в которых рассматривается алгоритм Doc2Vec. Например, Wang S. и Koopman R. [16] сравнивают подходы к векторному представлению документов Doc2Vec и Ariadne в контексте поиска информации. Эти подходы были протестированы на наборе документов, связанных с доказательной медициной. Однако результаты экспериментов показывают, что Ariadne работает так же хорошо, как и Doc2Vec в специфической задаче поиска информации.

Doc2Vec часто показывает низкую точность, если данные для обучения состоят из коротких предложений. Kurihara K. и др. [17] предлагают новый метод дополнения контекста коротких предложений для этапа обучения Doc2Vec. Этот метод использует идентификаторы целевой темы вместо идентификаторов предложений в качестве контекста. Они провели масштабный эксперимент на основе данных, связанных с обзорами фильмов, и доказали эффективность своего подхода.

Ни в одном из рассмотренных исследований не применяется подход, основанный на алгоритме Doc2Vec для семантического поиска в документации ПО на этапе сопровождения. Таким образом, отличительной чертой настоящего исследования является использование вышеуказанного алгоритма для выявления семантически связанных страниц документации.

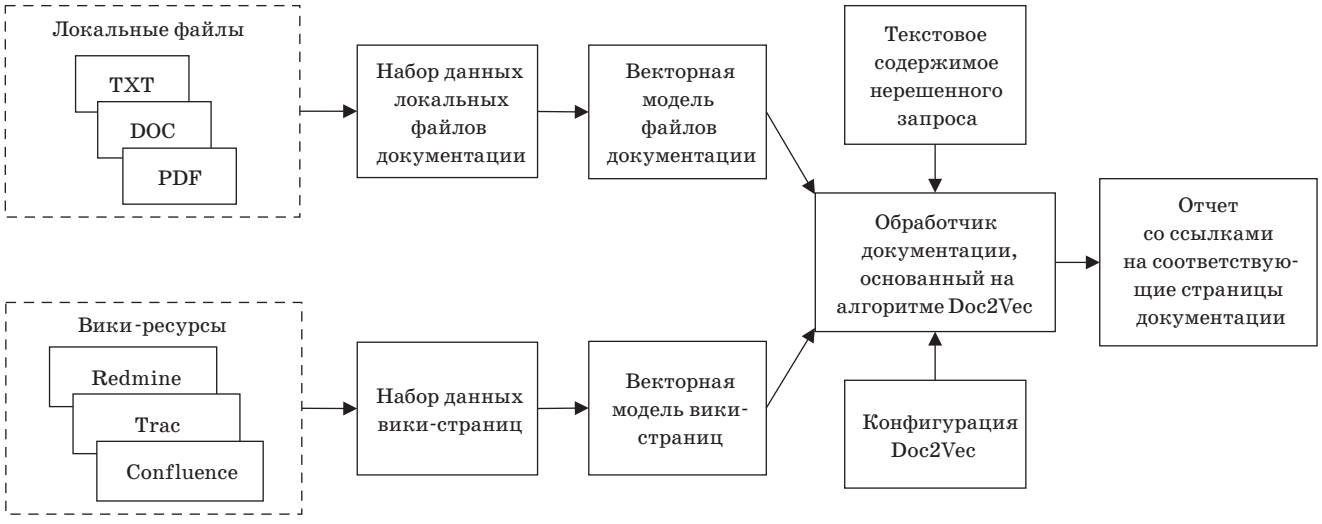
Предлагаемый подход

В данной работе предлагается снизить трудоемкость ручного анализа запросов заказчика за счет семантического поиска по документации. Суть ручного подхода состоит в том, чтобы получить список нерешенных запросов заказчика в системе отслеживания ошибок, а затем итеративно рассматривать и обрабатывать каждый запрос. Во время обработки запроса нередко приходится обращаться к документации программного продукта. Инженер-разработчик должен по ключевым словам искать конкретную информацию среди большого объема текстовых данных. Такой процесс неэффективен с точки зрения затрат времени, и его можно оптимизировать.

Концептуальная схема предлагаемого автоматизированного подхода к семантическому поиску по документации изображена на рис. 1.

На схеме видно, что предлагаемый подход к семантическому поиску по программной документации включает следующие этапы:

— создание наборов данных из локальных и удаленных ресурсов;



■ ***Рис. 1.*** Концептуальная схема предлагаемого подхода

■ ***Fig. 1.*** Conceptual schema of proposed approach

— обучение векторных моделей;

— применение алгоритма Doc2Vec к нерешенным запросам заказчика.

Существует два типа источников для создания наборов данных:

— локальные файлы документации различных форматов, таких как PDF, DOC (DOCX), RTF, TXT, PPT (PPTX) и т. д.;

— удаленные сетевые ресурсы, например вики-порталы (Confluence, Redmine, Trac и т. д.).

Структура набора данных следующая: каждая строка файла начинается с уникального идентификатора, который указывает на определенную страницу конкретного документа, затем идет разделяющий символ (например, «|»), а затем весь текстовый контент с конкретной страницы.

В дополнение к файлу набора данных создается файл метаданных. Он состоит из строк, которые разделены на три столбца: уникальный идентификатор страницы, затем имя документа, затем номер страницы в этом документе.

Файл метаданных необходим для того, чтобы после нахождения уникального идентификатора страницы было понятно, из какого документа и с какой страницы конкретный текст.

Затем для каждого набора данных создается векторная модель. После создания векторные модели используются алгоритмом Doc2Vec для поиска в документах страниц, которые семантически похожи на текстовое содержимое нерешенной проблемы.

В результате инструмент генерирует HTML-отчет, который содержит ссылки на конкретные страницы в локальных документах и на вики-сайтах.

Получив отчет с результатами, пользователь имеет возможность изучить соответствующие части документации, которые помогут быстро решить проблему заказчика. Нет необходимости искать нужную страницу в документе или в вики-системе.

Подводя итог, предлагаемый подход можно выразить формулой

$$U(L(T), I_i) = [Rate(t_0) \dots Rate(t_N)],$$

где $U(x, y)$ — функция применения векторной модели к текстовым данным, в результате чего получается массив наиболее похожих страниц из документации; $L(x)$ — функция обучения, которая применяется на текстовых наборах данных и результатом работы которой является векторная модель; T — массив текстовых данных из документации; I_i — текстовые данные, описывающие конкретный запрос заказчика; $Rate(x)$ — число, которое выражается в процентах и означает семантическое сходство t_i и I_i; t_i — одна из похожих по смыслу найденных страниц документации; N — количество найденных похожих страниц документации.

Процедура нахождения похожих текстов

Семантическая близость двух текстов определяется косинусным сходством их векторов. Косинусное сходство [18] — это мера, которая измеряет косинус угла между двумя ненулевыми векторами. Данная метрика определяет расположение одного вектора в пространстве относительно другого вектора. Два вектора с одинаковой ориентацией в пространстве имеют косинус-сходство 1, а два вектора, располагающиеся под углом 90° относительно друг друга, имеют сходство 0. Два диаметрально противоположных вектора имеют сходство –1. Косинус-сходство в основном используется в положительном векторном пространстве, в котором результат ограничен в пределах [0, 1]. Единичные векторы максимально «схожи», если они параллельны, и максимально «несхожи», если они ортогональны (перпендикулярны).

Коэффициент схожести *similarity* вычисляется по формуле

$$similarity = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}},$$

где A_i и B_i — компоненты векторов **A** и **B** соответственно.

Мера косинусного сходства применима для векторных пространств с любым числом измерений. Данная мера наиболее часто используется в многомерных положительных пространствах.

При векторном представлении текстовых данных каждое слово или документ сопоставляется со своим уникальным вектором. Косинусное сходство между двумя векторами слов или документов показывает вероятность семантической схожести этих двух слов или документов.

В предлагаемом подходе автоматизированного анализа запросов заказчика в качестве меры схожести двух текстов предлагается использовать косинусное сходство. Будем считать, что два текста являются похожими по смыслу, если косинусное сходство их векторных представлений больше или равно коэффициенту *threshold*. Данный коэффициент можно настраивать. Его увеличение может быть обусловлено желанием находить меньшее количество максимально схожих текстовых документов. В то же время уменьшение данного коэффициента приведет к большему количеству результатов поиска среди семантически схожих документов. В ходе оптими-

зации гиперпараметров алгоритма Doc2Vec было вычислено оптимальное значение коэффициента *threshold*, которое составило 0,8.

Реализация программного инструмента

Разработанный инструмент, реализующий предложенный автоматизированный подход, написан на языке Java 8 и включает несколько модулей:

— два коннектора к удаленным веб-ресурсам (Jira, Confluence);

— два обработчика (обработчик локальной документации и Confluence);

— генератор HTML-отчета;

— исполнитель обработчиков, который координирует работу всей системы.

Архитектура инструмента показана на рис. 2.

Коннекторы к Jira и Confluence разработаны и использованы для загрузки данных с веб-ресурсов Jira и Confluence. Подключение к этим сервисам осуществляется с помощью соответствующих интерфейсов REST API. Возможна анонимная или базовая аутентификация с использованием имени пользователя и пароля.

Jira-коннектор получает текстовое описание нерешенных запросов от Jira. Коннектор Confluence получает текстовое представление вики-страниц.

Библиотека Jira Rest Java Client (JRJC) [https://bitbucket.org/atlassian/jira-rest-java-client] используется в коннекторе Jira. Эта Java-библиотека позволяет подключаться к любому экземпляру Jira 4.2+ с помощью REST API. В настоящее время JRJC предоставляет тонкий слой абстракции поверх REST API, а также объектную модель Jira на стороне клиента. Эти объекты представляют сущности запроса: *Issue*, *Priority*, *Resolution*, *Status*, *User* и т. д.

Библиотека Confluence Rest Java Client (CRJC) [https://docs.atlassian.com/atlassian-confluence/5.9.2/] используется в коннекторе Confluence. Как и JRJC, CRJC обеспечивает тонкий слой абстракции поверх REST API.

Каждый обработчик отвечает за три процесса:

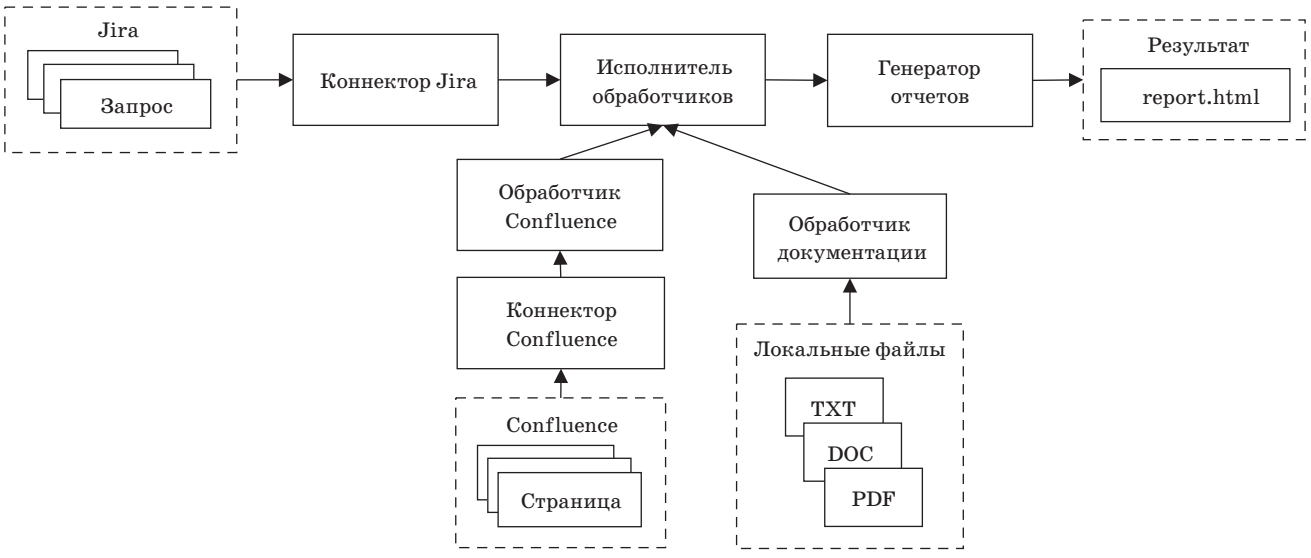
— создает структурированный набор данных из необработанных источников;

— создает векторную модель на основе полученного набора данных;

— использует созданную векторную модель для поиска страниц документации, которые связаны по смыслу с нерешенным запросом.

Обработчик документации использует локальные текстовые файлы для создания файла *Documentation_DataSet.txt*, а обработчик Confluence использует текстовое содержимое с вики-страниц для создания файла *Confluence_DataSet.txt*.

Структура набора данных для обработчика документации следующая: каждая строка файла начинается с уникального идентификатора, который указывает на конкретную страницу конкретного документа, затем идет символ разделительной вертикальной черты, а затем все текстовое содержимое с определенной страницы. Формат созданного файла набора данных можно увидеть на рис. 3, *а*. В дополнение к файлу набора данных создается файл метаданных (рис. 3, *б*).



■ ***Рис. 2.*** Архитектура программного инструмента

■ ***Fig. 2.*** The software tool architecture

а)
```
doc_1|Important Notice 2010 2020 Cloudera Inc All rights reserved Cl
ices processes or other information by trade name trademark manufact
in this document is subject to change without notice Cloudera shall
doc_2|Table of Contents Apache Kafka Guide 8 Ideal Publish Subscribe
doc_3|Single Cluster Scenarios 26 Leader Positions 26 In Sync Replic
Using Kafka with Apache Flume 45 Using Kafka with Apache Spark Strea
doc_4|Kafka Administration 60 Kafka Administration Basics 60 Broker
kens 78 Delegation Token Basics 79 Broker Configuration Settings 79
doc_5|Kafka Performance Tuning 86 Tuning Brokers 86 Tuning Producers
doc_6|Appendix Apache License Version 2 0 182
doc_7|Apache Kafka Guide Apache Kafka is a streaming message platfor
it of Unlimited Lookback A new Subscriber A1 can read Publisher A s
igh availability but different semantics and message delivery guaran
doc_8|The next few sections provide an overview of some of the more
messages sent to the topics and serves consumer requests Apache Kafk
```

б)
```
doc_0 cloudera-kafka-guide.pdf 1
doc_1 cloudera-kafka-guide.pdf 2
doc_2 cloudera-kafka-guide.pdf 3
doc_3 cloudera-kafka-guide.pdf 4
doc_4 cloudera-kafka-guide.pdf 5
doc_5 cloudera-kafka-guide.pdf 6
doc_6 cloudera-kafka-guide.pdf 7
doc_7 cloudera-kafka-guide.pdf 8
doc_8 cloudera-kafka-guide.pdf 9
doc_9 cloudera-kafka-guide.pdf 10
doc_10 cloudera-kafka-guide.pdf 11
```

■ ***Рис. 3.*** Часть файла набора данных (*а*) и метаданных (*б*) локальной документации
■ ***Fig. 3.*** Part of the documentation data set (*а*) and metadata (*б*) file

Структура набора данных для Confluence аналогична набору данных для файлов документации, но отличие состоит в том, что уникальный идентификатор каждой строки напрямую указывает на конкретный идентификатор вики-страницы, и в этом случае нет необходимости создавать файл метаданных.

Основой обработчиков является алгоритм Doc2Vec, который реализован в Java-библиотеке Deeplearning4j [https://deeplearning4j.org/]. Работа с алгоритмом делится на три этапа: подготовка данных, обучение и использование.

До передачи содержимого **_DataSet.txt* файлов на вход алгоритма Doc2Vec необходимо подготовить текстовые данные для создания векторной модели. Процесс подготовки состоит из токенизации [19], стемминга [20] и удаления стоп-слов. Программный инструмент использует популярную реализацию стеммера Портера [20, 21] для языка Java.

Файлы наборов данных, сгенерированные обработчиками документации и содержащие предварительно обработанные тексты страниц, используются для обучения. В результате обучения получается векторная модель, которая впоследствии сериализуется в файлы *Documentation_VectorModel.zip* и *Confluence_VectorModel.zip* соответственно. Векторы представляют численное значение «смысла» запросов, и, используя математические операции над векторами, можно найти сходство между различными запросами.

Все настройки для обучения содержатся в конфигурационном файле *doc2vec.properties*. Этот файл содержит следующие поля: 1) *minWordFrequency* — минимальная частота слов в наборе обучающих данных; 2) *iterations* — количество итераций обучения, выполненных для каждой части тренировочного корпуса; 3) *epochs* — количество итераций (эпох) по всему учебному корпусу; 4) *layerSize* — размер выходных векторов; 5) *learningRate* — начальная скорость обучения модели; 6) *windowSize* — размер окна контекста; 7) *sampling* — численная характеристика подвыборки [22]; 8) *threshold* — пороговое значение косинусного сходства.

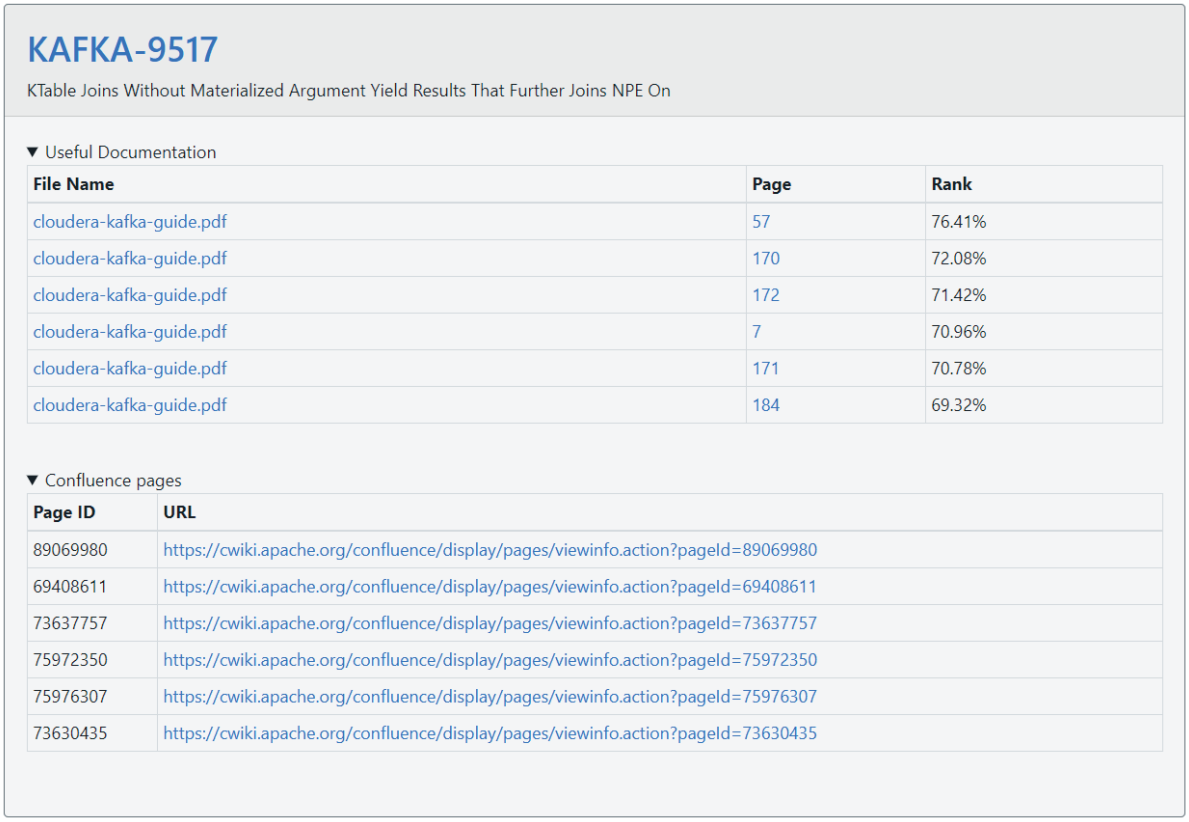
Архив *VectorModel.zip* содержит следующие файлы: *codes.txt* — коды для дерева Хаффмана [23]; *config.json* — настройки алгоритма Doc2Vec; *frequencies.txt* — метрики tf-idf [24] и bag-of-words [25]; *huffman.txt* — координаты точек дерева Хаффмана; *labels.txt* — список идентификаторов страниц документации в формате base64; *syn0.txt* — веса связей между входными и скрытыми слоями нейронной сети; *syn1.txt* — веса связей между скрытыми и выходными слоями нейронной сети.

После обучения алгоритмом Doc2Vec модель можно использовать для нахождения семантически близких векторов документов. Для этого в память загружается модель из файла **_VectorModel.zip*, в котором каждая страница представлена в виде числового вектора и связана с определенным идентификатором страницы в локальном документе или на вики-сайте. Затем текстовое содержимое неразрешенных запросов отправляется на вход алгоритма Doc2Vec.

Процесс поиска семантически связанной документации заключается в следующем. Сначала формируется числовой вектор из поступающего неразрешенного запроса. Затем этот вектор сравнивается с векторами страниц файлов документации или вики-страниц. В этом случае сходство текстовых данных определяется коэффициентом косинусной близости их векторных представлений. Чем больше значение коэффициента, тем более уверенно можно утверждать, что оба текста похожи.

Таким образом, результатом работы алгоритма является список с идентификаторами страниц документации, наиболее похожих на входные неразрешенные запросы. Все эти данные затем собираются в отчет и предоставляются пользователю (рис. 4).

Модуль исполнителя обработчиков координирует работу всех этих компонентов. Сначала он запускает коннектор Jira, получает список нерешенных запросов, а затем итеративно отправляет



KAFKA-9517

KTable Joins Without Materialized Argument Yield Results That Further Joins NPE On

▼ Useful Documentation

File Name	Page	Rank
cloudera-kafka-guide.pdf	57	76.41%
cloudera-kafka-guide.pdf	170	72.08%
cloudera-kafka-guide.pdf	172	71.42%
cloudera-kafka-guide.pdf	7	70.96%
cloudera-kafka-guide.pdf	171	70.78%
cloudera-kafka-guide.pdf	184	69.32%

▼ Confluence pages

Page ID	URL
89069980	https://cwiki.apache.org/confluence/display/pages/viewinfo.action?pageId=89069980
69408611	https://cwiki.apache.org/confluence/display/pages/viewinfo.action?pageId=69408611
73637757	https://cwiki.apache.org/confluence/display/pages/viewinfo.action?pageId=73637757
75972350	https://cwiki.apache.org/confluence/display/pages/viewinfo.action?pageId=75972350
75976307	https://cwiki.apache.org/confluence/display/pages/viewinfo.action?pageId=75976307
73630435	https://cwiki.apache.org/confluence/display/pages/viewinfo.action?pageId=73630435

■ ***Рис. 4.*** Отчет, содержащий ссылки на соответствующую входящему запросу документацию
■ ***Fig. 4.*** Final report with related documentation references

каждый запрос обоим обработчикам. Результаты обработки отправляются в генератор отчетов, который создает файл *report.html*.

Оценка точности модели Doc2Vec

Оптимизация гиперпараметров для обучающего алгоритма Doc2Vec проводилась методом случайного поиска (Random Search) с помощью программного пакета Deeplearning4j.

Для оптимизации сложных моделей (с более чем несколькими гиперпараметрами) случайный поиск превосходит по скорости и качеству другие классические методы, такие как поиск по решетке (Grid Search) и байесовскую оптимизацию [26].

Случайный поиск дополнительно сопровождался перекрестной проверкой на тренировочном наборе данных. Использование перекрестной проверки при оптимизации гиперпараметров позволяет произвести оценку эффективности выбранной модели с наиболее равномерным использованием имеющихся данных.

Тестовый набор данных представлен текстовым файлом, состоящим из множества строк. Нами рассматривается тестовый файл с количеством строк, равным 100. Каждая строка является тестовым шагом и разбита на три столбца специальным разделяющим символом. В первом столбце находится текстовый фрагмент исходного запроса. Во втором столбце находится текст близкого по смыслу запроса. А в третьем столбце, наоборот, находится текст запроса, который совершенно не связан с исходным.

Задача тестирования — определить, сколько строчек из 100 будут вычислены удачно. Вычисление заключается в следующем. Если алгоритм определяет, что первые два столбца похожи на более чем 80 %, а первый и третий столбец похожи менее чем на 20 %, то считаем, что данная строчка вычислена правильно. После вычисления всех строк суммируем количество правильных вычислений и делим на количество строк. Полученное число показывает точность векторной модели.

В результате тестирования и выбора гиперпараметров Doc2Vec выяснилось, что следующая конфигурация является оптимальной по качеству и времени выполнения алгоритма: *minWordFrequency* = 1; *iterations* = 5; *epochs* = 12; *layerSize* = 100; *learningRate* = 0,025; *windowSize* = = 5; *sampling* = 0, *threshold* = 0,8. При этом точ-

ность алгоритма на тестовом наборе данных составила 89,96 %.

Описание эксперимента

Разработанный инструмент был применен на проекте с открытым исходным кодом Apache Kafka [https://kafka.apache.org]. Любой инженер может найти ошибку в работе этой программы и зарегистрировать запрос, содержащий вопрос или описание ошибки, в соответствующей системе отслеживания ошибок Jira. Для эксперимента было получено 100 запросов из Jira.

Файл Apache Kafka Guide.pdf [https://docs.cloudera.com/documentation/enterprise/6/latest/PDF/cloudera-kafka.pdf] использовался в качестве локального файла документации. Confluence-пространство проекта Kafka [https://cwiki.apache.org/confluence/display/KAFKA] также использовалось в качестве удаленного вики-ресурса.

Инструмент развернут на платформе, которая является локальной рабочей станцией с операционной системой Windows 10 Enterprise x64, процессором Intel Core i7-4810MQ с тактовой частотой 2,80 ГГц и 16 ГБ оперативной памяти.

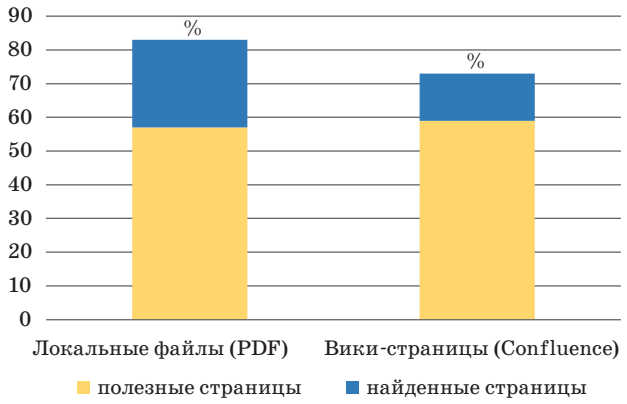
Файл *Documentation_DataSet.txt* (235 КБ) создан из PDF-файла документации. Количество строк в этом наборе данных равно количеству страниц в файле PDF и составляет 184 страницы. Количество строк в наборе данных *Confluence_DataSet.txt* (5,21 МБ) равно 765, что соответствует количеству загруженных вики-страниц. Затем из наборов данных были созданы векторные модели *Documentation_VectorModel.zip* (3,19 МБ) и *Confluence_VectorModel.zip* (35,5 МБ) соответственно. Создание наборов данных заняло 19 мин, а создание моделей — 12 мин.

В процессе работы над проектом Apache Kafka был проведен эксперимент, суть которого заключается в применении инструмента к нерешенным запросам и оценке его эффективности. Необходимо понять, в каком проценте случаев инструмент находит хотя бы одну соответствующую страницу в документации, а также в скольких случаях найденная страница является корректной и полезной для решения проблемы. Страница считается найденной, если семантическая близость между текстом запроса и текстом страницы составляет 80 %. В то же время найденная страница считается полезной, если она действительно может помочь в решении проблемы.

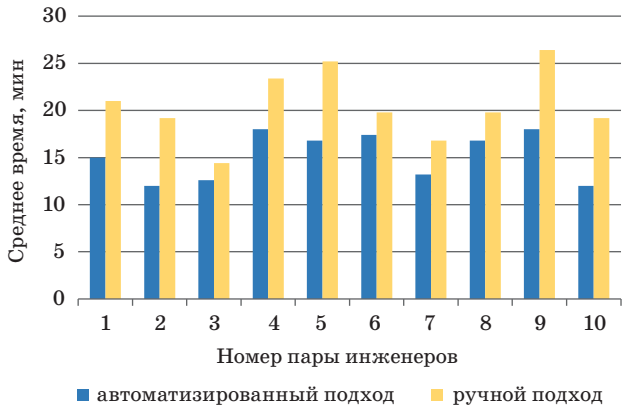
В рамках эксперимента было загружено и проанализировано 100 запросов из системы отслеживания ошибок Jira. Результаты анализа этих запросов представлены на рис. 5.

На диаграмме показано количество найденных и полезных страниц в документации. Были найдены 83 связанные страницы из локальных файлов, 57 из них оказались полезными для решения проблемы. Это означает, что точность алгоритма для локальных файлов составляет 69 %. Что касается вики-ресурсов, были найдены 73 связанные страницы, 59 из них оказались полезными для решения запроса. Точность в данном случае составляет 81 %.

Во второй части эксперимента сравнивалась эффективность ручного и автоматизированного подходов к поиску полезной документации. Для получения более уверенных результатов необходимо было провести этот эксперимент с большим количеством людей. Поэтому были приглашены 10 пар разработчиков для обработки 100 запросов. Во всех парах, участвовавших в эксперименте, оба инженера имели одинаковый опыт работы и квалификацию. Первый разработчик пытался



■ ***Рис. 5.*** Результаты эксперимента
■ ***Fig. 5.*** The experiment results



■ ***Рис. 6.*** Среднее время, затраченное на поиск полезной документации по одному запросу при ручном и автоматизированном подходе
■ ***Fig. 6.*** Average time spent to search for useful documentation for one request with the manual and automated approach

решить задачу вручную, а второй — в автоматизированном режиме.

Для сравнения времени поиска полезных страниц документации по каждому запросу использовался обычный таймер. Таймер включался, когда инженер начинал работу над новым запросом, и выключался, когда инженер находил хотя бы одну страницу документации, которая оказалась полезной для решения запроса. Результаты эксперимента показаны на рис. 6.

