균등함수들의 GAC에 관해서

손중제 김희진 김종덕 임종인
고려대학교 수학과

About Global Avalanche Characteristics Balanced Boolean functions

Son, Jung-Je Kim, Hee-Jean Kim, Jong-Deok Lim, Jong-In
Dept. of Mathematics, Korea University

November 7, 1997

Abstract

[6]에서 Zhang과 Zheng은 부울함수의 암호학적인 전역상관계수의 특성을 계산하기 위해서 GAC(Global Avalanche Characteristic)이라는 새로운 개념을 제시하였다. 그들은 GAC의 값들에 대한 측적을 위해서 2개의 단위를 제시했고 2개의 단위의 상한과 하한에 대 해서 계산했다. 그러나, 그들은 균등함수의 GAC의 하한은 향후의 연구과제로 남겨놓았다. 본 논문에서는 균등함수의 GAC의 하한에 대해서 계산했고, 연접의 방법에 의한 좋은 GAC 의 특성을 가지는 함수의 생성방법을 제시하였다.

1 Introduction

The Propagation Characteristic(PC) was introduced in [1] to study the dynamic behavior of a cryptographic Boolean function when the input to the function is modified. A function f on Z_2^n is said to satisfy the PC with respect to a vector $\alpha \in Z_2^n$ if complementing the vector α results in changing the output bit with the probability exactly one half.

The PC can be stated in terms of the autocorrelation function: for a function f on Z_2^n, its autocorrelation function is the real-valued function defined by $C_f(\alpha) = \sum_x (-1)^{f(x)\oplus f(x\oplus\alpha)}$, and f satisfies the PC with respect to a vector α if $C_f(\alpha) = 0$.

The PC is a very important concept in designing encryption algorithm and one-way hash function. However, the PC is a measure with a local flavor,

for it guarantees good propagation characteristics with respect to specific vectors. Functions satisfying the PC with respect to all vectors coincide with bent functions. Although bent functions have nice properties, they are not balanced and so can hardly be directly employed in practice[2].

In [6], Zhang and Zheng introduced the Global Avalanche Characteristic(GAC) to overcome those shortcomings of the PC. They also proposed the following two indicators related to GAC:

Definition 1 *For a Boolean function f on Z_2^n, the sum-of-squares indicator is defined by*

$$\sigma_f = \sum_{\alpha} C_f^2(\alpha),$$

and the absolute indicator is defined by

$$\Delta_f = \max_{\alpha \neq 0} |C_f(\alpha)|.$$

The smaller σ_f and Δ_f, the better the GAC of a function. Zhang and Zheng also derived bounds on two indicators:

$$2^{2n} \leq \sigma_f \leq 2^{3n}, \qquad 0 \leq \Delta_f \leq 2^{2n},$$

and also showed that two lower bounds are achieved by bent functions. However, an important issue to know how small two indicators can be for balanced Boolean functions is still open.

2 Bounds on σ_f and Δ_f for a balanced function f

In this section, we assume that f is a balanced Boolean function on Z_2^n. Let $\zeta_x = \#\{y \in Z_2^n | f(y) = f(x \oplus y)\}$. Then,

Lemma 1

$$\begin{array}{ll} (i) & \sum_x \zeta_x = 2^{2n-1}, \\ (ii) & \sum_x \zeta_x^2 = 4 \sum_x \{\sum_y f(y) f(x \oplus y)\}^2. \end{array}$$

Proof

Since $\zeta_x = \sum_y (2 f(y) f(x \oplus y) - f(y) - f(x \oplus y) + 1)$, we have

$$\begin{aligned} \sum_x \zeta_x &= \sum_x \sum_y (2 f(y) f(x \oplus y) - f(y) - f(x \oplus y)) + 2^{2n} \\ &= 2 \sum_y f(y) \sum_x f(x \oplus y) - 2^n \sum_y f(y) - \sum_x \sum_y f(x \oplus y) + 2^{2n} \\ &= 2^{2n-1}. \end{aligned}$$

The rest can be proved similarly.

Since $C_f(x) = 2\zeta_x - 2^n$, by Lemma 1,

Lemma 2

$$\sum_x C_f^2(x) = 16\sum_x(\sum_y f(y)f(x\oplus y))^2 - 2^{3n}.$$

The following is useful in proving our main theorem.

