Construction of Highly Nonlinear Injective S-boxes With Application to CAST-like Encryption Algorithms

A. M. Youssef, Z.G. Chen, and S.E. Tavares
Department Of Electrical and Computer Engineering
Queen's University, Kingston, Ontario, Canada, K7L 3N6
E-mail: tavares@ee.queensu.ca
http://adonis.ee.queensu.ca:8000

Abstract: In this paper we present two methods for constructing highly nonlinear injective s-boxes. Both of these methods, which are based on exponential sums, outperform previously proposed methods. In particular, we are able to obtain injective 8×32 s-boxes with nonlinearity equal to 80 and maximum XOR table entry of 2. We also re-evaluate the resistance of the CAST-like encryption algorithms constructed using randomly selected s-boxes to the basic linear cryptanalysis.

1. Introduction

Differential cryptanalysis [3] and linear cryptanalysis [7] are powerful cryptanalytic attacks on private-key block ciphers. The complexity of differential cryptanalysis depends on the size of the largest entry in the XOR table, the total number of zeroes in the XOR table, and the number of nonzero entries in the first column in that table [3], [12]. The complexity of linear cryptanalysis depends on the size of the largest entry in the linear approximation table (LAT) [8].

One way to reduce the size of the largest entry in the XOR table is to use injective substitution boxes (s-boxes) such that the number of output bits of the s-box is sufficiently larger than the number of input bits. In this way, it is very likely that the entries in the XOR distribution table of a randomly chosen injective s-box will have only small values, making the block cipher resistant to differential cryptanalysis. Some proposed block ciphers, such as CAST [1] and Blowfish [11], take advantage of this property.

On the other hand, Biham proved that if for an $n \times m$ s-box described by $f : Z_2^n \rightarrow Z_2^m$ we have $m \geq 2^n - n$, then at least one linear combination of the output bits must be an affine combination of the input bits and the block cipher can be trivially broken by linear cryptanalysis.

Highly nonlinear 8×32 s-boxes are particularly interesting because some practical ciphers are based on s-boxes with these dimensions [1], [11]. The expected value of the nonlinearity of randomly selected injective s-boxes is studied in [14] where it was found that the expected nonlinearity of a randomly selected 8×32 s-boxes is about 72. Mister and Adams [9] presented a construction method for injective s-boxes from bent functions. Although their basic objective was to obtain highly nonlinear 8×32 s-boxes, they were only able to achieve nonlinearity of 74 or less.

In this paper we present two construction methods for injective s-boxes. Both of these methods outperform previously proposed methods. In particular, we are able to obtain injective 8×32 s-boxes with nonlinearity equal to 80. We also re-evaluate the resistance of the CAST-like encryption algorithms to the basic linear cryptanalysis.

Throughout this paper we assume a field of characteristic 2. Our construction methods are based on the following Lemmas.

Lemma 1(Carlitz and Uchiyama bound [4])

If $F(\mathbf{x})$ is a polynomial over $GF(2^n)$ of degree r such that $F(\mathbf{x}) \neq G(\mathbf{x})^2 + G(\mathbf{x}) + \mathrm{b}$ for all polynomials $G(\mathbf{x})$ over $GF(2^n)$ and constants $\mathrm{b} \in GF(2^n)$, then

$$\left| \sum_{\mathbf{x}} (-1)^{Tr(F(\mathbf{x}))} \right| \leq (r-1)2^{n/2}. \tag{1}$$

where $Tr(\mathbf{x}) = \sum_{i=0}^{n-1} \mathbf{x}^{2^i}$ is the trace of $\mathbf{x} \in GF(2^n)$.

Lemma 2(Kloosterman sum [13],[4])

$$\left| \sum_{\mathbf{x} \neq 0} (-1)^{Tr\left(\mathbf{x} + \frac{c}{\mathbf{x}}\right)} \right| \leq 2^{n/2+1}. \tag{2}$$

Note that a function, F, over $GF(2^n)$ can also be expressed as a function over $GF(2)^n$, i.e., as n functions over $GF(2)$. For $x = \{x_1, \cdots, x_n\}$ let

$$f(x) = (f_1(x), \cdots, f_n(x))$$

be a function over $GF(2)^n$, let $\mathrm{B} = \{\alpha_1, \cdots, \alpha_n\}$ be any basis of $GF(2^n)$ over $GF(2)$, then the function

$$F(\mathbf{x}) = \sum_{i=1}^{n} f_i(x)\alpha_i \tag{3}$$

where $\mathbf{x} = \sum_{i=1}^{n} x_i \alpha_i \in GF(2^n)$. This means that there is a one-to-one correspondence between the functions of

