Unified Impossible Differential Cryptanalysis on Block Cipher Structures

Yiyuan Luo[1,2], Zhongming Wu[1] and Xuejia Lai[1]
[1]Department of Computer Science and Engineering
Shanghai Jiaotong University, China
luoyiyuan@sjtu.edu.cn
[2]Department of Electrical and Computer Engineering, University of Waterloo.

Abstract

In this paper, we propose a systematic search method for finding the impossible differential characteristic for block cipher structures, better than the $\mathcal{U}$-method introduced by Kim *et al* [6]. This method is referred as unified impossible differential (UID) cryptanalysis. We give practical UID cryptanalysis on some popular block ciphers and give the detailed impossible differential characteristics. On the generalized CAST-256 and generalized MARS block cipher structure, our results are better than the $\mathcal{U}$-method. On the Four-Cell, FOX64, our results are the same as previous best manual works. Thus UID method can be used as a tool for examining the security of a block cipher structure against impossible differential cryptanalysis.

1 Introduction

Impossible differential cryptanalysis was first proposed by Biham, *et al.* to attack Skipjack block cipher [1]. It is known as one of the most powerful attacks on block ciphers. It has drawn wide attention in block cipher design and analysis and many good results are achieved [1, 2, 14], just list a few.

Compared with ordinary differential cryptanalysis, impossible differential cryptanalysis considers the differences that are impossible at some intermediate state of the block cipher. Impossible differential characteristics is the differential characteristics with probability of 0. Specifically, when a pair of plaintexts satisfy the input difference of the characteristics, it is impossible for the intermediate states decrypted from ciphertexts by the right subkey to satisfy the output difference of characteristics. Thus the adversary can remove the wrong candidate subkey, and recover the right subkey.

Impossible differential attack is composed of two steps: retrieving the longest characteristics and recovering the subkeys. Retrieving the characteristics uses the idea of miss-in-the-middle, namely to find two differential characteristics with probability 1 from encryption and decryption, and connect them together when there are some inconsistencies. In the key recovering step, every candidate subkey is tested if there exists a pair of texts satisfying the input and output difference of the characteristics by using the key to decrypt.

The key step of impossible differential cryptanalysis is to retrieve the longest impossible differential characteristics. Impossible differential characteristics are in the form:

$$\{x_1, \cdots, x_n\}_n \nrightarrow_r \{y_1, \cdots, y_n\}_n$$

which means that when input difference is $\{x_1, \cdots, x_n\}_n$, the output difference after r rounds cannot be $\{y_1, \cdots, y_n\}_n$.

Suppose the block cipher has m rounds. Firstly, the adversaries choose several pairs of plaintexts which satisfy the input of the characteristic, then guess the last $m-r$ round subkeys and decrypt the corresponding ciphertexts to r-th round, and verify if the decrypted texts meet the output of the characteristic. Since the impossible differential characteristic holds with probability 1, one can conclude that the last $m-r$ round subkey is wrong if the decrypted texts meet the differences of characteristics. After selecting enough pairs of texts, only the right subkey remains.

Usually, the characteristics are retrieved manually by observing the structure of the block ciphers. In [6], Kim first introduced the $\mathcal{U}$-method to find longest impossible differential characteristics of various block ciphers. Although using $\mathcal{U}$-method, we can find the impossible differential characteristics automatically, there are some limitations of $\mathcal{U}$-method when retrieving the longest impossible differential characteristics:

- The characteristic matrix of the block cipher must have 1-Property[6]. So this method only can be applied some special block ciphers.
- Some information is lost during the calculating the impossible differential characteristics. so it might miss some characteristics.

In this paper, we present a unify impossible differential(UID) cryptanalysis to find the longest impossible differential characteristic which is automatically searched by computer. UID cryptanalysis has more flexible representation of characteristic matrices and more accurate middle status, thus UID cryptanalysis performs better than the original $\mathcal{U}$-method. We give some practical cryptanalysis on popular block cipher structures using UID. With UID cryptanalysis, we find the longer impossible differential characteristic on generalized CAST-256 [8] and generalized MARS [8] block cipher structures than the $\mathcal{U}$-method. For the block ciphers Four-Cell [4], FOX64 [5], our results are the same as previous best results obtained by case-by-case treatment. The detailed impossible differential characteristics are listed in Table 3. Our practical results show that UID cryptanalysis can be used as a tool to test the security of a block cipher structure against impossible differential cryptanalysis.

The rest of the paper is organized as follows. A detailed description of UID cryptanalysis is presented in Section 2. In Section 3, we give practical cryptanalysis of some popular block ciphers. In Section 4, a comparison with the $\mathcal{U}$-method is discussed. Section 5 provides conclusions.

