ISSN 2519-2310

COMPUTER SCIENCE AND CYBERSECURITY

ISSUE 4(4) 2016

V. N. Karazin Kharkiv National University Publishing

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені В.Н.КАРАЗІНА
ХАРЬКОВСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ имени В.Н. КАРАЗИНА
V.N. KARAZIN KHARKIV NATIONAL UNIVERSITY

КОМП'ЮТЕРНІ НАУКИ ТА КІБЕРБЕЗПЕКА
КОМПЬЮТЕРНЫЕ НАУКИ И КИБЕРБЕЗОПАСНОСТЬ
COMPUTER SCIENCE AND CYBERSECURITY
(CS&CS)

Issue 4(4) 2016

Заснований 2015 року

Міжнародний електронний науково-теоретичний журнал
Международный электронный научно-теоретический журнал
International electronic scientific journal

The journal publishes research articles on theoretical, scientific and technical problems of effective facilities development for computer information communication systems and on information security problems based on advanced mathematical methods, information technologies and technical means.

Journal is published quarterly.
Approved for placement on the Internet by Academic Council of the Karazin Kharkiv National University (December 26, 2016, protocol No. 17)

Editorial office:
Karazin Kharkiv National University
Svobody sq., 6, office 315a, Kharkiv, 61022, Ukraine
Phone: +38 (057) 705-10-83
E-mail: *cscsjournal@karazin.ua*
Web-page: *http://periodicals.karazin.ua/cscs (Open Journal System)*

Published articles have been internally and externally peer reviewed

TABLE OF CONTENTS

Issue 4(4) 2016

UDC 681.391

MATHEMATICAL AND PHYSICAL NATURE OF CHANNEL CAPACITY

S. Rassomakhin

V. N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine
rassomakhin@karazin.ua

Reviewer: Victor Krasnobayev, Dr., Full Professor, V.N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine;
krasnobayev@karazin.ua

Received on November 2016

Abstract. *The classic methodological approaches to the determination of channel capacity have been considered. The contradiction between analytical and geometric definitions of maximum achievable transmission rate has been shown. Objectivity of maximum likelihood rule usage in low-quality channels with low signal/noise ratio has been analyzed. The correct formulation of the mathematical and physical content of channel capacity has been made. Invariance of capacity to a noise distribution in continuous channels has been proved. The main causes of the crisis in the development of information transmission theories have been indicated.*

Keywords: *differential entropy, channel capacity, maximum likelihood rule, uncertainty sphere, random encoding.*

Introduction and problem formulation

Currently, the definitions of fundamental limits of speed, reliability of data transmission, channel capacity, value of signal/noise ratio, as the key indicator of predicted communication quality, have become the most extensively used categories in communication theory and its applications. The works of Kotelnikov [1] and Shannon [2], published in 1946–1948, are considered to be the discovery of the fundamental laws of compression, data transmission and marks the birth of information theory in its modern sense. The theory based on the deep intersection with probability theory, statistics, computer science and other fields of knowledge was the basis for the development of communications, data storage and processing, and other information technologies.

This theory can be defined as a science dealing with the study and optimization of information encoding/decoding algorithms in order to create economical and reliable ways of its transmission through communication channels and its memory storage. The theory has arisen from the needs of radio, radar, telephone, television and computer technology, and is the theoretical base for the construction of communication systems. This theory focuses on the problem of optimal (in terms of speed, reliability and efficiency) usage of available technical devices for transmission, transformation, distribution and storage of information. At present, by the depth and amount of the researches, information theory can be matched with many branches of mathematical physics.

Undoubtedly, the main category of modern information theory is the concept of noisy channel capacity defined by Shannon [2,6]. According to his interpretation, capacity is a boundary of the data transmission rate, which cannot be exceeded with any encoding/decoding methods under any high level of transmission reliability, but it can be approached arbitrarily close to by choosing the proper methods of encoding and decoding. Channel capacity was expressed in statistical terms by introducing mathematical characteristic of the joint probability distribution of two random variables, called the amount of information. It is equal to the maximum amount of information in the signal at the channel output relative to the signal at its input, where the maximum is taken over all probability distributions of the input signal. The amount of information, in its turn, is expressed through another value, which has long been used in thermodynamics – the entropy, and represents the difference between the entropy of the channel output signal and the conditional entropy, if the input signal is known. Methodological role of capacity is extremely high in information theory, because it is not only the basis for the coding theorem stated by Shannon, but also is instrumental in proving the majority of other fundamental theorems and the existing limits.

Despite the undeniable achievements in information theory, it has been criticized recently. The reason for this is not only a lack of practicality and constructiveness in various statements of theorems but, moreover, the theory development crisis is manifesting. Visible technological progress in communication services cannot hide the absence of significant increase in specific efficiency of telecommunication equipment. The channel and physical layer protocols of information transmission system (ITS) are rather expensive. Error correcting codes, which have history of theoretical and experimental studies that amounts to more than 70 years, almost are not used in practice. The reason is not only the computational complexity of constructing and decoding cumbersome constructions in high-speed channels, but also the unacceptability of substantial residual amount of erroneous decoding probability for a transmission of data and program texts. It can be said without exaggeration that the specific efficiency of telecommunications has not changed since the twenties of the last century. The development of technique and communication technology is purely extensive. Performance improvement is achieved mostly by the development of transceiver technological base, as well as the bandwidth expansion and transmitter power (which, actually, determines the mathematical definition of capacity). It has negative moral, material and ecological effects. The problem of electromagnetic compatibility is becoming all the more essential. Overloaded traditional radio frequency ranges and a small bandwidth of metallic communication lines have forced switch to the optical range (however its potential is not infinite). As a result, ITS has become less reliable and more expensive. Mobile technology is not undergoing radical changes, but only extensive modifications. Geostationary orbit of communication satellites is approaching the saturation limit. The increase in demand rate for communication services is exceeding the rate of ITS performance increment. This serves as a testament of the explicit crisis in the theory and practice of data transmission system construction.

The purpose of this work is to reveal three main causes of the crisis – the errors embodied in the "base" of information theory which are the root cause of its evolution dead end, in particular:

- proving obvious methodological errors in the existing definitions of continuous channels capacity;
- justifying incorrectness of the statement that capacity is the limit of attainable rates for any continuous channel models;
- «debunking" the view that the decision-making based on maximum likelihood rule is the best way to estimate noisy channel output state at low signal/noise ratio.

Undoubtedly, the issues being considered can be seen as debatable, especially out of context of newly obtained scientific results which are not the subject of this work and are waiting for being published. This paper should be considered as the motivations for searching fundamentally new solutions in the mathematical theory of communication, which correspond to the true physical content of information transmission process.

1 The differential entropy of continuous distributions and analytical determination of Gaussian channel capacity

Considering the work's subject, at first let's pay attention to some well-known facts. The first definition of the capacity of discrete binary channel without memory with symmetric transition graph determined by the error probability p_0, is given in [2], and uses statistical measure of uncertainty of discrete choice, called entropy:

$$C = V \cdot \max_{P(X)} \left\{ H(X) - H(X|Y) \right\}, \quad (1)$$

where X, Y – the messages at the input and output of a noisy channel; $P(X) = \{p(0), p(1)\}$ – probability distribution of binary alphabet symbol; V – the number of binary symbols transmitted through the channel per second;

$$H(X) = -\left\{ p(0) \log\left[p(0) \right] + p(1) \log\left[p(1) \right] \right\} \quad (2)$$

– the entropy (uncertainty) of a binary message source (if information is measured in bits – logarithm base equals two);

$$H(X|Y) = -\{p_0 \log[p_0] + (1-p_0)\log[1-p_0]\} \quad (3)$$

– the channel unreliability – the entropy (uncertainty) of noise. If the channel quality specified by the parameter p_0, s known, maximum (1) is achieved with equiprobable source symbols $p(0) = p(1) = 1/2$ and amounts to:

$$C = V\left[1 - H(X|Y)\right]. \quad (4)$$

The definition $I_b = C/V$ is often used for calculating the average amount of information, which a single binary symbol on the output of a discrete noisy channel contains. It is particularly used in assessing the index of specific effectiveness of ITS [3].

The equation (4) has been generalized for the case of non-binary channel without memory (see, for example, [4,10]). By now, in addition to the above cases for discrete channel models, analytical definitions of channel capacity with erasing and some "exotic" examples of asymmetric transition graphs discussed by C. Shannon in his original paper [2] are known.

By itself, any discrete channel model is a kind of an add-on to the model of continuous (in time and level) channel. The equations (1)-(4) are objectively understandable, are clear from the physical and mathematical point of view and will not be discussed further. They need to be considered in order to keep track of the continuity of the methodological approach used by Shannon for an analytic derivation of the continuous channel capacity equation. The class of continuous channels with defined capacity is narrowed to "Gaussian" [2, 5, 7, 8] in the current paradigm. Its incorrectness will be shown below.

For a continuous source, when the messages are selected from the infinite set, Shannon, following the logic of (1)-(4), introduces the concept of the entropy of a continuous distribution (often referred to as the differential entropy):

$$H(X) = -\int_{-\infty}^{\infty} f(x)\log\left[f(x)\right]dx, \quad (5)$$

where $f(x)$ – the probability density function (PDF) of continuous random variable x. Accordingly, the joint and conditional entropy of two statistically related random arguments which determine the input and output of a continuous channel are given by:

$$H(X,Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\left[f(x,y)\right]dx\,dy; \quad (6)$$

$$H(Y|X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\left[\frac{f(x,y)}{f(y)}\right]dx\,dy. \quad (7)$$

The main properties of the entropy of the continuous case (5) include the following:

1) for a given constraint on the average power σ^2 of the continuous process centered relatively to zero, the entropy (5) is maximal if this process is Gaussian, i.e.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (8)$$

in this case

$$\max_{f(x)}\left(H(x)\right) = \log\sqrt{2\pi e\sigma^2}; \quad (9)$$

2) unlike Shannon's discrete definition (5) – (9) [see. 2] the differential entropy measurement is relative to the given coordinate system, i.e., it is not absolute. This means that when the argument of the logarithm after calculating the integrals is less than unity, the differential entropy (5) – (8) can take on negative values! Such computing subjectivism has no a sensible physical interpretation till

now, and therefore, in most cases, simply is suppressed. Although Shannon tried to justify this fact asserting that, the possibility of negative differential entropy notwithstanding, the sum or the difference between two definitions of entropy is always positive [2]. However, such justification does not prevent the collapse, which will be shown below in the analytical determination of capacity by average mutual information (ratio of differential entropy).

In a continuous channel, the input source signals $x(t)$ are continuous functions of time, and the output signals – $y(t)=x(t)+\xi(t)$ are their implementations distorted by summing them with noise. The noise implementations $\xi(t)$ are also a continuous function of time. Continuous channel capacity is defined in [2] as the maximum (over all possible input distributions) of the function which essentially similar to the expression (1):

$$C=\frac{1}{T}\left(2FT\cdot\max_{f(x)}\left\{H(Y)-H(Y|X)\right\}\right), \tag{10}$$

where F – the frequency band which restricts the channel; T – duration of channel output observation; 2FT – number of degrees of freedom, defined on the duration T, as the number of independent measurements of function with a limited spectrum, defined by the sampling theorem [1,2]. In the formula (10) H(Y) – denotes the channel output entropy, and conditional entropy H(Y| X) defined by the expression (7). The difference, the maximum of which is sought in (10), is usually referred to as the average mutual information between the input and output per one channel usage:

$$I(X,Y)=H(Y)-H(Y|X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\log\frac{f(x,y)}{f(x)f(y)}. \tag{11}$$

Then for one channel usage:

$$C=\max_{f(x)}\left\{I(X,Y)\right\}. \tag{12}$$

It is convenient to consider the relationship of Shannon's information definitions for a continuous channel using the Venn diagram, shown in Fig. 1.

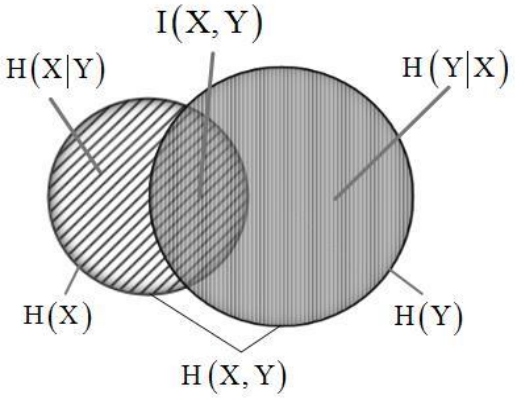


Fig. 1 – Relationship definitions of entropy for continuous channel

Therefore, the capacity of a continuous channel where noise is additive and not statistically associated with the signal, per one dimension equals the maximum of average mutual information for all variants of input distributions. [2,4,7,8] state that

$$C=2F\cdot\max_{f(x)}\left\{H(Y)-H(\xi)\right\}, \tag{13}$$

where $H(\xi)$ – the noise entropy. The theorem 16 in [2] postulates that when noise and signal are independent and additively interactive, the data transmission rate per one channel usage equals the difference between the channel output entropy and noise entropy:

$$R = H(Y) - H(\xi), \tag{14}$$

accordingly

$$C = \max_{f(x)} \{R\}. \tag{15}$$

The formulas (10), (12), (13) and (15) represent, in fact, various ways of defining the same physical magnitude for different types of measurements (total or per one channel usage). Let's continue the reasoning for the Gaussian channel in accordance with the logic of the presentation in [2], which is traditionally used in textbooks and monographs on information theory:

$$H(\xi) = \log \sqrt{2\pi eN}, \tag{16}$$

where N – the noise power. To maximize the rate, based on the properties (9), it is necessary to require that the source distribution is to be also Gaussian with the power S:

$$H(X) = \log \sqrt{2\pi eS}. \tag{17}$$

Since signal and noise are not linked statistically, due to the stability of normal distribution to the composition of any number of summable random variables [9], the distribution of their sum will be also normal with a total power which equals $(S+N)$

$$H(Y) = \log \sqrt{2\pi e(S+N)}. \tag{18}$$

As a result, we arrive at the well-known formula

$$C = F\left[\log\left(2\pi e(S+N)\right) - \log\left(2\pi eN\right)\right] \tag{19}$$

or

$$C = F \log\left(\frac{S+N}{N}\right). \tag{20}$$

It should be noted that the distribution of channel output is to be normal in the only case, when both *signal and noise are Gaussian.* The formula (20) being derived, only the signal and noise probability density functions has been used, and the methods of information receiving have not been mentioned, thereby this formula is referred to as *"Capacity of Gaussian channel"* [2–8].

Now let's focus on a strange behavior of the component analytical determination (19). To do this, we should recall that the minuend is the channel output entropy $H(Y)$, and the subtrahend is the noise entropy $H(\xi)$. What happens to the value C in case of the noise power decrease? It follows from (20) that if $F > 0$ then $\lim_{N \to 0} C = \infty$. At the same time the formula (19) shows that the capacity increases indefinitely not due to the growth channel output entropy (which, on the contrary, decreases), but due to the fact that the noise entropy (the subtrahend in (19)) tends to minus infinity:

$$\lim_{N \to 0} H(\xi) = -\infty. \tag{21}$$

This observation contradicts the physical meaning which is inherent in the definition of the difference (14). This change in the sign and adding of the subtrahend to the output entropy occurs already at "weak" noise: $N \le (2\pi e)^{-1}$. It is difficult to understand the physical meaning of this phenomenon. Although in the form (20) the capacity formula shows the monotonicity of the function $C(N)$ at $N \to 0$, that allows to explain this phenomenon by the difference between determining differential and discrete entropy, noted earlier. However, due to the lack of a clear physical interpretation of this phenomenon, correctness of the analytical derivation of capacity by using the concepts of the differential entropy and the average mutual information is doubtful.

As we will see later, attributing to this formula the ability to determine the upper limit of data transmission rates for the Gaussian channel is even more doubtful.

2 Geometric definition of capacity

After the publication of his work [2], a year later Shannon published a paper [6], which provides another method for determining capacity based on multi-dimensional geometric construction of the signal and noise space, represented in the "flat" image in Fig. 2. Any implementation of a continuous random signal, which has duration T, and which frequency spectrum is limited to F, is represented as a point in n = 2FT - dimensional space. If the transmission system is “good”, those points – S_i are uniformly distributed within the hyper sphere with the radius determined by the average signal power and the dimension of the space

$$r_S \approx \sqrt{nS} \tag{22}$$

and volume

$$V_S \approx \frac{\sqrt{\pi}^n}{\Gamma\left(n/2+1\right)}\left(\sqrt{nS}\right)^n, \tag{23}$$

where $\Gamma(n/2+1)$ – gamma function. For uniform distribution of signal points, an arbitrary choice of n coordinates – random variables with zero mean and variance, which equals S can be used. Providing the dimensions of space n increase unlimitedly, the distribution of points will monotonously approach the uniform. This asymptotic property of uniformity is the basis for the construction of random codes, almost any of which is "good" [7]. The random signal realization is a channel form of a codeword of a random code and can be obtained by two following ways:

$$S(t) = \sum_{i=0}^{2FT-1} s_i \frac{\sin(2\pi F\left(t - i\cdot\Delta t\right))}{2\pi F\left(t - i\cdot\Delta t\right)}, \quad \Delta t = 1/\left(2F\right); \tag{24}$$

$$S(t) = \sum_{i=1}^{2FT}\left\{s_{2(i-1)}\sin\left(2\pi F\frac{i}{T}\right) + s_{2(i-1)+1}\cos\left(2\pi F\frac{i}{T}\right)\right\}. \tag{25}$$

The formula (24) is an expansion of a random signal in the basis of the sinc-functions and has a continuous spectrum effectively bounded by the frequency F.

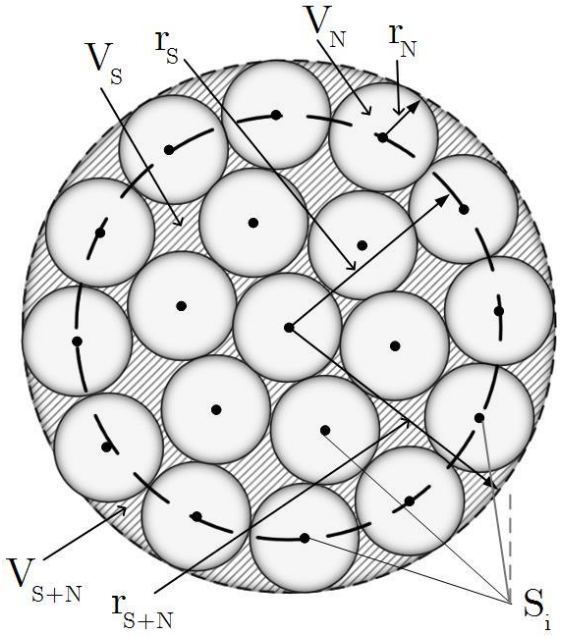


Fig. 2 – The geometric representation of the information transmission system space

For (25) the Fourier expansion in the orthogonal (on the interval T), harmonic basis is used. Thereby the realization S(t) is periodic on T, and if it is repeated indefinitely, it will have a discrete spectrum bounded above by F, non-zero measurements of which are arranged with a frequency step of $1/T$.

Both methods (24) and (25) may be used in the description of capacity attainment by means of coding given by Shannon in [2] (quoting 1): «…Let $m = 2^k$ samples of white noise be constructed, each of duration T. These are assigned binary numbers from 0 to m-1. At the transmitter, the message sequences are broken up into groups of k and for each group the corresponding noise sample is transmitted as the signal. At the receiver the m samples are known and the actual received signal (perturbed by noise) is compared with each of them. The sample which has the least R.M.S. discrepancy from the received signal is chosen as the transmitted signal and the corresponding binary number reconstructed. This process amounts to choosing the most probable (a posteriori) signal…». The formulas (24) and (25) in conjunction with the above quote is a description of the process of construction and decoding of a random code, where the decoding is performed according to the rule, which is traditionally called the *Rule of Maximum Likelihood* (MLR). With an unlimited increase in the length of the code block (synchronous increase parameters k and $n = 2TF$), if a noise is not too large, the probability of an error in the received codeword can be arbitrarily small. Thus, the geometric definition of capacity is the highest attainable rate of an arbitrary code which is decoded with the MLR and an arbitrarily low unreliability is provided.

In the geometric interpretation of the best code (Fig. 2), the point son channel output, $S_i, i \in [0, m-1]$, which correspond to the transmitted code words, are displaced under the influence of Gaussian noise within the spheres of uncertainty with the radius

$$r_N \approx \sqrt{nN} \tag{26}$$

and the volume

$$V_N \approx \frac{\sqrt{\pi}^n}{\Gamma(n/2+1)}\left(\sqrt{nN}\right)^n. \tag{27}$$

In accordance with the law of large numbers, when n increases, the probability of finding the displaced points outside the sphere with the radius $r_N + \varepsilon/\sqrt{n}$ tends to zero (ε – an arbitrary small value). Spheres of uncertainty become more delineated. Shannon compares them with regular billiard balls [2,6,8]. Since the signals of codewords and noise do not depend on each other, the total radius of hyper spherical space, which contains m spheres of uncertainty, is characterized by the radius and volume:

$$r_{S+N} \approx \sqrt{n(S+N)}, \tag{28}$$

$$V_{S+N} \approx \frac{\sqrt{\pi}^n}{\Gamma(n/2+1)}\left(\sqrt{n(S+N)}\right)^n. \tag{29}$$

With $n \to \infty$ and $\varepsilon/\sqrt{n} \to 0$ we can determine the maximum amount of non-overlapping spheres, which can be packed in the volume V_{S+N}, in such a way that there is practically no empty space between them:

$$m_C = V_{S+N}/V_N = \sqrt{\frac{S+N}{N}}^n = \sqrt{\frac{S+N}{N}}^{2FT}. \tag{30}$$

Let's recall that, if codewords are constructed in accordance with the rules (24) or (25), 2FT = n – the dimension of geometrical code space. Finding the logarithm (30) and averaging over the time T gives the maximum achievable code rate, or (*according to the modern information theory*) – channel capacity:

$$C = \frac{1}{T}\log m_C = F\log\left(\frac{S+N}{N}\right). \quad (31)$$

The results of (20) and (31) are the same apparently; allegedly it confirms the definition of C as the *maximum achievable information rate* in a channel with an additive noise and arbitrarily small unreliability. However, it should be noted, that according to the logic of the formulae derivation (31), (and of the quote discussed above as well), the value C is the *limit rate of the best code, when the Maximum Likelihood Rule is used in decoding*. If this was not true, and the receiver would not need to store samples of the signal realization segments in its memory in order to use them in the MLR comparisons (as it is described in the above quote from [2]) and when $N > S$, it would be sufficient to switch to the noise receiving (would there be any difference which of these two processes could be reliably distinguished from their mix?), in order to compensate the noise in the output mixture of the channel. We can refer to [8] or other works, which consider the physical and mathematical meaning of capacity, and see, that the value C, in the theorems proved by the author, is strictly an upper limit of rates for the codes in the Gaussian channel when the MLR is used, but not for the Gaussian channel itself, transmission and signal processing method notwithstanding. In the prevailing views on information theory there is no difference between these two concepts, because in the scheme of ITS, introduced by Shannon [2], the channel's encoder and decoder are present by default. The possibility to build an effective ITS, which does not use coding, is not considered at all! This contradicts the practical observation, noted in the introduction, that the error-correction codes are hardly used in the systems where mistakes are not allowed. To answer the question: *what the value C denotes: just the upper limit rate of information transmission over a channel with additive noise or the upper limit data transmission rate over the channel when the MLR is used for encoding and decoding* (unreliability is arbitrarily low in any case), let's turn to the analysis of mathematical and logical correctness of the reasoning used in the derivation of the formulas (20) and (31). For an objective analysis we need to change the conditions for which the formulas (20) and (31) have been obtained, i.e. consider the channel models different from the Gaussian one.

3 Comparison of the analytical and geometric definitions of capacity for non-Gaussian channel

Let's consider the following model of a continuous channel with the bandwidth limited value F, (where F - the frequency band which restricts the channel)and additive, stationary and signal-independent noise. Let the signal be Gaussian process with the probability density function:

$$f_1(x) = \frac{1}{\sqrt{2\pi S}}\exp\left(-\frac{x^2}{2S}\right), \quad (32)$$

with the mathematical expectation and variance

$$M[x] = 0, \quad D[x] = S. \quad (33)$$

The entropy of the signal is determined by the expression (17). The noise in the channel adds a random error to any signal measurement. This error has a uniform probability density in the range of $\left[-\frac{a}{2}, \frac{a}{2}\right]$, $a > 0$:

$$f_2(y) = \begin{cases} 1/a, & \text{при } y \in [-a/2, a/2]; \\ 0, & \text{при } |y| > a/2. \end{cases} \quad (34)$$

The corresponding numeric characteristics of distribution (34) are:

$$M(y) = 0, \quad D[y] = N = a^2/12. \quad (35)$$

The entropy of the noise is defined by the value:

$$H(N) = \log a. \tag{36}$$

In some cases, the exposure of the quantizer of level signal when it is measured with the values of the sampling interval $\Delta t = 1/2F$ and the limited (greater than zero) value a (the quantization step) can be described with such a noise model [3].

By the theorem 18 in [2] Shannon defines the limits of the capacity value for arbitrary non-Gaussian channel in the following form:

$$F\log\frac{S+N_1}{N_1} \le C \le F\log\frac{S+N}{N_1}, \tag{37}$$

Where N_1 – an entropy power, i.e., the power of equivalent Gaussian noise which has the same entropy as the original non-Gaussian noise do. For this model, we can calculate the entropy power by equating the values (16) and (36):

$$N_1 = a^2/(2\pi e) = \frac{12}{2\pi e}N. \tag{38}$$

Now let's calculate the capacity of the channel described, using an analytical approach (11) – (14). The channel output entropy, in this case, is the differential entropy of the process, which obtained by adding two independent processes:

– normal (signal) - with a mathematical expectation and variance (33);

– uniform (noise) - with a mathematical expectation and variance (35).

