The Improved 96th-Order Differential Attack on 11 Rounds of the Block Cipher CLEFIA

Yasutaka Igarashi, Seiji Fukushima, and Tomohiro Hachino
Kagoshima University, Kagoshima, Japan
Email: {igarashi, fukushima, hachino}@eee.kagoshima-u.ac.jp

Toshinobu Kaneko
Tokyo University of Science, Chiba, Japan
Email: kaneko@ee.noda.tus.ac.jp

Abstract—CLEFIA is a 128-bit block cipher proposed by Shirai of SONY et al. in 2007. Its key size is 128, 192, or 256 bits. The number of the round of data processing part depends on a key size, viz. it is 18, 22, or 26 rounds for 128, 192, or 256 bits of a key size, respectively. Such a characteristic of CLEFIA have been known that the 96th-order differential of 64 bits out of 128 bits of the 8th-round’s output is zero. With this characteristic, we reported the 96th-order differential attack on 11 rounds of CLEFIA that requires $2^{98.3}$ blocks of plain text and $2^{159}$ times of data encryption. In this paper, we reduce this number of the times of the encryption, (viz. computational complexity) by applying a partial sum technique proposed by Ferguson et al. With the technique, we sequentially derive a modulo 2 occurrence distribution of intermediate data of cryptanalysis. We also reduce the complexity by introducing a nested structure of iterative computations to the attack algorithm. As a result we reduce the complexity to $2^{106.6}$, which is 1/223.0 of the conventional complexity.

Index Terms—cryptanalysis, higher-order differential attack, block cipher, CLEFIA

I. INTRODUCTION

CLEFIA is a 128-bit block cipher proposed by Shirai of SONY et al. in 2007 [1], [2]. A data processing part of CLEFIA consists of some rounds of 4-branch type-2 generalized Feistel structure, where two similar nonlinear functions are placed in parallel. CLEFIA supports 128, 192, and 256 bits of secret keys. The number of round of data processing part depends on a key size, viz. it is 18, 22, or 26 rounds for 128, 192, or 256 bits of secret key, respectively. CLEFIA has been applied and evaluated for cryptographic techniques towards the revision of the e-Government Recommended Ciphers List in FY 2013 in Japan [3].

It has been known that the synchronous 8th-order differential of 24 bits out of 128 bits becomes zero at the 6th-round output of the data processing part [4]. It has been also known that the 96th-order differential of 64 bits out of 128 bits becomes zero at the 8th-round output [5]. Exploiting this property we reported the 96th-order differential attack [6] on 11 rounds of CLEFIA with $2^{98.3}$ blocks of plain text and $2^{159}$ times of data processing [7]. In the article we reduce the number of times of data processing for the 96th-order differential attack by using a modulo 2 occurrence distribution (MOD) of the intermediate data of data processing part. MOD is derived by using a partial sum technique proposed by Ferguson et al. [8]. Moreover we reduce the number of the times by applying a nested structure of iterative computation to the attack algorithm and optimizing the order of the nest.

As a result we show that the 96th-order differential attack can be performed with $2^{98.3}$ blocks of plain text and $2^{106.6}$ times of data processing, which is $1/2^{52.4}$ of computational complexity compared to the conventional attack. Note that the best known higher-order differential attack on CLEFIA is the 105th-order differential attack on 14 rounds of CLEFIA with $2^{108.3}$ blocks of plain text and $2^{223.0}$ times of data processing so far [9].

II. DATA PROCESSING PART OF CLEFIA

In this section we describe the data processing part of CLEFIA. We omit the detail specification of CLEFIA [2] being not required for the article.