На диаграмме видно, что автоматизированный подход требует значительно меньше времени, чем ручной. Среднее время, затраченное на поиск подходящих страниц документации при ручном и автоматизированном подходе, составило 15,2 и 20,5 мин соответственно, т. е. снижение времени составило 25,9 %.

Заключение

В работе рассмотрены существующие исследования в области семантического поиска в текстовых документах.

Предложен автоматизированный подход для снижения трудоемкости обработки запросов заказчиков. Этот подход основан на использовании алгоритма машинного обучения Doc2Vec, который решает проблему семантического поиска в документации к программному продукту.

Созданный инструмент был успешно протестирован на проекте Apache Kafka — было проанализировано 100 запросов. Показаны эффективность и преимущества его использования. Среднее время анализа документации сократилось по сравнению с традиционным ручным подходом на 25,9 %.

Литература

1. **Shylesh S.** A study of software development life cycle process models. *National Conference on Reinventing Opportunities in Management, IT, and Social Sciences*, 2017, pp. 534–541.
2. **Ogheneovo E. E.** On the relationship between software complexity and maintenance costs. *Journal of Computer and Communications*, 2014, vol. 2, no. 14, p. 1.
3. **Noei E., Zhang F., Wang S., Zou Y.** Towards prioritizing user-related issue reports of mobile applications. *Empirical Software Engineering*, 2019, vol. 24, no. 4, pp. 1964–1996.
4. **Filion L., Daviot N., Le Bel J., Gagnon M.** Using Atlassian tools for efficient requirements management: An industrial case study. *IEEE International Systems Conference*, April 2017, pp. 1–6.
5. **Pellegrini T.** Comparing SVM, Softmax, and shallow neural networks for eating condition classification. *16th Annual Conference of the International Speech Communication Association*, 2015, pp. 899–903.
6. **Maslova N., Potapov V.** Neural network Doc2Vec in automated sentiment analysis for short informal texts. *Lecture Notes in Computer Science*, 2017, vol. 10458, pp. 546–554.
7. **Kovalev A., Voinov N., Nikiforov I.** Using the Doc2Vec algorithm to detect semantically similar jira issues in the process of resolving customer requests. *Intelligent Distributed Computing XIII*, 2020, pp. 96–101. doi:10.1007/978-3-030-32258-8_11
8. **Aghajani E., Nagy C., Vega-Marquez O., Linares-Vasquez M., Moreno L., Bavota G., Lanza M.** Software documentation issues unveiled. *IEEE/ACM 41st International Conference on Software Engineering*, Montreal, QC, Canada, 2019, pp. 1199–1210.
9. **Liu M., Peng X., Meng X., Xu H., Xing S., Wang X., Liu Y., Lv G.** Source code based on-demand class documentation generation. *IEEE International Conference on Software Maintenance and Evolution*, Adelaide, Australia, 2020, pp. 864–865.
10. **Luciv D. V., Koznov D. V., Chernishev G. A., Terekhov A. N., Romanovsky K. Yu., Grigoriev D. A.** Detecting near duplicates in software documentation. *Program Comput Soft 44*, 2018, pp. 335–343.
11. **Калиниченко А. В.** Диалоговый метод автоматизации поиска семантически похожих документов. *Вестник ВГТУ*, Воронеж, 2012, т. 8, № 8, с. 15–17.
12. **Ensan F., Bagheri E.** Document retrieval model through semantic linking. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, February 2017, pp. 181–190.
13. **Kassim J. M., Rahmany M.** Introduction to semantic-search engine. *International Conference on Electrical Engineering and Informatics,* August 2009, vol. 2, pp. 380–386.
14. **Wei Xing, Croft W.** LDA-based document models for Ad-hoc retrieval. *Proceedings of the 29th Annual International ACM SIGIR*, 2006, pp. 178–185. doi:10.1145/1148170.1148204
15. **Ai Wang, Yao Dong Li, Wei Wang.** Cross language information retrieval based on LDA. *IEEE International Conference on Intelligent Computing and Intelligent Systems*, Shanghai, 2009, pp. 485–490. doi:10.1109/ICICISYS.2009.5358121
16. **Wang S., Koopman R.** Semantic embedding for information retrieval. *BIR@ECIR*, 2017, pp. 122–132.
17. **Kurihara K., Shoji Y., Fujita S., Durst M. J.** Target-topic aware Doc2Vec for short sentence retrieval from user generated content. *Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services*, 2019, pp. 463–467.
18. **Gunawan D., Sembiring C. A., Budiman M. A.** The implementation of cosine similarity to calculate text relevance between two documents. *Journal of Physics:*

Conference Series, March 2018, vol. 978, no. 1, article id. 012120.
19. **Ковалев А. Д., Никифоров И. В., Дробинцев П. Д.** Интеллектуальная обработка запросов заказчика на этапе сопровождения программного продукта. *Неделя науки СПбПУ: материалы научной конференции с международным участием*, Санкт-Петербург, 19–24 ноября 2018 г. СПб., 2019, с. 165–168.
20. **Farrar D., Hayes J. H.** A comparison of stemming techniques in tracing. *IEEE/ACM 10th International Symposium on Software and Systems Traceability (SST)*, May 2019, pp. 37–44.
21. **Литвинов М. Б., Воинов Н. В.** Система интеллектуального анализа заявок пользователей телекоммуникационных услуг. *Современные технологии в теории и практике программирования: сборник материалов конференции*, Санкт-Петербург, 23 апреля 2020 г. СПб., 2020, с. 159.
22. **Ji S., Satish N., Li S., Dubey P. K.** Parallelizing word2vec in shared and distributed memory. *Transactions on Parallel and Distributed Systems*, 2019, vol. 30, no. 9, pp. 2090–2100.
23. **Rahman M., Hamada M.** Burrows – Wheeler transform based lossless text compression using keys and huffman coding. *Symmetry*, 2020, vol. 12, no. 10, p. 1654.
24. **Qaiser S., Ali R.** Text mining: use of TF-IDF to examine the relevance of words to documents. *International Journal of Computer Applications*, 2018, vol. 181, no. 1, pp. 25–29.
25. **Ayadi W., Elhamzi W., Charfi I., Atri M.** A hybrid feature extraction approach for brain MRI classification based on Bag-of-words. *Biomedical Signal Processing and Control*, 2019, vol. 48, pp. 144–152.
26. **Bergstra J., Bengio Y.** Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 2012, vol. 13, no. 10, pp. 281–305.

UDC 004.416
doi:10.31799/1684-8853-2021-1-17-27

Automated approach to semantic search through software documentation based on Doc2Vec algorithm

A. D. Kovalev[a], Post-Graduate Student, Assistant Professor, orcid.org/0000-0002-4610-5524, kov3000@ya.ru
I. V. Nikiforov[a], PhD, Tech., Associate Professor, orcid.org/0000-0002-2330-2197
P. D. Drobintsev[a], PhD, Tech., Associate Professor, orcid.org/0000-0003-1116-7765
[a]Peter the Great St.Petersburg Polytechnic University, 29, Polytechnicheskaya St., 195251, Saint-Petersburg, Russian Federation

Introduction: An important stage in a software development life cycle is the support phase, when customers can contact the support service of the supplier company and request a solution to an issue encountered in the software. To solve the request, engineers often have to refer to the relevant documentation. In order to reduce the complexity of the maintenance phase, the search for the necessary documentation pages can be automated. **Purpose:** Development of an approach to semantic search through documentation using Doc2Vec machine learning algorithm in order to automate the solution of customer requests. **Results:** An approach is proposed to semantic search through text documentation files and wiki pages using Doc2Vec machine learning algorithm. The documentation pages with semantic similarities to the textual description of an unresolved customer request help the engineer to process the request more efficiently and rapidly. Based on the proposed approach, a software tool has been developed which provides the engineer with a report containing links to documentation pages semantically related to the unresolved request. During the configuration of this tool, the optimal parameters of the Doc2Vec algorithm were found, providing the necessary quality of the semantic search. The idea of the experiment was to apply the tool to unresolved requests and evaluate its effectiveness. The developed approach and software tool were successfully tested in an open source Apache Kafka project. In the course of the experiment, 100 requests from Jira bug tracking system were downloaded and analyzed. The experimental results show the advantage of using the tool in software product support. The average documentation analysis time has been reduced as compared to the traditional manual approach. **Practical relevance:** The research results were used to solve real customer requests. The developed approach and the software implemented on its basis can reduce the complexity of the maintenance phase.

Keywords — software maintenance, automation, Doc2Vec, machine learning, semantic search, documentation.

For citation: Kovalev A. D., Nikiforov I. V., Drobintsev P. D. Automated approach to semantic search through software documentation based on Doc2Vec algorithm. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 17–27 (In Russian). doi:10.31799/1684-8853-2021-1-17-27

References

1. Shylesh S. A study of software development life cycle process models. *National Conference on Reinventing Opportunities in Management, IT, and Social Sciences*, 2017, pp. 534–541.
2. Ogheneovo E. E. On the relationship between software complexity and maintenance costs. *Journal of Computer and Communications*, 2014, vol. 2, no. 14, p. 1.
3. Noei E., Zhang F., Wang S., Zou Y. Towards prioritizing user-related issue reports of mobile applications. *Empirical Software Engineering*, 2019, vol. 24, no. 4, pp. 1964–1996.
4. Filion L., Daviot N., Le Bel J., Gagnon M. Using Atlassian tools for efficient requirements management: An industrial case study. *IEEE International Systems Conference*, April 2017, pp. 1–6.
5. Pellegrini T. Comparing SVM, Softmax, and shallow neural networks for eating condition classification. *16th Annual Conference of the International Speech Communication Association*, 2015, pp. 899–903.
6. Maslova N., Potapov V. Neural network Doc2Vec in automated sentiment analysis for short informal texts.

Lecture Notes in Computer Science, 2017, vol. 10458, pp. 546–554.
7. Kovalev A., Voinov N., Nikiforov I. Using the Doc2Vec algorithm to detect semantically similar jira issues in the process of resolving customer requests. *Intelligent Distributed Computing XIII*, 2020, pp. 96–101. doi:10.1007/978-3-030-32258-8_11
8. Aghajani E., Nagy C., Vega-Marquez O., Linares-Vasquez M., Moreno L., Bavota G., Lanza M. Software documentation issues unveiled. *IEEE/ACM 41st International Conference on Software Engineering*, Montreal, QC, Canada, 2019, pp. 1199–1210.
9. Liu M., Peng X., Meng X., Xu H., Xing S., Wang X., Liu Y., Lv G. Source code based on-demand class documentation generation. *IEEE International Conference on Software Maintenance and Evolution*, Adelaide, Australia, 2020, pp. 864–865.
10. Luciv D. V., Koznov D. V., Chernishev G. A., Terekhov A. N., Romanovsky K. Yu., Grigoriev D. A. Detecting near duplicates in software documentation. *Program Comput Soft 44*, 2018, pp. 335–343.
11. Kalinichenko A. V. A dialogue technique for automation semantically similar documents search. *Vestnik VGTU*, Voronezh, 2012, vol. 8, no. 8, pp. 15–17 (In Russian).
12. Ensan F., Bagheri E. Document retrieval model through semantic linking. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, February 2017, pp. 181–190.
13. Kassim J. M., Rahmany M. Introduction to semanticsearch engine. *International Conference on Electrical Engineering and Informatics,* August 2009, vol. 2, pp. 380–386.
14. Wei Xing, Croft W. LDA-based document models for Ad-hoc retrieval. *Proceedings of the 29th Annual International ACM SIGIR*, 2006, pp. 178–185. doi:10.1145/1148170.1148204
15. Ai Wang, Yao Dong Li, Wei Wang. Cross language information retrieval based on LDA. *IEEE International Conference on Intelligent Computing and Intelligent Systems*, Shanghai, 2009, pp. 485–490. doi:10.1109/ICICISYS.2009.5358121
16. Wang S., Koopman R. Semantic embedding for information retrieval. *BIR@ECIR*, 2017, pp. 122–132.
17. Kurihara K., Shoji Y., Fujita S., Durst M. J. Target-topic aware Doc2Vec for short sentence retrieval from user generated content. *Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services*, 2019, pp. 463–467.
18. Gunawan D., Sembiring C. A., Budiman M. A. The implementation of cosine similarity to calculate text relevance between two documents. *Journal of Physics: Conference Series*, March 2018, vol. 978, no. 1, article id. 012120.
19. Kovalev A. D., Nikiforov I. V., Drobintsev P. D. Intelligent processing of customer requests at the stage of software product maintenance. *Materialy nauchnoy konferentsii s mezhdunarodnym uchastiyem "Nedelya nauki SPbPU"* [Proc. Int. Conf. "SPbPU Science Week"], Saint-Petersburg, 2018, pp. 165–168 (In Russian).
20. Farrar D., Hayes J. H. A comparison of stemming techniques in tracing. *IEEE/ACM 10th International Symposium on Software and Systems Traceability (SST)*, May 2019, pp. 37–44.
21. Litvinov M. B., Voinov N. V. System for intellectual analysis of requests from telecommunication services users. *Sbornik materialov konferentsii "Sovremennyye tekhnologii v teorii i praktike programmirovaniya"* [Proc. Conf. "Modern technologies in the theory and practice of programming"], Saint-Petersburg, 2020, p. 159 (In Russian).
22. Ji S., Satish N., Li S., Dubey P. K. Parallelizing Word2Vec in shared and distributed memory. *Transactions on Parallel and Distributed Systems*, 2019, vol. 30, no. 9, pp. 2090–2100.
23. Rahman M., Hamada M. Burrows – Wheeler transform based lossless text compression using keys and huffman coding. *Symmetry*, 2020, vol. 12, no. 10, p. 1654.
24. Qaiser S., Ali R. Text mining: use of TF-IDF to examine the relevance of words to documents. *International Journal of Computer Applications*, 2018, vol. 181, no. 1, pp. 25–29.
25. Ayadi W., Elhamzi W., Charfi I., Atri M. A hybrid feature extraction approach for brain MRI classification based on Bag-of-words. *Biomedical Signal Processing and Control*, 2019, vol. 48, pp. 144–152.
26. Bergstra J., Bengio Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 2012, vol. 13, no. 10, pp. 281–305.

УДК 004.056
doi:10.31799/1684-8853-2021-1-28-37

Научные статьи
Articles

Анализ моделей глубокого обучения для задач обнаружения сетевых аномалий интернета вещей

Д. А. Гайфулина[a], *младший научный сотрудник, orcid.org/0000-0002-5266-8649*
И. В. Котенко[a], *доктор техн. наук, профессор, orcid.org/0000-0001-6859-7120, ivkote@comsec.spb.ru*
[a]*Санкт-Петербургский федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ*

Введение: *актуальность решения задачи выбора моделей глубокого обучения для обнаружения аномалий в сетевом трафике интернета вещей связана с необходимостью анализировать большое число событий безопасности для выявления аномального поведения умных устройств. Мощной технологией анализа таких данных является машинное и, в частности, глубокое обучение.* ***Цель:*** *выработка рекомендаций по выбору моделей глубокого обучения для обнаружения аномалий в сетевом трафике интернета вещей.* ***Результаты:*** *проведен сравнительный анализ моделей глубокого обучения и предоставлены рекомендации по их использованию для обнаружения аномалий в сетевом трафике интернета вещей. В качестве базовых моделей глубокого обучения рассмотрены многослойный персептрон, сверточная нейронная сеть, рекуррентная нейронная сеть, блок долгой краткосрочной памяти, управляемый рекуррентный блок и комбинированная сверточно-рекуррентная нейронная сеть. Дополнительно осуществлен анализ следующих моделей традиционного машинного обучения: наивный байесовский классификатор, метод опорных векторов, логистическая регрессия, метод k-ближайших соседей, бустинг и случайный лес. Показателями эффективности обнаружения аномалий выступали следующие метрики: аккуратность, точность, полнота и F-мера, а также временные затраты на обучение модели. Построенные в процессе эксперимента модели глубокого обучения продемонстрировали более высокие показатели точности обнаружения аномалий в гетерогенном трафике большого объема, характерного для интернета вещей, по сравнению с методами традиционного машинного обучения. Выявлено, что с ростом числа слоев в нейронных сетях возрастает полнота обнаружения аномальных соединений, что улучшает распознавание неизвестных аномалий, но влечет за собой рост ложных срабатываний. Подготовка моделей традиционного машинного обучения в ряде случаев занимает меньшее время. Это связано с тем, что применение методов глубокого обучения требует большего количества ресурсов и вычислительных мощностей.* ***Практическая значимость:*** *полученные в исследовании результаты могут быть использованы для построения систем обнаружения сетевых аномалий в интернете вещей.*

Ключевые слова *— глубокое обучение, глубокие нейронные сети, обнаружение аномалий, интернет вещей, информационная безопасность.*

Для цитирования: Гайфулина Д. А., Котенко И. В. Анализ моделей глубокого обучения для задач обнаружения сетевых аномалий интернета вещей. *Информационно-управляющие системы*, 2021, № 1, с. 28–37. doi:10.31799/1684-8853-2021-1-28-37
For citation: Gaifulina D. A., Kotenko I. V. Analysis of deep learning models for network anomaly detection in Internet of Things. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 28–37 (In Russian). doi:10.31799/1684-8853-2021-1-28-37

Введение

В современном мире технология интернета вещей (Internet of Things — IoT) находит все большее применение в повседневной жизни человека. В общем виде интернет вещей представляет собой сеть распределенных устройств, которые связаны с окружающей средой при помощи датчиков, а также с программным обеспечением и серверами. Такие устройства также называют умными, или интеллектуальными. В то же время с ростом спроса на умные устройства и их доступности совершенствуются способы атак злоумышленников. Разнородность устройств и соединений, а также их ограничения на вычислительные ресурсы усложняют управление системами интернета вещей. В связи с этим обнаружение аномального поведения умных устройств порой сильно затруднено [1].

Для реагирования на угрозы безопасности необходимы инструменты анализа большого числа событий в системах интернета вещей, которые содержатся в сетевом трафике, логах и иных данных, объем которых порой очень велик. Помимо этого, разнородность источников и хранилищ информации приводит к высокой гетерогенности анализируемых данных. Машинное обучение и, в частности, глубокое обучение на данный момент являются мощными технологиями для анализа событий безопасности, обнаружения атак и аномального поведения умных устройств [2]. Ранее авторами был представлен системный анализ современных методов глубокого обучения, применяемых в задачах кибербезопасности [3, 4]. При этом сравнивать между собой различные модели глубоких нейронных сетей в научной литературе достаточно проблематично — в оценке эффективности применения моделей исследователи используют разные наборы данных или отличающиеся подмножества конкретного набора. Научная новизна проводимого исследования со-

стоит в предложенном сравнительном анализе моделей глубокого обучения различных классов и архитектур, основанном на оценке эффективности обнаружения аномалий в сетевом трафике интернета вещей с использованием единого программно-аппаратного обеспечения и одинаковых подмножеств набора данных для обучения и тестирования. Основной задачей проводимого исследования является выработка рекомендаций по выбору моделей глубокого обучения с высокими показателями эффективности обнаружения аномалий в сетевом трафике интернета вещей.

Классификация моделей глубокого обучения

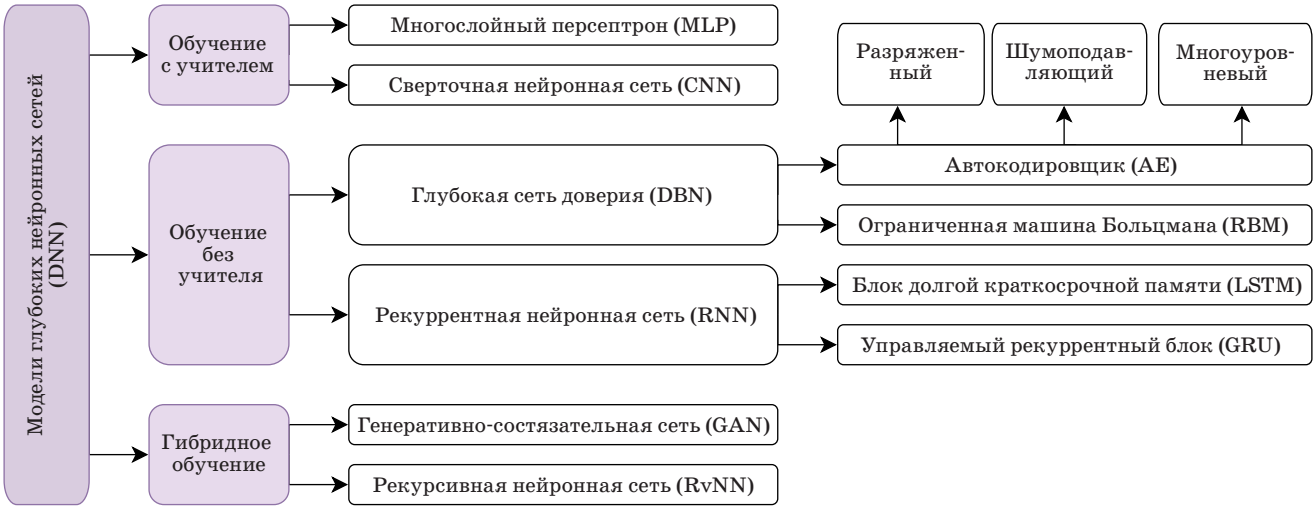
Глубокое обучение является частью семейства методов машинного обучения и основано на применении искусственных нейронных сетей. Глубокая нейронная сеть (Deep Neural Network — DNN) представляет собой нейронную сеть с несколькими слоями между входным и выходным слоями. Целью обучения DNN является нахождение корректного метода математических преобразований для превращения входных данных в выходные, независимо от линейной или нелинейной корреляции. Обучение может проходить как с учителем (supervised learning), так и без (unsupervised learning), а также при сочетании этих двух методов. Классификация наиболее распространенных моделей DNN по способу обучения представлена на рис. 1.

Многослойный персептрон (Multilayer Perceptron — MLP) является классом искусственных нейронных сетей прямого распространения (Feed Forward Neural Network — FFNN). При обнаружении вторжений, как правило, MLP используется для бинарной классификации сетевых соединений (нормальное или аномальное поведение). Взвешенные комбинации выходного слоя представляют собой прогноз, указывающий на принадлежность соединения к определенному классу.

Сверточная нейронная сеть (Convolutional Neural Network — CNN) используется обычно для эффективного распознавания образов, что также позволяет применять их для выявления вторжений. Заголовки пакетов сетевого трафика или последовательности системных вызовов кодируются в двумерную матрицу, которая является входными данными для CNN.

Методы обнаружения аномалий с использованием *глубоких сетей доверия* (Deep Belief Network — DBN), таких как *автокодировщики* (Autoencoder — AE) и *ограниченная машина Больцмана* (RBM), основаны на реконструкции данных, при которой определяется величина расхождения нормальных и аномальных данных. Для этого применяется распространение сигналов ошибки от выходов сети к входам для получения на выходе отклика, наиболее близкого к входному.

Связи между нейронами *рекуррентной нейронной сети* (Recurrent Neural Network — RNN) образуют направленный цикл. RNN может использовать свою внутреннюю память, такую как *блок долгой краткосрочной памяти* (Long Short-Term Memory — LSTM) или *управляемый рекуррентный блок* (Gated Recurrent Units — GRU). Использование RNN позволяет анализировать данные в виде временных рядов: сетевого трафика, последовательностей системных вызовов, журналов событий. Аномалия при этом может



■ ***Рис. 1.*** Классификация основных моделей DNN
■ ***Fig. 1.*** Classification of the main models of DNN

быть распознана как отклонение от предсказанного сетью последующего состояния.

В *генеративно-состязательных сетях* (Generative Adversarial Networks — GAN) используются модели генератора, анализирующего распределение реальных данных, и дискриминатора, оценивающего вероятность того, что входные данные поступают из реальных данных или из генератора. В подходах к обнаружению вторжений GAN применяются для исследования распределения нормальных данных, чтобы распознавать неизвестные аномалии.

Обзор существующих работ

Большая часть исследователей в области безопасности интернета вещей рассматривает методы глубокого обучения в рамках подходов к обнаружению атак и аномального поведения устройств [5, 6]. Основными преимуществами DNN по сравнению с методами традиционного машинного обучения являются высокая производительность и масштабируемость для растущего объема данных, а также возможность автоматически отбирать информативные признаки из необработанных данных.

В статье [7] исследуется потенциал рекуррентных нейронных сетей с LSTM для обнаружения вредоносных программ интернета вещей. Осуществляется сравнение разработанной модели с классификаторами, основанными на традиционных методах машинного обучения: методе опорных векторов (Support Vector Machine — SVM), наивном байесовском классификаторе (Naive Bayes), случайном лесе (Random Forest), бустинге (Adaptive Boosting — AdaBoost) и методе k-ближайших соседей (k-nearest neighbors algorithm — kNN). Анализ демонстрирует, что подход на основе глубокого обучения обеспечивает наилучший возможный результат. Сравнение с другими моделями глубокого обучения не производилось.

В исследовании [8] авторы предлагают собственную систему обнаружения аномалий для индустриальных систем интернета вещей с использованием автокодировщика и глубокой нейронной сети с прямой связью. Проводится сравнение созданной модели с характеристиками нескольких разработанных методов обнаружения аномалий, в том числе с глубокой сетью доверия [9], рекуррентной сетью [10], DNN [11] и Ensemble-DNN [12]. При этом приведенные модели оценивались на разных подмножествах исходных данных и с использованием разнородного аппаратного и программного обеспечения.

В статье [13] предлагается распределенная облачная среда глубокого обучения для обнаружения и предотвращения фишинговых и ботнет-атак на умные устройства. Разработанная модель RNN-LSTM сравнивается с моделями глубокого обучения, разработанными другими исследователями. Основным недостатком сравнительного анализа в данной работе является то, что рассматриваемые модели DNN оцениваются не на одинаковых наборах данных.

Авторы статьи [14] анализируют несколько методов глубокого обучения для обнаружения DDoS-атак: многослойный персептрон, сверточную нейронную сеть, RNN-LSTM и ансамбль CNN+LSTM. Проведено их сравнение с традиционными методами машинного обучения: методом опорных векторов, байесовским классификатором и случайным лесом. Авторы делают вывод о большей эффективности методов глубокого обучения, в особенности рекуррентных сетей.

В работе [15] также проводится систематическое сравнение CNN и RNN в системах обнаружения вторжений. Оцениваются следующие модели: базовая CNN (Basic CNN), CNN начальной архитектуры (Inception Architecture CNN), RNN-LSTM и управляемый рекуррентный блок. Авторы приходят к выводу, что CNN лучше подходит для бинарной классификации при обнаружении аномалий, а RNN лучше работают при обнаружении сложных атак в задачах мультиклассовой классификации.

Анализ указанных релевантных работ проведен по следующим атрибутам (табл. 1): сравнение моделей глубокого обучения (Г) между собой и с методами традиционного машинного обучения (М), использование метрик аккуратности (A), точности (P), полноты (R), F-меры (F) и временных затрат (T), используемый набор данных и анализируемые модели обучения.

Таким образом, особенностями предлагаемого исследования по сравнению с приведенными релевантными работами являются:

1) проведение эксперимента на едином наборе данных и с использованием одинакового программно-аппаратного обеспечения;

2) расширение сравнительной выборки моделей как глубокого обучения, так и традиционного машинного обучения;

3) введение оценки временных затрат на обучение модели, помимо таких показателей эффективности обнаружения аномалий, как аккуратность, точность, полнота и F-мера.

Выбор моделей глубоких нейронных сетей

В данном разделе проанализированы основные модели глубоких нейронных сетей:

— обучение с учителем — многослойный персептрон (MLP), сверточная нейронная сеть (CNN);

■ ***Таблица 1***. Анализ релевантных работ
■ ***Table 1***. Analysis of relevant works

Авторы, год	Метод		Показатель					Набор данных	Модели	Точность, %
	М	Г	A	P	R	F	T			
Haddad Pajouh, 2018 [7]	+	–	+	–	–	–	–	VirusTotal	RNN-LSTM	98
									SVM	82
									Naive Bayes	90
									Random Forest	92
									Ada Boost	93
									kNN	94
Muna, 2018 [8]	–	+	+	–	–	–	–	NSL-KDD	AE+FFNN	99
									DBN	95
									RNN	73
									DNN	76
									Ensemble-DNN	98
Parra, 2020 [13]	+	+	+	+	+	+	–	Собственный	RNN-LSTM	94
								CSIC 2010	DBN	95
									GRU	97
									SVM	99
Roopak, 2019 [14]	+	+	+	+	+	–	–	CICIDS 2017	MLP	86
									CNN	95
									LSTM	96
									CNN+LSTM	97
									SVM	95
									Naive Bayes	95
									Random Forest	94
Cui, 2018 [15]	–	+	+	+	+	+	–	ISCX2012	Basic CNN	94
									Inception Architecture CNN	95
									RNN-LSTM	93,7
									GRU	94

— обучение без учителя — рекуррентная нейронная сеть (RNN), блок долгой краткосрочной памяти (RNN-LSTM), управляемый рекуррентный блок (RNN-GRU);

— смешанное обучение — сверточно-рекуррентная сеть (CNN+RNN).

Методы глубокого обучения выбраны на основе проведения анализа существующих подходов к выявлению сетевых аномалий. Систематический анализ методов глубокого обучения, используемых в кибербезопасности, продемонстрировал, что данные модели дают хорошие результаты на практике [3, 4]. Анализ показателей обнаружения аномалий в сетевом трафике проводится не только между моделями разных классов, но и между моделями одного класса с различным количеством слоев и нейронов. Это позволяет экспериментально определить зависимость между структурой модели и ее производительностью.

Основные параметры выбранных моделей представлены в табл. 2. Архитектура сети описывается следующим образом: количество слоев и нейронов в каждом из них (h — скрытый слой, p — субдискретизирующий слой, c — сверточный слой, n — полносвязный слой).

