Improving the Upper Bound on the Maximum Differential and the Maximum Linear Hull Probability for SPN Structures and AES

Sangwoo Park[1], Soo Hak Sung[2], Sangjin Lee[3], and Jongin Lim[3]

[1] National Security Research Institute, Korea
psw@etri.re.kr
[2] Department of Applied Mathematics
Pai Chai University, Korea
sungsh@woonam.paichai.ac.kr
[3] Center for Information Security Technologies(CIST)
Korea University, Korea
{sangjin,jilim}@cist.korea.ac.kr

Abstract. We present a new method for upper bounding the maximum differential probability and the maximum linear hull probability for 2 rounds of SPN structures. Our upper bound can be computed for any value of the branch number of the linear transformation and by incorporating the distribution of differential probability values and linear probability values for S-box. On application to AES, we obtain that the maximum differential probability and the maximum linear hull probability for 4 rounds of AES are bounded by 1.144×2^{-111} and 1.075×2^{-106}, respectively.

1 Introduction

Differential cryptanalysis [2] and linear cryptanalysis [12] are the most well-known methods of analysing the security of block ciphers. Accordingly, the designer of block ciphers should evaluate the security of any proposed block cipher against differential cryptanalysis and linear cryptanalysis and prove that it is sufficiently invulnerable against them.

SPN(Substitution and Permutation Network) structure is one of the most commonly used structure in block ciphers. SPN structure is based on Shannon's principles of confusion and diffusion [3] and these principles are implemented through the use of substitution and linear transformation, respectively. AES [6, 14], Crypton [11], and Square [5] are block ciphers composed of SPN structures.

The security of SPN structures against differential cryptanalysis and linear cryptanalysis depends on the maximum differential probability and the maximum linear hull probability. Hong *et al.* proved the upper bound on the maximum differential and the maximum linear hull probability for 2 rounds of SPN structures with highly diffusive linear transformation [7]. Kang *et al.* generalized their result for any value of the branch number of the linear transformation [8].

In [10], Keliher *et al.* proposed a method for finding the upper bound on the maximum average linear hull probability for SPN structures. Application of

T. Johansson (Ed.): FSE 2003, LNCS 2887, pp. 247–260, 2003.

their method to AES yields an upper bound of 2^{-75} when 7 or more rounds are approximated. In [9], it was proposed that the improved upper bound on the maximum average linear hull probability for AES when 9 or more rounds are approximated is 2^{-92}. In [15], Park *et al.* proposed a method for upper bounding the maximum differential probability and the maximum linear hull probability for Rijndael-like structures. Rijndael-like structure is a special case of SPN structures. By applying their method to AES, they obtain that the maximum differential probability and the maximum linear hull probability for 4 rounds of AES are bounded by 1.06×2^{-96}.

In this paper, we present a new method for upper bounding on the maximum differential probability and the maximum linear hull probability for 2 rounds of SPN structures. Our upper bound can be computed for any value of the branch number of the linear transformation and by incorporating the distribution of differential probability values and linear probability values for S-box.

On application to AES, we obtain that the maximum differential probability and the maximum linear hull probability for 4 rounds of AES are bounded by 1.144×2^{-111} and 1.075×2^{-106}, respectively.

2 Backgrounds

One round of SPN structures generally consists of three layers of key addition, substitution, and linear transformation. On the key addition layer, round subkeys and round input values are exclusive-ored. Substitution layer is made up of n small nonlinear substitutions referred to as S-boxes, and the linear transformation layer is a linear transformation used in order to diffuse the cryptographic characteristics of the substitution layer. A typical example of one round of SPN structures is given in Figure 1.

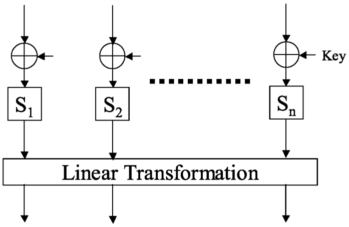

Fig. 1. One round of SPN structure.

On r rounds of SPN structures, the linear transformation of the last round, generally, is omitted, because it has no cryptographic significance. Therefore, 2 rounds of SPN structures is given in Figure 2.

S-boxes and linear transformations should be invertible in order to decipher. Therefore we assume that all S-boxes are bijections from Z_2^m to itself. Moreover, throughout this paper, we assume that round subkeys are independent and uniformly distributed.

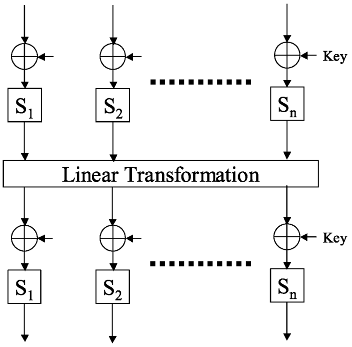

Fig. 2. 2 rounds of SPN structure.

