New construction of Boolean function with optimum algebraic immunity *

Yongjuan Wang Shuqin Fan Wenbao Han

Information research department , Information engineering university, zhengzhou, 450002

Abstract: Because of the algebraic attacks, a high algebraic immunity is now an important criteria for Boolean functions used in stream ciphers. In this paper, by using the relationship between some flats and support of a n variables Boolean function f, we introduce a general method to determine the algebraic immunity of a Boolean function and finally construct some balanced functions with optimum algebraic immunity.

Keywords: Boolean functions, Algebraic attack, Algebraic immunity, Affine subspace.

1 Introduction

Algebraic attack to LFSR-based stream cipher was proposed by Coutois and Meier[13] in 2003. Its main idea is to deduce the security of a stream cipher to solve an over-defined system of multivariate nonlinear equations whose unknowns are the bits of the initialization of the LFSR. By searching low degree annihilator, some LFSR-based stream ciphers such as Toyocrypt ,LILI-128[12, 15] and SFINKS etc were successfully attacked. This adds a new cryptographic property for designing Boolean functions which is known as algebraic immunity. A high algebraic immunity is now a necessary criteria for Boolean functions used in cryptosystems.

People are interested in constructing the Boolean functions with optimum algebraic immunity, which is $\lceil n/2 \rceil$ for n-variable Boolean function. In [8], an iterative construction of a $2k$-variable Boolean function with algebraic immunity provable equal to k was given. The produced functions have very high algebraic degree and there exists an algorithm giving a very fast way(whose complexity is linear in the number of variables) of computing the output to the function. But the function is not balanced and its nonlinearity is weak. In [1, 9, 14, 2], examples of symmetric functions achieving optimum algebraic immunity were given. Being symmetric, they present a risk if attacks using this peculiarity can be found in the future. Moreover, they do not have high nonlinearities either.

*Supported by 863 Program of China (No.2006AA01Z425) , NSF of China (90704003,60503011), 973 project of China(2007CB807902), Program for New Century Excellent Talents in University and the Science Foundation of Henan Province of China for Distinguished Young Scholars (0612000100)

Carlet in [6] introduce a general method to prove a given function has a prescribed algebraic immunity. By constructing a sequence of very simple flats which are defined by $x_j = \varepsilon, \varepsilon \in \{0, 1\}$, where j runs a subset of $\{0, 1 \cdots, n-1\}$, he give two algorithms to construct Boolean functions with optimum algebraic immunity. In his paper, there are two problems needed to be solved. One is the open problem Carlet gives that finding some flats inequivalent to the above flats to construct boolean functions with given algebraic immunity. Another is that the existence of some disjoint subset I, J of $\{1, 2, \cdots, \binom{n}{\frac{n}{2}}\}$ which satisfy some conditions is assumed without proof when Carlet construct the boolean function with optimum algebraic immunity in even number of variable.

In the present paper, we generalized the method proposed in [6], solved the open problem, i.e., presented some new kind of flats, proved the existence of I, J with optimal possible choice, and finally constructed some balanced boolean functions with maximum algebraic immunity in both odd and even number of variables,which are not affinely equivalent to majority functions.

2 Preliminaries

A Boolean function on n variables is a mapping from F_2^n onto F_2, the finite field with two elements. We denote by B_n the set of all n-variable Boolean functions. The basic representation of a Boolean function $f(x_1, \cdots, x_n)$ is by the output column of its truth table,i.e, a binary string of length 2^n,

$$f = [f(0, \cdots, 0), f(1, 0, \cdots, 0), \cdots, f(1, \cdots, 1)].$$

The Hamming weight $wt(f)$of a Boolean function f on n variables is the size of the support $supp(f) = \{x \in F_2^n : f(x) = 1\}$ of the function. We denote $1_f = supp(f)$, and $0_f = F_2^n \setminus supp(f)$. The support and offset of a vector $supp(a)$ is the situation number of value 1 and 0's (eg.$supp(10101) = \{1, 3, 5\}$, $off(10101) = \{2, 4\}$). We say that a Boolean function f is balanced if its truth table contains an equal number of 1's and 0's, that is, if its Hamming weight equals 2^{n-1}. But the truth table does not give an idea of the algebraic complexity of the function. Any Boolean function has an unique representation as a multivariate polynomial over F_2, called the algebraic normal form(ANF):

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

where the coefficients are in F_2 (without explanation, the notation + denotes the addition in F_2,i.e. the XOR). The algebraic degree $deg(f)$, is the number of variables in the highest order term with nonzero coefficient. A Boolean function is affine if it has degree at most 1 and the set of all affine function is denoted by A_n.

To be cryptographically secure[7], any Boolean function should have high algebraic degree, high nonlinearity, high order resilient, etc. Recently, it has been identified that any combining or filtering function should not have a low degree multiple. More precisely, it is shown in [13] that, given any n-variable Boolean function f, it is always possible to get a Boolean function g with degree at most $\lceil n/2 \rceil$ such that $f * g$ has degree at most $\lceil n/2 \rceil$. While choosing a function

f ,the cryptosystem designer should avoid that the degree of $f * g$ falls much below $\lceil n/2 \rceil$ with a nonzero function g whose degree is also much below $\lceil n/2 \rceil$. Otherwise, resulting low degree multivariate relations involving key bits and output bits of the combining or filter function f allow a very efficient attack. As observed in [16], it is enough to check that f and $f+1$ do not admit nonzero annihilators of low degrees.