2 Unified Impossible Differential Cryptanalysis

2.1 UID-identity

Impossible differential cryptanalysis is a chosen plaintext/ciphertext attack. The attacker knows the chosen difference variables during calculating the characteristics. In the impossible differential cryptanalysis, the plaintext of a block cipher can be divided into subblocks. Accordingly the differences of the input, output and internal status are treated as subblocks.

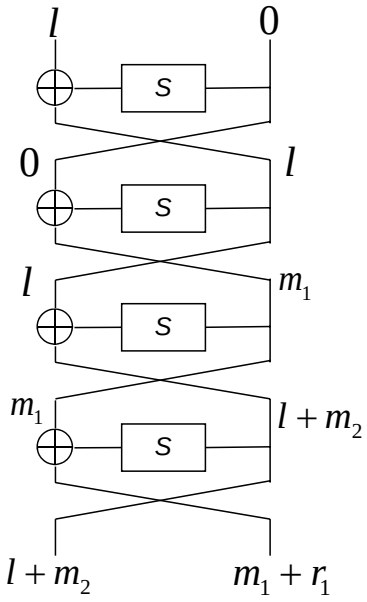
Many block ciphers use a small nonlinear bijection transform as the subblock (S-box) to implement confusion. There are some criteria established for S-box's properties [9]. Most block ciphers' S-boxes are designed under these criteria, so it becomes more and more difficult to find weaknesses in S-box. Impossible differential cryptanalysis tries to ignore weaknesses of the S-box. We usually discard the structure information of S-box and treat it as an ideal black box. Given an input difference, the output difference is uniformly distributed and cannot be determined. However, any permutation has the property that a non-zero difference input has a non-zero output difference.

We represents each subblock as a triple group, called a UID-identity.

Definition 1 *A UID-identity is a triple group denoted as $< L, M, R >$, where L, M and R are three sets of variables. We denote L, M or R as $\emptyset$ when it is an empty set for simplicity. The set L consists of known difference variables, that is, we know the exact difference of the variable in L. The set M consists of unknown non-zero difference variables, that is, the difference in every variable in M can be any non-zero value. The R is the set of unknown difference variables, i.e., the difference of the variable in R can be any value including zero.*

In the UID cryptanalysis, we denote those non-zero unknown difference variables in the M component, those completely unknown difference variables are denoted in the R component.

In Fig.1, we give an example to demonstrate how to use UID-identity to denote the intermediate difference of the Feistel structure. There are two subblocks (the left and the right) in the Feistel structure. Assume the difference



Round	UID-identity	
0	$< \{l\}, \emptyset, \emptyset >$	$< \emptyset, \emptyset, \emptyset >$
1	$< \emptyset, \emptyset, \emptyset >$	$< \{l\}, \emptyset, \emptyset >$
2	$< \{l\}, \emptyset, \emptyset >$	$< \emptyset, \{m_1\}, \emptyset >$
3	$< \emptyset, \{m_1\}, \emptyset >$	$< \{l\}, \{m_2\}, \emptyset >$
4	$< \{l\}, \{m_2\}, \emptyset >$	$< \emptyset, \{m_1\}, \{r_1\}, >$

Figure 1: Intermediate difference state of Feistel structure and their UID-identity. l, m_i and r are known difference, unknown non-zero difference and unknown difference respectively. Here we assume S is a bijection.

of the inputs at the first round is $(l, 0)$ where l is a known non-zero variable, that is the difference of the left is l and the difference of the right is 0, then the UID-identity of the left subblock will be $< \{l\}, \emptyset, \emptyset >$ and the UID-identity of the right subblock will be $< \emptyset, \emptyset, \emptyset >$.

1. After the first round, the difference becomes $(0, l)$, then the UID-identity of the left subblock becomes $< \emptyset, \emptyset, \emptyset >$ and the right subblock becomes $< \{l\}, \emptyset, \emptyset >$.

2. After the second round, the difference becomes (l, m_1) where m_1 is a new non-zero variable, since two inputs with non-zero difference l to a bijection S will result in two outputs with non-zero difference m_1. Then the UID-identity of the left becomes $< \{l\}, \emptyset, \emptyset >$ and the right becomes $< \emptyset, \{m_1\}, \emptyset >$.

3. After the third round, the difference becomes $(m_1, l + m_2)$ where m_2 is a new non-zero variable. Then the UID-identity of the left part becomes $< \emptyset, \{m_1\}, \emptyset >$ and the right part will be $< \{l\}, \{m_2\}, \emptyset >$ since we know the difference of the right is $l + m_2$ where l is a known and m_2 is a non-zero unknown variable.

