J. Korea Soc. Math. Educ. Ser. B: Pure Appl. Math.
6(1999), no. 2, 79–86

AN ALGORITHM FOR FINDING THE CORRELATION IMMUNE ORDER OF A BOOLEAN FUNCTION

MIN SURP RHEE, HYUN SOOK RHEE AND HYUNYONG SHIN

ABSTRACT. A Boolean function generates a binary sequence which is frequently used in a stream cipher. There are number of critical concepts which a Boolean function, as a key stream generator in a stream cipher, satisfies. These are nonlinearity, correlation immunity, balancedness, SAC(strictly avalanche criterion), PC(propagation criterion) and so on. In this paper we construct an algorithm for finding the correlation immune order of a Boolean function, and check how long to find the correlation immune order of a given Boolean function in our algorithm.

1. Introduction

A Boolean function on $GF(2)^n$ has been often used in cryptography and coding theory. In particular a Boolean function generates a binary sequence which is frequently used in stream cipher. There are number of critical properties which a Boolean function, as a key stream generator in a stream cipher, satisfies. These are nonlinearity, correlation immunity, balancedness, SAC(strictly avalanche criterion), PC(propagation criterion) and so on. In this paper we investigate correlation immunity.

A concept of correlation immunity is introduced by Blaser and Heinzmann [1]. In 1982 they observed that the cryptosystem could be breakable when the combining function leaks information about its component functions. Finding methods for an easy construction of correlation immune functions has been an active research area since the introduction of the notion by Siegenthaler [7], in 1984. Camion et al. [2] and Seberry et al. [6] construct a Boolean function which satisfies a given correlation immune order and a given algebraic order.

Received by the editors June 22, 1999, and in inrevised form November 13, 1999.
1991 *Mathematics Subject Classification.* 94A60.
Key words and phrases. correlation immune, Walsh-Hadarmard transform.
This paper was supported by the KOSEF, Project No. 97-01-00-13-01-5.

In this paper we construct an algorithm for finding the correlation immune order of a Boolean function and check how long to find the correlation immune order of a given Boolean function in our algorithm.

2. Basic definitions and preliminary results

Let $f: GF(2)^n \to GF(2)$ be a Boolean function, where $GF(2)^n$ is an n-dimensional vector space over the Galois field $GF(2)$. In other cases it may be more convenient to deal with $\hat{f}(\vec{x}) = (-1)^{f(\vec{x})}$ that takes the values in $-1, 1$.

A Boolean function $f : GF(2)^n \to GF(2)$ can be expressed as

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

which is called *an algebraic normal form* (ANF). A Boolean function f of the form

$$f(x_1, x_2, \cdots, x_n) = \lambda(x_1, x_2, \cdots, x_n) = \omega_0 + \omega_1 x_1 + \cdots + \omega_n x_n$$

(where $\omega_i \in GF(2)$) is called *an affine function.* In particular, an affine function with $\omega_0 = 0$ is called *a linear function.* Every linear function on $GF(2)^n$ is simply written as $\vec{\omega} \cdot \vec{x}$ for some $\vec{\omega} \in GF(2)^n$, where $\vec{\omega} \cdot \vec{x} = \omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n$. Here we write $L_{\vec{\omega}}(\vec{x})$ for a linear function $\vec{\omega} \cdot \vec{x}$.

It follows from above basic definitions that there are 2^{2^n} Boolean functions and 2^{n+1} affine functions on $GF(2)^n$.

Definition 1. Let $f : GF(2)^n \to GF(2)$ and $g : GF(2)^n \to GF(2)$ be Boolean functions. Then the *Hamming distance* of f and g is defined by

$$d(f, g) = \#\{\vec{x} \in GF(2)^n \mid f(\vec{x}) \neq g(\vec{x})\}.$$

For a Boolean function $f : GF(2)^n \to GF(2)$, the distance to affine function

$$N_f = \min_{\lambda \in A(n)} d(f, \lambda)$$

is called the *nonlinearity* of f, where $A(n)$ is the set of all affine functions.

