

On the Boolean functions With Maximum Possible Algebraic Immunity : Construction and A Lower Bound of the Count

Longjiang Qu, Guozhu Feng, and Chao Li

Dept of Mathematic and System Science,National University of Defence
Technology,Changsha,Hunan 410073, China

Abstract. This paper gives three construction methods which each can get a class of Boolean functions with maximum algebraic immunity from one such giving function. Our constructions get more functions than any previous construction. The cryptographic properties, such as balance, algebraic degree etc, of those functions are studied. It shows that we can construct Boolean functions with better cryptographic properties, which gives the guidance for the design of Boolean functions to resist algebraic attack, and helps to design good cryptographic primitives of cryptosystems. From these constructions, we get a lower bound of the count of Boolean functions which have maximum algebraic immunity, which shows that such boolean functions are numerous. As far as we know, this is the first bound about this count.

Key words: Algebraic Attack, Algebraic Degree, Algebraic Immunity, Annihilator, Balance, Boolean Functions, Nonlinearity

1 Introduction

Algebraic attack (that uses overdefined systems of multivariate equations to recover the secret key) has received a lot of attention recently [1,2,8,9,11-13,20,24] in studying security of the cryptosystems. This adds a new cryptographic property for designing Boolean functions to be used as building blocks in cryptosystems which is known as algebraic immunity [3-5, 7, 14, 15, 16, 23].

Given an n -variable Boolean function f , different cases related to low degree multiples of f have been studied in [12, 24]. The main objective is to find out minimum (or low) degree annihilators of f and $1+f$, i.e, to find out minimum (or low) degree n -variable nonzero functions g such that $f*g = 0$ and $(1+f)*g = 0$. To mount the algebraic attack, one needs the low degree linearly independent annihilators [12, 24] of f and $1+f$.

Though there are increasing interest in construction of Boolean functions with good annihilator immunity [3-5,7, 14, 15, 16], so far there is only two construction method [15,16] that can achieve the maximum possible annihilator immunity $\lceil \frac{n}{2} \rceil$ for an n -variable function. The heart of the construction in [15] was a function ϕ_{2k} on even $(2k)$ number of variables with maximum possible

annihilator immunity k . The main problem with ϕ_{2k} is that no clear intuition has been provided how one can land into such a complicated structure. Further, the other cryptographic properties, such as weight, nonlinearity or algebraic degree of the function ϕ_{2k} are yet to be answered and only a few experimental results have been provided in [15] for $n = 1, 2, \dots, 8$. Also the functions ϕ_{2k} are not balanced. [16] first explain a generic construction idea of functions with maximum algebraic immunity that comes from the basic theory, then study the cryptographic properties of the constructions, such as nonlinearity, algebraic degree etc. Both the two papers have the same shortcoming, they construct too few such functions, [15] gets only one high dimension function from a low dimension function, [16] provides only symmetric functions with maximum possible algebraic immunity, 1 for n odd and $2^{C_n^{\frac{n}{2}}}$ for n even, and [16] also points out that linear transformation can provide more such functions, but linear transformation don't change the algebraic degree and nonlinearity, so they can't improve the cryptographic properties of these functions. As provide so few Boolean functions, they're not good for cryptographic use.

In this paper we give three construction methods which each can get a class of Boolean functions with maximum algebraic immunity from one such giving function. Our constructions get more functions than any previous construction([15, 16]). The cryptographic properties, such as balance, algebraic degree etc, of those functions are studied. It shows that we can construct Boolean functions with better cryptographic properties. As we provides more functions, it's more free to choose functions with better cryptographic properties. This gives the guidance for the design of Boolean functions to resist algebraic attack, and helps to design good cryptographic primitives of cryptosystems. From these constructions, we get a lower bound of the count of Boolean functions which have maximum algebraic immunity, which shows that such boolean functions are numerous. As far as we know, this is the first bound about this count.

The organization of the paper is as follows. In the following section we give some preliminaries of this paper. In Section 3, we give three constructions to get a class of Boolean functions with maximum possible algebraic immunity from one such giving function. Their cryptographic properties are studied in Section 4. In Section 5, we discuss the count of the Boolean functions with maximum possible algebraic immunity and give a lower bound of that. Section 6 concludes the paper.

2 Preliminaries

A Boolean function on n variables may be viewed as a mapping from $V_n = \{0, 1\}^n$ into $V_1 = \{0, 1\}$ and define B_n as the set of all n -variable Boolean functions. One of the standard representation of a Boolean function $f(x_1, \dots, x_n)$ is by the output column of its truth table, i.e., a binary string of length 2^n ,

$$f = [f(0, 0, \dots, 0), f(1, 0, \dots, 0), f(0, 1, \dots, 0), \dots, f(1, 1, \dots, 1)]$$

The set of $x \in V_n$ for which $f(x) = 1$ (respectively $f(x) = 0$) is called the on set (respectively off set), denoted by $S_1(f)$ (respectively $S_0(f)$). We say that a Boolean function f is balanced if the truth table contains an equal number of 1's and 0's. The Hamming weight of a binary string S is the number of ones in the string. This number is denoted by $wt(S)$. The Hamming distance between two strings, S_1 and S_2 is denoted by $d(S_1, S_2)$ and is the number of places where S_1 and S_2 differ. Note that $d(S_1, S_2) = wt(S_1 + S_2)$ (by abuse of notation, we also use $+$ to denote the $GF(2)$ addition, i.e., the XOR).

