COMPUTATIONAL PROBLEMS OF ELECTRICAL ENGINEERING

Vol. 3, No. 1, 2013

PARAMETRIZED ORTHOGONAL TRANSFORMS FOR DATA ENCRYPTION

Dariusz Puchala, Kamil Stokfiszewski

Institute of Information Technology, Technical University of Lodz, Poland

dariusz.puchala@p.lodz.pl, dir.ippt@gmail.com, kamil.stokfiszewski@p.lodz.pl

Abstract: In this paper a scheme of data encryption and decryption that takes advantage of fast parametrized orthogonal transforms has been proposed. A way of mapping the private key to the values of transform parameters has been formulated and the effectiveness of the proposed scheme has been verified experimentally. Moreover, the authors propose the directions of further research and development of the considered data encryption scheme.

Key words: data encryption, fast parametrized linear transforms.

1. Introduction

In recent years increasing interest in orthogonal transformations for data encryption might be observed. As an example one might refer to publications [1], [2], and [3] where the authors proposed: a modification of the standard JPEG stream expanding the discrete cosine transform (DCT) operation to an image cyphering task by the modification of low-frequency DCT coefficients [1], the data hiding scheme realized with the use of parametrized Slant-Haddamard transforms [2] and also the data encryption algorithm operating in the discrete wavelet transform (DWT) domain designed for the use in JPEG2000 image compression standard. In the task of data encryption parametrized orthogonal transforms also play an important role, e.g. [4], [5], and [6], for which the proper modifications of their base functions' parameters is performed according to a specific data processing task. Such parametrization doesn't influence their computational complexity, which means that the discussed computational scheme remains fast and effective. In addition, many of the types of orthogonal transforms, including DCT and DWT, have their own hardware implementations, e.g. FPGA realizations. Those algorithms can be successfully utilized for data encryption tasks, but exclusively on the condition that they are properly parametrized. It's also worth noting that if one doesn't know the form of the coding transforms' parameters during the decoding process she/he is unable to construct an inverse transformation, which in turn would prevent accurate reconstruction of the original data. Taking into consideration the above facts it's justified to conclude that fast parametrized orthogonal transforms might be successfully used in real-time cyphering multimedia systems, for which the encryption and compression processes are invoked simultaneously. As a result it is possible to obtain algorithms, which are more computationally effective than the schemes utilizing compression and encryption processes in a sequential manner, such as DES [7], IDEA [8] or AES [9].

In this article the authors propose a simple and computationally effective data encryption scheme along with the method of mapping the private keys bits' to proper values of the transformations parameters, which utilizes fast parametrized orthogonal transforms. In addition the effectiveness of the proposed scheme is verified with modeled exemplary data.

2. Parametrized orthogonal transforms

In the proposed data encryption scheme the authors utilize parametrized orthogonal transforms, i.e. linear transforms $\mathbf{U}$. For the scheme the transforms, firstly, obey the following condition $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$, where $(\cdot)^T$ denotes matrix transposition and $\mathbf{I}$ stands for the identity matrix, secondly, they consist of parametrized base functions, i.e. functions which might be adapted through the proper change of their parameters. The next crucial requirement imposed upon the proposed transforms is that they have to be fast, i.e. their computational complexity has to be an order of magnitude less (in terms of number of addition and multiplication operations) than that resulting form direct matrix by vector multiplication operation, which means that the time required for calculating the given vector representation in the transformation domain would also be an order of magnitude less than the time required for the aforementioned direct operation. The above requirements are met by the matrices $\mathbf{U}$ which are constructed as products of sparse orthogonal matrices and permutation matrices. Sparse orthogonal matrices are have at most two non-zero elements in each of their rows. Moreover, in order to obtain orthogonality of such matrices their elements are taken as rotation coefficients in chosen $\Re^2$ subspaces. In such setup the rotation angles constitute the parameter set of the transformation $\mathbf{U}$. If the number of component matrices of a given $\mathbf{U}$

transform is sufficiently small, then such transform remains computationally effective. The above reasoning leads to the following general factorization of $\mathbf{U}$ matrices (see [5]):

$$\mathbf{U} = \mathbf{P}_L \left(\prod_{i=0}^{L-1} \mathbf{U}_i \mathbf{P}_i \right), \tag{1}$$

where $\mathbf{U}_i$ are $N \times N$-dimensional block-diagonal matrices whose blocks represent the rotation angles α_{ij} in chosen $\Re^2$ subspaces, i.e.

