SCIENCE CHINA
Information Sciences

• LETTER •

March 2017, Vol. 60 039101:1–039101:3
doi: 10.1007/s11432-015-0007-1

Improved linear (hull) cryptanalysis of round-reduced versions of SIMON

Danping SHI[1,2,3], Lei HU[1,2*], Siwei SUN[1,2], Ling SONG[1,2], Kexin QIAO[1,2] & Xiaoshuang MA[1,2]

[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]*University of Chinese Academy of Sciences, Beijing* 100093*, China*

Received June 24, 2015; accepted February 3, 2016; published online May 16, 2016

Citation Shi D P, Hu L, Sun S W, et al. Improved linear (hull) cryptanalysis of round-reduced versions of SIMON. Sci China Inf Sci, 2017, 60(3): 039101, doi: 10.1007/s11432-015-0007-1

Dear editor,

SIMON is a family of lightweight block ciphers proposed by the U.S. National Security Agency in 2013 [1]. It employs a Feistel structure, and performs well in hardware and software implementations. The algorithm has attracted much cryptanalytic interest, including differential analysis [2–6] and linear cryptanalysis [7].

Linear cryptanalysis [8] is an important technique for examining block ciphers. We have thoroughly investigated the properties of linear approximations of bitwise AND operations with dependent input bits. By using mixed-integer linear programming [4, 5] to automatically search for characteristics, we obtained improved linear characteristics and linear hulls for several versions of the SIMON family. Moreover, we present an improved linear hull analysis of some versions, which, to the best of our knowledge, represent the best linear cryptanalysis results for SIMON published so far. These results are summarized in Table 1.

Linear approximation of bitwise AND. The SIMON round function is $L^{r+1} = R^r \oplus K^r \oplus (L^r \lll 2) \oplus ((L^r \lll 1) \wedge (L^r \lll 8))$, $R^{r+1} = L^r$, where L^r is the left half of the n-bit input for the rth round, R^r denotes the right half, K^r is the subkey for the rth round, $\lll i$ is a left-circular shift by i bits, $\oplus$ denotes bitwise XOR, and $\wedge$ denotes bitwise AND.

We denote the nonlinear layer in the round function of SIMON by $f^N(L^r) = (L^r \lll 1) \wedge (L^r \lll 8)$. Let $X[j]$ be the $(j \mod n)$-th bit of X, where $X[1]$ is the most significant bit of X. Regarding each bitwise AND as a 2×1 S-box, the function f^N is composed of n 2×1 S-boxes with inputs $L^r[j+1]$ and $L^r[j+8]$:

$$f_j^N(L^r[j+1], L^r[j+8]) = L^r[j+1] \wedge L^r[j+8].$$

Suppose $g(x)$ is a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$, then $\alpha \cdot x \oplus \beta \cdot g(x)$ is one approximation, where $\cdot$ is inner product, $\alpha \in \mathbb{F}_2^n, \beta \in \mathbb{F}_2^m$ are input and output masks. The correlation of a boolean function f is defined by $|\Pr(f=0) - \Pr(f=1)|$. It can easily be calculated that the correlation of a linear approximation of a 2×1 S-box is 2^{-1} with a nonzero output mask, 0 with a nonzero input mask and zero output mask, and 1 with the all-zero input and output. Thus, the number of active S-boxes

* Corresponding author (email: hu@is.ac.cn)

The authors declare that they have no conflict of interest.

Table 1 Summary of linear cryptanalysis on SIMON

Version	#Round a)	Potential	Reference
SIMON32	13	$2^{-31.69}$	[7]
	13	$2^{-28.99}$	This paper
SIMON48	15	$2^{-44.11}$	[7]
	15	$2^{-42.28}$	This paper
SIMON64	21	$2^{-62.53}$	[7]
	21	$2^{-60.72}$	This paper
	22	$2^{-63.83}$	This paper

a) #Round: number of rounds for linear hulls.

is the sum of the Hamming weights of the output masks of the S-box layers for SIMON, where an S-box is considered to be active if it has a nonzero output mask.

Dependence of S-boxes. Let Y^r be the output value of the nonlinear function f^N in round r, and let I_1^r and I_8^r be the input masks and O^r the output mask. Suppose two S-boxes f_j^N and f_{j+7}^N are active in one round. We thus have the linear approximation $I_1^r[j]\cdot L^r[j+1]\oplus I_8^r[j]\cdot L^r[j+8]\oplus Y^r[j]\oplus I_1^r[j+7]\cdot L^r[j+8]\oplus I_8^r[j+7]\cdot L^r[j+15]\oplus Y^r[j+7]$ with correlation 0 or $\pm 2^{-1}$, instead of the $\pm 2^{-2}$ by the piling-up lemma [8]. The application of the piling-up lemma is invalid because of the dependence of the two S-boxes, namely, both take $L^r[j+8]$ as input. In the following, we scrutinize the relationship between the input variables of active S-boxes and the correlation of the corresponding approximation.