Any Boolean function has a unique representation as a multivariate polynomial over $GF(2)$, called the algebraic normal form (ANF),

$$f(x_1, \dots, x_n) = a_0 + \sum_{1 \leq i \leq n} a_i x_i + \sum_{1 \leq i < j \leq n} a_{i,j} x_i x_j + \dots + a_{12\dots n} x_1 x_2 \dots x_n$$

where the coefficients $a_0, a_i, a_{i,j}, \dots, a_{12\dots n}, \in \{0, 1\}$. The algebraic degree $deg(f)$, is the number of variables in the highest order term with nonzero coefficient. A Boolean function is affine if there exists no term of degree > 1 in the ANF and the set of all affine functions is denoted $A(n)$. An affine function with constant term equal to zero is called a linear function.

It is known that a Boolean function should be of high algebraic degree to be cryptographically secure [18]. Further, it has been identified recently, that it should not have a low degree multiple [12]. The algebraic attack (see [12, 24] and the references in these papers) is getting a lot of attention recently. To resist algebraic attacks, the Boolean functions used in the cryptosystems should be chosen properly.

Definition 1. [16] 1. Given $f \in B_n$, a nonzero function $g \in B_n$ is called an annihilator of f if $f * g = 0$. By $AN(f)$ we mean the set of annihilators of f .

2. Given $f \in B_n$, the algebraic immunity of f , denoted by $AI_n(f) = deg(g)$, where $g \in B_n$ is the minimum degree nonzero function such that either $f * g = 0$ or $(1 + f) * g = 0$.

It is known [12, 24] that for $f \in B_n$, $AI_n(f) \leq \lceil \frac{n}{2} \rceil$ and in [15,16] constructions achieving the maximum value were presented. In this paper we will present constructions to get more such functions.

The nonlinearity of an n -variable function f is the minimum distance from the set of all n -variable affine functions, i.e.,

$$nl(f) = \min_{g \in A(n)} (d(f, g))$$

Boolean functions used in cryptosystems must have high nonlinearity to prevent linear attacks [18].

Many properties of Boolean functions can be described by the Walsh transform. Let $x = (x_1, \dots, x_n)$ and $\omega = (\omega_1, \dots, \omega_n)$ both belonging to $V_n = \{0, 1\}^n$ and $x \bullet \omega = x_1 \omega_1 + \dots + x_n \omega_n$. Let $f(x)$ be a Boolean function on n variables. Then the Walsh transform of $f(x)$ is an integer valued function over $V_n = \{0, 1\}^n$

which is defined as

$$W_f(\omega) = \sum_{x \in \{0,1\}^n} (-1)^{f(x) + \omega x}$$

A Boolean function f is balanced iff $W_f(0) = 0$. The nonlinearity of f is given by $nl(f) = 2^{n-1} - \frac{1}{2} \max_{\omega \in \{0,1\}^n} |W_f(\omega)|$.

3 Constructions for maximum possible algebraic immunity

Let $f \in B_n$ and consider that f has an annihilator g of degree d . Let the ANF of $g = a_0 + \sum_{i=1}^n a_i x_i + \sum_{1 \leq i < j \leq n} a_{i,j} x_i x_j + \dots + \sum_{1 \leq i_1 < \dots < i_d \leq n} a_{i_1, \dots, i_d} x_{i_1} \dots x_{i_d}$. Note that $f(x) = 1$ implies $g(x) = 0$. So, we will be able to get linear equations from $g(x) = 0$ on the a's in ANF of g . That is we will get $wt(f)$ many homogeneous linear equations on the a's. Solving the system of linear homogeneous equations, we can find out annihilators g of degree $\leq d$ on nontrivial solutions. (In case of a trivial solution we will get all the a's equal to zero, i.e., $g(x) = 0$, which is not acceptable as we are interested in nonzero $g(x)$.) Here, we have $\sum_{i=0}^d \binom{n}{i}$ number of variables (the a's for the monomials up to degree d) and $wt(f)$ many number of equations. We get $wt(f)$ many homogeneous linear equations using the a's. Let us denote the coefficient matrix of this system of equations by $S_1^d(f)$, and the row coefficients as $\{u_j = (1, c_1, c_2, \dots, c_n, c_{i_1} c_{i_2}, \dots, c_{i_1} \dots c_{i_d}) | 1 \leq j \leq wt(f)\}$. Then u_j 's dimensions are all $\sum_{i=0}^d \binom{n}{i}$. $S_1^d(f)$ has $wt(f)$ many rows and $\sum_{i=0}^d \binom{n}{i}$ many columns, and it also can be seen as a vector set of u_j , that is $S_1^d(f) = \{(1, c_1, c_2, \dots, c_n, c_{i_1} c_{i_2}, \dots, c_{i_1} \dots c_{i_d}) | (c_1, c_2, \dots, c_n) \in \{0, 1\}^n, f(c_1, c_2, \dots, c_n) = 1\}$. Note r as the rank of the matrix $S_1^d(f)$, it's also the rank of the vector set $S_1^d(f)$, then it should have $r \leq \min\{wt(f), \sum_{i=0}^d \binom{n}{i}\}$.

Proposition 1. f has no annihilator of degree $\leq d$ if and only if $r = \sum_{i=0}^d \binom{n}{i}$.

Let $f' = 1 + f$, we can get the vector set $S_0^d(f) = \{(1, a_1, a_2, \dots, a_n, a_{i_1} a_{i_2}, \dots, a_{i_1} \dots a_{i_d}) | (a_1, a_2, \dots, a_n) \in \{0, 1\}^n, f(a_1, a_2, \dots, a_n) = 0\}$, which has $2^n - wt(f)$ vectors (each vector is $\sum_{i=0}^d \binom{n}{i}$ dimension). The rank of $S_0^d(f)$, $r' \leq \min\{2^n - wt(f), \sum_{i=0}^d \binom{n}{i}\}$.

Similarly, we have:

Proposition 2. $f' = 1 + f$ has no annihilator of degree $\leq d$ if and only if $r' = \sum_{i=0}^d \binom{n}{i}$.

