Tight bound between nonlinearity and algebraic immunity

Mikhail Lobanov

Mech. & Math. Department
Moscow State University
119992 Moscow, Russia
emails: misha_msu@mail.ru

Abstract

We obtain tight bound between nonlinearity and algebraic immunity of a Boolean function and construct balanced functions that achive this bound for all possible values of parameters.

Boolean functions have wide applications in cryptography. Recently, algebraic attacks against stream ciphers were invented that applied the requirement of high algebraic immunity in combinations with other requirements to Boolean functions exploited as nonlinear filters in stream ciphers (see, for example, [1, 5]). One more cryptographic important property of Boolean functions especially important in stream ciphers is nonlinearity. In this respect the problem of relations between nonlinearity and algebraic immunity of Boolean functions has an interest.

In [2] it was proved the lower bound for the nonlinearity of a Boolean function via its algebraic immunity.

In this paper we obtain stronger lower bound for the nonlinearity of a Boolean function via its algebraic immunity and construct balanced functions that achive this bound for all possible values of parameters.

It is well known that a Boolean function has the only representation by a polynomial.

Definition 1. *The degree* of a Boolean function is the length of the longest term in its polynomial (the number of variables in this term).

Definition 2. A Boolean function g over F_2^n is *an annihilator* of a Boolean function f over F_2^n if $fg = 0$.

Obviously, all annihilators of f form a linear subspace in the space of all Boolean functions of n variables.

Definition 3. *The algebraic immunity* $AI(F)$ of a Boolean function f over F_2^n is the degree of the Boolean function g over F_2^n where g is nonzero Boolean function of minimum degree such that $fg = 0$ or $(f+1)g = 0$.

It is known [1, 5] that for any f over F_2^n the inequality $AI(f) \leq \lceil \frac{n}{2} \rceil$ holds.

Definition 4. *The weight* $wt(x)$ of a vector x in F_2^n is the number of ones in x.

Definition 5. *The distance* between Boolean functions f_1 and f_2 is defined as $d(f_1, f_2) =\mid \{x \in F_2^n \mid f_1(x) \neq f_2(x)\} \mid$.

Definition 6. *The nonlinearity* $nl(f)$ of a Boolean function f over F_2^n is $\min\limits_{l,\ deg(l) \leq 1} d(f, l)$.

Definition 7. For any vector $u \in F_2^n$ the value

$$W_f(u) = \sum_{x \in F_2^n} (-1)^{f(x)+<u,x>}$$

is called *the Walsh coefficient* of f at u.

The nonlinearity is expressed via Walsh coefficients by the next formula:

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{u \in F_2^n} |W_f(u)|.$$

In [2] it was proved that if $nl(f) < \sum_{i=0}^{d} \binom{n}{i}$ then $AI(f) \leq d+1$. This is equivalent to the lower bound of nonlinearity

$$nl(f) \geq \sum_{i=0}^{AI(f)-2} \binom{n}{i}.$$

Definition 8. A Boolean function $f(x_1, \ldots, x_n)$ is called *self-dual* if $f(x_1 + 1, x_2 + 1, \ldots, x_n + 1) = f(x_1, \ldots, x_n) + 1$.

It is easy to see that if f is self-dual then the fact that f has not a nonzero annihilator of degree less than k follows that $f+1$ has not a nonzero annihilator of degree less than k too. Therefore the minimum degrees of nonzero annihilators of functions f and $f+1$ are the same. Thus, for the finding of algebraic immunity of a self-dual function f it is sufficient to consider only annihilators of the function f.

Lemma 1. *Any annihilator* $g(x_1, \ldots, x_n)$ *of the function* $l(x_1, \ldots, x_n)$, $\deg(l) = 1$, *can be represented in the form*

$$g(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)(l(x_1, \ldots, x_n) + 1)$$

where $\deg(f) = \deg(g) - 1$.

Proof. Because of affine equivalence without loss of generality it is possible to assume $l = x_1 + 1$.