Consider a Boolean function $f(x_1, x_2, \ldots, x_n) = L_x(x_1, x_2, \ldots, x_n) + B_x(x_1, x_2, \ldots, x_n)$ from $\mathbb{F}_2^n$ to $\mathbb{F}_2$, where $L_x(x_1, x_2, \ldots, x_n)$ is a linear combination of the first-degree terms and $B_x(x_1, x_2, \ldots, x_n)$ is the sum of quadratic terms $x_i x_j$ with $x_i, x_j \in \mathbb{F}_2$. A standard quadratic form $g(y_1, y_2, \ldots, y_n) = L_y(y_1, y_2, \ldots, y_n) + B_y(y_1, y_2, \ldots, y_n)$, retaining the same correlation, is obtained from $f(x_1, x_2, \ldots, x_n)$ by a nonsingular linear transformation $y = A \times x$, where $L_y(y_1, y_2, \ldots, y_n) = y_{j_1} + y_{j_2} + \cdots + y_{j_t}$ and $B_y(y_1, y_2, \ldots, y_n) = y_{i_1}y_{i_2} + y_{i_3}y_{i_4} + \cdots + y_{i_{2s-1}}y_{i_{2s}}$, with all subscripts $i_1, i_2, \ldots, i_{2s}$ having different values. The correlation of the standard quadratic form g is 0 if $\{j_1, j_2, \ldots, j_t\} \setminus \{i_1, i_2, \ldots, i_{2s}\}$ is nonempty or 2^{-s} if $\{j_1, j_2, \ldots, j_t\} \subseteq \{i_1, i_2, \ldots, i_{2s}\}$. In the latter case, therefore, having fewer variables in quadratic terms results in a greater correlation.

The linear approximation to the round function of SIMON is a quadratic function, denoted by G. Suppose $G(L^r[1], L^r[2], \ldots, L^r[n]) = L_G(L^r[1], \ldots, L^r[n]) + B_G(L^r[1], \ldots, L^r[n])$, where $L_G(L^r[1], \ldots, L^r[n]) = \sum_{j=1}^{n}(I_1^r[j]\cdot L^r[j+1] + I_8^r[j]\cdot L^r[j+8])$ and $B_G(L^r[1], \ldots, L^r[n]) = \sum_{j=1}^{n} O^r[j](L^r[j+1]\cdot L^r[j+8])$, and its correlation should be calculated following the above rules.

MILP modeling. Mixed-integer linear programming (MILP) modeling was investigated in [4] and further extended to automatically searching for linear characteristics and linear hulls in [5]. The modeling denotes each mask bit as a 0-1 variable and describes their propagation through the cipher as linear inequalities (constraints), which yields an optimized value of the number of active S-boxes. Specifically, the MILP modeling is as follows.

(1) Constraints for linear operations. Three constraints for linear operations can be directly obtained from [5, 9]:

(i) For each bitwise XOR operation, (α_1, α_2) and β denote the input masks and output mask of $\oplus$. The constraints on these mask bits are

$$\alpha_1 = \alpha_2 = \alpha_3. \tag{1}$$

(ii) For each branching in the cipher structure, let $(\alpha_1, \alpha_2, \alpha_3)$ denote the masks on three branches. The constraints on the masks are

$$\begin{cases} \tau \geqslant \alpha_1,\ \tau \geqslant \alpha_2,\ \tau \geqslant \alpha_3\,; \\ \alpha_1 + \alpha_2 + \alpha_3 \geqslant 2\tau\,; \\ \alpha_1 + \alpha_2 + \alpha_3 \leqslant 2\,, \end{cases} \tag{2}$$

where τ is a dummy variable.

(iii) For a left-circular shift by i bits, let $\mu = (\mu[1], \mu[2], \ldots, \mu[n])$ and $\nu = (\nu[1], \nu[2], \ldots, \nu[n])$ be the input and output masks. The constraints are

$$\nu[j] = \mu[j+i],\ j \in \{1, 2, \ldots, n\}\,. \tag{3}$$

(2) Constraints for S-boxes. The SIMON S-box is a bitwise AND, and to obtain valid linear characteristics, we only allow active S-boxes with nonzero output masks and inactive S-boxes with all-zero masks. This rule can be described as

$$\begin{cases} O^r[j] \geqslant I_1^r[j]\,, \\ O^r[j] \geqslant I_8^r[j]\,, \end{cases} \tag{4}$$

where the symbols are as defined above.

