Improved Upper Bounds of Differential and Linear Characteristic Probability for Camellia

Taizo Shirai, Shoji Kanamaru, and George Abe

Sony Corporation
7-35 Kitashinagawa 6-chome, Shinagawa-ku, Tokyo, 141-0001 Japan
{Taizo.Shirai, Shoji.Kanamaru, George.Abe}@jp.sony.com

Abstract. We discuss the security of the block cipher Camellia against differential attack and linear attack. The security of Camellia against these attacks has been evaluated by upper bounds of maximum differential characteristic probability (MDCP) and maximum linear characteristic probability (MLCP) calculated by the least numbers of active S-boxes which are found by a search method[2]. However, we found some truncated differential paths generated by the method have wrong properties. We show a new evaluation method for truncated differential and linear paths to discard such wrong paths by using linear equations systems and sets of nonzero conditions. By applying this technique to Camellia, we found tighter upper bounds of MDCP and MLCP for reduced-round Camellia. As a result, 10-round Camellia without FL/FL^{-1} has no differential and linear characteristic with probability higher than 2^{-128}.

1 Introduction

Camellia[1] is a block cipher which was suggested as a candidate for the NESSIE project[14] and recently selected for the 2nd phase of the project, and also suggested as a candidate for the CRYPTREC project in Japan[15]. The security of Camellia has been studied by many researchers [2, 4, 5, 7, 13]. Among them, designers of Camellia tried to evaluate its security against differential attack[3] and linear attack[11] by showing the upper bounds of maximum differential characteristic probability (MDCP) and maximum linear characteristic probability (MLCP) for each reduced round Camellia. Since both of maximum differential probability and maximum linear probability of Camellia's S-boxes are 2^{-6}, the MDCP and MLCP are upper-bounded by $(2^{-6})^d$ and $(2^{-6})^l$, respectively, where d is the least number of differentially active S-boxes, and l is the least number of linearly active S-boxes. The designers modified a search method for truncated differential probability[9, 10, 12] to count the least number of active S-boxes and then searched d and l [2]. The search algorithm works fast because it treats one byte differences at once by using truncated differences [8]. They also showed that 12-round Camellia and 11-round Camellia without FL/FL^{-1} has no differential or linear characteristic with probability higher than 2^{-128}.

But at the FSE2001, Kanda pointed out that the search algorithm using truncated differential theory is not enough to evaluate the security of Camellia because it has SPN-type round function [5]. In SPS-type round function like E2[6], each split data in F-functions are substituted by the second substitution layer before output. But SPN-type round function doesn't have such a second substitution layer. Therefore, if a truncated differential is applied to SPN round function, a strong relation appears between each byte in the output of the F-function. By approximating these relationships, he showed the upper bounds of truncated differential probability of Camellia.

In this paper, we investigate these relationships more strictly. And we show a new evaluation method exploiting the relations all over the cipher to count the least number of active S-box more strictly. Then we apply the proposed evaluation method to Camellia, and we show that the upper bounds of MDCP and MLCP for Camellia are tighter than ever. We reveal that 10-rounds Camellia without FL/FL^{-1} has no differential or linear characteristic with probability higher than 2^{-128}.

This paper is organized as follows: In Section 2 we give the description of Camellia and the SPN-type round function treated in this paper. In Section 3 we use an example to show a contradiction in the truncated differential path generated by Camellia's evaluation method. In Section 4 we present a new method to evaluate the validity of truncated differential paths. In Section 5 we apply the proposed method to Camellia in practice, and show the revised upper bound of MDCP and MLCP for Camellia. Section 6 summarizes our conclusion.

2 Preliminaries

2.1 Description of SPN Round Function and Camellia

Camellia [1] is a Feistel network block cipher with block size of 128 bits and applicable to 128, 192 and 256 key bits. The round numbers are determined by a key length, 18 rounds for 128 bits key, 24 rounds for 192 and 256 bits keys. The F-function of Camellia is composed of a so-called SPN(Substitution Permutation Network) type structure. Fig 1 shows the i-th round of a Feistel cipher with SPN round function to be treated in this paper.

The block size is $2 \times m \times n$ bits. $m \times n$ bits of key and $m \times n$ bits of data are inputted to the F-function in each round. The input data of F-function is split into m pieces of n bits data, represented by $X_i = (X_i[1], X_i[2], \ldots, X_i[m])$. After the key adding operation, each data is inputted to non-linear bijective function $S : \{0,1\}^n \rightarrow \{0,1\}^n, (1 \leq j \leq m)$. Then the output of S-functions $S_i = (S_i[1], S_i[2], \ldots, S_i[m])$ are inputted to linear transformation layer $P : \{0,1\}^{mn} \rightarrow \{0,1\}^{mn}$. $Y_i = (Y_i[1], Y_i[2], \ldots, Y_i[m])$ represents an output vector of the P-function at the i-th round. Y_i is also an output of the F-function at the i-th round.

