A method for obtaining lower bounds on the higher order nonlinearity of Boolean function

Mikhail Lobanov

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

Abstract

Obtainment of exact value or high lower bound on the r-th order nonlinearity of Boolean function is a very complicated problem (especial if $r > 1$). In a number of papers lower bounds on the r-th order nonlinearity of Boolean function via its algebraic immunity were obtain for different r. This bounds is rather high for function with maximum near maximum possible algebraic immunity. In this paper we prove theorem, which try to obtain rather high lower bound on the r-th order nonlinearity for many functions with small algebraic immunity.

Keywords: Boolean function, algebraic immunity, algebraic degree, nonlinearity, higher order nonlinearity, annihilator

A Boolean function of n variables is a function f: $F_2^n \to F_2$. The weight $wt(x)$ of a vector $x \in \mathbf{F}_2^n$ is the number of ones in x. The weight $wt(f)$ of f from $F_2^n into F_2$ is the number of vectors x from F_2^n, that $f(x) = 1$.

It is well known that a Boolean function can be uniquely represented as a polynomial

$$f(x_1, \ldots, x_n) = \bigoplus_{a_1,\ldots,a_n \in F_2^n} g(a_1, \ldots, a_n) x_1^{a_1} \cdots x_n^{a_n},$$

where g is a Boolean function too. This polynomial representation is called its algebraic normal form (ANF). The algebraic degree of f, denoted $\deg(f)$, is the length of the longest term in the polynomial of f.

The algebraic immunity of f is defined:

$$AI(f) = \min_{g \neq 0,\ gf \equiv 0 \text{ or } g(f+1) \equiv 0} deg(g).$$

It is known [3, 4] that for any f on $\mathbf{F}_2^n$ the inequality $AI(f) \leq \lceil \frac{n}{2} \rceil$ holds.

The nonlinearity of rth order $nl_r(f)$ of a Boolean function f over $\mathbf{F}_2^n$ is called the value $\min\limits_{l,\ deg(l) \leq r} d(f, l)$, where $d(f, l)$ is the Hamming distance.

In [1, 2] it was proved exact lower bound on first and second order nonlinearities via value of algebraic immunity. In [5] it was proved lower bound (not exact) on the r-th order nonlinearity.

$$nl_1(f) \geq 2 \sum_{i=0}^{AI(f)-2} \binom{n-1}{i}, nl_2(f) \geq \sum_{i=0}^{AI(f)-1} \binom{n}{i} - \sum_{i=0}^{AI(f)-1} 2^i \binom{n-2i-1}{AI(f)-1-i},$$

$$nl_r(f) \geq \sum_{i=0}^{AI(f)-r-1} \binom{n}{i} + \sum_{i=AI(f)-2r}^{AI(f)-r-1} \binom{n-r}{i}. \quad (1)$$

If $AI(f)$ is near to maximum possible, bounds are rather high, but they don't obtain so good results for function with low $AI(f)$. In [] it was proved lower bound on $nl_r(f)$ not via $AI(f)$, but via $\min_{g \neq 0,\ gf \equiv 0} deg(g)$ and $\min_{g \neq 0,\ g(f+1) \equiv 0} deg(g)$. This bound generalize bound (1) and it is better for some function with low algebraic immunity.

In this paper we prove theorem, that generalize author's method of obtaining of lower bounds on $nl_r(f)$ via $AI(f)$ from [2] and result from [6]. Bounds on $nl_r(f)$ obtaining from this theorem is higher for many functions.

Definition 1 *Let* $h(x_1, \ldots, x_n)$,*define* $An_k(h) = \{g(x_1, \ldots, x_n) | \, gh = 0, \, deg(g) \leq k\}$.

Definition 2 *Let* $C = \{\overline{x}_1, \ldots, \overline{x}_n\}$ *be some set of vectors in* $\mathbf{F}_2^n$. *For any given* k, $k \leq n$, *and for any vector* $x = (x_1, \ldots, x_n) \in \mathbf{F}_2^n$ *we correspond to* x *the uniform linear equation with the left side generated by the substitution of components of the vector* x *into the expression*

$$a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \cdots + \sum_{1 \leq i_1 \leq \ldots \leq i_k \leq n} a_{i_1 \ldots i_k} x_{i_1} \ldots x_{i_k}.$$

The right side of the equation is 0. *Then we call a* k*-rank of the set* C *the rank of the system of linear equations generated by such way from the vectors of the set* C. *Denote this rank by* $r_k(C)$.

