© International Journal of Combinatorial Optimization Problems and Informatics, Vol. 8, No. 1, Jan-April 2017, pp. 31-38. ISSN: 2007-1558

A Cryptosystem Based Upon Sums of Key Functions

Vyacheslav V. Kalashnikov[1,2,3], Viktor V. Avramenko[2], Nataliya I. Kalashnykova[4,*], and Vitaliy V. Kalashnikov[5]

[1]*Tecnológico de Monterrey (ITESM), Monterrey, Nuevo León, Mexico*
[2]*Sumy State University, Sumy, Ukraine*
[3]*Central Economics and Mathematics Institute (CEMI), Moscow, Russia*
[4]*Department of Physics and Mathematics (FCFM), Universidad Autónoma de Nuevo León (UANL), San Nicolás de los Garza, Nuevo León, Mexico*
[5]*Department of Economics, Universidad Autónoma de Nuevo León (UANL), Campus Mederos, Monterrey, Nuevo León, Mexico*

kalash@itesm.mx, avr@sumdu.edu.ua, nkalash2009@gmail.com, kalashnikov_de@yahoo.de

Abstract. This paper proposes an algorithm for designing a cryptosystem, in which the derivative disproportion functions are used. The symbols to be transmitted are encoded with the sum of at least two of these functions combined with random coefficients. A new algorithm is proposed for decoding the received messages by making use of important properties of the derivative disproportion functions. Numerical experiments are reported as demonstrating the algorithm's reliability and robustness.

Keywords: Cryptosystems, Derivative Disproportion Functions, Decoding Algorithms

1. Introduction

In the competitive world of today, the value of information is constantly increasing and therefore, it is necessary to encrypt this information in order to defend it from an unauthorized access. The latter aim has led to the widespread use of cryptographic techniques within information systems, the most famous of which are the Data Encryption Standard (DES) [1], Advanced Encryption Standard (AES) [2], and the Rivest-Shamir-Adleman (RSA) cryptosystem [3]. But the new powerful super-computers and the technologies of network and neural computing that have arisen since 2000, bring up the revision of the previous cryptographic systems that had been considered as absolutely reliable. Therefore, the development of new approaches to the creation of cryptosystems is relevant.

Almost all cryptosystems use integers as the keys. The greater the key length is, the more difficult it is to "hack" a cryptosystem by fitting the key or by solving a factorization problem. The transition from integers to real numbers, or even better, to real *functions* is expected to considerably complicate the task of code breaking (hacking) and thus to enhance the cryptosystem's reliability (resistance).

In this paper, we examine the capability of such an approach for classifying and declassifying of both an analog signal and the signal in the form of a sequence of symbols from the specified alphabet [5]. This cryptosystem is based on the use of derivative disproportion functions [5]–[6]. The input symbols are encoded by the sum of real functions (keys) multiplied by randomly selected coefficients. Due to the derivative disproportion functions an opportunity arises to recognize what functions had been used to encode the received signal. The latter allows one to decode the encrypted symbols *even though* the randomly selected factors multiplying the key functions *are unknown*.

[*]This research was financially supported by the SEP-CONACYT Grant CB-2013-01-221676

Received Jun 19, 2016 / Accepted Dec 10, 2016

The rest of the paper is arranged as follows. Sections 2 and 3 define the derivative disproportion functions and specify the problem. The decoding algorithm is described in Section 4, while Section 5 deals with the numerical examples and the results of numerical experiments. Sections 6 and 7 discuss the proposed cryptosystem's robustness and the requirements to the disproportion key functions, respectively. The paper is completed with the concluding remarks (Section 8) and the list of references.

2. Derivative Disproportion Functions

The new methods of classifying information can be developed on the basis of the use of derivative disproportion functions. Disproportion functions related to the derivatives and to the values were proposed and studied previously by the authors in [5]–[6].

The derivative disproportion functions are used to characterize real functions. They permit to obtain a quantitative assessment of the deviation of a numerical function from the power function $y = k \cdot x^n$ for a given value of the argument, regardless of the multiplier *k*. Here, $n \geq 1$ is an integer.

