Improved Higher-Order Differential Attacks on MISTY1

Achiya Bar-On(✉)

Department of Mathematics, Bar Ilan University,
52900 Ramat Gan, Israel
barona@macs.biu.ac.il

Abstract. MISTY1 is a block cipher designed by Matsui in 1997. It is widely deployed in Japan, and is recognized internationally as an European NESSIE-recommended cipher and an ISO standard. Since its introduction, MISTY1 was subjected to extensive cryptanalytic efforts, yet no attack significantly faster than exhaustive key search is known on its full version. The best currently known attack is a higher-order differential attack presented by Tsunoo et al. in 2012 which breaks a reduced variant of MISTY1 that contains 7 of the 8 rounds and 4 of the 5 FL layers in $2^{49.7}$ data and $2^{116.4}$ time.

In this paper, we present improved higher-order differential attacks on reduced-round MISTY1. Our attack on the variant considered by Tsunoo et al. requires roughly the same amount of data and only $2^{100.4}$ time (i.e., is 2^{16} times faster). Furthermore, we present the first attack on a MISTY1 variant with 7 rounds and all 5 FL layers, requiring $2^{51.4}$ data and 2^{121} time. To achieve our results, we use a new higher-order differential characteristic for 4-round MISTY1, as well as enhanced key recovery algorithms based on the *partial sums* technique.

1 Introduction

MISTY1 [10] is a 64-bit block cipher with 128-bit keys designed in 1997 by Matsui. In 2002, MISTY1 was selected by the Japanese government to be one of the CRYPTREC e-government ciphers, and since then, it became widely deployed in Japan. MISTY1 also gained recognition outside Japan, when it was selected to the portfolio of European NESSIE-recommended ciphers, and approved as an ISO standard in 2005. Furthermore, the block cipher KASUMI [1] designed as a slight modification of MISTY1 is used in the 3G cellular networks, which makes it one of the most widely used block ciphers today.

MISTY1 has an 8-round recursive Feistel structure, where the round function FO is in itself a 3-round Feistel construction, whose F-function FI is in turn a 3-round Feistel construction using 7-bit and 9-bit invertible S-boxes. The specific choice of S-boxes and the recursive structure ensure provable security against

A. Bar-On—This research was partially supported by the Israeli Ministry of Science, Technology and Space, and by the Check Point Institute for Information Security.

G. Leander (Ed.): FSE 2015, LNCS 9054, pp. 28–47, 2015.
DOI: 10.1007/978-3-662-48116-5_2

differential and linear cryptanalysis. In order to thwart other types of attacks, after every two rounds an FL function is applied to each of the two halves independently. The FL functions are key-dependent linear functions which play the role of whitening layers.

Since its introduction, MISTY1 was subjected to extensive cryptanalytic efforts using a variety of techniques, which resulted in numerous attacks on its reduced variants. The best currently known attacks are the following:

- A higher-order differential (HOD) attack on 6-round MISTY1 with 4 of the 5 FL layers, with a semi-practical complexity of $2^{49.4}$ chosen plaintexts and time [13].
- An impossible differential attack on 7-round MISTY1 with 3 FL layers, that requires 2^{58} known plaintexts and $2^{124.4}$ time [5].
- A zero-correlation linear attack on 7-round MISTY1 with 4 FL layers, that requires $2^{62.9}$ known plaintexts and 2^{118} time [15].
- A HOD attack on 7-round MISTY1 with 4 of the 5 FL layers, that requires $2^{49.7}$ chosen plaintexts and $2^{116.4}$ encryptions [13].
- A related-key differential attack on the full MISTY1, that requires 2^{61} chosen ciphertexts and $2^{90.9}$ encryptions, and applies under a weak key class assumption [9].
- A meet-in-the-middle attack which allows to speed up exhaustive key search on the full MISTY1 [6] by a factor of between 2 and 4.

Examination of the best currently known attacks on MISTY1 suggests that up to date, the technique that provided the strongest results against reduced-round MISTY1 is the higher-order differential attack. In this paper, we examine the currently known HOD attacks on MISTY1 thoroughly and show that they can be improved, both in the exact characteristic used for the attack and in the key-recovery algorithm. The results we obtain are the following:

1. The 44-order differential characteristic for 4-round MISTY1 introduced and deployed in [13] can be replaced by more efficient 43-order differentials. This allows to reduce the data and time complexities of the attacks of [13] on 6-round MISTY1 from $2^{49.4}$ to 2^{47}. As we explain in Sect. 4, the order of the differential cannot be reduced further unless an entirely different characteristic is introduced.
2. The time complexity of the attack of [13] on 7-round MISTY1 can be reduced by a factor of 2^{16} by using the *partial sums* technique [4], along with optimizations exploiting the exact structure of MISTY1.
3. Despite the fact that 7-round MISTY1 with all 5 FL layers uses 64 additional subkey bits (compared to the variant attacked in [13]), we can break this variant in data $2^{51.4}$ and time 2^{121} using a complex key-recovery procedure based on the partial sums technique.

The latter result is the first known attack on 7-round MISTY1 with all 5 FL functions present. A comparison of our attacks with the best previously known attacks on reduced-round MISTY1 is presented in Table 1.

The paper is organized as follows. In Sect. 2 we describe the structure of MISTY1 and introduce some notations that will be used throughout the paper. Since our attack is based heavily on the HOD attacks of [13,14], we describe these attacks briefly in Sect. 3. Our improved attack on 6-round MISTY1 is presented in Sect. 4. The attacks on 7-round MISTY1 with 4 and 5 FL layers are presented in Sects. 5 and 6, respectively. Finally, in Sect. 7 we summarize the paper.

Table 1. Summary of the best known single-key attacks on MISTY1

FO rounds	FL layers	Data complexity	Time complexity	Type
6	4	$2^{49.4}$	$2^{49.4}$	HOD attack [13]
6	4	2^{47}	2^{47}	HOD attack (Sect. 4)
7	3	2^{58} KP	$2^{124.4}$	ID attack [5]
7	4	$2^{62.9}$ KP	2^{118}	MZC attack [15]
7	4	$2^{49.7}$	$2^{116.4}$	HOD attack [13]
7	4	$2^{50.1}$	$2^{100.4}$	HOD attack (Sect. 5)
7	5	$2^{51.45}$	2^{121}	HOD attack (Sect. 6)

ID attack: Impossible Differential attack
HOD attack: Higher Order Differential attack
MZC attack: Multi-Dimensional Zero Correlation attack

2 Brief Description of MISTY1

MISTY1 is an 8-round Feistel construction, where the round function, FO, is in itself a variant of a 3-round Feistel construction, defined as follows. The input to FO is divided into two halves. The left one is XORed with a subkey, enters a keyed permutation FI, and the output is XORed with the right half. After the XOR, the two halves are swapped, and the same process (including the swap) is repeated two more times. After that, an additional swap and an XOR of the left half with a subkey is performed (see Fig. 1).

The FI function in itself also has a Feistel-like structure. Its 16-bit input is divided into two unequal parts – one of 9 bits, and the second of 7 bits. The left part (which contains 9 bits) enters an S-box, $S9$, and the output is XORed with the right 7-bit part (after padding the 7-bit value with two zeroes as the most significant bits). The two parts are swapped, the 7-bit part enters a different S-box, $S7$, and the output is XORed with 7 bits out of the 9 of the right part. The two parts are then XORed with a subkey, and swapped again. The 9-bit value again enters $S9$, and the output is XORed with the 7-bit part (after padding). The two parts are then swapped for the last time.