4. After the fourth round, the difference becomes $(l + m_2, m_1 + r_1)$ where r_1 is a unknown variable. r_1 can be zero since $l + m_2$ can be zero. Then the UID-identity of the left part becomes $< \{l\}, \{m_2\}, \emptyset >$ and right part becomes $< \emptyset, \{m_1\}, \{r_1\} >$

Any intermediate difference in the block cipher can be represented as a UID-identity. The value of the difference is calculated by sum (xor) the variables of three components. For example, in Fig.1, the UID-identity of the left subblock after the fourth round is $< \{l\}, \{m_2\}, \emptyset >$, then the difference here will be $l + m_2$.

In the following, we will define some operations on UID-identity.

Definition 2 *Let S_1 and S_2 are two sets, the symmetric difference of S_1 and S_2 is:*

$$S_1 \boxplus S_2 = (S_1 \setminus S_2) \cup (S_2 \setminus S_1)$$

Definition 3 *The addition of two UID identities $u =< L_1, M_1, R_1 >$ and $v =< L_2, M_2, R_2 >$ is defined as:*

$$u + v =< L_1 \boxplus L_2, M_1 \boxplus M_2, R_1 \boxplus R_2 >$$

Note that the addition on UID-identities is commutative.

Since the difference is calculated by xor the variables of three components in the UID-identity, if a variable occurs even times in the three components, it will not affect the addition of the differences.

Besides the addition operation, we define four kinds of transformations over UID-identity, which are listed in Table 1.

Table 1: Four Transformations over UID-identity

Trans.	Input	Output	Note
$\mathbb{0}$	$< L, M, R >$	$< \emptyset, \emptyset, \emptyset >$	Zero trans.
$\mathbb{1}$	$< L, M, R >$	$< L, M, R >$	Identical trans.
φ	$< \emptyset, \emptyset, \emptyset >$	$< \emptyset, \emptyset, \emptyset >$	Nonlinear trans.
	$< L, \emptyset, \emptyset >$	$< \emptyset, \{m_{new}\}, \emptyset >$	
	$< \emptyset, M, \emptyset >, \lvert M \rvert = 1$	$< \emptyset, \{m_{new}\}, \emptyset >$	
	$< L, M, R >$,otherwise	$< \emptyset, \emptyset, \{r_{new}\} >$	
ρ	$< L, M, R >$	$< \{l_{new}\}, \{m_{new}\}, \{r_{new}\} >$	Linear trans.

In the Fig.1, the bijection S is a nonlinear transform. If the UID-identity of the input is $< \{l\}, \emptyset, \emptyset >$, the UID-identity of the output of S will be $< \emptyset, \{m_{new}\}, \emptyset >$ where m_{new} represents a new variable not previous used. Since the UID-identity represents the difference between texts, two pairs of texts with same difference may have a different output difference, thus a new variable is needed.

In fact, most block cipher structures can be described as a composition of different transforms. In the next subsection we will show how to convert a block cipher in terms of the transformations defined above.

2.2 Matrice Representation of Block Ciphers

Most block ciphers are iterated by round functions. In UID cryptanalysis, we denote the round function in matrix form, then calculate the difference by multiplying those characteristics matrices.

Suppose there are n subblocks in the input and output of the round function, which are denoted by $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ respectively.

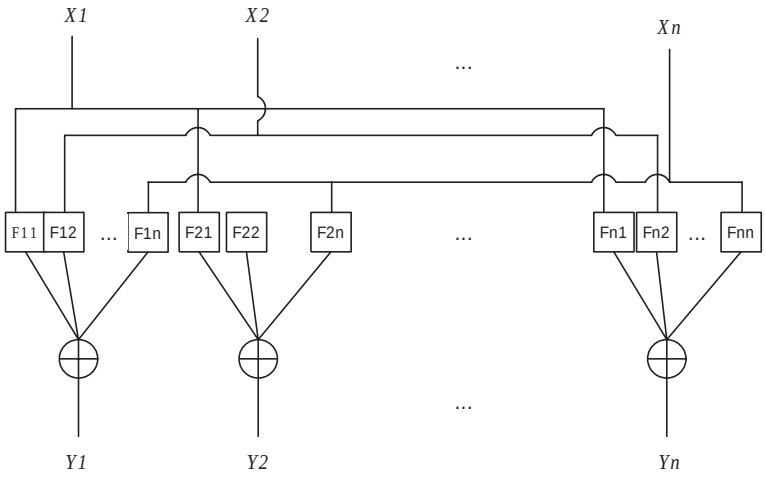


Figure 2: UID structure. The function $F_{i,j}$ are one of four UID Transformations

