The Relationship between the Construction and Solution of the MILP Models and Applications

Lingchen Li[1,2], Wenling Wu[1], and Yafei Zheng[1] and Lei Zhang[1]

[1] Institute of Software, Chinese Academy of Sciences, Beijing 100190, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
lilingchen@tca.iscas.ac.cn

Abstract. The automatic search method based on Mix-integer Linear Programming (MILP) is one of the most common tools to search the distinguishers of block ciphers. For differential analysis, the byte-oriented MILP model is usually used to count the number of differential active s-boxes and the bit-oriented MILP model is used to search the optimal differential characteristic. In this paper, we present the influences between the construction and solution of MILP models solved by Gurobi: 1). the number of variables; 2). the number of constraints; 3). the order of the constraints; 4). the order of variables in constraints. We carefully construct the MILP models according to these influences in order to find the desired results in a reasonable time.

As applications, we search the differential characteristic of PRESENT, GIFT-64 and GIFT-128 in the single-key setting. We do a dual processing for the constraints of the s-box. It only takes 298 seconds to finish the search of the 8-round optimal differential characteristic based on the new MILP model. We also obtain the optimal differential characteristic of the 9/10/11-round PRESENT. With a special initial constraint, it only takes 4 seconds to obtain a 9-round differential characteristic with probability 2^{-42}. We also get a 12/13-round differential characteristic with probability $2^{-58}/2^{-62}$. For GIFT-128, we improve the probability of differential characteristic of $9 \sim 21$ rounds and give the first attack on 26-round GIFT-128 based on a 20-round differential characteristic with probability $2^{-121.415}$.

Keywords: MILP, Gurobi, PRESENT, GIFT, Differential Cryptanalysis

1 Introduction

Differential Cryptanalysis is one of the most important and effective attacks on block ciphers which was first proposed by [1]. And then a lot of attacks were devised based on this method, such as related-key difference attack, truncated difference attack, impossible differential attack, rectangle attack and so on[2–5]. Evaluating the ability of a block cipher against the differential attack is a basic and important requirement.

The first step of the differential analysis is to find a difference characteristic with the high probability. The most classical method to search the optimal differential characteristic is the branch and bound algorithm [6], which has a high requirement for the programming ability. In recent years, the automatic search method based on the combination optimal problem is very efficient and has attracted wide attention. The search problem is described as an MILP, SMT/SAT or CP models and solved by the corresponding solvers [7–9]. Among them, the MILP method has been widely used in cryptanalysis. It can be used to the search of the differential characteristics, impossible differentials, divisible property and so on [10–12]. The Gurobi optimizer is recognized as the most efficient commercial solver [13]. In previous studies, the solution time is an important factor to evaluate the merits and demerits of different search methods. It seems that CP or SAT method outperforms the MILP method in [14, 15]. In this paper, we mainly focus on the influences between the construction and solution of MILP models solved by Gurobi. Actually, the solution time can be greatly improved with a carefully constructed MILP model.

Generally, the size of the MILP model related to the number of variables and constraints is proportionate to the solution time. In [16], they found that there is no clear positive correlation between the solution time and the number of inequalities of the differential distribution table (DDT) of the s-box. This phenomenon is much more than that one. The Gurobi optimizer is a commercial solver and the internal algorithm isn't published. Obviously, the search strategy of Gurobi is not the normal depth-first traversal. Usually, reducing the number of variables and constraints in MILP model or setting the special optimization parameters provided by Gurobi to modify the high-level solution strategy can accelerate the solution time [17]. In this paper, we observe the sensitivity of the Gurobi optimizer for the description of constraints in MILP model. The description of constraints is also an important influence to the solution time.

Our contribution. In this paper, we find that the influences between the construction and solution of MILP models solved by Gurobi are: 1). the number of variables; 2). the number of constraints; 3). the order of constraints; 4). the order of variables in constraints. We carefully construct a high quality model to search the differential characteristics. The solution time can be greatly improved than before. The inequalities of the differential propagation of the s-box without and with the probability information are added to the MILP model successively. The dual processing increases the number of constraints, but the solution time can be greatly reduced. As applications, we search the single-key differential characteristic of PRESENT, GIFT-64 and GIFT-128.

1. For PRESENT, the inequalities of the differential branch number of the s-box are added firstly. And then the inequalities with the probability information are considered too. In this model, the solution time of the search of the 8-round optimal differential characteristic is reduced from 4 days to 289 seconds. We also firstly obtain the exact bounds of the probability of the differential characteristic of $9 \sim 11$ rounds PRESENT.

2. For GIFT-64, it only takes 4 seconds to obtain the 9-round differential characteristic with probability 2^{-42} with a special initial constraint. We also can get a 12/13-round difference characteristic with probability $2^{-58}/2^{-62}$.
3. We improve the probability of the differential characteristic of the $9 \sim 21$ rounds GIFT-128. Based on a 20-round differential characteristic with probability $2^{-121.415}$, we can add 4-round at its beginning and 2-round at the end to attack 26-round GIFT-128. This is the best result for GIFT-128.

Organization. The paper is organized as follows. In Sect.2, we give a brief introduction of the automatic search tool based on MILP method. The relationship between the construction and solution of the MILP models solved by Gurobi is presented in Sect.3. We apply the high quality MILP model to PRESENT, GIFT-64 and GIFT-128 in Sect.4. Finally, we conclude this paper in Sect.5.

2 The automatic search tool based on MILP method

2.1 The construction and solution of MILP models

Mouha *et al.* [7] proposed the first framework to count the number of differential active s-boxes for AES. A MILP problem mainly contains two parts: the objective and the constraints. The objective can be set to the number of the active s-box or the differential probability. The constraints are based on the components of block ciphers, including the linear operations (*e.g.* XOR, MDS) and nonlinear operations(*e.g.* AND, S-box).

