Research and development of the mathematic models of cryptosystems based on the universal Diophantine language

V.O. Osipyan[*], *K.I.* Litvinov, and *A.S.* Zhuck

Kuban State University, Krasnodar, Russia

Abstract. This paper shows the objective necessity of improving the information security systems under the development of information and telecommunication technologies. The paper for the first time involves a new area of NP-complete problems from Diophantine analysis, namely, multi-degree systems of Diophantine equations of a given dimension and degree of Tarry-Escott type. Based on a fundamentally new number-theoretic method, a mathematical model of an alphabetic information security system (ISS) has been developed that generalizes the principle of building cryptosystems with a public key – the so called dissymmetric bigram cryptosystem. This implies to implement direct and inverse transformations according to a given algorithm based on a two-parameter solution of a multi-degree system of Diophantine equations. A formalized algorithm has been developed for the specified model of a dissymmetric bigram cryptosystem and a training example based on a normal multi-degree system of Diophantine equations of the fifth degree is presented.

1 Introduction

In terms of the rapid development of network and telecommunication technologies, including mobile communication technologies, robotic systems, the Internet of Things, digital economy and distributed ledger technology (blockchain), the problems of theory and practice of information security at all levels of its storage, processing and transmission through open communication channels are becoming more relevant. This fact stimulates scientific researches aimed at improving the existing software and hardware for information security and the development of new information security systems (ISS). Therefore, an important fundamental scientific problem of research is the development of number-theoretic methods and algorithms that make it possible to build a stable and efficient (from a practical point of view) mathematical model of the information security system ISS based on new theoretical results. At present, cryptography is an attribute of the development of information technologies and is becoming especially in demand.

Based on the theoretical sources of building the mathematical models of effective ISS or cryptosystems, we proceed from necessity to use complex NP-complete problems, solution of which would require a large amount of machine time and resources for an illegal user.

[*] Corresponding author: v.osippyan@gmail.com

Following C. Shannon [1], such problems include problems containing Diophantine difficulties, which allow modeling more stable mathematic models of ISS. Such ISS allow increasing the space of selected keys to a countable set.

The paper involves a new area of NP-complete problems from Diophantine analysis, namely the problem of the parametric solution of multi-degree systems of Diophantine equations (MSDE) of a given degree and dimension of Tarry-Escott type [2, 3]. The feature of such MSDE is that, in the general case, the algorithm of their parametric solution is unknown, based on the negative solution of Hilbert's tenth problem [4].

A new approach is proposed for the first time to build ISS mathematical models with a significant degree of efficiency and stability by combining two NP-complete problems [5, 6]: the non-standard knapsack problems proposed by the author [7, 8] and MSDE parameterization problems [9-11]. The mathematical model of a dissymmetric bigram cryptosystem (DBC) containing Diophantine difficulties, generalizing the principle of building the public key cryptosystems is given [6, 15].

The mathematical models described in this paper demonstrate the potential of using the universal Diophantine language for the development of cryptosystems with a high degree of reliability.

2 Description of some NP-complete problems using Diophantine equations

As known [4], the algebraic Diophantine equation (DE) is understood as a polynomial equation

$$D(x_1, x_2, \ldots, x_n) = 0, \tag{1}$$

the coefficients of which are integer numbers, and the solutions are also required to be found in integer numbers or non-negative integer numbers. The problem of solving DE (1) or systems of such equations is to find integer solutions or to prove that there are no such solutions. As a rule, the solutions of the equation (1) are given as an identity with one, two or more integer parameters [13-19].

So, for example, the polynomial Diophantine equation 47x–53y = 1 has the following one-parameter solution: x = 44+53n, y = 39+47n, where n is an integer numerical parameter, and the identity takes place: 47(44 + 53n) –53(39 + 47n) = 1, which holds for a countable number of n values.

The well-known DE of the second degree

$$x^2+y^2 = z^2 \tag{2}$$

has the following two-parameter solution (a, b∊Z are parameters) in the form of Pythagorean triplets:

$$x = a^2 – b^2, y = 2ab, z = a^2 + b^2.$$

Moreover, the identity

$$(a^2 – b^2)^2 +(2ab)^2= (a^2 + b^2)^2$$

shows that DE (2) has infinitely many solutions.

For practical applications, it is possible to narrow the set of numerical values both for the coefficients and for the variables included in equation (1), for example, to the set Zk= $\{0, 1, \cdots, k-1\}$, k≥ 2.

Let us consider some problems that are hard to calculable and let us show that each such problem can be modeled using some DE of the type (1). At the same time, the solution of such equation allows establishing the cipher of the corresponding cryptosystem.

The problem of solving a non-standard additive knapsack [7-9].

Let there be a set (knapsack) A $=\{a_1, a_2, \dots, a_n\}, a_i \in N$, i = 1 . . n and some natural number c. It is required to establish if for a given c such values $x_i \in Z_k = \{0, 1, \cdots, k-1\}$ exist for which a linear DE is performed:

$$\sum_{i=1}^{n} a_i x_i = \text{c}, \qquad (3)$$

which corresponds to the problem of solving a non-standard additive knapsack.