To calculate the entropy of the output channel H(Y) it is necessary to define the probability density function of the overall process f(y). The function, in this case, will be a composition of two distributions [9]:

$$f(z) = \int_{-\infty}^{\infty} f_1(w) f_2(z-w)\,dw. \tag{39}$$

Using (32) and (34) in (39) makes it possible to write:

$$f(z) = \int_{-a/2}^{a/2} \frac{1}{a}\left[(2\pi S)^{-1/2}\exp\left(\frac{-(z-w)^2}{2\cdot S}\right)\right]dw = \frac{1}{2a}\left[\operatorname{erf}\left(\frac{a+2z}{\sqrt{8\cdot S}}\right)+\operatorname{erf}\left(\frac{a-2z}{\sqrt{8\cdot S}}\right)\right], \tag{40}$$

where $\operatorname{erf}(A) = \frac{2}{\sqrt{\pi}}\int_0^A e^{-t}dt$.

Due to the independence of those two processes, numerical characteristics of the composition (40) are:

$$M[z] = M[x] + M[y] = 0; \quad D[z] = D[x] + D[y] = S + N. \tag{41}$$

The distribution (40) is not Gaussian, although is very similar to it. To make a comparison, Fig. 3 shows the probability density function (PDF) (40) and the similar PDF of the equipotent centered Gaussian process with the normalized dispersions $a = 2\sqrt{3}; \quad S = N = 1$.

Naturally, values of the differential entropy computed for PDF composition of two normal processes (formula (18)) and for the PDF composition of the case considered, are very similar as well. For example, for the values of numerical characteristics shown in Fig. 3 in (18) we have:

$$H(Y) = \log\sqrt{2\pi e\cdot 2} = 2{,}547.$$

Calculation of the entropy of the distribution (40) yields:

$$H'(Y) = -\int_{-\infty}^{\infty} f(z)\log f(z)dz = 2,544,$$

i.e., the entropy of the channel output with a uniform noise almost coincides with similar entropy of the Gaussian channel but it remains a bit smaller

$$H(Y) \approx H'(Y). \tag{42}$$

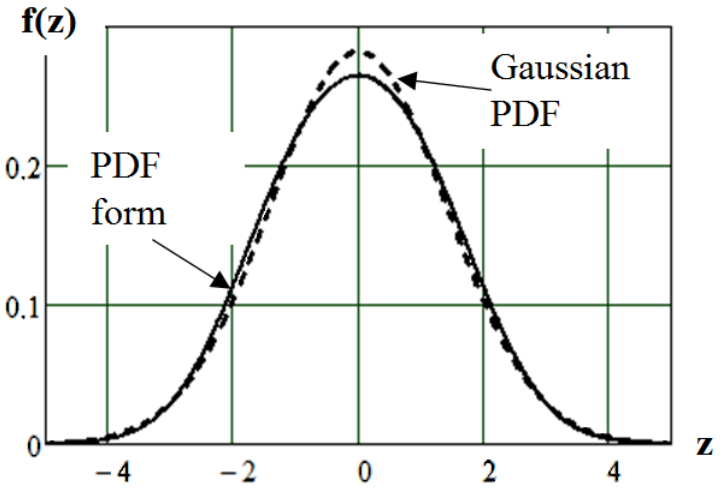


Fig. 3 – Comparison of Gaussian and composite PDFs

This result is a natural consequence of the central limit theorem of the probability theory [9]. We can write the expression for the analytical calculation of the capacity per one usage of this channel which has a uniform PDF of noise in the following form:

$$C' = H'(Y) - \log\sqrt{12 \cdot N}, \tag{43}$$

where value N determined by the formula (35).

The comparison of value (43) with the capacity per one usage of Gaussian channel, derived from (19), under the same energy conditions, gives

$$\frac{C'}{C} = \frac{H'(Y) - \log\sqrt{12 \cdot N}}{\log\sqrt{2\pi e(S+N)} - \log\sqrt{2\pi eN}}. \tag{44}$$

Example 1 For the case of equipotent signal and noise S=N=1, considered for the PDF in Fig. 3, we have:

– the entropy power determined in (38) $N_1 \approx 0,703$;

– the boundaries (37) defined by Shannon $0,638 \le C' \le 0,755$;

– the actual value calculated from (43) $C' \approx 0,751$;

– the Gaussian channel capacity, defined by the expression (19) under equivalent energy conditions

$$C = 0,5;$$

– the ratio of capacities, which defined by (44)

$$C'/C \approx 1,502.$$

The conclusion: the results of an analytical entropic definition of channel capacity with uniformly distributed noise lead to the following statement:

The channel capacity with uniformly distributed noise ***half as much again*** as the Gaussian channel capacity calculated for equipotent signal and noise! (45)

Now let's use the geometric method, discussed in Sec. 2, to determine channel capacity with uniformly distributed noise. To that end, we compare the geometric representation and the characteristics of uncertainty spheres of (Fig. 2), within which the signal points are shifted by the action of normal and uniform noise. Let's introduce the concept of normalized (to the dimension of the signal space n) displacement of a signal point under the influence of noise:

– for Gaussian noise
$$r_n = \left(\frac{1}{n} \sum_{i=1}^{n} \xi n_i^{\,2} \right)^{1/2}; \tag{46}$$

– for uniformly distributed noise
$$r_u = \left(\frac{1}{n} \sum_{i=1}^{n} \xi u_i^{\,2} \right)^{1/2}; \tag{47}$$

where ξn_i, ξu_i, $i \in [1, n]$ – random value i -th coordinates of additive noise for a normal and uniform noise respectively. The probability distribution densities of these quantities are determined by the formulas

$$f(\xi n) = \frac{1}{\sqrt{2\pi N}} \exp\left(-\frac{\xi n^2}{2N} \right); \tag{48}$$

$$f(\xi u) = \begin{cases} 1/\sqrt{12 \cdot N}, & \text{при } \xi u \in \left[-\sqrt{3 \cdot N}, \sqrt{3 \cdot N} \right]; \\ 0, & \text{при } |\xi n| > \sqrt{3 \cdot N}. \end{cases} \tag{49}$$

Normalized radii of uncertainty spheres $\overline{r_n}$ and $\overline{r_u}$ for two noise distributions, under consideration, are determined by the mathematical expectation of random variables (46) and (47), which are the functions of random summands having the PDF (48) and (49), and "delineation degree" of the spheres determined by their dispersion $D[r_n]$ и $D[r_u]$. The analytical result for normal noise is known [9]:

$$\overline{r_n} = M[r_n] = \sqrt{\frac{2N}{n}} \left[\Gamma\left(\frac{n+1}{2} \right) \Big/ \Gamma\left(\frac{n}{2} \right) \right]; \tag{50}$$

$$D[r_n] = N \left\{ 1 - \frac{2}{n} \left[\Gamma\left(\frac{n+1}{2} \right) \Big/ \Gamma\left(\frac{n}{2} \right) \right]^2 \right\}. \tag{51}$$

Analytical calculation of similar numerical characteristics for the uniformly distributed noise $\overline{r_u}$ and $D[r_u]$ is difficult because a multidimensional compositional PDF of a random variable (47) is a discontinuous (piecewise-linear) function. Therefore, these characteristics have been calculated by a statistical model. The results of the analytical and statistical research of the characteristics of uncertainty spheres for a normal and uniform distribution of the noise coordinates are illustrated in Fig. 4-6. The results are expectable due to the law of large numbers. Fig. 4 shows the "virtual" cross-sections by the plane of the multidimensional picture of the displacement points for a normal (left) and uniform (right) distribution of noise coordinates, and calculated for three different values of space dimension at the number of tests equal 10^6. The images of the cross-sections of the spheres are obtained under the even noise power (equals 1) and are normalized per space dimension for convenient comparison.

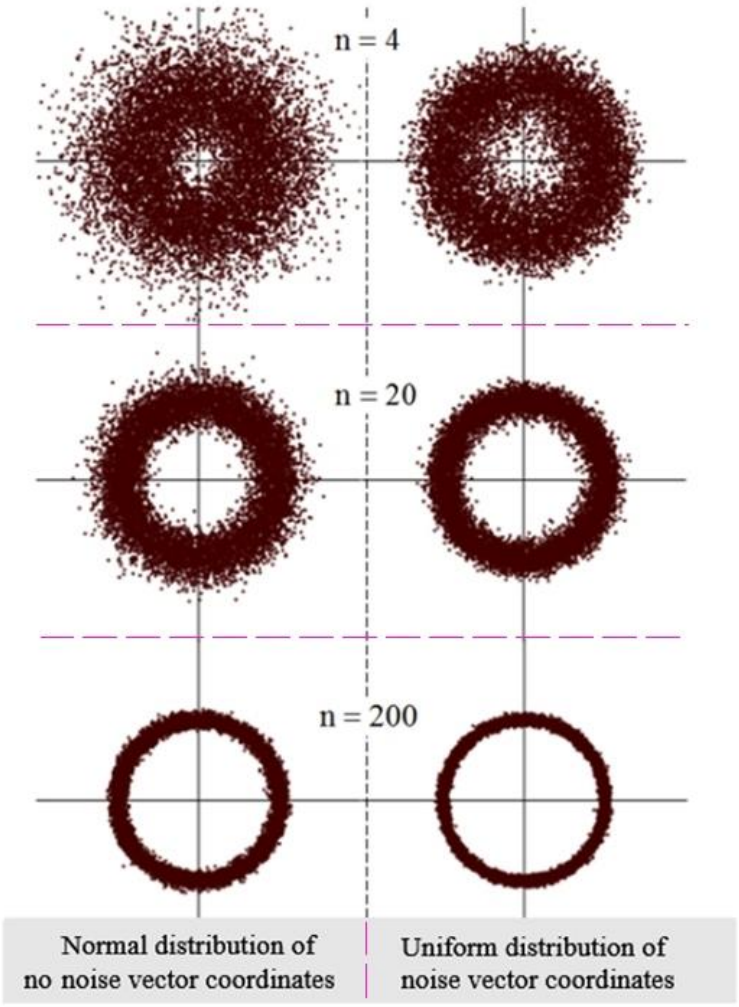


Fig. 4 – The projections of the normalized distributions of the vectors of Gaussian and uniform noise on the plane

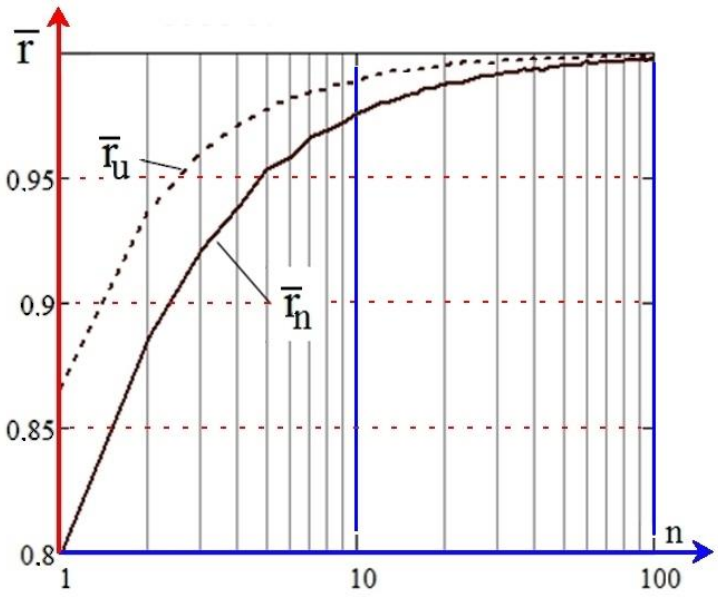


Fig. 5 – Dependence of the function $\overline{r}_u$ and $\overline{r}_n$ from the space dimension n for normal and uniform noise

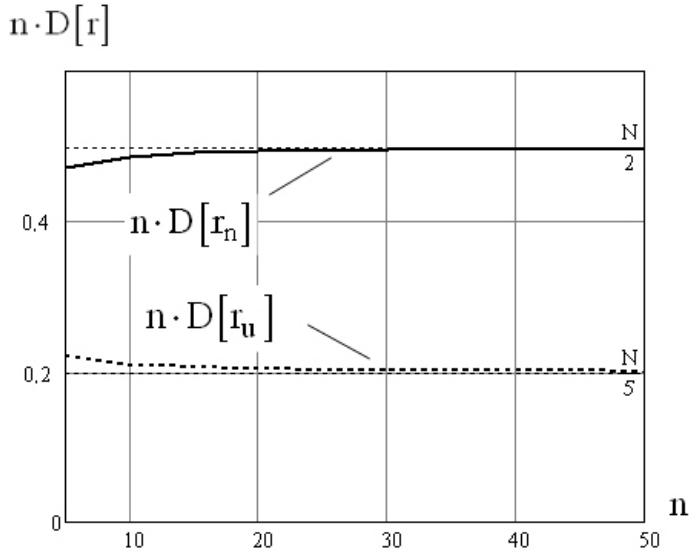


Fig. 6 – The dispersion of the radii of uncertainty spheres for the normal and uniform noise as a function of the dimension of space

The main conclusions from the results of the statistical experiment and analysis of the illustrations are following:

1) the spheres of normal and uniform noise have approximately equal average radii, while the value $\overline{r_u}$ tends to limiting and normalized value slightly faster than $\overline{r_n}$. This phenomenon is illustrated in Fig. 5;

2) the dispersion of scattering of the radius values for the sphere of uniform noise is smaller than the dispersion for the sphere of normal noise (the contours of the spheres on the right is more defined), and calculations yield the following limit ratio:

$$\lim_{n\to\infty}\left(D[r_n]/D[r_u]\right)=2,5, \tag{52}$$

i.e., the effective width of the "ring" of scattering for uniformly distributed noise less, on average, in $\sqrt{2,5}$ times(practically, in any space dimension n), than the same parameter for normal noise. Limiting absolute values of the dispersions for the radii of the spheres are:

$$\lim_{n\to\infty}\left\{n\cdot D[r_n]\right\}=N/2,\ \lim_{n\to\infty}\left\{n\cdot D[r_u]\right\}=N/5. \tag{53}$$

This phenomenon is illustrated by the graphs in Fig. 6 for $N=1$. For the normally distributed noise, the absolute dispersion of the radius increases and tends to the limit value "from below", but for uniformly distributed noise it decreases and tends to the limit value (53) "from above". Finally, we can draw the main and obvious conclusion:

3) the average radii of the uncertainty spheres for the types of the noise PDF under consideration coincide asymptotically:

$$\lim_{n\to\infty}\overline{r_n}=\lim_{n\to\infty}\overline{r_u}=\sqrt{n\cdot N}. \tag{54}$$

This result is a consequence of the law of large numbers. Of course, it can be generalized for any kind of centered PDF of signal and noise, i.e. for any continuous channel with additive noise, which are not statistically associated with signal. The parameters of the geometrical representations of ITS at $n\to\infty$ are affected only by the average power values of continuous signal and noise, but not the type of their distribution!

For similar reasons, the radius of the hyper sphere on non-Gaussian channel output space also coincides with the value determined by the expression (29), then using (27), (29) and (30) we arrive at the same value of channel capacity with uniformly distributed noise: $C'=C=F\log\frac{S+N}{N}$, which contradicts the definition (43) and the statement (45). Thus, two Shannon's works [2] and [6] published at one year interval contradict each other when being applied to a non-Gaussian channel. To the question "which of two methods of determining the capacity, the analytical (entropic) or geometrical, is correct?" – there is the definite answer: the geometrical one. The correctness of the geometrical approach can easily be verified by the statistical modeling of a random code [12]. Analytical

method gives the result which coincides with the result of the geometrical method in the only case when signal and noise are Gaussian processes. It's just a coincidence which can be explained by the properties of normal distribution, which has a special significance in the theory of probability and stochastic processes. Due to the mentioned reasons, the methodology of using "entropic power" and the boundaries defined by (37) are not correct.

The results of the analysis for both non-Gaussian and Gaussian channels (as, indeed, for any other model) have shown that these channels have the same capacity $C = C'$, the value of which depends only on the signal/noise ratio and the channel bandwidth. Therefore, the definition of C (20) as the limit of information transmission rate in a Gaussian channel with additive noise is not correct, to say the least.

The true physical meaning of capacity in the geometric derivation is to determine the *maximum of information transmission rate through a channel with any kind of additive noise* ***when the channel encoding and the maximum likelihood rule in decoding are used***.

Consequently, capacity is not a channel characteristic, it is the *natural limit which arises for any continuous channel model, as soon as we decide to use the encoding of information* (in the sense of making the decision according to the results of the comparison between the channel output and the known samples of valid signal realization). As a result, it is necessary to partition the signal space at the channel output into the fields of "similarity" which, in fact, are the spheres of uncertainty in the geometric representation in Fig. 2. These fields will not overlap as long as the noise power at a fixed transmitter power budget does not exceed a permissible value. This value does define the so-called capacity (actually, the limit rate of the best achievable code). The dominant axiomatic inevitability of code usage and the decision-making process based on the "the greatest similarity" principle are the source of fundamental limitations in the existing information theory paradigm. In other words, the scant achievements of the modern information transmission theory are the consequence of invariable usage of the so-called *maximum likelihood rule*.

In conclusion of this section we'd like to present some considerations as an additional argument for proving the incorrectness of the existing analytical definition of capacity as the maximum average mutual information, considered in Sec. 1. In the quotation from [2] (see Sec. 2), decoding is considered as the process of comparing the noise sample with one of $M = 2^k$ combinations of the source symbols. Therefore, obviously, the entropy of that sample can be defined correctly not by the formula (17), but as the uncertainty of discrete choice (according to the principle (2)), i.e.

$$H(X) = \log 2^k = k. \tag{55}$$

Therefore, it is this definition that should be used in the calculations (17) – (20). This leads to the another collapse, because in the same expression two different definitions of the entropy (for the discrete and the continuous choice) will be present which, according to Shannon, exist in different measurement systems.

4 The rule of maximum likelihood

Cramer Theorem (1740):
"There is no other method of treatment of the experimental results, which would give a better approximation to the truth than the maximum likelihood method."

The name of the rule (method) - the Maximum Likelihood Rule (MLR) is appropriate to its role in the statistical estimation of the random experience realizations and the decision-making processes under conditions of multiple-hypothesis. Modern information transmission paradigm in all known practical applications deals with the decision-making process concerning the noisy channel output state under the conditions of equiprobable hypotheses, i.e. all the source messages are assumed to be equally probable, and the effect of noise in the channel on them is assumed to be same (symmetric). This explains why other statistical methods and decision-making criteria are no alternative to

the MLR. Without much exaggeration we can say that the rule of maximum likelihood came to the statistical theory of communication from our life experience. We always try to hear the phrase in a disturbing noise or to recognize the object in low visibility conditions, subconsciously using the algorithm: "what (known to us) does it most look like?" This explains why the usage of the MLR in all standard applications of the information transmission theory is axiomatic.

The quotation from [2], which has been already referred to (see Sec. 2 of this paper), reflects the justifiable (taking into consideration our physiological experience) opinion of Shannon that the decoder on the channel output has to make a decision on the received codeword (signal) by comparing the proximity (in the mean square sense) of the received sample of a random process at the channel output with the samples available to the receiver.

The same approach can be observed in the description of the ideal (according to Kotelnikov) receiver for the non-coded modulation [1] (quotation 2): «… we assume that, depending on the total oscillation y(t), which affects the receiver input, it is certain to reproduce one of the possible message values $S_1(t),\ldots,S_m(t)$. … Obviously … full range of possible values y(t) can be divided into m non-overlapping areas. … The correct messages will be reproduced more or less frequently according to the configuration of the areas determined by the receiver. … We will call the receiver the ideal one when it is characterized by such (correctly selected) areas and thereby gives the minimum number of incorrectly reproduced messages when noise is applied».

Consequently, the basic postulate of the modern theory of potential noise immunity [1], as well as the error-correcting coding theory [2], is the rule of processing noisy signals (codes) based on the maximum likelihood (or the maximum similarity), which is used by the authors as the foundation for the further theories.

If the values of apriority probabilities of source messages are the same, the mathematical formulation of the MLR in the selection of k-th hypothesis from m alternatives is following:

$$\frac{f(S_k|y)}{f(S_i|y)} > 1, \text{ for all } i \in [1,m],\ i \neq k, \tag{56}$$

where $f(S_i|y)$ – the likelihood function recorded for message S_i. The problem of finding the most reliable solution comes to maximizing the likelihood function, and, in some cases, may have an analytical (non-exhaustive search) resolution based on methods of finding the extremum known from the mathematical analysis. In cases for a continuous channel (see the quotations 1 and 2 above), the likelihood function for the message S_i on the duration T can be expressed via the Euclidean (Hilbert) distance:

$$f(S_i|y) = \left\{ \int_T \left[S_i(t) - y(t) \right]^2 dt \right\}^{-1/2}. \tag{57}$$

In accordance with the maximum likelihood (maximum similarity) principle, the hypothesis, which has maximum of the function (57), is considered to be true [1,2]. Resorting to such a rule, we automatically introduce a limit on the permissible intensity of noise, i.e. *we limit from below S/N ratio at which the output signal point will not be outside its own area of similarity*. This process originates all the basic statements and, so-called, the fundamental limits of information transmission theory. These limits (the most important of which is, undoubtedly, channel capacity) are extremely rigid, unfortunately, and that is the reason for the scant achievements of the information transmission theory.

What is the value of probability P, which describes the similarity of the process at the channel output to the true transmitted message at the low S/N ratio? The answer is obvious – it is very small. Let assume that the channel alphabet allows you to send m different messages that may appear with an equal regularity. Then, for the fixed signal power S and increasing of noise power N it is true that:

$$\lim_{N\to\infty} P = m^{-1}; \quad \lim_{m\to\infty} P = 0. \tag{58}$$

With any heavy noise (if the rate is higher than channel capacity) the process at a channel output with high probability is not similar to the true transmitted message, since its representing point is equally likely to be in the area of similarity of almost any of the m possible messages. When signal points in n-dimensional space are packed most densely [11], the number of uncertainty spheres, which are adjacent to the similarity sphere of the true transmitted signal may be too large. It does not allow to create a multi-dimensional ordered manipulation codes (such as Gray code), which minimize the number of distorted binary symbols at errors of the true message transformation to the nearest to it in the ITS space. For example, at $n = 24$ there is the densest packing based on the Leech lattice and built with the Golay binary code [10, 11], in which the surface of one sphere is adjoined by 196560 surrounding spheres. If on the basis of this lattice any redundant $(24, k)$-codes with $k = 1, 2, \ldots, 18$, are constructed, it will be possible to provide mutual equidistance between all signal (code) points. Even if channel capacity is exceeded insignificantly (small overlapping of the uncertainty spheres), reception of any codeword on the channel output on the basis of MLR is almost equiprobable and practically independent from the transmitted word (message). In such conditions maximum likelihood rule usage certainly leads to an error in the reception. Therefore, there is a paradox and contradiction: on the one hand, MLR is the best way to receive, which minimizes the probability of errors at a low noise; on the other hand – the rule itself is the cause of limitations on the permissible rate and/or noise power. Can the decision-making rule be modified when we use encoding and probabilistic estimation of the channel output state?

5 Can codes work without Maximum Likelihood Rule?

It is convenient to estimate the possibility of changing the decision rule, when the true message is not considered to be the closest one to the realization on the channel output, with the help of the presentation of ITS space by Poisson field of points [12]. A random or ordered algebraic code being constructed, its codebook (a plurality of signal points) forms a random (Poisson) field of the points in a n-dimensional space, as following conditions are always satisfied:

1) at a fixed average power budget of the transmitter all the points of code words are placed in the limited volume of the multidimensional space, and with increasing n this placement asymptotically approaches a uniform (for random code) one, i.e. the density of the field of points is constant throughout the volume of code space;

2) the probability of occurrence of an arbitrary number of points in any volume of space does not depend on the quantity of points falling into any volumes which do not intersect the chosen one;

3) the probability of two or more points falling into the elementary volume is negligible in comparison with the probability of one point falling into it.

Let's assume that the transmission rate in an arbitrary Gaussian channel exceeds its capacity. In the geometrical representation it will lead to the mutual crossing of the uncertainty spheres which is shown for the fragment of channel output space in Fig. 7. To simulate the situation let's use the known [9] analytical description of PDF $\varphi(\Delta)$ of the random variable of the displacement under the noise influence $\Delta = \sqrt{n} \cdot r_n$ (here r_n is determined by the formula (46)):

$$\varphi(\Delta) = \frac{2\Delta^{n-1}}{\Gamma\left(\frac{n}{2}\right)\sqrt{2N}^{n}} \exp\left(-\frac{\Delta^2}{2N}\right). \tag{59}$$

The numerical characteristics $\varphi(\Delta)$ are derived from (50), (51) as follows:

$$M[\Delta] = \sqrt{n} \cdot \overline{r_n}, \; D[\Delta] = n \cdot D[r_n]. \tag{60}$$

Let the message, which corresponds to the point 1,be transmitted over the channel under the noise. Displacement caused by the noise is such that the point 2 is available for the receiver to observe at the channel output. Let's also assume that value of displacement is $\Delta = (M[\Delta] + \delta)$. Using

the MLR in this situation will identify the point 3 (which is the closest one to the received point 2) as true transmitted, which, obviously, leads to the error.

Let's modify the decision-making rule as follows: taking the point 2 observed at the channel output as the center and let's reconstruct the surface of the sphere with the radius M[Δ] around it. Then, checking all codebook points one by one we can identify the point which is the closest to this surface. This point will be considered as true transmitted. In accordance with the rule described in Fig. 7, the point 1 located at a distance δ from the surface of the auxiliary sphere is the true transmitted. This corresponds to the error-free receiver solution in this example. Let's name this decision-making rule the "Uncertainty Sphere Rule" (USR). According to this rule, not the message, which is the most similar one to the observed channel output realization, is considered to be the true, but the message, which is the nearest one to the surface of the sphere with the radius M[Δ] drawn around the observed output point.

Likelihood function of an arbitrary signal S_i for the USR can be written in the form:

$$f\left(S_i|y\right)=\left|\int_T\left[S_i\left(t\right)-y\left(t\right)\right]^2dt-M\left[\Delta\right]^2\right|^{-1/2}. \tag{61}$$

The signal having a maximum value of the function (61) is considered to be truly transmitted. The described rule will lead to the error-free decision only on condition that the auxiliary sphere around the received point (on the Fig. 7 – a point 2) has a radius which is exactly equal to the magnitude of the noise displacement of the transmitted point, i.e. if the noise power added to the transmitted signal (codeword) in a particular realization of the observed channel output is known precisely.