The byte- and bit-oriented models in differential analysis are presented as follows.

1. **The byte-oriented MILP model**. Model of this kind is usually used to account the number of differential active s-boxes of block ciphers. The constraints mainly come from the linear operations.
   (a) The XOR operation: $z = x \oplus y$. $(x, y, z) \notin \{(0,0,1),(0,1,0),(1,0,0)\}$ are the impossible differential patterns. Based on the logical condition, the constraints are:
   $$\begin{cases} x + y - z \geq 0 \\ x - y + z \geq 0 \\ -x + y + z \geq 0 \end{cases} \tag{1}$$
   (b) The linear operation with the branch number B: $(x_0, \ldots, x_{n-1}) \to (y_0, \ldots, y_{n-1})$.
   $$\begin{cases} \sum_i (x_i + y_i) - B \cdot d \geq 0 \\ d - x_i \geq 0, i \in \{0, \ldots, n-1\} \\ d - y_i \geq 0, i \in \{0, \ldots, n-1\} \end{cases} \tag{2}$$
   where d is a dummy variable taking values in $\{0, 1\}$.
   The s-box operation doesn't require any constraints. The objective in byte-oriented MILP model is the sum of the variables representing the input words of s-boxes $\sum x_i$.

2. **The bit-oriented MILP model**. This model is more detailed than the byte-oriented model.
   (a) The XOR operation: $z = x \oplus y$. The $(x, y, z) \notin \{(0,0,1),(0,1,0),(1,0,0),(1,1,1)\}$ are the impossible differential patterns. Based on the logical condition, the constraints are:

$$\begin{cases} x + y - z \geq 0 \\ x - y + z \geq 0 \\ -x + y + z \geq 0 \\ -x - y - z \geq -2 \end{cases} \tag{3}$$

   (b) The s-box operation: $(x_0, \ldots, x_{n-1}) \xrightarrow{s-box} (y_0, \ldots, y_{m-1})$. For this operation, we can further divide it into three classes:
      i. Without considering the differential propagation of the s-box, only the non-zero input difference must results in the non-zero output difference. By using this model, the differential properties of the linear layer of the block cipher can be analyzed.
      In this situation, a variable is introduced to mark a s-box, which means that $A = 1$ if the input word of the s-box is nonzero, otherwise $A = 0$. Then the constraints are:

$$\begin{cases} A - x_i \geq 0, i \in \{0, \ldots, n-1\} \\ x_0 + \ldots + x_{n-1} - A \geq 0 \end{cases} \tag{4}$$

      For the bijective property of the s-box, the nonzero input difference must result in nonzero output difference and vice versa. The constraints are:

$$\begin{cases} ny_0 + \ldots + ny_{m-1} - x_0 - \ldots - x_{n-1} \geq 0 \\ mx_0 + \ldots + mx_{n-1} - y_0 - \ldots - y_{m-1} \geq 0 \end{cases} \tag{5}$$

      The objective of this situation is $\sum A_i$.
      ii. Considering the possibility of the differential propagation patterns of the s-box. The impossible differential propagation patterns are excluded by inequalities which can be generated by the SAGE software [10]. The exact lower bounds of the differential active s-boxes can be obtained. The objective function is also minimize $\sum A_i$.
      iii. We can further consider the probability of the differential propagation of the s-box. The probability of the optimal difference characteristic can be obtained. In this model, it is necessary to introduce some variables to distinguish the probability. For example, if there are two nontrivial probability for a 4-bit s-box, two variables (p_0, p_1) need to be introduced:

$$(p_0, p_1) = \begin{cases} (0,0), \text{if } \Pr[(x_0,x_1,x_2,x_3) \to (y_0,y_1,y_2,y_3)] = 2^0 \\ (0,1), \text{if } \Pr[(x_0,x_1,x_2,x_3) \to (y_0,y_1,y_2,y_3)] = 2^{-2} \\ (1,0), \text{if } \Pr[(x_0,x_1,x_2,x_3) \to (y_0,y_1,y_2,y_3)] = 2^{-3} \end{cases} \tag{6}$$

This model is used to search the optimal differential characteristic of block ciphers. The objective function is minimize $\sum(3p_0 + 2p_1)$. The probability of the optimal differential characteristic is $2^{-min\sum(3p_0+2p_1)}$.

After defining the objective function and the constraints, we can construct the corresponding MILP instance which is often described in LP formate and solved by the Gurobi optimizer.

3 The observations of the construction and solution of MILP models solved by Gurobi

Reducing the solution time is what we've been looking for. However, there is no clear positive correlation between the number of inequalities of the DDT of s-box and the solution time as found in [16]. The Gurobi optimizer is one of the best performance solvers for MILP problems. While the internal search algorithms are not public. We study the relationship between the construction and solution of MILP models solved by Gurobi through several tests. We output all the solutions of each model by setting the parameter PoolSearchMode=2. We check the sequence of the solutions of each model to reflect the search strategy of Gurobi.

Obviously, the search strategy of Gurobi is not the normal depth-first traversal. The sequence of solutions of this model doesn't have any particular rule even in the empty model. Only variables need to be defined in the empty model. For example, the sequence of the empty model with $x_0x_1x_2x_3$ is $(0000, 1000, 0100, 1100, 0010, 1010, 0110, 1110, 0001, 1001, 0101, 1101, 0011, 1011, 0111, 1111)$ as shown in Fig.1. The order of the first half of the solutions is regular while the later is not.

