Tatra Mt. Math. Publ. **29** (2004), 69–77

TATRA MOUNTAINS
Mathematical Publications

A LOWER BOUND OF THE NUMBER OF HIGHLY NONLINEAR FUNCTIONS

ILDIKÓ LICSKÓ

ABSTRACT. This study provides an estimation for the lower bound of the number of highly nonlinear n-dimensional functions if n is odd. There is a relationship between highly nonlinear odd dimensional functions and bent functions in $(n+1)$-dimension, consequently a lower bound for $(n+1)$-dimensional bent functions is also given.

Introduction

Different types of ciphers use Boolean functions, LFSR based stream ciphers use them as a nonlinear combiner or a nonlinear filter. They appear in block ciphers in substitution boxes also. The design criteria for block ciphers correspond to some mathematical properties of the round function, which is used in an iterated block cipher. Boolean functions used in ciphers must satisfy some properties to resist different attacks. In particular, the use of highly nonlinear functions ensures a high resistance to linear attacks. The functions with maximal nonlinearity are called bent functions. They only exist in even number of variables. Despite of their good nonlinearity property, they are not used in cryptographic practice. The highly nonlinear functions of odd dimensional Boolean-space are preferred to bent functions since they have better cryptographic properties. A further advantage of highly nonlinear functions is that they guarantee a good resistance to the differential cryptanalysis.

1. Background

A mapping $f\colon \{0,1\}^n \to \{0,1\}$ is called a Boolean-function. Sometimes the mapping $\overline{f}\colon \{0,1\}^n \to \{-1,1\}$ is used instead of the Boolean-function. The

2000 Mathematics Subject Classification: 94A60, 06E30.
Keywords: S-box, nonlinearity of Boolean function, bent function.

relation between $\overline{f}$ and f can be described as follows:

$$\overline{f}(x) = (-1)^{f(x)}$$

or

$$\overline{f}(x) = 1 - 2f(x)\,.$$

In the following we use the notation $\overline{f}(x)$ to denote the whole $\{-1,1\}$ sequence generated by $\overline{f}(x)$, that means $\overline{f}(x)$ can be regarded as a vector having 2^n elements. The elements of set $\{0,1\}^n$ form a vector space and it is called Boolean-space. As the coordinates of these vectors are 0, 1 numbers, a vector can be regarded as an integer written in binary form. We can refer to a vector by an indexed name where the index is the integer corresponding to the vector.

The *weight* of a function $f(x)$ means the number of 1s in its truth-table:

$$w(f) = \sum_{x\in\{0,1\}^n} f(x)\,.$$

Function $f(x)$ is called *balanced* if the number of 1s and the number of 0s in its truth-table are equal, that is $\sum\limits_{x\in\{0,1\}^n} \overline{f}(x) = 0$, and the weight of a balanced function is $w(f) = 2^{n-1}$.

Ordinary operations are defined by components in $GF(2)$. If $a, b \in \{0,1\}^n$, $a = (a_0, a_1, \ldots, a_{n-1})$, $b = (b_0, b_1, \ldots, b_{n-1})$, the addition is $a \oplus b = (a_0 \oplus b_0,\ a_1 \oplus b_1, \ldots, a_{n-1} \oplus b_{n-1})$ and the scalar product is $ab = \bigoplus\limits_{i=0}^{n-1} a_i b_i$.

In the case when the Zhegalkin-polynom of a Boolean-function can be written in the form

$$f(x_0, x_1, \ldots, x_{n-1}) = a_0 x_0 \oplus a_1 x_1 \oplus \cdots \oplus a_{n-1} x_{n-1} \oplus c$$

with $a_i, c \in \{0,1\}$, $0 \le i \le n-1$, the function is called *affine* if $c = 1$ and *linear* if $c = 0$. A linear function can be regarded as a scalar product of a constant $a \in \{0,1\}^n$ and the x variable.