Функция активации нейронной сети определяет выходное значение в зависимости от результата взвешенной суммы входов и порогового значения [16]. Для всех моделей в качестве функции активации выходного слоя выбрана сигмои-

■ ***Таблица 2***. Параметры модели глубоких нейронных сетей

■ ***Table 2***. Parameters of the DNN

Обозначение	Архитектура	Функция активации
MLP-1	h(1024)-h(768)	ReLu, sigmoid
MLP-2	h(1024)-h(768)-h(512)	
MLP-3	h(1024)-h(768)-h(512)-h(256)	
MLP-4	h(1024)-h(768)-h(512)-h(256)-h(128)	
CNN-1	2c(64)-1p(2)-1n(128)	ReLu, sigmoid
CNN-2	2c(64)-1p(2)-2c(128)-1p(2)-1n(128)	
RNN-1	h(16)-h(16)-h(16)	Sigmoid
RNN-2	h(32)-h(32)-h(32) -h(32)	
RNN-LSTM-1	h(16)-h(16)-h(16)	
RNN-LSTM-2	h(32)-h(32)-h(32) -h(32)	
RNN-GRU-1	h(16)-h(16)-h(16)	
RNN-GRU-2	h(32)-h(32)-h(32) -h(32)	
CNN+LSTM-1	CNN(2c(64)-1p(2))-LSTM(h(128))	
CNN+LSTM-2	CNN(2c(64)-1p(2)-2c(128)-1p(2)- LSTM(h(128)	

да (sigmoid), и в ряде моделей она дополнена линейным выпрямителем (Rectified linear unit — ReLU) на скрытых слоях. Также применяется метод контроля емкости дропаут (dropout), позволяющий предотвратить переобучение нейронной сети [17].

Эксперименты

Оценка производительности любых систем обнаружения аномалий для интернета вещей требует наличия исходных данных, включающих в себя набор сетевых признаков, таких как признаки на основе номеров портов источника и назначения, полезной нагрузки (payload-based), поведения (behaviour based) и потока данных (flow-based).

В качестве экспериментальных данных для анализа моделей DNN в задачах обнаружения сетевых аномалий интернета вещей был выбран открытый набор данных UNSW-NB15 [18, 19], содержащий 2 540 044 записей — векторов признаков сетевых соединений TCP/IP и соответствующих им меток классов. В этом наборе данных сетевые пакеты включают как информацию о реальной нормальной активности сети, так и девять типов атак: фаззеры (Fuzzers), анализаторы (Analysis), бэкдоры (Backdoors), отказ в обслуживании (DoS), эксплоиты (Exploits), обобщенные (Generic), разведка (Reconnaissance), шелл-код (Shellcode) и черви (Worms). Данные UNSW-NB15 для обучения и тестирования систем обнаружения вторжений содержат 47 признаков, таких как IP-адреса, номера портов, байты транзакции и др. [20], и две метки класса — категорию атаки и метку аномальности соединения. Первые 35 признаков представляют собой интегрированную информацию из пакетов данных, а остальные определяются для сценариев подключения.

Обнаружение аномалий представляет собой процесс идентификации отклонений от нормального профиля системы. Таким образом, для обнаружения аномалий в сетевом трафике UNSW-NB15 используется бинарная классификация, и в качестве метки класса используется критерий аномальности соединения, где 0 соответствует нормальному профилю, а 1 — аномалии.

Анализ моделей DNN для задач обнаружения сетевых аномалий интернета вещей состоит из описанных ниже этапов.

Предобработка данных **(1)** заключается в преобразовании входного набора данных: 47 признаков сетевых соединений и метки класса — в форму, подаваемую на вход анализируемым моделям. К признакам номинального типа, таким как IP-адреса, название протокола и сервиса передачи данных, применяется горячее кодирование (one-hot encoding) — метод представления категориальных переменных в виде двоичных векторов. Далее производится нормализация значений всех признаков к диапазону [0...1]. Нормализация данных осуществляется, так как дисбаланс между значениями признаков может вызвать неустойчивость работы модели, ухудшить результаты обучения и замедлить процесс моделирования. В качестве данных для обучения моделей выбирается 80% исходного набора данных (1 547 081 запись), а для тестирования моделей — 20% (386 771 запись). Важной особенностью данного этапа исследований является отсутствие высокой сбалансированности нормального и аномального класса сетевых соединений, что наиболее близко к реальным условиям при возникновении аномалий в сетевом трафике. Так, в данном случае отношение аномальных данных к нормальным составляет 1:4. Обучающая и тестовая выборка являются однородными.

Обучение моделей **(2)** осуществляется на одинаковом тренировочном наборе данных, а обнаружение аномалий **(3)** — на одинаковом тестовом наборе данных. Для обучения и валидации моде-

лей глубокого обучения использовались следующие гиперпараметры: размер пакта (batch size) — 64, алгоритм оптимизации — adam, функция потерь (loss function) — binary cross-entropy.

Разработанные модели глубоких нейронных сетей и традиционного машинного обучения были реализованы с использованием Python 3.6, Tensorflow 2.1, Scikit-learn 0.23.2, Numpy 1.19.2, Pandas 1.1.3 и Scipy 1.5.2. Все эксперименты проводились на Acer Swift SF315-52G с процессором Intel Core i5 с тактовой частотой 1,8 ГГц, ОЗУ 8 ГБ и операционной системой Windows 10.

Оценка эффективности обнаружения аномалий **(4)** заключается в вычислении следующих метрик: аккуратности (A), точности (P), полноты (R), F-меры (F) и временных затрат на обучение (T). Аккуратность характеризует долю экземпляров сетевых соединений, по которым модель приняла правильное решение о принадлежности к нормальному или аномальному классу. Точность характеризует долю верно классифицированных экземпляров сетевых соединений относительно всех экземпляров сетевого трафика. Полнота характеризует долю найденных моделью экземпляров сетевых соединений, принадлежащих нормальному или аномальному классу относительно всех экземпляров сетевого трафика. F-мера представляет собой гармоническое среднее между точностью и полнотой.

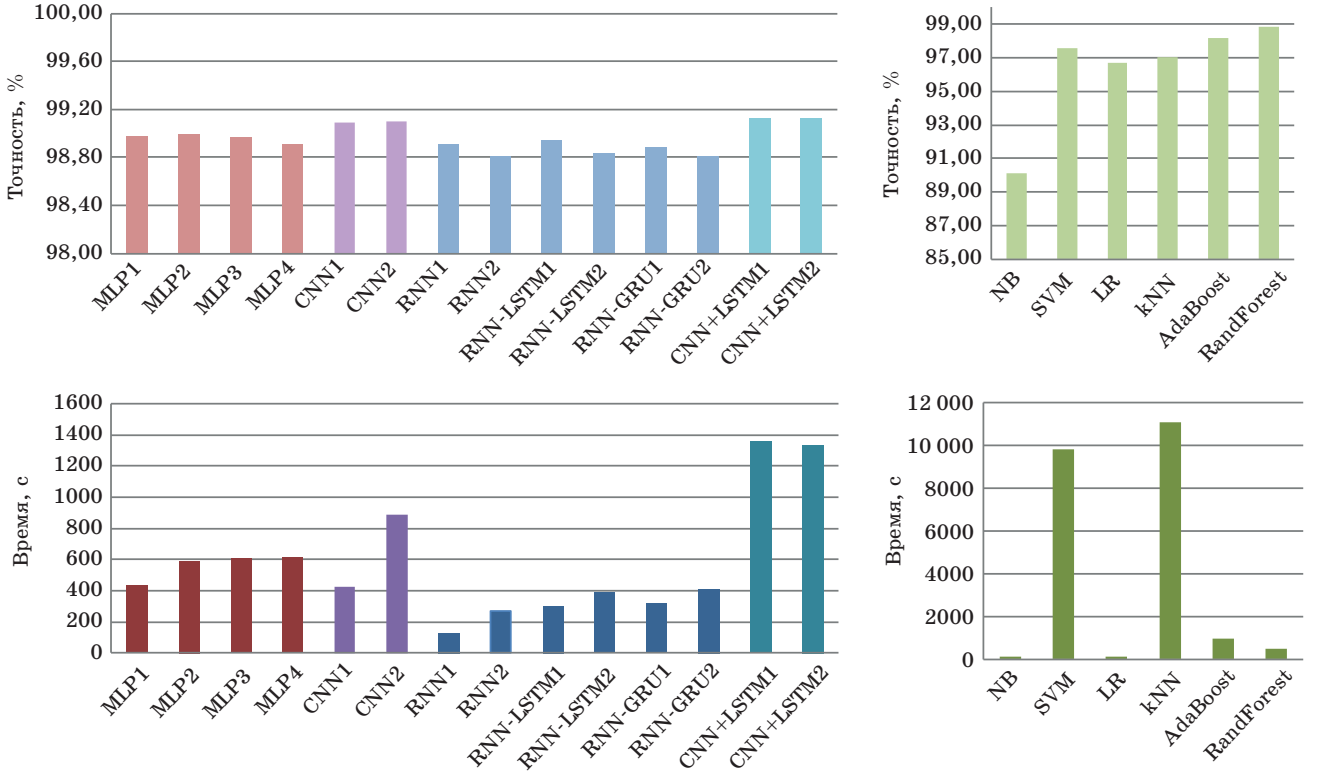
Результаты экспериментов по анализу моделей глубокого обучения представлены в табл. 3. Данные приводятся для первой эпохи обучения.

■ ***Таблица 3***. Анализ моделей глубокого обучения
■ ***Table 3***. Analysis of deep learning models

Модель	Оценка				
	A, %	P, %	R, %	F, %	T, с
MLP-1	98,976	93,365	98,465	95,847	430,946
MLP-2	98,996	93,612	98,384	95,939	595,026
MLP-3	98,965	93,198	**98,552**	95,800	606,213
MLP-4	98,913	92,665	**98,669**	95,573	617,150
CNN-1	**99,095**	**95,049**	97,741	**96,376**	423,574
CNN-2	**99,096**	94,713	98,086	**96,370**	892,186
RNN-1	98,912	93,621	97,688	95,612	**128,943**
RNN-2	98,810	92,199	98,295	95,149	**273,471**
RNN-LSTM-1	**98,938**	93,952	97,569	95,726	**301,561**
RNN-LSTM-2	98,840	92,629	98,106	95,289	388,294
RNN-GRU-1	98,883	93,212	97,867	95,483	317,043
RNN-GRU-2	98,808	92,104	**98,381**	95,139	412,727
CNN+LSTM-1	**99,126**	**95,932**	97,127	**96,526**	1355,819
CNN+LSTM-2	**99,124**	**95,972**	97,078	**96,522**	1334,130

■ ***Таблица 4***. Анализ моделей традиционного машинного обучения
■ ***Table 4***. Analysis of traditional machine learning models

Модель	Оценка				
	A, %	P, %	R, %	F, %	T, с
Naive Bayes	90,12	94,22	90,12	91,18	**130,36**
SVM	97,56	97,58	97,58	97,58	9786,67
Logistic Regression	96,73	96,77	96,73	96,74	**131,26**
kNN	97,07	97,07	97,07	97,07	11074,83
AdaBoost	98,2	98,2	98,2	98,2	940,17
Random Forest	**98,87**	**98,85**	**98,85**	**98,45**	468,05



■ ***Рис. 2.*** Сравнение точности и времени обучения моделей
■ ***Fig. 2.*** Comparison of accuracy and time of model learning

Дополнительно на тех же данных проанализированы следующие модели традиционного машинного обучения: наивный байесовский классификатор, метод опорных векторов, логистическая регрессия (Logistic Regression), метод k-ближайших соседей, бустинг и случайный лес. Результаты представлены в табл. 4.

Сравнение точности рассмотренных моделей и времени обучения для обнаружения аномалий продемонстрировано на рис. 2.

Анализ результатов экспериментов

Результаты проведенных экспериментов позволяют сделать вывод, что большинство моделей глубоких нейронных сетей обладает высокой точностью обнаружения аномалий в гетерогенном трафике большого объема для применения их на практике. Среди моделей традиционного машинного обучения сходной высокой точностью обладают ансамбли классификаторов, такие как AdaBoost (A = 98,2 %) и случайный лес (A = 98,87 %).

Среди моделей глубокого обучения с учителем лучшую точность обнаружения демонстрирует сверточная нейронная сеть (A = 99,1 %). При этом с увеличением числа слоев время обучения существенно возрастает, в отличие от показателя точности, изменяющегося не так сильно. Многослойный персептрон обладает наибольшей полнотой обнаружения аномалий (R = 98,67 %). Это значит, что данная модель распознает большее количество экземпляров аномальных соединений, что позволяет избежать их пропусков. С увеличением числа слоев точность многослойного персептрона ухудшается, а время, затраченное на обучение, возрастает (см. рис. 2). В данном эксперименте предпочтительной архитектурой многослойного персептрона является модель MLP-2 с точностью обнаружения аномалий A = 99 %.

Для моделей глубокого обучения без учителя, представленных рекуррентными сетями, лучшие результаты показывает блок долгой краткосрочной памяти (A = 98,938 %). Обучение данного вида моделей занимает наименьшее количество времени, следовательно, и меньшее количество вычислительных ресурсов, что нередко является существенным параметром для устройств интернета вещей. С увеличением числа слоев в архитектуре рекуррентных нейронных сетей возрастает полнота обнаружения аномальных соединений, но снижается точность, что связано с накоплением ошибок обучения. Таким образом, в модели обнаружения аномалий, настроенной на низкий коэффициент ложных срабатываний, точность будет

представлять собой более значимую характеристику, тогда как модель с высокой полнотой классификации предпочтительней для распознавания ранее неизвестных типов аномалий.

Наивысшей точностью обнаружения аномалий среди представленных моделей DNN обладает комбинированная сверточно-рекуррентная нейронная сеть (A = 99,13 %). При этом время обучения данной сети является самым продолжительным.

При сравнении между собой базовых моделей DNN разных классов можно установить, что различие в точности обнаружения аномалий не является весьма значительным — не более 1 %. Более разнящейся характеристикой является время обучения сети, которое также возрастает соответственно увеличению числа слоев. Стоит отметить, что подготовка моделей традиционного машинного обучения по большей части занимает меньшее количество времени, за исключением моделей опорных векторов и k-ближайших соседей. Это связано с тем, что применение методов глубокого обучения требует большего количества вычислительных мощностей.

Представлены рекомендации (табл. 5) для выбора наиболее предпочтительной модели глубокого обучения исходя из временных затрат на обучение и приоритета в обнаружении аномалий в сетевом трафике интернета вещей.

Для систем, работающих в режиме реального времени и часто обновляемых, скорость моделирования является значимой характеристикой и должна быть минимизирована. В то время как для некоторых систем, обучаемых офлайн, время моделирования может быть увеличено для более тщательной настройки и повышения эффективности функционирования.

■ ***Таблица 5***. Рекомендации по использованию моделей глубокого обучения

■ ***Table 5***. Recommendations for using deep learning models

Приоритет в обнаружении аномалий	Временные затраты	Рекомендация
Низкий коэффициент ложных срабатываний	Не имеют значения	Комбинированная нейронная сеть
	Минимизированы	Рекуррентная нейронная сеть, в частности блок долгой краткосрочной памяти
Распознавание неизвестных типов аномалий	Не имеют значения	Многослойный персептрон с большим количеством слоев
	Минимизированы	Управляемый рекуррентный блок с большим количеством слоев

Заключение

В данном исследовании представлен анализ базовых моделей глубокого обучения для задач обнаружения аномалий в сетевом трафике интернета вещей. Экспериментальная оценка моделей глубокого обучения проводилась с использованием единого программно-аппаратного обеспечения и одинаковых подмножеств набора данных UNSW-NB 15 для обучения и тестирования. В качестве базовых моделей глубоких нейронных сетей рассмотрены многослойный персептрон, сверточная нейронная сеть, рекуррентная нейронная сеть, блок долгой краткосрочной памяти, управляемый рекуррентный блок и комбинированная сверточно-рекуррентная нейронная сеть.

Построенные модели глубокого обучения продемонстрировали высокие показатели точности обнаружения аномалий — от 98,8 %. В работе представлены рекомендации для выбора наиболее предпочтительной модели глубокого обучения исходя из временных затрат на обучение модели и приоритета в обнаружении аномалий в сетевом трафике интернета вещей. При настройке модели обнаружения аномалий на низкий коэффициент ложных срабатываний точность будет представлять собой значимую характеристику. С увеличением числа слоев в архитектуре DNN возрастает полнота обнаружения аномальных соединений, что в свою очередь предпочтительней для распознавания ранее неизвестных типов аномалий. Увеличение числа слоев в модели обнаружения аномалий требует мощных вычислительных ресурсов центральных компонентов интернета вещей.

В дальнейшем предполагается продолжить анализ характеристик моделей DNN, применяемых в задачах кибербезопасности. Одним из направлений будущих работ является исследование влияния структуры сетевого трафика на показатели эффективности использования моделей глубокого обучения. На основании полученных результатов планируется разработать подход к выявлению и корреляции событий безопасности на базе методов глубокого обучения.

Финансовая поддержка

Работа выполнена при частичной финансовой поддержке проекта РФФИ 18-29-22034 мк и бюджетной темы 0073-2019-0002.

Литература

1. **Alrawais A., Alhothaily A., Hu C., Cheng X.** Fog computing for the internet of things: Security and privacy issues. *IEEE Internet Computing*, 2017, no. 21(2), pp. 34–42. doi:10.1109/MIC.2017.37
2. **Браницкий А. А., Котенко И. В.** Обнаружение сетевых атак на основе комплексирования нейронных, иммунных и нейро-нечетких классификаторов. *Информационно-управляющие системы*, 2015, № 4, с. 69–77. doi:10.15217/issn1684-8853.2015.4.69
3. **Гайфулина Д. А., Котенко И. В.** Применение методов глубокого обучения для решения задач кибербезопасности. Ч. 1. *Вопросы кибербезопасности*, 2020, № 3(37), с. 76–86. doi:10.21681/2311-3456-2020-03-76-86
4. **Гайфулина Д. А., Котенко И. В.** Применение методов глубокого обучения для решения задач кибербезопасности. Ч. 2. *Вопросы кибербезопасности*, 2020, № 4(38), с. 11–21. doi:10.21681/2311-3456-2020-04-11-21
5. **Al-Garadi M. A., Mohamed A., Al-Ali A., Du X., Guizani M.** A Survey of machine and deep learning methods for internet of things (IoT) security. *IEEE Communications Surveys & Tutorials*, 2020, vol. 22, no. 3, pp. 1646–1685. doi:10.1109/COMST.2020.2988293
6. **Левшун Д. С., Гайфулина Д. А., Чечулин А. А., Котенко И. В.** Проблемные вопросы информационной безопасности киберфизических систем. *Информатика и автоматизация*, 2020, т. 19, № 5, с. 1050–1088. doi:10.15622/ia.2020.19.5.6
7. **HaddadPajouh H., Dehghantanha A., Khayami R., Choo K. K. R.** A Deep recurrent neural network based approach for internet of things malware threat hunting. *Future Generation Computer Systems*, 2018, vol. 85, pp. 88–96. doi:10.1016/j.future.2018.03.007
8. **Muna Al H., Moustafa N., Sitnikova E.** Identification of malicious activities in industrial internet of things based on deep learning models. *Journal of Information Security and Applications*, 2018, vol. 41, pp. 1–11. doi:10.1016/j.jisa.2018.05.002
9. **Alom M. Z., Bontupalli V., Taha T. M.** Intrusion detection using deep belief networks. *2015 National Aerospace and Electronics Conference (NAECON)*, Dayton, 2015, pp. 339–344. doi:10.1109/NAECON.2015.7443094
10. **Yin C., Zhu Y., Fei J., He X.** A Deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 2017, vol. 5, pp. 21954–21961. doi:10.1109/ACCESS.2017.2762418
11. **Tang T. A., Mhamdi L., McLernon D., Zaidi S., Ghogho M.** Deep learning approach for network intrusion detection in software defined networking. *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2016, pp. 258–263. doi:10.1109/WINCOM.2016.7777224
12. **Ludwig S. A.** Intrusion detection of multiple attack classes using a deep neural net ensemble. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, 2017, pp. 1–7. doi:10.1109/SSCI.2017.8280825.
13. **Parra G. D. L. T., Rad P., Choo K. K. R., Beebe N.** Detecting internet of things attacks using distributed deep learning. *Journal of Network and Computer Applications*, 2020, vol. 163, pp. 102662. doi:10.1016/j.jnca.2020.102662
14. **Roopak M., Tian G. Y., Chambers J.** Deep learning models for cyber security in IoT networks. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, USA, 2019, pp. 0452–0457. doi:10.1109/CCWC.2019.8666588
15. **Cui J., Long J., Min E., Liu Q., Li Q.** Comparative study of CNN and RNN for deep learning based intrusion detection system. *International Conference on Cloud Computing and Security*, Springer, Cham, 2018, pp. 159–170. doi:10.1007/978-3-030-00018-9_15
16. **Nwankpa C., Ijomah W., Gachagan A., Marshall S.** Activation functions: Comparison of trends in practice and research for deep learning. *ArXiv preprint arXiv:1811.03378*, 2018. 20 p.
17. **Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdino R.** Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 2014, vol. 15, no. 1, pp. 1929–1958.
18. *UNSW-NB15 Dataset*. https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/ (дата обращения: 27.10.2020).
19. **Moustafa N., Slay J.** UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, 2015, pp. 1–6. doi:10.1109/MilCIS.2015.7348942
20. **Moustafa N., Turnbull B., Choo K. R.** An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet of Things Journal*, 2019, vol. 6, no. 3, pp. 4815–4830. doi:10.1109/JIOT.2018.2871719

UDC 004.056
doi:10.31799/1684-8853-2021-1-28-37

Analysis of deep learning models for network anomaly detection in Internet of Things

D. A. Gaifulina[a], Junior Researcher, orcid.org/0000-0002-5266-8649
I. V. Kotenko[a], Dr. Sc., Tech, Professor, orcid.org/0000-0001-6859-7120, ivkote@comsec.spb.ru
[a]St. Petersburg Federal Research Center of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: The article discusses the problem of choosing deep learning models for detecting anomalies in Internet of Things (IoT) network traffic. This problem is associated with the necessity to analyze a large number of security events in order to identify the abnormal behavior of smart devices. A powerful technology for analyzing such data is machine learning and, in particular, deep learning. **Purpose:** Development of recommendations for the selection of deep learning models for anomaly detection in IoT network traffic. **Results:** The main results of the research are comparative analysis of deep learning models, and recommendations on the use of deep learning models for anomaly detection in IoT network traffic. Multilayer perceptron, convolutional neural network, recurrent neural network, long short-term memory, gated recurrent units, and combined convolutional-recurrent neural network were considered the basic deep learning models. Additionally, the authors analyzed the following traditional machine learning models: naive Bayesian classifier, support vector machines, logistic regression, k-nearest neighbors, boosting, and random forest. The following metrics were used as indicators of anomaly detection efficiency: accuracy, precision, recall, and F-measure, as well as the time spent on training the model. The constructed models demonstrated a higher accuracy rate for anomaly detection in large heterogeneous traffic typical for IoT, as compared to conventional machine learning methods. The authors found that with an increase in the number of neural network layers, the completeness of detecting anomalous connections rises. This has a positive effect on the recognition of unknown anomalies, but increases the number of false positives. In some cases, preparing traditional machine learning models takes less time. This is due to the fact that the application of deep learning methods requires more resources and computing power. **Practical relevance:** The results obtained can be used to build systems for network anomaly detection in Internet of Things traffic.

Keywords — deep learning, deep neural networks, anomaly detection, Internet of Things, information security.

For citation: Gaifulina D. A., Kotenko I. V. Analysis of deep learning models for network anomaly detection in Internet of Things. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 28–37 (In Russian). doi:10.31799/1684-8853-2021-1-28-37

References

1. Alrawais A., Alhothaily A., Hu C., Cheng X. Fog computing for the internet of things: Security and privacy issues. *IEEE Internet Computing*, 2017, no. 21(2), pp. 34–42. doi:10.1109/MIC.2017.37
2. Branitskiy A. A., Kotenko I. V. Network attack detection based on combination of neural, immune and neuro-fuzzy classifiers. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 4, pp. 69–77 (In Russian). doi:10.15217/issn1684-8853.2015.4.69
3. Gaifulina D. A., Kotenko I. V. Application of deep learning methods in cybersecurity tasks. Part 1. *Voprosy kiberbezopasnosti*, 2020, no. 3(37), pp. 76–86 (In Russian). doi:10.21681/2311-3456-2020-03-76-86
4. Gaifulina D. A., Kotenko I. V. Application of deep learning methods in cybersecurity tasks. Part 2. *Voprosy kiberbezopasnosti*, 2020, no. 4(38), pp. 11–21 (In Russian). doi:10.21681/2311-3456-2020-04-11-21
5. Al-Garadi M. A., Mohamed A., Al-Ali A., Du X., Guizani M. A Survey of machine and deep learning methods for internet of things (IoT) security. *IEEE Communications Surveys & Tutorials*, 2020, vol. 22, no. 3, pp. 1646–1685. doi:10.1109/COMST.2020.2988293.
6. Levshun D., Gaifulina D., Chechulin A., Kotenko I. Problematic issues of information security of cyber-physical systems. *Informatics and Automation*, 2020, vol. 19, no. 5, pp. 1050–1088. doi:10.15622/ia.2020.19.5.6
7. HaddadPajouh H., Dehghantanha A., Khayami R., Choo K. K. R. A Deep recurrent neural network based approach for internet of things malware threat hunting. *Future Generation Computer Systems*, 2018, vol. 85, pp. 88–96. doi:10.1016/j.future.2018.03.007
8. Muna Al H., Moustafa N., Sitnikova E. Identification of malicious activities in industrial internet of things based on deep learning models. *Journal of Information Security and Applications*, 2018, vol. 41, pp. 1–11. doi:10.1016/j.jisa.2018.05.002
9. Alom M. Z., Bontupalli V., Taha T. M. Intrusion detection using deep belief networks. *2015 National Aerospace and Electronics Conference (NAECON)*, Dayton, 2015, pp. 339–344. doi:10.1109/NAECON.2015.7443094
10. Yin C., Zhu Y., Fei J., He X. A Deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 2017, vol. 5, pp. 21954–21961. doi: 10.1109/ACCESS.2017.2762418
11. Tang T. A., Mhamdi L., McLernon D., Zaidi S., Ghogho M. Deep learning approach for network intrusion detection in software defined networking. *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2016, pp. 258–263. doi:10.1109/WINCOM.2016.7777224
12. Ludwig S. A. Intrusion detection of multiple attack classes using a deep neural net ensemble. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, 2017, pp. 1–7. doi:10.1109/SSCI.2017.8280825
13. Parra G. D. L. T., Rad P., Choo K. K. R., Beebe N. Detecting internet of things attacks using distributed deep learning. *Journal of Network and Computer Applications*, 2020, vol. 163, pp. 102662. doi:10.1016/j.jnca.2020.102662
14. Roopak M., Tian G. Y., Chambers J. Deep learning models for cyber security in IoT networks. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, USA, 2019, pp. 0452–0457. doi:10.1109/CCWC.2019.8666588
15. Cui J., Long J., Min E., Liu Q., Li Q. Comparative study of CNN and RNN for deep learning based intrusion detection system. *International Conference on Cloud Computing and Security*, Springer, Cham, 2018, pp. 159–170. doi:10.1007/978-3-030-00018-9_15
16. Nwankpa C., Ijomah W., Gachagan A., Marshall S. Activation functions: Comparison of trends in practice and research for deep learning. *ArXiv preprint arXiv:1811.03378*, 2018. 20 p.
17. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdino R. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 2014, vol. 15, no. 1, pp. 1929–1958.
18. *UNSW-NB15 Dataset*. Available at: https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/ (accessed 27 October 2020).
19. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, 2015, pp. 1–6. doi:10.1109/MilCIS.2015.7348942
20. Moustafa N., Turnbull B., Choo K. R. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet of Things Journal*, 2019, vol. 6, no. 3, pp. 4815–4830. doi:10.1109/JIOT.2018.2871719

UDC 004.056.53
doi:10.31799/1684-8853-2021-1-38-44

Articles

Minimum-storage regenerating codes resistant to special adversary

S. A. Kruglik[a,b], *Junior Researcher, orcid.org/0000-0001-9557-5197, stanislav.kruglik@skoltech.ru*
[a]*Skolkovo Institute of Science and Technology, bld. 1, 30, Bolshoy Boulevard, 121205, Moscow, Russian Federation*
[b]*Moscow Institute of Physics and Technology, 9, Institutskiy Per., 141701, Dolgoprudny, Moscow region, Russian Federation*

Introduction: *To deal with temporally unavailable nodes in distributed storage system engineers apply special classes of erasure correction codes. These codes allow repairing temporally unavailable nodes by downloading a small amount of data from the remaining ones. At the same time, there are safety threats in the presence of an eavesdropper.* ***Purpose:*** *To consider a new mathematical model of eavesdropper that has limited access to all nodes in the system and develop codes resistant to it.* ***Methods:*** *Information-theoretic arguments and mixing information symbols with random ones by systematic Reed — Solomon code.* ***Results:*** *We introduced a new mathematical model of eavesdropper with limited access to all nodes in the distributed storage system. Note that the proposed eavesdropper is passive or, in other words, cannot change accessed data. In this paper, we derived parameters of optimal regenerating codes resistant to such adversary as well as give a technique to ensure the necessary resistance. As a result, we obtained the construction of optimal minimum storage regenerating codes resistant against such adversary.* ***Practical relevance:*** *Proposed constructions can provide resistance against a given adversary while ensuring effective data repair.*

Keywords *— distributed systems, MSR array codes, repair of a temporally unavailable node, mathematical model of system, resistance to adversary.*

For citation: Kruglik S. A. Minimum-storage regenerating codes resistant to special adversary. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 38–44. doi:10.31799/1684-8853-2021-1-38-44

Introduction

Distributed storage systems consisting of thousands of individual nodes that stores a portion of users information become de-facto the standard of modern data storage. High expansion of such systems is leveraged by the constant growth of amount of data stored by humanity. Leading technological companies such as Facebook or Google heavily rely on distributed storage systems [1, 2]. One of the most important problem of current version of such systems is drive failures that occur constantly. To handle it system designers employ erasure-correcting codes for efficient repair of temporally unavailable nodes. Despite several node failures are possible the most common scenario is one node failure and the main goal of research community is to develop codes that optimize the recovery of one node failure in different terms. These terms arose from the distributed nature of systems and the necessity to communicate data between several nodes [3, 4]. One of them, called locality, measures the efficiency of recovery in number of nodes accessed during this procedure [5]. Another one, called repair bandwidth, takes into account the total amount of data transmitted to accomplish the repair [6]. Codes optimized by the second measure are called the regenerating codes and are the main focus of this paper.

