Perfect Algebraic Immune Functions *

Meicheng Liu, Yin Zhang, and Dongdai Lin

SKLOIS, Institute of Information Engineering, CAS, Beijing 100093, P. R. China
meicheng.liu@gmail.com, zhangy@is.iscas.ac.cn, ddlin@iie.ac.cn

Abstract. A perfect algebraic immune function is a Boolean function with perfect immunity against algebraic and fast algebraic attacks. The main results are that for a perfect algebraic immune balanced function the number of input variables is one more than a power of two; for a perfect algebraic immune unbalanced function the number of input variables is a power of two. Also the Carlet-Feng function on 2^s+1 variables and the modified Carlet-Feng function on 2^s variables are shown to be perfect algebraic immune functions. Furthermore, it is shown that a perfect algebraic immune function behaves good against probabilistic algebraic attacks as well.

Keywords: Boolean functions, Algebraic immunity, Fast algebraic attacks, Probabilistic algebraic attacks

1 Introduction

The study of the cryptanalysis of the filter and combination generators of stream ciphers based on linear feedback shift registers (LFSRs) has resulted in a wealth of cryptographic criteria for Boolean functions, such as balancedness, high algebraic degree, high nonlinearity, high correlation immunity and so on. An overview of cryptographic criteria for Boolean functions with extensive bibliography is given in [5].

In recent years, algebraic and fast algebraic attacks [1, 8, 9] have been regarded as the most successful attacks on LFSR-based stream ciphers. These attacks cleverly use over-defined systems of multi-variable nonlinear equations to recover the secret key. Algebraic attacks make use of the equations by multiplying a non-zero function of low degree, while fast algebraic attacks make use of the equations by linear combination. Thus the algebraic immunity ($\mathcal{AI}$), the minimum algebraic degree of annihilators of f or $f+1$, was introduced in [24] to measure the ability of Boolean functions to resist algebraic attacks, while the notion of (e,d)-resistance against fast algebraic attacks was proposed in [16] for Boolean functions as a counterpart of fast algebraic attacks. It was shown in [8] that $\lceil \frac{n}{2} \rceil$ is maximum algebraic immunity of n-variable Boolean functions. The properties and constructions of Boolean functions with maximum algebraic immunity are researched in a large number of papers, e.g., [11, 12, 18, 19, 6]. While it is unclear what is maximum immunity to fast algebraic attacks despite that the attacks were publicly presented as early as 2003.

It has been demonstrated that the resistance of Boolean functions against fast algebraic attacks is not fully covered by algebraic immunity [10, 2, 20]. For determining the immunity against fast algebraic attacks, F. Armknecht et al. [2] introduced an effective algorithm and showed that a class of symmetric Boolean functions have poor resistance against fast algebraic attacks despite their resistance against algebraic attacks. Later M. Liu et al. [20] stated that almost all the symmetric Boolean functions behavior badly against fast algebraic attacks. Y. Du et al. [13] improved Armknecht's algorithm and got better computation complexity when deciding optimal possible resistance against fast algebraic attacks of Boolean functions. Also P.

* Supported by the National 973 Program of China under Grant 2011CB302400, the National Natural Science Foundation of China under Grants 10971246, 60970152, and 61173134, the Grand Project of Institute of Software of CAS under Grant YOCX285056 and the CAS Special Grant for Postgraduate Research, Innovation and Practice.

Rizomiliotis [27, 28] introduced a method to evaluate the behavior of Boolean functions against fast algebraic attacks using univariate polynomial representation.

A preprocessing of fast algebraic attacks on LFSR-based stream ciphers, which use a Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ as the filter or combination generator, is to find a function g of small degree such that the multiple gf has degree not too large. For any pair of positive integers (e, d) such that $e + d \geq n$, there is a non-zero function g of degree at most e such that gf has degree at most d [9]. Thus f has optimal possible resistance against fast algebraic attacks, if for any pair of positive integers (e, d) such that $e + d < n$ and $e < n/2$, there is no non-zero function g of degree at most e such that gf has degree at most d. Such functions are said to be perfect algebraic immune ($\mathcal{PAI}$). $\mathcal{PAI}$ functions also achieve maximum $\mathcal{AI}$. Note that one can use the fast general attack by splitting the function into two $f = h + l$ with l being the linear part of f [9]. In this case, e equals 1 and d equals the degree of the function f. Thus $\mathcal{PAI}$ functions have algebraic degree at least $n - 1$.

As mentioned above, an $\mathcal{PAI}$ function is a Boolean function with perfect immunity against algebraic and fast algebraic attacks. This raises the question whether there are $\mathcal{PAI}$ functions for arbitrary number of input variables. This open problem was also noticed in [6]. It seems that $\mathcal{PAI}$ functions are quite rare. It was observed in [6] that the Carlet-Feng function on 9 variables is $\mathcal{PAI}$. One can also check that the Carlet-Feng function on 5 variables is $\mathcal{PAI}$ as well (see also [13]). However, no function is shown to be $\mathcal{PAI}$ for arbitrary number of variables. M. Liu et al. [20] proved that no symmetric Boolean functions are $\mathcal{PAI}$. Recently, Y. Zhang et al. [30] showed that for an even number (but not a power of two) of input variables there do not exist $\mathcal{PAI}$ functions in the class of rotation symmetric functions.

In this paper, we study the upper bounds on the immunity to fast algebraic attacks. The immunity against fast algebraic attacks is related to a matrix thanks to Theorem 1 of [2]. By a simple transform on this matrix we obtain a symmetric matrix whose elements are the coefficients of the algebraic normal form of a given Boolean function. We improve the upper bounds on the immunity to fast algebraic attacks by proving that the symmetric matrix is singular. The results are that for an n-variable function, we have: (1) if n is a power of 2 then an $\mathcal{PAI}$ function has degree n; (2) if n is one more than a power of 2 then an $\mathcal{PAI}$ function has degree $n - 1$ (which is also balanced); (3) otherwise, the function is not $\mathcal{PAI}$. We then prove that the Carlet-Fent function, which has degree $n - 1$, is $\mathcal{PAI}$ for n equal to one more than a power of 2, and is almost $\mathcal{PAI}$ for the other cases. Also we prove that the modified Carlet-Fent function, which has degree n, is $\mathcal{PAI}$ for n equal to a power of 2, and is almost $\mathcal{PAI}$ for the other cases. The results show that our bounds on the immunity to fast algebraic attacks are tight, and that the Carlet-Feng functions are optimal against fast algebraic attacks as well as algebraic attacks.

It is a difficult challenge to determine the immunity of Boolean functions against probabilistic algebraic attacks. In 2003, N. Courtois and W. Meier [8] described the probabilistic scenario of algebraic attacks as follows:

S4 There exists a non-zero function g of low degree such that gf can be approximated by a function of low degree with probability $1 - \varepsilon$.

In 2005, A. Braeken and B. Preneel [3] generalized **S4** to the two scenarios:

S4a There exists a non-zero function g of low degree such that $gf = g$ on $\{x \mid f(x) = 0\}$ with probability $1 - \varepsilon$.

S4b There exists a non-zero function g of low degree such that $gf = 0$ on $\{x \mid f(x) = 1\}$ with probability $1 - \varepsilon$.