***Definition 1*.** The *n*-th *order derivative disproportion* of the function $y = y(x)$ with respect to *x* ($x \neq 0$) is defined as follows:

$$@d_x^{(n)} y = \frac{y}{x^n} - \frac{1}{n!} \cdot \frac{d^n y}{dx^n}. \quad \textbf{(1)}$$

In the particular case of $n = 1$ (order 1), Eq. (1) of the derivative disproportion is reduced to:

$$@d_x^{(1)} y = \frac{y}{x} - \frac{dy}{dx}. \quad \textbf{(2)}$$

As it could be expected, for the linear function $y = kx$ its disproportion of order 1 is zero for any value of the coefficient *k*. The symbol @ is chosen to designate the operation of determination of disproportion. The symbol "*d*" is selected to refer to the function's derivative as the main object of disproportion calculated. Finally, the left-hand side of Eq. (2) reads "a*t d one y with respect to x*".

If a function is specified in a parametric form, the *n*-th order derivative disproportion defined by Eq. (1) is determined by applying the rules of calculation of $d^n y / dx^n$ under the parametric dependence of *y* on *x*. In particular, the first-order derivative disproportion of the function defined parametrically as $y = \psi(t)$ and $x = \varphi(t)$ (where *t* is the parameter, and $\varphi(t) \neq 0, \varphi'(t) \neq 0$ for all *t*) has the form

$$@d_x^{(1)} y = @d_{\varphi(t)}^{(1)} \psi(t) = \frac{y}{x} - \frac{y'_t}{x'_t} = \frac{\psi(t)}{\varphi(t)} - \frac{\psi'(t)}{\varphi'(t)}. \quad \textbf{(3)}$$

It is clear that if $\psi(t) = k\varphi(t)$ for some constant *k*, its derivative disproportion defined by Eq. (3) equals zero on the common domain of the functions $y = \psi(t)$ and $x = \varphi(t)$.

Lemma 1. Every derivative disproportion function of order *n* boasts the following properties*:*

1. Multiplying the function *y* by any scalar *k* results in multiplying its derivative disproportion by the same scalar.
2. The order *n* derivative disproportion of a sum (difference) of functions equals the sum (difference) of their derivative disproportions.
3. For the linear function $y = kx$, its derivative disproportion of order 1 is zero for any value of the coefficient *k.*

Proof. It is readily verified by simple algebraic manipulations with the use of definition 1.

Remark 1. In other words, the operator $@d_x^{(n)}$ defined on the space $C^n(\Omega)$ of n times continuously differentiable real functions is linear over this space.

3. Problem Specification

Consider a communication system (channel) transmitting symbols (signals) encoded with a cryptosystem $\mathcal{K}$ based on the key functions $f_i = f_i(t)$, each defined on a (time) interval $t \in [0,T_i], T_i > 0, i = 1,\ldots,m$. The functions are assumed to be smooth and n times (continuously) differentiable. A symbol transmitted at the time moment t is encoded by the sum of (at least two) key functions with possible time delays (shifts) $\tau_i \in [0,T_i], i = 1,\ldots,m.$

For example, if the transmitted symbol is encoded by the (weighted) sum of two key functions f_p and $f_q, 1 \le p,q \le m,$ the signal transmitted through the communication channel has been encoded as

$$y(t) = k_p f_p(t+\tau_p) + k_q f_q(t+\tau_q), k_p > 0, k_q > 0. \quad (4)$$

It is assumed that an invader (intruder, hacker) who may have gotten an unauthorized access to the channel is not aware of either the key functions f_i or their time delays (shifts) τ_i, or the coefficients k_i, $i = p,q.$

On the receiver side of the communication system (channel), the complete list of key functions and their delays is known, but which of them (and with what coefficients) are involved in the received signal coded as in Eq. (4) is to be detected. The recognition of these functions and their coefficients in Eq. (4) permits one to decode the received symbol $y(t)$.

The problem of detecting both the key functions and their coefficients in Eq. (4) is solved by the algorithm proposed in the next section.

4. Algorithm Description

The problem in question is *not* easy to solve since the key functions and their coefficients can be detected only approximately. The received message $y(t)$ is expanded in time, so the exact or approximate derivatives of this function are needed. When the data are discrete, e.g., $y(t) = (y(t_0), y(t_1), \ldots, y(t_{N-1}))^T$, then the desired approximate "derivative" of the (discrete) function $y(t)$ is found by a special method, similar to that by Gregory-Newton (*cf.*, [4]).

