Analysis of Algebraic Immunity for Balanced Boolean Function
On 4-Variable

Hongli Liu
Institute of Mathematics and Statistics, Zhejiang University of Finance and Economics
18 XueYuan Street, Xiasha Higher Education Zone, Hangzhou 310018, China
Tel: 86-137-5827-3108 E-mail: ooolhl@163.com

Received: March 13, 2012 Accepted: March 28, 2012 Online Published: May 28, 2012
doi:10.5539/jmr.v4n3p13 URL: http://dx.doi.org/10.5539/jmr.v4n3p13

Abstract
Algebraic immunity and balancedness have been widely investigated as important cryptographic properties. Boolean functions with high algebraic immunity can resist algebraic attacks. So, Boolean functions which achieve maximum algebraic immunity and balancedness are our research objects. In this paper, we present a method to study the algebraic immunity of balanced Boolean functions on 4-variable from the rank of matrix, and indicate that all 4-variable balanced Boolean functions with algebraic degree not less than 2 have maximum algebraic immunity. Finally, we introduce two classes balanced functions on even variables which don’t achieve maximum algebraic immunity.

Keywords: Boolean function, Balanced function, Algebraic immunity, Annihilators

1. Introduction
Boolean functions are the core components of stream ciphers and block ciphers. Stream ciphers form an important class of symmetric-key encryption schemes, mainly designed for applications that require either low cost hardware implementation or an extremely high encryption rate. The most well studied models of stream ciphers are based on linear feedback shift registers (LFSR), namely the nonlinear combiners and the nonlinear filters, that consist of one or more LFSRs combined with a nonlinear Boolean function. Different criteria have been proposed for both the selection of the LFSRs and the nonlinear Boolean function, in order to resist attacks like the correlation attacks, time/memory/data trade-offs, and distinguishing attacks. Generally speaking, before 2003, cryptographic Boolean functions were usually required to be balanced, have high algebraic degree and high nonlinearity. Since 2003, the algebraic attacks proposed by Courtois and Meier have received a lot of attention in cryptanalysis, the main idea of which is to solve a system of low degree multivariate equations with unknown input keys. With this method, some cryptographic algorithms have been successfully attacked, such as Toyocrypt, LILI-128, SFINKS and so on. A new cryptographic property for designing Boolean functions to resist this kind of attacks, called algebraic immunity (AI), the algebraic immunity of an n-variable Boolean function is upper bounded by \( \left\lceil \frac{n}{2} \right\rceil \). Obviously, a Boolean function with maximum algebraic immunity is better to resist the algebraic attack.

At present, two classes Boolean functions possess maximum algebraic immunity. One class is Boolean function on even number of variables with high algebraic degree(approaching n), but not possessing balancedness. Another is symmetric Boolean functions. We wish that the Boolean functions have both maximum algebraic immunity and balancedness.

In this paper, we discuss the algebraic immunity of 4-variable balanced Boolean functions. The conclusion is that all 4-variable balanced Boolean functions which algebraic degree \( \geq 2 \) achieve maximum algebraic immunity. We wish that the method of matrix rank can provide an idea for the study of n-variable Boolean functions. At last, we indicate that two classes typical balanced Boolean functions can not achieve maximum algebraic immunity though definition.

2. Preliminaries
Let \( F_2 = \{0, 1\} \) be the binary field, \( F^n_2 \) be the \( n \)-dimensional vector space over \( F_2 \). A mapping from \( F^n_2 \) into \( F_2 \) is called a Boolean function in \( n \) variables, denoted by \( f(x_1, x_2, \cdots, x_n) \), or \( f(x) \) in brief. Let \( B_n \) be the set of all the \( n \)-variable Boolean functions. One of the representations of a Boolean function \( f(x_1, x_2, \cdots, x_n) \) is by its truth table, i.e., the binary sequence \( f = (v_1, v_2, \cdots, v_{2^n}) \), where the bits \( v_i \)'s are the values of \( f(x) \), when \( x \) runs through
the vectors $b_1 = (0, \cdots, 0), b_2 = (0, \cdots, 1, \cdots, b_{22} = (1, \cdots, 1, 1)$ of $F_2^n$ in lexicographical order. The algebraic degree of $f(x)$, denoted by $\deg(f)$, is defined to be the maximum degree appearing in the algebraic normal form (ANF).

