Practical Security Evaluation against Differential and Linear Cryptanalyses for Feistel Ciphers with SPN Round Function

Masayuki Kanda

NTT Information Sharing Platform Laboratories,
1-1-612A Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847, Japan
kanda@isl.ntt.co.jp

Abstract. This paper studies the upper bounds of the maximum differential and linear characteristic probabilities of Feistel ciphers with SPN round function. In the same way as for SPN ciphers, we consider the minimum number of differential and linear active s-boxes, which provides a measure of the upper bounds of these probabilities, in order to evaluate the security against differential and linear cryptanalyses. The purpose of this work is to clarify the (lower bound of) minimum numbers of active s-boxes in some consecutive rounds of Feistel ciphers, i.e., in three, four, six, eight, and twelve consecutive rounds, using differential and linear branch numbers $\mathcal{P}_d$, $\mathcal{P}_l$, respectively. Furthermore, we investigate the necessary condition for desirable P-functions, which means that the round functions are invulnerable to both differential and linear cryptanalyses. As an example, we show the round function of Camellia, which satisfies the condition.

1 Introduction and Motivation

The best known attacks are differential cryptanalysis [6] proposed by Biham and Shamir and linear cryptanalysis [13] proposed by Matsui. Since these cryptanalyses are the most powerful approaches known for attacking many symmetric block ciphers, designers should evaluate the security of any new proposed ciphers against differential and linear cryptanalyses. To do this it is necessary to determine the maximum differential and linear probabilities by a useful (and acceptable) method. Feistel ciphers are commonly analyzed by (a) the upper bounds of the maximum average of differential and linear hull probabilities or (b) the maximum differential and linear characteristic probabilities. SPN ciphers, on the other hand, are commonly analyzed by (c) the upper bounds of the maximum differential and linear characteristic probabilities. Recently, Hong et al. showed (a) the upper bounds of the maximum average of differential and linear hull probabilities of SPN ciphers [9].

With reference to method (a), Nyberg and Knudsen showed that the maximum average of differential and linear hull probabilities for r-round ($r \geq 4$) Feistel ciphers are bounded by $2p^2$, $2q^2$ if the maximum differential and linear

D.R. Stinson and S. Tavares (Eds.): SAC 2000, LNCS 2012, pp. 324–338, 2001.

probabilities of the round function are p, q, respectively[1] [18]. They stated that Feistel ciphers are *provably secure* against differential and linear cryptanalyses if these probabilities are sufficiently low. This means that they are theoretically invulnerable to differential and linear cryptanalyses, since these probabilities are the upper bounds of the average of differential and linear hull probabilities. However, this approach has one fatal disadvantage. That is, these probabilities settle at some constant value even if the number of rounds increases. Therefore, a round function has to yield extremely low maximum differential and linear probabilities. This imposes a hard restriction on designing the round function. As a matter of fact, for a commercial cipher, MISTY [15] is provably secure with respect to differential and linear cryptanalyses.

Method (b) has been used to estimate many (extended) Feistel ciphers such as DES [6,13] and FEAL [16,2]. Biham and Shamir claimed that the higher the differential characteristic probability is, the higher the success rate of differential cryptanalysis is. This is because they exploited a single path between plaintexts and ciphertexts which holds significant differential characteristic probability. Matsui also claimed the same for linear cryptanalysis. Thus, Feistel ciphers are *sufficiently secure* against differential and linear cryptanalyses if these probabilities are less than the security threshold. Strictly speaking, however, these probabilities only give the lower bounds of the maximum average of differential and linear hull probabilities, since this method does not consider multiple paths between the same plaintexts and ciphertexts [12,17].

For SPN ciphers, Rijmen et al. introduced the branch number $\mathcal{B}$ [19]. The number $\mathcal{B}$ is the minimum number of active s-boxes in two consecutive rounds of a non-trivial differential characteristic or a non-trivial linear trail. Since each active s-box reduces the differential and linear characteristic probabilities, the number $\mathcal{B}$ provides the upper bounds of the maximum differential and linear characteristic probabilities in two consecutive rounds. The security against differential and linear cryptanalyses is evaluated by piling up the number $\mathcal{B}$ every two rounds. It is noted that Knudsen proposed a very similar concept for Feistel ciphers [10]. He noted that Feistel ciphers are *practically secure* against differential and linear cryptanalyses if the upper bounds of the maximum differential and linear characteristic probabilities are less than the security threshold.

It is obvious that the upper bounds of the maximum differential and linear characteristic probabilities by method (c) lie between the upper bounds of the maximum average of differential and linear hull probabilities by method (a) and the maximum differential and linear characteristic probabilities by method (b). Moreover, for most ciphers, the maximum averages of differential and linear hull probabilities, which provide the actual invulnerability to differential and linear cryptanalyses, are much lower than the upper bounds of these probabilities if the number of rounds increases. Therefore, it is worth investigating the upper bounds of the maximum differential and linear characteristic probabilities.

[1] Aoki and Ohta showed that these probabilities are bounded by p^2, q^2 if the round function is bijective and $r \geq 3$ [3]

Knudsen discussed the upper bounds of the maximum differential and linear characteristic probabilities of general Feistel ciphers [10]. He showed that the upper bounds of these probabilities for $2r$-round Feistel ciphers are p^r, q^r if p, q are the maximum differential and linear probabilities of the round function, respectively[2]. His evaluation, unfortunately, did not take the interrelation between input and output data in consecutive rounds into consideration. That is, it is not always useful to evaluate the upper bounds of the maximum differential and linear characteristic probabilities, if the maximum differential and linear probabilities of the round function p, q are relatively high while those of some consecutive rounds are (sufficiently) low, such as DES [7].