Consider the representation of $g(x_1, \ldots, x_n)$ in the polynomial form. Since all annihilators of a function form a linear space, after the cancellation of all terms that contain x_1 we must obtain the function $g_1(x_2, \ldots, x_n)$ such that

$g_1 l = g_1(x_1+1) = 0$. Since g_1 does not depend on x_1 we have $g_1 = 0$. Hence, any term of g contains x_1, then

$$g(x_1,\dots,x_n) = x_1 f(x_1,\dots,x_n) = (l+1)f$$

where $\deg(f) = \deg(g) - 1$.□

Lemma 2. *Let $l(x_1,\dots,x_n)$ be a Boolean function,* $\deg(l) = 1$. *Then all annihilators of the function l of degree at most t form the linear space of dimension $\sum_{i=0}^{t-1}\binom{n-1}{i}$.*

Proof. Because of an affine equivalence, it is possible to assume $l = x_1 + 1$.

Consider an arbitrary annihilator $g(x_1,\dots,x_n)$ of the function $l(x_1,\dots,x_n)$ such that $\deg(g) \le t$. Consider the representation of $g(x_1,\dots,x_n)$ in the polynomial form. Since all annihilators of a function form a linear space, after the cancellation of all terms that contain x_1 we must obtain the function $g_1(x_2,\dots,x_n)$ such that $g_1 l = g_1(x_1+1) = 0$. Since g_1 does not depend on x_1 we have $g_1 = 0$. Hence, any term of g contains x_1, then

$$g(x_1,\dots,x_n) = x_1 f(x_2,\dots,x_n)$$

where $\deg(f) \le t-1$.

In addition, any function $g(x_1,\dots,x_n) = x_1 f(x_2,\dots,x_n)$, where $f(x_2,\dots,x_n)$ is an arbitrary Boolean function of $n-1$ variables and of degree at most $t-1$, is an annihilator of l of degree at most t. It follows the statement of Lemma. □

Remark. The proof of the next lemma it is possible to find in [4] but we give it here because of its simplicity.

Lemma 3. *If f is a Boolean function over F_2^n and $AI(f) > k$, then*

$$\sum_{i=0}^{k}\binom{n}{i} \le wt(f) \le \sum_{i=0}^{n-k-1}\binom{n}{i}.$$

Proof. We look for an annihilator of the function f by the method of indeterminate coefficients:

$$g = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{1\le i<j\le n} a_{ij}x_i x_j + \dots + \sum_{1\le i_1\le\dots\le i_k\le n} a_{i_1\dots i_k} x_{i_1}\dots x_{i_k},$$

$\deg(g) \le k$.

The function g is an annihilator of f if and only if $f(x) = 1$ follows $g(x) = 0$. Then in order to provide $AI(f) > k$, it is necessary that obtained homogeneous system of linear equations on $a_0, a_1, a_2, \dots$ has the only zero solution. For this it is necessary that the number of unknowns does not exceed the number of equations. The number of equations is $wt(f)$ whereas the number of unknowns is $\sum_{i=0}^{k}\binom{n}{i}$. Hence, the left inequality is proved. Applying the same reasoning to $f+1$ we obtain the right inequality. □

Theorem 1. *Let $f(x_1,\dots,x_n)$ be a Boolean function over F_2^n and $AI(f)=k$. Then*
$$nl(f)\geq 2^{n-1}-\sum_{i=k-1}^{n-k}\binom{n-1}{i}=2\sum_{i=0}^{k-2}\binom{n-1}{i}. \qquad (1)$$

Proof. For $k=1$ our bound gives $nl(f)\geq 0$. Assume $k\geq 2$.

Represent the nonlinearity of the function f in the form $nl(f)=2^{n-1}-\frac{\alpha}{2}$ where $\alpha=\max_{u\in F_2^n}|W_f(u)|$.