Next, we search all functions from $An_k(f)$ by the method of undefined coefficients:

$$g = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \cdots + \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1 \ldots i_k} x_{i_1} \ldots x_{i_k}.$$

The function g is an annihilator of f if and only if $f(x) = 1$ follows $g(x) = 0$. Thus, we obtain the system of linear equations.

Dimension of solution space of homogeneous linear system equals to rank of the system subtract number of variables.

$$\dim(An_k(f)) = \sum_{i=0}^{k} \binom{n}{i} - r_k(supp(f)). \quad (2)$$

Proposition 1 *Let f and f_0 be n-variable functions, $1 \le k_1, k_2 \le n$, $\dim(An_{k_1}(f)) \ge \dim(An_{k_1}(f_0))$ and $\dim(An_{k_2}(f+1)) \ge \dim(An_{k_2}(f_0+1))$. Then*

$$d(f, f_0) \ge \dim(An_{k_1}(f)) - \dim(An_{k_1}(f_0)) + \dim(An_{k_2}(f+1)) - \dim(An_{k_2}(f_0+1)).$$

Proof. From (2) we obtain $r_{k_1}(supp(f_0)) - r_{k_1}(supp(f)) = \dim(An_{k_1}(f)) - \dim(An_{k_1}(f_0))$. Hence, there exist at least $\dim(An_{k_1}(f)) - \dim(An_{k_1}(f_0))$ vectors where f_0 is equal to 1, and f is equal to 0.

Analogously, we obtain, that there exist at least $\dim(An_{k_2}(f+1)) - \dim(An_{k_2}(f_0+1))$ the number of vectors where f is 1 and f_0 is 0. □

Definition 3 *Let h be n-variable function. Define $B_{k_1,k_2}(h) = \{g(x_1, \ldots, x_n) | \deg(g) \le k_1, \deg(gh) \le k_2\}$.*

In [2] the following bound was proved

$$\dim(B_{k,k}(h)) \ge \sum_{i=0}^{k-r} \binom{n}{i} + \sum_{i=k-2r+1}^{k-r} \binom{n-r}{i}, \tag{3}$$

if $deg \le r$

Proposition 2 *Let $k_1 \ge k_2$, then*

$$\dim(An_{k_1}(f)) + \dim(An_{k_2}(f+1)) = \dim(B_{k_1,k_2}(f)).$$

Proof. It is sufficient to prove, that $B_{k_1,k_2}(f)$ is a direct sum of $An_{k_1}(f)$ and $An_{k_2}(f+1)$.

Because of $An_{k_1}(f) \cap An_{k_2}(f+1) = 0$ it is sufficient to prove, that some function from $B_{k_1,k_2}(f)$ is represented as a sum of functions from $An_{k_1}(f)$ and $An_{k_2}(f+1)$, and sum of some functions from $An_{k_1}(f)$ and $An_{k_2}(f+1)$ is a function from $B_{k_1,k_2}(f)$.

Let $g_1 \in An_{k_1}(f)$, a $g_2 \in An_{k_2}(f+1)$ then $g_1 + g_2 \in B_{k_1,k_2}(f)$. Genuinely:

$$\deg(g_1 + g_2) \le \max(\deg(g_1), \deg(g_2)) \le k_1,$$

$$\deg((g_1 + g_2)f) = \deg(g_1 f + g_2(f+1) + g_2) = \deg(g_2) \le k_2.$$

Inverse, let $g \in B_{k_1,k_2}(f)$, then $g(f+1) \in An_{k_1}(f)$ and $gf \in An_{k_2}(f+1)$. It follows from definition of space $B_{k_1,k_2}(f)$ and

$$gf(f+1) \equiv 0,$$

$$\deg(g(f+1) = \deg(gf + g) \le \max(\deg(gf), \deg(g)) \le k_1.$$

Because of $g = gf + g(f+1)$, some function $g \in B_{k_1,k_2}(f)$ can be represented as $g = g_1 + g_2$, where $g_1 = g(f+1) \in An_{k_1}(f)$ and $g_2 = gf \in An_{k_2}(f+1)$. □

As a simple corollary from Proposition 1 and 2 obtain:

Corollary 1 *Let f and f_0 be n-variable functions, $1 \leq k_2 \leq k_1 \leq n$, $\dim(An_{k_1}(f)) \geq \dim(An_{k_1}(f_0))$ and $\dim(An_{k_2}(f+1)) \geq \dim(An_{k_2}(f_0+1))$. Then $d(f,f_0) \geq \dim(B_{k_1,k_2}(f)) - \dim(B_{k_1,k_2}(f_0))$.*

A cause of constant approximating considered in its own right, we obtain from Corollary 1 the following bound on the r-th order nonlinearity.