On the other hand, in this paper, we would like to focus attention on the upper bounds of the maximum differential and linear characteristic probabilities for Feistel ciphers with SPN round function. Like SPN ciphers, Feistel ciphers with SPN round function only consist of s-boxes and bitwise exclusive-ORs. This means that the (lower bound of) minimum number of active s-boxes determines the upper bounds of the maximum differential and linear characteristic probabilities for not only SPN ciphers but also Feistel ciphers with SPN round function. This evaluation takes the interrelation between input and output data in some consecutive rounds into consideration, while Knudsen's evaluation doesn't. Accordingly, our motivation is to clarify the (lower bound of) minimum number of active s-boxes in some consecutive rounds of Feistel ciphers.

This paper is organized as follows. Section 2 introduces some notations and definitions. Previous works are shown in Sect. 3. In Sect. 4 and Sect. 5, the lower bounds of the minimum number of active s-boxes for differential and linear cryptanalyses are given, respectively, i.e., the upper bounds of the maximum differential and linear characteristic probabilities. The necessary condition for desirable P-functions is discussed in Sect. 6. Finally, we conclude in Sect. 7.

2 Preliminaries

2.1 Notations

$X = (x_1, \ldots, x_n)$, $x_i \in \mathbb{Z}_2^m$ $(1 \leq i \leq n)$:
vector X over $\mathrm{GF}(2^m)^n$ and element x_i of X over $\mathrm{GF}(2^m)$.
ΔX, ΓY : difference of X and mask value of Y, respectively.
$X \cdot \Gamma X$: parity of bitwise product X and ΓX.
$X \oplus Y$: bitwise exclusive-OR (XOR).
$X|Y$: concatenation between X and Y.
$\{S\}$, $\#\{S\}$: elements in set S and the number of elements in set S.

2.2 Model

Throughout this paper we consider Feistel ciphers with *mn-bit SPN round function* (See Fig. 1). Note that we neglect the effect of the round key hereafter

[2] Kanda et al. showed that the upper bounds of these probabilities for $3r$-round Feistel ciphers are p^{2r}, q^{2r} if the round function is bijective [11]

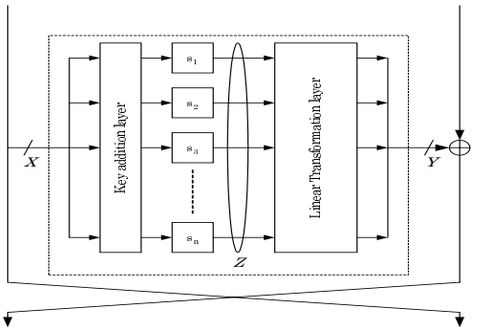

Fig. 1. SPN round function

since we assume that the round key, which is used within one round, consists of independent and uniformly random bits, and is bitwise XORed with data.

Notations describe the model as below.

S-function is a non-linear transformation layer with n parallel m-bit bijective s-boxes. That is,

$$\begin{aligned} S : (\mathbb{Z}_2^m)^n &\longrightarrow (\mathbb{Z}_2^m)^n \\ X = (x_1, \ldots, x_n) &\longmapsto Z = S(X) = (s_1(x_1), \ldots, s_n(x_n)) \end{aligned}$$

P-function is a linear transformation layer, i.e.,

$$\begin{aligned} P : (\mathbb{Z}_2^m)^n &\longrightarrow (\mathbb{Z}_2^m)^n \\ Z = (z_1, \ldots, z_n) &\longmapsto Y = P(Z) = (y_1, \ldots, y_n) \end{aligned}$$

Finally, the SPN round function can be described as follows.

$$\begin{aligned} F : (\mathbb{Z}_2^m)^n &\longrightarrow (\mathbb{Z}_2^m)^n \\ X = (x_1, \ldots, x_n) &\longmapsto Y = F(X) = P(S(X)) = (y_1, \ldots, y_n) \end{aligned}$$

Let $X^{(i)}$ be the input data to the i-th round function, and $Y^{(i)}$ be the i-round output data. The Feistel cipher is defined as:

$$X^{(i+1)} = X^{(i-1)} \oplus Y^{(i)} \; (1 \leq i \leq r),$$

where $(X^{(1)}|X^{(0)})$ is a plaintext and $(X^{(r)}|X^{(r+1)})$ is a ciphertext.

2.3 Definitions

We use the following definitions in this paper.

Definition 1. *For any given Δx, Δz, Γx, $\Gamma z \in \mathbb{Z}_2^m$, the differential and linear probabilities of each s-box s_i are defined as:*

$$DP^{s_i}(\Delta x \to \Delta z) = \frac{\#\{x \in \mathbb{Z}_2^m | s_i(x) \oplus s_i(x \oplus \Delta x) = \Delta z\}}{2^m}$$

$$LP^{s_i}(\Gamma z \to \Gamma x) = \left(2 \times \frac{\#\{x \in \mathbb{Z}_2^m | x \cdot \Gamma x = s_i(x) \cdot \Gamma z\}}{2^m} - 1\right)^2$$

Definition 2. *The maximum differential and linear probabilities of s-boxes are defined as:*

$$p_s = \max_i \max_{\Delta x \neq 0, \Delta z} DP^{s_i}(\Delta x \to \Delta z)$$
$$q_s = \max_i \max_{\Gamma x, \Gamma z \neq 0} LP^{s_i}(\Gamma z \to \Gamma x)$$

This means that p_s, q_s are the upper bounds of the maximum differential and linear probabilities for all s-boxes.

Definition 3. *A differential active s-box is defined as an s-box given a non-zero input difference, while a linear active s-box is defined as an s-box given a non-zero output mask value [11].*
Note: When an s-box is bijective, s-boxes given a non-zero output difference and a non-zero input mask value are also differential and linear active s-boxes, respectively.

Definition 4. *Let $X = (x_1, \ldots, x_n) \in \mathrm{GF}(2^m)^n$ then the Hamming weight of X is denoted by*

$$H_w(X) = \#\{i | x_i \neq 0\}.$$

This means that the Hamming weight of X equals the number of non-zero m-bit characters from $GF(2^m)$ of X.