The problem of solving a non-standard multiplicative knapsack [7-9].

The problem of solving a non-standard multiplicative knapsack can be considered in a similar way based on the following exponential DE:

$$\prod_{i=1}^{n} {a_i}^{x_i} = \text{c}. \qquad (4)$$

Note that the specified equalities (3) and (4) are Diophantine equations over the set Z_k with known c (cipher) and $a_i \in$ N, i = 1. .n. Based on these equalities, the mathematical models of non-standard additive and multiplicative knapsack cryptosystems have been developed [21].

The problem of natural numbers factorization [7-9].

For a given composite number n, find natural numbers p, q ≥ 2, such that n = pq. Note that this task has a high computational complexity based on which one of the most popular methods of public key cryptography, the RSA method, is built.

According to Lagrange theorem [18], every natural number is the sum of no more than four squares, and this fact is equivalent to the solvability in integers of DE:

$$n = (x_1a_1^2 + x_2a_2^2 + x_3a_3^2 + x_4a_4^2)\ (y_1b_1^2 + y_2b_2^2 + y_3b_3^2 + y_4b_4^2),\ a,\ b \in N_0,$$
$$x_i,\ y_j \in \{0,\ 1\}$$

The problem of decryption by RSA algorithm [22].

This problem consists in finding the residue x ϵ Zk encrypting the original text by its cipher c = x^e(modn), which is equivalent to the solvability of DE with respect to the variables x and y with known c and n:

$$x^e = c + n * y.$$

The problem of taking the discrete logarithm [6].

Definition. Let GF(p) be a Galois field of order p, and a, c ϵ GF(p). Any integer x, for which a^x= c (mod p), is called the discrete logarithm c for the base a, which is written as

$$x = \log_a c \ (\mathrm{mod}\ p) \text{ or } x = \log_a c + p * y.$$

This definition implies the following DE for the variables x and y:

$$a^x = c + p * y.$$

Note that calculating the discrete logarithm in GF(p) is a difficult task when p – 1 has a large prime factor.

The problem of quadratic residue in Galois field GF(p) [5].

This problem is hard to calculable, and it is reduced to the solvability of DE with respect to the variables x and y with known a and p:

$$x^2 = a + p * y\ .$$

3 Mathematical modeling of a dissymmetric bigram cryptosystem containing Diophantine difficulties

In the previous paragraph, we considered various cryptosystems and showed the corresponding DE, which can be represented in the form of more general Diophantine equation as follows:

$$D(x_1, x_2, \dots, x_n) = 0,$$

where D is an integral-valued function with integral-valued arguments.

The following multi-degree systems of Diophantine equations (MSDE) of m dimension and n order (or degree) will be of special interest in this paper [2, 3]:

$$X_1^k + X_2^k + \cdots + X_m^k = Y_1^k + Y_2^k + \cdots Y_m^k, \qquad k = 1..n$$

or in the short form:

$$X_1, X_2, ..., X_m \overset{n}{=} Y_1, Y_2, ..., Y_m. \qquad (5)$$

For the sake of brevity, we will present this record in the following form:

$$X \overset{n}{=} Y,$$

where $X = X_1, X_2, ..., X_m,\ Y = Y_1, Y_2, ..., Y_m,,$

and its parametric solution in the form as follows:

$$A \overset{n}{=} B,$$

where $\mathrm{A} = \mathrm{a}_1, \mathrm{a}_2, ..., \mathrm{a}_\mathrm{m},\ \mathrm{B} = \mathrm{b}_1, \mathrm{b}_2, ..., \mathrm{b}_\mathrm{m}, \mathrm{a}_\mathrm{i}, \mathrm{b}_\mathrm{i}$ are integer numeric values.

The multi-degree system (5) is called ideal or normal [3], if $\mathrm{m} = \mathrm{n} + 1$. For example, the following MSDE of dimension m= 6 and order n= 5 is normal

$$X_1, X_2, ..., X_6 \overset{5}{=} Y_1, Y_2, ..., Y_6$$

and has the following numerical solution:

$$1,\ 6,\ 7,\ 17,\ 18,\ 23 \overset{5}{=} 2,\ 3,\ 11,\ 13,\ 21,\ 22$$

and the following two-parameter solution:

$$a + b,\ a + 6b,\ a + 7b,\ a + 17b,\ a + 18b,\ a + 23b \overset{5}{=} a + 2b,\ a+3b,\ a+11b,\ a+13b,\ a+21b,\ a+22b, \qquad (6)$$

where a and b are arbitrary integer numbers.

We present the author’s mathematical model of dissymmetric bigram cryptosystem (DBC) based on two-parameter solution of MSDE containing Diophantine difficulties [10].