$$\mathbf{O}_{ij} = \begin{bmatrix} \cos\alpha_{ij} & \sin\alpha_{ij} \\ -\sin\alpha_{ij} & \cos\alpha_{ij} \end{bmatrix}, \tag{2}$$

and $\mathbf{P}_i$ are $N \times N$-dimensional permutation matrices. The transformation U constructed on the basis of the definition (1) is orthogonal as the product of orthogonal matrices. The angles α_{ij}, $i = 0, 1, \ldots, L$, $j = = 0, 1, \ldots, N/2$ (for even values of N) are the transformation parameters. Computational complexity of the constructed transform can be estimated as $l_{mul} = 2N \cdot L$ (in terms of number of multiplications) and $l_{add} = N \cdot L$ (in terms of number of addition operations). In practical tasks the most common assumption is to take $L = \log_2 N$ for N being the integral power of number 2, which results in computational complexity of the order $O(N \log_2 N)$.

For practical illustration of the factorization scheme (1) let us consider exemplary two-stage transform structure based on the proposed in [10] fast structures for calculation of the Fourier, Hartley, cosine and sine transformations. A data flow diagram of such a structure for $N = 8$ - point U transform is shown in Fig. 1.

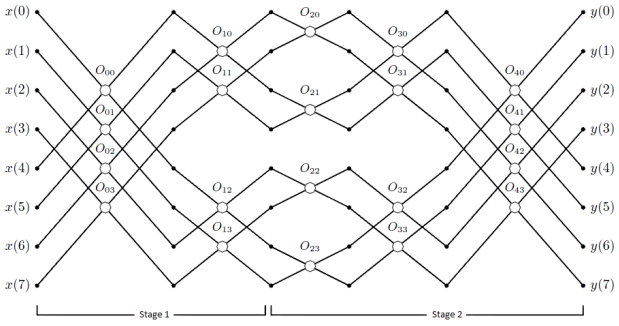


Fig. 1. Data flow diagram of fast parametrized orthogonal transform realized with a two-stage structure for $N = 8$.

Butterfly operators denoted in Fig. 1 as "∘" symbols represent rotation operations $\mathbf{O}_{ij}$ in $\Re^2$ subspaces of the forms given in (2). The $\mathbf{O}_{ij}$ operators constituting a single layer of the structure in Fig. 1 correspond to a single matrix $\mathbf{U}_i$. Additionally, the connections between $\mathbf{O}_{ij}$ operators located in consecutive layers correspond to $\mathbf{P}_i$ and $\mathbf{P}_{i+1}$ permutation matrices. Thus any transform realized with the use of the structure shown in Fig. 1 can be successfully described by formula (1). As a result, the structure depicted in Fig. 1 constitutes a graphical representation of a certain class of parametrized transforms. The number of parameters of the given transform realized with the use of the structure from Fig. 1 can be shown as $l_{par} = N(\log_2 N - 0.5)$.

The fast parametrized transform realized with the use of the structure depicted in Fig. 1 is subject to a great interest in data encryption tasks, since it is known [11], that despite the reduced number of connections constituting the structure it is still capable of realizing any given permutation of the elements of the input vector $\vec{x}$.

