Low Probability Differentials and the Cryptanalysis of Full-Round CLEFIA-128

Sareh Emami[2], San Ling[1], Ivica Nikolić[1]*, Josef Pieprzyk[2] and Huaxiong Wang[1]

[1] Nanyang Technological University, Singapore
[2] Macquarie University, Australia

Abstract. So far, low probability differentials for the key schedule of block ciphers have been used as a straightforward proof of security against related-key differential attacks. To achieve the resistance, it is believed that for cipher with k-bit key it suffices the upper bound on the probability to be 2^{-k}. Surprisingly, we show that this reasonable assumption is incorrect, and the probability should be (much) lower than 2^{-k}. Our counter example is a related-key differential analysis of the block cipher CLEFIA-128. We show that although the key schedule of CLEFIA-128 prevents differentials with a probability higher than 2^{-128}, the linear part of the key schedule that produces the round keys, and the Feistel structure of the cipher, allow to exploit particularly chosen differentials with a probability as low as 2^{-128}. CLEFIA-128 has 2^{14} such differentials, which translate to 2^{14} pairs of weak keys. The probability of each differential is too low for attacks, but the weak keys have a special structure which allows with a divide-and-conquer approach to gain advantage of 2^7 over generic attacks. We exploit the advantage and give a membership test for the weak-key class, provide analysis in the hashing mode, and show the importance for the secret-key mode. The proposed analysis has been tested with computer experiments on small-scale variants of CLEFIA-128. Our results do not threaten the practical use of CLEFIA.

Keywords: CLEFIA, cryptanalysis, weak keys, CRYPTREC, differentials

1 Introduction

CLEFIA [18] is a block cipher designed by Sony. It is advertised as a fast encryption algorithm in both software and hardware and it is claimed to be highly secure. The efficiency comes from the generalized Feistel structure and the byte orientation of the algorithm. The security is based

* The researcher is supported by the Singapore National Research Foundation Fellowship 2012 NRF-NRFF2012-06.

on the novel technique called Diffusion Switching Mechanism, which increases resistance against linear and differential attacks, in both single and related-key models. These and several other attractive features of CLEFIA-128 have been widely recognized, and the cipher has been submitted for standardization (and already standardized) by several bodies: CLEFIA was submitted as an encryption standard to IETF (Internet Engineering Task Force) [1], it is on the Candidate Recommended Ciphers List[3] of CRYPTREC (Japanese government standardization body), and it is one of the only two[4] lightweight block ciphers recommended by the ISO/IEC standard [13].

A significant body of analysis papers has been published on the round-reduced versions of CLEFIA [23, 24, 19, 22, 20, 15, 21, 14, 7], all for the single-key model. The analysis for the related-key model is missing. Often this type of attack can cover a higher number of rounds but requires the cipher to have a relatively simple and almost linear key schedule. CLEFIA, however, has a highly non-linear key schedule, equivalent roughly to 2/3 of the state transformation and designed with an intention to make the cipher resistant against related-key differential attacks. Using a widely accepted approach, the designers have proved that no such attack could exist as the key schedule has only low probability ($\leq 2^{-128}$ for CLEFIA with 128-bit keys) differential characteristics. Note, we will not try to exploit the fact that some characteristics can be grouped into a differential that has a much higher probability than the individual characteristics. Our results go a step further and we show that key schedule differentials with a probability as low as 2^{-128}, can still be used in attacks. This happens when they have a special structure, namely, the input/output differences of the differentials are not completely random, but belong to a set that, as in the case of CLEFIA-128, is described with a linear relation.

We exploit the special form of the key schedule: a large number of non-linear transformations at the beginning of the key schedule is followed by light linear transformations that are used to produce the round keys. In the submission paper of CLEFIA-128, the proof of related-key security is based only on the non-linear part as this part guarantees that the probability of any output difference is 2^{-128}. In contrast, our analysis exploits the linear part and we show that there are 2^{14} of the above low probability differences which, when supplied to the linear part, produce a special type of iterative round key differences. CLEFIA-128 is a Feistel cipher and, as shown in [6], iterative round key differences lead to an

[3] This is the final stage of evaluation, before becoming CRYPTREC standard.

[4] The second one is PRESENT [8].

iterative differential characteristic in the state that holds with probability 1. Therefore we obtain related-key differentials with probability 1 in the state and 2^{-128} in the key schedule. The low probability (2^{-128}) of each of the 2^{14} iterative round key differences means that for each of them there is only one pair of keys that produces such differences, or in total 2^{14} pairs for all of them – these pairs form the weak-key class of the cipher. When we target each pair independently, we cannot attack the cipher. However, the whole set of 2^{14} pairs has a special structure and we can target independently two smaller sets of sizes 2^7 and thus obtain the advantage of 2^7 over generic attacks. As we will see in the paper, the special structure of the weak key class is due to the linear part of the key schedule, therefore we exploit the weakness of this part twice (the first time for producing iterative round key differences).

We further analyze the impact of the 2^{14} pairs of keys and the advantage of 2^7 that we gain over generic attacks. First we show that CLEFIA-128 instantiated with any pair of weak keys can easily be attacked, namely we present a membership test for the weak class. That is, the cipher can be broken in the secret-key model if the key pair is chosen to be some of the 2^{14} special pairs. Next, for the hashing mode of CLEFIA-128, i.e. when the cipher is used in single-block-length hash constructions, we show that differential multicollisions [5] can be produced with a complexity lower than for an ideal cipher. Finally, we focus on distinguishing attacks in the framework, where the key is secret (and chosen uniformly at random from the set of all possible 2^{128} keys) but can be changed. We show that here the advantage 2^7 (and a weak-key class of 2^{14}) is insufficient to attack straightforwardly the cipher. However, constructions that internally use the cipher may be possible to attack. As the model of attacks under related weak keys is relatively new, neither strict bounds on complexity of attacks nor constructions resistant against such attacks are known. We formulate two open problems to tackle these critical questions and we conjecture that a construction, very similar to PRINCE [9] but with linear function in the key schedule replaced by a random permutation, could be a framework of great importance for related weak-key cryptanalysis. In such framework, the weak-key class of CLEFIA-128 could be used to show that the cipher is not ideal in the secret-key model.

The paper is organized as follows. We start with a description of CLEFIA-128 given in Section 2. We present the main results related to the analysis of the key schedule and the production of the class of 2^{14} pairs of weak-keys in Section 3. The distinguisher for the class, which is a differential membership test, is given in Section 4. We present the distinguisher for the cipher in Section 5 and apply it to the hashing mode and to the secret-key mode in Section 6. In Section 7 we conclude the paper.

2 Description of CLEFIA-128

CLEFIA is a 128-bit cipher that supports 128, 192, and 256-bit keys. We analyze CLEFIA with 128-bit keys that is referred as CLEFIA-128. Before we define the cipher, we would like to make an important note. To simplify the presentation, we consider CLEFIA-128 without whitening keys [5]. Our attack works for the original CLEFIA-128 and the analysis is given in Appendix C. We proceed now with a brief description of CLEFIA-128. It is an 18-round four-branch Feistel (see Fig. 3 of Appendix A) that updates two words per round. A definition of the state update function is irrelevant to our analysis (see [18] for a full description) and further we focus on the key schedule only.

