Differential Cryptanalysis of a Reduced-Round SEED

Hitoshi Yanami[1] and Takeshi Shimoyama[1]
[1]FUJITSU LABORATORIES LTD.
1-1, Kamikodanaka 4-Chome, Nakahara-ku, Kawasaki, 211-8588, Japan
{yanami, shimo}@flab.fujitsu.co.jp

June 20, 2002

Abstract

We analyze the security of the SEED block cipher against differential attacks. SEED is a 16-round Feistel cipher developed by the Korea Information Security Agency. The SEED proposers estimated their cipher against differential cryptanalysis in a self-estimation document and found a six-round differential characteristic with probability 2^{-130}. We present an improved method of examining the differential characteristics of SEED and show three six-round differential characteristics with probability 2^{-124}. These characteristics allow us to attack seven-round SEED, which surpasses the proposers estimation. Our differential attack needs 2^{126} chosen-plaintext pairs and 2^{126} computations of the F function to deduce the subkey used in the last round of seven-round SEED.
Keywords: symmetric block cipher, SEED, differential attack, characteristic, probability

1 Introduction

SEED is a block cipher which was developed by the Korea Information Security Agency [3]. The cipher has been attracting attention in South Korea as a next-generation cipher since it was proposed in 1998. In *The Analyses on SEED* [4], one of the self-estimation documents, the proposers have evaluated the security of SEED against several well-known attacks, including the security against differential cryptanalysis. In the document the proposers examined the differential characteristics of SEED and found a six-round differential characteristic with probability 2^{-130}. They investigated firstly a simplified version of SEED and then applied the characteristic with the highest probability they found to the original SEED.

Our aim in the present paper is to find out a differential characteristic which has a higher probability than the one found by the proposers. We directly deal with SEED itself and investigate differential characteristics that were not examined by the proposers' method to find a differential characteristics with a higher probability. We also illustrate a differential attack on seven-round SEED by utilizing our best differential characteristics. Our differential attack surpasses the conclusion stated in the proposers' self-estimation document.

The paper is organized as follows. We briefly describe the encryption algorithm for SEED in Section 2. In Section 3, we illustrate the proposers' method of investigating the differential characteristics and their result. In Section 4, we point out some defects in the proposers' method. In Section 5, our research method is presented and our result is shown. In Section 6,

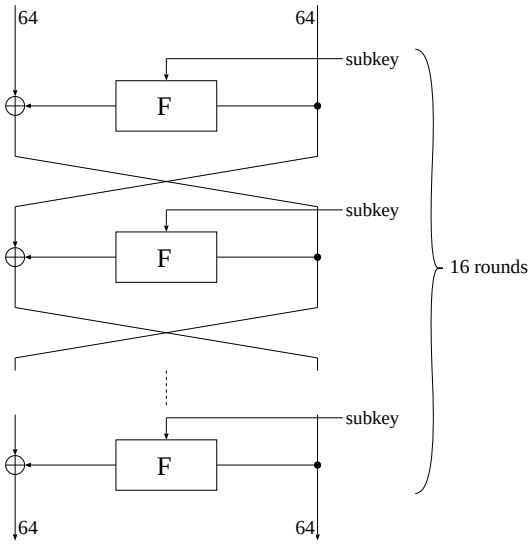

Figure 1: The SEED Feistel structure

we illustrate our differential attack on seven-round SEED. We summarize our paper in Section 7.

2 Description of the SEED Block Cipher

We present the algorithm for the SEED block cipher and then summarize the nonlinear functions in SEED. We remark that SEED supports only 128-bit user keys. Different from the AES, SEED does not support 192- and 256-bit user keys. The description of the SEED key schedule will be omitted here because our attack is independent of the structure of the SEED key schedule.

2.1 The SEED Encryption Algorithm

We briefly describe the SEED encryption algorithm. We refer to [3, 5] for the full description of SEED.

SEED is a block cipher that was proposed by the Korea Information Security Agency. The cipher, which has a 128-bit block size and supports only 128-bit user keys, has been submitted to ISO/IEC JTC1 SC27. SEED has been attracting attention in South Korea as a next-generation cipher since it was proposed in 1998.