The distance between two Boolean-functions f and g is

$$d(f,g) = w(f \oplus g) = \sum_{x\in\{0,1\}^n} \big(f(x) \oplus g(x)\big)\,,$$

while their correlation is defined as:

$$c(f,g) = \sum_{x\in\{0,1\}^n} (-1)^{f(x)\oplus g(x)} = 2^n - 2d(f,g)\,.$$

An m-order matrix is called *Hadamard-matrix* when $H = (h_{i,j})$, $h_{ij} \in \{-1,1\}$ $i,j = 0,1,\ldots,m-1$, $HH^T = mI$, H^T means the transposed matrix of the H matrix, and I is the identity matrix of order m. Let H_n denote the Hadamard-matrix of order 2^n generated by the following recursive process:

$$H_0 = 1, \quad H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}, \qquad n = 1, 2, \ldots$$

The rows of H_n are denoted by l_i, $i = 0, 1, \ldots, 2^{n-1}$. Row l_i can be regarded as a $\{-1, 1\}$ sequence of $\overline{g}_i(x)$, where $g_i(x) = ix$ is a linear function. The generated sequences of all linear functions in the n-dimensional Boolean space appear in the rows of H_n.

The Walsh-transformed form of an f function in the position $a = (a_0, a_1, \ldots, a_{n-1}) \in \{0, 1\}^n$ is

$$F(a) = \sum_{x \in \{0,1\}^n} f(x)(-1)^{ax},$$

while the Walsh-transformed form of $\overline{f}$ is specified as:

$$\overline{F}(a) = \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus ax}.$$

The integer value-set of $\{\overline{F}(a)\}$ for $a = 0, 1, \ldots, 2^n - 1$ is often referred to as the *Walsh spectrum* of f. The value $\overline{F}(a)$ is the correlation of function f with the linear function ax and it can be written as the scalar product $(\overline{f}(x) l_a)$.

A Boolean-function f is called *bent function*, if for all $a \in \{0, 1\}^n$ the elements of its Walsh spectrum have the constant absolute value and

$$\left|\overline{F}(a)\right| = \left| \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus ax} \right| = 2^{\frac{n}{2}}.$$

They exist in even-dimensional space only.

The non-linearity of a function f is the distance between f and the set of the affine functions denoted by N_f:

$$N_f = 2^{n-1} - \tfrac{1}{2} \max_{i \in \{0,1\}^n} \left(|\overline{f}(x) l_i| \right).$$

The minimum value of nonlinearity is 0, this is the nonlinearity of the affine functions, its maximum value can be reached when $\left|\overline{f}(x) l_i\right|$ is a constant and its value is: $2^{\frac{n}{2}}$. The bent functions have the maximum nonlinearity value, namely $N_f = 2^{n-1} - 2^{\frac{n}{2}-1}$. Despite of their good nonlinearity the bent functions are not preferred for cryptographic applications because they are never balanced or correlation-immune.

A Boolean-function $f\colon \{0, 1\}^n \to \{0, 1\}$ is called *highly nonlinear function* if n is odd and its Walsh-spectrum consists of 0 and $\pm C$ values and the number of non-zero elements is 2^{n-1} and C has the absolute value of $2^{\frac{n+1}{2}}$.

For cryptographic applications the highly nonlinear functions of odd dimensional Boolean-space are preferred to bent functions since they have better cryptographic properties. On the basis of [2], pairs of highly nonlinear functions can be formed so that the correlation of the members with a given linear function never takes the same absolute value. In this case we say that the set of linear functions having non-zero correlation with the first member is disjoint from the set of linear functions having non-zero correlation with the second member of the pair. Expressing it in formulas: n is odd, g_1 and $g_2\colon \{0,1\}^n \to \{0,1\}$ are highly nonlinear functions. g_1 and g_2 form a pair just described if and only if

$$\big(\overline{g}_1(x)l_i\big)^2 + \big(\overline{g}_2(x)l_i\big)^2 = 2^{n+1}$$

for each $i = 0, 1, \ldots, 2^n - 1$.

