A Note on Key Ranking for Optimal Collision Side-Channel Attacks

Cezary Glowacz

DT Security
cezary.glowacz@t-systems.com

Abstract. In [1] we studied collision side-channel attacks and derived an optimal distinguisher for key ranking. In this note we show that the problem of key ranking using this distinguisher is NP-hard and we provide estimates of lower bounds for secret key ranks in collision side-channel attacks.

Keywords: Collision Side-Channel Attacks · Key Ranking · Computational Complexity · Lower Bounds

1 Introduction

Side-channel attacks exploit measurable leakage signals produced by an underlying hardware platform during execution of cryptographic functions. Given an adequate stochastic model of the signals optimal strategies for key recovery can be derived. In [1] we studied collision side-channel attacks based on an optimal distinguisher for a key. The distinguisher is a statistic based on the measured values of leakage signals which allows to decide for any two key candidates the order of their a posteriori probabilities conditioned on the measured values, and it can be used to list the key candidates in a descending order. The attacker can use the list and try each key candidate starting with the first one for minimizing the expected number of trials until the secret key has been found. In security evaluations it is sufficient to know the position of the secret key, i.e. its rank, in the list even without actually creating the list. This allows to rate the best case effort needed to find the secret key. We show that the problem of key ranking using the optimal distinguisher is NP-hard and we provide estimates of lower bounds for secret key ranks in collision side-channel attacks.

The distinguisher $D_{opt.fun.gauss}$ and its objective function $D(k,x)$ which were derived in [1] for optimal collision side-channel attacks assuming a Gaussian noise and Gaussian leakage function values are restated in the following equations.

$$D_{opt.fun.gauss} = \underset{k \in (\mathbb{F}_2^n)^L}{\operatorname{argmax}} D(k,x)$$

$$D(k,x) = \sum_{q \in \mathbb{F}_2^n} (\sum_{l=1}^{L} x^{(q \oplus k^{(l)})(l)})^2 = \sum_{i=1}^{L} \sum_{j=1}^{L} D_{i,j}(k^{(i)} \oplus k^{(j)}, x)$$

$$D_{i,j}(d,x) = \sum_{q \in \mathbb{F}_2^n} x^{(q)(i)} x^{(q \oplus d)(j)}$$

The component $x^{(q)(l)}$ of $x \in (\mathbb{R}^{2^n})^L$ represents the measured value of a leakage signal during the calculation of the l-th out of L n-bit S-Boxes with the input data $q \in \mathbb{F}_2^n$. It is assumed that the actual input to the l-th S-Box is $q \oplus k^{\star(l)}$, where $k^{\star(l)} \in \mathbb{F}_2^n$ is the l-th sub-key of the secret key $k^\star \in (\mathbb{F}_2^n)^L$. The objective function $D(k,x)$ provides for each key candidate $k \in (\mathbb{F}_2^n)^L$ a value which is proportional to the a posteriori probability of k conditioned on the measured values x.

2 Key Ranking is NP-Hard

The key ranking problem for optimal collision side-channel attacks is stated as follows. Given the vector x of measured values of leakage signals calculate the number $R(x,r)$ of key candidates $k \in (\mathbb{F}_2^n)^L$ s.t. $D(k,x) \geq r$. Actually, in an attack the number $R(x, D(k^\star, x))$ is of interest.

We reduce the NP-complete partition problem[1] to the key ranking problem. Given a multiset instance $S = \{s_1, \ldots, s_L\} \subset \mathbb{N}$ of the partition problem we create in polynomial time an instance

$$x = ((x^{(0)(1)} = -s_1, x^{(1)(1)} = +s_1), \ldots, (x^{(0)(L)} = -s_L, x^{(1)(L)} = +s_L))$$

of the key ranking problem.

Let $\rho(0) = -1$ and $\rho(1) = 1$. Because $\forall_{k \in \mathbb{F}_2^L} D(k,x) \in \mathbb{N}$ we have $R(x,1) < 2^L$ if and only if there is a key $k \in \mathbb{F}_2^L$ for which $D(k,x) = \sum_{q \in \mathbb{F}_2} (\sum_{l=1}^{L} x^{(q \oplus k^{(l)})})^2 = 2(\sum_{l=1}^{L} \rho(k^{(l)}) s_l)^2 = 0$. It follows that $R(x,1) < 2^L$ if and only if the partition problem S has a solution. ■

3 Estimation of Lower Bounds for Secret Key Ranks

An estimate of a lower bound for the secret key rank can be obtained by estimating the number of keys in some key subset with ranks higher than or equal to the rank of the secret key.

Let $N_i = \{1, ..., i\} \subset \mathbb{N}$, $C_c = \{C \subseteq N_{L-1} \mid |C| = c\}$, $b \in N_{2^n - 1}$, $c \in N_{L-1}$ and $l \in N_{L-1}$.