Let S be an S-box with m input and output bits. Differential and linear probability of S are defined as in the following definition:

Definition 1. *For any given $a, b, \mathit{\Gamma}_a, \mathit{\Gamma}_b \in Z_2^m$, define differential probability $DP^S(a,b)$ and linear probability $LP^S(\mathit{\Gamma}_a, \mathit{\Gamma}_b)$ of S by*

$$DP^S(a,b) = \frac{\#\{x \in Z_2^m | S(x) \oplus S(x \oplus a) = b\}}{2^m}$$

and

$$LP^S(\mathit{\Gamma}_a, \mathit{\Gamma}_b) = \left(\frac{\#\{x \in Z_2^m | \mathit{\Gamma}_a \cdot x = \mathit{\Gamma}_b \cdot S(x)\}}{2^{m-1}} - 1 \right)^2,$$

respectively, where $x \cdot y$ denotes the parity(0 or 1) of bitwise product of x and y.

a and b are called input and output differences, respectively. Also, $\mathit{\Gamma}_a$ and $\mathit{\Gamma}_b$ are called input and output mask values, respectively. The strength of an S-box S against differential cryptanalysis is determined by the maximum differential probability, $\max_{a \neq 0, b} DP^S(a,b)$. The strength of an S-box S against linear cryptanalysis depends on the maximum linear probability, $\max_{\mathit{\Gamma}_a, \mathit{\Gamma}_b \neq 0} LP^S(\mathit{\Gamma}_a, \mathit{\Gamma}_b)$.

Definition 2. *The maximum differential probability p and the maximum linear probability q of S are defined by*

$$p = \max_{a \neq 0, b} DP^S(a,b)$$

and

$$q = \max_{\mathit{\Gamma}_a, \mathit{\Gamma}_b \neq 0} LP^S(\mathit{\Gamma}_a, \mathit{\Gamma}_b),$$

respectively.

The maximum differential probability p and the maximum linear probability q for a strong S-box S should be small enough for any input difference $a \neq 0$ and any output mask value $\mathit{\Gamma}_b \neq 0$.

Definition 3. *Differentially active S-box is defined as an S-box given a nonzero input difference and linearly active S-box is defined as an S-box given a nonzero output mask value.*

Since all S-boxes in substitution layer are bijective, if an S-box is differentially/linearly active, then it has a non-zero output difference/input mask value.

For SPN structures, there is a close relationship between the differential probability and the number of differentially active S-boxes. When the number of differentially active S-boxes is large, the differential probability becomes to be small, and when the number of differentially active S-boxes is small, the differential probability becomes to be big. Therefore, the concept of branch number was proposed [5]. We call it the branch number from the viewpoint of differential cryptanalysis, the minimum number of differentially active S-boxes of 2 rounds of SPN structures. Also, we call it the branch number from the viewpoint of linear cryptanalysis, the minimum number of linearly active S-boxes of 2 rounds of SPN structures.

The linear transformation $L : (Z_2^m)^n \longrightarrow (Z_2^m)^n$ can be represented by an $n \times n$ matrix $M = (m_{ij})$. We have $L(x) = Mx$, where $x \in (Z_2^m)^n$ and the addition is done through bitwise exclusive-or. For the block ciphers E2 [13] and Camellia [1], $m_{ij} \in Z_2$ and the multiplication is trivial. For the block cipher Crypton [11], $m_{ij} \in Z_2^m$ and the multiplication is the bitwise logical-and operation. For the block cipher Rijndael [6], $m_{ij} \in GF(2^m)$ and the multiplication is defined as the multiplication over $GF(2^m)$.

It is easy to show that $L(x) \oplus L(x^*) = L(x \oplus x^*)$ and $DP^L(a, L(a)) = 1$ [4].

Definition 4. *Let L be the linear transformation over $(Z_2^m)^n$. The branch number of L from the view point of differential cryptanalysis, β_d, is defined by*

$$\beta_d = min_{x \neq 0}\{wt(x) + wt(L(x))\},$$

where, $wt(x) = wt(x_1, x_2, \ldots, x_n) = \#\{1 \leq i \leq n | x_i \neq 0\}$.

Throughout this paper, we define $wt(x) = wt(x_1, x_2, \ldots, x_n) = \#\{1 \leq i \leq n | x_i \neq 0\}$ when $x = (x_1, x_2, \ldots, x_n)$. If $x \in Z_2^m$, then $wt(x)$ is the Hamming weight of x.

It can be proved that, if $m_{ij} \in Z_2$, then $LP^L(M^t \Gamma_b, \Gamma_b) = 1$. Therefore, we know that $LP^L(\Gamma_a, (M^{-1})^t \Gamma_a) = 1$. Also, if $m_{ij} \in GF(2^m)$, then it can be proved that $LP^L(\Gamma_a, C\Gamma_a) = 1$, for some $n \times n$ matrix C over $GF(2^m)$ [8]. Therefore, we can define the branch number β_l from the view point of linear cryptanalysis as follows:

$$\beta_l = \begin{cases} min_{\Gamma_a \neq 0}\{wt(\Gamma_a) + wt((M^{-1})^t \Gamma_a)\}, & \text{if } m_{ij} \in Z_2, 1 \leq i, j \leq n, \\ min_{\Gamma_a \neq 0}\{wt(\Gamma_a) + wt(C\Gamma_a)\}, & \text{if } m_{ij} \in GF(2^m), 1 \leq i, j \leq n. \end{cases}$$

3 Security of 2 Rounds of SPN Structures

In this section, we give an upper bound on the maximum differential probability for 2 rounds of SPN structure. We also give an upper bound on the maximum linear hull probability.

The following lemma can be considered as a generalized Cauchy-Schwarz inequality.