The Hamming weight of a Boolean function $f \in B_n$ is the number of nonzero coordinates in its truth table, denoted by $\text{wt}(f)$. The support of $f(x)$ is defined as the set $\text{supp}(f) = \{ x \in F_2^n | f(x) = 1 \}$. 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 $2^{n-1}$.

A nonzero $n$-variable Boolean function $g \in B_n$ is called an annihilator of $f$ if $f \cdot g = 0$, we denote the set of all annihilators of $f$ by $\text{Ann}(f) = \{ g \in B_n | f \cdot g = 0 \}$. The algebraic immunity (AI) of $f$ is defined as $AI(f) = \min \{ \deg(g) \ | \ 0 \neq g \in \text{Ann}(f) \cup \{ f + 1 \} \}$. It is known that for any $n$-variable function, the maximum possible AI is $\left\lceil \frac{n}{2} \right\rceil$. If $AI(f) = \left\lceil \frac{n}{2} \right\rceil$, we say it has the maximum algebraic immunity.

From $f(1 + f) = f + f^2 = 0$, it holds, $AI(f) \leq \deg(f)$ for every $0 \neq g \in B_n$. Thus, the algebraic degree of Boolean functions which achieving maximum algebraic immunity must $\geq \left\lceil \frac{n}{2} \right\rceil$. Among the set of 4-variable balanced Boolean functions, the Boolean functions which algebraic degree $\geq 2$ will achieve maximum algebraic immunity. As shown in previous researches, whether the 4-variable balanced Boolean functions achieve maximum algebraic immunity is closely related to the rank of matrix $B$, where

$$B = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 
\end{pmatrix}$$

is composed of arbitrarily 8 binary vectors $a_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4})^T, i = 1, 2, \cdots, 8$. When the rank of $B$ is 8, the Boolean functions whose support set is $a_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4})^T, i = 1, 2, \cdots, 8$ is optimum algebraic immunity.

We denote $y_i, i = 1, 2, 3, 4$ the 2, 3, 4, 5 row of matrix $B$, respectively. According to the linear algebra, we can obtain the following result.

**Proposition 1** The matrix $B$ has same rank, when $\text{wt}(y_i) = i$ and $8 - i, i = 0, 1, 2, 3, j = 1, 2, 3, 4$.

Since the two matrices of $\text{wt}(y_j) = i$ and $8 - i$ are interchangeable with primary transformation, we have that the rank of the two matrices is the same.

### 3. Main Results

In this section, we provide the judgement of 4-variable balanced Boolean function with maximum AI. First, we give a conclusion about the linear correlation of vectors over the binary field.

**Lemma 1** Let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be a set of linearly independent vectors over the binary fields, the vector $\beta$ is different from $\alpha_1, \alpha_2, \cdots, \alpha_n$, and $\beta$ is linearly independent with arbitrarily $n - 1$ vectors in $\alpha_1, \alpha_2, \cdots, \alpha_n$. Then, the necessary and sufficient condition for that $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta$ are linearly dependent is $\alpha_1 + \alpha_2 + \cdots + \alpha_n + \beta = 0$.

**Proof.** According to the definition of linear dependent $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta$ are linearly dependent if and only if $k_1 \alpha_1 + k_2 \alpha_2 + \cdots + k_n \alpha_n + b \beta = 0$, where $k_1, k_2, \cdots, k_n, b \in \{ 0, 1 \}$ not all 0.

If $b = 0$, then $k_1 = k_2 = \cdots = k_n = 0$, which contradicts with above facts. So, $b \neq 0$.

Suppose that $b \neq 0$, $k_1, k_2, \cdots, k_n$ are 0 at least one. Without loss of generality, let $k_1 = 0, k_2 = \cdots = k_n = 1$, then the corresponding vectors $\alpha_2, \cdots, \alpha_n, \beta$ are linear dependent, contradicts with the known condition. So, $k_1 = k_2 = \cdots = k_n = 1$. That is, $\alpha_1 + \alpha_2 + \cdots + \alpha_n + \beta = 0$.

From lemma 1 and primary transformation of matrix, we have that

**Lemma 2** The rank of matrix $B \leq 7$, if $\text{wt}(y_i) = 0$ or 8 for any $i = 1, 2, 3, 4$. 


Proof. We only prove \( wt(\gamma_1) = 0 \), and the proof of other cases is similar.