Let $K_{l,b} \subseteq \mathbb{F}_2^n \setminus \{k^{\star(l)}\}$ denote a set s.t. $|K_{l,b}| = b$ and

$$\forall_{d \in K_{l,b}} \forall_{d' \in \mathbb{F}_2^n \setminus \{k^{\star(l)}\} \setminus K_{l,b}} D_{l,L}(d \oplus k^{\star(L)}, x) \geq D_{l,L}(d' \oplus k^{\star(L)}, x).$$

Let $S_{b,c} = \{k \in (\mathbb{F}_2^n)^L \mid \exists_{C \in C_c} \forall_{l \in N_L} (l \notin C \wedge k^{(l)} = k^{\star(l)} \vee l \in C \wedge k^{(l)} \in K_{l,b})\}$ and let $s_{b,c} = |S_{b,c}|$.[2]

Let $t_{b,c}$ denote the number of keys $k \in S_{b,c}$ s.t. $D(k,x) \geq D(k^\star, x)$.

Then $r_{b,c} = max(2^n t_{b,c}, 2^n)$ is a lower bound for the secret key rank $R(x, D(k^\star, x))$.[3]

Let $\tilde{S}_{b,c}$ denote a multiset of $m = 2^{24}$ keys obtained by uniform random sampling the set $S_{b,c}$.[4]

Let $\tilde{t}_{b,c}$ denote the number of keys $k \in \tilde{S}_{b,c}$ s.t. $D(k,x) \geq D(k^\star, x)$.

Then $\tilde{r}_{b,c} = max(2^n \frac{\tilde{t}_{b,c}}{m} s_{b,c}, 2^n)$ is an estimate of the lower bound $r_{b,c}$.

If $\frac{t_{b,c}}{s_{b,c}} < \frac{\tilde{t}_{b,c}}{4m}$ then the probability that $\tilde{t}_{b,c} \geq \hat{t} = 32$ is less than $\hat{p} = 1.3 * 10^{-10}$.[5] We set $\tilde{r}_{b,c} = 2^n$ if $\tilde{t}_{b,c} < \hat{t}$. It is now expected that $log_2(r_{b,c}) \geq log_2(\tilde{r}_{b,c}) - 2$ in a fraction $1 - \hat{p}$ of estimations.

[1] See e.g. https://en.wikipedia.org/wiki/Partition_problem

[2] $|S_{b,c}| = \binom{L-1}{c} b^c$.

[3] For each $k = (k^{(1)}, \ldots, k^{(L-1)}, k^{\star(L)}) \in S(k^\star, x, b, c)$ and for each $d \in \mathbb{F}_2^n \setminus \{(0)^n\}$ there is a key $k' = (k^{(1)}, \ldots, k^{(L-1)}, k^{\star(L)} \oplus d) \notin S(k^\star, x, b, c)$ s.t. $D(k,x) = D(k',x)$; there are $2^n - 1$ such keys k'.

[4] First an element $C \in C_c$ and then for each $l \in C$ the sub-keys $k^{(l)} \in K_{l,b}$ are selected at random.

[5] For the calculation of the bound $\hat{p} = 1 - \sum_{i=0}^{\hat{t}-1} P(\frac{\hat{t}}{4}; i)$ an approximation of the binomial $B(m, \frac{\hat{t}}{4m}; i)$ by the Poisson distribution $P(\frac{\hat{t}}{4}; i)$ was used.

4 Simulation Results

We simulated attacks for a Gaussian noise and Gaussian leakage function values. The simulation parameters were n=8, $L = 16$, $k^\star = ((0)^n)^L$ and a noise variance σ^2. In a simulated attack we first created for each $q \in \mathbb{F}_2^n$ and each $l \in N_L$ realisations φ_q and $\eta_{q,l}$ of independent standard normal random variables. Then the vector x was created by setting its components $x^{(q)(l)} = \sqrt{2}\varphi_q + \sigma\eta_{q,l}$.

In each simulated attack a lower bound $r_{b,c}$ for the secret key rank $R(x, D(k^\star, x))$ was estimated according to section 3 with preselected values for the parameters b and c.[6] Attacks were simulated $1,000$ times for each of some noise variance values[7] and empirical quantiles of lower bound estimates were calculated. The results are shown in Table 1.

$\sigma^2/18.75$	1^{st}10-quantile	1^{st}5-quantile	median
1.0	8	8	11
1.5	16	18	25
2.0	27	36	53
2.5	52	56	72
3.0	69	73	88
3.5	83	87	101

Table 1. log_2 of empirical quantiles of lower bound estimates.

References

1. Cezary Glowacz and Vincent Grosso. Optimal collision side-channel attacks. Cryptology ePrint Archive, Paper 2019/828, 2019. https://eprint.iacr.org/2019/828.

[6] For each $b \in \{2^i - 1 \mid i \in N_n\}$ and for each $c \in N_{L-1}$ the estimates $\tilde{r}_{b,c}$ were calculated according to section 3 with $m = 2^{16}$ and $\hat{t} = 3$; then the values $\text{argmax}_{b,c}\, \tilde{r}_{b,c}$ were preselected.

[7] One value is $\sigma^2 = 18.75$ for which the optimal algorithm for collision side-channel attacks has a success rate of 0.1 in the same simulation set-up (see Fig. 1, [1]).