Lemma 3 *Let* $\sum_{i=1}^t x_i = X$. *Then*

$$\sum_{i=1}^t x_i^2 \geq 2X\lfloor X/t\rfloor - t(\lfloor X/t\rfloor)^2 + X - t\lfloor X/t\rfloor.$$

Proof
Let $\lfloor X/t\rfloor = M$ and $m_i = x_i - M$. Then,

$$\sum_{i=1}^t m_i = \sum_{i=1}^t (x_i - M) = X - tM \geq 0. \tag{1}$$

Therefore,

$$\begin{aligned}\textstyle\sum_{i=1}^t x_i^2 &= \textstyle\sum_{i=1}^t (M+m_i)^2 \\ &= tM^2 + 2XM - 2tM^2 + \textstyle\sum_{i=1}^t m_i^2 \\ &\geq tM^2 + 2XM - 2tM^2 + X - tM.\end{aligned}$$

We now derive a lower bound on σ_f.

Theorem 1 *For* $n \geq 3$, $\sigma_f \geq 2^{2n} + 2^{n+3}$.

Proof
By Lemma 2,

$$\begin{aligned}\textstyle\sum_x C_f^2(x) &= 16\textstyle\sum_x(\sum_y f(y)f(x\oplus y))^2 - 2^{3n} \\ &= 16\{\textstyle\sum_y(f^2(y)\sum_x f^2(x\oplus y)) \\ &\quad +2\textstyle\sum_{i<j} f(y_i)f(y_j)\sum_x f(x\oplus y_i)f(x\oplus y_j)\} - 2^{3n}.\end{aligned}$$

Let $B_f = \{y \in Z_2^n \mid f(y) = 1\}$. Since f is a balanced function, $\#B_f = 2^{n-1}$. Then

$$\begin{aligned}\sum_x C_f^2(x) &= 2^4\{\sum_{y\in B_f}\sum_x f^2(x\oplus y) + 2\cdot\sum_{i<j,y_i,y_j\in B_f}\sum_x f(x\oplus y_i)f(x\oplus y_j)\} - 2^{3n} \\ &= 2^{2n+2} + 2^5\cdot\sum_{i<j,y_i,y_j\in B_f}\sum_x f(x\oplus y_i)f(x\oplus y_j) - 2^{3n}. \end{aligned}\tag{2}$$

Let

$$Y_k = \{(y_l, y_m) \in B_f \times B_f \mid l < m \ \ and \ \ y_l \oplus y_m = y_k\},$$

for $k = 1, \cdots, 2^n - 1$. Since $y_l \oplus y_m = (x \oplus y_l) \oplus (x \oplus y_m)$, if $(y_l, y_m) \in Y_k$ and $f(x \oplus y_l) f(x \oplus y_m) = 1$, then either $(x \oplus y_l, x \oplus y_m) \in Y_k$ or $(x \oplus y_m, x \oplus y_l) \in Y_k$. Therefore,

$$\sum_{i<j, y_i, y_j \in B_f} \sum_x f(x \oplus y_i) f(x \oplus y_j) = 2 \sum_{k=1}^{2^n-1} \#Y_k^2.$$

Hence, by eq.(2),

$$\sum_x C_f^2(x) = 2^{2n+2} + 2^6 \sum_{k=1}^{2^n-1} \#Y_k^2 - 2^{3n}. \tag{3}$$

On the other hand, since $\sum_{k=1}^{2^n-1} \#Y_k = 2^{n-1}(2^{n-1} - 1)/2$, by Lemma 3, we have

$$\begin{aligned} \sum_k \#Y_k^2 &\geq (2^{n-3})^2 (2^n - 1 - 2^{n-3}) + (2^{n-3} - 1)^2 \cdot 2^{n-3} \\ &= 2^{3n-6} - 3 \cdot 2^{2n-6} + 2^{n-3}. \end{aligned} \tag{4}$$

Therefore, by eq.(3)

$$\sigma_f = \sum_x C_f^2(x) \geq 2^{2n+2} + 2^6 (2^{3n-6} - 3 \cdot 2^{2n-6} + 2^{n-3}) - 2^{3n} = 2^{2n} + 2^{n+3}.$$

For $n \geq 3$,

$$\begin{aligned} C_f(x) &= 2\zeta_x - 2^n \\ &= 2 \textstyle\sum_y (2 f(y) f(x \oplus y) - f(y) - f(x \oplus y) + 1) - 2^n \\ &= 4 \textstyle\sum_y f(y) f(x \oplus y) - 2^n. \end{aligned}$$

Since $\sum_y f(y) f(x \oplus y)$ is even, we have

Corollary 1 *For $n \geq 3$, $\Delta_f \geq 8$.*

If the equality of equation (4) holds, then by equation (1), we can see that there are $2^n - 2^{n-3} - 1$ x's with $\Delta_f(x) = 0$, and 2^{n-3} x's with $\Delta_f(x) = -8$, and $\Delta_f(0) = 2^n$. Thus we can derive a lower bound of Δ_f for balanced functions.