However, since it is impossible to know the exact power of the noise component in the particular received realization of a signal-noise mixture, then the auxiliary sphere can be outlined only by a radius which is equal to its mathematical expectation M[Δ]. This can lead to a wrong decision if any other codebook point will occur in the layer (with the thickness of δ) between two concentric spheres with radii $\Delta=\left(M\left[\Delta\right]+\delta\right)$ and M[Δ] (the hatched ring in Fig. 7).

On the basis of the Poisson field properties, the probability of a wrong decision can be calculated as a function of S, N, n. The occurrence of at least one code point within the space between two concentric spheres will lead to the error. For the Poisson field, the probability of this event is:

$$P\left(\lambda,\Delta\right)=1-\exp\left(-\lambda\left(m\right)\cdot V\left(\Delta\right)\right), \tag{62}$$

where λ – a field density, containing m points:

$$\lambda\left(m\right)=m/V_{S+N}; \tag{63}$$

the value V_{S+N} is defined by (29);

V(Δ) – the volume of a concentric layer around the auxiliary sphere:

$$V\left(\Delta\right)=\frac{\sqrt{\pi}^n}{\Gamma\left(\frac{n}{2}+1\right)}\begin{cases}\left\{M\left[\Delta\right]^n-\Delta^n\right\}, \text{ when } 0\le\Delta\le M\left[\Delta\right];\\ \left\{\Delta^n-M\left[\Delta\right]^n\right\}, \text{ when } \Delta>M\left[\Delta\right].\end{cases} \tag{64}$$

Using (62) - (64) and averaging the result in accordance with the distribution (59) we can calculate the probability of error in decoding by USR:

$$P_{er}=\int_0^\infty\varphi\left(\Delta\right)P\left(\lambda,\Delta\right)d\Delta. \tag{65}$$

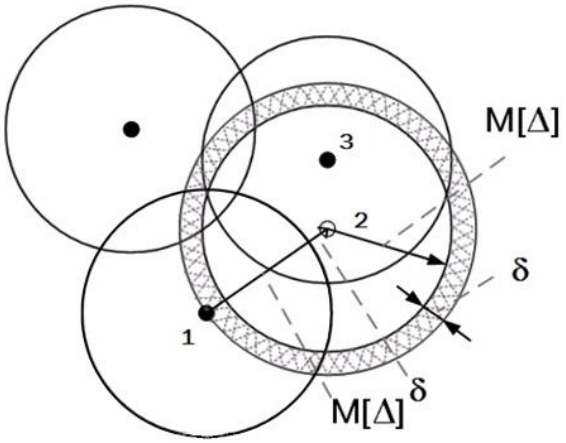


Fig. 7 – Geometric illustration the of uncertainty sphere rules (USR)

With the set values S, N and n, the number of signal points m_C in the code space which corresponds to the capacity is derived from (30), and the density of the field of points $\lambda(m_C)$ is calculated by the formula (63). Introducing the coefficient α changes of the transmission rate per one channel usage, we can model the situations, when the rate R exceeds the capacity C, which leads to intercrossing uncertainty spheres:

$$R > C \to R = \alpha \cdot C \to \alpha > 1 \Rightarrow m = (m_C)^{\alpha}; \quad (66)$$

or, by contrast, does not reach the channel capacity (uncertainty spheres do not intersect having a certain margin):

$$R < C \to R = \alpha \cdot C \to \alpha < 1 \Rightarrow m = (m_C)^{\alpha}. \quad (67)$$

For these expressions the argument, that regulates the simulated rate, is the number of points of different signals (codewords) for a fixed volume of signal space. For $m > m_C$ the channel capacity is exceeded, and for $m < m_C$ – the transmission rate does not reach the channel capacity. The coefficient α in (66) and (67) is located in the exponent because the transmission rate is measured by the logarithm of m.

The results of the calculation of a wrong decision probability (65) for USR with different values of the coefficient α and $S = N = 1$ are shownin Fig. 8.

Alas, the main conclusion from the analysis of the curves in Fig. 8 is disappointing – the USR (so attractive in the case of Fig. 7) leads to the same result as the MLR does! For $R > C$ the probability of error, when the space dimension (the number of degrees of freedom or the random code block length) increases, tends to unity monotonically.

When $R < C$ – the probability of error can became arbitrarily small with the corresponding increase of n. This result can be explained by the properties of multidimensional spheres, namely, almost all their volume is concentrated in a small area adjacent to the surface. In this area surrounding immediately the surface of the auxiliary sphere, PDF (defined by the expression (59)) reaches its highest values. Therefore, when the dimension of the space increases, the effective volume of layers around of the auxiliary sphere grows faster than the density of the field of points decreases. Certainly, we can try to formulate other criteria and decision-making rules, but, obviously, the results will be no better than the result of the MLR (*Cramer's theorem*). On the basis of the conducted simulation, it is possible to make an unambiguous conclusion – the MLR is the best and the only acceptable rule of statistical solutions for the codes – there are no alternatives to it.

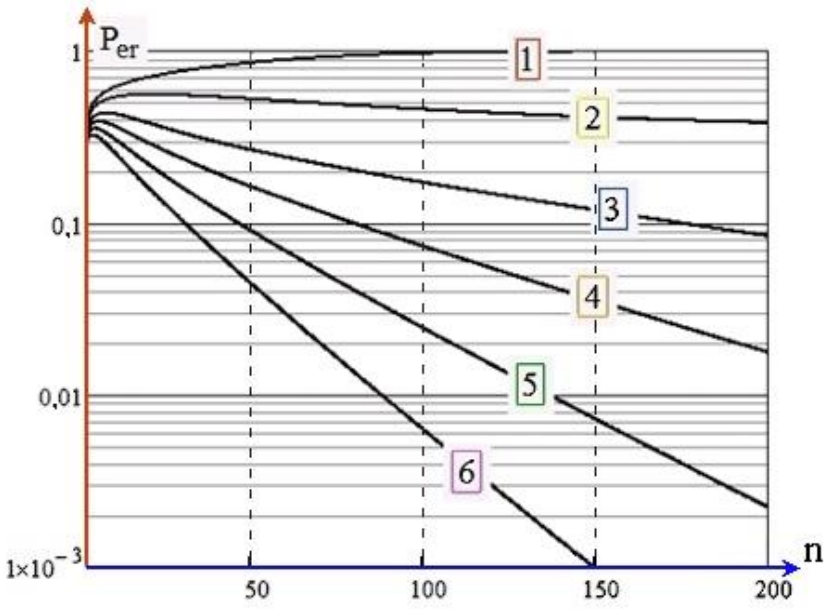


Fig. 8 – The dependencies $P_{er}(n)$ when using RSU and different rates
(1 → R = 1,05·C; 2 → R = 0,95·C; 3 → R = 0,8·C;
4 → R = 0,7·C; 5 → R = 0,6·C; 6 → R = 0,5·C)

6 Discussion of results and conclusions

The main conclusions from the results of the analysis of mathematical and physical nature of channel capacity, as well as from the modern information transmission theory contradictions are following:

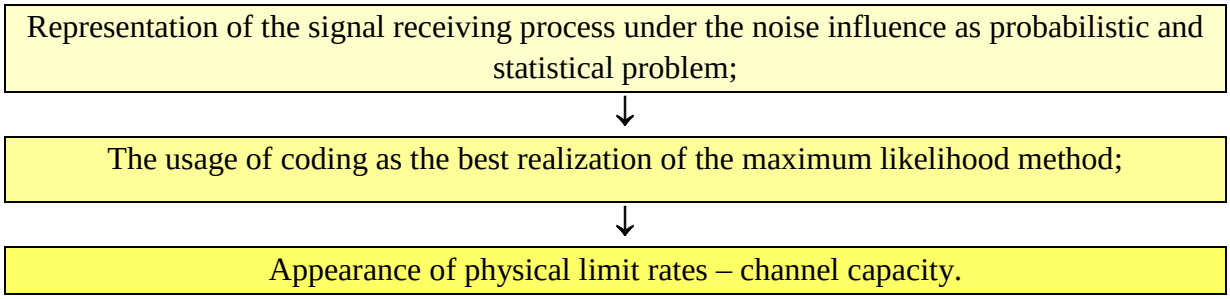
1. The probability-entropy approach to the analytical determination of capacity of continuous channels, which uses the concept of the average mutual information between input and output (5) – (15), can be considered as the correct one only in case when the distribution of the source and the noise is Gaussian (16), (17). Since the usage of this approach for non-Gaussian models of continuous channels leads to the erroneous results (39) – (45), then the unjustified conclusion about the impossibility of analytical determining the capacity for such models has been made in many published works.

2. The mathematical definition (31) describes correctly the channel capacity value for any continuous channel where noise is a stationary random process. The value of channel capacity is not affected by a noise distribution type and is determined only by the signal/noise ratio and channel bandwidth. Different noise distributions manifest only in changes in the speed of approaching to the capacity when the duration of the samples of random noise code sequences increases.

3. The correct geometric definition of channel capacity determines its physical nature as the limit of in-formation transmission rate in a channel with any kind of additive noise, when the coding/decoding is used and the maximum likelihood rule is applied in decoding. Channel capacity is the physical limit only for systems, which use the maximum likelihood method.

4. The maximum likelihood rule is the best and only decision-making rule for the decoding. At the same time capacity is an indirect determination of the lower boundary of the signal/noise ratio when the noise displacement of message points in the multidimensional space of the output channel is not outside of the fixed "are-as of similarity". The existence of these areas is defined by the maximum likelihood method nature. Thus, on the one hand, the maximum likelihood rule is the best rule of the statistical decision-making, and on the other hand, it causes the appearance of the physical limit – channel capacity. Abandoning the MLR usage, which causes the appearance of the physical limit of data rates, in case when the work of receiver consists in solving the probabilistic and statistical problem, is impossible!

5. These considerations give rise to the following logical causal chain of factors that has led to the crisis in the information transmission theory development:



As shown in this paper, we cannot break the chain represented above:

– it is impossible to exceed the capacity without abandoning the maximum likelihood method;

– it is impossible to abandon the maximum likelihood method when the work of receiver consists in solving the probabilistic problem.

Thus, the root of the considered problems is a probabilistic approach to receiving signals under the noise influence and even the most ambitious assumptions give no alternatives to it. However, it is not so. We do use not all the opportunities, which the nature offers for handling noisy digital signals in continuous channels. However, this will be discussed in the further publications.

References

[1] Kotel'nikov V. A. Teoriya potentsial'noi pomekhoustoichivosti / V. A.Kotel'nikov. – Moskva: Gosenergoizdat, 1956. – 152 s.
[2] Shannon C. E. A Mathematical Theory of Communication / C. E. Shannon // Bell Syst. Tech . – 1948, July-October. – Vol. 27. – P. 379 – 423; 623 – 656.
[3] Pomekhoustoichivost' i effektivnost' sistem peredachi informatsii / A. G. Zyuko, A. I. Fal'ko, I. P. Panfilov, V. L. Banket, P. V. Ivashchenko; pod red. A. G. Zyuko. – Moskva: Radio i svyaz', 1985. – 272 s.
[4] Gallager R. Teoriya informatsii i nadezhnaya svyaz' / R. Gallager; [per. s angl.; pod red. M. S. Pinskera i B. S. Tsybakova]. − Moskva: Sov. radio, 1974. − 720 s.
[5] Dobrushin R. L. Teoriya peredachi informatsii / R. L. Dobrushin, B. S.Tsybakov // Vestnik AN SSSR. – 1998. – № 6. – S. 76–81.
[6] Shannon C. E. Communication in the presence of noise / C. E. Shannon // Proc. IRE. – 1949, January. – Vol. 37. – P. 10 – 21.
[7] Verdu S. Fifty Years of Shannon Theory / S. Verdu // IEEE Transactions on Information Theory. – 1998, October. – Vol. 44. – № 6. – P. 2057 – 2078.
[8] Verdu S. A General Formula for Channel Capacity / S.Verdu, Te Sun Han // IEEE Transactions on Information Theory. – 1994, July. – Vol. 40. – № 4. – P. 1147 – 1157.
[9] Venttsel' E.S. Teoriya veroyatnostei i ee inzhenernye prilozheniya / E.S. Venttsel', L.A.Ovcharov. – Moskva: Vysshaya shkola, 2000. – 480 s.
[10] Khemming R.V. Teoriya kodirovaniya i teoriya informatsii / R.V.Khemming; per. s angl. – Moskva: Radio i svyaz', 1983. – 176 s.
[11] Konvei Dzh. Upakovki sharov, reshetki i gruppy / Dzh.Konvei, N.Sloen: [v 2 tomakh]. – Moskva: Mir, 1990.
[12] Rassomakhin S.G. Tekhnologiya psevdosluchainogo kodirovaniya v setevykh kommunikatsionnykh protokolakh kanal'nogo urovnya / S. G. Rassomakhin // Systemy obrobky informacii'. – 2012. – T.2. – Vyp. 3(101). – S. 206 – 211.

Рецензент: Віктор Краснобаєв, д.т.н., проф., Харківський національний університет імені В.Н. Каразіна, Харків, Україна. E-mail: krasnobayev@karazin.ua

Надійшло: Листопад 2016.

Автори: Сергій Рассомахін, д.т.н., проф., Харківський національний університет імені В.Н. Каразіна, Харків, Україна. E-mail: rassomakhin@karazin.ua

Математична і фізична природа пропускної здатності каналів.

Анотація: Розглянуто класичні методичні підходи до визначення пропускної здатності каналу (зв'язку). Показано протиріччя між аналітичним і геометричним визначенням максимально досяжної швидкості передачі. Аналізується об'єктивність

правила максимальної правдоподібності при його використанні для каналів з низьким відношенням сигнал/шум. Виконано коректне визначення математичної і фізичної сутності пропускної здатності каналу. Доведена інваріантність величини пропускної здатності до виду розподілу шуму в неперервних каналах. Наведено основні причини кризи в розвитку теорії передачі інформації.

Ключові слова: диференціальна ентропія, пропускна здатність каналу, правило максимальної правдоподібності, сфера невизначеності, випадкове кодування

Рецензент: Виктор Краснобаев, д.т.н., проф., Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина.
E-mail: krasnobayev@karazin.ua

Поступила: Ноябрь 2016.

Авторы: Сергей Рассомахин, д.т.н., проф., Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина.
E-mail: rassomakhin@karazin.ua

Математическая и физическая природа пропускной способности каналов.

Аннотация: Рассмотрены классические методические подходы к определению пропускной способности канала (связи). Показаны противоречия между аналитическим и геометрическим определением максимально достижимой скорости передачи. Анализируется объективность правила максимального правдоподобия при его использовании для каналов с низким отношением сигнал/шум. Выполнено корректное определение математической и физической сущности пропускной способности канала. Доказана инвариантность величины пропускной способности к виду распределения шума в непрерывных каналах. Приведены основные причины кризиса в развитии теории передачи информации.

Ключевые слова: дифференциальная энтропия, пропускная способность канала, правило максимального правдоподобия, сфера неопределенности, случайное кодирование

УДК 004.052.42

МЕТОД КОМПИЛЯЦИОННО-СЕМАНТИЧЕСКОЙ ВЕРИФИКАЦИИ ВРЕМЯПАРАМЕТРИЗОВАННЫХ МУЛЬТИПАРАЛЛЕЛЬНЫХ ПРОГРАММ

Елена Толстолужская, Дмитрий Толстолужский, Ольга Мороз

Харьковский национальный университет имени В. Н. Каразина, площадь Свободы, 4, г. Харьков, 61022, Украина.
tps@karazin.ua

Рецензент: Георгий Кучук, д.т.н., проф., Национальный технический университет «Харьковский политехнический институт», ул. Кирпичева, 21, г. Харьков, 61000, Украина.
kuchuk56@mail.ru

Поступила в ноябре 2016

Аннотация. *Приводится определение времяпараметризованных мультипараллельных программ, которые (в отличие от общепринятой трактовки параллельных программ) содержат спецификации моментов начала выполнения операций/функций, а также подмножества таких операций/функций. Обосновывается необходимость разработки новых методов верификации времяпараметризованных параллельных программ. Раскрываются этапы решения задачи компиляционно-семантической верификации времяпараметризованных мультипараллельных программ в интересах создания систем автоматического синтеза высокоэффективных параллельных программ для вычислительных систем различных классов. Приводится иллюстрирующий пример основных этапов метода.*

Ключевые слова: *времяпараметризованные мультипараллельные программы, верификация параллельных программ, компиляционно-семантическая верификация, семантико-числовые спецификации.*

1 Анализ проблематики и постановка задачи

Анализ литературы показывает, что одним из перспективных направлений развития технологий параллельного программирования являются исследования в области автоматического программирования времяпараметризованных мультипараллельных программ [1–3].

Времяпараметризованная мультипараллельная программа определяется (в отличие от принятой в настоящее время трактовки статических параллельных программ) как конструкция, которая содержит в явном виде спецификации следующих категорий информации:

– множество объектов – данных, над которыми должны выполняться действия (задаваемые составом операций/функций алгоритмического языка высокого уровня);

– множество действий (операций/функций), которые должны быть выполнены над данными для решения задачи;

– множество статических связей, задающих отношения упорядоченности операций/функций по данным и по управлению;

– упорядоченность операций/функций в динамике параллельного вычислительного процесса, задаваемую множеством моментов времени начала выполнения операций/функций;

– разделение множества операций/функций на временные фрагменты (множественные временные операторы (МВО)), включающие совокупность операций/функций, выполнение которых начинается одновременно в конкретный момент дискретного времени;

– разделение множества данных на фрагменты данных, поставленные в однозначное соответствие множественным временным операторам и используемые в соответствующие моменты дискретного времени;

– наличие информации о разбиении множества команд различных фрагментов на подмножества (нити), выполняемые соответствующими модулями/процессорами;

– наличие информации о единицах измерения физических величин данных.

Новизна класса времяпараметризованных мультипараллельных программ обуславливает необходимость решения задачи разработки новых методов верификации, ориентированных на применение формата структур семантико-числовой спецификации [1].

В качестве одного из этапов верификации времяпараметризованных мультипараллельных программ предлагается использовать компиляционно-семантическую верификацию. Основой компиляционно-семантической верификации является использование задаваемых пользователем единиц измерения («семантики») исходных данных и выходных результатов задач, автоматическое определение в процессе решения задач единиц измерения промежуточных и выходных результатов и их сравнение с единицами измерения выходных результатов, заданными пользователем.

Целью статьи является описание нового метода компиляционно-семантической верификации времяпараметризованных мультипараллельных программ в интересах создания систем автоматического синтеза высокоэффективных параллельных программ для вычислительных систем различных классов.

2 Основная часть

Исходные данные компиляционно-семантической верификации:

- исходная статическая Си-программа задачи;
- Си-текст временной мультипараллельной программы, соответствующей исходной Си-программе и удовлетворяющей требованиям/ограничениям пользователя;
- семантическая база данных (БД) *«SEM»* единиц измерения физических величин;
- семантика (единицы измерения) исходных данных и выходных Си-программы, задаваемые пользователем – структура *US_SEM*.

Требуется:

- проверить семантическую корректность исходной Си-программы задачи;
- проверить семантическую корректность времяпараметризованной мультипараллельной программы.
- проверить логическую эквивалентность временной мультипараллельной программы и исходной Си-программы задачи.

Выходными данными являются:

1. Семантико-числовая спецификация (СЧС) исходного кода программы (список единиц измерения исходных данных задачи, значений размерности результатов промежуточных вычислений и значений размерности результатов выполнения Си-программы).

2. Результаты проверки идентичности значений размерности результатов выполнения Си-программы и пользовательской семантической спецификации задачи.

3. Результаты проверки семантической корректности временной модели и текста времяпараметризованной мультипараллельной программы.

Основные этапы метода компиляционно-семантической верификации времяпараметризованных мультипараллельных программ приведены на рис. 1.

Рассмотрим содержание основных этапов метода компиляционно-семантической верификации.

Этап 1 (символ 2 на рис. 1) обеспечивает для Си-программы (см. рис. 2) синтез структур *BF* и *CF* семантико-числовой спецификации (табл. 3 представляет *BF*).

На этапе 2 (символ 3 рис. 1) обеспечивается графическая визуализация исходной Си-программы в виде Си-графа (см. рис. 2), исходя из структур СЧС *BF*, сформированных при выполнении 1-го этапа. Построение Си-графа осуществляется с помощью средств визуализации параллельных аппаратно-программных объектов, описанных в [1].

Третий этап (символ 4 на рис. 1) обеспечивает синтез единиц измерения данных, формируемых «внутренними» и «выходными» операторами P_j структур СЧС и Си-графа.

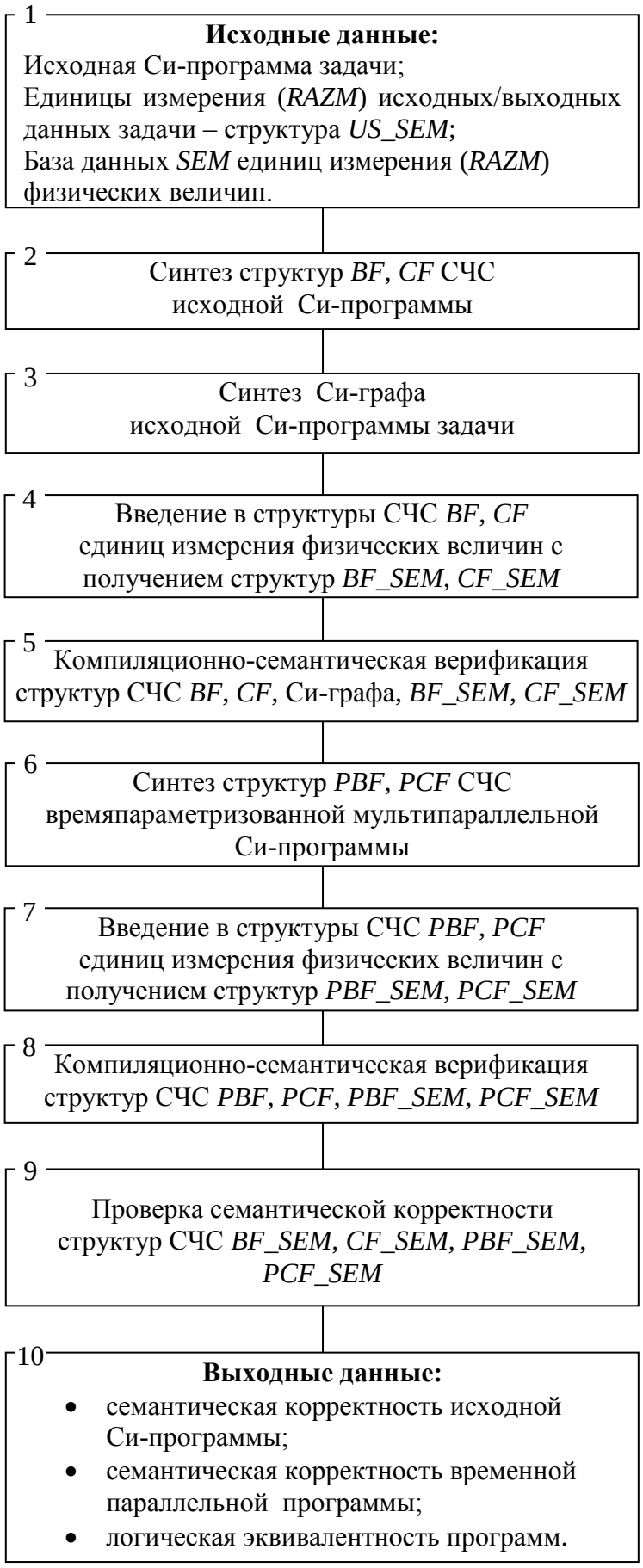


Рис. 1 Основные этапы метода компиляционно-семантической верификации

Синтез единиц измерения данных производится исходя из заданных пользователем единиц измерения исходных и выходных данных Си-программы задачи, типов операций, выполняемых операторами P_j, Си-программы, и общепринятой базы данных единиц измерения физических величин.

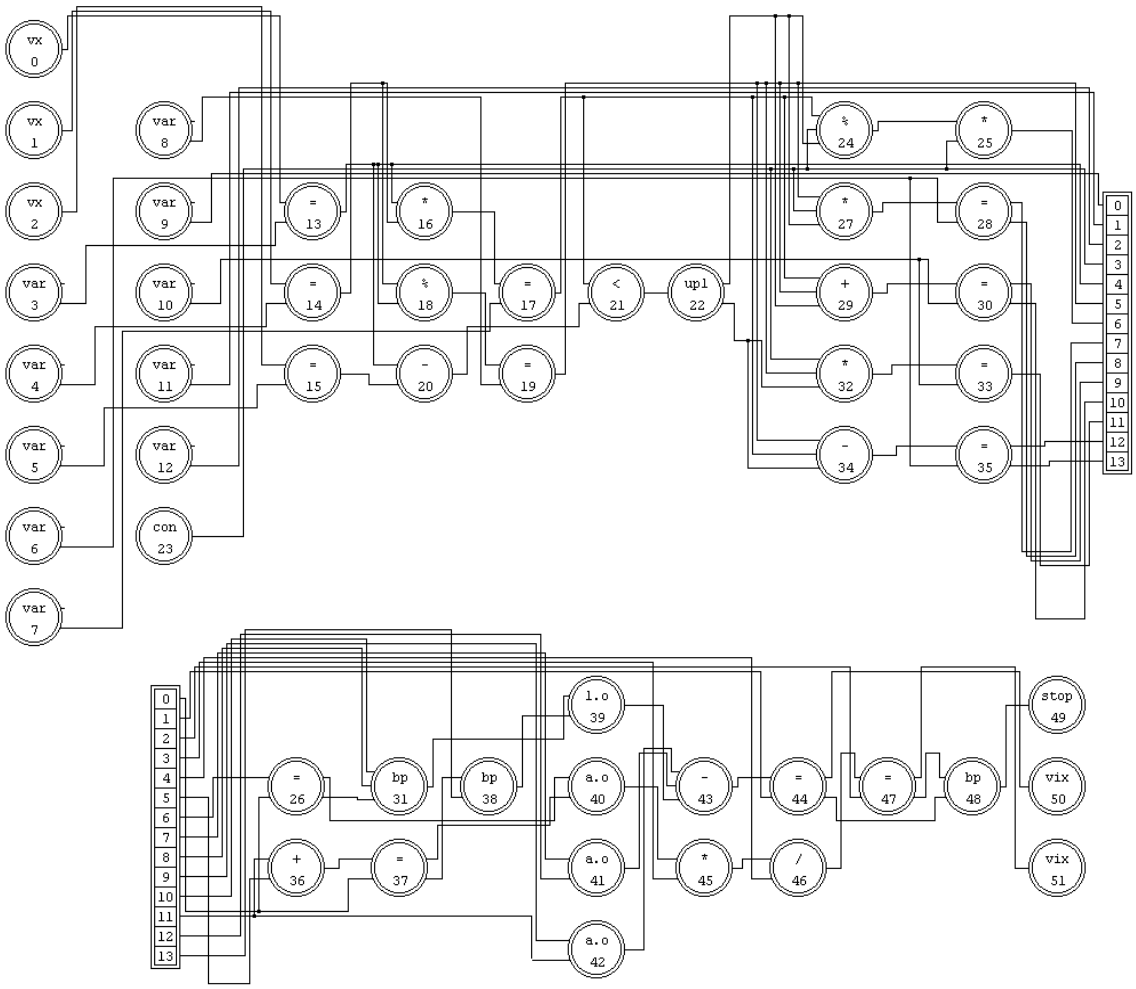


Рис. 2 – Си-граф исходной разветвляющейся Си-программы

Четвёртый этап (символ 5 на рис. 1) путем компиляционной верификации выполняет проверку синтеза структур СЧС Си-программы и соответствующего Си-графа. Методика компиляционной верификации рассмотрена в [1,4]. Кроме этого, 4-й этап обеспечивает проверку семантической корректности структур СЧС *BF_SEM*, *CF_SEM* исходной Си-программы. Проверка осуществляется путем сравнения рассчитанных единиц измерения данных, соответствующих выходным операторам синтезированной структуры *BF_SEM* исходной Си-программы, с единицами измерения выходных данных, заданными пользователем.