Our algorithm is quite complicated, and due to the space restriction, here we present its description for $m = 3$ only (the complete version can be found in [6] and other publications of the second author).

The main idea of the general algorithm is as follows: if the key functions' delays (shifts) $\tau_i, i = 1\ldots,m$ are known, we may represent the received message $y(t)$ as the sum of *all* key functions with yet unknown coefficients k_i:

$$y(t) = \sum_{i=1}^{m} k_i f_i(t+\tau_i). \quad (5)$$

Next, we have to detect these coefficients at the present moment *t*. The coefficients will be equal to zero for those functions that are not really involved in the encrypted signal Eq. (5).

As we mentioned above, the description of the algorithm will be given here only for the case $m = 3$. The algorithm consists of *m* steps (that is, 3 in our case).

Step 1. Select arbitrarily one of the key functions, for instance, the first one $f_1 = f_1(t+\tau_1)$. By making use of Eq. (3) calculate the derivative disproportion for the signal $y(t)$ and denote it as $F_{01}(t) := @d^{(1)}_{f_1} y(t)$. Besides, the derivative disproportions $F_{21}(t)$ and $F_{31}(t)$ are computed for the key functions $f_2(t+\tau_2)$ and $f_3(t+\tau_3)$ with respect to $f_1(t+\tau_1)$. Due to the linearity of operator @ (*see*, Remark 1), Eq. (5) yields (for $m = 3$):

$$F_{01}(t) \equiv @d^{(1)}_{f_1} y(t) = \frac{y(t)}{f_1(t+\tau_1)} - \frac{y'(t)}{f'_1(t+\tau_1)} = k_1 \cdot 0 + k_2 \left[\frac{f_2(t+\tau_2)}{f_1(t+\tau_1)} - \frac{f'_2(t+\tau_2)}{f'_1(t+\tau_1)} \right] +$$
$$+ k_3 \left[\frac{f_3(t+\tau_2)}{f_1(t+\tau_1)} - \frac{f'_3(t+\tau_2)}{f'_1(t+\tau_1)} \right] = k_2 @d^{(1)}_{f_1} f_2(t+\tau_2) + k_3 @d^{(1)}_{f_1} f_3(t+\tau_2) \equiv \quad (6)$$
$$\equiv k_2 F_{21}(t) + k_3 F_{31}(t).$$

Here, the first term on the right-hand side of the upper line of Eq. (6) is zero due to assertion 3 of Lemma 1.

Step 2. Again, pick up randomly one of the remaining derivative disproportions $F_{21}(t)$ and $F_{31}(t)$; let it be, for instance, $F_{21}(t)$. Now we compute the derivative disproportions of the functions $F_{01}(t)$ and $F_{31}(t)$ with respect to $F_{21}(t)$; denote them as $F_{0121}(t)$ and $F_{3121}(t)$, respectively.

Applying the operator of the derivative disproportion of order 1 to both sides of Eq. (6), then making use of its linearity and assertion 3 of Lemma 1, one easily gets

$$F_{0121}(t) \equiv \frac{F_{01}(t)}{F_{21}(t)} - \frac{F'_{01}(t)}{F'_{21}(t)} = k_2 \cdot 0 + k_3 \left[\frac{F_{31}(t)}{F_{21}(t)} - \frac{F'_{31}(t)}{F'_{21}(t)} \right] \equiv k_3 F_{3121}(t). \quad (7)$$

Step 3. The relationship given by Eq. (7) shows the linear dependence of the function F_{0121} on the function F_{3121}. Therefore, based on assertion 3 of Lemma 1, we conclude that the derivative disproportion function $F_{01213121}(t)$ of the function F_{0121} with respect to F_{3121} is *zero* for all feasible *t*:

$$F_{01213121}(t) \equiv @d^{(1)}_{F_{3121}} F_{0121}(t) = \frac{F_{0121}(t)}{F_{3121}(t)} - \frac{F'_{0121}(t)}{F'_{3121}(t)} = k_3 - k_3 = 0.$$