In our derivations, we consider a distributed storage system that stores in n nodes B independent random symbols uniformly distributed over the finite field $GF(q)$. Each of these nodes has a storage capacity of l symbols (also termed as sub-packetization level in corresponding literature). We encode B symbols by regenerating code in such a way that in case of one node failure the replacement node can repair its content (or function of it in case of functional repair) by connecting to any set of d helper nodes ($d > k - 1$) and downloading β symbols from each of them. The total amount of downloaded data $d\beta$ is termed as repair bandwidth. Also, regenerating code has such a property that any k nodes can recover all B message symbols. Note that in such a case we have to download all content from them.

In the initial paper on regenerating codes [6] authors utilizing network-flow graph established that parameters of these codes must satisfy the following bound

$$B \le \sum_{i=0}^{k-1} \min(l, (d-i)\beta). \qquad (1)$$

It can be deduced from the form of (1) that achieving equality in it while fixed parameters B, k, and d leads to the tradeoff between the repair bandwidth $d\beta$ and the sub-packetization level l. Two extreme points of this tradeoff determine two classes

of regenerating codes — minimum bandwidth regenerating (MBR) codes and minimum storage regenerating (MSR) codes. In the first case, we initially minimize bandwidth and after it minimize storage on each node. There are a lot of constructions of such codes in the literature, see [7–9] and references therein. Unfortunately, known constructions have code rate no more than 1/2 that restricts their practical applications. Another drawback of MBR codes is that there are no constructions with optimal access property, namely we have to access a large amount of data to accomplish the node repair process while transmitting only the function from them. In case of MSR codes that are the main focus of this paper we first minimize storage on each node and after it the bandwidth. These codes have many advantages over MBR codes, namely there are explicit constructions of high-rate MSR array codes as well as constructions of such codes with optimal-access property. The latter means that in case of node repair we only have to access helper node symbols transmitted to the replacement node. For more details, see papers [9–11] and references in them.

Despite importance of repairing the content of unavailable node, this paper focus on another aspect of distributed storage systems namely safety of stored data. Due to distributed nature of such systems and as a consequence, increasing use of untrusted node providers or communication channels, they are vulnerable to different type of attacks or data leakage [12–14]. In this paper, we focus on threats caused by eavesdropper that gains access to some portion of stored information. The considered eavesdropper also denoted below by E is passive, i. e. E cannot change accessed data. There are two popular approaches to preserve resistance against E. One of them is to use computational cryptography based on difficulty in the computation of some function. Deploying this approach needs to distribute keys as well as provide additional (typically hard) computations that make it irrelevant for distributed storage systems [15]. Another one is an information-theoretic approach in which we mix stored data with random symbols taken uniformly and independent from the same alphabet. In such a case, we ensure that eavesdropper gaining access to the limited number of symbols obtain no information about stored content. In other words, we ensure the zero-mutual information between stored content and information available to E [12]. In this paper we focus on information-theoretic approach only. Note that this problem formulation is highly connected with Wire-Tap Channel II in which eavesdropper has an access to any fixed size subset of symbols transmitted through a noiseless channel [16]. Proposed solution based on coset coding provide resistance against such E while ensuring reconstruction of all information content without the possibility of repair part of it. This fact makes it hard to generalize the given solution to the case of regenerating codes that support single node repair.

Recent papers within safety of regenerating codes focused on resistance against eavesdropper with full access to a limited number of nodes. Some papers also consider a stronger adversary with additional access to data transmitted during the repair. This eavesdropper model corresponds to the case then the adversary can control some subset of nodes. There exist corresponding bounds on the amount of information that can be safely stored in such systems as well as constructions attaining them. For more details, we refer to the papers [12, 13, 17].

In this paper, we continue our research initiated in [18] and consider a new mathematical model of eavesdropper that can access the limited number of symbols from each node in the distributed system. As before we aim to ensure zero mutual information between stored data and data available to E. We consider the minimum storage regenerating codes with optimal access property and derive the technique to make it resistant against given eavesdropper. Note that such consideration is enough natural as these codes ensure node recovery while accessing a small portion of symbols from any node in a given helper set.

The main contribution of this paper is as follows. We consider a new mathematical model of eavesdropper with limited access to all nodes in the distributed storage system, give a bound on parameters of regenerating codes resistant against such adversary as well as propose an explicit construction of MSR-array codes with optimal access property secure against it.

Preliminaries

Within this paper, we use the following notations. By $GF(q)$ we define the finite field with q elements and by $\mathbf{X} = (X_1, \ldots, X_n)^t$ the column vector with n elements over it. We denote the set of n positions as $[n] = \{0, 1, \ldots, n-1\}$ and define the restriction of column vector $\mathbf{X}$ to its subset T as $\mathbf{X}_T$. By superscript t we mean the transpose operations and by superscript s the parameters of safe version of code construction.

By $H(X)$ we define the entropy of discrete random variable X and by $I(X; Y) = H(X) - H(X|Y)$ the mutual information between discrete random variables X and Y. $H(X|Y)$ denote the conditional entropy of random variable X given random variable Y. The same is held for vectors consisting of discrete random variables.

Within this paper, we consider MSR-array codes with optimal access property proposed by Ye and

Barg in paper [11]. Such codes attain the extreme point in bound (1) and have the following parameters: $l = B/k$ and $d\beta = \frac{B}{k(d-k+1)}$. The code construction is explained in Construction 1.

Construction 1. Let us construct array code of length n, sub-packetization level $l = r^{n-1} = (n-k)^{n-1}$ land number of nodes necessary to recover information content k. The code is constructed over $GF(q)$ with size more than n and primitive element γ. We consider the case of $d = n-1$ that corresponds to the most common scenario of one node failure. The code is formed from $l \times n$ matrices over $GF(q)$ each encoding kl information symbols. Encoding procedures are defined using parity-check equations in the following form:

$$(\mathbf{C}_1, \ldots, \mathbf{C}_n) : \sum_{i=1}^{n} \mathbf{A}_{t,i}\mathbf{C}_i = 0, \; t = 1, \ldots, r, \qquad (2)$$

where $\mathbf{C}_i = (c_{i,0}, \ldots, c_{i,l-1})^t$ is a column vector that corresponds to l code symbols over $GF(q)$ stored on node i. $\mathbf{A}_{t,i} = \mathbf{A}_i^{t-1}$, where $t = 1, \ldots, r$ and $i = 1, \ldots, n$ are $l \times l$ matrices over $GF(q)$. Note that by forming the first k column vectors $\mathbf{C}_i$ from $B = kl$ information symbols we can determine the remaining $r = n-k$ column vectors. The specific code families can be obtained by choosing different forms of matrices $\mathbf{A}_1, \ldots, \mathbf{A}_n$ such that and $\mathbf{A}_i - \mathbf{A}_j$ is invertible and multiplication of two matrices has commutative property. In our case to obtain MSR codes with optimal access property we choose $\mathbf{A}_1, \ldots, \mathbf{A}_{n-1}$ to be permutation matrices and $\mathbf{A}_n$ to be an identity matrix. In such a case replacement node has to access $\frac{l}{d-k+1}$ symbols from each of d helper nodes to accomplish the node repair.

In such a case we can determine the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_{n-1}$ as follows:

$$\mathbf{A}_i = \sum_{a=0}^{l-1} \lambda_{i,a_i} \mathbf{e}_a \mathbf{e}_{a(i,a_i+1 \bmod r)}^t, \; i = 1, \ldots, n-1, \qquad (3)$$

where a_i denotes the i-th element from the right in r-ary representation $(a_{n-1}, \ldots, a_1)$ of a. By $a(i, u)$ we define a decimal element that coincides with a in all positions of r-ary representation except position i that is equal to u. $\mathbf{e}_0, \ldots, \mathbf{e}_{l-1}$ is standard basis of $GF(q^l)$ over $GF(q)$. As elements λ_{i,a_i} let us take $\lambda_{i,0} = \gamma^i$ and $\lambda_{i,u} = 1$ for $u = \{1, 2, \ldots, r-1\}$.

To define node repair procedure let us determine $\beta_{i,u,t}$ as follows:

$$\beta_{i,u,0} = 0;$$

$$\beta_{i,u,t} = \prod_{v=u}^{u+(t-1) \bmod r} \lambda_{i,v}, \; t = \{1, \ldots, r-1\}, \qquad (4)$$

where $u = \{0, \ldots, r-1\}$ and $\lambda_{i,v}$ are defined above. The repair of node $i = \{1, 2, \ldots, n-1\}$ can be done by accessing l/r symbols $\{c_{j,a}: j \neq i, a_i = 0\}$ from the remaining $n-1$ nodes and solving the following equations

$$\beta_{i,a_i,t} c_{i,a(i,a_i+t \bmod r)} = \\ = -c_{n,a} - \sum_{j \neq i,n} \beta_{j,a_j,t} c_{j,a(j,a_j+t \bmod r)}. \qquad (5)$$

The repair of node n can be done by accessing l/r symbols $\{c_{j,a}: j \neq n, a_1 + \ldots + a_{n-1} = 0 \bmod r\}$ and solving the following equations

$$c_{n,a} = -\sum_{i=1}^{n-1} \beta_{i,a_i,t} c_{i,a(i,a_i+t \bmod r)}. \qquad (6)$$

The reconstruction of information content can be accomplished by connecting to the set of any k nodes and downloading all information from them. In such a case from equations (5) we can form the system to define the symbols from the remaining $n-k$ nodes and recover users information as symbols from the first k nodes.

Eavesdropper model

In this paper, we consider a mathematical model of eavesdropper that can download up to t elements from each node in the set-up of the previous section. In other words, it means that $\mathbf{E}$ can accessed elements $\mathbf{C}_{i,E_i}$ where $E_i \subseteq [n]$, $(|E_i| < t+1)$ from each column vector $\mathbf{C}_i$ that represents the content stored on node i. We are focused on resistance against eavesdropper from an information-theoretic point of view that means that $\mathbf{E}$ does not gain any information about stored content $\mathbf{S}$ or, in other words, the mutual information between stored content and elements obtained from all servers by $\mathbf{E}$ is equal to zero. This can be written as

$$I(\mathbf{S}; \mathbf{C}_1, \ldots, \mathbf{C}_n) = 0. \qquad (7)$$

In information-theoretic approach we typically mix stored data with random symbols taken uniformly and independent from the same alphabet. There are two common ways to do it within distributed storage set up. The first of them is directly mixing information and random symbols utilizing storage codes. Note that typically it requires additional properties from code but allows to work within the same field. Another one is precoding information and random symbols by maximum rank distance codes, for example, Gabidulin code. In this paper we modify the last approach for our eavesdropper model, namely we encode information content of each node by Reed — Solomon

based scheme that allows recovering part of information content by accessing a limited number of symbols.

It's important to understand the bound on a message size that can be stored in such a system in presence of a given eavesdropper. In paper [17] by information-theoretic argument, it was proven that the number of information symbols B^s stored by regenerating code can be upper bounded as follows

$$B^s \le \sum_{i=1}^{k} \min(l-t,\ (d-i+1)\beta). \qquad (8)$$

Achieving equality in the bound (8) for a given B^s, k, d and t leads to the tradeoff between the repair bandwidth $d\beta$ and the sub-packetization level l. Let us explicitly find the values of MSR point that correspond to the case of minimizing l first and β after it. The corresponding relaxed optimization problem can be stated as

$$l^{\circ}(d,\ \beta) = \min l,$$
$$subject\ to: \sum_{i=1}^{k} \min\left(l-t,\ \left(1-\frac{i-1}{d}\right)d\beta\right) \ge B^s. \qquad (9)$$

Let us introduce $b_i = \left(1-\frac{k-i}{d}\right)d\beta$ and rewrite (9) as

$$l^{\circ}(d,\ \beta) = \min l,$$
$$subject\ to: \sum_{i=1}^{k} \min(l-t,\ b_i) \ge B^s. \qquad (10)$$

It can be easily seen that $C(l) = \sum_{i=1}^{k} \min(b_i,\ l-t)$ is a piecewise-linear function of l and has the following form:

$$c(1) = \begin{cases} 0 & l \in [0;\ t] \\ k(l-t) & l \in [t;\ b_1+t] \\ \vdots & \\ b_1+\ldots+b_{k-1}+l-t & l \in [b_{k-1}+t;\ b_k+t] \\ b_1+\ldots b_k & l \in (b_k+t,\ \infty) \end{cases}. \qquad (11)$$

This function is strictly monotone increasing on the segment $l \in [0,\ b_k+t]$. To find the extreme point of l such that $C(l) \ge B^s$ we simply take $l = C^{-1}(B^s)$ for the first non-zero value of $C(l)$, where $C^{-1}(*)$ is the inverse function of C. As a result, we receive

$$l = \frac{B^s}{k} + t. \qquad (12)$$

In this case $B^s = kb_1$ that leads to

$$\beta = \frac{B^s}{k(d-k+1)}. \qquad (13)$$

By the similar argument for MBR case we have

$$l = d\beta + t,$$
$$\beta = \frac{2B^s}{k(2d-k+1)}. \qquad (14)$$

Note that this optimization is the main focus of this paper. We shall say that code resistant against eavesdropper is MSR if its parameters coincide with (12) and (13). To construct it we modify Construction 1 of MSR codes without eavesdropper resistance.

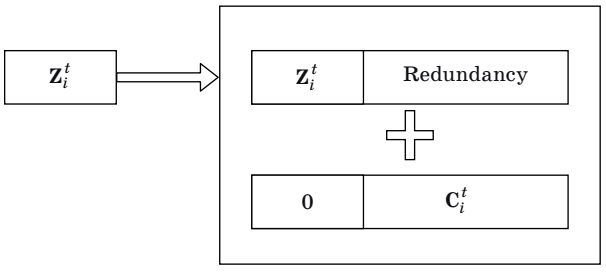
MSR-array codes resistant against eavesdropper

Let us construct MSR-array code resistant against eavesdropper with optimal access property utilizing previously introduced framework. For content $\mathbf{C}_i$ of each node i obtained by Construction 1 let us apply the modified safety scheme based on Reed — Solomon code which was introduced by the first time in paper [19]. Also, we mention paper [20] in which the similar schemes were investigated from another point of view. In that follows to ensure existence of Reed — Solomon codes we assume that we are working in $GF(q)$ with $q > \max(l+t,\ n)$. This scheme is depicted Figure.

In it, we first encode t uniformly and independently distributed random symbols $\mathbf{Z}_i^t = (z_{i,0},\ \ldots,\ z_{i,t-1})$ by systematic Reed — Solomon code of length $l+t$. After it, we add to the last l positions elements $\mathbf{C}_i^t = (c_{i,0},\ \ldots,\ c_{i,l-1})$ of the corresponding node. Defining the obtained row as $\mathbf{Y}_i^t = (y_{i,0},\ \ldots,\ y_{i,t+l-1})$ by the same argument as in [19] we can formally prove that

$$I(\mathbf{Y}_{i,E_i};\ \mathbf{C}_i) = 0 \qquad (15)$$

for any set of $E_i \subseteq [l+t]$ such that ($|E_i| < t+1$).



■ Safety scheme based on Reed — Solomon code

Remark 1. This fact can be understood form the point of view that in Reed — Solomon code any $t - 1$ or less code symbols does not give any information about stored content.

To recover any $0 < r < l + 1$ symbols of $\mathbf{C}_i$ we need to access the first t elements of $\mathbf{Y}_i$ that corresponds to $\mathbf{Z}_i$, encode them by the same Reed — Solomon code and substract necessary redundancy bits from corresponding elements of $\mathbf{Y}_i$. Based on these facts we can formulate the following theorem.

Theorem 1. Let $GF(q)$ be a finite field with $q > \max(l + t, n)$. Then MSR-array code of length n, sub-packetization level $l + t$, number of helper nodes $d = n - 1$ and number of nodes necessary to recover information content k resistant against eavesdropper with access to up to t symbols from any node can be defined by column vectors $\mathbf{Y}_i$. Each $\mathbf{Y}_i$ is formed from vectors $\mathbf{C}_i$ of array-codes from Construction 1 by the modified safety scheme based on Reed — Solomon code with independent and uniformly distributed random symbols $\mathbf{Z}_i$ for each node i.

Proof. From properties of used securing scheme we can write that $I(\mathbf{Y}_{i,E_i}; \mathbf{C}_i) = 0$ for any given $\mathbf{C}_i$ where $E_i \subseteq [l + t]$, ($|E_i| < t + 1$) defines the set of elements from node i available for the eavesdropper. As it holds for any given $\mathbf{C}_i$ and random symbols $\mathbf{Z}_i$ are independent, the elements $\mathbf{Y}_{i,E_i}$ are distributed uniformly and independent over all vectors of length $|E_i|$ over given field $GF(q)$. The last fact leads to $I(\mathbf{C}_j; \mathbf{Y}_{i,E_i}) = 0$. The resistance against eavesdropper means that $I(\mathbf{S}; \mathbf{Y}_{1,E_1}, \ldots, \mathbf{Y}_{n,E_n}) = 0$. As there is a bijection mapping between $\mathbf{C}_1$, ..., $\mathbf{C}_n$ and $\mathbf{S}$ this condition can be reformulated as $I(\mathbf{C}_1, \ldots, \mathbf{C}_n; \mathbf{Y}_{1,E_1}, \ldots, \mathbf{Y}_{n,E_n}) = 0$. Applying the facts above and the chain rule we can easily receive that

$$I(\mathbf{C}_1, \ldots, \mathbf{C}_n; \mathbf{Y}_{1,E_1}, \ldots, \mathbf{Y}_{n,E_n}) = H(\mathbf{Y}_{1,E_1}, \ldots, \mathbf{Y}_{n,E_n}) - \sum_{i=1}^{n} H(\mathbf{Y}_{i,E_i} | \mathbf{C}_i) \le \sum_{i=1}^{n} I(\mathbf{Y}_{i,E_i}; \mathbf{C}_i) = 0. \quad (16)$$

As node repair in Construction 1 is accomplished by downloading l/r symbols from each of $\mathbf{C}_i$ in our construction the replacement node can connect to first t symbols from each $\mathbf{Y}_i$. After it compute the redundancy of Reed — Solomon code, substract it from symbols of $\mathbf{Y}_i$ corresponding to symbols of $\mathbf{C}_i$ necessary for recovery and download only them. In such a case the repair bandwidth as well as sub-packetization level meets the corresponding extreme values (12) and (13). Note that after obtaining the content of failed node the replacement node has to apply to it safety scheme based on Reed — Solomon code.

Reconstruction of information content can be performed in the same way as in Construction 1. The only difference is that after connecting to k servers and downloading all information from them we have to compute the redundancy of Reed — Solomon code that encodes first t symbols and substract it from the last l symbols to obtain corresponding $\mathbf{C}_i$. This ends proof.

Example

To illustrate the proposed framework let us consider the following example. Let us consider $GF(8)$ constructed over primitive polynomial $\varphi(x) = x^3 + x + 1$ with root α. As array-code from Construction 1 let us take code with $n = 3$, $k = 1$, $r = 2$, $l = 4$. The first node stores information symbols while the last two nodes store parity-check symbols. Matrices that form parity-check equations (2) can be written as

$$A_1 = \begin{bmatrix} 0 & \alpha & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha \\ 0 & 0 & 1 & 0 \end{bmatrix}; A_2 = \begin{bmatrix} 0 & 0 & \alpha^2 & 0 \\ 0 & 0 & 0 & \alpha^2 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix};$$

$$A_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (17)$$

If we take as $\mathbf{C}_1 = (1, 1, 1, 1)^t$ when $\mathbf{C}_2 = (\alpha^3, 0, 0, 0)^t$ and $\mathbf{C}_3 = (\alpha, 1, 1, 1)^t$. Let us make obtained code resistant to eavesdropper by the scheme described in the previous section. In such a case, we receive $\mathbf{Y}_1 = (\alpha^6, \alpha^6, \alpha^2, 0, \alpha^6)^t$, $\mathbf{Y}_2 = (\alpha^5, 1, \alpha^5, \alpha^6, \alpha)^t$, $\mathbf{Y}_3 = (\alpha^3, \alpha^4, \alpha, \alpha^5, \alpha^2)^t$. If we have to recover the content of the first node from the remaining one we have to access $\mathbf{Y}_{2,\{0,1,3\}}$ and $\mathbf{Y}_{3,\{0,1,3\}}$. After it, we can find the redundancy of (5, 1) systematic Reed — Solomon code for information symbols $\mathbf{Y}_{2,0}$ and $\mathbf{Y}_{3,0}$. Receiving $(\alpha, \alpha^5, \alpha^6, \alpha)$ and $(\alpha^6, \alpha^3, \alpha^4, \alpha^6)$ as well as corresponding positions from $\mathbf{Y}_{2,\{1,3\}}$ and $\mathbf{Y}_{3,\{1,3\}}$ we obtain $\mathbf{C}_{2,\{0,2\}}$ and $\mathbf{C}_{3,\{0,2\}}$ that form the following parity-check equations

$$\begin{aligned} c_{1,0} + c_{2,0} + c_{3,0} &= 0; \\ \alpha c_{1,1} + \alpha^2 c_{2,2} + c_{3,0} &= 0; \\ c_{1,2} + c_{2,2} + c_{3,2} &= 0; \\ \alpha c_{1,3} + c_{2,0} + c_{3,2} &= 0 \end{aligned} \quad (18)$$

and determine $\mathbf{C}_1 = (1, 1, 1, 1)^t$. After it we have to apply to it introduced safety scheme and obtain $\tilde{\mathbf{Y}}_1 = (\alpha^2, \alpha^4, \alpha^6, \alpha, \alpha^4)^t$.

Conclusion

In this paper, we considered the new mathematical model of passive eavesdropper that has limited access to symbols from each node. We obtained the parameters of regenerating codes reaching extreme points of corresponding bound on the size of the stored message. Also, we proposed the construction of MSR-array codes resistant against the eavesdropper and illustrated the obtained construction by the corresponding example. In further research, we will consider the hybrid eavesdropper model that has a limited access to all nodes together with full access to a small subset of them.

Funding

The reported study was funded by RFBR, projects no. 19-01-00364, 19-37-90022, 20-07-00652 and joint RFBR and JSPS project no. 20-51-50007.

Acknowledgments

Author thanks A. Frolov and G. Kabatiansky for introducing this problem to him and numerous fruitful discussions during work on this paper.

References

1. Aftab U., Siddiqui G. F. Big data augmentation with data warehouse: A survey. *2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018, pp. 2785–2794. doi:10.1109/BigData.2018.8622206
2. Chun B.-G., Dabek F., Haeberlen A., Sit E., Weatherspoon H., Kaashoek M. F., Kubiatowicz J., Morris R. Efficient replica maintenance for distributed storage systems. *3rd Symposium on Networked Systems Design & Implementation*, USENIX Association, 2006, pp. 45–58.
3. Balaji S. B., Krishnan M. N., Vajha M., et al. Erasure coding for distributed storage: an overview. *Science China Information Sciences*, 2018, vol. 61, pp. 1–45. doi:10.1007/s11432-018-9482-6
4. Kruglik S., Frolov A. An information-theoretic approach for reliable distributed storage systems. *Journal of Communications Technology and Electronics*, 2020, vol. 65, no. 12, pp. 1505–1516. doi:10.1134/S1064226920120116
5. Yekhanin S. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, 2012, vol. 6, no. 3, pp. 139–255. doi:10.1561/0400000030
6. Dimakis A. G., Godfrey P. B., Wu Y., Wainwright M. J., Ramchandran K. Network coding for distributed storage systems. *IEEE Transactions on Information Theory*, 2010, vol. 56, no. 9, pp. 4539–4551. doi:10.1109/TIT.2010.2054295
7. Han Y. S., Pai H., Zheng R., Varshney P. K. Update-efficient error-correcting product-matrix codes. *IEEE Transactions on Communications*, 2015, vol. 63, no. 6, pp. 1925–1938. doi:10.1109/TCOMM.2015.2424416
8. Lin S., Chung W. Novel repair-by-transfer codes and systematic exact-MBR codes with lower complexities and smaller field sizes. *IEEE Transactions on Parallel and Distributed Systems*, 2014, vol. 25, no. 12, pp. 3232–3241. doi:10.1109/TPDS.2013.2297109
9. Rashmi K. V., Shah N. B., Kumar P. V. Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction. *IEEE Transactions on Information Theory*, 2011, vol. 57, no. 8, pp. 5227–5239. doi:10.1109/TIT.2011.2159049
10. Li J., Tang X., Tian C. A generic transformation for optimal repair bandwidth and rebuilding access in MDS codes. *2017 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 1623–1627. doi:10.1109/ISIT.2017.8006804
11. Ye M., Barg A. Explicit constructions of high-rate MDS array codes with optimal repair bandwidth. *IEEE Transactions on Information Theory*, 2017, vol. 63, no. 4, pp. 2001–2014. doi:10.1109/TIT.2017.2661313
12. Kadhe S., Sprintson A. Security for minimum storage regenerating codes and locally repairable codes. *2017 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 1028–1032. doi:10.1109/ISIT.2017.8006684
13. Rawat A. S., Koyluoglu O. O., Silberstein N., Vishwanath S. Secure locally repairable codes for distributed storage systems. *2013 IEEE International Symposium on Information Theory*, IEEE, 2013, pp. 2224–2228. doi:10.1109/ISIT.2013.6620621
14. Agarwal A., Mazumdar A. Security in locally repairable storage. *IEEE Transactions on Information Theory*, 2016, vol. 62, no. 11, pp. 6204–6217. doi:10.1109/TIT.2016.2605118
15. Bian J., Luo S., Li Z., Yang Y. Optimal weakly secure minimum storage regenerating codes scheme. *IEEE Access*, 2019, vol. 7, pp. 151120–151130. doi:10.1109/ACCESS.2019.2947248
16. Ozarow L. H., Wyner A. D. Wire-tap channel II. *AT&T Bell Lab Technical Journal*, 1984, vol. 63, pp. 2135–2157. doi:10.1002/j.1538-7305.1984.tb00072.x
17. Rashmi K. V., Shah N. B., Ramchandran K., Kumar P. V. Information-theoretically secure erasure codes for distributed storage. *IEEE Transactions on Information Theory*, 2018, vol. 64, no. 3, pp. 1621–1646. doi:10.1109/TIT.2017.2769101
18. Kruglik S. Secure MBR array codes in the presence of special type eavesdropper.

19. *Internet of Things, Smart Spaces, and Next Generation Networks and Systems. NEW2AN 2020, ruSMART 2020. Lecture Notes in Computer Science*, Springer, 2020, vol. 12526, pp. 1–11. doi:10.1007/978-3-030-65729-1_5
20. Huang W. Coding for security and reliability in distributed system Ph.D. dissertation, California Institute of Technology, 2017. Available at: paradise.caltech.edu/papers/thesis016.pdf (accessed 16 January 2021).
21. Holzbaur L., Kruglik S., Frolov A., Wachter-Zeh A. Secrecy and accessibility in distributed storage. *2020 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2020, pp. 1–6. doi:10.1109/GLOBECOM42002.2020.9322434

УДК 004.056.53
doi:10.31799/1684-8853-2021-1-38-44

Коды с минимальным хранением, устойчивые к атакам специального типа

С. А. Круглик[а,б], младший научный сотрудник, orcid.org/0000-0001-9557-5197, stanislav.kruglik@skoltech.ru
[а]Сколковский институт науки и технологий, Большой б-р, 30, стр. 1, Москва, 121205, РФ
[б]Московский физико-технический институт, Институтский пер., 9, Долгопрудный, Московская обл., 141701, РФ

Введение: для борьбы с временным или постоянным выходом из строя серверов распределенной системы хранения информации применяются специальные классы кодов, исправляющих стирания. Данные коды позволяют восстановить информацию с временно недоступного узла путем скачивания малого объема информации с других узлов. При этом возникают угрозы защищенности хранимых данных. **Цель:** введение новой математической модели, в которой злоумышленник имеет доступ к небольшому числу символов с каждого узла, и разработка соответствующих кодов, устойчивых к атакам злоумышленника. **Методы:** теоретико-информационный анализ и перемешивание информационных символов со случайными с помощью систематического кода Рида — Соломона. **Результаты:** введена новая математическая модель злоумышленника в распределенной системе хранения информации, имеющего доступ к малому числу символов с каждого узла. Отметим, что рассматривается модель пассивного злоумышленника — «подслушивателя», не способного каким-либо образом видоизменять полученные им данные. Найдены характеристики оптимальных кодов, устойчивых к выходу из строя серверов в распределенной системе хранения информации при наличии злоумышленника, а также построены оптимальные коды-массивы с минимальным хранением, устойчивые к атакам такого рода. **Практическая значимость:** представленная конструкция позволяет сохранить защищенность данных при обеспечении эффективного восстановления пользовательской информации.

Ключевые слова — распределенная система, коды-массивы с минимальным хранением, восстановление недоступного узла, математическая модель системы, устойчивость к действиям злоумышленника.

Финансовая поддержка

Исследование выполнено при поддержке РФФИ в рамках научных проектов № 19-01-00364, 19-37-90022, 20-07-00652, а также РФФИ и ЯОПН в рамках научного проекта № 20-51-50007.