Fig. 1 shows an 11-round data processing part of CLEFIA where input plain text and output cipher text are 128 bits represented by $X_i$ and $C^{ji}(i = 0, 1, 2, 3)$, respectively. A bit size of $X_i$ and $C^{ji}(j = 1, 2, \ldots, 11)$ represents 32-bit output of the $j$th round given by

$$
C^{ji} = (C_{ij}^{(0)} || C_{ij}^{(1)} || C_{ij}^{(2)} || C_{ij}^{(3)})^T
$$

(1)

where $C^{(k)}(k = 0, 1, 2, 3)$ is an 8-bit data. The superscript $T$ represents transposition of a vector or a matrix. The symbol $\oplus$ represents an XOR operation. $WK_i(i = 0, 1)$ is a 32-bit whitening key, and $RK_i(i = 0, 1, 2, \ldots, 21)$ is a 32-bit round key used in $F_i(i = 0, 1)$. $F_i$ is a bijective nonlinear function with 32-bit input/output (I/O) adopting SP structure. 4-branch 32-bit data $X$ is mixed by XOR and substituted by a nonlinear function. Such a structure is called 4-branch generalized Feistel.
Fig. 1. 11-round data processing part of CLEFIA.

Fig. 2. (a) $F_0$ and (b) $F_1$.

Fig. 2 shows (a) $F_0$ and (b) $F_1$ where its input and output data are represented by an 8-bit $x$ and an 8-bit $y$ ($i = 0, 1, 2, 3$), respectively. $RK_{ki}$ ($i = 0, 1, 2, 3$) is an 8-bit key satisfying the following equation: $RK_i = (RK_{i,0}, RK_{i,1}, RK_{i,2}, RK_{i,3})^T$. $S_0$ and $S_1$ are bijective nonlinear functions with 8-bit I/O. Their output is represented by $x'_i$. $M_0$ and $M_1$ are 4x4 nonsingular matrices given by

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} = M_i \begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} \quad (i = 0, 1),$$

where “$a$” is a hexadecimal number. The multiplications of a vector and a matrix are performed in $GF(2^8)$ defined by the primitive polynomial $z^8 + z^4 + z^3 + z^2 + 1$.

III. THE 96TH-ORDER DIFFERENTIAL CHARACTERISTICS OF CLEFIA AND ITS ATTACK EQUATION

In this section we show the 96th-order differential characteristics of CLEFIA and its attack equation.

First we focus on an input plain text (see Fig. 1). We set $X_0$ or $X_2$ to take arbitrary constant, and input the 96th-order differential to the three remaining $X_i$. Namely, all the $2^{96}$ kinds of data from 0 to $2^{96} - 1$ are put into the three $X_i$. At this time the 96th-order differentials of $C_{10}$ and $S_0(C_{00} \oplus RK_{16,0})$ becomes the same value [7] (see Fig. 3 that shows the equivalently modified data processing part from the 9th round to the 11th round). $M_0^{-1}$ is the inverse matrix of $M_0$ that is given by $M_0^{-1} = M_0$. There is the product of the matrices $M_0^{-1}$ and $M_1$ at the right side of the 11th round that is given by

$$M_0^{-1} \cdot M_1 = \begin{pmatrix} 25 & 2e & 22 & 28 \\ 2e & 25 & 28 & 22 \\ 22 & 28 & 25 & 2e \\ 28 & 22 & 2e & 25 \end{pmatrix}$$

By using the 96th-order differential characteristics the following attack equations can be derived:

$$\sum_{x' \oplus y'} S_i(C_{00} \oplus RK_{16,0}) = \sum_{x' \oplus y'} C_{10}^{(9)}$$

(6)

$$C_{10}^{(9)} = C_{00}^{(10)} = W_{20}^{(11)} \oplus 28 S_2(C_{23}^{(11)} \oplus RK_{21,1})$$

(7)

$$W_{20}^{(11)} = W_{20}^{(10)} \oplus 22 S_1(C_{22}^{(11)} \oplus RK_{21,2})$$

(8)

$$W_{00}^{(11)} = W_{00}^{(10)} \oplus 2e S_1(C_{21}^{(11)} \oplus RK_{21,1})$$

(9)

$$W_{20}^{(11)} = 25 S_1(C_{20}^{(11)} \oplus RK_{21,0})$$

(10)

$$W_{20}^{(11)} = 2C_3^{(11)} + 25 S_1(C_{21}^{(11)} \oplus RK_{21,1})$$

(11)

$$W_{00}^{(11)} = W_{20}^{(10)} \oplus a S_1(C_{23}^{(10)} \oplus RK_{19,3})$$

(12)