When we have such pairs of highly nonlinear functions it is possible to generate further pairs of highly nonlinear functions fulfilling the conditions shown before. The following functions

$$\overline{h}_1(i) = \frac{\big(\overline{g}_1(x)l_i\big) + \big(\overline{g}_2(x)l_i\big)}{2^{\frac{n+1}{2}}}$$

and

$$\overline{h}_2(i) = \frac{\big(\overline{g}_1(x)l_i\big) - \big(\overline{g}_2(x)l_i\big)}{2^{\frac{n+1}{2}}}$$

defined for each $i = 0, 1, \ldots, 2^n - 1$ result in two $\{-1, 1\}$ sequences belonging to a pair of highly nonlinear functions $h_1, h_2\colon \{0,1\}^n \to \{0,1\}$ that satisfy the conditions: $\big(\overline{h}_1(x)l_i\big)^2 + \big(\overline{h}_2(x)l_i\big)^2 = 2^{n+1}$.

Furthermore the concatenation of the truth-tables belonging to the members of such a function-pair generates the truth-table of a bent function in the Boolean-space of $(n+1)$ dimension and the truth-table of any bent function in the $(n+1)$ dimensional Boolean-space can be cut into two parts being truth-tables of highly nonlinear functions that form such pair in the n dimension.

2. Estimation of the number of functions

In the 3-dimensional Boolean-space one can generate highly nonlinear functions by a very simple process.

The truth-tables of the linear functions in 3-dimension are:

	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$0x$	0	0	0	0	0	0	0	0
$1x$	0	1	0	1	0	1	0	1
$2x$	0	0	1	1	0	0	1	1
$3x$	0	1	1	0	0	1	1	0
$4x$	0	0	0	0	1	1	1	1
$5x$	0	1	0	1	1	0	1	0
$6x$	0	0	1	1	1	1	0	0
$7x$	0	1	1	0	1	0	0	1

There are 8 linear functions and let us consider one of them and make a change of two function-values in its truth-table. This process results highly non-linear functions. The situation is the same when we alter 6 function-values in the truth-table of the same linear function. The result is almost the same when we alter 4 function-values, but in this case we get linear functions also.

The modified truth-table of the linear function $0x$ and the correlation of the resulted function with linear functions are shown in the table below.

Source lin. fun	Modified function								Correlation with							
	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	l_0	l_1	l_2	l_3	l_4	l_5	l_6	l_7
$0x$	1	1	0	0	0	0	0	0	4	0	-4	0	-4	0	-4	0
$0x$	1	0	1	0	0	0	0	0	4	-4	0	0	-4	-4	0	0
$0x$	1	0	0	1	0	0	0	0	4	0	0	-4	-4	0	0	-4
$0x$	1	0	0	0	1	0	0	0	4	-4	-4	-4	0	0	0	0
$0x$	1	0	0	0	0	1	0	0	4	0	-4	0	0	-4	0	-4
$0x$	1	0	0	0	0	0	1	0	4	-4	0	0	0	0	-4	-4
$0x$	1	0	0	0	0	0	0	1	4	0	0	-4	0	-4	-4	0
$0x$	0	1	1	0	0	0	0	0	4	0	0	4	-4	0	0	4
$0x$	0	0	1	1	1	1	1	1	-4	0	4	0	4	0	4	0
$0x$	0	1	0	1	1	1	1	1	-4	4	0	0	4	4	0	0
$0x$	0	1	1	0	1	1	1	1	-4	0	0	4	4	0	0	4
$0x$	0	1	1	1	0	1	1	1	-4	4	4	4	0	0	0	0
$0x$	0	1	1	1	1	0	1	1	-4	0	4	0	0	4	0	4
$0x$	0	1	1	1	1	1	0	1	-4	4	0	0	0	0	4	4
$0x$	0	1	1	1	1	1	1	0	-4	0	0	4	0	4	4	0
$0x$	1	0	0	1	1	1	1	1	-4	0	0	-4	4	0	0	-4
$0x$	0	0	0	0	1	1	1	1	0	0	0	0	8	0	0	0
$0x$	0	0	0	1	0	1	1	1	0	4	4	0	4	0	0	-4
$0x$	0	0	0	1	1	0	1	1	0	0	4	-4	4	4	0	0
$0x$	0	0	0	1	1	1	0	1	0	4	0	-4	4	0	4	0
$0x$	0	0	0	1	1	1	1	0	0	0	0	0	4	4	4	-4
$0x$	0	0	1	0	0	1	1	1	0	0	4	4	4	-4	0	0

Doing the changes for all linear functions in all possible way we get all of the highly nonlinear functions with their negated version and the linear function more than once. When we select the different highly non-linear functions we get 112 pieces of them, so it is the total number of 3-dimensional highly nonlinear functions. This number contains the negated versions also.

