A NEW CRITERION FOR THE DESIGN OF 8×8 S-BOXES IN PRIVATE-KEY CIPHERS

Jianhong Xu and Howard M. Heys

Electrical Engineering
Memorial University of Newfoundland
St. John's, Newfoundland, Canada A1B 3X5

ABSTRACT

In this paper, we examine the security of the class of substitution-permutation private-key block ciphers with respect to linear and differential cryptanalysis. A new S-box nonlinearity criterion is proposed and it is shown that S-boxes satisfying this criterion and having good diffusion improve remarkably the ability of an SPN to resist linear cryptanalysis and differential cryptanalysis.

1. INTRODUCTION

A basic substitution-permutation encryption network (SPN) consisting of a number of rounds of substitutions (S-boxes) connected by bit permutations is an implementation of a private-key block cipher [1]. The SPN structure is directly based on the concepts of "confusion" and "diffusion" introduced by Shannon [2]. Letting N represent the block size of a basic SPN composed of R rounds of $n \times n$ S-boxes, a simple example of an SPN with $N = 16$, $n = 4$, and $R = 3$ is illustrated in Figure 1. Keying the network can be realized by XORing the key bits with the data bits before each round of substitution and after the last round. In this paper, we consider specifically 64-bit ciphers based on 8×8 S-boxes and which use a permutation that has the i-th output bit of the j-th S-box connected to the j-th input bit of the i-th S-box.

Linear cryptanalysis, suggested by Matsui [3], is a known plaintext attack which uses knowledge of plaintext-ciphertext pairs to break the cipher. Differential cryptanalysis, as introduced by Biham and Shamir [4], is a chosen plaintext attack which examines the changes in the ciphertext in response to controlled changes in the plaintext input. To attack an SPN, linear cryptanalysis makes use of a highly probable linear approximation, while differential cryptanalysis is implemented by finding a highly probable differential characteristic. Both a highly probable linear expression and a highly probable characteristic are achieved by exploiting the local properties of the network, specifically the S-box properties. Thus the design of S-boxes is crucial to the strength of an SPN.

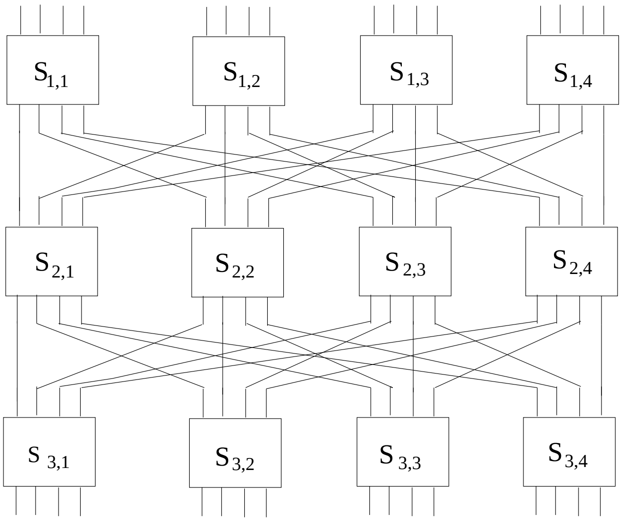

Figure 1: SPN with $N = 16$, $n = 4$, and $R = 3$

In this work, a new S-box nonlinearity criterion is proposed. S-boxes satisfying this criterion and having good diffusion [5] improve the ability of an SPN to resist linear cryptanalysis and differential cryptanalysis noticeably.

2. BACKGROUND

In the application of linear cryptanalysis to SPNs, the best R-round linear approximation of the form

$$P_{i_1} \oplus \cdots \oplus P_{i_\alpha} \oplus C_{j_1} \oplus \cdots \oplus C_{j_\beta} = K_{k_1} \oplus \cdots \oplus K_{k_\gamma} \quad (1)$$

is of interest, where P_i, C_j, and K_k represent a plaintext, ciphertext, and key bit, respectively.

This linear approximation is derived by combining a number of probable linear expressions of S-boxes from different rounds such that any intermediate terms (i.e., terms that are not plaintext, ciphertext, or key terms) are eliminated. As we will see later, the number of known plaintexts required to attack an SPN using basic linear attack is related to p, the probability of an S-box linear approximation. The probability p that a linear expression holds for an S-box is defined as $p = NS(\alpha,\beta)/2^n$, where $NS(\alpha,\beta)$ is defined as follows.

Definition 1 [3]: For a given $n \times n$ S-box, S, $NS(\alpha,\beta)$ is defined as the number of inputs to S, where a mod-2 linear combination of the input bits specified by vector α is equal to a mod-2 linear combination of output bits specified by vector β. In particular,

$$NS(\alpha,\beta) = \#\{x \in \{0,1\}^n | L(x, s(x)) \text{ holds}\} \quad (2)$$

where $s(x)$ is the output of the S-box corresponding to input x, and $L(x, s(x))$ is the following linear expression:

$$\bigoplus_{i=0}^{n-1} \alpha[i] \cdot x[i] = \bigoplus_{i=0}^{n-1} \beta[i] \cdot s(x) \quad (3)$$

where the symbol $x[i]$ represents the i-th bit of vector x. When $|p - 1/2|$ is small for all α and β, the nonlinearity of the S-box is said to be high.