SEED has a classical Feistel structure and its total number of rounds is 16 (Fig. 1). The round function in SEED is called the F function; it is a 64-bit input/output function. The structure of the F function (Fig. 2) is similar to that of the FO function used in the MISTY block cipher. The F function in SEED uses the 32-bit ADD operation in three places, while the FO function in MISTY adopted XOR operation. The use of ADDs makes examination of the differential characteristics complicated. Calculation should be carefully and efficiently carried out when we examine the differential characteristics of SEED.

The F function in SEED contains three G functions. The G function is a 32-bit input/output function, consisting of four S-boxes and a linear transformation (Fig. 2). The input of the G function is divided into four 8 bits, each of which will be called a (8-bit) *word.*

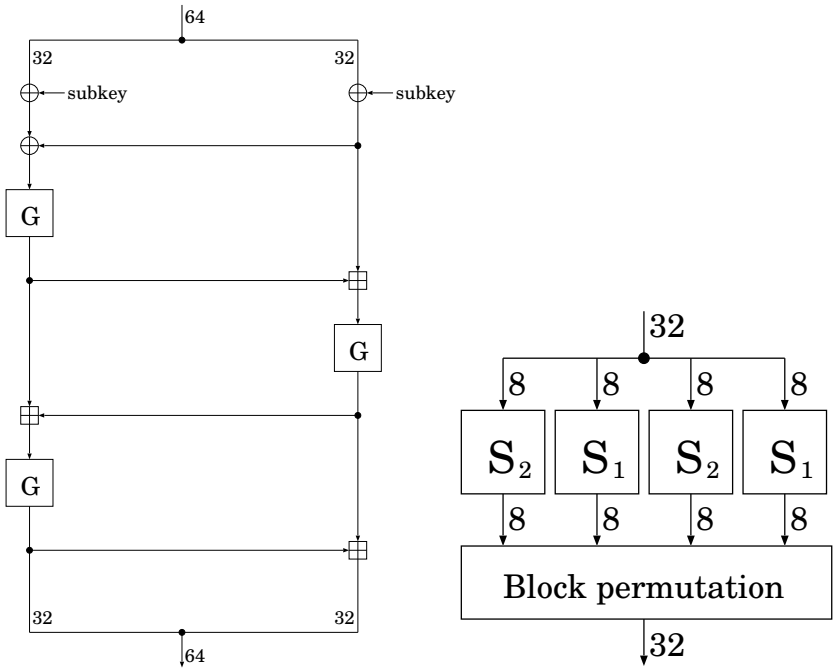

Figure 2: The F function (left) and the G function

Each word enters an S-box, either S_1 or S_2, depending on the position of the word. The four output words are put into a linear transformation, whose output becomes the output of the G function. The proposers named this 32-bit linear transformation the *block permutation*. The output (d', c', b', a') of the block permutation for the input (d, c, b, a) is represented in the following formulae:

$$\begin{aligned} a' &= (a \wedge m_0) \oplus (b \wedge m_1) \oplus (c \wedge m_2) \oplus (d \wedge m_3), \\ b' &= (a \wedge m_1) \oplus (b \wedge m_2) \oplus (c \wedge m_3) \oplus (d \wedge m_0), \\ c' &= (a \wedge m_2) \oplus (b \wedge m_3) \oplus (c \wedge m_0) \oplus (d \wedge m_1), \\ d' &= (a \wedge m_3) \oplus (b \wedge m_0) \oplus (c \wedge m_1) \oplus (d \wedge m_2), \end{aligned}$$

where $m_0 =$ `0xfc`, $m_1 =$ `0xf3`, $m_2 =$ `0xcf` and $m_3 =$ `0x3f` are all 8-bit constants, and $\wedge$ represents the bitwise-AND operation. Note that a bit in an input word influences only the corresponding bits in the four output words. We show how a single bit in the leftmost word is diffused by the block permutation in Table 1.

2.2 The Nonlinear Functions in SEED

We summarize the nonlinear functions in SEED and the differential probabilities for these functions.