These functions can be ordered in 14 groups, from which 7 pairs of function-groups can be formed. All of the couples have the following property. Let g_1 be a function in the first member and let g_2 be a function from the second member of the couple, the correlation of g_1 with a given linear function and the correlation of g_2 with the same linear function never takes the same absolute value.

A couple of function-sets is shown below.
The first member of the couple is:

Source lin. fun	Modified function									Correlation with							
	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7		l_0	l_1	l_2	l_3	l_4	l_5	l_6	l_7
$0x$	0	0	1	1	1	1	1	1		-4	0	4	0	4	0	4	0
$0x$	1	1	0	0	1	1	1	1		-4	0	-4	0	4	0	-4	0
$0x$	1	1	1	1	0	0	1	1		-4	0	4	0	-4	0	-4	0
$0x$	1	1	1	1	1	1	0	0		-4	0	-4	0	-4	0	4	0
$0x$	1	1	0	0	0	0	0	0		4	0	-4	0	-4	0	-4	0
$0x$	0	0	1	1	0	0	0	0		4	0	4	0	-4	0	4	0
$0x$	0	0	0	0	1	1	0	0		4	0	-4	0	4	0	4	0
$0x$	0	0	0	0	0	0	1	1		4	0	4	0	4	0	-4	0

The second member of the couple is:

Source lin. fun	Modified function									Correlation with							
	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7		l_0	l_1	l_2	l_3	l_4	l_5	l_6	l_7
$1x$	0	1	1	0	1	0	1	0		0	-4	0	4	0	4	0	4
$1x$	1	0	0	1	1	0	1	0		0	-4	0	-4	0	4	0	-4
$1x$	1	0	1	0	0	1	1	0		0	-4	0	4	0	-4	0	-4
$1x$	1	0	1	0	1	0	0	1		0	-4	0	-4	0	-4	0	4
$1x$	1	0	0	1	0	1	0	1		0	4	0	-4	0	-4	0	-4
$1x$	0	1	1	0	0	1	0	1		0	4	0	4	0	-4	0	4
$1x$	0	1	0	1	1	0	0	1		0	4	0	-4	0	4	0	4
$1x$	0	1	0	1	0	1	1	0		0	4	0	4	0	4	0	-4

The number of possible function-pairs is 448. Considering all possible pairs one can construct 896 pieces of bent functions in 4-dimension. This number contains all possible negated versions, consequently we do not have to count them extra.

Unfortunately the process described above can not be used in a higher dimension.

In the case of a higher dimension we can apply the following construction to estimate the minimum number of highly nonlinear functions in odd dimension or the minimum number of bent functions in one-level higher dimension. This example can be found in [6]. Let e_i be the ith row of H_n. Hence $e_0, e_1, \dots, e_{2^n-1}$ are the $\{-1,1\}$ sequences of all the 2^n linear functions in $\{0,1\}^n$. Note that the length of each e_i is 2^n. Thus one can see that the concatenation of any 2^{n-1} different e_i of length 2^n is the sequence of a function on $\{0,1\}^{2n-1}$ and the result of the concatenation is a $\{-1,1\}$ sequence of a highly nonlinear function in $(2n-1)$-dimension.

Lemma 2.0.1. *Let* $g_i\colon \{0,1\}^n \to \{0,1\}$, $i = 0,1,\ldots,2^n-1$, *and* $G_j\colon \{0,1\}^{2n-1} \to \{0,1\}$, $j = 0,1,\ldots,2^{2n-1}-1$, *be linear functions defined by* $\overline{g}_i = l_i$, $\overline{G}_j = L_j$, *where* l_i *is the* i*th row of* H_n *and* L_j *is the* j*th row of* H_{2n-1}. *Let* $f\colon \{0,1\}^{2n-1} \to \{0,1\}$ *be a function such that* $\overline{f}(x)$ *is produced by the concatenation of* (2^{n-1}) *different* l_i. *Then function* f *is a highly nonlinear function in the* $(2n-1)$*-dimensional Boolean-space. The correlation of* f *with* G_j *is non-zero when* L_j *contains an* l_i *that occurs in* $\overline{f}(x)$ *and is zero when* L_j *does not contain an* l_i *which appears in* $\overline{f}(x)$.