Definition 1 Given $f \in B_n$, define $AN(f) = \{g \in B_n | f * g = 0\}$. Any function $g \in AN(f)$ is called an annihilator of f.

Definition 2 Given $f \in B_n$, the algebraic immunity of f is the minimum degree of all nonzero annihilator of f or $f+1$. We denote it by $AI(f)$.

Note that $AI(f) \leq deg(f)$ since $f * (f+1) = 0$ and $AI(f) \leq \lceil n/2 \rceil$. If a function has low nonlinearity, then it must have a low value of AI, this implies that a function with good value of AI have good nonlinearity(see[22]). If a function has optimal algebraic immunity $\lceil n/2 \rceil$ with n odd number of variable, then it is balanced. Hence, the AI property take care of three fundamental properties of Boolean function: balancedness, algebraic degree and nonlinearity.

Definition 3 Let V be a linear subspace of F_2^n with dimension k, s is a non-zero vector of F_2^n. We call the set $\{s+v, v \in V\}$ a k dimension flat (affine subspace).

3 The main result

Theorem 1 Let $f \in B_n$ and k be any positive integer such that $k \leq \lceil n/2 \rceil$. Suppose that there exists a sequence of flats $(A_i)_{1 \leq i \leq r}$ with dimensions $k + d_i (d_i \geq 0)$, such that:

1) $\forall i \leq r, \mid A_i \backslash [1_f \cup \bigcup_{i' < i} A_{i'}] \mid \leq 2^{d_i}$;

2) $0_f \subseteq \bigcup_{1 \leq i \leq r} A_i$.

Then f have no non-zero annihilator of degree strictly less than k.

We can easily get the following Corollary which is Proposition 1 in [6].

Corollary 1 [6] Let $f \in B_n$ and k be any positive integer such that $k \leq \lceil n/2 \rceil$. Suppose that there exists a sequence of flats $(A_i)_{1 \leq i \leq r}$ with dimensions at least k, such that:

1) $\forall i \leq r, \mid A_i \backslash [1_f \cup \bigcup_{i' < i} A_{i'}] \mid \leq 1$;

2) $0_f \subseteq \bigcup_{1 \leq i \leq r} A_i$.

Then f have no non-zero annihilator of degree strictly less than k.

In order to prove the theorem, we need the following lemmas:

Lemma 1 [17] Let $f \in B_n$. Suppose that $wt(f) \geq 2^n - 2^{n-d}$, any annihilator of f has its algebraic degree at least d.

Lemma 2 Let $f \in B_n$ and L be a flat of dimension t, such that $|L \backslash 1_f| \leq 2^{t-d}$ for some integer $1 \leq d \leq t$. Then for any non-zero annihilator g of f,

(1) Either g has its algebraic degree at least d.

(2) Or $g|_L = 0$.

Proof. Let g be an annihilator of f, i.e $gf = 0$. We have $(g|_L)(f|_L) = 0$. As we know, $f|_L$ can be viewed as a Boolean function of t variables. Since $|L\backslash 1_f| \leq 2^{t-d}$, we have $|1_f|_L| \geq 2^t - 2^{t-d}$. Suppose $g|_L \neq 0$, applying Lemma 1, we have $\deg(g|_L) \geq d$,which means $\deg(g) \geq d$. □

Now we turn to prove Theorem 1:

Proof of Theorem 1: We only need to prove that any annihilator g of f whose degree is at most $k-1$ satisfy that $g|_{A_i} = 0$ for every $1 \leq i \leq r$ by induction. Suppose that g is a nonzero annihilator of f whose degree is at most $k-1$.

When $i = 1$, we have $|A_1\backslash 1_f| \leq 2^{d_1}$, which means $|1_f|_{A_1}| \geq 2^{k+d_1} - 2^{d_1}$. From Lemma 2, we have $g|_{A_1} = 0$.

Next we prove $g|_{A_i} = 0$ for $i \geq 2$ by induction. Suppose that for $i' < i$, we have $g|_{A_{i'}} = 0$, i.e., $g|_{\bigcup_{i'<i} A_{i'}} = 0$. Then

$$| A_i\backslash[1_f \cup \bigcup_{i'<i} A_{i'}] | \leq 2^{d_i} \Leftrightarrow |[A_i\backslash 1_f] \cap [A_i\backslash \bigcup_{i'<i} A_{i'}]| \Leftrightarrow [A_i \cap 0_f] \cap [A_i\backslash \bigcup_{i'<i} A_{i'}]| \leq 2^{d_i} \quad (2)$$

Since $1_g \subseteq 0_f$, we have$|[A_i \cap 1_g] \cap [A_i\backslash \bigcup_{i'<i} A_{i'}]| \leq 2^{d_i}$. On the other hand, since $g|_{\bigcup_{i'<i} A_{i'}} = 0$, the set

$$|[A_i \cap 1_g] \cap [A_i\backslash \bigcup_{i'<i} A_{i'}]| = |[A_i \cap 1_g]| \leq 2^{d_i}.$$