Lemma 1. *Let* $\{x_i^{(j)}\}_{i=1}^n, 1 \leq j \leq m$, *be sequence of real numbers. Then the following inequality is satisfied.*

$$\sum_{i=1}^n |x_i^{(1)} x_i^{(2)} \cdots x_i^{(m)}| \leq \left(\sum_{i=1}^n |x_i^{(1)}|^m\right)^{\frac{1}{m}} \left(\sum_{i=1}^n |x_i^{(2)}|^m\right)^{\frac{1}{m}} \cdots \left(\sum_{i=1}^n |x_i^{(m)}|^m\right)^{\frac{1}{m}}.$$

Proof. We will prove the result by using mathematical induction. For $m = 2$, the result is trivial. Assume that the result holds for $m - 1$. We have, by the Hölder's inequality, that

$$\sum_{i=1}^n |x_i^{(1)} \cdots x_i^{(m-1)} x_i^{(m)}| \leq \left(\sum_{i=1}^n |x_i^{(1)} \cdots x_i^{(m-1)}|^{\frac{m}{m-1}}\right)^{\frac{m-1}{m}} \left(\sum_{i=1}^n |x_i^{(m)}|^m\right)^{\frac{1}{m}}.$$

By the induction hypothesis, the right hand side is bounded by

$$\left(\sum_{i=1}^n |x_i^{(1)}|^m\right)^{\frac{1}{m}} \cdots \left(\sum_{i=1}^n |x_i^{(m-1)}|^m\right)^{\frac{1}{m}} \left(\sum_{i=1}^n |x_i^{(m)}|^m\right)^{\frac{1}{m}}.$$

Thus, the result is proved.

From Lemma 1, we get the following lemma.

Lemma 2. *Let* $\{x_i^{(j)}\}_{i=1}^n, 1 \leq j \leq m$, *be sequence of real numbers. Then the following inequality is satisfied.*

$$\sum_{i=1}^n |x_i^{(1)} \cdots x_i^{(m)}| \leq \max\{\sum_{i=1}^n |x_i^{(1)}|^m, \cdots, \sum_{i=1}^n |x_i^{(m)}|^m\}.$$

Theorem 1. *Let* β_d *be the branch number of the linear transformation* L *from the viewpoint of differential cryptanalysis. Then, the maximum differential probability for 2 rounds of SPN structure is bounded by*

$$\max\left\{\max_{1\leq i\leq n} \max_{1\leq u\leq 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(u,j)\}^{\beta_d}, \max_{1\leq i\leq n} \max_{1\leq u\leq 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(j,u)\}^{\beta_d}\right\}.$$

Proof. Let $a = (a_1, \cdots, a_n)$, $b = (b_1, \cdots, b_n)$ be the input difference and output difference, respectively, for 2 rounds of SPN structure. Since $DP^L(\alpha, L(\alpha)) = 1$, the differential probability $DP_2(a, b)$ is given as

$$DP_2(a,b) = \sum_x \left(\prod_{i=1}^n DP^{S_i}(a_i, x_i)\right) \left(\prod_{j=1}^n DP^{S_j}(y_j, b_j)\right),$$

where $y = L(x), x = (x_1, \cdots, x_n)$, and $y = (y_1, \cdots, y_n)$. Without loss of generality, we assume that $a_1 \neq 0, \cdots, a_k \neq 0, a_{k+1} = 0, \cdots, a_n = 0$, $b_1 \neq 0, \cdots, b_l \neq 0, b_{l+1} = 0, \cdots, b_n = 0$. Note that if $\alpha = 0, \beta \neq 0$ or $\alpha \neq 0, \beta = 0$, then $DP^{S_i}(\alpha, \beta) = 0$. Hence, it is enough to consider the following x(and $y = L(x)$) only in the above summation.

$$x_1 \neq 0, \cdots, x_k \neq 0, x_{k+1} = 0, \cdots, x_n = 0,$$

$$y_1 \neq 0, \cdots, y_l \neq 0, y_{l+1} = 0, \cdots, y_n = 0.$$

We let the solutions of the above system be as follows:

t	x_1	$\cdots$	x_k	y_1	$\cdots$	y_l
1	$x_1^{(1)}$	$\cdots$	$x_1^{(k)}$	$y_1^{(1)}$	$\cdots$	$y_1^{(l)}$
2	$x_2^{(1)}$	$\cdots$	$x_2^{(k)}$	$y_2^{(1)}$	$\cdots$	$y_2^{(l)}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
δ	$x_\delta^{(1)}$	$\cdots$	$x_\delta^{(k)}$	$y_\delta^{(1)}$	$\cdots$	$y_\delta^{(l)}$

Then the maximum differential probability $DP_2(a, b)$ can be written as

$$DP_2(a, b) = \sum_{t=1}^{\delta} \left(\prod_{i=1}^{k} DP^{S_i}(a_i, x_t^{(i)}) \right) \left(\prod_{j=1}^{l} DP^{S_j}(y_t^{(j)}, b_j) \right).$$

By the definition of branch number, it follows that $k + l \geq \beta_d$. We divide the proof into two cases: $k + l = \beta_d$ and $k + l > \beta_d$.

(Case 1: $k + l = \beta_d$). In this case, we have that, for each $i(1 \leq i \leq k)$, $x_1^{(i)}, \cdots, x_\delta^{(i)}$ are distinct, because L is linear and $k + l = \beta_d$. If, for some $i(1 \leq i \leq k)$, $x_1^{(i)}, \cdots, x_\delta^{(i)}$ are not distinct, then there exist a pair $(x_J^{(i)}, x_{J'}^{(i)})$ such that $x_J^{(i)} = x_{J'}^{(i)}$, where $x_J^{(i)}$ is i-th component of x and $x_{J'}^{(i)}$ is i-th component of x', respectively. Therefore, i-th component of $x \oplus x'$ is equal to zero. Since $L(x) \oplus L(x') = L(x \oplus x')$, this is a contradiction of the definition of branch number. We also have that, for each $j(1 \leq j \leq l)$, $y_1^{(j)}, \cdots, y_\delta^{(j)}$ are distinct.

