Multidimensional Zero-Correlation Linear Cryptanalysis of the Block Cipher KASUMI

Wentan Yi* **and Shaozhen Chen**

State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China

Abstract. The block cipher KASUMI is widely used for security in many synchronous wireless standards. It was proposed by ETSI SAGE for usage in 3GPP (3rd Generation Partnership Project) ciphering algorithms in 2001. There are a great deal of cryptanalytic results on KASUMI, however, its security evaluation against the recent zero-correlation linear attacks is still lacking so far. In this paper, we select some special input masks to refine the general 5-round zero-correlation linear approximations combining with some observations on the FL functions and then propose the 6-round zero-correlation linear attack on KASUMI. Moreover, zero-correlation linear attacks on the last 7-round KASUMI are also introduced under some weak keys conditions. These weak keys take more than half of the whole key space.

The new zero-correlation linear attack on the 6-round needs about $2^{107.8}$ encryptions with $2^{59.4}$ known plaintexts. For the attack under weak keys conditions on the last 7 round, the data complexity is about $2^{62.1}$ known plaintexts and the time complexity $2^{125.2}$ encryptions.

Keywords: KASUMI, Zero-correlation linear cryptanalysis, Cryptography.

1 Introduction

With the rapid growth of wireless services, various security algorithms have been developed to provide users with effective and secure communications. The KASUMI developed from a previous block cipher known as MISTY1[10], which was chosen as the foundation for the 3GPP confidentiality and integrity algorithm[14]. Nowadays, it is widely used in UMTS, GSM and GPRS mobile communications. KASUMI adopts the basic Feistel structure and has eight rounds. It accepts a 64-bit block and a 128-bit key.

Up to now, a great deal of attention was paid to KASUMI and many cryptanalytic methods were used to evaluate its security, such as differential cryptanalysis[6], integral-interpolation attack[11], higher order differential attack[12][13], sandwich attack[7] and impossible differential attack[8]. In the past years, higher order differential attack[12][13]and integral-interpolation attack[5] were applied to analyze variants of KASUMI. Kühn [9] presented an

* Corresponding authors.
E-mail addresses: nlwt8988@gmail.com.

Attack Type	Rounds	Date	Time	Source
Higher-Order Differential	5	$2^{22.1}$CP	$2^{60.7}$Enc	[6]
Higher-Order Differential	5	$2^{28.9}$CP	$2^{31.2}$Enc	[7]
Integral-Interpolation	6	2^{48}CP	$2^{126.2}$Enc	[5]
Impossible Differential	6	2^{55}CP	2^{100}Enc	[9]
Multidimensional Zero-Correlation	6	$2^{59.4}$KP	$2^{107.8}$Enc	Sect.[4]
Impossible Differential	7	$2^{52.5}$CP	$2^{114.3}$Enc	[14]
Impossible Differential	7	2^{62}CP	$2^{115.8}$Enc	[14]
Multidimensional Zero-Correlation	7	$2^{62.1}$KP	$2^{125.2}$Enc	Sect.[5]

CP refers to the number of chosen plaintexts.
KP refers to the number of known plaintexts.
Enc refers to the number of encryptions.

Table 1: The key schedule of KASUMI

impossible differential attack on a 6-round version at EuroCrypt 2001. Later, Jia et al[8] refined the impossible differential by selecting some special input differential values and extended the 12-years old impossible differential attack on 6-round KASUMI to 7 rounds at SAC 2013. In the related-key setting, attacks on full 8-round [5] KASUMI were presented using related-key booming and rectangle attack and the complexes are about $2^{78.7}$ and $2^{76.1}$ encryptions, respectively. At Crypto 2010, a new strategy called sandwich attacks [7] belonging to a formal extension of booming attacks on the full KASUMI was obtained.

Bogdanov and Rijmen[1] proposed zero-correlation linear cryptanalysis. It is a novel promising attack technique for block ciphers and has its theoretical foundation in the availability of numerous key-independent unbiased linear approximations with correlation zero for many ciphers. However, the initial distinguisher of [1] had some limitations in terms of data complexity, which needs at least half of the codebook. In FSE 2012, Bogdanov and Wang [2] proposed a more data-efficient distinguisher by making use of multiple linear approximations with correlation zero. The date complexity is reduced, however, the distinguisher relies on the assumption that all linear approximations with correlation zero are independent. At AsiaCrypt 2012[3], a multidimensional distinguisher has been constructed for the zero-correlation property, which removed the unnecessary independency assumptions on the distinguishing side. Recently, multidimensional zero-correlation linear cryptanalysis has been using in the attack of block cipher CAST-256[3], CLEFIA[4], HIGHT[15] and E2[16] successfully.

In this paper, we evaluate the security of KASUMI with respect to the recent technique of zero-correlation linear cryptanalysis. Our contributions can be summarized as follows.

1. We reveal some 5-round linear approximations of correlation zero in KASUMI. For the zero-correlation linear approximations of 5-round KASUMI: $(\bar{a}, 0) \overset{\text{5-round}}{\nrightarrow} (\bar{a}, 0)$, if we take all non-zero values for $\bar{a}$, then there are so many guessed subkey bits involved in the key recovery process that the time complexity will be greater than exhaustive search. Therefore, in order to reduce the number of guessed subkey bits, we only use some special linear approximations.

Based on some observations on FL function, we give some conditions the special linear approximations should satisfy.

