Improved Integral Attacks on MISTY1*

Xiaorui Sun and Xuejia Lai

Department of Computer Science
Shanghai Jiao Tong University
Shanghai, 200240, China
sunsirius@sjtu.edu.cn, lai-xj@cs.sjtu.edu.cn

Abstract. We present several integral attacks on MISTY1 using the FO Relation. The FO Relation is a more precise form of the Sakurai-Zheng Property such that the functions in the FO Relation depend on 16-bit inputs instead of 32-bit inputs used in previous attacks, and that the functions do not change for different keys while previous works used different functions.

We use the FO Relation to improve the 5-round integral attack. The data complexity of our attack, 2^{34} chosen plaintexts, is the same as previous attack, but the running time is reduced from 2^{48} encryptions to $2^{29.58}$ encryptions. The attack is then extended by one more round with data complexity of 2^{34} chosen plaintexts and time complexity of $2^{107.26}$ encryptions. By exploring the key schedule weakness of the cipher, we also present a chosen ciphertext attack on 6-round MISTY1 with all the FL layers with data complexity of 2^{32} chosen ciphertexts and time complexity of $2^{126.09}$ encryptions. Compared with other attacks on 6-round MISTY1 with all the FL layers, our attack has the least data complexity.

1 Introduction

The MISTY1 algorithm is a block cipher with a 64-bit block size and a 128-bit key size proposed by Matsui [8]. It was recommended by the European NESSIE project and the CRYPTREC project, and became an ISO standard in 2005. The cipher generally uses an 8-round Feistel structure with a round function FO. Before each odd round and after the last round, there is an additional FL layer.

Many cryptanalysis results on MISTY1 have been published [1, 2, 3, 4, 5, 6, 7, 10, 11]. The integral attack on 5 rounds with all but the last FL layers [4] requires 2^{34} chosen plaintexts, and has a time complexity of 2^{48} encryptions. The impossible differential attack on 6 rounds with all the FL layers [2] requires 2^{51} chosen plaintexts, and has a time complexity of $2^{123.4}$ encryptions. With all the FL functions absent, the impossible differential attack [2] could break the 7 rounds with data complexity of $2^{50.2}$ known plaintexts and time complexity of $2^{114.1}$ encryptions.

* This work was supported by NSFC Grant No.60573032, 60773092 and 11th PRP of Shanghai Jiao Tong University.

M.J. Jacobson Jr., V. Rijmen, and R. Safavi-Naini (Eds.): SAC 2009, LNCS 5867, pp. 266–280, 2009.

In this paper, we present several integral attacks using a more precise form of the variant Sakurai-Zheng Property for the round function FO. We call this new property the FO Relation. Sakurai-Zheng Property was founded by Sakurai and Zheng in [9]. Knudsen and Wagner used a variation of this property for the FO function to attack the 5-round MISTY1 [4]. Compared with the variant Sakurai-Zheng property, there are two merits of the FO Relation: the inputs of the functions in the FO Relation are shortened from 32 bits to 16 bits, and these functions do not change for different keys while the previous property used different functions for different keys.

We use this new relation to improve the integral attack [4] on 5 rounds with all but the last FL layers. The data complexity of our improved attack is 2^{34} chosen plaintexts, and the time complexity of the attack is $2^{29.58}$ encryptions. Compared with the 5-round integral attack [4], the time complexity of our attack is reduced from 2^{48} encryptions to $2^{29.58}$ encryptions with the same data complexity.

Next, we extend the 5-round attack by one more round. Using the equivalent description of the FO function [5, 11] and the FO Relation, we modify the partial decryption process of computing the required intermediate values to reduce the key bits needed.The data complexity of this 6-round attack is 2^{34} chosen plaintexts, and the time complexity of the attack is $2^{107.26}$ encryptions.

Table 1. Attacks on MISTY1

Rounds	Attack	FL functions	Data	Time	Ref.
5	Higher-Order Differential	None	$2^{10.5}$ CP	2^{17}	[1]
6	Impossible Differential	None	2^{54} CP	2^{61}	[5]
6	Impossible Differential	None	2^{39} CP	2^{106}	[5]
6	Impossible Differential	None	2^{39} CP	2^{85}	[7]
7	Impossible Differential	None	$2^{50.2}$ KP	$2^{114.1}$	[2]
4	Impossible Differential	Most	2^{23} CP	$2^{90.4}$	[5]
4	Impossible Differential	Most	2^{38} CP	2^{62}	[5]
4	Collision Search	Most	2^{20} CP	2^{89}	[5]
4	Collision Search	Most	2^{28} CP	2^{76}	[5]
$4^{\dagger}$	Slicing	All	$2^{22.25}$ CP	2^{45}	[6]
4	Slicing	All	$2^{27.2}$ CP	$2^{81.6}$	[6]
4	Impossible Differential	All	$2^{27.5}$ CP	2^{116}	[6]
5^{*}	Integral	Most	2^{34} CP	2^{48}	[4]
$5^{\dagger}$	Impossible Differential	All	2^{38} CP	$2^{46.45}$	[2]
6	Impossible Differential	All	2^{51} CP	$2^{123.4}$	[2]
5^{*}	Integral	Most	2^{34} CP	$2^{29.58}$	Section 4
6	Integral	Most	2^{34} CP	$2^{107.26}$	Section 5
6	Integral	All	2^{32} CC	$2^{126.09}$	Section 6

KP - Known Plaintext CP - Chosen Plaintext CC - Chosen Ciphertext

None - the version of MISTY1 without all the FL layers

Most - the version of MISTY1 without the final FL layer

All - the version of MISTY1 with all the FL layers

† - the attack retrieves 41.36 bits of information about the key.

* - the attack retrieves 50 bits of information about the key.

We also provide an attack on 6 rounds with all the FL layers. The attack is a chosen ciphertext attack starting from the FL_3, FL_4 layer to the end of the cipher. We explore the key schedule weakness to speed up the computation of the required intermediate values. The data complexity of the attack is 2^{32} chosen ciphertexts, and the time complexity of the attack is $2^{126.09}$ encryptions. Compared with other attacks on 6 rounds with all the FL layers, our attack has the least data complexity. The summarization of our attacks and previous attacks is listed in Table 1, where the data complexity is measured by the number of plaintexts/ciphertexts and the time complexity is measured by the number of encryptions needed in the attack.

