On the Security of the CAST Encryption Algorithm

H. M. Heys and S. E. Tavares

Department of Electrical and Computer Engineering
Queen's University
Kingston, Ontario, Canada
email: tavares@ee.queensu.ca

Abstract — In this paper we examine a new private key encryption algorithm referred to as CAST. Specifically, we investigate the security of the cipher with respect to linear cryptanalysis. From our analysis we conclude that it is easy to select S-boxes so that an efficient implementation of the CAST algorithm is demonstrably resistant to linear cryptanalysis.

I. Introduction

The CAST encryption algorithm [1][2] belongs to a class of private key block ciphers which are composed of substitution boxes (S-boxes) with fewer input bits than output bits. Recently, a software implementation of the CAST cipher has been developed for application in computer security products [3]. In [2], it is suggested that, with an appropriate number of substitution rounds, the CAST cipher is resistant to differential cryptanalysis [4]. In this paper, we show that the CAST cipher, with an appropriate number of rounds, is resistant to linear cryptanalysis [5].

Similarly to the Data Encryption Standard [6] and other proposed block ciphers, the CAST algorithm consists of a series of rounds of substitutions in order to achieve the "confusion" and "diffusion" principles suggested by Shannon [7]. The basic algorithm structure is illustrated in Figure 1. The algorithm encrypts by dividing the N-bit plaintext input block in half. The right half-block, $\mathbf{R}_1$, is transformed by a round function F and then XORed bit-by-bit to the left half-block, $\mathbf{L}_1$. The right and left halves are then swapped. This is repeated for the number of rounds in the cipher, R. Consequently, the algorithm may be viewed as the following iterated operation:

$$\begin{aligned}\mathbf{R}_{i+1} &= F(\mathbf{R}_i, \mathbf{K}_i) \oplus \mathbf{L}_i \\ \mathbf{L}_{i+1} &= \mathbf{R}_i .\end{aligned} \tag{1}$$

This format is identical to all so-called DES-like ciphers. The novelty of CAST is the round function F. The round function is implemented using S-boxes of dimension $m \times n$ where $n = N/2$ and $m < n$.

The example CAST system presented in [2] uses four 8×32 S-boxes to implement a 64–bit block cipher. The inputs to the four S-boxes are determined by XORing a 32–bit sub-key[1] $\mathbf{K}_i$ to the right half-block and then dividing the half-block into four 8–bit groups. The 32–bit outputs of the four S-boxes are XORed together to produce the output of the function F. Letting $S_j(\mathbf{X})$ represent the 32–bit output of the j-th S-box given an 8–bit input $\mathbf{X}$, the operation of the round function may be represented by

$$F(\mathbf{R}_i, \mathbf{K}_i) = \bigoplus_{j=1}^{4} S_j\left(\mathbf{R}_i^{(j)} \oplus \mathbf{K}_i^{(j)}\right) \tag{2}$$

where $\mathbf{R}_i^{(j)}$ is the j-th byte of $\mathbf{R}_i$ and $\mathbf{K}_i^{(j)}$ is the j-th byte of $\mathbf{K}_i$.

In the following section we consider the application of linear cryptanalysis to the CAST algorithm.

II. Linear Cryptanalysis of CAST

Linear cryptanalysis [5] is one of the most powerful attacks on private key block ciphers. It has been shown that DES is theoretically susceptible to the attack using 2^{47} known plaintexts.

To attack a block cipher using a basic linear cryptanalysis technique, the cryptanalyst is interested in the best linear approximation of the form:

$$P_{i_1} \oplus ... \oplus P_{i_\gamma} \oplus C_{j_1} \oplus ... \oplus C_{j_\zeta} = K_{k_1} \oplus ... \oplus K_{k_\theta} . \tag{3}$$

Let the probability that equation (3) is satisfied be represented by p_L. If the magnitude $|p_L - 1/2|$ is large enough and sufficient plaintext-ciphertext pairs

[1] The sub-keys are determined by a key scheduling algorithm. For details of the CAST key scheduling algorithm see [8].