0-7803-3716-6 /97/$5.00 © 1997 IEEE

$GF(2^n)$ and those of $GF(2)^n$ under a chosen basis of $GF(2^n)$ over $GF(2)$. If we let $\{\alpha_1^*, \cdots, \alpha_n^*\}$ be the dual basis of B, then each component of $f(x)$ can be expressed as

$$f_i(x) = Tr(F(\mathrm{x})\alpha_i^*). \tag{4}$$

Nonlinearity: The nonlinearity of the function $f(x) = (f_1(x), \cdots, f_n(x))$ is defined as the minimum Hamming distance between the set of affine functions and every nonzero linear combination of the output coordinates of f, i.e.,

$$\mathcal{NL}_f = \min_{a,b,w} \#\{x \in \mathrm{Z}_2^n | a \cdot f(x) \neq w \cdot x \oplus b\}, \tag{5}$$

where $a \in \mathrm{Z}_2^n$, $w \in \mathrm{Z}_2^n \backslash \{0\}$, $b \in \mathrm{Z}_2$ and $w \cdot x$ denotes the dot product between w and x over Z_2.

From the above, one can easily prove that the nonlinearity of the function f that corresponds to the function F is given by

$$\begin{aligned} \mathcal{NL}(f) &= \min_{\mathbf{c}\neq 0,\mathbf{w},\mathrm{b}} d(Tr(\mathbf{c}F(\mathbf{x})), Tr(\mathbf{wx}) \oplus b) \\ &= 2^{n-1} - \max_{\mathbf{c}\neq 0,\mathbf{w}} \left| \sum_{\mathbf{x}} (-1)^{Tr(\mathbf{c}F(\mathbf{x})+\mathbf{wx})} \right|, \end{aligned} \tag{6}$$

where $\mathbf{c}, \mathbf{w} \in GF(2^n)$, $b \in GF(2)$ and

$$d(f_1, f_2) = \#\{x \in \mathrm{Z}_2^n \mid f_1(x) \neq f_2(x)\}. \tag{7}$$

2. Results

2.1 Construction Method I

This construction method is based on the observation that highly nonlinear injective s-boxes may be obtained by adding the coordinate functions of highly nonlinear bijective s-boxes.

Given the distinct bijective functions F_i, $1 \leq i \leq M$, over $GF(2^n)$ an injective s-box G with n inputs and nM outputs can be obtained by setting

$$G = (F_1 || F_2 || \cdots || F_M). \tag{8}$$

In this method we use the inversion mapping proposed by Nyberg [10]

$$F_i(\mathrm{x}) = \begin{cases} (\mathrm{x} + \mathrm{a}_i)^{-1}, & \mathrm{x} \neq \mathrm{a}_i, \\ 0, & \mathrm{x} = \mathrm{a}_i, \end{cases} \tag{9}$$

where $\mathrm{x} \in GF(2^n)$.

Using Lemma 2, it is easy to see that the nonlinearity of the function F_i is lower bounded by $2^{n-1} - 2^{n/2}$.

Experimental results show that injective 8 × 16 s-boxes constructed by this method always have nonlinearity of 96. For 8 × 24 and 100 random choices of a_i pairs, we found 57, 40 and 3 s-boxes with $\mathcal{NL} = 86, 84$ and 80 respectively. The only 8 × 32 s-box tested to date has nonlinearity of 76.

Conjecture:

The nonlinearity of the function

$$G = \left(\mathrm{x}^{-1} || (\mathrm{x} + \mathrm{a})^{-1}\right), \mathrm{a} \neq 0, \mathrm{x} \in GF(2^n) \tag{10}$$

obtained using construction method I, is bounded by

$$\mathcal{NL}_G \geq 2^{n-1} - 2^{n/2+1}. \tag{11}$$

We verified this conjecture for $n \leq 10$ where we found that this bound is tight for even n.

