Spectral and Nonlinear Properties of the Sum of Boolean Functions

O.N. Zhdanov¹, A.V. Sokolov²

¹Siberian State University of Science and Technology named after Academician M.F.Reshetnev ²Odessa National Polytechnic University

onzhdanov@mail.ru

Abstract— Boolean functions are the mathematical basis of modern cryptographic algorithms. However, in practice, a set of interrelated Boolean functions is often used to construct a cryptographic algorithm. This circumstance makes the task of research of cryptographic quality, in particular, the distance of the nonlinearity of the sum of few Boolean functions important. The nonlinearity distance of a Boolean function is determined by the maximum value of its Walsh-Hadamard transform coefficients. In this paper, we proposed a formula that is the equivalent of the summation of Boolean functions in the Walsh-Hadamard transform domain. The application of this formula, as well as the Walsh-Hadamard spectral classification made it possible to determine the structure of Walsh-Hadamard transform coefficients, and the distance of the nonlinearity when summing the Boolean functions lengths N=8 and N=16, indicating valuable practical application for information protection.

Index Terms — Boolean Function; Walsh-Hadamard Trans-Form; Distance of Nonlinearity; Cryptography.

I. INTRODUCTION AND BASIC DEFINITIONS

Design of a modern telecommunication system is unthinkable without the use of access control technology and information protection. Modern information security technologies are largely based on the use of cryptographic algorithms, which must be constantly improved in order to be resistant to the modern types of cryptanalysis attacks.

Further development of cryptographic algorithms and therefore, the security of modern telecommunication systems today is largely depending on the further improvement of the mathematical apparatus of Boolean functions in the sense of increasing their cryptographic quality.

Boolean function is the most important object in modern cryptography. Boolean functions have found their numerous applications in the construction of cryptographic S-boxes, which are the main component of modern block symmetric cryptographic algorithms [1]. Boolean functions are also a main component of pseudo-random key sequence generators, which form the basis of modern stream encryption algorithms [2]. In the cryptographic transforms, a set of Boolean functions that satisfies certain quality criteria is used. Often the quality of each of the components of the set is insufficient, and it is necessary to consider their compatibility [3].

One of the main criteria for the cryptographic quality of Boolean functions is the distance of nonlinearity [4][5][6]. There are several ways to measure the distance of nonlinearity of a Boolean function, one of which is a spectral method based on the research of Walsh-Hadamard transform coefficients. The spectral coefficients of the Walsh-Hadamard transform $W_f(\omega)$ of a Boolean function

f(x) of k variables can be represented in the matrix form

$$W_f(\omega) = fA_N, \ \omega = 0, 1, ..., N-1,$$
 (1)

where, A_N is the Walsh-Hadamard matrix of order $N = 2^k$, which is constructed in accordance with the recurrent rule [7]

$$\mathbf{A}_{2^{k}} = \begin{bmatrix} A_{2^{k-1}} & A_{2^{k-1}} \\ A_{2^{k-1}} & -A_{2^{k-1}} \end{bmatrix}.$$
 (2)

To estimate the distance of nonlinearity of Boolean function we have the formula

$$N_{f} = 2^{k-1} - \frac{1}{2} \max_{\omega \in \mathbb{Z}_{2}^{k}} |W_{f}(\omega)|.$$
(3)

Note that for the Walsh-Hadamard transform coefficients (1), Parseval's equality is valid [8]

$$\sum_{\omega \in \mathbb{Z}_{2}^{k}} \left(W_{f}\left(\omega\right) \right)^{2} = 2^{2k} .$$

$$\tag{4}$$

Since the number of all coefficients is equal to 2^k , it follows from the equality that the maximum absolute value of the Walsh-Hadamard coefficient cannot be smaller than the value $2^{\frac{k}{2}}$. It is established that the minimum value of Walsh-Hadamard transform coefficients is equal to $2^{\frac{k}{2}}$ only if the absolute values of all Walsh-Hadamard transform coefficients are equal to each other. In this case, the distance of nonlinearity (3) is maximal, the Boolean function is called a bent-function, and its truth table is a bent-sequence [9]. Bent-functions, as nonlinearity nonpareil of the entire set of Boolean functions, have found their application in the form of the main components of many modern systems of information transmission and processing. High values of the non-linearity of Boolean bent-functions, and hence their maximum distance from the set of affine functions, make them widely used in modern cryptographic algorithms and their components [10,11].

The uniformity of the absolute values of the Walsh-Hadamard spectral coefficients causes the use of bentfunctions in systems with code division multiple access (CDMA), allowing to reduce the PAPR (Peak-to-Average Power Ratio) of the signals transmitted in the system to the lowest possible value [12,13].