Definition 2. A Boolean function $f : GF(2)^n \to GF(2)$ is called the *k-th order correlation immune function* if

$$P(f(x_1, x_2, \cdots, x_n)=1 | x_{i_1} = a_1, x_{i_2} = a_2, \cdots, x_{i_k} = a_k) = P(f(x_1, x_2, \cdots, x_n)=1)$$

for any $i_1, i_2, i_3, \cdots, i_k$ with $1 \leq i_1 < i_2 < i_3 < \cdots < i_k \leq n$ and $a_1, a_2, a_3, \cdots, a_k \in GF(2)$. In this case k is called the *correlation immune order* (c. i. order) of f. A Boolean function $f : GF(2)^n \to GF(2)$ is said to be *balanced* if

$$\#\{\vec{x} \in GF(2)^n | f(\vec{x}) = 0\} = \#\{\vec{x} \in GF(2)^n | f(\vec{x}) = 1\}.$$

The following theorem explains the relations between algebraic degree and the correlation immune order.

Theorem 3 [7]. *If $f(x_1, x_2, \cdots, x_n)$ is the k-th order correlation immune function, where $1 \leq k \leq n$, then no product of $n - k + 1$ or more variables can be present in the algebraic normal form of f. Moreover if f is balanced, then no product of $n - k$ variables can be present in the algebraic normal form of f unless $k = n - 1$.*

Since correlation immunity is an important concept in a stream cipher system, it is worthwhile to get the correlation immune order of a given Boolean function. We will use the following concept to get the correlation immune order of a given Boolean function.

Definition 4. For a Boolean function f on $GF(2)^n$, the *Walsh-Hadamard transform* F of f is the real-valued function over $GF(2)^n$ defined as

$$F(\vec{\omega}) = \sum_{\vec{x} \in GF(2)^n} (-1)^{\vec{x} \cdot \vec{\omega}} f(\vec{x}).$$

Xiao and Massey [9] has given an equivalent condition for correlation immunity. That is,

Theorem 5 [9]. *Let $f : GF(2)^n \to GF(2)$ be a Boolean function and $\hat{F}$ the Walsh transform of $\hat{f}$. Then f is a k-th order correlation immune function if and only if $\hat{F}(\vec{\omega}) = 0$ for all $\vec{\omega} \in GF(2)^n$ with $1 \leq W(\vec{\omega}) \leq k$, where $W(\vec{\omega})$ is the Hamming weight i.e., the number of components in $\vec{\omega}$ which is 1.*

Let all elements of $GF(2)^n$ give the natural order like

$$\vec{\alpha}_0 = (0 \cdots 00), \quad \vec{\alpha}_1 = (0 \cdots 01), \quad \vec{\alpha}_2 = (0 \cdots 10),$$
$$\vec{\alpha}_3 = (0 \cdots 11), \quad \cdots, \quad \vec{\alpha}_{2^n-2} = (1 \cdots 10), \quad \vec{\alpha}_{2^n-1} = (1 \cdots 11).$$

Let $[f]$ and $[F]$ be the column matrices of the function values of a Boolean function f and the Walsh-Hadamard transform F, respectively. That is,

$$[f] = \begin{pmatrix} f(\vec{\alpha}_0) \\ f(\vec{\alpha}_1) \\ f(\vec{\alpha}_2) \\ \vdots \\ f(\vec{\alpha}_{2^n-1}) \end{pmatrix}, \qquad [F] = \begin{pmatrix} F(\vec{\alpha}_0) \\ F(\vec{\alpha}_1) \\ F(\vec{\alpha}_2) \\ \vdots \\ F(\vec{\alpha}_{2^n-1}) \end{pmatrix}.$$

A matrix H_n which can be recursively defined as

$$H_n = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \bigotimes H_{n-1}, \quad H_0 = [1]$$

is called the *Walsh-Hadamard matrix of order n*, where $\bigotimes$ denotes the Kronecker product between matrices.

Theorem 6 [5]. *In above notations* $[F] = H_n[f]$.