We will use the following notation to represent a differential probability. Given a pair of input differences $(\Delta A, \Delta B)$ and an output difference ΔC, we denote the differential probability for these differences through the ADD operation $\boxplus$ by $\mathrm{DP}[(\Delta A, \Delta B) \xrightarrow{\boxplus} \Delta C]$. More precisely, the differential probability is defined by the following:

$$\begin{aligned} &\mathrm{DP}[(\Delta A, \Delta B) \xrightarrow{\boxplus} \Delta C] \\ &\stackrel{\text{def}}{=} \mathrm{Prob}_{X,Y}[(X \boxplus Y) \oplus ((X \oplus \Delta A) \boxplus (Y \oplus \Delta B)) = \Delta C]. \end{aligned}$$

Table 1: Diffusion of a single bit through the block permutation

Input	Output
0x8000 0000	0x8080 8000
0x4000 0000	0x4040 4000
0x2000 0000	0x0020 2020
0x1000 0000	0x0010 1010
0x0800 0000	0x0800 0808
0x0400 0000	0x0400 0404
0x0200 0000	0x0202 0002
0x0100 0000	0x0101 0001

A similar notation will be used for the differential probability through an S-box.

SEED has two types of nonlinear functions: the 8-bit S-box tables and the 32-bit ADD operation. Four S-boxes are in the G function and three ADDs in the F function. We state the differential probabilities for both functions.

The S-boxes. SEED uses two 8-bit S-boxes, S_1 and S_2. Both S-boxes are optimized against differential cryptanalysis. Given a pair of nonzero input and output differences, the differential probability for the pair through an S-box is either 2^{-6} or 2^{-7} (or 0). Moreover, given a nonzero input difference, there exists only one output difference for which the differential probability is 2^{-6}.

The 32-bit ADD operation. SEED uses the 32-bit ADD operation in three places in the F function instead of the usual 32-bit XOR operation.

In case of the XOR operation, the behavior of input/output differences is quite obvious; for any pair of input differences $(\Delta A, \Delta B)$, the output difference results in, by definition, $\Delta A \oplus \Delta B$ with probability 1. Compared with the XOR operation, it is much difficult to compute the differential probability with respect to the ADD operation. For a pair of input differences, the possible output differences may vary. We will treat the differential probability with respect to the ADD operation later in Section 4.

3 Proposers' Estimation

We describe how the proposers investigated the differential characteristics of SEED and show their result. SEED uses the ADD operation, which is hard to be dealt with when the differential characteristics are examined. The proposers circumvented the difficulty as follows. They firstly changed SEED into a simplified version, in which every ADD operation was replaced by the XOR operation. They called the version the *modified SEED*. They examined the differential characteristics of the modified SEED, and then, for the best characteristic found by their search, they calculated the probability at each of the ADDs to obtain the exact differential probability for the characteristic.

Their best result for six and seven rounds is summarized as follows:

Number of rounds	Modified SEED	SEED
6	2^{-102}	2^{-130}
7	2^{-108}	2^{-144}.

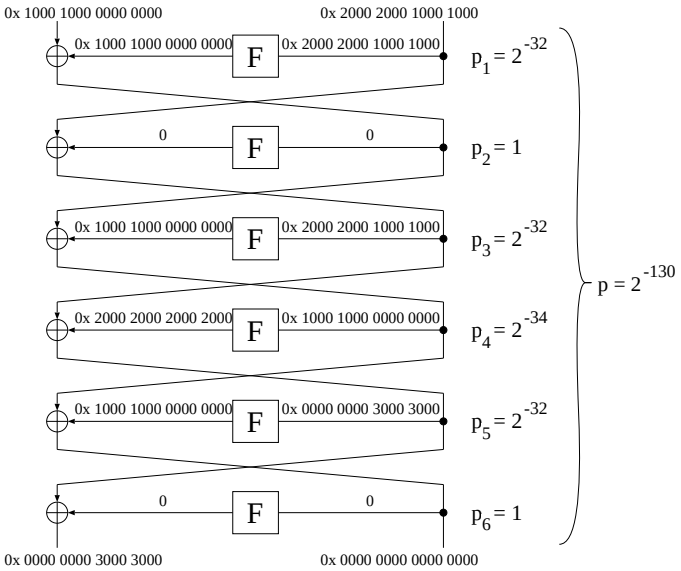

Figure 3: The proposers' best 6-round differential characteristic

The best six-round differential characteristic, which has a probability of 2^{-130}, is shown in Fig. 3. Their result shows that the highest differential probability is less than 2^{-128} for six-round SEED or more. Note that there exists a five-round differential characteristic with probability higher than 2^{-128}; eliminate the first round in the best six-round characteristic, and a five-round characteristic with probability 2^{-98} is obtained. The proposers also stated in [4] that the diffusion of the block permutation in the G function prevents attackers from applying elimination attacks more than one round to SEED.