The probability for the scenario **S4a** is equal to $p = 1 - \frac{\mathrm{d}(gf,g)}{2^n - \mathrm{wt}(f)}$, and equal to $p = 1 - \frac{\mathrm{d}(gf,0)}{\mathrm{wt}(f)}$ for the scenario **S4b**. Then $p = \max\{1 - \frac{\mathrm{d}(gf,g)}{2^{n-1}}, 1 - \frac{\mathrm{d}(gf,0)}{2^{n-1}}\}$ for a balanced function. The authors also gave an example of a function on 6 variables for which probabilistic algebraic attacks perform better than classical algebraic attacks if the length of the LFSR is less than or equal to 18. Later, E. Pasalic [26] claimed that from time complexity point of view deterministic algebraic attacks are in general more efficient than probabilistic ones for practical sizes L (e.g. $L = 256$) of LFSR in the context of their application to certain LFSR-based stream ciphers under an assumption that the minimum distance of the code derived by shortening Reed-Muller code (which depends on the filter function) meets the Gilbert-Varshamov (GV) bound. Nevertheless, one should still verify whether the structure of the function itself allows a low-degree approximation that is satisfied with high probability. This raises the question whether there exist Boolean functions using as filters in an LFSR-based nonlinear filter generator for which probabilistic algebraic attacks outperform deterministic ones for practical sizes L. M. Liu gave a response in [21] that there do exist.

In this paper, it is shown that for a Boolean function with maximum $\mathcal{AI}$ probabilistic algebraic attacks are worse than exhaustive search in the context of the nonlinear filter generator if the length of the LFSR is greater than or equal to 46. Since an $\mathcal{PAI}$ function achieves maximum $\mathcal{AI}$, it behaves good against probabilistic algebraic attacks.

The remainder of this paper is organized as follows. In Section 2 some basic concepts are provided. Section 3 studies the upper bounds on the immunity of Boolean functions against fast algebraic attacks while Section 4 shows that the Carlet-Feng functions and their modifications achieve these bounds. Section 5 states that an $\mathcal{PAI}$ function is immune to probabilistic algebraic attacks. Section 6 concludes the paper.

2 Preliminary

Let $\mathbb{F}_2$ be the binary field. An n-variable Boolean function is a mapping from $\mathbb{F}_2^n$ into $\mathbb{F}_2$. Denote by $\mathbf{B}_n$ the set of all n-variable Boolean functions. An n-variable Boolean function f can be uniquely represented as its truth table, i.e., a binary string of length 2^n,

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

The support of f is given by $\mathrm{supp}(f) = \{x \mid f(x) = 1\}$. The Hamming weight of f, denoted by $\mathrm{wt}(f)$, is the number of ones in the truth table of f. An n-variable function f is said to be balanced if its truth table contains equal number of zeros and ones, that is, $\mathrm{wt}(f) = 2^{n-1}$. The Hamming distance between n-variable functions f and g, denoted by $\mathrm{d}(f, g)$, is the number of $x \in \mathbb{F}_2^n$ at which $f(x) \neq g(x)$. It is well known that $\mathrm{d}(f, g) = \mathrm{wt}(f + g)$.

An n-variable Boolean function f can also be uniquely represented as a multivariate polynomial over $\mathbb{F}_2$,

$$f(x) = \sum_{c \in \mathbb{F}_2^n} a_c x^c,\ a_c \in \mathbb{F}_2,\ x^c = x_1^{c_1} x_2^{c_2} \cdots x_n^{c_n},\ c = (c_1, c_2, \cdots, c_n),$$

called the algebraic normal form (ANF). The algebraic degree of f, denoted by $\deg(f)$, is defined as $\max\{\mathrm{wt}(c) \mid a_c \neq 0\}$.

The Boolean function f can also be uniquely represented as a mapping from $\mathbb{F}_{2^n}$ into $\mathbb{F}_2$:

$$f(x) = \sum_{i=0}^{2^n-1} f_i x^i,\ f_i \in \mathbb{F}_{2^n}, \tag{1}$$

where $f^2(x) \equiv f(x)(\mathrm{mod}\, x^{2^n} - x)$. Expression (1) is called the univariate polynomial representation of the function f. It is well known that $f^2(x) \equiv f(x)(\mathrm{mod}\, x^{2^n} - x)$ if and only if $a_0, a_{2^n-1} \in \mathbb{F}_2$ and for $1 \le i \le 2^n - 2,\ a_{2i \bmod (2^n-1)} = a_i^2$.

The algebraic degree of the function f equals $\max_{a_i \neq 0} \mathrm{wt}(i)$, where $i = \sum_{k=0}^{n-1} i_k 2^k$ is considered as $(i_1, i_2, \cdots, i_n) \in \mathbb{F}_2^n$.

Let α be a primitive element of $\mathbb{F}_{2^n}$. The a_i's of Expression (1) are given by $a_0 = f(0)$, $a_{2^n-1} = f(0) + \sum_{j=0}^{2^n-2} f(\alpha^j)$ and

$$f_i = \sum_{j=0}^{2^n-2} f(\alpha^j)\alpha^{-ij}, \text{ for } 1 \leq i \leq 2^n - 2. \tag{2}$$

For more details with regard to the representation of Boolean functions, we refer to [5].

The algebraic immunity of Boolean functions is defined as follows.

Definition 1 *[24] The algebraic immunity of a function $f \in \mathbf{B}_n$, denoted by $\mathcal{AI}(f)$, is defined as*

$$\mathcal{AI}(f) = \min\{\deg(g) \mid gf = 0 \text{ or } g(f+1) = 0, 0 \neq g \in \mathbf{B}_n\}.$$

If the Boolean function f admits a function g of small degree such that the multiple gf has reasonable degree, then fast algebraic attacks are feasible. Thus the immunity against fast algebraic attacks is related to the degree e of g and the degree d of gf with $e < d$. It was shown in [9] that for an n-variable function f and any positive integer e with $e < n/2$, there is a non-zero function g of degree at most e such that gf has degree at most $n - e$. There are several notions of the immunity of Boolean functions against fast algebraic attacks in previous literatures, such as [16, 25]. The perfect algebraic immune function we define below is actually a Boolean function which is algebraic attack resistant (see [25]) and has degree at least $n - 1$. The condition that the degree is more than or equal to $n - 1$ is necessary for perfect algebraic immune function because for a function of degree less than $n - 1$ the fast general attack uses $e = 1$ and $d = \deg(f) < n - 1$, which implies $e + d < n$.

Definition 2 *Let f be an n-variable Boolean function. The function f is called perfect algebraic immune ($\mathcal{PAI}$) if for any positive integers $e < n/2$ it is necessary that the product gf has degree at least $n - e$ for any non-zero function g of degree at most e.*

A $\mathcal{PAI}$ function also achieves maximum $\mathcal{AI}$. As a matter of fact, if a function does not achieves maximum $\mathcal{AI}$, then it admits a non-zero function g of degree less than $n/2$ such that $gf = 0$ or $gf = g$, and it is not $\mathcal{PAI}$.

3 The immunity of Boolean functions against fast algebraic attacks

In this section, we present the upper bounds on the immunity of Boolean functions against fast algebraic attacks. We first recall the previous results for determining the immunity against fast algebraic attacks, then state our bounds.