Every two rounds, starting before the first one, each of the two 32-bit halves enters an FL layer. The FL layer is a simple linear transformation. Its input is divided into two halves of 16 bits each, the AND of the left half with a subkey

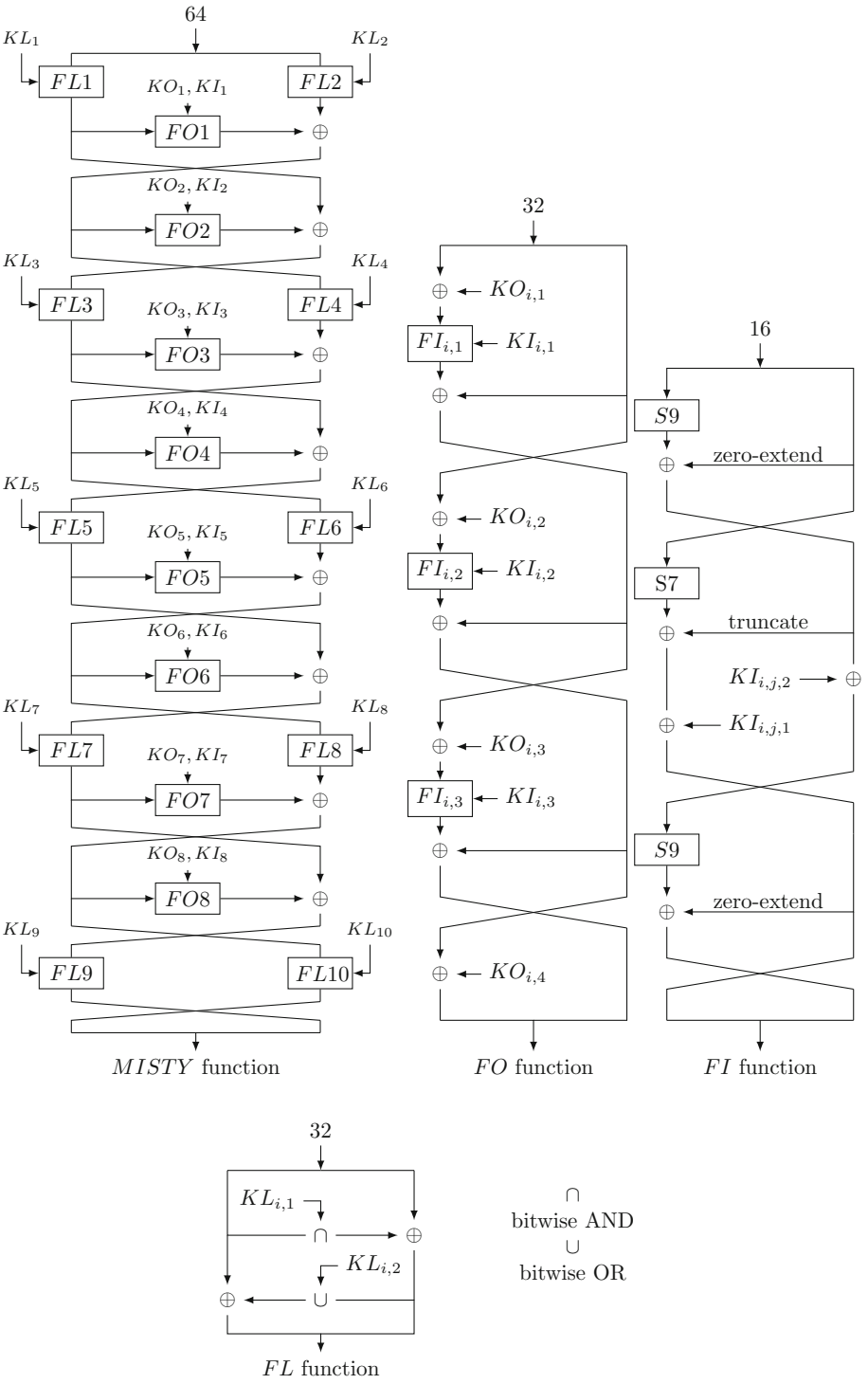


Fig. 1. Outline of MISTY1

Table 2. The key schedule of MISTY1

$KO_{i,1}$	$KO_{i,2}$	$KO_{i,3}$	$KO_{i,4}$	$KI_{i,1}$	$KI_{i,2}$	$KI_{i,3}$	$KL_{i,1}$	$KL_{i,2}$
K_i	K_{i+2}	K_{i+7}	K_{i+4}	K'_{i+5}	K'_{i+1}	K'_{i+3}	$K_{\frac{i+1}{2}}$ (odd i) $K_{\frac{i}{2}+4}$ (even i)	$K_{\frac{i+1}{2}+6}$ (odd i) $K_{\frac{i}{2}+2}$ (even i)

is XORed to the right half, and the OR of the updated right half with another subkey is XORed to the left half. We outline the structure of MISTY1 and its parts in Fig. 1.

The key schedule of MISTY1 takes the 128-bit key, and treats it as eight 16-bit words $K_1, K_2, \ldots, K_8$. From this sequence of words, another sequence of eight 16-bit words is generated, according to the rule $K'_i = FI_{K_{i+1}}(K_i)$.

In each round, seven words are used as the round subkey, and each of the FL functions accepts two subkey words. We give the exact key schedule of MISTY1 in Table 2.

2.1 Notations Used in the Paper

Throughout the paper, we use the following notations for intermediate values during the MISTY1 encryption process.

- The plaintext and the ciphertext are denoted, as usual, by P and $C = E(P)$.
- The input of the i'th round ($1 \leq i \leq 8$) is denoted by X_i. If we want to emphasize that the intermediate value corresponds to the plaintext P, we denote it by $X_i(P)$.
- For odd rounds, we denote by X'_i the intermediate value after application of the FL functions.
- The output of the FO function of round i is denoted Out_i.
- For any intermediate value Z, $Z[k-l]$ denotes bits from k to l (inclusive) of Z.
- For any intermediate value Z, the right and left halves of Z are denoted by Z_R and Z_L, respectively.

3 Brief Description of the HOD Attacks of Tsunoo Et Al. [13,14] on Reduced-Round MISTY1

In this section we present a brief description of the attacks of Tsunoo et al. [13, 14], that serve as the basis for our results.

3.1 General Outline of the Attack

The higher order differential attack was presented by Knudsen [7] in 1994 (see also [8]). The basic idea behind the attack is as follows.

Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. Suppose that the degree of f (as a multivariate polynomial) is d and that $V \subseteq \{0,1\}^n$ is a vector subspace of dimension k. Then it is easy to show that Equation (1) holds, independently of x.

$$\bigoplus_{y \in V} f(x \oplus y) = \begin{cases} \text{const} & \text{if } k = d \\ 0 & \text{if } k > d. \end{cases} \tag{1}$$

Suppose now that for some block cipher E, the mapping from the plaintext to a single bit in some intermediate state, that is, $X_i[j]$, can be represented by a (key-dependent) Boolean function f_K of a low degree. Then by Eq. (1), we have

$$\bigoplus_{y \in V} X_i[j](x \oplus y) = \begin{cases} \text{const}(K) & \text{if } k = d \\ 0 & \text{if } k > d. \end{cases} \tag{2}$$

Note that const(K) depends only on K and not on the choices of x and V. Equation (2) can be used to mount the following attack on E. Denote by E_1^{-1} the Boolean function that represents the mapping from the ciphertext of E to the intermediate state bit $X_i[j]$. Then Eq. (2) can be rewritten as

$$\bigoplus_{y \in V} E_1^{-1}(E(x \oplus y)) = \begin{cases} \text{const}(K) & \text{if } k = d \\ 0 & \text{if } k > d. \end{cases} \tag{3}$$

The adversary asks for the encryption of several structures of plaintexts of the form $\{x \oplus y | y \in V\}$, where x is arbitrary and V is an arbitrary vector subspace of degree d, partially decrypts the corresponding ciphertexts (by guessing the key material used in E_1^{-1}), and checks whether Eq. (3) holds. The dimension d of V is called *the order of the differential*.

As guessing the key material used in E_1^{-1} may be very time consuming, various other techniques are used to approach Eq. (3). The technique used in [13, 14] is linearization, which allows to exploit the low algebraic degree of a single MISTY1 round.

In the linearization method, we first express $X_i[j]$ as a multivariate polynomial $f'(C, K)$ in the ciphertext bits and the key bits, where ciphertext bits are treated as constants and key bits are treated as variables. This transforms Eq. (3) into a polynomial equation in the key bits. Then, we linearize the equation by replacing each non-linear expression in the key bits (e.g., $k_1 \cdot k_2$) with a new variable. In such a way, Eq. (3) for each structure of plaintexts contributes a linear equation, where the total amount m of variables is the number of non-linear terms in f'. If the equations we obtain are independent (which is usually the case), m equations are sufficient for obtaining a unique solution. Every extra equation can be used as a filtering condition. Hence, if the amount of key bits used in E_1^{-1} is s, then $m+s$ structures are sufficient for determining all of them.

In summary, the HOD attack of [13,14] on MISTY1 consists of three stages. The first stage is detecting a HOD of order as small as possible that "predicts" a bit as close as possible to the ciphertext. The second stage is to create a system

of linear equations by linearization of the corresponding function E_1^{-1}. The third stage is solving the equation system.

Before presenting the three stages in some more detail, we introduce a notation that will be used throughout the paper to describe higher-order differentials of MISTY1.

Consider a partial encryption of MISTY1, which starts at the state X_j and ends at the state X_k, where we are interested only in the bits $X_k[\ell - m]$. Denote this encryption function by $E'_K : X_j \to X_k[\ell - m]$. We denote by $V^{(d)}_{X_j[i_1,\ldots i_d]} X_k[\ell - m]$ the d'th order differential that starts in X_j, where $V = span\{e_{i_1}, \ldots e_{i_d}\}$ is the vector subspace of $\{0,1\}^{64}$ spanned by the unit vectors $e_{i_1}, \ldots e_{i_d}$. In other words,

$$V^{(d)}_{X_j[i_1,\ldots i_d]} X_k[\ell - m] = \bigoplus_{y \in span\{e_{i_1},\ldots e_{i_n}\}} E'_K(x \oplus y).$$

3.2 A 44'th Order Differential for 4-round MISTY1

The 44'th order differential of 4-round MISTY1 found in [13] is a culmination of a series of observations.

The basic observation is a 7'th order differential of 3-round MISTY1 without FL functions discovered by Babbage and Frisch [2].

Theorem 1. *For any three consecutive rounds of MISTY1 without FL functions, the equation*

$$V^{(7)}_{X_i[0-6]} X_{i+3}[57 - 63] = 0x6d$$

holds, independently of the (fixed) value of the key and of the (constant) value of $X_i[7 - 63]$.

As noted in [2], the theorem fails for MISTY1 with FL layers. In order to overcome this obstacle, Tsunoo et al. [14] suggested the notion of *neighbor* bit positions.

Definition 2. For any intermediate state Z of MISTY1, the *neighbor* of the bit position $Z[i]$ is the bit position $Z[i + 16]$.

Tsunoo et al. showed that when we accompany each bit position in the 7'th order differential with its neighbor position, the resulting differential can bypass the FL layers.

Theorem 3. *For any three consecutive rounds of MISTY1 with FL functions that start at an even round $2j$, the equation*

$$V^{(14)}_{X_{2j}[0-6,16-22]} X'_{2j+3}[57 - 63] = 0$$

holds, independently of the (fixed) value of the key and of the (constant) value of $X_{2j}[7 - 15, 23 - 63]$.

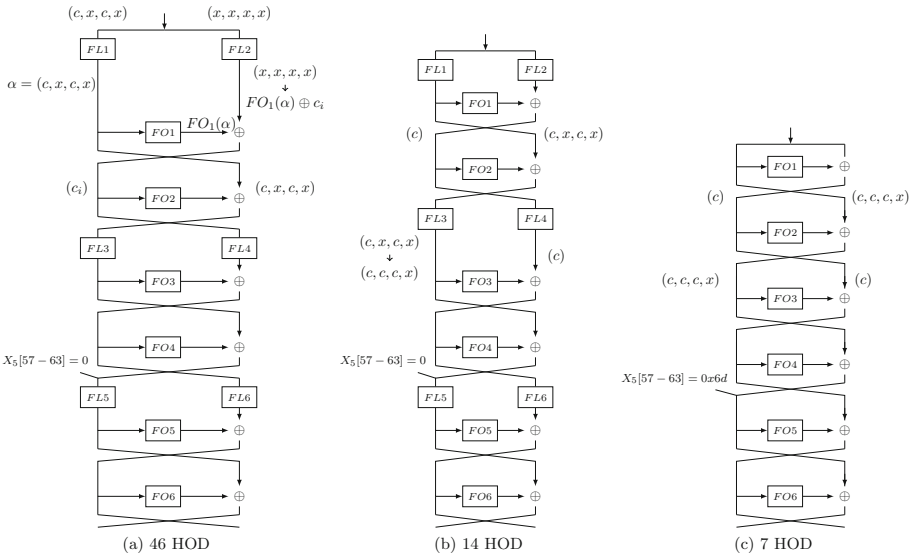
Due to the Feistel structure of MISTY1, the 3-round 14'th order differential can be extended to a 4-round 46'th order differential by taking all the 2^{32} possible values in the previous round. We obtain:

Theorem 4. *For any four consecutive rounds of MISTY1 with FL functions that start at an odd round* $2j+1$*, the equation*

$$V^{(46)}_{X_{2j+1}[0-31,32-38,48-54]} X'_{2j+5}[57-63] = 0$$

holds, independently of the (fixed) value of the key and of the (constant) value of $X_{2j+1}[39-47, 55-63]$.

The 7'th, 14'th and 46'th order differentials are illustrated in Fig. 2. The proof of the theorems can be found in [2,14].



The form $(*,*,*,*)$ represents division of a 32-bit state into sets of $(9,7,9,7)$ bits. The x's denote bit positions that are active in the differential. The c's denote "constant" bits. In Figure 2a, the symbol c_i, for $1 \leq i \leq 2^{32}$, denotes a 32-bit value that is constant for each sub-structure of size 2^{14}.