2. We propose a multidimensional zero-correlation linear attack on 6-round of KASUMI. To my knowledge, there are no linear attacks on KASUMI so far and we bridge this gap, if we treat the zero-correlation linear attack as a special case of linear attacks.

3. We provide a key-recovery attack on 7-round KASUMI(round 2 to round 8) under some weak key conditions. We assume that the second keys of FL function in round 2 and round 8 have the some value in at least 8 bit positions. The purpose is to make a balance between selecting out enough linear approximations and more master key satisfying the assumption.

The paper is organized as follows. We give a brief description of the block cipher KASUMI and outlines the ideas of multidimensional zero-correlation linear cryptanalysis in Section 2. Some observations on FL function are shown in Section 3. Section 4 and Section 5 illustrate our attacks on 6-round and the last 7-round KASUMI. We conclude in Section 6.

2 Preliminarise

2.1 Description of KASUMI

The KASUMI algorithms [14] are symmetric block ciphers with a block size of 64 bits and a key size of 128 bit. We give a brief description of KASUMI in this section.

KASUMI is a Feistel structure with 8 round, see Fig. 1 (a) for an illustration. The round function consists of an FL function and an FO function. The FL function is a simple key-dependent boolean function, depicted in Fig. 1 (c). Let the inputs of the FL function of the i-th round be $XL_i = XL_{i,l}\|XL_{i,r}, KL_i = (KL_{i,1}, KL_{i,2})$, the output be $YL_i = YL_{i,l}\|YL_{i,r}$, where $XL_{i,l}$,$XL_{i,r}$,$YL_{i,l}$ and $YL_{i,r}$ are 16-bit integers. We define the FL function as follows:

$$YL_{i,r} = ((XL_{i,l} \cap KL_{i,1}) \lll 1) \oplus XL_{i,r},$$

$$YL_{i,l} = ((YL_{i,r} \cup KL_{i,2}) \lll 1) \oplus XL_{i,l},$$

where $\cap$ and $\cup$ denote bitwise AND and OR respectively, $x \ggg i$ implies that x rotates left by i bits, $\oplus$ denotes the bitwise exclusive-or (XOR), $\|$ represents the concatenation, and FL_i denote the FL function of i-th round with subkey KL_i.

The FO function is depicted in Fig. 1 (b), which is another three-round Feistel structure consisting of three FI functions and key mixing stages. Let $XO_i = XO_{i,l}\|XO_{i,r}$, $KO_i = (KO_{i,1}, KO_{i,2}, KO_{i,3})$, $KI_i = (KI_{i,1}, KI_{i,2}, KI_{i,3})$ be the inputs of the FO function of i-th round, and $YO_i = YO_{i,l}\|YO_{i,r}$ be the corresponding output, where $XO_{i,l}$,$XO_{i,r}$,$YO_{i,l}$,$YO_{i,r}$ and $\overline{XI_{i,3}}$ are 16-bit integers. Then the FO function has the form

$$\overline{XI_{i,3}} = FI((XO_{i,l} \oplus KO_{i,1}), KI_{i,1}) \oplus XO_{i,r},$$

$$YO_{i,l} = FI((XO_{i,r} \oplus KO_{i,2}), KI_{i,2}) \oplus \overline{XI_{i,3}},$$

Algorithm 1 The KASUMI block cipher

Require: 64-bit plaintext $P = (L_0, R_0)$; main key K,
Ensure: 64-bit ciphertext $C = (L_8, R_8)$.
1: Derive round keys KO_i, KI_i and KL_i $(1 \leq i \leq 8)$ from K.
2: for $j = 1$ to 8 do
3: if j is odd, do
4: $L_j = FO(FL(L_{j-1}, KL_j), KO_j, KI_j) \oplus R_{j-1}, R_j = L_{j-1}$,
5: else, do :
6: $L_j = FL(FO(L_{j-1}, KO_j, KI_j), KL_j) \oplus R_{j-1}$, $R_j = L_{j-1}$.
7: end for
8: return $C = (L_8, R_8)$.

$$YO_{i,r} = FI((XI_{i,3} \oplus KO_{i,3}), \overline{KI_{i,3}}) \oplus YO_{i,l}.$$

For simplicity, FO_i denotes the FO function of i-th round.

The FI function uses two S-boxes S_7 and S_9 which are permutations of 7-bit to 7-bit and 9-bit to 9-bit respectively. Suppose the inputs of the j-th FI function of the i-th round are $XI_{i,j}$, $KI_{i,j}$ and the output is $YI_{i,j}$, where $XI_{i,j}$ and $YI_{i,j}$ are 16-bit integers. We define half of FI function as $\overline{FI}$, which is a 16-bit to 16-bit permutation. The structure of $\overline{FI}$ and FI is depicted in Fig. 1 (c). $\overline{YI_{i,j}} = FI(XI_{i,j})$ is defined as

$$\overline{YI_{i,j}}[0 \sim 8] = S_9(XI_{i,j}[7 \sim 15]) \oplus XI_{i,j}[0 \sim 6],$$

$$\overline{YI_{i,j}}[9 \sim 15] = S_7(XI_{i,j}[0 \sim 6]) \oplus \overline{YI_{i,j}}[0 \sim 6],$$

where $z[i_1 \sim i_2]$ denotes the $(i_2 \sim i_1)$bits from the i_1-th bit to i_2-th bit of z, and 0 is the least significant bit. The FI function is simplified as

$$YI_{i,j} = FI(XI_{i,j}, KI_{i,j}) = \overline{FI}((\overline{FI}(XI_{i,j}) \oplus KI_{i,j}) \lll 7),$$

and we denote $FI_{i,j}$ as the j-th FI function of the i-th round with subkey $KI_{i,j}$.