In general, any linear transformation P can be represented by a matrix form. In this paper, we specially concentrate on a linear transformation P which can be represented by a square matrix of order m over $GF(2^n)$ produced by some irreducible polynomial f defined in the cipher. Let $P_{mat} = (a_{ij})$ be a matrix

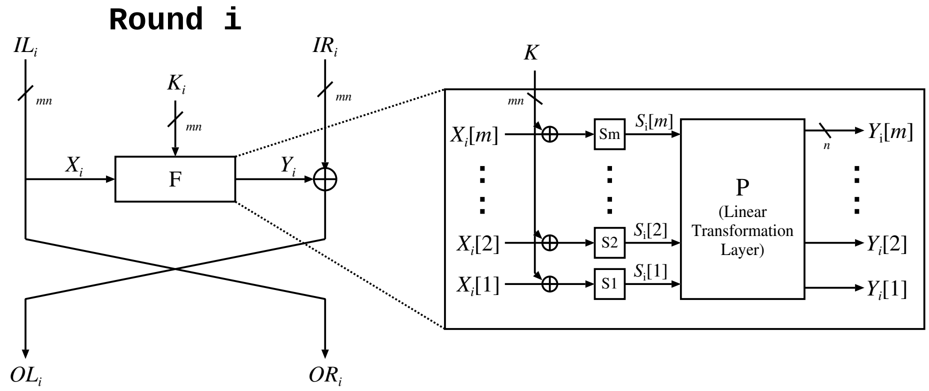

Fig. 1. Feistel Network with SPN round function

of order m over $GF(2^n)$ corresponding to the P-function. By using P_{mat}, the relation between S_i and Y_i can be described as follows:

$$\begin{pmatrix} Y[1] \\ Y[2] \\ \vdots \\ Y[m] \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix} \begin{pmatrix} S[1] \\ S[2] \\ \vdots \\ S[m] \end{pmatrix}$$

In the context of the matrix multiplication, each n bits of data are comprehended as a bit representation of an element in $GF(2^n)$. Throughout this paper, we treat n bits of data as a bit representation of an element in $GF(2^n)$. Thus the exclusive-or operation $\oplus$ and add operation $+$ of n bits of data are same.

In the case of Camellia, the round function can be viewed as $n = 8, m = 8$ SPN-type F-function, and the P-function is represented by the following matrix:

$$P_{Camellia} = \begin{pmatrix} 1&0&1&1&0&1&1&1 \\ 1&1&0&1&1&0&1&1 \\ 1&1&1&0&1&1&0&1 \\ 0&1&1&1&1&1&1&0 \\ 1&1&0&0&0&1&1&1 \\ 0&1&1&0&1&0&1&1 \\ 0&0&1&1&1&1&0&1 \\ 1&0&0&1&1&1&1&0 \end{pmatrix}$$

All elements in Camellia's matrix can be represented by only two elements, 0 and 1, in $GF(2^8)$.

Camellia also has a key scheduling algorithm, a key whitening layer and key-dependent linear functions FL and FL^{-1} which are inserted every 6 rounds. Details of these are described in [1]. In this paper, we assume that round keys are independent and uniformly random.

2.2 Definitions

We use the following definitions in this paper.

Definition 1. (Active S-box)
An S-box which has non-zero input difference is called **differentially active S-box**, *and an S-box which has non-zero output linear mask is called* **linearly active S-box**.

Definition 2. (χ function)
For any difference $\Delta X \in \{0,1\}^n$, a function $\chi : \{0,1\}^n \to \{0,1\}$ is defined as follows:

$$\chi(\Delta X) = \begin{cases} 0 \; if \;\; \Delta X = \{0\}^n \\ 1 \; if \;\; \Delta X \neq \{0\}^n \end{cases}$$

For any differential vector $\Delta X = (\Delta X[1], \Delta X[2], \ldots, \Delta X[m])$, $\Delta X[i] \in \{0,1\}^n$ truncated difference of ΔX is defined as

$$\delta X = \chi(\Delta X) = (\chi(\Delta X[1]), \chi(\Delta X[2]), \ldots, \chi(\Delta X[m]))$$

Definition 3. (truncated differential probability of F-function)*[5] Let $\delta X, \delta Y$ be input truncated difference and output truncated difference of F-function , respectively. Then the truncated differential probability $p_F(\delta X \to \delta Y)$ is defined as follows:*

$$p_F(\delta X \to \delta Y) = \frac{\displaystyle\sum_{\chi(\Delta X)=\delta X} \sum_{\chi(\Delta Y)=\delta Y} \Pr_{X \in (\{0,1\}^n)^m}[F(X) \oplus F(X \oplus \Delta X) = \Delta Y]}{\#\{\Delta X | \chi(\Delta X) = \delta X\}}$$

where $\#\{A\}$ denotes the number of elements in set A.

Specially, in the case of SPN-type round function defined above and the output difference of the S-function is assumed to be uniform[9, 10, 12], the truncated differential probability $p'_F(\delta X \to \delta Y)$ can be defined approximately as follows:

$$p'_F(\delta X \to \delta Y) = \frac{\#\{\Delta S | \chi(\Delta S) = \delta X, \chi(P(^t\Delta S)) = \delta Y\}}{\#\{\Delta S | \chi(\Delta S) = \delta X\}}$$

3 Wrong Truncated Differential Paths

The algorithm used in Camellia's evaluation counts the least number of active S-boxes for any round of Camellia[2]. Since the algorithm has exploited an approximation at the XOR operation in each round function, some of truncated differential paths generated by the algorithm cannot exist in reality.

Before proposing a new judgment method for the existence of truncated differential paths, we show an example of such wrong truncated differential paths and consider contradictions between truncated differences.

A truncated differential path is defined as follows:

Definition 4. A truncated differential path
For an r round Feistel block cipher, let δPR be a truncated difference of the right half of plain text and δCR be a truncated difference of the right half of cipher text. For each i-th round, let $\delta X_i, \delta Y_i$ be input truncated difference and output truncated difference of F-function, respectively. Then the truncated differential path TP is represented by:

$$TP = (\delta PR, \delta X_1, \delta Y_1, \ldots, \delta X_r, \delta Y_r, \delta CR) \in (\{0,1\}^m)^{2r+2}$$

The truncated difference of left half of plain text δPL and cipher text δCL can be represented by $\delta X_1, \delta X_r$, respectively.