$$W_{20}^{(10)} = W_{10}^{(10)} \oplus 2 S_1(C_{22}^{(10)} \oplus RK_{19,2})$$

(13)

$$W_{00}^{(10)} = W_{00}^{(10)} \oplus 8 S_0(C_{20}^{(10)} \oplus RK_{19,3})$$

(14)

$$W_{00}^{(10)} = C_{30}^{(10)} \oplus S_0(C_{20}^{(10)} \oplus RK_{19,3})$$

(15)
\[ C_2^{(10)} = U_2^{(11)} \oplus (6,4,2,1)^7 S_1(C_{03}^{(11)} \oplus RK_{20,3}) \quad (16) \]

\[ U_2^{(11)} = U_4^{(11)} \oplus (4,6,1,2)^7 S_0(C_{02}^{(11)} \oplus RK_{20,2}) \quad (17) \]

\[ U_i^{(11)} = U_0^{(11)} \oplus (2,1,6,4)^7 S_i(C_{01}^{(11)} \oplus RK_{20,1}) \quad (18) \]

\[ U_0^{(11)} = C_i^{(11)} \oplus (1,2,4,6)^7 S_0(C_{00}^{(11)} \oplus RK_{20,0}) \quad (19) \]

\[ U_{i,j}^{(j)} = (U_{i,j}^{(j)})^7, U_{i,j}^{(j)}, U_{i,j}^{(j)}, U_{i,j}^{(j)} \]

\[ W_{i,j}^{(j)} = (W_{i,j}^{(j)})^7, W_{i,j}^{(j)}, W_{i,j}^{(j)}, W_{i,j}^{(j)} \]

\[ C^{(j)}(X \oplus \Delta X) = \begin{cases} C_0^{(j)} \\ C_1^{(j)} \\ C_2^{(j)} \\ C_3^{(j)} \end{cases} \quad (22) \]

where

\[ C^{(j)}(X \oplus \Delta X) \] is the \( j \)-th-round 128-bit output corresponding to the input plain text \((X \oplus \Delta X)\) where \( \Delta X \) is an input difference. \( U_{i,j}^{(j)} \) and \( W_{i,j}^{(j)} \) are 32-bit intermediate data of the equivalent circuit shown in Fig. 4 and Fig. 5 where the hexadecimal number in a box represents multiplication. Because (6) is 8 sets of Boolean equation, it holds with probability \( 2^{-8} \) even if a guessed key is false. On the other hand it holds with probability 1 if the guessed key is correct. Therefore attacker can recover the total 104 bits of the round keys, \( RK_{16,0}, RK_{19,0}, RK_{20,0}, \) and \( RK_{21,0} \) in (6)-(22) by analyzing (6)-(22) for sufficient number of times.

IV. ATTACK ALGORITHM AND COST ESTIMATION

In this section we show the attack algorithm to recover the total 104 bits of round keys, and estimate the cost of the attack. We reduce the computational complexity of the attack by exploiting MOD for intermediate data of (6)-(22), which is sequentially derived by using a partial sum technique proposed by Ferguson et al. [8]. The advantages of using MOD are as follows. Even number of XOR operations of a certain variable \( x \) is zero, while odd number of them is \( x \). Therefore even number of XOR operations of \( x \) becomes unnecessary and odd number of them can be substituted with \( x \) by using MOD. These advantages result in complexity reduction.
Next we show the attack algorithm exploiting MOD as follows.

A. Attack Algorithm

0) Make MOD of the total 72-bit data, $C_{00}^{(11)}$, $C_{01}^{(11)}$, $C_{20}^{(11)}$ (name as MOD1) and MOD of the total 40-bit data, $C_{21}^{(11)}$ and $C_{30}^{(11)}$ (name as MOD2) from the 2^{96} blocks of cipher text $c_{i}^{(11)}(X@\Delta X)$ where $\Delta X \in V^{96}$. The number of elements of MOD1 is at most 2^{52} and the average is 2^{47}. The number of elements of MOD2 is at most 2^{40} and the average is 2^{39}.