As already noted above, the feature of MSDE is in the fact that general non-exhaustive search methods of solving are unknown for any m and n [4]. At the same time, for individual values m and n, these MSDE allow parameterization for one, two or more parameters, from which it is possible to obtain specific solutions in integers or natural numbers $a_1, a_2, . . . ,a_m, b_1, b_2, . . . , b_m$ such that the equalities are met [10, 13]

$$a_1, a_2, ..., a_m \overset{n}{=} b_1, b_2, ..., b_m. \qquad (7)$$

Note that according to the found solution of MSDE (7), it is not possible to restore the numerical values of its parameters in an acceptable time. In addition, in practice, the calculations are performed for sufficiently large natural numbers, so that standard computing tools are often inapplicable. Therefore, to develop an effective ISS based on parametric solutions of MSDE, depending on the dimension m and the degree n, it is necessary to take into account either the complexity of the solution of the system (5), or the solutions themselves, or both at the same time.

As known [21], the mathematical model of an arbitrary alphabetic cryptosystem may be presented in the form of the following tuple:

$$\textstyle\sum_0 = \langle M^*, Q, C^*, E(m), D(c) | V(E(m), D(c))\rangle,$$

where M^* is a set of all messages $\mathrm{m} = \mathrm{m}_1\, \mathrm{m}_2 \ldots \mathrm{m}_\mathrm{k}$ (plain texts) over the alphabet M; Q is a set of all numeric equivalents of elementary messages m_i (in particular, letters or concatenation of letters from the alphabet M); C^* is a set of all cryptograms $\mathrm{c} = \mathrm{c}_1\, \mathrm{c}_2 \ldots \mathrm{c}_\mathrm{k}$ over the alphabet C; E(m) is the algorithm for the direct transformation of the plain text $\mathrm{m} = \mathrm{m}_1\mathrm{m}_2 \ldots \mathrm{m}_\mathrm{k}$; D(c) is the algorithm for the inverse transformation of the cryptogram $\mathrm{c} = \mathrm{c}_1\, \mathrm{c}_2 \ldots \mathrm{c}_\mathrm{k}$; and V(E(m), D(c)) is an unambiguity link between algorithms E(m) and D(c).

Let us consider the mathematical model of alphabetic DBC, the synthesis of which is based on Diophantine difficulties arising in the parametric solution of high-degree MSDE [16-22]. For the sake of brevity, we will illustrate an approach for building the mathematical model of DBC based on two-parameter solution of given MSDE. First of all, we determine the dimension l and the order k, and then its two-parameter solution (a and b are parameters) of MSDE (5) (see the example (6)) in the form:

$$X_i = v_i(a, b) = v_i, \qquad i = 1..l,$$

$$Y_i = v_j(a,b) = v_j,\ j = (l+1)..2l,$$

which may be presented in the form of the following ordered set of length 2l:

$$V^{2l} = v_1, v_2, \dots, v_{2l},$$

for which the following equalities hold for all values of the interval $1..\mathrm{k}$:

$$\mathrm{v}_1, \mathrm{v}_2, \dots, \mathrm{v}_\mathrm{l} \overset{\mathrm{k}}{=} \mathrm{v}_{\mathrm{l}+1}, \mathrm{v}_{\mathrm{l}+2}, \dots, \mathrm{v}_{2\mathrm{l}}. \tag{8}$$

Then, for fixed degree $\mathrm{d}, 1 \le \mathrm{d} \le \mathrm{k}$, we generate a direct transformation (encryption) function according to the given algorithm $\mathrm{E(m)}$ as follows:

$$E(m_{2i-1}m_{2i}) = C_L(a,b) = v_1^d + v_2^d + \dots + v_r^d = c_i, r < 2l \tag{9}$$

assuming that a is a cipher of an elementary message of bigram $\mathrm{m}_{2\mathrm{i}}\mathrm{m}_{2\mathrm{i}-1}$, and b is a private key. Accordingly, we will define D(c) – the inverse transformation (decryption) algorithm of the cryptogram c based on the following ratio:

$$C_R(a,b) = v_{r+1}^d + v_{r+2}^d + \dots + v_{2l}^d = c_i, \tag{10}$$

where $\mathrm{D(c_i)}$ is a solution of equation $\mathrm{v}_{\mathrm{r}+1}^\mathrm{d} + \mathrm{v}_{\mathrm{r}+2}^\mathrm{d} + \dots + \mathrm{v}_{2\mathrm{l}}^\mathrm{d} = \mathrm{c_i}$.

The number of terms in the right part (10) of the inverse transformation function $\mathrm{C_R(a,b)}$ can be minimized, for example, to one term [24].

3.1 Formalized algorithm of direct and inverse transformations

Let us define a formalized algorithm of direct and inverse transformations of the developed mathematical model. To demonstrate the algorithm, we will also accompany it with a training example at each stage of operation:

- algorithm for encryption and decryption of the original message bigrams

First of all, we will define the logic of encryption of the original message bigrams into a sequence of characters, which will be used in the direct and inverse transformations processes [6].

