# Improved Impossible Differential Cryptanalysis on SMS4

Gaoli Wang<sup>1,2</sup>

<sup>1</sup>School of Computer Science and Technology
Donghua University
Shanghai, China
wanggaoli@dhu.edu.cn

<sup>2</sup>State Key Laboratory of Information Security
Institute of Software, Chinese Academy of Sciences
Beijing, China

Abstract—This paper presents an improved impossible differential attack on the block cipher SMS4 which is used in WAPI (the Chinese WLAN national standard). Combining with some new observations, this paper can filter out the wrong keys more efficiently, and present an impossible differential attack on 17-round SMS4, which updates the best known impossible differential attacks on reduced SMS4.

Keywords-block cipher; SMS4; cryptanalysis; impossible differential attack

#### I. INTRODUCTION

SMS4 [1] was released as the symmetric-key encryption standard of Wireless Local Area Network (WLAN) by China in February 2006. It is a 32-round block cipher with a 128-bit block size and a 128-bit user key. There have been several attacks on reduced SMS4.

Some research has been devoted to the security of SMS4. The differential fault analysis on SMS4 was given in [2]. Reference [3] presented an integral attack on 13 rounds SMS4. The rectangle attack on 14 rounds and the impossible differential attack on 16 rounds SMS4 were presented in [4]. Reference [5] noticed that there are some flaws in reference [4], and made a more comprehensive analysis and further improved the results. The rectangle attack on 16 rounds and the differential attack on 21 rounds SMS4 were presented in [6]. The rectangle attack on 18 rounds, the differential attack and linear attack on 22 rounds SMS4 were presented in [7]. An improved differential attack on 22 rounds SMS4 was presented in [8]. A linear cryptanalysis on 23 rounds SMS4 was presented in [9].

Impossible differential cryptanalysis [10] is a sieving attack which uses differentials with probability 0 to eliminate the wrong keys and filter out the right key candidate. With some new tricks, reference [11] can filter out the wrong keys more efficiently, and improve the impossible differential attack on the block cipher CLEFIA. In this paper, the attacker presents some new observations and improvements so as to improve the impossible differential attacks on reduced SMS4. The complexities of

our new attack along with some previously known impossible differential attacks against reduced SMS4 are summarized in Table I .

TABLE I. SUMMARY OF THE PREVIOUS IMPOSSIBLE DIFFERENTIAL ATTACKS AND OUR NEW ATTACK

Number	Data	Time	Attack	Source
of Rounds	Complexity	Complexity	Type	
16	2 <sup>105,a</sup>	2 <sup>107,a</sup> Enc	Imp. Diff	Ref. [4]
16	2117.06	2 <sup>132.06</sup> MA	Imp. Diff	Ref. [5]
17	2117	2 <sup>132</sup> MA	Imp. Diff	This
				Paper

Enc - Encryptions, MA - Memory Accesses.

## II. DESCRIPTION OF SMS4

## A. Notations

The following notations are used throughout this paper.

- 1.  $(X_0, X_1, X_2, X_3)$  denotes 128-bit block composed of four 32-bit word.
- 2.  $W \le s$  denotes the bit rotation of the word w by s positions to the left.
- 3.  $e_j$  denotes a word whose all positions except the j-th bit are zero.
- 4.  $e_{i1,...,ij}$  denotes  $e_{i1} \oplus ... \oplus e_{ij} (0 \le i_1, ..., i_j \le 31)$ .
- 5.  $\Delta$  a denotes a  $\oplus$  a<sup> $\prime$ </sup>.

Note that the words and blocks are in an order as follows. The most significant bit of a 32-bit word is the leftmost bit numbered 0, and the least significant bit is the rightmost bit numbered 31. The most significant byte of a 32-bit word is the leftmost byte numbered 0, and the least significant byte is numbered 3. The most significant word of a 128-bit block is the leftmost word numbered 0, and the least significant word is numbered 3.

## B. Data Processing Part of SMS4

SMS4 [1] is a 128-bit block cipher with the key length of 128 bits, and has a total of 32 rounds. It employs an unbalanced Feistel structure with four data lines, and the width of each data line is 32 bits.



<sup>&</sup>lt;sup>a</sup>As noted in Reference [5], these figures are underestimated.