However, the extreme nonlinearity of bent-functions has a reverse side. Here is an example showing the existence of a bent-function application problem, and thereby confirming the actuality of the topic. We choose bent-sequences of the length N = 16

Boolean functions (5) have a uniform spectrum $W_{B1,B2} = \{\pm 4\}$, and accordingly, the maximum possible distance of nonlinearity $N_f = 6$. We calculate the sum of two bent-functions (5)

Both functions (5) have high cryptographic quality, but their sum (6), as it is easy to see, has a nonlinearity distance $N_f = 0$, hence it is inadmissible to use it in nonlinear cryptographic applications. Along with such a (negative) example, it is easy to show the example of Boolean functions, the sum of which is of high quality.

There are many modern researches devoted to the problems of using certain Boolean function as a single unit. In practice, cryptographic Boolean functions are rarely used as single units. This makes the problem of their compatibility in terms of maintaining a high level of nonlinearity especially relevant. However, this problem has been insufficiently researched in the literature. Thus, a complete understanding of the nonlinear properties of the sum of two Boolean functions is needed.

The purpose of this paper is to research the dependence of the spectral properties of the sum of two Boolean functions from the spectral properties of the summands.

II. AN ANALOGUE OF THE SUMMATION OPERATION IN THE DOMAIN OF THE WALSH-HADAMARD TRANSFORM COEFFICIENTS

The structure of the spectral vector $W_f(\omega)$ is rigidly determined by the properties of the Walsh-Hadamard transform. For example, for a Boolean function $f_{155} = [--+--++-]$ the spectral vector has the form $W_{155} = [-2 \ 2 \ -2 \ -6 \ -2 \ 2 \ -2 \ 2]$.

Thus, the problem of determining the nonlinearity of the sum of Boolean functions is equivalent to the problem of finding its Walsh-Hadamard transform coefficients. Despite the ubiquitous use of the summation operation modulo 2, the equivalent of this operation in the domain of Walsh-Hadamard transform coefficients is unknown in the literature today. In this case, this paper proposed an investigation of the behavior of the coefficients of the Walsh-Hadamard transform, while adding Boolean functions modulo 2. Firstly, we consider the Boolean functions of length N = 2, and then gradually increase the length in order to prepare a transition to the general formula for summing the Walsh-Hadamard spectral coefficients.

Therefore, we let the two Boolean functions $f = \{f_1 \ f_2\}$ and $g = \{g_1 \ g_2\}$ to be given. Then, in accordance with (1.1), we write down their Walsh-Hadamard transform coefficients

$$W_{f} = [f_{1} \quad f_{2}] \begin{bmatrix} + & + \\ + & - \end{bmatrix} =$$

$$= \{f_{1} + f_{2} \quad f_{1} - f_{2}\} = \begin{bmatrix} w_{f1} & w_{f2} \end{bmatrix};$$

$$W_{g} = [g_{1} \quad g_{2}] \begin{bmatrix} + & + \\ + & - \end{bmatrix} =$$

$$= \{g_{1} + g_{2} \quad g_{1} - g_{2}\} = \begin{bmatrix} w_{g1} & w_{g2} \end{bmatrix}.$$
(7)

Similarly, we can write the Walsh-Hadamard transform coefficients for the sum of the considered functions $f \oplus g$ above an alphabet $\{0,1\}$ that is equivalent to their product fg over an alphabet $\{+1,-1\}$

$$W_{fg} = \begin{bmatrix} f_1 g_1 & f_2 g_2 \end{bmatrix} \begin{vmatrix} + & + \\ + & - \end{vmatrix} =$$

$$= \begin{bmatrix} f_1 g_1 + f_2 g_2 & f_1 g_1 - f_2 g_2 \end{bmatrix} = \begin{bmatrix} w_{fg1} & w_{fg2} \end{bmatrix}.$$
(8)

From the comparative analysis of (7) and (8), it is not difficult to express the coefficients of the Walsh-Hadamard transform of the sum of two Boolean functions in terms of the transform coefficients of each of them

$$W_{fg} = \begin{bmatrix} w_{fg1} & w_{fg2} \end{bmatrix} = = \frac{1}{2} \begin{bmatrix} w_{f1}w_{g1} + w_{f2}w_{g2} & w_{f1}w_{g2} + w_{f2}w_{g1} \end{bmatrix}.$$
 (9)

