New Construction of Single- and Multi-output Boolean Functions with High Algebraic Immunity

Yefeng He and Wenping Ma

***Abstract*—In order to respond to algebraic attacks, it is important to construct Boolean functions with high algebraic immunity of the graphs. In this paper, the complicated problem of algebraic immunity of the graph is converted to the simpler problem of annihilators of the single-output assistant function. Based on this, we propose a new method for constructing single- and multi-output Boolean functions with high algebraic immunity of the graphs. This method can also give many more general single-output Boolean functions with maximum algebraic immunity.**

***Index Terms*—Boolean function, algebraic attack, algebraic immunity, multivariate equations.**

I. Introduction

Recently, *algebraic attacks* [1], [2] have been paid a lot of attention to by researchers for symmetric ciphers. The idea behind the algebraic attack is to express the cipher as a system of multivariate equations whose solution gives the secret key. The complexity of the attack depends on the degree of these equations. This adds a new cryptographic property for designing single-and multi-output Boolean functions, which is known as *algebraic immunity AI* [3]. A high algebraic immunity is now a necessary criterion for single- and multi- output Boolean functions used as building blocks in cryptographic systems (like, e.g., filtering function, combining function and S-box) [4]-[8].

On the other hand, it is also important to know whether there exist nontrivial low degree annihilating relations between input- and output bits. Corresponding to this, Armknecht and Krause [9] gave the concept of the *algebraic immunity* $AI(gr(f))$ *of the graphs* of Boolean functions, which was a further important design parameter of cryptographic functions. Based on matroid union, they also presented a polynomial time algorithm which obtained multi-output functions with $AI(gr(f)) > d$, where d is a positive integer. So far, it is the only construction in this direction. Meanwhile, they also obtained single-output functions f with maximum $AI(f)$ and $AI(gr(f))$. These single-output functions are either symmetric or almost symmetric.

In this paper, we obtain a sufficient and necessary condition for $AI(gr(f)) \geq d$ by considering the annihilators of the single-output assistant Boolean function. Based on the condition, we obtain a new method for constructing multi-output Boolean functions f with $AI(gr(f)) \geq d$. Using the method, we can obtain many more general single-output Boolean functions with maximum $AI(f)$ and $AI(gr(f))$.

Manuscript received October 29, 2009. This work was supported in part by the National Science Foundation of China under Grant 6077 3002, the 863 Program under Grant 2007AA01Z472, the 111 Project under Grant B08038, and the Project sponsored by SRF for ROCS, SEM.

Y. He is with the Ministry of Education Key Lab. of Computer Network and Information Security, Xidian University, Xi'an, 710071, China. She is also with the school of Communication and Information Engineering, Xi'an Institute of Post and Telecommunications, Xi'an, 710121, China (phone: 86-029-88591689; fax: 86-029-88201174; e-mail: yefenghe2008@hotmail.com).

W. Ma is with the Ministry of Education Key Lab. of Computer Network and Information Security, Xidian University, Xi'an, 710071, China (e-mail: wp_ma@hotmail.com).

II. Preliminaries

A multi-output Boolean function is a mapping $f: F_2^n \to F_2^m$. If $m = 1$, then it is a single-output Boolean function. Any single-output Boolean function f has a unique representation as a multivariate polynomial over F_2, called the *algebraic normal form* (ANF):

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

where the coefficients are in F_2. The algebraic degree $\deg(f)$ is the number of variables in the highest order term with nonzero coefficient. The *Hamming Weight* $wt(f)$ of a Boolean function f is the size of the set $1_f = \{x \in F_2^n \mid f(x) = 1\}$. We denote $0_f = \{x \in F_2^n \mid f(x) = 0\}$. The *support* $\mathrm{supp}(a)$ of a vector a is the set of all situation numbers of value 1's. A Boolean function is said to be *balanced* if its Hamming weight equals 2^{n-1}.

Definition 1: Let f be a single-output Boolean function. Any function g is called an *annihilator* of f if $fg = 0$. The *algebraic immunity* of f is the minimum degree of all nonzero annihilators of f and $f+1$. We denote it by $AI(f)$.

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

It was proved in [10] that $AI(f) \le \lceil n/2 \rceil$. If a function has maximum algebraic immunity with n odd, then it is balanced. Moreover, A.Canteaut also observed the following result.