Definition 4 (UID-Structure) *A round function can be transformed to a UID-Structure if the output* $(Y_1, \ldots, Y_n)$ *can be expressed in terms of* $(X_1, \ldots, X_n)$ *in the following form:*

$$\begin{aligned} Y_1 &= F_{11}(X_1) \oplus F_{12}(X_2) \oplus \ldots \oplus F_{1n}(X_n) \\ Y_2 &= F_{21}(X_1) \oplus F_{22}(X_2) \oplus \ldots \oplus F_{2n}(X_n) \\ &\vdots \quad\quad \vdots \\ Y_n &= F_{n1}(X_1) \oplus F_{n2}(X_2) \oplus \ldots \oplus F_{nn}(X_n) \end{aligned}$$

where $F_{11}, \ldots, F_{nn}$ *are transformations defined in Table 1.*

The computational graph of a UID structure is shown in Fig.2. If a round function can be transformed to a UID structure, we can find the impossible differential characteristics with UID method. In the following parts, we assume the block cipher has already been transformed into the UID structure.

Definition 5 *Assume a round function has n data subblocks, and denote the input and output of the a UID structure as* $(X_1, \cdots, X_n)$ *and* $(Y_1, \cdots, Y_n)$*. The encryption characteristic matrix* $\mathcal{E}$ *is an* $n \times n$ *matrix defined as:*

$$\mathcal{E} = \begin{pmatrix} F_{11} & F_{21} & \cdots & F_{n1} \\ F_{12} & F_{22} & \cdots & F_{n2} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ F_{1n} & F_{2n} & \cdots & F_{nn} \end{pmatrix}$$

Similarly, the decryption characteristic matrix $\mathcal{D}$ *is defined by the round function of decryption.*

For example, the round function of Feistel structure can be transformed to a UID structure, where the $\mathcal{E}$ and $\mathcal{D}$ of Feistel structure are:

$$\mathcal{E} = \begin{pmatrix} \mathbb{0} & \mathbb{1} \\ \mathbb{1} & \varphi \end{pmatrix}, \mathcal{D} = \begin{pmatrix} \varphi & \mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix}$$

The characteristic matrix describes the relations between input difference variables and output difference variables. As the definition above, there are totally four kinds of transformations in the matrix. It is unknown whether these four kinds of transformations are sufficient to describe any round function, but most current block cipher structures can be transformed to a UID structure.

The first step for converting the characteristic matrix is to define the subblock size. Usually, the size of subblock is the same as the size of 'S-Box' in the round function.

When analyze the structure of a round function, we usually transform the round function into a composition of several UID structures, and for each structure a characteristic matrix is formed. For example, we design a simple round function as Fig.3:

The characteristic matrices of this example is:

$$\mathcal{E}_1 \cdot \mathcal{E}_2 = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} \begin{pmatrix} \varphi & \varphi \\ \mathbb{0} & \mathbb{1} \end{pmatrix}$$

In the example, we divide the round function into the composition of two functions. The first function is $(Y1, Y2) = (X1 \oplus X2, X2)$ and the second function is $(Y1, Y2) = (F(X1), F(X1) \oplus X2)$. Consequently the UID structure of the first and second function are $\mathcal{E}_1$ and $\mathcal{E}_2$, defined as above. Note that we cannot multiply the two matrices together since the multiplication of transformation is not defined.

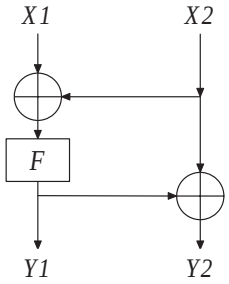


Figure 3: A round function can be divided into 2 UID structures

2.3 Searching Impossible Differential Characteristics

The meaning of three sets L, M and R in the UID-identities are known difference, unknown non-zero difference and unknown difference respectively. Given an input or output difference structure $\Delta = (x_1, \cdots x_n)$, the corresponding UID vector U can be written as $\{u_1, \cdots, u_n\}$ where $u_i =< \{l_i\}, 0, 0 >$.