Though the proposers did not mention explicitly, their result indicates that with a five-round characteristic and a one-round elimination attack, an attacker is able to attack six-round SEED, and the differential attack is not applicable for seven-round SEED or more.

4 Defects in the Proposers' Method

As was shown in the previous section, the SEED proposers investigated the differential characteristics via the modified SEED. We indicate two defects in their investigation.

The Output Differences of the ADD Operation. Since the proposers firstly examined the modified SEED, in which every ADD operation was changed into the XOR operation, they only took up the XORed value $\Delta A \oplus \Delta B$ as an output difference for a pair of input differences $(\Delta A, \Delta B)$. The other possible output differences were ignored.

Let us show an example by using the 8-bit ADD operation. Let $(\Delta A, \Delta B) = (\mathtt{11000100}_2, \mathtt{01000000}_2)$ be a pair of input differences. We list the output differences ΔC such that

$$\mathrm{DP}[(\Delta A, \Delta B) \xrightarrow{\boxplus} \Delta C] > 0$$

in Table 2. The proposers' method would consider ΔC as $\mathtt{10000100}_2$, the XORed value of ΔA and ΔB. There are, in fact, a lot of possible output differences for an input pair. In general, the differential probability is the highest when the output difference equals the XORed value, and decreases according as it deviates from the XORed value. Though the proposers always

Table 2: An example of input/output differences (the 8-bit ADD operation)

	Difference	Probability
ΔA	$\mathtt{11000100}_2$	
ΔB	$\mathtt{01000000}_2$	
	$\mathtt{10000100}_2$	2^{-2}
	$\mathtt{00000100}_2$	2^{-2}
	$\mathtt{10001100}_2$	2^{-3}
	$\mathtt{00001100}_2$	2^{-3}
ΔC	$\mathtt{10011100}_2$	2^{-4}
	$\mathtt{00011100}_2$	2^{-4}
	$\mathtt{01111100}_2$	2^{-4}
	$\mathtt{10111100}_2$	2^{-5}
	$\mathtt{00111100}_2$	2^{-5}

chose one of the output differences with the highest probability at every ADD operation, a better characteristic might be found if all the possible output differences were examined.

Exceptional Behavior of the Most Significant Bit. The second problem in the proposers' method is that they did not pay any attention to a peculiarity of the most significant bit in calculating a differential probability with respect to the ADD operation. A better result would be obtained if this bit were exceptionally treated in examination. It was shown in [6] that, for given three 32-bit differences ΔA, ΔB and ΔC,

$$\mathrm{DP}[(\Delta A, \Delta B) \stackrel{\boxplus}{\rightarrow} \Delta C] = 2^s,$$

provided that $\mathrm{DP}[(\Delta A, \Delta B) \stackrel{\boxplus}{\rightarrow} \Delta C] > 0$, where the integer s is given by

$$s = \#\{i | 0 \le i < 31, \text{not } ((\Delta A)_i = (\Delta B)_i = (\Delta C)_i)\}.$$

The above formula implies that the differential probability is independent of the state of the most significant bit ($i = 31$) if the probability turns out to be positive. For example, the following two sets of differences have the same differential probability 2^{-1}:

$$\mathrm{DP}[(\mathtt{0x\,0000\,8000}, \mathtt{0x\,0000\,8000}) \stackrel{\boxplus}{\rightarrow} \mathtt{0}] = 2^{-1};$$
$$\mathrm{DP}[(\mathtt{0x\,8000\,8000}, \mathtt{0x\,8000\,8000}) \stackrel{\boxplus}{\rightarrow} \mathtt{0}] = 2^{-1}.$$

Slightly complicated formulae are required to state a necessary and sufficient condition for $\mathrm{DP}[(\Delta A, \Delta B) \stackrel{\boxplus}{\rightarrow} \Delta C]$ to be positive. We only show the result by Lipmaa and Moriai. We refer to [6] for the proof.

Let ΔA, ΔB and ΔC be 32-bit differences. Then

$$\mathrm{DP}[(\Delta A, \Delta B) \stackrel{\boxplus}{\rightarrow} \Delta C] > 0$$
$$\Longleftrightarrow$$
$$\mathrm{eq}(\Delta A \ll 1, \Delta B \ll 1, \Delta C \ll 1) \wedge (\Delta A \oplus \Delta B \oplus \Delta C \oplus (\Delta B \ll 1)) = 0,$$

where $\Delta A \ll 1$ represents the left shift of ΔA by one bit and $\mathrm{eq}(x, y, z)$ is the 32-bit value whose i-th bit $\mathrm{eq}(x, y, z)_i$ equals 1 if and only if the three bit values x_i, y_i and z_i are the same.

