Constructing Mixed-integer Programming Models whose Feasible Region is Exactly the Set of All Valid Differential Characteristics of SIMON

Siwei Sun[1,2], Lei Hu[1,2], Meiqin Wang[3], Peng Wang[1,2], Kexin Qiao[1,2],
Xiaoshuang Ma[1,2], Danping Shi[1,2], Ling Song[1,2], Kai Fu[3]

[1]State Key Laboratory of Information Security, Institute of Information Engineering,
Chinese Academy of Sciences, Beijing 100093, China
[2]Data Assurance and Communication Security Research Center,
Chinese Academy of Sciences, Beijing 100093, China
[3]Key Laboratory of Cryptologic Technology and Information Security,
Ministry of Education, Shandong University, Jinan 250100, China
{sunsiwei,hulei}@iie.ac.cn, mqwang@sdu.edu.cn,
{wpeng,qiaokexin,maxiaoshuang,shidanping,songling}@iie.ac.cn

Abstract. In IACR ePrint 2014/747, a method for constructing mixed-integer linear programming (MILP) models whose feasible regions are exactly the sets of all possible differential (or linear) characteristics for a wide range of block ciphers is presented. These models can be used to search for or enumerate differential and linear characteristics of a block cipher automatically. However, for the case of SIMON (a lightweight block cipher designed by the U.S. National Security Agency), the method proposed in IACR ePrint 2014/747 is not *exact* anymore. That is, the feasible region of the MILP model constructed for SIMON contains invalid differential characteristics due to the dependent input bits of the AND operations, and these invalid characteristics must be filtered out by other methods. This is a very inconvenient process and reduces the level of automation of the framework of MILP-based automatic differential analysis. In this paper, by using quadratic constraints or constraints from the H-representation of a specific convex hull, we give a method for constructing mixed-integer (non)linear programming models whose feasible regions are exactly the sets of all valid differential characteristics for SIMON. The technique presented in this paper may be also useful for other ciphers. How to construct an MILP model whose feasible region is exactly the set of all linear characteristics of SIMON is still an open problem.

Keywords: Automatic cryptanalysis, Related-key differential attack, Mixed-integer Linear Programming, Convex hull

1 Differential Behavior of the AND Operations of SIMON with Dependent Input Bits

We will focus on the case of SIMON32 [1] with block size 32 bits, for other cases the method is similar. The nonlinear layer of SIMON32 can be described by a non-linear function $F : \mathbb{F}_2^{16} \to \mathbb{F}_2^{16}$, such that

$$F(x) = (x <<< 1) \cdot (x <<< 8),\ x = (x_0, \cdots, x_{15}) \in \mathbb{F}_2^{16}.$$

where $\cdot$ is the bitwise AND operation.

Let $\Delta = (\Delta_0, \cdots, \Delta_{15}) \in \mathbb{F}_2^{16}$, and $\delta = (\delta_0, \cdots, \delta_{15}) \in \mathbb{F}_2^{16}$, then the differential $\Delta \to \delta$ is a valid differential pattern for F if and only if there exists $x \in \mathbb{F}_2^{16}$ such that

$$F(x) + F(x + \Delta) = \delta$$

That is

$$((x + \Delta) <<< 1) \cdot ((x + \Delta) <<< 8) = \delta$$

Write it bitwisely, we claim that the differential $\Delta \to \delta$ is valid if and only if the following system of equations of x_i has a solution