Remark: We can prove that $\mathcal{NL}_G \geq 2^{n-1} - (2^{n/2+1} + 1)$ by proving that for $\mathrm{a} \neq 0$ we have

$$\left| \sum_{\mathrm{x} \in GF(2^n) \backslash \{0,\mathrm{a}\}} (-1)^{Tr\left(\mathrm{x} + \frac{\mathrm{c}}{\mathrm{x}} + \frac{\mathrm{d}}{\mathrm{x}+\mathrm{a}}\right)} \right| \leq 2^{n/2+2}, \tag{12}$$

which seems to be a hard problem.

2.2 Construction Method II

This method is also based on the observation that highly nonlinear injective s-boxes may be obtained by adding the coordinate functions of highly nonlinear s-boxes (not necessary bijective).

If the concatenated functions F_i's are distinct polynomials over $GF(2^n)$ such that

$$\mathbf{wx} + \sum_{i=1}^{M} \mathrm{a}_i F_i(\mathrm{x}) \neq U(\mathrm{x})^2 + U(\mathrm{x}) + \mathrm{b} \tag{13}$$

for all polynomials $U(\mathrm{x})$ over $GF(2^m)$ and constants $\mathrm{a}_i \in GF(2^n) \setminus \{0\}$, $\mathrm{b}, \mathbf{w} \in GF(2^n)$ then the Carlitz and Uchiyama bound can be used to provide a lower bound the nonlinearity of the resulting s-box as follows.

If $r = \max_i (degree(F_i))$ then the nonlinearity of the resulting function is lower bounded by

$$\mathcal{NL}_G \geq 2^{n-1} - 2^{n/2-1}(r-1). \tag{14}$$

Using the Carlitz and Uchiyama bound it is easy to check that the nonlinearity of the function $F(\mathrm{x}) = \mathrm{x}^3, \mathrm{x} \in GF(2^8)$ is lower bounded by 112. Also the nonlinearity of the function $G(\mathrm{x}) = (\mathrm{x}^3 || \mathrm{x}^5), \mathrm{x} \in GF(2^8)$ is lower bounded by 96.

Our basic result is based on the experimental observation that the Carlitz and Uchiyama bound is not tight for higher values of r. In this paper we consider the following five constructions

$$\begin{aligned} G_1 &= (\mathrm{x}^5 || \mathrm{x}^7 || \mathrm{x}^{11} || \mathrm{x}^{13}), \\ G_2 &= (\mathrm{x}^3 || \mathrm{x}^7 || \mathrm{x}^{11} || \mathrm{x}^{13}), \\ G_3 &= (\mathrm{x}^3 || \mathrm{x}^5 || \mathrm{x}^{11} || \mathrm{x}^{13}), \\ G_4 &= (\mathrm{x}^3 || \mathrm{x}^5 || \mathrm{x}^7 || \mathrm{x}^{13}), \\ G_5 &= (\mathrm{x}^3 || \mathrm{x}^5 || \mathrm{x}^7 || \mathrm{x}^{11}). \end{aligned} \tag{15}$$

Using the Carlitz and Uchiyama bound we have

$$\mathcal{NL}_{G_i} \geq \begin{cases} 32, & i = 1, 2, 3, 4, \\ 48, & i = 5. \end{cases} \tag{16}$$

Experimental results shows that

$$\mathcal{NL}_{G_i} = \begin{cases} 72, & i = 1,2, \\ 80, & i = 3,4,5. \end{cases} \qquad (17)$$

Let $F_i = \mathbf{x}^3, \mathbf{x}^5, \mathbf{x}^7, \mathbf{x}^{11}, \mathbf{x}^{13}$ for $i = 1,2,3,4,5$ respectively. By noting that F_i is bijective for $i = 3,4,5$ then any construction that includes any of F_i, $i = 3,4$ or 5 will be injective. In fact, it is not hard to see that all the s-boxes below are injective.

Table 1 below shows the nonlinearity and the maximum XOR table entry, XOR^*, for the injective 8×16 s-boxes constructed by this method by concatenating F_i with F_j. Table 2 shows similar results for the 8×24 s-boxes.

i,j	1,2	1,3	1,4	1,5	2,3	2,4	2,5	3,4	3,5	4,5
$\mathcal{NL}$	96	96	80	80	96	96	96	88	96	96
XOR^*	2	2	2	2	4	4	4	4	4	4

Table 1 $\mathcal{NL}, XOR^*$ for 8×16 S-boxes Obtained Using Construction Method II

i,j,k	1,2,3	1,2,4	1,2,5	1,3,4	1,3,5	1,4,5	2,3,4	2,3,5	2,4,5	3,4,5
$\mathcal{NL}$	96	80	80	80	80	80	88	80	88	80
XOR^*	2	2	2	2	2	2	4	4	4	4

Table 2 $\mathcal{NL}, XOR^*$ for 8×24 S-boxes Obtained Using Construction Method II

For the 8×32 s-boxes, $XOR^* = 2$ for all the constructions in (15) except for G_1 where it is equal to 4 which may limit the usefulness of G_1.