If $\max_{u\in F_2^n}|W_f(u)|$ is achieved at zero vector, then f or $f+1$ has the weight $\frac{2^n-\alpha}{2}$. Then in accordance with Lemma 3 we have
$$\frac{2^n-\alpha}{2}\geq\sum_{i=0}^{k-1}\binom{n}{i}.$$
Therefore, $\alpha\leq\sum_{i=k}^{n-k}\binom{n}{i}\leq 2\sum_{i=k-1}^{n-k}\binom{n-1}{i}$. From here we obtain the required bound on the nonlinearity.

If $\max_{u\in F_2^n}|W_f(u)|$ is not achieved at zero vector, then there exists the function $l(x_1,\dots,x_n),\deg(l)=1$, such that $d(f,l)=\frac{2^n-\alpha}{2}$. The functions f and l have the same values at $\frac{2^n+\alpha}{2}$ vectors. Suppose that among these vectors there exist exactly β vectors x where $f(x)=1$, then there exist exactly $2^{n-1}-wt(f)-\frac{\alpha}{2}+\beta$ vectors where $f=0$ and $l=1$.

Then
$$wt(f(l+1))=wt(f)-\beta \qquad (2)$$
and
$$wt((f+1)l)=2^{n-1}-wt(f)-\frac{\alpha}{2}+\beta. \qquad (3)$$
The right side in (2) is decreasing in β whereas the right side in (3) is increasing in β. The equality as achieved for $\beta=wt(f)-2^{n-2}+\frac{\alpha}{4}$. It follows that
$$\min(wt(f(l+1)),wt((f+1)l))\leq 2^{n-2}-\frac{\alpha}{4}.$$
If $wt(f(l+1))<wt((f+1)l)$ then define $f_1=f,l_1=l+1$, otherwise define $f_1=f+1,l_1=l$.

Input the function $f_2=f_1l_1$. Then $wt(f_2)\leq 2^{n-2}-\frac{\alpha}{4}$.

We look for annihilators g of the function f_2 of degree at most $k-2$ by the method of indeterminate coefficients:
$$g=a_0+\sum_{i=1}^{n}a_ix_i+\sum_{1\leq i<j\leq n}a_{ij}x_ix_j+\dots+\sum_{1\leq i_1\leq\ldots\leq i_{k-2}\leq n}a_{i_1\ldots i_{k-2}}x_{i_1}\ldots x_{i_{k-2}}.$$
A function g is the annihilator of f if and only if $f(x)=1$ follows $g(x)=0$.

Hence, we obtain the homogeneous system of at most $2^{n-2}-\frac{\alpha}{4}$ linear equations on $\sum_{i=0}^{k-2}\binom{n}{i}$ unknowns. The space of solutions of this system has the dimension at least $\sum_{i=0}^{k-2}\binom{n}{i}-(2^{n-2}-\frac{\alpha}{4})$.

By Lemma 2 the dimension of the space of annihilators of the function l_1 of degree at most $k-2$ is $\sum_{i=0}^{k-3}\binom{n-1}{i}$.

If $\sum_{i=0}^{k-2}\binom{n}{i}-(2^{n-2}-\frac{\alpha}{4})>\sum_{i=0}^{k-3}\binom{n-1}{i}$ then there exists the function f_3, $\deg(f_3)\le k-2$, such that $f_2f_3=0$ but $f_3l_1\ne 0$. Then f_3l_1 is the annihilator of f_1, in addition $\deg(f_3l_1)\le k-1$ that contradicts to $AI(f)=k$.

It follows $\sum_{i=0}^{k-2}\binom{n}{i}-(2^{n-2}-\frac{\alpha}{4})\le\sum_{i=0}^{k-3}\binom{n-1}{i}$,

$$\frac{\alpha}{4}\le 2^{n-2}-\frac{1}{2}\sum_{i=k-2}^{n-k+1}\binom{n-1}{i}+2^{n-2}-\left(2^{n-1}-\frac{1}{2}\sum_{i=k-1}^{n-k+1}\binom{n}{i}\right),$$

$$\frac{\alpha}{4}\le\frac{1}{2}\left(\sum_{i=k-1}^{n-k+1}\left(\binom{n-1}{i}+\binom{n-1}{i-1}\right)-\sum_{i=k-2}^{n-k+1}\binom{n-1}{i}\right)=\frac{1}{2}\sum_{i=k-1}^{n-k}\binom{n-1}{i}.$$