Definition 6 *Suppose $U = \{u_1, \cdots u_n\}$ is an n-dimension vector of UID-identity. The multiplication of U and an encryption (decryption) characteristic matrix $\mathcal{E}(\mathcal{D})$ is define as:*

$$U \cdot \mathcal{E} = (\sum_{i=1}^{n} u_i \cdot \mathcal{E}_{i1}, \sum_{i=1}^{n} u_i \cdot \mathcal{E}_{i2}, \ldots, \sum_{i=1}^{n} u_i \cdot \mathcal{E}_{in})$$

For example, the input difference of the Feistel round function is $(l, 0)$, then $U = (< \{l\}, \emptyset, \emptyset >, < \emptyset, \emptyset, \emptyset >)$.

$$\begin{aligned} U \cdot \mathcal{E} &= (< \{l\}, \emptyset, \emptyset >, < \emptyset, \emptyset, \emptyset >) \cdot \begin{pmatrix} \mathbb{0} & \mathbb{1} \\ \mathbb{1} & \varphi \end{pmatrix} \\ &= (< \emptyset, \emptyset, \emptyset >, < \{l\}, \emptyset, \emptyset >) \end{aligned}$$

After the encryption function and decryption function are converted into matrices, denoted as $\mathcal{E}_1 \cdots \mathcal{E}_i$ and $\mathcal{D}_1 \cdots \mathcal{D}_j$, we can compute the UID vectors of the intermediate differences after i-round encryption from the input difference vector U_X^0:

$$U_X^i = (U_X^0 \cdot \mathcal{E}_1) \cdot \cdots \cdot \mathcal{E}_i)$$

and the UID vector of the intermediate difference after j-round decryption from the out difference vector U_Y^0:

$$U_Y^j = (U_Y^0 \cdot \mathcal{D}_1) \cdot \cdots \cdot \mathcal{D}_j)$$

Note that the multiplication between UID vectors and characteristic matrix has not association laws,

$$U \cdot \mathcal{E}_1 \cdot \mathcal{E}_2 \neq U \cdot (\mathcal{E}_1 \cdot \mathcal{E}_2)$$

The last step of finding impossible differential characteristics is detecting the inconsistency of UID vectors from encrypt and decrypt processes. The inconsistency of two UID vectors is defined as follows.

Definition 7 *Two UID vectors $U = (u_1, u_2, \ldots, u_n)$ and $V = (v_1, v_2, \ldots, v_n)$ are inconsistent if and only if there exists a subset $I \subseteq \{1, 2, \ldots, n\}$, such that the sum of UID-identities in the subset are inconsistent:*

$$\sum_{i \in I} u_i \neq_{inconsist} \sum_{i \in I} v_i$$

Two UID-identities, denoted as $< L_1, M_1, R_1 >$ and $< L_2, M_2, R_2 >$, are inconsistent if and only if one of following condition satisfied:

- $|L_1| \neq |L_2|, |M_1| = |M_2|, |R_1| = |R_2|$. *The known difference is inequal.*
- $|L_1| = |L_2|, |M_1 \boxplus M_2| = 1, |R_1| = |R_2|$. *There exists a unknown non-zero variable equals to zero.*

In the UID cryptanalysis, firstly we choose an input UID vector U_X^0 and an output UID vector U_Y^0, then we calculate U_X^i from U_X^0 forwardly and U_Y^j from U_Y^0 inversely, if U_X^i and U_Y^j are inconsistent then we get an impossible differential characteristic.

After achieving a maximum $i+j$ such that U_X^i and U_Y^j are inconsistent, we find the longest impossible differential characteristic based on given input difference U_X^0 and output difference U_Y^0. To find the longest impossible difference of a block cipher, we enumerate every possible UID vector of input U_X and output U_Y, and find the maximum $i+j$. Since the adversary knows the input and output difference of block cipher, UID-identities in the input and output of the encryption algorithm contain L component only. Take the 2-subblock Feistel structure as an example, the possible input/output UID vectors are listed in Table 2.

Table 2: Possible input/output UID vectors

UID vector	difference
$(<\emptyset,\emptyset,\emptyset>,<\emptyset,\emptyset,\emptyset>)$	$(0,0)$
$(<\{l_1\},\emptyset,\emptyset>,<\emptyset,\emptyset,\emptyset>)$	$(\Delta_1,0)$
$(<\emptyset,\emptyset,\emptyset>,<\{l_1\},\emptyset,\emptyset>)$	$(0,\Delta_1)$
$(<\{l_1\},\emptyset,\emptyset>,<\{l_1\},\emptyset,\emptyset>)$	(Δ_1,Δ_1)
$(<\{l_1\},\emptyset,\emptyset>,<\{l_2\},\emptyset,\emptyset>)$	(Δ_1,Δ_2)

It is easy to know that the total number N of possible input-output UID-vectors has the relation $N \leq (2n+2)!$. Since n is small, it is feasible to search by the computer. We also remove trivial impossible differential characteristics:

$$\{0,\cdots,x,\cdots,0\} \nrightarrow_\infty \{0,\cdots,0\}$$
$$\{0,\cdots,0\} \nrightarrow_\infty \{0,\cdots,x,\cdots,0\}$$

3 Practical Results on Block Ciphers

In this section, we will give a detailed analysis of block cipher Four-Cell with UID cryptanalysis, and list our results for some popular block cipher structures, such as Gen-CAST, Gen-MARS, Gen-RC6 [8], SMS4 [13] and FOX64 [5].