Let $L_i||R_i = \big((L_{i,l}\|L_{i,r})\|(R_{i,l}\|R_{i,r})\big)$ be the input of the i-th round, and then the round function is defined as

$$L_i = FO(FL(L_{i-1}, KL_i), KO_i, KI_i) \oplus R_{i-1}, R_i = L_{i-1},$$

where $i = 1, 3, 5, 7$, and when $i = 2, 4, 6, 8$,

$$L_i = FL(FO(L_{i-1}, KO_i, KI_i), KL_i) \oplus R_{i-1},\ R_i = L_{i-1}.$$

Here, L_0, R_0, L_8, R_8 are the plaintext and ciphertext respectively, and L_{i-1}, R_{i-1} denote the left and right 32-bit halves of the i-th round input. The KASUMI cipher can be described in Algorithm 1.

Round	$KL_{i,1}$	$KL_{i,2}$	$KO_{i,1}$	$KO_{i,2}$	$KO_{i,3}$	$KI_{i,1}$	$KI_{i,2}$	$KI_{i,3}$
1	$k_1 \lll 1$	k'_3	$k_2 \lll 5$	$k_6 \lll 8$	$k_7 \lll 13$	k'_5	k'_4	k'_8
2	$k_2 \lll 1$	k'_4	$k_3 \lll 5$	$k_7 \lll 8$	$k_8 \lll 13$	k'_6	k'_5	k'_1
3	$k_3 \lll 1$	k'_5	$k_4 \lll 5$	$k_8 \lll 8$	$k_1 \lll 13$	k'_7	k'_6	k'_2
4	$k_4 \lll 1$	k'_6	$k_5 \lll 5$	$k_1 \lll 8$	$k_2 \lll 13$	k'_8	k'_7	k'_3
5	$k_5 \lll 1$	k'_7	$k_6 \lll 5$	$k_2 \lll 8$	$k_3 \lll 13$	k'_1	k'_8	k'_4
6	$k_6 \lll 1$	k'_8	$k_7 \lll 5$	$k_3 \lll 8$	$k_4 \lll 13$	k'_2	k'_1	k'_5
7	$k_7 \lll 1$	k'_1	$k_8 \lll 5$	$k_4 \lll 8$	$k_5 \lll 13$	k'_3	k'_2	k'_6
8	$k_8 \lll 1$	k'_2	$k_1 \lll 5$	$k_5 \lll 8$	$k_6 \lll 13$	k'_4	k'_3	k'_7

$x \lll i$: x rotates left by i bits. $k'_i = k'_i \oplus c_i$, where the c_is are fixed constants.

Table 2: The key schedule of KASUMI

The key schedule of KASUMI is much simpler than the original key schedule of MISTY1. The 128-bit key K is divided into eight 16-bit words: $(k_1, k_2, ..., k_8)$, i.e., $K = (k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8)$. In each round, eight key words are used to compute the round subkeys, which are made up of three parts KL_i, KO_i and KI_i. Here, $KL_i = (KL_{i,1}, KL_{i,2}), KO_i = (KO_{i,1}, KO_{i,2}, KO_{i,3})$ and $KI_i = (KI_{i,1}, KI_{i,2}, KI_{i,3})$. We summarize the details of the key schedule of KASUMI in Tab. 2.

2.2 Zero-correlation cryptanalysis

In this section, we briefly recall the basic concepts of zero-correlation linear cryptanalysis based on [1], [2] and [3]. Linear cryptanalysis is based on linear approximations determined by input mask a and output mask β. A linear approximation $a \to \beta$ of a vectorial function f has a correlation denoted by

$$C(\beta \cdot f(x), a \cdot x) = 2Pr_x(\beta \cdot f(x) \oplus a \cdot x = 0) - 1,$$

where we denote the scalar product of binary vectors by $a \cdot x = \oplus_{i=1}^{n} a_i x_i$.

In zero-correlation linear cryptanalysis, the distinguisher uses linear approximations with zero correlation for all keys while the classical linear cryptanalysis utilizes linear approximations with correlation as far from zero as possible. Bogdanov et al. [3] proposed a multidimensional zero-correlation linear distinguisher using l zero-correlation linear approximations and requiring $O(2^n/\sqrt{l})$ known plaintexts, where n is the block size of a cipher.