Denoted by $\mathcal{W}_i$ the set $\{x \in \mathbb{F}_2^n \mid \mathrm{wt}(x) \leq i\}$ in lexicographic order and by $\overline{\mathcal{W}}_i$ the set $\{x \in \mathbb{F}_2^n \mid \mathrm{wt}(x) \geq i+1\}$ in reverse lexicographic order. For $x \in \mathbb{F}_2^n$, let $\bar{x} = (x_1 + 1, \ldots, x_n + 1)$. If x is the j-th element of $\mathcal{W}_i$, then $\bar{x}$ is the j-th element of $\overline{\mathcal{W}}_i$. Here are some additional notational conventions: for $y, z \in \mathbb{F}_2^n$, let $z \subset y$ be an abbreviation for $\mathrm{supp}(z) \subset \mathrm{supp}(y)$, where $\mathrm{supp}(x) = \{i \mid x_i = 1\}$, and let $y \cap z = (y_1 \wedge z_1, \ldots, y_n \wedge z_n)$, $y \cup z = (y_1 \vee z_1, \ldots, y_n \vee z_n)$, where $\wedge$ and $\vee$ are the AND and OR operations respectively. We can see that $z \subset y$ if and only if $y^z = y_1^{z_1} y_2^{z_2} \cdots y_n^{z_n} = 1$.

Let g be a non-zero function of algebraic degree at most e such that $h = gf$ has algebraic degree at most d. Let

$$f(x) = \sum_{c \in \mathbb{F}_2^n} f_c x^c, \; f_c \in \mathbb{F}_2,$$

$$g(x) = \sum_{z \in \mathcal{W}_e} g_z x^z, \;\; g_z \in \mathbb{F}_2,$$

and

$$h(x) = \sum_{y \in \mathcal{W}_d} h_y x^y, \;\; h_y \in \mathbb{F}_2.$$

For $y \in \overline{\mathcal{W}}_d$, we have $h_y = 0$ and therefore

$$0 = h_y = \sum_{c \in \mathbb{F}_2^n} \sum_{\substack{c \cup z = y \\ z \in \mathcal{W}_e}} f_c g_z = \sum_{z \in \mathcal{W}_e} g_z \sum_{\substack{c \cup z = y \\ c \in \mathbb{F}_2^n}} f_c. \tag{3}$$

The above equations on g_z's are homogeneous linear. Denote the coefficient matrix of the equations by $V(f; e, d)$, which is a $\sum_{i=d+1}^{n} \binom{n}{i} \times \sum_{i=0}^{e} \binom{n}{i}$ matrix with the ij-th element equal to

$$v_{yz} = \sum_{\substack{c \cup z = y \\ c \in \mathbb{F}_2^n}} f_c = \sum_{\substack{y \cap \bar{z} \subset c \subset y \\ z \subset y}} f_c = y^z \sum_{y \cap \bar{z} \subset c \subset y} f_c \tag{4}$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$. Then f admits no non-zero function g of algebraic degree at most e such that $h = gf$ has algebraic degree at most d if and only if the rank of the matrix $V(f; e, d)$ equals the number of g_z's which is $\sum_{i=0}^{e} \binom{n}{i}$, i.e., $V(f; e, d)$ has full column rank (see also [2, 13]).

Theorem 1 *[2, 13] Let $f \in \mathbf{B}_n$. Then there exists no non-zero function g of degree at most e such that the product gf has degree at most d if and only if the matrix $V(f; e, d)$ has full column rank.*

Now we show that performing some column operations on the matrix $V(f; e, d)$ creates a matrix with f_c's as its elements.

Lemma 2 $\sum_{z^* \subset z} v_{yz^*} = f_{y \cap \bar{z}}$.

Proof. Note that $c \cup z = y$ if and only if $c \subset y, z \subset y$ and $y \subset c \cup z$, that is, $y^c = 1, y^z = 1$ and $(c \cup z)^y = 1$. By (4) we have

$$\begin{aligned}
\sum_{z^* \subset z} v_{yz^*} &= \sum_{z^* \subset z} \sum_{c \cup z^* = y} f_c \\
&= \sum_{z^* \subset z} \sum_{c} y^c y^{z^*} (c \cup z^*)^y f_c \\
&= \sum_{c} y^c f_c \sum_{z^* \subset z} y^{z^*} (c \cup z^*)^y \\
&= \sum_{c \subset y} f_c \sum_{\substack{z^* \subset z \cap y \\ y \subset c \cup z^*}} 1 \\
&= \sum_{c \subset y} f_c \sum_{y \cap \bar{c} \subset z^* \subset z \cap y} 1 \\
&= \sum_{c \subset y, c = y \cap \bar{z}} f_c \\
&= f_{y \cap \bar{z}}.
\end{aligned}$$

$\square$

By Lemma 2 we know that the matrix $V(f;e,d)$ can be transformed into a matrix with the ij-th element equal to

$$w_{yz} = f_{y\cap\bar{z}}$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$. Denote the new matrix by $W(f;e,d)$. The ji-th element of $W(f;e,d)$ is equal to

$$w_{\bar{z}\bar{y}} = f_{\bar{z}\cap\bar{\bar{y}}} = f_{y\cap\bar{z}} = w_{yz}$$

since $\bar{z}$ is the j-th element in $\overline{\mathcal{W}}_d$ and $\bar{y}$ is the i-th element in $\mathcal{W}_e$ by the definition of $\overline{\mathcal{W}}_d$ and $\mathcal{W}_e$. Recall that $V(f;e,d)$ and $W(f;e,d)$ are $\sum_{i=d+1}^{n}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrices. Therefore the matrix $W(f;e,n-e-1)$, denoted by $W(f;e)$, is a symmetric $\sum_{i=0}^{e}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrix.

Theorem 3 *Let $f \in \mathbf{B}_n$ and $f(x) = \sum_{c\in\mathbb{F}_2^n} f_c x^c$. Then there exists no non-zero function g of degree at most e such that the product gf has degree at most d if and only if $W(f;e,d)$ has full column rank.*

Proof. Lemma 2 shows that $V(f;e,d)$ and $W(f;e,d)$ have the same rank. Then the theorem follows from Theorem 1. □

Next we concentrate on the upper bounds with respect to the immunity of Boolean functions against fast algebraic attacks. As mentioned in Section 2, for an n-variable function f and any positive integer e with $e < n/2$, there is a non-zero function g of degree at most e such that gf has degree at most $n-e$. This can also be explained by Theorem 1 or Theorem 3. The matrices $V(f;e,n-e)$ and $W(f;e,n-e)$ always have not full column rank since they are $\sum_{i=0}^{e-1}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrices. Then from Theorem 3 the bounds on the immunity to fast algebraic attacks are related to the question whether the symmetric matrix $W(f;e)$ is invertible.

Before stating our main results, we list a useful lemma about the determinant of a symmetric matrix over a field with characteristic 2.

Lemma 4 *Let A be a symmetric $m \times m$ matrix over a field with characteristic 2, and*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{12} & a_{12}^2 & a_{23} & \cdots & a_{2m} \\ a_{13} & a_{23} & a_{13}^2 & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & a_{3m} & \cdots & a_{1m}^2 \end{pmatrix}. \tag{5}$$

If $a_{11} = (m+1) \bmod 2$, then $\det(A) = 0$.