Note $d_0 = \lceil \frac{n}{2} \rceil$, $r_0 = \sum_{i=0}^{d_0} \binom{n}{i}$, $I = S_0^{d_0}(f) \cup S_1^{d_0}(f)$, then $I = \{(1, c_1, c_2, \dots, c_n, c_{i_1} c_{i_2}, \dots, c_{i_1} \dots c_{i_{\lceil \frac{n}{2} \rceil}}) | (c_1, c_2, \dots, c_n) \in \{0, 1\}^n\}$, and $\forall u \in I$, $dim(u) = r_0$, I is a subset of $V_{r_0} = \{0, 1\}^{r_0}$. Then from Proposition 1 and Proposition 2, we should have:

Proposition 3. Let $f \in B_n$, then $AI_n(f) = \lceil \frac{n}{2} \rceil$ if and only if $r(S_0^{d_0}(f)) = r(S_1^{d_0}(f)) = r_0$.

As $S_0^{d_0}(f) \cup S_1^{d_0}(f) = I$, $S_0^{d_0}(f) \cap S_1^{d_0}(f) = \emptyset$, the problem to construct a boolean function with maximum algebraic immunity is the problem, as Proposition 3 shows, to cut I into two disjoint subsets whose ranks are both r_0 ; and the count of boolean functions with maximum algebraic immunity, is the count of different cut methods of I .

Proposition 4. [6] *Let $f \in B_n$ (n odd) be balanced function and it does not have any annihilator with algebraic degree $< \lceil \frac{n}{2} \rceil$. Then $1 + f$ has no annihilator with algebraic degree $< \lceil \frac{n}{2} \rceil$. Consequently, $AI_n(f) = \lceil \frac{n}{2} \rceil$.*

Proposition 5. *Let $f \in B_n$ (n odd), then $AI_n(f) = \lceil \frac{n}{2} \rceil$ if and only if f is balanced and $r(S_1^{d_0}(f)) = 2^{n-1}$.*

Proof. When n is odd, $d_0 = \lceil \frac{n}{2} \rceil = \frac{n-1}{2}$, $r_0 = \sum_{i=0}^{d_0} \binom{n}{i} = 2^{n-1}$.

If $AI_n(f) = \lceil \frac{n}{2} \rceil$, f must to be balanced and $r(S_1^{d_0}(f)) = r(S_0^{d_0}(f)) = r_0 = 2^{n-1}$;

If $r(S_1^{d_0}(f)) = 2^{n-1}$, then f does not have any annihilator with algebraic degree $< \lceil \frac{n}{2} \rceil$; As f is balanced, then by above proposition, we have $AI_n(f) = \lceil \frac{n}{2} \rceil$. \square

So, when n is odd, the problem to construct a boolean function with maximum algebraic immunity is the problem to choose 2^{n-1} distinct elements from I to form a r_0 rank subset; and the count of boolean functions with maximum algebraic immunity, is the count of different choose methods of the subsets.

Lemma 1. *Let $S_1 = \{u_1, \dots, u_{r_0}\} \subseteq I$, $S_2 = \{v_1, \dots, v_{r_0}\} \subseteq I$, $S_1 \cap S_2 = \emptyset$, $r(S_1) = r(S_2) = r_0$, then for any $u_i \in S_1$, there have at least one $v_j \in S_2$, and note $S'_1 = S_1 \setminus \{u_i\} \cup \{v_j\}$, $S'_2 = S_2 \setminus \{v_j\} \cup \{u_i\}$, then $r(S'_1) = r(S'_2) = r_0$.*

Proof. As $u_i, v_l (1 \leq l \leq r_0)$ s' dimension is r_0 , and $r(S_2) = r_0$, so u_i can be linear expressed by $v_l (1 \leq l \leq r_0)$. Let $u_i = \sum_{l=1}^{r_0} b_l v_l = \sum_{b_l=1, l=1}^{r_0} v_l = v_{l_1} + \dots + v_{l_t}$, then there must have at least one element in $\{v_{l_1}, \dots, v_{l_t}\}$ that is linear independent with vector set $S_1 \setminus \{u_i\}$. If not, all of $\{v_{l_1}, \dots, v_{l_t}\}$ are linearly dependent with vector set $S_1 \setminus \{u_i\}$, as $r(S_1 \setminus \{u_i\}) = r_0 - 1$, $\{v_{l_1}, \dots, v_{l_t}\}$ are all can be linearly expressed by vector set $S_1 \setminus \{u_i\}$, but $u_i = v_{l_1} + \dots + v_{l_t}$, so u_i can be linearly expressed by vector set $S_1 \setminus \{u_i\}$, this is conflict with $r(S_1) = r_0$.

Get any element in $\{v_{l_1}, \dots, v_{l_t}\}$ that is linear independent with vector set $S_1 \setminus \{u_i\}$, and note it as v_j , then $r(S'_1) = r_0$.

For $r(S'_2) = r_0$, one can only noticed from $u_i = v_{i_1} + \dots + v_{i_t}$, then v_j can be linearly expressed by S'_2 , as $r(S_2) = r_0$, so we must have $r(S'_2) = r_0$. \square

Lemma 2. *Let $S_1, S_2 \subseteq I$, $|S_1| = t_1 \geq r_0$, $|S_2| = t_2 \geq r_0$, $S_1 \cap S_2 = \emptyset$, $r(S_1) = r(S_2) = r_0$, then for any $u \in S_1$, there have at least one $v \in S_2$, and note $S'_1 = S_1 \setminus \{u\} \cup \{v\}$, $S'_2 = S_2 \setminus \{v\} \cup \{u\}$, then $r(S'_1) = r(S'_2) = r_0$.*

Proof. From u , one can extend it to a maximum linearly independent vector group, thus the result can be got from Lemma1. \square

