Improved Linear Cryptanalysis of SMS4 Block Cipher

Joo Yeon Cho and Kaisa Nyberg

Nokia and Aalto University

SKEW 2010, Feb. 17, 2011

Outline

1. Multidimensional Linear Attack: Algorithm Aspect

2. Cryptanalysis of SMS4 Block Cipher: Approach and Results

Multidimensional Linear Attack Algorithm 1

Step 1 Choose a certain number (say, m) of linearly independent approximations.

$$U_i \cdot P \oplus V_i \cdot C = W_i \cdot K, \quad 0 \leq i \leq m-1$$

where U_i, V_i and W_i denote linear masks.

Step 2 Generate $2^m - 1$ linear approximations by combining m approximations. Their correlations are denoted as $c_1, \cdots, c_{2^m-1}$. The capacity $\sum_i c_i^2$ is expected to be high.

Multidimensional Linear Attack Algorithm 1

Step 3 Suppose $G = (g_0, \ldots, g_{m-1})$ where $g_i = W_i \cdot K$. For each value of G, create its probability distribution

$$p_G = (p_{0,G}, \ldots, p_{2^m-1,G})$$

where

$$p_{i,G} = 2^{-m} \sum_{j=0}^{2^m-1} (-1)^{j\cdot(i \oplus G)} c_j$$

Step 4 Measure the frequency of the vectors $(g_0, \ldots, g_{m-1})$ where $g_i = U_i \cdot P \oplus V_i \cdot C$. Obtain the empirical probability distribution $q_K = (q_{0,K}, \ldots, q_{2^m-1,K})$. K is unknown.

Multidimensional Linear Attack Algorithm 1

Step 6 Compute the log-likelihood ratio (LLR) between p_G and q_K

$$LLR(p_G, q_K) = \sum_{i=0}^{2^m-1} q_{i,K} \log p_{i,G} + m.$$

where $u = (u_0, \ldots, u_{2^m-1})$ is the uniform distribution.

Step 7 Choose the G such that $\max_G LLR(p_G, q_K)$ as the right key.

Multidimensional Linear Attack Algorithm 2

1. Suppose l is the length of the guessed key. Measure $q_\kappa = (q_{\kappa,0}, \ldots, q_{\kappa,2^m-1})$ for $\kappa \in [0, 2^l - 1]$.
2. Choose κ and G such that $\max_\kappa \max_G LLR(p_G, q_\kappa)$ as the right key values.
3. Recover $(l + m)$ bits information of the secret key.

Convolution Method: Reducing Time Complexity

- It was proposed at CT-RSA 2010 by Hermelin and Nyberg.
- Instead of using *LLR*-statistics, the statistical decision can be equivalently made by computing

$$D_G = \sum_{i=0}^{2^m-1} (-1)^{i \oplus G} \hat{c_i} \times c_i$$

where $\hat{c_0}, \ldots, \hat{c_{2^m-1}}$ are the empirically measured correlations of $2^m - 1$ linear approximations.

Convolution Method: Reducing Time Complexity

- The *LLR*-statistic requires around $2^m \cdot 2^m$ on-line computation efforts.
- Convolution method requres $m \times 2^m$ operations by FFT algorithm. The correct key is recovered by choosing G such that D_G is maximal.
- We can further reduce the complexity by choosing only $M(< 2^m - 1)$ significant correlations.

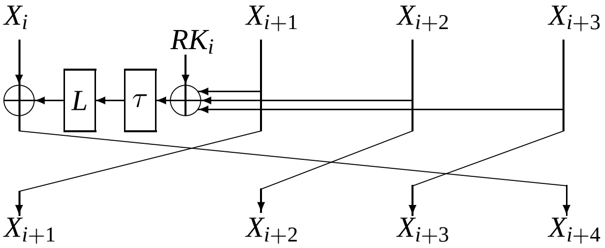
SMS4

SMS4 is

- a Chinese block cipher designed for Wireless LAN WAPI (Wired Authentication and Privacy Infrastructure).
- a generalized Feistel block cipher taking 128-bit input, 128-bit output and 128-bit key.
- is composed of 32 rounds.

Detailed specification is available at IACR ePrint Archive.

Round Function of SMS4



$$X_{i+4} = X_i \oplus L(\tau(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus RK_i)), \quad X_i, RK_i \in \mathbb{F}_2^{32}$$

Round Function

1. Let S denote the 8×8 S-box of SMS4. The non-linear transformation τ is defined as

$$\tau(A) = S(a_0)||S(a_1)||S(a_2)||S(a_3)$$

   where $||$ stands for the concatenation.

2. The linear transformation L is defined as

$$L(X) = X \oplus (X \lll 2) \oplus (X \lll 10) \oplus (X \lll 18) \oplus (X \lll 24)$$

   where $X \lll n$ denotes the left-rotated X by n-bit.