For the i-th round, shown as Fig.1, let $\delta IL_i, \delta IR_i, (2 \leq i \leq r)$ be the left and right input truncated differences, respectively. And let δOL_i, $\delta OR_i, (1 \leq i \leq r-1)$ be left and right output truncated differences, respectively. Then there are following relations between these truncated differences and truncated differential path: $\delta IR_i = \delta X_{i-1}$, $\delta IL_i = \delta OR_i = \delta X_i$, $\delta OL_i = \delta X_{i+1}$.

3.1 Algorithm to Find Truncated Differential Paths

We show an outline of an algorithm used in the designer's evaluation to count the least number of active S-boxes of Camellia without FL/FL^{-1}[2]. This algorithm is modified version of Matsui's algorithm[9, 10] which is originally developed to estimates truncated differential probabilities of E2[6].

Algorithm
INPUT: a round number N
OUTPUT: a table of the least number of active S-boxes for all pattern of output truncated difference $\delta CR, \delta CL$ for N round Camellia without FL/FL^{-1}.

1. Make a table (F-table) of truncated differential probability of F-function for all δX and δY. The number of entries in the F-table is 2^{16}.
2. Using the F-table, make a table (R-table) of truncated differential probability of the round function for all $\delta IL, \delta IR$ and δCL. The number of entries in the R-table is 2^{24}.
3. (Inductive step) Using a table of the least number of active S-boxes for $N-1$ round of Camellia and R-table, calculate the least number of the active S-boxes for N round's output $\delta CL, \delta CR$ and store it in a table. The number of entries in the table is 2^{16}.

This algorithm only counts the least number of active S-boxes for each output truncated difference. To recover and to evaluate truncated differential paths which have the least number of active S-boxes, we have modified the algorithm additionally as follows.

- In Step 2, store all candidates of the output truncated difference of F-function.
- In Step 3, when the N round's output is searching, store all candidates of the $N-1$ round's output truncated difference which realize the least number of active S-boxes.

3.2 Connection Errors in Truncated Differential Paths

We show an example of a truncated differential path holding wrong properties. The search algorithm for the least number of active S-boxes found a Camellia's truncated differential path whose consecutive 3 rounds have truncated differential as in Fig. 2.

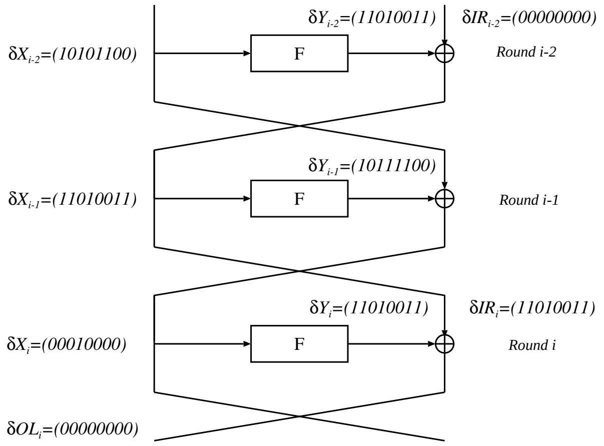

Fig. 2. A Example of Wrong Truncated Differential Path

The search algorithm exploits the following XOR rule of truncated difference based on a property of truncated difference.

XOR rule of truncated difference

$$\delta OL[j] = \begin{cases} 0 & if \quad \delta IR[j] = \delta Y[j] = 0 \\ * & if \quad \delta IR[j] = \delta Y[j] = 1 \\ 1 & \qquad else \end{cases} \qquad (1)$$

for $(1 \leq j \leq m)$. Value $*$ means arbitrary selection of $0, 1$.

Obeying above rule, the XOR operation of $\delta IR_i = (11010011)$ and $\delta Y_i = (11010011)$ generates $\delta OL_i = (**0*00**)$. Thus $\delta OL_i = (00000000)$ can be obtained by the search algorithm.

But by considering a relationship of each output difference of F-function δY_i and δY_{i-2}, we can prove such an XOR operation cannot be realized.

Proof. From $\delta OL_i = (00000000)$ and $\delta IR_{i-2} = (00000000)$, we get an equation of differences $\Delta Y_i = \Delta Y_{i-2}$. And from the linear relation $\Delta Y = P(\Delta S)$, then $\Delta S_i[j] = \Delta S_{i-2}[j]$. But considering given input truncated differences of S4 in both round, we get $\Delta S_i[4] \neq \Delta S_{i-2}[4]$ because $\delta X_i[4] = 1$ and $\delta X_{i-2}[4] = 0$. This is contradiction. Also in S1, S3, S5, S6, there are contradictions. Thus such a path cannot exist in reality. $\square$

From above observations, wrong truncated paths have some contradictions between input and output truncated differences of F-functions for every two round. In the next section, we propose an algorithm which determines if the target truncated differential path is wrong or not by checking solutions of the linear equations system constructed from the path.

4 Validity Checking Method for Truncated Differential Paths

We show a new algorithm to check the validity of truncated differential paths. This algorithm exploits a method of checking solutions of a linear equations system which is constructed from a truncated differential path. The basic steps of checking the validity of a truncated differential path is as follows:

- Represent all differences in Feistel network by linear forms using output difference of S-boxes
- Construct two linear equations systems and two sets of nonzero conditions using truncated differential values
- Check solutions of each linear equations system under corresponding nonzero conditions, and determine the validity of the truncated differential path.