Which means $|1_{g+1}|_{A_i}| \geq 2^{k+d_i} - 2^{d_i}$. From Lemma 2 again, we have $g|_{A_i} = 0$. This means for $1 \leq i \leq r$, $g|_{A_i} = 0$. On the other hand, from condition 2), we have $g = 0$ over F_2 and thus reach a contradiction. So f have no non-zero annihilator of degree strictly less than k. This finishes the proof. □

We obtain by applying Theorem 1 to f and to $f+1$ (exhibiting a sequence of flats $(A_i)_{1\leq i\leq r}$ for f and a sequence of flats $(A'_i)_{1\leq i\leq r'}$ for $f+1$) a sub-class of the class of functions with algebraic immunity at least k. First we give the example of the majority function,

$$f(x) = \begin{cases} 0 & if wt(x_1, \cdots, x_n) \leq \lfloor n/2 \rfloor; \\ 1 & if wt(x_1, \cdots, x_n) > \lfloor n/2 \rfloor. \end{cases}$$

In [6], Carlet use Corollary 1 to explain why the majority function has optimum algebraic immunity. We brief the example for our later convenience.

Example 1 [6] Let f be the majority function with n variables. Let $A_{j'}$'s be the vector spaces $\{x \in F_2^n | supp(x) \subseteq supp(a)\}$ where a ranges over the set of vectors of weights at least $k = \lceil n/2 \rceil$, the order being by increasing weights (with any order for vectors of the same weight), and let the A_i's be the flats $\{x \in F_2^n | supp(a) \subseteq supp(x)\}$ where a ranges over the set of vectors of weights at most $n-k$, the order being by decreasing weights. It is easy to see that for every i, the set $|A_i\backslash[1_f \cup \cup_{i'<i} A_{i'}]| = 1$ if A_i has dimension k and otherwise $A_i\backslash \cup_{i'<i} A_{i'}$ equals the singleton containing the vector of minimum weight in A_i. Similar results can be got for A'_j's. From Corollary 1, we have $AI(f) = [n/2]$.

Noticing that the constructed sequence of flats in Example 1 are all the simplest possible ones that the flats A_i''s are the vector spaces of equations $x_j = 0$ (where j ranges over a set depending on i and of size at most $\lfloor n/2 \rfloor$) and the flats A_i's are their translates by the vector $(1, ..., 1)$. Carlet gave the following open problem in [6].

Open Problem [6] Find some flats inequivalent to the flats in Example 1, e.g., some flats have some equations of the form $x_j + x_k = \varepsilon$ or $x_j = \varepsilon$, where $\varepsilon = 0, 1$ to construct boolean functions with given algebraic immunity.

In the following example and examples in the next two sections, we will give some flats having equations of the form $x_j + x_k = \varepsilon$, and thus solve the above open problem.

Example 2 Let f be the majority function with n (odd) variables, $k = \lceil n/2 \rceil$ and $d = \binom{n}{k}$. For $1 \leq i \leq d$, let A_i be the flat $\{x \in F_2^n | supp(a) \subseteq supp(x)\}$ of dimension k where a ranges over the set of vectors of weight $n - k$(whatever the order is).The set $A_i \setminus [1_f \cup \bigcup_{i'<i} A_i]$ is a singleton containing the vector of minimum weight in A_i.

Now we give some other flats different from those flats in Example 1. Let $N = \{1, 2, \cdots, n\}$, denote $P(N)$ the power set of N. For $\{\alpha_1, \alpha_2, \cdots, \alpha_{n-k-2}, \beta_1, \beta_2\} \in P(N)$, define flat

$$A_{(\alpha_1,\alpha_2,\cdots,\alpha_{n-k-2},\beta_1,\beta_2)} = \{(x_1, \cdots, x_n) \in F_2^n : x_{\alpha_i} = \epsilon, x_{\beta_1} + x_{\beta_2} = 1, 1 \leq i \leq n-k-2, \epsilon \in F_2\}. \tag{3}$$

The flats of form Eq.(3) have dimension $k + 1$. It is easy to see that we can choose a sequence of flats of form Eq.(3) $A_{d+1}, \cdots, A_{d+r}$(following $A_1, \cdots, A_d$, with the order of $A_i (i > d)$ being by decreasing of minimum weight vectors in flats and with any order for the same minimum weight) and $A_{d+r+1} = F_2^n$ which satisfy that $|A_i \setminus [1_f \cup \bigcup_{i'<i} A_i]| \leq 2$ for $d + 1 \leq i \leq d + r + 1$. For example, we can choose all $A_{(\alpha_1,\alpha_2,\cdots,\alpha_{n-k-2},\beta_1,\beta_2)}$ with $(\alpha_1, \alpha_2, \cdots, \alpha_{n-k-2}, \beta_1, \beta_2)$ running across all possible ones. From above, we constructed a sequence of flats which satisfy the two conditions in Theorem 1, and thus f have no non-zero annihilator of degree strictly less than $\lfloor n/2 \rfloor$. For n odd, this means $AI(f) = \lfloor n/2 \rfloor$. We can similarly deal with the case of even n.

