Best Information is Most Successful
CHES, Atlanta, GA, USA,
Aug 27, 2019
Setup ............ remember CHES 2014 [HRG14]?

Good Is Not Good Enough
Deriving Optimal Distinguishers from Communication Theory

Annelie Heuser\(^1\)*, Olivier Rioul\(^1\), and Sylvain Guilley\(^{1,2}\)

\(^1\) Télécom ParisTech, Institut Mines-Télécom, CNRS LTCI, Department Comelec46 rue Barrault, 75634 Paris Cedex 13, France
firstname.lastname@telecom-paristech.fr

\(^{2}\) Secure-IC S.A.S.,
80 avenue des Buttes de Coësmes, 35700 Rennes, France
Setup .............. remember CHES 2014 [HRG14]?
The attacker makes \( q \) queries \( X = (X_1, \ldots, X_q) \) which depend on the secret \( K \) and on the text \( T \) through a sensitive variable \( Y \), and estimates the secret using a \textit{distinguisher} \( \hat{K} = \mathcal{D}(X, T) \).
Side-Channel Analysis Setup

The attacker makes \( q \) queries \( X = (X_1, \ldots, X_q) \) which depend on the secret \( K \) and on the text \( T \) through a sensitive variable \( Y \), and estimates the secret using a *distinguisher* \( \hat{K} = \mathcal{D}(X, T) \).

---

- Any noisy measurement channel;
- *Countermeasures* can protect \( Y = \text{random funct. of } (K, T) \).
Test and evaluation tool (ISO/IEC 19790 & 15408)

Catalyzr®, Virtualyzr®, Analyzr® tools.
Side-Channel Attacks on Hardware

Best attack (MAP, ML)

The best distinguisher maximizes likelihood for uniformly distributed $K$ [HRG14]:

$$\hat{K} = \mathcal{D}(X, T) = \arg \max_{k \in K} P(X|T, k) \quad \text{where } X = \text{noisy } Y$$
Side-Channel Attacks on Hardware

Best attack (MAP, ML)

The best distinguisher maximizes likelihood for uniformly distributed $K$ [HRG14]:

$$\hat{K} = \mathcal{D}(X, T) = \arg\max_{k \in \mathcal{K}} \mathbb{P}(X|T, k) \quad \text{where } X = \text{noisy } Y$$

This is a *template* attack which requires estimation of unknown conditional distributions with a leakage *model*, e.g.,

$$Y(K, T) = w_H(S_{\text{box}}(T \oplus K)) \quad \text{(unprotected)}$$

$$Y(K, T) = \left[w_H(S_{\text{box}}(T \oplus K) \oplus M), w_H(M)\right] \quad \text{(masked)}$$
Side-Channel Attacks on Hardware

Best attack (MAP, ML)

The best distinguisher maximizes likelihood for uniformly distributed $K$ [HRG14]:

$$\hat{K} = D(\mathbf{X}, \mathbf{T}) = \arg\max_{k \in \mathcal{K}} \mathbb{P}(\mathbf{X}|\mathbf{T}, k) \quad \text{where } \mathbf{X} = \text{noisy } \mathbf{Y}$$

This is a template attack which requires estimation of unknown conditional distributions with a leakage model, e.g.,

$$\mathbf{Y}(K, \mathbf{T}) = w_H(S_{\text{box}}(\mathbf{T} \oplus K)) \quad \text{(unprotected)}$$

$$\mathbf{Y}(K, \mathbf{T}) = \left[ w_H(S_{\text{box}}(\mathbf{T} \oplus K) \oplus \mathbf{M}), w_H(\mathbf{M}) \right] \quad \text{(masked)}$$

Many practical attacks exist (CPA, MIA, KSA, M. Learning)
Side-Channel Attacks on Hardware

Best attack (MAP, ML)

The best distinguisher maximizes likelihood for uniformly distributed $K$ [HRG14]:

$$\hat{K} = D(X, T) = \arg \max_{k \in \mathcal{K}} P(X|T, k)$$

where $X = \text{noisy } Y$

This is a template attack which requires estimation of unknown conditional distributions with a leakage model, e.g.,

$Y(K, T) = w_H(S_{box}(T \oplus K))$ (unprotected)

$Y(K, T) = \left[w_H(S_{box}(T \oplus K) \oplus M), w_H(M)\right]$ (masked)

Many practical attacks exist (CPA, MIA, KSA, M. Learning)

The attacker will eventually always succeed as $q \to \infty$. 
The Defender (Chip Designer)’s Viewpoint

Question
Assuming any possible attack, possibly with an omniscient attacker, (which knows everything except $K$ (Kerckhoffs principle), noise and masks) what is the least number of queries to achieve a given key recovery success rate?

$$q(P_s) = \min \{ q \text{ s.t. } \mathbb{P}(\hat{K} = K) \geq P_s \}$$
The Defender (Chip Designer)’s Viewpoint

Question
Assuming any possible attack, possibly with an omniscient attacker, (which knows everything except $K$ (Kerckhoffs principle), noise and masks) what is the least number of queries to achieve a given key recovery success rate?