We will exploit the particular behavior of the most significant bit to obtain a characteristic with a high probability.

5 The Differential Characteristics of SEED

Taking the observations in the previous section into consideration, we investigate the differential characteristics of SEED. We directly deal with the SEED cipher itself to examine differential characteristics that have not been researched by the proposers. We firstly present our research method and then show our result.

5.1 Our Research Method

We present our research method in detail. To utilize the exceptional behavior of the most significant bit, we restrict the leftmost word to either `0x00` or `0x80`. Moreover, considering the influence of the block permutation, we also restrict the other words to being in the same form, i.e., we will only deal with the 32-bit differences of the form $\Delta A = (\Delta A_3, \Delta A_2, \Delta A_1, \Delta A_0)$, $\Delta A_i \in \{\mathtt{0x00}, \mathtt{0x80}\}$, $i = 0, \ldots, 3$. We describe such a type of difference as *fundamental.* Note that if a fundamental difference is input to the block permutation, the corresponding output difference is also fundamental.

We exceptionally allow non-fundamental differences for the output of the ADD operation to examine all the possible output differences. We are sure to make such a non-fundamental difference back into a fundamental state at the S-boxes in the G function immediately after the ADD operation; otherwise the difference would be badly transformed by the following block permutation. We show an example of a differential characteristic for the G function containing a non-fundamental difference in Fig. 4. The output difference `0x 0100 0000` in the ADD operation is non-fundamental and it is restored to `0x 8000 0000` by the S-boxes.

In investigating the differential characteristics, we pre-compute, for every pair of fundamental input differences $(\Delta A, \Delta B)$, the output differences ΔC's satisfying

$$\mathrm{DP}[(\Delta A, \Delta B) \xrightarrow{\boxplus} \Delta C] > 0.$$

As was shown in [6], all such output differences can be inductively constructed by using the following formulae:

$$\begin{array}{l}(\Delta C)_0 = (\Delta A)_0 \oplus (\Delta B)_0; \\ \text{For } 1 \le i \le 31, \\ (\Delta C)_i = \begin{cases} (\Delta A)_i \oplus (\Delta B)_i \oplus (\Delta B)_{i-1}, \text{ if } (\Delta C)_{i-1} = (\Delta A)_{i-1} = (\Delta B)_{i-1}, \\ \text{arbitrary value, otherwise.} \end{cases}\end{array}$$

We summarize our research method of finding a differential characteristic with a high probability.

1. For each pair of fundamental differences $(\Delta A, \Delta B)$, make the list $\{\Delta C\}$ of output differences such that $\mathrm{DP}[(\Delta A, \Delta B) \xrightarrow{\boxplus} \Delta C] > 0$.

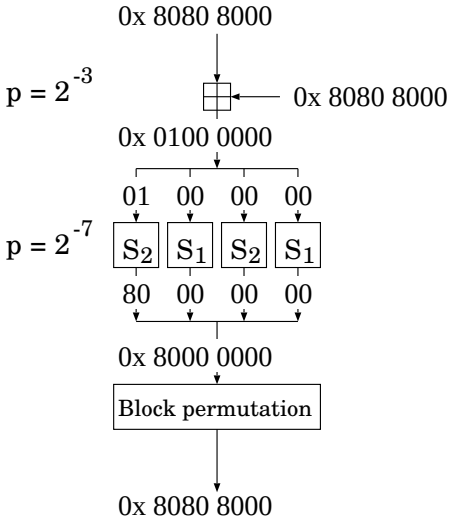

Figure 4: A characteristic containing a non-fundamental difference

2. By using the list made in the first process, make the list of fundamental input/output differences that have a positive probability through the ADD operation together with the G function. Calculate the differential probability for each triplet of differences and store the value.

3. Make a similar list for the F function.

4. Based on the list made in the third prosess, examine inductively whether or not a difference is connected to another one up to the prescribed number of rounds, say, n. If a connected chain of differences with length n (a n-round differential characteristic) is found, record the characteristic with its probability.