Let's give the original message m over the alphabet M – capital letters of the English 27-letter alphabet from A to Z and a space with the set Q of all numerical equivalents of q bigrams of elementary messages $\mathrm{m_i}$ from M^* with elementary numerical equivalents from 0 to 26.

We define a numerical equivalent $\widetilde{\mathrm{q_\imath}}$ of bigram $\mathrm{m}_{2\mathrm{i}-1}\mathrm{m}_{2\mathrm{i}}$ of the message m consisting of two letters $\mathrm{m}_{2\mathrm{i}-1}$ and $\mathrm{m}_{2\mathrm{i}}$ with numeric equivalents $\mathrm{q}_{2\mathrm{i}-1}$ and $\mathrm{q}_{2\mathrm{i}} \in \mathrm{Q}$ as an integer:

$$\widetilde{\mathrm{q}_\imath} = 27\mathrm{q}_{2\mathrm{i}-1} + \mathrm{q}_{2\mathrm{i}} \in \{0,1,\dots,728\}$$

(previously, the original message m is divided into bigrams with addition of a space if m contains an odd number of elementary messages, in particular, letters). Thus, a message m of length t is divided into $\lceil\frac{\mathrm{t}}{2}\rceil$ numbers corresponding to the bigrams of the original message.

Inf_Coding algorithm. {The encryption algorithm of the original message}

Input: the original message m of even length t.

Output: a set of numbers $\widetilde{\mathrm{Q}} = \{\widetilde{\mathrm{q}_1}, \dots, \widetilde{\mathrm{q}_{\lceil\frac{\mathrm{t}}{2}\rceil}}\}$.

Method:

- splitting m into separate blocks by 2 characters – bigrams;
- matching each letter $\mathrm{m_i}$ in the bigram of its ordinal number $\mathrm{q_i}$ in the alphabet M;
- calculation of the numerical equivalent of the bigram $\tilde{\mathrm{q}} = 27\mathrm{q}_1 + \mathrm{q}_2$ for each bigram;
- formation of a sequence of numerical equivalents of bigrams $\widetilde{\mathrm{Q}} = \{\widetilde{\mathrm{q}_1}, \dots, \widetilde{\mathrm{q}_{\lceil\frac{\mathrm{t}}{2}\rceil}}\}$.

De_Coding algorithm. {The decryption algorithm of the ciphertext}

Input: a set of numbers $\widetilde{\mathrm{Q}} = \{\widetilde{\mathrm{q}_1}, \dots, \widetilde{\mathrm{q}_{\lceil\frac{\mathrm{t}}{2}\rceil}}\}$.

Output: the original message m of length t.

Method:

- calculation of the initial numerical equivalents of characters from the value of the element $\tilde{q}$ of each bigram according to the rule:

$q_1 = \tilde{q}$ div 27 – the integer part of the division by 27;

$q_2 = \tilde{q}$ mod 27 – remainder of the division by 27;

- we match the symbol m_i from the alphabet M for each calculated numerical equivalent of symbol q_i ;
- then we sequentially form the message $m = \{m_1 m_2 \dots m_t\}$.

b) key and transformation functions generation

For key generation it is necessary to choose the solution of equation (5) after defining parameters l and k based on the theorems 1–6 given in the source [7]. It will allow one to get necessary solution with the specified parameters based on some initial equation. The generating such initial equation is still an open question, which will be considered in the following works.

In this example we will use the following vectors, defined by (8). Let us suppose that the following equally matched vectors are defined:

$A^l = (a_1, \dots, a_l),\ B^l = (b_1, \dots., b_l), A^l \overset{k}{=} B^l, 1 < k < l.$

Based on them we will build parametrical solution of multi-degree Diophantine equation (5) with parameters a and b according to the following rule:

$$v_i = \begin{cases} a_i a + b_i b, i = 1..l, \\ b_{i-l} a + a_{i-l} b,\ \ i = (l+1)..2l. \end{cases}$$

Using this (8) parametric solution, we will define the function of direct transformation $C_L(a, b)$ of the plain text m and the function of inverse transformation $C_R(a, b)$ of cryptogram c:

$C_L(a, b) = v_1(a, b)^d + \dots + v_l(a, b)^d - v_{l+1}(a, b)^d - \dots - v_{2l-1}(a, b)^d$,

$C_R(a, b) = v_{2l}(a, b)^d, 1 < d \leq k.$

The parameter a is numerical equivalent of the original elementary message, in particular letter or letter concatenation under the alphabet M, the parameter b is randomly selected and it is the cryptosystem key.

Gen_Keys algorithm. {Key and transformation functions generation algorithm}

Input: parameters l and k.

Output: transformation functions $C_L(a, b)$, $C_R(a, b)$ and the secret key b.