The round numbers are denoted by 1 to 32, and  $RK_i \in \mathbb{Z}_2^{32}$  ( $1 \le i \le 32$ ) are round subkeys produced by the key scheduling part. Let  $P^{i-1} = (X_{i-4}, X_{i-3}, X_{i-2}, X_{i-1})$  denotes the four-word input of the i-th round, and  $P^i = (X_{i-3}, X_{i-2}, X_{i-1}, X_i)$  denotes the four-word output of the i-th round ( $1 \le i \le 32$ ). Then for a 128-bit plaintext  $P = (P_0, P_1, P_2, P_3)$ , the encryption procedure of SMS4 is as follows:

- 1. Input the plaintext  $P^0=(X_{-3}, X_{-2}, X_{-1}, X_0)=P=(P_0, P_1, P_2, P_3)$ .
- 2. For i=1,2,...,32,  $X_i=X_{i-4} \oplus T(X_{i-3} \oplus X_{i-2} \oplus X_{i-1} \oplus RK_i)$ =  $X_{i-4} \oplus L(S(X_{i-3} \oplus X_{i-2} \oplus X_{i-1} \oplus RK_i))$

 $P^i \!\! = \! (X_{i\text{--}4},\, X_{i\text{--}3},\, X_{i\text{--}2},\, X_{i\text{--}1}),$ 

3. Output the ciphertext  $C=R(P^{32})=R(X_{29}, X_{30}, X_{31}, X_{32})=(X_{32}, X_{31}, X_{30}, X_{29}),$ 

where the nonlinear confusion function S applies the same 8×8 S-box (described in [1]) four times in parallel to a 32-bit input, and L is the linear transformation defined as: L(x)=x  $\oplus$  (x<<<2)  $\oplus$  (x<<<10)  $\oplus$  (x<<<18)  $\oplus$  (x<<<24),  $x \in Z_2^{32}$ .

Suppose that all the round subkeys are independent of each other, and omit the description of the key scheduling part.

# III. SOME PROPERTIES OF SMS4

This section describes some properties of SMS4, which are important to our attack.

Reference [4] presented a 12-round impossible differential which resulted in the attack on 16-round SMS4. By exploring more observations, the attacker presents an improved attack on SMS4 utilizing the same impossible differential and some tricks in [11].

Proposition 1. (Impossible Differential of 12-round SMS4 [4]) For 12-round SMS4, given a plaintext pair with difference ( $e_{\Gamma}$ ,  $e_{\Gamma}$ ,  $e_{\Gamma}$ ,0), where  $\Gamma$  is an arbitrary but nonempty subset of the set  $\{0, \ldots, 15\}$ , the output difference can't be equal to  $\{0, e_{\Gamma}, e_{\Gamma}, e_{\Gamma}\}$ . The 12-round impossible differential can be denoted by

$$(e_{\Gamma}, e_{\Gamma}, e_{\Gamma}, 0) \xrightarrow{can't} (0, e_{\Gamma}, e_{\Gamma}, e_{\Gamma}).$$

The correctness of Proposition 1 can be referred to [4]. *Proposition 2*. For the  $8\times8$  S-Box, there exists 127 possible input differences for any nonzero output difference, of which 1 input difference occurs with probability  $2^{-6}$ , and each of the other 126 input differences occurs with probability  $2^{-7}$ .

This proposition can be verified by a simple computer program.

Let SET( $\Delta$  In,  $\Delta$  Out)={  $a \in \mathbb{Z}_2^8 \mid S(a) \oplus S(a \oplus \Delta In) = \Delta$  Out}. Illuminated by Proposition 2 in [11], we present the following proposition.

Proposition 3. For the function T, let  $In=(In_0,In_1,In_2,In_3)$  and

 $In' = (In'_0, In'_1, In'_2, In'_3)$  be two 32-bit inputs, and  $\Delta \operatorname{Out} = (\Delta \operatorname{Out}_0, \Delta \operatorname{Out}_1, \Delta \operatorname{Out}_2, \Delta \operatorname{Out}_3)$  be the difference of the corresponding output. For  $0 \le i \le 3$ , if the SET( $\operatorname{In}_i \oplus In'_i$ ,  $\Delta \operatorname{Out}_i$ ) is not null, then  $2^4$  32-bit round subkey RK=(RK<sub>0</sub>, RK<sub>1</sub>, RK<sub>2</sub>, RK<sub>3</sub>) involved in T can be deduced with about one T-computation.