We treat the zero-correlation linear approximations available as a linear space spanned by m base zero-correlation linear approximations such that all $l = 2m - 1$ non-zero linear combinations of them have zero correlation. For each of the 2^m data values $z \in F_2^m$, the attacker initializes a counter $V[z]$, $z = 0, 1, ..., 2m - 1$ to value zero. Then, for each distinct plaintext, the attacker computes the corresponding data value in F_2^m by evaluating the m basis linear approximations and increments the counter $V[z]$ of this data value by one. Then

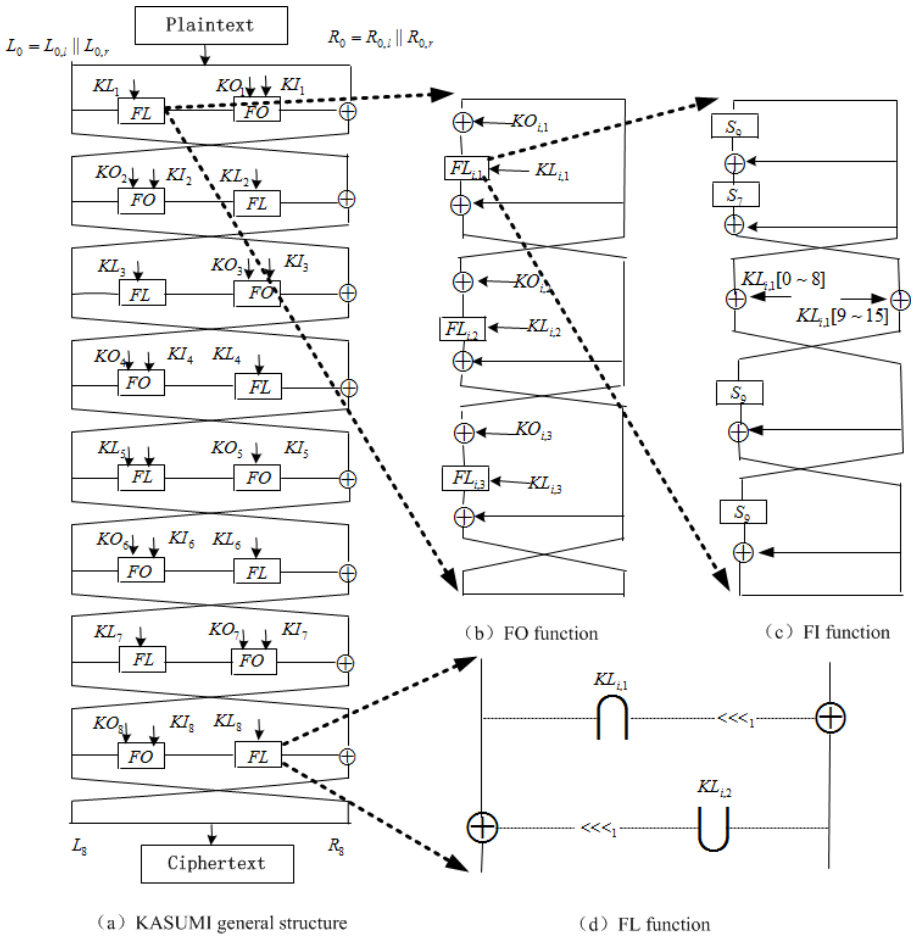

Figure 1: The structure and building blocks of KASUMI

the attacker computes the statistic T:

$$T = \sum_{i=0}^{2^m-1} \frac{(v[z] - N2^{-m})^2}{N2^{-m}(1-2^{-m})}.$$

The statistic T follows a $\mathcal{X}^2$ -distribution with mean $\mu_0 = (l-1)\frac{2^n-N}{2^n-1}$ and variance $\sigma_0^2 = 2(l-1)\left(\frac{2^n-N}{2^n-1}\right)^2$ for the right key guess, while for the wrong key guess, it follows a X^2-distribution with mean $\mu_1 = l-1$ and variance $\sigma_1^2 = 2(l-1)$.

If we denote the probability of false positives and the probability of false negatives to distinguish between a wrong key and a right key as β_0 and β_1, respectively, and we consider the decision threshold $\tau = \mu_0 + \sigma_0 z_{1-\beta_0} = \mu_1 - \sigma_1 z_{1-\beta_1}$, then the number of known plaintexts N should be about

$$N = \frac{(2^n-1)(z_{1-\beta_0} + z_{1-\beta_1})}{\sqrt{(l-1)/2} + z_{1-\beta_0}} + 1,$$

where $z_{1-\beta_0}$ and $z_{1-\beta_1}$are the respective quantiles of the standard normal distribution, See [3] for detail.

3 Some Observations of KASUMI

Let $M = (m_0, m_1, ..., m_{l-1})$, $x = (x_0, x_1, ..., x_{l-1})$, $\neg M = (\neg m_0, \neg m_1, ..., \neg m_{l-1})$, $M \cup x = (m_0 \cup x_0, m_1 \cup x_1, ..., m_{l-1} \cup x_{l-1})$, $M \cap x = (m_0 \cap x_0, m_1 \cap x_1, ..., m_{l-1} \cap x_{l-1})$ and $M \diamond x = (m_0 x_0, m_1 x_1, ..., m_{l-1} x_{l-1})$. We describe some observations on the FL functions, which are used in our cryptanalysis of KASUMI.

Observation 1. *Let M be a l-bit value and $h_1(x) = M \cup x$, $h_2(x) = M \cap x$. Then $C(\beta \cdot h_1(x), a \cdot x) \neq 0$ if and only if $a = \neg M \diamond \beta$ and $C(\beta \cdot h_2(x), a \cdot x) \neq 0$ if and only if $a = M \diamond \beta$, see Figure 2 (a)(b).*

We only consider the function $h_1(x)$. For any $i \in (0, l-1)$, if $m_i = 0$, then the input mask a_i is the same with the output mask β_i. If $m_i = 1$, then $a_i = 0$, no matter what the value β_i takes. The following result can be deduced from Observation 1.