A 128-bit master key K is input to a 12-round Feistel $GFN_{4,12}$(with the same round function as the one in the state, refer to Fig. 3 of Appendix A) resulting in a 128-bit intermediate key L. All the 36 round keys[6] $RK_i, i = 0, \ldots, 35$ are produced by applying a linear transformation to the master key K and the intermediate key L as shown below ($\oplus$ stands for the XOR operation and $||$ is concatenation):

$$\begin{array}{lll}
RK_0||RK_1||RK_2||RK_3 & \leftarrow L & \oplus S_1, \\
RK_4||RK_5||RK_6||RK_7 & \leftarrow \Sigma(L) \oplus K & \oplus S_2, \\
RK_8||RK_9||RK_{10}||RK_{11} & \leftarrow \Sigma^2(L) & \oplus S_3, \\
RK_{12}||RK_{13}||RK_{14}||RK_{15} & \leftarrow \Sigma^3(L) \oplus K & \oplus S_4, \\
RK_{16}||RK_{17}||RK_{18}||RK_{19} & \leftarrow \Sigma^4(L) & \oplus S_5, \\
RK_{20}||RK_{21}||RK_{22}||RK_{23} & \leftarrow \Sigma^5(L) \oplus K & \oplus S_6, \\
RK_{24}||RK_{25}||RK_{26}||RK_{27} & \leftarrow \Sigma^6(L) & \oplus S_7, \\
RK_{28}||RK_{29}||RK_{30}||RK_{31} & \leftarrow \Sigma^7(L) \oplus K & \oplus S_8, \\
RK_{32}||RK_{33}||RK_{34}||RK_{35} & \leftarrow \Sigma^8(L) & \oplus S_9,
\end{array}$$

[5] There are four whitening keys: two are added to the plaintext, and two to the ciphertext.

[6] Two round keys are used in every round, thus there are $2 \cdot 18 = 36$ keys in total.

where S_i are predefined 128-bit constants, and Σ is a linear function defined further. In short, each four consecutive round keys RK_{4i}, RK_{4i+1}, RK_{4i+2}, RK_{4i+3} are obtained by XOR of multiple applications of Σ to L, possibly the master key K, and the constant S_i. The resulting 128-bit sequence is divided into four 32-bit words and each is assigned to one of the round key words. The linear function Σ (illustrated in Fig. 1) is a simple 128-bit permutation used for diffusion. The function $\Sigma : \{0,1\}^{128} \rightarrow \{0,1\}^{128}$ is defined as follows:

$$X_{128} \rightarrow Y_{128}$$
$$Y = X[120-64]X[6-0]X[127-121]X[63-7],$$

where $X[a-b]$ is a bit sequence from the a-th bit to the b-th bit of X.

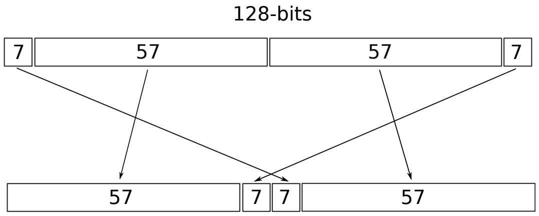

Fig. 1. The function Σ. The numbers denote the size of the bit sequence.

We would like to make a note about the notations of XOR differences used throughout the paper. To emphasize that a difference is in the word X, we use ΔX, otherwise, if it irrelevant or clear from the context we use simply Δ.

3 Weak Keys for CLEFIA-128

In the related-key model, the security of a cipher is analyzed by comparing two encryption functions obtained by two unknown but related keys. Given a specific relation[7] between keys, if the pair of encryption functions differs from a pair of random permutations, then the cipher has a weakness and can be subject to related-key attacks. Sometimes these attacks succeed only when the pairs of related keys belong a relatively small subset of all possible pairs of keys. The subset is called the *weak-key class* of the cipher and the number of pairs of keys is the size of the class.

[7] Some relations are prohibited as they lead to trivial attacks, see [3] for details.

We will show that a weak-key class in CLEFIA-128 consists of pairs of keys $(K, \tilde{K} = K \oplus \mathcal{L}_1(D))$, where D can take approximately 2^{14} different 128-bit values, such that for *any* plaintext P, the following relation holds:

$$E_K(P) \oplus E_{\tilde{K}}(P \oplus \mathcal{L}_2(D)) = \mathcal{L}_3(D), \tag{1}$$

where $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$ are linear functions defined below. The property can be seen as a related-key differential, with the difference $\mathcal{L}_1(D)$ for the master key, $\mathcal{L}_2(D)$ for the plaintext and $\mathcal{L}_3(D)$ for the ciphertext. From Equation (1), it follows that once D is defined, the probability of the differential is precisely one.

In the state of CLEFIA-128, the probability of a differential characteristic is one if for each Feistel round, there is no incoming difference for the non-linear round function. This happens when the differences in the state and in the round key cancel each other. Consequently, the input difference to the round function becomes zero[8]. An illustration of the technique for four rounds of CLEFIA-128 is given in Fig. 2. Notice that the input state difference at the beginning of the first round $(\Delta_1, \Delta_2, \Delta_3, \Delta_4)$ is the same as the output difference after the fourth round, i.e. it is iterative with the period of 4 rounds. Therefore, we will obtain a differential characteristic with probability 1 (in the state) for the full-round CLEFIA-128 if *we can produce 4-round iterative round key differences.*

Each round of the state uses two round keys, thus the above 4-round iterative characteristic requires the round key differences to have a period of 8, i.e. $\Delta RK_i = \Delta RK_{i+8}$. Moreover, an additional condition has to hold. Note that in Fig. 2, the differences in the consecutive round keys are $(\Delta_1, \Delta_3, \Delta_2, \Delta_4, \Delta_3, \Delta_1, \Delta_4, \Delta_2)$, that is among the 8 round key differences, the first four are different, while the remaining four are only permutations of the first. These two conditions can be summarized as follows:

Condition 1 - For all i, it should hold $\Delta RK_i = \Delta RK_{i+8}$.

Condition 2 - For all i divisible by 8, it should hold $\Delta RK_i = \Delta RK_{i+5}$, $\Delta RK_{i+1} = \Delta RK_{i+4}$, $\Delta RK_{i+2} = \Delta RK_{i+7}$, $\Delta RK_{i+3} = \Delta RK_{i+6}$. This can be rewritten as
$(\Delta RK_{i+4}, \Delta RK_{i+5}, \Delta RK_{i+6}, \Delta RK_{i+7}) = \pi(\Delta RK_i, \Delta RK_{i+1}, \Delta RK_{i+2}, \Delta RK_{i+3})$,
where π is 4-word permutation $(0, 1, 2, 3) \rightarrow (1, 0, 3, 2)$.

Further we show how to find the set of differences for which the two conditions hold.

[8] A similar idea is given in [6].

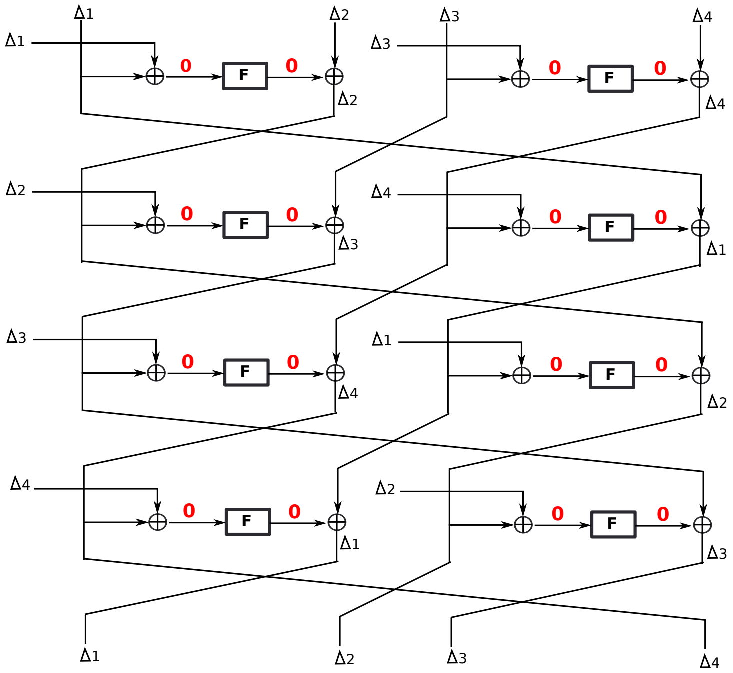

Fig. 2. Iterative related-key differential characteristic for 4 rounds of the `CLEFIA-128` that is true with probability 1. The symbols $\Delta_1, \Delta_2, \Delta_3, \Delta_4$ denote word differences.