Table 3 shows the best s-box nonlinearity obtained by different methods

	8×16	8×24	8×32
Random	*87*	*80*	*73*
Mister and Adams [9]	—	—	*74*
Method I	*96*	*86*	*76*
Method II	*96*	*96*	*80*

Table 3 Best S-box Nonlinearity Obtained by Different Construction Methods

The construction methods proposed in this paper can be extended to other highly nonlinear mappings such as those proposed in [5],[10].

In order to frustrate possible algebraic attacks, the four 8×32 s-boxes should be generated using different irreducible polynomials. When using s-boxes obtained from construction method II, we recommend that the bytes XORed together should correspond to different degrees. For example, if G_5 is used, then the 8×32 s-boxes may be constructed as follows

$$\begin{aligned} s_1 &= (\mathbf{x}^3 \| \mathbf{x}^5 \| \mathbf{x}^7 \| \mathbf{x}^{11}), \\ s_2 &= (\mathbf{x}^5 \| \mathbf{x}^7 \| \mathbf{x}^{11} \| \mathbf{x}^3), \\ s_3 &= (\mathbf{x}^7 \| \mathbf{x}^{11} \| \mathbf{x}^3 \| \mathbf{x}^5), \\ s_4 &= (\mathbf{x}^{11} \| \mathbf{x}^3 \| \mathbf{x}^5 \| \mathbf{x}^7), \end{aligned} \qquad (18)$$

such that the exponents form a Latin square.

3. Comments on the Security of CAST-like Encryption Algorithms

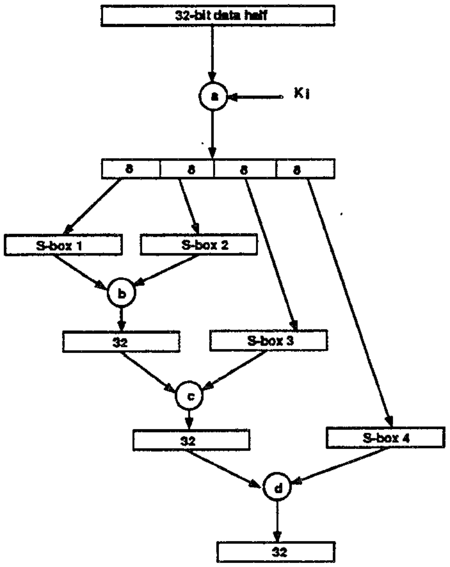

Figure 1 CAST Round Function

Figure 1 shows the CAST round function. In this paper we assume that operations a, b, c and d are XOR addition of 32 bit quantities.

The resistance of CAST-like encryption algorithms [1] constructed using randomly generated s-boxes against the basic linear cryptanalysis [8] was studied in [6]. The number of known plaintexts, N_p, in a basic linear attack (Algorithm 1 in [8]) required to give a 97.7% confidence of getting the right key bit is approximately given by

$$N_p \approx \left| p_l - \frac{1}{2} \right|^{-2}, \qquad (19)$$

$$\left| p_l - \frac{1}{2} \right| \leq 2^{\gamma-1} \left| p_s - \frac{1}{2} \right|^{\gamma}, \qquad (20)$$

where γ is the number of s-boxes involved in the R-round linear approximation expression, $\left| p_s - \frac{1}{2} \right| = \frac{2^{n-1} - \mathcal{NL}_S}{2^n}$ and n is the number of input bits to the s-box. The bounds for N_p can be calculated by re-evaluating the expressions in [6] using the new 8×32 s-boxes nonlinearity. However, a better bound can be obtained by considering the nonlinearity of the resulting 32×32 s-boxes. The nonlinearity, $\mathcal{NL}_S$ can be bounded using the nonlinearity, $\mathcal{NL}_{s_i}, 1 \leq i \leq 4$, of the four 8×32 used in its construction as follows

$$\mathcal{NL}_S \geq 2^{32-1} - \frac{1}{2} \prod_{i=1}^{4} (2^8 - 2\mathcal{NL}_{s_i}). \qquad (21)$$