Proof: Because the linear transformation L is invertible, the input difference  $\Delta \operatorname{Out}^{-1}$  of L can be easily computed, i.e.,

$$\Delta \operatorname{Out}^{-1} = \operatorname{L}^{-1}(\Delta \operatorname{Out}) = (\Delta \operatorname{Out}_0^{-1}, \Delta \operatorname{Out}_1^{-1}, \Delta \operatorname{Out}_2^{-1},$$

 $\Delta Out_3^{-1}$  ), where  $\Delta Out_0^{-1}$ ,  $\Delta Out_1^{-1}$ ,  $\Delta Out_2^{-1}$  and  $\Delta Out_3^{-1}$  are four 8-bit output XOR of the four S-boxes respectively. Therefore, we get the input XOR and the corresponding output XOR for each of the four S-boxes. If the SET(In<sub>i</sub>  $\oplus$   $In_i^{'}$ ,  $\Delta Out_i^{-1}$ ) is not null for  $0 \le i \le 3$ , it is easy to get the inputs to each S-box by searching the XOR distribution table of the S-box. By Proposition 2, it is easy to know that there are 1 input difference of the S-box with probability 126/127, and 2 input differences with probability 1/127.

For  $0 \le i \le 3$ , denote the two inputs of each S-box as Input<sub>i</sub> and  $Input'_i$ . Then two 8-bit round subkeys  $RK_i$  and  $RK'_i$  can be obtained from the following equations:

$$RK_i = Input_i \oplus In_i$$
,  $RK'_i = Input_i \oplus In'_i$ .

Thus the  $2^4$  32-bit round subkey RK can be deduced. Obviously the time complexity is about one *T*-computation.

# IV. IMPOSSIBLE DIFFERENTIAL ATTACK ON 17-ROUND SMS4

The paper [4] presented a 12-round impossible differential and showed an impossible differential attack on 16-round SMS4. The data and time complexities claimed in [4] are 2<sup>105</sup> chosen plaintexts and 2<sup>107</sup> encryptions respectively. Paper [5] noticed that the attacks in [4] have some flaws and their complexity analysis is inaccurate. A more comprehensive analysis of 16-round impossible differential attack was given in [5], and the data and time complexities are 2<sup>117,06</sup> chosen plaintexts and 2<sup>132,06</sup> memory accesses respectively.

This section describes an impossible differential attack on 17-round SMS4 from round 1 to round 17 with two additional rounds before the differential and three additional rounds after the differential. This attack uses the 12-round impossible differential presented in [4] from round 3 to round 14. Choose  $\Gamma \subseteq \{0,...,15\}$  to reduce the data and time complexities of the attack. See Table II for the following attack.

TABLE II. IMPOSSIBLE DIFFERENTIAL ATTACK ON 17-ROUND SMS4

Round(i)	$\Delta X_{i-3}$	$\Delta X_{i-2}$	$\Delta X_{i-1}$	$\Delta X_{i}$
0	Δ X3	Δ X .2	$e_{\Gamma}$	$e_{\Gamma}$
1	Δ X .2	$e_{\Gamma}$	$e_{\Gamma}$	$e_{\Gamma}$
2	$e_{\Gamma}$	$e_{\Gamma}$	$e_{\Gamma}$	0
•••				
14	0	$e_{\Gamma}$	$e_{\Gamma}$	$e_{\Gamma}$
15	$e_{\Gamma}$	$e_{\Gamma}$	$e_{\Gamma}$	$\Delta X_{15}$
16	$e_{\Gamma}$	$e_{\Gamma}$	$\Delta X_{15}$	$\Delta x_{16}$
17 <sup>b</sup>	$e_{\Gamma}$	$\Delta X_{15}$	$\Delta X_{16}$	$\Delta x_{17}$

<sup>b</sup>For simplicity, we omit the R transformation in the last step of the SMS4 algorithm. The plaintext difference is  $(\Delta X_{-3}, \Delta X_{-2}, e_{\Gamma}, e_{\Gamma}) \in \sum_{l}(I)$ . The ciphertext difference is  $(e_{\Gamma}, \Delta X_{15}, \Delta X_{16}, \Delta X_{17}) \in \sum_{l}(I)$ .