Condition 1. From the definition of the key schedule

$$\begin{aligned} RK_{8i+0}||RK_{8i+1}||RK_{8i+2}||RK_{8i+3} &\leftarrow \Sigma^{2i}(L) &&\oplus S_{2i+1} \\ RK_{8i+8}||RK_{8i+9}||RK_{8i+10}||RK_{8i+11} &\leftarrow \Sigma^{2i+2}(L) &&\oplus s_{2i+3}, \end{aligned}$$

it follows that Condition 1 for the first 4 (out of 8) round key differences in an octet of round keys can be expressed as

$$\Delta L = \Sigma^2(\Delta L). \tag{2}$$

We will obtain the same equation if we consider the remaining 4 round key differences. To satisfy Condition 1, we have to find possible values for ΔL such that Equation (2) holds. This can be achieved easily as (2) is a system of 128 linear equations with 128 unknowns (refer to the definition of Σ), and has solutions of the form (expressed as concatenation of bit sequences):

$$\Delta L = a_1 a_2 t b_2 b_1 b_2 b_1 b_2 b_1 b_2 a_2 a_1 a_2 a_1 a_2 a_1 a_2 t b_1 b_2, \tag{3}$$

where a_1, a_2 are any 7-bit values, t is the most significant bit of a_1 and the 7-bit values b_1, b_2 are defined as $tb_2b_1 = a_1a_2t$. Thus there are $2^7 \cdot 2^7 = 2^{14}$ solutions.

Condition 2. From the definition of the key schedule

$$\begin{aligned} RK_{8i+0}||RK_{8i+1}||RK_{8i+2}||RK_{8i+3} &\leftarrow \Sigma^{2i}(L) & \oplus S_{2i+1}, \\ RK_{8i+4}||RK_{8i+5}||RK_{8i+6}||RK_{8i+7} &\leftarrow \Sigma^{2i+1}(L) \oplus K & \oplus S_{2i+2}, \end{aligned}$$

we see that Condition 2 can be expressed as

$$\pi(\Delta L) = \Sigma(\Delta L) \oplus \Delta K,$$

where π is 4-word permutation $(0, 1, 2, 3) \to (1, 0, 3, 2)$. Thus when ΔL is fixed (to one of the values from (3)), the difference in the master key ΔK can be determined as

$$\Delta K = \pi(\Delta L) \oplus \Sigma(\Delta L). \tag{4}$$

Summary. We have shown above that Conditions 1 and 2 can be achieved simultaneously as there are 2^{14} values for ΔL_i (see Equation (3)) with corresponding values of ΔK_i (see Equation (4)). It means that given the difference in the master key ΔK_i and the difference of the intermediate key ΔL_i (i.e. the differential in the 12-round Feistel $GFN_{4,12}$ of the key schedule is $\Delta K_i \to \Delta L_i$), the differences in the round keys are going to be of the requested form as shown below:

$$\begin{aligned} \Delta RK_0||\Delta RK_1||\Delta RK_2||\Delta RK_3 &= \Delta_1||\Delta_3||\Delta_2||\Delta_4, \\ \Delta RK_4||\Delta RK_5||\Delta RK_6||\Delta RK_7 &= \Delta_3||\Delta_1||\Delta_4||\Delta_2, \\ &\dots \\ \Delta RK_{28}||\Delta RK_{29}||\Delta RK_{30}||\Delta RK_{31} &= \Delta_3||\Delta_1||\Delta_4||\Delta_2, \\ \Delta RK_{32}||\Delta RK_{33}||\Delta RK_{34}||\Delta RK_{35} &= \Delta_1||\Delta_3||\Delta_2||\Delta_4, \end{aligned}$$

where $\Delta_1||\Delta_3||\Delta_2||\Delta_4 = \Delta L_i$. As a result, we have obtained the necessary differences in the round keys and we can use the 4-round iterative characteristic from Fig. 2.

Now we can easily specify the description of the weak-key class given by Equation (1). The value of D coincides with the values of ΔL from Equation (3). Therefore the first linear function $\mathcal{L}_1$ is defined as $\mathcal{L}_1(D) = \pi(D) \oplus \Sigma(D)$. The input difference in the plaintext is the same as the input difference in the first four round keys (which is again ΔL), but the order of the words is slightly different – instead of $(\Delta_1, \Delta_3, \Delta_2, \Delta_4)$ it is $(\Delta_1, \Delta_2, \Delta_3, \Delta_4)$, see Fig. 2. Hence, we introduce the 4-word permutation $\pi_2 : (0,1,2,3) \to (0,2,1,3)$ that corrects the order. With this notation, the second linear function $\mathcal{L}_2$ is defined as $\mathcal{L}_2(D) = \pi_2(D)$. Finally, $\mathcal{L}_3$ is defined similarly. `CLEFIA-128` has 18 rounds, thus the last 4-round iterative characteristic (for the rounds 17,18) will be terminated after the second round, with an output difference $(\Delta_2, \Delta_3, \Delta_4, \Delta_1)$. It differs from ΔL only in the order of the four words, hence we introduce $\pi_3 : (0,1,2,3) \to (3,1,0,2)$ and conclude that $\mathcal{L}_3(D) = \pi_3(D)$.

In the weak-key class the pairs of keys are defined as $(K, K \oplus \pi(D) \oplus \Sigma(D))$ and for any plaintext P, it holds

$$E_K(P) \oplus E_{K \oplus \pi(D) \oplus \Sigma(D)}(P \oplus \pi_2(D)) = \pi_3(D). \qquad (5)$$

A pair of keys belongs to this class if for any of the 2^{14} values $D = \Delta L$ defined by Equation (3), the 12-round Feistel $GFN_{4,12}$ in the key schedule, on input difference $\Delta K = \pi(\Delta L) \oplus \Sigma(\Delta L)$ gives the output difference ΔL, i.e. $GFN_{4,12}(K \oplus \pi(\Delta L) \oplus \Sigma(\Delta L)) \oplus GFN_{4,12}(K) = \Delta L$. Therefore not all of the keys K have a related key and form a pair in the weak-key class, but only those for which the differential in the Feistel permutation holds.

We deal with a 12-round Feistel permutation and thus the probability of the differential $\pi(\Delta L) \oplus \Sigma(\Delta L) \to \Delta L$ is low. We assume it is 2^{-128} (as proven by the designers), which is the probability of getting fixed output difference from a fixed input difference in a random permutation. However, even when we model the Feistel permutation by a random one, *there still exist* 2^{14} *key schedule differentials that have a probability of* 2^{-128} *and that result in iterative round key differences.*

In `CLEFIA-128`, there are 2^{128} possible keys K, and therefore for a specific value of D, the number of related keys $(K, K \oplus \pi(D) \oplus \Sigma(D))$ is the same. The probability of the differential in the Feistel permutation is 2^{-128}, thus among all of the pairs, only one will pass the differential. However, there are 2^{14} possible values for D, hence the size of the weak-key class is 2^{14}.

4 Membership Test Distinguisher

An attack technique that succeeds when the related keys belong to the weak-key class is called a membership test. For the weak-key class of `CLEFIA-128`, the membership test will be a differential distinguisher that succeeds always and whose data, time and memory complexities are equal to 2^8. That is to say that we can decide with probability 1 whether the underlying cipher is `CLEFIA-128` or other (possibly ideal) cipher.