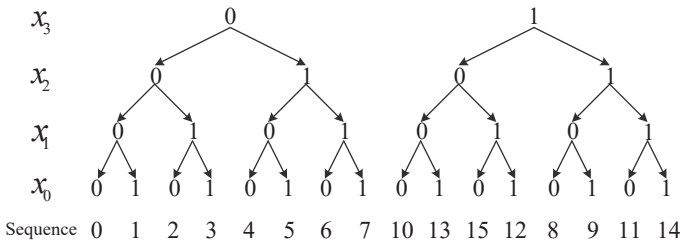

Fig. 1. The sequence of the empty model with $x_0x_1x_2x_3$

The advanced searching algorithms in Gurobi bring this phenomenon. It's the key to improve the solution time. We need to know what can change the search strategy and how to carefully construct a high quality MILP model. It's

the kernel in this paper. Next, we mainly focus on the relationship between the construction and solution of MILP models. We can change the number and order of variables and constraints. Actually, all of them would influence the search strategy of Gurobi. The growth of the number of variables will expand the search space which isn't a wise choice. So we only focus on the remaining three situations with the following tests. All models in tests have an empty objective function.

Test 1 The influence of the number of constraints. The inequalities in the first column of Appendix A are about the DDT of the GIFT s-box. This constraints are added in the two models. A redundant constraint $(-1, -1, -1, 1, 0, 0, -1, 0, 3)$ is later added at the end in the second model. Obviously, the solutions of two models are the same. We solve two models by Gurobi and output the whole solutions one by one. The test results show that the sequences of the solutions of two models are different after the 95-*th* solution which is equal to (01000111) and (11100001) respectively. Even we describe the constraints of the GIFT s-box twice, the sequence of solution is changed after the 20-*th* solution.

Test 2 The influence of the order of constraints. We rotate down the constraints in the first column of Appendix A to produce a total of 21 models. The results show that the sequence of the solutions is different to the models with the number of rotation $(0 \sim 4, 17 \sim 20)$ and the number of rotation $(5 \sim 16)$. We may obtain more kinds of sequences if the order of constraints is set randomly.

Test 3 The influence of the order of variables in constraints. We change the order of variables in constraints by the left rotation. The constraints also come from the first column of Appendix A. For example, the constraint $a_0x_0 + a_1x_1 + \ldots + a_{n-1}x_{n-1} \geq C$ is rotated left by one position to get $a_1x_1 + \ldots + a_{n-1}x_{n-1} + a_0x_0 \geq C$. Eight models can be constructed by this change. The sequence of solutions for each model is not the same as shown in results.

As shown above, the Gurobi optimizer is very sensitive to the construction of MILP models. This sensitivity may bring unexpected advantages or disadvantages. In practice, we want to optimize the solution of the model by adjusting the description of the constraints in the model, so that we can obtain a optimal or higher quality solution in less time. We present some applications to show the feasibility of this way in next section.

The computation is performed on PC (Intel(R) Core(TM) i7-7500U CPU, 2.70 GHz, 8.00GB RAM, 4 cores, window10) with the optimizer Gurobi7.5.2.

4 Applications

4.1 Application to PRESENT

PRESENT is an Ultra-Lightweight block cipher proposed by Bogdanov *et al.*[18]. Its a 32-round SP-network with block size 64-bit and key size 80-bit or 128-bit.

Each round of PRESENT consists of three operations: sBoxLayer, pLayer and addRoundKey. The round function is shown in Fig.2. The sBoxLayer adopts the same sixteen 4-bit s-box, which is the only nonlinear component of this cipher.

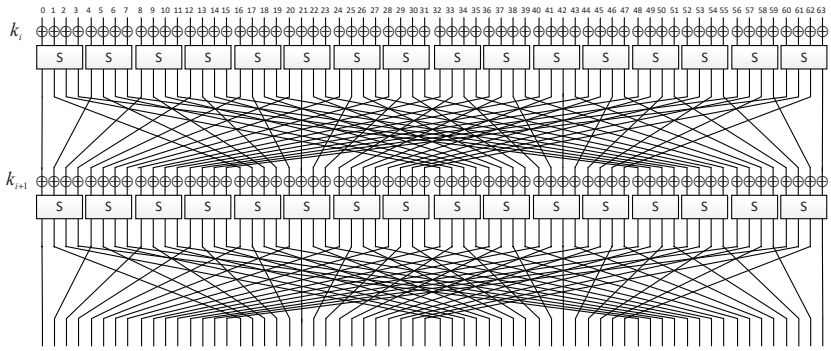

Fig. 2. The round function of PRESENT

For PRESENT, it's enough to add the constraints of the s-box operation. Based on the MILP method, the lower bounds of the number of differential active s-boxes under the single-key setting are given in [19]. The mainly constrains in the model are the differential branch number of s-box which is equal to 3. Then the inequalities of the DDT of the s-box are included to search the optimal difference characteristic [20]. The probability information is added according to the equation (6). It takes 4 days to search the 8-round optimal probability differential characteristic.

We carefully construct a new model to reduce the solution time. In section 3, we can change the search strategy of the solver by adjusting the number of constraints. Therefore, we try to do the dual processing for the s-box. That means the constraints of the differential branch number of the s-box are added in the model and then the constraints of the DDT of the s-box in Appendix A are added later. Superficially, the increase of the constraints leads to a bigger model file. While the solution time of the new model can be greatly improved. The solution time of the search of the optimal difference characteristic of the 8-round PRESENT is about 298 seconds. In addition, we obtain the optimal difference characteristics of $9 \sim 11$ rounds PRESENT in a reasonable time, as shown in the Table 1.

4.2 Application to GIFT-64

GIFT is called " A small PRESENT " which has a extremely good performance [21]. It has two versions, namely GIFT-64 and GIFT-128. GIFT-64 is a 28-round SPN cipher with 64-bit block size. GIFT-128 is a 40-round SPN cipher with 128-bit block size. Both versions have a key size of 128-bit. Similar to

Table 1. The results of the best single-key differential characteristic of PRESENT

Rounds	$log_2 Pr$	Time	
		This paper	[20]
8	-32	298s	4d
9	-36	184s	-
10	-41	11h	-
11	-46	2d	-

PRESENT, each round of GIFT consists of three operations: SubCells, PermBits and AddRoundKey. The round function of GIFT-64 is shown in Fig.3.