3 Previous Works – the Security of SPN Ciphers

As mentioned above, the security of most SPN ciphers against differential and linear cryptanalyses is evaluated using the (lower bound of) minimum number of differential and linear active s-boxes, which are a measure of the upper bounds of differential and linear characteristic probabilities [19,8,5]. To determine the (lower bound of) minimum number of active s-boxes, Rijmen et al. defined the branch number $\mathcal{B}$ [19].

Definition 5. *In SPN ciphers, the differential branch number $\mathcal{B}_d$ is defined as:*

$$\mathcal{B}_d = \min_{\Delta X \neq 0} (H_w(\Delta X) + H_w(\theta(\Delta X))),$$

where ΔX is an input difference into the diffusion layer and $\theta(\Delta X)$ is an output difference from the layer.

Note that ΔX is also an output difference from a substitution layer and $\theta(\Delta X)$ is also an input difference to the next substitution layer. Since s-boxes are bijective, $H_w(\Delta X)$ equals the number of differential active s-boxes in the substitution layer and $H_w(\theta(\Delta X))$ equals that in the next substitution layer. That is, if n_d is the minimum number of differential active s-boxes in two consecutive rounds, then $n_d = \mathcal{B}_d$. Thus, it turns out that the minimum number of differential active s-boxes in $2r$-round SPN ciphers is lower bounded by $r\mathcal{B}_d$, and the following theorem is obtained.

Theorem 1. *The maximum differential characteristic probability for $2r$-round SPN cipher, $p_d^{(2r)}$, is upper bounded by $p_s^{(r\mathcal{B}_d)}$.*

From the duality between differential characteristics and linear approximations [4,14], the following definition and theorem also are established.

Definition 6. *The linear branch number $\mathcal{B}_l$ is defined as:*

$$\mathcal{B}_l = \min_{\Gamma Y \neq 0} (H_w(\theta^*(\Gamma Y)) + H_w(\Gamma Y)),$$

where ΓY is an output mask value of the diffusion layer θ and $\theta^*(\Gamma Y)$ is an input mask value of the layer. θ^* is the diffusion function of mask values concerning the layer.

Theorem 2. *The maximum linear characteristic probability for $2r$-round SPN cipher, $q_l^{(2r)}$, is upper bounded by $q_s^{(r\mathcal{B}_l)}$.*

4 Upper Bound of Differential Characteristic Probability

In this section, we investigate the upper bound of differential characteristic probability of Feistel cipher with SPN round function. In the same way as in the previous section, our goal is to clarify the (lower bound of) minimum number of differential active s-boxes in some consecutive rounds of Feistel cipher.

First, we show the useful lemma concerning the hamming weight for Feistel ciphers.

Lemma 1. *In Feistel ciphers, the following relationship holds.*

$$H_w(\Delta Y^{(i)}) = H_w(\Delta X^{(i-1)} \oplus \Delta X^{(i+1)}) \leq H_w(\Delta X^{(i-1)}) + H_w(\Delta X^{(i+1)})$$

Proof.

$$\begin{aligned} H_w(\Delta Y^{(i)}) &= H_w(\Delta X^{(i-1)} \oplus \Delta X^{(i+1)}) \\ &= \#\{s|\Delta x_s^{(i-1)} \neq 0 \text{ and } \Delta x_s^{(i+1)} = 0\} \\ &\quad +\#\{t|\Delta x_t^{(i-1)} = 0 \text{ and } \Delta x_t^{(i+1)} \neq 0\} \\ &\quad +\#\{u|\Delta x_u^{(i-1)} \neq 0 \text{ and } \Delta x_u^{(i+1)} \neq 0 \text{ and } x_u^{(i-1)} \neq x_u^{(i+1)}\} \\ &\leq H_w(\Delta X^{(i-1)}) + \#\{t|\Delta x_t^{(i-1)} = 0 \text{ and } \Delta x_t^{(i+1)} \neq 0\} \\ &\leq H_w(\Delta X^{(i-1)}) + H_w(\Delta X^{(i+1)}) \end{aligned}$$

Q.E.D.

Since there is a linear transformation layer (P-function) in the SPN round function, we will define the differential branch number $\mathcal{P}_d$ in the same way as in the previous section. Note that it is obvious that if S-function is bijective then $H_w(\Delta X) = H_w(\Delta Z)$, since Δz_i also becomes a non-zero output difference through the differential active s_i-box.

Definition 7. *If S-function is bijective, the differential branch number $\mathcal{P}_d$ is defined as follows.*

$$\mathcal{P}_d = \min_{\Delta X \neq 0} (H_w(\Delta X) + H_w(\Delta Y))$$

Here, we will define the upper bound of the maximum differential characteristic probability of Feistel cipher with SPN round function in the same way as used for the SPN cipher. That is, the upper bound of the probability is shown by the (lower bound of) minimum number of differential active s-boxes.

Definition 8. *Assume Feistel cipher with SPN round function. Let $H_w(\Delta X^{(i)})$ be the number of the ith-round differential active s-boxes, then the differential characteristic probability of the r-round Feistel cipher, $p_d^{(r)}$, satisfies the following relationship.*

$$p_d^{(r)} \leq p_s^{\min_{(\Delta X^{(0)}, \Delta X^{(1)}, \ldots, \Delta X^{(r+1)}) \neq (0,0,\ldots)} \sum_{i=1}^{r} H_w(\Delta X^{(i)})}$$

From this definition, clarifying the upper bound of the maximum differential characteristic probability becomes equivalent to showing the (lower bound of) minimum number of differential active s-boxes. To discuss the minimum number easily after this, it is denoted as follows.

$$\mathcal{D}^{(r)} = \min_{(\Delta X^{(0)}, \Delta X^{(1)}, \ldots, \Delta X^{(r+1)}) \neq (0,0,\ldots)} \sum_{i=1}^{r} H_w(\Delta X^{(i)})$$

Hereafter, because of limitations of space, we assume P-function is bijective. Note that this leads to $\mathcal{P}_d \geq 2$.