На 5-м этапе (символ 6 на рис. 1), аналогично этапу 1 (символ 2 на рис. 1), выполняется синтез структур *PBF*, *PCF* СЧС времяпараметризованной мультипараллельной программы.

На этапе 6 (символ 7 на рис. 1) осуществляется введение в структуры СЧС *PBF*, *PCF* единиц измерения физических величин с получением структур *PBF_SEM*, *PCF_SEM*.

На этапе 8 (символ 8, рис. 1) осуществляется компиляционная верификация структур СЧС *PBF*, *PCF* с одновременной проверкой семантической корректности расширенных структур СЧС *PBF_SEM*, *PCF_SEM* времяпараметризованной мультипараллельной программы. Проверка проводится путем сравнения рассчитанных единиц измерения данных (соответствую-

щих выходным операторам синтезированной структуры *PBF_SEM*) с единицами измерения выходных данных, заданными пользователем.

Проиллюстрируем содержание основных этапов метода компиляционно-семантической верификации с помощью Си-программы, представленной на рис. 3. В состав исходных данных входят также фрагмент базы данных единиц измерения физических величин (см. табл. 1) и перечень данных, задаваемых пользователем (см. табл. 2).

```
#include <stdio.h>
void main(void)
{
int a,b,c,r, k,l,m,p, s,t;
        scanf("%d %d %d\n",&a,&b,&c);
        k = a * b;
        l = b % a;
        if(k < a-c)
        { m = (k % 2) * 2;
         r = l * 2;
        p = k + l;
        }
        else
        { p = 2 * l;
         r = l - k;
         m = p + l;
        }
        s = p - r;
        t = (m * 2) / a;
        printf ("%4d %4d\n",s,t);
}
```

Рис. 3 – Исходная Си-программа

Базовая структура *BF* (см. табл. 3) описывает номера и состав операторов P_j задачи (массив *N*), их типы (массив *TYP*), число входных (массив *SJD*) и выходных (массив *WJD*) связей каждого оператора P_j, идентификаторы операторов (массив *RES*), указатели на начало цепочек сопряженных и внешних операторов (массивы *NSJ* и *NWJ*) для каждого оператора Си-программы. Синтез структур *BF* и *CF* СЧС осуществляется в соответствии с методикой, изложенной в [1].

Таблица 1 – Фрагмент семантической базы данных *«SEM»*

KOD_RAZM	*RAZM*	*KOD_RAZM*	*SEM*
2	м	14	рад/с
3	кг	15	м/(с*с)
4	с	16	рад/(с*с)
5	А	17	1/м
6	К	18	кг/(м*м*м)
7	моль	19	м*м*м/кг
8	кд	20	А/(м*м)
9	рад	21	А/м
10	ср	22	моль/(м*м*м)
11	м*м	23	1/с
12	м*м*м	24	м*м/с
13	м/с	25	кд/(м*м)

Таблица 2 – Структура US_SEM единиц измерения, задаваемых пользователем

N_OP	*REZ*	*VHOD_VIH*	*KOD_SEM*
3	a	0	2
4	b	0	2
5	c	0	2
11	s	1	2
12	t	1	11

Таблица 3 – Базовая структура BF операторов исходной Си-программы

N	*MET*	*TYP*	*NSJ*	*SJD*	*BJ*	*NWJ*	*WJD*	*MP1*	*MP2*	*VH*	*VIH*	*RES*	*N*	*MET*	*TYP*	*NSJ*	*SJD*	*BJ*	*NWJ*	*WJD*	*MP1*	*MP2*	*VH*	*VIH*	*RES*
0	0	58	–1	0	0	0	1	0	0	0	1	a_in	26	0	12	24	2	1	48	2	0	0	2	2	=
1	0	58	–1	0	0	1	1	0	0	0	1	b_in	27	0	3	26	3	1	50	1	0	0	3	1	*
2	0	58	–1	0	0	2	1	0	0	0	1	c_in	28	0	12	29	2	1	51	2	0	0	2	2	=
3	0	47	–1	0	0	3	1	0	0	0	2	a	29	0	1	31	3	1	53	1	0	0	3	1	+
4	0	47	–1	0	0	4	1	0	0	0	2	b	30	0	12	34	2	1	54	2	0	0	2	2	=
5	0	47	–1	0	0	5	1	0	0	0	2	c	31	0	50	36	3	1	56	1	3	0	3	1	bp
6	0	47	–1	0	0	6	2	0	0	0	2	r	32	2	3	39	3	2	57	1	0	0	3	1	*
7	0	47	–1	0	0	8	1	0	0	0	2	k	33	0	12	42	2	2	58	2	0	0	2	1	=
8	0	47	–1	0	0	9	1	0	0	0	2	l	34	0	2	44	3	2	60	1	0	0	3	1	–
9	0	47	–1	0	0	10	2	0	0	0	2	m	35	0	12	47	2	2	61	2	0	0	2	2	=
10	0	47	–1	0	0	12	2	0	0	0	2	P	36	0	1	49	2	2	63	1	0	0	2	1	+
11	0	47	–1	0	0	14	1	0	0	0	2	s	37	0	12	51	2	2	64	2	0	0	2	2	=
12	0	47	–1	0	0	15	1	0	0	0	2	t	38	0	50	53	2	2	66	1	3	0	2	1	bp
13	0	12	0	2	0	16	4	0	0	2	1	=	39	3	54	55	2	3	67	1	0	0	2	1	l.o
14	0	12	2	2	0	20	2	0	0	2	1	=	40	0	53	57	2	3	68	1	0	0	2	1	a.o
15	0	12	4	2	0	22	1	0	0	2	1	=	41	0	53	59	2	3	69	1	0	0	2	1	a.o
16	0	3	6	2	0	23	1	0	0	2	1	*	42	0	53	61	2	3	70	1	0	0	2	1	a.o
17	0	12	8	2	0	24	4	0	0	2	1	=	43	0	2	63	3	3	71	1	0	0	3	1	–
18	0	5	10	2	0	28	1	0	0	2	1	%	44	0	12	66	2	3	72	2	0	0	2	2	=
19	0	12	12	2	0	29	5	0	0	2	1	=	45	0	3	68	2	3	74	1	0	0	2	1	*
20	0	2	14	2	0	34	1	0	0	2	1	–	46	0	4	70	2	3	75	1	0	0	2	1	/
21	0	25	16	2	0	35	1	0	0	2	1	<	47	0	12	72	2	3	76	2	0	0	2	2	=
22	0	51	18	1	0	36	5	1	2	1	2	upl	48	0	50	74	2	3	78	1	980	0	2	1	bp
23	0	57	–1	0	1	41	5	0	0	0	1	C2_	49	980	49	76	1	4	–1	0	0	0	1	0	stop
24	1	5	19	3	1	46	1	0	0	3	1	%	50	0	48	77	1	4	–1	0	0	0	1	0	s out
25	0	3	22	2	1	47	1	0	0	2	1	*	51	0	48	78	1	4	–1	0	0	0	1	0	t_out

Результаты компиляционной верификации структур семантико-числовой спецификации Си-программы и Си-графа *(View of test results)* представлены на рис.4.

Результаты синтеза семантико-числовой спецификации *BF_SEM* Си-программы представлены в табл. 4.

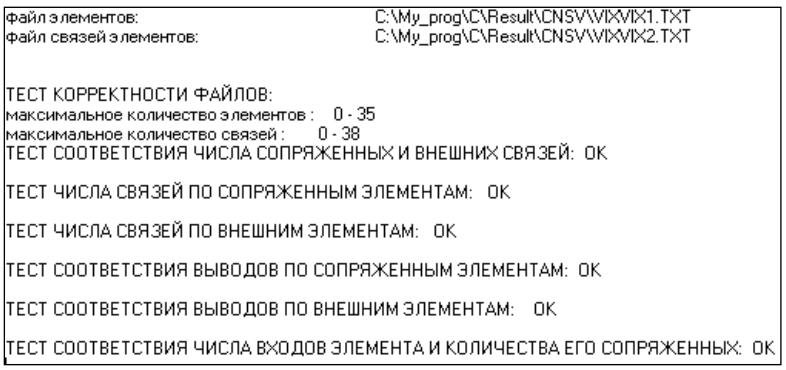


Рис. 4 – Результаты компиляционной верификации

Таблица 4 – Структура *BF_SEM* – результат синтеза единиц измерения операторов исходной Си-программы

N	*TYP*	*NSJ*	*SJD*	*BJ*	*NWJ*	*WJD*	*RES*	*SEM*	*N*	*TYP*	*NSJ*	*SJD*	*BJ*	*NWJ*	*WJD*	*RES*	*SEM*
0	58	–1	0	0	0	1	a_in	*м*	26	12	24	2	1	48	2	=	*м*м*
1	58	–1	0	0	1	1	b_in	*м*	27	3	26	3	1	50	1	*	*безразм.*
2	58	–1	0	0	2	1	c_in	*м*	28	12	29	2	1	51	2	=	*безразм.*
3	47	–1	0	0	3	1	a	*м*	29	1	31	3	1	53	1	+	*м*м*
4	47	–1	0	0	4	1	b	*м*	30	12	34	2	1	54	2	=	*м*м*
5	47	–1	0	0	5	1	c	*м*	31	50	36	3	1	56	1	bp	*безразм.*
6	47	–1	0	0	6	2	r	*безразм.*	32	3	39	3	2	57	1	*	*безразм.*
7	47	–1	0	0	8	1	k	*м*м*	33	12	42	2	2	58	2	=	*безразм.*
8	47	–1	0	0	9	1	1	*безразм.*	34	2	44	3	2	60	1	–	*м*м*
9	47	–1	0	0	10	2	m	*м*м*	35	12	47	2	2	61	2	=	*м*м*
10	47	–1	0	0	12	2	P	*м*м*	36	1	49	2	2	63	1	+	*безразм.*
11	47	–1	0	0	14	1	s	*м*	37	12	51	2	2	64	2	=	*безразм.*
12	47	–1	0	0	15	1	t	*м*м*	38	50	53	2	2	66	1	bp	*безразм.*
13	12	0	2	0	16	4	=	*м*	39	54	55	2	3	67	1	l.o	*безразм.*
14	12	2	2	0	20	2	=	*м*	40	53	57	2	3	68	1	a.o	*м*м*
15	12	4	2	0	22	1	=	*м*	41	53	59	2	3	69	1	a.o	*безразм.*
16	3	6	2	0	23	1	*	*м*м*	42	53	61	2	3	70	1	a.o	*м*м*
17	12	8	2	0	24	4	=	*м*м*	43	2	63	3	3	71	1	–	*м*м*
18	5	10	2	0	28	1	%	*безразм.*	44	12	66	2	3	72	2	=	*м*м*
19	12	12	2	0	29	5	=	*безразм.*	45	3	68	2	3	74	1	*	*м*м*
20	2	14	2	0	34	1	–	*м*	46	4	70	2	3	75	1	/	*м*
21	25	16	2	0	35	1	<	*безразм.*	47	12	72	2	3	76	2	=	*м*
22	51	18	1	0	36	5	upl	*безразм.*	48	50	74	2	3	78	1	bp	*безразм.*
23	57	–1	0	1	41	5	C2_	*безразм.*	49	49	76	1	4	–1	0	stop	*безразм.*
24	5	19	3	1	46	1	%	*м*м*	50	48	77	1	4	–1	0	s_out	*м*м*
25	3	22	2	1	47	1	*	*м*м*	51	48	78	1	4	–1	0	t_out	*м*

В данной таблице приняты следующие обозначения: массив *RAZM* – единицы измерения исходных данных и данных – результатов выполнения операций: «м» – метр; «нет» – отсутствие вычисленного значения выходной переменной; «б/раз» – безразмерная величина; «м*м», «м*м*м» – синтезированные единицы измерения производных величин.

Таблица 4 отображает результаты автоматической проверки идентичности значений размерности результатов выполнения Си-программы и пользовательской семантической спецификации задачи.

3 Выводы

В настоящее время во многих областях науки и техники находят все более широкое применение времяпараметризованные мультипараллельные программы, что обусловило актуальность вопросов верификации и тестирования программ данного класса.

Рассмотренный метод компиляционно-семантической верификации времяпараметризованных мультипараллельных программ, в отличие от уже известных методов, обеспечивает учет целой совокупности групп факторов, оказывающих существенное влияние на эффек-

тивность программных средств параллельных вычислительных систем. К таким факторам следует отнести:

- использование реального времени в качестве одного из основных параметров формального синтеза параллельных программ и соответствующих им временных параллельных процессов;
- явное отражение в конструкциях параллельных программ состава фактически используемых методов параллельной обработки данных;
- явное отражение в конструкциях параллельных программ единиц измерения (семантики) обрабатываемых данных;
- поддержку (в явном виде) структурами временных параллельных программных и аппаратных продуктов требований архитектурной и/или проблемной ориентации;
- поддержку (в явном виде) структурами параллельных программных и аппаратных продуктов требований и ограничений пользователей *(например, обеспечение требуемого времени выполнения программы, заданной тактовой частоты обработки данных цифровым устройством и т. п.)*.

Рассмотренный в данной статье метод позволяет автоматизировать процесс верификации времяпараметризованных мультипараллельных программ, что уменьшает общие трудозатраты на их разработку.

Ссылки

[1] Polyakov G. A. Sintez i analiz parallel'nykh protsessov v adaptivnykh vremyaparametrizovannykh vychislitel'nykh sistemakh: monografiya / G. A. Polyakov, S. I. Shmatkov, E. G. Tolstoluzhskaya, D. A. Tolstoluzhskii. – Khar'kov: KhNU im. V. N. Karazina, 2012. – 670 s.

[2] Voevodin V. V. Parallel'nye vychisleniya / V. V. Voevodin, Vl. V. Voevodin. – Sankt-Peterburg: BKhV – Peterburg, 2004. – 608 s.

[3] Kalbertson R. Bystroe testirovanie / Robert Kalbertson, Braun Kris, Kobb Geri: per. s angl. – Moskva: Izdatel'skii dom «Vil'yams», 2002. – 384 s.

[4] Polyakov G. A. Kompilyatsionnaya metodika verifikatsii statiko-dinamicheskikh ob"ektov avtomaticheskogo proektirovaniya mul'tiparallel'nykh tsifrovykh ustroistv. / G. A. Polyakov, D. A. Tolstoluzhskii // Prikladnaya radioelektronika. – 2005. – T.1. – № 2. – S. 37–41.

Reviewer: Georgiy Kuchuk, Doctor of Technical Sciences, Full Professor, Professor of the Department of Computer Science and Programming, National Technical University "Kharkiv Polytechnic Institute", st. Kirpichova, 21, Kharkiv, Ukraine.
E-mail: kuchuk56@mail.ru

Received: November 2016.

Authors: Olena Tolstoluzka, Doctor of Technical Sciences, V. N. Karazin Kharkiv National University, Svobody Sq., 4, Kharkiv, 61022, Ukrain.
Email: tps@karazin.ua
Dmitriy Tolstoluzkiy, engineer, V. N. Karazin Kharkiv National University, Svobody Sq., 4, Kharkiv, 61022, Ukrain.
Email: tps@karazin.ua
Olga Moroz, senior lecturer, V. N. Karazin Kharkiv National University, Svobody Sq., 4, Kharkiv, 61022, Ukrain.
Email: tps@karazin.ua

Compilations method and semantic verification time parameterized of multiparallel programs.

Abstract: It provides definitions vremyaparametrizovannyh multiparallelnyh programs that (in contrast to the conventional treatment of parallel programs) contain moments of the start of operations specifications / features, as well as a subset of such operations / functions. The necessity of development of new methods of verification vremyaparametrizovannyh parallel programs. Reveals the stages of solving the problem, a compilation of semantic verification vremyaparametrizovannyh multiparallelnyh programs for the creation of automated synthesis systems, high-performance parallel programs for computing systems of various classes. We present an example illustrating the basic steps of the method.

Keywords: time parameterizing, multiparallelism, parallel programs, verification of parallel programs, compilation-semantic verification, semantic-numerical specification.

Рецензент: Георгій Кучук, доктор технічних наук, професор, професор кафедри обчислювальної техніки та програмування, Національний технічний університет «Харківський політехнічний інститут», вулиця Кирпичова, 21, Харків, Україна, 61000.
E-mail: kuchuk56@mail.ru

Надійшло: Листопад 2016.

Автори:
Олена Толстолузька, доктор технічних наук, старший науковий співробітник, Харківський національний університет імені В. Н. Каразіна, майдан Свободи 4, м. Харків, 61022, Україна.
E-mail: tps@karazin.ua
Дмитро Толстолузький, інженер, Харківський національний університет імені В. Н. Каразіна, майдан Свободи 4, м. Харків, 61022, Україна.
E-mail: tps@karazin.ua
Ольга Мороз, старший викладач, Харківський національний університет імені В. Н. Каразіна, майдан Свободи 4, м. Харків, 61022, Україна.
E-mail: tps@karazin.ua

Метод компіляційно-семантичної верифікації часопараметризованих мультипаралельних програм.

Анотація. Наводиться визначення часопараметризованих мультипаралельних програм, які (на відміну від загальноприйнятого трактування паралельних програм) містять специфікації моментів початку виконання операцій/функцій, а також підмножини таких операцій/функцій. Обґрунтовується необхідність розробки нових методів верифікації часопараметризованих паралельних програм. Розкриваються етапи вирішення завдання компіляційного-семантичної верифікації часопараметризованих мультипаралельних програм в інтересах створення систем автоматичного синтезу високоефективних паралельних програм для обчислювальних систем різних класів. Наводиться приклад, що ілюструє основні етапи методу.

Ключові слова: часопараметризовані мільтипаралельні програми, верифікація паралельних програм, компіляційно-семантична верифікація, семантично-числові специфікації.

UDC 004.94:53.01.07

WILL THE ARTIFICIAL INTELLIGENCE HELP US?

Vladimir Kuklin

V.N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine
kuklinvm1@gmail.com

Reviewer: Alexandr Kuznetsov Dr., Full Professor, V.N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine
kuznetsov@karazin.ua

Received on November 2016

Abstract.. *Discussed what can help us (humanity) artificial intelligence. The unification of artificial neural networks and decision-making expert systems based on the logic has discussed. The integration of formed (human) concepts of the system of fuzzy logic and artificial neural networks, allowed us to understand what is happening in the problem-solving process of neural network. The human brain is MEGA processor, therefore, all the efforts of researchers should be focused on the development of MEGA processor systems of new generation. Noted that for the implement intelligent system similar to the human brain, it is necessary to ensure her connection with the outside world and the ability of self-study.*

Key words*: neural networks, expert systems, megaprocessor systems.*

1 Introduction

The theme of this publication: the ways of development the artificial intelligence and a bit on what we are to expect from it in the nearest future. Over time humans have more or less figured out how to deal with hard manual labor. As a result various mechanisms and appliances have appeared. Gradually, the mankind gets rid of not qualified manual work.

The price to pay for this is a sedentary lifestyle, the absence of physical workload, and as a result a lot of health issues. At present the world many is trying to abandon already and from intense intellectual work. Even not so much intellectual, but rather simply require to think about. Arguing that the need and of it get rid. Up till now, it has been possible to get rid of the monotonous (and therefore exhausting), moderately challenging intellectual work. Too, with help different of machines and gadgets. However at many, there is a temptation to go further and liberate humanity from was intellectual work at forever. Better even not to reflect than us have to sacrifice in this case. The fact is that many people wants to get rid of work completely and dedicate himself to only for various games and entertainment. Therefore, there is such a strong interest in artificial intelligence systems, which actually originated from the simple desire to play chess with the computer. Intrigued by this idea, talented humans created quite a lot of useful things, what nowadays is interpreted as a breakthrough on a path to progress. Though, perhaps, it's not a breakthrough but rather a breakaway. They had to figure out how the human being thinks, how one ought to think in principle, and how to teach the machine to do the same. Indeed, how does a human being get to know the world? The key element of the cerebral cortex is neurons, which are cells with a paradoxical reaction to irritation. Weak irritation makes them enraged whereas strong irritation leaves them almost indifferent. A strong response of one neuron subdues its neighbors. It is interesting that the human being is paradoxical in the same degree. Powerful noise doesn't prevent you from falling asleep whereas a word said in whisper while you sleep would make you wake up in fright. The extreme shock is capable to immobilize human and other beings, suppressing them desire and ability to escape.

The cerebral cortex is a multi-storey building, where the top looks like a service floor, i.e. a huge bunch of "cables" each of which connects neurons to one another. Any external disturbance that comes to the brain via these "cables", causes alarm among neurons. This alarm can be increased or reduced additionally by chemical components injected into the blood. Hence, emerge centers of ex-

citation recognized as reflections of various patterns. When these visual imagery partially coincide (on 60-70%) with imagery previously stored in the human memory, happens them identification.

The mechanism of motor reaction goes off. It is no wonder that people want to create some artificial mechanism acting in a similar way. The idea of the device with the ability to recognize the input images is the idea of perception, neurocomputer and neural network.

At first, the man had decided to take away the ability of higher powers to predict future, realizing that each sequence of operations and actions can be interpreted as a computation. And then he wanted to create something in his own likeness.

The fact is that the early successes of developers of the artificial intelligence have lowered the demands to its level. Nowadays, the main feature of the artificial intelligence is considered to be the ability to find solutions independently. Hence, *the natural intelligence has become an important, but only a special case of the intellect in general* [1]. Each set of interacting agents having the ability to perceive and respond to the environment changes may be regarded as the intelligent system. The result of this interaction can be considered as a solution of the problem (or even realization of some actions).

2 The appearance of the perceptron, neurocomputer and neural network

It all started with Frank Rosenblatt's Perceptron Mark I, which recognized with some errors the letters of the alphabet as early as 1960 [2].

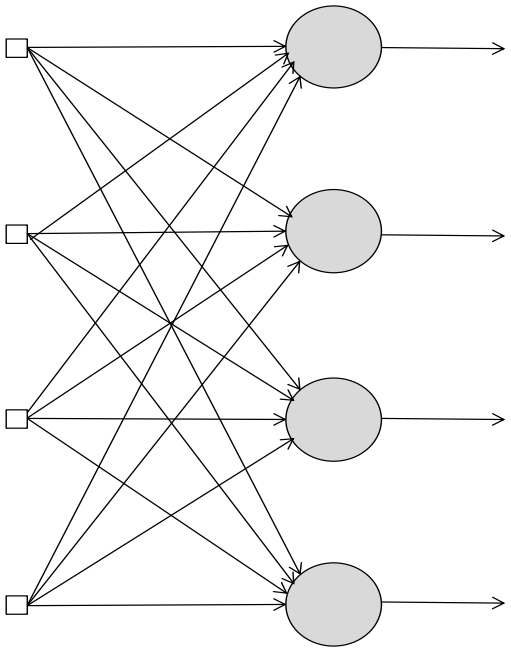


Fig. 1 - The simple perceptron

As usual for science, ten years later, M. Minsky, who had been a schoolmate of Rosenblatt, strongly criticized the capacities of the perceptron, which led to delay in the development of neurocomputers (Fig.2,3) in the following decades [3]. But later, people still decided to use the perceptron as intended, because nothing else had occurred to them. The neural network is the web composed of neurons, which we have to teach. It is necessary to give it several tasks with solutions and adjust network settings so that the network would be able to generate these solutions independently. Though there are networks that do not require training. But there is a lot of the problems with trained networks too. If the tests show that the number of errors decreases after training, then it means we got lucky, the training is successful. But when the number of errors grows, the network just remembers all that it was told during the training, and really learned nothing (the politically-correct term for it is overtraining). Then the network is considered as untutorable and it should be rejected.

There is another reason. When the network is rather powerful, it is capable of providing its own solution, which the experimentalists did not expect. The network violates their view of the world in this case. Such cleverness is discouraged and this excessively smart network is rejected too. The rational response to this situation is to increase the quantity and quality of training tasks to keep the stability of the world view. The hypothesis "*of necessary development* (*www.csd.univer.kharkov.ua/content/files/cat16/zarajenie_razumom.pdf*)": *Expanding intelligent system requires a greater amount of the basic knowledge to prevent the destruction of the current world view.* Human beings demonstrate the same thing: if the student after studies answers "perfectly well" looking honestly in your eyes, but can't solve any problem on this subject, then the experienced teacher knows that the person in front of him is "untutorable". But sometimes the examiners make a mistake and take such student for "untutorable" if he suggests other continuations and variants hidden from the examiner.