Fig. 2. Higher order differentials in MISTY1

The last observation, made by Tsunoo et al. in [13], is that each of the 14'th order differentials for 3-round MISTY1 presented above contains 28 12'th order differentials that also can be shown to sum up to zero (due to a non-maximal algebraic degree of the underlying function). These are all possible differentials obtained by taking any six of the 7 bits $X_{2j}[0-6]$, along with their neighbors. As some of these 12'th order differentials are linearly dependent, it turns out that only 22 of them can be used in parallel.

The observation of Tsunoo et al. in [13] allows to reduce the data and time complexities of the HOD attacks based on the 46'th round differential of 4-round MISTY1 by a factor of 22. Indeed, while in the previous attacks, each data structure of 2^{46} plaintexts contributes a single linear equation, Tsunoo et al.'s observation allows to use it to obtain 22 linearly independent equations (each coming from an extension by one round of a 12'th order differential). This obviously reduces the data complexity by a factor of 22, and if partial encryption/decryption of all the data is the most time consuming operation in the attack (as is the case for the attacks of [14], as we shall see below), the time complexity of the attacks is reduced by the same factor.

Given the progress in the size of HODs from paper to paper, it seems reasonable to check whether the differential can be improved yet another time by dividing the basic 12'th order differential to smaller ones. In this paper, we show that the answer is positive, to some extent. Namely, we show that when considering MISTY1 without FL functions, the 12'th order differential can be divided into several 11'th order differentials. While these differentials do not hold for MISTY1 with FL functions, we show that they can be applied in a more sophisticated way, and result in a reduction in the data and time complexities of the attacks on 6-round MISTY1 by a factor of 5. As far as we checked, our results cannot be pushed further, in the sense that no other sub-differential of the 12'th order differential of [14] sums up to zero, even for 3-round MISTY1 without FL functions.

3.3 HOD Attacks on 6-Round and 7-Round MISTY1

The attack of Tsunoo et al. [14] on 6-round MISTY1 uses the 46'th order differential illustrated in Fig. 2a. The equation given by the differential is $V^{(46)}_{P[0-31,32-38,48-54]}X_5[57-63]=0$. In order to bypass the layer $FL5$, we note that if one of the bits in $KL_{5,2}[9-15]$ is equal to 1, say $KL_{5,2}[15]=1$, then by the structure of FL, we obtain $V^{(46)}_{P[0-31,32-38,48-54]}X'_5[63]=0$, and similarly for the other bits. Since for a vast majority of the keys, at least one bit of $KL_{5,2}[9-15]$ equals 1, we can repeat the attack 7 times, each time assuming that one of the bits equals 1, so that with an overwhelming probability the attack will succeed in one of the times.

Assume, w.l.o.g., that $KL_{5,2}[15]=1$, and hence, we have the equation $V^{(46)}_{P[0-31,32-38,48-54]}X'_5[63]=0$. By the Feistel structure, it follows that $V^{(46)}_{P[0-31,32-38,48-54]}X_6[31]=0$. Note that $X_6[31]=FL7^{-1}(C)[63]\oplus Out_6[31]$. Since $FL7^{-1}$ acts as 16 parallel applications of a function from two bits to two bits, $FL7^{-1}(C)[63]$ for all ciphertexts can be computed easily given a guess of two key bits. Hence, all we need to do in order to check whether the differential is satisfied is to compute

$$\bigoplus_{y\in span\{e_0,e_1,\ldots,e_{38},e_{48},\ldots,e_{54}\}} Out_6[31](x\oplus y).$$

It turns out that when we represent the bit $Out_6[31]$ as a function of X_{6L}, its degree as a multilinear polynomial is only 3, and moreover, many of the possible second and third degree terms do not appear. As a result, after the linearization of this function we obtain only 189 variables.[1] Hence, 220 linear equations are sufficient to filter out all wrong suggestions of the subkey used in round 6. As each structure of size 2^{46} contributes 22 equations as described above, 10 structures, or $2^{49.4}$ chosen plaintexts are sufficient for the attack.

The time complexity of solving the equations (in all 7 attack attempts together) is at most $7 \cdot 220^3 = 2^{26.2}$ operations. As for the time required for creating the linear equation, the most naive way is to consider the ciphertexts one by one and check the contribution of each ciphertext to the equations. Even in this way, the time complexity of the attack is dominated by encrypting the plaintexts. (As we show in Sect. 4, this part can be performed very efficiently using the *partial sums* technique). Hence, the attack requires $2^{49.4}$ chosen plaintexts and its time complexity is $2^{49.4}$ encryptions.

The attack of Tsunoo et al. [14] on 7-round MISTY1 without the last FL layer is an easy extension of the 6-round attack. Most naively, one guesses all key material used in round 7, peels off the 7th round and applies the 6-round attack. As shown in [14], it is sufficient to guess 75 of the 96 key bits used in round 7, as the remaining key bits can be absorbed into the linear equations, at the price of slightly increasing the number of variables. In total, the data complexity is increased to $2^{49.7}$ chosen plaintexts, and the time complexity is $2^{49.4} \cdot 2^{75} = 2^{116.4}$ encryptions.

4 Improved Attack on 6-Round MISTY1 Using a 43'th Order Differential

In this section we show that the attack of Tsunoo et al. on 6-round MISTY1 can be improved by a factor of 5 in the data and time complexities, by using an improved higher-order differential, along with a refined key recovery technique.

In order to find out whether the differential of Tsunoo et al. can be improved, we first examined the simpler variant of MISTY1 without the FL functions and used an experimental approach. We considered the 12'th order differentials on 3-round MISTY1 used in [13], e.g., $V^{(12)}_{X_{2j}[0-5,16-21]} X'_{2j+3}[57-63]$, and checked whether replacing $V = span\{e_0, e_1, \ldots, e_5, e_{16}, \ldots, e_{21}\}$ with any of its subspaces of the form $V' = span\{e_{i_1}, \ldots e_{i_d}\}$ yields a higher-order differential. The experiment was performed for all the options of 6 out the 7 bits $X_{2j}[0-6]$ with their neighbors. For each such choice of a 12'th order differential, several random keys and random constants were taken.

The experiments showed the existence of 11'th order differentials, of the form $V^{(11)}_{X_{2j}[S,S']} X'_{2j+3}[57-63]$, where S is any subset of $[0-6]$ of size 6, and S' consists

[1] It should be noted that the number of variables depends on the exact bit in $X_6[25-31]$ that is analyzed. As each of the 7 bits is analyzed in one of the 7 applications of the attack, we use the maximal possible number of variables throughout the paper, as a worst-case assumption.

of 5 among the 6 neighbors of the elements of S. It turns out that all differentials of this form are indeed 11'th order differentials for 3-round MISTY1 without FL functions. On the other hand, the experiment showed that all other subspaces V' do not yield HODs, which implies that our improved differential cannot be improved further, unless entirely different HODs are used.

The obvious obstacle in exploiting the 11'th order differential is that it cannot bypass FL layers. However, it turns out that we can overcome this obstacle, using a careful key guessing procedure that exploits the exact structure of the FL's. We start with a 12'th order differential of 3-round MISTY1, like those used in [13] and show how to divide it into two 11'th order differentials. For sake of simplicity, we exemplify the process for the differential $V^{(12)}_{X_2[0-5,16-21]} X'_5[57-63]$.