Theorem 1. Let $f \in B_n$, $AI_n(f) = \lceil \frac{n}{2} \rceil$, let $u = (1, a_1, a_2, \dots, a_n, a_{i_1} a_{i_2}, \dots, a_{i_1} \dots a_{i_{\lceil \frac{n}{2} \rceil}}) \in S_1^{\lceil \frac{n}{2} \rceil}(f)$, for any element $v = (1, b_1, b_2, \dots, b_n, b_{i_1} b_{i_2}, \dots, b_{i_1} \dots b_{i_{\lceil \frac{n}{2} \rceil}}) \in S_0^{\lceil \frac{n}{2} \rceil}(f)$, let

$$\begin{aligned} & g_{(b_1, b_2, \dots, b_n)}(x_1, x_2, \dots, x_n) \\ &= \begin{cases} f(x_1, x_2, \dots, x_n) + 1, & (x_1, x_2, \dots, x_n) = (a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n) \\ f(x_1, x_2, \dots, x_n), & \text{else} \end{cases} \end{aligned}$$

then there exist at least one $g_{(b_1, b_2, \dots, b_n)}(x_1, x_2, \dots, x_n)$ such that $AI_n(g) = \lceil \frac{n}{2} \rceil$.

Proof. As $AI_n(f) = \lceil \frac{n}{2} \rceil$, we have $r(S_1^{\lceil \frac{n}{2} \rceil}(f)) = r(S_0^{\lceil \frac{n}{2} \rceil}(f)) = r_0$, then by Lemma 2, we know that there exist at least one $g_{(b_1, b_2, \dots, b_n)}(x_1, x_2, \dots, x_n)$ such that $r(S_1^{\lceil \frac{n}{2} \rceil}(g)) = r(S_0^{\lceil \frac{n}{2} \rceil}(g)) = r_0$, so we have $AI_n(g) = \lceil \frac{n}{2} \rceil$. \square

Theorem 2. Let $f \in B_n$ (n odd), $AI_n(f) = \lceil \frac{n}{2} \rceil$. Let $S_0^{\lceil \frac{n}{2} \rceil}(f) = \{u_1, \dots, u_{r_0}\}$, $S_1^{\lceil \frac{n}{2} \rceil}(f) = \{v_1, \dots, v_{r_0}\}$, $u_s = (1, a_1^s, a_2^s, \dots, a_n^s, a_{i_1}^s a_{i_2}^s, \dots, a_{i_1}^s \dots a_{i_{\lceil \frac{n}{2} \rceil}}^s) \in S_0^{\lceil \frac{n}{2} \rceil}(f)$, $s = 1, 2$, for any two elements $v^s = (1, b_1^s, b_2^s, \dots, b_n^s, b_{i_1}^s b_{i_2}^s, \dots, b_{i_1}^s \dots b_{i_{\lceil \frac{n}{2} \rceil}}^s) \in S_1^{\lceil \frac{n}{2} \rceil}(f)$, $s = 1, 2$. Let

$$\begin{aligned} & g_{(b_1^s, b_2^s, \dots, b_n^s)}(x_1, x_2, \dots, x_n) \\ &= \begin{cases} f(x_1, x_2, \dots, x_n) + 1, & (x_1, x_2, \dots, x_n) = (a_1^s, a_2^s, \dots, a_n^s), (b_1^s, b_2^s, \dots, b_n^s), s = 1, 2 \\ f(x_1, x_2, \dots, x_n), & \text{else} \end{cases} \end{aligned}$$

then there exist at least one $g_{(b_1^s, b_2^s, \dots, b_n^s)}(x_1, x_2, \dots, x_n)$ such that $AI_n(g) = \lceil \frac{n}{2} \rceil$.

Proof. As $AI_n(f) = \lceil \frac{n}{2} \rceil$, n odd, we have f is balanced, and so g is also balanced, by Proposition 3, to show $AI_n(g) = \lceil \frac{n}{2} \rceil$, $r(S_1^{\lceil \frac{n}{2} \rceil}(g)) = r_0$ is enough.

As $AI_n(f) = \lceil \frac{n}{2} \rceil$, we have $r(S_1^{\lceil \frac{n}{2} \rceil}(f)) = r_0$. As all of u_i, v_i ($1 \leq i \leq r_0$)'s dimension is r_0 , so u^s can be linear expressed by v_i ($1 \leq i \leq r_0$). Assume we have

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} a_1 & \dots & a_{r_0} \\ b_1 & \dots & b_{r_0} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_{r_0} \end{pmatrix}$$

As $u_1 \neq u_2$, there have at least one k ($1 \leq k \leq r_0$), $a_k \neq b_k$. Assume $a_k = 1, b_k = 0$, As the first elements of all u_i, v_i ($1 \leq i \leq r_0$) are all 1, so the weights of a_i, b_i ($1 \leq i \leq r_0$) are both odd, so there have at least one $l \neq k$ ($1 \leq l \leq r_0$), $b_l = 1$. For the vector set $S_1^{\lceil \frac{n}{2} \rceil}(f) \setminus \{v_k, v_l\} \cup \{u_1, u_2\}$, express them by $S_1^{\lceil \frac{n}{2} \rceil}(f)$,

then follow Theorem 2, we have $v_k, v_l \in S_1^{\lceil \frac{n}{2} \rceil}(f)$, so plus $(1, 0, 0)^T$, assume it's the j column, then certainly $j \neq k, l$, we have $v_j, v_k, v_l \in S_1^{\lceil \frac{n}{2} \rceil}(f)$ which satisfies the condition.

If we have one column $(0, 1, 0)^T$ or $(0, 0, 1)^T$, similarly we can come to this conclusion. Now assume we have only $(0, 0, 0)^T$, $(1, 1, 0)^T$, $(1, 0, 1)^T$, $(0, 1, 1)^T$, $(1, 1, 1)^T$. It can be noticed that any three elements of $(1, 1, 0)^T$, $(1, 0, 1)^T$, $(0, 1, 1)^T$, $(1, 1, 1)^T$ is linearly independent, so if we have three different elements of these four elements, we can get this result.