5-Round Characteristic

1. Let $\gamma \in \mathbb{F}_2^{32}$ be a linear mask.
2. Get two rounds linear approximations

$$\gamma \cdot X_{i+4} = \gamma \cdot (X_i \oplus X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus RK_i)$$

   and

$$\gamma \cdot X_{i+5} = \gamma \cdot (X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus X_{i+4} \oplus RK_{i+1}).$$

3. By adding two approximations, we get

$$\gamma \cdot (X_i \oplus X_{i+5}) = \gamma \cdot (RK_{i+1} \oplus RK_i)$$

   with the correlation of $\rho^2(\gamma, \gamma)$.

18-Round Characteristic

1. Add three consecutive 5-round characteristics:

   $$\gamma \cdot X_5 \oplus \gamma \cdot X_{20} = \gamma \cdot (RK_5 \oplus RK_6 \oplus RK_{10} \oplus RK_{11} \oplus RK_{15} \oplus RK_{16})$$

   with the correlation of $\rho^6(\gamma, \gamma)$.
2. This is a 18-round characteristic from Round 3 to Round 20

   $$(X_2, X_3, X_4, X_5) \to (X_{20}, X_{21}, X_{22}, X_{23})$$

Best Linear Approximations

There are 24 linear approximations holding with the highest correlations of $2^{-9.19}$.

set	γ	$L_2(\gamma)$	set	γ	$L_2(\gamma)$
$\mathcal{A}_0$	0x0011ffba	0x0084be2f	$\mathcal{A}_1$	0xba0011ff	0x2f0084be
	0x007905e1	0x005afbc6		0xe1007905	0xc6005afb
	0x00edca7c	0x0083ffaa		0x7c00edca	0xaa0083ff
	0x007852b3	0x00582b15		0xb3007852	0x1500582b
	0x00a1b433	0x00f1027a		0x3300a1b4	0x7a00f102
	0x00fa7099	0x00d20b1d		0x9900fa70	0x1d00d20b
$\mathcal{A}_2$	0xffba0011	0xbe2f0084	$\mathcal{A}_3$	0x11ffba00	0x84be2f00
	0x05e10079	0xfbc6005a		0x7905e100	0x5afbc600
	0xca7c00ed	0xffaa0083		0xedca7c00	0x83ffaa00
	0x52b30078	0x2b150058		0x7852b300	0x582b1500
	0xb43300a1	0x027a00f1		0xa1b43300	0xf1027a00
	0x709900fa	0x0b1d00d2		0xfa709900	0xd20b1d00

Mapping L_2

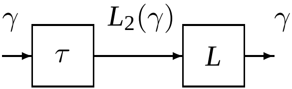
- The mapping L_2 is defined to satisfy the following equation:

$$\gamma \cdot L(x) = L_2(\gamma) \cdot x$$

  for $x \in GF(2^{32})$.

- Linear approximation of the round function is

$$\begin{aligned}
\gamma \cdot (X_{i+4} \oplus X_i) &= \gamma \cdot L(\tau(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus RK_i)) \\
&= L_2(\gamma) \cdot \tau(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus RK_i) \\
&\approx \gamma \cdot (X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus RK_i)
\end{aligned}$$



Our Observations

- Let $\mathcal{A}_0$ be a set of linear masks which is defined as

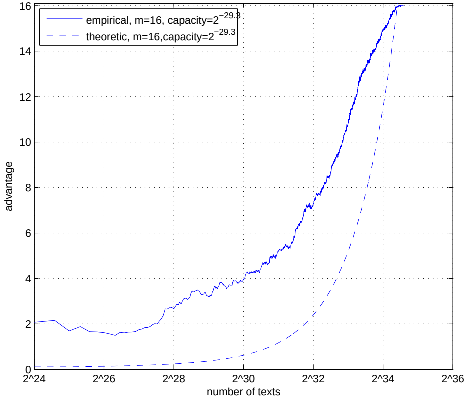
$$\mathcal{A}_0 = \{a | 0 \le a < 2^{24}, 0 \le L_2(a) < 2^{24}\}.$$

- There are 52744 non-zero linear approximations in $\mathcal{A}_0$.
- All the non-zero approximations can be generated by using 16 independent approximations.
- The capacity of those probability distribution is around $2^{-29.3}$. Note that the square of correlation of the strongest approximation is $2^{-36.76}$.

Experiments on 5-Round Characteristic

The data complexity for MA1 is calculated as

$$N_{MA1} = \frac{(\Phi^{-1}(P_S) + \Phi^{-1}(1 - 2^{-a}))^2}{Capacity}$$



20-Round Linear Characteristic

- Re-use 18-round characteristic from Round 5 to Round 22:

$$(X_4, X_5, X_6, X_7) \rightarrow (X_{22}, X_{23}, X_{24}, X_{25})$$

- Add 2-round linear characteristic from Round 3 to Round 4 with linear masks α, β.

$$\alpha \cdot X_2 \oplus \beta \cdot (X_3 \oplus X_4 \oplus X_5 \oplus RK_2) = \alpha \cdot X_6$$
$$\gamma \cdot X_3 \oplus \alpha \cdot (X_4 \oplus X_5 \oplus X_6 \oplus RK_3) = \gamma \cdot X_7$$

and the correlation is $\rho(\beta, \alpha)\rho(\alpha, \gamma)$.