It is clear that the only obstacle we have to bypass is the layer $FL3$. Our goal is to define the structure in X_2 (i.e., the input of the differential) in such a way that the corresponding structure after $FL3$ will be $V^{(11)}_{X'_{3,L}[0-5,16-20]}$. If this is achieved, the continuation of the differential will hold like in the differentials found for MISTY1 without the FL functions.

As FL acts like 16 2-bit to 2-bit invertible functions applied in parallel, it is clear that the structure in X_2 contains $V = span\{e_0, e_1, \ldots, e_4, e_{16}, \ldots, e_{20}\}$, and it is only left to determine which two of the four elements $0, e_5, e_{21}, e_5 \oplus e_{21}$ we should add. Now, we observe that there are only three possible pairs of elements (along with their complements), and one of them must lead to the desired form after the FL. Hence, it is sufficient to try 3 structures in X_2 to ensure that the HOD equation holds for one of them. Note that the structures we use are not standard HODs, as they do not correspond to an affine subspace.

In order to exploit all possible 11'th order differentials, we have to try 3 options for each of the 6 pairs of neighbor bit positions $(0, 16), (1, 17), \ldots, (5, 22)$, and thus, to repeat the attack 3^6 times. As we show below, all the steps of the attack can be performed very efficiently, such that even when they are repeated 3^6 times, the overall time complexity is still dominated by encrypting the plaintexts. We obtain 12 11'th order differentials, but due to linear dependence, we can use only 7 of them simultaneously. By using the same arguments (with $FL3 \circ FL1$ in place of $FL3$), we can divide the 44'th order differential of 4-round MISTY1 used in [13] into 12 43'th order differentials and use 7 of them simultaneously.

It is important to note that the "correct" structure in X_2 depends only on the secret key bits of the FL's, and hence, is common to *all* structures. We also note that an alternative way to overcome the FL layers is to guess the relevant subkey bits (e.g., bits $KL_{3,1}[5]$ and $KL_{3,2}[5]$ in the above example). However, such a guess for all 6 relevant pairs of neighbor positions requires to repeat the attack 4^6 times if the 12'th order differential is used, and 16^6 times if the 44'th order differential is used (as $FL3 \circ FL1$ has 4 key bits in each 2-bit to 2-bit function). Hence, our strategy of overcoming FL is significantly more efficient.

Now, we consider the time complexity of the improved attack. The stage of solving the linear equation system requires now at most $7 \cdot 220^3 \cdot 3^6 = 2^{35.6}$ operations, which is negligible compared to the time required for encrypting the

plaintexts. We show now that the stage of constructing the linear equations can be also performed very efficiently, using the *partial sums* technique.

We observe that the 188 coefficients of the linear equations (except for the constant coefficient) can be divided into two groups of 130 and 58 coefficients, such that the first group depends only on bits $X_6[0-6]$ and their neighbors, and the second group depends only on bits $X_6[7-15]$ and their neighbors. Such a "separation" property of the MISTY1 round function was already used in [3,11,12]. As a result, we can compute these sets of coefficients separately.

Consider the computation of the 130 coefficients that depend only on $X_6[0-6,16-22]$ for a single structure of size 2^{43}. The basic observation used in *partial-sum* techniques is that if for two ciphertexts, the corresponding values of $X_6[0-6,16-22]$ are equal, then the contributions of these ciphertexts to $\bigoplus_{y \in span\{e_0,e_1,\ldots,e_{38},e_{48},\ldots,e_{54}\}} Out_6[31](x \oplus y)$ cancel each other. Hence, before computing the contribution of each ciphertext to the coefficients, we can reduce the structure into a list of size 2^{14} that represents the information on which of the 2^{14} possible values of $X_6[0-6,16-22]$ appears an odd number of times in the structure. Furthermore, this reduction (or most of it, to be precise) can be performed before the guess of the exact structures in X_2, and hence it has to be performed only once for each structure. After the reduction is performed, we go over all 2^{14} values in the list and collect their contributions to the coefficients of the equations. (As shown in the next section, this part can also be performed more efficiently). The total time complexity of this step for each structure is

$$7 \cdot 3^6 \cdot (130 \cdot 2^{14} + 58 \cdot 2^{18}) \ll 2^{43},$$

and hence, the overall time complexity is dominated by the encryption of the plaintexts.

Summarizing the attack, the data and time complexities of the attack are 2^{47} chosen plaintexts and time, an improvement by a factor of 5 in both data and time complexities over the results of Tsunoo et al. (Note that the improvement is only by a factor of 5 and not by a factor of 7, since in order to exploit the structures optimally, we have to use "full" structures of size 2^{46}. Since a single structure is not sufficient, we must use two structures, and thus, the data complexity is 2^{47}.)

5 Improved Attack on 7-Round MISTY1 with 4 *FL* Layers

In this section we describe an attack on 7-round MISTY1 with all FL layers except the last layer ($FL9, FL10$), that improves the attack presented in [14].

As the attack on 6-round MISTY1, our attack is based on the 46'th order differential $V^{(46)}_{P[0-31,32-38,48-54]} X_6[25-31] = 0$ and the attack equation derived from it:

$$\bigoplus_{y \in span\{e_0,e_1,\ldots,e_{38},e_{48},\ldots,e_{54}\}} Out_6[31](x \oplus y).$$