3. Data cyphering scheme with the use of orthogonal transforms

Let $\vec{x}$ be an N-element input data vector. Such a vector represents the message that will undergo the encryption process. The encryption process itself is realized as the product of the form $\vec{y} = \mathbf{U}\vec{x}$, where $\mathbf{U}$ denotes the parametrized orthogonal transform and $\vec{y}$ is an N-element transformation output vector called a cryptogram. To ensure the required protection of the cyphered data, the form of the encryption transform $\mathbf{U}$ (and thus the form of the decryption transform $\mathbf{U}^T$) is described by the sequence of bits constituting a private key K. The proposed mapping of the private key's individual bits to the values of the transform parameters is described later in this section.

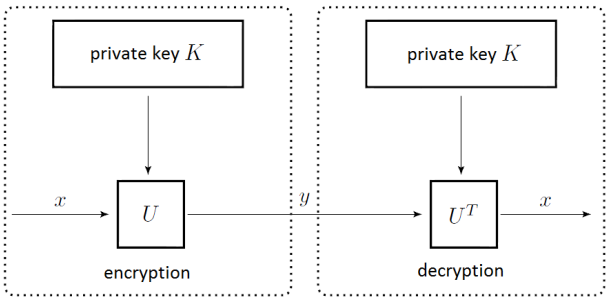


Fig. 2. Block flowchart of the cyphering and decryption processes utilizing the private key K and realized with the use of orthogonal transform U.

Decryption of the cyphered message takes place with the use of the inverse transformation $\mathbf{U}^T$ built on the basis of the private key K and takes the form of the product $\vec{z} = \mathbf{U}^T \vec{y}$, where $\vec{z}$ denotes an N-element real vector. If during the encryption and decryption processes

one uses the same private key then the message obtained as a result of the decription process is identical (up to the accuracy of computational errors) with the input message, i.e. the following equality holds $\vec{x} = \vec{z}$. On the other hand, in case of different private keys used during the cyphering and decryption processes, we demand that $\vec{x} \neq \vec{z}$ must hold. Fig. 2 shows the proposed cyphering/decryption scheme in the form of a block flowchart.

The way of mapping the private key's K values to individual values of α_{ij} parameters of the U transform for $i = 0, 1, \ldots, L-1$ and $j = 0, 1, \ldots, N/2-1$ might, in the most simple case, take the following form: one might divide (perform discretization) the $[0, 2\pi)$ angles' variability interval to equal length subintervals and assign to each of such subintervals a single representant, e.g. in the form of subintervals' starting point values, afterwards she/he may divide the sequence of bits constituting the private key K into a number of subsequences, which after the transformation from a binary to decimal form would indicate discretized values of individual angles, i.e. the transform parameters. Next, we assign the individual transform parameters α_{ij} to the private key's bit subsequences of a constant length k_0. Doing so would cause the discretization subintervals of the $[0, 2\pi)$ interval to be also of constant length taking the following form $d\alpha = 2\pi / 2^{k_b}$ and the discrete values of those parameters would then be calculated according to the following formula:

$$\alpha_{ij} = k_{ij} d\alpha , \tag{3}$$

where k_{ij} denotes the decimal value of the key's K bit subsequence corresponding to a given transform parameter. Proceeding in such manner would cause the values of k_{ij} for each $\{i\, j\}$ pair to be integer numbers taken form the $0,1,\ldots,2^{k_b} - 1$ set. The private key K is of the following form $K = b_{l_{key}-1} b_{l_{key}-2} \ldots b_1 b_0$ of the bit sequence of the length $l_{key} = k_b\, l_{par}$, wherein we assign to the individual parameters α_{ij} bit subsequences of the same length, e.g. according to the formula: $b_{k_b(l+1)-1}\, b_{k_b(l+1)-2} \ldots b_{k_b l+1}\, b_{k_b l}$, where $l = (N/2)\, i + j$ for $i = 0, 1, \ldots, L-1$ and $j = 0, 1, \ldots, N/2-1$, respectively.