From Lemma 2, $DP_2(a, b)$ is bounded by

$$\begin{aligned} &\max \left\{ \sum_{t=1}^{\delta} \{DP^{S_1}(a_1, x_t^{(1)})\}^{\beta_d}, \cdots, \sum_{t=1}^{\delta} \{DP^{S_k}(a_k, x_t^{(k)})\}^{\beta_d}, \right. \\ &\qquad \left. \sum_{t=1}^{\delta} \{DP^{S_1}(y_t^{(1)}, b_1)\}^{\beta_d}, \cdots, \sum_{t=1}^{\delta} \{DP^{S_l}(y_t^{(l)}, b_l)\}^{\beta_d} \right\} \\ &\leq \max \left\{ \max_{1 \leq i \leq n} \max_{1 \leq u \leq 2^m - 1} \sum_{j=1}^{2^m - 1} \{DP^{S_i}(u, j)\}^{\beta_d}, \right. \\ &\qquad \left. \max_{1 \leq i \leq n} \max_{1 \leq u \leq 2^m - 1} \sum_{j=1}^{2^m - 1} \{DP^{S_i}(j, u)\}^{\beta_d} \right\}. \end{aligned}$$

(Case 2: $k+l > \beta_d$). In this case, $x_1^{(i)}, \cdots, x_\delta^{(i)}$ or $y_1^{(j)}, \cdots, y_\delta^{(j)}$ are not necessarily dintinct. However, when we consider the subset of solutions such that $k+l-\beta_d$ components are fixed($x_1 = i_1, \ldots, x_p = i_p$, $y_1 = j_1, \ldots, y_q = j_q$), each of the other β_d components has distinct values, where $0 \leq p \leq k-1$, $0 \leq q \leq l-1$, and $p+q = k+l-\beta_d$. We denote this subset of solutions by $A_{i_1,\ldots,i_p,j_1,\ldots,j_q}$. Note that $A_{i_1,\ldots,i_p,j_1,\ldots,j_q}$ could be the empty set. As in the case 1(or by Lemma 2), we obtain that

$$\begin{aligned}
&\sum_{(x,y)\in A_{i_1,\ldots,i_p,j_1,\ldots,j_q}} \left(\prod_{i=1}^{k} DP^{S_i}(a_i, x_i)\right)\left(\prod_{j=1}^{k} DP^{S_j}(y_j, b_j)\right) \\
&= DP^{S_1}(a_1, i_1)\cdots DP^{S_p}(a_p, i_p) DP^{S_1}(j_1, b_1)\cdots DP^{S_q}(j_q, b_q) \times \\
&\quad \sum_{(x,y)\in A_{i_1,\ldots,i_p,j_1,\ldots,j_q}} \left(\prod_{i=p+1}^{k} DP^{S_i}(a_i, x_i)\right)\left(\prod_{j=q+1}^{k} DP^{S_j}(y_j, b_j)\right) \\
&\leq DP^{S_1}(a_1, i_1)\cdots DP^{S_p}(a_p, i_p) DP^{S_1}(j_1, b_1)\cdots DP^{S_q}(j_q, b_q) \times \\
&\quad \max\left\{ \max_{1\leq i\leq n} \max_{1\leq u\leq 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(u, j)\}^{\beta_d}, \right. \\
&\quad\quad \left. \max_{1\leq i\leq n} \max_{1\leq u\leq 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(j, u)\}^{\beta_d} \right\} \\
&=: p_{i_1,\ldots,i_p,j_1,\ldots,j_q}
\end{aligned}$$

Thus $DP_2(a, b)$ is bounded by

$$\begin{aligned}
&\sum_{i_1=1}^{2^m-1} \cdots \sum_{i_p=1}^{2^m-1} \sum_{j_1=1}^{2^m-1} \cdots \sum_{j_q=1}^{2^m-1} p_{i_1,\ldots,i_p,j_1,\ldots,j_q} \\
&= \max\left\{ \max_{1\leq i\leq n} \max_{1\leq u\leq 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(u, j)\}^{\beta_d}, \right. \\
&\quad\quad \left. \max_{1\leq i\leq n} \max_{1\leq u\leq 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(j, u)\}^{\beta_d} \right\}.
\end{aligned}$$

From Cases 1 and 2, the result is proved.

Corollary 1. *Let β_d be the branch number of the linear transformation L from the viewpoint of differential cryptanalysis. Then the maximum differential probability for 2 rounds of SPN structure is bounded by $p^{\beta_d - 1}$, where p is the maximum differential probability for the S-boxes.*

Proof. By Theorem 1, the maximum differential probability for 2 rounds of SPN structure is bounded by

$$p^{\beta_d - 1} \times \max\left\{ \max_{1\le i\le n} \max_{1\le u\le 2^m-1} \sum_{j=1}^{2^m-1} DP^{S_i}(u,j), \right.$$
$$\left. \max_{1\le i\le n} \max_{1\le u\le 2^m-1} \sum_{j=1}^{2^m-1} DP^{S_i}(j,u) \right\} = p^{\beta_d-1}.$$

Theorem 2. *Let β_l be the branch number of the linear transformation L from the viewpoint of the linear cryptanalysis. The maximum linear hull probability for 2 rounds of SPN structure is bounded by*

$$\max\left\{ \max_{1\le i\le n} \max_{1\le u\le 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(u,j)\}^{\beta_l}, \max_{1\le i\le n} \max_{1\le u\le 2^m-1} \sum_{j=1}^{2^m-1} \{DP^{S_i}(j,u)\}^{\beta_l} \right\}.$$

Corollary 2. *Let β_l be the branch number of the linear transformation L from the viewpoint of linear cryptanalysis. Then the maximum linear hull probability for 2 rounds of SPN structure is bounded by $q^{\beta_l - 1}$, where q is the maximum linear hull probability for the S-boxes.*

Hong *et al.* proved Corollary 1 and 2 when $\beta_l = n+1$ or n [7]. Kang *et al.* proved them for any value of the branch number of the linear transformation [8].