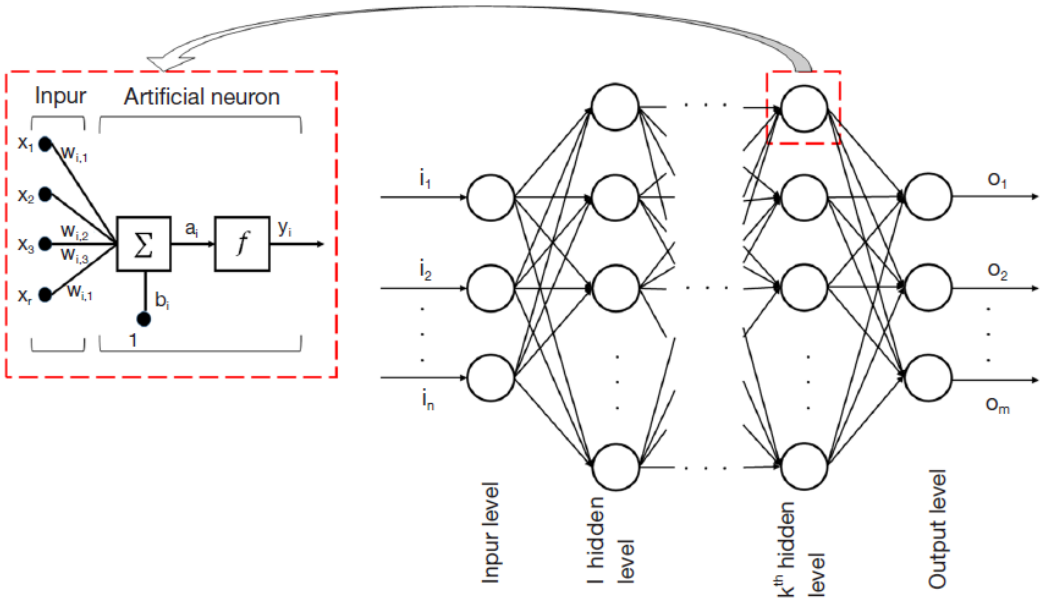


Fig. 2 - Artificial neuron and network

($x_1,\ldots,x_n$ - input signals coming from other neurons; $w_1,\ldots,w_n$ - synaptic weights of a neuron; b - the threshold value (threshold); y - the output of the neuron; $f = \varphi(v)$ - activation function)

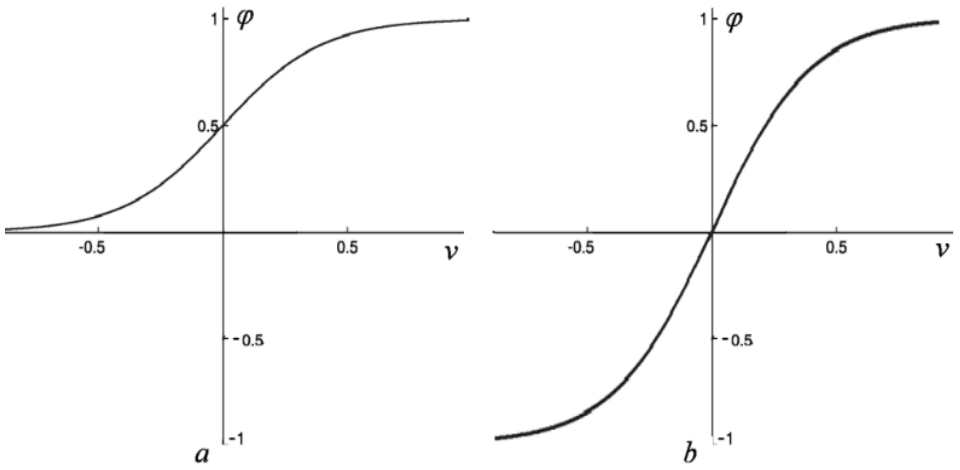


Fig. 3 - Type of the activates

The examiner may be a prisoner of stereotypes imposed by the training program or simply might not be adequately equipped to deal with discussion about new approaches to possible solutions of the problem. But now even the highly advanced neural networks cannot match the human intelligence in the slightest degree.

In addition, M. Minsky was right in some way because a neurocomputer with the insufficient number of neurons is not capable of solving a specific class of problems. This is a lower limit. But there is an upper limit. An excessively extended neural network can produce an undesirable result. Such a computer will revise results all the time, get snarled in its output and will turn out to be completely useless (the loss of the solution uniqueness).The same can be observed on human beings. The intellectual who is instructed to solve a simple problem, will torment the problem originators with doubts and arguments, so all simple tasks should be entrusted to "the men of action", not burdened with extensive education and intellectual abilities.

All investigators have been aware of the difficulty to understand how neural web produces knowledge. But as always, people need a result. Only enthusiastic mathematicians wish to go deeply into the needless details. Indeed, customers need a device that can do something useful, but they are too busy to investigate how it works. It is sufficient for them to obtain the user's manual. Enthusiastic people, who are not aware of what is happening in the "black box", i.e. neurocomputer, have made a lot of such devices by using the methods of selection and sorting. These neural networks learned that what their stubborn creators required, but for the rest neither the first nor the second were fit. Such is the apotheosis of empiricism.

3 Continuation of the history of the artificial intelligence development

So, Allen Newell [4,5], after a late fascination with chess, has created a program (1954), based on the methods of Claude Shannon [6], which produced a long line of followers. However, the mathematical logic, which had been developed long ago by efforts of Friedrich Ludwig Gottlob Frege and his contemporaries, was not put into practice, which is not surprising. Games seemed to be more interesting. However, a method suitable for solving the problem of creating chess programs without the use of a correct mathematical formalism, was proposed by a real inventor – Alan Turing [7]. The efforts by the RAND Corporation employees John Shaw and Herbert Simon, supported by de Groot and his colleagues, resulted in the development of the IPL language (1956), the predecessor of the LISP computer programming language, which was created by John McCarthy (1960). Incidentally, LISP (LIst Processing Language) [8,9] is based on a system of Lambda calculus formulated for the first time by Alonzo Church [10]. The use of formal mathematical logic to represent and execute computer programs, proposed by John Alan Robinson (1965) proved to be revolutionary [11].

1. Two expressions $P(a)$ and $P(c) \Rightarrow G(d)$;
2. $P(c) \Rightarrow G(d)$ can be replaced $\sim P(c) \vee G(d)$;
3. When the unification of the variables $a \to c$;
4. Using the procedure of resolution $P(a \to c) \vee \sim P(c) \vee G(d)$;
5. Remove tautology $P(c) \vee \sim P(c)$;
6. Get the third expression $G(d)$.

Based on this method, albeit with a number of differences, Alain Kolmerauer [12] developed in 1971 the PROLOG language which uses the first-order predicate logic. By the way, the Resolution method was applied earlier by young mathematician Jacques Herbrand. Thus, in the early 70s, the programming languages for artificial intelligence had been created (LISP, Prologue, PLANNER, REDUCE etc., oriented towards various tasks).

Using the artificial intelligence programming languages based on mathematics, one can review the entire solution of the task and correct it, if necessary. Such systems are referred to as expert systems. Of course, this is different from the neural systems, where one can not understand what's going on inside the "black box" – , i.e. neurocomputer. At that time, the logic programming expert

systems possessed one essential shortage – they operated with only two logic values "true" and "false". Sadly, that's just not the way life works.

4 Historical association of artificial neural networks and expert systems of decision making on the basis of logic

It was necessary to go to the structure of fuzzy concepts. Therefore, the theory of fuzzy sets and fuzzy logic, proposed to the mankind by Lotfi Zadeh in 1965 [13], turns out to be very useful. For example

$$A \cap B \quad \mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x));$$
$$A \cup B \quad \mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x));$$

where $\mu(x)$ - show belonging to this type.

In 1993, Bart Kosko [14] proved the important theorem that any mathematical system can be approximated by fuzzy logic. The integration of fuzzy logic concepts, formulated by mathematicians, with artificial neural networks was initially performed by J.-S. Roger Jang from the Taiwan University (the neuro-fuzzy system refers to combinations of artificial neural networks and fuzzy logic) [15]. This allowed us to understand what is happening within the neural network while solving the task. Thus, a *historic unification of artificial neural networks and decision-making expert systems based on the logic has happened.*

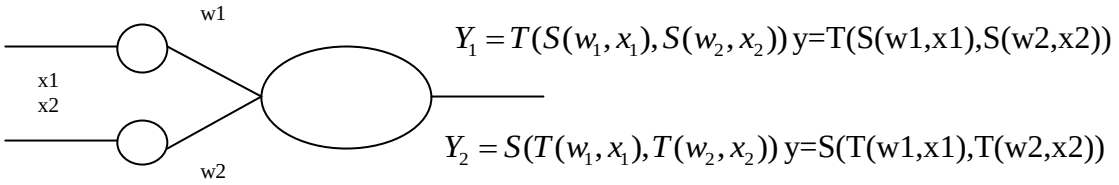


Fig.4 - Fuzzy neurons, where $T(w_1, x_1) = \min(w_1, x_1)$, $S(w_1, x_1) = \max(w_1, x_1)$

5 The problem of learning and prospects

The human beings fill their databases and knowledge bases during all their his life. However, when we want to create an artificial intelligent system and going to fill data and knowledge bases, that's when the problems arise. At first, a human being possesses a set of knowledge that he believes to be known by default. All this knowledge should be spoon-fed and explained to the artificial intelligent system. Secondly, databases should be filled by experts and their job should be highly paid for. Third, the time necessary for database filling, checking and rechecking is rather considerable. All this is somewhat disappointing.

But, already today neural networks allow us to establish the reasons for the poor performance of car engines, a neurocomputers allow to identify the similarity of different objects and processes. Expert systems provide pilots warning function and give its recommendations to action. These systems assist aircraft pilots, train drivers, crews of ships and car drivers, but the final decision is usually left up to the person. Thus in the short term, we can look forward to the development of artificial intelligence, efficiency and work speed of which will be comparable with the human brain.

6 Conclusions

The rate of human decision-making often far exceeds processing speed data of modern artificial neural networks and expert systems (especially for mode multi-task). Here the matter is likely that the human brain is MEGA processor. Simultaneous processing by this huge number of processes accelerates the decision-making.

Therefore, it's obvious that, all the efforts of researchers should be focused on the development of MEGA processor systems of new generation. A Multi-processor system should possess an individual memory for each processor and a shared memory, as an analogue of libraries in the human society.

The Supercomputer Tianhe-2 located in National Supercomputer Center in Guangzhou demonstrates performance of 33.86 petaflops (almost 34 quadrillion (thousand trillion) floating point operations per second). In Britain, the Cray XC40TM, one of the fastest supercomputers in the world, will be launched soon. It has 480,000 cores, 2 million gigabytes of memory and can store up to 17 million gigabytes of data. At its peak, it is able to make 16,000 trillion calculations per second. The best achievement of IBM is the neuro-synaptic processor True North – 16 million digital neurons and 4 billion synapses. But it is still not enough ...

To create the artificial intelligent system, it is necessary to provide its connection to the outside world and the ability to self-learn. For example, as the IBM company is developing the interactive artificial intelligence system based on the Watson supercomputer that has access to the technologies of the cloud system Watson Developer Cloud. Besides, developers, while examining the intelligent systems, have realized the benefits of the network structure, each element of which has:

- independent access to external data and general information;
- internal memory, in addition to libraries;
- sufficient autonomy.

Besides, tend to use the their synergy.

References

[1] McCarthy J. A proposal for the Dartmouth summer research project on artificial intelligence / J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon. – Dartmouth , 1955.

[2] Rosenblatt F. Perceptron simulation experiments / F. Rosenblatt // Proceedings of the IRE. – 1960, March. – Vol.18. – №3. – P.301 – 309. – [also Project Para Technical Report VG-1196-G-3, CAL June 1959].

[3] Minsky M. Perceptrons. An Introduction to Computational Geometry / M. Minsky, S. Papert. – Cambridge, Mass.: M.I.T. Press, 1969. – 258 p.

[4] Newell A. The chess machine: an example of dealing with a complex task by adaptation / A. Newell // ACM. Proceedings of the Western Joint Computer Conference (1955, March 1-3) . – 1955. – P.101 – 108.

[5] Newell A. GPS, program that simulates human thought / A. Newell, H.A. Simon // Defense Technical Information Center. – 1961. – №10. – P. 109 – 124.

[6] Shannon C.E. A Mathematical Theory of Communication / C.E. Shannon // The Bell System Technical Journal. – 1948. – V.27. – P.379 – 423; 623 – 656.

[7] Turing A. M. On computable numbers, with an application to the Entscheidungs problem / A. M. Turing // Proc. of the London Math. Soc. Ser. 2. – [1936-1937] . – Vol. 42. – P. 230 – 265.

[8] Newell A. Programming the Logic Theory Machine / A. Newell and F.C. Shaw // Proceedings of the Western Joint Computer Conference (1957, Feb.) . – 1957. – P. 230 – 240.

[9] Lisp 1.5 Programmer's Manual / J. McCarthy, P. Abrahams, D. Edwards, et al. – Cambrige, Massachusetts : MIT Press, 1962.

[10] Church A. Introdaction to mathematical logic. Vol.1. / A. Church. – Princeton: Princeton University Press, 1956. – 485 p.

[11] Robinson J. A. A Machine-Oriented Logic Based on the Resolution Principle / John Alan Robinson // Communications of the ACM. – 1965. – № 5. – P. 23 – 41.

[12] Colmerauer A. Un système de communication en français / Alain Colmerauer, Henry Kanoui, Robert Pasero et Philippe Roussel // Rapport préliminaire de fin de contrat IRIA, Groupe Intelligence Artificielle, Faculté des Sciences de Luminy, Université Aix-Marseille II, France, October 1972.

[13] Zadeh L. A. Fuzzy sets / Lotfi A. Zadeh // Information and Control. – 1965. – Vol.8. – P. 338 – 353; Fuzzy sets and systems // System Theory [Fox J., editor] . – Brooklyn, NY: Polytechnic Press, 1965. – P. 29 – 39.

[14] Kosko B. Fuzzy systems as universal approximation / B. Kosko // IEEE Transactions on Computers. – 1994 . – Vol. 43 . – № 11. – P. 1329 – 1333.

[15] Jang J.S.R. ANFIS: adaptive-network-based fuzzy inference system / J.S.R. Jang // IEEE transactions on systems, man, and cybernetics. – 1993. – Vol. 23. – № 3. – P. 665 – 685.

Рецензент: Александр Кузнецов, д.т.н., проф., Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина.
E-mail: kuznetsov@karazin.ua

Поступила: Ноябрь 2016.

Автор:
Владимир Куклин, д.ф.-м.н., проф., зав. кафедры, Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина.
E-mail: kuklinvm1@gmail.com

Будет ли искусственный интеллект нам помогать?

Аннотация. Обсуждается чем может нам (человечеству) помочь искусственный интеллект. Рассмотрено важное объединение систем на основе нейронных сетей и экспертных систем на базе математической логики. Объединение сформированных (человеком) понятий системы нечеткой логики с искусственными нейронными сетями позволило понять, что происходит в процессе решения задачи нейронной сетью. Так как человеческий интеллект это мегапроцессорная система, то подчеркнуто, что основные усилия следует направить на создание мегапроцессорных систем новых поколений. Отмечено, что для реализации интеллектуальной системы, аналогичной мозгу человека, необходимо обеспечить ее связь с внешним миром и возможность самообучения.

Ключевые слова: нейронные сети, экспертные системы, мегапроцессорные системы.

Рецензент: Олександр Кузнецов, д.т.н., проф., Харківський національний університет імені В. Н. Каразіна, Харків, Україна.
E-mail: kuznetsov@karazin.ua

Надійшло: Листопад 2016.

Автор:
Володимир Куклін, д.ф.-м.н., проф., завідувач кафедри, Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: kuklinvm1@gmail.com

Буде штучний інтелект нам допомагати?

Анотація. Обговорюється чім може нам (людству) допомогти штучний інтелект. Розглянуто важливе об'єднання систем на основі нейронних мереж і експертних систем на базі математичної логіки. Об'єднання сформованих (людиною) понять системи нечіткої логіки з штучними нейронними мережами дозволило зрозуміти, що відбувається в процесі рішення задачі нейронною мережею. Так як людський інтелект це мегапроцесорна система, то підкреслено, що основні зусилля слід спрямувати на створення мегапроцесорних систем нових поколінь. Відзначено, що для реалізації інтелектуальної системи, що аналогічна мозку людини, необхідно забезпечити її зв'язок із зовнішнім світом і можливість самонавчання.

Ключові слова: нейронні мережі, експертні системи, мегапроцесорні системи.

УДК 004.056.55

АЛГЕБРАИЧЕСКИЙ ИММУНИТЕТ НЕЛИНЕЙНЫХ УЗЛОВ СИММЕТРИЧНЫХ ШИФРОВ

А. Кузнецов[1], Ю. Горбенко[2], И. Белозерцев[3], А. Андрушкевич[4], А. Нарежний[5]

[1,3,4,5] Харьковский национальный университет имени В.Н. Каразина, пл. Свободы, 4, г. Харьков, 61022, Украина
kuznetsov@karazin.ua, ivanbelozersevv.jw@gmail.com, hitori26@mail.ru, o.nariezhnii@karazin.ua

[2] АО «Институт информационных технологий», ул. Бакулина, 12, г. Харьков, 61166, Украина
gorbenkou@iit.kharkov.ua

Рецензент: Антон Алексейчук, д.т.н., доцент, Институт специальной связи и защиты информации национального технического университета Украины «КПИ», пр. Победы, 37, г. Киев, 03056, Украина.
alex-dtn@ukr.net

Поступила в декабре 2016

Аннотация. *Исследуются методы вычисления алгебраической иммунности криптографических булевых функций и нелинейных узлов замен (подстановок) симметричных шифров. Приводятся результаты сравнительного анализа алгебраической иммунности нелинейных узлов симметричных шифров.*

Ключевые слова*: симметричные шифры, алгебраический иммунитет, нелинейные узлы замены.*

1 Введение

Криптографическое преобразование играет важную роль в обеспечении безопасности современных информационных систем и технологий [1, 2]. Симметричные шифры, в силу свой простоты, эффективности и многофункциональности, применяются практически во всех современных криптопротоколах, а также как составная часть других криптографических примитивов: в хешировании, формировании псевдослучайных последовательностей, генерации паролей и пр. Следовательно, анализ и исследование методов синтеза симметричных криптопримитивов, разработка и теоретическое обоснование критериев и показателей эффективности, в том числе отдельных узлов современных шифров, является важной и актуальной научно-технической задачей.

Ключевым компонентом современных симметричных шифров являются нелинейные узлы (нелинейные подстановки, таблицы замен, S-блоки), которые выполняют функции скрытия статистических связей открытого текста и шифртекста, перемешивания и рассеивания данных, внесения нелинейности в процедуру зашифрования для противостояния различным криптоаналитическим и статистическим атакам. Таким образом, от показателей эффективности нелинейных узлов (сбалансированности, нелинейности, автокорреляции, корреляционной иммунности и пр.) непосредственно зависят эффективность симметричного шифра, его устойчивость к большинству известных криптографических атак и уровень обеспечиваемой им безопасности информационных технологий.

Отдельные показатели эффективности нелинейных узлов симметричных шифров рассмотрены в [3-9]. Понятие алгебраического иммунитета впервые введено в работах [10, 11] для оценки стойкости булевых функций к т.н. алгебраическому криптоанализу, предложенному в работе [12]. В работе [13] эти положения были обобщены для булевых отображений (S-блоков), для вычисления алгебраического иммунитета используется математический аппарат базисов Грёбнера.

В данной работе рассматриваются различные методы расчета алгебраического иммунитета, изучается их взаимосвязь и приводятся результаты сравнительных исследований алгебраической иммунности нелинейных узлов наиболее известных современных симметричных шифров.

2 Алгебраический иммунитет булевых функций

Понятие алгебраического иммунитета было впервые введено в работах [10,11] и подробно рассмотрено в диссертационной работе [14]. Для дальнейшего изложения материала введём необходимые определения и обозначения, придерживаясь ранее принятых в [14] формулировок.

Пусть $GF(2)$ – двоичное поле и $GF(2)^n$ – n-мерное векторное пространство над $GF(2)$.

Булева функция $f(x)$ от n переменных – это отображение $f(x): GF(2)^n \to GF(2)$, где $x = (x_1, ..., x_n)$.

Таблица истинности булевой функции $f(x)$ от n переменных – это двоичный выходной вектор значений функции, который содержит 2^n элементов, каждый элемент принадлежит множеству {0, 1}.

Алгебраическая нормальная форма (*полином Жегалкина*) булевой функции $f(x)$ от n переменных записывается в виде:

$$f(x) = a_0 \oplus a_1 x_1 \oplus a_2 x_2 \oplus ... \oplus a_n x_n \oplus a_{12} x_1 x_2 \oplus a_{13} x_1 x_3 \oplus ... \oplus a_{(n-1)n} x_{n-1} x_n \oplus ... \oplus a_{123...n} x_1 x_2 x_3 ... x_n,$$

где коэффициенты $a_i \in \{0,1\}$ и каждая булева функция реализуется полиномом Жегалкина единственным образом, т.е. каждое представление $f(x)$ соответствует уникальной таблице истинности.

Алгебраическая степень $Deg(f)$ булевой функции $f(x)$ – число переменных в самом длинном слагаемом алгебраической нормальной формы функции, имеющем ненулевой коэффициент a_i. При этом считаем $Deg(0) = 0$.

Обозначим через V_n множество всех отображений $GF(2)^n \to GF(2)$, т.е. это множество всех возможных булевых функций $f(x)$ от n переменных.

Множество V_n в будем рассматривать и как кольцо булевых функций и как векторное (линейное) пространство над двоичным полем, т.е. $V_n = GF(2)^{2^n}$.

Булева функция $g \in V_n$ называется *аннигилятором* функции $f \in V_n$, если

$$f \cdot g = 0$$

или

$$(f + 1) \cdot g = 0.$$

Множество различных аннигиляторов булевой функции $g(x)$ образует линейное пространство, которое обозначим как

$$Ann(f) = \{g \in V_n \mid f \cdot g = 0\}.$$

Линейное пространство аннигиляторов степени $\le d$ обозначим

$$A_d^n(f) = \{g \in V_n \mid f \cdot g = 0, Deg(g) \le d\} \subset Ann(f).$$

Понятие аннигиляторов булевых функций тесно связано с оценкой эффективности алгебраического криптоанализа поточных шифров [10]. В частности, при использовании фильтрующего генератора (см. Рис. 1) псевдослучайных последовательностей (ПСП) поиск начального состояния регистра сдвига с линейной обратной связью (РСЛОС) сопряжен с понижением степени совместной системы полиномиальных булевых уравнений.

Алгоритм алгебраического криптоанализа, предложенный в [10], позволяет, при определенных условиях, по части перехваченной выходной последовательности (ПСП) находить начальное состояние РСЛОС с временной сложностью $O\left((S_n^d)^3\right)$, где

$$S_n^d = \sum_{i=0}^{d} \frac{n!}{i!(n-i)!}$$

и d - наименьшая степень ненулевого аннигилятора фильтрующей булевой функции $f(x)$ или ее инверсии $f(x)+1$.

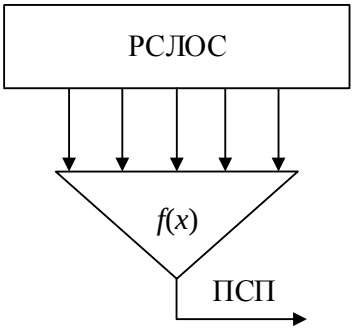


Рис. 1 – Структурная схема фильтр-генератора ПСП

Таким образом, задачей алгебраического криптоанализа является поиск ненулевых аннигиляторов или, по крайней мере, оценка их минимальной степени. С этой целью в работе [11] введено определение *алгебраической иммунности* $AI(f)$ булевой функции $f \in V_n$:

$$AI(f) = \min\{Deg(g) \mid g \in Ann(f) \text{ или } g \in Ann(f+1)\}.$$

Величина $AI(f)$ численно равна минимальной степени такой булевой функции $g \in V_n$, что $f \cdot g = 0$ или $(f+1) \cdot g = 0$.

Используя введенное выше понятие линейного пространства аннигиляторов степени $\le d$ запишем:

$$AI(f) = \min\{d \mid A_d^n(f) \neq 0 \text{ или } A_d^n(f+1) \neq 0\}, \quad (1)$$

т.е. для оценки алгебраической иммунности булевой функции $f \in V_n$ достаточно найти ненулевой базис пространства аннигиляторов наименьшей степени d.

Величина d позволяет количественно оценить сложность алгебраического криптоанализа и, при достаточно большом d, гарантировать устойчивость поточного криптоалгоритма к алгебраической атаке.

Алгоритм вычисления алгебраической иммунности булевых функций. Один из возможных алгоритмов расчета алгебраической иммунности булевых функций представлен в диссертационной работе [14]. Он основан на построении базиса линейного пространства аннигиляторов $A_d^n(f)$ заданной степени d. Итеративно увеличивая d и повторяя построение базиса пространства $A_d^n(f)$, оценку $AI(f)$ получим используя формулу (1), т.е. через ненулевой базис аннигиляторов наименьшей степени.

Для изложения сути алгоритма необходимо ввести следующие дополнительные обозначения.

Моном (одночлен) относительно переменных $x_1,...,x_n$ запишем в виде

$$x^u = \prod_{i=1}^{n} x_i^{u_i} = \begin{cases} x_i, u_i = 1, \\ 1, u_i = 0, \end{cases}$$

где вектора $x, u \in V_2^n$, $x = (x_1,...,x_n)$, $u = (u_1,...,u_n)$.

Степень одночлена x^u определяется весом Хемминга (числом ненулевых координат) $w_h(u)$ вектора $u = (u_1,...,u_n)$, т.е.

$$Deg(x^u) = w_h(u).$$

С учетом этих обозначений булеву функцию $f(x)$ в алгебраической нормальной форме (в форме полинома Жегалкина) запишем в виде

$$f(x)=\sum_{u\in GF(2)^n} a_u x^u,\ a_u\in GF(2). \quad (2)$$

Функцию (*аннигилятор*) $g\in A_d^n(f)$ также представим в виде полинома Жегалкина

$$g(x)=\sum_{v\in GF(2)^n:w_h(v)\le d} b_v x^v, \quad (3)$$

где $b_v\in GF(2)$ – неизвестные коэффициенты аннигилятора, $w_h(v)$ – вес Хемминга вектора $v=(v_1,...,v_n)$.