Now, we can carry out similar research for Boolean functions of length N = 4. Let there be two given Boolean functions $f = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 \end{bmatrix}$ and $g = \begin{bmatrix} g_1 & g_2 & g_3 & g_4 \end{bmatrix}$. Then, in accordance with (1), we can write down their Walsh-Hadamard transform coefficients

In a similar way, we can write down the Walsh-Hadamard transform coefficients of the sum of two Boolean functions:

$$W_{fg} = \begin{bmatrix} fg_1 & fg_2 & fg_3 & fg_4 \end{bmatrix} \begin{vmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{vmatrix} = \\ = \begin{bmatrix} fg_1 + fg_2 + fg_3 + fg_4 \\ fg_1 - fg_2 + fg_3 - fg_4 \\ fg_1 + fg_2 - fg_3 - fg_4 \\ fg_1 - fg_2 - fg_3 - fg_4 \end{bmatrix}^T .$$
(11)

Comparing expression (10) and (11), we can express the spectral coefficients of the new Boolean function fg in terms of the spectral coefficients of the original functions:

$$W_{fg} = \begin{bmatrix} w_{fg1} & w_{fg2} & w_{fg3} & w_{fg4} \end{bmatrix} = \begin{bmatrix} w_{f1}w_{g1} + w_{f2}w_{g2} + w_{f3}w_{g3} + w_{f4}w_{g4} \\ w_{f1}w_{g2} + w_{f2}w_{g1} + w_{f3}w_{g4} + w_{f4}w_{g3} \\ w_{f1}w_{g3} + w_{f2}w_{g4} + w_{f3}w_{g1} + w_{f4}w_{g2} \\ w_{f1}w_{g4} + w_{f2}w_{g3} + w_{f3}w_{g2} + w_{f4}w_{g1} \end{bmatrix}^{T}.$$
(12)

Thus, formula (12) connects the coefficients of the Walsh-Hadamard transform of the sum modulo 2 of two Boolean functions and the Walsh-Hadamard transform coefficients of each of the Boolean functions separately. In [14], a fundamental dyadic shift operator was derived, which, as shown by research, it can be applied to simplify and generalize formula (9) and (12).

Definition [14]. The operator of the dyadic shift (dyadic permutation) is represented by the following construction:

$$\mathbf{Dyad}(N) = \begin{bmatrix} \mathbf{Dyad}(N/2), & \mathbf{Dyad}(N/2) + N/2 \\ \mathbf{Dyad}(N/2) + N/2, & \mathbf{Dyad}(N/2) \end{bmatrix}, \quad (13)$$

where **Dyad**(2) = $\begin{bmatrix} 1, 2\\ 2, 1 \end{bmatrix}$.

Thus, in accordance with (13), for the order N = 4, we have the following dyadic shift matrix:

$$\mathbf{Dyad}(4) = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix}.$$
 (14)

It is easy to see that the indices w_{gi} from (12) completely coincide with the dyadic permutations (14). Thus, we can write down the general formula for the spectrum of the sum of the two Boolean functions.

Proposition. The transform coefficients of the sum of the two Boolean functions can be represented in terms of the transform coefficients of each of them, as follows:

$$w_{fg\ k} = \frac{1}{N} \sum_{i=1}^{N} w_{f\ i} w_{g\ \mathbf{Dyad}(N)_{k,i}},$$
 (15)

or, equivalently, in a matrix form:

$$W_{fg} = \frac{1}{N} W_f \cdot W_g \left(\mathbf{Dyad}(N) \right), \qquad (16)$$

where the parentheses denote the application of the dyadic shift operator (13) to the transform coefficients. For all possible pairs of the Boolean functions of two, three, and four variables, the validity of the formula (16) is established with the help of direct calculations.

We note the practical significance of the case of Boolean functions of four variables. Thus, in the algorithm GOST 28147-89, which is considered to be cryptographically strong now, the substitution boxes are constructed from Boolean functions of four variables. Earlier, in [3], a reasonable method for selecting such substitution boxes was proposed. The method was implemented in software, and it was justified in terms of the nonlinearity of Boolean functions.

We also note that formula (16) allows us to find the Walsh-Hadamard transform coefficients of the sum of any number of Boolean functions. Thus, the application of formula (16) greatly simplifies the construction of substitution boxes for the algorithm GOST 28147-89, and with simple changes, for other algorithms of this class.

It seems important to us to prove algebraically formula (16) for an arbitrary number of variables. However, it is beyond the scope of this paper.