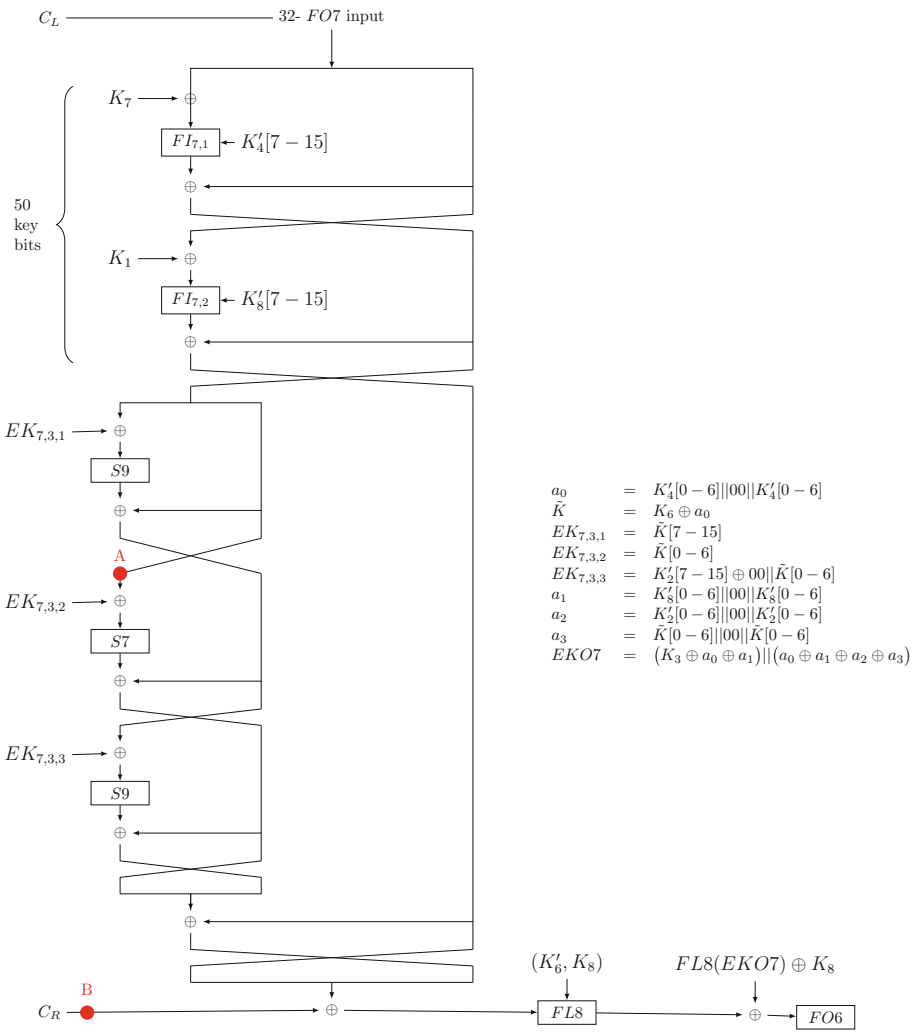


Fig. 3. Reference figure for the attack on 7-round MISTY1 without the last FL layer

However, in the case of the 7-round MISTY1 we have to guess some key material before creating the linear system. Since the attack procedure is a bit complex, we illustrate it in Fig. 3 which includes the order of the steps, as well as equivalent subkeys that we will use.

As noted in Sect. 4, the 188 coefficients of the linear equations (except for the constant coefficient) can be divided into two groups of 130 and 58 coefficients, that can be calculated separately. We describe the calculations for a single structure that corresponds to a 44'th order differential (recall that by [13], each 46'th round differential can be used to construct 22 such structures).

Calculation of the Coefficients Related to the Bits of $X_6[7-15]$. The procedure consists of several steps.

1. We guess the 50 key bits marked in the figure and decrypt all 2^{44} ciphertexts through the first two FI layers of round 7. At this stage, we note that if two intermediate values agree on 34 bit positions, which are the 16 bits of the input to $FI_{7,3}$ and the 18 bits $C_R[7-15, 23-31]$, then their contributions to the sum $\bigoplus_{y \in span\{e_0, e_1, \ldots, e_{38}, e_{48}, \ldots, e_{54}\}} Out_6[31](x \oplus y)$ cancel each other. Hence, as in the 6-round attack, we can reduce the list of ciphertexts into a list of size 2^{34} that shows for each of the 2^{34} values of these 34 bits, whether it appears an odd number of times in the structure.
2. We guess 18 bits of $EK_{7,3,1}, EK_{7,3,3}$ and partially encrypt our list of 2^{34} values through $FI_{7,3}$. After this stage, we note that if two intermediate values agree on 25 bits positions (7 bits in the input to $S7$ of $FI_{7,3}$, marked by A, and 18 bits in C_R, marked by B), then their contributions to the sum cancel out. Hence, we can further reduce the list to 2^{25} values.
3. At this stage, when only one $S7$ (along with a 7-bit subkey) is left, we do not guess the remaining subkey, but rather represent the sum in the attack equation as a function of the current intermediate state and linearize this representation. The number of variables we obtain is $771 \approx 2^{9.6}$, and we can compute the coefficients efficiently using precomputed tables of size $2^{25} \cdot 771$ that store the contribution of each possible value of the 25 bits to the 771 coefficients.
4. We guess the 7 key bits of $EK_{7,3,2}$ and reduce the number of variables to 153. (Note that the number of variables is much higher than 58. This happens since when we guess part of the key after we constructed the equations, we cannot unify linearly dependent variables, and thus, the total number of variables is increased). This reduction can be performed by direct calculation in time complexity of $2^{9.6} \cdot 153$ for each key guess.

We note that guessing the key of the second $S9$ in $FI_{7,3}$ can be done in time complexity of 2^5 (instead of 2^9) by a somewhat complex procedure described in the full version of the paper. Combining all parts of the algorithm together, the time complexity of this part is

$$\begin{aligned} T_{X_6[7-15]} &= 2^{50} \cdot 2^{44} + 2^{50+9} \cdot 2^{34} + 2^{50+9+5} \cdot 2^{34} \\ &\quad + 2^{68} \cdot 2^{25} \cdot \frac{2^{9.6}}{2^6} \cdot 7 + 2^{75} \cdot 2^{9.5} \cdot 153 \\ &= 2^{94} + 2^{93} + 2^{98} + 2^{99.4} + 2^{92.3} \\ &\approx 2^{99.9} \end{aligned}$$

operations.

Calculation of the Coefficients Related to the Bits $X_6[0-6]$. In this case, the procedure is simpler:

1. We guess the 50 key bits marked in the figure and partially decrypt all 2^{44} ciphertexts through the first two FI functions of round 7 (like in the previous case). At this point, if two intermediate values agree on 30 bit positions, which are the 16 bits of the input to $FI_{7,3}$ and the 14 bits $C_R[0-6, 16-22]$, then their contributions to the sum cancel each other. Thus, we can reduce the data to a list of length 2^{30}.
2. We guess the 16 bits of $EK_{7,3,1}, EK_{7,3,2}$, and reduce the data list to a size of 2^{23}, and then guess the 9 bits of $EK_{7,3,3}$ and further reduce the list to only 2^{14} values.
3. We calculate the 130 coefficients of the linear equation using precomputed tables of size $2^{14} \cdot 130$.

The time complexity for this part is

$$\begin{aligned} T_{X_6[0-6]} &= 2^{50} \cdot 2^{44} + 2^{50+16} \cdot 2^{30} + 2^{66+9} \cdot 2^{23} + 2^{74} \cdot 2^{14} \cdot \frac{130}{2^6} \cdot 7 \\ &\approx 2^{94} + 2^{96} + 2^{98} + 2^{93.8} \approx 2^{98.5} \end{aligned}$$

operations.

Combining the Calculations. By combining the sets of coefficients computed in the two calculations described above, we create a system of linear equations. Since we guess 75 key bits overall, we have $7 \cdot 2^{75}$ linear systems to create and solve. The maximum number of variables for each system is $130 + 153 = 283$.

To filter out wrong keys and add a safety factor, we take $283 + 75 + 10 \approx 2^{8.55}$ structures of 44'th order differentials. As each structure corresponding to a 46'th order differential contains 22 structures of 44'th order differentials, we need $\frac{283+75+10}{22} \approx 17 \approx 2^{4.1}$ structures of 46'th order differentials to complete the attack. (Note that in this attack we cannot use our improved 43'th order differentials, since repeating the attack several times (as required for them) would increase the time complexity.) Thus, the total data complexity is $D = 2^{4.1} \cdot 2^{46} = 2^{50.1}$.