Persume we have only two elements of those four elements, assume they are $(1, 1, 0)^T$, $(1, 1, 1)^T$, then plus $(0, 0, 0)^T$, we should have $u_1 = u_2$, conflicted! then we assume they are $(1, 1, 0)^T$, $(1, 0, 1)^T$, then plus $(0, 0, 0)^T$, we should have $u_1 = u_2 + u_3$, this is impossible as the first elements of u_1, u_2, u_3 are all 1. So we have at least three different elements of above four elements.

As above, M is invertible at any instance, so we can come to the conclusion. \square

Construction 1 Let $f \in B_n$, $AI_n(f) = \lceil \frac{n}{2} \rceil$, let $u = (1, a_1, a_2, \dots, a_n, a_{i_1}a_{i_2}, \dots, a_{i_1} \dots a_{i_{\lceil \frac{n}{2} \rceil}}) \in S_1^{\lceil \frac{n}{2} \rceil}(f)$ then there exist an element $v = (1, b_1, b_2, \dots, b_n, b_{i_1}b_{i_2}, \dots, b_{i_1} \dots b_{i_{\lceil \frac{n}{2} \rceil}}) \in S_0^{\lceil \frac{n}{2} \rceil}(f)$ (one can get such an element by following the steps of the theorem 1's proof), let

$$g(x_1, x_2, \dots, x_n) = \begin{cases} f(x_1, x_2, \dots, x_n) + 1, & (x_1, x_2, \dots, x_n) = (a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n) \\ f(x_1, x_2, \dots, x_n), & \text{else} \end{cases}$$

then $AI_n(g) = \lceil \frac{n}{2} \rceil$.

Construction 2 Let $f \in B_n$ (n odd), $AI_n(f) = \lceil \frac{n}{2} \rceil$. Let $u^s = (1, a_1^s, a_2^s, \dots, a_n^s, a_{i_1}^s a_{i_2}^s, \dots, a_{i_1}^s \dots a_{i_{\lceil \frac{n}{2} \rceil}}^s) \in S_1^{\lceil \frac{n}{2} \rceil}(f)$, $s = 1, 2$. Then there exist two elements $v^s = (1, b_1^s, b_2^s, \dots, b_n^s, b_{i_1}^s b_{i_2}^s, \dots, b_{i_1}^s \dots b_{i_{\lceil \frac{n}{2} \rceil}}^s) \in S_0^{\lceil \frac{n}{2} \rceil}(f)$, $s = 1, 2$ (one can get such two elements by following the steps of the theorem 2's proof). Let

$$g(x_1, x_2, \dots, x_n) = \begin{cases} f(x_1, x_2, \dots, x_n) + 1, & (x_1, x_2, \dots, x_n) = (a_1^s, a_2^s, \dots, a_n^s), (b_1^s, b_2^s, \dots, b_n^s), s = 1, 2 \\ f(x_1, x_2, \dots, x_n), & \text{else} \end{cases}$$

then $AI_n(g) = \lceil \frac{n}{2} \rceil$.

Construction 3 Let $f \in B_n$ (n odd), $AI_n(f) = \lceil \frac{n}{2} \rceil$. Let $u^s = (1, a_1^s, a_2^s, \dots, a_n^s, a_{i_1}^s a_{i_2}^s, \dots, a_{i_1}^s \dots a_{i_{\lceil \frac{n}{2} \rceil}}^s) \in S_1^{\lceil \frac{n}{2} \rceil}(f)$, $s = 1, 2, 3$. Then there exist three elements $v^s = (1, b_1^s, b_2^s, \dots, b_n^s, b_{i_1}^s b_{i_2}^s, \dots, b_{i_1}^s \dots b_{i_{\lceil \frac{n}{2} \rceil}}^s) \in S_0^{\lceil \frac{n}{2} \rceil}(f)$, $s = 1, 2, 3$ (one can get

such three elements by following the steps of the theorem 3's proof). Let

$$g(x_1, x_2, \dots, x_n) = \begin{cases} f(x_1, x_2, \dots, x_n) + 1, & (x_1, x_2, \dots, x_n) = (a_1^s, a_2^s, \dots, a_n^s), (b_1^s, b_2^s, \dots, b_n^s), s = 1, 2, 3 \\ f(x_1, x_2, \dots, x_n), & \text{else} \end{cases}$$

then $AI_n(g) = \lceil \frac{n}{2} \rceil$.

4 Balance and Algebraic Degree of Our Constructions

This part we will discuss the cryptographic properties of the Boolean functions which we constructed in last section.

Construction 1 interchange one element of $S_1(f)$ with one element of $S_0(f)$, Construction 2 interchange two elements of $S_1(f)$ with two elements of $S_0(f)$, Construction 3 interchange three elements of $S_1(f)$ with three elements of $S_0(f)$, so they both keep the weight of the function, thus surely keep the balance.

Proposition 6. Let $\Delta_2(x_1, \dots, x_n) \in B_n$, $wt(\Delta_2(x_1, \dots, x_n)) = 2$, then $deg(\Delta_2(x_1, \dots, x_n)) = n - 1$.