Method:

- selection of the equation (1) solution with parameters l and k as

$A^l \overset{k}{=} B^l,\ 1 < k < l,\ A^l = (a_1, \dots, a_l),\ B^l = (b_1, \dots., b_l);$

- calculation of the functions $v_i(a, b)$ according to the following rule:

$$v_i = \begin{cases} a_i a + b_i b, i = 1..l, \\ b_{i-l} a + a_{i-l} b,\ \ i = (l+1)..2l; \end{cases}$$

- calculation of the transformation functions $C_L(a, b)$ and $C_R(a, b)$ according to the following rule:

$C_L(a, b) = v_1(a, b)^d + \dots + v_l(a, b)^d - v_{l+1}(a, b)^d - \dots - v_{2l-1}(a, b)^d, \qquad 1 < d \leq k,$
$C_R(a, b) = v_{2l}(a, b)^d;$

- selection of the value for the secret key b.

c) Direct transformation

Let us encrypt the original message m of length t according to the rule from the point a). Thus, we get the sequence $\widetilde{Q} = \{\widetilde{q_1}, \dots, \widetilde{q_{\lceil \frac{t}{2} \rceil}}\}$ of bigrams numerical equivalents of the message m. Then the direct transformation of the original text using key b will look like a sequence $C = \{c_1, \dots, c_{\lceil \frac{t}{2} \rceil}\}$ defined as follows:

$$\begin{cases} c_1 = C_L(\widetilde{q_1}, b) = v_1(\widetilde{q_1}, b)^d + \cdots - v_{2l-1}(\widetilde{q_1}, b)^d, \\ c_2 = C_L(\widetilde{q_2}, b) = v_1(\widetilde{q_2}, b)^d + \cdots - v_{2l-1}(\widetilde{q_2}, b)^d, \\ \cdots \\ c_{\lceil \frac{t}{2} \rceil} = C_L\left(\widetilde{q_{\lceil \frac{t}{2} \rceil}}, b\right) = v_1\left(\widetilde{q_{\lceil \frac{t}{2} \rceil}}, b\right)^d + \cdots - v_{2l-1}\left(\widetilde{q_{\lceil \frac{t}{2} \rceil}}, b\right)^d. \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} c_1 = (a_1\widetilde{q_1} + b_1 b)^d + \cdots - (b_{2l-1}\widetilde{q_1} + a_{2l-1} b)^d, \\ c_2 = (a_1\widetilde{q_2} + b_1 b)^d + \cdots - (b_{2l-1}\widetilde{q_2} + a_{2l-1} b)^d, \\ \cdots \\ c_{\lceil \frac{t}{2} \rceil} = \left(a_1 \widetilde{q_{\lceil \frac{t}{2} \rceil}} + b_1 b\right)^d + \cdots - \left(b_{2l-1}\widetilde{q_{\lceil \frac{t}{2} \rceil}} + a_{2l-1} b\right)^d. \end{cases}$$

The sequence $C = \{c_1, \dots, c_{\lceil \frac{t}{2} \rceil}\}$ is the result of direct transformation of the original text m.

Encryption algorithm. {The direct transformation algorithm}

Input: the original message m and the direct transformation key *$(C_L(a, b), b)$.*

Output: the sequence $C = \{c_1, \dots, c_{\lceil \frac{t}{2} \rceil}\}$.

Method:

- formation of the message m’ according to the following rule: if the length t of the original message m is an even number, then m’ = m, otherwise m’ = m + " ";
- encryption algorithm of the original message m’;
- calculation of the sequence $C = \{c_1, \dots, c_{\lceil \frac{t}{2} \rceil}\}$ based on the sequence of numerical equivalents of bigrams;
- $\widetilde{Q} = \{\widetilde{q_1}, \dots, \widetilde{q_{\lceil \frac{t}{2} \rceil}}\}$ according to the rule $c_i = C_L(\widetilde{q_i}, b)$ separately for each bigram.

d) Inverse transformation

For inverse transformation of the ciphertext c, one should use the function $C_R(a, b)$. In this case, the inverse transformation algorithm for the number c will consist in solving the equation $C_R(a, b) = c$ with respect to the unknown a.

$$C_R(a, b) = c \Leftrightarrow v_{2l}(a, b)^d = c \Leftrightarrow (b_l a + a_l b)^d = c \Leftrightarrow$$

$$\Leftrightarrow b_l a + a_l b = \sqrt[d]{c} \Leftrightarrow a = \frac{\sqrt[d]{c} - a_l b}{b_l}. \tag{9}$$

Then the inverse transformation will consist in applying formula (9) to the elements of the sequence $C = \{c_1, \dots, c_{\lceil \frac{t}{2} \rceil}\}$, and further in dividing the obtained elements $\widetilde{q_i}$ into q_{2i-1} and q_{2i}.