Corollary 2 *For $n \geq 3$, $\Delta_f \geq 8$.*

In the case of $n = 3$, the lower bound on σ_f fits. But in the case of $n = 4$ and $n = 5$, the lower bounds on σ_f have gaps with real minimal values.(For $n = 4$ and $n = 5$, 256 and 384, respectively.) But it is for the value 2^6, the coefficient of $\sum_{k=1}^{2^n-1} \#C_k^2$ in (4). In the view of $\#C_k^2$, it has small difference only 4 and 6 for $n = 4$ and $n = 5$, respectively. So, in a large dimension, it is thought that the lower bound will be more similar with the real minimum of σ_f. Table1 shows the values of σ_f and $\#C_k$ for balanced function f having real minimum value and our lower bound when $n = 3$ and $n = 4$.

n	σ_f	$(f(\alpha_0), \cdots, f(\alpha_{2^n-1})$	$(\#C_1, \cdots, \#C_{2^n-1})$
			minimum value
3	128	(1,1,1,0,1,0,0,0)	(1,1,1,1,1,1,1,0)
			our lower bound
3	128		(1,1,1,1,1,1,0)
			minimum value
4	640	(1,1,1,1,1,0,1,0,1,1,0,0,0,0,0,0)	(3,3,2,2,2,2,2,2,2,2,2,1,1,1,1)
			our lower bound
4	384		(2,2,2,2,2,2,2,2,2,2,2,2,2,1,1)
			minimum value
5	1664	(0,0,1,1,1,1,0,1,1,1,1,0,1,1,0,0, 1,1,1,1,1,0,0,0,1,0,0,0,0,0,0,0)	(6,4,4,4,4,4,4,4,4,4,4,4,3,3,4,4,4, 4,3,3,4,4,4,4,4,4,4,4,4,4,3,3)
			our lower bound
5	1280		(4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4, 4,4,4,4,4,4,4,4,4,4,4,4,3,3,3,3)

Table 1: Balanced functions having lower bound on σ_f of dimension $n = 3$, 4, 5

3 Balanced Boolean functions having Good properties on GAC

Lemma 4 $g(u, x_1, \cdots, x_n) = (1 \oplus u)f_1(x_1, \cdots, x_n) \oplus uf_2(x_1, \cdots, x_n)$

$$< \xi_1, l > \leq P_1 \text{ and } < \xi_2, l > \leq P_2.$$

for $\xi_i = (-1)^{f_i(x)}$ *and* l *: affine sequence of length* 2^n*. Then,*

$$N_g \geq 2^n - \frac{1}{2}(P_1 + P_2).$$

Proof
By properties of Walsh-Hadamard transform,

$$H_{n+1} \cdot g(u, x_1, \cdots, x_n) = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix} \cdot ((-1)^{f_1(x_1,\cdots,x_n)}(-1)^{f_2(x_1,\cdots,x_n)}).$$

Using above assumption, we can show

$$N_g \geq 2^n - \frac{1}{2}(P_1 + P_2).$$

We can generalize above lemma to concatenating 2^t's Boolean function on Z_2^n $f_0, \cdots, f_{2^t-1}$ to Boolean function on Z_2^{n+t}.

Theorem 2 *Let ξ_i be the sequence of f_i such that $< \xi_i, l > \leq P_i$, $\forall 0 \leq i \leq 2^t - 1$ for all affine sequence l of length 2^n. Let g be a Boolean function such that*

$$\begin{aligned} g(y, x) &= \oplus_{i=0}^{2^t-1} D_{\alpha_i}(y) f_i(x). \\ D_{\alpha_i}(y) &= (y_1 \oplus i_1 \oplus 1) \cdots (y_t \oplus i_t \oplus 1), \text{for } \alpha_i = (i_1, \cdots, i_t). \end{aligned}$$

Then, $N_g \geq 2^{n+t-1} - \frac{1}{2}\sum_0^{2^t-1} P_i$.