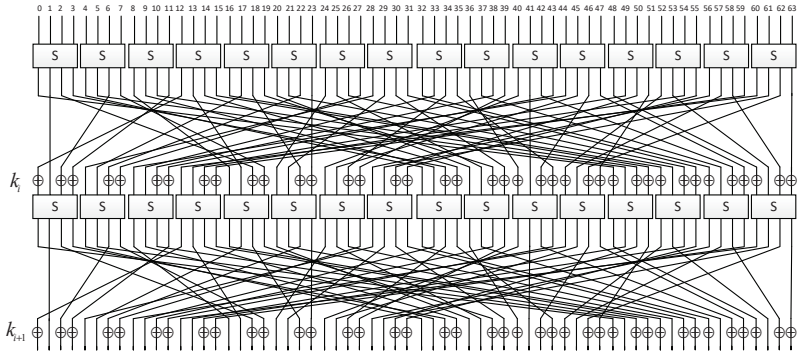

Fig. 3. The round function of GIFT-64

For GIFT-64, the design document gave the lower bounds of the differential active s-boxes based on the MILP method [21]. The difference of the 14-round of GIFT-64 is lower than 2^{-63} estimated by using the 9-round difference with probability $2^{-44.415}$. Subsequently, Zhu *et.al.* [22] got the 4-round iterative differential characteristic based on the MILP method and the two-step search strategy, and the 12/13 rounds differential characteristic with probability $2^{-60}/2^{-64}$.

The dual processing for the s-box is also valid to other block ciphers. However, the differential branch number of the GIFT s-box is equal to 2, which is a trivial property. Similar to PRESENT, we add the inequalities of the DDT of s-box without the probability information and then the constraints with probability information according to the equation (7). There are tree nontrivial probability values of the GIFT s-box, so we need to introduce three variables to distinguish them:

$$(p_0, p_1, p_2) = \begin{cases} (0,0,0), \text{if } \Pr[(x_0,x_1,x_2,x_3) \to (y_0,y_1,y_2,y_3)] = 2^0 \\ (0,0,1), \text{if } \Pr[(x_0,x_1,x_2,x_3) \to (y_0,y_1,y_2,y_3)] = 2^{-1.415} \\ (0,1,0), \text{if } \Pr[(x_0,x_1,x_2,x_3) \to (y_0,y_1,y_2,y_3)] = 2^{-2} \\ (1,0,0), \text{if } \Pr[(x_0,x_1,x_2,x_3) \to (y_0,y_1,y_2,y_3)] = 2^{-3} \end{cases} \tag{7}$$

The objective function is minimize $\sum(3p_0 + 2p_1 + 1.415p_2)$. Since 3 variables are required to distinguish the probability information. It's more difficult to solve the MILP model of GIFT than PRESENT. The solution time of the search of the optimal differential characteristic of 9-round GIFT-64 is still intolerable under the new model. As we known, the Gurobi optimizer usually returns a high quality feasible solution soon, while takes a long time to prove whether the solution is optimal. According to Test 3, we try to construct different models by changing the order of variables in the initial constraint and solve them with 10 seconds. Here, the initial constraint is the input difference of the first round:

$$x_0 + x_1 + \ldots + x_{63} \geq 1$$

Obviously, 64 models can be obtained through the left rotation of the variables in the initial constraint. The Gurobi is very sensitive to this minor change in the new model. It is worth noting that this sensitivity will greatly weaken if the MILP model only contains the constraints of the s-box with the probability information. The current solutions returned to us by Gurobi given by different models with 10 seconds are different. Actually, we also set the optimization parameter MIPFocus=2 to find a higher quality solution as quickly as possible. The results of the search of differential characteristic of 9-round GIFT-64 is shown in Fig.4. When the number of left rotation is equal to 24, the solver returns a feasible solution with 42 which appeared in the 4-*th* seconds. However, a feasible solution is equal to 87.83 which cannot be improved to 42 in a reasonable time under the default initial constraint.

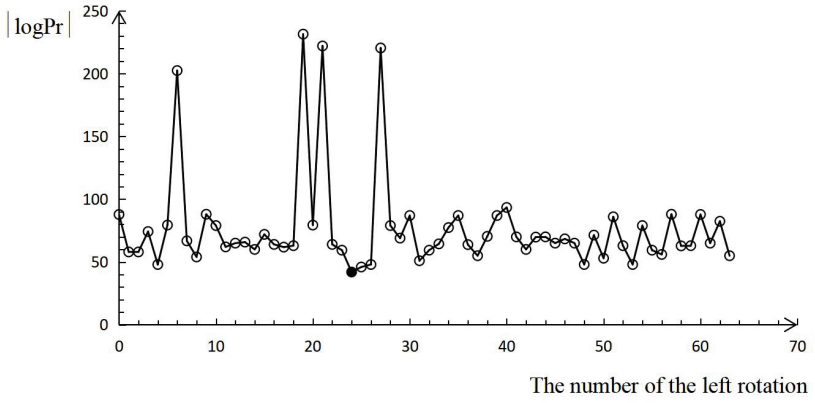

Fig. 4. The results of the 9-round GIFT-64

The 9-round differential characteristic with probability 2^{-42} is shown in Table 2. We also obtain some 4-round iterative differential characteristics with probability 2^{-20}, *e.g.* $(0000\,0000\,0202\,0000) \rightarrow (0000\,0000\,0202\,0000)$. And we can

Table 2. The differential characteristic of 9-round GIFT-64

Rounds	Differential	$log_2 Pr$
input	0000 000c 0000 0006	1
1	0000 0000 0202 0000	-4
2	0000 0050 0000 0050	-8
3	0000 0000 0000 0202	-14
4	0000 0005 0000 0005	-18
5	0000 0000 0202 0000	-24
6	0000 0050 0000 0050	-28
7	0000 0000 0000 0202	-34
8	0000 0005 0000 0005	-38
9	0080 0404 0202 0101	-42

generate a 12/13 rounds differential characteristics with probability $2^{-58}/2^{-62}$ in Appendix B.