Для цитирования: Kruglik S. A. Minimum-storage regenerating codes resistant to special adversary. *Информационно-управляющие системы*, 2021, № 1, с. 38–44. doi:10.31799/1684-8853-2021-1-38-44
For citation: Kruglik S. A. Minimum-storage regenerating codes resistant to special adversary. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 38–44. doi:10.31799/1684-8853-2021-1-38-44

УДК 623.612
doi:10.31799/1684-8853-2021-1-45-54

Научные статьи
Articles

Потенциальная помехоустойчивость когерентного приема четырехпозиционного фазоманипулированного радиосигнала в присутствии когерентной гармонической помехи

В. В. Звонарев[a]***,*** *канд. техн. наук, начальник лаборатории, orcid.org/0000-0003-1172-2239, zvonarevvitalii@yandex.ru*
А. С. Попов[a]***,*** *доктор техн. наук, профессор, orcid.org/0000-0001-5962-0587*
[a]*Военно-космическая академия им. А. Ф. Можайского, Ждановская наб., 13, Санкт-Петербург, 197198, РФ*

Введение: *использование известных методик расчета помехоустойчивости приема радиосигнала при наличии, например, гармонических помех приводит к существенно отличающимся численным значениям. Каждая подобная методика расчета имеет свой алгоритм вывода результирующей формулы, и эти выводы основаны на уровне «инженерной строгости».* ***Цель:*** *на основе линейного преобразования координат разработать корректную методику расчета вероятности ошибки корреляционного приема четырехпозиционного фазоманипулированного радиосигнала в присутствии когерентной гармонической помехи.* ***Методы:*** *представление четырехмерной плотности вероятности вектора выходных напряжений корреляторов демодулятора в четырехкратном интеграле произведением одномерных плотностей вероятности в пространстве собственных векторов ковариационной матрицы, в котором две плотности вероятности представляют собой дельта-функции Дирака. Четырехкратный интеграл приводится к двукратному с новыми пределами интегрирования, определяемыми из уравнений плоскостей, ограничивающих область интегрирования в этом пространстве.* ***Результаты:*** *выполнен вывод формул для точного расчета средних вероятностей символьных и битовых ошибок когерентного приема четырехпозиционного фазоманипулированного радиосигнала в присутствии когерентной гармонической помехи. По выведенным точным формулам построены графики зависимостей средних вероятностей символьных и битовых ошибок от отношения сигнал/шум для заданного отношения помеха/шум и заданного сдвига фазы помехи относительно фазы сигнала. Исследовано влияние энергетических соотношений сигнала и помехи, а также фазового сдвига помехи на вероятности символьной и битовой ошибки. Установлено, что влияние неэнергетического параметра эквивалентно приводит к изменению энергетических соотношений.* ***Практическая значимость:*** *результаты могут быть использованы при оценке эффективности связи в условиях воздействия помех. Применение разработанной методики позволит точно определить энергетические характеристики радиоканала, обеспечивающие требуемое качество приема передаваемых сообщений при наличии гармонической помехи.*

Ключевые слова *— вероятность символьной и битовой ошибки, гармоническая помеха, четырехпозиционная фазовая манипуляция, помехоустойчивость.*

Для цитирования: Звонарев В. В., Попов А. С. Потенциальная помехоустойчивость когерентного приема четырехпозиционного фазоманипулированного радиосигнала в присутствии когерентной гармонической помехи. *Информационно-управляющие системы*, 2021, № 1, с. 45–54. doi:10.31799/1684-8853-2021-1-45-54

For citation: Zvonarev V. V., Popov A. S. Potential interference immunity of coherent reception of quadruple phase-manipulated radio signal in the presence of coherent harmonic interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 45–54 (In Russian). doi:10.31799/1684-8853-2021-1-45-54

Введение

Определение помехоустойчивости приема радиосигнала с наиболее применяемыми видами манипуляции при наличии помех различного характера является одной из основных задач в теории и практике передачи дискретных сообщений в инфокоммуникационных радиосистемах, особенно для фазоманипулированных полосно-эффективных сигналов [1–14]. Формализация постановки задачи и методы ее решения могут существенно отличаться в зависимости от принятого алгоритма обработки сигнала и структуры системы модулятор-демодулятор. Истинность решения определяется правильной постановкой задачи, адекватностью моделей сигнала и помехи, а также корректностью использования математического аппарата в методике ее решения.

Радиосигналы с четырехпозиционной фазовой манипуляцией (ФМ-4) находят широкое применение в современных цифровых навигационных, спутниковых и других коммуникационных системах. Наиболее обоснованными и употребительными показателями качества передачи цифровой информации при этом являются средние вероятности символьной и битовой ошибки приема [3, 4, 6, 7, 11, 14], которые представляют собой, как известно, дополнения к единице для вероятностей правильного приема или достоверности приема. Задачи расчета помехоустойчивости приема сигналов с многопозиционной фазовой манипуляцией в присутствии эффективной для

ФМ-4 [13, 15] гармонической помехи решались во многих работах, например [11, 12, 16–19]. Однако в настоящее время задача получения точной формулы для вероятностей символьной и битовой ошибок остается актуальной.

Для расчета помехоустойчивости когерентного приема радиосигнала ФМ-4 в присутствии когерентной гармонической помехи рассмотрен упрощенный вид сигнала ФМ-4 на k-м тактовом интервале длительности T. Модель рассматриваемых сигнала и помехи можно записать следующим образом:

$$s_i(t) = A_c \cos(\omega_0 t + \varphi_i),\ \varphi_i = (i-1)\frac{\pi}{2},$$
$$t \in \left[(k-1)T,\ kT\right],\ i \in \{1,\ 2,\ 3,\ 4\};$$
$$s_п(t) = A_п \cos(\omega_0 t + \varphi_п),$$

где $A_c = \sqrt{2P_c}$ — амплитуда сигнала; P_c — мощность сигнала; $A_п = \sqrt{2P_п}$ — амплитуда помехи; $P_п$ — мощность гармонической помехи; ω_0 — частота несущего колебания; $\varphi_п$ — сдвиг фазы помехи относительно фазы сигнала.

На вход демодулятора поступает аддитивная смесь сигнала и помех [1–6, 11]:

$$u^j(t) = s_j(t) + s_п(t) + n(t).$$

Здесь j — номер позиции принятого информационного символа; $n(t)$ — шумовая помеха, моделируемая белым гауссовым шумом (БГШ) с корреляционной функцией:

$$\langle n(t)n(t')\rangle = \frac{N_0}{2}\delta(t-t'),$$

где N_0 — односторонняя спектральная плотность БГШ; $\delta(t - t')$ — дельта-функция Дирака [12, 20].

Вывод формулы для расчета помехоустойчивости когерентного приема радиосигнала ФМ-4 при воздействии когерентной гармонической помехи

В соответствии с алгоритмом демодуляции напряжение на выходе i-го коррелятора, представленного на рис. 1, при приеме j-й позиции сигнала ξ_i^j в момент отсчета T вычисляется по формуле [20, 21]

$$\xi_i^j = \frac{1}{N_0}\int_0^T \left[s_j(t) + s_п(t) + n(t)\right] s_i(t)\,\mathrm{d}t =$$
$$= \frac{1}{N_0}\int_0^T s_j(t)s_i(t)\,\mathrm{d}t + \frac{1}{N_0}\int_0^T s_п(t)s_i(t)\,\mathrm{d}t +$$
$$+ \frac{1}{N_0}\int_0^T n(t)s_i(t)\,\mathrm{d}t,$$

где $i, j \in \{1, 2, 3, 4\}$.

Необходимо отметить, что помеховая и случайная составляющие величины ξ_i^j не зависят от номера (позиции) информационного символа принимаемого сигнала, а определяются только номером коррелятора i. Значение случайной составляющей вычисляется по формуле

$$\zeta_i = \frac{1}{N_0}\int_0^T n(t)s_i(t)\,\mathrm{d}t.$$

Правило решения при оптимальном приеме многопозиционного ФМ-радиосигнала в данном случае может быть представлено в следующем формализованном виде [20]:

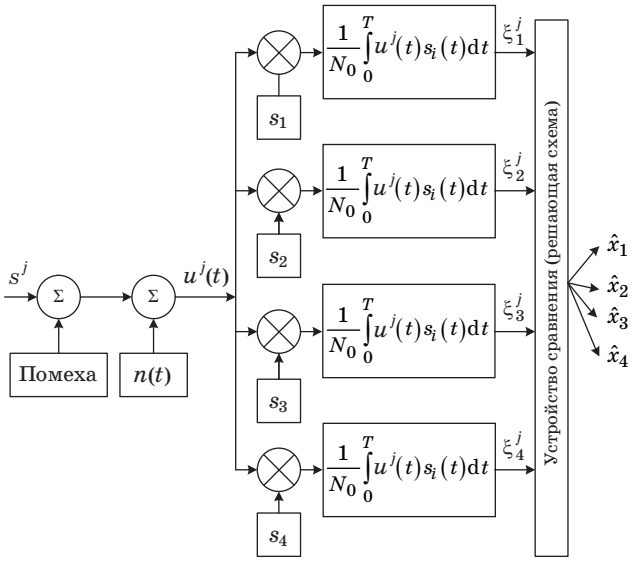
$$\xi_j^j > \xi_i^j;\ j \neq i;\ i, j \in \{1,\ 2,\ 3,\ 4\}.$$

Напряжения с выходов вычислителей-интеграторов сравниваются в решающем устройстве (см. рис. 1), и выбирается номер вычислителя с максимальным напряжением в момент отсчета T.

Обозначим вектор напряжений с выходов корреляторов следующим образом:

$$\boldsymbol{\xi}^j = \left(\xi_1^j,\ \xi_2^j,\ \xi_3^j,\ \xi_4^j\right)^{\mathrm{T}},$$

где индекс T — знак транспонирования.



■ ***Рис. 1.*** Упрощенная функциональная схема корреляционного демодулятора: s_1, s_2, s_3, s_4 — копии принимаемых радиосигналов информационных позиций

■ ***Fig. 1.*** The simplified functional diagram of the correlation demodulator: s_1, s_2, s_3, s_4 — copies of the accepted radio signals of information positions

Соответственно $\boldsymbol{\zeta}$ — вектор напряжений случайных составляющих вектора $\boldsymbol{\xi}^j$ — примет вид

$$\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \zeta_3, \zeta_4)^{\mathrm{T}}.$$

Математические ожидания вектора $\boldsymbol{\xi}^j$ можно представить в виде

$$\left\langle \boldsymbol{\xi}^j \right\rangle = \boldsymbol{\xi}^j - \boldsymbol{\zeta} = \frac{1}{N_0}\int_0^T s_j(t)\mathbf{s}\mathrm{d}t + \frac{1}{N_0}\int_0^T s_{\text{п}}(t)\mathbf{s}\mathrm{d}t. \quad (1)$$

Здесь $\mathbf{s} = (s_1, s_2, s_3, s_4)^{\mathrm{T}}$ — вектор копий принимаемых радиосигналов информационных позиций.

Тогда из (1) можно выразить вектор $\boldsymbol{\xi}^j$:

$$\boldsymbol{\xi}^j = \left\langle \boldsymbol{\xi}^j \right\rangle + \boldsymbol{\zeta}.$$

Нетрудно понять, что $\boldsymbol{\zeta}$ и $\boldsymbol{\xi}^j$ есть гауссовы случайные векторы. Отсюда совместная плотность вероятности (ПВ) составляющих вектора $\boldsymbol{\xi}^j$ есть четырехмерная гауссова: $\omega_{\xi^j}\left(\xi_1^j, \xi_2^j, \xi_3^j, \xi_4^j\right)$.

Общая формула многомерной ПВ $\omega_4(\boldsymbol{\xi}^j)$ вектора $\boldsymbol{\xi}^j$ в векторно-матричном представлении имеет вид [20]

$$\omega_4\left(\boldsymbol{\xi}^j\right) = \frac{1}{(2\pi)^{-2}\left|\mathbf{K}_{\xi^j}\right|}\times$$

$$\times \exp\left[-\frac{1}{2}\left(\boldsymbol{\xi}^j - \left\langle \boldsymbol{\xi}^j \right\rangle\right)^{\mathrm{T}} \mathbf{K}_{\xi^j}^{-1}\left(\boldsymbol{\xi}^j - \left\langle \boldsymbol{\xi}^j \right\rangle\right)\right],$$

где $\mathbf{K}_{\xi}^{j} = \left\langle \left(\boldsymbol{\xi}^j - \left\langle \boldsymbol{\xi}^j \right\rangle\right)\left(\boldsymbol{\xi}^j - \left\langle \boldsymbol{\xi}^j \right\rangle\right)^{\mathrm{T}} \right\rangle$ — ковариационная матрица вектора $\boldsymbol{\xi}^j$.

Как известно, вероятность того, что случайная величина ξ_j будет больше каждого из множества чисел $\{\xi_1, \ldots, \xi_{j-1}, \xi_{j+1}, \ldots, \xi_M\}$, равна

$$P\left\{\xi_j > \xi_i;\ j \neq i;\ j,\ i \in [1, \ldots, M]\right\} =$$

$$= \int_0^{\xi_j} \mathrm{d}\xi_1 \ldots \int_{-\infty}^{\xi_1} \mathrm{d}\xi_{j-1} \int_{-\infty}^{\infty} \mathrm{d}\xi_j \int_{-\infty}^{\xi_j} \mathrm{d}\xi_{j+1} \ldots \times$$

$$\times \int_{-\infty}^{\infty} \omega_{\xi^j}\left(\xi_1, \ldots, \xi_{j-1}, \xi_j, \xi_{j+1}, \ldots, \xi_M\right)\mathrm{d}\xi_M. \quad (2)$$

В данном случае $M = 4$.

Ковариационная матрица для всех индексов $j \in \{1, 2, 3, 4\}$ и определяется выражением

$$\mathbf{K}_{\xi^j} = \left\langle \left(\boldsymbol{\xi}^j - \left\langle \boldsymbol{\xi}^j \right\rangle\right)\left(\boldsymbol{\xi}^j - \left\langle \boldsymbol{\xi}^j \right\rangle\right)^{\mathrm{T}} \right\rangle = \left\langle \boldsymbol{\zeta}\boldsymbol{\zeta}^{\mathrm{T}} \right\rangle = \mathbf{K}_{\zeta}.$$

Если принимаемая смесь радиосигнала, помехи и шума $u^j(t) = s_j(t) + s_{\text{п}}(t) + n(t)$ содержит сигнал $s_j(t) = s(x_j, t)$, а значит, был передан информационный символ x_j, то вероятность (2) интерпретируется как вероятность правильного приема символа x_j, т. е.

$$P_{\text{пр}j} = P\left\{\xi_j^j > \xi_i^j;\ j \neq i;\ j,\ i \in [1, 2, 3, 4]\right\} = P\left(\hat{x}_j / x_j\right).$$

С учетом априорных вероятностей P_j передачи символов $\{x_1, x_2, x_3, x_4\}$ или сигналов $\{s_1(t), s_2(t), s_3(t), s_4(t)\}$ (при $j \in \{1, 2, 3, 4\}$) средняя вероятность ошибки приема символа находится по формуле полной вероятности:

$$P_{\text{ош.ср}} = \sum_{j=1}^{4} P_j\left(1 - P_{\text{пр}j}\right). \quad (3)$$

При априорной равновероятности формулу (3) можно представить в следующем виде:

$$P_{\text{ош.ср}} = \frac{1}{4}\sum_{j=1}^{4}\left(1 - P_{\text{прав}j}\right).$$

Для расчета средней вероятности ошибки приема символа необходимо определить:

— матрицу преобразования вектора $\boldsymbol{\xi}^j$ в новую систему координат, в которой ковариационная матрица $\mathbf{K}_{\xi}^{j}$ становится диагональной;

— вероятности правильного приема каждого информационного символа.

Определение матрицы преобразования в системе координат

Система сигналов в данной статье при ФМ-4 является биортогональной [1, 3, 6, 20, 21]. Ковариационная матрица $\mathbf{K}_{\zeta}$ является особенной (сингулярной, вырожденной). В соответствии с известным порядком вычисления ранга матрицы и ее определителя можно показать, что ранг матрицы равен двум, а ее определитель равен нулю ($\det \mathbf{K}_{\zeta} = |\mathbf{K}_{\zeta}| = 0$).

Поэтому интеграл (2) не может быть вычислен на ЭВМ [21]. При расчетах осуществляется переход из пространства вектора $\boldsymbol{\xi}$ в пространство собственных векторов $\boldsymbol{\eta}$ ковариационной матрицы $\mathbf{K}_{\zeta}$.

Матрица $\mathbf{K}_{\zeta}$ в данном случае вычисляется по формуле

$$\mathbf{K}_{\zeta} = \frac{h_{\text{с}}^2}{2}\mathbf{A},$$

где $h_{\text{с}}^2 = E_{\text{с}}/N_0$ — отношение сигнал/шум [1–5], $E_{\text{с}} = P_{\text{с}}T$ — энергия сигнала на длительности информационного символа, а матрица $\mathbf{A}$ имеет вид

$$\mathbf{A}=\begin{bmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\\ -1 & 0 & 1 & 0\\ 0 & -1 & 0 & 1\end{bmatrix}.$$

Собственные числа матрицы **A** находятся из известного определения характеристического уравнения матрицы **A** [15, 21] $\det(\mathbf{A}-\lambda\mathbf{I})=0$, имеющего вид $\left[(1-\lambda)^2-1\right]^2=0$. Решениями этого уравнения являются $\lambda_1=\lambda_2=0$ и $\lambda_3=\lambda_4=2$.

Ковариационная матрица в новой системе координат $\boldsymbol{\eta}$ будет иметь вид

$$\mathbf{K}_\eta=h_c^2\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}.$$

Найдем матрицу преобразования системы координат $\boldsymbol{\xi}$ в новую систему координат $\boldsymbol{\eta}$.

Собственные векторы матрицы **A** находятся из известного определения $(\mathbf{A}-\lambda\mathbf{I})\boldsymbol{\xi}=\mathbf{0}$. Один из удобных вариантов собственных векторов имеет вид

$$\xi_1=\begin{bmatrix}0\\1\\0\\1\end{bmatrix};\ \xi_2=\begin{bmatrix}1\\0\\1\\0\end{bmatrix};\ \xi_3=\begin{bmatrix}0\\1\\0\\-1\end{bmatrix};\ \xi_4=\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}.$$

Все векторы попарно ортогональны, т. е. $\xi_j^{\mathrm{T}}\xi_i=0$ при $j\neq i$, где $j, i\in\{1, 2, 3, 4\}$. Нормы всех векторов одинаковы: $\|\xi_j\|=\sqrt{2}$. Ортонормированные векторы имеют вид

$$\mathbf{v}_1=\begin{bmatrix}0\\\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{bmatrix};\ \mathbf{v}_2=\begin{bmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\\0\end{bmatrix};\ \mathbf{v}_3=\begin{bmatrix}0\\\frac{1}{\sqrt{2}}\\0\\-\frac{1}{\sqrt{2}}\end{bmatrix};\ \mathbf{v}_4=\begin{bmatrix}\frac{1}{\sqrt{2}}\\0\\-\frac{1}{\sqrt{2}}\\0\end{bmatrix}.$$

Таким образом, матрица преобразований **V** системы координат $\boldsymbol{\xi}$ в новую систему координат $\boldsymbol{\eta}$ будет иметь вид

$$\mathbf{V}=\begin{bmatrix}0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0\end{bmatrix}.$$

Непосредственным вычислением можно установить, что для матрицы **V** справедливы следующие равенства: $\mathbf{V}=\mathbf{V}^{\mathrm{T}}=\mathbf{V}^{-1}$. Следовательно, матрица преобразования Якоби равна самой матрице **V**. Преобразования системы координат $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ здесь симметричны, т. е. $\boldsymbol{\eta}=\mathbf{V}\boldsymbol{\xi}$ и $\boldsymbol{\xi}=\mathbf{V}\boldsymbol{\eta}$. Симметричны также преобразования математических ожиданий случайных векторов $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$, т. е. $\langle\boldsymbol{\eta}\rangle=\mathbf{V}\langle\boldsymbol{\xi}\rangle$ и $\langle\boldsymbol{\xi}\rangle=\mathbf{V}\langle\boldsymbol{\eta}\rangle$.

Расчет вероятности правильного приема информационных символов

Рассмотрим вектор $\langle\boldsymbol{\xi}^j\rangle$ для первой позиции, где $s_1(t)=A_c\cos\omega_0(t)$. Непосредственными вычислениями получим вектор математических ожиданий вектора $\boldsymbol{\xi}^1$ для первой позиции принимаемого сигнала:

$$\left\langle\boldsymbol{\xi}^1\right\rangle=\begin{bmatrix}h_c^2+h_ch_п\cos\varphi_п\\ h_ch_п\sin\varphi_п\\ -\left(h_c^2+h_ch_п\cos\varphi_п\right)\\ h_ch_п\sin\varphi_п\end{bmatrix},$$

где $h_п^2=E_п/N_0$ — отношение помеха/шум [1–5], $E_п=P_пT$ — энергия помехи на длительности информационного символа.

Используя матрицу преобразования **V**, получаем вектор математических ожиданий в новой системе координат $\left\langle\boldsymbol{\eta}^1\right\rangle$:

$$\left\langle\boldsymbol{\eta}^1\right\rangle=\begin{bmatrix}0\\0\\\sqrt{2}h_ch_п\sin\varphi_п\\ \sqrt{2}\left(h_c^2+h_ch_п\cos\varphi_п\right)\end{bmatrix}.$$

Отсюда видно, что вероятностная мера задана на двумерной плоскости $(\eta_3 0\eta_4)$ в новой системе координат $\boldsymbol{\eta}$. В этой системе координат четырехмерная ПВ $\omega_4(\boldsymbol{\eta})$ представляет собой произведение четырех одномерных ПВ, а именно:

$$\omega_4\left(\boldsymbol{\eta}^1\right)=\omega_1\left(\eta_1^1\right)\omega_1\left(\eta_2^1\right)\omega_1\left(\eta_3^1\right)\omega_1\left(\eta_4^1\right).$$

Первые два сомножителя имеют дисперсию, равную нулю, и представляют собой дельта-функцию Дирака:

$$\omega_1\left(\eta_1^1\right)=\delta\left(\eta_1^1-\left\langle\eta_1^1\right\rangle\right),\ \omega_1\left(\eta_2^1\right)=\delta\left(\eta_2^1-\left\langle\eta_2^1\right\rangle\right).$$

Кроме того, имеем $\left\langle\eta_1^1\right\rangle=\left\langle\eta_2^1\right\rangle=0$. Интегрирование дельта-функции в данном случае приводит

к единице, и четырехкратный интеграл становится двукратным.

В общем случае формулу (2) перепишем для обеих систем координат $\boldsymbol{\xi}^1$ и $\boldsymbol{\eta}^1$ при приеме первого информационного символа:

$$P_{\text{прав1}} = \int_{-\infty}^{\infty} d\xi_1 \int_{-\infty}^{\xi_1^1} d\xi_2 \int_{-\infty}^{\xi_1^1} d\xi_3 \int_{-\infty}^{\xi_1^1} \omega_4\left(\boldsymbol{\xi}^1\right) d\xi_4 =$$

$$= \int_{\eta_{1\text{н}}^1}^{\eta_{1\text{в}}^1} \omega_1\left(\eta_1^1\right) d\eta_1 \int_{\eta_{2\text{н}}^1}^{\eta_{2\text{в}}^1} \omega_1\left(\eta_2^1\right) d\eta_2 \times$$

$$\times \int_{\eta_{3\text{н}}^1}^{\eta_{3\text{в}}^1} \omega_1\left(\eta_3^1\right) d\eta_3 \int_{\eta_{4\text{н}}^1}^{\eta_{4\text{в}}^1} \omega_1\left(\eta_4^1\right) d\eta_4,$$

где $\eta_{i\text{н}}^1$ и $\eta_{i\text{в}}^1$ — нижний и верхний пределы интегрирования, которые надо найти.

Первые два сомножителя равны единице, поэтому выражение упрощается:

$$P_{\text{прав1}} = \int_{\eta_{3\text{н}}}^{\eta_{3\text{в}}} d\eta_3 \int_{\eta_{4\text{н}}}^{\eta_{4\text{в}}} \omega_1(\eta_3)\omega_1(\eta_4) d\eta_4. \qquad (4)$$

Найдем нижние и верхние пределы интегрирования выражения (4). Для этого необходимо определить уравнения плоскостей, ограничивающих область интегрирования в четырехмерной системе координат $\boldsymbol{\xi}$ [20]:

$$\xi_2 - \xi_1 = 0;\ \xi_3 - \xi_1 = 0;\ \xi_4 - \xi_1 = 0. \qquad (5)$$

Используя преобразования $\boldsymbol{\xi} = \mathbf{V}\boldsymbol{\eta}$, определяем значения составляющих вектора $\boldsymbol{\xi}$ через составляющие вектора $\boldsymbol{\eta}$. Тогда имеем:

$$\xi_1 = \frac{1}{\sqrt{2}}(\eta_2 + \eta_4);\ \xi_2 = \frac{1}{\sqrt{2}}(\eta_1 + \eta_3);$$

$$\xi_3 = \frac{1}{\sqrt{2}}(\eta_2 - \eta_4);\ \xi_4 = \frac{1}{\sqrt{2}}(\eta_1 - \eta_3).$$

Подставляя данные выражения в уравнение (5), получаем:

$$\begin{cases} \xi_2 - \xi_1 = \dfrac{1}{\sqrt{2}}(\eta_1 + \eta_3 - \eta_2 - \eta_4) = 0; \\ \xi_3 - \xi_1 = \dfrac{1}{\sqrt{2}}(\eta_2 - \eta_4 + \eta_2 - \eta_4) = 0; \\ \xi_4 - \xi_1 = \dfrac{1}{\sqrt{2}}(\eta_1 - \eta_3 - \eta_2 - \eta_4) = 0. \end{cases}$$

Получаем уравнения плоскостей в четырехмерном пространстве, ограничивающих область интегрирования в системе координат $\boldsymbol{\eta}$:

$$\begin{cases} \eta_4 = 0; \\ \eta_1 - \eta_2 + \eta_3 - \eta_4 = 0; \\ \eta_1 - \eta_2 - \eta_3 - \eta_4 = 0. \end{cases}$$

Далее найдем уравнения следов пересечения этих плоскостей на плоскости ($\eta_3 0 \eta_4$). Для этого приравниваем значения координат η_1 и η_2 нулю и получаем $\eta_3 = \eta_4$, $\eta_3 = -\eta_4$.

Пределы интегрирования для (4) имеют значения $\eta_{4\text{н}} = 0$, $\eta_{4\text{в}} = \infty$, $\eta_{3\text{н}} = -\eta_4$, $\eta_{3\text{в}} = \eta_4$. В формуле (4) одномерные ПВ определены выражениями

$$\omega_1(\eta_3) = \frac{1}{\sqrt{2\pi}h_\text{с}} \exp\left(-\frac{\left(\eta_3^1 - \sqrt{2}h_\text{с}h_\text{п}\sin\varphi_\text{п}\right)^2}{2h_\text{с}^2} \right);$$

$$\omega_1(\eta_4) = \frac{1}{\sqrt{2\pi}h_\text{с}} \exp\left(-\frac{\left(\eta_4^1 - \sqrt{2}\left(h_\text{с}^2 + h_\text{с}h_\text{п}\cos\varphi_\text{п}\right)\right)^2}{2h_\text{с}^2} \right).$$

Для получения окончательной формулы для расчета в одномерные ПВ необходимо ввести нормированные переменные [20]

$$x = \frac{\eta_4 - \langle\eta_4\rangle}{h_\text{с}};\ y = \frac{\eta_3 - \langle\eta_3\rangle}{h_\text{с}}.$$

Вычислим пределы интегрирования в переменных x и y, т. е. $x_\text{н}$, $x_\text{в}$, $y_\text{н}$, $y_\text{в}$:

$$\begin{cases} x_\text{н} = -\dfrac{\langle\eta_4\rangle}{h_\text{с}} = -\sqrt{2}\left(h_\text{с} + h_\text{п}\cos\varphi_\text{п}\right);\ x_\text{в} = \infty; \\ y_\text{н} = \dfrac{-\eta_4 - \langle\eta_3\rangle}{h_\text{с}} = -\dfrac{\eta_4}{h_\text{с}} - \sqrt{2}h_\text{п}\sin\varphi_\text{п}; \\ y_\text{в} = \dfrac{\eta_4 - \langle\eta_3\rangle}{h_\text{с}} = \dfrac{\eta_4}{h_\text{с}} - \sqrt{2}h_\text{п}\sin\varphi_\text{п}. \end{cases}$$

Определим слагаемое $\dfrac{\eta_4}{h_\text{с}}$ в удобном виде из выражения нормировки:

$$x = \frac{\eta_4 - \langle\eta_4\rangle}{h_\text{с}} \ldots \rightarrow \ldots \frac{\eta_4}{h_\text{с}} = x + \frac{\langle\eta_4\rangle}{h_\text{с}}.$$

Тогда можно показать, что нижняя $y_\text{н}$ и верхняя $y_\text{в}$ границы определяются следующими выражениями:

$$\begin{cases} y_\text{н} = -x - \sqrt{2}\left[h_\text{с} + h_\text{п}\left(\cos\varphi_\text{п} + \sin\varphi_\text{п}\right)\right]; \\ y_\text{в} = x + \sqrt{2}\left[h_\text{с} + h_\text{п}\left(\cos\varphi_\text{п} - \sin\varphi_\text{п}\right)\right]. \end{cases}$$

Нормированные ПВ для переменных x и y равны:

$$\begin{cases} \omega(x)=\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{x^2}{2}\right); \\ \omega(y)=\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{y^2}{2}\right). \end{cases}$$

Полученные значения ПВ подставляются в формулу (3) для вычисления вероятности правильного приема первого информационного символа $P_{\text{прав1}}$:

$$P_{\text{прав1}}=\frac{1}{2\pi}\int\limits_{-\sqrt{2}(h_{\text{с}}+h_{\text{п}}\cos\varphi_{\text{п}})}^{\infty}\mathrm{d}x\times$$
$$\times\int\limits_{-x-\sqrt{2}\left[h_{\text{с}}+h_{\text{п}}(\cos\varphi_{\text{п}}+\sin\varphi_{\text{п}})\right]}^{x+\sqrt{2}\left[h_{\text{с}}+h_{\text{п}}(\cos\varphi_{\text{п}}-\sin\varphi_{\text{п}})\right]}\exp\left(-\frac{x^2+y^2}{2}\right)\mathrm{d}y.$$

Выпишем формулы для расчета вероятности правильного приема второго, третьего и четвертого информационных символов, полученные по приведенной методике:

$$P_{\text{прав2}}=\frac{1}{2\pi}\int\limits_{-\sqrt{2}(h_{\text{с}}+h_{\text{п}}\sin\varphi_{\text{п}})}^{\infty}\mathrm{d}x\times$$
$$\times\int\limits_{-x-\sqrt{2}\left[h_{\text{с}}+h_{\text{п}}(\cos\varphi_{\text{п}}+\sin\varphi_{\text{п}})\right]}^{x+\sqrt{2}\left[h_{\text{с}}-h_{\text{п}}(\cos\varphi_{\text{п}}-\sin\varphi_{\text{п}})\right]}\exp\left(-\frac{x^2+y^2}{2}\right)\mathrm{d}y;$$

$$P_{\text{прав3}}=\frac{1}{2\pi}\int\limits_{-\sqrt{2}(h_{\text{с}}-h_{\text{п}}\cos\varphi_{\text{п}})}^{\infty}\mathrm{d}x\times$$
$$\times\int\limits_{-x-\sqrt{2}\left[h_{\text{с}}-h_{\text{п}}(\cos\varphi_{\text{п}}-\sin\varphi_{\text{п}})\right]}^{x+\sqrt{2}\left[h_{\text{с}}-h_{\text{п}}(\cos\varphi_{\text{п}}+\sin\varphi_{\text{п}})\right]}\exp\left(-\frac{x^2+y^2}{2}\right)\mathrm{d}y;$$

$$P_{\text{прав4}}=\frac{1}{2\pi}\int\limits_{-\sqrt{2}(h_{\text{с}}-h_{\text{п}}\sin\varphi_{\text{п}})}^{\infty}\mathrm{d}x\times$$
$$\times\int\limits_{-x-\sqrt{2}\left[h_{\text{с}}+h_{\text{п}}(\cos\varphi_{\text{п}}-\sin\varphi_{\text{п}})\right]}^{x+\sqrt{2}\left[h_{\text{с}}-h_{\text{п}}(\cos\varphi_{\text{п}}+\sin\varphi_{\text{п}})\right]}\exp\left(-\frac{x^2+y^2}{2}\right)\mathrm{d}y.$$

В частном случае, когда отношение помеха/шум будет равно нулю ($h_{\text{п}}=0$), имеем $P_{\text{прав1}}=P_{\text{прав2}}=P_{\text{прав3}}=P_{\text{прав4}}$, и формула для расчета средней вероятности правильного приема информационного символа примет следующий вид:

$$P_{\text{прав}}=\frac{1}{2\pi}\int\limits_{-\sqrt{2}h_{\text{с}}}^{\infty}\mathrm{d}x\int\limits_{-x-\sqrt{2}h_{\text{с}}}^{x+\sqrt{2}h_{\text{с}}}\exp\left(-\frac{x^2+y^2}{2}\right)\mathrm{d}y. \quad (6)$$

Таким образом, в отсутствие помехи средняя вероятность правильного приема символа равна вероятности одной позиции, например: $P_{\text{прав}}=P_{\text{прав1}}$.