1) Guess 8-bit $RK_{00,0}$, and make MOD of the total 64-bit data, $C_{00}^{(11)}$, $C_{01}^{(11)}$, $C_{02}^{(11)}$, $C_{03}^{(11)}$, $U_{0}^{(11)}$, and $C_{30}^{(10)}$ (MOD3) from MOD1 through the equivalent circuit shown in Fig. 4 with at most 2^{52} times and the average 2^{47} times of $S_{0}$ operation and the following multiplications, which corresponds to (19). The number of elements of MOD3 is at most 2^{60} and the average is 2^{56}.

2) Guess 8-bit $RK_{20,1}$, and make MOD of the total 56-bit data, $C_{02}^{(11)}$, $C_{03}^{(11)}$, $U_{1}^{(11)}$, and $C_{30}^{(10)}$ (MOD4) from MOD3 through the equivalent circuit shown in Fig. 4 with at most 2^{48} times and the average 2^{43} times of $S_{1}$ operation and the following multiplications, which corresponds to (18). The number of elements of MOD4 is at most 2^{60} and the average is 2^{55}.

3) Guess 8-bit $RK_{20,2}$, and make MOD of the total 48-bit data, $C_{03}^{(11)}$, $U_{2}^{(11)}$, and $C_{30}^{(10)}$ (MOD5) from MOD4 through the equivalent circuit shown in Fig. 4 with at most 2^{46} times and the average 2^{41} times of $S_{0}$ operation and the following multiplications, which corresponds to (17). The number of elements of MOD5 is at most 2^{58} and the average is 2^{53}.

4) Guess 8-bit $RK_{20,3}$, and make MOD of the total 40-bit data, $C_{10}^{(10)}$ and $C_{30}^{(10)}$ (MOD6) from MOD5 through the equivalent circuit shown in Fig. 4 with at most 2^{44} times and the average 2^{39} times of $S_{1}$ operation and the following multiplications, which corresponds to (16). The number of elements of MOD6 is at most 2^{56} and the average is 2^{51}.

5) Guess 8-bit $RK_{19,0}$, and make MOD of the total 32-bit data, $C_{21}^{(10)}$, $C_{22}^{(10)}$, $C_{23}^{(10)}$, and $W_{00}^{(10)}$ (MOD7) from MOD6 through the equivalent circuit shown in Fig. 5 with at most 2^{42} times and the average 2^{37} times of $S_{1}$ operation corresponding to (15). The number of elements of MOD7 is at most 2^{54} and the average is 2^{48}.

6) Guess 8-bit $RK_{21,0}$, and make MOD of the total 32-bit data, $C_{21}^{(10)}$, $C_{22}^{(10)}$, $C_{23}^{(10)}$, and $W_{00}^{(10)}$ (MOD8) from MOD2 through the equivalent circuit shown in Fig. 5 with at most 2^{40} times and the average 2^{35} times of $S_{1}$ operation and the following multiplication, which corresponds to (10). The number of elements of MOD8 is at most 2^{52} and the average is 2^{47}.

7) Guess 8-bit $RK_{19,1}$, and make MOD of the total 24-bit data, $C_{22}^{(10)}$, $C_{23}^{(10)}$, and $W_{10}^{(10)}$ (MOD9) from MOD7 through the equivalent circuit shown in Fig. 5 with at most 2^{38} times and the average 2^{33} times of $S_{0}$ operation and the following multiplication, which corresponds to (14). The number of elements of MOD9 is at most 2^{24} and the average is 2^{23}.

8) Guess 8-bit $RK_{21,1}$, and make MOD of the total 24-bit data, $C_{22}^{(10)}$, $C_{23}^{(10)}$, and $W_{10}^{(10)}$ (MOD10) from MOD8 through the equivalent circuit shown in Fig. 5 with at most 2^{36} times and the average 2^{31} times of $S_{0}$ operation and the following multiplication, which corresponds to (9). The number of elements of MOD10 is at most 2^{24} and the average is 2^{23}.

9) Guess 8-bit $RK_{19,2}$, and make MOD of the total 16-bit data, $C_{21}^{(10)}$ and $W_{20}^{(10)}$ (MOD11) from MOD9 through the equivalent circuit shown in Fig. 5 with at most 2^{24} times and the average 2^{19} times of $S_{1}$ operation and the following multiplication, which corresponds to (13). The number of elements of MOD11 is at most 2^{20} and the average is 2^{17}.