The paper is organized as follows: In Section 2 we give a brief description of MISTY1 block cipher. We present the FO Relation in Section 3, and then use this new property to improve the integral attack on 5-round MISTY1 with all but the last FL layers in Section 4. Section 5 extends the 5-round attack to 6 rounds with all but the last FL layers. In Section 6 we present the attack on 6 rounds with all the FL layers. Section 7 concludes this paper.

2 The MISTY1 Block Cipher

MISTY1 is a block cipher with a 64-bit block size and a 128-bit key size. Let P and C denote the 64 bit plaintext and ciphertext, respectively. We use the superscript without brackets to distinguish the values corresponding to different plaintexts, e.g. C^1 and C^2 denote the ciphertexts for P^1 and P^2 respectively. The superscript with brackets denotes the bits of the words, e.g. $C^{1(1\ldots7)}$ denote the left 7 bits of C^1. The subscript(without brackets) is used to distinguish the intermediate values for different rounds.

MISTY1 has a recursive structure. As shown in Figure 1(a), the cipher generally uses an 8-round Feistel structure with a round function FO. In each round, the left 32-bit part is functioned with the FO function, and then is XORed with the right 32 bits. This new 32-bit value is the left 32-bit input of the next round and the right 32-bit input of the next round is the original 32 left bits. Before each odd round and after the last round, there is an additional FL layer. We also use subscripted FO(or FL) to distinguish FO(or FL) functions with different keys, e.g, FO_3 denote FO function keyed with AKO_3 and AKI_3.

In the original specification of MISTY1 [8], the round function FO uses a 112-bit key. Several equivalent descriptions of the FO function [5,11] have been proposed, which use less key bits. Here we use the equivalent FO description similar to [5] as Figure 1(b) and (c). The round function FO_i itself has a 3-round Feistel-like structure. The 32-bit input of FO_i is divided into two blocks of 16 bits, denoted as IL_i and IR_i, respectively. In each round, the left 16-bit part is XORed with a subkey $AKO_{i,1}$ and then functioned with FI using a 9-bit key $AKI_{i,1}$. The output of the FI is XORed with the right 16 bits. The left 16 bits and the right 16 bits are then swaped. The same procedure is repeated three times, the left 16 bits and the right 16 bits after the third round are denoted as ML_i and MR_i. OL_i is the XOR of ML_i and the subkey $AKO_{i,4}$, and OR_i

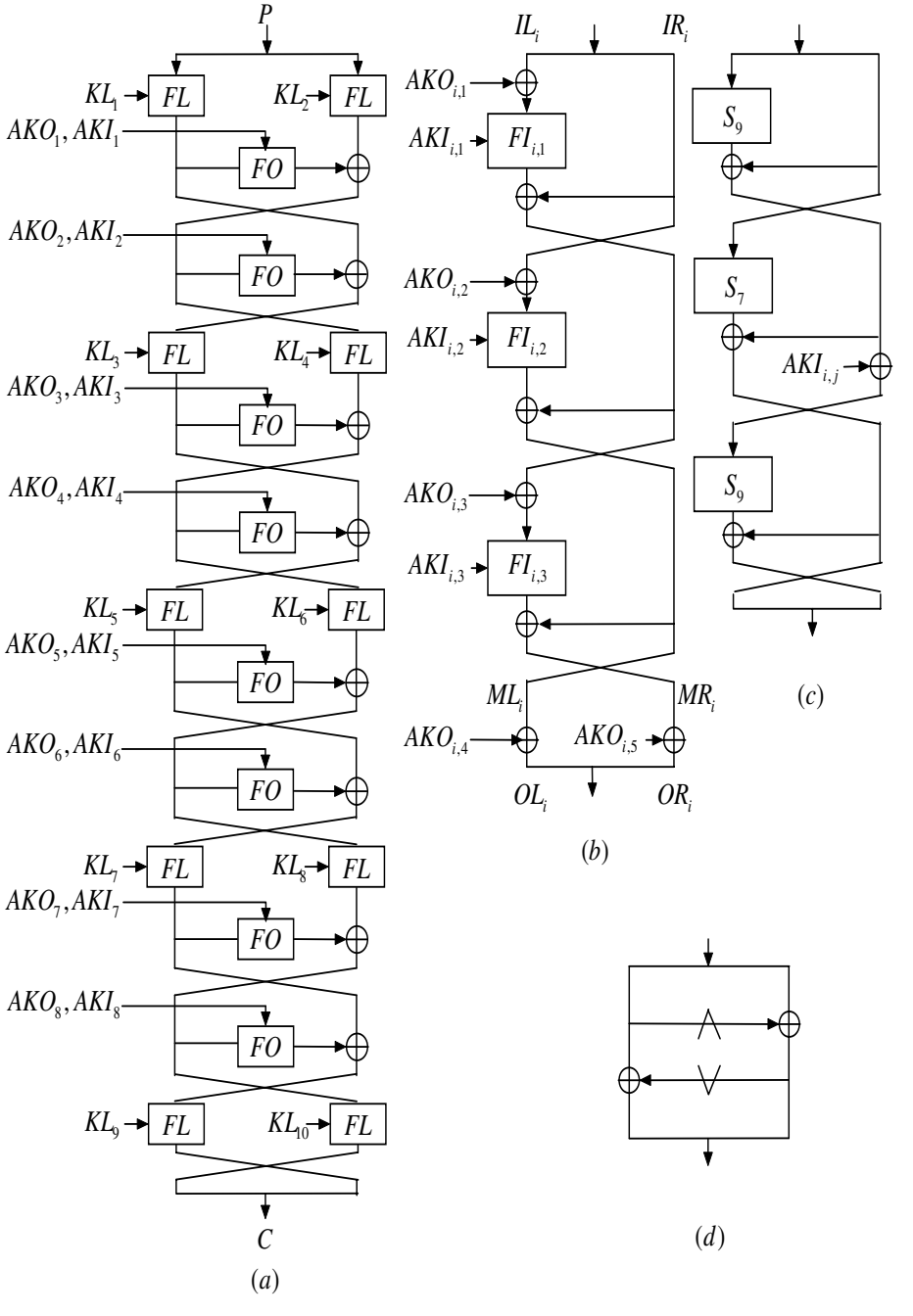

Fig. 1. (a) MISTY1 general structure (b) FO function (c) FI function (d) FL function