Given a pair of weak keys $(K, K \oplus \pi(D) \oplus \Sigma(D))$, it is easy to distinguish `CLEFIA-128` (see Equation (5)) with only a single pair of related plaintexts $(P, P \oplus \pi_2(D))$ but D has to be known. If it is unknown, we will have to try all 2^{14} possible values of D (as D coincides with one of ΔL_i). Consequently, we are going to end up with a brute force attack on the space of weak keys. To address this problem, we have to be able to detect the correct value of ΔL efficiently.

Finding the correct ΔL_i can be performed much faster if we take into account the additional properties of the difference in the intermediate key. All 2^{14} values of ΔL_i (see Equation (3)) can be defined as XOR of two elements from two different sets each of cardinality 2^7 as shown below

$$\begin{aligned}\Delta L_i = \Delta L_i(a_1, a_2) =& a_1 a_2 t b_2 b_1 b_2 b_1 b_2 b_1 b_2 a_2 a_1 a_2 a_1 a_2 a_1 a_2 t b_1 b_2 = \\ =& G^1(a_1) \oplus G^2(a_2), \\ & a_1 = 0, \ldots, 2^7 - 1, a_2 = 0, \ldots, 2^7 - 1,\end{aligned}$$

where $G^1(a_1)$ is a 128-bit word that is the same as ΔL on the bits that depend on a_1 and has 0's for the bits that depend on a_2 while $G^2(a_2)$ is the opposite, i.e. coincides with ΔL on bits for a_2 and has 0's for bits that depend on a_1[9].

Using the representation helps to detect the correct ΔL by finding collisions on two specific sets. Assume the pair $(K, \tilde{K} = K \oplus \pi(\Delta L) \oplus \Sigma(\Delta L))$ belongs to the weak-key class. For a randomly chosen plaintext P, let us define two pools, each with 2^7 chosen plaintexts:

$$\begin{aligned} P_i^1 &= \pi_2(P \oplus G^1(a_1^i)), a_1^i = 0, 1, \ldots, 2^7 - 1, \\ P_i^2 &= \pi_2(P \oplus G^2(a_2^i)), a_2^i = 0, 1, \ldots, 2^7 - 1. \end{aligned}$$

Next, we obtain two pools of ciphertexts with $(K, \tilde{K})$ as encryption keys, i.e. $C_i^1 = E_K(P_i^1), C_i^2 = E_{\tilde{K}}(P_i^2)$. Finally, we compute two sets V^1, V^2:

$$\begin{aligned} V^1 &= \{V_i^1 | V_i^1 = \pi_2^{-1}(P_i^1) \oplus \pi_3^{-1}(C_i^1)\}, \\ V^2 &= \{V_i^2 | V_i^2 = \pi_2^{-1}(P_i^2) \oplus \pi_3^{-1}(C_i^2)\}. \end{aligned}$$

The crucial observation is that the sets V^1 and V^2 will always collide, i.e. there exist V_i^1 and V_j^2 such that $V_i^1 = V_j^2$. This comes from the following

[9] Recall that each bit of b_1, b_2, t is equal to a single bit of either a_1 or a_2.

sequence:

$$
\begin{aligned}
&V_i^1 \oplus V_j^2 = \\
&= \pi_2^{-1}(P_i^1) \oplus \pi_3^{-1}(C_i^1) \oplus \pi_2^{-1}(P_j^1) \oplus \pi_3^{-1}(C_j^2) = \\
&= \pi_2^{-1}(P_i^1 \oplus P_j^2) \oplus \pi_3^{-1}(E_K(P_i^1) \oplus E_{\tilde{K}}(P_j^2)) = \\
&= \pi_2^{-1}(\pi_2(G^1(a_1^i) \oplus G^2(a_2^i))) \oplus \pi_3^{-1}(E_K(P_i^1) \oplus E_{\tilde{K}}(P_i^1 \oplus \pi_2(G^1(a_1^i) \oplus G^2(a_2^i)))) = \\
&= \Delta L' \oplus \pi_3^{-1}(E_K(P_i^1) \oplus E_{\tilde{K}}(P_i^1 \oplus \pi_2(\Delta L'))),
\end{aligned}
$$

where $\Delta L' = G^1(a_1^i) \oplus G^2(a_2^i)$. Note that $\Delta L'$ can take all possible 2^{14} values (as a_1^i, a_2^j take all 2^7 values), and therefore for some particular i, j, it must coincide with ΔL. In such case, the difference in the plaintext is $\pi_2(\Delta L)$, and thus for the ciphertext we obtain

$$E_K(P_i^1) \oplus E_{\tilde{K}}(P_i^1 \oplus \pi_2(\Delta L)) = \pi_3(\Delta L)$$

Then $V_i^1 \oplus V_j^2 = \Delta L \oplus \pi_3^{-1}(\pi_3(\Delta L)) = 0$.

The possibility to create the sets independently and then to find a collision between them is the main idea of the distinguishing membership test on CLEFIA-128. It works according to the following steps.

1. Choose at random a plaintext P.
2. Create a pool of 2^7 plaintexts $P_i^1 = \pi_2(P \oplus G^1(a_1^i))$ and ask for the corresponding ciphertext C_i^1 obtained with encryption under the first key, i.e. $C_i^1 = E_K(P_i^1)$. Compute the set V^1 composed of elements $V_i^1 = \pi_2^{-1}(P_i^1) \oplus \pi_3^{-1}(C_i^1)$.
3. Create a pool of 2^7 plaintexts $P_i^2 = \pi_2(P \oplus G^2(a_2^i))$ and ask for the corresponding ciphertext C_i^2 obtained with encryption under the second key, i.e. $C_i^2 = E_{\tilde{K}}(P_i^2)$. Compute the set V^2 composed of elements $V_i^2 = \pi_2^{-1}(P_i^2) \oplus \pi_3^{-1}(C_i^2)$.
4. Check for collisions between V^1 and V^2. If such a collision exists, then output that the examined cipher is CLEFIA-128. Otherwise, it is an ideal cipher.

The total data complexity of the membership test is $2^7 + 2^7 = 2^8$ plaintexts. The time complexity of each of the steps 2,3 is 2^7 encryptions, while the collision at step 4 can be found with 2^7 operations and 2^7 memory that is used to store one of the sets V^1 or V^2. Therefore, given a pair of keys from the weak-key class, we can distinguish CLEFIA-128 in 2^8 data, time and memory.

To confirm the correctness of the distinguisher, we implemented it for a small-scale variant of CLEFIA-128. Each word was shrunk to 8-bit value, thus the whole state became 32 bits. The Sbox from AES was taken as the round function F, and random 8-bit values were chosen as constants. The chunks in the linear function Σ were taken of size $5, 11$ (compared to the $7, 57$ in the original version). The expected size of the weak-key class in this toy version is 2^{10} (because $X = \Sigma^2(X)$ has 2^{10} solutions), while in practice we obtained $960 = 2^{9.9}$ solutions. For a random key pair chosen from this class, we were able to distinguish the cipher after 2^6 encryptions which confirms our findings to a large extend.

5 Distinguisher for `CLEFIA-128`

The weak-key class can be used to distinguish the cipher when the oracle can be asked to change the pair of related keys. After repeating this step certain number of times, if the oracle is `CLEFIA-128`, it will hit a pair from the weak-key class which then will be used with the membership test to distinguish the cipher from ideal. However, the relation between the two keys is not fixed (the XOR difference can take 2^{14} possible values) thus a straightforward application of the above idea will fail due to the low probability of randomly hitting the weak-key class. We know that any weak key happens with probability 2^{-128}.