Now one can use relations from Eq. (6) and Eq. (7) in the converse order and calculate the desired values of the unknown coefficients k_i. Indeed, first from Eq. (7), one readily gets

$$k_3 = \frac{F_{0121}}{F_{3121}}; \quad (8)$$

the latter, in its turn, combined with Eq. (6) implies:

$$k_2 = \frac{F_{01} - k_3 F_{31}}{F_{21}}. \quad (9)$$

Finally, by substituting the just found coefficients k_2 and k_3 in Eq. (5), one deduces the value of k_1:

$$k_1 = \frac{y(t) - k_2 f_2(t+\tau_2) - k_3 f_3(t+\tau_3)}{f_1(t+\tau_1)}. \quad (10)$$

The algorithm stops after having decoded the received message $y(t)$ by having detected the unknown coefficients associated with the involved key functions. All coefficients related to the non-used key functions are zero.

Remark 2. As it can be easily concluded, the knowledge of the list of involved functions and their delay (shift) values τ_i is indispensable for the implementation of this simplified version of the decoding algorithm. The more sophisticated procedures that may be needed to decipher the received message in the lack of such important information are described in [6].

5. Numerical Examples and Experiments

In order to illustrate the cryptosystem operation, let us consider the binary coding in the form of an arbitrary sequence of symbols "0", "1", space "_", and a transition to the new line (paragraph return) "\". For this model, only three real key functions are employed. The symbols being transmitted are encoded by the (weighted) sum of at least two of these functions multiplied by random factors (coefficients). The time delays (shifts) of the standard functions with respect to the current time *t* are assumed to be *zero*. The communication system (channel) can transmit only binary code symbols. Therefore, if there appears any other symbol apart from those listed above, it is perceived as a paragraph return.

To develop the numerical methods calculating the approximate derivatives, it is necessary to control the signal $y(t)$ within the interval containing at least 10 (discrete) points of the time variable *t*. In fact, the number of points in this interval may vary (the greater this number of points, the higher the cryptosystem's stability (resistance)), but in our case, it is selected constant and equals 75 (*cf*., again, [4]).

In order to simulate the operations of the cryptosystem, the following three functions are employed as key functions:

$$f_1(t) = 100\sin\left((\alpha_1 - \beta_1)t\right)\cos(15\beta_1 t);$$

$$f_2(t) = 100\exp(-0.1\alpha_2 t)\sin(10\beta_2 t)\cos\left((\alpha_2 + \beta_2)t\right);$$

$$f_3(t) = 100\exp(-0.1\alpha_3 t)\sin(400\beta_3 t),$$

where $\alpha_1 = 1$; $\alpha_2 = 0.12$; $\alpha_3 = 0.5$; $\beta_1 = 0.1$; $\beta_2 = 1.2$; $\beta_3 = 0.7$.

The coefficients k_1, k_2, and k_3, with which the key functions encode the signal $y(t)$ by Eq. (5) before its transmission, have been selected randomly by making use of a generator of pseudo-random numbers with the uniform distribution law from zero to 10 (for each symbol). However, only when encoding a symbol '1', $y(t)$ includes the entire (weighted) sum of all three key functions and therefore, their coefficients k_1, k_2, and k_3 are not equal to zero. When encrypting '0', we put $k_1 = 0$, and while encoding a space, we set $k_3 = 0$. Finally, if another symbol or the paragraph return is encoded, then $k_2 = 0$.

At any given time moment, the receiver tries to identify the involved key functions and calculate the unknown weights (coefficients) k_i, $i = 1,2,3$, by making use of the formulas from Eq. (8) – Eq. (10). Thereafter, the received message is decoded.

When any text is encrypted with the application of derivative disproportion functions, it is recommended to always introduce at least two random letters before the transmission of the binary code.

Figures 1, 2, and 3 show the diagrams of the signal $y(t)$ transmitted via the communication channel. Various examples of the cryptosystem operation when the same symbols are transmitted, as well as when the binary symbols are alternated, were treated.