Proof. Let S_m be the symmetric group of degree m. Then

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_m}\prod_{i=1}^{m} a_{i,\sigma(i)} \\
&= \sum_{\sigma\in S_m,\sigma^2=1}\prod_{i=1}^{m} a_{i,\sigma(i)} + \sum_{\sigma\in S_m,\sigma^2\neq 1}\prod_{i=1}^{m} a_{i,\sigma(i)} \\
&\left(\text{since } \prod_{i=1}^{m} a_{i,\sigma(i)} = \prod_{i=1}^{m} a_{\sigma(i),i} = \prod_{i=1}^{m} a_{\sigma(i),\sigma^{-1}(\sigma(i))} = \prod_{i=1}^{m} a_{i,\sigma^{-1}(i)}\right) \\
&= \sum_{\sigma\in S_m,\sigma^2=1}\prod_{i=1}^{m} a_{i,\sigma(i)}.
\end{aligned}$$

If m is odd, then $a_{11} = 0$ and therefore

$$\begin{aligned}
\det(A) &= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1\\ \sigma(1)=j}} a_{1j} \prod_{i=2}^{m} a_{i,\sigma(i)} \\
&= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1\\ \sigma(1)=j}} a_{1j}^2 \prod_{\substack{2\le i\le m\\ i\neq j}} a_{i,\sigma(i)} \\
&\quad \text{(for odd } m \text{ and } \sigma^2 = 1\text{, there is } j' \neq j \text{ such that } \sigma(j') = j'\text{)} \\
&= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1\\ \sigma(1)=j,\sigma(j')=j'}} a_{1j}^2 a_{1j'}^2 \prod_{\substack{2\le i\le m\\ i\neq j,j'}} a_{i,\sigma(i)} \\
&\quad \text{(there is } \sigma' \neq \sigma \text{ with } \sigma'(1) = j',\ \sigma'(j) = j \text{ such that } \prod_{i=1}^{m} a_{i,\sigma'(i)} = \prod_{i=1}^{m} a_{i,\sigma(i)}\text{)} \\
&= 0.
\end{aligned}$$

If m is even, then $a_{11} = 1$ and therefore

$$\begin{aligned}
\det(A) &= \sum_{\substack{\sigma^2=1\\ \sigma(1)=1}} \prod_{i=2}^{m} a_{i,\sigma(i)} + \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1\\ \sigma(1)=j}} a_{1j}^2 \prod_{\substack{2\le i\le m\\ i\neq j}} a_{i,\sigma(i)} \\
&= \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1\\ \sigma(1)=1,\sigma(j)=j}} a_{1j}^2 \prod_{\substack{2\le i\le m\\ i\neq j}} a_{i,\sigma(i)} + \sum_{j=2}^{m} \sum_{\substack{\sigma^2=1\\ \sigma(1)=j}} a_{1j}^2 \prod_{\substack{2\le i\le m\\ i\neq j}} a_{i,\sigma(i)} \\
&= 0.
\end{aligned}$$

□

By the above lemma it holds that $\det(A) = \det(A^{(1,1)})$ if $a_{11} = m \bmod 2$, where $A^{(i,j)}$ is the $(m-1)\times(m-1)$ matrix that results from A by removing the i-th row and the j-th column.

Theorem 5 *Let $f \in \mathbf{B}_n$ and f_{2^n-1} be the coefficient of the monomial $x_1x_2\cdots x_n$ in the ANF of f. Let e be a positive integer less than $n/2$. If $f_{2^n-1} = \binom{n-1}{e} + 1 \bmod 2$, then there exists $g \neq 0$ with degree at most e such that gf has degree at most $n-e-1$.*

Proof. According to Theorem 3 we need to prove that the square matrix $W(f;e)$ is singular when $f_{2^n-1} = \binom{n-1}{e} + 1 \bmod 2$. Let ω_{ij} be the ij-th element of $W(f;e)$. Since $\mathbf{1} = (1,1,\cdots,1)$ and $\mathbf{0} = (0,0,\cdots,0)$ are the first elements in $\overline{\mathcal{W}}_{n-e-1}$ and $\mathcal{W}_e$ respectively, by the definition of $W(f;e)$ we have $\omega_{11} = w_{\mathbf{1},\mathbf{0}} = f_{2^n-1}$. Because $\sum_{i=0}^{e}\binom{n}{i} = \sum_{i=1}^{e}\binom{n-1}{i} + \sum_{i=1}^{e}\binom{n-1}{i-1} + 1 \equiv \binom{n-1}{e} (\bmod 2)$, we know $\omega_{11} = \sum_{i=0}^{e}\binom{n}{i} + 1 \bmod 2$ when $f_{2^n-1} = \binom{n-1}{e} + 1 \bmod 2$. As mentioned previously, $W(f;e)$ is a symmetric $\sum_{i=0}^{e}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrix over $\mathbb{F}_2$. We wish to show that $W(f;e)$ has the form of (5). By the definition of $W(f;e)$ we have $\omega_{1i}^2 = \omega_{1i} = w_{\mathbf{1}z} = f_{\mathbf{1}\cap\bar{z}} = f_{\bar{z}} = f_{\bar{z}\cap\bar{z}} = w_{\bar{z}z} = \omega_{ii}$ where $\bar{z}$ is the i-th element in $\overline{\mathcal{W}}_{n-e-1}$ and z is the i-th element in $\mathcal{W}_e$. It follows from Lemma 4 that the matrix $W(f;e)$ is singular. □

Corollary 6 *Let n be an even number and $f \in \mathbf{B}_n$. If f is balanced, then there exists a non-zero function g with degree at most 1 such that the product gf has degree at most $n-2$.*

Proof. If f is balanced, then $f_{2^n-1} = 0$. For even n, it holds that $\binom{n-1}{1} + 1 \equiv 0 (\mathrm{mod}\, 2)$. Therefore the result follows from Theorem 5. □

From Corollary 6 it seems that for the number n of input variables, odd numbers are better than even ones from a cryptographic point of view (since cryptographic functions are always balanced).

Lucas' theorem states that for positive integers m and i, the following congruence relation holds:

$$\binom{m}{i} \equiv \prod_{k=0}^{s} \binom{m_k}{i_k} (\mathrm{mod}\, 2),$$

where $m = \sum_{k=0}^{s} m_k 2^k$ and $i = \sum_{k=0}^{s} i_k 2^k$ are the binary expansion of m and i respectively. It means that $\binom{n-1}{e} \bmod 2 = 1$ if and only if $e \subset n-1$. Note that $f_{2^n-1} = 1$ if and only if $\deg(f) = n$. Theorem 5 shows that for an n-variable function f with degree n and for $e \subset n-1$, there is a non-zero function g with degree at most e such that gf has degree at most $n-e-1$, and that for an n-variable function f with degree less than n and for $e \not\subset n-1$, there is a non-zero function g with degree at most e such that gf has degree at most $n-e-1$. For the case $n-1 \notin \{2^s, 2^s-1\}$, there are integers e, e^* with $0 < e, e^* < n/2$ such that $e \subset n-1$ and $e^* \not\subset n-1$, and thus an n-variable function is not $\mathcal{PAI}$. This shows that for a $\mathcal{PAI}$ function the number n of input variables is one more than or equal to a power of 2. For $n-1 = 2^s$ (resp. 2^s-1), it holds that $e \not\subset n-1$ (resp. $e \subset n-1$) for positive integer $e < n/2$, and thus an n-variable function with degree equal to n (resp. less than n) is not $\mathcal{PAI}$. Recall that a function on odd number of variables with maximum $\mathcal{AI}$ is always balanced and a $\mathcal{PAI}$ function also achieves maximum $\mathcal{AI}$. The fact that a function has an odd weight if and only if it has degree n gives: if $n = 2^s + 1$ then a $\mathcal{PAI}$ function is balanced; if $n = 2^s$ then a $\mathcal{PAI}$ function has degree n and hence is unbalanced. Consequently the following theorem is obtained.