4.1 Linear Forms Representation of Feistel Block Cipher

All differences in a Feistel cipher with SPN round function defined in Section 2 can be represented in linear form using differences $\Delta PL, \Delta PR$ and ΔS_i. To show this, we divide the Feistel block cipher in two parts **the left chain** and **the right chain** as shown in Fig.3.

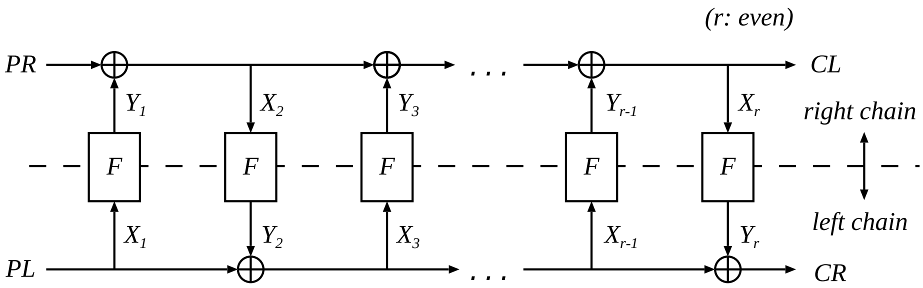

Fig. 3. The left chain and The right chain

Without loss of generality, we assume that the round number r is even.

The left chain is a data path which starts from left half of plain text PL, while the right chain starts from right half of plain text PR. In each chain, output of

F-functions Y_i's are added to the ongoing data. The left chain includes the data $X_{2i-1}, Y_{2i}, CR (1 \le i \le \frac{r}{2})$, The right chain includes $PR, Y_{2i-1}, X_{2i} (1 \le i \le \frac{r}{2})$. From these definitions we can show the following lemma.

Lemma 1.
Let $\Delta X_1[j](= \Delta PL[j]), \Delta PR[j] (1 \le j \le m)$ be variables which are set differences of plain text. And Let $\Delta S_i[j]$ be variables which are set differences of the output of j-th S-box Sj for i-th round $(1 \le j \le m, 1 \le i \le r)$, respectively. All differential data in the left chain are represented by linear form of elements in $\Delta X_i, \Delta S_{2i} (1 \le i \le \frac{r}{2})$. All differential data in the right chain are represented by linear form of elements in $\Delta PR, \Delta S_{2i-1} (1 \le i \le \frac{r}{2})$.

Proof. In the right chain, we show that $\Delta PR, \Delta Y_{2i-1} and \Delta X_{2i}$ are represented by linear form. ΔPR are monomials. And using elements in linear transformation matrix P_{mat}, differences of output of F-functions $\Delta Y_{2i-1}[j] (1 \le i \le r/2, 1 \le j \le m)$ can be described as following linear form:

$$\Delta Y_{2i-1}[j] = \sum_{k=1}^{m} a_{jk} \Delta S_{2i-1}[k] \tag{2}$$

And the difference of the input data of the $2i$-th round $\Delta X_{2i}[j] (1 \le i \le r/2, 1 \le j \le m)$ has the following equation:

$$\Delta X_{2i}[j] = \Delta PR[j] + \sum_{l=1}^{i} \Delta Y_{2l-1}[j] \tag{3}$$

By combining (2) and (3), $\Delta X_{2i}[j]$ has the following linear form:

$$\Delta X_{2i}[j] = \Delta PR[j] + \sum_{l=1}^{i} \sum_{k=1}^{m} a_{jk} \Delta S_{2l-1}[k] \tag{4}$$

All differential data in the left chain are also represented in same manner. □

From the viewpoint of linear form representation, the sets of variables used in both chains are completely separated. The left chain uses a set of differential variables $(\Delta X_1, \Delta S_2, \Delta S_4, \ldots, \Delta S_r)$, and the right chain uses a set of differential variables $(\Delta PR, \Delta S_1, \Delta S_3, \ldots, \Delta S_{r-1})$. Thus, we can treat both chains separately. Then we can describe the relation between the variables and the differences in the right chain as follows.

$$\begin{pmatrix} {}^t\Delta PR \\ {}^t\Delta Y_1 \\ {}^t\Delta X_2 \\ {}^t\Delta Y_3 \\ {}^t\Delta X_4 \\ \vdots \\ {}^t\Delta Y_{r-3} \\ {}^t\Delta X_{r-2} \\ {}^t\Delta Y_{r-1} \\ {}^t\Delta X_r \end{pmatrix} = \begin{pmatrix} I & 0 & 0 & \ldots & 0 & 0 \\ 0 & P_{mat} & 0 & \ldots & 0 & 0 \\ I & P_{mat} & 0 & \ldots & 0 & 0 \\ 0 & 0 & P_{mat} & \ldots & 0 & 0 \\ I & P_{mat} & P_{mat} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & P_{mat} & 0 \\ I & P_{mat} & P_{mat} & \ldots & P_{mat} & 0 \\ 0 & 0 & 0 & \ldots & 0 & P_{mat} \\ I & P_{mat} & P_{mat} & \ldots & P_{mat} & P_{mat} \end{pmatrix} \begin{pmatrix} {}^t\Delta PR \\ {}^t\Delta S_1 \\ {}^t\Delta S_3 \\ \vdots \\ {}^t\Delta S_{r-3} \\ {}^t\Delta S_{r-1} \end{pmatrix}$$

I denotes a unit matrix of order m, 0 in the above matrix denotes zero matrix of order m. The total number of linear forms are $m(r+1)$, and the total number of variables are $m(r/2+1)$.