Функция g принадлежит пространству $A_d^n(f)$ только в том случае, если для любого $x\in GF(2)^n$ выполняется равенство $f(x)\cdot g(x)=0$.

Подставив (2) и (3) получим:

$$f(x)\cdot g(x)=\left(\sum_{u\in GF(2)^n} a_u x^u\right)\left(\sum_{v\in GF(2)^n:w_h(v)\le d} b_v x^v\right)=\sum_{u\in GF(2)^n}\left(\sum_{v\in GF(2)^n:w_h(v)\le d} a_u b_v x^{u\vee v}\right)=0,$$

где $u\vee v=(u_1\vee v_1,...,u_n\vee v_n)$, $\vee$ – дизъюнкция (логическая операция ИЛИ).

После группировки слагаемых по общему множителю, получим равенство:

$$\sum_{w\in GF(2)^n}\left(\sum_{a_u,b_v:a_u\vee b_v=w} a_u b_v\right)x^w=0, \quad (4)$$

которое выполняется для любого $w\in GF(2)^n$. Следовательно, имеем систему линейных однородных уравнений

$$\begin{cases}\sum\limits_{a_u,b_v:a_u\vee b_v=w} a_u b_v=0,\ \forall w\in GF(2)^n\end{cases} \quad (5)$$

относительно неизвестных коэффициентов b_v аннигилятора $g(x)$.

Решение данной системы уравнений (например, методом Гаусса) задает базис пространства $A_d^n(f)$.

Пример. Для $n=2$ и $d=1$ имеем:

$$f(x)=a_{00}+a_{10}x_1+a_{01}x_2+a_{11}x_1x_2,$$
$$g(x)=b_{00}+b_{10}x_1+b_{01}x_2.$$

После подстановки в $f(x)\cdot g(x)=0$ получим

$$f(x)\cdot g(x)=a_{00}b_{00}+(a_{00}b_{10}+a_{10}b_{10}+a_{10}b_{00})x_1+(a_{00}b_{01}+a_{01}b_{01}+a_{01}b_{00})x_2+$$
$$+(a_{10}b_{01}+a_{01}b_{10}+a_{11}b_{00}+a_{11}b_{10}+a_{11}b_{01})x_1x_2=0,$$

откуда имеем систему линейных однородных уравнений:

$$\begin{cases}a_{00}b_{00}=0,\\ a_{00}b_{10}+a_{10}b_{10}+a_{10}b_{00}=0,\\ a_{00}b_{01}+a_{01}b_{01}+a_{01}b_{00}=0,\\ a_{10}b_{01}+a_{01}b_{10}+a_{11}b_{00}+a_{11}b_{10}+a_{11}b_{01}=0\end{cases}$$

относительно неизвестных b_{00},b_{10},b_{01} – коэффициентов функции $g(x)$.

Тогда, например, для функции $f(x)=x_1+x_2$ (т.е. при $a_{00}=a_{11}=0$ и $a_{10}=a_{01}=1$) получим систему:

$$\begin{cases} b_{10} + b_{00} = 0, \\ b_{01} + b_{00} = 0, \\ b_{01} + b_{10} = 0, \end{cases}$$

которой удовлетворяет только два решения:

$$b_{00} = b_{10} = b_{01} = 0, \text{ т.е. } g(x) = 0,$$
$$b_{00} = b_{10} = b_{01} = 1, \text{ т.е. } g(x) = 1 + x_1 + x_2.$$

Непосредственная проверка показывает, что $g(x) = 1 + x_1 + x_2$ действительно является аннигилятором функции $f(x) = x_1 + x_2$:

$$f(x) \cdot g(x) = (x_1 + x_2)(1 + x_1 + x_2) = x_1 + x_2 + x_1 + x_1x_2 + x_1x_2 + x_2 = 0.$$

Обобщая вышеизложенное, определим основные шаги **алгоритма поиска базиса пространства аннигиляторов** [14].

Вход: $n \in \mathrm{N}, d \in \{1, ..., n\}$, функция $f(x)$ (заданная списком одночленов x^u с ненулевыми коэффициентами a_u в (2)).

Выход: Линейное пространство $A_d^n(f)$, заданное в виде параметрического семейства многочленов Жегалкина от n булевых переменных степени $\le d$.

Шаг 1. Представляем функции $f(x)$ и $g(x)$ в виде сумм (2) и (3), соответственно.

Шаг 2. Раскрываем скобки в произведении $f(x) \cdot g(x)$ и, группируя слагаемые $a_u b_v x^w$ путем сортировки по $a_u \vee b_v = w$, получаем уравнение (4).

Шаг 3. Составляем систему линейных однородных уравнений (5).

Шаг 4. Находим общее решение системы (5) в параметрическом виде и подаем на выход алгоритма.

В работе [14] приводится оценка $O\left(m \cdot \left(S_n^d\right)^3\right)$ битовой сложности рассмотренного алгоритма, где m – количество ненулевых коэффициентов a_u в (2).

Используя приведенный алгоритм поиска базиса пространства аннигиляторов можно вычислить алгебраическую иммунность булевой функции $f(x)$ последовательно перебирая все значения $d > 0$ до тех пор, пока не получим ненулевое пространство аннигиляторов $A_d^n(f)$ или $A_d^n(f+1)$.

Минимальное значение $d > 0$, для которого $A_d^n(f) \neq 0$ и/или $A_d^n(f+1) \neq 0$ соответствует значению алгебраической иммунности булевой функции $f(x)$.

Алгоритм вычисления алгебраической иммунности $AI(f)$.

Вход: $n \in \mathrm{N}$, функция $f(x)$ (заданная списком одночленов x^u с ненулевыми коэффициентами a_u в (2)).

Выход: Значение алгебраической иммунности $AI(f)$.

Шаг 1. Присваиваем $d = 1$.

Шаг 2. Вычисляем пространство аннигиляторов $A_d^n(f)$ и $A_d^n(f+1)$.

Шаг 3. Если $A_d^n(f) = 0$ и $A_d^n(f+1) = 0$, то присваиваем $d = d + 1$ и переходим к Шагу 2.

Шаг 4. Если $A_d^n(f) \neq 0$ и/или $A_d^n(f+1) \neq 0$, то присваиваем $AI(f) = d$ и подаем на выход алгоритма.

3 Алгебраический иммунитет булевых отображений (S-блоков)

Понятие алгебраической иммунности булевых функций в [13] обобщено на случай булевых отображений $F: GF(2)^n \to GF(2)^n$ (*векторных булевых функций*), которые реализуются узлами замен (*таблицами подстановок, S-блоками*) блочных симметричных шифров.

Для определения алгебраической иммунности $AI(F)$ воспользуемся терминами и определениями из работы [15].

Зафиксируем натуральные числа n, m и некоторое поле K. Рассмотрим конечную систему S из m алгебраических уравнений

$$\begin{cases} P_1(x_1, x_2, ..., x_n) = 0, \\ P_2(x_1, x_2, ..., x_n) = 0, \\ ... \\ P_m(x_1, x_2, ..., x_n) = 0 \end{cases} \quad (6)$$

от переменных $x_1, x_2, ..., x_n$ с коэффициентами над полем K.

Пусть $K[x_1, x_2, ..., x_n]$ – множество всех многочленов от переменных $x_1, x_2, ..., x_n$ с коэффициентами над полем K. На этом множестве определены операции сложения и умножения, а само множество называют *кольцом многочленов*. Это кольцо коммутативно (т.е. для любых элементов $a, b \in K[x_1, x_2, ..., x_n]$ выполняется равенство $a \cdot b = b \cdot a$), с единицей (для всех $a \in K[x_1, x_2, ..., x_n]$ выполняется равенство $a \cdot e = a$, где $e = 1$).

Непустое подмножество I коммутативного кольца с единицей R называется <u>*идеалом*</u> в R (обозначается как $I \lhd R$), если выполняются следующие два условия:

– для любых элементов $a, b \in I$ элемент $a - b \in I$;

– для любых $a \in I$ и $c \in R$ элемент $a \cdot c \in R$.

Элементы $a_1, a_2, ..., a_k$ составляют <u>*базис идеала*</u>

$$I = (a_1, a_2, ..., a_k) = \{a_1 \cdot r_1 + a_2 \cdot r_2 + ... + a_k \cdot r_k; r_1, r_2, ..., r_k \in R\} \subseteq R.$$

Принято считать, что идеал $I \lhd R$ <u>*допускает конечный базис*</u>, если в нем найдутся такие элементы $a_1, a_2, ..., a_k$, что $I = (a_1, a_2, ..., a_k)$.

Фундаментальная ***теорема Гилберта о базисе*** утверждает, что каждый идеал $I \lhd K[x_1, x_2, ..., x_n]$ допускает конечный базис, т.е. найдутся такие $f_1(x_1, x_2, ..., x_n)$, $f_2(x_1, x_2, ..., x_n)$, …, $f_k(x_1, x_2, ..., x_n) \in I$, что

$$I = (f_1, f_2, ..., f_k) = \{f_1 \cdot r_1 + f_2 \cdot r_2 + ... + f_k \cdot r_k; r_1, r_2, ..., r_k \in K[x_1, x_2, ..., x_n]\}.$$

С системой S (6) свяжем идеал I, порожденный многочленами $P_1(x_1, x_2, ..., x_n)$, $P_2(x_1, x_2, ..., x_n)$, …, $P_m(x_1, x_2, ..., x_n)$, отвечающим уравнениям системы:

$$I(S) = (P_1, P_2, ..., P_m) = \{P_1 \cdot r_1 + P_2 \cdot r_2 + ... + P_m \cdot r_m; r_1, r_2, ..., r_m \in K[x_1, x_2, ..., x_n]\}.$$

Если $F \in I(S)$, то тогда для каждого решения $(X_1, X_2, ..., X_n)$ системы (6) будет выполняться равенство

$$\begin{aligned} &F(X_1, X_2, ..., X_n) = \\ &= P_1(X_1, X_2, ..., X_n) \cdot r_1(X_1, X_2, ..., X_n) + P_2(X_1, X_2, ..., X_n) \cdot r_2(X_1, X_2, ..., X_n) + ... + \\ &+ P_m(X_1, X_2, ..., X_n) \cdot r_m(X_1, X_2, ..., X_n) = \\ &= 0 \cdot r_1(X_1, X_2, ..., X_n) + 0 \cdot r_2(X_1, X_2, ..., X_n) + ... + 0 \cdot r_m(X_1, X_2, ..., X_n) = 0. \end{aligned}$$

Если $\{P_1, P_2, ..., P_m\}$ и $\{\overline{P}_1, \overline{P}_2, ..., \overline{P}_k\}$ – два базиса одного идеала I, тогда системы алгебраических уравнений

$$\begin{cases} P_1(x_1, x_2, ..., x_n) = 0, \\ P_2(x_1, x_2, ..., x_n) = 0, \\ ... \\ P_m(x_1, x_2, ..., x_n) = 0, \end{cases} \qquad \begin{cases} \overline{P}_1(x_1, x_2, ..., x_n) = 0, \\ \overline{P}_2(x_1, x_2, ..., x_n) = 0, \\ ... \\ \overline{P}_k(x_1, x_2, ..., x_n) = 0 \end{cases}$$

эквивалентны, т.е. множества их решений совпадают.

Следовательно, множество решений системы алгебраических уравнений однозначно определяется идеалом системы, а различные базисы одного идеала отвечают эквивалентным системам [15].

Предположим, что имеется некоторый многочлен $h(x_1, x_2, ..., x_n) \in K[x_1, x_2, ..., x_n]$ и требуется за конечное число шагов выяснить, принадлежит ли он идеалу $I \lhd K[x_1, x_2, ..., x_n]$, заданному своим базисом $I = (f_1, f_2, ..., f_m)$. Другими словами, нужно решить т.н. *задачу вхождения*: – выяснить, существуют ли такие многочлены $r_1(x_1, x_2, ..., x_n)$, $r_2(x_1, x_2, ..., x_n)$, …, $r_m(x_1, x_2, ..., x_n)$, что $h = f_1 \cdot r_1 + f_2 \cdot r_2 + ... + f_m \cdot r_m$ и $h \in I = (f_1, f_2, ..., f_m)$.

Задачу вхождения решают посредством упрощения выражения для $h(x_1, x_2, ..., x_n)$ используя т.н. *редукцию многочлена*.

Запишем многочлен $h(x_1, x_2, ..., x_n)$ в виде суммы: $h = h_C + h_M$, где h_C – старший одночлен (моном), а h_M - сумма оставшихся одночленов в h. Предположим также, что h_C делится на старший член f_{iC} одного из многочленов f_i, т.е. $h_C = f_{iC} \cdot Q$ и $h = f_{iC} \cdot Q + h_M$ для некоторого одночлена Q. Тогда *операция редукции* задается выражением

$$h_1 = h - f_i \cdot Q = f_{iC} \cdot Q + h_M - f_{iC} \cdot Q - f_{iM} \cdot Q = h_M + (-f_{iM}) \cdot Q, \qquad (7)$$

где f_{iM} - сумма оставшихся одночленов в $f_i = f_{iC} + f_{iM}$. При этом старший член многочлена h_1 меньше старшего члена многочлена h. Если многочлен h принадлежит идеалу $I = (f_1, f_2, ..., f_m)$, тогда и редуцированный многочлен h_1 также будет принадлежать этому идеалу. Действительно, если $h \in (f_1, f_2, ..., f_m)$, тогда $h - h_1 = f_i Q \in (f_1, f_2, ..., f_m)$. Следовательно, задачу вхождения теперь можно решать уже не для многочлена h, а для редуцированного многочлена h_1. Если за конечное число редукций (7) многочлен h сведется (редуцируется) к нулю (ноль принадлежит любому идеалу), тогда $h \in (f_1, f_2, ..., f_m)$.

Базис $f_1, f_2, ..., f_m$ идеала $I = (f_1, f_2, ..., f_m)$ называется ***базисом Грёбнера*** этого идеала, если всякий многочлен $h \in I$ редуцируется к нулю при помощи $f_1, f_2, ..., f_m$. Иначе: набор многочленов $f_1, f_2, ..., f_m$ является базисом Грёбнера в идеале $I = (f_1, f_2, ..., f_m)$, если для любого $h \in I$ одночлен h_C делится на один из одночленов $f_{1C}, f_{2C}, ..., f_{mC}$ [15].

Для операции редукции многочленов используется понятие старшего одночлена (монома). Другими словами, предполагается, что на множестве всех одночленов кольца $K[x_1, x_2, ..., x_n]$ задан *линейный порядок* (мономиальное упорядочение $\prec$), удовлетворяющий следующим свойствам [16]:

- из $x^u \prec x^v$ следует, что $x^w \cdot x^u \prec x^w \cdot x^v$ для любых одночленов x^u, x^v, x^w (одночлены определены как в (2), т.е. $x, u, v, w \in V_2^n, x = (x_1, ..., x_n), u = (u_1, ..., u_n), v = (v_1, ..., v_n), w = (w_1, ..., w_n)$);

- $1 \preceq x^v$ для любого одночлена x^u.

В качестве примеров мономиального упорядочения приведем:

- *словарный (лексикографический) порядок (lex)*: $x^u \prec_{\text{lex}} x^v$, если существует такое i, что $u_i < v_i$ и $u_j = v_j$ для $j < i$ (сперва упорядочиваем переменные в одночленах в требуемом алфавитном порядке, а потом смотрим до первого различия в одночленах);

- *степенно-словарный порядок (deglex)*: $x^u \prec_{\text{deglex}} x^v$, если $w_h(u) < w_h(v)$ или $w_h(u) = w_h(v)$, но при этом $x^u \prec x^v$ в словарном порядке (упорядочиваем по сумме степеней, в случае равенства сумм сравниваем по словарному порядку);

- *степенной обратный словарный порядок (degrevlex)*: $x^u \prec_{\text{degrevlex}} x^v$, если $w_h(u) < w_h(v)$ или $w_h(u) = w_h(v)$, но при этом $x^u \succ_{\text{lex}} x^v$ в словарном порядке (упорядочиваем по сумме степеней, в случае равенства сумм сравниваем по обратному словарному порядку).

Решение задачи вхождения, *т.е. определение принадлежности многочлена* h *идеалу* $I = (f_1, f_2, ..., f_m)$, заключается в построении всех возможных редукций h с помощью элементов базиса Грёбнера идеала I. Многочлен h принадлежит идеалу $I = (f_1, f_2, ..., f_m)$ тогда и только тогда, когда в результате редукции получен нуль [15].

Для каждого идеала $I \lhd K[x_1, x_2, ..., x_n]$ существует базис Грёбнера, а само построение базиса Грёбнера основано на разрешении *зацеплений* [15].

Многочлены f_i и f_j имеют зацепление, если их старшие члены делятся одновременно на некоторый одночлен ω, отличный от константы. Пусть $f_{iC} = \omega \cdot q_1$, $f_{jC} = \omega \cdot q_2$, где ω – наибольший общий делитель старших одночленов f_{iC} и f_{jC}. Рассмотрим многочлен $F_{i,j} = f_i \cdot q_2 - f_j \cdot q_1 \in I$ и редуцируем его с помощью базиса $f_1, f_2, ..., f_m$ до тех пор, пока это возможно. Если полученный в результате многочлен $F'_{i,j} \equiv 0$, тогда говорят, что зацепление разрешимо. Иначе, добавим к базису $f_1, f_2, ..., f_m$ идеала I полученный многочлен $f_{m+1} = F'_{i,j}$, после чего процедуру поиска и редуцирования зацеплений продолжим. После редуцирования конечного числа зацеплений получим набор $f_1, f_2, ..., f_m, f_{m+1}, ..., f_M$, в котором каждое зацепление разрешимо.

В соответствии с ***бриллиантовой леммой*** базис $f_1, f_2, ..., f_m$ идеала $I \lhd K[x_1, x_2, ..., x_n]$ является базисом Грёбнера только тогда, когда в нем нет неразрешимых зацеплений [15].

Разрешение зацеплений позволяет определить эффективный алгоритм построения базиса Грёбнера идеала $I = (f_1, f_2, ..., f_m)$ (***алгоритм Бухбергера***).

Шаг 1. Проверяем наличие зацеплений в наборе $f_1, f_2, ..., f_m$. Если зацеплений нет, тогда набор $f_1, f_2, ..., f_m$ является базисом Грёбнера идеала $I = (f_1, f_2, ..., f_m)$. Если зацепление есть, то осуществляется переход к Шагу 2.

Шаг 2. По найденному зацеплению многочленов f_i и f_j составляем многочлен $F_{i,j} = f_i \cdot q_2 - f_j \cdot q_1$ и редуцируем его с помощью набора $f_1, f_2, ..., f_m$ пока это возможно. Если многочлен $F_{i,j}$ редуцировался к ненулевому многочлену f_{m+1}, то переходим к Шагу 3, иначе – к Шагу 4.

Шаг 3. Добавляем многочлен f_{m+1} к набору $f_1, f_2, ..., f_m$ и переходим к Шагу 4.

Шаг 4. Ищем ранее не рассмотренное зацепление и переходим к Шагу 2. Если все зацепления рассмотрены, то выводим полученный набор $f_1, f_2, ..., f_m, f_{m+1}, ..., f_M$, в котором все зацепления разрешимы. Это и есть базис Грёбнера идеала $I = (f_1, f_2, ..., f_m)$.

В настоящее время (2016 год) известны и другие алгоритмы построения базиса Грёбнера, например, алгоритмы F4, F5 [17,18].

Базис Грёбнера можно *упростить* следующими способами [15].

1. *Минимизация базиса Грёбнера*. Если f_i и f_j два элемента базиса Грёбнера, причем их старшие члены f_{iC} и f_{jC} делятся друг на друга, например, $f_{jC} \mid f_{iC}$, тогда многочлен f_i можно удалить из набора $f_1, f_2, ..., f_m$. Базис Грёбнера называют *минимальным*, если f_{iC} не делится на f_{jC} для всех $i \neq j$.
2. *Редуцирование базиса Грёбнера*. Если некоторый член q многочлена f_i делится на старший член многочлена f_j, тогда редуцируем q с помощью f_j и результат редукции запишем вместо члена q в многочлен f_i. При этом базис Грёбнера останется базисом Грёбнера, число элементов базиса не изменится, однако степени многочленов $f_1, f_2, ..., f_m$ понижаются. Базис Грёбнера называют *редуцированным*, если ни один член многочлена f_i не делится на старший член многочлена f_j для всех $i \neq j$.

Минимальный редуцированный базис Грёбнера идеала $I \lhd K[x_1, x_2, ..., x_n]$ определен однозначно (с единичными коэффициентами при старших степенях элементов базиса), т.е. не зависит от выбора исходного базиса идеала $I = (f_1, f_2, ..., f_m)$ и от последовательности проводимых операций (но зависит от упорядочения переменных $x_1, x_2, ..., x_n$) [15].

Понятие *минимального редуцированного базиса Грёбнера* было использовано в работе Жан-Шарля Фожера (Jean-Charles Faugère) [13] с целью определения алгебраической иммунности S-блоков *(нелинейных узлов усложнения)* блочных симметричных шифров.

Рассмотрим нелинейный узел (S-блок) блочного симметричного шифра (см. Рис. 2), который реализует булево отображение $S : GF(2)^n \to GF(2)^m$ [1-9].

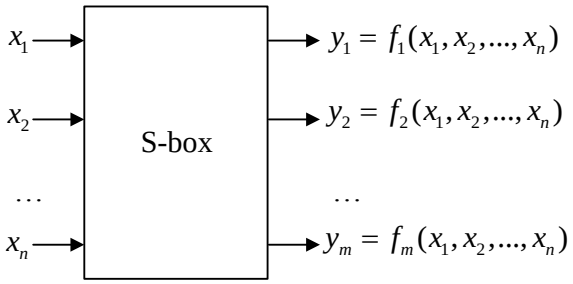


Рис. 2 – Структурная схема нелинейного узла блочного симметричного шифра

S-блок задается системой алгебраических уравнений над двоичным полем:

$$\begin{cases} f_1(x_1, x_2, ..., x_n) = y_1, \\ f_2(x_1, x_2, ..., x_n) = y_2, \\ ... \\ f_m(x_1, x_2, ..., x_n) = y_m, \end{cases} \quad (7)$$

т.е. совокупностью булевых многочленов

$$\begin{aligned} & y_1 - f_1(x_1, x_2, ..., x_n), \\ & y_2 - f_2(x_1, x_2, ..., x_n), \\ & \quad ..., \\ & y_m - f_m(x_1, x_2, ..., x_n) \end{aligned} \quad (8)$$

в кольце $K[x_1, x_2, ..., x_n, y_1, y_2, ..., y_m]$ от переменных $x_1, x_2, ..., x_n, y_1, y_2, ..., y_m$ с коэффициентами над полем $K = GF(2)$.

С системой уравнений (7), алгебраически задающих структуру S-блока, свяжем идеал I, порожденный многочленами (8):

$$I(S) = (y_1 - f_1(x_1, x_2, ..., x_n), y_2 - f_1(x_1, x_2, ..., x_n), ..., y_m - f_m(x_1, x_2, ..., x_n)) =$$
$$= \{(y_1 - f_1) \cdot r_1 + (y_2 - f_1) \cdot r_2 + ... + (y_m - f_m) \cdot r_m; r_1, r_2, ..., r_m \in GF(2)[x_1, x_2, ..., x_n, y_1, y_2, ..., y_m]\}.$$

Алгебраическая иммунность нелинейного узла блочного симметричного шифра определяется как минимальная степень многочлена P из идеала $I(S)$ [13]:

$$AI(S) = \min\{\deg(P), P \in I(S) \lhd GF(2)[x_1, x_2, ..., x_n, y_1, y_2, ..., y_m]\}, \quad (9)$$

причем минимальный редуцированный базис Грёбнера идеала $I(S)$ при степенном обратном словарном упорядочении (degrevlex) содержит линейный базис полиномов P из $I(S)$, таких, что $AI(S) = \deg(P)$. Другими словами, для вычисления алгебраической иммунности $AI(S)$ достаточно построить минимальный редуцированный базис Грёбнера идеала $I(S)$, заданного уравнениями (8) и найти многочлен минимальной степени среди элементов этого базиса. Значение минимальной степени и является значением алгебраической иммунности $AI(S)$ узла замен блочного симметричного шифра.

Связь алгебраической иммунности S-блока (9) и булевой функции (1) показана на стр. 337 в работе [19].

Рассмотрим булеву функцию $f_S(x_1, x_2, ..., x_n, y_1, y_2, ..., y_m) : GF(2)^{2n} \to GF(2)$, значения которой определим следующим образом:

$$f_S(x_1, x_2, ..., x_n, y_1, y_2, ..., y_m) = \begin{cases} 1, \forall i, j : f_i(x_1, x_2, ..., x_n) = y_j, \\ 0, \exists i, j : f_i(x_1, x_2, ..., x_n) \neq y_j. \end{cases}$$

Множество решений уравнения

$$f_S(x_1, x_2, ..., x_n, y_1, y_2, ..., y_m) - 1 = 0$$

совпадает с множеством решений системы (7). Следовательно, имеем различные базисы $(f_S - 1)$ и $(y_1 - f_1, y_2 - f_1, ..., y_m - f_m)$ одного идеала эквивалентных систем, т.е.

$$I(f_S - 1) = I(y_1 - f_1, y_2 - f_1, ..., y_m - f_m).$$

Идеал пространства аннигиляторов $Ann(f_S)$ в кольце $GF(2)[x_1, x_2, ..., x_n, y_1, y_2, ..., y_m]$ совпадает с идеалом $I(f_S - 1)$, следовательно, алгебраическая иммунность (9) булевого отображения $S : GF(2)^n \to GF(2)^m$ совпадает с минимальной степенью ненулевых полиномов, принадлежащих аннигилятору функции f_S:

$$AI(S) = \min\{D\deg(g) \mid g \in Ann(f_S)\}.$$

Таким образом, любой S-блок можно однозначно описать булевой функцией [19], а алгебраическую иммунность этой функции можно вычислить, например, при помощи алгоритма, рассмотренного выше, в пункте 2.