Lemma 2. *The minimum number of differential active s-boxes in any three consecutive rounds satisfies $\mathcal{D}^{(3)} \geq 2$.*

Proof. If $\Delta X^{(i)} = 0$, then $\Delta Y^{(i)} = 0$ and $\Delta X^{(i-1)} = \Delta X^{(i+1)} \neq 0$. This leads to $\mathcal{D}_1^{(3)} = 2 \times H_w(\Delta X^{(i-1)}) \geq 2$. On the other hand, If $\Delta X^{(i)} \neq 0$, it follows that $\mathcal{D}_2^{(3)} \geq H_w(\Delta X^{(i)}) + H_w(\Delta Y^{(i)}) \geq \mathcal{P}_d$, since Lemma 1 shows $H_w(\Delta X^{(i-1)}) + H_w(\Delta X^{(i+1)}) \geq H_w(\Delta Y^{(i)})$.

Q.E.D.

Lemma 3. *The minimum number of differential active s-boxes in any four consecutive rounds satisfies $\mathcal{D}^{(4)} \geq \mathcal{P}_d$.*

Proof. Without loss of generality, we assume that the four consecutive rounds run from the first round to the fourth round.

At no time do both input differences into any consecutive two rounds equal zero. In addition, by the assumption, at no time also do both input differences of every two rounds equal zero. Thus we only consider the six following cases concerning input differences into the consecutive four rounds.

(1) $\Delta X^{(1)} \neq 0$, $\Delta X^{(2)} \neq 0$, $\Delta X^{(3)} \neq 0$, $\Delta X^{(4)} \neq 0$

(2) $\Delta X^{(1)} = 0$, $\Delta X^{(2)} \neq 0$, $\Delta X^{(3)} \neq 0$, $\Delta X^{(4)} \neq 0$
(3) $\Delta X^{(1)} \neq 0$, $\Delta X^{(2)} = 0$, $\Delta X^{(3)} \neq 0$, $\Delta X^{(4)} \neq 0$
(4) $\Delta X^{(1)} \neq 0$, $\Delta X^{(2)} \neq 0$, $\Delta X^{(3)} = 0$, $\Delta X^{(4)} \neq 0$
(5) $\Delta X^{(1)} \neq 0$, $\Delta X^{(2)} \neq 0$, $\Delta X^{(3)} \neq 0$, $\Delta X^{(4)} = 0$
(6) $\Delta X^{(1)} = 0$, $\Delta X^{(2)} \neq 0$, $\Delta X^{(3)} \neq 0$, $\Delta X^{(4)} = 0$

In case (1), by Lemma 2, $\mathcal{D}_1^{(4)} = \mathcal{D}_2^{(3)} + H_w(\Delta X^{(4)}) \geq \mathcal{P}_d + H_w(\Delta X^{(4)}) \geq \mathcal{P}_d + 1$.

In case (2), $\Delta X^{(1)} = 0$ leads to $\Delta Y^{(2)} = \Delta X^{(3)}$. Thus, $\mathcal{D}_2^{(4)} = H_w(\Delta X^{(2)}) + H_w(\Delta Y^{(2)}) + H_w(\Delta X^{(4)}) \geq \mathcal{P}_d + H_w(\Delta X^{(4)}) \geq \mathcal{P}_d + 1$.

Similarly, in cases (3), (4), and (5), we get $\mathcal{D}^{(4)} \geq \mathcal{P}_d + 1$.

In case (6), by Lemma 2, $\mathcal{D}_6^{(4)} = \mathcal{D}_2^{(3)} \geq \mathcal{P}_d$.

Q.E.D.

From the above proof, the following corollary is obtained.

Corollary 1. *The minimum number of differential active s-boxes in any four consecutive rounds satisfies*
(i) $\mathcal{D}^{(4)} \geq \mathcal{P}_d$*, if and only if the input differences in both the first round and the fourth round are zero.*
(ii) $\mathcal{D}^{(4)} \geq \mathcal{P}_d + 1$ *in the other cases.*

Lemma 4. *The minimum number of differential active s-boxes in any six consecutive rounds satisfies* $\mathcal{D}^{(6)} \geq \mathcal{P}_d + 2$.

Proof.
- If $\Delta X^{(2)} \neq 0$ and $\Delta X^{(5)} \neq 0$, by Lemma 2, $\mathcal{D}_1^{(6)} = \mathcal{D}_2^{(3)} + \mathcal{D}_2^{(3)} \geq 2 \times \mathcal{P}_d$.
- If $\Delta X^{(2)} = \Delta X^{(5)} = 0$, we get $\Delta X^{(1)} = \Delta X^{(3)}$ and $\Delta Y^{(3)} = \Delta X^{(4)} = \Delta X^{(6)}$. Thus, $\mathcal{D}_2^{(6)} = 2 \times (H_w(\Delta X^{(3)}) + H_w(\Delta X^{(4)})) = 2 \times (H_w(\Delta X^{(3)}) + H_w(\Delta Y^{(3)})) \geq 2 \times \mathcal{P}_d$
- If $\Delta X^{(2)} = 0$ and $\Delta X^{(5)} \neq 0$, or $\Delta X^{(2)} \neq 0$ and $\Delta X^{(5)} = 0$, then $\mathcal{D}_3^{(6)} = \mathcal{D}_1^{(3)} + \mathcal{D}_2^{(3)} \geq \mathcal{P}_d + 2$ by Lemma 2.

Q.E.D.

Lemma 5. *The minimum number of differential active s-boxes in any eight consecutive rounds satisfies* $\mathcal{D}^{(8)} \geq 2 \times \mathcal{P}_d + 1$.

Proof. Again, corollary 1 shows that, in any four consecutive rounds, the minimum number of differential active s-boxes satisfies (i) $\mathcal{D}^{(4)} \geq \mathcal{P}_d$, if and only if the input differences in both the first round and the fourth round are zero, and $\mathcal{D}^{(4)} \geq \mathcal{P}_d + 1$ in the other cases.