Differential cryptanalysis is dependent on the existence of a highly probable $(R-1)$-round characteristic. The existence of a highly probable $(R-1)$-round characteristic is determined by two factors [5]: (1) the distribution of S-box XOR difference pairs $(\Delta x, \Delta y)$, where Δx is input XOR difference of 2 input vectors, x_1 and x_2, (i.e., $\Delta x = x_1 \oplus x_2$), and Δy is output XOR difference of an S-box, (i.e., $\Delta y = s(x_1) \oplus s(x_2)$), and (2) the diffusion of bit changes within the network.

As shown in [5], for an SPN, S-boxes with good diffusion properties can increase the diffusion of bit changes within the network.

Definition 2 [5]: An S-box satisfies a diffusion order of λ, $\lambda \geq 0$, if, for $wt(\Delta x) > 0$,

$$wt(\Delta y) > \begin{cases} \lambda + 1 - wt(\Delta x) & , wt(\Delta x) < \lambda + 1 \\ 0 & , wt(\Delta x) \geq \lambda + 1 \end{cases} \quad (4)$$

where Δx and Δy denote the input XOR difference and the corresponding output XOR difference of an S-box respectively, and $wt(\cdot\ )$ refers to the *Hamming weight* of the specified argument.

3. S-BOX DESIGN CONSTRAINTS

In this section, constraints on S-boxes that effectively strengthen an SPN against linear cryptanalysis and differential cryptanalysis are proposed.

3.1. Nonlinearity Requirement

In the linear cryptanalysis of an SPN, a cryptanalyst is interested in finding a linear approximation which is deduced by combining a number of probable linear expressions of the involved S-boxes. Suppose there are δ S-boxes involved in the derivation of a linear approximation of the overall cipher, and the probable linear expression of the i-th S-box holds with probability p_i, then, assuming the S-box inputs are independent, according to Lemma 3 in [3] the linear approximation holds with probability

$$p_L = 1/2 + 2^{\delta-1} \prod_{i=1}^{\delta} (p_i - 1/2). \quad (5)$$

Also, it is shown in [3] that, for a basic linear attack (algorithm 1), the number of known plaintexts required to guess a correct key bit in equation (1) is approximated by N_L, where

$$N_L = |p_L - 1/2|^{-2}. \quad (6)$$

By rewriting expression (5) as

$$p_L = 1/2 + 1/2 \prod_{i=1}^{\delta} (2p_i - 1) \quad (7)$$

it is evident that p_L is determined by two factors:

1. δ, the number of S-boxes involved in the overall cipher linear approximation, and
2. p_i, the probability with which the linear expression of the i-th S-box is satisfied.

In [5], $|p_L - 1/2|$ was bounded by considering δ and $|p_i - 1/2|$ separately. The probability p_i of a linear expression of an S-box was bounded with

$$|p_i - 1/2| \leq |p_\varepsilon - 1/2| \tag{8}$$

where p_ε represents the probability of the best linear approximation of any S-box in the network and it can be determined by

$$|p_\varepsilon - 1/2| = (2^{n-1} - NL_{min})/2^n \tag{9}$$

where n is the size of an S-box and NL_{min} is the lowest nonlinearity of an S-box [5], i.e., for all S-boxes

$$|2^{n-1} - NS(\alpha, \beta)| \leq 2^{n-1} - NL_{min}. \tag{10}$$

Actually, after studying the structure of a linear path in an SPN, Theorem 1 is found and a new bound for p_i is then established.

Theorem 1 *Let β, $2 \leq \beta \leq 2n$, represent the number of input bits plus output bits in a linear approximation of an S-box. Assume all S-boxes use a β-term linear approximation of a specific β. Then the best cipher linear approximation must involve $\beta/2$ S-boxes on average per round.*

Proof: Consider an S-box in the r-th round. Since each input bit or output bit of an S-box connects to a different S-box in the previous or next round, based on the assumption, the number of involved S-boxes in the previous plus the number of involved S-boxes in the next round must be at least β.

Since we are interested in the best cipher linear approximation, if a scenario in which the number of involved S-boxes in the previous and next round is β exists, then the theorem is proven. It is trivial to show that such scenarios do exist. □

Now for a given value of β, bound the linearity property of an S-box by

$$\eta_i(\beta) \leq \eta(\beta) \tag{11}$$

where $\eta_i(\beta) = |p_i(\beta) - 1/2|$ with $p_i(\beta)$ representing the probability for a β-term linear approximation of the i-th S-box in the cipher approximation and $\eta(\beta) = |p_\varepsilon(\beta) - 1/2|$ with $p_\varepsilon(\beta)$ representing the probability of the best β-term linear approximation of any S-box in the network.