Учитывая симметрию области определения второго интеграла, формулу (6) можно упростить и записать следующим образом:

$$P_{\text{прав}}=\frac{1}{\pi}\int\limits_{-\sqrt{2}h_{\text{с}}}^{\infty}\mathrm{d}x\int\limits_{0}^{x+\sqrt{2}h_{\text{с}}}\exp\left(-\frac{x^2+y^2}{2}\right)\mathrm{d}y. \quad (7)$$

Формула (7) может быть представлена и в другом, равносильном виде [21, 22]:

$$P_{\text{прав}}=\left[1-Q(h_{\text{с}})\right]^2, \quad (8)$$

где $Q(x)=\dfrac{1}{\sqrt{2\pi}}\int\limits_{x}^{\infty}\exp\left(-\dfrac{t^2}{2}\right)\mathrm{d}t$ — функция Гаусса.

Средняя по символам вероятность символьной ошибки приема определяется выражением

$$P_{\text{ош.симв}}=1-\frac{1}{4}\sum_{j=1}^{4}P_{\text{прав}j}.$$

Графики зависимостей $P_{\text{ош.симв}}(E_b/N_0)$, где E_b — энергия одного бита, при разных уровнях гармонической помехи $h_{\text{п}}$ и значениях ее фазового сдвига $\varphi_{\text{п}}$ относительно фазы несущего колебания сигнала построены на рис. 2. Для расчета $P_{\text{ош.симв}}(E_b/N_0)$ были приняты следующие исходные данные:

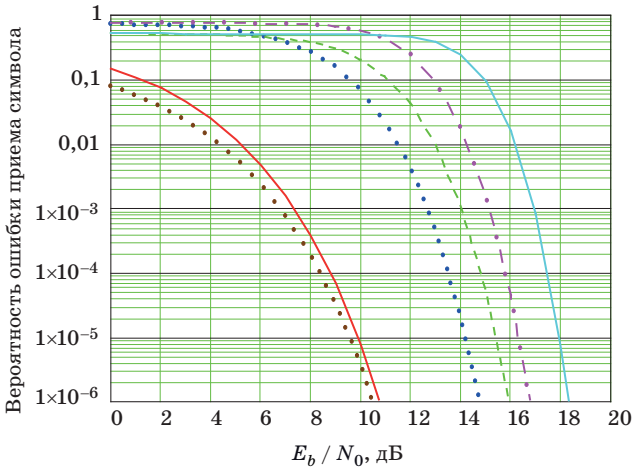
— отношение помеха/шум $h_{\text{п}}$ составило 0, 3 и 5 (в отсутствие помехи 9,5 и 14 дБ соответственно);

— сдвиг фазы помехи $\varphi_{\text{п}}$ равен 0 и $\pi/4$.

Фазовый сдвиг помехи может принимать различные, в том числе случайные, значения. Поэтому целесообразно рассмотреть также усредненные по фазовому сдвигу помехи $\varphi_{\text{п}}$ вероятности символьной и битовой ошибок. Примем, что случайная величина $\varphi_{\text{п}}$ имеет равномерное распределение в пределах $(-\pi, \pi)$.

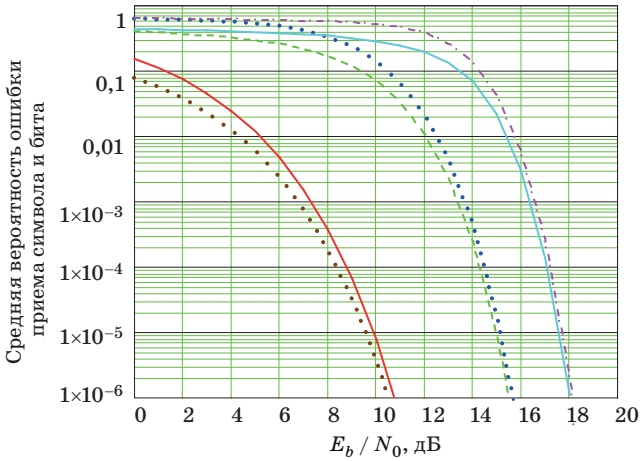
Формулы для расчета ошибок приема первого и второго битов, а также средней вероятности ошибки приема бита с учетом кода отображения Грея получены на основе использования матрицы переходных условных вероятностей [22] и построения вероятностного пространства на множестве совместных событий. Подробное изложение этой методики для QPSK занимает немалый объем и будет представлено в отдельной статье.

Графики зависимостей средней вероятности битовой и символьной ошибки от отношения сигнал/шум при заданных значениях отношения помеха/шум при тех же исходных данных, что и для рис. 2, но усредненные по $\varphi_{\text{п}}$, представлены на рис. 3.



■ ***Рис. 2.*** Зависимость вероятности ошибки приема символа от отношения сигнал/шум

■ ***Fig. 2.*** Dependences of the probability of a symbol reception error on the signal-to-noise ratio



■ ***Рис. 3.*** Зависимость средней вероятности битовой и символьной ошибки от отношения сигнал/шум

■ ***Fig. 3.*** Dependence of the average probability of bit and character errors on the signal-to-noise ratio

Как видно из графиков, увеличение фазового сдвига помехи приводит к увеличению вероятности ошибки приема символа, а увеличение отношения помеха/шум — к необходимости увеличивать отношение сигнал/шум для обеспечения требуемой средней вероятности ошибки приема символа и бита.

Заключение

Представленный расчет помехоустойчивости когерентного приема четырехпозиционного фазоманипулированного радиосигнала в присутствии когерентной гармонической помехи позволяет точно определить вероятность символьной и битовой ошибки при заданных значениях уровня помехи и сдвига фазы помехи относительно фазы сигнала и их усредненные по фазе помехи величины.

Наличие когерентной гармонической помехи существенно ухудшает качество приема информации, полностью нарушая прием информации. Кривые имеют пороговый характер при значениях уровня помехи $h_п \geq h_с$.

При некоторых значениях сдвига фазы помехи относительно фазы несущего колебания сигнала происходит заметное (до 2 дБ) увеличение вероятности символьной ошибки. Таким образом, показано, как промежуточное значение фазового сдвига несущего колебания помехи, не совпадающее с фазой информационной позицией сигнала, дополнительно увеличивает значение вероятности ошибки приема. Это позволяет утверждать, что влияние неэнергетического параметра эквивалентно приводит к изменению энергетических соотношений.

Представленные на рис. 3 результаты показывают, что когерентная гармоническая помеха больше оказывает влияние на среднюю по фазе помехи вероятность ошибки приема символа при заданных значениях отношения сигнал/шум.

Установлено, что вероятности ошибки приема первого и второго битов при QPSK, усредненные по фазовому сдвигу гармонической когерентной помехи при отображении по Грею, имеют одинаковую величину во всех точках кривой. Отсюда следует, что справедлива формула

$$P_{\text{ош.бит}} = 1 - \sqrt{P_{\text{прав.симв}}},$$

где $P_{\text{ош.бит}}$ — средняя вероятность битовой ошибки; $P_{\text{прав.симв}}$ — средняя вероятность правильного приема символа.

Влияние величины фазового сдвига несущего колебания двоичной фазоманипулированной помехи на качество приема сигнала рассматривалось в работе [22] с применением другой (векторной) модели сигнала и помехи.

Литература

1. **Беккиев А. Ю., Борисов В. И.** Оценка помехозащищенности каналов радиосвязи в условиях действия помех от средств радиоэлектронной борьбы. *Радиотехника и электроника*, 2019, т. 64, № 9, с. 891–901. doi:10.1134/S0033849419080035
2. **Куликов Г. В., Лелюх А. А., Граченко Е. Н.** Помехоустойчивость когерентного приемника сигналов с квадратурной амплитудной манипуляцией при наличии ретраслированной помехи. *Радиотехника и электроника*, 2020, т. 65, № 8, с. 804–808. doi:10.31857/S0033849420070074
3. **Бучинский Д. И., Вознюк В. В., Фомин А. В.** Исследование помехоустойчивости приемника сигналов с многопозиционной фазовой манипуляцией к воздействию помех с различной структурой. *Труды Военно-космической академии им. А. Ф. Можайского*, 2019, вып. 671, с. 120–127.
4. **Паршуткин А. В., Маслаков П. А.** Исследование помехоустойчивости современных стандартов спутниковой связи к воздействию нестационарных помех. *Труды СПИИРАН*, 2017, № 4(53), с. 159–177. doi:https://doi.org/10.15622/sp.53.8
5. **Бондаренко В. Н., Гарифуллин В. Ф., Краснов Т. В., Феоктистов Д. С., Богатырев Е. В.** Помехоустойчивость корреляционного приемника MSK-BOC сигнала к сосредоточенной помехе. *Успехи современной радиоэлектроники*, 2017, № 12, с. 71–74.
6. **Волхонская Е. В., Коротей Е. В., Власова К. В., Рушко М. В.** Модельное исследование помехоустойчивости приема радиосигналов с QPSK, BPSK, 8PSK и DBPSK. *Известия КГТУ*, 2017, № 46, с. 165–174.
7. **Chernoyarov O. V., Glushkov A. N., Litvinenko V. P., Litvinenko Yu. V., Matveev B. V.** Digital demodulator of the quadrature amplitude modulation signals. *Measurement Science Review*, 2018, vol. 18, no. 6, pp. 236–242. http://www.measurement.sk/2018/msr-2018-0032.pdf (дата обращения: 12.10.2019). doi:10.1515/msr-2018-0032
8. **Glushkov A. N., Litvinenko V. P., Matveev B. V., Chernoyarov O. V., Salnikova A. V.** Basic algorithm for the coherent digital processing of the radio signals. *2015 International Conference on Space Science & Communication*, Langkawi, Malaysia, IEEE, pp. 389–392. https://www.researchgate.net/publication/308835672_Basic_algorithm_for_the_coherent_digital_processing_of_the_radio_signals. (дата обращения: 10.05.2019). doi:10.1109/IconSpace.2015.7283834
9. **Ezeagwu C. O., Chukwuneke N. S., Emeka S. C., Ezeribe B. E.** Comparative study of the performance of different digital modulation schemes. *International Journal of Advances in Scientific Research and Engineering*, 2018, vol. 4, no. 3, pp. 149–154. doi:dx.doi.org/10.7324/IJASRE.2018.32639
10. **Anmol Kumar, Rajdeep Kaur.** PSO-based NBI resistant asynchronous MCCDMA multiuser detector. *International Journal of Intelligent Systems and Applications*, 2016, no. 10, pp. 60–67. doi:10.5815/ijisa.2016.10.07
11. **Куликов Г. В., Нгуен Ван Зунг, Нестеров А. В., Лелюх А. А.** Помехоустойчивость приема сигналов с многопозиционной фазовой манипуляцией в присутствии гармонической помехи. *Наукоемкие технологии*, 2018, № 11, с. 32–38. doi:10.18127/j19998465-201811-06
12. **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*, Екатеринбург, 2019, № 1 (79), ч. 1, с. 53–59. https://research-journal.org/technical/analiz-metodov-rascheta-bitovoj-veroyatnosti-oshibki-pri-kogerentnom-prieme-signalov-s-m-ichnoj-fazovoj-manipulyaciej/ (дата обращения: 10.05.2019). doi: https://doi.org/10.23670/IRJ.2019.79.1.009
13. **Абед А. Х., Жуков В. М.** Анализ помехоустойчивости радиостанции при воздействии организованных помех. *Вестник Тамбовского государственного технического университета*, 2016, т. 22, № 1, с. 53–57. doi:10.17277/vestnik.2016.01
14. **Савищенко Н. В., Лебеда Е. В.** Вероятности ошибки когерентного приема многопозиционных сигналов в канале с общими гамма- или *K*-замираниями и белым шумом. *Информационно-управляющие системы*, 2019, № 1, с. 76–88. doi:10.31799/1684-8853-2019-1-76-88
15. **Kirillov S. N., Lisnichuk A. A.** Analysis of narrow-band interference effect on cognitive radio systems based on synthesized four-position radio signals. *Proc. 14th International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering, APEIE 2018*, Novosibirsk, IEEE, 2018, pp. 50–54. https://ieeexplore.ieee.org/document/8545965 (дата обращения: 10.05.2019). doi:10.1109/APEIE.2018.8545965
16. **Сидоркина Ю. А., Сизых В. В., Шахтарин Б. И., Шевцев В. А.** Схема Костаса при воздействии аддитивных гармонических помех и широкополосного шума. *Радиотехника и электроника*, 2016, т. 61, № 7, с. 671–680. doi:10.7868/S003384941607010X
17. **Куликов Г. В., Нестеров А. В., Лелюх А. А.** Помехоустойчивость приема сигналов с квадратурной амплитудной манипуляцией в присутствии гармонической помехи. *Журнал радиоэлектроники*, 2018, № 11. http://jre.cplire.ru/jre/nov18/9/text.pdf (дата обращения: 10.05.2019). doi:10.30898/1684-1719.2018.11.9
18. **Ложкин К. Ю., Стиценко А. И.** Помехоустойчивость некогерентного и когерентного приема ДФРМ-сигнала в условиях воздействия фазоманипулированной, гармонической или гауссовской помех. *Журнал Сибирского федерального университета. Серия: Техника и технологии*, 2017, т. 10, № 2, с. 260–270. http://elib.sfu-kras.ru/bitstream/handle/

2311/32132/12_Lozhkin.pdf?sequence=1&isAllowed=y (дата обращения: 10.05.2019). doi:10.17516/1999-494X-2017-10-2-260-270
19. **Ложкин К. Ю.** Помехоустойчивость приема OFDM-сигнала с однократной фазовой манипуляцией и корректирующим кодированием на фоне полигармонической помехи. *Информация и космос*, 2018, № 2, с. 37–43.
20. **Звонарев В. В., Попов А. С.** Методика расчета потенциальной помехоустойчивости оптимального когерентного приема многопозиционного фазоманипулированного радиосигнала с белым шумом. *Радиотехника*, 2019, т. 83, № 4, с. 79–83. doi:10.18127/j00338486-201904-79
21. **Proakis J. G.** *Digital communications*. McGraw-Hill, Book Company, 1995. 800 p.
22. **Савищенко Н. В.** *Многомерные сигнальные конструкции: их частотная эффективность и потенциальная помехоустойчивость приема*. СПб., СПбГПУ, 2005. 420 с.

UDC 623.612
doi:10.31799/1684-8853-2021-1-45-54

Potential interference immunity of coherent reception of quadruple phase-manipulated radio signal in the presence of coherent harmonic interference

V. V. Zvonarev[a], PhD, Tech., Head of a Laboratory, orcid.org/0000-0003-1172-2239, zvonarevvitalii@yandex.ru
A. S. Popov[a], Dr. Sc., Tech., Professor, orcid.org/0000-0001-5962-0587
[a]A. F. Mozhaiskiy Military Space Academy, 13, Zhdanovskaia Emb., 197198, Saint-Petersburg, Russian Federation

Introduction: The known methods for calculating the interference immunity of radio signal reception in the presence of, for example, harmonic interference, often lead to significantly different numerical values. Each calculation technique of this type has its own algorithm for the resulting formula output, and these conclusions are based on a different level of "engineering rigor". **Purpose:** To develop, on the basis of linear transformation of coordinates, a correct method for calculating the error probability in the correlating reception of a four-fold phase-manipulated radio signal in the presence of coherent harmonic interference. **Methods:** Four-dimensional probability density of a vector of output voltages of the demodulator correlators in a four-fold integral was represented by a product of one-dimensional probability densities in the space of eigenvectors of the covariance matrix, in which two probability densities are Dirac delta functions. The quadruple integral is brought to double, with new integration limits defined from the plane equations bounding the integration region in this space. **Results:** Formulas were derived for accurate calculation of average probabilities of symbol and bit errors in coherent reception of a four-fold phase-manipulated radio signal in the presence of coherent harmonic interference. The derived exact formulas were used to plot the dependencies of the average probabilities of symbol and bit errors on the signal-to-noise ratio for the given interference-to-noise ratio and the given interference phase shift relative to the signal phase. It has been studied how the energy ratios of the signal and interference, as well as the interference phase shift, affect the probabilities of symbol and bit errors. It was found that the influence of a non-energy parameter equivalently leads to a change in the energy ratios. **Practical relevance:** The results of the research can be used in assessing the communication efficiency under interference. The developed technique will allow you to accurately determine the energy characteristics of a radio channel providing the required quality for the reception of transmitted messages in presence of harmonic interference.

Keywords — symbol and bit error probability, harmonic interference, quadruple phase manipulation, interference immunity.

For citation: Zvonarev V. V., Popov A. S. Potential interference immunity of coherent reception of quadruple phase-manipulated radio signal in the presence of coherent harmonic interference. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 45–54 (In Russian). doi:10.31799/1684-8853-2021-1-45-54

References

1. Bekkiev A. Y., Borisov V. I. Estimation of the anti-interference ability of radio communication channels under electronic warfare conditions. *Journal of Communications Technology and Electronics*, 2019, vol. 64, no. 9, pp. 962–972. doi:10.1134/S0033849419080035
2. Kulikov G. V., Lelyuh A. A., Grachenko E. N. Noise immunity of coherent signal receiver with quadrature amplitude modulation in the presence of relayed interference. *Journal of Communications Technology and Electronics*, 2020, vol. 65, no. 8, pp. 934–938. doi:10.31857/S0033849420070074
3. Buchinskii D. I., Voznuk V. V., Fomin A. V. Research of noise stability of the receiver with MPSK modulation under the influence of interference with different structure. *Proceedings of the Military Space Academy named after A. F. Mozhaisky*, Saint-Petersburg, 2019, iss. 671, pp. 120–127 (In Russian).
4. Parshutkin A. V., Maslakov P. A. Study of the noise immunity of modern standards of satellite communications to the impact of non-stationary interference. *SPIIRAS Proceedings*, 2017, no. 4, pp. 159–177 (In Russian). doi:https://doi.org/10.15622/sp.53.8
5. Bondarenko V. N., Garifullin V. F., Krasnov T. V., Feoktistov D. S., Bogatyrev E. V. Interference immunity of the correlation receiver MSK-BOC signal to a concentrated interference. *Journal Achievements of Modern Radioelectronics*, 2017, no. 12, pp. 71–74 (In Russian).
6. Volhonskaya E. V., Korotei E. V., Vlasova K. V., Rushko M. V. Simulation study of the noise resistence of radiosignals reception with QPSK, BPSK, 8PSK and DBPSK. *Izvestiya KGTU*, 2017, no. 46, pp. 166–174 (In Russian).
7. Chernoyarov O. V., Glushkov A. N., Litvinenko V. P., Litvinenko Yu. V., Matveev B. V. Digital demodulator of the quadrature amplitude modulation signals. *Measurement Science Review*, 2018, vol. 18, no. 6, pp. 236–242. Available at: http://www.measurement.sk/2018/msr-2018-0032.pdf (accessed 12 October 2019). doi:10.1515/msr-2018-0032
8. Glushkov A. N., Litvinenko V. P., Matveev B. V., Chernoyarov O. V., Salnikova A. V. Basic algorithm for the coherent digital processing of the radio signals. *2015 International Conference on Space Science & Communication*, Langkawi, Malaysia, IEEE, pp. 389–392. Available at: https://www.researchgate.net/publication/308835672_Basic_algorithm_for_the_coherent_digital_processing_of_

the_radio_signals (accessed 10 May 2019). doi:10.1109/IconSpace.2015.7283834
9. Ezeagwu C. O., Chukwuneke N. S., Emeka S. C., Ezeribe B. E. Comparative study of the performance of different digital modulation schemes. *International Journal of Advances in Scientific Research and Engineering*, 2018, vol. 4, no. 3, pp. 149–154. doi:dx.doi.org/10.7324/IJASRE.2018.32639
10. Anmol Kumar, Rajdeep Kaur. PSO-based NBI resistant asynchronous MCCDMA multiuser detector. *International Journal of Intelligent Systems and Applications*, 2016, no. 10, pp. 60–67. doi:10.5815/ijisa.2016.10.07
11. Kulikov G. V., Nguyen Van Dung, Nesterov A. V., Lelyuh A. A. Noise immunity of reception of signals with multiple phase shift keying in the presence of harmonic interference. *Science Intensive Technologies*, 2018, no. 11, pp. 32–38 (In Russian). doi:10.18127/j19998465-201811-06
12. 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. *International Research Journal*, 2019, no. 1 (79), pp. 53–59. Available at: https://research-journal.org/technical/analiz-metodov-rascheta-bitovoj-veroyatnosti-oshibki-pri-kogerentnom-priemesignalov-s-m-ichnoj-fazovoj-manipulyaciej/ (accessed 10 May 2019) (In Russian). doi:https://doi.org/10.23670/IRJ.2019.79.1.009
13. Abed A. Kh., Zhukov V. M. The analysis of radio station noise immunity under the influence of transmission noise. *Transactions of the TSTU*, 2016, vol. 22, no. 1, pp. 53–57 (In Russian). doi:10.17277/vestnik.2016.01
14. Savischenko N. V., Lebeda E. V. Multi-position signal coherent reception error probability in a channel with generalized gamma or *K* fading and white noise. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 1, pp. 76–88 (In Russian). doi:10.31799/1684-8853-2019-1-76-88
15. Kirillov S. N., Lisnichuk A. A. Analysis of narrow-band interference effect on cognitive radio systems based on synthesized four-position radio signals. *Proc. 14th International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering, APEIE 2018*, Novosibirsk, IEEE, 2018, pp. 50–54. Available at: https://ieeexplore.ieee.org/document/8545965 (accessed 10 May 2019). doi:10.1109/APEIE.2018.8545965
16. Sidorkina Y. A., Shakhtarin B. I., Sizykh V. V., Shevtsev V. A. Costas circuit under the action of additive harmonic interferences and wideband noise. *Journal of Communications Technology and Electronics*, 2016, vol. 61, no. 7, pp. 807–816. doi:10.30898/1684-1719.2020.4.9
17. Kulikov G. V., Nesterov A. V., Lelyuh A. A. Interference immunity of reception of signals with quadrature amplitude shift keying in the presence of harmonic interference. *Journal of Radio Electronics*, 2018, no. 11. Available at: http://jre.cplire.ru/jre/nov18/9/text.pdf. (accessed 10 May 2019) (In Russian). doi:10.30898/1684-1719.2018.11.9
18. Lozhkin K. Yu., Stitsenko A. I. The immunity of non-coherent and coherent reception of a signal DQPSK in the conditions of influence of PSK, a harmonic interference or Gaussian noise. *Journal of Siberian Federal University. Engineering & Technologies*, 2017, vol. 10, no. 2, pp. 260–270. Available at: http://elib.sfu-kras.ru/bitstream/handle/2311/32132/12_Lozhkin.pdf?sequence=1&isAllowed=y (In Russian). doi:10.17516/1999-494X-2017-10-2-260-270
19. Lozhkin K. Yu. Jamming resistance of the OFDM signal with single phase-shift keying and corrective coding in polyharmonic interference environment. *Information and Space*, 2018, no. 2, pp. 37–43 (In Russian).
20. Zvonarev V. V., Popov A. S. Methodical approach to estimation of the potential noise stability of optimum coherent reseption of the multiposition phase-shift keyd radio signal with the white noise. *Radioengineering*, 2019, vol. 83, no. 4, pp. 79–83 (In Russian). doi:10.18127/j00338486-201904-79
21. Proakis J. G. *Digital communications*. McGraw-Hill, Book Company, 1995. 800 p.
22. Savischenko N. V. *Mnogomernye signalnye konstruktsii: ih chastotnaya effektivnost i potentsialnaya pomehoustqichivost priema* [Multidimensional signal structures: their frequency efficiency and potential noise immunity of reception]. Saint-Petersburg, Politekhnicheskii universitet Publ, 2005. 420 p. (In Russian).

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

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

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

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

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

UDC 621.372
doi:10.31799/1684-8853-2021-1-55-65

Articles

Theoretical foundations of digital vector Fourier analysis of two-dimensional signals padded with zero samples

O. V. Ponomareva[a], *Dr. Sc., Tech., Professor, orcid.org/0000-0002-7311-3108, ponva@mail.ru*
A. V. Ponomarev[a], *PhD, Econ., Associate Professor, orcid.org/0000-0002-3746-9289*
[a]*Kalashnikov Izhevsk State Technical University, 7, Studencheskaya St., 426069, Izhevsk, Russian Federation*

Introduction: *The practice of using Fourier-processing of finite two-dimensional signals (including images), having confirmed its effectiveness, revealed a number of negative effects inherent in it. A well-known method of dealing with negative effects of Fourier-processing is padding signals with zeros. However, the use of this operation leads to the need to provide information control systems with additional memory and perform unproductive calculations.* ***Purpose:*** *To develop new discrete Fourier transforms for efficient and effective processing of two-dimensional signals padded with zero samples.* ***Method:*** *We have proposed a new method for splitting a rectangular discrete Fourier transform matrix into square matrices. The method is based on the application of the modulus comparability relation to order the rows (columns) of the Fourier matrix.* ***Results:*** *New discrete Fourier transforms with variable parameters were developed, being a generalization of the classical discrete Fourier transform. The article investigates the properties of Fourier transform bases with variable parameters. In respect to these transforms, the validity has been proved for the theorems of linearity, shift, correlation and Parseval's equality. In the digital spectral Fourier analysis, the concepts of a parametric shift of a two-dimensional signal, and a parametric periodicity of a two-dimensional signal have been introduced. We have estimated the reduction of the required memory size and the number of calculations when applying the proposed transforms, and compared them with the discrete Fourier transform.* ***Practical relevance:*** *The developed discrete Fourier transforms with variable parameters can significantly reduce the cost of Fourier processing of two-dimensional signals (including images) padded with zeros.*

Keywords *— discrete Fourier transform, two-dimensional signal, Fourier processing, effects of discrete Fourier transform, basis, variable parameter.*

For citation: Ponomareva O. V., Ponomarev A. V. Theoretical foundations of digital vector Fourier analysis of two-dimensional signals padded with zero samples. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 55–65. doi:10.31799/1684-8853-2021-1-55-65

Introduction

Fourier-processing of finite discrete two-dimensional (FDTD) signals (including images) in informational control (IC) systems is the most important method for studying processes and analyzing information [1–8]. The theoretical basis of Fourier-processing of FDTD signals is two-dimensional direct and inverse discrete Fourier transforms (2D DFT, 2D IDFT) [9–15] which can be represented in form of:

— algebraic form 2D DFT

$$S_{N_1,N_2}(k_1,\, k_2) = \frac{1}{N_1 N_2}\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} x(n_1,\, n_2) W_{N_1}^{k_1 n_1}\cdot W_{N_2}^{k_2 n_2},$$

where $S_{N_1,N_2}(k_1,\, k_2)$ are coefficients (bins) 2D DFT; $k_1 = \overline{0,\, (N_1-1)}$, $k_2 = \overline{0,\, (N_2-1)}$ are spatial frequencies; $x(n_1,\, n_2)$ is 2D signal; $n_1 = \overline{0,\, N_1-1}$, $n_2 = \overline{0,\, N_2-1}$; $W_{N_1}^{k_1 n_1} = \exp\left(-j\frac{2\pi}{N_1}k_1 n_1\right)$; $W_{N_2}^{k_2 n_2} = \exp\left(-j\frac{2\pi}{N_2}k_2 n_2\right)$;

— algebraic form 2D IDFT

$$x(n_1,\, n_2) = \sum_{k_1=0}^{N_1-1}\sum_{k_2=0}^{N_2-1} S_{N_1,N_2}(k_1,\, k_2) W_{N_1}^{-k_1 n_1}\cdot W_{N_2}^{-k_2 n_2}.$$

The practice of using DFT and 2D DFT, on the one hand, confirmed their efficiency, on the other hand, revealed a number of effects: aliasing effect, scalloping effect, picket fence effect, negatively affecting on the results of analysis and information processing [16–21].