Since there is no case in which both input differences into any two consecutive rounds are zero at the same time, the input differences in both the fourth and fifth rounds cannot be zero. That is, the eight consecutive rounds cannot be divided into two cases (i). Thus, $\mathcal{D}^{(8)} \geq \mathcal{P}_d + (\mathcal{P}_d + 1) \geq 2 \times \mathcal{P}_d + 1$.

Q.E.D.

Lemma 6. *The minimum number of differential active s-boxes in any twelve consecutive rounds satisfies* $\mathcal{D}^{(12)} \geq 3 \times \mathcal{P}_d + 1$.

Proof. $\mathcal{D}^{(12)}$ can be converted to three expressions, i.e., $4 \times \mathcal{D}^{(3)}$, $2 \times \mathcal{D}^{(6)}$, and $\mathcal{D}^{(8)} + \mathcal{D}^{(4)}$. Since $\mathcal{D}^{(12)}$ satisfies the three evaluations at the same time, $\mathcal{D}^{(12)} = \max\{4 \times \mathcal{D}^{(3)}, 2 \times \mathcal{D}^{(6)}, \mathcal{D}^{(8)} + \mathcal{D}^{(4)}\} \geq \mathcal{D}^{(8)} + \mathcal{D}^{(4)} \geq 3 \times \mathcal{P}_d + 1$.

Q.E.D.

From the proofs of above-mentioned lemmas, the useful theorem for the $4r$-round Feistel ciphers is established as follows.

Theorem 3. *The minimum number of differential active s-boxes* $\mathcal{D}^{(4r)}$ *for* $4r$*-round Feistel ciphers with SPN round function satisfies* $\mathcal{D}^{(4r)} \geq r \times \mathcal{P}_d + \lfloor r/2 \rfloor$.

Knudsen argued that for a Feistel cipher to be practically secure against differential and linear cryptanalyses, the upper bounds of the maximum differential and linear characteristic probabilities must be less than the security threshold. Generally speaking, the security threshold is equated to the inverse of the number of all plaintext blocks, i.e., 2^{-64} for 64-bit ciphers and 2^{-128} for 128-bit ciphers.

For example, let the maximum differential probability of an 8-bit s-box be $p_s = 2^{-6}$ and the differential branch number be $\mathcal{P}_d = 5$. It follows that 18-round Feistel ciphers, such as Camellia [1], are practically secure against differential cryptanalysis because of the following corollary.

Corollary 2. *Assuming that the round function consists of s-boxes yielding the maximum differential probability* $p_s = 2^{-6}$ *and P-function yielding the differential branch number* $\mathcal{P}_d = 5$*, then a 128-bit Feistel cipher with more than 16-rounds has no effective differential characteristic.*

Proof. By Definition 8 and Theorem 3, $p_d^{(16)} \leq (2^{-6})^{4\times5+2} = 2^{-132} < 2^{-128}$.

Q.E.D.

5 Upper Bound of Linear Characteristic Probability

In this section, the upper bound of linear characteristic probability is derived in the same way as in the previous section. That is, our goal is to clarify the (lower bound of) minimum number of linear active s-boxes in some consecutive rounds of Feistel cipher using the duality of differential characteristic and linear approximation.

First, the following theorem is established.

Theorem 4. *Consider a Feistel cipher with SPN round function. If the linear transformation layer P (P-function) is bijective, the cipher can be transformed into a Feistel cipher with the PSN round function.*

Proof. From the assumption that P-function is bijective, let describe $P(Z)$ as the transformation of Z by the P-function, and $P^{-1}(Z)$ as that by the inverse function of the P-function.

As mentioned above, in a Feistel cipher with SPN round function, the equation, $X^{(i+1)} = X^{(i-1)} \oplus P(S(X^{(i)}))$, is satisfied. Now, let $V^{(i)} = P^{-1}(X^{(i)})$. The above equation can be transformed as follows, since $C = A \oplus P(B) \Leftrightarrow C = P(P^{-1}(A) \oplus B)$ for any (A, B, C).

$$\begin{aligned} X^{(i+1)} = X^{(i-1)} \oplus P(S(X^{(i)})) &\Leftrightarrow X^{(i+1)} = P(P^{-1}(X^{(i-1)}) \oplus S(X^{(i)})) \\ &\Leftrightarrow P(V^{(i+1)}) = P(V^{(i-1)} \oplus S(P(V^{(i)}))) \\ &\Leftrightarrow V^{(i+1)} = V^{(i-1)} \oplus S(P(V^{(i)})) \end{aligned}$$

The equation, $V^{(i+1)} = V^{(i-1)} \oplus S(P(V^{(i)}))$, denotes a Feistel cipher with the PSN round function. Accordingly, the ciphertext $(X^{(r)}, X^{(r+1)})$ obtained by applying a Feistel cipher with SPN round function to a plaintext $(X^{(1)}, X^{(0)})$ is equivalent to the result of changing the plaintext $(X^{(1)}, X^{(0)})$ to $(V^{(1)}, V^{(0)})$ by the P^{-1}-function first, then getting $(V^{(r)}, V^{(r+1)})$ from $(V^{(1)}, V^{(0)})$ from the Feistel cipher with PSN round function, and finally transforming it into the ciphertext $(X^{(r)}, X^{(r+1)})$ by the P-function.

Q.E.D.

Starting with the duality between differential characteristic and linear approximation, we will define the linear branch number $\mathcal{P}_l$, which is similar to the differential branch number $\mathcal{P}_d$. Hereafter, we assume P-function is bijective.

Definition 9. *The linear branch number $\mathcal{P}_l$ is defined as:*

$$\mathcal{P}_l = \min_{\Gamma Y \neq 0} (H_w(P^*(\Gamma Y)) + H_w(\Gamma Y)) = \min_{\Gamma Y \neq 0} (H_w(\Gamma Z) + H_w(\Gamma Y)),$$

where ΓY, ΓZ is an output mask value and an input mask value of the P-function, respectively, and P^* is a diffusion function of mask values concerning the P-function.

Next, we will define the upper bound of the linear characteristic probability of a Feistel cipher with SPN round function. That is, the upper bound of the probability is shown by the (lower bound of) minimum number of linear active s-boxes.