4 Security of AES

AES is a block cipher composed of SPN structures and its linear transformation consists of ShiftRows transformation and MixColumns transformation.

Let $\pi : (Z_2^8)^{16} \longrightarrow (Z_2^8)^{16}$ be the ShiftRows transformation of AES. Let $x = (x_1,x_2,x_3,x_4) = (x_{11},x_{12},x_{13},x_{14},\ x_{21},\ \ldots,\ x_{34},\ x_{41},x_{42},x_{43},x_{44})$ be the input of π. Figure 3 illustrate the ShiftRows transformation π of AES.

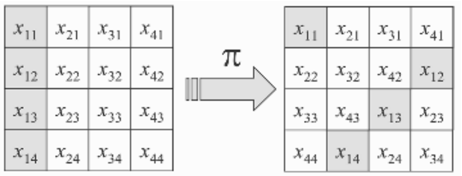
x_{11} x_{21} x_{31} x_{41}
x_{12} x_{22} x_{32} x_{42}
x_{13} x_{23} x_{33} x_{43}
x_{14} x_{24} x_{34} x_{44}
π
x_{11} x_{21} x_{31} x_{41}
x_{22} x_{32} x_{42} x_{12}
x_{33} x_{43} x_{13} x_{23}
x_{44} x_{14} x_{24} x_{34}

Fig. 3. ShiftRows transformation of AES.

Let $y = (y_1,y_2,y_3,y_4) = (y_{11},y_{12},y_{13},y_{14},\ y_{21},\ \ldots,\ y_{34},\ y_{41},y_{42},y_{43},y_{44})$ be the output of π. It is easy to check that, for any $i (i = 1,2,3,4)$, each byte of y_i comes from different x_i. For example, for $y_1 = (y_{11}, y_{12}, y_{13}, y_{14}) = (x_{11}, x_{22}, x_{33}, x_{44})$, x_{11} is a byte coming from x_1. Furthermore, x_{22}, x_{33} and x_{44} are elements of x_2, x_3 and x_4, respectively.

The MixColumns transformation of AES operates on the state column by column, treating each column as a four-term polynomial. Let $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ be the MixColumns transformation of AES. Let $y = (y_1, y_2, y_3, y_4) = (y_{11}, y_{12}, y_{13}, y_{14}, y_{21}, \ldots, y_{34}, y_{41}, y_{42}, y_{43}, y_{44})$ be the input of θ and $z = (z_1, z_2, z_3, z_4) = (z_{11}, z_{12}, z_{13}, z_{14}, z_{21}, \ldots, z_{34}, z_{41}, z_{42}, z_{43}, z_{44})$ be the output of θ, respectively. Each of θ_i can be written as a matrix multiplication as follows:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \end{pmatrix} = \begin{pmatrix} 02\ 03\ 01\ 01 \\ 01\ 02\ 03\ 01 \\ 01\ 01\ 02\ 03 \\ 03\ 01\ 01\ 02 \end{pmatrix} \cdot \begin{pmatrix} z_{i1} \\ z_{i2} \\ z_{i3} \\ z_{i4} \end{pmatrix}.$$

In the matrix multiplication, the addition is just bitwise exclusive-or and the multiplication is defined as the multiplication over $GF(2^8)$. We can consider each θ_i as a linear transformation and we know that the branch number of each θ_i is 5.

In [15], the upper bound on the maximum differential probability for 2 rounds of Rijndael-like structure is obtained as follows:

Definition 5. *Rijndael-like structures are the block ciphers composed of SPN structures satisfying the followings:*

(i) Their linear transformation has the form $(\theta_1, \theta_2, \theta_3, \theta_4) \circ \pi$.
(ii) (The condition of π) Each of bytes of y_i comes from each different x_i, where $x = (x_1, x_2, x_3, x_4)$ is input of π and $y = (y_1, y_2, y_3, y_4)$ is output of π, respectively.
(iii) (The condition of $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$) When we consider each of θ_i as a linear transformation, the followings hold:

$$\beta_d^{\theta_1} = \beta_d^{\theta_2} = \beta_d^{\theta_3} = \beta_d^{\theta_4} \text{ and } \beta_l^{\theta_1} = \beta_l^{\theta_2} = \beta_l^{\theta_3} = \beta_l^{\theta_4}.$$

Definition 6. *For $x = (x_1, \ldots, x_n)$, the pattern of x, γ_x, is defined by $\gamma_x = (\gamma_1, \ldots, \gamma_n) \in Z_2^n$, where, if $x_i = 0$, then $\gamma_i = 0$, and if $x_i \neq 0$, then $\gamma_i = 1$.*

Theorem 3 ([15]).

$$DP_2(a, b) \leq \begin{cases} p^{wt(\gamma_{\pi(a)})(\beta_d - 1)}, & \text{if } \gamma_{\pi(a)} = \gamma_b, \\ 0, & \text{otherwise.} \end{cases}$$

By Theorem 3, the upper bound on the maximum differential probability for 2 rounds of Rijndael-like structures is $p^{\beta_d - 1}$. By applying Theorem 3 to AES, it is obtained that the maximum differential probability for 2 rounds of AES is bounded by 2^{-24}, because $\beta_d = 5$, $p = 2^{-6}$. Note that this result depends on the maximum differential probability of S-box.