Theorem 1 *Let for $f(x_1,\ldots,x_n)$ the following inequations are true $\min_{1\leq deg(g)\leq r}\dim(An_{k_1}(g)) \geq \dim(An_{k_1}(f))$ and $\min_{1\leq deg(g)\leq r}\dim(An_{k_1}(g)) \geq \dim(An_{k_1}(f+1))$, then*

$$nl_r(f) \geq \min\big(\min_{deg(g)\leq r}\dim(B_{k_1,k_2}(g)) - \dim(B_{k_1,k_2}(f)), wt(f), wt(f+1)\big),$$

if $k_1 \geq k_2$, and

$$nl_r(f) \geq \min\big(\min_{deg(g)\leq r}\dim(B_{k_2,k_1}(g)) - \dim(B_{k_2,k_1}(f+1)), wt(f), wt(f+1)\big),$$

if $k_1 < k_2$.

Next can be useful to check the assumptions of 1.

Proposition 3 *The following inequation is true*

$$\min_{1\leq deg(g)\leq r}\dim(An_k(g)) \geq \sum_{i=0}^{k-r}\binom{n-r}{i}.$$

Theorem 1 generalizes the corresponding results from [6, 5, 2]; therefore, the estimates obtained with its help, are necessarily not less sharp. In what follows, we shall show that, in fact, for some particular functions, this theorem implies stronger estimates for $nl_r(f)$ than those from [6, 5, 2].

For $n = 4k+1$ let us define the function $f_n(x_1,\ldots,x_n)$:

$$f_n(x_1,\ldots,x_n) = \begin{cases} 0, & \text{if} \quad wt(x_1,\ldots,x_n) \leq 2k, \\ 1, & \text{if} \quad wt(x_1,\ldots,x_n) > 2k, \end{cases}$$

Consider the function $f = f_n \cdot (x_1 \vee x_2 \vee \ldots \vee x_k) \vee x_1x_2\ldots x_k$. It is easy to prove that $AI(f) = k$. From [5, 2] and from [6] in view (3) the same bound can be deduced:

$$nl_r(f) \geq \min_{deg(g)\leq r}\dim(B_{k-1,k-1}(g)) \geq \sum_{i=0}^{k-r-1}\binom{n}{i} + \sum_{i=k-2r}^{k-r-1}\binom{n-r}{i}.$$

At the same time, using Theorem 1 and Proposition 3, we obtain the significantly stronger bound

$$nl_r(f) \geq \min_{deg(g)\leq r}\dim(B_{2k,2k}(g)) - 2\sum_{i=0}^{k}\binom{n-k}{i} \geq$$

$$\geq \sum_{i=0}^{2k-r}\binom{n}{i} + \sum_{i=2k-2r+1}^{2k-r}\binom{n-r}{i} - 2\sum_{i=0}^{k}\binom{n-k}{i}.$$

References

[1] M.S.Lobanov, Exact relation between nonlinearity and algebraic immunity // Discrete Mathematics and Applications, 16 :5, 2006. — P. 453 -460

[2] M.S.Lobanov, Exact relations between nonlinearity and algebraic immunity // Journal of Applied and Industrial Mathematics, 3:3, 2009. — P. 367 - 376

[3] N.Courtois and W.Meier. Algebraic attacks on stream ciphers with linear feedback // Advances in cryptology, EUROCRYPT 2003. — Berlin/Heidelberg: Springer Verl., 2003. — P. 345-359. (Lecture Notes in Computer Science; Vol. 2656).

[4] W.Meier, E.Pasalic and C.Carlet. Algebraic attacks and decomposition of Boolean functions // Advances in Cryptology — EUROCRYPT 2004. — Berlin/Heidelberg: Springer Verl., 2004. — P. 474-491. (Lecture Notes in Computer Science; Vol. 3027).

[5] S.Mesnager, Improving the lower bound on the higher order nonlinearity of Boolean functions with prescribed algebraic immunity, IEEE Trans. Information Theory 54 (8),2008. — P. 3656-3662.

[6] P.Rizomiliotis. Improving the high order nonlinearity lower bound for Boolean functions with given algebraic immunity, Discrete Appl. Math., 158:18,2010. — P. 2049-2055