(3) Constraints dealing with dependence of S-boxes. Given the masks (I_1^r, I_8^r) and O^r for the nonlinear function. Suppose fewer variables in the quadratic terms of a function lead to fewer variables in the quadratic terms of its standard quadratic form. Thus fewer variables in the quadratic terms of approximation function G lead to a greater correlation if the correlation is nonzero. To indicate the number of variables in quadratic terms of the linear approximation of f^N, n new 0-1 variables $V^r[j]$ $(j \in \{1, 2, \ldots, n\})$ indicating whether $L^r[j]$ exists in the quadratic terms

are introduced in each round. The constraints are

$$\begin{cases} V^r[j] \geqslant O^r[j-1], \\ V^r[j] \geqslant O^r[j-8], \\ V^r[j] \leqslant O^r[j-1] + O^r[j-8]. \end{cases} \tag{5}$$

Thus, $\sum_{j=1}^{n} V^r[j]$ is the number of variables appearing in quadratic terms in the linear approximation for one round and $\sum_r \sum_{j=1}^{n} V^r[j]$ is the number of variables appearing in the overall linear approximation.

(4) Objective function. To obtain the minimum number of linearly active S-boxes, the objective function is set to minimize the sum of all output mask bits of S-box layers as in [5]. However, it is the correlation of the linear characteristics that determines the effectiveness of the linear cryptanalysis. Therefore, considering the influence of the dependence of active S-boxes on correlation, our objective function is one that minimizes $\sum_r \sum_{j=1}^{n} V^r[j]$.

With the modeling defined above, we sought out better linear characteristics and linear hulls for SIMON with Gurobi solver. The linear characteristics found may have a correlation of 0 due to the dependence of active S-boxes. It is therefore imperative to test for this. For each linear characteristic, we obtain an accurate absolute value of linear approximation's correlation for each round by use of the nonsingular transform method described above.

(5) Constraints for linear hulls. To obtain linear hulls, we set the input and output mask bits to that of one obtained linear characteristics as added constraints to above modeling and then seek as many as possible others that form a linear hull. The details can be found in the supporting materials.

Conclusion. We have considered the interdependence of S-boxes by evaluating correlations in a linear approximation. All of the results from the analysis using our method represent, to the best of our knowledge, an improvement over those in the literature. Moreover, the 21- and 22-round linear hulls for SIMON64 can be used to mount the best attack that we are aware of. The specific results are provided in the supporting materials.

Acknowledgements This work was supported by National Basic Research Program of China (Grant No. 2013CB834203), National Natural Science Foundation of China (Grant Nos. 61472417, 61472415, 61402469), Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA06010702), and State Key Laboratory of Information Security, Chinese Academy of Sciences.

Supporting information The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Beaulieu R, Shors D, Smith J, et al. The SIMON and SPECK families of lightweight block ciphers. IACR Cryptology ePrint Archive, Report 2013/404. http://eprint.iacr.org/
2 Abed F, List E, Wenzel J, et al. Differential cryptanalysis of round-reduced Simon and Speck. In: Fast Software Encryption. Berlin: Springer, 2014. 525–545
3 Biryukov A, Roy A, Velichkov V. Differential analysis of block ciphers SIMON and SPECK. In: Fast Software Encryption. Berlin: Springer, 2014. 546–570
4 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. In: Advances in Cryptology — ASIACRYPT 2014. Berlin: Springer, 2014. 158–178
5 Sun S W, Hu L, Wang M Q, et al. Automatic enumeration of (related-key) differential and linear characteristics with predefined properties and its applications. IACR Cryptology ePrint Archive, Report 2014/747. http://eprint.iacr.org/
6 Wang N, Wang X, Jia K, et al. Improved differential attacks on reduced SIMON versions. IACR Cryptology ePrint Archive, Report 2014/448. http://eprint.iacr.org/
7 Alizadeh J, Alkhzaimi H A, Aref M R, et al. Improved linear cryptanalysis of round reduced SIMON. IACR Cryptology ePrint Archive, Reprot 2014/681. http://eprint.iacr.org/
8 Matsui M. Linear cryptanalysis method for DES cipher. In: Advances in Cryptology — EUROCRYPT'93. Berlin: Springer, 1994. 386–397
9 Bogdanov A, Rijmen V. Linear hulls with correlation zero and linear cryptanalysis of block ciphers. Design Code Cryptogr, 2014, 70: 369–383