The Walsh-Hadamard matrix of order n can be written as

$$\begin{aligned} H_n &= [(-1)^{\vec{\alpha}_j \cdot \vec{\alpha}_i}] \\ &= \begin{bmatrix} (-1)^{(00\cdots0)\cdot(00\cdots0)} & \cdots & (-1)^{(11\cdots0)\cdot(00\cdots0)} & (-1)^{(11\cdots1)\cdot(00\cdots0)} \\ (-1)^{(00\cdots0)\cdot(00\cdots1)} & \cdots & (-1)^{(11\cdots0)\cdot(00\cdots1)} & (-1)^{(11\cdots1)\cdot(00\cdots1)} \\ \vdots & & \vdots & \vdots \\ (-1)^{(00\cdots0)\cdot(11\cdots1)} & \cdots & (-1)^{(11\cdots0)\cdot(11\cdots1)} & (-1)^{(11\cdots1)\cdot(11\cdots1)} \end{bmatrix}. \end{aligned}$$

Let $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \cdots, \vec{e}_n\}$ be the standard basis for $GF(2)^n$. Now, for each vector $\vec{e}_i$ we define the row vector $\vec{B}_i$ by

$$((-1)^{(00\cdots0)\cdot\vec{e}_i}, \cdots, (-1)^{(11\cdots0)\cdot\vec{e}_i}, (-1)^{(11\cdots1)\cdot\vec{e}_i})$$

Then each row vector $\vec{B}_i$ is a row of H_n and for any $\vec{\alpha} = (\vec{\omega}_1, \vec{\omega}_2, \vec{\omega}_3, \cdots, \vec{\omega}_n) \in GF(2)^n$, we can express the row $\vec{l}$ in H_n corresponding to $\vec{\alpha}$ as

$$\vec{l} = \bigodot_{\substack{\omega_j = 1 \\ 1 \le j \le n}} \vec{B}_j$$

where $\vec{B}_i \odot \vec{B}_j = (b_{i_1} \cdot b_{j_1}, b_{i_2} \cdot b_{j_2}, b_{i_3} \cdot b_{j_3}, \cdots, b_{i_{2^n}} \cdot b_{j_{2^n}})$. Therefore we have the following theorem:

Theorem 7. *Every row of Walsh-Hadamard matrix H_n can be represented by combining a finite number of row vectors $\vec{B}_i$'s. That is,*

$$H_n = \begin{bmatrix} \vec{l}_0 \\ \vec{l}_1 \\ \vdots \\ \vec{l}_{2^n-1} \end{bmatrix} = \begin{bmatrix} (1,1,\cdots,1) \\ \vec{B}_1 \\ \vdots \\ \vec{B}_1 \odot \vec{B}_2 \odot \cdots \odot \vec{B}_n \end{bmatrix}.$$

Now, we prove Theorem 8 by a series of above three theorems.

Theorem 8. *A Boolean function $f : GF(2)^n \to GF(2)$ is k-th order correlation immune if and only if $\hat{F}(\vec{\alpha}) = \vec{l}[\hat{f}] = 0$ for any $\vec{\alpha}$ with $1 \leq W(\vec{\alpha}) \leq k$, where*

$$[\hat{f}] = (\hat{f}(\vec{\alpha}_0), \hat{f}(\vec{\alpha}_1), \hat{f}(\vec{\alpha}_2), \cdots, \hat{f}(\vec{\alpha}_{2^n-1}))$$

and the low $\vec{l}$ in H_n corresponding to $\vec{\alpha}$.

Proof. It follows from Theorem 5 that a Boolean function $f : GF(2)^n \to GF(2)$ is k-th order correlation immune if and only if $\hat{F}(\vec{\alpha}) = 0$ for any $\vec{\alpha}$ with $1 \leq W(\vec{\alpha}) \leq k$. Hence we get Theorem 8 from Theorems 6 and 7. □

3. An algorithm for finding a correlation immune order

In this chapter, we explain the algorithm for finding the correlation immune order of a given algebraic normal form of f.