Actually, we need not to choose across all $A_{(\alpha_1,\alpha_2,\cdots,\alpha_{n-k-2},\beta_1,\beta_2)}$ to construct the sequence of flats we need. For example, we consider the majority function of 7 variables, whose 1_f is the set of vectors of weight at least 4. For $a_i (1 \leq i \leq \binom{7}{4} = 35)$ ranging over vectors of weight 3,the flats $A_i = \{x | supp(a_i) \subseteq supp(x)\}$ have dimension 4. Then we classify the 7 variables vectors of weight 2 into 3 categories by the indices of circular transaction of vectors (see[23]). For a given vector $(a_1, a_2, \cdots, a_n) \in F_2^n$, we define

$$\rho_n^k(a_1, a_2, \cdots, a_n) = (a_{k+1(modn)}, a_{k+2(modn)}, \cdots, a_{k(modn)}), \ where \ k = 0, \cdots, n-1. \tag{4}$$

And the set $G_{(a_1,a_2,\cdots,a_n)} = \{\rho_n^k(a_1, a_2, \cdots, a_n), k = 0, \cdots, n-1.\}$. Then we have :

$$G_{(1010000)} = \{(1010000), (0101000)(0010100), (0001010), (0000101), (1000010), (0100001)\}$$

$$G_{(1100000)} = \{(1100000), (0110000)(0011000), (0001100), (0000110), (0000011), (1000001)\}$$

$$G_{(1001000)} = \{(1001000), (0100100)(0010010), (0001001), (1000100), (0100010), (0010001)\}$$

Define

$$A_{35+i} = \{x \in F_2^7 | x_i = 1, x_{(i+1)mod7} + x_{(i+2)mod7} = 1\}, \text{ for } i = 1, \cdots, 7,$$

$$A_{42+i} = \{x \in F_2^7 | x_i = 1, x_{(i+3)mod7} + x_{(i+4)mod7} = 1\}, \text{ for } i = 1, 2, 3,$$

and

$$A_{46} = \{x \in F_2^7 | x_4 = 1, x_6 + x_7 = 1\}.$$

For $35 < i < 46$, the flats A_i has dimension 5 , the sets $A_i \setminus [1_f \cup \bigcup_{i'<i} A_{i'}]$ have two elements with weight 2, and all of them are disjoint. the set $A_{46} \setminus [1_f \cup \bigcup_{i'<46} A_{i'}]$ is singleton, the element is (0001001). Next we choose

$$A_{46+i} = \{x \in F_2^7 | x_{2i-1} = 0, x_{2i} + x_{2i+1} = 1\}, i = 1, 2, 3. \text{ and } A_{50} = F_2^n.$$

For $1 \leq i \leq 50$, denote

$$B_i = A_i \setminus [1_f \cup \bigcup_{i'<i} A_{i'}].$$

Then $B_i = \{a_i\}$ for $1 \leq i \leq 35$, and is a set with two vectors of weight 2 for $35 < i < 46$. For $i = 46$, B_i is a singleton because only one vector (0001001) of weight 2 is left. For $46 < i < 50$, the set B_i have two vectors of weight 1; To satisfy the relationship $0_f \subseteq [\bigcup_{i \leq 50} A_i]$, we define the last flat to be the full space and $B_{50} = \{(100000), (0000000)\}$.

Thus we show 7 variables majority function has maximum AI by constructing a sequence of flats $A_i (1 \leq i \leq 50)$ which satisfy the two conditions of Theorem 1.

By the result of Theorem 1, we give a method to construct Boolean functions with optimum AI in the next two sections, which are not equivalent to majority functions.

4 Constructing functions with optimum algebraic immunity in odd number of variables

In [3], Canteaut has observed that, if a balanced function f in an odd number n of variables admits no non-zero annihilator of degree at most $\frac{n-1}{2}$, then it has optimum algebraic immunity $\frac{n+1}{2}$(this means that we do not need to check also that $f + 1$ has no non-zero annihilator of degree at most $\frac{n-1}{2}$ for showing that f has optimum algebraic immunity). If a function has optimal algebraic immunity $\lceil n/2 \rceil$, then it is balanced. We deduce the following corollary of Theorem 1:

Corollary 2 Let n be odd and let $A_i (i = 1, \cdots, r)$ be a sequence of flats of F_2^n with dimension $k + d_i, 0 \leq d_i \leq \lfloor n/2 \rfloor$, and such that, for every $1 \leq i \leq r$, the set $A_i \setminus \bigcup_{j<i} A_j$ is non-empty. Then for any choice of $B_i \subseteq A_i \setminus \bigcup_{j<i} A_j$, such that $|B_i| \leq 2^{d_i}$ and $\Sigma_{i=1}^r |B_i| = 2^{n-1}$, the balanced function with support $B = \bigcup_{1 \leq i \leq r} B_i$ and the function with support $F_2^n \setminus \{B\}$ both have optimum algebraic immunity $\lceil n/2 \rceil$.