Theorem 7 *Let $f \in \mathbf{B}_n$. If f is perfect algebraic immune, then n is one more than or equal to a power of 2. Further, if f is balanced and perfect algebraic immune, then n is one more than a power of 2; if f is unbalanced and perfect algebraic immune, then n is a power of 2.*

4 The immunity of Boolean functions against fast algebraic attacks using univariate polynomial representation

In this section we focus on the immunity of Boolean functions against fast algebraic attacks using univariate polynomial representation and show that the bounds presented in the previous section can be achieved.

Recall that $\mathcal{W}_e$ is the set $\{x \in \mathbb{F}_2^n | \,\mathrm{wt}(x) \le e\}$ in lexicographic order and $\overline{\mathcal{W}}_d$ is the set $\{x \in \mathbb{F}_2^n | \,\mathrm{wt}(x) \ge d+1\}$ in reverse lexicographic order. Hereinafter, an element $z = (z_1, z_2, \cdots, z_n)$ in $\mathcal{W}_e$ or $\overline{\mathcal{W}}_d$ is considered as an integer $z_1 + z_2 2 + \cdots + z_n 2^{n-1}$ from 0 to $2^n - 1$, and the operations "+" and "−" may be considered as addition and subtraction operations modulo $2^n - 1$ respectively if there is no ambiguity.

Let f, g, h be n-variable functions and g of algebraic degree at most e satisfy that $h = gf$ has algebraic degree at most d. Let

$$f(x) = \sum_{k=0}^{2^n-1} f_k x^k, \;\; f_k \in \mathbb{F}_{2^n},$$

$$g(x) = \sum_{z \in \mathcal{W}_e} g_z x^z, \;\; g_z \in \mathbb{F}_{2^n},$$

and

$$h(x) = \sum_{y \in \mathcal{W}_d} h_y x^y, \ h_y \in \mathbb{F}_{2^n}$$

be the univariate polynomial representation of f, g and h respectively. For $y \in \overline{\mathcal{W}}_d$, we have $h_y = 0$ and thus

$$0 = h_y = \sum_{\substack{k+z=y \\ z \in \mathcal{W}_e}} f_k g_z = \sum_{z \in \mathcal{W}_e} f_{y-z} g_z. \tag{6}$$

The above equations on g_z's are homogeneous linear. Denote the coefficient matrix of the equations by $U(f; e, d)$, which is a $\sum_{i=d+1}^{n} \binom{n}{i} \times \sum_{i=0}^{e} \binom{n}{i}$ matrix with the ij-th element equal to

$$u_{yz} = f_{y-z} \tag{7}$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$. Then f admits no non-zero function g of algebraic degree at most e such that $h = gf$ has algebraic degree at most d if the rank of the matrix $U(f; e, d)$ equals the number of g_z's which is $\sum_{i=0}^{e} \binom{n}{i}$, i.e., $U(f; e, d)$ has full column rank.

Theorem 8 *Let $f \in \mathbf{B}_n$. Then there exists no non-zero function g of degree at most e such that the product gf has degree at most d if the matrix $U(f; e, d)$ has full column rank.*

Remark 1. If the matrix $U(f; e, d)$ has not full column rank, then Equations (6) have non-zero solutions, but the function g may be not a Boolean function.

Remark 2. The matrix $U(f; e, n-e-1)$, denoted by $U(f; e)$, is symmetric since

$$u_{\bar{z}\bar{y}} = f_{\bar{z}-\bar{y}} = f_{(2^n-1-z)-(2^n-1-y)} = f_{y-z} = u_{yz}.$$

Further, we have

$$u_{y\bar{y}} = f_{y-\bar{y}} = f_{y-(2^n-1-y)} = f_{2y} = f_y^2 = u_{y,0}^2,$$

and therefore $U(f; e)$ has the form of (5).

4.1 Carlet-Feng functions

The class of Carlet-Feng functions was first presented in [14] and further studied by C. Carlet and K. Feng [6]. Such functions have maximum algebraic immunity and good nonlinearity. It was observed through computer experiments by Armknecht's algorithm [2] that the functions also have good behavior against fast algebraic attacks. In [28], P. Rizomiliotis determined the immunity of the Carlet-Feng functions against fast algebraic attacks by computing the linear complexity of a sequence, which is more efficient than Armknecht's algorithm. In this section, we further discuss the immunity of the Carlet-Feng functions against fast algebraic attacks and prove that the functions achieve the bound of Theorem 5.

Let n be an integer and α a primitive element of $\mathbb{F}_{2^n}$. Let $f \in \mathbf{B}_n$ and

$$\mathrm{supp}(f) = \{\alpha^l, \alpha^{l+2}, \cdots, \alpha^{l+2^{n-1}-1}\}, 0 \le l \le 2^n - 2. \tag{8}$$

Then $\mathcal{AI}(f) = \lceil \frac{n}{2} \rceil$ according to [14, 6].

A similar proof of [6, Theorem 2] applies to the following result. Here we give a proof for self-completeness.

Proposition 9 *Let $f(x)=\sum_{i=0}^{2^n-1} f_i x^i (f_i \in \mathbb{F}_{2^n})$ be the univariate representation of the function f of (8). Then $f_0=0$, $f_{2^n-1}=0$, and for $1\le i\le 2^n-2$,*

$$f_i=\frac{\alpha^{-il}}{1+\alpha^{-i/2}}.$$

Hence the algebraic degree of f is equal to $n-1$.

Proof. We have $f_0=f(0)=0$ and $f_{2^n-1}=0$ since f has even Hamming weight and thus algebraic degree less than n. For $1\le i\le 2^n-2$, by Equality (2) we have

$$\begin{aligned} f_i=\sum_{j=0}^{2^n-2} f(\alpha^j)\alpha^{-ij} &= \sum_{j=l}^{l+2^{n-1}-1}\alpha^{-ij}=\alpha^{-il}\sum_{j=0}^{2^{n-1}-1}\alpha^{-ij} \\ &=\alpha^{-il}\frac{1+\alpha^{-i2^{n-1}}}{1+\alpha^{-i}}=\alpha^{-il}\frac{1+\alpha^{-i/2}}{1+\alpha^{-i}}=\frac{\alpha^{-il}}{1+\alpha^{-i/2}}. \end{aligned}$$

It is clear that $f_{2^n-2}\neq 0$ and therefore f has algebraic degree $n-1$. □

Remark 3. For the function f of (8), the ij-th element of the matrix $U(f;e,d)$ $(e\le d)$ is equal to

$$u_{yz}=f_{y-z}=\frac{\alpha^{-yl}\alpha^{zl}}{1+\alpha^{-y/2}\alpha^{z/2}},\ \text{for } (i,j)\neq(1,1),$$

where y is the i-th element in $\overline{\mathcal{W}}_d$ and z is the j-th element in $\mathcal{W}_e$ (since it holds for $(i,j)\neq(1,1)$ that $1\le y-z\le 2^n-2$ when $e\le d$).