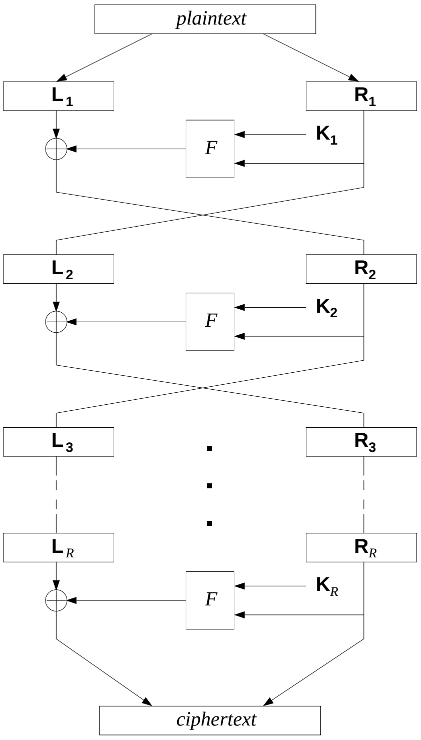

Figure 1. CAST Encryption Algorithm

are known, the equivalent of one key bit, expressed by the XOR sum of the key bits on the right side of equation (3) may be determined as the value that most often satisfies equation (3).

An appropriate linear equation is derived by combining a number of linear approximations for the S-boxes of different rounds such that intermediate terms (i.e., terms that are not plaintext, ciphertext, or key terms) are cancelled. Let the best linear approximation of an S-box be satisfied with probability p_ϵ. If the number of S-box linear approximations combined to give the overall expression is α, it can be shown [5][9] that

$$|p_L - 1/2| \lesssim 2^{\alpha-1}|p_\epsilon - 1/2|^\alpha. \qquad (4)$$

From [5], the number of known plaintexts, N_L, required to give a 97.7% confidence of the correct key bit is given roughly by

$$N_L \approx |p_L - 1/2|^{-2}. \qquad (5)$$

It is obvious that N_L can be increased by decreasing $|p_L - 1/2|$. Hence, selecting S-boxes for which $p_\epsilon \rightarrow 1/2$ will clearly aid in resisting the attack. As well, increasing the number of S-box terms, α, involved in the system linear approximation, increases the number of known plaintexts required in the cryptanalysis.

Consider the following nonlinearity measures based on the Hamming distance to the nearest affine function.[2] Let the distance between two m-bit functions, f and g, be given by

$$d(f,g) = \sum_{\mathbf{X}\in\{0,1\}^m} [f(\mathbf{X}) \oplus g(\mathbf{X})]. \qquad (7)$$

Then the nonlinearity of an m-bit boolean function f is defined to be

$$\mathcal{N}(f) = \min_{g\in\mathcal{L}} \; d(f,g) \qquad (8)$$

where $\mathcal{L}$ is the set of all m-bit affine boolean functions.

The nonlinearity of an $m \times n$ S-box S is given as the minimum nonlinearity over all non-zero linear combinations of the S-box output functions:

$$\mathcal{N}(S) = \min_{a_1,\ldots,a_n\in\{0,1\},\; all\ a_i\neq 0} \mathcal{N}\left(\bigoplus_{i=1}^{n} a_i f_i\right) \qquad (9)$$

where f_i represents the m-bit function of the i-th output bit of the S-box.

For a network in which the S-boxes have nonlinearities satisfying $\mathcal{N}(S) \geq \mathcal{N}_{min}$, we have

$$|p_\epsilon - 1/2| = \frac{2^{m-1} - \mathcal{N}_{min}}{2^m}. \qquad (10)$$

Using the result of following theorem, it is possible to determine an approximate lower bound on the number of known plaintexts, N_L, required for linear cryptanalysis if a lower bound on the nonlinearity of the S-boxes is known.

Theorem 1: Consider a 64–bit, R-round CAST cipher which uses four 8×32 S-boxes. The number of known plaintexts, N_L, required in the cryptanalysis satisfies

$$N_L \gtrsim \frac{2^{2-4R}}{|p_\epsilon - 1/2|^{4R}}. \qquad (11)$$

[2] An m-bit affine function is defined to be a function of the form

$$f(\mathbf{X}) = a_0 \oplus a_1 X_1 \oplus \ldots \oplus a_m X_m \qquad (6)$$

where $\mathbf{X} = [X_1 \ldots X_m]$ represents the m-bit input and $a_i \in \{0,1\}$, $0 \leq i \leq m$.