Proof. L_j is the jth row of H_{2n-1}. Starting from H_n let us use the recursive process to generate H_{2n-1}:

$$H_n \to H_{n+1} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix} \to \cdots H_{2n-1} = \begin{bmatrix} H_{2n-2} & H_{2n-2} \\ H_{2n-2} & -H_{2n-2} \end{bmatrix},$$

$$H_n = \begin{bmatrix} l_0 \\ \vdots \\ l_{2^n-1} \end{bmatrix}, \quad H_{2n-1} = \begin{bmatrix} L_0 \\ \vdots \\ L_{2^{2n-1}-1} \end{bmatrix}.$$

Because of the construction any L_j row is actually generated by the concatenation of the sequences l_p or $-l_p$, where $p = j \mod 2^n$. Let us consider the generated sequence of f: $\overline{f}(x) = (l_{i_1}, l_{i_2}, \ldots, l_{i_{2^{n-1}}})$, $\{i_1, i_2, \ldots, i_{2^{n-1}}\} \subset \{0,1,\ldots,2^n-1\}$ and $L_j = (m_1 l_p, m_2 l_p, \ldots, m_{2^{n-1}} l_p)$, $m_r \in \{-1,1\}$, for every $r = 1,2,\ldots,2^{n-1}$. The correlation of function f and the linear function G_j is, as a matter of fact, the scalar product of the two sequences $\overline{f}(x)$ and L_j, that means

$$\overline{f}(x)L_j = \sum_{r=1}^{2^{n-1}} \left(l_{i_r}(l_p m_r)\right) = \sum_{r=1}^{2^{n-1}} (l_{i_r} l_p) m_r\,.$$

H_n has the property that the scalar product of two rows takes the value of 0 if they are different, and it takes the value of 2^n if they are the same. In the light of this the scalar product $\overline{f}(x)L_j$ takes the value of 2^n or -2^n, if and only if any $l_{i_r} = l_p$. Due to the construction of H_{2n-1}, a given row l_p comes (2^{n-1})-times among the rows of H_{2n-1}. $\overline{f}(x)$ is built by concatenated of (2^{n-1}) different sequences, consequently there are $2^{n-1}2^{n-1} = 2^{2n-2} = 2^{(2n-1)-1}$ such G_j functions, whose correlation with f takes the value of 2^n or -2^n. □

Theorem 2.0.2. *The number of highly nonlinear functions* $f\colon \{0,1\}^{2n-1} \to \{0,1\}$ *is at least*

$$2\frac{(2^n)!}{(2^{n-1})!}\,.$$

P r o o f . There are 2^n pieces of linear functions in the n-dimensional Boolean space, from those it is possible to choose 2^{n-1} pieces in $\binom{2^n}{2^{n-1}}$ different ways. The number of possible permutations of the 2^{n-1} selected rows is $(2^{n-1})!$, consequently the number of functions belonging to a choice is $(2^{n-1})!$. According to Lemma 2.0.1 each of the constructed functions is highly nonlinear function in the $(2n-1)$-dimensional space the number of which is:

$$\binom{2^n}{2^{n-1}}(2^{n-1})! = \frac{(2^n)!(2^{n-1})!}{(2^{n-1})!(2^n-2^{n-1})!} = \frac{(2^n)!}{(2^{n-1})!}.$$

When we take into consideration the fact, that the negation does not change the absolute-value of the correlation, we can multiply this number by 2. □