Our idea is to generate the data (pairs of plaintext-ciphertext) from independent keys and then look for a special set of differences among all the data. To achieve this we use the linear relations given by Equation (5). In the membership test, we use the fact that each difference ΔL_i is an XOR of two elements (defined as $G^1(a_1)$ and $G^2(a_2)$) from sets of size 2^7, i.e. $\Delta L = G^1(a_1) \oplus G^2(a_2)$. A similar fact holds for ΔK:

$$\begin{aligned}\Delta K = \pi(\Delta L) \oplus \Sigma(\Delta L) &= \pi(G^1(a_1) \oplus G^2(a_2)) \oplus \Sigma(G^1(a_1) \oplus G^2(a_2)) = \\ &= [\pi(G^1(a_1)) \oplus \Sigma(G^1(a_1))] \oplus [\pi(G^2(a_2)) \oplus \Sigma(G^2(a_2))] = \\ &= T^1(a_1) \oplus T^2(a_2),\end{aligned}$$

where $T^1(a_1) = \pi(G^1(a_1)) \oplus \Sigma(G^1(a_1)), T^2(a_2) = \pi(G^2(a_2) \oplus \Sigma(G^2(a_2))$ are two linear functions (as π, Σ, G^1, G^2 are linear), and therefore the difference in the keys is an XOR of two sets as well.

Using this idea, we can describe the distinguisher for `CLEFIA-128` as follows:

1. Ask the secret key K to be fixed and randomly choose a plaintext P.
2. Create 2^7 plaintexts $P_i^1 = \pi_2(P \oplus G^1(a_1^i))$ and ask for the 2^7 corresponding ciphertexts C_i^1 obtained under the secret keys $K \oplus T^1(a_1^i)$,

i.e. $C_i^1 = E_{K \oplus T^1(a_1^i)}(P_i^1)$. Compute the set V^1 composed of elements $V_i^1 = \pi_2^{-1}(P_i^1) \oplus \pi_3^{-1}(C_i^1)$.

3. Create 2^7 plaintexts $P_i^2 = \pi_2(P \oplus G^2(a_2^i))$ and ask for the 2^7 corresponding ciphertexts C_i^2 obtained under the secret keys $K \oplus T^2(a_2^i)$, i.e. $C_i^2 = E_{K \oplus T^2(a_2^i)}(P_i^2)$. Compute the set V^2 composed of elements $V_i^2 = \pi_2^{-1}(P_i^2) \oplus \pi_3^{-1}(C_i^2)$.
4. Check for collisions between V^1 and V^2. If such a collision exists, then confirm the key pair is weak on an additional pair of plaintexts. If so, then output that the examined cipher is CLEFIA-128.
5. Go to Step 1 if Steps 1-4 are repeated less than 2^{114} times. Otherwise, output that the examined cipher is ideal.

The attack works in 2^{122} time and data, 2^7 memory, and it requires encryption under 2^{122} different secret keys. The validity of the distinguishing attack is proven in Appendix B and tested on the small-scale CLEFIA-128, described in the previous section. As the weak-key class in this toy version has $2^{9.9}$ pairs, the size of the state and the key is 32 bits, the expected complexity of the distinguisher is $2^{32-4.95+1} = 2^{28.05}$. Our experiments confirmed this finding: the average complexity over 100 trials was $2^{28.3}$.

6 Applications and Open Problems

Let us examine implications of our findings. In particular, we look at cryptographic constructions based on block ciphers and their security level when the block cipher is instantiated by CLEFIA-128.

First we focus on cryptographic hashing. More precisely, we consider hashing based on single-block-length[10] modes, where a compression function is built from a block cipher. If the compression function uses CLEFIA-128 then we can find a pair of weak keys in 2^{122} time using the described distinguisher. Once such pair (K_1, K_2) is found, we can produce any number of differential multicollisions [5]:

$$E_{K_1}(P_1^i) = C_1^i, E_{K_1}(P_2^i) = C_2^i, P_1^i \oplus P_2^i = \Delta_P, C_1^i \oplus C_2^i = \Delta_C, i = 1, 2, \ldots, .$$

Note that we do not need to encrypt plaintext pairs as it is sufficient to take pairs with the input difference $\Delta P = \Delta L$ (when the difference in the key pair is $\pi(\Delta L) \oplus \Sigma(\Delta L)$) and then the ciphertext difference must be $\Delta C = \pi_3(\Delta L)$. Consequently, we can produce an arbitrary number of differential multicollisions with the complexity 2^{122}. Note that the

[10] The state and key sizes in CLEFIA-128 coincide, thus we can construct only single-block-length compression functions.

proven lower bound (see [5]) in the case of ideal cipher is 2^{128}. We stress that the compression functions of all 12 modes investigated by Preneel et al. [17], including the popular Davies-Meyer, Matyas-Meyer-Oseas modes, are vulnerable to differential multicollisions due to the fact that all three differences (for plaintext, key, and ciphertext) are fixed by the distinguisher.

A distinguisher for the hashing based on `CLEFIA-128` has already been presented by Aoki at ISITA'12 [2]. It works in the framework of middletext distinguishers [16] (open-key version of the integral attack), where the attacker starts with a set of particularly chosen states in the middle of the cipher, then from them (and the knowledge of the key) produces the set of plaintexts and the set of ciphertexts, and finally shows that these two sets have some property that cannot be easily reproduced if the cipher was ideal. For `CLEFIA-128`, Aoki shows how to choose 2^{112} starting middle states that result in 17-round middletext distinguisher, and then adds one more round where he uses subkey guesses, to obtain the 18-round distinguisher. We want to point out that there is a substantial difference, between our result and that of Aoki. We do not fix the values neither of the plaintexts nor of the ciphertexts, and our distinguisher works as long as the pair of plaintexts has the required difference – the values can be arbitrary and even secret.

These findings suggest that our distinguisher can be used in the secret-key mode (in addition to the weak-key secret mode). However, even though the complexity is below the generic (our distinguisher has complexity 2^{122}, whereas the generic complexity is 2^{128}), the caveat is in the number of queries under different keys: to launch the distinguisher we need 2^{122} oracles queries under 2^{122} different keys. Most of the published related-key attacks require smaller number of queries under different keys, thus they seem valid without a proof in the generic case. Nevertheless, in case when this number is large, the generic lower bound is unknown. Thus we want to propose the following open problem:

Open problem 1 *For an arbitrary block cipher $E_K(P)$ with k-bit key and n-bit state, what is the lower bound on the time complexity T and the number of queries under different keys DQ, required to distinguish (with a significant advantage) the cipher from ideal.*

Biham [4] proposed a generic attack with $T = DQ = 2^{k/2}$ or more general $T \cdot DQ = 2^k$. The steps of the attack are as follows: randomly choose a plaintext P, encrypt P offline under $2^{\frac{k}{2}}$ different keys, and ask online for the encryptions of P under $2^{\frac{k}{2}}$ different keys. If the oracle is the target cipher then a collision will occur between the offline and online ciphertexts, which would result in a distinguisher[11].

There is no proof that Biham's attack is optimal, thus the lower bound (Open Problem 1) is still unknown. However, this generic attack already provides the bound $2^{k/2}$ and as a result, a weak-key class of a size smaller than $2^{\frac{k}{2}}$, cannot be used to attack the cipher. It is enough to observe that to exploit the weak-key class, any attack first requires a hit in the class and thus at least $2^{k-\frac{k}{2}} = 2^{\frac{k}{2}}$ queries under different keys. It means that our distinguisher with 2^{122} different keys, cannot be used to attack `CLEFIA-128` itself.

The secret-key security of `CLEFIA-128`, however, can be analyzed in a framework, where the cipher is used as a part of a larger construction. One such example is PRINCE [9] with a 128-bit key block cipher based on the 64-bit key PRINCE^{core} cipher and defined as

$$PRINCE_{K_1,K_2}(P) = PRINCE^{core}_{K_1}(P \oplus K_2) \oplus \Lambda(K_2),$$

where Λ is a linear function. The authors gave security proofs of the construction in the single-key model when PRINCE^{core} is assumed to be ideal. However, the construction is insecure for related-key attacks as trivial distinguishers [12] exist.