By applying our result to Theorem 3, new upper bound on the maximum differential probability for 2 rounds of AES can be obtained. We apply Theorem 1 to 2 rounds of AES. Let S be the S-box of AES. If nonzero $a \in Z_2^8$ is fixed, and b varies over Z_2^8, then the distribution of differential probability of S-box, $DP^S(a,b)$ is independent of a, and is given in Table 1. In Table 1, ρ_i is the differential probability and π_i is the number of occurrences of ρ_i. If nonzero $b \in Z_2^8$ is fixed, and a varies over Z_2^8, then the same distribution is obtained.

Table 1. The distribution of differential probability for AES S-box.

i	1	2	3
ρ_i	2^{-6}	2^{-7}	0
π_i	1	126	129

From Theorem 1 and Table 1, we have $DP_2^{\theta_i}(a,b) \leq \sum_{j=1}^{255}\{DP^S(1,j)\}^5 \approx 1.23 \times 2^{-28}$.

Theorem 4. *When $\gamma_{\pi(a)} = \gamma_b$, the upper bound of the maximum differential probability of 2 rounds of AES is as following:*

$$DP_2(a,b) \leq (1.23 \times 2^{-28})^{wt(\pi(a))}.$$

Therefore, the maximum differential probability of 2 rounds of AES is bounded by 1.23×2^{-28}.

To compute the upper bound on the maximum differential probability for 4 rounds of AES, we need the following notations:

- $x^{(i)} = (x_1^{(i)}, \ldots, x_4^{(i)}) = (x_{11}^{(i)}, x_{12}^{(i)}, x_{13}^{(i)}, x_{14}^{(i)}, \ldots, x_{41}^{(i)}, x_{42}^{(i)}, x_{43}^{(i)}, x_{44}^{(i)})$: the input of π at i-th round.
- $y^{(i)} = (y_1^{(i)}, \ldots, y_4^{(i)}) = (y_{11}^{(i)}, y_{12}^{(i)}, y_{13}^{(i)}, y_{14}^{(i)}, \ldots, y_{41}^{(i)}, y_{42}^{(i)}, y_{43}^{(i)}, y_{44}^{(i)})$: the output of π at i-th round, i.e. the input of θ at i-th round.
- $z^{(i)} = (z_1^{(i)}, \ldots, z_4^{(i)}) = (z_{11}^{(i)}, z_{12}^{(i)}, z_{13}^{(i)}, z_{14}^{(i)}, \ldots, z_{41}^{(i)}, z_{42}^{(i)}, z_{43}^{(i)}, z_{44}^{(i)})$: the output of θ at i-th round.

Theorem 5. *The differential probability for 4 rounds of AES is bounded by 1.144×2^{-111}.*

Proof. We compute the upper bound on $DP_4(a,b)$ for the value of $wt(\gamma_{\pi(a)})$ and $wt(b)$. Since $\beta_d = 5$, if $wt(\gamma_{\pi(a)}) + wt(b) \leq 4$, then $DP_4(a,b) = 0$. Therefore, it is sufficient to compute the upper bound on $DP_4(a,b)$, when $wt(\gamma_{\pi(a)}) + wt(b) \geq 5$.

(Case 1: $wt(\gamma_{\pi(a)}) = 4$). By Theorem 4,

$$\begin{aligned} DP_4(a,b) = \sum_{x^{(2)}} DP_2(a,x^{(2)}) DP_2(z^{(2)},b) &\leq \max_{x^{(2)}} DP_2(a,x^{(2)}) \\ &\leq (1.23 \times 2^{-28})^4 \approx 1.144 \times 2^{-111}. \end{aligned}$$

(Case 2: $wt(b)=4$). By Theorem 4,

$$\begin{aligned} DP_4(a,b) &= \sum_{x^{(2)}} DP_2(a,x^{(2)})DP_2(z^{(2)},b) \le \max_{z^{(2)}} DP_2(z^{(2)},b) \\ &\le (1.23\times 2^{-28})^4 \approx 1.144\times 2^{-111}. \end{aligned}$$

(Case 3: $wt(\gamma_{\pi(a)})=2$ and $wt(b)=3$). We assume that $\gamma_{\pi(a)}=(1,1,0,0)$ and $\gamma_b=(1,1,1,0)$. Then we can represent $DP_4(a,b)$ as follows:

$$\begin{aligned} DP_4(a,b) &= \sum_{x^{(2)}} DP_2(a,x^{(2)})DP_2(z^{(2)},b) \\ &= \sum_{i=1}^{4} \sum_{x^{(2)},wt(z^{(2)})=i} DP_2(a,x^{(2)})DP_1(z^{(2)},b) =: I+II+III+IV. \end{aligned}$$