In the input of round 2, for every T, there are  $127^2$  input differences that may lead to  $e_{\Gamma}$ , and they can be generated by  $127^6$  input differences of round 1, which is denoted by the set  $\sum_1(\Gamma)$  for each T. Similarly, there are  $127^2$  output differences of round 15, and  $127^6$  possible output differences of round 16, and  $127^{10}$  possible output differences of round 17, which is denoted by the set  $\sum_2(\Gamma)$  for each T.

Sieving the candidate pairs: A structure composed of  $2^{96}$  plaintexts is defined as follows:  $Struc = \{P = (P_0, P_1, (\widetilde{P}_2, b), (\widetilde{P}_3, c)) \mid b, c \text{ are fixed, } P_0, P_1 \in \mathbb{Z}_2^{32} \text{ and } \widetilde{P}_2, \widetilde{P}_3 \in \mathbb{Z}_2^{16} \text{ are non-zero}\}.$  Each structure suggested  $(2^{96})^2/2=2^{191}$  pairs.  $(\Delta \widetilde{P}_2, 0)=(\Delta \widetilde{P}_3, 0)$  holds with probability  $2^{-16}$ .  $\Sigma_1(\Gamma)$  is composed of  $127^6\approx 2^{42}$  possible input differences for each  $(\Delta \widetilde{P}_2, 0)=(\Delta \widetilde{P}_3, 0)=e_{\Gamma}$ . Therefore, the probability of a plaintext pair (P, P') with  $P \oplus P' \in \Sigma_1(\Gamma)$  is  $2^{42}/2^{64} \times 2^{-16} = 2^{-38}$ , and  $2^{191} \times 2^{-38} = 2^{153}$  pairs satisfy the target plaintext difference.

Note that once the plaintext pair is fixed, T is also fixed, similar to  $\sum_{1}(\Gamma)$ ,  $\sum_{2}(\Gamma)$  is composed of  $127^{10}\approx 2^{70}$  possible output differences for each T. Therefore, the probability of a ciphertext pair (C, C') with  $C \oplus C' \in \sum_{2}(\Gamma)$  is  $2^{70}/2^{128} = 2^{-58}$ , and the number of pairs for a given structure with the target ciphertext difference is  $2^{153}\times 2^{-58}=2^{95}$ .

In this attack, about  $2^{116}$  such plaintext pairs are necessary to sieve the right key. So we choose  $2^{21}$  such structures. The attacker needs to explore a fast algorithm to obtain the  $2^{116}$  pairs. Follow the notations in [4] and [5], and the definitions of  $\Theta(e_{\Gamma}), \gamma(e_{\Gamma}, m \in \Theta(e_{\Gamma}))$  and  $\Pi(e_{\Gamma}, m \in \Theta(e_{\Gamma}), n \in \gamma(e_{\Gamma}, m))$  can refer to [4]. As the definition in [5],  $\widetilde{e}_{\tau}$  is the least significant two bytes of  $e_{\Gamma}$  for each  $\Gamma$ .

The ciphertext is denoted by  $C = ((\widetilde{C}_0, \hat{C}_0), C_1, C_2, C_3) \, .$ 

1. Insert every plaintext-ciphertext pair

$$(P = (P_0, P_1, (\widetilde{P}_2, b), (\widetilde{P}_3, c)), C = ((\widetilde{C}_0, \hat{C}_0), C_1, C_2, C_3))$$
 indexed by  $\widetilde{P}_2 \parallel \widetilde{P}_3 \parallel \widetilde{C}_0 \parallel \hat{C}_0$  into a hash table.

2. For every non-empty bin satisfying  $\widetilde{P}_2 < \widetilde{P}_2^{/}$  , do the following:

(a) Search the corresponding bin with

$$\widetilde{P}_2 \oplus \widetilde{P}_2^{\prime} = \widetilde{P}_3 \oplus \widetilde{P}_3^{\prime} = \widetilde{C}_0 \oplus \widetilde{C}_0^{\prime}, \hat{C}_0 = \hat{C}_0^{\prime}$$

Denote  $(\widetilde{P}_2 \oplus \widetilde{P}_2^{\prime}, 0)$  as  $e_{\Gamma}$ .

(b) For all possible combinations of entries, pick the plaintext pairs for which the following conditions are satisfied:

$$P_1 \oplus P_1 \in \Theta(e_{\Gamma}), P_0 \oplus P_0 \in \gamma(e_{\Gamma}, P_1 \oplus P_1),$$

$$C_1 \oplus C_1 \in \Theta(e_{\Gamma}), C_2 \oplus C_2 \in \gamma(e_{\Gamma}, C_1 \oplus C_1),$$

$$C_3 \oplus C_3 \in \Pi(e_{\Gamma}, C_1 \oplus C_1, C_2 \oplus C_2).$$

If one of them fails, do not check the remaining conditions.