Lemma 10 *Let A be an $m\times m$ matrix with the ij-th element $a_{ij}=(1+\beta_i\gamma_j)^{-1}$ in a field K of characteristic 2, $\beta_i,\gamma_j\in K$ and $\beta_i\gamma_j\neq 1$, $1\le i,j\le m$. Then the determinant of A is equal to*

$$\prod_{1\le i<j\le m}(\beta_i+\beta_j)(\gamma_i+\gamma_j)\prod_{1\le i,j\le m}a_{ij}.$$

Furthermore, all the minors of A are non-zero if $\beta_i\neq\beta_j$ and $\gamma_i\neq\gamma_j$ for $i\neq j$.

Proof. The second half part of this lemma is derived from the first half part. The proof of the first half part is given by induction on m. First we can check that the statement is certainly true for $m=1$. Now we verify the induction step. Suppose that it holds for $m-1$. Thus we suppose that

$$\det(A^{(1,1)})=\prod_{2\le i<j\le m}(\beta_i+\beta_j)(\gamma_i+\gamma_j)\prod_{2\le i,j\le m}a_{ij},$$

where $A^{(i,j)}$ is the $(m-1)\times(m-1)$ matrix that results from A by removing the i-th row and the j-th column.

We wish to show that it also holds for m. Let $B=(b_{ij})_{m\times m}$ with $b_{1j}=a_{1j}$ and for $i>1$,

$$\begin{aligned} b_{ij}&=a_{ij}+a_{11}^{-1}a_{i1}a_{1j} \\ &=\frac{1}{1+\beta_i\gamma_j}+(\frac{1}{1+\beta_1\gamma_1})^{-1}\cdot\frac{1}{1+\beta_i\gamma_1}\cdot\frac{1}{1+\beta_1\gamma_j} \\ &=\frac{(1+\beta_i\gamma_1)(1+\beta_1\gamma_j)+(1+\beta_1\gamma_1)(1+\beta_i\gamma_j)}{(1+\beta_i\gamma_j)(1+\beta_i\gamma_1)(1+\beta_1\gamma_j)} \\ &=\frac{\beta_i\gamma_1+\beta_1\gamma_j+\beta_1\gamma_1+\beta_i\gamma_j}{(1+\beta_i\gamma_j)(1+\beta_i\gamma_1)(1+\beta_1\gamma_j)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{(\beta_1+\beta_i)(\gamma_1+\gamma_j)}{(1+\beta_i\gamma_j)(1+\beta_i\gamma_1)(1+\beta_1\gamma_j)} \\
&= a_{ij}\cdot(\beta_1+\beta_i)a_{i1}\cdot(\gamma_1+\gamma_j)a_{1j}.
\end{aligned}
$$

Let

$$P = \mathrm{diag}(1, (\beta_1+\beta_2)a_{21}, \cdots, (\beta_1+\beta_m)a_{m1})$$

and

$$Q = \mathrm{diag}(1, (\gamma_1+\gamma_2)a_{12}, \cdots, (\gamma_1+\gamma_2)a_{1m})$$

where $\mathrm{diag}(x_1, \cdots, x_m)$ denotes a diagonal matrix whose diagonal entries starting in the upper left corner are $x_1, \cdots, x_m$. Then

$$B = P\begin{pmatrix} a_{11} & * \\ 0 & A^{(1,1)} \end{pmatrix} Q.$$

Hence

$$
\begin{aligned}
\det(A) &= \det(B) \\
&= \det(P)\cdot a_{11}\det(A^{(1,1)})\cdot\det(Q) \\
&= \left(\prod_{i=2}^{m}(\beta_1+\beta_i)a_{i1}\right)\cdot a_{11}\det(A^{(1,1)})\cdot\left(\prod_{j=2}^{m}(\gamma_1+\gamma_j)a_{1j}\right) \\
&= \prod_{1\le i<j\le m}(\beta_i+\beta_j)(\gamma_i+\gamma_j)\prod_{1\le i,j\le m}a_{ij}.
\end{aligned}
$$

It has now been proved by mathematical induction that the first half part of this lemma holds for all positive integers m. □

Lemma 11 *Let $A = (a_{ij})_{m\times m}$ and $B = (b_{ij})_{m\times m}$ be $m\times m$ matrices with $a_{ij} = \beta_i\gamma_j b_{ij}$ and $\beta_i \neq 0, \gamma_j \neq 0$ for $1 \le i,j \le m$. Then $\det(A)\neq 0$ if and only if $\det(B)\neq 0$.*

Proof. Let $P = \mathrm{diag}(\beta_1, \beta_2, \cdots, \beta_m)$ and $Q = \mathrm{diag}(\gamma_1, \gamma_2, \cdots, \gamma_m)$. Then $A = PBQ$ and hence $\det(A) = \det(B)\prod_{i=1}^{m}\beta_i\gamma_i$, which proves this lemma. □

Proposition 12 *For the function f of (8), $U(f;e)$ is invertible for $e \not\subset n-1$ and $U(f;e,n-e-2)$ has full column rank for $e \subset n-1$.*

Proof. Let $U = U(f;e)$. We have $U_{11} = f_{2^n-1} = 0$. Note that U is a symmetric matrix of order $\sum_{i=0}^{e}\binom{n}{i}$ with the form of (5) (see Remark 2). For the case $e \not\subset n-1$, by Lucas' theorem we have $\binom{n-1}{e} \bmod 2 = 0$ and therefore $\sum_{i=0}^{e}\binom{n}{i} \bmod 2 = U_{11}$. By Lemma 4 it holds that $\det(U) = \det(U^{(1,1)})$. Remark 3 shows that the ij-th element of $U^{(1,1)}$ is

$$U_{ij}^{(1,1)} = \frac{\alpha^{-yl}\alpha^{zl}}{1+\alpha^{-y/2}\alpha^{z/2}},$$

where y is the $(i+1)$-th element in $\overline{\mathcal{W}}_d$ and z is the $(j+1)$-th element in $\mathcal{W}_e$. Let U^* be a $(\sum_{i=0}^{e}\binom{n}{i}-1)\times(\sum_{i=0}^{e}\binom{n}{i}-1)$ matrix with the ij-th element equal to

$$U_{ij}^* = \frac{1}{1+\alpha^{-y/2}\alpha^{z/2}}.$$

Lemma 11 shows that if $\det(U^*)\neq 0$ then $\det(U^{(1,1)})\neq 0$. Since $y \in \overline{\mathcal{W}}_d \setminus \{2^n-1\}$ and $z \in \mathcal{W}_e\setminus\{0\}$, it is derived from Lemma 10 that $\det(U^*)\neq 0$. Thus $\det(U) = \det(U^{(1,1)})\neq 0$.