$$\begin{cases} \widetilde{q_1} = \dfrac{\sqrt[d]{c_1} - a_l b}{b_l}, \\ \widetilde{q_2} = \dfrac{\sqrt[d]{c_2} - a_l b}{b_l}, \\ \cdots \\ \widetilde{q_{\lceil \frac{t}{2} \rceil}} = \dfrac{\sqrt[d]{c_{\lceil \frac{t}{2} \rceil}} - a_l b}{b_l}. \end{cases} \Rightarrow \begin{cases} q_1 = \widetilde{q_1} div\ 27, \\ q_2 = \widetilde{q_1} mod\ 27, \\ \cdots \\ q_{t-1} = \widetilde{q_{\lceil \frac{t}{2} \rceil}} div\ 27, \\ q_t = \widetilde{q_{\lceil \frac{t}{2} \rceil}} mod\ 27. \end{cases}$$

Then, we assign the symbols $m_1, \dots, m_t$ from the alphabet M to the elements $q_1, \dots, q_t$ and get the plain text $m_1, \dots, m_t$ as a result of the inverse transformation.

Decryption algorithm. {The inverse transformation algorithm}

Input: the sequence $C = \{c_1, \dots, c_{\lceil \frac{t}{2} \rceil}\}$ and the inverse transformation key $(C_R(a, b), b)$.

Output: the original message m.

Method:

- calculation of the sequence $\widetilde{\mathrm{Q}} = \{\widetilde{\mathrm{q}_1}, \dots, \widetilde{\mathrm{q}_{\lceil \frac{t}{2} \rceil}}\}$ based on the result of direct transformation $\mathrm{C} = \{\mathrm{c}_1, \dots, \mathrm{c}_{\lceil \frac{t}{2} \rceil}\}$ according to the rule $\widetilde{\mathrm{q}_\mathrm{l}} = \frac{\sqrt[\mathrm{d}]{\mathrm{c}_\mathrm{i}} - \mathrm{a}_\mathrm{l}\mathrm{b}}{\mathrm{b}_\mathrm{l}}$ separately for each element of the sequence.
- the original message decryption algorithm for $\widetilde{\mathrm{Q}} = \{\widetilde{\mathrm{q}_1}, \dots, \widetilde{\mathrm{q}_{\lceil \frac{t}{2} \rceil}}\}$;
- if the received message m’ ends with “ “, then to receive m we discard it; otherwise, m = m’.

3.2 DBC training example based on the normal MSDE of five order

Let us consider DBC training example based on two-parameter solution of the normal MSDE of six dimension and five order.

Based on the paragraph 4, we have in stages:

- Encryption of the original message bigrams

Let the original message be m = HELLOWORLD. Let us transform it into the sequence $\tilde{Q} = \{\widetilde{q_1}, \dots, \widetilde{q_{\lceil \frac{t}{2} \rceil}}\}$:

$$\begin{cases} m_1 = H \\ m_2 = E \\ m_3 = L \\ m_4 = L \\ m_5 = O \\ m_6 = W \\ m_7 = O \\ m_8 = R \\ m_9 = L \\ m_{10} = D \end{cases} \Rightarrow \begin{cases} q_1 = 8 \\ q_2 = 5 \\ q_3 = 12 \\ q_4 = 12 \\ q_5 = 15 \\ q_6 = 23 \\ q_7 = 15 \\ q_8 = 18 \\ q_9 = 12 \\ q_{10} = 4 \end{cases} \Rightarrow \begin{cases} \widetilde{q_1} = 27q_1 + q_2 = 221 \\ \widetilde{q_2} = 27q_3 + q_4 = 336 \\ \widetilde{q_3} = 27q_5 + q_6 = 428 \\ \widetilde{q_4} = 27q_7 + q_7 = 423 \\ \widetilde{q_5} = 27q_9 + q_8 = 328 \end{cases}$$

The sequence $\widetilde{\mathrm{Q}} = \{221, 336, 428, 423, 328\}$ will be used for the direct transformation.

- Key and transformation functions generation

For simplicity and demonstration of DBC based on two-parameter solution of the normal MSDE of five order, we will take small values both for the original solution and for the key.

Let $\mathrm{l} = 6, \mathrm{k} = 5$ and the original solution $\mathrm{A}^\mathrm{l} \overset{\mathrm{k}}{=} \mathrm{B}^\mathrm{l}$ has the following form:

$$1, 6, 7, 17, 18, 23 \; =^5 \; 2, 3, 11, 13, 21, 22.$$

Then the parametric solutions $\mathrm{v}_\mathrm{i}(\mathrm{a}, \mathrm{b})$ are defined as follows:

$$\begin{cases} v_1(a,b) = a + 2b, & v_5(a,b) = 18a + 21b, & v_9(a,b) = 11a + 7b, \\ v_2(a,b) = 6a + 3b, & v_6(a,b) = 23a + 22b, & v_{10}(a,b) = 13a + 17b, \\ v_3(a,b) = 7a + 11b, & v_7(a,b) = 2a + b, & v_{11}(a,b) = 21a + 18b, \\ v_4(a,b) = 17a + 13b, & v_8(a,b) = 3a + 6b, & v_{12}(a,b) = 22a + 23b. \end{cases}$$

Based on this parametric solution, we will define $\mathrm{C_L}(\mathrm{a}, \mathrm{b})$ and $\mathrm{C_R}(\mathrm{a}, \mathrm{b})$:

$\mathrm{C_L}(\mathrm{a}, \mathrm{b}) = (\mathrm{a} + 2\mathrm{b})^5 + \dots + (23\mathrm{a} + 22\mathrm{b})^5 - (2\mathrm{a} + \mathrm{b})^5 - \dots - (21\mathrm{a} + 18\mathrm{b})^5$, $\mathrm{C_R}(\mathrm{a}, \mathrm{b}) = (22\mathrm{a} + 23\mathrm{b})^5$.