3. Output (P, P').

The time complexity to sieve the candidate pairs is as follows. Step 1 takes  $2^{96}$  memory accesses for each structure. There are  $2^{96}$  plaintext-ciphertext pairs in a structure, the expected number of entries in each of the  $2^{64}$  bins is  $2^{32}$ . Performing one-time Step 2(b) would require  $2^{32}$  memory accesses.  $e_{\Gamma}$  can take about  $2^{16}$  different values, as a result, the total number of memory accesses of  $2^{21}$  structures is  $2^{21} \times 2^{64}/2 \times 2^{16} \times 2^{32} = 2^{132}$ .

Attack Procedure to recover the subkeys  $(RK_1, RK_{2,0}, RK_{2,1}, RK_{17}, RK_{16}, RK_{15,0}, RK_{15,1})$ :

For every selected pair (P, P'), the wrong subkeys  $(RK_1, RK_{2,0}, RK_{2,1}, RK_{17}, RK_{16}, RK_{15,0}, RK_{15,1})$  resulting in the impossible differential are computed as follows:

Step 1. Compute the subkey RK<sub>1</sub> which produces the partial encryption of the pair to match  $\Delta P^1 = (\Delta X_{-2}, \Delta X_{-1}, \Delta X_0,$ 

 $\Delta X_1$ )=(  $\Delta X_{-2}$ ,  $e_{\Gamma}$ ,  $e_{\Gamma}$ ,  $e_{\Gamma}$ ). From the algorithm of SMS4, the attacker only needs to compute RK<sub>1</sub> that cause  $\Delta X_1 = e_{\Gamma}$ .

(a) Since  $X_1 = X_{-3} \oplus T(X_{-2} \oplus X_{-1} \oplus X_0 \oplus RK_1)$ , then

 $\Delta$  T(  $X_{-2} \oplus X_{-1} \oplus X_0 \oplus RK_1$ )=(  $\Delta$   $X_1 \oplus \Delta X_{-3}$ )=

 $\mathbf{e}_{\Gamma} \oplus \Delta X_{-3}$  . The inputs of T are  $X_{-2} \oplus X_{-1} \oplus \mathbf{X}_{0}$ 

and  $X_{-2}^{'} \oplus X_{-1}^{'} \oplus X_{0}^{'}$ , so  $2^{4}$  32-bit RK<sub>1</sub> can be calculated with one *T*-computation by Proposition 3.

(b) Encrypt the plaintext pair (P, P') by RK<sub>1</sub> to get the outputs of round 1, and denote them by  $(P^1=(X_{-2}, X_{-1}, X_0, X_1), P^{1/}=(X_{-2}', X_{-1}', X_0', X_1'))$ .

Step 2. Because the last two bytes of  $e_{\Gamma}$  are zero, only 16-bit subkey (RK<sub>2,0</sub>, RK<sub>2,1</sub>) is related to the condition  $\Delta X_2 = 0$ . From  $X_2 = X_{-2} \oplus L(S(X_{-1} \oplus X_0 \oplus X_1 \oplus RK_2))$ , we get

 $S(X_{-1} \oplus X_0 \oplus X_1 \oplus RK_2) \oplus S(X'_{-1} \oplus X'_0 \oplus X'_1 \oplus RK_2)$  $=L^{-1}(\Delta X_2 \oplus \Delta X_{-2}) = L^{-1}(\Delta X_{-2}).$ 

As the two inputs for S-box are  $X_{-1} \oplus X_0 \oplus X_1$  and

 $X_{-1}' \oplus X_0' \oplus X_1'$ ,  $2^2$  16-bit  $RK_{2,0}$ ,  $RK_{2,1}$  can be computed with less than one *T*-computation by Proposition 3.