To eliminate these negative effects of DFT and 2D DFT, the zero-padding operation (ZPO) the FDTD signal has found wide application. ZPO can significantly reduce the impact of negative effects on the results of Fourier-processing [22, 23]. However, effective use of ZPO requires solving the problem of Fourier-processing of FDTD of this kind of signals. The essence of the problem lies in the fact that in Fourier-processing of signals subjected to ZPO, on the one hand, it is necessary to provide the corresponding IC systems with a significant amount of additional memory, on the other hand, the IC systems must perform unproductive computations with zero samples, which significantly increases time of Fourier-processing. The paper pro-

poses and investigates new discrete Fourier transforms, which allow efficient and effective analysis and processing of two-dimensional signals padded with zero samples.

The role of the zero-padding operation of FDTD signals in two-dimensional Fourier-processing

The systems analysis of Fourier-processing theory of FDTD signals made it possible to formulate its axiomatic basic provisions:

— determination of FDTD signals on a finite two-dimensional reference plane, which is interpreted as a two-dimensional fundamental period $SA_{N_1\times N_2}$ (2D period). 2D period is set by horizontal and vertical periods;

— determination of the shift of a two-dimensional discrete signal in the form of a cyclic shift, carried out by cyclic permutation of its samples on the final reference area $SA_{N_1\times N_2}$;

— definition of a complete two-dimensional basis system

$$\text{def}_{N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2) = W_{N_1}^{k_1 n_1}\cdot W_{N_2}^{k_2 n_2},$$

where $n_1 = \overline{0, N_1 - 1};\quad n_2 = \overline{0, N_2 - 1};\quad k_1 = \overline{0, (N_1 - 1)};$ $k_2 = \overline{0, (N_2 - 1)}.$

As a result of the discreteness and periodicity of 2D signals in the spatial domain, the periodicity and discreteness of 2D Fourier spectra in the spatial-frequency domain, the mathematical operations of convolution and correlation are cyclical. However, the analysis, design, and modeling of isoplanatic systems requires the results of linear operations with 2D signals.

The method, which allows obtaining the results of linear operations using cyclic operations, consists in expanding the reference regions with zero samplex of the convoluted signals by applying ZPO to them.

If the reference area $SA_{V_1\times V_2}$ of signal $x(n_1,\, n_2)$ and the reference area $SA_{Q_1\times Q_2}$ of signal $y(n_1,\, n_2)$ are specified, then the size of the reference area, padded with zeros to obtain linear convolution $h_{linear}(n_1,\, n_2)$, should be

$$SA_{(V_1+Q_1)\times(V_2+Q_2)},$$

where $n_1 = \overline{0,\, (V_1 + Q_1 - 1)};\; n_2 = \overline{0,\, (V_2 + Q_2 - 1)}.$

And the size of the reference area for obtaining linear correlation $C_L(n_1,\, n_2)$ should be

$$SA_{2V_1\times 2V_2},$$

where $n_1 = \overline{0,\, (2V_1 - 1)};\; n_2 = \overline{0,\, (2V_2 - 1)}.$

Therefore, the algorithm for obtaining 2D linear convolution based on 2D cyclic convolution consists of the following operations:

1. Pad 2D signals $x(n_1,\, n_2)$ and $y(n_1,\, n_2)$ with Q_1, Q_2 and V_1, V_2 zero samples respectively, which sets new 2D signals $x_0(n_1,\, n_2)$, $y_0(n_1,\, n_2)$ with horizontal N_2 and vertical N_1 periods according to the ratios

$$N_1 \geq (V_1 + Q_1 - 1);\; N_2 \geq (V_2 + Q_2 - 1).$$

2. Perform 2D DFT of 2D signals $x_0(n_1,\, n_2)$ and $y_0(n_1,\, n_2)$:

$$x_0(n_1,\, n_2) \xrightarrow{F} X_{0,N_1,N_2}(k_1,\, k_2);$$

$$y_0(n_1,\, n_2) \xrightarrow{F} Y_{0,N_1,N_2}(k_1,\, k_2),$$

where $\xrightarrow{F}$ is the 2D DFT execution symbol.

3. Perform 2D IDFT product

$$X_{0,N_1,N_2}(k_1,\, k_2)\cdot Y_{0,N_1,N_2}(k_1,\, k_2).$$

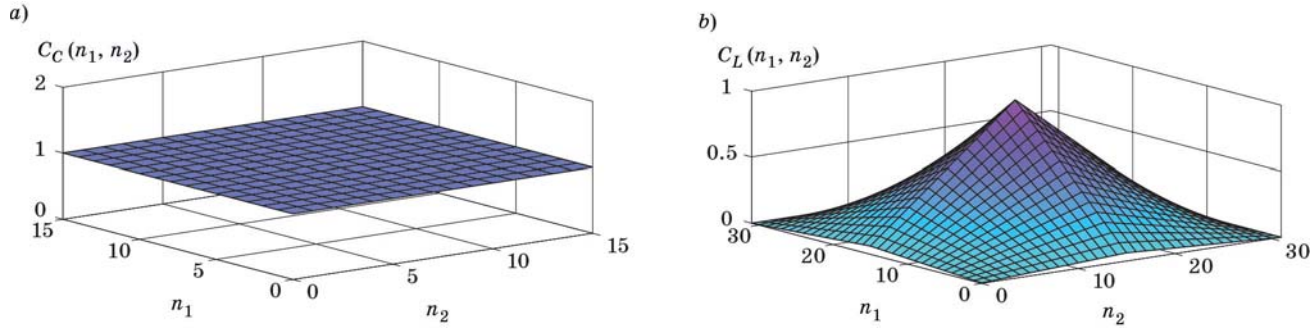
The algorithm for obtaining a linear 2D correlation function based on a cyclic 2D correlation function is easy to obtain from the previous algorithm. Fig. 1, *a* and *b* illustrates the differences between cyclic $C_C(n_1,\, n_2)$ and linear $C_L(n_1,\, n_2)$ correlation functions of a finite unit 2D signal.

According to the two-dimensional version of the Wiener — Khinchin theorem, Fourier transform of the linear 2D correlation function allows one to obtain the energy spectrum of a 2D signal. There is a so-called direct method for obtaining the energy spectrum of a 2D signal $x(n_1,\, n_2)$, bypassing the stage of obtaining the correlation function:

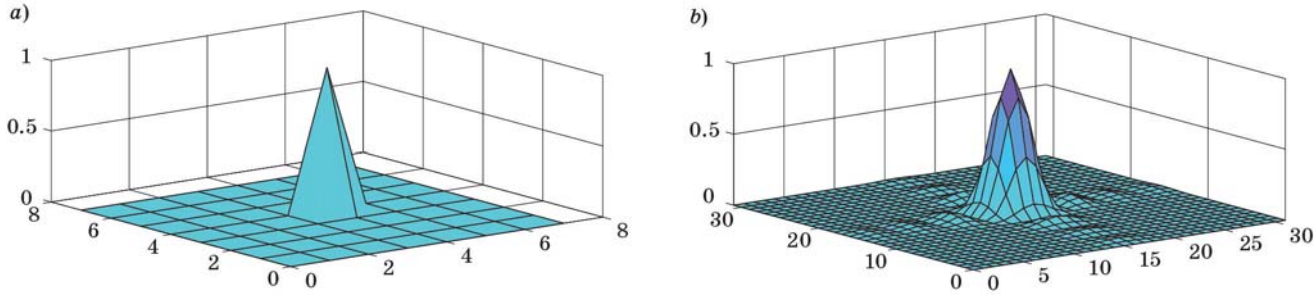
$$G_{N_1,N_2}(k_1,\, k_2) = N_1 N_2 \left| S_{N_1,N_2}(k_1,\, k_2) \right|^2.$$

A significant drawback of this definition of the energy spectrum of a 2D signal $x(n_1,\, n_2)$ is insufficient detailing $G_{N_1,N_2}(k_1,\, k_2)$, for example, to fulfill the conditions of Pugachev canonical signal decomposition. The method of increasing the detail $G_{N_1,N_2}(k_1,\, k_2)$ is carried out by padding the 2D signal $x(n_1,\, n_2)$ with zeros at least twice. Fig. 2, *a* and *b* illustrates the detailing of the energy spectrum of a finite single 2D signal.

As noted in the introduction, the effective application of the ZPO requires a solution to the problem of Fourier-processing of FDTD signals padded with zero samples. The essence of the problem lies in the fact that in Fourier-processing of signals subjected to the ZPO, on the one hand, it is necessary to provide the corresponding IC system with a significant additional amount of RAM (storage), on the other hand, IC system must perform a lot of non-productive calculations with zero samples, which significantly increases the time of Fourier-processing.



■ ***Fig. 1.*** Cyclic (*a*) and linear (*b*) correlation 2D functions of a finite single 2D signal



■ ***Fig. 2.*** Energy spectrum of a finite single 2D signal with zero frequency in the center of the spectrum: *a* — a finite single 2D signal $N_1 = 8$, $N_2 = 8$; *b* — a finite single 2D signal, padded with zero samples to $N_1 = 32$, $N_2 = 32$

Let us consider a generalization of 2D DFT in the form of a 2D DFT with a variable parameter, which makes it possible to efficiently analyze and process two-dimensional signals subjected to ZPO.

Two-dimensional DFT with variable parameter

Let two 2D signals be given: a signal $\mathbf{X}_{N_1 \times N_2}$ and a signal $\mathbf{O}_{N_1 \times N_2}$ with zero samples.

To perform the linear transformations considered in the previous section, it is necessary to pad (supplement) the horizontal period of the 2D signal $\mathbf{X}_{N_1 \times N_2}$ with $(r_2 - 1)$ zero matrices $\mathbf{O}_{N_1 \times N_2}$, which leads to a block matrix:

$$\mathbf{X}_{N_1 \times (N_2 r_2)} = \begin{matrix} 0 & 1 & . & . & (r_2 - 1) \end{matrix} \\ \left[\mathbf{X}_{N_1 \times N_2} \quad \mathbf{O}_{N_1 \times N_2} \quad . \quad . \quad \mathbf{O}_{N_1 \times N_2} \right]. \tag{1}$$

Taking into account the separability property of the 2D DFT, Fourier transform of a signal $\mathbf{X}_{N_1 \times (N_2 r_2)}$ in matrix form can be represented as

$$\mathbf{S}^{k_1, k_2}_{N_1 \times (N_2 r_2)} = \frac{1}{N_1 N_2} \mathbf{F}^{(2)}_{N_1 \times N_1} \left[\mathbf{X}_{N_1 \times (N_2 r_2)} \mathbf{F}^{(1)}_{(N_2 r_2) \times (N_2 r_2)} \right], \tag{2}$$

where

$$\mathbf{F}^{(2)}_{N_1 \times N_1} = \begin{matrix} \\ 0 \\ 1 \\ . \\ . \\ (N_1 - 1) \\ k_1 \end{matrix} \begin{matrix} \begin{matrix} 0 & 1 & . & . & (N_1 - 1) \end{matrix} \\ \begin{bmatrix} W_{N_1}^{0 \cdot 0} & W_{N_1}^{0 \cdot 1} & . & . & W_{N_1}^{0 \cdot (N_1 - 1)} \\ W_{N_1}^{1 \cdot 0} & W_{N_1}^{1 \cdot 1} & . & . & W_{N_1}^{1 \cdot (N_1 - 1)} \\ . & . & . & . & . \\ . & . & . & . & . \\ W_{N_1}^{(N_1 - 1) \cdot 0} & W_{N_1}^{(N_1 - 1) \cdot 1} & . & . & W_{N_1}^{(N_1 - 1)(N_1 - 1)} \end{bmatrix} \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ n_1 \end{matrix};$$

$$\mathbf{X}_{N_1\times(N_2r_2)} = \begin{array}{c} \\ 0 \\ 1 \\ \cdot \\ \cdot \\ (N_1-2) \\ (N_1-1) \\ n_1 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & .. & (N_2-1) & N_2 & .. & (N_2r_2-1) \end{array} \\ \begin{bmatrix} x(0,\,0) & .. & x(0,\,(N_2-1)) & 0 & .. & 0 \\ x(1,\,0) & .. & x(1,\,(N_2-1)) & 0 & .. & 0 \\ \cdot & .. & \cdot & \cdot & .. & \cdot \\ \cdot & .. & \cdot & \cdot & .. & \cdot \\ x((N_1-2),\,0) & .. & x((N_1-2),\,(N_2-1)) & 0 & .. & 0 \\ x((N_1-1),\,0) & .. & x((N_1-1),\,(N_2-1)) & 0 & .. & 0 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ n_2 \end{array};$$

$$\mathbf{F}^{(1)}_{(N_2r_2)\times(N_2r_2)} = \begin{array}{c} \\ 0 \\ 1 \\ \cdot \\ \cdot \\ (N_2r_2-1) \\ n_2 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & . \; . & (N_2r_2-1) \end{array} \\ \begin{bmatrix} W_{N_2}^{0\cdot 0} & W_{N_2}^{0\cdot 1} & . \; . & W_{N_2}^{0\cdot(N_2r_2-1)} \\ W_{N_2}^{1\cdot 0} & W_{N_2}^{1\cdot 1} & . \; . & W_{N_2}^{1\cdot(N_2r_2-1)} \\ \cdot & \cdot & . \; . & \cdot \\ \cdot & \cdot & . \; . & \cdot \\ W_{N_2}^{(N_2r_2-1)\cdot 0} & W_{N_2}^{(N_2r_2-1)\cdot 1} & . \; . & W_{N_2}^{(N_2r_2-1)(N_2r_2-1)} \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ k_2 \end{array}. \quad (3)$$

Let us interrogate the structure of matrix equation (2). It is easy to see that the multiplication of matrices $\mathbf{X}_{N_1\times(N_2r_2)}$ and $\mathbf{F}^{(1)}_{(N_2r_2)\times(N_2r_2)}$ leads to a rectangular matrix $\mathbf{Y}_{N_1\times(N_2r_2)}$. A matrix $\mathbf{Y}_{N_1\times(N_2r_2)}$ can be interpreted as the product of a matrix $\mathbf{X}_{N_1\times N_2}$ by a matrix $\mathbf{F}^{(1)}_{N_2\times(N_2r_2)}$:

$$\mathbf{Y}_{N_1\times(N_2r_2)} = \mathbf{X}_{N_1\times N_2}\cdot\mathbf{F}^{(1)}_{N_2\times(N_2r_2)},$$

where

$$\mathbf{X}_{N_1\times N_2} = \begin{array}{c} \\ 0 \\ 1 \\ \cdot \\ \cdot \\ (N_1-2) \\ (N_1-1) \\ n_1 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & \ldots & (N_2-1) \end{array} \\ \begin{bmatrix} x(0,\,0) & x(0,\,1) & \ldots & x(0,\,(N_2-1)) \\ x(1,\,0) & x(1,\,1) & \ldots & x(1,\,(N_2-1)) \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ x((N_1-2),\,0) & x((N_1-2),\,1) & \ldots & x((N_1-2),\,(N_2-1)) \\ x((N_1-1),\,0) & x((N_1-1),\,1) & \ldots & x((N_1-1),\,(N_2-1)) \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ n_2 \end{array}; \quad (4)$$

$$\mathbf{F}^{(1)}_{N_2\times(N_2r_2)} = \begin{array}{c} \\ 0 \\ 1 \\ \cdot \\ \cdot \\ (N_2-1) \\ n_2 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & . \; . & (N_2r_2-1) \end{array} \\ \begin{bmatrix} W_{N_2}^{0\cdot 0} & W_{N_2}^{0\cdot 1} & . \; . & W_{N_2}^{0\cdot(N_2r_2-1)} \\ W_{N_2}^{1\cdot 0} & W_{N_2}^{1\cdot 1} & . \; . & W_{N_2}^{1\cdot(N_2r_2-1)} \\ \cdot & \cdot & . \; . & \cdot \\ \cdot & \cdot & . \; . & \cdot \\ W_{N_2}^{(N_2-1)\cdot 0} & W_{N_2}^{(N_2-1)\cdot 1} & . \; . & W_{N_2}^{(N_2-1)(N_2r_2-1)} \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ k_2 \end{array}. \quad (5)$$

Comparing matrices $\mathbf{F}^{(1)}_{(N_2r_2)\times(N_2r_2)}$ (3) and $\mathbf{F}^{(1)}_{N_2\times(N_2r_2)}$ (5), we may see that the matrix $\mathbf{F}^{(1)}_{N_2\times(N_2r_2)}$ is the result of truncating $N_2(r_2 - 1)$ the rows of the matrix $\mathbf{F}^{(1)}_{(N_2r_2)\times(N_2r_2)}$. According to [22], "we denote the set of matrix column numbers by A...:

$$A : A = \{0,\, 1,\, 2,\, \ldots,\, (N_2r_1-1)\}.$$

We apply to the set A the relation of comparability modulo r_2.

It is known that the relation of comparability in modulus m is an equivalence relation and has the properties of reflexivity, symmetry and transitivity."

Also we know that "the relation of comparability modulo r_2 divides the set A into r_2 classes of residues modulo r_2:

$$A_0 = \{0,\ r_2,\ \ldots,\ (N_2 - 1)r_2\};$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$A_{(r_2-1)} = \{(r_2 - 1),\ \ldots,\ (N_2 r_2 - 1)\};$$
$$A_i \neq \varnothing;\ A_i \bigcap_{i \neq j} A_j = \varnothing;\ \bigcup_{i=0}^{r_1 - 1} A_i = A. \qquad (6)$$

The matrix $\mathbf{F}^{(1)}_{N_2 \times (N_2 r_2)}$, applying the partition (6) of the set A into r_2 residue classes modulo r_2, can be represented in the form of r_2 matrices of size $N_2 \times N_2$" [22]:

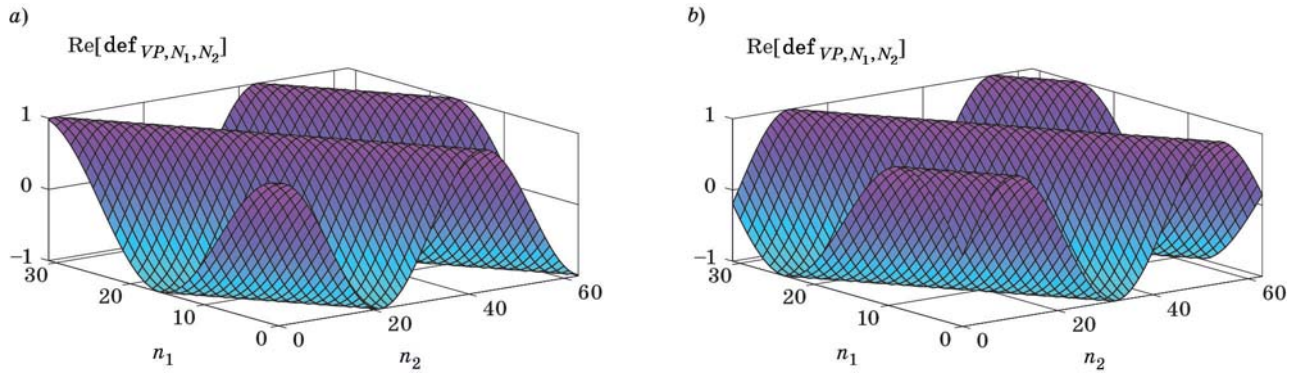
$$\mathbf{F}^{(1)}_{N_2 \times N_2, \theta_2} = \begin{array}{c} \\ 0 \\ 1 \\ . \\ . \\ (N_2 - 1) \\ n_2 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & . & . & (N_2 - 1) & \\ \end{array} \\ \begin{bmatrix} W_{N_2}^{0\cdot(0+\theta_2)} & W_{N_2}^{0\cdot(1+\theta_2)} & . & . & W_{N_2}^{0\cdot(N_2-1+\theta_2)} \\ W_{N_2}^{1\cdot(0+\theta_2)} & W_{N_2}^{1\cdot(1+\theta_2)} & . & . & W_{N_2}^{1\cdot(N_2-1+\theta_2)} \\ . & . & . & . & . \\ . & . & . & . & . \\ W_{N_2}^{(N_2-1)(0+\theta_2)} & W_{N_2}^{(N_2-1)(1+\theta_2)} & . & . & W_{N_2}^{(N_2-1)(N_2-1+\theta_2)} \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \\ \\ k_2 \end{array}, \qquad (7)$$

where $\theta_2 = 0;\ 1/r_2,\ \ldots,\ (r_2 - 1)/r_2$.

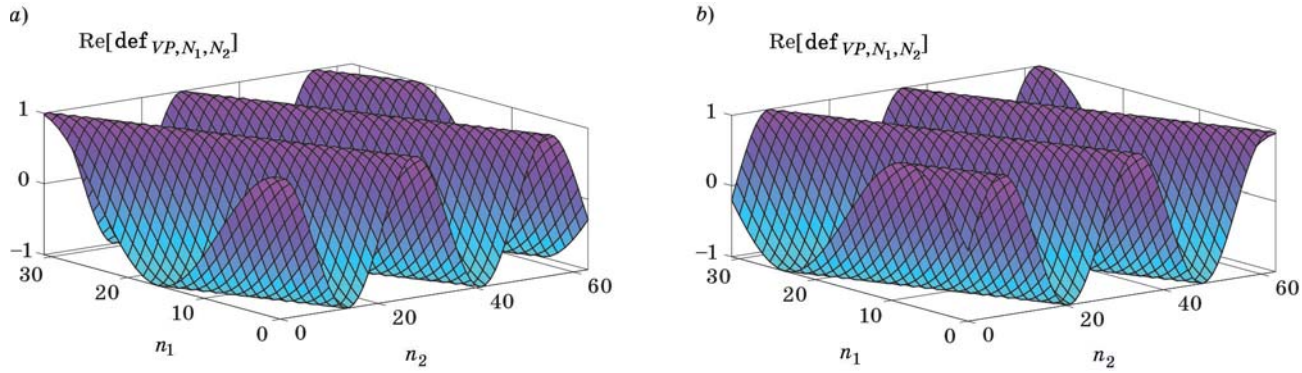
Discrete two-dimensional exponential functions of the form

$$\mathrm{def}_{HP,N_1,N_2}(k_1,\ n_1,\ k_2,\ n_2,\ \theta_2) = W_{N_1}^{k_1 n_1} \cdot W_{N_2}^{(k_2+\theta_2)n_2} = \left[\exp\left(-j\frac{2\pi}{N_1}k_1 n_1\right)\right] \cdot \left[\exp\left(-j\frac{2\pi}{N_2}(k_2+\theta_2)n_2\right)\right] =$$
$$= \left[\cos\left(\frac{2\pi}{N_1}k_1 n_1\right) - j\sin\left(\frac{2\pi}{N_1}k_1 n_1\right)\right] \cdot \left[\cos\left(\frac{2\pi}{N_2}(k_2+\theta_2)n_2\right) - j\sin\left(\frac{2\pi}{N_2}(k_2+\theta_2)\right)n_2\right] =$$
$$= \cos\left(\frac{2\pi}{N_1}k_1 n_1 + \frac{2\pi}{N_2}(k_2+\theta_2)n_2\right) - j\sin\left(\frac{2\pi}{N_1}k_1 n_1 + \frac{2\pi}{N_2}(k_2+\theta_2)n_2\right), \qquad (8)$$

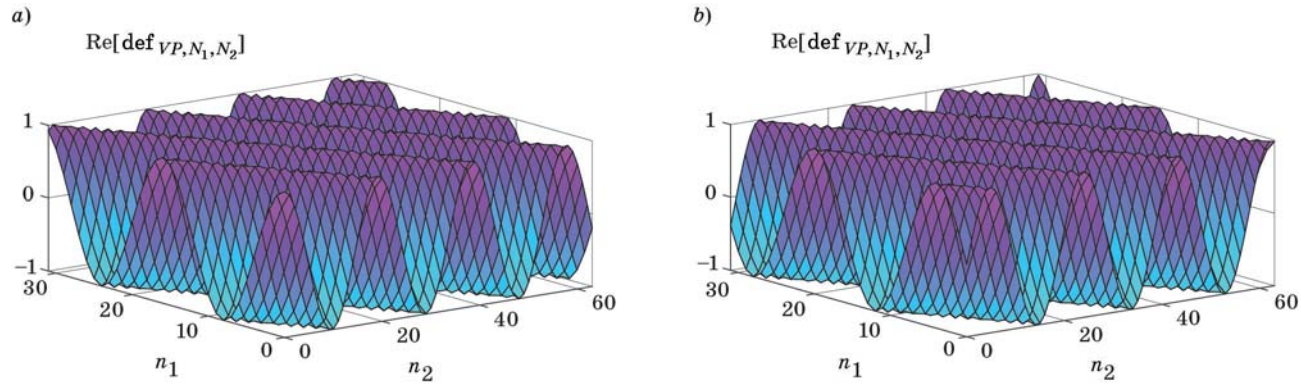
where $k_1 = \overline{0,\ N_1 - 1};\ k_2 = \overline{0,\ N_2 - 1};\ 0 \le \theta_2 < 1$, will be called two-dimensional discrete exponential functions with a variable parameter — 2D DEF-VP (Figs. 3–5).



■ ***Fig. 3.*** Two-dimensional exponential function with variable parameter at $N_1 = 32$, $N_2 = 64$; $k_1 = 1$, $k_2 = 1$; $\theta_2 = 1/2$: *a* — a real part; *b* — an imaginary part



■ ***Fig. 4.*** Two-dimensional exponential function with variable parameter at $N_1 = 32$, $N_2 = 64$; $k_1 = 1$, $k_2 = 2$; $\theta_2 = 1/3$: *a* — a real part; *b* — an imaginary part



■ ***Fig. 5.*** Two-dimensional exponential function with variable parameter at $N_1 = 32$, $N_2 = 64$; $k_1 = 2$, $k_2 = 3$; $\theta_2 = 1/3$: *a* — a real part; *b* — an imaginary part

The introduction of discrete exponential functions with a variable parameter makes it possible to generalize the concept of periodicity of the DEF-VP system. Recall that the periodicity of the DEF system in the classical DFT is understood as a periodic continuation of the DEF system outside the interval of N samples. Moreover, the system of discrete basis functions in the classical DFT does not contain discontinuities. In the case of discrete Fourier transform with a variable parameter (DFT-VP) (9), for the DEF-VP system to be inseparable, the periodicity should be understood as parametric periodicity. The parametric periodicity of discrete exponential functions with a variable parameter is understood as their periodic continuation with rotation in complex space by an angle of $2\pi\theta$. Note that the introduced concept of parametric periodicity is valid for 1D and 2D real and complex functions.

Consider the main properties of two-dimensional discrete exponential functions of 2D DEF-VP.

Main properties of 2D DEF-VP

Each of the two-dimensional discrete exponential functions with a variable parameter has its own spatial frequencies k_1, k_2, which determine its place in a particular basic system. The set of 2D DEF-VP makes its basic system of two-dimensional discrete Fourier transform with a variable parameter (2D DFT-VP) in space $\mathbf{I}_2^N$.

For each value of the parameter θ_2 we can say that:

1. 2D DEF-VP are complex functions by definition.

2. The basis system 2D DEF-VP is a generalization of the basis system 2D DEF and is equal to it at $\theta_2 = 0$.

3. 2D DEF-VP are two-dimensional functions of four equivalent variables k_1, k_2, n_1, n_2, and one variable parameter θ_2:

$$\mathrm{def}_{HP,N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2,\, \theta_2) = W_{N_1}^{k_1 n_1} \cdot W_{N_2}^{(k_2+\theta_2)n_2}.$$

4. 2D DEF-VP are periodic in variables k_1 and n_1 with a period N_1 and a variable with a period N_2:

$$\mathrm{def}_{HP,N_1,N_2}((k_1 \pm lN_1),\, (n_1 + qN_1),\, (k_2 \pm mN_2),$$
$$n_2,\, \theta_2) = \mathrm{def}_{HP,N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2,\, \theta_2),$$

where l, m, q are integers.

5. 2D DEF-VP are parametrically periodic in a variable n_2 with a period N_2:

$$\mathrm{def}_{HP,N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2 \pm pN_2,\, \theta_2) = \mathrm{def}_{VP,N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2,\, \theta_2)\cdot W_{N_1}^{\theta_2 N_2 p},$$

where p is integer.