To fix this security weakness, we alter the PRINCE construction. We replace the linear function Λ with a random permutation Q and obtain a cipher $\mathrm{FXR}_{K_1,K_2}(P)$ with $2n$-bit keys based on a cipher $E_K(P)$ with n-bit keys:

$$\mathrm{FXR}_{K_1,K_2}(P) = E_{K_1}(P \oplus K_2) \oplus Q(K_2).$$

In the single-key model, the proof of security of FXR is identical to the proof of PRINCE. The related-key distinguishers applicable to PRINCE, cannot be applied to FXR as a difference in K_2 has to go through the random permutation Q. To recover K_2 one can try to launch an attack similar to the attack of Daemen [10] on Even-Mansour [11], but this would require first a guess of the key K_1. However, guessing K_1 (used in the smaller cipher) immediately results in an attack with complexity 2^n, which is already worse (or at least equal) to Biham's attack. Thus we conjecture that this construction significantly increases (from $2^{n/2}$ to 2^n) the complexity of the generic attack:

[11] A false positive can easily be discarded by encrypting a few more plaintexts with the same key.

[12] A difference Δ in the plaintext P and in the key K_2 results in a difference $\Lambda(\Delta)$ in ciphertext C.

Open problem 2 (Conjecture) *If the bound of Biham's attack is optimal, then the simultaneous lower bound on the time complexity and different key queries required to distinguish the FXR construction instantiated with n-bit state/key ideal cipher, is $T = DQ = 2^n$.*

If the conjecture is correct, then weak-key classes of any size could be used to show weakness of cipher in the secret-key model. For instance, when CLEFIA-128 is used in FXR, then we can distinguish FXR with 2^{122} time complexity and different key queries, whereas the case of ideal cipher would require $2^n = 2^{128}$, hence CLEFIA-128 would be insecure in the secret-key model.

7 Conclusion

We have presented a cryptanalysis of the full-round CLEFIA-128. The analysis shows existence of a weak-key class that consists of 2^{14} pairs of related keys. We have shown how to exploit the pairs in four different scenarios, depending on the model (hashing mode or secret-key mode) and on the type of pairs of keys (randomly chosen pair or weak-key pair) that are implemented by the oracle. In the hashing mode (or open-key mode in general) we have shown that when the pair belongs to the weak-key class, then the encryption functions can be distinguished with only two queries to the oracle. To find one pair we need around 2^{122} encryptions, and such pair can be used to produce differential multicollisions faster than the generic 2^{128}. In the secret-key model with a pair from the weak-key class, we can distinguish CLEFIA-128 from an ideal cipher with 2^8 time and data complexity, compared to the generic 2^{14}. In the same model, but with keys chosen randomly, we need 2^{122} time, data, and queries under different keys to distinguish the cipher. Here the generic attack performs better, however, we have shown that potentially there are constructions, where our distinguisher has lower complexity. The main ideas of the analysis in all four scenarios have been verified with computer experiments on small-scale variants of CLEFIA-128.

The analysis is invariant of important security features that presumably increase the strength of a cipher. First, the non-linear part of the key schedule can be any random permutation (not necessarily a 12-round Feistel). Our attacks would still work as we do not need high probability differentials for this permutation. Next, the state update functions (in CLEFIA-128 F_0, F_1 are one round substitution-permutation networks) can be arbitrary functions or permutations, including several layers of SP – the difference never goes into them, hence, the probability of the characteristic in the state would stay 1. Finally, the number of rounds in CLEFIA-128 plays absolutely no role in our analysis – even if CLEFIA-128 had 1000 rounds, the complexity of the attacks would stay the same.

To prevent future attacks as ours, we have to clearly understand what are the main drawbacks of the design. The weak-key class and the three attack invariances are results of these drawbacks (not their cause) and provide clues on what the actual cause might be. The invariance of the state update function is due to the Feistel structure of the cipher – this construction can lead to probability 1 characteristics as it can cancel round key and state differences. To maintain the cancellation through arbitrary number of rounds (invariance of the number of rounds), the round key differences have to be iterative. The key schedule prevents high probability iterative (or any fixed value) differences as they have to be produced from a difference in the key that goes initially through a 12-round Feistel modeled as random permutation. The Feistel, however, produces low probability (2^{-128}) differences (invariance of the random permutation), and 2^{14} of them become iterative round key differences due to the linear function used after the Feistel. That is, because of the linear function, with 2^{-128} we can have a special type of differences in 36 rounds keys (1152 bits !). Therefore, the analysis of CLEFIA-128 holds due to the Feistel structure of the cipher and the weak linear function that is used to produce the round keys.

To conclude, our work shows that *low probability differentials (around 2^{-k} for a cipher with k-bit key and n-bit state) for the key schedule of Feistel ciphers, cannot be used as a sole proof of resistance against related-key differential attacks.* A safe upper bound on the probability of such differentials, which proves and provides security against related-key attacks, is not 2^{-k} but 2^{-2k-n} – this comes from the fact that there can be as many as 2^{2k} pairs of weak keys, and their combined probability should be below 2^{-n}.

References

1. RFC 6114. CLEFIA. http://www.rfc-editor.org/rfc/rfc6114.txt.
2. K. Aoki. A middletext distinguisher for full CLEFIA-128. In *ISITA*, pages 521–525. IEEE, 2012.
3. M. Bellare and T. Kohno. A theoretical treatment of related-key attacks: RKA-PRPs, RKA-PRFs, and applications. In E. Biham, editor, *EUROCRYPT*, volume 2656 of *Lecture Notes in Computer Science*, pages 491–506. Springer, 2003.
4. E. Biham. How to decrypt or even substitute DES-encrypted messages in 2^{28} steps. *Inf. Process. Lett.*, 84(3):117–124, 2002.