Observation 2. *If the output mask of FL function is (a, a'), where $a'[i] = a[i-1] \neg KL_2[i]$, $i = 1, 2, ...l-1$, and $a'[0] = a[l] \neg KL_2[l]$, that is, $a' = (a \lll 1) \diamond KL_2$, then the input mask of FL function is $(a, 0)$, see Figure 2(c).*

Base on Observation 2 and the structure of round functiom of the KASUMI block cipher, we have the following two results.

Observation 3. *If the input mask of FL_6 function is (a, a'), where $a' = (a \lll 1) \diamond \neg KL_{6,2}$, then the input masks of $FO_{6,l}$ and $FO_{6,r}$ function only depend on the 64-bit subkey k_1, k_4, k_5 and k_3, see Figure 3.*

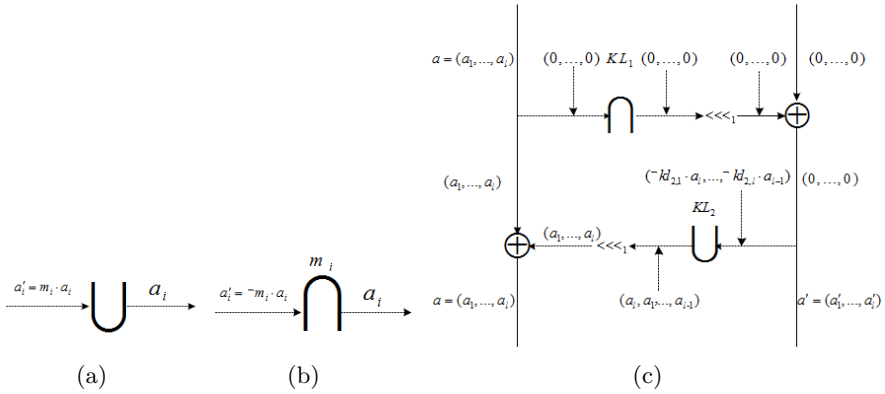

Figure 2: Property of OR, AND and FL function

Observation 4. *Let (a, a') be the output mask of FL_2 and FL_8 functions, where $a[i-1] = 0$, if $KL_{2,2}[i] \oplus KL_{8,2}[i] = 1$, $i = 1, 2, ...l-1$ and $a[l] = 0$, if $KL_{2,2}[0] \oplus KL_{8,2}[0] = 1$, and let $a' = (a \lll 1) \diamond \neg KL_{2,2} \diamond \neg KL_{8,2}$, then the output mask of $FI_{2,3}$ and $FI_{8,3}$ be zero, and the input masks of $FO_{2,l}$, $FO_{2,r}$, $FO_{8,l}$ and $FO_{8,r}$ functions depend on the 96-bit subkey k_3, k_6, k_7, k_5, k_1 and k_4, see Figure 4.*

4 Key-recovery attack on the 6 Rounds of KASUMI

The generic 5-round zero-correlation linear approximations of Feistel structure was introduced by Bogdanov and Rijmen in [1], which is: $(\bar{a}, 0) \overset{\text{5-Round}}{\nrightarrow} (\bar{a}, 0)$, where $\bar{a}$ is a 32-bit non-zero value. Combined with the Feistel structure of the round function, some special values of input mask $\bar{a}$ are selected to attack the 6-round version of KASUMI. We mount the 5-round zero-correlation linear approximations from round 1 to round 5, and extend one round backward. We select the 5-round zero-correlation linear approximations as:

$$(a \| a', 0) \overset{\text{5-Round}}{\nrightarrow} (a \| a', 0),$$

where a is 16-bit non-zero value and $a' = (a \lll 1) \diamond \neg KL_{6,2}$. The choice is to minimize the key words guessing during the attack on 6 rounds of KASUMI. Based on observations 3, we know that, if the input mask of the first round is selected as above, $FI_{6,3}$ and $FO_{6,3}$ are not involved in the computation, which can help us to reduce the complexity of the attack. The zero-correlation linear attack on the 6-round variant of KASUMI is demonstrated as follows, see also Fig. 3.

In our attack, we guess the subkey and evaluate the linear approximation $(a, a')^T \cdot \big((L_{0,l}, L_{0,r}) \oplus (R_{5,l}, R_{5,r})\big) = 0$, that is

$$(a, a') \cdot (L_{0,l} \oplus L_{6,l}, L_{0,r} \oplus L_{6,r}) \oplus a \cdot \big(FI(L_{5,l} \oplus (k_1 \lll 5), k'_4) \oplus L_{5,r} \oplus FI(L_{5,r} \oplus (k_5 \lll 8), k'_3)\big) = 0,$$

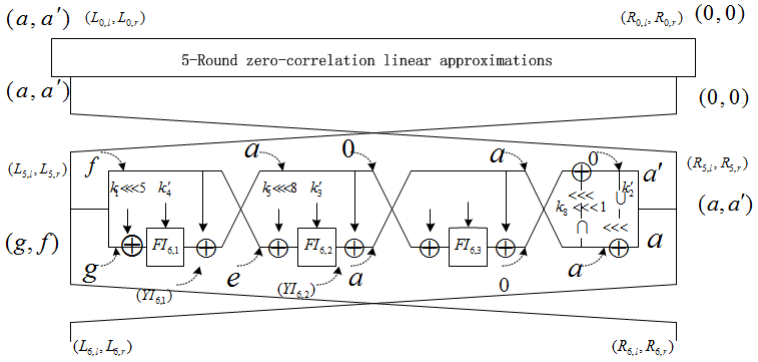

Figure 3: Multidimensional Zero-correlation attack on 6-round KASUMI

where $a' = (a \lll 1) \diamond \neg k'_8$. Then the key-recovery attack on 6-round KASUMI is proceeded with partial-sum technique as follows.