4.3 Application to GIFT-128

For GIFT-128, the design document gave the lower bounds of the active S-box of $1 \sim 9$ rounds GIFT-128 based on the MILP method [21]. They obtained a 9-round differential characteristic with probability $2^{-46.99}$. In [22], they searched the optimal differential characteristic under some limitations. The r-round search is divided into $r_i (i = 1, 2, \ldots, t)$ rounds sub-ciphers, $\sum_1^t r_i = r$. The probability thresholds of each sub-cipher are set in advance. The output difference of the r_i-round sub-cipher is regarded as the input difference of the r_{i+1}-round sub-cipher. They obtained the differential characteristic of 18-round of GIFT-128 with probability 2^{-109} and gave a 23-round differential attack.

Under the new model, we search the optimal differential characteristic of GIFT-128. We divide the r-round search to two parts r_f and r_b, $r = r_f + r_b$. The hamming weight of the input (or output) difference is equal to 1 in r_b (or r_f)-round search. For GIFT-128, we set a active position to 1 and others to 0. A triple (ip, r_f, r_b) marks the search patterns, where ip represents the active position, r_f and r_b represent the number of round extended forward and backward respectively. For $ip \in \{0, \ldots, 127\}$, we search the optimal differential characteristic of $r_f \in \{2, 3, 4\}$ and $r_b \in \{3, 4, 5, 6\}$. We choose 16 positions $ip \in \{4, 12, 19, 23, 27, 31, 36, 44, 51, 55, 59, 63, 68, 76, 100, 108\}$ which have the 9-round differential characteristics with probability 2^{-47} when $r_f = 3$ and $r_b = 6$ to further do the high rounds' search. The results is showed in Table 3.

The differential attack on 26-round GIFT-128. Based on the 20-round differential characteristic (0000 0000 0000 0000 0000 0000 0000 00a0) $\rightarrow$ (0000 0000 4001 0000 2000 0000 1004 0000) with probability $2^{-121.415}$, a 26-round differential attack can obtained by extending forward 4-round and 2-round respectively.

The detail key-recovery process is shown in Table 5. In the data collection phase, we would build 2^n structures. Each structure traverses 64 bits undetermined in $\triangle X_S^1$. For such a pair, it has average probability of $2^{n+127-34-23-7-121.415}$

Table 3. The results of the singe-key differential characteristics for GIFT-128

Rounds	Searching Patterns	$log_2 Pr$	
		This paper	[22]
9	$(ip_a, 2, 7)$	-45.415	-47
10	$(ip_a, 2, 8)$	-49.415	-
11	$(ip_b, 2, 9)$	-54.415	-
12	$(ip_b, 3, 9)$	-60.415	-62.415
13	$(ip_b, 3, 10)$	-67.83	-
14	$(ip_a, 4, 10)$, $(ip_b, 3, 11)$	-79	-85
15	$(ip_a, 2, 13)$	-86	-
16	$(ip_a, 3, 13)$	-91	-
17	$(ip_a, 4, 13)$	-97	-
18	$(ip_a, 2, 16)$	-103.415	-109
19	$(ip_a, 5, 14)$	-115	-
20	$(ip_a, 2, 18)$	-121.415	-
21	$(ip_a, 3, 18)$	-126.415	-

[1] $ip_a \in \{4, 12, 36, 44, 68, 76, 100, 108\}$
[2] $ip_b \in \{19, 23, 27, 31, 51, 55, 59, 63\}$

Table 4. The 21-round differential characteristic with $2^{-126.415}$ for GIFT-128

Rounds	Differential	$log_2 Pr$
0	0000 0000 0000 0000 0000 0000 1060 0000	1
1	0000 0000 0000 0000 0000 0000 0000 00a0	-5
2	0000 0001 0000 0000 0000 0000 0000 0000	-7
3	0800 0000 0000 0000 0000 0000 0000 0000	-10
4	2000 0000 1000 0000 0000 0000 0000 0000	-12
5	4040 0000 2020 0000 0000 0000 0000 0000	-17
6	5050 0000 0000 0000 5050 0000 0000 0000	-25
7	0000 0000 0000 0000 0000 0000 a000 a000	-37
8	0000 0000 0000 0000 0000 0011 0000 0000	-41
9	0000 0800 0000 0800 0000 0000 0000 0000	-47
10	0202 0000 0101 0000 0000 0000 0000 0000	-51
11	0000 0000 5050 0000 0000 0000 5050 0000	-61
12	0000 0000 0000 0000 0000 0000 00a0 00a0	-73
13	0000 0011 0000 0000 0000 0000 0000 0000	-77
14	00800 0000 0800 0000 0000 0000 0000 0000	-83
15	2020 0000 1010 0000 0000 0000 0000 0000	-87
16	5050 0000 0000 0000 5050 0000 0000 0000	-97
17	0000 0000 a000 a000 0000 0000 0000 0000	-109
18	0000 0000 0000 0000 0011 0000 0000 0000	-113
19	0000 0000 0000 c000 0000 6000 0000 0000	-119
20	0004 0000 0000 0200 0000 0000 0000 0000	-123
21	0000 0000 4001 0000 2000 0000 1004 0000	-126.415

$= 2^{n-58.415}$ to meet the differential in 5-*th* round. Here we choose $n = 60.415$ to expect 4 pairs remaining for the right subkey guesses, the data complexity is $2^{124.415}$. According to the key expansion algorithm in Appendix C, the bits involved are 109 bits. It remains $2^{107.415}$ pairs which are filtered according to X_P^{26}. The time complexity is $2^{124.415}$ bounded by the date complexity and the memory complexity is 2^{109} bits.