- By combining two approximations, we get

$$\alpha \cdot X_2 \oplus (\beta \oplus \gamma) \cdot X_3 \oplus (\alpha \oplus \beta) \cdot (X_4 \oplus X_5) \oplus \gamma \cdot X_{22}$$
$$= \beta \cdot RK_2 \oplus \alpha \cdot RK_3 \oplus \gamma \cdot (RK_7 \oplus RK_8 \oplus RK_{12} \oplus RK_{13} \oplus RK_{17} \oplus RK_{18})$$

with the correlation of $\rho(\beta, \alpha)\rho(\alpha, \gamma)\rho^6(\gamma, \gamma)$.

Evaluation of $\rho(\gamma, \gamma)$

Suppose $\gamma \in \mathcal{A}_0$ and $0 \leq \alpha < 2^{24}$.

$\lvert\rho(\gamma,\gamma)\rvert$	Number of approx.	$\lvert\rho(\alpha,\gamma)\rvert$	Number of approx.
$2^{-9.19}$	6	$2^{-9.0}$	125
$2^{-9.39}$	11	$2^{-9.10}$	0
$2^{-9.42}$	15	$2^{-9.20}$	1200
$2^{-9.58}$	12	$2^{-9.30}$	0
$2^{-9.61}$	76	$2^{-9.40}$	6540
$2^{-9.68}$	7	$2^{-9.50}$	0
$2^{-9.80}$	120	$2^{-9.60}$	21376
$2^{-9.83}$	89	$2^{-9.70}$	1800
$2^{-9.87}$	56	$2^{-9.80}$	47088

Target key

- Since the most significant 8 bits of γ are zero and $0 \leq L_2(\gamma) < 2^{24}$, it is sufficient to guess the lower 24 bits for RK_{22}.
- Since $0 \leq \alpha < 2^{24}$ and $0 \leq L_2(\alpha) < 2^{32}$, we need to guess 32 bits of RK_0 and RK_1.
- Hence, the target key length is $32 \cdot 2 + 24 = 88$ bits.

Probability Distribution and Capacity

- Let us define $\mathcal{M}$ as

$$\mathcal{M} = \{(\alpha, \beta) \mid (\rho(\beta, \alpha)\rho(\alpha, \gamma))^2 > \delta\}.$$

where δ denote a threshold value.

- The capacity of the probability distribution is calculated as

$$C_p = \sum_{\gamma \in \mathcal{A}_0} C_{\mathcal{M}}(\gamma)$$

where

$$C_{\mathcal{M}}(\gamma) = \sum_{(\alpha,\beta) \in \mathcal{M}} \rho^2(\beta, \alpha)\rho^2(\alpha, \gamma)\rho^{12}(\gamma, \gamma).$$

Evaluation of the number of linear approximations and capacity

- We chose $m = 34$ and $M = 2^{24.7}$.
- Then, the capacity of the 20-round characteristic is $C_p = 2^{-119.7}$.
- The data complexity required for the full advantage ($a = 88$) of the attack is around $N_{MA2} = (88 + 34)/2^{-119.7} = 2^{126.6}$ with $Ps = 0.95$.

δ	M	C_p
$2^{-36.0}$	$125 = 2^{7.0}$	$2^{-135.6}$
$2^{-36.4}$	$2075 = 2^{11.0}$	$2^{-131.9}$
$2^{-36.8}$	$14615 = 2^{13.8}$	$2^{-129.5}$
$2^{-37.2}$	$62476 = 2^{15.9}$	$2^{-127.7}$
$2^{-37.6}$	$211462 = 2^{17.7}$	$2^{-126.2}$
$2^{-38.0}$	$1696134 = 2^{20.7}$	$2^{-123.0}$
$2^{-38.4}$	$4249383 = 2^{22.0}$	$2^{-122.0}$
$2^{-38.8}$	$10655129 = 2^{23.4}$	$2^{-121.3}$
$2^{-39.2}$	$31530029 = 2^{24.7}$	$2^{-119.7}$
$2^{-39.6}$	$75192630 = 2^{26.2}$	$2^{-119.0}$

Comparison of data and time complexity of the attacks against reduced-round SMS4

round	data	time	memory	method
22	$2^{118.4}$	2^{117}	2^{112}	Linear
22	2^{117}	$2^{112.3}$	2^{110}	Differential
23	$2^{126.6}$	$2^{127.4}$	$2^{120.7}$	MultiDim. Linear (this paper)

Conclusion and Future Work

1. We showed how the multidimensional linear cryptanalysis could improve the previous linear attack on the reduced version of SMS4.
2. We also demonstrated that the convolution method could reduce the time complexity of multidimensional linear attack.
3. $m = 34$ is still not optimal. It might be reduced.

Thank you for your attention