The time complexity is composed from encryption of the required data, creation of the linear systems and their solution. The time required for encryption of the data is negligible. The time of creation the linear systems is

$$2^{8.55} \cdot (T_{X_6[7-15]} + T_{X_6[0-6]}) \approx 2^{8.55} \cdot 2^{100.3} = 2^{108.85}.$$

The time of solving the 2^{75} linear systems is $2^{75} \cdot (2^{8.55})^3 \approx 2^{100.65}$ operations. Hence, the total time complexity of the attack is $T = 2^{108.85} + 2^{100.65} \approx 2^{108.85}$ simple operations. Assuming that each round of MISTY1 encryption is comparable to 50 simple operations like was assumed in [14], the time complexity is

$$T = \frac{2^{108.85}}{7 \cdot 50} \approx 2^{100.4}$$

7-round MISTY1 encryptions.

6 New Attack on 7-Round MISTY1 with All FL Layers Present

In this section we describe an attack on 7-round MISTY1 with all FL functions, which is the first attack on this variant that is significantly faster than exhaustive key search. The attack uses the same 44'th order differential and the same division into two types of linear coefficients like the attack presented in Sect. 5. However, in order to handle the 64 subkey bits that are added in this variant, we must perform a more careful procedure, that also takes into consideration the exact MISTY1 key schedule. As in the previous attack, we describe the calculations made for a single structure that corresponds to a 44'th order differential. The reference figure to this attack is Fig. 4.

Calculation of the Coefficients Related to the Bits of $X_6[7-15]$. The procedure consists of several steps.

1. We guess the 57 key bits of $K_1, K_7, K_8, K'_4[7-15]$ and partially decrypt all 2^{44} ciphertexts through $FL10$ and the first two FI functions of round 7. At this stage, we can reduce the data to a list of length 2^{43} (the 43 relevant bits correspond to 18 bits in C_R, 9 bits in the point B and the 16-bit input to $FI_{7,3}$).
2. We guess the 9 bits of $EK_{7,3,1}$. After this guess, the size of the list remains 2^{43} as before, but now the 43 bits correspond to 18 bits in C_R, 9 bits in B and all bits of A.
3. At this point, we perform linearization. Due to the amount of key material which we haven't guessed yet, the maximal possible number of variables is $2713 \approx 2^{11.4}$. Using directly a precomputed table for computing the coefficients requires a table of size $2^{43} \cdot 2^{11.4}$. Instead of this table, we will use three smaller tables. We note that out of the 2713 coefficients, there are 2269 coefficients in which only the bits of A, B and $C_R[7-15]$ are involved and 424 coefficients in which only the bits in A, B and $C_R[24-31]$ are involved. Only 20 variables are left which depend on all the 43 bits. Hence, we can use for the computation three smaller tables of sizes $2^{34} \cdot 2269, 2^{34} \cdot 424$, and $2^{43} \cdot 20$. Hence, the memory complexity required for the linearization is $2^{34} \cdot 2269 + 2^{34} \cdot 424 + 2^{43} \cdot 20 \approx 2^{47.66}$.
4. We guess the $16+9$ key bits of $EK_{7,3,2}, EK_{7,3,3}$ and $K'_3[7-15]$. Using the guessed key bits, and the fact that K_8 and KL_8 are known, we can reduce the number of variables. (As in Sect. 5, guessing the key at this point forces us to not unify linearly dependent variables.) The number of the new variables is only 213. This transformation is done by a direct calculation in time complexity of $2^{11.4} \cdot 213$ for each key guess (there is 2^{91} guesses at this point).

In total, the time complexity of this part is

$$T_{X_6[7-15]} = 2^{57} \cdot 2^{44} + 2^{57+9} \cdot 2^{43} + 2^{66} \cdot 2^{43} \cdot \frac{2^{11.4}}{2^6} \cdot 7 + 2^{91} \cdot 2^{11.4} \cdot 213$$

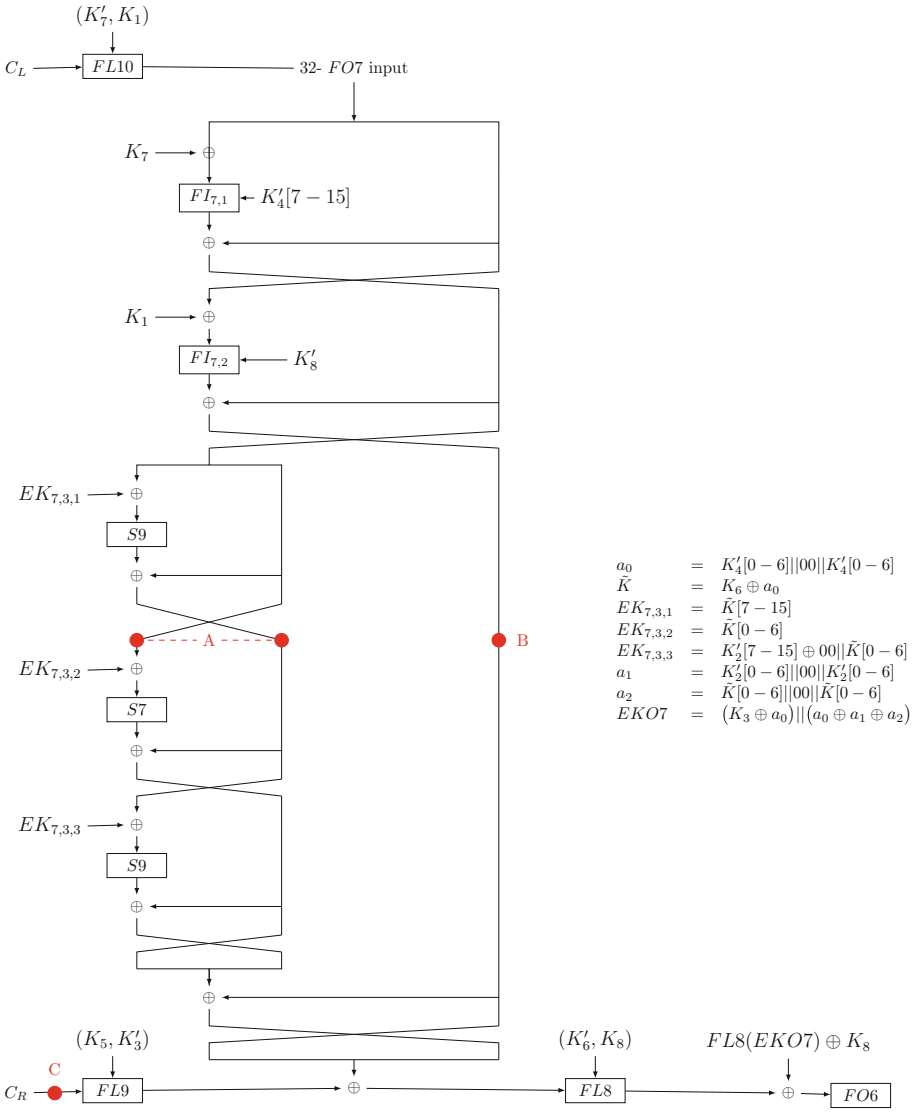


Fig. 4. Reference figure for the 7-round with all FL's functions

$$= 2^{101} + 2^{109} + 2^{117.2} + 2^{110.14}$$
$$\approx 2^{117.22}$$

operations.