THEOREM 2.0.3. *The number of bent functions $f\colon \{0,1\}^{2n} \to \{0,1\}$ is at least:*

$$4(2^n)!.$$

P r o o f . When we construct a highly nonlinear function as described in Lemma 2.0.1, every time we select a half of the n-dimensional linear functions we actually create two sets of functions. Let us consider two elements from both sets, denote them h_1, $h_2\colon \{0,1\}^{2n-1} \to \{0,1\}$ and let $G_j\colon \{0,1\}^{2n-1} \to \{0,1\}$ be a linear function, for which $\overline{G}_j = L_j$ is a row in H_{2n-1} produced by l_p. If the correlation of h_1 with G_j is 0 then $\overline{h}_1(x)$ does not contain l_p but it occurs in $\overline{h}_2(x)$, consequently the correlation of h_2 with G_j is non-zero. So we get:

$$\left(\overline{h}_1(x)L_j\right)^2 + \left(\overline{h}_2(x)L_j\right)^2 = 2^{2(2n-1)-(2n-1-1)} = 2^{2n}.$$

That means h_1 and h_2 are a pair of highly nonlinear functions proper to construct bent functions in one-level higher dimension.

Under the conditions of Theorem 2.0.2 one can select $\frac{(2^n)!}{2(2^{n-1})!}$ pairs of function-sets. Both of them contain $(2^{n-1})!$ pieces of highly nonlinear functions, that means from a couple of function-sets it is possible to form $\left((2^{n-1})!\right)^2$ pairs of highly nonlinear functions. The linear functions in the $(2n-1)$-dimensional Boolean-space make it possible to create a further pair. For each $i = 0, 1, \ldots, 2^{2n-1}-1$,

$$\overline{h}_1(i) = \frac{(\overline{g}_1(x)L_i) + (\overline{g}_2(x)L_i)}{2^n}$$

and

$$\overline{h}_2(i) = \frac{(\overline{g}_1(x)L_i) - (\overline{g}_2(x)L_i)}{2^n}$$

are the generated sequences of such pairs generated by the use of the linear functions in shifted order.

Altogether $\left((2^{n-1})!\right)^2 + \left((2^{n-1})!\right)^2$ pairs of functions can be selected from a couple of function-sets. Multiplying this by the number of the pairs of function-sets we receive:

$$\frac{(2^n)!}{2\left((2^{n-1})!\right)^2}\left((2^{n-1})!\right)^2 2 = (2^n)!\,.$$

Regarding the possible negated versions of the used functions we have to multiply this expression by 4, so our proposition is proved. □

Summary. The results presented in this paper provide an estimate of the minimum number of highly nonlinear functions of odd dimension and the bent functions in the one-level higher dimension.

REFERENCES

[1] LICSKÓ, I.: *On highly nonlinear functions*, Ann. Univ. Sci. Budapest, Sect. Comp. **21** (2002), 165–175.

[2] LICSKÓ, I.: *Characteristics of highly nonlinear functions*, in: Hajducrypt '02, Debrecen, 2002.

[3] PÁSZTOR-VARGA, K.: *An optimizing method of Boolean functions based on structural features of the Boolean algera of Boolean functions* (Hungarian, English summary), Alkalmaz. Mat. Lapok **13** (1987–1988), 69–76.

[4] SEBERRY, J.—ZHANG, X. M.—ZHENG, Y.: *Nonlinearity and propagation characteristics of balanced Boolean functions*, Inform. Comput. **119** (1995), 1–13.

[5] VAJDA, I.—BUTTYÁN, L.—SZEKERES, B.: *Az algoritmikus adatvédelem módszerei* (*Methods of algorithmic data protection*), Technical University of Budapest, (non-published manuscript).

[6] ZHANG, X. M.—ZHENG, Y.: *New lower bounds on nonlinearity and a class of highly nonlinear functions*, in: Information Security and Privacy, Second Australian Conference Sidney, Australia, July 1997; Lecture Notes in Comput. Sci., Vol. 1270, Springer-Verlag, Berlin, 1997, pp. 147–158.

Received August 14, 2003

Budapesti Gazdasági Föiskola Kereskedelmi
Vendéglátóipari és Idegenforgalmi Kar
Informatikai Intézet
V. Alkotmány utca 9–11
H–1054 Budapest
HUNGARY
E-mail: licsko@mailbox.hu