5. Output the characteristic with the highest probability.

5.2 Our Best Differential Characteristics

By using our research method, we have found three six-round differential characteristics with probability 2^{-124} and a differential characteristic with probability 2^{-128}. Each of these probabilities is higher than 2^{-130}, the highest found by the proposers. We show a differential characteristic with probability 2^{-124} in Fig. 5 and with probability 2^{-128} in Fig. 7, respectively.

The S-boxes in the second to fourth F functions in Fig. 5 utilizes the following differential probability:

$$\mathrm{DP}[\mathtt{0x80} \xrightarrow{S_1} \mathtt{0x80}] = \mathrm{DP}[\mathtt{0x80} \xrightarrow{S_2} \mathtt{0x80}] = 2^{-7}.$$

At every ADD operation, the highest probability 2^{-1} except 1 are recorded, which is attributed to the use of the most significant bit. We show the differences in the second F function in Fig. 6 as an illustration.

The other two characteristics with probability 2^{-124} have similar differences as in Fig. 5. We put $\Delta L = \mathtt{0x\,8000\,0080\,0000\,0000}$, $\Delta M = \mathtt{0x\,0000\,0000\,8000\,0080}$ and $\Delta N = \mathtt{0x\,8000\,0080\,8000\,0080}$. By cyclically shifting ΔL, ΔM and ΔN, we obtain the other

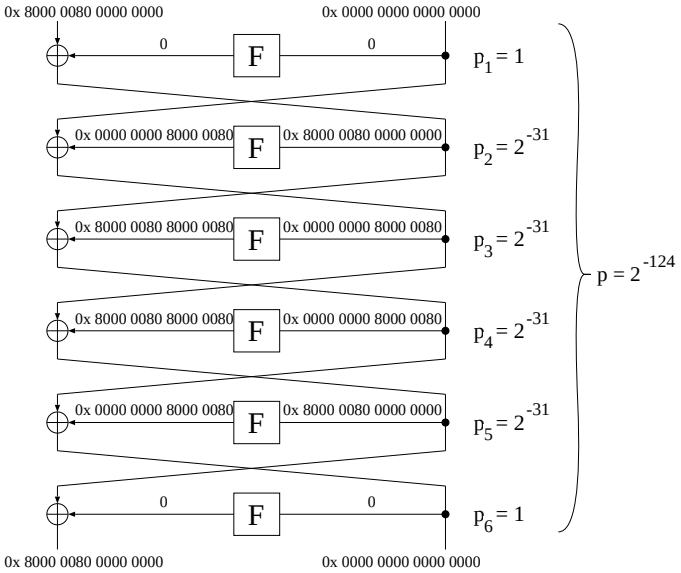

Figure 5: Our best 6-round differential characteristic

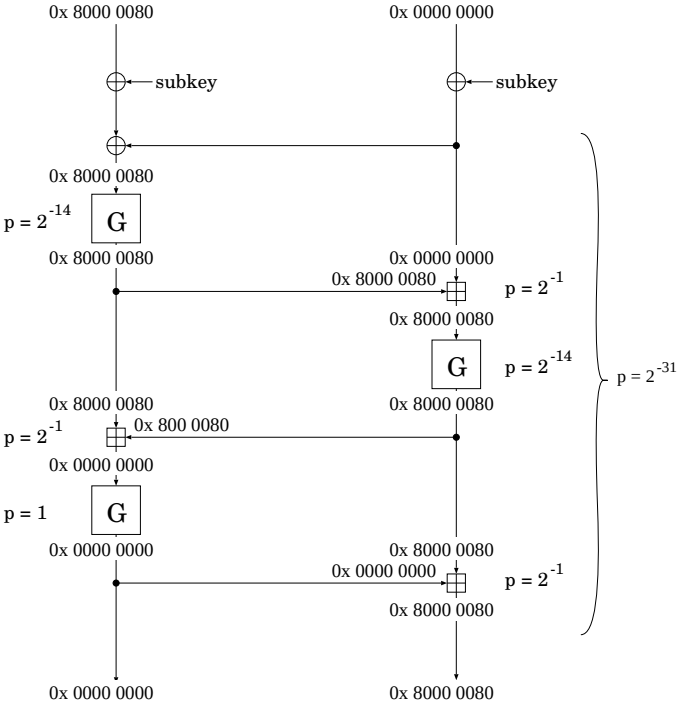

Figure 6: The differences in the 2nd F function in Fig. 5

characteristics with the same probability. We will use these three characteristics in the next section to reduce the number of plaintext pairs in attacking seven-round SEED.