Therefore, $nl(f)\ge 2^{n-1}-\sum_{i=k-1}^{n-k}\binom{n-1}{i}$.$\Box$

Corollary 1. *If n odd and $AI(f(x_1,\dots,x_n))=\lceil\frac{n}{2}\rceil$ then*

$$nl(f)\ge 2^{n-1}-\binom{n-1}{\frac{n-1}{2}}. \tag{4}$$

Note that in [4] it was constructed the function of odd number n of variables with the algebraic immunity $\lceil\frac{n}{2}\rceil$ and nonlinearity $nl(f)=2^{n-1}-\binom{n-1}{\frac{n-1}{2}}$. Our Corollary 1 clarifies that this function achieves our bound (4), i. e. among all functions with maximum possible algebraic immunity this function has the worst possible nonlinearity. The calculation of its nonlinearity in [4] is quite difficult and takes some pages. Now the lower bound for the function from [4] follows immediately from our Corollary 1. At the same time the upper bound for the nonlinearity of the function from [4] will follow from our Theorem 2 since this function is a particular case of our functions $f_{n,k}$ appeared in the proof of our Theorem 2. Note also that in [3] for the constructed there the function f with odd number n of variables and the algebraic immunity $\lceil\frac{n}{2}\rceil$ it was proved the lower bound of nonlinearity $nl(f)\ge 2^{n-1}-\binom{n-1}{\frac{n-1}{2}}$ that coincides with our bound in Corollary 1 for all functions with such number of variables and such algebraic immunity.

Corollary 2. *If n even and $AI(f(x_1,\dots,x_n))=\lceil\frac{n}{2}\rceil$ then*

$$nl(f)\ge 2^{n-1}-\binom{n}{\frac{n}{2}}.$$

Note that in [4] the bound of our Corollary 2 was proved for very narrow class of functions.

Theorem 2. *The bound (1) in Theorem 1 is unimprovable for any n and any $k\le\lceil\frac{n}{2}\rceil$. Moreover, for any admissible parameters n and k there exists a balanced function that achieves this bound.*

Proof. Show that the bound (1) in Theorem 1 is unimprovable presenting for any n and any $k \le \lceil \frac{n}{2} \rceil$ the balanced function $f(x_1, \dots, x_n)$ such that $AI(f) = k$ and $nl(f) = 2^{n-1} - \sum_{i=k-1}^{n-k} \binom{n-1}{i}$.

Define the function $f_{n,k}$ by the next way:

$$f_{n,k}(x_1, \dots, x_n) = \begin{cases} 0, & \text{if} \quad wt(x_1, \dots, x_n) < k, \\ 1, & \text{if} \quad wt(x_1, \dots, x_n) > n - k, \\ x_1, & \text{if} \quad k \le wt(x_1, \dots, x_n) \le n - k. \end{cases}$$

Now prove that for any n and any $k \le \lceil \frac{n}{2} \rceil$ we have $AI(f_{n,k}) = k$ and $nl(f_{n,k}) = 2^{n-1} - \sum_{i=k-1}^{n-k} \binom{n-1}{i}$.

It is easy to see that $f(x_1 + 1, x_2 + 1, \dots, x_n + 1) = f(x_1, \dots, x_n) + 1$, i. e. $f_{n,k}$ is a self-dual function. Hence, the function $f_{n,k}$ is a balanced function.

Since $f_{n,k}$ is self-dual, in order to prove $AI(f) \ge k$, it is sufficient to prove that $f_{n,k} + 1$ has not a nonzero annihilator of degree less than k.