Proposition 1 [1]: Let f be a single-output Boolean function on n variables, where n is odd. If f is balanced and it does not have any annihilator g with $\deg(g) < \lceil n/2 \rceil$, then $f+1$ has no annihilator g' with $\deg(g') < \lceil n/2 \rceil$. That is $AI(f) = \lceil n/2 \rceil$.

Let V be a linear subspace of F_2^n with dimension k and s be a nonzero vector of F_2^n. We call the set $\{s+v \mid v \in V\}$ a k dimension flat (affine subspace). Based on the flat theory, C.Carlet [11] obtained a sufficient condition for a function f to have no nonzero annihilator of degree strictly less than d. In [12], Y.J. Wang generalized the result.

Proposition 2: Let f be a single-output Boolean function on n variables and $d \le \lceil n/2 \rceil$ be a positive integer. Suppose that there exists a sequence of flats $(A_i)_{1\le i\le r}$ with dimensions $d+k_i (k_i \ge 0)$, such that

1) $\forall i \le r$, $|A_i \setminus [1_f \cup \bigcup_{j<i} A_j]| \le 2^{k_i}$;

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

Then f has no nonzero annihilator of degree strictly less than d.

A function $p: F_2^n \to F_2$ is said to be an annihilator of $S \subseteq F_2^n$ if $p(x) = 0$ for all $x \in S$. The algebraic immunity of S is defined by the minimum degree of all nonzero annihilators of S. Thus, the algebraic immunity of a single-output function f is equal to the minimum of $AI(0_f)$ and $AI(1_f)$. This definition can be easily generalized to a multi-output function f, whose algebraic immunity is defined by the minimum of $AI(z_f)$ over all $z \in F_2^m$.

Since the graph of a function $f: F_2^n \to F_2^m$ is also a set $gr(f) = \{(x, f(x)), x \in F_2^n\} \subseteq F_2^{n+m}$, the algebraic immunity of the graph is defined as follows.

Definition 2: Given a function $f: F_2^n \to F_2^m$, the *algebraic immunity* $AI(gr(f))$ *of the graph* is defined by the minimum degree of all nonzero annihilators of $gr(f)$.

For any function $f: F_2^n \to F_2^m$, it holds that $AI(f) \le AI(gr(f)) \le AI(f)+m$ and $AI(gr(f)) \le d_0$, where $d_0 = \min\{d \mid \sum_{i=0}^{d} C_{n+m}^i > 2^n\}$ [9].

III. Constructing Multi-output Boolean Functions with Large $AI(gr(f))$

The aim of this section is to construct multi-output Boolean functions f with $AI(gr(f)) \ge d$. To make the problem simple, we construct a single-output assistant Boolean function. Given a multi-output Boolean function $f: F_2^n \to F_2^m$, or $f(x) = (x_{n+1}, \cdots, x_{n+m})$. Let

$$F(X) = \begin{cases} 1, & \text{if } X \in gr(f) \\ 0, & \text{if } X \notin gr(f) \end{cases} \quad (2)$$

where $X = (x_1, \cdots, x_n, x_{n+1}, \cdots, x_{n+m})$. Note that $wt(F) = 2^n$, since $|gr(f)| = 2^n$. By considering the annihilators of the function F, we can obtain a sufficient and necessary condition for $AI(gr(f)) \ge d$.

Theorem 1: Let d be a positive integer and $d \le d_0$. $AI(gr(f)) \ge d$ holds if and only if F has no nonzero annihilator of degree strictly less than d.

Proof: Given a function $G: F_2^{n+m} \to F_2$, then G is an annihilator of the set $gr(f)$ if and only if G is an annihilator of the function F. So we have $AI(gr(f)) = AI(1_F)$.

By Theorem 1 and Proposition 2, we can get a new method for constructing multi-output functions f with $AI(gr(f)) \ge d$ $(d \le d_0)$. The construction in detail is given as follows.

Construction 1

Let n, m be positive integers and $d \le d_0, k_i \ge 0$.

(1) Let $(A_i)_{1\le i\le r}$ be a sequence of flats of F_2^{n+m} with dimensions $d+k_i$, and such that the set $A_i \setminus \bigcup_{j<i} A_j$ is non-empty for every $1 \le i \le r$.