Step 3. Compute the subkey RK<sub>17</sub> which causes the partial decryption of the ciphertext pair (C=P<sup>17</sup>=(X<sub>14</sub>, X<sub>15</sub>, X<sub>16</sub>, X<sub>17</sub>), C'=P<sup>17/</sup>=(X<sup>1</sup><sub>14</sub>, X<sup>1</sup><sub>15</sub>, X<sup>1</sup><sub>16</sub>, X<sup>1</sup><sub>17</sub>)) to match  $\Delta$  P<sup>16</sup>=(e<sub>Γ</sub>, e<sub>Γ</sub>,  $\Delta$  X<sub>15</sub>,  $\Delta$  X<sub>16</sub>).

(a) According to  $X_{17} = X_{13} \oplus T(X_{14} \oplus X_{15} \oplus X_{16} \oplus RK_{17})$ , we get that  $\Delta T(X_{14} \oplus X_{15} \oplus X_{16} \oplus RK_{17}) = (\Delta X_{17} \oplus X_{18} \oplus X_{$ 

 $\Delta X_{13}$ ) =  $e_{\Gamma} \oplus \Delta X_{17}$ . Thus  $2^4$  RK<sub>17</sub> can be derived with one *T*-computation by Proposition 3.

(b) Decrypt the ciphertext pair (C, C') by RK<sub>17</sub> to get the outputs of round 16, and denote them by ( $P^{16}=(X_{13}, X_{14}, X_{15}, X_{16}), P^{16'}=(X_{13}, X_{14}', X_{15}', X_{16}')$ ).

Step 4(a) For  $\Delta$  P<sup>16</sup>= (e<sub> $\Gamma$ </sub>, e<sub> $\Gamma$ </sub>,  $\Delta$  X<sub>15</sub>,  $\Delta$  X<sub>16</sub>), since X<sub>16</sub>=X<sub>12</sub>  $\oplus$  T(X<sub>13</sub>  $\oplus$  X<sub>14</sub>  $\oplus$  X<sub>15</sub>  $\oplus$  RK<sub>16</sub>), then  $\Delta$  T(X<sub>13</sub>  $\oplus$  X<sub>14</sub>  $\oplus$  X<sub>15</sub>  $\oplus$  RK<sub>16</sub>) = ( $\Delta$  X<sub>16</sub>  $\oplus$   $\Delta$  X<sub>12</sub>) = e<sub> $\Gamma$ </sub>  $\oplus$   $\Delta$  X<sub>16</sub>. Thus 2<sup>4</sup> RK<sub>16</sub> can be derived with one *T*-computation by Proposition 3.

(b) Decrypt  $(P^{16}, P^{16/})$  by  $RK_{16}$  to get the outputs of round 15, and denote them by  $(P^{15}=(X_{12}, X_{13}, X_{14}, X_{15}), P^{15/}=(X_{12}, X_{13}, X_{14}, X_{15}))$ .

Step 5. For  $\Delta$  P<sup>15</sup>= (e<sub> $\Gamma$ </sub>, e<sub> $\Gamma$ </sub>, e<sub> $\Gamma$ </sub>,  $\Delta$  X<sub>15</sub>), because X<sub>15</sub>=X<sub>11</sub>  $\oplus$  L(S(X<sub>12</sub>  $\oplus$  X<sub>13</sub>  $\oplus$  X<sub>14</sub>  $\oplus$  RK<sub>15</sub>), then S(X<sub>12</sub>  $\oplus$  X<sub>13</sub>  $\oplus$  X<sub>14</sub>  $\oplus$  RK<sub>15</sub>)  $\oplus$  S(X'<sub>12</sub>  $\oplus$  X'<sub>13</sub>  $\oplus$  X'<sub>14</sub>  $\oplus$  RK<sub>15</sub>)=L<sup>-1</sup>( $\Delta$  X<sub>15</sub>  $\oplus$   $\Delta$  X<sub>11</sub>) = L<sup>-1</sup>( $\Delta$  X<sub>15</sub>).

Because  $\Delta X_{12} \oplus \Delta X_{13} \oplus \Delta X_{14} = e_{\Gamma} \oplus e_{\Gamma} \oplus e_{\Gamma} = e_{\Gamma}$ , and the last two bytes of  $e_{\Gamma}$  are zero, it only needs to derive 16-bit subkey (RK<sub>15,0</sub>, RK<sub>15,1</sub>). Similar to Step 2,  $2^2$  16-bit (RK<sub>15,0</sub>, RK<sub>15,1</sub>) can be computed with less than one *T*-computation by Proposition 3.