Definition 10. *Assume a Feistel cipher with SPN round function. If $H_w(\Gamma Z^{(i)})$ is the number of the ith-round linear active s-boxes, then the linear characteristic probability of the r-round Feistel cipher satisfies the following relationship.*

$$p_l^{(r)} \leq p_s^{\min_{(\Gamma Y^{(0)}, \ldots, \Gamma Y^{(r)}, \Gamma Y^{(r+1)}) \neq (\ldots, 0, 0)} \sum_{i=1}^{r} H_w(\Gamma Z^{(i)})},$$

where $\Gamma Z^{(i)} = P^*(\Gamma Y^{(i)})$ and P^* is the diffusion function of mask values concerning the P-function.

From this definition, clarifying the upper bound of the linear characteristic probability becomes equivalent to determining the (lower bound of) minimum number of linear active s-boxes. To discuss the minimum number easily after this, we denote it as follows.

$$\mathcal{L}^{(r)} = \min_{(\Gamma Y^{(0)},\ldots,\Gamma Y^{(r)},\Gamma Y^{(r+1)})\neq(\ldots,0,0)} \sum_{i=1}^{r} H_w(\Gamma Z^{(i)})$$

Theorem 5. *Assume a Feistel cipher with SPN round function. If both S-function and P-function are bijective, then $\mathcal{L}^{(r)}$ and $\mathcal{P}_l$ also satisfy Lemma 2 to Lemma 6 and Theorem 3.*

Proof. Because of the bijective P-function, a Feistel cipher with SPN round function is transformed into one with PSN round function by Theorem 4. The cipher can be described as:

$$V^{(i+1)} = V^{(i-1)} \oplus S(P(V^{(i)})) = V^{(i-1)} \oplus S(X^{(i)}) = V^{(i-1)} \oplus Z^{(i)},$$

where $V^{(i)} = P^{-1}(X^{(i)})$, $Z^{(i)} = S(X^{(i)})$.

From the duality between differential characteristic and linear approximation, the linear approximation of the round function of the transformed cipher can be expressed as follows using the concatenation rules [4,14].

$$\Gamma V^{(i)} = \Gamma Z^{(i-1)} \oplus \Gamma Z^{(i+1)} = P^*(\Gamma X^{(i)})$$

By the way, since S-function is bijective, $H_w(\Gamma X) = H_w(\Gamma Z)$ because Γx_i is a non-zero input mask value of a linear active s_i-box. Therefore, the linear branch number $\mathcal{P}_l$ is redefined as:

$$\mathcal{P}_l = \min_{\Gamma X\neq 0}(H_w(P^*(\Gamma X)) + H_w(\Gamma X)) = \min_{\Gamma Z\neq 0}(H_w(\Gamma V) + H_w(\Gamma Z))$$

Accordingly, if $\Delta X^{(i)}$ and $\Delta Y^{(i)}$ are exchanged for $\Gamma Z^{(i)}$ and $\Gamma V^{(i)}$, respectively, it turns out that all proofs are satisfied in the same way as for Lemma 2 to Lemma 6 and Theorem 3.

Q.E.D.

For example, the P^*-function of Camellia can be expressed as:

$$P^*_{Camellia} = \begin{bmatrix} 0&1&1&1&0&1&1&1 \\ 1&0&1&1&1&0&1&1 \\ 1&1&0&1&1&1&0&1 \\ 1&1&1&0&1&1&1&0 \\ 1&1&0&0&1&0&1&1 \\ 0&1&1&0&1&1&0&1 \\ 0&0&1&1&1&1&1&0 \\ 1&0&0&1&0&1&1&1 \end{bmatrix}$$

Thus, it is easily seen that $\mathcal{P}_l = 5$, and the following corollary is obtained.

Corollary 3. *Camellia with reduced to 16-rounds (without FL- and FL^{-1}-functions) has no effective linear approximation.*

Proof. The maximum linear probability of Camellia's s-boxes is $q_s = 2^{-6}$. From Theorem 5 and $\mathcal{P}_l = 5$, the maximum linear characteristic probability of Camellia with reduced to 16-rounds is also upper bounded by 2^{-132}.

Q.E.D.

6 Necessary Condition for Desirable P-Functions

In this section, we consider the necessary condition for desirable P-functions. Here, "desirable" means that the round functions are invulnerable to linear cryptanalysis as well as differential cryptanalysis.

Obviously, the condition is $\mathcal{P}_d = \mathcal{P}_l$ from Sect. 4 and Sect. 5. Thus, we investigate P-functions wherein $\mathcal{P}_d = \mathcal{P}_l$.

Theorem 6. *Assume that P-function is bijective and is expressed as an $n \times n$ matrix P over $\mathrm{GF}(2)^m$. When the P-function satisfies $[y_i]^t = [p_{ij}][z_j]^t$, the following relations are satisfied.*

$$[\Delta y_i]^t = [p_{ij}][\Delta z_j]^t,\ [\Gamma z_i]^t = [p_{ij}]^t[\Gamma y_j]^t = [p_{ji}][\Gamma y_j]^t,$$

where $[x_i]$ denotes the vector (or matrix) of X and $[x_i]^t$ denotes the transposed vector (or matrix) of X.