If \( wt(\gamma_1) = 0 \), then

\[
B = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
a_{12} & a_{22} & a_{32} & a_{42} & a_{52} & a_{62} & a_{72} & a_{82} \\
a_{13} & a_{23} & a_{33} & a_{43} & a_{53} & a_{63} & a_{73} & a_{83} \\
a_{14} & a_{24} & a_{34} & a_{44} & a_{54} & a_{64} & a_{74} & a_{84} \\
a_{11}a_{12} & a_{12}a_{22} & a_{12}a_{32} & a_{12}a_{42} & a_{12}a_{52} & a_{12}a_{62} & a_{12}a_{72} & a_{12}a_{82} \\
a_{11}a_{13} & a_{12}a_{23} & a_{12}a_{33} & a_{12}a_{43} & a_{12}a_{53} & a_{12}a_{63} & a_{12}a_{73} & a_{12}a_{83} \\
a_{11}a_{14} & a_{12}a_{24} & a_{12}a_{34} & a_{12}a_{44} & a_{12}a_{54} & a_{12}a_{64} & a_{12}a_{74} & a_{12}a_{84} \\
a_{12}a_{13} & a_{12}a_{23} & a_{12}a_{33} & a_{12}a_{43} & a_{12}a_{53} & a_{12}a_{63} & a_{12}a_{73} & a_{12}a_{83} \\
a_{12}a_{14} & a_{12}a_{24} & a_{12}a_{34} & a_{12}a_{44} & a_{12}a_{54} & a_{12}a_{64} & a_{12}a_{74} & a_{12}a_{84} \\
a_{13}a_{14} & a_{13}a_{24} & a_{13}a_{34} & a_{13}a_{44} & a_{13}a_{54} & a_{13}a_{64} & a_{13}a_{74} & a_{13}a_{84}
\end{pmatrix}
\]

Obviously, the rank of matrix \( B \) \( \leq 7 \) since \( B \) has 4 zero rows.

From proposition 1, matrix of \( wt(\gamma_i) = 8 \) and matrix of \( wt(\gamma_i) = 0 \) have the same rank.

Lemma 3 The rank of matrix \( B \) is 8, when \( wt(\gamma_i) = 1, 2, 3 \) or \( 5, 6, 7 \) for any \( i = 1, 2, 3, 4 \).

Proof. From the proposition 1, without loss of generality, we will discuss the case of \( wt(\gamma_1) = 1, 2, 3 \), respectively.

Case 1. The rank of matrix \( B \) is 8, when \( wt(\gamma_1) = 1 \).

When \( \gamma_1 = 1 \),

\[
B = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
a_{12} & a_{22} & a_{32} & a_{42} & a_{52} & a_{62} & a_{72} & a_{82} \\
a_{13} & a_{23} & a_{33} & a_{43} & a_{53} & a_{63} & a_{73} & a_{83} \\
a_{14} & a_{24} & a_{34} & a_{44} & a_{54} & a_{65} & a_{74} & a_{84} \\
a_{11}a_{12} & a_{12}a_{22} & a_{12}a_{32} & a_{12}a_{42} & a_{12}a_{52} & a_{12}a_{62} & a_{12}a_{72} & a_{12}a_{82} \\
a_{11}a_{13} & a_{12}a_{23} & a_{12}a_{33} & a_{12}a_{43} & a_{12}a_{53} & a_{12}a_{63} & a_{12}a_{73} & a_{12}a_{83} \\
a_{11}a_{14} & a_{12}a_{24} & a_{12}a_{34} & a_{12}a_{44} & a_{12}a_{54} & a_{12}a_{64} & a_{12}a_{74} & a_{12}a_{84} \\
a_{12}a_{13} & a_{12}a_{23} & a_{12}a_{33} & a_{12}a_{43} & a_{12}a_{53} & a_{12}a_{63} & a_{12}a_{73} & a_{12}a_{83} \\
a_{12}a_{14} & a_{12}a_{24} & a_{12}a_{34} & a_{12}a_{44} & a_{12}a_{54} & a_{12}a_{64} & a_{12}a_{74} & a_{12}a_{84} \\
a_{13}a_{14} & a_{13}a_{24} & a_{13}a_{34} & a_{13}a_{44} & a_{13}a_{54} & a_{13}a_{64} & a_{13}a_{74} & a_{13}a_{84}
\end{pmatrix}
\]

From lemma 1, it is clear that the sum of column vectors for matrix is not equal to 0. Thus, the rank of matrix \( B \) is 8.