Proof. The proof is direct and thus omitted. □

Let n be odd, and $d = \binom{n}{\frac{n+1}{2}}$. Let $a_1, \cdots, a_d$ be an ordering of the set of all vectors of weight $\frac{n+1}{2}$. For every a_i, let

$$A_i = \{x \in F_2^n : supp(x) \subseteq supp(a_i)\},$$

where the dimension of A_i is $k = \frac{n+1}{2}$ and the corresponding $d_i = 0$. It is easy to see that for $1 \leq i \leq d$, $A_i \setminus \bigcup_{j<i} A_j$ is nonempty. We choose $b_i \in A_i \setminus \bigcup_{j<i} A_j$ arbitrarily and let $B_i = \{b_i\}$. So we have $1 = |B_i| \leq 2^0$. On the other hand, $\bigcup_{1 \leq i \leq d} A_i$ contains all vectors $\leq \frac{n+1}{2}$.

Next we will try to define some new kind of flats A_{d+i} with dimension greater than $\frac{n+1}{2}$ such that $A_{d+i} \setminus \bigcup_{j<d+i} A_j$ is non-empty. Based on the circular transaction of indices as in Example 2, we choose flats A_{d+i} of form Eq.(3) such that $A_{d+i} \setminus [1_f \cup \bigcup_{j<d+i} A_j]$ is non-empty, with the order being by increasing of maximum weight vectors in flats and with any order for the same maximum weight. The sequence of flats satisfies:

1) $dim(A_{d+i}) = k + 1$.

2) $|A_{d+i} \setminus \bigcup_{j<d+i} A_j| \leq 2$.

Let $B_i = A_{d+i} \setminus \bigcup_{j<d+i} A_j$ for $i > d$. Suppose we find enough flats such that $B = \bigcup B_i$ has cardinality 2^{n-1}, since we have

$$A_i \setminus [\bigcup_{j<i} A_j \cup (F_2^n \setminus B)] = B \cap (A_i \setminus \bigcup_{j<i} A_j) = B_i \tag{5}$$

for every i, and $B \subseteq \bigcup A_i$. The balanced function with support $F_2^n \setminus B$ has no annihilator of degree strictly less than$\lceil n/2 \rceil$ and has maximum algebraic immunity. Consequently, the function with support B has maximum algebraic immunity too.

Example 3 Let $n = 5$ is odd, and let $a_1, \cdots, a_{10}$ be all vectors with weight 3. Let

$$A_i = \{x \in F_2^n : supp(x) \subseteq supp(a_i)\}, 1 \leq i \leq 10$$

be flats of dimension 3, let $B_i = \{b_i\}$, where $b_i \in A_i \setminus (\bigcup_{j<i} A_j)$. Let

$$\bigcup_{i \leq 10} B_i = \{(0,0,0,0,1), (0,1,0,1,0), (1,1,0,0,0), (1,0,0,0,1),$$

$$(1,0,0,1,0), (1,0,1,0,1), (0,1,0,1,1), (1,0,1,1,0), (0,1,1,0,1), (1,1,0,1,0)\}$$

Next we list some flats of dimension larger than 3,

$$A_{11} = \{x \in F_2^n | x_1 + x_2 = 1\},$$

$$A_{12} = \{x \in F_2^n | x_3 + x_4 = 1\},$$

$$A_{13} = F_2^n.$$

Take $B_{11} = \{(1,0,1,1,1), (0,1,1,1,1)\}$, $B_{12} = \{(1,1,0,1,1), (1,1,1,0,1)\}$, $B_{13} = \{(1,1,1,1,1), (1,1,1,1,0)\}$. It is trivial to prove a balanced function with support $B = \bigcup_{i=1}^{13} B_i$ has optimum algebraic immunity 3.

We see that $k = 3$ in Example 3 and $card(A_i \setminus [B \cup \bigcup_{i'<i} A_{i'}]) = 2$ for flats with dimension 4. The bound of Condition 1) of Theorem 1 is tight, the needed number of flats 13 is less than the number $2^{5-1} = 16$ in Example 1. Considering the open problem, we find a sequence of flats having equations of form $x_j + x_k = \epsilon, (\epsilon \in F_2)$ or $x_j = \epsilon$ to construct balanced Boolean functions with maximum algebraic immunity in odd number of variables which are not affine equivalent to majority function.

5 Constructing balanced functions with maximum AI in even number of variables

In this section, we want to construct even variables balanced Boolean functions with maximum AI. In Corollary 3 of [6], Carlet gave a method to construct Boolean function with maximum AI in even variables, where the existence of some disjoint subset I, J of $\{1, 2, \cdots, \binom{n}{\frac{n}{2}}\}$ which satisfy some conditions is assumed without proof. In this section, we proved the existence of I, J with optimal possible choice, improved the result and constructed some balanced Boolean function with maximum AI in even number of variables.

Let n be even and let $a_1, \cdots, a_{\binom{n}{\frac{n}{2}}}$ be an ordering of the set of all vectors with weight $n/2$ in F_2^n. For every $i \in S = \{1, \cdots, \binom{n}{\frac{n}{2}}\}$, denote A_i the flat $\{x \in F_2^n : supp(x) \subseteq supp(a_i)\}$ and A'_i the flat $\{x \in F_2^n : supp(x) \supseteq supp(a_i)\}$.