Table 2. The Key Schedule for MISTY1

Subkey	Correspondence
$KL_{i,1}$	$K_{\frac{i+1}{2}}$(odd i) or $K'_{\frac{i}{2}+2}$(even i)
$KL_{i,2}$	$K'_{\frac{i+1}{2}+6}$(odd i) or $K_{\frac{i}{2}+4}$(even i)
$AKO_{i,1}$	K_i
$AKO_{i,2}$	K_{i+2}
$AKO_{i,3}$	$K_{i+7} \oplus K'^{(1\ldots7)}_{i+5}||00||K'^{(1\ldots7)}_{i+5}$
$AKO_{i,4}$	$K_{i+4} \oplus K'^{(1\ldots7)}_{i+5}||00||K'^{(1\ldots7)}_{i+5} \oplus K'^{(1\ldots7)}_{i+1}||00||K'^{(1\ldots7)}_{i+1}$
$AKO_{i,5}$	$K'^{(1\ldots7)}_{i+5}||00||K'^{(1\ldots7)}_{i+5} \oplus K'^{(1\ldots7)}_{i+1}||00||K'^{(1\ldots7)}_{i+1} \oplus K'^{(1\ldots7)}_{i+3}||00||K'^{(1\ldots7)}_{i+3}$
$AKI_{i,1}$	$K'^{(8\ldots16)}_{i+5}$
$AKI_{i,2}$	$K'^{(8\ldots16)}_{i+1}$
$AKI_{i,3}$	$K'^{(8\ldots16)}_{i+3}$

is the XOR of MR_i and $AKO_{i,5}$. The output of FO_i is $OL_i||OR_i$ (|| denotes concatenation). All the $AKO_{i,k}(1 \leq k \leq 5)$ are 16-bit subkeys, and all the $AKI_{i,k}(1 \leq k \leq 3)$ are 9-bit subkeys, hence the FO_i function uses a 107-bit key. Since $AKO_{i,5}^{(8\ldots9)}$ is zero and $AKO_{i,5}^{(1\ldots7)}$ is equal to $AKO_{i,5}^{(10\ldots16)}$, the FO_i function actually takes a 98-bit key.

The FI function also has a 3-round Feistel-like structure. In the first round, the left 9-bit input enters a S-box $S9$, and then is XORed with the right 7-bit input(padded two zero bits left to the 7 bits). Swap the left and the right parts. In the second round, the left 7-bit part enters a S-box $S7$ and then is XORed with the right 9 bits(truncated the left 2 bits). The right 9-bit part is XORed with $AKI_{i,1}$ and then is swaped with the left 7 bits. The third round of FI is the same as the first round.

In the FL layer, the left 32 bits and the right 32 bits are put into the FL functions. In each FL function, the 32-bit input is divided into two blocks of 16 bits. The left 16-bit part is ANDed with the subkey $KL_{i,1}$, and then XORed with the right 16 bits to produce the right 16-bit output. This right 16-bit output is ORed with the subkey $KL_{i,2}$ and then XORed with left 16 bits to produce the left 16-bit output.

The key schedule of MISTY1 divides the 128-bit key into eight blocks of 16-bit words K_1, K_2, ..., K_8. Another eight 16-bit words are computed by $K'_i = FI_{K_{i+1}}(K_i)$. The correspondence of these 16 bit words and the subkeys used in the encryption is listed in Table 2.

3 The *FO* Relation

The following proposition on the FO function, which is a variant for Sakurai-Zheng Property [9], is presented in [4].

Proposition 1 ([4]). [1] *For the FO function of round i, the following equation holds*

$$OL_i^{(1...7)} = f_{AKO_{i,1}}(IL_i||IR_i) \oplus g_{AKO_{i,2}}(IL_i||IR_i) \oplus k \tag{1}$$

where $f_{AKO_{i,1}}$ and $g_{AKO_{i,2}}$ are functions related to the subkeys $AKO_{i,1}$ and $AKO_{i,2}$, respectively, and k is a constant related to the key used in this FO_i function.

As shown in Figure 2, ML_i is the XOR of the values corresponding to the point α and β, hence $OL_i^{(1...7)} = \alpha \oplus \beta \oplus AKO_{i,4}^{(1...7)}$. Since the left 7 bits of the values at α and β are not related to $AKI_{i,1}$ and $AKI_{i,2}$, Equation (1) holds by letting $f_{AKO_{i,1}}(IL_i||IR_i)$ correspond to the left 7 bits of the value at α, $g_{AKO_{i,2}}(IL_i||IR_i)$ correspond to the left 7 bits of the value at β and the key-related constant k correspond to $AKO_{i,4}^{(1...7)}$. Expanding $f_{AKO_{i,1}}$ and $g_{AKO_{i,2}}$, Equation (1) is

$$OL_i^{(1...7)} = [FI^{(1...7)}(IL_i \oplus AKO_{i,1}) \oplus IR_i^{(1...7)}] \oplus [FI^{(1...7)}(IR_i \oplus AKO_{i,2})] \oplus AKO_{i,4}^{(1...7)} \tag{2}$$

where $FI^{(1...7)}$ denotes the partial FI function which inputs 16-bit input of FI, and outputs the left 7-bit output of FI. By identical transformation, Equation (2) can be rewritten as follows:

$$\begin{aligned} OL_i^{(1...7)} =& [FI^{(1...7)}(IL_i \oplus AKO_{i,1})] \oplus [FI^{(1...7)}(IR_i \oplus AKO_{i,2}) \oplus (IR_i \oplus AKO_{i,2})^{(1...7)}] \\ & \oplus [AKO_{i,4}^{(1...7)} \oplus AKO_{i,2}^{(1...7)}] \end{aligned} \tag{3}$$

Let $f_{AKO_{i,1}}(IL_i||IR_i)$ be $FI^{(1...7)}(IL_i \oplus AKO_{i,1})$, $g_{AKO_{i,2}}(IL_i||IR_i)$ be $FI^{(1...7)}(IR_i \oplus AKO_{i,2}) \oplus (IR_i \oplus AKO_{i,2})^{(1...7)}$ and the key-related constant k be $AKO_{i,4}^{(1...7)} \oplus AKO_{i,2}^{(1...7)}$, Equation (1) still holds, but $f_{AKO_{i,1}}$ is not related to IR_i and $g_{AKO_{i,2}}$ is not related to IL_i. Hence Proposition 1 can be refined as follows:

Lemma 1 (the FO Relation). *For the FO function of round i, the following equation holds*

$$OL_i^{(1...7)} = f(IL_i \oplus AKO_{i,1}) \oplus g(IR_i \oplus AKO_{i,2}) \oplus k \tag{4}$$

where f and g are two fixed functions, and k is a constant related to the key used in this FO_i function.