1. Allocate a counter vector $V[L_{5,l}|L_{5,r}|L_{5,r} \oplus R_{5,l} \oplus L_{0,l}|R_{5,r} \oplus L_{0,r}]$ of size 64 where each element is 8-bit length and initialize to zero. In this step, about 2^{64} plaintext-ciphertext pairs are divided into 2^{64} different state. The expected pairs for each state is around one. So the assumption V as a 8-bit counter is sufficient.

2. Guess all possible values of 16 subkey bits k'_2.

3. For N plaintext-ciphertext pairs, extract the 48-bit value

$$i = (L_{5,l}|L_{5,r}|X^1)$$

where $X^1 = L_{5,r} \oplus L_{6,l} \oplus L_{0,l} \oplus \Big(\big(\neg k'_2 \diamond (L_{6,r} \oplus L_{0,r})\big) \ggg 1\Big)$ and increment the counter x_i according to the value of i.

4. Guess all possible values of 32 master key bits k_4 and k_1, partially encrypt $L_{5,l}$ to get $YI_{6,1}$. Let $X^2 = X^1 \oplus YI_{6,1}$. Add one to the corresponding $i = (L_{5,r}|X^2)$. The time complexity of Step 4 is no more than $2^{16} \times 2^{32} \times 2^{48} \times \frac{1}{3} \times \frac{1}{6}$ 6-round encryptions.

5. Guess all possible values of 32 master key bits k_5 and k_3, partially encrypt $L_{5,r}$ to get $YI_{6,2}$. Let $X^3 = X^2 \oplus YI_{6,2}$. Add one to the corresponding $i = (X^3)$. The time complexity of Step 5 is no more than $2^{16} \times 2^{32} \times 2^{32} \times 2^{32} \times \frac{1}{3} \times \frac{1}{6}$ 6-round encryptions.

6. After Step 5, 80 master key bits have been guessed and the parity of $a \cdot X^3$ could be evaluated for all zero-correlation linear approximations.

7. Allocate a counter vector $V[z]$ of size 2^{16} where each element is 64-bit length for 16-bit z (z is the concatenation of evaluations of 16 basis zero-correlation masks).

8. For 2^{16} values of X, evaluate all basis zero-correlation masks on X and put the evaluations to the vector z, then add the corresponding $V[z] : V[z] += V[X]$.

9. Compute $T = N2^{16} \sum_{z=0}^{2^{16}-1} (\frac{v[z]}{N} - \frac{1}{2^{16}})$, if $T \leq \tau$,then the guessed key is a possible key candidate. As there are 48 master key bits that we havent guessed, we do exhaustive search for all keys conforming to this possible key candidate.

Step	Guess	Computed States	Counter-Size
1	k_2	$x_i^1 = (L_{5,l}\|L_{5,r}\|X^1)$	48
2	k_4,k_1	$x_i^2 = (L_{5,r}\|X^2)$	32
3	k_5,k_3	$x_i^3 = (X^3)$	16

Table 3: Partial decryption on 6-Round KASUMI

In this attack, we set the type-I error probability $\beta_0 = 2^{-2.7}$ and the type-II error probability $\beta_1 = 2^{-82}$. We have $z_{1-\beta_0} \approx 1$, $z_{1-\beta_1} \approx 10$, $n = 64$, $l = 2^{16} - 1$. The date complex N is about $2^{59.4}$ and the decision threshold $\tau \approx 2^{15.9}$.

There are 80-bit master key value guessed during the encryption phase, and only the right key candidates survive in the wrong key filtration. The complexity of Step 3 is no more than $N2^{16}$ 6-round KASUMI encryptions and the complexity of Step 5 is about $2^{107.8}$ 6-round KASUMI encryptions which is also the dominant part of our attack. In total, the data complexity is about $2^{59.4}$ known plaintexts, the time complexity is about $2^{107.8}$ 6-round KASUMI encryptions and the memory requirement are 2^{64} 8-bit for counters.

5 Key-recovery attack on the last 7 round KASUMI

In this section, we describe our attacks on the last 7 round of KASUMI. We mount the 5-round zero-correlation linear approximations from round 3 to round 7, and extend one round forward and backward respectively. We assume that the subkeys k_2' and k_4' have the same value at least 8 bit positions among the 16 bits positions, then based on Observation 4, we know there are a least 2^8 input masks and It is easy to know that more than half of the master keys space has this property. In the attack, we also select some special input masks to reduce number of guessed key bits.

In our attack, we guess the subkey and evaluate the linear approximation $(a, a')^T \cdot \big((L_{2,l}, L_{2,r}) \oplus (R_{7,l}, R_{7,r})\big) = 0$, that is

$$(a, a') \cdot \big((R_{1,l}, R_{1,r}) \oplus (L_{8,l}, L_{8,r})\big) \oplus a \cdot \big(FI(L_{1,l} \oplus (k_3 \lll 5), k_6') \oplus L_{1,r} \oplus FI(L_{1,r} \oplus k_5', k_7 \lll 8) \oplus FI(L_{7,l} \oplus (k_1 \lll 5), k_4') \oplus L_{7,r} \oplus FI(L_{7,r} \oplus (k_5 \lll 8), k_3')\big) = 0,$$

where $a[i-1] = 0$, if $k_4'[i] \oplus k_2'[i] = 1$, $i = 1, 2, ...15$ and $a[15] = 0$, if $k_4'[0] \oplus k_2'[0] = 1$ and we let $a' = a \ggg 1 \diamond \neg k_2' \diamond \neg k_4'$.