We know that $wt(y_i^{(2)}) \le wt(x^{(2)}) = wt(\gamma_{\pi(a)}) = 2$ and $wt(z_i^{(2)}) = wt(x_i^{(3)}) \le wt(b)=3$. Since $\beta_d^{\theta_i}=5$, we obtain that $wt(y_i^{(2)})=2$ and $wt(z_i^{(2)})=3$, where $y_i^{(2)}$ and $z_i^{(2)}$ are the nonzero components of $y^{(2)}$ and $z^{(2)}$, respectively. Note that $y_i^{(2)}$ is the input mask of θ_i and $z_i^{(2)}$ is the output mask of θ_i. Now, we compute the value of I. We can represent I as follows:

$$\begin{aligned} I &= \sum_{x^{(2)},\gamma_{y^{(2)}}=(1,0,0,0)} DP_2(a,x^{(2)})DP_2(y^{(2)},b) \\ &\quad + \sum_{x^{(2)},\gamma_{y^{(2)}}=(0,1,0,0)} DP_2(a,x^{(2)})DP_2(y^{(2)},b) \\ &\quad + \sum_{x^{(2)},\gamma_{y^{(2)}}=(0,0,1,0)} DP_2(a,x^{(2)})DP_2(y^{(2)},b) \\ &\quad + \sum_{x^{(2)},\gamma_{y^{(2)}}=(0,0,0,1)} DP_2(a,x^{(2)})DP_2(y^{(2)},b) \\ &=: I_1+I_2+I_3+I_4 \end{aligned}$$

At first, we compute the value of I_1. Since $\gamma_{x^{(2)}}=\gamma_{\pi(a)}=(1,1,0,0)$, $\gamma_{z^{(2)}}=(1,0,0,0)$, and, $wt(y_1^{(2)})=2$, from the definition of π, we obtain that $x^{(2)}=(x_{11}^{(2)},0,0,0,\ 0,0,0,x_{24}^{(2)},\ 0,0,0,0,\ 0,0,0,0)$. Furthermore, since $\gamma_{z_1^{(2)}}=\gamma_{x_1^{(3)}}$, $\gamma_z^{(3)}=\gamma_y^{(3)}=\gamma_b=(1,1,1,0)$, and, $wt(z_1^{(2)})=3$, we obtain that $z^{(2)}=(z_{11}^{(2)},z_{12}^{(2)},z_{13}^{(2)},0,\ 0,0,0,0,\ 0,0,0,0,\ 0,0,0,0)$. $(x_{11}^{(2)},0,0,x_{24}^{(2)})$ and $(z_{11}^{(2)},z_{12}^{(2)},z_{13}^{(2)},0)$ are the nonzero input mask and output mask of θ_1, respectively. Since $\beta_d^{\theta_1}=5$, each of $x_{11}^{(2)}$, $x_{24}^{(2)}$, $z_{11}^{(2)}$, $z_{12}^{(2)}$, $z_{13}^{(2)}$ is of distinct value. Therefore, we can establish the following:

$$\begin{aligned} I_1 &= \sum_{x^{(2)},\gamma_y^{(2)}=(1,0,0,0)} DP_2^{\theta_1}(a_1^*,(x_{11}^{(2)},0,0,0))DP_2^{\theta_2}(a_2^*,(0,0,0,x_{24}^{(2)}))DP_2(y^{(2)},b) \\ &\le P^4 \sum_{x_{11}^{(2)}} DP_2^{\theta_1}(a_1^*,(x_{11}^{(2)},0,0,0)), \end{aligned}$$

where $P = 1.23 \times 2^{-28}$, the upper bound of $DP_2^{\theta_i}(a,b)$. By applying the same method, the upper bounds of I_2, I_3 and I_4 can be determined.

$$\begin{aligned} I \le P^4 \Bigg(&\sum_{x_{11}^{(2)}} DP_2^{\theta_i}(a_1^*,(x_{11}^{(2)},0,0,0)) + \sum_{x_{12}^{(2)}} DP_2^{\theta_i}(a_1^*,(0,x_{12}^{(2)},0,0)) \\ &+ \sum_{x_{13}^{(2)}} DP_2^{\theta_i}(a_1^*,(0,0,x_{13}^{(2)},0)) + \sum_{x_{14}^{(2)}} DP_2^{\theta_i}(a_1^*,(0,0,0,x_{14}^{(2)})) \Bigg). \end{aligned}$$

Using the same method, we arrive at the followings:

$$\begin{aligned} II &\le P^4 \sum_{wt(x_1^{(2)})=2} DP_2^{\theta_1}(a_1^*,x_1^{(2)}) \\ III &\le P^4 \sum_{wt(x_1^{(3)})=2} DP_2^{\theta_1}(a_1^*,x_1^{(2)}) \\ IV &\le P^4 \sum_{wt(x_1^{(4)})=2} DP_2^{\theta_1}(a_1^*,x_1^{(2)}) \end{aligned}$$

Therefore,

$$\begin{aligned} DP_4(a,b) \le I + II + III + IV \le P^4 \sum_{x_1^{(2)}} DP_2^{\theta_i}(a_1^*,x_1^{(2)}) = P^4 \\ \le (1.23 \times 2^{-28})^4 \approx 1.144 \times 2^{-111}. \end{aligned}$$

(Case 4: $wt(\gamma_{\pi(a)}) = 3$ and $wt(b) = 2$). The proof is similar to that of Case 3 and we arrive at the following:

$$DP_4(a,b) \le (1.23 \times 2^{-28})^4 \approx 1.144 \times 2^{-111}.$$

(Case 5: $wt(\gamma_{\pi(a)}) = 3$ and $wt(b) = 3$). The proof is similar to that of Case 3 and we arrive at the following:

$$DP_4(a,b) \le (1.23 \times 2^{-28})^4 \approx 1.144 \times 2^{-111}.$$

The distribution of linear probability value $LP^S(a,b)$ for AES S-box is given in the Table 2. In the table, ρ_i is the linear probability value and ϕ_i is the number of occurence of ρ_i.

From Theorem 2 and Table 2, we have $LP_2^{\theta_i}(a,b) \le \sum_{j=1}^{255}\{LP^S(1,j)\}^5 \approx 1.44 \times 2^{-27}$. Using the similar method as in Theorem 5, we can compute the upper bound on the linear hull probability for 4 rounds of AES.