Proposition 1 (see [6] corollary 3) The denotations are as above, let I, J, K be three disjoint subsets of $\{1, \cdots, \binom{n}{n/2}\}$. Assume that, for every $i \in I$, there exists a vector $b_i \neq a_i$ such that $b_i \in A_i \setminus \bigcup_{i' \in I, i' < i} A_{i'}$. Assume that, for every $i \in J$, there exists a vector $c_i \neq a_i$ such that $c_i \in A'_i \setminus \bigcup_{i' \in J, i' < i} A'_{i'}$. Then the function whose support equals:

$$\{x \in F_2^n | wt(x) > 2\} \cup \{b_i, i \in I\} \cup \{a_i, i \in J \cup K\} \setminus \{c_i, i \in J\}$$

has algebraic immunity $n/2$.

We examine the existence and choice of I, J, K, and find that there exist two disjoint subsets I, J of S with same cardinality $\frac{|S|}{2}$, such that for every $i \in I$, there exists a vector $b_i \neq a_i$ such that $b_i \in A_i \setminus \bigcup_{i' \in I, i' < i} A_{i'}$ and for every $i \in J$, there exists a vector $c_i \neq a_i$ such that $c_i \in A'_i \setminus \bigcup_{i' \in J, i' < i} A'_{i'}$.

Proposition 2 Let the notations be defined as above. There exist two disjoint subsets I, J of S with the same cardinality $|S|/2$, such that for every $i \in I$, $|A_i \setminus \bigcup_{i' \in I, i' < i} A_{i'}| \geq 2$ and $i \in J$, $|A'_i \setminus \bigcup_{i' \in J, i' < i} A'_{i'}| \geq 2$.

Proof. Define $B_i = A_i \setminus \cup_{i' < i} A_{i'}$. We have $B_i \subseteq A_i$, $a_i \in B_i$, and

$$B_i = A_i \setminus \cup_{i' < i} B_{i'} = (((A_i \setminus B_{i-1}) \setminus B_{i-2}) \setminus \cdots) \setminus B_1. \tag{6}$$

From the definition, $B_i \cap B_j = \Phi,\ for\ i \neq j,\ i, j \in S$, the order of $B_{i'}$ in Eq.(6) is changeable. Now we want to find a subset I $(with\ |I| = \frac{1}{2}|S|)$ of S to make $\bigcup_{i \in I} B_i = \{x \in F_2^n | wt(x) < n/2\} \cup \{a_i, i \in I\}$.

In fact, it is feasible to take B_i to be a flat with dimension $n/2 - k (0 \leq k < n/2)$ by carefully choosing A_i(where i ranges over half of S) with the order being by increasing of k. So we have $|B_i| = 2^{n/2-k} \geq 2$.

At first, we randomly choose a vector of weight $n/2$, which we denote by a_1 again, $B_1 = A_1$ is a flat of dimension $n/2$. We denote $l_{n/2} = 1$. Then we choose another vector of weight $n/2$ as a_2, such that $|supp(a_2) \setminus supp(a_1)| = 1$. Denote $L = \{1, \cdots, n\}$. Suppose $\{l_2\} = supp(a_2) \setminus supp(a_1)$, then $B_2 = A_2 \setminus A_1 = \{x_{l_2} = 1, x_j = 0, j \in L \setminus supp(a_2)\}$ is a flat of dimension $n/2 - 1$.

We continue to find new vector with $|supp(a_i) \setminus \cup_{i'<i} supp(a_{i'})| = 1$ until $\bigcup_i supp(a_i) = L$. It is easy to see that we can choose $l_{n/2-1} = n/2$ number of such a_i. Suppose $\{l_i\} = supp(a_i) \setminus \cup_{i'<i} supp(a_{i'})$. Then every flat

$$B_i = A_i \setminus [\bigcup_{i'<i} A_{i'}] = \{x_{l_i} = 1, x_j = 0, j \in L \setminus supp(a_i)\}$$

has dimension $n/2 - 1$. Next we continue to construct B_i of dimension $n/2 - 2$.

Define

$$L^2_{a_i} = \{(i_1, i_2) | i_1, i_2 \in supp(a_i), i_1 < i_2\}, \text{ for } 1 \leq i \leq \binom{n}{n/2}$$

and

$$L_2 = \{(i, j) | 1 \leq i < j \leq n\}.$$

Suppose that we can find a new vector a_i satisfying $|L^2_{a_i} \setminus \bigcup_{i'<i} L^2_{a_{i'}}| = 1$, and $(i_1, i_2) = L^2_{a_i} \setminus \bigcup_{i'<i} L^2_{a_{i'}}$. We choose the above a_i until $\bigcup_{i'<i} L^2_{a_i} = L_2$, and

$$B_i = \{x_{i_1} = 1, x_{i_2} = 1, x_j = 0, j \in L \setminus supp(a_i)\}.$$

It is easy to see that we can choose $l_{n/2-2} = \binom{n}{2} - (\binom{n/2}{2} + \binom{n/2-1}{1} n/2)$ number of such B_i's of dimension $n/2 - 2$.