Then the key-recovery attack on 7-round KASUMI is proceeded with partial-sum technique as follows.

1. Allocate a counter vector $N[L_{1,l}|L_{1,r}|L_{7,l}|L_{7,r}|R_{1,l} \oplus L_{8,l} \oplus L_{1,r} \oplus L_{7,r}|R_{1,r} \oplus L_{8,r}]$ of size 96 where each element is 8-bit length and initialize to zero.

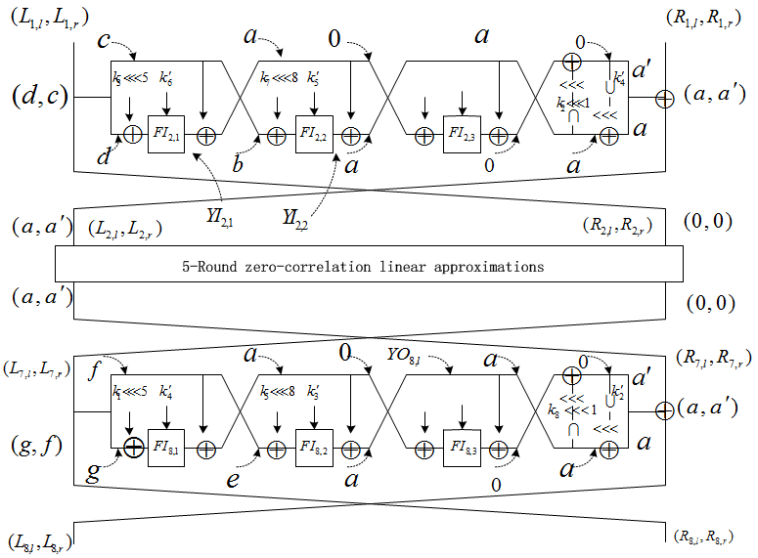

Figure 4: Multidimensional Zero-correlation attack on KASUMI reduced to rounds 2-8

2. Guess all possible values of 16 master key bits k'_4.

3. For N plaintext-ciphertext pairs, extract the 80-bit value

$$i = (L_{1,l}|L_{1,r}|L_{7,l}|L_{7,r}|R_{1,l}|Y^1)$$

where $Y^1 = R_{1,l} \oplus L_{8,l} \oplus L_{1,r} \oplus L_{7,r} \oplus \Big(\big(\neg k'_4 \diamond (R_{1,r} \oplus L_{8,r}) \big) \ggg 1 \Big)$ and increment the counter x_i according to the value of i.

4. Guess all possible values of 48 master key bits k_1, k_5, k_3, partially encrypt $L_{7,l}$ and $L_{7,r}$ to get $YO_{8,l}$. Let $Y^2 = Y^1 \oplus YO_{8,l}$ and add one to the corresponding $i = (L_{1,l}|L_{1,r}|Y^2)$. The time complexity of Step 4 is no more than $N \times 2^{16} \times 2^{48} \times \frac{1}{3} \times \frac{1}{7}$ 7-round encryptions.

5. Guess all possible values of 16 master key bits k_6, partially encrypt $L_{1,l}$ to get $YI_{2,1}$. Let $Y^3 = Y^2 \oplus YI_{2,1}$. Add one to the corresponding $i = (L_{1,r}, Y^3)$. The time complexity of Step 5 is no more than $2^{16} \times 2^{64} \times 2^{48} \times \frac{1}{3} \times \frac{1}{7}$ 7-round encryptions.

6. Guess all possible values of 16 master key bits k_7, partially encrypt $L_{1,r}$ to get $YI_{2,2}$. Let $Y^4 = Y^4 \oplus YI_{2,2}$. Add one to the corresponding $i = (Y^4)$. The time complexity of Step 5 is no more than $2^{16} \times 2^{16} \times 2^{16} \times 2^{48} \times 2^{32} \times \frac{1}{3} \times \frac{1}{7}$ 7-round encryptions.

7. Guess 2^{15} number of k'_2 under weak key condition. k_2 has the same value with k_4 in at least 8-bit positions and we call those bit positions be active bit positions. We let the masks a be 0 or 1 in the first 8 active bit positions, and be 0 in others. there are 2^8 masks.

8. Allocate a counter vector $N[z]$ of size 2^8 where each element is 64-bit length for 8-bit z (z is the concatenation of evaluations of 8 basis zero-correlation masks).

9. For 2^{16} values of Y^4, evaluate 8 basis zero-correlation masks on Y^4 and put the evaluations to the vector z, then add the corresponding $N[z] : N[z] += N[Y^4]$.

10. Compute $T = N2^8 \sum_{z=0}^{2^8-1} (\frac{v[z]}{N} - \frac{1}{2^8})$, if $T \leq \tau$,then the guessed key is a possible key candidate. As there are 48 master key bits that we havent guessed, we do exhaustive search for all keys conforming to this possible key candidate.