Proof. First, since $y_i = \bigoplus_{j=1}^{n}(p_{ij} \cdot z_j)$,

$$\begin{aligned}\Delta y_i = y_i \oplus y_i' &= \left(\bigoplus_{j=1}^{n}(p_{ij} \cdot z_j)\right) \oplus \left(\bigoplus_{j=1}^{n}(p_{ij} \cdot z_j')\right) \\ &= \bigoplus_{j=1}^{n}(p_{ij} \cdot z_j \oplus p_{ij} \cdot z_j') \\ &= \bigoplus_{j=1}^{n}(p_{ij} \cdot (z_j \oplus z_j')) = \bigoplus_{j=1}^{n}(p_{ij} \cdot \Delta z_j)\end{aligned}$$

Thus, $[\Delta y_i]^t = [p_{ij}][\Delta z_j]^t$ is satisfied.
Second, since the P-function is bijective, $Z \cdot \Gamma Z = Y \cdot \Gamma Y$. Then,

$$\begin{aligned}Y \cdot \Gamma Y &= \bigoplus_{j=1}^{n}\left(\left(\bigoplus_{i=1}^{n}(p_{ji} \cdot z_i)\right) \cdot \Gamma y_j\right) = \bigoplus_{j=1}^{n}\left(\bigoplus_{i=1}^{n}(p_{ji} \cdot z_i \cdot \Gamma y_j)\right) \\ &= \bigoplus_{i=1}^{n}\left(\bigoplus_{j=1}^{n}((p_{ji} \cdot \Gamma y_j) \cdot z_i)\right) = \bigoplus_{i=1}^{n}\left(\left(\bigoplus_{j=1}^{n}(p_{ji} \cdot \Gamma y_j)\right) \cdot z_i\right)\end{aligned}$$

On the other hand, since $Z \cdot \Gamma Z = \bigoplus_{i=1}^{n} (z_i \cdot \Gamma z_i)$, it is obvious that $\Gamma z_i = \bigoplus_{j=1}^{n} (p_{ji} \cdot \Gamma y_j) = \bigoplus_{j=1}^{n} (p_{ij}^t \cdot \Gamma y_j)$. Thus, $[\Gamma z_i]^t = [p_{ji}][\Gamma y_j]^t = [p_{ij}]^t[\Gamma y_j]^t$ is satisfied.

Q.E.D.

Theorem 7. *Assume that I' is a set of matrices that consist of only one 1-element and $(m-1)$ 0-elements in each line and row, i.e., the matrices generated by only interchanging lines and/or rows of unit matrix.*

If a bijective P-function can be expressed as an $n \times n$ matrix P over $\mathrm{GF}(2)^m$ such that $P^t \cdot P \in I'$ or $P^t = I_2 \cdot P \cdot I_1$ where $I_1, I_2 \in I'$, then the P-function satisfies $\mathcal{P}_d = \mathcal{P}_l$.

Proof. By Theorem 6, if the P-function can be expressed as an $n \times n$ matrix P over $\mathrm{GF}(2)^m$, then

$$\mathcal{P}_d = \min_{\Delta Z \neq 0} (H_w(\Delta Z) + H_w(P(\Delta Z))),\ \mathcal{P}_l = \min_{\Gamma Y \neq 0} (H_w(P^t(\Gamma Y)) + H_w(\Gamma Y))$$

(i) In the case of $P^t \cdot P = I^* \in I'$, let $\Gamma Y = P(\Gamma W)$.

Since the P-function is bijective, it is guaranteed that $\{\Gamma Y\} = \{\Gamma W\}$. Thus,

$$\begin{aligned} \mathcal{P}_l &= \min_{\Gamma W \neq 0} (H_w(P^t(P(\Gamma W))) + H_w(P(\Gamma W))) \\ &= \min_{\Gamma W \neq 0} (H_w(I^* \cdot \Gamma W)) + H_w(P(\Gamma W))). \end{aligned}$$

Here, because $I^* \in I'$, $I^* \cdot \Gamma W$ leads to another vector simply by interchanging the elements of ΓW. Thus, $H_w(\Gamma W) = H_w(I^* \cdot \Gamma W)$. As a result, $\exists(\Delta Z, \Gamma W)$, s.t. $\mathcal{P}_d = \mathcal{P}_l$.

(ii) In the case of $P^t = I_2 \cdot P \cdot I_1$ where $I_1, I_2 \in I'$, as mentioned above, since I_1 and I_2 lead to another vector simply by interchanging the elements, $H_w(\Delta X) = H_w(I_1(\Delta X))$ and $H_w(I_2(\Gamma W)) = H_w(\Gamma W)$. Now, let $\Delta Z = I_1(\Delta X)$. Since I_1 is bijective, it is guaranteed that $\{\Delta X\} = \{\Delta Z\}$. Thus,

$$\begin{aligned} \mathcal{P}_d &= \min_{\Delta X \neq 0} (H_w(I_1(\Delta X)) + H_w(P \cdot I_1(\Delta X))) \\ &= \min_{\Delta X \neq 0} (H_w(\Delta X) + H_w(P \cdot I_1(\Delta X))). \end{aligned}$$

On the other hand,

$$\begin{aligned} \mathcal{P}_l &= \min_{\Gamma Y \neq 0} (H_w(I_2 \cdot P \cdot I_1(\Gamma Y)) + H_w(\Gamma Y)) \\ &= \min_{\Gamma Y \neq 0} (H_w(P \cdot I_1(\Gamma Y)) + H_w(\Gamma Y)). \end{aligned}$$

As a result, $\exists(\Delta X, \Gamma Y)$, s.t. $\mathcal{P}_d = \mathcal{P}_l$.

Q.E.D.

For example, the relationship between P-function and P^*-function of Camellia is shown as follows. Thus Theorem 7 indicates that the P-function of Camellia is "desirable."

$$P^*_{Camellia} = P^t_{Camellia} = I^* \cdot P_{Camellia} \cdot I^*,$$

because

$$P_{Camellia} = \begin{bmatrix} 0&1&1&1&1&0&0&1 \\ 1&0&1&1&1&1&0&0 \\ 1&1&0&1&0&1&1&0 \\ 1&1&1&0&0&0&1&1 \\ 0&1&1&1&1&1&1&0 \\ 1&0&1&1&0&1&1&1 \\ 1&1&0&1&1&0&1&1 \\ 1&1&1&0&1&1&0&1 \end{bmatrix}, \quad P^*_{Camellia} = \begin{bmatrix} 0&1&1&1&0&1&1&1 \\ 1&0&1&1&1&0&1&1 \\ 1&1&0&1&1&1&0&1 \\ 1&1&1&0&1&1&1&0 \\ 1&1&0&0&1&0&1&1 \\ 0&1&1&0&1&1&0&1 \\ 0&0&1&1&1&1&1&0 \\ 1&0&0&1&0&1&1&1 \end{bmatrix}, \quad I^* = \begin{bmatrix} 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \end{bmatrix}$$

7 Conclusion

This paper studied the upper bounds of the maximum differential and linear characteristic probabilities of Feistel ciphers with SPN round function. In the same way as for SPN ciphers, we considered the minimum number of differential and linear active s-boxes, which are a measure of the upper bounds of these probabilities, in order to evaluate security against differential and linear cryptanalyses. The advantage of this method is that it considers the interrelation between input and output data in consecutive rounds, unlike Knudsen's estimation.