4 Значения алгебраической иммунности нелинейных узлов современных шифров

В данной работе проведены сравнительные исследования алгебраической иммунности нелинейных узлов современных симметричных шифров. В качестве объектов исследования выбраны широко известные и стандартизированные на национальном и/или международном уровне блочные симметричные криптопреобразования:

- криптоалгоритм AES (стандартизирован в США, как федеральный стандарт обработки данных FIPS-197 [20], а также на международном уровне в качестве блочного шифра в ISO/IEC 18033-3 [21]);

- криптоалгоритм Camellia (стандартизирован на международном уровне в качестве блочного шифра в ISO/IEC 18033-3 [21]);
- криптоалгоритм CAST (стандартизирован на международном уровне в качестве блочного шифра в ISO/IEC 18033-3 [21]);
- криптоалгоритм SEED (стандартизирован на международном уровне в качестве блочного шифра в ISO/IEC 18033-3 [21]);
- криптоалгоритм «Калина» (национальный стандарт Украины ДСТУ 7624:2014 [22]);
- криптоалгоритм «Кузнечик» (стандартизирован в России как ГОСТ 34.12-2015 [23]);
- алгоритм «BelT» симметричного шифрования и контроля целостности Республики Беларусь (стандартизирован в СТБ 34.101.31-2011 [24]);
- криптографическая хеш-функция Whirlpool, основанная на использовании блочных симметричных криптопреобразований (стандартизирована на международном уровне в ISO/IEC 10118-3:2004 [25]).

Для вычисления алгебраического иммунитета использовалось выражение (9).

Для непосредственных вычислений использован пакет прикладного программного обеспечения Magma [26], который реализует широкий спектр функций, связанных с алгеброй, теорией групп, колец и полей, теорией чисел и многими другими разделами математики.

Исследуемые узлы замен, кроме S-блока хеш-функции Whirlpool, были подробно рассмотрены в работе [9]. В таблице 1 приведены некоторые результаты этих исследований.

Таблица 1 – Криптографические свойства нелинейных узлов блочных шифров

	B	**N**	**A**	**AD**	**PC**	**CI**	**AI**
AES	+	112	32	7	0	0	2
SEED	–	110	40	7	0	0	2
CAST-128	–	120	0	4	8	0	2
Camellia	+	112	32	7	0	0	2
«Калина»	+	104	72	7	0	0	3
«Кузнечик»	+	102	72	7	0	0	3
«BelT»	+	104	72	7	0	0	3
Whirlpool	+	95	80	7	0	0	3
Принятые обозначения:		*B – сбалансированность;*			*AD – алгебраическая степень;*		
		N – нелинейность;			*PC – критерий распространения;*		
		A – автокорреляция;			*CI – корреляционный иммунитет.*		

В последней колонке «AI» Табл. 1 приведены значения алгебраической иммунности нелинейных узлов замены современных шифров. Эти данные получены по формуле (9) посредством построения базисов Грёбнера идеалов $I(S)$, заданных совокупностями многочленов (8) из уравнений (7) соответствующих S-блоков.

Полученные результаты позволяют судить о недостаточной алгебраической иммунности нелинейных узлов блочных шифров, которые были разработаны в конце 90-х – начале 2000-х годов. Рассмотренные алгоритмы (AES, SEED, CAST-128, Camellia), представленные в современном международном стандарте ISO/IEC 18033-3, обладают сравнительно низкой алгебраической иммунностью и потенциально могут рассматриваться в качестве реальных кандидатов на построение эффективных алгебраических атак.

Напротив, блочные симметричные криптоалгоритмы «Калина», «Кузнечик», «BelT», а также криптографическая функция хеширования Whirlpool, разработаны с учетом возможного применения алгебраических атак. Нелинейные узлы замен этих алгоритмов обладают высокой алгебраической иммунностью и, по всей видимости, останутся устойчивыми к новым методам алгебраического криптоанализа.

5 Выводы

1. Методы алгебраического криптоанализа, уже с момента появления первых публикаций [27,28], превратились из абстрактных и малоприменимых математических идей в развитый и широко обсуждаемый в научном сообществе раздел современной криптологии. На сегодняшний день в этой области знаний проводится огромное число исследовательских проектов и, очевидно, что уже в ближайшие годы следует ожидать появления эффективных вычислительных алгоритмов алгебраического криптоанализа современных симметричных шифров.

2. В данной работе были рассмотрены лишь отдельные аспекты алгебраического криптоанализа, в частности, исследованы методы вычисления алгебраической иммунности нелинейных узлов симметричных шифров. Это понятие, впервые введенное для поточных криптоалгоритмов в работах [10, 11], было обобщено в [13] на случай булевых отображений, т.е. для нелинейных узлов с произвольной размерностью входов-выходов. Алгебраическая иммунность, в некотором смысле, характеризует сложность решения системы уравнений, описывающих нелинейный узел и, таким образом, позволяет получить представление об устойчивости симметричного шифра к алгебраическому криптоанализу. В частности, в работе [10] предложен алгоритм алгебраического криптоанализа поточных шифров, построенных по схеме фильтр-генератора. Сложность реализации этого алгоритма является функцией от значения алгебраической иммунности криптографической булевой функции.

3. Вычисление алгебраической иммунности нелинейного узла в общем случае сопряжено с построением базиса Грёбнера идеала кольца многочленов, заданного многочленами из уравнений блока подстановок. Эта задача решается вычислительно эффективными алгоритмами Бухбергера, F4, F5 и пр. [15-18]. Кроме того, рассмотренные математические методы могут также использоваться и для поиска эффективных алгебраических атак [19], что подтверждает перспективность и актуальность проводимых работ в данной области.

4. В данной работе приведены значения алгебраического иммунитета для узлов замен некоторых образцов современных шифров. Установлено, что криптоалгоритмы, разработанные на рубеже 90-х – начала 2000-х годов, не обладают предельными значениями алгебраической иммунности и потенциально могут рассматриваться, как кандидаты для реализации эффективных алгебраических атак. В тоже время блочные шифры последнего поколения («Калина», «Кузнечик», «BelT») учитывают возможность потенциального применения алгебраического криптоанализа и обладают предельными значениями алгебраического иммунитета.

5. Перспективным направлением являются исследования методов алгебраического криптоанализа, в частности, применение технологий квантовых вычислений для решения систем алгебраических уравнений, описывающих симметричный шифр. По мнению авторов данной работы, именно в этом направлении исследований следует ожидать наиболее значимые и интересные научные результаты.

Ссылки

[1] Menezes A. J. Handbook of Applied Cryptography / Alfred J. Menezes, Paul C. van Oorschot, Scott A. Vanstone. – CRC Press, 1997. – 794 p.

[2] Gorbenko I.D. Prykladna kryptologija. Teorija. Praktyka. Zastosuvannja: pidruchnyk dlja vyshhyh navch. zakladiv / I.D. Gorbenko, Ju.I. Gorbenko. – Kharkiv: Vyd-vo «Fort», 2013. – 880 s.

[3] Preneel B. Analysis and Design of Cryptographic Hash Functions [Electronic resource]. – Way of access: homes.esat.kuleuven.be/~preneel/phd_preneel_feb1993.pdf.

[4] Carlet C. Vectorial boolean functions for cryptography. – Cambridge: Cambridge Univ. Press. – 95 p. [Electronic resource]. – Way of access: www.math.univ-paris13.fr/~carlet/chap-vectorial-fcts-corr.pdf.

[5] Carlet C. Boolean functions for cryptography and error correcting codes. – Cambridge : Cambridge Univ. Press, 2007. – 148 p. [Electronic resource]. – Access mode: www1.spms.ntu.edu.sg/~kkhoongm/chap-fcts-Bool.pdf.

[6] Zepeng Z. On correlation properties of Boolean functions / Zhuo Zepeng, Zhang Weiguo // Chinese Journal of Electronics. – 2011. – Vol.20. – №1. – P.143-146.

[7] O'Connor L. An analysis of a class of algorithms for S-box construction / L. O'Connor // J. Cryptology. – 1994. – P. 133-151.

[8] Clark J.A., Jacob J.L., Stepney S. The Design of S-Boxes by Simulated Annealing / J.A. Clark, J.L. Jacob, S. Stepney // New Generation Computing. – 2005. – Issue 23(3). – P.219–231.

[9] Kuznetsov A.A. Analiz i sravnitel'nye issledovaniya nelineinykh uzlov zameny sovremennykh blochnykh simmetrichnykh shifrov / A.A. Kuznetsov, I.N. Belozertsev, A.V. Andrushkevich // Prikladnaya radioelektronika. – 2015. – T.14. – №4. – S. 343 – 350.
[10] Courtois N. Algebraic Attacks on Stream Ciphers with Linear Feedback / N. Courtois, W. Meier // Eurocrypt 2003: LNCS. – 2003. – Vol.2656. – P. 345-359.
[11] Meier W. Algebraic Attacks and Decomposition of Boolean Functions / W. Meier, E. Pasalic, C. Carlet // Eurocrypt 2004: LNCS. – 2004. – Vol.3027. – P. 474-491.
[12] Courtois N. Cryptanalysis of Block Ciphers with Overdefined Systems of Equations / Nicolas Courtois, Josef Pieprzyk // LNCS. – 2002. – Vol.2501. – P.267–287.
[13] Ars G. Algebraic Immunities of functions over finite fields / Gw´enol´e Ars, Jean-Charles Faug`ere // RR-5532: [Research Report]. – INRIA, 2005. – P.17.
[14] Baev V. V. Effektivnye algoritmy polucheniya otsenok algebraicheskoi immunnosti bulevykh funktsii: dissertatsiya na soiskanie uchenoi stepeni kandidata fiziko-matematicheskikh nauk : 01.01.09 / Baev Vladimir Valer'evich; [Mesto zashchity: Mosk. gos. un-t im. M.V. Lomonosova. Fak. vychislit. matematiki i kibernetiki]. – Moskva, 2008. – 101 s.
[15] Arzhantsev I.V. Bazisy Grebnera i sistemy algebraicheskikh uravnenii / I.V. Arzhantsev. // Sovremennaya matematika: Letnyaya shkola (Dubna, iyul' 2002). – Moskva: MTsNMO, 2003. – 68 s.
[16] Zlobin A.I. Komp'yuternaya algebra v sisteme Sage: uchebnoe posobie / A.I. Zlobin, O.V. Sokolova. – Moskva: MGTU im. Baumana, 2011. – 55 s.
[17] Faugère J.-C. A new efficient algorithm for computing Gröbner bases / J.-C. Faugère // Journal of Pure and Applied Algebra: [F4]. –1999. – Issue 139 (1). – P.61–88.
[18] Faugère J.-C. A new efficient algorithm for computing Gröbner bases without reduction to zero / J.-C. Faugère // Proceedings of the International Symposium on Symbolic and algebraic computation (ISSAC, 2002, July): [F5] . – 2002 . – P.75–83.
[19] Gröbner Bases, Coding, and Cryptography / Massimiliano Sala, Teo Mora, Ludovic Perret, Shojiro Sakata, Carlo Traverso. – Berlin: Springer-Verlag Heidelberg. – 426 p.
[20] FIPS 197. National Institute of Standards and Technology: Advanced Encryption Standard. – 2001 [Electronic resource]. – Way of access: http://www.nist.gov/aes. – Title from the screen.
[21] ISO/IEC 18033-3. Information technology – Security techniques – Encryption algorithms, Part 3: Block ciphers. – 80 p.
[22] DSTU 7624:2014. Informacijni tehnologii'. Kryptografichnyj zahyst informacii'. Algorytm symetrychnogo blokovogo peretvorennja. – Kyi'v: Minekonomrozvytku Ukrai'ny, 2015. – 238 s.
[23] GOST R 34.12-2015. Informatsionnaya tekhnologiya. Kriptograficheskaya zashchita informatsii. Blochnye shifry. – Moskva: Standartinform, 2015. – 25 s.
[24] STB 34.101.31-2011. Informatsionnye tekhnologii i bezopasnost'. Kriptograficheskie algoritmy shifrovaniya i kon-trolya tselostnosti. – Minsk: Gosstandart, 2011. – 32 s.
[25] ISO/IEC 10118-3:2004. Information technology – Security techniques – Hash-functions – Part 3: Dedicated hash-functions. – 94 p.
[26] Magma Computational Algebra System [Electronic resource]. – Way of access: http://magma.maths.usyd.edu.au/magma. – Title from the screen.
[27] Courtois N. Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations / Nicolas Courtois, Alexander Klimov, Jacques Patarin, Adi Shamir // Proceedings of the 19th international conference on Theory and application of cryptographic techniques EUROCRYPT'00. – 2000. – P. 392 – 407.
[28] Courtois N. Cryptanalysis of Block Ciphers with Overdefined Systems of Equations / Nicolas Courtois, Josef Pieprzyk // Advances in cryptology (ASIACRYPT, 2002). – 2002. – P.267-287.
[29] Pyshkin A. Algebraic Cryptanalysis of Block Ciphers Using Grobner Bases: Dissertation zur Erlangung des Grades Doktor rerum naturalium / Andrey Pyshkin; [Technischen Universitat Darmstadt]. – Darmstadt, 2008. – 118 p.

Reviewer: Anton Alekseychuk, Doctor of Sciences (Engineering), Associate Prof., Institute of Special Communication and Information Security, National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, Ukraine.
E-mail: alex-dtn@ukr.net

Received: December 2016.

Authors:
Alexandr Kuznetsov, Doctor of Sciences (Engineering), Full Prof., V.N. Karazin Kharkiv National University, Kharkiv, Ukraine.
E-mail: kuznetsov@karazin.ua
Yuriy Gorbenko, Ph.D., Senior Researcher, Institute of Information Technology (IIT), Kharkiv, Ukraine.
E-mail: YuGorbenko@iit.kharkov.ua
Ivan Belozertsev, student, V.N. Karazin Kharkiv National University, Kharkiv, Ukraine.
E-mail: ivanbelozersevv.jw@gmail.com
Alina Andrushkevich, Junior Researcher, V.N. Karazin Kharkiv National University, Kharkiv, Ukraine.
E-mail: hitori26@mail.ru
Aleksey Naregniy, Ph.D., Senior Researcher, V.N. Karazin Kharkiv National University, Kharkiv, Ukraine.
E-mail: o.nariezhnii@karazin.ua

The algebraic immunity of nonlinear nodes symmetric ciphers.

Abstract. Researched methods for computing algebraic immunity cryptographic Boolean functions and nonlinear knots of replacements (substitutions) of symmetric ciphers. The presented results of a comparative analysis algebraic immunity of non-linear nodes of symmetric ciphers.

Keywords: symmetric ciphers, algebraic immunity, nonlinear replacement nodes.

Рецензент: Anton Alekseychuk, Doctor of Sciences (Engineering), Associate Prof., National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, Ukraine.
E-mail: alex-dtn@ukr.net

Надійшло: Грудень 2016.

Автори:
Олександр Кузнецов, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: kuznetsov@karazin.ua
Юрій Горбенко, к.т.н., с.н.с., АТ «Інститут інформаційних технологій» (ІІТ), Харків, Україна.
E-mail: YuGorbenko@iit.kharkov.ua
Іван Білозерцев, студент, Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: ivanbelozersevv.jw@gmail.com
Аліна Андрушкевіч, м.н.с, Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: hitori26@mail.ru
Олексій Нарежній, к.т.н., с.н.с., Харківський національний університет імені В.Н. Каразіна, Харків, Україна.
E-mail: o.nariezhnii@karazin.ua

Алгебраїчний імунітет нелінійних вузлів симетричних шифрів.

Анотація. Досліджуються методи обчислення алгебраїчної імунності криптографічних булевих функцій і нелінійних вузлів замін (підстановок) симетричних шифрів. Наводяться результати порівняльного аналізу алгебраїчної імунності нелінійних вузлів симетричних шифрів.

Ключові слова: симетричні шифри, алгебраїчний імунітет, нелінійні вузли заміни.

UDC 004.056.55

IMPLEMENTING NTRU-similar ALGORITHM ON THE BASIS OF NTRUPrime

Ivan Gorbenko[1], Olena Kachko[2], Gleb Naumenko[3]

[1] V.N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine
gorbenkoi@iit.kharkov.ua

[2,3] Kharkiv National University of Radioelectronic, Nauka Ave, 14 Kharkov, 61166 Ukraine
kachko@iit.com.ua, naumenko.gs@gmail.com

Reviewer: Roman Oliynikov, Dr., Full Professor, V.N. Karazin Kharkiv National University, Svobody sq., 4, Kharkov, 61022, Ukraine
roliynykov@gmail.com

Received on December 2016

Abstract. *The modern attacks uses special structures of the rings in the NTRU similar algorithms. The article was proposed post-quantum parameters NTRUPrime without these structures. Have investigated the possibility of using these parameters for encryption on the part of the most important characteristic which distinguish NTRU methods from the rest of algorithms, namely, speed regulation characteristics. In fact, standard ANSI X9.98 – 2010 is used for this, but with NTRUPrime mathematics. Use AVX2 commands for multiplication of polynomials and effective implementation of necessary operations yielded minimal reduction in performance*

Keywords: *ANSI X9.98 – 2010, NTRUPrime, Encription Speed, Decryption Speed.*

1 Introduction

In connection with problems concerned with scope for using modern asymmetric algorithms under the conditions of the availability of quantum computers, cryptographic algorithms, the reliability of which is based on lattices [1], are of the greatest interest. It is this class that NTRU similar algorithms belong to. The standard of asymmetric encryption ANSI X9.98 – 2010 [2] has a number of drawbacks [3]. The work [3,4] offers new parameters that is, a different polynomial and different modulus for use (NTRU Prime, hereinafter referred to as NTRUPrime). We have investigated the possibility of using these parameters for encryption on the part of the most important characteristic which distinguish NTRU methods from the rest of algorithms, namely, speed regulation characteristics. In fact, standard ANSI X9.98 – 2010 is used for this, but with NTRUPrime mathematics. This is the third paper out of a series of works devoted to this class of algorithms [5,6]. Further analysis on the part of cryptographic security (*further - cryptosecurity*), including that under the conditions of using quantum computers will be carried out in the next works of the series.

2 Distinctions of classic NTRU-method from NTRUPrime and their analyses in view of computing complexity [1]

1. The field $(Z/qZ)[X]/(X^N - X - 1)$ is used in place of the ring $(Z/qZ)[X]/(X^N - 1)$. The value of 739 is suggested as N.

2. The value of 9829 is used instead of the value of the parameter q = 2048, which is employed to calculate the modulus of the polynomial coefficients. The values of parameters N and q define the name of the method, which was given by the authors [3], namely, Streamlined NTRU Prime 9829^{739}.

3. The number of 1 and -1 in a private key (the value d_f), and in a blinding polynomial (the value d_r) depend upon N and the required cryptosecurity, but for all the parameters $d_f = d_r$. For NTRUPrime $d_f = d_r = 202$.

[1] - We use the notation parameters adopted in the NTRU standard [2].

4. The full set of parameters for the classic NTRU ensure security from 112 to 256. The chosen set of parameters (N=739, q = 9829), as the authors [3] state, ensures the value of security no less than 128.

Using the ring in the classic method in carrying out multiplication (the most complex operation in encryption - decryption) was just realized with the help of calculating the coefficient index according to the modulus N, the use of the field requires multiplication according to the modulus $X^N - X - 1$.

The value q defines the modulus for calculating all the polynomial coefficients. Calculating by the modulus q = 2048 does not require using the operation of division, q = 9829 requires the said operation. Besides, the value q specifies the length of a public key. The public key is a polynomial of power N, whose boundary coefficients values are defined by the value q. The public key length is $11N$ bits for q = 2048, it is $14N$ for q = 9829. We obtain 10346 bits or 1294 bytes for N = 739. The authors [3] suggest using a combined radix (2-10) for assuming a public key in a packaged format, than the public key requires 1232 bytes. The saving of less than 5% of address space for assuming the public key will lead to time consumption for setting this key, the more so, as it is an outside user public key, and such a transformation must be performed for each new user.

The preliminary analysis of the main parameters show that the computing complexity for NTRUPrime is anticipated to be more, than that for the classic NTRU. It is this analysis that is realized in the paper.

3 Calculating additional parameters for NTRUPrime

In connection with the need for implementing the classic NTRU for new parameters, it is necessary to calculate additional parameters which are used to implement the classic NTRU.

This parameter and formulae for their calculation are given in Table 1.

Basic operations

Operations of converting between byte strings and polynomials, operations of polynomial multiplication and inversion calculation are used as basic operations. The conversion algorithms are given in [2].

3.1 Operations of polynomial multiplication

Their own multiplication algorithms are used for encryption and decryption:

for encryption: c = a01_1 * b;

for decryption and generation of a public key: c = 3(a01_1 * b) + b;

where:

a01_1 is a polynomial in field $\frac{\left(\frac{Z}{3Z}\right)[X]}{(X^N - X - 1)}$;

b, c is a polynomial in field $\frac{\left(\frac{Z}{qZ}\right)[X]}{(X^N - X - 1)}$.

The work [3] proposes using Karatsuba, multiplication methods and their combinations. We have verified the efficiency of these methods for the polynomials with N = 739 and have not achieved the operation speed increase. This is due to a great number of zero elements in the polynomial coefficients (not less than $N/3$) as well as the actual lack of multiplication operations, an addition operation is employed in place of them.

We used the algorithm of multiplication of polynomials [4]. Operations AVX2 are used instead of SSE operations.

This algorithm has been redesigned with the new modules:

We used the algorithm of multiplication of polynomials [4]. Operations AVX2 are used instead of SSE operations.

Table 1 - Additional parameters for NTRUPrime [2]

Designation	Purpose	Calculation formula	Value for NTRUPrime
d_G	Amount 1(-1) in polynomial G	$d_G = N/3$	246
bLen	Length of a random sequence. It is defined by cryptosecurity	$bLen = S$	192
maxMsgLenBytes	Maximum length of a message for encryption. It is determined taking into account *bLen* and the manner of encoding bytes of a message.	$\frac{\frac{3(N-1)}{2} - bLen}{8} - 1$	113
Llen	Number of bytes to define the length of a message to be encrypted	$\lceil Log_{256} \max MsgLenBytes \rceil$	1
Buffer LenBits	The length of the encoded data buffer	*bLen + (Llen + maxMsgLenBytes)*8*	1104
dm0	Number of 1(-1) in a polynomial, which is created after transforming the message being encrypted	$dm0 = d_f$	202
Hash HashLen	Hash being used [3]	*Depends upon crypto security*	SHA256, 256
C	Number of bits for defining the number of a polynomial coefficient	More $\lceil Log_2 N \rceil$	11
minCallsR	Minimum number of calling up a hash function to create a blinding polynomial	More $\left\lceil \frac{4 d_r c}{HashLen} \right\rceil$	35
minCallsMask	Minimum number of calling up a hash function to mask a message	More $\left\lceil \frac{16N}{5 * HashLen} \right\rceil + 1$	11

Remark: [2] - *Here we use the notation parameters adopted in the NTRU standard [2];*

[3] - *The authors [3] declare cryptosecurity of over 128 for the parameters chosen, therefore the cryptosecurity value is chosen to be 192, and the appropriate hash function is SHA256.*

This algorithm has been redesigned with the new modules:
$x^N - x - 1$ modulo reduction is performed using code:

```
    int j, k = N;
    c[0] += c[N];
    for (j = 1; j < N-1; j++, k++);
          c [j] += c [k] + c[k + 1];
c [N-1] += c[k];
```

This operation requires only N additions. Loop may be performed in parallel.
Q modulo reduction is performed with AVX2 operations.

3.2 Inversion of the polynomials

The inversion operation is applied to calculate the public key that is defined by the formula

$$h = f^{-1}gp,$$

where:

$f = 3F+1$, F is a polynomial with coefficients -1, 0, 1. The number 1, -1 is the same and is equal to d_f;

g - is a polynomial with coefficients -1, 0, 1. The number -1 is equal to $d_g = N/3$, the number 1 equals $d_g + 1$;

$p = 3$.

As it is stated in the work [7] the algorithm "almost inversion" is the most efficient one among the considered algorithms of calculating inversion for the classic NTRU, but, unfortunately, this algorithm cannot be used for calculating inversion with reference to the polynomial $x^p - x - 1$ so as with reference to q that is not the power 2. In this case it is necessary to apply Extended Euclidean Algorithm that is optimized at the expense of replacing data swapping by address swapping.

4 Algorithms of encryption and decryption

4.1 Encryption algorithm

Algorithm 23 and Algorithm 24, respectively, are defined in [2].

The main stages of these algorithms and their optimization are considered next.

Components:

Parameters *N*, *q*.

Additional parameters (Table 1).

Input:

The message *m*, which is a byte string of l bytes length,

The public key *h*.

Output:

Chiphertext e, which is a byte string or Error, if length l exceeds the parameter maxMsgLenBytes.

The algorithm of encryption is composed of the following steps.

Step 1. Forming byte string *M* of bufferLenBits/8 bytes length as:

$$b \,||\, octL \,||\, m \,||\, p0,$$

where:

b is a random bit string of length bLen;

octL, *m* are the length of a message (octL) and the message itself (*m*). To specify message length the parameter Llen is used;

p0 is the number of zero bytes, which add a message m to the maximum length. It is calculated by the formula *maxMsgLenBytes + 1 – l.*

To optimize this step padding byte string *p0* with zeroes is not obligatory it is sufficient to add one 0.

Step 2. Transformation of the obtained byte string M into the polynomial MTrin with coefficients {-1, 0, 1}. To optimize this step we transform 6-bit sequences in place of converting 3-bit ones. As a result, we obtain 4 bytes that comply with 4 MTrin coefficients. To replace the 6-bit data an array of constant, with a size of 64 elements is used that has the form:

```
0x00000000,     //{0, 0, 0, 0},
0x01000000,     //{0, 0, 0, 1},
0xFF000000,     //{0, 0, 0, -1},
…
```

If we do not obtain all the coefficients as a result of processing the byte string M, the remaining coefficients are filled with zeroes.

Step 3. Forming the byte string *sData* with the format:

OID || *m* || *b* || *hTrunc*,

where:

OID - is an identificator that is taken from parameters;

m - is a message for encryption (byte string);
b - is a random bit string of length bLen;
hTrunc - is a part of a packaged public key of length bLen bits.

To optimize this step, a common memory is used for the byte strings *sData* and M, which enables not to copy messages for forming the byte string *sData*. The public key is written in a container at its setting, which enables not to realize its transformation in a byte string at forming *sData*.