Following the above idea, we can similarly define $L^k_{a_i}$ and L_k respectively until $k = n/2 - 1$. We choose the new vector a_i satisfying $|L^k_{a_i} \setminus \bigcup_{i'<i} L^k_{a_{i'}}| = 1$ until $\bigcup_{i'<i} L^k_{a_{i'}} = L_k$ and the corresponding

$$B_i = \{x_{i_1} = \cdots = x_{i_k} = 1, x_j = 0, j \in L \setminus supp(a_i), \text{ where } \{i_1, \cdots, i_k\} = L^k_{a_i} \setminus \bigcup_{i'<i} L^k_{a_{i'}}\}$$

is a flat of dimension $n/2 - k$. By induction, we have

$$l_{n/2-k} = \binom{n}{k} - \sum_{i=0}^{k-1} \binom{n/2 - i}{k - i} l_{n/2-i}$$

number of flats with dimension $n/2 - k, (k = 0, 1, \cdots, n/2 - 1)$.

From the configuration of B_i, we can examine that $\bigcup_{i \in I} B_i = \{x | wt(x) < n/2\} \cup \{a_i, i \in I\}$. The number of flats equals $\sum_{k=0}^{n/2-1} l_{n/2-k} = \frac{1}{2}\binom{n}{n/2}$, which explains the existence of I .

For given subset I, we have $J = S \setminus I$. Let $C_i = A'_i \setminus \bigcup_{i' \in J, i'<i} A'_{i'}$, we need to discuss the size of set C_i. Because for arbitrary $s \in \{1, \cdots, n\}$, s appears in the offset of $a_i, i \in J$ the same times with the support of $a_i, i \in I$. We can find a right order of A'_i to make $|C_i| \geq 2$, and $\bigcup_{i \in J} C_i = \{x | wt(x) > n/2\} \cup \{a_i, i \in J\}$. The process is similar to the choice the order of $A_i, i \in I$, with the difference that we discuss the order of the flats by the offset of $a_i, i \in J$ instead of by the support of $a_i, i \in I$, where the offset of a vector a is $\{1, \cdots, n\} \setminus supp(a)$ $(a \in F_2^n)$, which we denote by $off(a)$. □

Next we give an example of Proposition 2.

Example 4 Let $n = 6$, then $|S| = 20$. For simplicity, we use the support to express the vector. Let $supp(I) = \{supp(a_i), i \in I\}$, $off(I) = \{off(a_i), i \in I\}$, we choose the subset I, J as follow:

$$supp(I) = \{(456), (345), (234), (123), (134), (145), (156), (125), (256), (356)\},$$

$$supp(J) = \{(246), (346), (136), (135), (356), (146), (126), (235), (245), (124)\}.$$

The corresponding offset of $a_i (i \in J)$ is :

$$off(J) = \{(135), (125), (245), (246), (124), (235), (345), (146), (136), (356)\}.$$

We can examine that for given $A_i, i \in I(\ A'_i, i \in J)$, whose order is as above, satisfies that $|B_i| \geq 2, i \in I(|C_i| \geq 2, i \in J)$.

The existence of I, J with cardinality $\frac{|S|}{2}$ allows us to construct balanced Boolean function with maximum AI in even variables. We get the following result:

Corollary 3 Let n is even, and the denotations are as above. The balanced function whose support equals:

$$\{x \in F_2^n | wt(x) > \frac{n}{2}\} \cup \{b_i, i \in I\} \cup \{a_i, i \in J\} \setminus \{c_i, i \in J\} \tag{7}$$

has algebraic immunity $n/2$.

Example 5 Let $n = 6$, I, J is as Example 4. So we choose

$$\{b_i, i \in I\} = \{(4), (3), (23), (13), (14), (15), (16), (25), (26), (36)\}$$

$$\{a_i, i \in J\} = \{(246), (346), (136), (135), (356), (146), (126), (235), (245), (124)\}$$

$$\{c_i, i \in J\} = \{(123456), (13456), (1236), (12345), (2356), (1456), (1256), (2345), (1245), (1234)\}.$$

We can examine that the balance function with support as form (7) has optimum AI.

By Proposition 2 (the choice of I, J), we can get a lower bound of the balanced Boolean function with optimum AI we constructed in Corollary 3.

Proposition 3 Let n be even and the notations are as above. By the construction of Corollary 3, we can get at least

$$N = 2(\prod_{k=1}^{n/2} (2^k - 1)^{l_k})^2$$

number of balanced boolean functions with maximum algebraic immunity.

Proof. From the Eq.(7), we know that the number of functions is decided by the size of selecting sets of $\{b_i, i \in I\}$ and $\{c_i, i \in J\}$, i.e., the size of B_i, C_i respectively.

In fact, given a flat $B_i, i \in I$ of dimension $k, 1 \leq k \leq n/2$, B_i consists of 2^k vectors. There are $2^k - 1$ possible choices of b_i satisfying $b_i \neq a_i$. On the other hand, we have l_k number of B_i with dimension k , so the set $\{b_i, i \in I\}$ have $\prod_{k=1}^{n/2} (2^k - 1)^{l_k}$ different choices. From the proof of Proposition 2 we know that $\{c_i, i \in J\}$ has the same number of choices. So we have at

least $N = (\prod_{k=1}^{n/2}(2^k-1)^{l_k})^2$ number of functions of form (7). Furthermore, if f is a function with maximum algebraic immunity , then $f+1$ has maximum algebraic immunity too. So we finally have at least $2(\prod_{k=1}^{n/2}(2^k-1)^{l_k})^2$ balanced Boolean functions with $AI(f) = n/2$ by the construction of Corollary 3. □