Let us consider an example. Let there be given two Boolean functions of length N = 16

In accordance with (1), we find their Walsh-Hadamard transform coefficients

$$W_{f} = \begin{bmatrix} -6 - 6 \ 2 \ 2 - 2 \ 6 - 2 \ 6 \\ -2 \ 6 \ 6 - 2 \ 2 \ 2 \ 2 \ 2 \end{bmatrix};$$
(18)
$$W_{g} = \begin{bmatrix} 4 \ 0 \ 8 - 4 \ 0 \ 4 \ 4 \ 0 \\ 0 - 4 - 4 \ 0 \ 4 \ 8 \ 0 - 4 \end{bmatrix}.$$

In accordance with (16), we find the spectrum of a Boolean function $f \oplus g$



Calculating the coefficients of the Walsh-Hadamard transform of the sum of two Boolean functions provides the same result:

$$W_{fg} = (f \oplus g) \rightarrow \{+1, -1\} \cdot H_{16} =$$

= [+--++++---+--] \cdot H_{16} = (20)
= [-2 2 2 -2 -2 2 -6 6 6 2 2 6 -2 -6 2 6].

III. NONLINEARITY OF THE SUM OF TWO BOOLEAN FUNCTIONS AND THE RESULTS OF MODELING

The results of the experiments carried out in [15] made it possible to establish that the entire set of Walsh-Hadamard transform vectors is divided into classes depending on the elementary structure of the vector.

Definition [15]. The elementary structure of a spectral vector W_i is the set of absolute values of its spectral components.

For example, the vector W_{155} considered above has an elementary structure $\{6(1), 2(7)\}$, where the number in parentheses indicates the number of times the specified spectral component appears in the spectral vector.

Thus, for N = 8, a set of vectors W_i , $i = 0, 1, ..., 2^8 - 1$ is divided into three equivalent classes (Table 1).

Table 1 The classification of the complete set of spectral vectors W_i of length N = 8

Number of	A set of absolute	Distance of	The
a class of	values of the	nonlinearity	cardinality
spectral	spectral		of class
vectors	components		
1	$\{\pm 8(1), 0(7)\}$	0	16
2	$\{\pm 6(1), \pm 2(7)\}$	1	128
3	$\{\pm 4(4), 0(4)\}$	2	112

Thus, when summing two Boolean functions f and g of the length N = 8, the resulting Boolean function $h = f \oplus g$ will belong to one of the classes, as indicated in Table 1. In this case, the dependence of the spectral class (and, respectively, the value of nonlinearity) of the resulting Boolean function from the summands is established, which is shown in the form of a table of class numbers.

Table 2 The result of addition of classes of spectral vectors of Boolean functions of length N = 8

length $N = 8$					
Œ)	1	2	3	
1		1	2	3	
2		2	1,3	2	
3		3	2	1,3	

In [15], a classification of spectral vectors W_i , $i = 0, 1, ..., 2^{16} - 1$ of length N = 16, was also carried out, resulting in the formation of eight equivalent classes, as shown in Table 3.

Table 3 Classification of the complete set of spectral vectors W_i of length N = 16

Number of a class of spectral	A set of absolute values of the spectral components	Distance of nonlinearity	The cardinality of class	
vectors	I I I I I			
1	{16(1),0(15)}	0	32	
2	{14(1), 2(15)}	1	512	
3	$\{12(1), 4(7), 0(8)\}$	2	3840	
4	{10(1), 6(3), 2(12)}	3	17920	
5	$\{8(2), 4(8), 0(6)\}$	4	26880	
6	$\{8(4), 0(12)\}$	4	1120	
7	$\{6(6), 2(10)\}$	5	14336	
8	{4(16)}	6	896	

It is not difficult to find the addition table of classes of vectors of length N = 16 (Table 4), based on the conduct of the computational experiments in accordance with (2.10),

Table 4 The result of classes addition of Boolean functions spectral vectors of length N = 16

\oplus	1	2	3	4	5	6	7	8
1	1	2	3	4	5	6	7	8
2	2	1,3	2,4	3,5,6	4	4,7	6, 8	3,5,6
3	3	2,4	1,3,	2,4,7	3,6	3,6,	4,7	6
			5,6			5, 8		
4	4	3,5,	2,4,	1,3,6,	2,4,7	2,4,7	3,6, 8	7,4
		6	7	5, 8				
5	5	4	3,6	2,4,7	1,5,8	3,6	4,7	8 ,5
6	6	4,7	3,6,	2,4,7	3,6	1,3,5,	2,4,7	6,3
			5, 8			6, 8		
7	7	6,8	4,7	3,6, 8	4,7	2,4,7	1,3,5,	4,
							6, 8	7,2
8	8	3,	6	7,4	8,5	6,3	4,7,2	1,
		5,6						5, 8

We especially note that, in accordance with (1.3), the nonlinearity of a Boolean function is completely determined by the maximum value of its Walsh-Hadamard transform coefficients. Thus, drawn from Table 2 and Table 4, we can have a priori of the information about the nonlinearity of the resulting Boolean function when two arbitrary Boolean functions are added.