For the case $e \subset n-1$, we consider the $\sum_{i=0}^{e+1}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrix $U(f; e, n-e-2)$. Let U^{**} be the $\sum_{i=0}^{e}\binom{n}{i} \times \sum_{i=0}^{e}\binom{n}{i}$ matrix that results from $U(f; e, n-e-2)$ by removing the first $\binom{n}{e+1}$ rows. A similar proof of $\det(U^*) \neq 0$ also applies to $\det(U^{**}) \neq 0$. Then $U(f; e, n-e-2)$ has full column rank. □

Theorem 13 *Let e be a positive integer less than $n/2$ and f be the function of (8). Then f admits no non-zero function g with degree at most e such that gf has degree at most $n-e-1$ if $\binom{n-1}{e} \equiv 0 (\bmod 2)$, and admits no non-zero function g with degree at most e such that gf has degree at most $n-e-2$ if $\binom{n-1}{e} \equiv 1 (\bmod 2)$.*

Proof. It is derived from Theorem 8 and Proposition 12. □

Corollary 14 *Let $n = 2^s + 1$ and $f \in \mathbf{B}_n$ be the function of (8). Then f is $\mathcal{PAI}$.*

Proof. It is obtained from Theorem 13. □

Theorem 13 states that the Carlet-Feng functions achieve the bounds of Theorem 5 and thus the bounds of Theorem 5 are tight for the functions with algebraic degree equal to $n-1$, while Corollary 14 states that the Carlet-Feng function on $2^s + 1$ variables is $\mathcal{PAI}$.

Next we consider the Boolean functions with algebraic degree equal to n.

Let n be an integer and α a primitive element of $\mathbb{F}_{2^n}$. Let $f \in \mathbf{B}_n$ and

$$\mathrm{supp}(f) = \{\alpha^l, \alpha^{l+2}, \cdots, \alpha^{l+2^{n-1}-2}\}, 0 \leq l \leq 2^n - 2. \tag{9}$$

Then $\mathcal{AI}(f) = \lfloor \frac{n}{2} \rfloor$ according to [6].

A similar proof of Proposition 9 applies to the following result.

Proposition 15 *Let $f(x) = \sum_{i=0}^{2^n-1} f_i x^i (f_i \in \mathbb{F}_{2^n})$ be the univariate representation of the function f of (9). Then $f_0 = 0$, $f_{2^n-1} = 1$, and for $1 \leq i \leq 2^n - 2$,*

$$f_i = \frac{\alpha^{-i(l+\frac{1}{2})}}{1 + \alpha^{-i/2}}.$$

Hence the algebraic degree of f is equal to n.

A similar proof of Proposition 12 also applies to the following theorem.

Proposition 16 *For the function f of (9), $U(f; e)$ is invertible for $e \subset n-1$ and $U(f; e, n-e-2)$ has full column rank for $e \not\subset n-1$.*

Theorem 17 *Let e be a positive integer less than $n/2$ and f be the function of (9). Then f admits no non-zero function g with degree at most e such that gf has degree at most $n-e-1$ if $\binom{n-1}{e} \equiv 1 (\bmod 2)$, and admits no non-zero function g with degree at most e such that gf has degree at most $n-e-2$ if $\binom{n-1}{e} \equiv 0 (\bmod 2)$.*

Proof. It is derived from Theorem 8 and Proposition 16. □

Corollary 18 *Let $n = 2^s$ and $f \in \mathbf{B}_n$ be the function of (9). Then f is $\mathcal{PAI}$.*

Proof. It is obtained from Theorem 17. □

Theorem 17 states that the modified Carlet-Feng functions achieve the bounds of Theorem 5 and thus the bounds of Theorem 5 are tight for the functions with algebraic degree equal to n, while Corollary 18 states that the modified Carlet-Feng function on 2^s variables is $\mathcal{PAI}$.

Consequently, as mentioned above, the bounds of Theorem 5 are tight and there exist $\mathcal{PAI}$ functions on 2^s and $2^s + 1$ variables.

5 The immunity of Boolean functions with maximum $\mathcal{AI}$ against probabilistic algebraic attacks

This section mainly focuses on the time complexities of probabilistic algebraic attacks on an LFSR-based nonlinear filter generator with the filter function achieving maximum $\mathcal{AI}$.

Let p be the probability for **S4a** or **S4b**. Then an overdetermined system of nonlinear equations with degree r is obtained where each equation holds with probability p. One can use the linearization algorithm to solve the system, where $R = \sum_{i=0}^{r} \binom{L}{i}$ equations are used and hold with probability p^R. Then the time complexity of probabilistic algebraic attacks is $p^{-R}R^w$, where $w \approx 2.807$ is the exponent of the Gaussian reduction.

In the affine case probabilistic algebraic attacks relate to the (fast) correlation attacks [3], so we always consider the nonlinear case here. Recall that the maximum $\mathcal{AI}$ of an n-variable function is $\lceil \frac{n}{2} \rceil$. Then, for the case $r \geq \lceil \frac{n}{2} \rceil$, deterministic algebraic attacks can be used. Therefore hereinafter we always assume that $2 \leq r \leq \lceil \frac{n}{2} \rceil - 1$.

Let g be a non-zero function with degree at most r. For a balanced function we know that the maximum probability for applying **S4a** or **S4b** is

$$p_{\max} = 1 - \frac{\min\{\mathrm{d}(gf, g), \mathrm{d}(gf, 0)\}}{2^{n-1}}.$$

According to [4, Proposition 5] (see also [21]) we have

$$d_r = \min\{\mathrm{d}(gf, g), \mathrm{d}(gf, 0)\} \geq \sum_{i=0}^{\mathcal{AI}(f)-r-1} \binom{n-r}{i}.$$

Since $\mathcal{AI}(f) \leq \lceil \frac{n}{2} \rceil$, we have $2\mathcal{AI}(f) - 2r - 1 < n - r$ and therefore for $r \leq \mathcal{AI}(f) - 1$,

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

Then for a function with maximum $\mathcal{AI}$ we have $d_r \geq 2^{n-2r-2}$ and therefore

$$p_{\max} = 1 - \frac{d_r}{2^{n-1}} \leq 1 - 2^{-2r-1}.$$

It is well known that the real function $1 - x - e^{-x}$ is decreasing when $x \geq 0$. Hence we have

$$p_{\max} \leq 1 - 2^{-2r-1} \leq e^{-2^{-2r-1}}$$

and the time complexity of probabilistic algebraic attacks

$$p^{-R}R^w \geq p_{\max}^{-R} \geq \left(e^{-2^{-2r-1}}\right)^{-R} = e^{R/2^{2r+1}} \geq 2^{1.44R/2^{2r+1}} \geq 2^{1.44\binom{L}{r}/2^{2r+1}}.$$

For $r \leq L/5$, we have

$$\frac{1}{2^{2r+1}} \binom{L}{r} \geq \frac{1}{2^{2r-1}} \binom{L}{r-1},$$

and it then holds that

$$p^{-R} \geq 2^{1.44\binom{L}{r}/2^{2r+1}} \geq 2^{1.44\binom{L}{2}/2^5}. \tag{10}$$

Corollary 9 of [22, Page 310] states that for $0 < \mu < 1/2$,

$$\sum_{i=0}^{\mu L} \binom{L}{i} \geq \frac{2^{H_2(\mu)L}}{\sqrt{8L\mu(1-\mu)}},$$

where $H_2(\mu) = -\mu \log_2 \mu - (1-\mu)\log_2(1-\mu)$. For $L/5 < r < L/2$, it follows that

$$R^w \geq \left(\sum_{i=0}^{L/5} \binom{L}{i}\right)^{2.807} \geq \left(\frac{2^{H_2(1/5)L}}{\sqrt{32L/25}}\right)^{2.807} \geq \frac{2^{2.02L}}{1.42L^{1.41}}. \tag{11}$$

From (10) and (11) we can calculate that for $L \geq 46$,

$$p^{-R} R^w \geq 2^L.$$

Consequently, we know that probabilistic algebraic attacks are worse than exhaustive key search in the context of their application to the nonlinear filter generator if the filter function achieves maximum $\mathcal{AI}$ and the size L of the LFSR is greater than or equal to 46. Since an $\mathcal{PAI}$ function has maximum $\mathcal{AI}$, the function also behaves good against probabilistic algebraic attacks.