Table 2. The distribution of linear probability values for AES S-box.

i	1	2	3	4	5	6	7	8	9
ρ_i	$(\frac{8}{64})^2$	$(\frac{7}{64})^2$	$(\frac{6}{64})^2$	$(\frac{5}{64})^2$	$(\frac{4}{64})^2$	$(\frac{3}{64})^2$	$(\frac{2}{64})^2$	$(\frac{1}{64})^2$	0
ϕ_i	5	16	36	24	34	40	36	48	17

Theorem 6. *The linear probability of 4 rounds of AES is bounded by* $(1.44 \times 2^{-27})^4 \approx 1.075 \times 2^{-106}$.

We know that the differential probabilities for $r (r \geq 5)$ rounds of AES are smaller than or equal to the maximum differential probability for 4 rounds of AES.

$$DP_5(a, b) = \sum_{x^{(4)}} DP_4(a, x^{(4)}) DP_1(z^{(4)}, b) \leq \max_{x^{(4)}} DP_4(a, x^{(4)}).$$

Therefore, the upper bound on the maximum differential probability in Theorem 5 is the upper bound for $r (r \geq 5)$ rounds of AES. Similarly, the maximum linear hull probability for 4 rounds of AES in Theorem 6 is the upper bound for $r (r \geq 5)$ rounds of AES.

5 Conclusion

In this paper, we have obtained a new upper bound on the maximum differential probability and the maximum linear hull probability for 2 rounds of SPN structure. Our upper bound can be computed for any value of the branch number of the linear transformation. By applying this result, we have proved that the maximum differential probability for 4 rounds of AES is bounded by 1.144×2^{-111}. Also, we have proved that the maximum linear hull probability for 4 rounds of AES is bounded by 1.075×2^{-106}.

References

1. Kazumaro Aoki, Tetsuya Ichikawa, Masayuki Kanda, Mitsuru Matsui, Shiho Moriai, Junko Nakajima, and Toshio Tokita. Camellia: A 128-bit block cipher suitable for multiple platforms - design and analysis. In Douglas R. Stinson and Stafford Tavares, editors, *Selected Areas in Cryptography*, volume 2012 of *Lecture Notes in Computer Science*, pages 39–56. Springer, 2000.
2. Eli Biham and Adi Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, 4(1):3–72, 1991.
3. C.E.Shannon. Communication Theory of Secrecy System. *Bell System Technical Journal*, 28:656–715, October 1949.
4. Joan Daemen, René Govaerts, and Joos Vandwalle. Correlation matrices. In Bart Preneel, editor, *Fast Software Encryption, Second International Workshop*, volume 1008 of *Lecture Notes in Computer Science*, pages 275–285. Springer, 1994.

5. Joan Daemen, Lars R. Knudsen, and Vincent Rijmen. The block cipher square. In Eli Biham, editor, *Fast Software Encryption, 4th International Workshop*, volume 1267 of *Lecture Notes in Computer Science*, pages 149–165. Springer, 1997.
6. Joan Daemen and Vincent Rijmen. Rijndael, AES Proposal. *http://www.nist.gov/aes*, 1998.
7. Seokhie Hong, Sangjin Lee, Jongin Lim, Jaechul Sung, Donghyeon Cheon, and Inho Cho. Provable security against differential and linear cryptanalysis for the SPN structure. In Bruce Schneier, editor, *Fast Software Encryption, 7th International Workshop*, volume 1978 of *Lecture Notes in Computer Science*, pages 273–283. Springer, 2000.
8. Ju-Sung Kang, Seokhie Hong, Sangjin Lee, Okyeon Yi, Choonsik Park, and Jongin Lim. Practical and provable security against differential and linear cryptanalysis for substitution-permutation networks. *ETRI Journal*, 23(4):158–167, 2001.
9. Liam Keliher, Henk Meijer, and Stafford Tavares. Improving the upper bound on the maximum average linear hull probability for Rijndael. In Serge Vaudenay and Amr M. Youssef, editors, *Selected Areas in Cryptography, 8th Annual International Workshop*, volume 2259 of *Lecture Notes in Computer Science*, pages 112–128. Springer, 2001.
10. Liam Keliher, Henk Meijer, and Stafford Tavares. New method for upper bounding the maximum average linear hull probability for SPNs. In Birgit Pfitzmann, editor, *Advances in Cryptology - Eurocrypt 2001*, volume 2045 of *Lecture Notes in Computer Science*, pages 420–436. Springer-Verlag, Berlin, 2001.
11. Chae Hoon Lim. CRYPTON, AES Proposal. *http://www.nist.gov/aes*, 1998.
12. Mitsuru Matsui. Linear cryptanalysis method for DES cipher. In Tor Helleseth, editor, *Advances in Cryptology - Eurocrypt'93*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer-Verlag, Berlin, 1994.
13. NTT-Nippon Telegraph and Telephone Corporation. E2: Efficient Encryption algorithm, AES Proposal. *http://www.nist.gov/aes*, 1998.
14. National Institute of Standards and Technology. FIPS PUB 197 : Advanced Encryption Standard(AES), November 2001.
15. Sangwoo Park, Soo Hak Sung, Seongtaek Chee, E-Joong Yoon, and Jongin Lim. On the security of Rijndael-like structures against differential and linear cryptanalysis. In Yuliang Zheng, editor, *Advances in Cryptology - Asiacrypt 2002*, volume 2501 of *Lecture Notes in Computer Science*, pages 176–191. Springer, 2002.