Let us consider an example: Let the Boolean functions of length N = 16, which belong to the spectral class No. 5 and No. 7 (Table 3), corresponding to the values of nonlinearity 4 and 5, respectively. As a result, in accordance with Table 4, we can get a new Boolean function whose elementary structure belongs to class No. 4 or No. 7, which corresponds to a nonlinearity distance value of 3 or 5.

Thus, Table 2 and Table 4 give complete information about the possible values of the nonlinearity of the resulting Boolean function drawn from the form of the elementary structure of the Walsh-Hadamard transforms of the summands.

IV. CONCLUSION

On the basis of the use of the regular dyadic shift operator, the formula for summing the Walsh-Hadamard transform coefficients is obtained. This formula is the direct equivalent in the Walsh-Hadamard transform domain of the addition operation modulo two in the time domain.

Computational experiments were carried out for practically significant lengths of Boolean functions N = 8 and N = 16 made it possible to understand how the elementary structure of Walsh-Hadamard transform coefficients is changed during summation of Boolean functions. Subsequently, it leads to the understanding of the changes in the distance of the nonlinearity of Boolean functions when summing them.

Considering it is common to use several Boolean functions simultaneously in practical applications of information protection, the obtained results are of practical interest from the point of view of choosing compatible constructions of Boolean functions to preserve their overall level of the nonlinearity distance.

REFERENCES

- B. Schneier, Applied Cryptography. Second edition, New York, John Wiley & Sons(1996).
- [2] I.V. Agafonova, Cryptographic properties of non-linear Boolean functions, Seminar on discrete harmonic analysis and geometeric modeling, St. Petersburg DHA & CAGD (2007), 1–24.
- [3] O.N. Zhdanov, The method of selection of key information for the block cipher algorithm, Moscow, INFRA-M(2013), 90.

- [4] Rodier, F. (2003). On the nonlinearity of Boolean functions. In Proceedings of WCC2003, Workshop on coding and cryptography (pp. 397-405).
- [5] A.V. Sokolov, New methods for synthesizing nonlinear transform of modern ciphers, Lap Lambert Academic Publishing (2015), 100.
- [6] Canteaut, A., Carlet, C., Charpin, P., & Fontaine, C. (2000, May). Propagation characteristics and correlation-immunity of highly nonlinear Boolean functions. In International Conference on the Theory and Applications of Cryptographic Techniques (pp. 507-522). Springer, Berlin, Heidelberg.
- [7] M.I. Mazurkov, Broadband radio communication systems, Odessa, Science and Technology(2010), 340.
- [8] M.I. Mazurkov, A.V. Sokolov, The regular rules of constructing the complete class of bent-sequences of length 16, Proceedings of ONPU (2013), No.2(41), 231–237.
- [9] O.S. Rothaus, On "bent" functions J. Comb. Theory Ser. A, USA: Academic Press Inc (1976), No.20(3), 300–305.
- [10] A.V. Sokolov, Constructive method for the synthesis of nonlinear Sboxes satisfying the strict avalanche criterion, Radioelectronics and Communications Systems(2013), 56, 8, 415–423.
- [11] M.I. Mazurkov, N.A. Barabanov, A.V. Sokolov, The key sequences generator based on bent functions dual couples, Proceedings of ONPU(2013), No.3(42), 150–156.
- [12] K. G. Paterson, Sequences For OFDM and Multi-code CDMA: two problems in algebraic coding theory, Sequences and their applications. Seta(2001,2002). Second Int. Conference. Proc. Berlin: Springer, 46–71.
- [13] Mazurkov M.I. Synthesis method for families of constant amplitude correcting codes based on an arbitrary bent-square / M.I. Mazurkov, A.V. Sokolov, I.V. Tsevukh. — Journal of Telecommunication, Electronic and Computer Engineering (JTEC). — Vol. 2. — No.9. — P. 99-103.
- [14] M.I. Mazurkov, A.V. Sokolov, Fast orthogonal transforms based on bent-sequences, Informatics and mathematical methods in simulation(2014), No. 1, 5–13.
- [15] A.V. Sokolov, N.A. Barabanov, Algorithm for removing the spectral equivalence of component Boolean functions of Nyberg-design Sboxes. Radioelectronics and Communications
- [16] Systems(2015), Vol. 58, Vo. 5, 220-227.