Upper Bound for Algebraic Immunity on a Subclass of Maiorana McFarland Class of Bent Functions

Kishan Chand Gupta
Centre for Applied Cryptographic Research
University of Waterloo
Waterloo, ON, N2L 3G1, CANADA
E-mail:kgupta@math.uwaterloo.ca

Yassir Nawaz and Guang Gong
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, ON, N2L 3G1, CANADA
E-mail:{ynawaz, ggong}@uwaterloo.ca

Abstract

Studying algebraic immunity of Boolean functions is recently a very important research topic in cryptography. It is recently proved by Courtois and Meier that for any Bollean function of n-variable the maximum algebraic immunity is $\lceil \frac{n}{2} \rceil$. We provide a simple proof that a large subclass of Maiorana McFarland bent functions on n-variable have a surprisingly low level of algebraic immunity $\leq \lceil \frac{n}{4} \rceil + 2$. To the best of our knowledge we provide for the first time a new upper bound for algebraic immunity for a nontrivial class of Boolean functions. We also discuss that this result has some fascinating implications.

Keywords : Algebraic Immunity, bent functions, nonlinearity.

1 Introduction

Algebraic attack has recently become an important tool in cryptanalysing different stream and block cipher systems. A Boolean function, when used in some cryptosystem, should be designed properly to resist this kind of attack. The cryptographic property of a Boolean function, that resists algebraic attack, is known as Algebraic Immunity ($\mathcal{AI}$). The attack uses overdefined systems of multivariate linear equations to recover the secret key [3, 4, 5, 6, 12]. Given a Boolean function f on n-variables, different kinds of scenarios related to low degree multiples of f have been studied in [5, 12]. The core of the analysis is to find out minimum (or low) degree annihilators of f and $1+f$. To mount the algebraic attack, one needs only the low degree linearly independent annihilators [5, 12] of f and $1+f$. It is known that [5, 12] for any n variable Boolean function f, $\mathcal{AI}_n(f) \leq \lceil \frac{n}{2} \rceil$.

The Maiorana McFarland class $\mathcal{M}$ [11, 9] of bent functions is the set of all the Boolean functions on $\mathbb{F}_2^n = \{(x, y), x, y \in \mathbb{F}_2^{\frac{n}{2}}$, of the form:

$$f(x, y) = x \cdot \pi(y) \oplus g(y) \tag{1}$$

where π is any permutation and g is any $\frac{n}{2}$-variable Boolean function. The number of Maiorana McFarland functions equals $(2^{\frac{n}{2}}! \times 2^{2^{\frac{n}{2}}})$ and this is asymptotically equivalent to $(\frac{2^{\frac{n}{2}+1}}{e})^{2^{\frac{n}{2}}} \sqrt{(}2\pi 2^{\frac{n}{2}})$. Let $n = 2m$. We consider a subclass of Maiorana McFarland function where π is taken to be

$$\pi(y) = (y_1, \cdots, y_m) \times A \tag{2}$$

where A is an invertible matrix given by

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2m} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}$$

and the elements of $A, a_{ij} \in \mathbb{F}_2$. Note that $x \cdot \pi(y)$ is quadratic. Also the number of functions in this subclass is given by $(2^{\frac{n}{2}} - 1)(2^{\frac{n}{2}} - 2) \cdots (2^{\frac{n}{2}} - 2^{\frac{n}{2}-1}) \times 2^{2^{\frac{n}{2}}}$.

2 Preliminaries

Let $\mathbb{F}_2 = GF(2)$ be the finite field of two elements. We consider the domain of an n-variable Boolean function to be the vector space $(\mathbb{F}_2^n, \oplus)$ over $\mathbb{F}_2$, where $\oplus$ is used to denote the addition operator over both $\mathbb{F}_2$ and the vector space $\mathbb{F}_2^n$. An n-variable Boolean function is a map $f : \mathbb{F}_2^n \to \mathbb{F}_2$. and define B_n as the set of all n-variable Boolean functions.

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. (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),

$$f(x_1, \ldots, x_n) = a_0 + \sum_{1 \le i \le n} a_i x_i + \sum_{1 \le i < j \le n} a_{i,j} x_i x_j + \ldots + a_{1,2,\ldots,n} x_1 x_2 \ldots x_n,$$

where the coefficients $a_0, a_i, a_{i,j}, \ldots, a_{1,2,\ldots,n} \in \mathbb{F}_2$. The algebraic degree, $\deg(f)$, is the number of variables in the highest order term with non zero 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.

Towards defining algebraic immunity [5, 12, 7], one needs to consider the multiples of both f and $1 + f$. In this direction we present the following definition

Definition 2.1 *Given $f \in B_n$ the set of n-variable Boolean function, the algebraic immunity is denoted by $\mathcal{AI}_n(f) = \deg(g)$, where $g \in B_n$, is the minimum degree nonzero function such that either $f * g = 0$ or* $(1 + f) * g = 0$

Later we also need the following definition related to the annihilator set of a function.

Definition 2.2 *Given $f \in B_n$, define $AN(f) = \{g \in B_n | g \text{ nonzero}, \ f * g = 0\}$.*

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 crypto systems must have high nonlinearity to prevent linear attacks [10].