Taking into consideration the aforementioned fast structure depicted in Fig. 1 we are able to easily calculate exemplary lengths of the private keys obtained for different transforms' lengths N assuming k_b to be equal to, for example, 8 bits. In such a case we obtain: $l_{par} = 20$ and $l_{key} = 160$ bits, in case of $N = 16$ we have $l_{par} = 56$ and $l_{key} = 448$, for $N = 32$ the values of l_{par} and l_{key} are equal to 144 and 1152, respectively, whereas for $N = 64$ the number of transform parameters l_{par} is equal to 352 while the key length $l_{key} = 448$ is as large as 2816 bits.

3. Experimental studies

The subject of the experimental research is statistical verification of the effectiveness of the proposed encryption scheme, covering such aspects as: the probability distribution of selecting the keys with specified Hamming distances from a fixed key K, the distribution of the expected value of the mean square error (MSE) as a function of the Hamming distance between the private keys, and the probability distribution of MSE of signal reconstruction during random trials of guessing the private key K. For that purpose, we have carried a number of experimental studies based on the first order Markov signal with variance $\sigma^2 = 1$ and the correlation coefficient $\rho = 0.9$. In the encryption and decryption process, we have used the fast two-step orthonormal transform with the structure shown in Fig. 1 (for $N = 8$) and private keys of the length $l_{key} = 160$ bits. The results are shown in Fig. 3, Fig. 4, and Fig. 5.

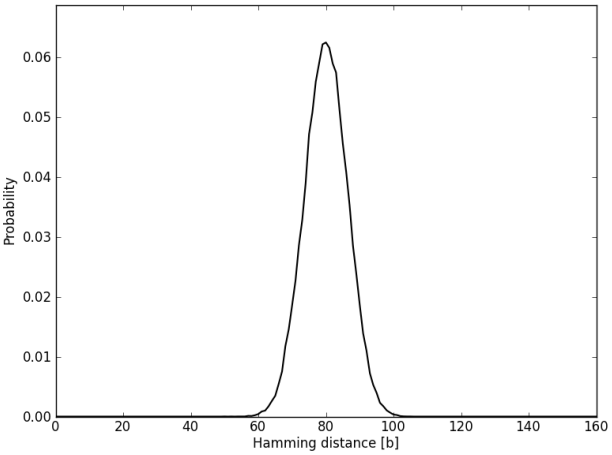


Fig. 3. Probability distribution of selecting the keys with the specified Hamming distances from any fixed key K.

Fig. 3 shows the probability distribution of selecting the private keys (an uniform distribution of the keys generation has been assumed) that have a specified Hamming distance from an arbitrarily chosen key K. It is simple to show that the aforementioned distribution is the binomial Bernoulli distribution with the probability of success $1/2$, i.e. it is described by the formula $p_k = \binom{N}{k} \cdot 2^{-N}$, where k is the Hamming distance

and p_k is the probability of selecting a key that is k bits distant. The analysis of p_k distribution reveals that for $l_{key} = 160$ the probability of selecting a key that is distant by 70 to 90 bits from any key K is close to 0.9. This means in general that, when trying to guess the key K, it is highly probable to select a key that is distant from K by a number of bits close to one half of the total length of the key. It should also be noted that for the Bernoulli distribution its expected value and standard deviation are equal to $N/2$ and $\sqrt{N/4}$, respectively (in the considered experiment those values are 80 and ≈ 6).

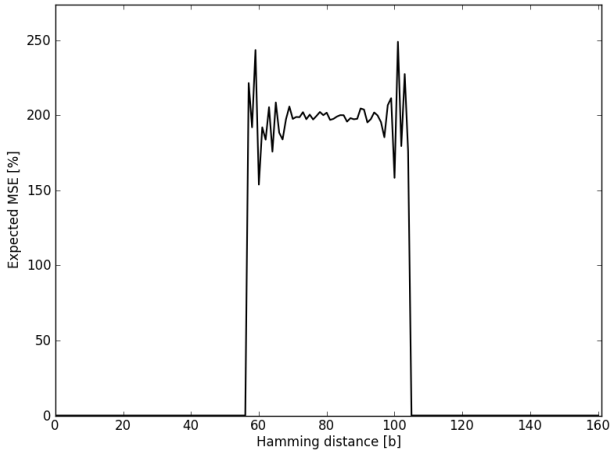


Fig. 4. Probability distribution of the expected value of the mean square error of signal reconstruction as a function of the Hamming distance between private keys.