$$\begin{cases} \delta_0 = \Delta_1 \cdot x_8 + \Delta_8 \cdot x_1 \\ \delta_1 = \Delta_2 \cdot x_9 + \Delta_9 \cdot x_2 \\ \delta_2 = \Delta_3 \cdot x_{10} + \Delta_{10} \cdot x_3 \\ \delta_3 = \Delta_4 \cdot x_{11} + \Delta_{11} \cdot x_4 \\ \delta_4 = \Delta_5 \cdot x_{12} + \Delta_{12} \cdot x_5 \\ \delta_5 = \Delta_6 \cdot x_{13} + \Delta_{13} \cdot x_6 \\ \delta_6 = \Delta_7 \cdot x_{14} + \Delta_{14} \cdot x_7 \\ \delta_7 = \Delta_8 \cdot x_{15} + \Delta_{15} \cdot x_8 \\ \delta_8 = \Delta_9 \cdot x_0 + \Delta_0 \cdot x_9 \\ \delta_9 = \Delta_{10} \cdot x_1 + \Delta_1 \cdot x_{10} \\ \delta_{10} = \Delta_{11} \cdot x_2 + \Delta_2 \cdot x_{11} \\ \delta_{11} = \Delta_{12} \cdot x_3 + \Delta_3 \cdot x_{12} \\ \delta_{12} = \Delta_{13} \cdot x_4 + \Delta_4 \cdot x_{13} \\ \delta_{13} = \Delta_{14} \cdot x_5 + \Delta_5 \cdot x_{14} \\ \delta_{14} = \Delta_{15} \cdot x_6 + \Delta_6 \cdot x_{15} \\ \delta_{15} = \Delta_0 \cdot x_7 + \Delta_7 \cdot x_0 \end{cases} \tag{1}$$

2 Constructing Exact Mixed-integer Programming Models for SIMON

In the work of IACR ePrint 2014/747 [2], the MILP models generated for SIMON only have variables for the differences (δ_i and Δ_i) and the active AND operations. To generate exact models for SIMON, we need to introduce a new set of variables ($x_i, 0 \leq i \leq 15$) for every round of SIMON, and include the constraints listed in (1) into the model. However, the equations in (1) are quadratic constraints and we do not know how to solve the resulting quadratic integer programming model. In the following, we show how to convert it into a set of linear constraints by the convex hull computation technique.

Taking the first equation $\delta_0 = \varDelta_1 \cdot x_8 + \varDelta_8 \cdot x_1$ in (2) for example, let $Sol(\delta_0 = \varDelta_1 \cdot x_8 + \varDelta_8 \cdot x_1)$ be the set of all 0-1 solutions for this equation. Then

$$Sol(\delta_0 = \varDelta_1 \cdot x_8 + \varDelta_8 \cdot x_1)$$

can be treated as a subset of $\{0,1\}^5 \in \mathbb{R}^5$. The vectors $(\delta_0, \varDelta_1, x_8, \varDelta_8, x_1)$ in $Sol(\delta_0 = \varDelta_1 \cdot x_8 + \varDelta_8 \cdot x_1)$ are given below

```
[ 0, 0, 0, 0, 0 ]
[ 0, 0, 0, 0, 1 ]
[ 0, 0, 0, 1, 0 ]
[ 0, 0, 1, 0, 0 ]
[ 0, 0, 1, 0, 1 ]
[ 0, 0, 1, 1, 0 ]
[ 0, 1, 0, 0, 0 ]
[ 0, 1, 0, 0, 1 ]
[ 0, 1, 0, 1, 0 ]
[ 0, 1, 1, 1, 1 ]
[ 1, 0, 0, 1, 1 ]
[ 1, 0, 1, 1, 1 ]
[ 1, 1, 0, 1, 1 ]
[ 1, 1, 1, 0, 0 ]
[ 1, 1, 1, 0, 1 ]
[ 1, 1, 1, 1, 0 ]
```