The data in left chain $(\Delta X_1, \Delta Y_2, \Delta X_3, \ldots, \Delta Y_r, \Delta CR)$ can be expressed by using $(\Delta X_1, \Delta S_2, \Delta S_4, \ldots, \Delta S_r)$ in the same way. Thus, all differences in the Feistel cipher with SPN round function defined in this paper have been represented by linear forms.

4.2 Evaluation Algorithm for Truncated Differential Paths

Let TP be a r round truncated differential path. To evaluate TP, we check whether there exist any differential characteristic which follow the truncated differential path TP.

For the right chain, truncated differences $\delta X_1, \delta X_3, \ldots, \delta X_{r-1}$ in TP determine whether S-boxes, whose outputs are added to the right chain, are active or not. If truncated difference $\delta X_i[j] = 0$ the output differences of j-th S-box in round i is always 0. Thus we can remove differential variables of such non active S-boxes from all linear forms. And if $\delta PR[i] = 0$, we can also remove the differential variable $\Delta PR[i]$ in linear forms. In this way, we obtain reduced linear forms.

Next we construct linear equations systems and sets of linear conditions. In the right chain, truncated differences $\delta Y_1, \delta Y_3, \ldots, \delta Y_{r-1}$ determine whether differential data of F-function's output have a differential value or not. Let $lf_{Y_i[j]}$ be a reduced linear form of differential data $\Delta Y_i[j]$. If $\delta Y_i[j] = 0$, we get a linear equation $lf_{Y_i[j]} = 0$. And if $\delta Y_i[j] = 1$, we get a nonzero condition $lf_{Y_i[j]} \neq 0$. The other reduced linear forms $\Delta PR, \Delta X_2, \Delta X_4, \ldots, \Delta X_r$ in the right chain can be sorted according to their truncated differential values in same way. Letting z be the total number of 0's in truncated differences $\delta PR, \delta Y_{2i-1}, \delta X_{2i}, (1 \leq i \leq r/2)$, we get a linear homogeneous equations system which has z equations and a set of nonzero conditions which has $m(r+1) - z$ conditions for the right chain.

Definition 5. *Let $\mathcal{F}$ be a r round Feistel cipher with SPN round function defined in Section 2, and let TP be a truncated differential path of $\mathcal{F}$. We define LES_{Right}, LES_{Left} as equation systems made of TP for the right chain and the left chain, respectively. And we define NC_{Right}, NC_{Left} as sets of nonzero conditions made of TP for the right chain and the left chain, respectively.*

Then we evaluate both chains by using $LES_{Right}, LES_{Left}, NC_{Right}, NC_{Left}$. Moreover, we use the following useful definition.

Definition 6. (reduced row-echelon matrix) *A matrix is a reduced row-echelon matrix if*

- *All rows of zero (if exists) are at the bottom of the matrix.*
- *The first nonzero number in a row is a 1 (leading 1).*
- *Each leading 1 is to the right of the leading 1's in the rows above it.*
- *Each column that contains a leading 1 has zeros everywhere else.*

Any matrix can be transformed into an unique reduced row-echelon matrix by performing a finite sequence of elementally row operations (sweep out method). Letting E_M be the set of all matrices which are obtained by performing elementally row operations on M, every matrix in E_M has the same reduced row-echelon matrix M'.

The following important properties related to a reduced row-echelon matrix and a homogeneous linear equations system are obtained.

Property 1. (partial solution) Let $\boldsymbol{Mx} = \mathbf{0}$ be a matrix and vector representation of a linear homogeneous equations system with a vector of variables $\boldsymbol{x} = {}^t(x_1, x_2, \ldots, x_k)$. And let $\boldsymbol{M'}$ be a reduced row-echelon matrix obtained from $\boldsymbol{M}$. If there is any row containing only a leading 1 in $\boldsymbol{M'}$. Letting i be the column index of the leading 1, the system has a solution $x_i = 0$.

Property 2. (property of kernel) Let $\boldsymbol{Mx} = \mathbf{0}$ be a matrix and vector representation of a linear homogeneous equations system. And let $Z = \{\boldsymbol{z}|\boldsymbol{Mz} = \mathbf{0}\}$ be a set of solutions. And let $\boldsymbol{M'}$ be a reduced row-echelon matrix obtained from $\boldsymbol{M}$. Let $\boldsymbol{v_1, \ldots, v_r}$ be each nonzero row vector in $\boldsymbol{M'}$. Let S be a subspace spanned by $\boldsymbol{(v_1, \ldots, v_r)}$. Then $\boldsymbol{v_S} \in S, \boldsymbol{z} \in Z$ satisfies $\boldsymbol{v_S} \cdot \boldsymbol{z} = 0$. ($\cdot$ denotes inner products of vectors.)

Using the above properties we construct an evaluation algorithm as follows.