- Direct transformation

Let us take $\mathrm{b} = 3$ as a key. Then let us calculate the direct transformation (encryption) of the sequence elements:

$$\tilde{Q} = \{221, 336, 428, 423, 328\}:$$

$$\begin{cases} c_1 = C_L(221{,}3) = (221 + 2 * 3)^5 + \cdots - (21 * 221 + 18 * 3)^5, \\ c_2 = C_L(336{,}3) = (336 + 2 * 3)^5 + \cdots - (21 * 221 + 18 * 3)^5, \\ c_3 = C_L(428{,}3) = (428 + 2 * 3)^5 + \cdots - (21 * 221 + 18 * 3)^5, \Rightarrow \\ c_4 = C_L(423{,}3) = (423 + 2 * 3)^5 + \cdots - (21 * 221 + 18 * 3)^5, \\ c_5 = C_L(328{,}3) = (328 + 2 * 3)^5 + \cdots - (21 * 221 + 18 * 3)^5. \end{cases}$$

$$\Rightarrow \begin{cases} c_1 = 2\ 915\ 244\ 687\ 863\ 990\ 000, \\ c_2 = 23\ 119\ 860\ 000\ 976\ 300\ 000, \\ c_3 = 76\ 769\ 140\ 112\ 716\ 400\ 000, \\ c_4 = 72\ 419\ 643\ 402\ 099\ 600\ 000, \\ c_5 = 20\ 518\ 602\ 614\ 059\ 500\ 000. \end{cases}$$

$\mathrm{C} = \{2\ 915\ 244\ 687\ 863\ 990\ 000, 23\ 119\ 860\ 000\ 976\ 300\ 000,$
$76\ 769\ 140\ 112\ 716\ 400\ 000, 72\ 419\ 643\ 402\ 099\ 600\ 000,$
$20\ \ 518\ 602\ 614\ 059\ 500\ 000\}$ is the result of direct transformation.

- Inverse transformation

For inverse transformation (decryption) of the original message, one should use the function $\mathrm{C_R}(\mathrm{a}, \mathrm{b})$. Let us calculate the values $\widetilde{\mathrm{q_i}}$ of the original text:

$$\begin{cases} \widetilde{\mathrm{q_1}} = (\sqrt[5]{2915244687863990000} - 23 * 3)/22 = 221, \\ \widetilde{\mathrm{q_2}} = (\sqrt[5]{23119860000976300000} - 23 * 3)/22 = 336, \\ \widetilde{\mathrm{q_3}} = (\sqrt[5]{76769140112716400000} - 23 * 3)/22 = 428, \Rightarrow \\ \widetilde{\mathrm{q_4}} = (\sqrt[5]{72419643402099600000} - 23 * 3)/22 = 423, \\ \widetilde{\mathrm{q_5}} = (\sqrt[5]{20518602614059500000} - 23 * 3)/22 = 328. \end{cases}$$

$$\Rightarrow \begin{cases} q_1 = 221\ div\ 27 = 8 \\ q_2 = 221\ mod\ 27 = 5 \\ q_3 = 336\ div\ 27 = 12 \\ q_4 = 336\ mod\ 27 = 12 \\ q_5 = 428\ div\ 27 = 15 \\ q_6 = 428\ mod\ 27 = 23 \\ q_7 = 423\ div\ 27 = 15 \\ q_8 = 423\ mod\ 27 = 18 \\ q_9 = 328\ div\ 27 = 12 \\ q_{10} = 328\ mod\ 27 = 4 \end{cases} \Rightarrow \begin{cases} m_1 = H \\ m_2 = E \\ m_3 = L \\ m_4 = L \\ m_5 = O \\ m_6 = W \\ m_7 = O \\ m_8 = R \\ m_9 = L \\ m_{10} = D. \end{cases}$$

Thus, we have got the original message:

$$\mathrm{m} = \ \mathrm{HELLOWORLD}.$$

4 Conclusion

Thus, for practical applications we should select convenient multi-degree system of Diophantine equation (MSDE) and the corresponding modified ratios (9), (10). In DBC example considered above, we chose a simple case of the direct transformation function, however, for practical applications it is possible to suggest a difficult algorithm for selection of the specified function.