In Theorem 7 we showed that every row $\vec{l}$ of Walsh-Hadamard marix H_n is generated by the vectors $\vec{B}_1, \vec{B}_2, \vec{B}_3, \cdots, \vec{B}_n$.

Our algorithm is the following:

Algorithm

Step 1 Input n which is the number of variables and the algebraic normal form of a function $f(\vec{x})$.

Step 2 Calculate and $[f]$ and $[\hat{f}]$.

Step 3 From $i = 0$ to n, generate standard n row vectors $\vec{B}_i$

Step 4 From $i = 0$ to n, calculate $\vec{B}_i \cdot [\hat{f}]$.

Step 5 If $\sum_{i=1}^{n} \vec{B}_i \cdot [\hat{f}] = 0$ then $k = 1$ and go to step 6.

Otherwise, $k = 0$ and print "the correlation immune order f of is 0", and stop.

Step 6 For $0 \leq i < n$, $\vec{H}_i = \vec{B}_i$
and $Number = n$.

Step 7 If $k < n$
FOR $i \geq n$,
FOR $0 \leq j < n-1$
FOR $j < k < n$
Put $\vec{l}_i = \vec{B}_i \cdot \vec{B}_k$
and $Number + 1 \rightarrow Number$
else the "correlation immune order of f is k".

Step 8 If $\vec{l}_i \cdot [\hat{f}] = 0$ from $i = 0$ to $Number$ then $k = k + 1$ and go to step 6,
else print "the correlation immune order of f is k" and Stop.

Example 1. The dimension n is 7 and $f(\vec{x}) = x_1 + x_2 + x_3 + x_4x_5x_6x_7$ then the result is as follows (Figure 1).

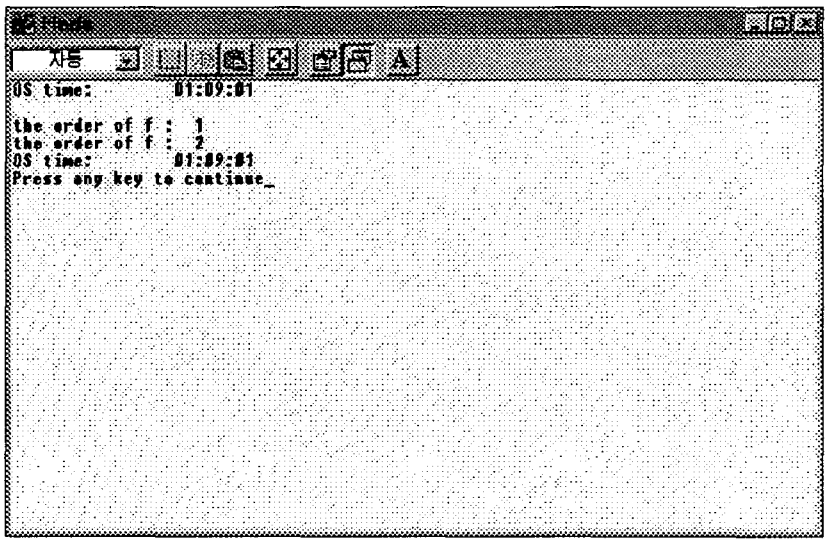

Figure 1.

Jun [3] used Walsh-Hadamard matrix to find the correlation immune order (c. i. order) of a given algebraic normal form of f. They use the whole Walsh-Hadamard

matrix H_n in all cases. They calculate $[F] = H_n[f]$ and find the c. i. order k such that $F(\vec{\omega}) = 0$ for all $0 < W(\vec{\omega}) \leq k$ and $F(\vec{\omega}) \neq 0$ for some $W(\vec{\omega}) = k + 1$. Our algorithm to find the c. i. order of a Boolean function is based on Theorem 8. Now, we compare our Algorithm (ourAL) with JunAL [3] using our PC which is Pentium III 450 MHZ with 256 MB RAM. We have the following tables (Tables 2 and 3):

Table 1.