So far, we can filter out  $2^{16}$  wrong subkeys (RK<sub>1</sub>, RK<sub>2,0</sub>, RK<sub>2,1</sub>, RK<sub>17</sub>, RK<sub>16</sub>, RK<sub>15,0</sub>, RK<sub>15,1</sub>) which support the impossible differential with about 5 *T*-computations. Thus, for every pair, a wrong subkey (RK<sub>1</sub>, RK<sub>2,0</sub>, RK<sub>2,1</sub>, RK<sub>17</sub>, RK<sub>16</sub>, RK<sub>15,0</sub>, RK<sub>15,1</sub>) survives with probability  $1-2^{16}/2^{128}=1-2^{-112}$ . After analyzing  $2^{116}$  pairs, the number of the remaining subkeys is  $2^{128} \times (1-2^{-112})^{2^{116}} \approx 2^{105}$ . Then we can get the right subkey (RK<sub>1</sub>, RK<sub>2,0</sub>, RK<sub>2,1</sub>, RK<sub>17</sub>, RK<sub>16</sub>, RK<sub>15,0</sub>, RK<sub>15,1</sub>) by  $2^{105}$  computations.

Complexity Evaluation

The number of chosen plaintexts is about  $2^{21} \times 2^{96} = 2^{117}$ . The time complexity is about  $2^{116} \times 5 \approx 2^{118.4}$  *T*-computations, in addition to  $2^{132}$  memory accesses in the preliminary

elimination process to sieve the candidate pairs. Our impossible differential attack on 17-round SMS4 is faster than exhaustively search.

#### V. CONCLUSION

This paper presents a chosen-plaintext attack on 17-round SMS4. The attacker explores some techniques such as differential cryptanalysis to reduce the computational complexity, and break more rounds. In future work, we plan to take the key scheduling algorithm of SMS4 into consideration to analyze the security of SMS4 against impossible differential attack.

#### ACKNOWLEDGMENT

This work was supported in part by "Chen Guang" project (supported by Shanghai Municipal Education Commission and Shanghai Education Development Foundation); the Fundamental Research Funds for the Central Universities.

#### REFERENCES

- Office of State Commercial Cryptography Administration.: P.R. China, The SMS4 Block Cipher (in Chinese), Archive available at: http://www.oscca.gov.cn/UpFile/200621016423197990.pdf
- [2] L. Zhang and W. L. Wu, "Differential Fault Analysis on SMS4. Chinese Journal of Computers", 2006, Vol. 29, No. 9, pp. 1596-1602.
- [3] F. Liu, W. Ji, L. Hu, J. Ding, S. Lv, A. Pyshkin, and R. Weinmann, "Analysis of the SMS4 Block Cipher", Proc. ACISP 2007, LNCS 4586, pp. 158-170, 2007.
- [4] J. Lu, "Attacking Reduced-Round Versions of the SMS4 Block Cipher in the Chinese WAPI Standard", Proc. ICICS 2007, LNCS 4861, pp. 306-318, 2007.
- [5] D. Toz and O. Dunkelman, "Analysis of Two Attacks on Reduced-Round Versions of the SMS4", Proc. ICICS 2008, LNCS 5308, pp. 141-156, 2008.
- [6] L. Zhang, W. Zhang, and W. Wu, "Cryptanalysis of Reduced-Round SMS4 Block Cipher", Proc. ACISP 2008, LNCS 5107, pp. 216-229, 2008
- [7] T. Kim, J. Kim, S. Hong, and J. Sung, "Linear and Differential Cryptanalysis of Reduced SMS4 Block Cipher", Cryptology ePrint Archive: Report 2008/281.
- [8] W. Zhang, W. Wu, D. Feng, and B. Su, "Some New Observations on the SMS4 Block Cipher in the Chinese WAPI Standard", Proc. ISPEC 2009, LNCS 5451, pp. 324-335, 2009.
- [9] J. Etrog and M. J. B. Robshaw, "The Cryptanalysis of Reduced-Round SMS4", Proc. SAC 2008, LNCS 5381, pp. 51-65, 2009.
- [10] E. Biham, A. Biryukov, and A. Shamir, "Cryptanalysis of Skipjack Reduced to 31 Rounds", Proc. EUROCRYPT 1999, LNCS 1592, pp. 12-23, 1999.
- [11] W. Wang and X. Wang, "Improved Impossible Differential Cryptanalysis of CLEFIA", Cryptology ePrint Archive: Report 2007/466.