The paper develops the mathematical model of DBC containing Diophantine difficulties that appeared when solving MSDE of given dimension and order. As it follows from the mentioned above, to determine the numerical equivalents of the elementary messages a legal user solves a simple equation of given degree, while an illegal user solves a multivariative MSDE of given dimension and order.

A new approach to the development of DBC based on the parametric solution of MSDE is proposed that generalizes the principle of building the public key cryptosystems, namely, for direct and inverse transformations of the processed information, the parametric solution of MSDE is divided into two parts: one part is used according to the given algorithm for the direct transformation of the plain text, and the other part is used for the reverse transformation of the closed text using blocks of given length, for example, bigrams.

In conclusion, we note that, in the general case, the problems associated with Diophantine equations systems are difficult to solve [4], and the general non-exhaustive search methods of their solution are not known for any Diophantine equations of predetermined degree and complexity. Therefore, following C. Shannon [1], these problems can be taken as a basis for development of similar DBC.

The solution of the tasks set in the paper will make it possible to get scientific and technical base for development and further implementation of stable and efficient mathematical models of alphabetic ISS and to give new impulse for development of mathematical modeling of the cryptosystems containing Diophantine difficulties.

The paper is supported by the grant No. 19-01-00596 of the Russian Foundation for Basic Research (RFBR).

References

1. C. Shannon, Bell System Techn. J., **28(4)**, 656 (1949)
2. H.L. Dorwart, O. E. Brown, Amer. Math. Monthly **44**, 613 (1937)
3. J. Chernick, Amer. Monthly **5(44n.10),** 626 (1937)
4. Ju. V. Matijasevich, *Hilbert's 10th problem*, 224 p. (M.: Izdatelstvo “Fiziko-matematicheskaja literatura” VO Nauka, 1993)
5. M. E. Gurari, O. N. Soloviev, *NP-complete set of theoretical problems: Sat. Doc.10th Anne.* AFM. Symp. On the theory of computing. New York, pp. 205 (1978)
6. N. Koblitz, *A Course in Number Theory and Cryptography*, 235 p. (New York: Springer-Verlag, 1987)
7. V. O. Osipyan, *Construction of alphabetical cryptosystems of data protection on the basis of equal power packs with Diophantine problems*, ACM, 124 (2012)
8. V. O. Osipyan, Mathematical modeling of cryptosystems based on the Diophantine problem with the gamma superposition method. Proceedings of the 8th International Conference on Information Security and Networks ACM, 338 (2015)
9. V. O. Osipyan, K.I. Litvinov, A mathematical model of the cryptosystem based on the linear Diophantine equation. SIN '18: Proceedings of the 11th International Conference on Security of Information and Networks, Article No.: 15, pp 1 (2018) https://doi.org/10.1145/3264437.3264464
10. V. O. Osipyan, K.I. Litvinov, Development of information security system mathematical models by the solutions of the multigrade Diophantine equation systems. SIN'19: Proceedings of the 12th International Conference on Security of Information and Networks September 2019, Article No.: 18, 1 (2019) https://doi.org/10.1145/3357613.3357624
11. V. O. Osipyan, K.I. Litvinov, Development of the mathematic model of dissymmetric bigram cryptosystem based on a parametric solution family of multi-degree system of Diophantine equations SIN’2020: 13th International Conference on Security of Information and Networks November 2020 Article No.: 25, 1 (2020) https://doi.org/10.1145/3433174.3433596
12. A. Salomaa, *Public Key cryptography*, 318 p. (Moscow, World Publ., 1995)

13. K. Ireland, M. Rosen, A Classical Introduction to Modern Number Theoryv, 84 (Springer-Verlag New York Heidelberg Berlin, 1982)
14. R. D. Carmichael, *Theory of numbers and Diophantine Analysis*, 118p. (New York, 1959)
15. W. Serpinski, *Elementary Theory of numbers*, 480 p. (Hafner Publishing, 1964)
16. J. W. S. Cassels, Acta Arithmetica **6**, 47 (1960)
17. A. Gloden, *Mehgradige Gleichungen, Groningen*, 104 (1944)
18. L. E. Dickson, *History of theory of numbers. Diophantine Analysis*. N.Y., vol. 2 (1971)
19. L. J. Mordell, *Diophantine equations*, 312 p. (London – New York, Acad. Press, 1969)
20. S. H. Lin, S. S. Cheng, R. T. S. Li, IEEE transactions on computers **44(1),** (1995)
21. V. O. Osipyan, *Development of mathematical models of information security systems containing Diophantine difficulties*: monograph, Rossiyskoy Federatsii, Kubanskiy gosudarstvennyy universitet, 180 p (Krasnodar: Kubanskiy gos. un-t, 2021). ISBN 978-5-8209-1960-2
22. A. P. Alferov, A. Yu. Zubov, A. S. Kuzmin, A. V. Cheremushkin, Foundations of cryptography, 2nd ed. Moscow, 480 p. (Helios ARV Publ., 2002)