Proof. Let $\Delta_2(x_1, \dots, x_n)$ is 1 at point (a_1, \dots, a_n) and (b_1, \dots, b_n) , then

$$\begin{aligned} \Delta_2(x_1, \dots, x_n) &= (x_1 + a_1 + 1) \cdots (x_n + a_n + 1) + (x_1 + b_1 + 1) \cdots (x_n + b_n + 1) \\ &= \sum_{i=1}^n (a_i + b_i) \prod_{j=1, j \neq i}^n x_j + \cdots \end{aligned}$$

Because a_i can't all equal to b_i , then at least one $\prod_{j=1, j \neq i}^n x_j$ is exist, thus we have $deg(\Delta_2(x_1, \dots, x_n)) = n - 1$. \square

Then for the functions we constructed by Construction 1, we should have:

1. If $deg(f) < n - 1$, then $deg(g) = n - 1$;
2. If $deg(f) = n$, then $deg(g) = n$;
3. If $deg(f) = n - 1$, then $deg(g) \leq n - 1$;

As we can see, in most instances, new functions by our construction 1 have better algebraic agree than the initial function.

Lemma 3. [21] Let $f \in B_n$, $deg(f) = d$, then $2^{n-d} \leq wt(f) \leq 2^n - 2^{n-d}$.

Proposition 7. Let $\Delta_4(x_1, \dots, x_n) \in B_n$, $wt(\Delta_4(x_1, \dots, x_n)) = 4$, then $n - 2 \leq deg(\Delta_4(x_1, \dots, x_n)) \leq n - 1$.

Proof. First it should have $deg(\Delta_4(x_1, \dots, x_n)) \leq n - 1$ as $wt(\Delta_4(x_1, \dots, x_n))$ is even. Then by the above lemma, we should have $deg(\Delta_4(x_1, \dots, x_n)) \geq n - 2$. This comes to the result. \square

Proposition 8. Let $\Delta_6(x_1, \dots, x_n) \in B_n$, $wt(\Delta_6(x_1, \dots, x_n)) = 6$, then $n - 2 \leq deg(\Delta_6(x_1, \dots, x_n)) \leq n - 1$.

Proof. Similarly it should have $\deg(\Delta_6(x_1, \dots, x_n)) \leq n - 1$ as $wt(\Delta_6(x_1, \dots, x_n))$ is even. Then by the above lemma, we should have $\deg(\Delta_6(x_1, \dots, x_n)) \geq n - 2$. This comes to the result. \square

Then for the functions we constructed by Construction 2, 3, we should have:

1. If $\deg(f) < n - 2$, then $\deg(g) = n - 1$ or $n - 2$;
2. If $\deg(f) = n$, then $\deg(g) = n$;
3. If $\deg(f) = n - 1$ or $n - 2$, then $\deg(g) \leq n - 1$;

For the functions constructed by construction 3 in Dalai[16], their algebraic degree are $2^{\lfloor \log_2 n \rfloor}$. And Dalai[16] showed that linear transformation can provide more boolean functions with maximum algebraic immunity, but linear transformation don't change the algebraic degree.

Let $t = \lfloor \log_2 n \rfloor$, then for a function g we constructed in Construction 1:

1. If $n > 2^t + 1$, then $\deg(g) = n - 1 > 2^t$;
2. If $n = 2^t$, then $\deg(g) = n = 2^t$;
3. If $n = 2^t + 1$, then $\deg(g) \leq n - 1 = 2^t$;

For a function g we constructed in Construction 2,3:

1. If $n > 2^t + 2$, then $\deg(g) \geq n - 2 > 2^t$;
2. If $n = 2^t$, then $\deg(g) = n = 2^t$;
3. If $n = 2^t + 1$, then $\deg(g) \leq n - 1 = 2^t$;
4. If $n = 2^t + 2$, then $\deg(g) \leq n - 1 = 2^t + 1$;

As we can see, in most instances, new functions by our constructions have better algebraic degree than the functions in Dalai[16]. If the initial function have a good algebraic degree, as we constructed a large class of functions, among them there must have some functions which have as high algebraic degree as the initial function. As in most instances, the degree of the initial Boolean function is changed, so they are not the linear transformation of the initial function. Thus we provide many more functions than Dalai[16], and in most instances, we get many functions with higher algebraic degree.

If we have a boolean function with maximum algebraic immunity, but we don't be satisfied with it's other cryptographic properties, we can construct a large class of functions with maximum algebraic immunity from this function, among which we can choose them freely, according to different cryptographic properties. So our constructions give the guidance for the design of Boolean functions to resist algebraic attack, and help to design good cryptographic primitives of cryptosystems.

5 An lower bound of the number of the Boolean functions with maximum algebraic immunity

By our construction, and by Dalai[16] Construction 2, we can get an lower bound of the count of the Boolean functions that have the maximum algebraic immunity. As far as we know, this is the first bound about this count.

First we show the Construction by Dalai[16]:

Construction 4 [16] Let $f \in B_n$,

1. If n is odd then

$$f(x_1, \dots, x_n) = \begin{cases} 0, & \text{for } wt(x_1, \dots, x_n) \leq \lceil \frac{n}{2} \rceil \\ 1, & \text{for } wt(x_1, \dots, x_n) \geq \lceil \frac{n}{2} \rceil \end{cases}$$

2. If n is even then

$$f(x_1, \dots, x_n) = \begin{cases} 0, & \text{for } wt(x_1, \dots, x_n) < \lceil \frac{n}{2} \rceil \\ 1, & \text{for } wt(x_1, \dots, x_n) > \lceil \frac{n}{2} \rceil \\ b \in \{0, 1\}, & \text{for } wt(x_1, \dots, x_n) = \frac{n}{2} \end{cases}$$

Theorem 4. Note $S_n = \{f \in B_n \mid AI_n(f) = \lceil \frac{n}{2} \rceil\}$,

1. If n is odd then

$$|S_n| \geq \frac{1}{3}2^{n-1}(2^{2n-2} + 5) + 2, n \geq 5$$

2. If n is even then probably

$$|S_n| \geq (2^{C_n^{\frac{n}{2}} + 1}) \cdot \left[\frac{(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{2^{n-1}} + 1 \right], n \geq 4$$

Proof. Use our constructions on Dalai[16]'s Construction:

1. For n is odd, there is only one function in Dalai's construction, note it as f_0 , then it's weight is 2^{n-1} . For f_0 , use our construction 1, we can get new functions, each of them is 2 distance with f_0 ; use our construction 2, we can get $C_{2^{n-1}}^2$ more functions, each of them is 4 distance with f_0 ; use our construction 3, we can get $C_{2^{n-1}}^3$ more functions, each of them is 6 distance with f_0 ; together we have

$$\begin{aligned} & C_{2^{n-1}}^1 + C_{2^{n-1}}^2 + C_{2^{n-1}}^3 \\ &= 2^{n-1} + \frac{1}{2}2^{n-1}(2^{n-1} - 1) + \frac{1}{6}2^{n-1}(2^{n-1} - 1)(2^{n-1} - 2) \\ &= \frac{1}{6}2^{n-1}[6 + (2^{n-1} - 1)(2^{n-1} + 1)] \\ &= \frac{1}{6}2^{n-1}(2^{2n-2} + 5) \end{aligned}$$

functions, their algebraic immunity are all $\lceil \frac{n}{2} \rceil$. For $1 + f_0$, use our construction 1, 2, and 3, we can get $\frac{1}{6}2^{n-1}(2^{2n-2} + 5)$ more functions, their algebraic immunity are all $\lceil \frac{n}{2} \rceil$. And their distance to $1 + f_0$ is 2 or 4 or 6, to f_0 is $2^n - 2$ or $2^n - 4$ or $2^n - 6$. So when $n \geq 5$, as the above functions' distance to f_0 is 2 or 4 or 6, these functions are distinct with above functions. So adding $f_0, 1 + f_0$, together we have $\frac{1}{3}2^{n-1}(2^{2n-2} + 5) + 2$ functions, their algebraic immunity are all $\lceil \frac{n}{2} \rceil$. This gives the first part of our result.

2. For n is even, let f is a function from Dalai's construction, then we have $(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}) \leq wt(f) \leq (2^{n-1} + \frac{1}{2}C_n^{\frac{n}{2}})$. Let f_t be any function from Dalai's

construction with weight t , use our Construction 1, we choose any element of $\{(x_1, \dots, x_n) | wt(x_1, \dots, x_n) < \lceil \frac{n}{2} \rceil\}$ and inverse its value, then by Theorem 1, there exists at least one element of $\{(x_1, \dots, x_n) | f(x_1, \dots, x_n) = 1\}$ satisfy the theorem's condition, we concern the elements of the set $\{(x_1, \dots, x_n) | wt(x_1, \dots, x_n) > \lceil \frac{n}{2} \rceil\}$, as there are t elements of $\{(x_1, \dots, x_n) | f(x_1, \dots, x_n) = 1\}$, assume each element have the same probability, then with $\frac{2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}}{t}$ probability, we can find that element in the set $\{(x_1, \dots, x_n) | wt(x_1, \dots, x_n) > \lceil \frac{n}{2} \rceil\}$ (As there don't have only one element which satisfy the theorem's condition, the probability we can find an element in that set will larger than this value), so we can get approximately $(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}) \cdot (\frac{2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}}{t}) = \frac{(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{t}$ such functions from this f_t . Those functions are all weight t , and 2 distance to f_t , one is in the set $\{(x_1, \dots, x_n) | wt(x_1, \dots, x_n) > \lceil \frac{n}{2} \rceil\}$, the other is in the set $\{(x_1, \dots, x_n) | wt(x_1, \dots, x_n) < \lceil \frac{n}{2} \rceil\}$. As we have $C_n^{\frac{n}{2}}$ weight t functions in Dalai's Construction. Note we only count the functions which inverse values in the set $\{(x_1, \dots, x_n) | wt(x_1, \dots, x_n) \neq \lceil \frac{n}{2} \rceil\}$, those functions must be distinct from each other, so we have nearly $\frac{(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{t} \cdot C_n^{\frac{n}{2}}$ weight t functions, their algebraic immunity arrival the maximum value.

Note when $t > 2^{n-1}$, we can inverse the element of the set $\{(x_1, \dots, x_n) | wt(x_1, \dots, x_n) > \lceil \frac{n}{2} \rceil\}$ firstly, then the probability we need is $\frac{2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}}{2^n - t}$, not $\frac{2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}}{t}$, which become a little larger than before. For $f'_t = 1 + f_t$, use our construction, we can get the same number functions as f_t , their weight are all $2^n - t$. Note when $n \geq 4$, f'_t and the functions constructed by it can't be same with f_{2^n-t} and the functions constructed by it, those distinct functions are all distinct from Dalai[16]'s construction. As there are $2C_n^{\frac{n}{2}}$ functions in Dalai[16]'s construction, plus their complements, $2C_n^{\frac{n}{2}} + 1$ functions.

Then we have:

$$\begin{aligned}
|S_n| &\geq 2\{2 \sum_{t=2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}}^{2^{n-1} - 1} (\frac{(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{t} \cdot C_n^{\frac{n}{2}}) + \frac{(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{2^{n-1}} \cdot C_n^{\frac{1}{2}C_n^{\frac{n}{2}}}\} + 2C_n^{\frac{n}{2}} + 1 \\
&> \frac{2(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{2^{n-1}} (\sum_{t=2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}}}^{2^{n-1} - 1} 2C_n^{\frac{n}{2}} + C_n^{\frac{1}{2}C_n^{\frac{n}{2}}}) + 2C_n^{\frac{n}{2}} + 1 \\
&= \frac{(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{2^{n-1}} (2C_n^{\frac{n}{2}} + 1) + 2C_n^{\frac{n}{2}} + 1 \\
&= (2C_n^{\frac{n}{2}} + 1) \cdot [\frac{(2^{n-1} - \frac{1}{2}C_n^{\frac{n}{2}})^2}{2^{n-1}} + 1]
\end{aligned}$$

Thus prove the theorem. \square

The construction in [16] provides only one symmetric Boolean function when n is odd, and $2C_n^{\frac{n}{2}}$ symmetric Boolean functions when n is even. The construction in [15] can provide only one high dimension maximum algebraic immunity Boolean function from a low dimension maximum algebraic immunity Boolean

function, this number is very small, because the count of low dimension Boolean functions with maximum algebraic immunity is much smaller than that of high dimensions. Our constructions, as Theorem 4 shows, can provide much more functions than the former two constructions. And among these functions, we can find some that have good cryptographic properties, which is good for cryptographic use.