Case 2. The rank of matrix \( B \) is 8, when \( wt(\gamma_1) = 2 \).
When $\gamma_1 = 2$,

\[
B = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
a_{12} & a_{22} & a_{32} & a_{42} & a_{52} & a_{62} & a_{72} & a_{82} \\
a_{13} & a_{23} & a_{33} & a_{43} & a_{53} & a_{63} & a_{73} & a_{83} \\
a_{14} & a_{24} & a_{34} & a_{44} & a_{54} & a_{64} & a_{74} & a_{84} \\
a_{1112} & a_{1212} & a_{1312} & a_{1412} & a_{1115} & a_{1215} & a_{1315} & a_{1415} \\
a_{1113} & a_{1213} & a_{1313} & a_{1413} & a_{1116} & a_{1216} & a_{1316} & a_{1416} \\
a_{1114} & a_{1214} & a_{1314} & a_{1414} & a_{1117} & a_{1217} & a_{1317} & a_{1417} \\
a_{1113} & a_{1213} & a_{1313} & a_{1413} & a_{1118} & a_{1218} & a_{1318} & a_{1418} \\
a_{1114} & a_{1214} & a_{1314} & a_{1414} & a_{1119} & a_{1219} & a_{1319} & a_{1419} \\
\beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_8
\end{pmatrix}
\]

It is clear that the set of vectors $\beta_1, \beta_2, \cdots, \beta_7$ is linearly independent. If $\beta_1, \beta_2, \cdots, \beta_7, \beta_8$ are linearly dependent, then $\beta_8$ can be linearly represented by $\beta_1, \beta_2, \cdots, \beta_7$. The features of vectors determine the representation can be shown as following: $\beta_8 = l_1\beta_1 + l_2\beta_2 + \cdots + l_7\beta_7$, where $l_1, l_2, \cdots, l_7 \in \{0, 1\}$. Therefore, $\beta_1 = \beta_8$. It is inconsistent with the known facts. Thus, $\beta_1, \beta_2, \cdots, \beta_7, \beta_8$ are linearly independent.

Case 3. The rank of matrix $B$ is 8, when $wt(\gamma_1) = 3$.

It is easy to proof.

**Lemma 4** When $wt(\gamma_1) = 4$, the necessary and sufficient condition of $r(B) < 8$ is that the sum of subset of $\alpha_i (i = 1, 2, 3, 4)$ which satisfy the $j$th $(j = 2, 3, 4)$ component is 0 and 1 is zero, respectively.

**Proof.** In fact, for any $j = 2, 3, 4$, when $wt(\gamma_j) \neq 4$, the rank of matrix $B$ is 8. Thus, we consider only the case of $wt(\gamma_i) = 4, i = 2, 3, 4$, when $wt(\gamma_1) = 4$. Here,

\[
B = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
a_{12} & a_{22} & a_{32} & a_{42} & a_{52} & a_{62} & a_{72} & a_{82} \\
a_{13} & a_{23} & a_{33} & a_{43} & a_{53} & a_{63} & a_{73} & a_{83} \\
a_{14} & a_{24} & a_{34} & a_{44} & a_{54} & a_{64} & a_{74} & a_{84} \\
0 & 0 & 0 & 0 & a_{53} & a_{63} & a_{73} & a_{83} \\
0 & 0 & 0 & 0 & a_{54} & a_{64} & a_{74} & a_{84} \\
a_{12a13} & a_{22a23} & a_{32a33} & a_{42a43} & a_{52a53} & a_{62a63} & a_{72a73} & a_{82a83} \\
a_{12a14} & a_{22a24} & a_{32a34} & a_{42a44} & a_{52a54} & a_{62a64} & a_{72a74} & a_{82a84} \\
a_{13a14} & a_{23a24} & a_{33a34} & a_{43a44} & a_{53a54} & a_{63a64} & a_{73a74} & a_{83a84} \\
\beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_8
\end{pmatrix}
\]