Algorithm
INPUT: $LES_{Right}, NC_{Right}, LES_{Left}, NC_{Left}$
OUTPUT: "wrong path" or "OK"

1. Represent linear equations system LES_{Right} in matrix and vector form as follows:
$$\boldsymbol{M_{Right}}\ \boldsymbol{x_{Right}} = \mathbf{0}$$
2. Perform elementary row operations on matrix $\boldsymbol{M_{Right}}$, and get a reduced row-echelon matrix $\boldsymbol{M'_{Right}}$
3. If there is any row containing only leading 1 in $\boldsymbol{M'_{Right}}$, output "wrong path"; else go to the next step.
4. For each nonzero condition $nc \in NC_{Right}$ do the following:
   (a) Represent nonzero condition nc by an inner product of vectors: $\boldsymbol{n} \cdot \boldsymbol{x_{Right}}$
   (b) Let $\boldsymbol{v_1, \ldots, v_r}$ be nonzero row vectors in $\boldsymbol{M'_{Right}}$. If vector $\boldsymbol{n}$ can be represented by a linear combination of $\boldsymbol{v_1, \ldots, v_r}$,
   $$\boldsymbol{n} = c_1\boldsymbol{v_1} + \ldots + c_r\boldsymbol{v_r}$$
   where $c_i \in GF(2^n), (1 \leq i \leq r)$,
   then output "wrong path"; else go to the next step.
5. For LES_{Left} and NC_{Left}, apply the same steps Step 1. $\sim$ Step 4.
6. Output "OK"

In Step 3. if such a row containing only a leading 1 exists, the corresponding variable in $\boldsymbol{x_{Right}}$ has to be always 0 (property 1). But all variables in $\boldsymbol{x_{Right}}$ have been selected to have a nonzero value, the solution cannot be realized.

In Step 4.(b) the check whether vector $\boldsymbol{n}$ can be represented as a linear combination can be achieved by easy row operations of the matrix because $\boldsymbol{v_1, \ldots, v_r}$ are linearly independent and the vectors have the 4th condition of row echelon matrix. And if such a linear combination exists, nc must have difference 0 in spite of nonzero condition (property 2). This contradiction means the path is wrong.

4.3 Complexity

Let r be the round number of truncated differential path, and let m be a number of split data in F-function. Each chain has a $m(r+1) \times mr$ matrix of linear forms. If LES_{Right} has x rows, then NC_{Right} has $m(r+1) - x$ rows. Complexity to obtain reduced row-echelon matrix of LES_{Right} is proportional to xm^2r^2. And complexity to check NC_{Right} rows is proportional to $mrx(m(r+1) - x)$. Since we assume that $x \approx mr/c$ with some constant c, the total complexity of the proposed method is $O(m^3r^3)$.

4.4 Expansion to Truncated Linear Paths

The proposed method is also applicable to evaluation of truncated linear paths. Let ΓS be an input linear mask and let ΓY be an output linear mask of linear transformation layer P. Then there is a diffusion function P^* satisfying $\Gamma S = P^*(\Gamma Y)$ determined by P. If the diffusion function P^* can be expressed as a square matrix of order m on $GF(2^n)$, the proposed algorithm can be applied to truncated linear paths of the Feistel cipher exploiting the dual property between differential and linear mask[5].

In case of Camellia, matrix P^* can be expressed as ${}^tP_{Camellia}$ which is a square matrix of degree m on $GF(2^n)$. Thus, our method can be applied to evaluation for both truncated differential and linear paths for Camellia.

5 Evaluation for Camellia

We applied the proposed method to block cipher Camellia to reevaluate the upper bound of MDCP and MLCP. To obtain corrected values of the least numbers of active S-boxes, we also extracted truncated paths which have more than the least number of active S-boxes searched by the algorithm used in designer's (previous) evaluation. We used the following definitions and lemma to extract such truncated paths.

Definition 7. *Let $L_{act}(C, R)$ be the least number of active S-boxes of R round truncated path with output truncated difference C in the previous evaluation.*

The algorithm for counting the least number of active S-boxes outputs $L_{act}(C,R)$ for all $C \in \{0,1\}^{16}$ with any round number R.

Definition 8. *Let* $Cipher(ACT,R) = \{c \mid L_{act}(c,R) = ACT\}$. *And let* $Paths(C,ACT,R)$ *be a set of all* R *round truncated differential paths which have* ACT *active S-boxes with output truncated difference* C *in the previous evaluation. And let* $Paths(ACT,R)$ *be a set of all* R *round truncated differential paths which have* ACT *active S-boxes in the previous evaluation.*

Lemma 2. *Letting* α *be the least number of active S-boxes for* r *round Camellia, truncated differential paths which have* $\beta(\geq \alpha)$ *active S-boxes can be written as follows:*

$$Paths(\beta,r) = \{p \mid p \in Paths(c,\beta,r), c \in \bigcup_{i=\alpha}^{\beta} Cipher(i,r)\} \tag{5}$$

Proof. Truncated differential paths which have β active S-boxes are also written as follows:

$$Paths(\beta,r) = \{p \mid p \in Paths(c,\beta,r), c \in \{0,1\}^{16}\} \tag{6}$$

Let $D = \bigcup_{i=\beta+1}^{\infty} Cipher(i,r)$. For $d \in D$, we get following relations:

$$Paths(d,\beta,r) = \emptyset \tag{7}$$

Thus, we can reduce the candidates for c in (6) then obtain the equation (5). □

By using the above lemma, we obtain truncated paths which have variable number of active S-boxes. To obtain $Paths(c,\beta,r)$ of $c \in Cipher(\gamma,r)$ where $\beta > \gamma$, we extract paths by considering replacement of internal r' round paths $Paths(c',\gamma',r')$ in $Paths(c,\gamma,r)$ as subset of $Paths(\gamma'+\beta-\gamma,r')$, which are connectable to $r'+1$ round's output, for $(1 \leq r' \leq r-1)$.

5.1 Result

We show the result of our re-estimation of the upper bound of MDCP and MLCP in Table 1∼3.

