Encipher of information on the basis of geometrical presentations

A.S. Petrov, A.D. Plotnikov[*]

Abstract

In this paper, we examine a ciphertext on the basis of using geometrical objects. Each symbol normative alphabet is determined as a point on the plane. We consider possible ways for presentation of these points.

1 Statement of the problem

Let there is some alphabet $A = \{a_1, a_2, \ldots, a_L\}$ and there is a plaintext

$$M = (m_1, m_2, \ldots, m_N),$$

where $m_k \in A$ $(k = 1, 2, \ldots, N)$.

The alphabet A is called *normative*, and symbols (letters) of the normative alphabet is presented by natural numbers 1,2, ..., L. As a rule, these numbers are represented in the binary scale of notation.

Formally, the plaintext M is the ordered excerpts with repetition of elements of L-set A. The number of such excerpts [1]:

$$L^N$$

and this number determines complexity of random writing M of symbols of the normative alphabet A.

[*]Department "Computing systems and networks" East-Ukrainian national university named for V. Dal.

For the information hiding, the plaintext M is enciphered by means of the key K and, as a result, the source text is transformed into a ciphertext $C = \phi(M, K)$.

We can assume that any symmetric cruptosystem can be examined as a multipole, which transforms N Boolean variables $m_k \in \{0; 1\}$ into N Boolean functions $c_k \in \{0; 1\}$. Transformation of data on every iteration of encryption can be presented by a system of Boolean functions. To impede opening the key, this Boolean functions must be *nonlinear*. This is reached by embedding so-called S-boxes.

Besides, to increase security of the symmetric cryptosystems, key length is increased. However, processing power of computers grows steadily, therefore, key length grows also. For example, key length of cryptosystem AES can be 256 bits, and Blowfish can be 448 bits. This competition between the processing power and key length can complicate a cryptosystem considerably.

Aim of this paper be to examine organization of non-linearity in the process of encryption due to presentation of the plaintext by means of the geometrical objects. This presentation can be used with any existing symmetric cryptosystem or instead of S-boxes.

2 Some variants of the use of geometrical objects

We will notice, as a value of the symbol $m_k \in M$ is determined by a value of a function $f(k, i)$, then plaintext M can be given by points on the plane $k0i$. Clear, what such presentation is a graphic interpretation of M.

We will consider such possibilities.

Let a normative alphabet be symbols (letters) of the Roman alphabet, which correspond the followings natural numbers:

Normative alphabet	A	B	C	D	E	F	G	H	I
Numerical equivalents	1	2	3	4	5	6	7	8	9
Normative alphabet	J	K	L	M	N	O	P	Q	R
Numerical equivalents	10	11	12	13	14	15	16	17	18
Normative alphabet	S	T	U	V	W	X	Y	Z	_
Numerical equivalents	19	20	21	22	23	24	25	26	27

We will notice that the numerical equivalents do not contain a zero. It is important for further.

Then the phrase "I_LOVE_MY_MOTHER" can be presented the followings pairs:

$$(1,9), (2,27), (3,12), (4,15), (5,22), (6,5), (7,27), (8,13),$$

$$(9,25), (10,27), (11,13), (12,15), (13,20), (14,8), (15,5), (16,18). \quad (1)$$

Clearly we can omit of first the components of the pairs and can consider as a sequence of second components of the ordered pair:

$$(9, 27, 12, 15, 22, 5, 27, 13, 25, 27, 13, 15, 20, 8, 5, 18). \quad (2)$$

The obtained co-ordinates of two-dimensional space can be presented as intersections some geometrical lines. In simplest case — as intersections lines on a plane. In future, for illustration of results of the use of geometrical approach, we will use, mainly, the straight lines. We will consider an example.

We suppose that the required lines pass through nearby points in the obtained sequence. The first and last points are also considered nearby. As is generally known, the normative equation of a straight line has form: $y = ax + b$ and is determined coefficients a and b fully. We will find these coefficients of a straight lines in our example.

Equation of a straight line, passing through two points (x_1, y_1) and (x_2, y_2), looks like

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}$$

or

$$y = \frac{y_2 - y_1}{x_2 - x_1}x - \frac{y_2 - y_1}{x_2 - x_1}x_1 + y_1.$$

Here, $y \neq 0$ in our case (see above).

For the first pair of points (1,9) and (16,18) we have:

$$y = 0.6x + 8.4,$$

that is, here $a = 0.6$, $b = 8.4$.

For the next pair of points (1,9) and (2,27):

$$y = 18x - 10,$$

that is, $a = 1$, $b = 29$.

Continuing, we will obtain the followings results:

Pair of points	(1,9) (2,27)	(2,27) (3,12)	(3,12) (4,15)	(4,15) (5,22)	(5,22) (6,5)	(6,5) (7,27)	(7,27) (8,13)	(8,13) (9,25)
a	18.00	-15.00	3.00	7.00	-17.00	22.00	-14.00	12.00
b	-9.00	57.00	3.00	-13.00	107.00	-127.00	125.00	-83
Pair of points	(9,25) (10,27)	(10,27) (11,13)	(11,13) (12,15)	(12,15) (13,20)	(13,20) (14,8)	(14,8) (15,5)	(15,5) (16,18)	(16,18) (1,9)
a	2.00	-14.00	2.00	5.00	-12.00	-3.00	13.00	0.60
b	7.00	167.00	-9.00	-45.00	176.00	50.00	-190	8.40)

Finding intersections nearby straight lines, we will restoration the initial plaintext.