The differential characteristic shown in Fig. 7 has a non-fundamental difference in the third round. We show the differences in the third F function in Fig. 8. Though occurring such a difference is intriguing, the characteristic is not applicable to a differential attack because its probability does not exceed 2^{-128}.

6 A Differential Attack on 7-Round SEED

The six-round differential characteristics with probability 2^{-124} found by our research enable us to attack seven-round SEED. We illustrate our differential attack on seven-round SEED. By utilizing the best characteristics together with a one-round elimination attack, we are able to deduce the subkey in the last round. We used the following algorithm to infer the subkey in the last round. (See Fig. 9.)

- We start by collecting $4 \cdot 2^{124} = 2^{126}$ pairs[1] (P, P^*) of plaintexts such that $P \oplus P^* =$ 0x8000 0080 0000 0000 0000 0000 0000 0000, the input difference of the characteristic shown in Fig. 9. We denote the respective corresponding ciphertexts of P and P^* by C and C^*.
- Initialize a 2^{64} array of counters corresponding to the 64 subkey bits in the last round.
- Exclude the plaintext pairs whose ciphertext difference is not equal to 0x8000 0080 0000 0000, the right 64-bit output difference of the characteristic shown in Fig. 9.
- Compute the output difference in the last F function for each remaining pair and for every 64-bit subkey value. (We expect that about $2^{126}/2^{64} = 2^{62}$ pairs pass the third process.) Increment the corresponding counter by 1 if the difference equals the left 64 bits in $C \oplus C^*$.
- After the counting, we deduce that the counter with the highest number indicates the right subkey.

We remark on the number of plaintexts and computations we require. Differential cryptanalysis is categorized chosen-plaintext attacks; the attacker prepares a number of plaintext pairs with a designated XORed value. Since our attack in the present paper needs to prepare 2^{126} chosen-plaintext pairs, our attack would seem to requires 2^{127} plaintexts. We are able to reduce the number to 2^{126} by a well-known trick shown in [2]. We use the notations ΔL, ΔM and ΔN as in the previous section. The equation $\Delta L \oplus \Delta M = \Delta N$ holds for these three differences. When, for three plaintexts P, P^* and P^{**}, the equations $P \oplus P^* = (\Delta L, 0)$ and $P \oplus P^{**} = (\Delta M, 0)$ hold, the equation $P^* \oplus P^{**} = (\Delta N, 0)$ also holds. Thus three plaintext pairs are obtained from these three plaintexts. Sharing the plaintexts and applying a similar algorithm to each of these three characteristics with probability 2^{-124} enable us to reduce the number of plaintexts and computations of the F function by half.

[1]Since the signal-to-noise value is large ($S/N = 2^4$), the subkey might be deduced from about $4 \cdot p^{-1}$ plaintext pairs, where p represents the differential probability for the characteristics [1].