The FO Relation can be viewed as an improvement of Proposition 1. There are two folds of the improvement:

[1] The description of this proposition in [4] uses the subkeys $KO_{i,1}$ and $KO_{i,2}$ corresponding to the original form of the FO function described in [8]. Here we use the subkeys $AKO_{i,1}$ and $AKO_{i,2}$ as described in Section 2. The proposition does not change because $AKO_{i,1} = KO_{i,1}$ and $AKO_{i,2} = KO_{i,2}$.

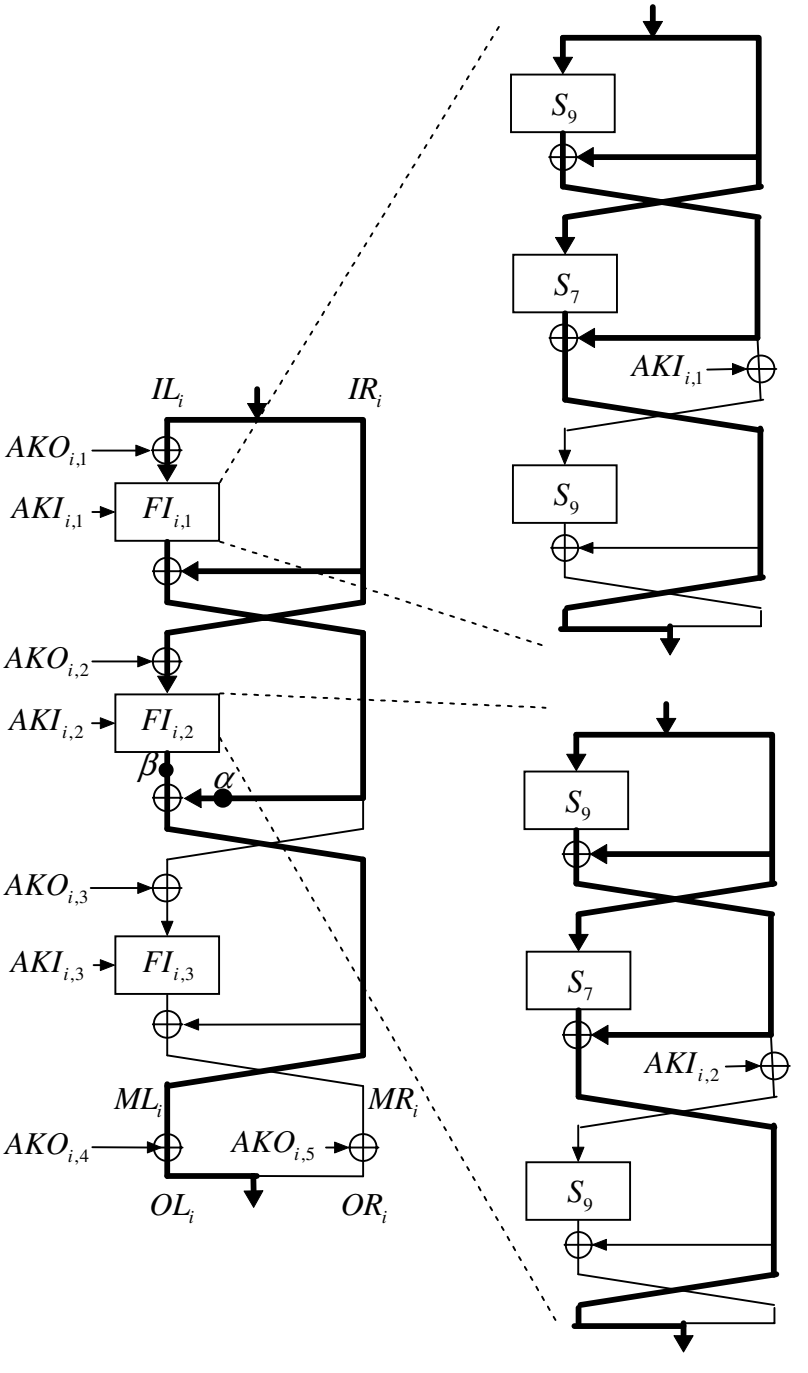

Fig. 2. The FO Relation. The thick lines denote the pathes related to the calculation of $OL_i^{(1\ldots7)}$.

1. The functions f and g used in Lemma 1 rely only on the 16-bit partial input of the FO function instead of the whole 32-bit input used in Proposition 1 (the original Sakurai-Zheng Property [9] is similar to this form, however, Proposition 1 for the FO function proposed in [4] does not have this property).
2. The functions f and g are not related to the subkeys $AKO_{i,1}$ and $AKO_{i,2}$. The Subkeys $AKO_{i,1}$ and $AKO_{i,2}$ are moved into the inputs of the functions f and g.

These two merits will benefit our attack.

Based on the FO Relation, the following theorem can be obtained:

Theorem 1. *Let $IL^1||IR^1$, $IL^2||IR^2$, ...,$IL^{2n}||IL^{2n}$ denote $2n$ inputs of the FO function of round i for some even number $2n$, the following equation holds:*

$$\bigoplus_{j=1}^{2n} OL_i^{j(1...7)} = \bigoplus_{j=1}^{2n} f(IL_i^j \oplus AKO_{i,1}) \oplus \bigoplus_{j=1}^{2n} g(IR_i^j \oplus AKO_{i,2}) \tag{5}$$

This theorem indicates that to obtain the value of $\bigoplus_{j=1}^{2n} OL_i^{j(1...7)}$, we can treat the left 16 bits and the right 16 bits separately to compute the value of $\bigoplus_{j=1}^{2n} f(IL_i^j \oplus AKO_{i,1})$ and $\bigoplus_{j=1}^{2n} g(IR_i^j \oplus AKO_{i,2})$. Based on this theorem, we are ready to present our attacks.