Tables show the evaluation results of truncated differential paths and truncated linear paths, respectively. The first row of each table indicates round number of reduced round Camellia without FL/FL^{-1}(we call it Camellia*). The rows below the first are divided into three parts. The upper part, the middle part and the lower part include information about the least number, the secondary least number, and the tertiary or the tertiary and the forth least number of active S-boxes, respectively. Each part has three rows. The first one indicates the number of active S-boxes, the second one indicates the number of truncated differential or linear paths which are evaluated to have the above number of active S-boxes by the previous(old) evaluation method. The last row of each part

Round	1	2	3	4	5	6	7	8	9	10	11	12
Active S-box	0	1	2	7	9	11	13	15	18	21	22	25
Path Number	255	16	8	1368	160	8	16	48	48	320	4	64
OK	255	16	8	1328	80	0	0	0	0	0	0	0
Active S-box	1	2	3	8	10	12	14	16	19	22	23	26
Path Number	19440	292	0	9817	5564	3136	1160	368	2428	14732	224	1736
OK	19440	292	0	7431	4408	1008	196	36	0	156	0	0
Active S-box	2	3	4	9	11	13	15	17	20	23	24	27
Path Number	182228	5000	112	114256	66624	41576	16748	4360	42456	302784	4784	38672
OK	182228	5000	112	72432	17772	4336	1568	328	812	2176	16	0

Table 1. Check Result of Truncated Differential Paths (1∼12 rounds)

Round	1	2	3	4	5	6	7	8	9	10	11	12
Active S-box	0	1	2	6	9	11	13	14	18	20	22	25
Path Number	255	16	8	64	480	8	16	4	180	16	4	128
OK	255	16	8	16	240	0	0	0	0	0	0	0
Active S-box	1	2	3	7	10	12	14	15	19	21	23	26
Path Number	19440	292	0	1952	8732	4540	1208	24	4968	1040	448	3188
OK	19440	292	0	1656	4864	1948	192	0	144	0	0	0
Active S-box	2	3	4	8	11	13	15	16/17	20	22	24	27
Path Number	182228	4960	112	9733	76820	52072	21956	184/8142	63128	28640	10640	83800
OK	182228	4960	112	7327	26748	5680	5920	0/1560	4992	792	36	0

Table 2. Check Result of Truncated Linear Paths (1∼12 rounds)

indicates the number of truncated differential or linear paths which are evaluated as “OK” by proposed method.

In 6 $\sim$ 12 rounds Camellia*, truncated differential and linear paths which are evaluated to have the least number of active S-boxes are all evaluated as wrong. At some rounds, all truncated paths with the secondary or the tertiary number of active S-boxes are evaluated as wrong.

NEW

	Differential		Linear	
	Camellia	Camellia*	Camellia	Camellia*
1	$1_{[0]}$	$1_{[0]}$	$1_{[0]}$	$1_{[0]}$
2	$2^{-6}{}_{[1]}$	$2^{-6}{}_{[1]}$	$2^{-6}{}_{[1]}$	$2^{-6}{}_{[1]}$
3	$2^{-12}{}_{[2]}$	$2^{-12}{}_{[2]}$	$2^{-12}{}_{[2]}$	$2^{-12}{}_{[2]}$
4	$2^{-42}{}_{[7]}$	$2^{-42}{}_{[7]}$	$2^{-36}{}_{[6]}$	$2^{-36}{}_{[6]}$
5	$2^{-54}{}_{[9]}$	$2^{-54}{}_{[9]}$	$2^{-54}{}_{[9]}$	$2^{-54}{}_{[9]}$
6	$\star 2^{-72}{}_{[12]}$	$\star 2^{-72}{}_{[12]}$	$\star 2^{-72}{}_{[12]}$	$\star 2^{-72}{}_{[12]}$
7	$\star 2^{-78}{}_{[13]}$	$\star 2^{-84}{}_{[14]}$	$\star 2^{-78}{}_{[13]}$	$\star 2^{-84}{}_{[14]}$
8	$\star 2^{-78}{}_{[13]}$	$\star 2^{-96}{}_{[16]}$	$\star 2^{-78}{}_{[13]}$	$\star 2^{-102}{}_{[17]}$
9	$\star 2^{-84}{}_{[14]}$	$\star 2^{-120}{}_{[20]}$	$\star 2^{-84}{}_{[14]}$	$\star 2^{-114}{}_{[19]}$
10	$\star 2^{-114}{}_{[19]}$	$\star 2^{-132}{}_{[22]}$	$\star 2^{-108}{}_{[18]}$	$\star 2^{-132}{}_{[22]}$
11	$\star 2^{-126}{}_{[21]}$	$\star 2^{-144}{}_{[24]}$	$\star 2^{-126}{}_{[21]}$	$\star 2^{-144}{}_{[24]}$
12	$\star 2^{-144}{}_{[24]}$	–	$\star 2^{-144}{}_{[24]}$	–