Fig. 4 presents the results obtained in the second part of the experiment, i.e. the expected values of the MSE as a function of the Hamming distance between keys. The error value is expressed here as a percentage of the total energy of the signal and can take values between 0% and 400%. Based on the resulting plot it can be concluded that the expected value of the MSE in the "experimentally probable" range of the Hamming distance is close to 200%. For distances falling below ≈ 56 and above ≈ 104 bits, the probability of selecting the keys distant by such numbers of bits is so small that in the experiment (10^6 trials) we failed to observe such key pairs. In addition, large fluctuations of the expected value of the MSE observable close to the ends of the mentioned range appear presumably also due to small number of randomly selected pairs of keys distant by such bit values. Fig. 5 presents the probabilities of the MSE of signal reconstruction expressed as a percentage of the total signal energy for the whole possible range of its variation, i.e. from 0% to 400%. The experimentally obtained results show that the most probable value of MSE is close to 200%. In addition, in the total number of 10^6 trials, we can observe MSE values in the range from $\approx 40\%$ up to $\approx 360\%$.

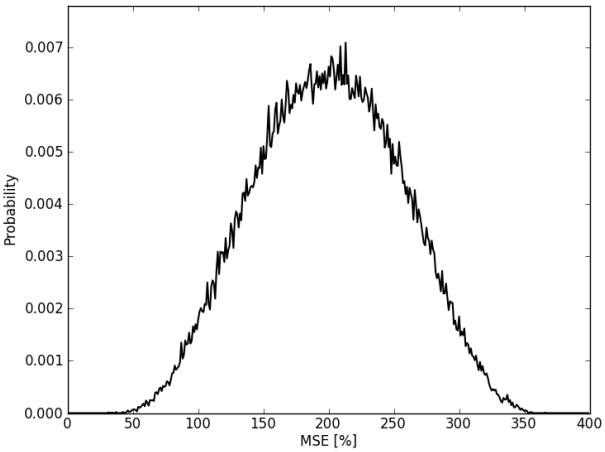


Fig. 5. Probability distribution of MSE of signal reconstruction during random trials of guessing a private key K.

4. Summary and directions for future research

This paper proposes a scheme of data encryption using a private key and parametrized orthogonal transforms. It also provides a way to map the bits of a private key to individual parameters of transformation. The proposed encryption scheme implies the possibility of using fast transforms with computationally efficient structures in the role of encryption and decryption transformations. It should be noted that fast transforms have a relatively large number of parameters even for small transform lengths, which directly translates into large bit-lengths of private keys. In view of the above, and taking into account the results of experimental studies that indicate a very low probability of random guessing of the encryption key and large expected values of signal reconstruction error obtained with different private keys used at the encryption and decryption stages, it can be concluded that the proposed scheme can be very attractive from a practical point of view.

The directions of future research on the proposed encryption scheme may involve a thorough analysis, and in particular the theoretical one, of the following issues:

1. an ambiguous description of U transform which relies on the possibility of finding two or more keys that provide equivalent cryptographic transformations. For example, let us consider the structure shown in Fig. 1. It is well known [11] that such a structure is able to realize any permutation in N - element set. The total number of possible keys depending on the value of N for k_b is equal to $2^{l_{key}} = 2^{N(\log_2 N - 0.5)} = N^N / 2^{N/2}$. However, assuming, e.g. $N = 8$, we receive 1,048,576 possible keys, while 8!=40,320. This means that each private key is accompanied by an average number of 25 equivalent private keys,

2. application of the proposed encryption scheme in conjunction with conventional data compression techniques based on orthonormal transformations. The result of such conjunction will be a fast and efficient algorithm that provudes the possibility of simultaneous encryption and compression of data for real-time multimedia systems. It can be expected, however, that the "requirements" of encryption and compression stages will be mutually exclusive.