5. A. Biryukov, D. Khovratovich, and I. Nikolic. Distinguisher and related-key attack on the full AES-256. In S. Halevi, editor, *CRYPTO*, volume 5677 of *Lecture Notes in Computer Science*, pages 231–249. Springer, 2009.
6. A. Biryukov and I. Nikolic. Complementing Feistel ciphers. *FSE 2013*. To appear.
7. A. Bogdanov, H. Geng, M. Wang, L. Wen, and B. Collard. Correlation linear cryptanalysis with FFT and improved attacks on ISO standards Camellia and CLEFIA. *Selected Areas of Cryptography 2013*. To appear.
8. A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw, Y. Seurin, and C. Vikkelsoe. PRESENT: An ultra-lightweight block cipher. In P. Paillier and I. Verbauwhede, editors, *CHES*, volume 4727 of *Lecture Notes in Computer Science*, pages 450–466. Springer, 2007.
9. J. Borghoff, A. Canteaut, T. Güneysu, E. B. Kavun, M. Knezevic, L. R. Knudsen, G. Leander, V. Nikov, C. Paar, C. Rechberger, P. Rombouts, S. S. Thomsen, and T. Yalçin. PRINCE - a low-latency block cipher for pervasive computing applications - extended abstract. In X. Wang and K. Sako, editors, *ASIACRYPT*, volume 7658 of *Lecture Notes in Computer Science*, pages 208–225. Springer, 2012.
10. J. Daemen. Limitations of the Even-Mansour construction. In Imai et al. [12], pages 495–498.
11. S. Even and Y. Mansour. A construction of a cioher from a single pseudorandom permutation. In Imai et al. [12], pages 210–224.
12. H. Imai, R. L. Rivest, and T. Matsumoto, editors. *Advances in Cryptology - ASIACRYPT '91, International Conference on the Theory and Applications of Cryptology, Fujiyoshida, Japan, November 11-14, 1991, Proceedings*, volume 739 of *Lecture Notes in Computer Science*. Springer, 1993.
13. ISO/IEC 29192-2. Information technology - Security techniques - Lightweight cryptography - Part 2: Block ciphers. `http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56552`.
14. Y. Li, W. Wu, and L. Zhang. Improved integral attacks on reduced-round CLEFIA block cipher. In S. Jung and M. Yung, editors, *WISA*, volume 7115 of *Lecture Notes in Computer Science*, pages 28–39. Springer, 2011.
15. H. Mala, M. Dakhilalian, and M. Shakiba. Impossible differential attacks on 13-round CLEFIA-128. *J. Comput. Sci. Technol.*, 26(4):744–750, 2011.
16. M. Minier, R. C.-W. Phan, and B. Pousse. Distinguishers for ciphers and known key attack against Rijndael with large blocks. In B. Preneel, editor, *AFRICACRYPT*, volume 5580 of *Lecture Notes in Computer Science*, pages 60–76. Springer, 2009.
17. B. Preneel, R. Govaerts, and J. Vandewalle. Hash functions based on block ciphers: A synthetic approach. In D. R. Stinson, editor, *CRYPTO*, volume 773 of *Lecture Notes in Computer Science*, pages 368–378. Springer, 1993.
18. T. Shirai, K. Shibutani, T. Akishita, S. Moriai, and T. Iwata. The 128-bit blockcipher CLEFIA (Extended Abstract). In A. Biryukov, editor, *FSE*, volume 4593 of *Lecture Notes in Computer Science*, pages 181–195. Springer, 2007.
19. B. Sun, R. Li, M. Wang, P. Li, and C. Li. Impossible differential cryptanalysis of CLEFIA. *IACR Cryptology ePrint Archive*, 2008:151, 2008.
20. X. Tang, B. Sun, R. Li, and C. Li. Impossible differential cryptanalysis of 13-round CLEFIA-128. *Journal of Systems and Software*, 84(7):1191–1196, 2011.
21. C. Tezcan. The improbable differential attack: Cryptanalysis of reduced round CLEFIA. In G. Gong and K. C. Gupta, editors, *INDOCRYPT*, volume 6498 of *Lecture Notes in Computer Science*, pages 197–209. Springer, 2010.
22. Y. Tsunoo, E. Tsujihara, M. Shigeri, T. Saito, T. Suzaki, and H. Kubo. Impossible differential cryptanalysis of CLEFIA. In K. Nyberg, editor, *FSE*, volume 5086 of *Lecture Notes in Computer Science*, pages 398–411. Springer, 2008.

23. W. Wang and X. Wang. Improved impossible differential cryptanalysis of CLEFIA. *IACR Cryptology ePrint Archive*, 2007:466, 2007.

24. W. Zhang and J. Han. Impossible differential analysis of reduced round CLEFIA. In M. Yung, P. Liu, and D. Lin, editors, *Inscrypt*, volume 5487 of *Lecture Notes in Computer Science*, pages 181–191. Springer, 2008.

A Specification on `CLEFIA-128`

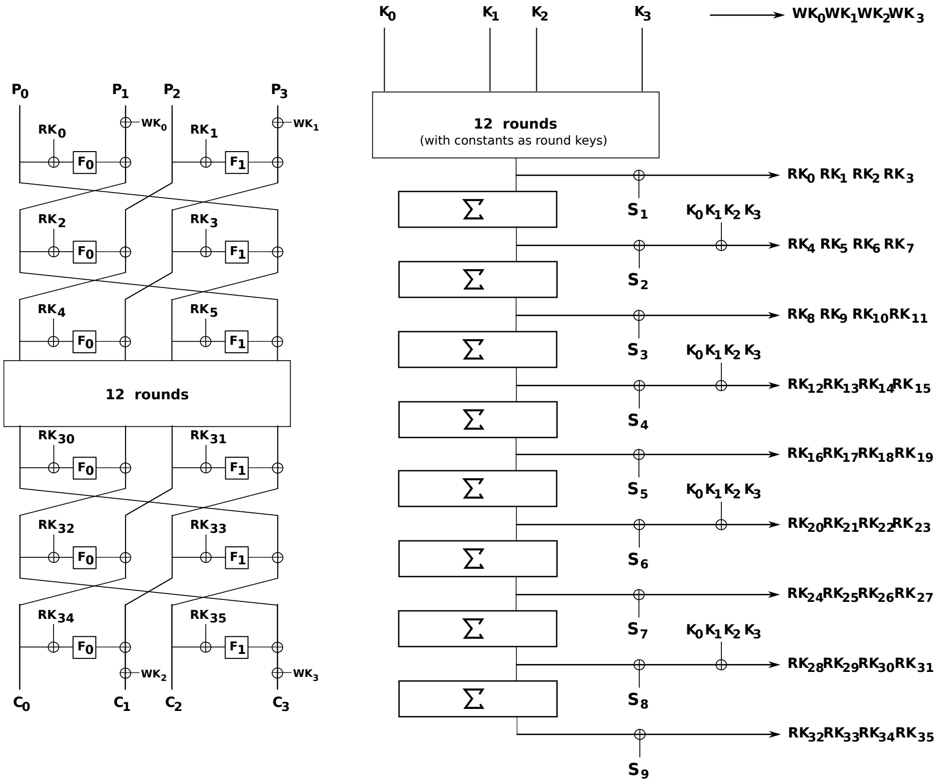

Fig. 3. The encryption function of `CLEFIA-128` at the left, and the key schedule at the right. P_0, P_1, P_2, P_3 are 32-bit plaintext words, C_0, C_1, C_2, C_3 are the ciphertext words, K_0, K_1, K_2, K_3 are the key words, RK_i, WK_j are the round and whitening keys, respectively, and S_i are 128-bit constants. Finally, F_0, F_1 are the two state update functions, while Σ is a linear function (permutation).

B Proof of Correctness for the Distinguisher

Assume that the underlying cipher is `CLEFIA-128`. The encryptions at steps 2 and 3, are done under 2^7 different keys, or in total, 2^{14} pairs of keys. The difference in each key pair is

$$[K \oplus T^1(a_1^i)] \oplus [K \oplus T^2(a_2^j)] = T^1(a_1^i) \oplus T^2(a_2^j) = \Delta K_l,$$

i.e. each key pair has a difference that corresponds to one of the required key differences ΔK. After 2^{114} repetitions of steps 1-4, one of the $2^{14+114} = 2^{128}$ key pairs (with some difference $\Delta K = T^1(a_1^i) \oplus T^2(a_2^j)$) will pass the 12-round Feistel differential (which holds with probability 2^{-128}), and produce the required difference $\Delta L = G^1(a_1^i) \oplus G^2(a_2^j)$ in the intermediate key, which in turn will result in iterative differences in the round keys. Let us take a closer look at the difference in the state of `CLEFIA-128`. As the initial difference in the round keys is ΔL, in order to use the probability 1 characteristic in the state, the difference in the plaintexts should be $\pi_2(\Delta L)$. From the description of the distinguisher (steps 2,3), it follows that the difference in the plaintexts is

$$\begin{aligned} P_i^1 \oplus P_j^2 &= \pi_2(P \oplus G^1(a_1^i)) \oplus \pi_2(P \oplus G^2(a_2^j)) = \\ &= \pi_2(G^1(a_1^i) \oplus G^2(a_2^j)) = \pi_2(\Delta L). \end{aligned}$$

Therefore, by (5) the difference at the ciphertexts will be $\pi_3(\Delta L)$. As in the membership test, the sets V_1, V_2 will collide.

C Analysis of `CLEFIA-128` with Whitening Keys

The whitening keys are the four words $WK_i, i = 0, 1, 2, 3$, defined as $WK_0||WK_1||WK_2||WK_3 = K$, i.e. they are the words of the master key K. The first two are XOR-ed to the second and the fourth plaintext words, and the remaining two to the second and the fourth ciphertext words (see Fig 3).