Subsequently, letting $\eta_L = |p_L - 1/2|$, based on Theorem 1, we have

$$\eta_L = 2^{\delta-1} \prod_{i=1}^{\delta} \eta_i(\beta) \leq 2^{\delta-1} \prod_{i=1}^{\delta} \eta(\beta). \tag{12}$$

Hence, since $\delta = \beta/2 \cdot R$ where R is the number of rounds in the SPN,

$$\eta_L \leq 1/2(2\eta(\beta))^{\beta/2 \cdot R}. \tag{13}$$

According to (13), to prevent a cryptanalyst from using some specific β-term linear expressions to obtain a linear approximation with a higher probability, a straightforward way is to establish the equation

$$(2\eta(2))^2 = (2\eta(\beta))^\beta, \tag{14}$$

for all values of $\beta > 2$. Constraint (14) can be used to minimize the upper bound on η_L by selecting S-boxes with a small 2-term linear approximation bound and weighting the effects of the linear approximations of different β to provide a uniform upper bound on η_L across all values of β.

However, we have observed, by experimentation, that the relation

$$2\eta_i(\beta) \leq \begin{cases} (2\eta(2))^{2/\beta} & ,2 \leq \beta \leq 3 \\ (2\eta(2))^{2/4} & ,4 \leq \beta \leq 2n \end{cases} \tag{15}$$

leads to a tighter bound on the probability of the best linear approximation, and is reasonable to adopt as the constraint put on an S-box.

In setting the bound for the probability of the best linear approximation, the number of involved linear expressions of S-boxes (i.e., the number of involved S-boxes) needs to be calculated based on the network structure (i.e. permutations), and this number can be related to an *equivalent* number of 2-term S-boxes involved in a linear approximation according to constraint (15).

Using this approach we have written a program to calculate the *equivalent* number of 2-term S-boxes involved in a linear approximation for all possible cases in which the number of S-boxes in each round is some value between 1 and 8. It is found that, for an 8-round SPN, with constraint (15) the smallest number of equivalent 2-term S-boxes involved in a linear approximation is 22/3. For example, one scenario is that from the 1st to 8-th round, the number of actual S-boxes is 1,1,1,2,2,1,1, and 1, respectively.

Let us calculate the number of known plaintexts required in the basic linear attack. As mentioned above, under condition (15) the equivalent number of 2-term linear expressions involved in the best linear approximation of an 8-round SPN is 22/3. Therefore, according to (13), for an 8-round SPN $|p_L - 1/2| \leq 1/2(2\eta(2))^{22/3}$. This signifies that in the basic linear attack the number of plaintexts required to deduce one equivalent bit of key is at least $4/(2\eta(2))^{44/3}$. From the results of our experiments, 8×8 S-boxes satisfying (15) with $2\eta(2) = 1/8$ can be achieved. Hence, if an 8-round SPN is constructed using 8×8 S-boxes satisfying (15) with $2\eta(2) = 1/8$, $|p_L - 1/2|$ is 2^{-23} and it requires at least 2^{46} known plaintexts to determine one key bit using the basic linear attack.

In contrast, in [5], (8) is used to bound the probability of a linear expression of an S-box, with the value of $2|p_\varepsilon - 1/2| = 1/4$. Since the minimum number of S-boxes involved in a linear approximation is 8, the resulting $|p_L - 1/2|$ is 2^{-17} as determined by (7) and the number of required plaintexts in a basic linear attack is at least 2^{34}.

3.2. Diffusion Order Requirement

S-boxes with a high diffusion order can enhance the ability of an SPN to resist differential cryptanalysis [2]. By using the *depth-first-search* algorithm in [5], S-boxes are examined for the relationship between their nonlinearity property and diffusion order. It is determined that S-boxes with diffusion order of 1 which satisfy the suggested nonlinearity requirement with a small value of $2\eta(2)$ are easily found. For example, the proportion of S-boxes satisfying (15) with $2\eta(2) = 1/8$ and selected from randomly generated S-boxes with diffusion order of 1 is 0.267.

4. CONCLUSION

In accordance with the new nonlinearity requirement suggested in this paper, S-boxes whose fewer-term linear approximations are highly nonlinear are found. These S-boxes can be selected from the S-boxes with diffusion order of 1. Thus the ability to resist linear cryptanalysis and differential cryptanalysis of an SPN that is constructed from these S-boxes is improved remarkably.

5. REFERENCES

[1] H. Feistel.“ Cryptography and computer privacy”, *Scientific American*, 228, pp. 15-23, 1973.

[2] C.E. Shannon. “Communication theory of secrecy systems”, *Bell systems Technical Journal*, vol.28, pp.656-715, 1949.

[3] M. Matsui. “Linear cryptanalysis method for DES cipher”, *Proceedings of Eurocrypt’93*, Springer -verlag, Berlin, pp. 386-397.

[4] E.Biham and A.Shamir.“Differential Cryptanalysis of DES-like Cryptosystems”,*Journal of Cryptology*, vol. 4, no. 1, pp. 3-72 ,1991.

[5] H. M. Heys and S. E. Tavares. “The Design of Substitution Permutation Network Ciphers Resistant to Cryptanlysis”, *Journal of Cryptology*, Vol.9, no.1, pp.1-19, 1996.