5 Conclusion

In this paper, firstly we present the relationship between the construction and solution of the MILP models solved by the Gurobi optimizer. The results show that Gurobi is very sensitive to the changes of the number and order of constraints and variables. We carefully construct the MILP model to search the differential characteristics of PRESENT, GIFT-64 and GIFT-128. The solution time of the MILP model with the dual processing of the differential propagation of the s-box has a great improvement. We firstly obtain the bounds of the probability of differential characteristics of the $9 \sim 11$ rounds PRESENT in a reasonable time. With a special initial constraint, we only need 4 seconds to get the differential characteristics of 9-round GIFT-64 with probability 2^{-42}. In addition, We improve the probability of the differential characteristics of $9 \sim 21$ rounds GIFT-128 and obtain the first differential attack on 26-round GIFT-128.

The techniques to construct the MILP model, especially the dual processing of the constraints of s-box, are not only suitable for the block ciphers in our paper. In the future work, we will try to apply them to other ciphers.

Additional information. Recently, two papers have been posted to Cryptology ePrint Archive[23, 24]. We have independently reached the similar results, but the core idea is different. Compared with them, we want to emphasize:

1. We focus on the relationship between the construction and solution of MILP models. The solution time of a carefully constructed MILP model can be greatly improved. This has been reflected in the applications in our paper.
2. The probability of the differential characteristics of $9 \sim 11$ rounds PRESENT is the optimal.
3. The search of the differential characteristics of $9 \sim 21$ rounds GIFT-128 is more comprehensive and we give the first differential attack on 26-round GIFT-128.

Here, we present our work to others for reference. And we will go on to improve our results combined with other techniques in future.

References

1. Biham E, Shamir A. Differential cryptanalysis of DES-like cryptosystems[J]. Journal of CRYPTOLOGY, 1991, 4(1): 3-72.
2. Biham E. New types of cryptanalytic attacks using related keys[C]//Workshop on the Theory and Application of of Cryptographic Techniques. Springer Berlin Heidelberg, 1993: 398-409.

Table 5. The differential attack on 26-round GIFT-128

input	???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ????
	???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ????
$\triangle X_S^1$	0??0 00?? ?00? ??00 0??0 00?? ?00? ??00 0??0 00?? ?00? ??00 0??0 00?? ?00? ??00
	0??0 00?? ?00? ??00 0??0 00?? ?00? ??00 0??0 00?? ?00? ??00 0??0 00?? ?00? ??00
$\triangle X_P^1$	???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ????
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
$\triangle X_K^1$	???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ???? ????
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
$\triangle X_S^2$	0??0 00?? ?00? ??00 0??0 00?? 100? ?100 0??0 00?? ?00? ??00 0??0 00?? ?00? ??00
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
$\triangle X_P^2$	???? ???? ???? ???? 0000 0000 0000 0000 ???? 11?? ???? ???? 0000 0000 0000 00000
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 000
$\triangle X_K^2$	???? ???? ???? ???? 0000 0000 0000 0000 ???? 11?? ???? ???? 0000 0000 0000 00000
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 000
$\triangle X_S^3$	?000 0?00 00?0 000? 0000 0000 0000 0000 ?000 0100 00?0 000? 0000 0000 0000 0000
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
$\triangle X_P^3$	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
	0000 0000 0000 0000 0000 0000 0000 0000 ???? 0000 ?1?? 0000 0000 0000 0000 0000
$\triangle X_K^3$	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
	0000 0000 0000 0000 0000 0000 0000 0000 ???? 0000 ?1?? 0000 0000 0000 0000 0000
$\triangle X_S^4$	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
	0000 0000 0000 0000 0000 0000 0000 0000 1000 0000 0010 0000 0000 0000 0000 0000
$\triangle X_P^4$	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 1010 0000
$\triangle X_K^4$	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 1010 0000
5-*th* round	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
input	0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 1010 0000
24-*th* round	0000 0000 0000 0000 0000 0000 0000 0000 0100 0000 0000 0001 0000 0000 0000 0000
output	0010 0000 0000 0000 0000 0000 0000 0000 0001 0000 0000 0100 0000 0000 0000 0000
$\triangle X_S^{25}$	0000 0000 0000 0000 0000 0000 0000 0000 ???1 0000 0000 ???? 0000 0000 0000 0000
	???? 0000 0000 0000 0000 0000 0000 0000 ???? 0000 0000 ???1 0000 0000 0000 0000
$\triangle X_P^{25}$	0000 0000 ??00 0000 0?00 0000 ??00 0000 0000 0000 0??0 0000 00?0 0000 0??0 0000
	0000 0000 00?1 0000 000? 0000 00?? 0000 0000 0000 ?00? 0000 ?000 0000 ?001 0000
$\triangle X_K^{25}$	0000 0000 ??00 0000 0?00 0000 ??00 0000 0000 0000 0??0 0000 00?0 0000 0??0 0000
	0000 0000 00?1 0000 000? 0000 00?? 0000 0000 0000 ?00? 0000 ?000 0000 ?001 0000
$\triangle X_S^{26}$	0000 0000 ???? 0000 ???? 0000 ???? 0000 0000 0000 ???? 0000 ???? 0000 ???? 0000
	0000 0000 ???? 0000 ???? 0000 ???? 0000 0000 0000 ???? 0000 ???? 0000 ???? 0000
$\triangle X_P^{26}$	000? 0?0? 000? 0?0? 000? 0?0? 000? 0?0? ?000 ?0?0 ?000 ?0?0 ?000 ?0?0 ?000 ?0?0
	0?00 0?0? 0?00 0?0? 0?00 0?0? 0?00 0?0? 00?0 ?0?0 00?0 ?0?0 00?0 ?0?0 00?0 ?0?0
$\triangle X_K^{26}$	000? 0?0? 000? 0?0? 000? 0?0? 000? 0?0? ?000 ?0?0 ?000 ?0?0 ?000 ?0?0 ?000 ?0?0
	0?00 0?0? 0?00 0?0? 0?00 0?0? 0?00 0?0? 00?0 ?0?0 00?0 ?0?0 00?0 ?0?0 00?0 ?0?0