(2) For every i, we choose $B_i \subseteq A_i \setminus \bigcup_{j<i} A_j$ such that $|B_i| \le 2^{k_i}$ and $\sum_{i=1}^{r} |B_i| = 2^{n+m} - 2^n$. Let $B = \bigcup_{1\le i\le r} B_i$. If we have $(x_{i_1}, \cdots, x_{i_n}) \ne (x_{j_1}, \cdots, x_{j_n})$ for any $X_i = (x_{i_1}, \cdots, x_{i_n}, x_{i_{n+1}}, \cdots, x_{i_{n+m}})$, $X_j = (x_{j_1}, \cdots, x_{j_n}, x_{j_{n+1}}, \cdots, x_{j_{n+m}}) \in F_2^{n+m} \setminus B$ while $i \ne j$, then B is needed. Otherwise, we will choose a new set B.

(3) Let $1_F = F_2^{n+m} \setminus B$.

(4) For any $(x_1, \cdots, x_n) \in F_2^n$, we choose $(x_1, \cdots, x_n, x_{n+1}, \cdots, x_{n+m}) \in 1_F$. Let $z = (z_1, \cdots, z_m) = f(x) = f(x_1, \cdots, x_n) = (x_{n+1}, \cdots, x_{n+m})$.

So the function $z = f(x)$ is needed.

Proposition 3: Let $f(x)$ be obtained from construction 1, then we have $AI(gr(f)) \ge d$.

Proof: Since

ISBN: 978-988-18210-4-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$A_i \setminus [1_F \cup \bigcup_{j<i} A_j] = A_i \setminus [(F_2^{n+m} \setminus B) \cup \bigcup_{j<i} A_j] = B \cap (A_i \setminus \bigcup_{j<i} A_j) = B_i ,$$

we have $|A_i \setminus [1_F \cup \bigcup_{j<i} A_j]| = |B_i| \le 2^{k_i}$ for every i. On the other hand, we have $0_F = B \subseteq \bigcup_{1\le i\le r} A_i$. According to Theorem 1 and Proposition 2, we obtain $AI(gr(f)) \ge d$.

If $d = d_0$, we can get a function f for which $AI(gr(f))$ is maximum. An example is given as follows.

Example 1: Let $n = 4$ and $m = 2$. Obviously, $d_0 = 2$.

Let $a_1, \cdots, a_{15}$ be all vectors with weight 4, since $C_{n+m}^4 = C_6^4 = 15$. Let $A_i = \{x \in F_2^6 \mid \mathrm{supp}(a_i) \subseteq \mathrm{supp}(x)\}$ be flats of dimension 2, and let $B_i = \{b_i\}$, where $b_i \in A_i \setminus \bigcup_{j<i} A_j$. Then

$$\bigcup_{1\le i\le 15} B_i = \{(111100), (111010), (111001), (110110), (110101), (110011), (101110), (101101), (101011), (011110), (011101), (011011), (100111), (010111), (001111)\}$$

Next we list some flats of dimension 4 (or $d_0 + 2$),

$A_{16} = \{x \in F_2^6 \mid x_1 = 1, x_2 + x_3 = 1\}$,
$A_{17} = \{x \in F_2^6 \mid x_2 = 1, x_3 + x_4 = 1\}$,
$A_{18} = \{x \in F_2^6 \mid x_3 = 1, x_4 + x_5 = 1\}$,
$A_{19} = \{x \in F_2^6 \mid x_4 = 1, x_5 + x_6 = 1\}$,
$A_{20} = \{x \in F_2^6 \mid x_5 = 1, x_6 + x_1 = 1\}$,
$A_{21} = \{x \in F_2^6 \mid x_6 = 1, x_1 + x_2 = 1\}$,
$A_{22} = \{x \in F_2^6 \mid x_1 = 0, x_2 + x_3 = 1\}$,
$A_{23} = \{x \in F_2^6 \mid x_3 = 0, x_4 + x_5 = 1\}$.