Acquiring answers to the questions arising from the above issues will allow us to develope an encryption algorithm that is more efficient and capable of using in conjunction with well-known and popular data compression techniques.

References

[1] L. Krikor, S. Baba, T. Arif, and Z. Shaaban, "Image encryption using DCT and stream cipher", *European Journal of Scientific Research*, vol. 32, № 1, pp. 47-57, 2009.

[2] J. Xie, S. Agaian, and J. Noonan, "Secure information hiding algorithm using parametric Slant-Hadamard transforms", in *Proc. of Mobile Multimedia/Image Processing, Security, and Applications*, vol. 6982, 2008.

[3] A. Pande and J. Zambreno, "The secure wavelet transform", *Journal of Real-Time Image Processing*, Springer, 2010.

[4] S. Agaian, K. Tourshan, and J. P. Noonan, "Parametric Slant-Hadamard transforms with applications", *IEEE Signal Processing Letters*, vol. 9, № 11, pp. 375-377, November 2002.

[5] S. Minasyan, J. Astola, and D. Guevorkian, "On unified architectures for synthesizing and implementa-tion of fast parametric transforms", in *Proc. 5th International Conference, Information, Communication and Signal Processing*, pp. 710-714, December 2005.

[6] S. Bouguezel and M. O. Ahmad, "A new class of reciprocal-orthogonal rarametric transforms", *IEEE Trans. On Circuits and Systems*, vol. 56, № 4, pp. 795-805, April 2009.

[7] "Data encryption standard", FIPS PUB 46, National Bureau of Standards, Jan. 1997.

[8] X. Lai, *On the design and security of block cipher*, Konstanz, Germany: Hartung-Gorre, 1992.

[9] R. Anderson, E. Biham, and L. Knudsen, "Serpent: A Proposal for the Advanced Encryption Standard", AES submission, 1998.

[10] M. M. Yatsymirskyy and R. I. Liskevytch, "Lattice structures for Fourier, Hartley, cosine and sine transformations", *Modeling and Information Technolo-gies*, Ukrainian Academy of Sciences, vol. 2, pp. 173-181, 1999. (Ukrainian)

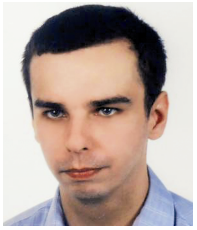
[11] M. M. Yatsymirskyy, "Encryption on the base of FFT algorithm graph", *Journal of East Ukrainian National University*, № 9, pp. 24-29, 2010. (Ukrainian)

ПАРАМЕТРИЗОВАНІ ОРТОГОНАЛЬНІ ПЕРЕТВОРЕННЯ ДЛЯ ШИФРУВАННЯ ДАНИХ

Даріуш Пухала, Каміль Стокфішевскі

У даній роботі пропонується схема шифрування й дешифрування даних, яка використовує переваги швидкого параметричного ортогонального перетворення. Сформульовано спосіб відображення секретного ключа до значень параметрів перетворення. Ефективність запропонованої схеми підтверджено експериментально. Крім того, автори пропонують напрямки подальших досліджень і розвитку розглянутої схеми шифрування даних.

Dariusz Puchala – Ph.D. in Com-puter Science, employee at the Institute of Information Technology, Technical Uni-versity of Lodz, Poland. His main fields of scientific interests include: digital signal and image processing, mathemati-cal foundations of data compression, adaptive fast Fourier-type and fast para-metric linear transformations, problems of parallel and distributed computing.



Kamil Stokfiszewski – Ph.D. in Computer Science, employee at the Insti-tute of Information Technology, Tech-nical University of Lodz, Poland. His scientific interests include: digital signal and image processing, image compres-sion, artificial neural networks for signal processing, parallel computing.