Proof (Sketch): Since an output bit of the round function F is the result of the XOR of the corresponding output of all 4 S-boxes, a linear approximation of F must involve linear approximations of all 4 S-boxes. Assume that the half-block inputs to round i, $\mathbf{L}_i$ and $\mathbf{R}_i$, are known. Then the linear approximation of $\mathbf{L}_{i+2} = \mathbf{R}_{i+1}$ must involve the 4 S-boxes of round i and the linear approximation of $\mathbf{R}_{i+2}$ must involve at least the 4 S-boxes of round $i+1$. Hence, a 2–round linear approximation must involve at least 4 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 S-boxes involved in an R-round linear approximation must be at least $\alpha = 4 \cdot (R/2) = 2R$. Substituting into (4) and (5) results in (11). □

Consider a 64–bit, R-round CAST cipher with a 64–bit key. Assume that the 8×32 S-boxes have nonlinearities greater than or equal to $\mathcal{N}_{min} = 64$. (In Section III we discuss the generation of such S-boxes.) In this case, $|p_\epsilon - 1/2| = 2^{-2}$. The number of plaintexts required for linear cryptanalysis is listed in Table 1 for various values of R. For comparison we have included the corresponding values for DES. Note that the CAST algorithm is comparable in size to DES and is significantly more resistant to linear cryptanalysis.

It should be noted that the bound for N_L arises by determining the number of known plaintexts required to determine one key bit. To determine all key bits using linear cryptanalysis, the number of plaintexts will be significantly larger then the bounds listed. For example, for $R = 16$, $N_L >> 2^{66}$. Clearly, in comparison with a theoretical security level of 2^{64} (based on the number of encryptions required in an exhaustive key search), a 16–round CAST network constructed using S-boxes with $\mathcal{N}(S) \geq 64$ is secure against linear cryptanalysis in the strictest theoretical sense. As well, a 12–round CAST cipher requires much more than 2^{50} known plaintexts, an impractical memory requirement for a cryptanalyst and we can therefore state that such a cipher is secure against a practical linear cryptanalysis. The bound on the 8–round cipher appears low where 64–bit key security is required. However, it should be noted that it would be a very difficult task for a cryptanalyst to find a linear approximation close to the lower bound. Certainly for systems with security requirements on the order of 40 key bits or less, an 8–round cipher would be adequately secure against linear cryptanalysis.

Number of Rounds, R	Required Known Plaintexts, N_L	
	CAST	DES
8	2^{34}	2^{22}
12	2^{50}	2^{34}
16	2^{66}	2^{47}

Table 1. Number of Required Plaintexts in Linear Attack ($\mathcal{N}_{min} = 64$)

III. Selection of Highly Nonlinear S-boxes

In this section we discuss the likelihood that a CAST cipher designer can find S-boxes which satisfy a suitable level of nonlinearity. The design procedure outlined in [2] for constructing S-boxes involves the use of maximally nonlinear “bent” functions for the n output functions of the S-boxes to ensure high nonlinearity for the individual output functions. It is further recommended that the designer verify the nonlinearity of the S-box by checking all linear combinations of output functions.

In the following analysis we shall examine the likelihood of an $m \times n$ S-box being selected which will display good nonlinearity of $\mathcal{N}(S) \geq 2^{m-2}$. Specifically, for a 64–bit cipher, from the development of the previous section, we are interested in the generation of 8×32 S-boxes which have $\mathcal{N}(S) \geq 64$. To make the analysis tractable we shall assume that the 2^n functions generated from considering the linear combinations of the n S-box outputs are all randomly generated and independent. Hence, we ignore the constraints of using bent functions and that the 2^n functions are derived from a set of n functions.[3]

Given a function g, consider the probability of randomly selecting f so that $d(f,g) < 2^{m-2}$. The probability of selecting f so that $d(f,g) = 2^{m-2}$ is given by

$$P\big(d = 2^{m-2}\big) = \binom{2^m}{2^{m-2}} / 2^{2^m}. \qquad (12)$$

[3] The independence assumption cannot be strictly correct. Consider, for example, that the linear combination of two linear functions is itself linear. However, it is extremely unlikely that any of the S-box functions are linear and we shall therefore assume that all functions (and their linear combinations) are selected from a set of randomly generated functions.