A parametric shift of a two-dimensional signal $X_{N_1\times N_2}$ in the horizontal direction is understood as a two-dimensional cyclic parametric shift of the form of

$$\mathbf{C}_{H.Sh} = \begin{matrix} 0 \\ 1 \\ 2 \\ . \\ . \\ (N_2-1) \end{matrix}\begin{bmatrix} [\mathbf{X}_{N_1\times N_2}]^{\mathrm{T}}\cdot\mathbf{H}^0_{H.Sh,N_2\times N_2} \\ [\mathbf{X}_{N_1\times N_2}]^{\mathrm{T}}\cdot\mathbf{H}^1_{H.Sh,N_2\times N_2} \\ [\mathbf{X}_{N_1\times N_2}]^{\mathrm{T}}\cdot\mathbf{H}^2_{H.Sh,N_2\times N_2} \\ . \\ . \\ [\mathbf{X}_{N_1\times N_2}]^{\mathrm{T}}\cdot\mathbf{H}^{(N_2-1)}_{H.Sh,N_2\times N_2} \end{bmatrix},$$

where $\mathbf{H}^0_{H.Sh,N_2\times N_2}$ is two-dimensional identity matrix, expression $\mathbf{H}^m_{H.Sh,N_2\times N_2}$, $m=\overline{0,\,(N_2-1)}$ means raising to the power m of the matrix of two-dimensional parametric shift:

$$\mathbf{H}_{H.Sh,N_2\times N_2,\theta_2} = \begin{matrix} \\ 0 \\ 1 \\ 2 \\ . \\ . \\ . \\ (N_2-2) \\ (N_2-1) \\ n_2 \end{matrix}\begin{matrix} \begin{matrix} 0 & 1 & 2 & . & . & . & (N_2-2) & (N_2-1) & n_2 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 1 & . & . & . & & 0 & \\ 0 & 0 & 0 & . & . & . & & 0 & \\ 0 & 0 & 0 & . & . & . & & 0 & \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & 0 & 1 & \\ \exp(-2\pi\theta_2) & 0 & 0 & . & . & . & 0 & 0 & \end{bmatrix} \end{matrix}.$$

6. The basis system 2D DEF-VP in the variables k_1, k_2 is not multiplicative:

$$\mathrm{def}_{HP,N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2,\, \theta_2)\cdot\mathrm{def}_{HP,N_1,N_2}(k_3,\, n_1,\, k_4,\, n_2,\, \theta_2) \neq$$
$$\neq \mathrm{def}_{HP,N_1,N_2}((k_1+k_3)_{\mathrm{mod}\,N_1},\, n_1,\, (k_2+k_4)_{\mathrm{mod}\,N_2},\, n_2,\, \theta_2).$$

7. The basis system 2D DEF-VP in the variables n_1, n_2 is multiplicative:

$$\mathrm{def}_{HP,N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2,\, \theta_2)\cdot\mathrm{def}_{H,N_1,N_2}(k_1,\, n_3,\, k_2,\, n_4,\, \theta_2) =$$
$$= \mathrm{def}_{HP,N_1,N_2}(k_1,\, (n_1+n_3)_{\mathrm{mod}\,N_1},\, k_2,\, (n_2+n_4)_{\mathrm{mod}\,N_2},\, \theta_2).$$

8. Average value of 2D DEF-IP with spatial frequencies $k_1\neq 0$, $k_2\neq 0$ is equal to zero:

$$\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1}\mathrm{def}_{HP,N_1,N_2}(k_1,\, n_1,\, k_2,\, n_2,\, \theta_2) = \sum_{n_1=0}^{N_1-1} W_{N_1\times N_2}^{N_2 k_1 n_1}\left[\sum_{n_2=0}^{N_2-1} W_{N_1\times N_2}^{N_2(k_2+\theta_1)n_2}\right] =$$

$$= \left[\frac{1-W_{N_1}^{k_1 N_1}}{1-W_{N_1}^{k_1}}\right]\left[\frac{1-W_{N_2}^{(k_2+\theta_2)N_2}}{1-W_{N_2}^{k_2}}\right] = 0.$$

9. The basic system 2D DEF-VP is an orthogonal basis system with respect to variables k_1, k_2:

$$\frac{1}{N_1N_2}\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} W_{N_1\times N_2}^{(N_2k_1n_1+N_1(k_2+\theta_2)n_2)} \times$$

$$\times W_{N_1\times N_2}^{(N_2k_3n_1+N_1(k_4+\theta_2)n_2)*} = \begin{cases} 1, \text{ if } k_1=k_3,\ k_2=k_4 \\ 0, \text{ if } k_1\neq k_3,\ k_2\neq k_4 \end{cases};$$

and with respect to variables n_1, n_2:

$$\frac{1}{N_1N_2}\sum_{k_1=0}^{N_1-1}\sum_{k_2=0}^{N_2-1} W_{N_1\times N_2}^{(N_2k_1n_1+N_1(k_2+\theta_2)n_2)} \times$$

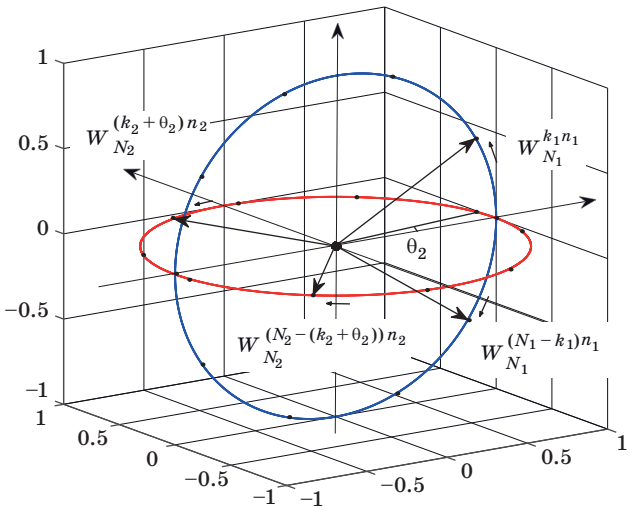
$$\times W_{N_1\times N_2}^{(N_2k_1n_3+N_1(k_2+\theta_2)n_4)*} = \begin{cases} 1, \text{ if } n_1=n_3,\ n_2=n_4 \\ 0, \text{ if } n_1\neq n_3,\ n_2\neq n_4 \end{cases},$$

where the symbol * means complex conjugation.

10. 2D DEF-VP can be represented by two unit vectors, which represent $W_{N_1}^{k_1n_1}$ and $W_{N_2}^{(k_2+\theta_2)n_2}$. The unit vectors rotate discontinuously (discretely) in perpendicular complex planes. On the interval N_1, the unit vector which displays $W_{N_1}^{k_1n_1}$, passes the angle of $2\pi k_1$ radians, making k_1 revolutions, and on the interval N_2, the unit vector, representing $W_{N_2}^{(k_2+\theta_2)n_2}$, passes the angle $2\pi(k_2+\theta_2)$ radians, making $(k_2+\theta_2)$ revolutions. The unit vectors representing the complex conjugate DEF-VP:

$$W_{N_1}^{-k_1n_1} = W_{N_1}^{(N_1-k_1)n_1} \text{ and}$$

$$W_{N_2}^{-(k_2+\theta_2)n_2} = W_{N_2}^{(N_2-(k_2+\theta_2))n_2}$$



■ ***Fig. 6.*** Representation of a two-dimensional discrete exponential function with a variable parameter in the spatial domain

and make (N_1-k_1) and $(N_2-(k_2+\theta_2))$ revolutions respectively.

Figure 6 illustrates such a 2D DEF-VP representation, where the angles $2\pi k_1/N_1$ and $2\pi(k_2+\theta_2)/N_2$ are marked with the corresponding points.

11. The basis system 2D DEF-VP is complete in space $\mathbf{I}_2^N$.

Expansion in basis systems of the form (8) is defined as a 2D DFT-VP. Algebraic form of 2D DFT-VP

$$S_{N_1,N_2}(k_1,\ k_2,\ \theta_2) =$$

$$= \frac{1}{N_1N_2}\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} x(n_1,\ n_2)W_{N_1}^{k_1n_1}\cdot W_{N_2}^{(k_2+\theta_2)n_2}, \quad (9)$$

where k_1, k_2 are spatial frequencies, $k_1=\overline{0,\ (N_1-1)}$, $k_2=\overline{0,\ (N_2-1)}$; θ_2 is a parameter, $0\le\theta_2<1$; $x(n_1,\ n_2)$ — two-dimensional signal, $n_1=\overline{0,\ N_1-1}$, $n_2=\overline{0,\ N_2-1}$; $S_{N_1,N_2}(k_1,\ k_2,\ \theta_2)$ are bins of 2D DFT-VP (two-dimensional vector spatial-frequency spectrum of the signal $x(n_1,\ n_2)$ in the basic 2D DEF-VP system).

The algebraic form of direct 2D DFT-VP, taking into account the property of separability of the kernel (core) of 2D DFT-VP, can be represented as

$$S_{N_1,N_2}(k_1,\ k_2,\ \theta_2) =$$

$$= \frac{1}{N_1}\sum_{n_1=0}^{N_1-1} W_{N_1}^{k_1n_1}\left[\frac{1}{N_2}\sum_{n_2=0}^{N_2-1} x(n_1,\ n_2)W_{N_2}^{(k_2+\theta_2)n_2}\right]. \quad (10)$$

It can be seen that formula (10) makes it possible to step-by-step calculation of the direct 2D DFT-VP by the method of sequential calculation of two DFT-P (parametric DFT). Note that the calculation of the DFT-P can be carried out by methods of parametric fast Fourier transform (FFT-P) [1].

There is an inverse 2D DFT-VP (2D IDFT-VP):

$$x(n_1,\ n_2) =$$

$$= \sum_{k_1=0}^{N_1-1}\sum_{k_2=0}^{N_2-1} S_{N_1,N_2}(k_1,\ k_2,\ \theta_2)W_{N_1}^{-k_1n_1}\cdot W_{N_2}^{-(k_2+\theta_2)n_2},$$

where $n_1=\overline{0,\ N_1-1}$; $n_2=\overline{0,\ N_2-1}$.

Using the separability property of the 2D DFT-VP kernel, we can introduce the matrix form of the direct 2D DFT-VP:

$$\mathbf{S}_{N_1\times N_2,\theta_2} = \frac{1}{N_1}\mathbf{F}_{N_1\times N_1}^{(2)}\cdot\frac{1}{N_2}\left[\mathbf{X}_{N_1\times N_2}\cdot\mathbf{F}_{N_2\times N_2,\theta_2}^{(1)}\right],$$

where $0\le\theta_2<1$;

$$\mathbf{F}^{(2)}_{N_1\times N_1} = \begin{array}{c} \\ 0 \\ 1 \\ . \\ . \\ (N_1-1) \\ k_1 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & . & . & (N_1-1) & \\ \end{array} \\ \begin{bmatrix} W_{N_1}^{0\cdot 0} & W_{N_1}^{0\cdot 1} & . & . & W_{N_1}^{0\cdot(N_1-1)} \\ W_{N_1}^{1\cdot 0} & W_{N_1}^{0\cdot 1} & . & . & W_{N_1}^{1\cdot(N_1-1)} \\ . & . & . & . & . \\ . & . & . & . & . \\ W_{N_1}^{(N_1-1)\cdot 0} & W_{N_1}^{(N_1-1)\cdot 1} & . & . & W_{N_1}^{(N_1-1)(N_1-1)} \end{bmatrix} \\ n_1 \end{array};$$

$\mathbf{X}_{N_1\times N_2}$ is given by relation (4);

$$\mathbf{F}^{(1)}_{N_2\times N_2,\theta_2} = \begin{array}{c} \\ 0 \\ 1 \\ . \\ . \\ (N_2-1) \\ n_2 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & . & . & (N_2-1) & \\ \end{array} \\ \begin{bmatrix} W_{N_2}^{0\cdot(0+\theta_2)} & W_{N_2}^{0\cdot(1+\theta_2)} & . & . & W_{N_2}^{0\cdot(N_2-1+\theta_2)} \\ W_{N_2}^{1\cdot(0+\theta_2)} & W_{N_2}^{1\cdot(1+\theta_2)} & . & . & W_{N_2}^{1\cdot(N_2-1+\theta_2)} \\ . & . & . & . & . \\ . & . & . & . & . \\ W_{N_2}^{(N_2-1)(0+\theta_2)} & W_{N_2}^{(N_2-1)(1+\theta_2)} & . & . & W_{N_2}^{(N_2-1)(N_2-1+\theta_2)} \end{bmatrix} \\ k_2 \end{array}. \quad (11)$$

We note the difference between matrices (11) and (7), which lies in the nature of the parameter θ_2 change.

The inverse 2D DFT-VP in matrix form is determined by the matrix equation

$$\mathbf{X}_{N_1\times N_2} = \frac{1}{N_1}\mathbf{F}^{(2)*}_{N_1\times N_1}\cdot\frac{1}{N_2}[\mathbf{S}_{N_1\times N_2,\theta_2}\cdot\mathbf{F}^{(1)*}_{N_2\times N_2,\theta_2}],$$

where $0 \le \theta_2 < 1$.

It can be shown that the theorems of linearity, shift, correlation and Parseval's equality are valid for 2D DFT-VP. For 2D DFT-VP, similar to 2D DFT, the concepts of power spectrum $P_{N_1,N_2}(k_1, k_2, \theta_2)$ and energy spectrum $G_{N_1,N_2}(k_1, k_2, \theta_2)$ can be introduced

$$P_{N_1,N_2}(k_1, k_2, \theta_2) = \left|S_{N_1,N_2}(k_1, k_2, \theta_2)\right|^2;\ G_{N_1,N_2}(k_1, k_2, \theta_2) = P_{N_1,N_2}(k_1, k_2, \theta_2)/\Delta f;\ \Delta f = 1/(N_1N_2).$$

Let us estimate the efficiency of increasing the detailing of the two-dimensional energy spectrum $G_{N_1,N_2}(k_1, k_2)$ using 2D DFT-VP in comparison with the classical 2D DFT.

Evaluation of the efficiency of Fourier-processing of signals padded with zero samples in 2D DFT-VP basis

The increase in the detailing of the two-dimensional energy spectrum $G_{N_1,N_2}(k_1, k_2)$ by r_2 times is carried out by padding the horizontal period of the 2D signal $X_{N_1\times N_2}$ with $(r_2 - 1)$ zero matrices $\mathbf{O}_{N_1\times N_2}$ (1). Padding the horizontal period of a 2D signal $X_{N_1\times N_2}$ with $(r_2 - 1)$ zero matrices $\mathbf{O}_{N_1\times N_2}$ makes it possible to obtain a new 2D signal $X_{N_1\times(N_2r_2)}$ from a 2D signal $X_{N_1\times N_2}$.

Applying the 2D DFT in algebraic form to the 2D signal $X_{N_1\times(N_2r_2)}$, we obtain the number of coefficients (bins) of 2D DFT $S_{N_1,N_2r_2}(k_1, k_2)$, which is r_2 times greater than with 2D DFT of the signal $X_{N_1\times N_2}$. However, obtaining a r_2 times more detailed energy spectrum $G_{N_1,N_2r_2}(k_1, k_2)$ by a method based on the separability of the 2D DFT kernel, will require additional $(r_2 - 1)N_1N_2$ cells for storing zero samples and implementing $N_1N_2r_2(N_1 + N_2r_2)$ additional complex operations.

Obtaining an r_2 times detailed energy spectrum $G_{N_1,N_2r_2}(k_1, k_2)$ by a method based on the separability property of the 2D DFT-VP kernel does not require additional RAM (storage) for storing zero samples and requires $N_1N_2r_2(N_1 + N_2)$ complex operations. Thus, the use of 2D DFT-VP instead of the classic 2D DFT allows:

— decrease number of complex operations by $\gamma = \dfrac{N_1 + N_2r_2}{N_1 + N_2}$ times;

— decrease storage size by r_2 times;

— parallelize the process of detailing the two-dimensional energy spectrum $G_{N_1,N_2}(k_1, k_2)$, thus reducing the execution time of the 2D DFT by r_2 times.

Conclusions

Discrete Fourier transforms with a variable parameter have been developed. These transforms make it possible to efficiently process two-dimensional signals, the horizontal periods of which are padded with zero samples. The generalization of classical two-dimensional discrete Fourier transform is based on a new method of splitting the rectangular matrix of discrete Fourier transform into square matrices. The splitting of rectangular matrices into square matrices is carried out by using the ordering of the columns of rectangular matrices using the equivalence relation — the relation of comparability in modulus. The properties of the bases of the proposed transformations are investigated. The validity for Fourier transforms with variable parameters of the following theorems is proved: linearity, shift, correlation, and Parseval's equality.

New concepts of digital spectral Fourier analysis are introduced: the concept of parametric shift of two-dimensional signal and parametric periodicity of two-dimensional signal. The estimation of the reducing the amount of RAM (random access memory) needed and the number of calculations when applying the proposed transforms is carried out in comparison with the application of classical two-dimensional discrete Fourier transform to 2D signals padded with zero samples. Developed two-dimensional discrete Fourier transform with variable parameters can significantly reduce the cost of Fourier-processing of two-dimensional signals (including images), padded with zero samples. In addition, the developed transforms also allow parallelizing the process, thus significantly reducing Fourier-processing time. Note one more application of developed two-dimensional discrete Fourier transform with variable parameters: determination of the parameters of 2D hidden periodicities by varying the parameter θ_2.

References

1. Ponomarev A. V. *Systems Analysis of Discrete Two-Dimensional Signal Processing in Fourier Basis*. In: *Advances in Signal Processing. Theories, Algorithms, and System Control-7*. Favorskaya M. N., Jain L. C. (eds). Springer, Cham. Vol. 184. Pp. 87–96. doi.org/10/1007/978-3-030-40312-6_7
2. Ponomareva O. V., Ponomarev A. V., Smirnova N. V. *Sliding Spatial Frequency Processing of Discrete Signals*. In: *Advances in Signal Processing. Theories, Algorithms, and System Control-8*. Favorskaya M. N., Jain L. C. (eds). Springer, Cham. Vol. 184. Pp. 97–110. doi.org/10/1007/978-3-030-40312-6_8
3. Favorskaya M., Savchina E., Popov A. Adaptive visible image watermarking based on Hadamard transform. *IOP Conference Series: Materials Science and Engineering*, 2018, vol. 450, no. 5, MIST Aerospace, pp. 052003.1-052003.6. doi:10.1088/1757-899X/450/5/052003
4. Klionskiy D. M., Kaplun D. I., Geppener V. V. Empirical more decomposition for signal preprocessing and classification of intrinsic mode functions. *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)*, 2018, vol. 28, no. 1, pp. 122–132. doi:10.1134/S1054661818010091
5. Batishchev V. I., Volkov I. I., Zolin A. G. Using a stochastic basis in signal and image recovery problems. *Optoelectronics, Instrumentation and Data Processing*, 2017, vol. 53, no. 4, pp. 414–420.
6. Bakulin M. G., Vityazev V. V., Shumov A. P., Kreyndelin V. B. Effective signal detection for the spatial multiplexing mimo systems. *Telecommunications and Radio Engineering*, 2018. vol. 77, no. 13, pp. 1141–1158. doi.org/10.1615/TelecomRadEng.v77.i13.30
7. Lerner I. M., Il'In G. I., Il'In V. I. To the matter of optimal transfer characteristics of linear selective systems of communication channel with memory and apsk-n. *2019 Systems of Signal Synchronization, Generating and Processing in Telecommunications, Synchroinfo*, 2019, pp. 8814277. doi.org/10.1109/SYNCHROINFO.2019.8814021
8. Kulikovskikh I., Prokhorov S. Psychological perspectives on implicit regularization: a model of retrieval-induced forgetting (RIF). *Journal of Physics: Conference Series*, 2018, pp. 012079. doi:10.1088/1742-6596/1096/1/012079
9. Lysenko N., Labkov G. Applying of Kutter-Jor-dan-Bossen steganographic algorithm in video sequences. *2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2017, pp. 695–696. doi:10.1109/EICon-Rus.2017.7910651
10. Favorskaya M. N., Buryachenko V. V. *Authentication and Copyright Protection of Videos under Transmitting Specifications*. In: *Computer Vision in Advanced Control Systems-5*. Favorskaya M. N., Jain L. C. (eds). Springer, Cham, 2020. Vol. 175. Pp. 119–160. doi.org/10.1007/978-3-030-33795-7_5
11. Khanyan G. S. Sampling theorem in frequency domain for the finite spectrum. *Proceedings of 2018 IEEE East-West Design and Test Symposium, EWDTS 2018*, 2018, pp. 8524822. doi:10.1109/EWDTS.2018.8524822
12. Petrovsky N. A., Rybenkov E. V., Petrovsky A. A. Two-dimensional non-separable quaternionic paraunitary filter banks. *22nd IEEE Signal Processing: Algorithms, Architectures, Arrangements, and Applica-*

tions, *SPA 2018*, 2018, pp. 120–125. doi:10.23919/SPA.2018.8563311
13. Likhttsinder B. Conditional average value of queues in queuing systems with bath request flows. *2017 4th International Scientific-Practical Conference Problems of Infocommunications Science and Technology, PIC S and T 2017 — Proceedings*, 2018, pp. 49–52. doi:10.1109/INFOCOMMST.2017.8246347
14. Khanyan G. S. Frequency domain sampling theorem for an infinite spectrum. *DSPA: Voprosy primeneniya cifrovoj obrabotki signalov*, 2018, vol. 8, no. 2, pp. 56–61 (In Russian).
15. Bakulin M. G., Vityazev V. V., Shumov A. P., Kreyndelin V. B. Effective signal detection for the spatial multiplexing mimo systems. *Telecommunications and Radio Engineering*, 2018. vol. 77, no. 13, pp. 1141–1158. doi.org/10.1615/TelecomRadEng.v77.i13.30
16. Smirnova N. V., Ponomareva O. V. Vector and spectral digital signal processing in musical acoustics using the parametric discrete Fourier transform. *Digital Signal Processing*, 2019, no. 2, pp. 3–11 (In Russian).
17. Klionskiy D. M., Geppener V. V. Application of iterative procedures for signal processing. *Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2018*, 2018, pp. 1090–1093. doi:10.1109/EIConRus.2018.8317280
18. Bakulin M. G., Vityazev V. V., Shumov A. P., Kreyndelin V. B. Effective signal detection for the spatial multiplexing mimo systems. *Telecommunications and Radio Engineering*, 2018, vol. 77, no. 13, pp. 1141–1158.
19. Prozorov D., Tatarinova A. Comparison of grapheme-to-phoneme conversions for spoken document. *2019 IEEE East-West Design and Test Symposium, EWDTS 2019*, 2019, pp. 8884449. doi:10.1109/EWDTS.2019.8884449
20. Prozorov D., Trubin I. Detection of a signal in the simo system with spatial correlation of noise. *2018 Proceedings of 7th Mediterranean Conference on Embedded Computing, MECO 2018 — Including ECYPS 2018*, 2018, pp. 1–5. doi:10.1109/MECO.2018.8405965
21. Urakov A., Gurevich K., Alies M., Reshetnikov A., Kasatkin A., Urakova N. The tissue temperature during injection of drug solution into it as an integral indicator of rheology. *Journal of Physics: Conference Series. 4th International Conference on Rheology and Modeling of Materials, IC-RMM 2019*, 2020, vol. 1527, pp. 012003. doi:10.1088/1742-6596/1527/1/012003
22. Ponomarev A. V. Fundamentals of the theory of two-dimensional digital signal processing in Fourier bases with variable parameters. *Digital Signal Processing*, 2019, no. 2, pp. 12–20 (In Russian).
23. Ponomareva N. V. Problems of computer spectral signal processing in musical acoustics. *Intelligent Systems in Manufacturing*, 2018, vol. 16, no. 1, pp. 26–33 (In Russian). doi:10.22213/24-10-9304-2018-1-26-33

УДК 621.372
doi:10.31799/1684-8853-2021-1-55-65

Теоретические основы цифрового векторного фурье-анализа двумерных сигналов, дополненных нулевыми отсчетами

О. В. Пономарева[a], доктор техн. наук, профессор, orcid. org/0000-0002-7311-3108, ponva@mail.ru
А. В. Пономарев[a], канд. экон. наук, доцент, orcid.org/0000-0002-3746-9289
[a]Ижевский государственный технический университет им. М. Т. Калашникова, Студенческая ул., 7, Ижевск, 426069, РФ

Введение: практика применения фурье-обработки финитных двумерных сигналов (в том числе изображений), подтвердив ее действенность, выявила и ряд присущих ей отрицательных эффектов. Известный метод борьбы с негативными эффектами фурье-обработки — операция дополнения сигналов нулями. Однако применение этой операции приводит к необходимости обеспечения информационно-управляющих систем дополнительной памятью и проведения ими непроизводительных вычислений. **Цель:** разработка новых дискретных преобразований Фурье для эффективной и результативной обработки двумерных сигналов, дополненных нулями. **Методы:** предложен новый метод разбивки прямоугольной матрицы дискретного преобразования Фурье на квадратные матрицы. Метод основан на применении для упорядочения строк (столбцов) матрицы Фурье отношения сравнимости по модулю. **Результаты:** разработаны новые дискретные преобразования Фурье с варьируемыми параметрами, являющиеся обобщением классического дискретного преобразования Фурье. Исследованы свойства базисов преобразований Фурье с варьируемыми параметрами. Для данных преобразований доказаны теоремы линейности, сдвига, корреляции и равенство Парсеваля. В цифровой спектральный фурье-анализ введены понятия параметрического сдвига двумерного сигнала и параметрической периодичности двумерного сигнала. Выполнена оценка сокращения требуемого объема памяти и числа вычислений в случае применения предложенных преобразований. Проведено их сравнение с дискретным преобразованием Фурье. **Практическая значимость:** разработанные дискретные преобразования Фурье с варьируемыми параметрами позволяют существенно сократить затраты на фурье-обработку двумерных сигналов (в том числе изображений), дополненных нулями.

Ключевые слова — дискретное преобразование Фурье, двумерный сигнал, фурье-обработка, эффекты дискретного преобразования Фурье, базис, варьируемый параметр.

Для цитирования: Ponomareva O. V., Ponomarev A. V. Theoretical foundations of digital vector Fourier analysis of two-dimensional signals padded with zero samples. *Информационно-управляющие системы*, 2021, № 1, с. 55–65. doi:10.31799/1684-8853-2021-1-55-65
For citation: Ponomareva O. V., Ponomarev A. V. Theoretical foundations of digital vector Fourier analysis of two-dimensional signals padded with zero samples. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 55–65. doi:10.31799/1684-8853-2021-1-55-65

СВЕДЕНИЯ ОБ АВТОРАХ

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

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 группы Ли, квантовая запутанность, комбинаторика.
Эл. адрес: djokovic@uwaterloo.ca

ЗВОНАРЕВ Виталий Валерьевич

Начальник лаборатории, старший научный сотрудник Военного института (научно-исследовательского) Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.
В 2007 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Сети связи и системы коммутации».
В 2017 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 50 научных публикаций.
Область научных интересов — радиоканалы связи в условиях замираний.
Эл. адрес:
zvonarevvitalii@yandex.ru

ГАЙФУЛИНА Диана Альбертовна

Младший научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского федерального исследовательского центра РАН.
В 2017 году окончила бакалавриат Оренбургского государственного университета, в 2019 году — магистратуру Университета ИТМО по специальности «Информационная безопасность».
Является автором более 27 научных публикаций.
Область научных интересов — информационная безопасность, киберфизические системы, интеллектуальный анализ данных, обнаружение аномалий в среде передачи данных.
Эл. адрес:
gaifulina@comsec.spb.ru

ДРОБИНЦЕВ Павел Дмитриевич

Доцент, директор Высшей школы программной инженерии Института компьютерных наук и технологий Санкт-Петербургского политехнического университета Петра Великого.
В 2003 году окончил Санкт-Петербургский государственный политехнический университет по специальности «Автоматизированные системы управления».
В 2006 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 92 научных публикаций.
Область научных интересов — программная инженерия, автоматизация, верификация, тестирование.
Эл. адрес:
drob@ics2.ecd.spbstu.ru

КОВАЛЕВ Артем Дмитриевич

Аспирант, ассистент Высшей школы программной инженерии Института компьютерных наук и технологий Санкт-Петербургского политехнического университета Петра Великого.
В 2018 году окончил Санкт-Петербургский политехнический университет Петра Великого по специальности «Программная инженерия».
Является автором восьми научных публикаций.
Область научных интересов — автоматизация поддержки программного обеспечения, машинное обучение, семантический поиск.
Эл. адрес: kov3000@ya.ru

КОТЕНКО Игорь Витальевич

Профессор, главный научный сотрудник, заведующий лабораторией проблем компьютерной безопасности Санкт-Петербургского федерального исследовательского центра РАН.
В 1983 году окончил Военно-космическую академию им. А. Ф. Можайского по специальности «Математическое обеспечение автоматизированных систем управления», в 1987 году — Военную академию связи по специальности «Инженерная автоматизированных систем управления».
В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 500 научных публикаций.
Область научных интересов — безопасность компьютерных сетей, обнаружение компьютерных атак, межсетевые экраны и др.
Эл. адрес: ivkote@comsec.spb.ru

НИКИФОРОВ Игорь Валерьевич

Доцент Высшей школы программной инженерии Института компьютерных наук и технологий Санкт-Петербургского политехнического университета Петра Великого.
В 2011 году окончил Санкт-Петербургский государственный политехнический университет по специальности «Программное обеспечение вычислительной техники и автоматизированных систем».
В 2014 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 50 научных публикаций.
Область научных интересов — большие данные, машинное обучение, распределенные вычисления, верификация.
Эл. адрес:
i.nikiforov@ics2.ecd.spbstu.ru

ПОНОМАРЕВ Алексей Владимирович

Доцент кафедры языковой подготовки в профессиональной сфере Ижевского государственного технического университета им. М. Т. Калашникова.
В 2001 году окончил Ижевский механический институт по специальности «Прикладная математика», в 2003 — по специальности «Финансы и кредит».
В 2004 году защитил диссертацию на соискание ученой степени кандидата экономических наук.
Является автором более 60 научных публикаций.
Область научных интересов — системный анализ, управление и обработка информации, теория цифровой обработки двумерных сигналов, математические и инструментальные методы экономики и управления народным хозяйством.
Эл. адрес: palexizh@gmail.com

КРУГЛИК Станислав Александрович

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

ОЛЕНЕВ Валентин Леонидович

Директор Института высокопроизводительных компьютерных и сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2007 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Информатика и вычислительная техника».
В 2012 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 100 научных публикаций и трех патентов на изобретения.
Область научных интересов — встроенные системы, бортовые космические и авиационные сети и др.
Эл. адрес:
Valentin.Olenev@guap.ru

ПОНОМАРЕВА Ольга Владимировна

Профессор кафедры приборов и методов измерений, контроля, диагностики Ижевского государственного технического университета им. М. Т. Калашникова.
В 1976 году окончила Ижевский механический институт по специальности «Электронные вычислительные машины».
В 2016 году защитила диссертацию на соискание ученой степени доктора технических наук.
Является автором более 150 научных публикаций.
Область научных интересов — теория цифровой обработки сигналов и изображений; методы и средства измерений, контроля и диагностики и др.
Эл. адрес: ponva@mail.ru

ПОПОВ Александр Сергеевич

Профессор, старший научный сотрудник Военного института (научно-исследовательского) Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.
В 1971 году окончил Ленинградскую военную инженерную краснознаменную академию им. А. Ф. Можайского по специальности «Радиотехнические системы».
В 1989 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 150 научных публикаций.
Область научных интересов — пространственно-временная обработка сигналов, статистический синтез алгоритмов приема при наличии помех.
Эл. адрес: arahar@mail.ru

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