4 Improved Integral Attack on 5-Round MISTY1

The integral attack on 5-round MISTY1 with all but the last FL layers, which was proposed in [4], uses the following four-round integral:

Proposition 2 ([4]). *Consider a structure (named integral structure) of 2^{32} plaintexts where the left 32 bits are held constant and the right 32 bits take on all possible values. The four round integral after FL_6(the XOR of all the 2^{32} 32-bit corresponding intermediate values of the structure after FL_6) is equal to zero.*

The main idea of the previous attack is to partially decrypt the encryptions of the structure and check whether Proposition 2 holds. Proposition 1 shown in Section 3 is used for fast checking whether the left seven bits of the integral are equal to zero predicated by Proposition 2.

We improve the above attack by using the FO Relation, which provides a more efficient method for checking the left seven bits of the integral than Proposition 1. The improved attack is as follows:

1. Ask for the encryptions of four different integral structures. Each structure includes all plaintexts that have the same left 32 bits and all possible right 32 bits.
2. For encryptions of each integral structure:
   (a) For every possible $AKO_{5,1}$, compute the value of $\bigoplus_{j=1}^{2^{32}} f(IL_5^j \oplus AKO_{5,1})$.

(b) For every possible $AKO_{5,2}$, compute the value of $\bigoplus_{j=1}^{2^{32}} g(IR_5^j \oplus AKO_{5,2})$.

(c) Discard all the $AKO_{5,1}, AKO_{5,2}$ pairs such that $\bigoplus_{j=1}^{2^{32}} f(IL_5^j \oplus AKO_{5,1}) \oplus \bigoplus_{j=1}^{2^{32}} g(IR_5^j \oplus AKO_{5,2})$ does not equal to $\bigoplus_{j=1}^{2^{32}} C^{j(1\ldots7)}$.

3. For the remaining $AKO_{5,1}$, $AKO_{5,2}$ pairs, guess all the possible values of $AKI_{5,1}, AKI_{5,2}$ to get full 16 bit $\bigoplus_{j=1}^{2^{32}} OL^j$, discard all guesses such that Proposition 2 is not satisfied.

For each integral structure in Step 2(a), directly computing $\bigoplus_{j=1}^{2^{32}} f(IL_5^j \oplus AKO_{5,1})$ for each possible $AKO_{5,1}$ takes about 2^{48} encryptions. We develop one one technique when implementing this step to reduce the time needed from 2^{48} encryptions to $2^{26.58}$ encryptions for each integral structure as follows.

There are only 2^{16} possible different IL_5 values. For one 16-bit value that occurs an even number of times in all IL_5^j $(1 \leq j \leq 2^{32})$, the XOR of all the corresponding $f(IL_5^j \oplus AKO_{5,1})$ is zero. Hence, in Step 2(a) the attack first counts the occurrences of each 16-bit value in all IL_5^j. Then for each guessed $AKO_{5,1}$, using the 16-bit values that occur odd times in all IL_5^j to compute $\bigoplus_{j=1}^{2^{32}} f(IL_5^j \oplus AKO_{5,1})$.

For an integral structure in Step 2(a), counting the occurrences of every possible 16 bits among all IL_5^j can be accomplished by 2^{32} simple instructions. Since each simple instruction takes about 2^{-6} encryptions, the workload of the counting is 2^{26} encryptions. There are expected 2^{15} different 16 bit values which occur odd times in all IL_5^j. So, for one fixed $AKO_{5,1}$ we could compute $\bigoplus_{j=1}^{2^{32}} f(IL_5^j \oplus AKO_{5,1})$ by computing 2^{15} times function f. If we precompute all the possible value of function f(the time for this preprocess is neglectable compared with the total time complexity), it is possible to use table look-up to speed up. Since one table look-up takes no more than 2^{-6} encryptions, the running time for one integral is no more than $2^{26} + 2^{15} \cdot 2^{-6} \cdot 2^{16} = 2^{26.58}$ encryptions for an integral structure. Hence, Step 2(a) needs about $2^{26.58} \cdot 4 = 2^{28.58}$ encryptions for all the four integral structures. By using similar technique, Step 2(b) also needs about $2^{28.58}$ encryptions.

In Step 2(c), each guess of $AKO_{5,1}$, $AKO_{5,2}$ pair has a probability of 2^{-7} passing the check of an integral structure. For each integral structure, if we generate all the values could pass this check and then check whether they have already been discarded, there are at most $2^{32} \cdot 2^{-7} = 2^{25}$ candidates need to check. Since checking one pair needs only one table look-up, this step needs about $2^{25} \cdot 2^{-6} \cdot 4 = 2^{21}$ encryptions, which is neglectable compared with the time used in Step 2(a) and 2(b).

After checking of four integral structures, the probability of one $AKO_{5,1}$, $AKO_{5,2}$ pair not being discarded is 2^{-28}, thus there are about 2^4 such pairs entering Step 3. Using similar technique of Step 2, it is possible to finish this step within $2^4 \cdot 2^{19} = 2^{23}$ encryptions. After Step 3, only the correct guess remains and the wrong guesses are all discarded with high probability.

As shown above, the total time needed is dominated by Step 2(a) and Step 2(b). Hence, the time complexity of this attack is about $2^{28.58} + 2^{28.58} = 2^{29.58}$ encryptions.

5 Attack on 6-Round MISTY1 without the Last *FL* Layer

In this section, we extend the improved 5-round integral attack to 6-round without the last FL layer. To apply the method of the 5-round integral attack, the 6-round attack needs to recover the actual value of the input of FO_5, which means the attack needs to partially decrypt the sixth round. However, directly guessing 98 key bits used in FO_6 will make the attack slower than exhaustive key search. To reduce the time needed, we start from the following observation.