3. Knudsen L R. Truncated and higher order differentials[C]//International Workshop on Fast Software Encryption. Springer, Berlin, Heidelberg, 1994: 196-211.
4. Kim J, Hong S, Sung J, et al. Impossible differential cryptanalysis for block cipher structures[C]//International Conference on Cryptology in India. Springer, Berlin, Heidelberg, 2003: 82-96.
5. Wagner D. The boomerang attack[C]//International Workshop on Fast Software Encryption. Springer, Berlin, Heidelberg, 1999: 156-170.
6. Matsui M. On correlation between the order of S-boxes and the strength of DES[C]//Workshop on the Theory and Application of of Cryptographic Techniques. Springer Berlin Heidelberg, 1994: 366-375.
7. Mouha N, Wang Q, Gu D, et al. Differential and linear cryptanalysis using mixed-integer linear programming[C]//International Conference on Information Security and Cryptology. Springer Berlin Heidelberg, 2011: 57-76.
8. Mouha,N., Preneel,B.Towards finding optimal differential characteristics for ARX: Application to Salsa20. Cryptology ePrint Archive, Report 2017/139, 2017
9. Grault D, Lafourcade P, Minier M, et al. Revisiting AES Related-Key Differential Attacks with Constraint Programming[J]. IACR Cryptology ePrint Archive, 2017.
10. Sun S, Hu L, Wang P, et al. Automatic security evaluation and (related-key) differential characteristic search: application to SIMON, PRESENT, LBlock, DES (L) and other bit-oriented block ciphers[C]//International Conference on the Theory and Application of Cryptology and Information Security. Springer, Berlin, Heidelberg, 2014: 158-178.
11. Sasaki Y, Todo Y. New impossible differential search tool from design and cryptanalysis aspects[C]//Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2017: 185-215.
12. Xiang Z, Zhang W, Bao Z, et al. Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block ciphers[C]//Advances in CryptologyCASIACRYPT 2016: 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part I 22. Springer Berlin Heidelberg, 2016: 648-678.
13. Gurobi Optimization. Gurobi optimizer reference manual. 2013. http://www.gurobi.com.
14. Sun S, Gerault D, Lafourcade P, et al. Analysis of AES, SKINNY, and others with constraint programming[J]. IACR Transactions on Symmetric Cryptology, 2017.
15. Sun L, Wang W, Wang M. Automatic search of bit-based division property for ARX ciphers and word-based division property[C]//International Conference on the Theory and Application of Cryptology and Information Security. Springer, Cham, 2017: 128-157.
16. Sasaki Y, Todo Y. New Algorithm for Modeling S-box in MILP Based Differential and Division Trail Search[C]//International Conference for Information Technology and Communications. Springer, Cham, 2017: 150-165.
17. Li L., Wu W., Zhang L. Improved Automatic Search Tool for Bit-Oriented Block Ciphers and Its Applications. In: Qing S., Mitchell C., Chen L., Liu D. (eds) Information and Communications Security. ICICS 2017. Lecture Notes in Computer Science, vol 10631. Springer, Cham.
18. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: an ultra-lightweight block cipher. In: Cryptographic Hardware and Embedded Systems - CHES 2007, 9th International Workshop, Vienna, Austria, September 10-13, 2007, Proceedings. pp. 450-466 (2007), https://doi.org/10.1007/978-3-540-74735-2-31

19. Sun S., Hu L., Song L., Xie Y., Wang P. (2014) Automatic Security Evaluation of Block Ciphers with S-bP Structures Against Related-Key Differential Attacks. In: Lin D., Xu S., Yung M. (eds) Information Security and Cryptology. Inscrypt 2013. Lecture Notes in Computer Science, vol 8567. Springer, Cham.
20. Sun S, Hu L, Wang M, et al. Towards finding the best characteristics of some bit-oriented block ciphers and automatic enumeration of (related-key) differential and linear characteristics with predefined properties. Cryptology ePrint Archive, Report 2014/747, 2014.
21. Banik S, Pandey S K, Peyrin T, et al. GIFT: a small PRESENT[C]//International Conference on Cryptographic Hardware and Embedded Systems. Springer, Cham, 2017: 321-345.
22. Zhu B, Dong X, Yu H. MILP-based Differential Attack on Round-reduced GIFT.Cryptology ePrint Archive, Report 2018/390, 2018.
23. Zhou C., Zhang W., Ding T., et al.Improving the MILP-based Security Evaluation Algorithms against Differential Cryptanalysis Using Divide-and-Conquer Approach. Cryptology ePrint Archive, Report 2019/019, 2019.
24. Liu Y., Liang H., Li M., et al.STP Models of Optimal Differential and Linear Trail for S-box Based Ciphers. Cryptology ePrint Archive, Report 2019/025, 2019.