Now, we can compute the critical set $\mathcal{O}$ of the H-representation of the convex hull of $Sol(\delta_0 = \varDelta_1 \cdot x_8 + \varDelta_8 \cdot x_1)$. The H-representation of the convex hull is given below

```
An inequality (0, -1, 0, 0, 0) V + 1 >= 0
An inequality (0, 0, -1, 0, 0) V + 1 >= 0
An inequality (0, 0, 0, -1, 0) V + 1 >= 0
An inequality (-1, 1, 0, 0, 1) V + 0 >= 0
An inequality (-1, 0, 1, 0, 1) V + 0 >= 0
An inequality (-1, 0, 0, 0, 0) V + 1 >= 0
An inequality (0, 0, 0, 0, 1) V + 0 >= 0
An inequality (1, -1, -1, 0, 1) V + 1 >= 0
An inequality (0, 1, 0, 0, 0) V + 0 >= 0
An inequality (0, 0, 0, 0, -1) V + 1 >= 0
An inequality (-1, 1, 0, 1, 0) V + 0 >= 0
An inequality (1, 0, 0, 0, 0) V + 0 >= 0
An inequality (-1, 0, 1, 1, 0) V + 0 >= 0
An inequality (0, 0, 0, 1, 0) V + 0 >= 0
An inequality (1, 0, 1, -1, -1) V + 1 >= 0
An inequality (0, 0, 1, 0, 0) V + 0 >= 0
An inequality (1, -1, -1, 1, 0) V + 1 >= 0
An inequality (1, 1, 0, -1, -1) V + 1 >= 0
An inequality (-1, -1, -1, -1, -1) V + 4 >= 0
```

Where $V = (\delta_0, \varDelta_1, x_8, \varDelta_8, x_1)^T$, and the method for computing the critical set is presented in IACR ePrint 2014/747 [2]. The critical set $\mathcal{O}$ is

```
(-1, 0, 1, 0, 1, 0)
(-1, 1, 0, 1, 0, 0)
(1, -1, -1, 1, 0, 1)
(1, 1, 0, -1, -1, 1)
(1, 0, 1, -1, -1, 1)
(-1, -1, -1, -1, -1, 4)
(-1, 1, 0, 0, 1, 0)
(-1, 0, 1, 1, 0, 0)
(1, -1, -1, 0, 1, 1)
```

$\mathcal{O}$ is a set of 9 linear inequalities involving the 5 variables: $\delta_0, \varDelta_1, x_8, \varDelta_8, x_1$. Then we can add this set of linear constraints to the overall MILP model. Note that such linear constraints must be added into the model for every equation in (1). Now, we come to an MILP model for SIMON whose feasible region is exactly the set of all differential characteristics for SIMON.

3 How to Use This Technique in Practice

Compared with the models generated in IACR ePrint 2014/747, the new models contains more variables and constraints which will make it more difficult to solve. So, we suggest that we should first try to find a good differential characteristic with input difference α and output difference β by the method presented in IACR ePrint 2014/747. According to our experimental experience, we will get a valid characteristic with a very high chance. Then when we want to enumerate the characteristics in the differential $\alpha \rightarrow \beta$, we fix the variables in the MILP model according to the the input and output differences and limit the number of active AND operations, and add the new constraints described in this paper to the MILP model. Now, we can enumerate all differential characteristics of this differential with the predefined properties by finding all solutions of the MILP model. Since the number of active AND is limited, lots of variables in the model will just disappear (when it is assigned a value of 0), and the model will be not very difficult to solve.

4 Open Problem

How to construct an MILP model whose feasible region is exactly the set of all linear characteristics of SIMON.

5 Acknowledgment

We thanks Qingju Wang *et al.* and Stefan Kölbl *et al.* for the fruitful discussion, and recently, Stefan Kölbl told us that they have a method for constructing a Satisfiability Modulo Theory (SMT) model whose solution set is exactly the set of all valid characteristics, and some of their results have been presented at ESC 2015 [3].

References

1. Ray Beaulieu, Douglas Shors, Jason Smith, Stefan Treatman-Clark, Bryan Weeks and Louis Wingers. The SIMON and SPECK families of lightweight block ciphers. IACR Cryptology ePrint Archive, Report 2013/404, 2013. http://eprint.iacr.org/2013/404.
2. Siwei Sun, Lei Hu, Meiqin Wang, Peng Wang, Kexin Qiao, Xiaoshuang Ma, Danping Shi, Ling Song, Kai Fu. 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. http://eprint.iacr.org/2014/747.
3. Gregor Leander, Stefan Kölbl, Tyge Tiessen. Insights in the Simon Round Function. https://www.cryptolux.org/mediawiki-esc2015/images/9/92/Simon_esc2015.pdf.