We focused on the minimum number of active s-boxes in some consecutive rounds of Feistel ciphers, i.e., in three, four, six, eight, and twelve consecutive rounds, since they can determine the upper bounds of the maximum differential and linear probabilities using the differential and linear branch numbers $\mathcal{P}_d$, $\mathcal{P}_l$, respectively. These numbers provide the avalanche effects of P-functions with regard to differential and linear characteristics. As a result, we clarified that the lower bounds of the minimum number of differential (resp. linear) active s-boxes are 2, $\mathcal{P}_d$ ($\mathcal{P}_l$), $\mathcal{P}_d + 2$ ($\mathcal{P}_l + 2$), $2\mathcal{P}_d + 1$ ($2\mathcal{P}_l + 1$), and $3\mathcal{P}_d + 1$ ($3\mathcal{P}_l + 1$), respectively. The interesting result is that the lower bound of the minimum number of active s-boxes is proportional to the branch number every fourth round, while it seems to be every third round at first glance. Furthermore, this means that, if the branch number is the same, a $2r$-round Feistel cipher has almost same invulnerability to differential and linear cryptanalyses as a r-round SPN cipher in terms of the upper bounds of the maximum differential and linear probabilities.

Finally, we investigated the necessary condition for desirable P-functions, which means that the round functions are invulnerable to both differential and linear cryptanalyses. In addition, we showed the example of the round function of Camellia, which satisfies the condition.

References

1. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima, and T. Tokita, "*Camellia*: A 128-bit Block Cipher Suitable for Multiple Platforms – Design and Analysis –," *Selected Areas in Cryptography — 7th Annual International Workshop, SAC2000*, LNCS in this proceeding.

2. K. Aoki, K. Kobayashi, and S. Moriai, "Best Differential Characteristic Search of FEAL," *Fast Software Encryption — 4th International Workshop, FSE'97*, LNCS **1267**, pp.41–53, 1997.
3. K. Aoki, and K. Ohta, "Strict Evaluation of the Maximum Average of Differential Probability and the Maximum Average of Linear Probability," *IEICE Transactions Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E80-A, No. 1, pp. 2–8, 1997.
4. E. Biham, "On Matsui's Linear Cryptanalysis," *Advances in Cryptology — EUROCRYPT'94*, LNCS **950**, pp.341–355, 1995.
5. E. Biham, R. Anderson, and L. R. Knudsen, "Serpent: A New Block Cipher Proposal," *Fast Software Encryption — 5th International Workshop, FSE'98*, LNCS **1372**, pp.222–238, 1998.
6. E. Biham, and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *Journal of Cryptology*, Vol.4, No.1, pp.3–72, 1991.
7. Data Encryption Standard, FIPS-PUB-46, 1977.
8. J. Daemen, L. Knudsen, and V. Rijmen, "The block cipher SQUARE," *Fast Software Encryption — 4th International Workshop, FSE'97*, LNCS **1267**, pp.54–68, 1997.
9. S. Hong, S. Lee, J. Lim, J. Sung, and D. Cheon, "Provable Security against Differential and Linear Cryptanalysis for the SPN structure," *Fast Software Encryption Workshop 2000*, 2000. (LNCS to appear).
10. L. R. Knudsen, "Practically Secure Feistel Ciphers," *Fast Software Encryption — Cambridge Security Workshop*, LNCS **809**, pp.211–221, 1994.
11. M. Kanda, Y. Takashima, T. Matsumoto, K. Aoki, and K. Ohta, "A strategy for constructing fast round functions with practical security against differential and linear cryptanalysis," *Selected Areas in Cryptography — 5th Annual International Workshop, SAC'98*, LNCS **1556**, pp.264–279, 1999.
12. X. Lai, J. L. Massey, and S. Murphy, "Markov ciphers and differential cryptanalysis," *Advances in Cryptology — EUROCRYPT'91*, LNCS **547**, pp.17–38, 1991.
13. M. Matsui, "Linear cryptanalysis method for DES cipher," *Advances in Cryptology — EUROCRYPT'93*, LNCS **765**, pp.386–397, 1994.
14. M. Matsui, "On Correlation Between the Order of S-boxes and the Strength of DES," *Advances in Cryptology — EUROCRYPT'94*, LNCS **950**, pp.366–375, 1995.
15. M. Matsui, "New Block Encryption Algorithm MISTY," *Fast Software Encryption — 4th International Workshop, FSE'97*, LNCS **1267**, pp.54–68, 1997.
16. S. Moriai, K. Aoki, and K. Ohta, "The Best Linear Expression Search of FEAL," IEICE Transactions Fundamentals of Electronics, Communications and Computer Sciences, Vol. E79-A, No. 1, pp. 2–11, 1996.
17. K. Nyberg, "Linear Approximation of Block Ciphers," *Advances in Cryptology — EUROCRYPT'94*, LNCS **950**, pp.439–444, 1995.
18. K. Nyberg, and L. R. Knudsen, "Provable Security Against a Differential Attack," *Journal of Cryptology*, Vol. 8 No. 1, pp. 27–37, 1995.
19. V. Rijmen, J. Daemon, B. Preneel, A. Bosselaers, and E. D. Win, "The cipher SHARK," *Fast Software Encryption — Third International Workshop*, LNCS **1039**, pp.99–111, 1996.