(unit: sec)

n \ c. i. order / algorithm		1	2	3	4	5	6	7	8	9	10	11
11	JunAL	5	5	5	4	5	5	5	5	5	5	
	ourAL	1	2	3	4	5	7	7	7	7	7	
12	JunAL	20	20	19	20	20	20	19	19	19	19	20
	ourAL	3	8	12	17	22	28	33	39	44	47	50

Table 2.

(unit: sec)

n \ c. i. order	1	2	3	4	5	···	12	13	14
13	19	33	49	70	89	···	238		
14	73	131	205	293	397	···	1293	1292	
15	345	636	991	1402	1888	···	···	···	6500

Remark. We can analyze the following facts from two algorithms:

(1) If $n \leq 10$, then both algorithms are very fast. So, two algorithms are similar.

(2) If $11 \leq n \leq 12$, then the speed of our algorithm depends on k, while the speed of JunAL is independent of k. From Table 1, we know that our algorithm is better than JunAL when k is small. In fact it follows from Theorem 3 that the greater the c. i. order is, the less the algebraic degree is. The concept of nonlinearity is related to the concept of an algebraic degree. Since a binary sequence which can be used in a stream cipher system needs some scale of nonlinearity, our algorithm is not bad.

(3) If $n \geq 13$, then JunAL is not available to find the c. i. order of a given Boolean function. But we can find the c. i. order of a given Boolean function with our algorithm (Table 2). Also, we can find the c. i. order of a given

Boolean function with $n \leq 20$. In fact, we have spent 51095 seconds to find the c. i. order of a given Boolean function of c. i. order 2 with $n = 18$.

REFERENCES

1. W. Blaser and P. Heinzmann, *New cryptographic device with high security using public key distribution*, IEEE Student Papers Contest 1979–1980 (pp.145–153), IEEE, 1982.
2. P. Camion, C. Carlet, P. Charpin and N. Sendrier, *On correlation-immune functions*, Advances in Cryptology—CRYPTO '91 (pp. 87–100), Springer-Verlag, 1992.
3. Y. Jun, *Algorithms for generation nonlinear combining functions*, Master's Thesis, Korea National University of Education, 1999.
4. W. Meier and O. Staffelbach, *Nonlinearity criteria for cryptographic functions*, Advances in Cryptology—EUROCRYPT '89 (pp. 549–562), Springer-Verlag, 1990.
5. B. Preneel, W. van Leewijk, L. van Linden, R. Govaerts and J. Vandewalle, *Propagation characteristics of Boolean functions*, Advances in Cryptology—EUROCRYPT '90 (pp. 161–173), Springer-Verlag, 1991.
6. J. Seberry, X. Zhang and Y. Zheng, *On constructions and nonlinearity of correlation immune functions for cryptographic applications*, Advances in Cryptology—EUROCRYPT '93 (pp. 181–200), Springer-Verlag, 1994.
7. T. Siegenthaler, *Correlation-immunity of nonlinear combining functions for cryptographic applications*, IEEE Trans. Information Theory **30** (1984), 776–780.
8. A. F. Webster and S. E. Tavares, *On the design of S-Boxes*, Advances in Cryptology—CRYPTO '85 (pp. 523–534), Springer-Verlag, 1986.
9. G. Xiao and J. Massey, *A spectral characterization of correlation immune combining functions*, IEEE Trans. Inform. Theory **34** (1988), no. 3, 569–571.

(M. S. RHEE) DEPARTMENT OF MATHEMATICS, DANKOOK UNIVERSITY, CHEONAN, CHUNGNAM 330-714, KOREA.
E-mail address: msrhee@anseo.dankook.ac.kr

(H. S. RHEE) DEPARTMENT OF MATHEMATICS, DANKOOK UNIVERSITY, CHEONAN, CHUNGNAM 330-714, KOREA.

(H. SHIN) DEPARTMENT OF MATHEMATICS EDUCATION, KOREA NATIONAL UNIVERSITY OF EDUCATION, CHUNGBUK 363-791, KOREA.
E-mail address: shin@cc.knue.ac.kr