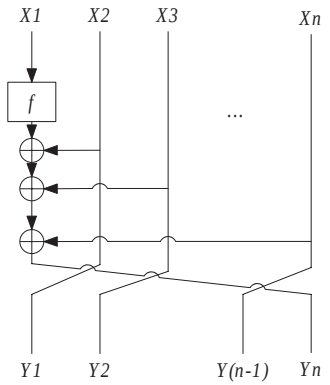


Figure 4: Round function of n-Cell GF-NLFSR

Choy *et al.* proposed a new structure called generalized feistel non-linear feedback shift register (GF-NLFSR) [4]. A new block cipher called Four-Cell was designed based on the 4-cell GF-NLFSR. The round function of n-cell GF-NLFSR is depicted in Fig.4.

The $\mathcal{E}$ and $\mathcal{D}$ of Four-Cell are:

$$\mathcal{E} = \begin{pmatrix} \mathbb{0} & \mathbb{0} & \mathbb{0} & \varphi \\ \mathbb{1} & \mathbb{0} & \mathbb{0} & \mathbb{1} \\ \mathbb{0} & \mathbb{1} & \mathbb{0} & \mathbb{1} \\ \mathbb{0} & \mathbb{0} & \mathbb{1} & \mathbb{1} \end{pmatrix}, \mathcal{D}_{\mathbb{1}}\mathcal{D}_2 = \begin{pmatrix} \mathbb{1} & \mathbb{1} & \mathbb{0} & \mathbb{0} \\ \mathbb{1} & \mathbb{0} & \mathbb{1} & \mathbb{0} \\ \mathbb{1} & \mathbb{0} & \mathbb{0} & \mathbb{1} \\ \mathbb{1} & \mathbb{0} & \mathbb{0} & \mathbb{0} \end{pmatrix} \begin{pmatrix} \varphi & \mathbb{0} & \mathbb{0} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} & \mathbb{0} & \mathbb{0} \\ \mathbb{0} & \mathbb{0} & \mathbb{1} & \mathbb{0} \\ \mathbb{0} & \mathbb{0} & \mathbb{0} & \mathbb{1} \end{pmatrix}$$

With UID cryptanalysis, we found one 18-round impossible differential characteristic: $\{x,0,0,0\} \nrightarrow_{18} \{y,y,0,0\}$. The UID-identities of the middle states of the characteristic are shown in Table 3. In these states, the 4th component

Table 3: UID Method on Four-Cell

R	X_1	X_2	X_3	X_4
0	$\{l_1\},\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$
1	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1\},\emptyset$
2	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1\},\emptyset$	$\emptyset,\{m_1\},\emptyset$
3	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1\},\emptyset$	$\emptyset,\{m_1\},\emptyset$	$\emptyset,\emptyset,\emptyset$
4	$\emptyset,\{m_1\},\emptyset$	$\emptyset,\{m_1\},\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$
5	$\emptyset,\{m_1\},\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1,m_3\},\emptyset$
6	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1,m_3\},\emptyset$	$\emptyset,\{m_1,m_3,m_5\},\emptyset$
7	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1,m_3\},\emptyset$	$\emptyset,\{m_1,m_3,m_5\},\emptyset$	$\emptyset,\{m_5\},\emptyset$
8	$\emptyset,\{m_1,m_3\},\emptyset$	$\emptyset,\{m_1,m_3,m_5\},\emptyset$	$\emptyset,\{m_5\},\emptyset$	$\emptyset,\emptyset,\emptyset$
9	$\emptyset,\{m_1,m_3,m_5\},\emptyset$	$\emptyset,\{m_5\},\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1,m_3\},\{r_4\}$
10	$\emptyset,\{m_5\},\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1,m_3\},\{r_4\}$	$\emptyset,\{m_1,m_3,m_5\},\{r_4,r_6\}$
11	$\emptyset,\emptyset,\emptyset$	$\emptyset,\{m_1,m_3\},\{r_4\}$	$\emptyset,\{m_1,m_3,m_5\},\{r_4,r_6\}$	$\emptyset,\{m_5,m_6\},\{r_6\}$
12	$\emptyset,\{m_1,m_3\},\{r_4\}$	$\emptyset,\{m_1,m_3,m_5\},\{r_4,r_6\}$	$\emptyset,\{m_5,m_6\},\{r_6\}$	$\emptyset,\{m_6\},\emptyset$
12	$\emptyset,\emptyset,\{r_1\}$	$\emptyset,\{m_4\},\emptyset$	$\emptyset,\{m_2\},\emptyset$	$\emptyset,\emptyset,\emptyset$
13	$\emptyset,\{m_4\},\emptyset$	$\emptyset,\{m_2\},\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$
14	$\emptyset,\{m_2\},\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$
15	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\{l_2\},\emptyset,\emptyset$
16	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\{l_2\},\emptyset,\emptyset$	$\{l_2\},\emptyset,\emptyset$
17	$\emptyset,\emptyset,\emptyset$	$\{l_2\},\emptyset,\emptyset$	$\{l_2\},\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$
18	$\{l_2\},\emptyset,\emptyset$	$\{l_2\},\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$	$\emptyset,\emptyset,\emptyset$