Write the possible annihilator g of the function $f + 1$ of degree at most $k - 1$ by means of indeterminate coefficients:

$$g = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{1 \le i < j \le n} a_{ij} x_i x_j + \dots + \sum_{1 \le i_1 \le \dots \le i_{k-1} \le n} a_{i_1 \dots i_{k-1}} x_{i_1} \dots x_{i_{k-1}}.$$

The function g is the annihilator of $f_{n,k} + 1$ if and only if $f(x) + 1 = 1$ follows $g(x) = 0$. We obtain the system of homogeneous linear equations on the coefficients of the function g:

$$g(x_1, \dots, x_n) = 0$$

for all vectors x such that $wt(x) \le k - 1$.

Since $g(0, \dots, 0) = 0$, we have $a_0 = 0$. Since $g(x) = 0$ if $wt(x) = 1$, we have $a_i = a_0 = 0$. Applying the induction on the weight of vectors we obtain that all coefficients of g are zeros, hence, $g \equiv 0$. Thus, $AI(f_{n,k}) \ge k$. At the same time it is easy to see that $g(x_1, \dots, x_n) = (x_1 + 1) \dots (x_k + 1)$ is the annihilator of $f_{n,k}$ of degree k. Therefore, $AI(f_{n,k}) = k$.

Calculate the Walsh coefficient of the function $f_{n,k}$ at the vector $(1, 0, \dots, 0)$ using the self-duality of $f_{n,k}$:

$$W_{f_{n,k}}(1, 0, \dots, 0) = \sum_{(x_1, \dots, x_n) \in F_2^n} (-1)^{f_{n,k}(x_1, \dots, x_n) + x_1} =$$

$$= 2^n - 2wt(f_{n,k}(x_1, \dots, x_n) + x_1) =$$

$$= 2^n - 2(wt(f_{n,k}(0, x_2, \dots, x_n)) + wt(f_{n,k}(1, x_2 \dots, x_n) + 1)) =$$

$$= 2^n - 4wt(f_{n,k}(0, x_2, \dots, x_n)) = 2^n - 4 \sum_{i=n-k+1}^{n-1} \binom{n-1}{i} = 2 \sum_{i=k-1}^{n-k} \binom{n-1}{i}.$$

Hence, $nl(f_{n,k}) \leq 2^{n-1} - \sum_{i=k-1}^{n-k} \binom{n-1}{i}$. Above we proved that $AI(f_{n,k}) = k$, hence, by Theorem 1 we have $nl(f_{n,k}) \geq 2^{n-1} - \sum_{i=k-1}^{n-k} \binom{n-1}{i}$, it follows $nl(f_{n,k}) = 2^{n-1} - \sum_{i=k-1}^{n-k} \binom{n-1}{i}$.□

The author is deeply grateful to his scientific supervisor Prof. Yuriy Tarannikov for the formulation of the problem, attention to the work and valuable advices.

References

[1] N.Courtois and W.Meier. Algebraic attacks on stream ciphers with linear feedback. In Anvances in Cryptology — EUROCRYPT 2003, number 2656 in Lecture Notes in Computer Science, pages 345–359. Springer-Verlag, 2003.

[2] D.K.Dalai, K.C.Gupta and S.Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. Indocrypt 2004, Chennai, India, December 20–22, pages 92–106, volume 3348 in Lecture Notes in Computer Science, Springer-Verlag, 2004.

[3] D.K.Dalai, K.C.Gupta and S.Maitra. Cryptographically Significant Boolean Functions: Constraction and Analysis in terms of Algebraic Immunity. FSE 2005, pages 98–111, volume 3557 in Lecture Notes in Computer Science, Springer-Verlag, 2005.

[4] D.K.Dalai, S.Maitra, S.Sarkar. Basic Theory in Construction of Boolean Functions with Maximum Possible Annihilator Immunity. Cryptology ePrint archive (http://eprint.iacr.org/), Report 2005/229.

[5] W.Meier, E.Pasalic and C.Carlet. Algebraic attacks and decomposition of Boolean functions. In Advances in Cryptology — EUROCRYPT 2004, number 3027 in Lecture Notes in Computer Science, pages 474–491. Springer-Verlag, 2004.