The vectors $\beta_1, \beta_2, \cdots, \beta_7$ are linearly independent, $\beta_8$ is linearly independent with arbitrarily 6 vectors of $\alpha_1, \alpha_2, \cdots, \alpha_n$. So, $\beta_1, \beta_2, \cdots, \beta_8$ are linearly dependent if and only if $\beta_1 + \beta_2 + \cdots + \beta_8 = 0$. Through the discussion of values of $a_{12}, a_{22}, a_{32}, a_{42}$, we have the following conclusion:
\[ \beta_1 + \beta_2 + \cdots + \beta_k = 0 \] if and only if the sum of vectors is 0, and the vectors are subset of \( a_i (i = 1, 2, 3, 4) \) which satisfy the \( j \)th \((j = 2, 3, 4)\) component is 0 and 1. Actually, if the sum of the elements of \( \gamma_1 \) where the elements of \( \gamma_2 \) is 1 is not 0. Then, the sum of the elements of the 10th row of matrix \( B \) is not 0, because the 10th row elements of matrix \( B \) are generated respectively by the product of \( \gamma_2 \) and \( \gamma_4 \) elements. The results are obtained. By observing the functions corresponding to the matrixes such that \( \text{wt}(\gamma_1) = 0 \) and the matrixes which described in lemma 4, we can find that the functions are 4-variable linear functions. Overall, all of 4-variable balanced Boolean functions are optimum algebraic immunity except linear functions. This is consistent with the conclusion that the algebraic degree of \( n \)-variable optimum algebraic immunity Boolean functions \( \geq \left\lceil \frac{n}{2} \right\rceil \). Then, we have the following theorem:

**Theorem 1** The 4-variable balanced Boolean functions which algebraic degree \( \geq 2 \) are optimum algebraic immunity.

**Proof.** We known that any binary string can form a Boolean function. The Boolean functions of support \( a_i, i = 1, 2, \cdots , 8 \) which generate matrix \( B \) are 4-variable balanced functions. So, there are relationship between 4-variable balanced Boolean function and matrix \( B \). When the rank of matrix \( B \) is 8, the functions which corresponding to matrix \( B \) achieve maximum algebraic immunity. Through calculation, the Boolean function which corresponding to the matrix \( B \) satisfy \( \text{rank}(B) < 8 \) are linear functions. Consequently, the degree of functions which corresponding to \( \text{rank}(B) = 8 \) are 2 minimum, and they are optimum algebraic immunity.

Inspired by theorem 1, the algebraic immunity of 2m-variable balanced Boolean functions which algebraic degree \( \geq m \) was discussed, and obtained the following conclusion:

**Theorem 2** For \( f = x_1x_2 \cdots x_m + g(x_{m+1}, x_{m+2}, \cdots , x_{2m}) \) \((m \geq 4)\), if \( g \) is balanced function with deg \((g) \leq (m - 2)\) or \( g = x_{m+1}x_{m+2} \cdots x_{2m-1} + x_{2m} \), then the Boolean function \( f \) is not optimum algebraic immunity.

**Proof.** If \( \text{deg}(g) \leq (m - 2) \), let \( h = (x_1 + 1)(g + 1) \), we have \( fg = 0 \). Since \( \text{deg}(h) = m - 1 < m \), the Boolean function \( f \) is not optimum algebraic immunity.

If \( g = x_{m+1}x_{m+2} \cdots x_{2m-1} + x_{2m} \), let \( h = (x_1 + 1)(x_{m+1} + 1)(x_{2m} + 1) \), obviously, \( fg = 0 \). Since \( \text{deg}(h) = 3 < m \), the Boolean function \( f \) is not optimum algebraic immunity.

**Theorem 3** For \( f = x_1x_2 \cdots x_t + g(x_{t+1}, x_{t+2}, \cdots x_{2m})(m \geq 4, t \geq m + 1) \), if \( g \) is balanced function, then the Boolean function \( f \) is not optimum algebraic immunity.

**Proof.** Let \( h = (x_1 + 1)(g + 1) \), we have \( fh = 0 \). So, \( f \) is not optimum algebraic immunity since \( \text{deg}(h) < m \).

4. Conclusion

The rank of matrix can be used to judgement the maximum algebraic immunity of balanced Boolean function. Based on the case of 4-variable, we can popularize the method to 6-variable or more variables. Since the calculation of matrix rank will be difficult with the increase of rows and columns, the algebraic immunity of Boolean functions corresponding to the matrix will be difficult to decide.

References