Many properties of Boolean functions can be described by the Walsh transform. Let $x = (x_1, \ldots, x_n)$ and $\omega = (\omega_1, \ldots, \omega_n)$ both belonging to $\mathbb{F}_2^n$ and $x \cdot \omega = x_1\omega_1 + \ldots + x_n\omega_n$. Let $f(x)$ be a Boolean function on n variables. Then the *Walsh transform* of $f(x)$ is an integer valued function over $\mathbb{F}_2^n$ which is defined as

$$W_f(\omega) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+x\cdot\omega}.$$

A Boolean function f is balanced iff $W_f(0) = 0$. The nonlinearity of f is given by $nl(f) = 2^{n-1} - \frac{1}{2}\max_{\omega\in\mathbb{F}_2^n} |W_f(\omega)|$. The maximum achievable nonlinearity of an n-variable Boolean function is $2^{n-1} - 2^{\frac{n}{2}-1}$. Functions achieving this value of nonlinearity are called bent and exist only when n is even. In other words, for even number of variables, a Boolean function is called bent if $W_f(\omega) = \pm 2^{\frac{n}{2}}$ for all $\omega \in \mathbb{F}_2^n$.

3 Existence of very low degree annihilator for a Subclass of Maiorana McFarland bent functions

It has been studied using statistical techniques in [12] that any randomly chosen balanced function on large number of variables will have good algebraic immunity with very high probability. in [1] it is shown that most of the Boolean functions on large number of variables are of high algebraic degree and of high nonlinearity. in general. However, when considering a specific construction technique, the number of functions constructed by that method is much lower than the total space of Boolean functions and generally such statistical analysis may not work.

Here we show, as far as algebraic immunity is concerned, a subclass of Maiorana McFarland bent function is possibly the worst. Towards this we prove a simple yet pretty astonishing result.

Theorem 3.1 *The $\mathcal{AI}$ of a subclass of n-variable Maiorana McFarland bent function $f(x,y) = x\cdot\pi(y) \oplus g(y)$, where $f(x,y)$ is defined in equation 1 and π is defined in equation 2, is at most $\lceil\frac{n}{4}\rceil + 2$.*

Proof : Let $f(x,y) = x\cdot\pi(y) \oplus g(y)$ be any function. For simplicity we denote $x\cdot\pi(y)$ by h and and $g(y)$ by g. So we have $f(x,y) = h \oplus g$ Now consider $(h \oplus g)(AN(h) * AN(g))$ which is 0. As g is $\frac{n}{2}$ variable Boolean function we can always find $\deg(AN(g)) \leq \lceil\frac{n}{4}\rceil$, and take $AN(h) = 1 \oplus h$ and as h is quadratic so is $AN(h)$. So $\deg(AN(h) * AN(g)) \leq \lceil\frac{n}{4}\rceil + 2$ ■

The bound is tight and depends on g. If we take g to be zero polynomial then algebraic immunity is 2. So we can say $2 \leq \mathcal{AI} \leq \lceil\frac{n}{4}\rceil + 2$.

4 Implication

It is known that [5, 12] for $f \in B_n$, $\mathcal{AI}_n(f) \leq \lceil\frac{n}{2}\rceil$. There exist a construction method [8] that achieves maximum possible algebraic immunity $\lceil\frac{n}{2}\rceil$. So this bound is tight. But to the best of our knowledge no one has provided any bound on the Algebraic Immunity of a special class of Boolean function. This may be crucial in cipher design. From probability analysis in [2, 1] a cryptographer might select a function with large number of variables to get high algebraic immunity since there is no efficient method to check algebraic immunity of a boolean function with large number of variables (i.e., > 20). This subclass of Maiorana McFarland class of functions is an example which should make cryptographer cautious in selecting a boolean function.

References

[1] C. Carlet. On the Degree, Nonlinearity, Algebraic Thickness, and Nonnormality of Boolean Functions, With Developments on Symmetric Functions. In *IEEE Transactions on Information Theory 50(9): 2178-2185 (2004)*.

[2] C. Carle and P. Gaborit. Hyper-bent functions and cyclic codes. *Journal of Comb. Theory, Series A, To apper.*

[3] J. H. Cheon and D. H. Lee. Resistance of S-boxes against Algebraic Attacks. In *FSE 2004*, number 3017 in Lecture Notes in Computer Science, pages 83–94. Springer Verlag, 2004.

[4] 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.

[5] 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.

[6] 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.

[7] D. K. Dalai, K. C. Gupta and S. Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. In *Progress in Cryptology - INDOCRYPT 2004*, pages 92–106, number 3348, Lecture Notes in Computer Science, Springer-Verlag.

[8] D. K. Dalai, K. C. Gupta and S. Maitra. Cryptographically Significant Boolean functions: Construction and Analysis in terms of Algebraic Immunity. In *FSE 2005*, To be published in Lecture Notes in Computer Science, Springer-Verlag.

[9] J. F. Dillon. *Elementary Hadamard Difference sets.* Ph. D. Thesis, Univ. of Maryland, 1974.

[10] C. Ding, G. Xiao, and W. Shan. *The Stability Theory of Stream Ciphers.* Number 561 in Lecture Notes in Computer Science. Springer-Verlag, 1991.

[11] R. L. McFarland. A family of noncyclic difference sets. *Journal of Comb. Theory, Series A* no. 15, pages 1–10.1973.

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