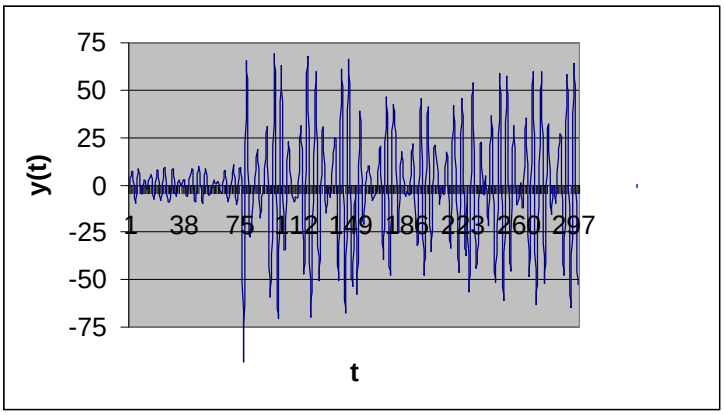

Fig. 1. The signal corresponding to the serial transmission of four symbols '0'.

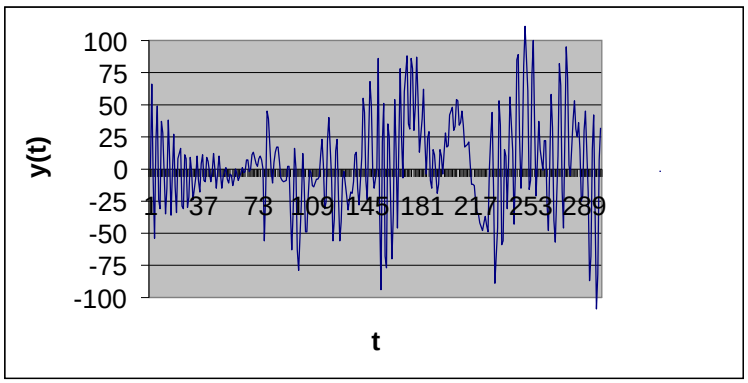

Fig. 2. The signal corresponding to the serial transmission of four symbols '1'.

Besides, the case when ASCII- codes of symbols A, B, C, D, O are transmitted, was tested. The corresponding codes were as follows:

01000001 01000010 01000011
01000100 01001111.

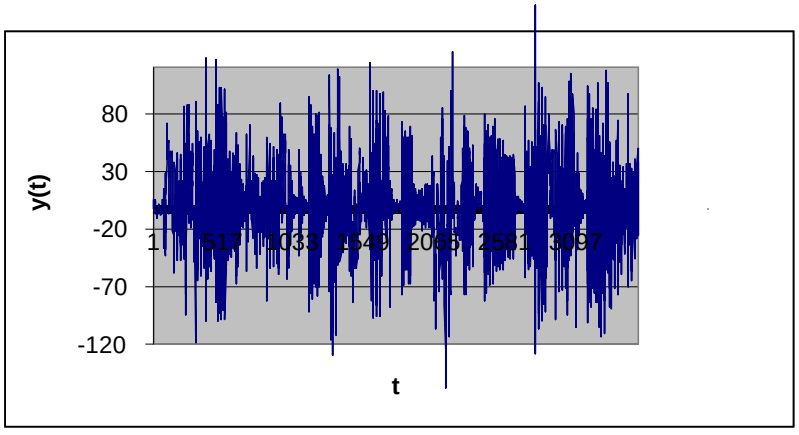

Fig. 3. The signal corresponding to the serial transmission of ASCII- codes of symbols A, B, C, D, O.

The cryptosystem's operation can be illustrated by the last (third) example. The following message was encrypted:

ar01000001 01000010q 01000011
01000100
01001111

The following message was obtained after decoding:

01000001 01000010
01000011
01000100
01001111

According to the operational algorithm, each of the letters led to the appropriate paragraph return.

In all cases, the received message was deciphered exactly as what was transmitted. At the same time, as it can be seen from the figures, it is quite difficult to reveal a message born by the transmitted signal through the communication channel unless the decoding algorithm is applied.

6. Robustness of the Cryptosystem

The cryptosystem's robustness (stability) depends on the choice of the key functions as well as on their total number. The more components are included in the signal, the more difficult the task of deciphering becomes in case it's been intercepted as a result of a hacker attack. Obviously, it is necessary not only to identify the type and the number of key functions but also to fit the coefficients involved.