As shown in Figure 3, the input of FO_5, $IL_5||IR_5$, can be written as $C^{(1\ldots16)} \oplus ML_6 \oplus AKO_{6,4}||C^{(17\ldots32)} \oplus MR_6 \oplus AKO_{6,5}$. The corresponding form of Equation (5) is then

$$\bigoplus_{j=1}^{2n} OL_5^{j(1\ldots7)} = \bigoplus_{j=1}^{2n} f(C^{j(1\ldots16)} \oplus ML_6^j \oplus AKO_{6,4} \oplus AKO_{5,1}) \oplus \bigoplus_{j=1}^{2n} g(C^{j(17\ldots32)} \oplus MR_6^j \oplus AKO_{6,5} \oplus AKO_{5,2}) \tag{6}$$

where $AKO_{5,1}$ and $AKO_{5,2}$ are then replaced by $AKO_{5,1} \oplus AKO_{6,4}$ and $AKO_{5,2} \oplus AKO_{6,5}$, respectively, and the input of FO_5 is then replaced by $C^{(1\ldots16)} \oplus ML_6||C^{(17\ldots32)} \oplus MR_6$, because the subkeys $AKO_{6,4}$, $AKO_{6,5}$, $AKO_{5,1}$ and $AKO_{5,2}$ are not related to compute ML_6 and MR_6.

To compute the intermediate values $C^{(1\ldots16)} \oplus ML_6$ and $C^{(17\ldots32)} \oplus MR_6$, $AKO_{6,1}$, $AKO_{6,2}$, $AKO_{6,3}$, $AKI_{6,1}$, $AKI_{6,2}$ and $AKI_{6,3}$ are required. These six subkeys, which are corresponding to K_6, K_8, $K_5 \oplus K_3'^{(1\ldots7)}||00||K_3'^{(1\ldots7)}$, $K_3'^{(8\ldots16)}$, $K_7'^{(8\ldots16)}$, $K_1'^{(8\ldots16)}$, only take 75 key bits. The attack can be described as follows:

1. Ask for the encryptions of four different integral structures. Each structure includes all plaintexts that have the same left 32 bits and all possible right 32 bits.

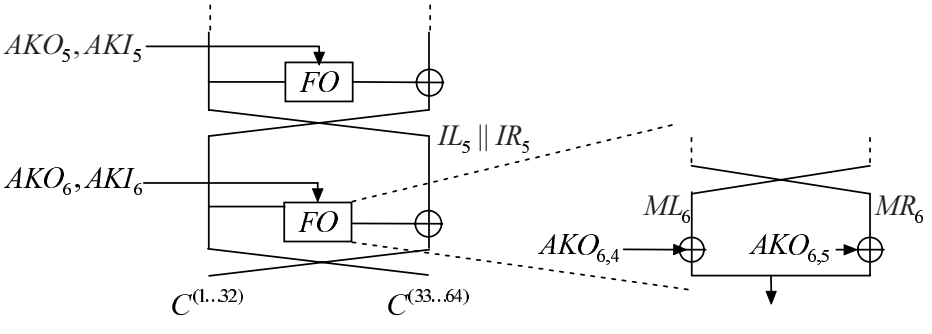

Fig. 3. Partial decryption in the attack on 6-round MISTY1 with all but the last FL layers

2. Guess 75 key bits, and partially decrypt all the 2^{34} encryptions to obtain the value of $C^{(1...16)} \oplus ML_6$ and $C^{(17...32)} \oplus MR_6$.
3. For each integral structure:
   (a) For every possible $AKO_{6,4} \oplus AKO_{5,1}$, compute the value of $\bigoplus_{j=1}^{2^{32}} f(C^{j(1...16)} \oplus ML_6^j \oplus AKO_{6,4} \oplus AKO_{5,1})$.
   (b) For every possible $AKO_{6,5} \oplus AKO_{5,2}$, compute the value of $\bigoplus_{j=1}^{2^{32}} f(C^{j(17...32)} \oplus MR_6^j \oplus AKO_{6,5} \oplus AKO_{5,2})$.
   (c) Discard all $AKO_{6,4} \oplus AKO_{5,1}$, $AKO_{6,5} \oplus AKO_{5,2}$ pairs such that $\bigoplus_{j=1}^{2^{32}} f(C^{j(1...16)} \oplus ML_6^j \oplus AKO_{6,4} \oplus AKO_{5,1}) \oplus \bigoplus_{j=1}^{2^{32}} g(C^{j(17...32)} \oplus MR_6^j \oplus AKO_{6,5} \oplus AKO_{5,2})$ does not equal to $\bigoplus_{i=1}^{2^{32}} C^{j(33...39)}$
4. For the guessed keys that are not discarded, exhaustively search for remaining key bits.

For each guessed 75-bit key, Step 2 partially decrypts 2^{34} encryptions. Each partial decryption takes no more than 1/4 encryption. So this step needs about 2^{32} encryptions for each guessed 75-bit key.

As shown in Section 4, Both Step 3(a) and 3(b) need $2^{28.58}$ encryptions for the four integral structures, and Step 3(c) needs neglectable time compared with Step 3(a) and 3(b).

There are expected 2^4 out of 2^{32} possible $AKO_{5,1} \oplus AKO_{6,4}$, $AKO_{5,2} \oplus AKO_{6,5}$ pairs entering Step 4 for each guessed 75 bits in Step 2. For each guess entering Step 4, the attack exhaustively searches the 2^{21} possible remaining key bits. So the running time of this step is about 2^{25} encryptions for each guessed 75-bit key in Step 2.

As a result, the total time complexity of this attack is $2^{75} \cdot (2^{32} + 2^{28.58} + 2^{28.58} + 2^{25}) = 2^{107.26}$ encryptions.

6 Attack on 6-Round MISTY1 with All FL Layers

In this section, we extend the attack presented in last section to 6-round with all FL layers. If we simply extend the MISTY1 used in last section with FL_9 and FL_{10}, the attack will slower than exhaustive key search. By exploring the key schedule algorithm, we can perform a chosen ciphertext attack on last 6 round MISTY1 block cipher with all FL functions. The encryption then starts before the FL_3, FL_4 layer and ends at the end of the cipher.