(When f_i's are all bent function, $N_g \geq 2^{n+t-1} - 2^{\frac{1}{2}n+t-1}$.)

Let f and h are Boolean functions in Z_2^n such that f is a balanced function satisfying $< f, l > \leq P$ for all affine functions l on Z_2^n and h is a bent function. Then we can construct an Balanced Boolean function g on Z_2^{n+2}:

$$g(x_1, \cdots, x_n, x_{n+1}, x_{n+2}) = f||h||f||\bar{h},$$

The algebraic normal form of g is

$$g = (1 \oplus x_{n+1})(1 \oplus x_{n+2}) f \oplus x_{n+1}(1 \oplus x_{n+2}) h \oplus (1 \oplus x_{n+1}) x_{n+2} f \oplus x_{n+1} x_{n+2} \bar{h}.$$

Since

$$\Delta_g(\alpha) = \begin{cases} 2(\Delta_f(\alpha) + \Delta_h(\alpha)) & \text{when } \alpha = (\alpha_n, 0, 0) \text{ for } \alpha_n \in Z_2^n, \\ 0 & \text{when } \alpha = (\alpha_n, 1, 0) \text{ for } \alpha_n \in Z_2^n, \\ 2(\Delta_f(\alpha) - \Delta_h(\alpha)) & \text{when } \alpha = (\alpha_n, 0, 1) \text{ for } \alpha_n \in Z_2^n, \\ 0 & \text{when } \alpha = (\alpha_n, 1, 1) \text{ for } \alpha_n \in Z_2^n, \end{cases},$$

we can get σ_g as follows:

$$\begin{aligned}\sigma_g &= \sum_{\alpha \in Z_2^{n-2}} \triangle_g^2(\alpha) \\ &= 4\{\sum_{\alpha_n \in Z_2^n}(\triangle_f(\alpha_n) + \triangle_h(\alpha_n)^2 + \sum_{\alpha_n \in Z_2^n}(\triangle_f(\alpha_n) - \triangle_h(\alpha_n)^2\} \\ &= 8\sum_{\alpha_n \in z_2^n}(\triangle_f^2(\alpha_n) + \triangle_h^2(\alpha_n)) \\ &= 8(\sigma_f + 2^{2n})\end{aligned}$$

Using of the above lemma, we get

$$N_g \geq 2^{n+1} - 2^{\frac{1}{2}n} - P.$$

If we apply f of dimension 4 in Table1 on the above formula of construction g, recursively, we can have the same result $2^{2n} + 2^{\frac{3}{2}+3} - 2^{\frac{3}{2}+1}$ on σ_g of functions made by modification of bent functions at [6]

References

[1] B. Preneel, W.V. Leekwijck, L.V. Linden, R. Govaerts, and J. Vandewalle, *"Propagation characteristics of Boolean function"*, Advances in Cryptology–EuroCrypt'90, pp. 161–173, Springer–Verlag, 1991.

[2] B. Preneel, *"Analysis and design of cryptographic hash functions"*, Ph.D. Dissertation, Katholieke Universiteit Leuven, 1993.

[3] O. S. Rothaus, *"On bent functions"*, Journal of Combinatorial Theory, ser. A, vol 20, pp. 300-305, 1976.

[4] J. Seberry, X-M. Zhang and Y. Zheng, *"Nonlinearly balanced Boolean functions and their propagation characteristics"*, Advances in Cryptology–Crypto'93, Pre-Proceedings, pp. 6.1–6.12, 1994.

[5] J. Seberry, Xian-Mo Zhang, Yuliang Zheng, *"Nonlinearity and Propagation Characteristics of Balanced Boolean Functions"*, Information and Computation, vol. 119, pp. 1–13, 1995.

[6] X-M. Zhang and Y. Zheng, *"Gac-the Criterion for Global Avalanche Characteristics of Cryptographic Functions"*, J. Universal Computer Science, vol. 1, no. 5, pp. 320–337, 1995.

[7] X-M. Zhang and Y. Zheng, *"Auto-Correlation and New Bounds on the Nonlienarity of Boolean Functions"*, Advances in Cryptology–EuroCrypt'96, pp. 294–306, Springer–Verlag, 1996.