Take

$B_{16} = \{(110000), (110100), (110001), (101000)\}$
$B_{17} = \{(011000), (011010), (111000), (010110)\}$
$B_{18} = \{(001100), (001101), (111101), (101010)\}$
$B_{19} = \{(100110), (111110), (000101), (100101)\}$
$B_{20} = \{(000011), (010011), (000111), (011111)\}$
$B_{21} = \{(100001), (100011), (101111), (010001)\}$
$B_{22} = \{(010000), (010100), (001000), (001011)\}$
$B_{23} = \{(000100), (100010)\}$,

such that $|B_i| \le 2^2 = 4$. Let $A_{24} = F_2^6$, and take $B_{24} = \{(001010), (000010), (000001)\}$. Let $B = \bigcup_{1\le i\le 24} B_i$, then $|B| = 48$. Since $1_F = F_2^6 \setminus B$, we have

$$1_F = \{(111111), (111011), (110111), (101100), (011100), (110010), (101001), (100100), (011001), (010101), (001110), (100000), (010010), (001001), (000110), (000000)\}$$

Obviously, for any $X_i = (x_{i_1}, \cdots, x_{i_6})$, $X_j = (x_{j_1}, \cdots, x_{j_6}) \in F_2^6 \setminus B$, we have $(x_{i_1}, \cdots, x_{i_4}) \ne (x_{j_1}, \cdots, x_{j_4})$ while $i \ne j$. For any $(x_1, \cdots, x_4) \in F_2^4$, we choose $(x_1, \cdots, x_6) \in 1_F$. Let $z = (z_1, z_2) = f(x_1, \cdots, x_4) = (x_5, x_6)$, then $AI(gr(f)) = 2$. So we can obtain a multi-output Boolean function with maximum $AI(gr(f))$.

IV. Constructing Single-output Boolean Functions of Maximum Algebraic Immunity

When $m = 1$, we can obtain Boolean functions f with maximum $AI(f)$ and $AI(gr(f))$ by Construction 1. If n is even, we have $AI(f) \le n/2$ and $AI(gr(f)) \le d_0 = n/2 + 1$.

Proposition 4: Let $f(x)$ be a single-output Boolean function on n variables which is obtained from Construction1, where n is even and $d = n/2 + 1$. So we have $AI(f) = n/2$ and $AI(gr(f)) = n/2 + 1$.

Proof: Because $n+1$ is odd, we have $\lceil (n+1)/2 \rceil = n/2 + 1$. Considering the Boolean function $F + 1$, we have $1_{(F+1)} = 0_F = F_2^{n+1} \setminus B$. Since

$$A_i \setminus [1_{(F+1)} \cup \bigcup_{j<i} A_j] = A_i \setminus [(F_2^{n+1} \setminus B) \cup \bigcup_{j<i} A_j] = B \cap (A_i \setminus \bigcup_{j<i} A_j) = B_i$$

we have $|A_i \setminus [1_{(F+1)} \cup \bigcup_{j<i} A_j]| = |B_i| \le 2^{k_i}$. On the other hand, we also have $0_{(F+1)} = B \subseteq \bigcup_{1\le i\le r} A_i$. In accordance with Proposition 2, $F + 1$ has no nonzero annihilator of degree strictly less than $n/2 + 1$.

Furthermore, the Boolean function F is balanced, since $wt(F) = 2^n = 2^{(n+1)-1}$. Then F has no nonzero annihilator of degree strictly less than $n/2 + 1$ by Proposition 1. By Theorem 1, we have $AI(gr(f)) \ge n/2 + 1$. Since $AI(gr(f)) \le n/2 + 1$, we obtain $AI(gr(f)) = n/2 + 1$. From the relationship between $AI(f)$ and $AI(gr(f))$, we get $AI(f) = n/2$.

Example 2: Let $n = 4$ is even and $m = 1$.