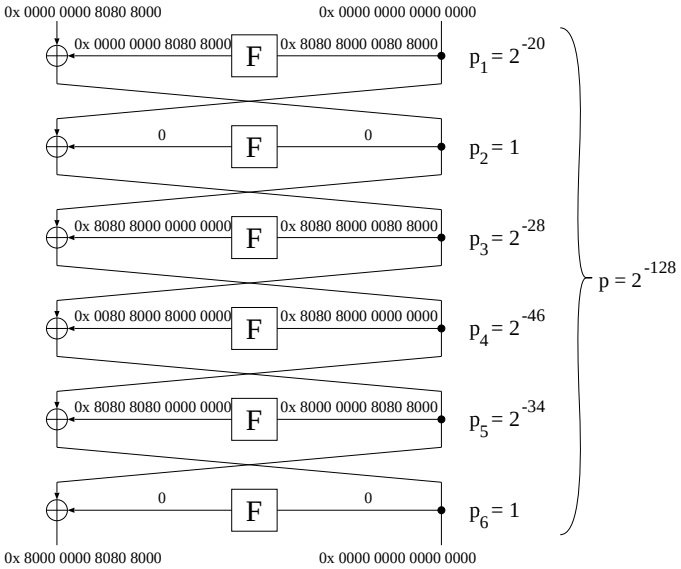

Figure 7: Our second-best 6-round differential characteristic

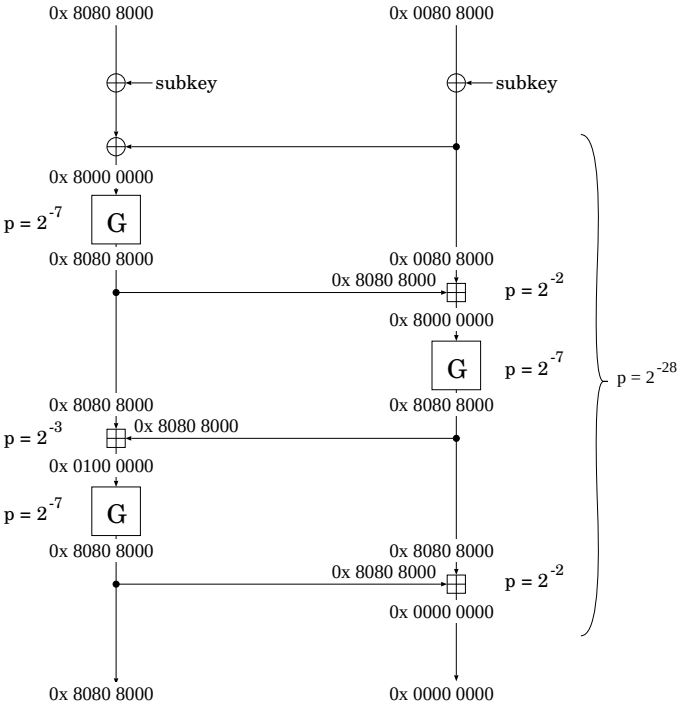

Figure 8: The differences in the 3rd F function in Fig. 7

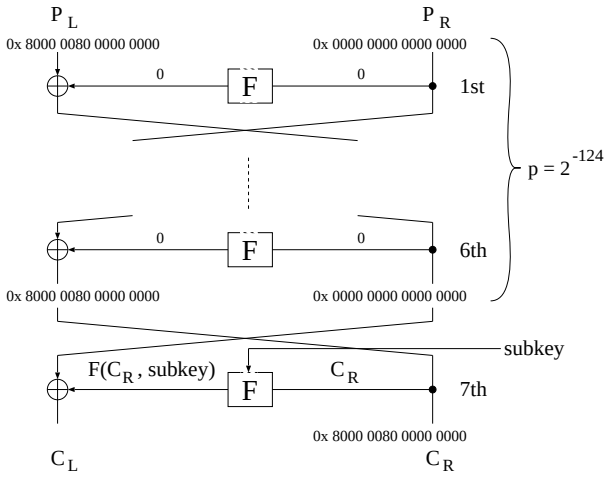

Figure 9: A differential attack on 7-round SEED

7 Conclusions

We have investigated the differential characteristics of the SEED block cipher. Improving the proposers' research method, we found four six-round differential characteristics that have higher probabilities than the proposers'. Our best six-round characteristics have a probability of 2^{-124}. By utilizing them, we were able to make a differential attack on seven-round SEED. Our differential attack needs 2^{126} chosen-plaintext pairs and 2^{126} computations of the F function to deduce the subkey in the last round. Although our result is beyond the proposers' self-estimation, the full-round SEED cipher, which has 16 rounds, seems to have enough security against differential cryptanalysis.

References

[1] E. Biham and A. Shamir, *Differential Cryptanalysis of DES-like Cryptosystems,* Journal of Cryptology, Vol. 4, No. 1, 3-72, 1991.

[2] E. Biham and A. Shamir, *Differential Cryptanalysis of Feal and N-Hash,* EUROCRYPT '91, Lecture Notes in Computer Science **547**, 1-16, 1991.

[3] Korean Information Security Agency, *A Design and Analysis of SEED,* 1998 (in Korean). (`http://www.kisa.or.kr/technology/sub1/128-seed.pdf`)

[4] Korean Information Security Agency, *ANNEX: The Analyses on SEED,* seed_analysis.doc, 2000.
(`http://www.kisa.or.kr/seed/algorithm.htm`)

[5] Korean National Body, *Contribution for Korean Candidates of Encryption Algorithm (SEED),* ISO/IEC JTC1 SC27 N2563, seed_english.doc, 2000.
(`http://www.kisa.or.kr/seed/algorithm.htm`)

[6] H. Lipmaa and S. Moriai, *Efficient Algorithms for Computing Differential Properties of Addition,* Preproceedings of FSE2001, 347-361, 2001.