To index the whitening words, we define two linear functions on 128-bit words (or four 32-bit words). Assume X is 128-bit word, such that $X = a|b|c|d$, where a, b, c, d are 32-bit words. Then $l(X) : \{0,1\}^{128} \to \{0,1\}^{128}$ is defined as $l(X) = l(a|b|c|d) = 0|a|0|b$. Similarly $r(X) : \{0,1\}^{128} \to \{0,1\}^{128}$ is defined as $r(X) = r(a|b|c|d) = 0|c|0|d$.

Now we can easily specify the weak-key class:

- the key difference remains the same,

– the plaintext difference, instead of $\pi_2(\Delta L)$, should be $\pi_2(\Delta L) \oplus l(\Delta K)$,
– the ciphertext difference, instead of $\pi_3(\Delta L)$, should be $\pi_3(\Delta L) \oplus r(\Delta K)$.

As $\Delta K = \pi(\Delta L) \oplus \Sigma(\Delta L)$, it follows that the weak-key class for the original `CLEFIA-128` is defined as 2^{14} pairs of keys $(K, K \oplus \pi(\Delta L) \oplus \Sigma(\Delta L))$ such that for any plaintext P holds:

$$E_K(P) \oplus E_{K \oplus \pi(\Delta L) \oplus \Sigma(\Delta L)}(P \oplus \pi_2(\Delta L) \oplus l(\pi(\Delta L) \oplus \Sigma(\Delta L))) = \pi_3(\Delta L) \oplus r(\pi(\Delta L) \oplus \Sigma(\Delta L)).$$

Let us focus on the membership test. We define the plaintexts pools as:

$$P_i^1 = P \oplus \pi_2(G^1(a_1^i)) \oplus l(T^1(a_1^i)), a_1^i = 0, 1, \ldots, 2^7 - 1,$$
$$P_i^2 = P \oplus \pi_2(G^2(a_2^i)) \oplus l(T^2(a_2^i)), a_2^i = 0, 1, \ldots, 2^7 - 1.$$

This way, the difference between each two plaintext from two different pools is $\pi_2(\Delta L') \oplus l(\Delta K)$, i.e. it is as required by the class.

To define the sets V^1, V^2 that lead to a collision, first we have to understand how a collision can occur. In the previous membership test (on `CLEFIA-128` without whitening keys), we used the trick that the difference in both the plaintext and the ciphertext is ΔL, but with permuted words (that is why we applied π_2^{-1}, π_3^{-1}). Here it is not the same: in the plaintext the difference is ΔL and two more words of ΔK, while in the ciphertext it is ΔL and the remaining two words of ΔK. Hence, XOR of these values does not trivially produce zero as the two words from l and the two from r are different.

Nevertheless, we can achieve collisions. Assume $\Delta L = a|b|c|d$. Then the difference Δ_P in the plaintext is

$$\begin{aligned}\Delta_P =& \pi_2(a|b|c|d) \oplus l(\pi(a|b|c|d) \oplus \Sigma(a|b|c|d)) = \\ & a|c|b|d \oplus l(b|a|d|c) \oplus l(\Sigma(a|b|c|d)) = \\ & a|c + b|b|d + a \oplus l(\Sigma(a|b|c|d)).\end{aligned}$$

Note, $l(\Sigma(a|b|c|d)$ has zeros at the first and at the third words.

Similarly, the difference Δ_C in the ciphertext is

$$\begin{aligned}\Delta_C =& \pi_3(a|b|c|d) \oplus r(\pi(a|b|c|d) \oplus \Sigma(a|b|c|d)) = \\ & c|b|d|a \oplus r(b|a|d|c) \oplus r(\Sigma(a|b|c|d)) = \\ & c|b + d|d|a + c \oplus r(\Sigma(a|b|c|d)).\end{aligned}$$

Again, in the sum r influences only the second and the fourth word.

Let us introduce a function f, that acts on the four 32-bit words of a 128-bit state and it XORs the first word to the fourth word, and the third word to the second word, i.e. $f(x|y|z|t) = (x|y+z|z|t+x)$. Then

$$f(\varDelta_P) = a|c|b|d \oplus l(\varSigma(a|b|c|d)),$$
$$f(\varDelta_C) = c|b|d|a \oplus r(\varSigma(a|b|c|d)).$$

The function $\varSigma$ is linear and therefore $\varSigma(a|b|c|d) = \varSigma(a|0|0|0)+\varSigma(0|b|0|0)+\varSigma(0|0|c|0)+\varSigma(0|0|0|d)$. Let us denote these four values with $\varSigma_a, \varSigma_b, \varSigma_c$, and $\varSigma_d$. Furthermore, with superscripts we denote the four 32-bit words of $\varSigma_x$, e.g. $\varSigma_a^2$ is the second (most significant) word of $\varSigma_a$. This allows us to remove the functions l, r from the terms, and as a result we obtain

$$f(\varDelta_P) = a|c+\varSigma_a^1+\varSigma_b^1+\varSigma_c^1+\varSigma_d^1|b|d+\varSigma_a^2+\varSigma_b^2+\varSigma_c^2+\varSigma_d^2,$$
$$f(\varDelta_C) = c|b+\varSigma_a^3+\varSigma_b^3+\varSigma_c^3+\varSigma_d^3|d|a+\varSigma_a^4+\varSigma_b^4+\varSigma_c^4+\varSigma_d^4.$$

Next, we define a function $g(x|y|z|t)$ that from x, z computes $\varSigma_x^1, \ldots, \varSigma_x^4$, $\varSigma_z^1, \ldots, \varSigma_z^4$ and it adds $\varSigma_x^4, \varSigma_z^4$ to the first word, $\varSigma_x^1, \varSigma_z^1$ to the second, $\varSigma_x^3, \varSigma_z^3$ to the third, and $\varSigma_x^2, \varSigma_z^2$ to the fourth. Similarly, for $\varDelta_C$ we define $h(x|y|z|t)$ that from x, z computes $\varSigma_x^1, \ldots, \varSigma_z^4$ and it adds $\varSigma_x^1, \varSigma_z^1$ to the first word, $\varSigma_x^3, \varSigma_z^3$ to the second, $\varSigma_x^2, \varSigma_z^2$ to the third, and $\varSigma_x^4, \varSigma_z^4$ to the fourth. Thus we get

$$g(f(\varDelta_P)) = a+\varSigma_a^4+\varSigma_b^4|c+\varSigma_c^1+\varSigma_d^1|b+\varSigma_a^3+\varSigma_b^3|d+\varSigma_c^2+\varSigma_d^2,$$
$$h(f(\varDelta_C)) = c+\varSigma_c^1+\varSigma_d^1|b+\varSigma_a^3+\varSigma_b^3|d+\varSigma_c^2+\varSigma_d^2|a+\varSigma_a^4+\varSigma_b^4.$$

Obviously $h(f(\varDelta_C)) = \pi_4(g(f(\varDelta_P)))$, where $\pi_4(0,1,2,3) \rightarrow (3,0,1,2)$. Therefore the sets V_1, V_2 are defined as:

$$V^1 = \{V_i^1 | V_i^1 = \pi_4(g(f(P_i^1))) \oplus h(f(C_i^1))\},$$
$$V^2 = \{V_i^2 | V_i^2 = \pi_4(g(f(P_i^2))) \oplus g(f(C_i^2))\},$$

and a collision between this two sets suggests that $\varDelta L'$ coincides with $\varDelta L$. Thus the membership test for `CLEFIA-128` with whitening keys has the same complexity as before (without whitening).

For the distinguisher of `CLEFIA-128` with whitening keys, we proceed the same way as in the attack on `CLEFIA-128` without whitening keys, but look for collisions on the sets V_1, V_2 defined as above.

We stress out once again that the above analysis was confirmed as well with our computer simulation on the small scale variant of `CLEFIA-128`. The distinguisher required on average $2^{28.2}$ time and data, and 2^5 memory to detect the cipher.