6 Conclusion

In this paper we give three construction methods which each can get a class of Boolean functions with maximum algebraic immunity from one such giving function. Our constructions get more functions than any previous construction. The cryptographic properties, such as balance, algebraic degree etc, of those functions are studied. It shows that we can construct Boolean functions with better cryptographic properties, which gives the guidance for the design of Boolean functions to resist algebraic attack, and helps to design good cryptographic primitives of cryptosystems. From these constructions, we get a lower bound of the count of Boolean functions which have maximum algebraic immunity, which shows that such boolean functions are numerous. As far as we know, this is the first bound about this count.

References

1. F. Armknecht. Improving Fast Algebraic Attacks. In FSE 2004, number 3017 in Lecture Notes in Computer Science, pages 65-82. Springer Verlag, 2004.
2. L. M. Batten. Algebraic Attacks over $GF(q)$. In Progress in Cryptology - INDOCRYPT2004, pages 84-91, number 3348, Lecture Notes in Computer Science, Springer-Verlag.
3. A. Botev. On algebraic immunity of some recursively given sequence of correlation immune functions. In Proceedings of XV international workshop on Synthesis and complexity of control systems, Novosibirsk, October 18-23, 2004, pages 8-12 (in Russian).
4. A. Botev. On algebraic immunity of new constructions of filters with high nonlinearity. In Proceedings of VI international conference on Discrete models in the theory of control systems, Moscow, December 7-11, 2004, pages 227-230 (in Russian).
5. A. Botev and Y. Tarannikov. Lower bounds on algebraic immunity for recursive constructions of nonlinear filters. Preprint 2004.
6. A. Canteaut. Open problems related to algebraic attacks on stream ciphers. In WCC 2005, pages 1-10, invited talk.
7. C. Carlet. Improving the algebraic immunity of resilient and nonlinear functions and constructing bent functions. IACR ePrint server, <http://eprint.iacr.org>, 2004/276.
8. J. H. Cheon and D. H. Lee. Resistance of S-boxes against Algebraic Attacks. In FSE2004, number 3017 in Lecture Notes in Computer Science, pages 83-94. Springer Verlag, 2004.
9. J. Y. Cho and J. Pieprzyk. Algebraic Attacks on SOBER-t32 and SOBER-128. In FSE2004, number 3017 in Lecture Notes in Computer Science, pages 49-64. Springer Verlag, 2004.

10. G. M. Constantine. Combinatorial Theory and Statistical Design. John Wiley Sons,1987.
11. N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. In Advances in Cryptology - ASIACRYPT 2002, number 2501 in Lecture Notes in Computer Science, pages 267-287. Springer Verlag, 2002.
12. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. In Advances in Cryptology - EUROCRYPT 2003, number 2656 in Lecture Notes in Computer Science, pages 345-359. Springer Verlag, 2003.
13. N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. In Advances in Cryptology - CRYPTO 2003, number 2729 in Lecture Notes in Computer Science, pages 176-194. Springer Verlag, 2003.
14. D. K. Dalai, K. C. Gupta and S. Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. In INDOCRYPT 2004, pages 92-106, number 3348, Lecture Notes in Computer Science, Springer-Verlag.
15. D. K. Dalai, K. C. Gupta and S. Maitra. Cryptographically Significant Boolean functions: Construction and Analysis in terms of Algebraic Immunity. In FSE 2005. Tobe published in Lecture Notes in Computer Science, Springer-Verlag.
16. D. K. Dalai, S. Maitra and S. Sarkar. Basic Theory in Construction of Boolean Functions with Maximum Possible Annihilator Immunity. www.eprint.org/2005
17. J. F. Dillon. Elementary Hadamard Difference sets. PhD Thesis, University of Maryland,1974.
18. C. Ding, G. Xiao, and W. Shan. The Stability Theory of Stream Ciphers. Number 561 in Lecture Notes in Computer Science. Springer-Verlag, 1991.
19. I. Krasikov. On integral zeros of Krawtchouk polynomials. Journal of Combinatorial Theory, Series A, 74:71-99, 1996.
20. D. H. Lee, J. Kim, J. Hong, J. W. Han and D. Moon. Algebraic Attacks on Summation Generators. In FSE 2004, number 3017 in Lecture Notes in Computer Science, pages34-48. Springer Verlag, 2004.
21. F. J. MacWilliams and N. J. A. Sloane. The Theory of Error Correcting Codes. NorthHolland, 1977.
22. S. Maitra. Boolean functions with important cryptographic properties. PhD Thesis,Indian Statistical Institute, 2000.
23. S. Maitra and P. Sarkar. Maximum Nonlinearity of Symmetric Boolean Functions on Odd Number of Variables. IEEE Transactions on Information Theory, 48(9):2626-2630,September 2002.
24. W. Meier, E. Pasalic and C. Carlet. Algebraic attacks and decomposition of Boolean functions. In Advances in Cryptology - EUROCRYPT 2004, number 3027 in LectureNotes in Computer Science, pages 474-491. Springer Verlag, 2004.
25. P. Savicky. On the bent Boolean functions that are symmetric. European Journal of Combinatorics, 15:407-410, 1994.