However, the presentation of the initial plaintext by coefficients of the intersecting straight lines requires passing to redundant information. To pass the plaintext, it is necessary to pass two identical arrays of data, one of which contains the values of coefficients a, and other — b. At the described approach we must send coefficients for the straight lines, the number of which is equal to the common number of symbols in the initial plaintext. That is, the number of the straight lines is equal N — to length of the plaintext M.

It will obtain an analogical result if instead of the straight lines we will be use an ellipse, also passing through two given points. Equation of such ellipse looks like:

$$y^2 = x^3 + ax + b.$$

In this case, all said above is valid, but the formulas of calculation for ellipses are more complexity.

To decrease length of arrays, containing the values of coefficients a and b, it is possible, if symbols of the initial plaintext split on pairs. If the initial plaintext has odd length, we can add symbol of interval in the end or beginning the plaintext.

Because the phrase "I_LOVE_MY_MOTHER", in which symbols of the normative alphabet are transferable by their numerical equivalents, looks like: (9,27,12,15,22,5,27,13,25,27,13,15,20,8,5,18). Then the partition on pairs looks so:

$$((9, 27), (12, 15), (22, 5), (27, 13), (25, 27), (13, 15), (20, 8), (5, 18)) \quad (3)$$

Examining every pair in the partition (3) as a co-ordinates of point on a plane, one can present these points as intersections of the straight lines, passing through nearby points. As well as in the previous example we suppose that the first and last points determine a straight line also. Now, we must give information about straight lines, the number of which is half less, than in the previous case.

However, the equation of a straight line, passing, for example, through the pair of points (12,1) and (12,19), is $x = 12$, but it can not be presented in the form $y = ax + b$. By the way, application of an elliptic curve ($y^2 = x^3 + ax + b$) leads to the same result.

The indicated difficulty can be overcame, supposing that every equation of a straight line, connecting nearby points, it is presented in a general form:

$$Ax + By + C = 0. \quad (4)$$

Then formulas for the calculation of coefficients of general equation of the line (4), passing through two set points (x_1, y_1) and (x_2, y_2), look like:

$$A = y_2 - y_1; \quad B = x_1 - x_2; \quad C = y_1(x_2 - x_1) - x_1(y_2 - y_1).$$

For our plaintext "I_LOVE_MOTHER", the coefficients of intersecting straight lines are presented in the following table.

Pairs of points	(9,27) (12,15)	(12,15) (22,5)	(22,5) (27,13)	(27,13) (25,27)	(25,27) (13,15)	(13,15) (20,8)	(20,8) (5,18)	(5,18) (9,27)
A	-12	-10	8	14	-12	-7	10	9
B	-3	-10	-5	2	12	-7	15	-4
C	189	270	-151	-404	-24	196	-320	27

In this approach, the information about the intersecting straight lines can be passed by an arrays of coefficients, length each of which is half of length of the initial plaintext.

In common case, we can split symbols of the initial plaintext into groups from p symbols and each such group we examine as a point of p-dimensional space.

Theorem 1 *The most dimension of space, in which it is possible to examine an initial plaintext, is equal $p = [N/3]$, where $[a]$ is the least integer, exceeding a.*

To use the offered approach, it is necessary, obviously, to pass information at least three points. Q.E.D.

We will consider another approach to presentation of information, using the approximation theory of functions. We can present points of space, corresponding the initial plaintext, by the polynomial of Lagrange. Because in this case, the polynomial has a degree n, where n is the number of points, it is expedient the initial plaintext split on groups, for example, for four points. In this case the plaintext will be presented next tables, corresponding each of groups.

The combined table for these groups looks like:

i	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
x_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y_i	9	27	12	15	22	5	27	13	15	20	8	5	18	27	27	27

From the first table we obtain a polynomial:

$$P_1(x) = -93 + 161x - 67.5x^2 + 8.5x^3.$$

From the second table we have:

$$P_2(x) = 184.2 - 207.72x + 48.7x^2 - 3.18x^3.$$

The third table gives a polynomial:

$$P_3 = -55.67 + 82.91x - 12.77x^2 + 0.52x^3.$$

A fourth table gives a polynomial:

$$P_4 = 15.92 + 2.22x - 0.15x^2 + 0.003x^3.$$

Thus, instead of the sequence (2) we have a sequence

$$(-93; 161; -67.5; 8.5; 184.2; -207.72; 48.7; -3.18;$$

$$-55.67; 82.91; -12.77; 0.52; 15.92; 2.22; -0.15; 0.003).$$

It is not difficult verify that for the calculation of true sequence (3) on the built polynomials, it is necessary to find the integer values of co-ordinates of the corresponding points on the rules of rounding off.

It seems that similar approach, apparently, is not effective.

3   Conclusion

As follows from the considered variants of application of geometrical presentations for enciphering of texts, merit of this method consists of difficulty of analysis of ciphertext, because the use is practically impossible Boolean functions for description of procedure of enciphering declaration. Lack consists of that it requires excessiveness at the use. Clear, that such approach requires additional researches.

References

[1] A. Kaufmann. Introduction a la combinatorique en vue des applications. Paris: Dunod, 1968.