In this attack, we also make use of the 4-round integral. Since the attack is a chosen ciphertext attack, the four round integral corresponds to the XOR of all the 2^{32} 32-bit intermediate values of the integral structure before FL_5. We also use Theorem 1 for fast checking the integral. As shown in Figure 4, the Equation (5) of Theorem 1 for the forth round can be rewritten as:

$$\bigoplus_{j=1}^{2n} OL_4^{j(1...7)} = \bigoplus_{j=1}^{2n} f(D^{j(33...48)} \oplus ML_3^j \oplus AKO_{3,4} \oplus AKO_{4,1}) \oplus \bigoplus_{j=1}^{2n} g(D^{j(49...64)} \oplus MR_3^j \oplus AKO_{3,5} \oplus AKO_{4,2}) \tag{7}$$

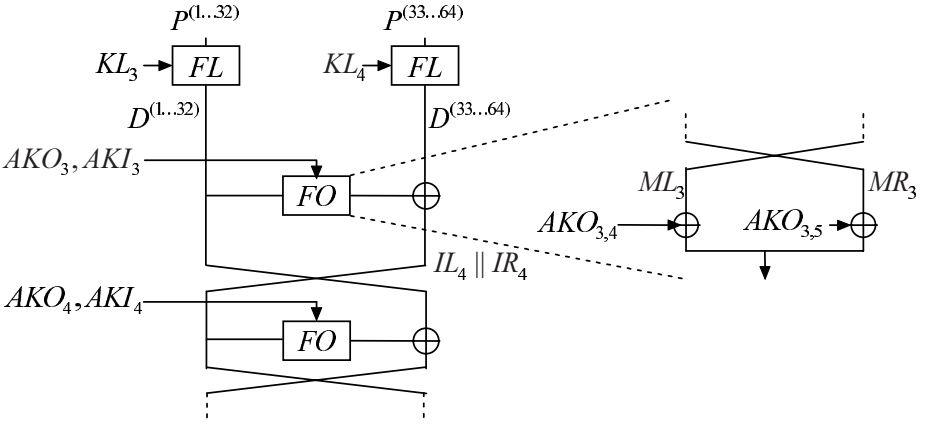

Fig. 4. Partial decryption in attack on 6-round MISTY1 with all layers

where D denotes the result of the plaintext P passing through the first FL_3 and FL_4 layer as shown in Figure 4.

We need to obtain the values of $D^{(33...48)} \oplus ML_3$ and $D^{(49...64)} \oplus MR_3$ from decryptions. For one decryption, if the attack partially decrypts for these values directly, it needs to guess at least total 105 key bits. Such a guess together with 2^{32} partial decryption will make the attack slower than exhaustive key search. However, to check Equation (7), we could obtain all the 2^{32} $D^{j(33...48)} \oplus ML_3^j$ to compute $\bigoplus_{j=1}^{2^{32}} f(D^{j(33...48)} \oplus ML_3^j \oplus AKO_{3,4} \oplus AKO_{4,1})$ and obtain all the 2^{32} $D^{j(49...64)} \oplus MR_3^j$ to compute $\bigoplus_{j=1}^{2^{32}} g(D^{j(49...64)} \oplus MR_3^j \oplus AKO_{3,5} \oplus AKO_{4,2})$ separately.

By exploring the key schedule weakness, we notice that none of the two processes needs all the 105 key bits. To obtain $D^{(33...48)}$, the attack needs $KL_{4,1}$ and $KL_{4,2}$, which are K_4' and K_6. To obtain ML_3 from the plaintext P, the attack needs only the subkeys $KL_{3,1}$, $KL_{3,2}$, $AKO_{3,1}$, $AKO_{3,2}$, $AKI_{3,1}$ and $AKI_{3,2}$, which correspond to K_2, K_8', K_3, K_5, $K_8'^{(8...16)}$, $K_4'^{(8...16)}$. Thus, only K_4', K_2, K_8', K_3, K_5 and K_6 are required.

Proposition 3. *For one decryption, computing $D^{(33...48)} \oplus ML_3$ from plaintext needs only 96 key bits.*

Consider the process of computing $D^{(49...64)} \oplus MR_3$. To obtain $D^{(49...64)}$, the attack only needs to know $KL_{4,1}(K_4')$ but not $KL_{4,2}(K_6)$. To obtain MR_3, the attack needs only $KL_{3,1}$, $KL_{3,2}$, $AKO_{3,1}$, $AKO_{3,2}$, $AKO_{3,3}$, $AKI_{3,1}$, $AKI_{3,2}$ and $AKI_{3,3}$, which correspond to K_2, K_8', K_3, K_5, $K_2 \oplus K_8'^{(1...7)}||00||K_8'^{(1...7)}$, $K_8'^{(8...16)}$, $K_4'^{(8...16)}$, $K_6'^{(8...16)}$, as Table 2.

Proposition 4. *For one decryption, computing $D^{(49...64)} \oplus MR_3$ from plaintext needs only 89 key bits.*

When computing the value of $\bigoplus_{j=1}^{2^{32}} f(D^{j(33\ldots48)} \oplus ML_3^j \oplus AKO_{3,4} \oplus AKO_{4,1})$ for an integral structure, the $AKO_{4,1}$ and $AKO_{3,4}$ used correspond to K_4 and $K_7 \oplus K_8'^{(1\ldots7)}||00||K_8'^{(1\ldots7)} \oplus K_4'^{(1\ldots7)}||00||K_4'^{(1\ldots7)}$. The subkey K_7 is not included in the guessed 96 key bits. Hence K_7 should be guessed after obtaining all the $D^{(33\ldots48)} \oplus ML_3^j$.

To obtain $\bigoplus_{j=1}^{2^{32}} g(D^{j(49\ldots64)} \oplus MR_3^j \oplus AKO_{3,5} \oplus AKO_{4,2})$, the attack still needs to guess $K_6 \oplus K6'^{(1\ldots7)}||00||K6'^{(1\ldots7)}$. This guess can be done after obtaining all the $D^{(49\ldots64)} \oplus MR_3^j$.