How difficult it is to fit their values can be judged from the fact that in the given example, it suffices to apply $sin(9.9999\beta_2 t)$ instead of the present $sin(10\beta_2 t)$ in $f_2(t)$, or to select 400.0001 instead of the current 400 in $f_3(t)$, so that the code word consisting of four consecutive 0's is "decoded" as four 1's. This simple example confirms that any attempts on part of a hacker to "guess" the coded word by an exhaustive search for the coefficients (weight) even after having detected the key functions used, is almost always doomed to fail.

Another instance: the replacement of $\alpha_1 = 1$ with $\alpha_1 = 0.99$ in $f_1(t)$ has resulted in the distorted reception of the sole line 1100000001101000000110100000000000100000000000 without breaks in contrast to the three-line original message boasting with spaces as well.

The cited examples show that it is quite difficult just to fit the coefficients by a simple guess, to say nothing about the necessity to determine the number of functions and to fit their types.

It should be also noted that the same symbol is encoded differently depending on its position (location). Besides, one should pay attention to the fact that in this case, the frequency analysis cannot be applied for the unauthorized access and decoding.

All the above-mentioned facts show that the cryptosystems based on the (weighted) sum of real key functions are sufficiently resistant to hacking (cryptographically secure).

7. Requirements for Key Functions

1. Each (one real variable) key function has to be real-valued and sufficiently smooth (*n* times continuously differentiable).

2. The key function and its derivatives up to order *n* must *not* be constant.

3. The key functions should *not* asymptotically approach a constant value within its domain (e.g., like the function $x^{-\alpha}$, $\alpha > 0$, for the large values of *x*).

4. The collection of key functions must be selected so that to exclude the possibility that the value of one function at some point be negligible (too small by its absolute values) as compared to the values of other functions at the same point; that is, every function has to make a quite significant "contribution" to the (weighted) sum of all key functions.

5. The key functions cannot be identical.

8. Concluding Remarks

We propose a cryptosystem where (one) real variable functions are used as the keys. An example is provided to illustrate the operation of such a system where symbols are encoded by the (weighted) sum of the key functions with random coefficients. The decoding is fulfilled with the aid of the first order derivative disproportion functions calculated for the received signal and the key functions.

For a practical application of such cryptosystems, one should bear in mind that in the process of calculation of the coefficients during decoding, a division by small numbers, or a ratio of two numbers both close to zero may happen. This can lead to the information distortion. Therefore, the encrypted message must be decoded before it is transmitted via a communication channel. If necessary, the message should be encoded once again with other coefficients (weights) generated randomly for every key function.

Acknowledgments

This research was partly financially supported by the SEP-CONACYT Grant CB-2013-01-221676.

The authors would also like to express their profound gratitude to the two anonymous reviewers whose valuable comments and suggestions have helped them considerably improve the text and the structure of the paper.

References

1. U.S. Department of Commerce/National Institute of Standards and Technology Data Encryption Standard (DES) Federal Information, Processing Standards Publication, 46-3, October 25 (1999). http://csrc.nist.gov/publications/fips/fips46-3/fips46-3.pdf
2. Federal Information Processing Standards Publication 197, November 26 (2001), Specification for the ADVANCED ENCRYPTION STANDARD (AES). http://crsc.nist.gov/publications/fips/fips197/fips197.pdf
3. Rivest, R., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public – key cryptosystems, Communications of the ACM, 21(2), 120-126 (1978). DOI: 10.1145/359340.359342
4. Khan, I.R., Ohba, R., and Hozumi, N.: Mathematical proof of closed form expressions for finite difference approximations based on Taylor series, Journal of Computational and Applied Mathematics, 150(3), 303-309 (2003).
5. Avramenko, V.V., Zabolotny, M.I.: A Way of Data Coding, Patent UA H04L 9/00 №42957, Ukraine (2009).
6. Avramenko, V.V., Karpenko, A.P.: Recognition of fragments of given standards in an analyzed signal with the aid of disproportionality functions, Transactions of Sumy State University (SumDU), 34(1), 96-101 (2002).