Calculation of the Coefficients Related to the Bits $X_6[0-6]$. In this case, the procedure is simpler:

1. We guess the 57 key bits of $K_1, K_7, K_8, K'_4[7-15]$ and partially decrypt all the 2^{44} ciphertexts through $FL10$ and the first two FI layers of round 7. At this stage, the data can be reduced to a list of size 2^{37}, where the 37 bits are 14 bits of C_R, 7 bits in B and the input to $FI_{7,3}$.
2. We guess the 25 bits of $EK_{7,3,i} \quad i = 1, 2, 3$. After this guess, the size of the list is reduced to 2^{28}, which corresponds to 14 bits in C_R (after FL_9) and 14 bits before FL_9.
3. At this point we perform linearization. The maximum number of variables is $684 \approx 2^{9.42}$. We calculate them using a precomputed table of size $2^{28} \cdot 684$.

The time complexity for this part is

$$\begin{aligned} T_{X_6[0-6]} &= 2^{57} \cdot 2^{44} + 2^{57+25} \cdot 2^{37} + 2^{82} \cdot 2^{28} \cdot \frac{2^{9.42}}{2^6} \cdot 7 \\ &\approx 2^{101} + 2^{119} + 2^{116.22} \approx 2^{119.2} \end{aligned}$$

operations.

Combining the Calculations We create the system of linear equations using the two previous calculations. We have $7 \cdot 2^{91}$ linear systems to create and solve. The maximum number of variable for each system is $213 + 684 = 897$. (Note that this time, the number of variables is significantly larger than in the previous attacks, due to the amount of key material which we absorb into the equations.)

To filter out wrong keys and add a safety factor, we take $897 + 91 + 10 \approx 2^{9.97}$ structures that correspond to 44'th order differentials. Therefore, we need $\frac{897+91+10}{22} \approx 43.6 \approx 2^{5.45}$ structures of 46'th order differentials, which means that the total data complexity is

$$D = 2^{5.45} \cdot 2^{46} = 2^{51.45}.$$

The time complexity is composed from encryption of the required data, creation of linear systems and their solution. The time of data encryption is negligible. The time of creation the linear systems is $2^{9.97} \cdot (T_{X_6[7-15]} + T_{X_6[0-6]}) \approx 2^{9.97} \cdot 2^{119.56} = 2^{129.53}$ operations. The time of solving the 2^{91} linear systems is $2^{91} \cdot (2^{9.97})^3 \approx 2^{120.91}$ operations. The total time complexity is thus $T = 2^{129.53} + 2^{120.91} \approx 2^{129.53}$ simple operations. Assuming that each round of MISTY1 encryption is comparable to 50 simple operations, the time complexity is

$$T = \frac{2^{129.53}}{7 \cdot 50} \approx 2^{121}$$

7-round encryptions.

7 Summary and Conclusions

In this paper we investigated higher-order differential attacks on MISTY1. We improved the 44'th order differential used in the best previously known attack of

Tsunoo et al. [13] into a 43'th order differential, and used it to reduce the data and time complexities of the best known attack on 6-round MISTY1 from $2^{49.3}$ to $2^{46.5}$. We gave evidence that our 43'th order differential cannot be further improved using current techniques.

We also considered the best known higher-order differential attacks on 7-round MISTY1, also by Tsunoo et al. We showed that by using the partial sums technique and other techniques, the time complexity of the attack can be reduced from $2^{116.4}$ to $2^{100.4}$ – a reduction by a factor of 2^{16}. Finally, we presented an attack on 7-round MISTY with all FL functions present that requires $2^{51.5}$ chosen plaintexts and 2^{121} encryptions. This is the first known attack on a variant of MISTY1 with all FL layers.

As a problem for further research, it will be interesting to find out whether our techniques can be used also to improve higher-order differential attacks on KASUMI. It seems that the case of KASUMI will be harder, due to the higher algebraic degree of the modified FI function KASUMI uses.

References

1. 3rd Generation Partnership Project. Specification of the 3GPP Confidentiality and Integrity Algorithms - Document 2: KASUMI Specification (Release 6). Technical report 3GPP TS 35.202 V6.1.0 (2005–09), September 2005
2. 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)
3. Dunkelman, O., Keller, N.: Practical-time attacks against reduced variants of MISTY1. IACR Cryptol. ePrint Arch. **2013**, 431 (2013)
4. Ferguson, N., Kelsey, J., Lucks, S., Schneier, B., Stay, M., Wagner, D., Whiting, D.L.: Improved cryptanalysis of rijndael. In: Schneier, B. (ed.) FSE 2000. LNCS, vol. 1978, pp. 213–230. Springer, Heidelberg (2001)
5. Jia, K., Li, L.: Improved impossible differential attacks on reduced-round MISTY1. In: Lee, D.H., Yung, M. (eds.) WISA 2012. LNCS, vol. 7690, pp. 15–27. Springer, Heidelberg (2012). http://dblp.uni-trier.de/db/conf/wisa/wisa2012.html#JiaL12
6. Jia, K., Hongbo, Y., Wang, X.: A meet-in-the-middle attack on the full kasumi. IACR Cryptol. ePrint Arch. **2011**, 466 (2011)
7. Knudsen, L.R.: Truncated and higher order differentials. In: Preneel, B. (ed.) FSE 1994. LNCS, vol. 1008, pp. 196–211. Springer, Heidelberg (1995)
8. Lai, X.: Higher order derivatives and differential cryptanalysis. In: Blahut, R.E., Costello Jr., D.J., Maurer, U., Mittelholzer, T. (eds.) Communications and Cryptography. The International Series in Engineering and Computer Science, vol. 276, pp. 227–233. Springer, Heidelberg (1994)
9. Lu, J., Yap, W.-S., Wei, Y.: Weak keys of the full MISTY1 block cipher for related-key differential cryptanalysis. In: Dawson, E. (ed.) CT-RSA 2013. LNCS, vol. 7779, pp. 389–404. Springer, Heidelberg (2013)
10. Matsui, M.: New block encryption algorithm MISTY. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 54–68. Springer, Heidelberg (1997)
11. Sakurai, K., Zheng, Y.: On non-pseudorandomness from block ciphers with provable immunity against linear cryptanalysis. In: Proceedings of AAECC 1999. LNCS, vol. 1719, pp. 19–24. Springer (1999)

12. Sun, X., Lai, X.: Improved integral attacks on MISTY1. In: Jacobson Jr., M.J., Rijmen, V., Safavi-Naini, R. (eds.) SAC 2009. LNCS, vol. 5867, pp. 266–280. Springer, Heidelberg (2009)
13. Tsunoo, Y., Saito, T., Kawabata, T., Nakagawa, H.: Finding higher order differentials of MISTY1. IEICE Trans. **95**–**A**(6), 1049–1055 (2012)
14. Tsunoo, Y., Saito, T., Shigeri, M., Kawabata, T.: Higher order differential attacks on reduced-round MISTY1. In: Lee, P.J., Cheon, J.H. (eds.) ICISC 2008. LNCS, vol. 5461, pp. 415–431. Springer, Heidelberg (2009)
15. Yi, W., Chen, S.: Multidimensional zero-correlation linear attacks on reduced-round MISTY1. CoRR, abs/1410.4312 (2014)

http://www.springer.com/978-3-662-48115-8

Fast Software Encryption
22nd International Workshop, FSE 2015, Istanbul, Turkey, March 8-11, 2015, Revised Selected Papers
Leander, G. (Ed.)
2015, XI, 600 p. 131 illus., Softcover
ISBN: 978-3-662-48115-8