Using Stirling's approximation of $k! \approx \sqrt{2\pi k}(k/e)^k$, (12) can be approximated by

$$P\left(d = 2^{m-2}\right) \approx \left[\pi\left(2^{m-1} - 2^{m-3}\right)\right]^{-1/2} \cdot \frac{2^{\left(3m2^{m-2} - 2^{m-1}\right)}}{\left(2^m - 2^{m-2}\right)^{\left(2^m - 2^{m-2}\right)}}. \qquad (13)$$

It can be shown that, for reasonable values of m, the probability of $d(f,g) < 2^{m-2}$ is bounded as

$$P\left(d < 2^{m-2}\right) < P\left(d = 2^{m-2}\right). \qquad (14)$$

Since any two linear m-bit functions are exactly a distance of 2^{m-1} apart, the events that a function f is less than a distance of 2^{m-2} from two different linear functions are mutually exclusive. Hence,

$$P\left(\mathcal{N}(f) < 2^{m-2}\right) < \rho \qquad (15)$$

where $\rho = 2^{m+1} \cdot P\left(d = 2^{m-2}\right)$ and we have used the fact that there are 2^{m+1} m-bit affine functions.

Finally, in our analysis we use the assumption that the 2^n functions determined from all linear combinations of the n output functions of an S-box may be considered independently in an analysis of their nonlinearities and that the probability distribution of the nonlinearity of each function is the same as that of a randomly generated function. As a result, the probability that an S-box has a nonlinearity that is less than or equal to 2^{m-2} is given by

$$P\left(\mathcal{N}(S) < 2^{m-2}\right) < 1 - (1-\rho)^{2^n}. \qquad (16)$$

Assuming that $2^n \cdot \rho << 1$, this may be approximated by

$$P\left(\mathcal{N}(S) < 2^{m-2}\right) \lesssim 2^n \cdot \rho. \qquad (17)$$

Consider the CAST algorithm with $m = 8$ and $n = 32$. In this case, from (13), $P(d = 64) \approx 2^{-52}$. Therefore, $\rho \approx 2^{-43}$, resulting in $P(\mathcal{N}(S) < 64) \lesssim 2^{-11}$. Hence, we expect at least 99.95% of all randomly generated 8×32 S-boxes to have nonlinearities of at least 64. This implies that most S-boxes have suitable nonlinearities and, therefore, selecting candidate S-boxes and then screening those candidates to eliminate any with low nonlinearities as suggested in [2] is a reasonable S-box design procedure.

IV. Summary

Our analysis of the CAST encryption algorithm has determined a theoretical bound on the security against linear cryptanalysis given the minimum nonlinearity of the S-boxes used. The results suggest that a 64–bit CAST cipher with a 64–bit key based on 8×32 S-boxes is secure against linear cryptanalysis with a moderate number of rounds. Further analysis suggests that sufficiently nonlinear S-boxes are easy to find by simple random generation. We conclude that construction of efficient block ciphers resistant to linear cryptanalysis is straightforward using the CAST encryption algorithm design procedures.

References

[1] C. M. Adams, *A Formal and Practical Design Procedure for Substitution-Permutation Network Cryptosystems*. PhD thesis, Queen's University at Kingston, Canada, 1990.

[2] C. M. Adams and S. E. Tavares, "Designing S-boxes resistant to differential cryptanalysis," *Proceedings of 3rd Symposium on the State and Progress of Research in Cryptography,* Rome, Italy, pp. 181–190, Feb. 1993.

[3] B. O'Higgins, "BNR leads industry in client/server network security," *Telesis*, pp. 79–80, May 1994.

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

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

[6] "National Bureau of Standards - Data Encryption Standard," *Federal Information Processing Standard Publication 46*, 1977.

[7] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[8] C. M. Adams, "Simple and effective key scheduling for symmetric ciphers," *presented at Workshop on Selected Areas in Cryptography (SAC '94),* Queen's University, Kingston, Canada, May 1994.

[9] H. M. Heys and S. E. Tavares, "The design of product ciphers resistant to differential and linear cryptanalysis," *presented at CRYPTO '93,* Santa Barbara, Calif., Aug. 1993.