of $U_X^{12} =< \emptyset,\{m_6\},\emptyset >$ and $U_Y^6 =< \emptyset,\emptyset,\emptyset >$ are inconsistent. Thus an impossible differential characteristic is found. The result is the same as the best impossible differential cryptanalysis of Four-Cell. [12, 12]

Besides the Four-Cell cipher, we also give the UID cryptanalysis results to some popular block cipher structures, as listed in Table 4.

Table 4: UID method on popular block cipher structures

Block Cipher	Subblock	UID method	Characteristic	comment
Four-Cell[4]	4	18 round	$\{x,0,0,0\} \nrightarrow_{18} \{y,y,0,0\}$	the same as[12]
Gen CAST-256[8]	4	16 round	$\{0,0,0,x\} \nrightarrow_{16} \{y,0,0,x\}$	this paper
Gen MARS[8]	4	9 round	$\{0,0,0,x\} \nrightarrow_{9} \{x,y,0,0\}$	this paper
Gen RC6[8]	4	9 round	$\{0,0,x,0\} \nrightarrow_{9} \{0,y,0,0\}$	[6]
SMS4[13]	4	11 round	$\{x,x,x,0\} \nrightarrow_{11} \{0,x,x,x\}$	[7]
FOX64[5]	8	4 round	$\{0,x,0,x,0,x,0,x\} \nrightarrow_4$ $\{y_1,y_2,y_1,y_3,y_1,y_2,y_1,y_3\}$	the same as [11]

In Table 4, the UID result on Four-Cell is the same as [12] and the UID result on FOX64 is the same as [11]. In [6], Kim *et al.* present a 15-round impossible differential characteristics on Gen CAST-256 and 7-round impossible differential characteristics on Gen MARs. As in Table 4, we found a 16-round impossible differential characteristics on Gen CAST-256 and 9-round differential characteristics on Gen MARS, which are better than Kim *et al.*'s results. On the Gen RC6 block cipher structure, Kim *et al.* present a 17-round impossible differential characteristics which is better than ours. We doubt this result since we can not reproduce it, both by Kim *et al.*'s $\mathcal{U}$-method and manual work. For the block cipher SMS4, we found a 11-round impossible differential characteristics, which is shorter than Lu's result. This is because Lu used the details of the S-Box to exhaustive search the impossible differential, while our method consider only the block cipher structures.

4 A Comparison with $\mathcal{U}$-method

There are several limitations of $\mathcal{U}$-method:

a). *The characteristic matrix of block cipher must have 1-Property*.

If the number of entry 1 in each column of the characteristics matrix is zero or one, then the matrix is a 1-property matrix. $\mathcal{U}$-method can not determine the result of addition of two known differences, thus the method works only on those block ciphers whose characteristics matrices have 1-property. In our method, we express the intermediate states directly by variables. And the round function can be decomposed into several stages, and the characteristic matrix is generated for each stage; hence our method is flexible when representing the round function.

b). *Some information is lost during the calculating the impossible differential characteristics in the $\mathcal{U}$-method.*

For example, denote l_i as a known difference, and m_j as a unknown non-zero difference. Assume the difference in a intermediate state is $s_1 = l_1 + m_1$ and difference in another intermediate state is $s_2 = m_1$. In $\mathcal{U}$-method, s_1 is denoted as '2*' and s_2 is denoted as '1'; the sum of them is $2^* + 1 = 3$, which means unknown difference. Comparied with the $\mathcal{U}$-method, our method denotes s_1 as $< \{l_1\}, \{m_1\}, \emptyset >$ and s_2 as $< \{0\}, \{m_1\}, \emptyset >$, the sum of these UID-identities is $< \{l_1\}, \emptyset, \emptyset >$, which means a known difference.

c). *$\mathcal{U}$-method can not determine some kinds of inconsistencies*.