A The inequalities of the s-box

Table 6. The inequalities of the s-box

The GIFT s-box		The PRESENT s-box
$(x_0x_1x_2x_3y_0y_1y_2y_3c)$	$(x_0x_1x_2x_3y_0y_1y_2y_3p_0p_1p_2c)$	$(x_0x_1x_2x_3y_0y_1y_2y_3p_0p_1c)$
1 -1 -1 -1 0 -1 -1 0 4	0 0 0 0 0 0 0 0 -1 -1 -1 1	3 -2 -3 1 6 -1 -4 -3 3 7 0
-1 -2 0 -1 2 -1 -1 -1 5	0 0 0 0 0 0 0 0 0 0 0 1 0	1 2 2 0 -1 0 -1 0 -1 1 0
-1 1 -3 2 -1 0 -3 -3 8	-1 1 0 2 -1 0 -3 -3 5 8 5 0	0 0 0 1 0 0 0 1 1 -1 0
-1 1 0 -1 -1 0 -1 -1 4	1 -1 0 -2 1 0 -2 -2 5 6 5 0	1 2 -2 -4 -3 -4 -2 1 -1 12 0
-1 2 0 0 2 -1 2 1 0	-2 0 -1 -2 0 -2 1 -4 10 7 5 0	3 -3 -2 1 -4 -1 6 -3 3 7 0
-2 -1 2 -2 -1 -2 -1 2 7	4 -1 4 2 -1 1 -1 2 0 -2 -1 0	1 4 -3 2 -2 3 1 -2 2 2 0
3 1 -1 -1 2 3 1 -1 0	1 1 1 0 -1 -1 1 -1 2 1 0 0	1 -1 -1 -4 -2 4 -2 0 3 6 0
3 2 0 -1 3 2 -1 -1 0	1 1 0 2 3 1 2 2 -5 -3 -5 0	1 -1 -1 1 -1 0 -1 -1 -2 5 0
-1 -1 1 1 0 0 0 -1 2	0 2 -1 -3 1 2 -1 2 3 0 -2 0	-2 -1 -1 -2 1 2 1 2 -1 4 0
-1 -1 2 -1 -2 2 2 -1 4	-1 -2 3 -2 -3 3 3 -2 7 3 -2 0	0 -1 -1 -1 2 -2 2 -2 -1 5 0
1 3 3 2 -1 -1 2 -1 0	-1 -1 0 -1 1 -1 1 1 2 4 0 0	-1 1 1 3 2 -1 2 2 -1 -1 0
-2 2 1 2 0 0 1 1 0	-4 -1 -4 -1 -2 4 -1 3 11 6 3 0	-3 2 2 -2 1 -1 1 -4 1 6 0
2 -1 -1 2 0 0 2 1 0	0 2 1 3 2 0 1 1 -4 -2 -4 0	0 1 1 2 3 3 3 1 1 -5 0
2 1 3 2 -1 -1 -1 1 0	-3 -2 -3 4 0 1 -5 -2 11 10 12 0	1 -3 4 2 1 3 -2 -2 2 2 0
-2 -1 -2 -1 -2 2 -1 2 7	3 1 -1 -1 2 3 1 -1 0 -1 0 0	1 -2 2 -4 -2 -4 -3 1 -1 12 0
1 -1 1 0 -1 1 -1 0 2	-1 1 2 -1 -2 -2 -2 4 6 3 -1 0	3 1 1 0 2 3 2 4 2 -6 0
0 3 -1 -1 2 3 -1 3 0	-2 0 1 4 4 -1 2 -2 0 3 1 0	-4 1 -2 -2 4 -1 -3 -1 2 8 0
2 -1 1 1 1 -1 0 2 0	1 -1 -1 3 -3 0 1 0 4 1 1 0	-2 -1 -1 2 -3 0 -3 -2 -1 10 0
1 -2 0 0 2 1 2 1 0	1 -1 0 -2 4 -1 -2 -1 5 3 5 0	-4 -2 1 -2 -3 -1 4 -1 2 8 0
-3 -2 -1 2 1 -1 -2 -1 7	2 -2 -1 -1 -1 -1 -1 -1 6 7 5 0	
-1 -1 -2 -2 -1 -2 2 -1 8		

B The differential characteristics of GIFT-64

Table 7. The differential characteristic of 12-round GIFT-64

Rounds	Differential	$log_2 Pr$
input	0000 000c 0000 0006	1
1	0000 0000 0202 0000	-4
2	0000 0050 0000 0050	-8
3	0000 0000 0000 0202	-14
4	0000 0005 0000 0005	-18
5	0000 0000 0202 0000	-24
6	0000 0050 0000 0050	-28
7	0000 0000 0000 0202	-34
8	0000 0005 0000 0005	-38
9	0000 0000 0202 0000	-44
10	0000 0050 0000 0050	-48
11	0000 0000 0000 0202	-54
12	0000 0005 0000 0005	-58

Table 8. The differential characteristic of 13-round GIFT-64

Rounds	Differential	$log_2 Pr$
input	0000 000c 0000 0006	1
1	0000 0000 0202 0000	-4
2	0000 0050 0000 0050	-8
3	0000 0000 0000 0202	-14
4	0000 0005 0000 0005	-18
5	0000 0000 0202 0000	-24
6	0000 0050 0000 0050	-28
7	0000 0000 0000 0202	-34
8	0000 0005 0000 0005	-38
9	0000 0000 0202 0000	-44
10	0000 0050 0000 0050	-48
11	0000 0000 0000 0202	-54
12	0000 0005 0000 0005	-58
13	0808 0404 0202 0101	-62

C The key schedule of GIFT-128

For GIFT-128, four 16-bit words of the key state are extracted as the round key $RK = U \parallel V$. $U \leftarrow k_5 \parallel k_4$, $V \leftarrow k_1 \parallel k_0$. The key state is then updated as follows:

$$k_7 \parallel k_6 \parallel \ldots \parallel k_1 \parallel k_0 \parallel \leftarrow k_1 \ggg 2 \parallel k_0 \ggg 12 \parallel \ldots \parallel k_3 \parallel k_2$$

Where $\ggg i$ is an i bits right rotation with a 16-bit word.