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 propose a heuristic estimation procedure for key ranking based on this distinguisher, and provide estimates of lower bounds for secret key ranks in collision side channel attacks.

Keywords: Collision Side-Channel Attacks · Key Ranking

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 an attack on the secret key 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 propose a heuristic estimation procedure for key ranking based on the optimal distinguisher, and provide estimates of lower bounds for secret key ranks in collision side channel attacks.

2 Background

The distinguisher $D_{opt.fun.gauss}$ and its objective function $D(k)$ 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:

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

$$D(k) = \sum_{i=1}^{L} \sum_{j=1}^{L} D_{i,j}(k^{(i)} \oplus k^{(j)}), \text{ and}$$

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

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 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 $D(k)$ provides for each key candidate $k \in (\mathbb{F}_2^n)^L$ a value which is strictly increasing with the a-posteriori probability of k conditioned on the measured values x.

3 Estimation of Lower Bounds for Secret Key Ranks

An estimate of a lower bound for the secret key rank can be obtained by sampling some key subset and counting the keys in the sample with ranks lower than or equal to the rank of the secret key.

Let $N_i = \{1, \dots, i\} \subset \mathbb{N}$, and let $b \in N_{2^n-1}$, $c \in N_{L-1}$, $m, \hat{u} \in \mathbb{N}^+$ and $\hat{t} \in N_m$ denote the estimation paramaters.

Let K_l denote a set of b largest wrt. $\sum_{i=1}^{L} D_{l,i}(d \oplus k^{*(i)})$ (Variant I) or wrt. $D_{l,L}(d \oplus k^{*(L)})$ (Variant II) elements d from $\mathbb{F}_2^n \setminus \{k^{\star(l)}\}$, $C = \{N \subseteq N_{L-1} \mid |N| = c\}$, $S = \{k \in (\mathbb{F}_2^n)^L \mid \exists_{N \in C} \forall_{l \in N_L} (l \notin N \wedge k^{(l)} = k^{\star(l)} \vee l \in N \wedge k^{(l)} \in K_l)\}$ and $t = |\{k \in S \mid D(k) \geq D(k^\star)\}|$. Then $r = 2^n max(t, 1)$ is a lower bound for the secret key rank $|\{k \in (\mathbb{F}_2^n)^L \mid D(k) \geq D(k^\star\}|$.[1]

Let $\tilde{t}$ denote a number of keys k found in m samples from the set S s.t. $D(k) \geq D(k^\star)$.[2] Then $\tilde{r} = 2^n max(\frac{\tilde{t}}{m}|S|, 1)$ is an estimate of the lower bound r.

Let $B(n, p; k)$ denote the binomial distribution. Let $\gamma(n, p; k) = \sum_{i=0}^{k-1} B(n, p; i)$. When $\tilde{t} \geq \hat{t} \geq 3$ and $\hat{u} \geq 2$ then the confidence level $\gamma(m, \frac{\tilde{t}}{\hat{u}m}, \tilde{t})$ for $\frac{t}{|S|} > \frac{\tilde{t}}{\hat{u}m}$[3] is at least $\gamma(m, \frac{\hat{t}}{\hat{u}m}, \hat{t})$.[4] We set $\tilde{r} = 2^n$ when $\tilde{t} < \hat{t}$. Now the confidence level for $log_2(r) \geq log_2(\tilde{r}) - log_2(\hat{u})$ is at least $\gamma(m, \frac{\hat{t}}{\hat{u}m}, \hat{t})$. E.g. $\gamma(2^{24}, \frac{32}{4*2^{24}}, 32) = 1 - 1.3 * 10^{-10}$.

[1] The reason for the factor 2^n is: for each $k \in S$ there are $2^n - 1$ keys $k' = k \oplus d^L \notin S$ with $d \in \mathbb{F}_2^n \setminus \{(0)^n\}$ s.t. $D(k') = D(k)$.

[2] The set S is sampled uniformly at random; for each sample first $N \in C$ and then for each $l \in N$ the sub-keys $k^{(l)} \in K_l$ are selected uniformly at random.

[3] See 2.3, [2].

[4] $\frac{d}{dp}\gamma(n, p; k) = -k\binom{n}{k}p^{k-1}(1-p)^{n-k} < 0$ (see 2, [2]) and $\frac{d^2}{dp^2}\gamma(n, p; k) = k\binom{n}{k}p^{k-2}(1-p)^{n-k-1}((n-1)p - (k-1)) < 0$ for $p < \frac{k-1}{n-1}$ hence
$\gamma(n, p + \Delta_p; k+1) > \gamma(n, p; k) + \Delta_p \frac{\partial}{\partial p}\gamma(n, p + \Delta_p; k) + \binom{n}{k}(p + \Delta_p)^k(1 - (p + \Delta_p))^{n-k}$
$= \gamma(n, p; k) - \Delta_p k\binom{n}{k}(p + \Delta_p)^{k-1}(1 - (p + \Delta_p))^{n-k} + \binom{n}{k}(p + \Delta_p)^k(1 - (p + \Delta_p))^{n-k}$
$= \gamma(n, p; k) + \binom{n}{k}(p + \Delta_p)^{k-1}(1 - (p + \Delta_p))^{n-k}(p + \Delta_p(1 - k))$ for $p + \Delta_p < \frac{k-1}{n-1}$.

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 for the secret key rank was estimated[5] with selected parameters b, c[6] and with $m = 2^{24}$ and $\hat{t} = 32$. Attacks were simulated $1,000$ times for each noise variance and empirical quantiles of lower bound estimates were calculated. The results are shown in Table 1.

$\sigma^2/18.75$	1^{st} decile	2^{nd} decile	median
1.0	8[7]	8	14
1.5	22	28	38
2.0	41	47	59
2.5	56	67	84
3.0	74	83	99
3.5	85	94	108[8]

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

5 Summary

We proposed a heuristic estimation procedure for key ranking, and provided estimates of lower bounds for secret key ranks in collision side channel attacks.

References

1. Glowacz, C., Grosso, V.: Optimal collision side-channel attacks. Cryptology ePrint Archive, Paper 2019/828 (2019), https://eprint.iacr.org/2019/828
2. Scholz, F.: Confidence bounds & intervals for parameters relating to the binomial, negative binomial, poisson and hypergeometric distributions with applications to rare events (2019), https://faculty.washington.edu/fscholz/DATAFILES498B2008/ConfidenceBounds.pdf

With $n = m, p = \frac{\tilde{t}}{\hat{u}m}, \Delta_p = \frac{1}{\hat{u}m}, k = \tilde{t} > \hat{t} \geq 3$ and $\hat{u} \geq 2$ we then have $\gamma(m, \frac{\tilde{t}+1}{\hat{u}m}; \tilde{t}+1) > \gamma(m, \frac{\tilde{t}}{\hat{u}m}; \tilde{t}) > \cdots > \gamma(m, \frac{\hat{t}}{\hat{u}m}; \hat{t})$.

[5] We used Variant I for $\sigma^2/18.75 \leq 2.0$ and Variant II for $\sigma^2/18.75 > 2.0$.

[6] For each $b \in N_{2^n-1}$ and each $c \in N_{L-1}$ estimates $\tilde{r}$ were calculated with $m = 2^{18}$ and $\hat{t} = 3$ and a pair b, c with largest value of $\tilde{r}$ was selected.

[7] For $\sigma^2 = 1.0 * 18.75$ 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]).

[8] For $\sigma^2 = 3.5 * 18.75$ the empirical median of secret key rank estimates is 2^{109} (each of 1,000 estimates was obtained in our simulation set-up using 2^{24} uniform random samples from the set of 2^{128} keys).