We know that $d_0 = n/2 + 1 = 3$. Let $a_1, \cdots, a_{10}$ be all vectors with weight 3, and let $A_i = \{x \in F_2^5 \mid \mathrm{supp}(x) \subseteq \mathrm{supp}(a_i)\} (1 \le i \le 10)$ be flats of dimensions 3. Let $B_i = \{b_i\}$, where $b_i \in A_i \setminus \bigcup_{j<i} A_j$. Let

$$\bigcup_{1\le i\le 10} B_i = \{(11100),(11010),(10110),(01110),$$
$$(11001),\ (10101),(10011),(01101),$$
$$(01011),(00111)\}$$

Next we list some flats of dimension 4,

$A_{11} = \{x \in F_2^5 \mid x_1 + x_2 = 1\}, A_{12} = \{x \in F_2^5 \mid x_3 + x_4 = 1\}$

Take $B_{11} = \{(10001),(01001)\}, B_{12} = \{(00100),(00010)\}$. Let $A_{13} = F_2^5$ and take $B_{13} = \{(11111),(00000)\}$. Let $1_F = B = \bigcup_{1\le i\le 13} B_i$. For any $(x_1,\cdots,x_4) \in F_2^4$, we choose $(x_1,\cdots,x_4,x_5) \in 1_F$. Let $z = f(x) = f(x_1,\cdots,x_4) = x_5$. So we obtain a Boolean function f with maximum $AI(f)$ and $AI(gr(f))$.

V. Conclusion

In this paper, we obtain a new method for constructing multi-output Boolean functions f with $AI(gr(f)) \ge d$. Using the method, we can obtain many more general single-output Boolean functions with maximum $AI(f)$ and $AI(gr(f))$. The results provide more available Boolean functions used as building blocks in cryptographic systems. However, it is necessary to be further studied whether these functions can fulfill additional criteria such as high nonlinearity and balancedness.

Acknowledgment

The authors would like to thank the referees for many valuable suggestions and comments.

References

[1] A. Canteaut, "Open problems related to algebraic attacks on stream ciphers," *Workshop on Coding and Cryptography 2005*, Berlin: Springer-Verlag, 2006, vol. 3969, pp.120-134.

[2] S. Z. Al-Hinai, E. Dawson, M.Henricksen, and L.Simpson, "On the security of the LILI family of stream ciphers against algebraic attacks," *ACISP 2007*, Berlin: Springer-Verlag, 2007, vol.4586, pp.11–28.

[3] W.Meier, E.Pasalic, and C.Carlet, "Algebraic attacks and decomposition of Boolean functions," *Advances in Cryptology –Eurocrypt 2004*, Berlin: Springer-Verlag, 2004, vol.3027, pp. 474–491.

[4] D. K. Dalai, S. Maitra, and S. Sarkar, "Basic theory in construction of Boolean functions with maximum possible annihilator immunity," *Des. Codes. Cryptogr.*, vol. 40, no.1, 2006, pp. 41-58.

[5] N. Li, L.J. Qu, W.F. Qi, G.Z. Feng, C.Li, and D.Q. Xie, "On the construction of Boolean functions with optimal algebraic immunity," *IEEE Trans. Inf. Theory*, vol. 54, Mar. 2008, pp. 13 30-1333.

[6] J. H. Cheon and D. H. Lee, "Resistance of S-boxes against algebraic attacks," *Workshop on Fast Software Encryption 2004*, Berlin: Springer-Verlag, 2004, vol.3017, pp.83–94.

[7] Y. Nawaz, K.C.Gupta, and G. Gong, "Algebraic Immunity of S-boxes based on power mappings: analysis and construction," *IEEE Trans. Inf. Theory*, vol. 55, Sep. 2009, pp.4263-4273.

[8] S. Fischer and W. Meier, "Algebraic immunity of S-boxes and augmented functions," *Workshop on Fast Software Encryption 2007*, Berlin: Springer-Verlag, 2007, vol. 4593, pp.366-381.

[9] F. Armknecht and M. Krause, "Constructing single- and multi-output Boolean functions with maximal algebraic immunity," *ICALP* 2006, Berlin: Springer-Verlag, Part II, vol. 4052, pp.180–191.

[10] C. Carlet, D. K. Dalai, K. C. Gupta, and S. Maitra, "Algebraic immunity for cryptographically significant Boolean functions: analysis and construction," *IEEE Trans. Inf. Theory*, vol. 52, Jul. 2006, pp.3105- 3121.

[11] C. Carlet. (2006, April 16). A method of construction of balanced functions with optimum algebraic immunity [Online]. Available: http://eprint.iacr.org/ 2006/149.

[12] Y.J. Wang, S.Q. Fan, and W.B. Han. (2008, April 16). New construction of Boolean function with optimum algebraic Immunity [Online]. Available: http://eprint.iacr.org/ 2008/176.