6 Conclusion

In this paper, several open problems about the immunity of Boolean functions against algebraic attacks have been solved: maximum immunity to fast algebraic attacks, immunity of the Carlet-Feng functions against fast algebraic attacks, and resistance of Boolean functions with maximum algebraic immunity against probabilistic algebraic attacks. It seems that for a balanced function the number n of input variables equal to 2^s+1 is optimal in terms of immunity of fast algebraic attacks. The Carlet-Feng functions previously shown to have maximum algebraic immunity and good nonlinearity are proved to be optimal against fast algebraic attacks among the balanced functions. This is the first time that a function is shown to have such cryptographic property.

Acknowledgement

Meicheng Liu thanks Dingyi Pei for many enlightening conversations on the resistance of Boolean functions against algebraic attacks.

References

1. F. Armknecht. Improving fast algebraic attacks. In: B. Roy and W. Meier (eds.) FSE 2004. LNCS vol. 3017, pp. 65–82. Berlin, Heidelberg: Springer, 2004.
2. F. Armknecht, C. Carlet, P. Gaborit, et al. Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. In: S. Vaudenay (eds.) EUROCRYPT 2006. LNCS vol. 4004, pp. 147–164. Berlin, Heidelberg: Springer, 2006.
3. A. Braeken and B. Preneel. Probabilistic algebraic attacks. In 10th IMA International Conference on Cryptography and Coding, 2005, Lecture Notes in Computer Science, Vol. 3796, pp. 290–303. Springer-Verlag, 2005.
4. C. Carlet. On the higher order nonlinearities of algebraic immune functions. CRYPTO 2006, LNCS 4117, 584–601. Berlin, Heidelberg: Springer, 2007.
5. C. Carlet. Boolean functions for cryptography and error correcting codes. In: Y. Crama, P. Hammer, eds. Boolean Methods and Models in Mathematics, Computer Science, and Engineering, pp. 257–397. Cambridge: Cambridge University Press, 2010.
6. C. Carlet and K. Feng. An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity. In: ASIACRYPT 2008, LNCS vol. 5350, 425–440. Berlin, Heidelberg: Springer, 2008.
7. G. Cohen, I. Honkala, S. Litsyn and A. Lobstein. Covering codes. North-Holland, 1997.
8. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. Advances in Cryptology-EUROCRYPT 2003, LNCS 2656, 345–359. Berlin, Heidelberg: Springer, 2003.
9. N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. Advances in Cryptology-CRYPTO 2003, LNCS 2729, 176–194. Berlin, Heidelberg: Springer, 2003.

10. N. Courtois. Cryptanalysis of Sfinks. ICISC 2005, Lecture Notes in Computer Science, Volume 3935, 261–269. Berlin, Heidelberg: Springer, 2006.
11. D. K. Dalai, S. Maitra, and S. Sarkar. Basic theory in construction of Boolean functions with maximum possible annihilator immunity. Designs, Codes and Cryptography, vol. 40, no. 1, 41–58, 2006.
12. D. K. Dalai, K. C. Gupta, and S. Maitra. Results on algebraic immunity for cryptographically significant Boolean functions. INDOCRYPT 2004, LNCS 3348, 92–106. Berlin, Heidelberg: Springer, 2005.
13. Y. Du, F. Zhang and M. Liu. On the Resistance of Boolean Functions against Fast Algebraic Attacks. To appear in ICISC 2011.
14. K. Feng, Q. Liao, J. Yang. Maximal values of generalized algebraic immunity. Designs, Codes and Cryptography 50(2), 243–252 (2009)
15. S. Fischer and W. Meier. Algebraic Immunity of S-boxes and Augmented Functions. In: Biryukov, A. (ed.) FSE 2007. LNCS, vol. 4593, pp. 366–381. Springer, Heidelberg (2007)
16. G. Gong. Sequences, DFT and resistance against fast algebraic attacks. SETA 2008, LNCS, Vol.5203, pp. 197–218, 2008.
17. P. Hawkes and G. Rose. Rewriting Variables: the Complexity of Fast Algebraic Attacks on Stream Ciphers, in Crypto 2004, LNCS 3152, pp. 390–406. Springer, 2004.
18. N. Li, L. Qu, W. Qi, et al. On the construction of Boolean Functions with optimal algebraic immunity. IEEE Trans. Inform. Theory, vol. 54, no. 3, 1330–1334, 2008.
19. N. Li, W. Qi. Construction and analysis of Boolean functions of $2t+1$ variables with maximum algebraic immunity. ASIACRYPT 2006, LNCS 4284, pp. 84–98. Berlin, Heidelberg: Springer, 2006.
20. M. Liu, D. Lin, D. Pei. Fast algebraic attacks and decomposition of symmetric Boolean functions. IEEE Transaction on Information Theory, vol. 57, no. 7, pp. 4817–4821, 2011.
21. M. Liu, D. Lin, D. Pei. Results on the immunity of Boolean functions against probabilistic algebraic attacks. U. Parampalli, P. Hawkes (Eds.): ACISP 2011, LNCS 6812, pp. 34–46. Berlin, Heidelberg: Springer, 2011.
22. F.J. MacWilliams and N.J.A. Sloane. The theory of error correcting codes. New York: North-Holland, 1977.
23. S. Mesnager. Improving the Lower Bound on the Higher Order Nonlinearity of Boolean Functions With Prescribed Algebraic Immunity. IEEE Transactions on Information Theory, vol. 54, no. 8, 3656–3662, 2008.
24. W. Meier, E. Pasalic, and C. Carlet. Algebraic attacks and decomposition of Boolean functions. Advances in Cryptology-EUROCRYPT 2004, LNCS 3027, 474–491. Berlin, Heidelberg: Springer, 2004.
25. E. Pasalic. Almost fully optimized infinite classes of Boolean functions resistant to (fast) algebraic cryptanalysis. ICISC 2008, LNCS 5461, 399–414. Berlin, Heidelberg: Springer, 2009.
26. E. Pasalic. Probabilistic versus deterministic algebraic cryptanalysis – a performance comparison. IEEE Transactions on Information Theory, vol. 55, no. 11, pp. 5233–5240, Nov. 2009.
27. P. Rizomiliotis. On the resistance of Boolean functions against algebraic attacks using univariate polynomial representation. IEEE Transaction on Information Theory, vol. 56, NO. 8, pp. 4014–4024, 2010.
28. P. Rizomiliotis. On the security of the Feng-Liao-Yang Boolean functions with optimal algebraic immunity against fast algebraic attacks. Designs, Codes and Cryptography, vol. 57, no. 3, pp. 283-292, 2010.
29. X. Zhang, J. Pieprzyk, and Y. Zheng. On algebraic immunity and annihilators. ICISC 2006. In: Lecture Notes in Computer Science, vol. 4296, pp. 65–80. Berlin, Heidelberg: Springer, 2006.
30. Y. Zhang, M. Liu, D. Lin. On the immunity of rotation symmetric Boolean functions against fast algebraic attacks. [Online]. Available at http://eprint.iacr.org/2012/111.pdf