6 Conclusion

In this paper, we give a method (Theorem 1) to determine the algebraic immunity of a Boolean function, which can be viewed as a generalization of the main result of [6](Corollary 1). As an application, we solved the open problem in [6] by presenting a sequence of new flats having equation of form $x_j + x_k = \epsilon$, and $x_i = \epsilon(\epsilon \in F_2)$ which satisfy the two conditions of Theorem 1. We construct two kinds of balanced Boolean functions with maximum AI which are not affinely equivalent to majority function. For the even case, we proved the existence of I, J with optimal possible choice, and give a lower bound of the number of balanced functions with maximum AI we constructed. By introducing the new kind flats, we enlarge the size of $A_{k+i}\backslash[supp(f) \cup \bigcup_{j<k+i} A_j]$, and naturally reduce the number of flats to contain 0_f. Furthermore, there are still some problems need to be studied such as constructing more functions by the results in this paper, and whether these functions can achieve high nonlinearities and be robust against fast algebraic attacks etc.

Reference

[1] A.Braeken, J.Lano and B. Preneel. On the algebraic immunity of symmetric Boolean functions[A]. Indocrypt 2005[C], LNCS 3797, Springer Verlag, 2005, 35-48.

[2] A.Canteaut and M. Videau. Symmetric Boolean functions[J]. IEEE Transactions on Information Theory, 2005, IT-51(8): 2791-2811

[3] A.Canteaut.Open problems related to algebraic attacks on stream ciphers[R]. Workshop on coding and cryptograpgy, WCC 2005, 1-10, invited talk.

[4] A.Shamir. Stream cipher:dead or alive[A]. Advances in Cryptology-ASIACRYPT 2004[C]. LNCS 3329, Berlin: Springer-Verlag, 2004, 78.

[5] C.Carlet, D.Dalai, K.Kupta and S.Maitra. Algebraic immunity for cryptographically signigicant Boolean functions: Analysisi and Construction[J]. IEEE Transactions on Information Theory, 2006[36], IT-52(7): 3105-3121.

[6] C.Carlet A method of construction of balanced functions with optimum algebraic immunity[OL]. Cryptology e-print Archive, 2006.

[7] C.Ding, G.Xiao and W.Shan. The stability Theory of Stream Cipher[M]. LNCS 561 , Berlin: Springer-Verlag, 1991.

[8] D.K.Dalai,K.C.Gupta and S.Maitra. Cryptographically Significant Boolean functions: Construction and Analysis in terma of algebraic immunity[A]. Fast Software Encryption 2005[C].LNCS 3557, Springer Verlag, 2005, 98-111.

[9] D.K.Dalai,K.C. Gupta and S.Maitra. Notion ofalgebraic immunity and its evaluation related tofast algebraic attacks[R]. 2nd international workshop on Boolean functions: Aryptography and applictions, BFCA 2006, University of Ronen, France, March 13-15, 2006. Cryptology ePrint Archive, eprint.iacr.org. Report 2006/018, January 2006.

[10] Jin Hong,Dong Hoon Lee, Yonggin Yeom and Daewan Han. A new class of single cycle T-function[A].Fast Software Encryption 2005[C]. LNCS 3557, Berlin:Springer-Verlag, 2005

[11] Martin Boesgaard, Mette Vesterager, Thomas Pedesen,Jesper Chritiansen, Ove Cavenius. Rabbit:A new highperformance stream cipher[A]. Fast Software Encryption 2003[C]. LNCS 2887, Berlin: Springer-Verlag, 2003, 307-329.

[12] M Mihaljevie, H Imai. Cryptanalysis of Toyocrypt-HSIstream cipher IEICE Tran sactions on Fundamentals, vo1.F-B5-A 66-73[OL] **http://www.cs1.esat.sony CO jp /atl/ papers/ IEICE-jan02.pdf.**

[13] N.T. Coutois,Wili Meier. Algebraic attacks on stream ciohers with linear feedback[A]. Adcances in Cryptology-EROCRYPT 2003[C], LNCS 2656,Berlin: Springer-verlag, 2003, 346-359.

[14] N. Ferguson, D. Whiting and B. Schneier et al. Helix: Fast Encryption and Authentication in a Single Cryptographic Primitive[A].Fast Software Encryption 2003 Workshop[C]. LNCS 2887. Berlin Heiderberg: Springer-Verlag, 2004, 330-346

[15] Steve Babbage.Cryptanalysis of LILI-128. Nessie project internal report[OL]. http ://www. Cosic.esat.kuleuven.ac.be/nessie/reports/, January 2001.22.

[16] W.Meier, E. Pasilic and C.Carlet. Algebraic attacks and decomposition of Boolean functions[A]. Advances in Cryptology-Eurocrypt 2004[C]. LNCS 3027,Springer Verlag, 2004, 471-491.

[17] F.Armknecht, On the existence of low-degree equations for algebraic attack[OL], Cryptology eprint Archive, 2004/185.

[18] J.Pieprzyk and C.X.Qu. Fast hashing and rotation-symmetric functions[J]. Journal of Universal Computer Science , 1999, 20-31.