Step 4. Forming a *blinding polynomial r*. The formation algorithm (Algorithm 18 [2]) uses IGF algorithm to form a bit string, which is further used to define the numbers (indices) of the polynomial coefficients that possess the value 1, and then -1. Defines two algorithms of forming IGF [2]: IGF-2 (Algorithm 20) and IGF-RBG (Algorithm 21). We use Algorithm 20. According to algorithm 20, the hash function is called minCallsR times. For each call the byte string *sData* is added by the call number. To optimize this algorithm the hash calculation step for constant part is performed one time.

Step 5. The multiplication operation $R = r*h$ is the most time-consuming one.

Step 6. Calculating the bit string oR4. We calculate Ri according to the modulus 4 for each coefficient *R*. Then we perform packing of the obtained coefficients: 4 coefficients per byte. Finally, we obtain the bit string oR4.

Step 7. MGF transforming for the byte string oR4 with allowance for the parameter minCallsMask and obtaining the polynomial *mask*. The byte string is formed similar to Step 4. But oR4 is used as an input bit string and minCallsMask is applied as a call up number. The obtained byte string is used to form the polynomial. 5 polynomial coefficients are formed out of each byte whose value does not exceed 3^5. To optimize the coefficients calculation an array of constants of 273 sequence long is used.

The beginning of the sequence is:

```
{ 0, 0, 0, 0, 0 };
{ 1, 0, 0, 0, 0 };
{ -1, 0, 0, 0, 0 };
…
```

Step 8. Calculating the ciphertext

```
r1 = (r + Mtrin) %3
if (r1.count (1) < df or r1.count (-1) < df or r1.count (0) < df ) goto Step 1
rh := rh + r1
```

To optimize step all the necessary operations are carried out for each coefficient.

Step 9. Transforming the polynomial rh into the byte string em.

4.2 Decryption algorithm

It is made up of the following Steps.

Parameters. *N*, *q*

Input. Ciphertext (byte string em) and its length (len), private key f.

Output. OpenText (byte string m) and its length (l) or Error.

Step 1. Transforming byte string em into the polynomial e.

Step 2. Set cm'= f * e.

Step 3. Set d := cm' mod 3.

Step 4. if (d.count (-1) < df or d.count (-1) < df of d.count (0) < df)
Return Error;

Step 5. Set coR4 := ConvertToBytes((e - cm') mod 4).[4]

Step 6. Generating the mask polynomial. MGF transformating for the bit string coR4 in view of the parameter minCallsMask and obtaining the polynomial mask (see Step 7 of the encryption algorithm) and generating cMtrin.

Step 7. Transforming cMtrin into the bit string.

[4] - For optimizing steps 3-5 are realized as one step for each polynomial coefficient.

An index table is used for optimizing:

Input		0	0x0001	0x00FF	0x0100	0x0101	0x01FF	0xFF00	0xFF01
Output	0	3	6	1	4	7	2	5	

If the element 0xFFFF is available among the input elements, then return Error.

Step 8. Defining an open text and its length.

In realizing this Step it is necessary to check the length of an open text (*it cannot exceed the maximum possible one*) and the availability of the required number of zero elements at the end.

5 Experimental results

All the experiments have been performed on the computer Intel® Core™ i5-4400 CPU - 3.10 GHz, OS Windows 7 (64 bit).

With the help of profiling the functions have been defined that require the maximum time at encryption and decryption. These are functions of polynomial multiplication.

The time of realizing the functions of multiplication for polynomials with coefficients (-1, 0, 1) for encryption and coefficients (-3, 0, 3) with the exception of the first coefficient for decryption as well as the functions of encryption and decryption has been determined. For comparison similar results for the classic NTRU and NTRUPrime are given. The parameters which satisfy the following criteria were chosen for the classic NTRU out of all the parameters:

- cryptosecurity level S = 192;
- the value N closest to 739;
- ratio of the number of nonzero elements to N closest to $404/739 \approx 0{,}55$.

$N = 677$, $d_f = 157$ are the closest parameters that satisfy all the conditions.

Table 2 shows the result of the calculation experiment. These results are shown for the implementation [4] (first column), and realization of our (second column).

Table 2 – Results of the calculation experiment

	MUL (ms)		Encryption (ms)		Decryption (ms)	
Classic NTRU	0,021	0,0132	0,054	0,0394	0,076	0,0232
NTRUPrime		0,0288		0,055		0,04

6 Conclusions

Since the advent of quantum computers, the NTRU cryptosystem, that is based on cryptosecurity of lattice [1], will become one of the most perspective asymmetric cryptography methods satisfying security and operation speed conditions.

The classic NTRU [2] has been used since 2010, had shortcomings that can be eliminated due to applying the new parameters [3].

A possibility of using the parameters from [3] to the algorithm [2] has been show and implementing the suggested combination scheme has been carried out in the paper.

The results obtained are given in the table 2. Our implementation is superior to the implementation of [4]. The classic method provides better speed characteristics than the new one by a factor of practically 2, which has been anticipated. Indeed, the use of the module, inconvenient in terms of computational complexity, leads to this result. But in view of the fact that the NTRU algorithm is tens and hundreds of times smaller as the existing asymmetric methods of encryption in speed characteristics, such a deceleration not being critical.

The cryptosecurity of the suggested methods and possibility of using other parameters to increase cryptoresistance will be considered in detail in subsequent works.

References

[1] Report on Post-Quantum Cryptography [Electronic resource]. – Way of access: http://nvlpubs.nist.gov/nistpubs/ir/2016/NIST.IR.8105.pdf.
[2] American National Standard X9.98-2010. Lattice-Based Polynomial Public Key Encryption Algorithm Part 1: Key Establishment: Part 2: Data Encryption, 2010.
[3] J. Daniel. Bernstein, Chitchanok Chuengsatiansup, Tanja Lange and Christine van Vredendaal1. NTRUPrime [Electronic resource]. – Access mode: https://ntruprime.cr.yp.to/ntruprime-20160511.pdf.
[4] [Electronic resource]. – Access mode: https://github.com/NTRUOpenSourceProject/ntru-crypto.
[5] Gorbenko I.D. Features of Parameters Calculation for NTRU Algorithm / I.D. Gorbenko, O.G. Kachko, K.A. Pogrebnyak // Applied Radio Electronics: Sci. Journ. – 2015. – Vol. 14. – № 3. – P. 272-277.
[6] Kachko O.G. Analysis, Estimates and Suggestions With Respect To System parameters Generation Method in NTRU – Similar Assymetric Systems / O.G. Kachko, K.A. Pogrebnyak, L.V. Makytonina. // Radio Eengineering: Sci. Journ. – 2016. – Vol. 186. – P. 103 – 110 [in Ukrainian].
[7] Kachko E.G. Research of methods of calculating of inversion in NTRU Algorithm / E.G. Kachko, D.S. Balagura, K.A. Pogrebnyak, Yu.I. Gorbenko // Applied Radio Electronics: Sci. Journ. – 2013. – Vol. 12. – № 2. – P. 254 – 257 [in Russian].

Рецензент: Роман Олійников, д.т.н., проф., Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна.
E-mail: roliynykov@gmail.com

Надійшло: Грудень 2016.

Автори:
Іван Горбенко, д.т.н., проф., лауреат Державної премії України, Харківський національний університет імені В.Н. Каразіна, Харків, Україна. E-mail: gorbenkoi@iit.kharkov.ua

Олена Качко, к.т.н., проф., Харківський національний університет радіоелектроніки (ХНУРЕ), Харків, Україна.
E-mail: kachko@iit.com.ua

Глеб Науменко, студент, факультет комп'ютерних наук**,** Харківський національний університет радіоелектроніки, Харків, Україна. E-mail: naumenko.gs@gmail.com

Реалізація NTRU –подібного алгоритму на базі NTRUPrime.

Анотація. Сучасні атаки використовують спеціальну структуру кільця в NTRU-подібних алгоритмах. Постквантові NTRUPrime не використовують такого кільця. В роботі досліджується можливість використання цих параметрів для шифрування з боку найважливішої характеристики, яка відрізняє NTRU методи від решти методів несиметричного шифрування, а саме швидкодії. Фактично стандарт ANSI X9.98 – 2010 використовується для цього, але з NTRUPrime математикою. Застосування AVX2 команд для множення поліномів та ефективна реалізація необхідних операцій забезпечили мінімальне зменшення продуктивності.

Ключові слова: стандарт ANSI X9.98 – 2010, NTRUPrime, швидкість шифрування, швидкість дешифрування.

Рецензент: Роман Олейников, д.т.н., проф., Харьковский национальный университет имени В.Н. Каразина, Харьков, Украина. E-mail: roliynykov@gmail.com

Поступила: Декабрь 2016.

Авторы:
Иван Горбенко, д.т.н., проф., лауреат Государственной премии Украины, Харьковский национальный университет имени В. Н. Каразина, Харьков, Украина. E-mail: gorbenkoi@iit.kharkov.ua

Елена Качко, к.т.н., проф., Харьковский национальный университет радиоэлектроники (ХНУРЭ), Харьков, Украина.
E-mail: kachko@iit.com.ua

Глеб Науменко, студент, факультет компьютерных наук, Харьковский национальный университет радиоэлектроники, Харьков, Украина. E-mail: naumenko.gs@gmail.com

Реализация NTRU-подобного алгоритма на основе NTRUPrime.

Аннотация. Современные атаки используют специальную структуру кольца в NTRU подобных алгоритмах. Пост квантовый NTRUPrime не использует такого кольца. В работе исследуется возможность применения этих параметров для шифрования с точки зрения самой важной характеристики, которая отличает NTRU методы от остальных методов несимметричного шифрования, а именно быстродействия. Фактически стандарт ANSI X9.98 – 2010 используется для этого, но с NTRUPrime математикой. Применение AVX2 команд для умножения полиномов и эффективная реализация необходимых операций обеспечили минимальное уменьшение производительности.

Ключевые слова: стандарт ANSI X9.98 – 2010, NTRUPrime, скорость шифрования, скорость дешифрования.

Vesa A. Niskanen
Ph.D., Adjunct Professor,
Department of Economics & Management, University of Helsinki,
P.O. Box 27 (Latokartanonkaari 9), 00014 Helsinki, Finland
E-mail: vesa.a.niskanen@helsinki.fi

Igor Romenskiy
Doktor für physikalische-mathematische Wissenschaften, GFaI Gesellschaft zur Förderung angewandter Informatik e.V.,
Volmerstraße 3, 12489 Berlin, Deutschland
E-mail: iromensky@mail.ru

Alexey Stakhov
Doctor of Sciences (Engineering), Full Professor,
Academicians of the Academy of Engineering Sciences of Ukraine,
International Club of the Golden Section,
6 McCreary Trail, Bolton, Ont., L7E 2C8, Canada
E-mail: goldenmuseum@rogers.com

Vadim Geurkov
Ph.D., Associate Professor,
Department of Electrical and Computer Engineering Ryerson University,
350 Victoria Street, Toronto, Ontario, M5B 2K3, Canada
E-mail: vgeurkov@ee.ryerson.ca

Fionn Murtagh
Ph.D., Professor,
Department of Computing and Mathematics, University of Derby,
Kedleston Road, Derby DE22 1GB, UK
Email: f.murtagh@derby.ac.uk
Department of Computing, Goldsmiths, University of London,
New Cross, London SE14 6NW, UK
E-mail: f.murtagh@gold.ac.uk

C. Pandu Rangan
Ph.D., FNAE, Senior Professor,
Department of Computer Science and Engineering,
Indian Institute of Technology,
Madras, Chennai - 600036, India
E-mail: prangan55@gmail.com

Håvard Raddum
Ph.D.,
Simula Research Laboratory, P.O. Box 134, 1325 Lysaker, Norway
E-mail: haavardr@simula.no

Oleksandr Kazymyrov
Ph.D.,
EVRY Norge AS,
Snarøyveien 30A, 1360 Fornebu, Norway
E-mail: oleksandr.kazymyrov@evry.com

Mikołaj Karpiński
Doctor of Sciences (Engineering), Full Professor,
University of Bielsko-Biala,
ul. Willowa 2, 43-309 Bielsko-Biala, Poland
E-mail: mkarpinski@ath.bielsko.pl

Веса А. Нисканен
доктор філософії, ад'юнкт професор,
департамент економіки та менеджменту, Університет Гельсінкі,
P.O. Бокс 27 (Латокартанонкаарі 9), 00014 Гельсінкі, Фінляндія
E-mail: vesa.a.niskanen@helsinki.fi

Ігор Роменський
доктор фізико-математичних наук,
GFaI - Спілка з просування прикладної інформатики,
Фольмерштрассе 3, 12489 Берлін, Німеччина
E-mail: iromensky@mail.ru

Олексій Стахов
доктор технічних наук, професор, академік Академії інженерних наук України,
Міжнародний Клуб Золотого Перетину,
6 МаккКрері Трейл, Болтон, Онтаріо, L7E 2C8, Канада
E-mail: goldenmuseum@rogers.com

Вадим Геурков
доктор філософії, доцент,
факультет електротехніки та обчислювальної техніки університету Раєрсон,
350 Вікторія-стріт, Торонто, Онтаріо, M5B 2K3, Канада
E-mail: vgeurkov@ee.ryerson.ca

Фінн Мерта
доктор філософії, професор,
факультет обчислювальної математики університету Дербі,
Кедлестон Роад, Дербі DE22 1GB, Великобританія
Email: f.murtagh@derby.ac.uk
факультет обчислень Голдсмітського коледжу Лондонського університету,
Нью-Крос, Лондон SE14 6NW, Великобританія
E-mail: f.murtagh@gold.ac.uk

С. Панду Ранган
доктор філософії, FNAE, старший викладач,
факультет комп'ютерних наук та інженерії Індійського технологічного інституту,
Мадрас, Ченнаї - 600036, Індія
E-mail: prangan55@gmail.com

Ховард Радум
доктор філософії,
науково-дослідна лабораторія Симула, P.O. Бокс 134, 1325, Лісакер, Норвегія
E-mail: haavardr@simula.no

Олександр Казіміров
доктор філософії,
EBPI Норге АС,
Снаройвиен 30A, 1360 Форнебу, Норвегія
E-mail: oleksandr.kazymyrov@evry.com

Микола Карпінський
доктор технічних наук, професор,
Університет Бельсько-Бяла,
вул. Віллова 2, 43-309 Бельсько-Бяла, Польща
E-mail: mkarpinski@ath.bielsko.pl

Volodymyr Khoma
Doctor of Sciences (Engineering), Full Professor,
Institute «Automatics and Informatics», The Opole University of Technology,
76 Prószkowska ul., 45-758 Opole, Poland
E-mail: xoma@wp.pl

Joanna Świątkowska
Ph.D., CYBERSEC Programme Director,
Senior Research Fellow of the Kosciuszko Institute,
Feldmana ul. 4/9-10, 31-130 Kraków,
Poland
E-mail: joanna.swiatkowska@ik.org.pl

Nick Bilogorskiy
Director of Security Research,
Cyphort, 5451 Great America Parkway, Suite 225,
Santa Clara, California 95054, USA
E-mail: nick@novaukraine.org

Richard Kemmerer
Ph.D., Professor,
Computer Science Department, University of California,
Santa Barbara, CA 93106, USA
E-mail: kemm@cs.ucsb.edu

Dimiter Velev
Ph.D., Professor,
Department of Information Technologies and Communications, Faculty of Applied Informatics and Statistics, University of National and World Economy,
„8-ми декември“ st., UNSS - Studentski grad, 1700
Sofia, Bulgaria
E-mail: dgvelev@unwe.bg

Robert Brumnik
Ph.D., Professor Assistant,
GEA College, Dunajska cesta 156, 1000 Ljubljana, Slovenia
Metra Engineering Ltd, Špruha 19, 1236 Trzin, Slovenia
E-mail: robert.brumnik@metra.si

Stephan Dempe
Ph.D., Professor,
Department of Mathematics and Computer Science,
Technical University Bergakademie Freiberg, Germany
Akademischestrafte 6, D-09596, Freiberg,
Germany
E-mail: dempe@math.tu-freiberg.de

Ludmila Babenko
Doctor of Sciences (Engineering), Full Professor,
Institute of Computer Technologies and Information Safety of Southern Federal University
Chekhov str., 2, Taganrog, Rostov obl., Russia
E-mail: blk@tsure.ru

Valeriy Zadiraka
Doctor of Sciences (Physics and Mathematics),
Full Professor, Academician of the National Academy of Sciences of Ukraine, Glushkov Institute of Cybernetics (GIC) of National Academy of Sciences of Ukraine,
40 Glushkov av., Kyiv, 03187, Ukraine
E-mail: zvk140@ukr.net

Володимир Хома
доктор технічних наук, професор,
Інститут «Автоматика та інформатика», Технологічний університет Ополе,
вул. Пружовська 76, 45-758 Ополе, Польща
E-mail: xoma@wp.pl

Джоана Святковська
доктор філософії, директор програми CYBERSEC,
старший науковий співробітник Інституту Костюшки
вул. Фельдман 4 / 9-10, 31-130 Краків,
Польща
E-mail: joanna.swiatkowska@ik.org.pl

Нік Білогорський
директор з досліджень безпеки,
Цифорт, 5451 Гріт Америка Парквей, Люкс 225,
Санта-Клара, Каліфорнія 95054, США
E-mail: nick@novaukraine.org

Річард Кеммерер
PhD., професор,
факультет інформатики, Каліфорнійський університет,
Санта-Барбарі, CA 93106, США
E-mail: kemm@cs.ucsb.edu

Дімітер Велев
доктор філософії, професор,
кафедра інформаційних технологій і комунікацій,
факультет прикладної інформатики та статистики,
Університет національної та світової економіки,
вул. "8-ми декември", UNSS - Студентські град, 1700
Софія, Болгарія
E-mail: dgvelev@unwe.bg

Роберт Брумнік
доктор філософії, доцент,
GEA коледж, Дунайська цеста 156, 1000 Любляна, Словенія
Метра Інжиніринг ЛТД, Шпруха 19,1236 Тржин, Словенія
E-mail: robert.brumnik@metra.si

Стефан Демп
доктор філософії, професор,
факультет математики та інформатики, технічний університет Фрайберзької Гірничої Академії, Німеччина
Akademischestrafte 6, D -09596, Фрайберг, Німеччина
E-mail: dempe@math.tu-freiberg.de

Людмила Бабенко
доктор технічних наук, професор,
Інститут комп'ютерних технологій та інформаційної безпеки Південного федерального університету
вул. Чехова 2, Таганрог, Ростовська обл., Росія
E-mail: blk@tsure.ru

Валерій Задірака
доктор технічних наук, професор,
академік Національної академії наук України,
Інститут кібернетики імені В.М. Глушкова Національної академії наук України,
проспект Академіка Глушкова, 40, Київ, 03187, Україна
E-mail: zvk140@ukr.net

Ludmila Kovalchuk
Doctor of Sciences (Engineering), Associate Professor,
Department of mathematical methods of information security Institute of Physics and Technology,
National Technical University of Ukraine
"Kyiv Polytechnic Institute"
37, Prospect Peremohy, 03056, Kyiv-56, Ukraine
E-mail: lusi.kovalchuk@gmail.com

Людмила Ковальчук
доктор технічних наук, доцент,
кафедра математичних методів захисту інформації фізико-технічного інституту
національного технічного університету України «КПІ»,
03056, Київ, пр. Перемоги, 37, НТУУ "КПІ"
E-mail: lusi.kovalchuk@gmail.com

Anton Alekseychuk
Doctor of Sciences (Engineering), Associate Professor,
Department of application of means of cryptographic and technical defense of information, Institute of Special Communication and Information Security,
National Technical University of Ukraine
"Kyiv Polytechnic Institute"
37, Prospect Peremohy, 03056, Kyiv-56, Ukraine
E-mail: alex-dtn@ukr.net

Антон Олексійчук
доктор технічних наук, доцент,
кафедра застосування засобів криптографічного та технічного захисту інформації Інституту спеціального зв'язку та захисту інформації національного технічного університету України «КПІ»,
03056, Київ, пр. Перемоги, 37, НТУУ "КПІ"
E-mail: alex-dtn@ukr.net

Volodymyr Maxymovych
Doctor of Sciences (Engineering), Full Professor,
Institute of Computer Technologies, Automation and Metrology (ICTA), Lviv Polytechnic National University,
12 Bandera st., Lviv, 79013, Ukraine
E-mail: vmax@polynet.lviv.ua

Володимир Максимович
доктор технічних наук, професор,
Інститут комп'ютерних технологій, автоматики та метрології Національного університету «Львівська політехніка»,
вул. Степана Бандери, 12, м. Львів, 79013, Україна
E-mail: vmax@polynet.lviv.ua

Oleksiy Borysenko
Doctor of Sciences (Engineering), Full Professor,
Sumy State University,
2, Rymskogo-Korsakova st., 40007 Sumy, Ukraine
E-mail: 5352008@ukr.net

Олексій Борисенко
доктор технічних наук, професор,
Сумський державний університет,
вул. Римського-Корсакова, 2, 40007 Суми, Україна
E-mail: 5352008@ukr.net

Anatoliy Biletsky
Doctor of Sciences (Engineering), Full Professor,
Institute of Air Navigation, National Aviation University,
Kosmonavta Komarova av. 1, Kyiv, 03058, Ukraine
E-mail: abelnau@ukr.net

Анатолій Білецький
доктор технічних наук, професор,
навчально-науковий інститут аеронавігації національного авіаційного університету,
пр. Космонавта Комарова 1, Київ, 03058, Україна
E-mail: abelnau@ukr.net

Sergii Kavun
Doctor of Sciences (Economics), Ph.D. (Engineering),
Full Professor, Department of Information Technologies,
Kharkiv Educational and Research Institute
of the University of Banking,
Peremogy av. 55, Kharkiv, 61174, Ukraine
E-mail: kavserg@gmail.com

Сергій Кавун
доктор економічних наук, кандидат технічних наук,
професор, кафедра інформаційних технологій,
Харківський навчально-науковий інститут
ДВНЗ "Університет банківської справи",
пр. Перемоги 55, м. Харків, 61174, Україна
E-mail: kavserg@gmail.com

Vyacheslav Kharchenko
Doctor of Sciences (Engineering), Full Professor,
Academician of the Academy of Applied Radioelectronics Sciences, N.Ye. Zhukovskiy National Aerospace University – Kharkiv Aviation Institute (KhAI),
17 Chkalov st., 61070, Kharkiv, Ukraine
E-mail: v_s_kharchenko@ukr.net

В'ячеслав Харченко
доктор технічних наук, професор,
академік Академії наук прикладної радіоелектроніки,
Національний аерокосмічний університет
ім. М. Є. Жуковського,
вул. Чкалова, 17, 61070, м. Харків, Україна
E-mail: v_s_kharchenko@ukr.net

Valentin Lazurik
Doctor of Sciences (Physics and Mathematics),
Full Professor, V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: vtlazurik@karazin.ua

Валентин Лазурик
доктор фізико-математичних наук, професор,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: vtlazurik@karazin.ua

Volodymyr Kuklin
Doctor of Sciences (Physics and Mathematics), Full Professor, V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: kuklinvm1@gmail.com

Володимир Куклін
доктор фізико-математичних наук, професор,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: kuklinvm1@gmail.com

Ivan Gorbenko
Doctor of Sciences (Engineering), Full Professor, Academician of the Academy of Applied Radioelectronics Sciences, V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: gorbenkoi@iit.kharkov.ua

Іван Горбенко
доктор технічних наук, професор, академік Академії наук прикладної радіоелектроніки,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: gorbenkoi@iit.kharkov.ua

Victor Krasnobayev
Doctor of Sciences (Engineering), Full Professor, Honourable Inventor of Ukraine,
Honourable Radio Specialist of the USSR,
V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: v.a.krasnobaev@gmail.com

Віктор Краснобаєв
доктор технічних наук, професор, заслужений винахідник України, почесний радист СРСР,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: v.a.krasnobaev@gmail.com

Irina Lisitska
Doctor of Sciences (Engineering), Full Professor, Corresponding Member of the Academy of Applied Radioelectronics Sciences,
V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: lisitska@karazin.ua

Ірина Лисицька
доктор технічних наук, професор,
член-кореспондент Академії наук прикладної радіоелектроніки, Харківський національний університет імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: lisitska@karazin.ua

Oleksandr Potii
Doctor of Sciences (Engineering), Full Professor, Academician of the Academy of Applied Radioelectronics Sciences, V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: potav@ua.fm

Олександр Потій
доктор технічних наук, професор,
академік Академії наук прикладної радіоелектроніки,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: potav@ua.fm

Viktor Dolgov
Doctor of Sciences (Engineering), Full Professor, Academician of the Academy of Applied Radioelectronics Sciences, V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: dolgovvi@mail.ru

Віктор Долгов
доктор технічних наук, професор,
академік Академії наук прикладної радіоелектроніки,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: dolgovvi@mail.ru

Roman Oliynikov
Doctor of Sciences (Engineering), Full Professor,
V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: roliynykov@gmail.com

Роман Олійников
доктор технічних наук, професор,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: roliynykov@gmail.com

Volodymyr Mashtalir
Doctor of Sciences (Engineering), Full Professor,
V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: mashtalir@kture.kharkov.ua

Володимир Машталір
доктор технічних наук, професор,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: mashtalir@kture.kharkov.ua

Grygoriy Zholtkevych
Doctor of Sciences (Engineering), Full Professor,
V. N. Karazin Kharkiv National University,
Svobody sq., 4, Kharkiv, 61022, Ukraine
E-mail: g.zholtkevych@karazin.ua

Григорій Жолткевич
доктор технічних наук, професор,
Харківський національний університет
імені В.Н. Каразіна,
майдан Свободи 4, м. Харків, 61022, Україна
E-mail: g.zholtkevych@karazin.ua

Статті пройшли внутрішнє та зовнішнє рецензування.

Наукове видання

КОМП'ЮТЕРНІ НАУКИ ТА КІБЕРБЕЗПЕКА

Випуск 4(4) 2016

Міжнародний електронний науково-теоретичний журнал

Англійською, українською, російською мовами

Комп'ютерне верстання – Федоренко В.В.

61022, Харків, майдан Свободи, 6
Харківський національний університет імені В.Н. Каразіна

V. N. Karazin Kharkiv National University Publishing