The attack is as follows:

1. Ask for decryptions of one integral structure in which all ciphertexts have the same left 32 bits and all possible right 32 bits.
2. For each guess of the 80-bit K_2, K_8', K_3, K_5, K_4':
   (a) Compute the value of $\bigoplus_{j=1}^{2^{32}} D^{j(1\ldots7)}$.
   (b) Guess 16-bit K_6 and obtain all the 2^{32} $D^{j(33\ldots48)} \oplus ML_3^j$. Continue to guess 16 bit words K_7 and compute $\bigoplus_{j=1}^{2^{32}} f(D^{j(33\ldots48)} \oplus ML_3^j \oplus AKO_{3,4} \oplus AKO_{4,1})$.
   (c) Guess 9-bit $K_6'^{(8\ldots16)}$ and obtain all the 2^{32} $D^{j(49\ldots64)} \oplus MR_3^j$. Continue to guess 16 bit words $K_6 \oplus K_6'^{(1\ldots7)}||00||K_6'^{(1\ldots7)}$ and obtain $\bigoplus_{j=1}^{2^{32}} g(D^{j(49\ldots64)} \oplus MR_3^j \oplus AKO_{3,5} \oplus AKO_{4,2})$.
   (d) Discard all guesses of K_6, K_7, $K_6'^{(8\ldots16)}$ and $K_6 \oplus K_6'^{(1\ldots7)}||00||K_6'^{(1\ldots7)}$ such that Equation (7) does not hold or the guesses that result in conflict (K_6 and K_7 do not produce K_6' corresponding to $K_6'^{(8\ldots16)}$ and $K_6 \oplus K_6'^{(1\ldots7)}||00||K_6'^{(1\ldots7)}$).
3. For the guesses not discarded, exhaustively search for the remaining key bits.

In Step 2(a), the computation of $\bigoplus_{j=1}^{2^{32}} D^{j(1\ldots7)}$ takes no more than $2^{32} \cdot 1/4 = 2^{30}$ encryptions for each guessed 80 key bits.

In Step 2(b), for each guessed 16-bit K_6, the calculation of 2^{32} $D^{j(33\ldots48)} \oplus ML_3^j$ takes no more than $2^{32} \cdot 1/4 = 2^{30}$ encryptions. For each K_7, the calculation of $\bigoplus_{j=1}^{2^{32}} f(D^{j(33\ldots48)} \oplus ML_3^j \oplus AKO_{3,4} \oplus AKO_{4,1})$ takes about 2^{15} table look-ups, which is equivalent to about $2^{15} \cdot 2^{-6} = 2^9$ encryptions. Hence, the running time of Step 2(b) is no more than $2^{16} \cdot (2^{30} + 2^{16} \cdot 2^9) = 2^{46.04}$ encryptions for each guessed 80 key bits. Similarly, Step 2(c) needs about $2^{39.04}$ encryptions for each guessed 80 key bits.

In Step 2(d), for each K_6 and K_7, we calculate the value of $K_6'^{(8\ldots16)}$ and $K_6 \oplus K_6'^{(1\ldots7)}||00||K_6'^{(1\ldots7)}$, and then check whether Equation (7) holds. Hence Step 2(d) checks 2^{32} guesses, and the time needed is also neglectable to Step 2(b). For each guess of the 32 bit K_6 and K_7, the probability of satisfying Equation (7) is 2^{-7}. Hence, for each guessed 80 bits in Step 2, there are expected 2^{25} guesses out of 2^{32} possible K_6, K_7 entering Step 3. We notice that for each guess entering Step 3, the attack still needs to exhaustively search for the remaining 16 key bits. Therefore, Step 3 takes 2^{41} encryptions for each guess in Step 2.

The running time of the whole attack is dominated by Step 2(b). The time complexity of the attack is $(2^{30} + 2^{46.04} + 2^{39.04} + 2^{41}) \cdot 2^{80} = 2^{126.09}$ encryptions.

7 Conclusion

In this paper, we presented several integral attacks on reduced MISTY1 block cipher. Our attack improved the 5-round integral attack presented in [4] with the use of the FO Relation. We also extended the attack to 6-round with all FL layers by exploring the key schedule algorithm.

The existence of the FO Relation stems from the structure of the FO function and the fact that the key is XORed in the FO function; the resulting diffusion effect is too weak to defeat popular cryptanalysis techniques, such as differential cryptanalysis and integral cryptanalysis.

Our attack also indicates that the correspondence between subkeys used and the 128-bit key might be simple. Further exploration of this weakness of the key schedule is still worthy studying.

Acknowledgement

The authors would like to acknowledge Zheng Gong and Ruoyao Shi for their helpful advices.

References

1. Babbage, S., Frisch, L.: On MISTY1 Higher Order Differential Cryptanalysis. In: Won, D. (ed.) ICISC 2000. LNCS, vol. 2015, pp. 22–36. Springer, Heidelberg (2001)
2. Dunkelman, O., Keller, N.: An Improved Impossible Differential Attack on MISTY1. In: Pieprzyk, J. (ed.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 441–454. Springer, Heidelberg (2008)
3. Hatano, Y., Tanaka, H., Kaneko, T.: An Optimized Algebraic Method for Higher Order Differential Attack. In: Fossorier, M.P.C., Høholdt, T., Poli, A. (eds.) AAECC 2003. LNCS, vol. 2643, pp. 61–70. Springer, Heidelberg (2003)
4. Knudsen, L.R., Wagner, D.: Integral Cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 112–127. Springer, Heidelberg (2002)
5. Kühn, U.: Cryptanalysis of Reduced-Round MISTY. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 325–339. Springer, Heidelberg (2001)
6. Kühn, U.: Improved Cryptanalysis of MISTY1. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 61–75. Springer, Heidelberg (2002)
7. Lu, J., Kim, J., Keller, N., Dunkelman, O.: Improving the Efficiency of Impossible Differential Cryptanalysis of Reduced Camellia and MISTY1. In: Malkin, T.G. (ed.) CT-RSA 2008. LNCS, vol. 4964, pp. 370–386. Springer, Heidelberg (2008)
8. Matsui, M.: New Block Encryption Algorithm MISTY. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 54–68. Springer, Heidelberg (1997)
9. Sakurai, K., Zheng, Y.: On Non-Pseudorandomness from Block Ciphers with Provable immunity Against Linear Cryptanalysis. In: Fossorier, M.P.C., Imai, H., Lin, S., Poli, A. (eds.) IEICE Trans. Fund., vol. E80-A(1), pp. 19–24 (1997)

10. Tanaka, H., Hatano, Y., Sugio, N., Kaneko, T.: Security Analysis of MISTY1. In: Kim, S., Yung, M., Lee, H.-W. (eds.) WISA 2007. LNCS, vol. 4867, pp. 215–226. Springer, Heidelberg (2008)
11. Tanaka, H., Hisamatsu, K., Kaneko, T.: Strength of MISTY1 without FL Function for Higher Order Differential Attack. In: Fossorier, M.P.C., Imai, H., Lin, S., Poli, A. (eds.) AAECC 1999. LNCS, vol. 1719, pp. 221–230. Springer, Heidelberg (1999)