10) Guess 8-bit $RK_{21,2}$, and make MOD of the total 16-bit data, $C_{23}^{(10)}$ and $W_{20}^{(10)}$ (MOD12) from MOD10 through the equivalent circuit shown in Fig. 5 with at most 2^{24} times and the average 2^{19} times of $S_{1}$ operation and the following multiplication, which corresponds to (8). The number of elements of MOD12 is at most 2^{20} and the average is 2^{16}.

11) Guess 8-bit $RK_{19,3}$, and make MOD of the total 8-bit data $C_{00}^{(9)}$ (MOD13) from MOD11 through the equivalent circuit shown in Fig. 5 with at most 2^{16} times and the average 2^{15} times of $S_{0}$ operation and the following multiplication, which corresponds to (12). The number of elements of MOD13 is at most 2^{16} and the average is 2^{11}.

12) Guess 8-bit $RK_{21,3}$, and compute the right-hand side of (6) via (7) from MOD12 through the equivalent circuit shown in Fig. 5 with at most 2^{16} times and the average 2^{15} times of $S_{0}$ operation and the following multiplication.

13) Guess 8-bit $RK_{16,0}$, and compute the left-hand side of (6) from MOD13 with at most 2^{16} times and the average 2^{15} times of $S_{0}$ operation. The guessed keys at the steps 1-13 are the candidate of correct key if (6) holds, otherwise they are false keys.

Attacker guesses all the 2^{104} kinds of the round key at the steps 1-13. Attacker executes step 0 one time. Steps 1-13 are executed by using a nested structure of loop iterations.

Step 1 is an outermost loop, and step 13 is an innermost loop.

First, attacker confirms the authenticity of the total 2^{104} sets of the guessed key by computing (6). And then, the number of candidate keys is reduced to 2^{96} (= 2^{104} \times 2^8).

Second, attacker confirms the authenticity of 2^{96} sets of the guessed keys by computing (6) again where the plain text $X$ is different from the one at the first time. And then, the number of candidate keys is reduced to 2^{88}. Attacker repeats this confirmation 14 times, and then the average number of candidate keys is reduced to 2^{7} (= 2^{104} \times (2^{8})^{14}) where the last key shall be a correct key. Because attacker has to compute the 96bit-order differential to prepare one set of (6), the number of chosen plain texts to prepare 14 different sets of (6) is given by $D$ as follows:
The maximum number ($T_{\text{max}}$) of data processing of 11-round CLEFIA, which include 88 sets of $S_i$ ($i = 0, 1$), for the attack algorithm are given by

$$T_{\text{max}} = \sum_{i=0}^{11} 2^{3i} T_i / 88 \approx 2^{106.6} \quad (24)$$

$$T_i = 2^8 (2^{72} + T_{i1}) \quad (25)$$

$$T_i = 2^8 (2^{48} + T_{i2}) \quad (26)$$

$$T_i = 2^8 (2^{24} + T_{i3}) \quad (27)$$

$$T_9 = 2^8 (2^{24} + T_{10}) \quad (28)$$

$$T_{10} = 2^8 (2^{24} + T_{11}) \quad (29)$$

$T_i$ ($i = 1, 2, \ldots, 13$) is the number of $S_i$ ($i = 0, 1$) operations for the nested loops from step $i$ to step 13.

V. CONCLUSIONS

We have studied the 96th-order differential attack on 11 rounds of CLEFIA. We reduced the number of times of data processing for the attack by applying partial sum technique proposed by Ferguson et al. With the technique, we sequentially derived MOD of intermediate data of the data processing part. Moreover we reduced the number of the times by applying a nested structure of iterative computation to the attack algorithm and optimizing the order of the nest. As a result we showed that the 96th-order differential attack can be performed with $2^{99.8}$ blocks of plain text and $2^{106.6}$ times of data processing, which is $1/52.4$ of computational complexity compared to the conventional attack. The future work is to apply our attack algorithm to [4].

REFERENCES