$\mathcal{U}$-method considers only the inconsistency by the corresponding component of vectors. Our method considers the inconsistency of the sum of several corresponding components, which has more capability to detect the conflict. For example, the difference of input after a rounds is $U_X^a = (m_1, m_1)$ and the difference of output before b rounds is $U_Y^b = (l_1 + m_2, m_2)$. In the $\mathcal{U}$-method, U_X^a is denoted as $\{1, 1\}$ while $U_X^b X$ is denoted as $\{2^*, 1\}$. There is no conflict. In our method, U_X^a and U_Y^b are denoted as $\{< \emptyset, \{m_1\}, \emptyset >, < \emptyset, \{m_1\}, \emptyset >\}$ and $\{< \{l_1\}, \{m_2\}, \emptyset >, < \emptyset, \{m_2\}, \emptyset >\}$ respectively, thus an inconsistency of these two vectors is found:

$$< \emptyset, \{m_1\}, \emptyset > + < \emptyset, \{m_1\}, \emptyset > \neq < \{l_1\}, \{m_2\}, \emptyset > + < \emptyset, \{m_2\}, \emptyset > .$$

UID cryptanalysis is suitable for those block ciphers whose nonlinear transformation is a permutation and especially use for the Feistel or extended Feistel structure. Since in the UID cryptanalysis, we assume that for nonlinear transformation in the block cipher, non-zero input difference results in non-zero output difference. There exist some S-boxes of block ciphers are not permutations, such as the S-box in DES. Thus both $\mathcal{U}$-method and UID method are useless in this case.

5 Conclusion

Inspired by the work [6] of automatically retrieving the impossible differential characteristics, we made some improvement based on $\mathcal{U}$-method and proposed unified impossible differential cryptanalysis on block cipher structures. By UID cryptanalysis, we found improved impossible differential cryptanalysis characteristics with the generalized CAST-256 and generalized MARS block cipher structure, which are better than Kim *et al.*'s $\mathcal{U}$ method. On the block cipher Four-Cell and FOX64, our results are just the same as previous' manual work. Thus, UID cryptanalysis can be used as a unify tool to evaluate the vulnerability of new block ciphers against impossible differential cryptanalysis.

References

[1] E. Biham, A. Biryukov and A. Shamir, Cryptanalysis of Skipjack reduced to 31 rounds using impossible diffrentials, *EUROCRYPT'99*, LNCS 2595, pp:12-23, 1999.

[2] E. Biham, A. Biryukov and A. Shamir, Miss in the middle attacks on IDEA. *FSE'99*, LNCS 1636, pp. 124-138, 1999.

[3] J. Cheon, M. Kim, K. Kim and J. Lee. Improved impossible differential cryptanalysis of Rijndael and Crypton, *ICISC'01*, LNCS 2288, pp 39-49, 2001.

[4] J. Choy, G. Chew, K. Khoo and H. Yap: Cryptographic properties and application of a generalized unbalanced feistel network structure. *ACISP'2009*, LNCS 5594, pp.73-89, 2009.

[5] P. Junod and S. Vaudenay, Fox: A new family of block ciphers, *SAC'04*, LNCS 2595, pp 131-146, 2004.

[6] J. Kim, S. Hong, J. Sung, S. Lee and J. Lim: Impossible differential cryptanalysis for block cipher structures, *INDOCRYPT 2003*, LNCS 2904, pp. 82-96, 2003.

[7] J. Lu, Attacking reduced-round versions of the SMS4 block cipher in the chinese WAPI standard, *ICICS'07*, LNCS 4861, pp. 306-318, 2007.

[8] S. Moriai and S. Vaudenay, On the pseudorandomness of Top-Level schemes of block ciphers, *ASIACRYPT'00*, LNCS 1976, pp 289-302, 2000.

[9] K. Nyberg, Differentially uniform mappings for cryptography, *Eurocrypt'93*, LNCS 765, pp. 55-64, 1994.

[10] D. Toz and O. Dunkelman, Analysis of two attacks on reduced-round versions of the SMS4 , *ICICS'08*, LNCS 5308, pp. 141-156, 2008.

[11] Z. Wu, Y. Luo, X. Lai and B. Zhu. Improved cryptanalysis of FOX block cipher, *The First International Conference on Trusted Systems (INTRUST 2009)*. To appear.

[12] W. Wu, L. Zhang, L. Zhang and W. Zhang: Security analysis of the GF-NLFSR structure and Four-Cell block cipher, *Cryptology ePrint Archive, Report 2009/346*, 2009.

[13] Specication of SMS4, block cipher for WLAN products SMS4 (in Chinese), http://www.oscca.gov.cn/UpFile/200621016423197990.pdf.

[14] W. Zhang, W. Wu, and D. Feng, New results on impossible differential cryptanalysis of reduced AES, *ICISC'07*, LNCS 4817, pp:239-250, 2007.