OLD

	Differential		Linear	
	Camellia	Camellia*	Camellia	Camellia*
1	$1_{[0]}$	$1_{[0]}$	$1_{[0]}$	$1_{[0]}$
2	$2^{-6}{}_{[1]}$	$2^{-6}{}_{[1]}$	$2^{-6}{}_{[1]}$	$2^{-6}{}_{[1]}$
3	$2^{-12}{}_{[2]}$	$2^{-12}{}_{[2]}$	$2^{-12}{}_{[2]}$	$2^{-12}{}_{[2]}$
4	$2^{-42}{}_{[7]}$	$2^{-42}{}_{[7]}$	$2^{-36}{}_{[6]}$	$2^{-36}{}_{[6]}$
5	$2^{-54}{}_{[9]}$	$2^{-54}{}_{[9]}$	$2^{-54}{}_{[9]}$	$2^{-54}{}_{[9]}$
6	$2^{-66}{}_{[11]}$	$2^{-66}{}_{[11]}$	$2^{-66}{}_{[11]}$	$2^{-66}{}_{[11]}$
7	$2^{-72}{}_{[12]}$	$2^{-78}{}_{[13]}$	$2^{-72}{}_{[12]}$	$2^{-78}{}_{[13]}$
8	$2^{-72}{}_{[12]}$	$2^{-90}{}_{[15]}$	$2^{-72}{}_{[12]}$	$2^{-84}{}_{[14]}$
9	$2^{-78}{}_{[13]}$	$2^{-108}{}_{[18]}$	$2^{-78}{}_{[13]}$	$2^{-108}{}_{[18]}$
10	$2^{-108}{}_{[18]}$	$2^{-126}{}_{[21]}$	$2^{-102}{}_{[17]}$	$2^{-120}{}_{[20]}$
11	$2^{-120}{}_{[20]}$	$2^{-132}{}_{[22]}$	$2^{-120}{}_{[20]}$	$2^{-132}{}_{[22]}$
12	$2^{-132}{}_{[22]}$	–	$2^{-132}{}_{[22]}$	–

Note: The numbers in brackets are the number of active S-boxes.

Table 3. Revised Upper bounds MDCP and MLCP

The left side of table 3 shows new upper bounds of MDCP and MLCP for Camellia and Camellia* while the right side shows the old result in [2]. Letting a be the number of active S-boxes, each probability is calculated by $(2^{-6})^a$. In NEW table, $\star$ denotes the upper bounds which are updated. All upper bounds of more than 6 round Camellia and Camellia* are as $2^{-6} \sim 2^{-18}$ times low as previous upper bounds. The horizontal lines between upper bounds denotes probability 2^{-128}. Thus, we obtain the new result that any characteristics of 10 rounds Camellia* cannot be distinguishable from random permutations.

6 Conclusions and Future Research

In this paper we proposed a new algorithm to evaluate truncated differential/linear paths whether they are wrong or not. This algorithm is applicable to Feistel ciphers whose data are represented by linear forms of the output difference of S-boxes and the difference of plain text. By applying the proposed algorithm to evaluate the security of Camellia, we found tighter upper bounds of MDCP and MLCP because the least number of active S-boxes are updated. It is revealed that Camellia has a stronger immunity against differential attack and linear attack than before.

Note that even though a truncated differential path is not judged as a wrong path by proposed algorithm, it is not guaranteed that the path exists in reality. Because we have only excluded wrong paths in the context of truncated differential, we haven't taken into consideration of the variety of output differences of active S-boxes for fixed input difference. The direction of next research contains to find the MDCP, MLCP and paths which realize them by combining our proposed evaluation method and the properties of S-boxes.

References

1. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima and T. Tokita, "Specification of Camellia - a 128-bit Block Cipher," submitted to the First Open NESSIE Workshop, 13-14 November 2000, Leuven, Belgium -available at http://cryptonessie.org.
2. 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 Area in Cryptography, SAC 2000, LNCS2012, pp.39-56, 2000.
3. E. Biham and A. Shamir," Differential Cryptanalysis of DES-like Cryptosystems," CRYPTO '90, LNCS 537, pp.2-21, 1991.
4. Y. He and S. Qing,"Square Attack on Reduced Camellia Cipher," Information and Communications Security, ICICS 2001, LNCS 2229, pp.238-245, 2001.
5. M. Kanda and T. Matsumoto, "Security of Camellia against Truncated Differential Cryptanalysis," Fast Software Encryption, FSE2001, to appear
6. M. Kanda, S. Moriai, K. Aoki, H. Ueda, Y. Takashima, K. Ohta and T. Matsumoto, "E2- A New 128-Bit Block Cipher," IEICE Transactions Fundamentals of Electronics, Communications and Computer Sciences, Vol. E83-A, No.1, pp.48-59, 2000.
7. T. Kawabata and T.Kaneko, "A study on higher order differential attack of Camellia," In Proceedings of the 2nd NESSIE workshop, 2001.
8. L.R.Knudsen, "Truncated and Higher Order Differentials," Fast Software Encryption - Second International Workshop, LNCS 1008, pp.196-211, 1995.
9. M.Matsui, "Differential Path Search of the Block Cipher E2," Technical Report ISEC99-19, IEICE, 1999.(written in Japanese)
10. M. Matsui and T. Tokita,"Cryptanalysis of Reduced Version of the Block Cipher E2," Fast Software Encryption, FSE'99, LNCS 1636, 1999.
11. M. Matsui, "Linear Cryptanalysis of the Data Encryption Standard," EUROCRYPT '93, LNCS 765, pp.386-397, 1994.
12. S. Moriai, M. Sugita, K. Aoki and M. Kanda, "Security of E2 against Truncated Differential Cryptanalysis," Selected Areas in Cryptography, SAC'99, LNCS 1758, pp.106-117, 2000.
13. M. Sugita, K. Kobara and H. Imai,"Security of Reduced Version of the Block Cipher Camellia against Truncated and Impossible Differential Cryptanalysis," ASISCRYPT 2001, LNCS 2248, pp.193-207.
14. New European Schemes for Signatures, Integrity, and Encryption, http://www.cryptonessie.org
15. CRYPTREC project, http://www.ipa.go.jp/security/enc/CRYPTREC/.