The exact nonlinearity can be efficiently calculated using the Walsh transforms [2] of the four 8×32 s-boxes. Since an R-round linear approximation must involve at least as many s-boxes as $R/2$ iterations of the best 2-round approximation, the number of 32×32 s-boxes involved in an R-round linear approximation is at least $R/2$ and hence we have

$$N_p \approx \frac{\left|\frac{1}{2} - \frac{\mathcal{NL}_S}{2^{32}}\right|^{-R}}{2^{R-2}} \qquad (22)$$

where $\mathcal{NL}_S$ is the nonlinearity of the 32×32 s-box.

An important observation, which was overlooked in [9] is that the nonlinearity of the 32×32 s-box depends not only on the nonlinearity of 8×32 s-boxes used in the construction, but it also depends on how the four 8×32 s-boxes interact together. This means that improving the nonlinearity of the individual 8×32 s-boxes does not always guarantee improving the resistance of the cipher to the basic linear cryptanalysis. For example, when combining the output of the four CAST s-boxes [1] (each with nonlinearity 74) by XOR, the resulting 32×32 s-box has nonlinearity $2,132,774,912$ which is less than $2,133,721,088$, the nonlinearity we got by combining four randomly selected s-boxes with nonlinearity less than 74. Using these randomly selected s-boxes, we have $N_p \approx 2^{60}$ for $R = 8$ which is much higher than $N_p \approx 2^{34}$ estimated in [6].

Finally, we note that the primary motive for this work is to obtain highly nonlinear injective s-boxes. We are not proposing the use of such s-boxes in CAST-like ciphers before examining their other cryptographic properties. In fact, we believe that randomly selected 8×32 s-boxes are a good choice for CAST-like ciphers.

References

[1] C. M. Adams. Constructing symmetric ciphers using the CAST design procedure. *to appear (also available through www.entrust.com)*, July 4,1996.

[2] N. Ahmed and K.R. Rao. *Orthogonal Transforms for Digital Signal Processing*. Springer-Verlag, New York, 1975.

[3] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology, vol. 4, no. 1, pp. 3–72*, 1991.

[4] L. Carlitz and S. Uchiyama. Bounds for exponential sums. *Duke Math. Journal, Vol. 24, pp. 37–41*, 1957.

[5] X. Chang, Z. Dai, and G. Gong. Some cryptographic properties of exponential functions. *Advances in Crytology: Proc. of ASIACRYPT '94, Springer-Verlag, pp. 415–418*, 1995.

[6] J. Lee, H.M. Heys, and S.E. Tavares. On the resistance of the CAST encryption algorithm to differential and linear cryptanalysis. *accepted for publication*, 1996.

[7] M. Matsui. The first experimental cyptanalysis of the Data Encryption Standard. *Advances in Cryptology: Proc. of CRYPTO '94, Springer-Verlag, Berlin. pp. 1–11*, 1994.

[8] M. Matsui. Linear cryptanalysis method for DES cipher. *Advances in Cryptology: Proc. of EUROCRYPT '93, Springer-Verlag, Berlin. pp. 386–397*, 1994.

[9] S. Mister and C. Adams. Practical s-box design. *Workshop on Selected Areas in Cryptography, SAC '96, Workshop Record, pp. 61–76*, 1996.

[10] K. Nyberg. Differentially uniform mappings for cryptography. *Advances in Cryptology: Proc. of EUROCRYPT '93, Springer-Verlag, Berlin, pp55–64*, 1994.

[11] B. Schneier. Description of a new variable-length key, 64–bit block cipher (Blowfish). *Proc. of Fast Software Encryption Workshop, LNCS 809, Springer-Verlag, Berlin, pp. 191–204*, 1994.

[12] J. Seberry, X. Zhang, and Y. Zheng. Systematic generation of cryptographically robust s-boxes. *1st ACM Conference on Computer and Communications Security, Fairfax, Virginia, pp.172–182*, November, 1993.

[13] A. Weil. On some exponential sums. *Proceedings of the National Academy of Sciences, Vol. 34, pp.204–207*, 1948.

[14] A.M. Youssef, S.E. Tavares, S. Mister, and C.M. Adams. Linear approximation of injective s-boxes. *Electronics letters, Vol.31, No. 25, pp 2165–2166*, 1995.