In this attack, we set the type-I error probability $\beta_0 = 2^{-2.7}$ and the type-II error probability $\beta_1 = 2^{-10}$. We have $z_{1-\beta_0} \approx 1$, $z_{1-\beta_1} \approx 3.09$, $n = 64$, $l = 2^8 - 1$. The date complex N is about $2^{62.1}$ and the decision threshold $\tau \approx 2^{7.6}$.

There are 2^{111} master key value guessed during the encryption and decryption phase, and $2^{111} \cdot 2^{-10} = 2^{101}$ key candidates survive in the wrong key filtration. Step 10 needs about $2^{128} \cdot 2^{-10} = 2^{118}$ 7-round KASUMI encryption. The complexity of the dominant Step 5, 6, 7 is about $2^{123.61}$, $2^{123.61}$ 7-round KASUMI encryptions and 2^{127} memory accesses. If we assume that one time of memory accesses is equivalent to one FI function operator, then, the total complexity is about $2^{125.2}$ 7-round KASUMI encryptions with $2^{62.1}$ known plaintexts.

6 Conclusion

In this paper, we evaluate the security of KASUMI with respect to the novel technique of multidimensional zero-correlation cryptanalysis. We refine the zero-correlation linear approximations by selecting some special input masks. Besides, we give some observations on the FL function with some special input masks, with which we give the first multidimensional zero-correlation attack on the 6 round and the last 7 round of KASUMI block cipher. The two attacks need $2^{107.8}$ encryptions with $2^{59.4}$ chosen plaintexts and $2^{125.2}$ encryptions with $2^{62.1}$ known plaintexts, respectively.

References

[1] Bogdanov, A., Rijmen, V.: Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers. Designs, Codes and Cryptography, Springer, US, 2012, pp.1-15.

[2] Bogdanov, A., Wang, M.: Zero Correlation Linear Cryptanalysis with Reduced Data Complexity, in: A. Canteaut (Ed.), FSE 2012, in: Lect.Notes Com put. Sci., vol. 7549, Springer, Heidelberg, 2012, pp. 29-48.

[3] Bogdanov, A., Leander, G., Nyberg, K., Wang, M. : Integral and multidimensional linear distinguishers with correlation zero, in: X. Wang,K. Sako (Eds.), AsiaCrypt 2012, in: Lect. Notes Comput. Sci., vol. 7658, Springer, Heidelberg, 2012, pp. 24C262.

[4] Bogdanov, A., Geng, H., Wang, M., Wen, L., Collard, B.: Zero-correlation linear cryptanalysis with FFT and improved attacks on ISO standards Camellia and CLEFIA, in: T. Lange, K. Lauter, P. Lisonek (Eds.), SAC13, in: Lect. Notes Comput. Sci., Springer-Verlag, 2013, in press.

[5] Biham, E., Dunkelman, O., Keller, N.: A Related-Key Rectangle Attack on the Full KASUMI. In: Roy, B. (ed.) ASIACRYPT 2005. LNCS, vol. 3788, pp. 443C461. Springer, Heidelberg (2005).

[6] Blunden, M., Escott, A.: Related Key Attacks on Reduced Round KASUMI. In: Matsui, M. (ed.) FSE 2001. LNCS, vol. 2355, pp. 277C285. Springer, Heidelberg (2002).

[7] Dunkelman, O., Keller, N., Shamir, A.: A Practical-Time Related-Key Attack on the KASUMI Cryptosystem Used in GSM and 3G Telephony. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 393C410. Springer, Heidelberg (2010).

[8] Jia, K., Li, L., Rechberger, C., Chen,J., Wang, X.: Improved Cryptanalysis of the Block Cipher KASUMI. SAC 2013, Lecture Notes in Computer Science, Volume 7707, 2013, pp 222-233.
[9] Kühn, U.: Cryptanalysis of Reduced-Round MISTY. In: Pfitzmann, B. (ed.) EU- ROCRYPT 2001. LNCS, vol. 2045, pp. 325C339. Springer, Heidelberg (2001).
[10] Matsui, M.: New Block Encryption Algorithm MISTY. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 54C68. Springer, Heidelberg (1997).
[11] Sugio, N., Aono, H., Hongo, S., Kaneko, T.: A Study on Integral-Interpolation Attack of MISTY1 and KASUMI. In: Computer Security Symposium 2006, pp.173C178 (2006) (in Japanese).
[12] Sugio, N., Aono, H., Hongo, S., Kaneko, T.: A Study on Higher Order Differential Attack of KASUMI. IEICE Transactions 90-A(1), 14C21 (2007).
[13] Sugio, N., Tanaka, H., Kaneko, T.: A Study on Higher Order Differential Attack of KASUMI. In: 2002 International Symposium on Information Theory and its Applications (2002).
[14] 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, 3G Security, Specification of the 3GPP Confidentiality and Integrity Algorithms; Document 2: KASUMI Specification, V3.1.1 (2001).
[15] Wen, L., Wang, M., Bogdanov, A., Chen,H.: Multidimensional Zero-Correlation Attacks on Lightweight Block Cipher HIGHT: Improved Cryptanalysis of an ISO Standard. Information Processing Letters 114(6), pp. 322-330, Elsevier, 2014.
[16] Wen, L.,Wang, M., Bogdanov, A.: Multidimensional Zero-Correlation Linear Cryptanalysis of E2. Africacrypt'14, Lecture Notes in Computer Science (LNCS), Springer-Verlag, 2014, to appear.