$$q(P_s) = \min\{q \text{ s.t. } \mathbb{P}(\hat{K} = K) \geq P_s\}$$

Practical significance:
- any attacker with budget $< q(P_s)$ cannot recover the key with probability $> P_s$;
- when $q > q(P_s)$, there only might be an attack with success $P_s$. 
Information Theoretic Background

Notations:

- $H$ is Shannon entropy, e.g., $H(K) = n$ bit
- $H_2(p) = -p \log p - (1 - p) \log(1 - p)$ is the binary entropy
- $D(P_A \parallel P_B)$ is the Kullback-Leibler divergence
- $D_2(p_A \parallel p_B) = p_A \log \frac{p_A}{p_B} + (1 - p_A) \log \frac{1-p_A}{1-p_B}$ is the binary divergence
- $I(A; B) = D(P_{A,B} \parallel P_A \otimes P_B)$ is mutual info btw $A$ and $B$
- $I(A; B \mid C)$ is mutual info btw $A$ and $B$ conditioned by $C$

DPI: Data Processing Inequality

- $A \rightarrow B \rightarrow C \rightarrow D : I(B; C) \geq I(A; D)$
- $P_A \rightarrow Q_A$ and $P_B \rightarrow Q_B$ for same processing: $D(P_A \parallel P_B) \geq D(Q_A \parallel Q_B)$
Application of Data Processing Inequality

First, we notice that:

\[ I(K; \hat{K}) = D(P_{K, \hat{K}} \| P_K \otimes P_{\hat{K}}) \]

\[ \geq D(P(K = \hat{K}) \| P'(K = \hat{K})) \quad // \text{DPI for } f : (K, \hat{K}) \mapsto 1_{K=\hat{K}} \]

\[ = P_s \log \left( \frac{P_s}{1/2^n} \right) + P_e \log \left( \frac{P_e}{1 - 1/2^n} \right) \]

\[ = n - H_2(P_s) - P_e \log(2^n - 1). \quad // \text{Fano’s inequality} \]

Since \( K \rightarrow Y \rightarrow X \rightarrow \hat{K} \) for a given \( T \) is a Markov chain:

\[ I(K; \hat{K}) \leq I(X; Y \mid T). \]
Application of Data Processing Inequality

First, we notice that:

\[
I(K; \hat{K}) = D(\mathbb{P}_K, \mathbb{P}_{\hat{K}} \| \mathbb{P}_K \otimes \mathbb{P}_{\hat{K}})
\]

\[
\geq D_2(\mathbb{P}(K = \hat{K}) \| \mathbb{P}'(K = \hat{K})) \quad // \text{DPI for } f : (K, \hat{K}) \mapsto 1_{K=\hat{K}}
\]

\[
= \mathbb{P}_s \log \frac{\mathbb{P}_s}{1/2^n} + \mathbb{P}_e \log \frac{\mathbb{P}_e}{1 - 1/2^n}
\]

\[
= n - H_2(\mathbb{P}_s) - \mathbb{P}_e \log(2^n - 1). \quad // \text{Fano’s inequality}
\]

Since \(K \rightarrow Y \rightarrow X \rightarrow \hat{K}\) for a given \(T\) is a Markov chain:

\[
I(K; \hat{K}) \leq I(X; Y \mid T).
\]
Fundamental Lower Bound on $I(X; Y \mid T)$

Proposition
For any $n$-bit key $K$:

$$n - H_2(\mathbb{P}_s) - (1 - \mathbb{P}_s) \log_2(2^n - 1) \leq I(X; Y \mid T).$$
Fundamental Lower Bound on $I(X; Y | T)$

Proposition
For any $n$-bit key $K$:

$$n - H_2(P_s) - (1 - P_s) \log_2(2^n - 1) \leq I(X; Y | T).$$

$I(X; Y | T)$ depends on $q$
Fundamental Lower Bound on $I(X; Y \mid T)$

**Proposition**

For any $n$-bit key $K$:

$$n - H_2(P_s) - (1 - P_s) \log_2(2^n - 1) \leq I(X; Y \mid T).$$

- $I(X; Y \mid T)$ depends on $q$
- When $q = 0$ (blind attacker) $I(X; Y \mid T) = 0$ and $P_s = 1/2^n$. 
Fundamental Lower Bound on $I(X;Y \mid T)$

Proposition
For any $n$-bit key $K$:

$$n - H_2(\mathbb{P}_s) - (1 - \mathbb{P}_s) \log_2(2^n - 1) \leq I(X;Y \mid T).$$

- $I(X;Y \mid T)$ depends on $q$
- When $q = 0$ (blind attacker) $I(X;Y \mid T) = 0$ and $\mathbb{P}_s = 1/2^n$.
- In the context of cryptanalysis, $\mathbb{P}_s$ should be high enough (divide and conquer approach, e.g., 16 bytes for AES [NIS01]).
  In such regime, Fano’s inequality is fairly tight.
First Upper Bound on $I(X; Y | T)$

Linear Bound
For $q$ queries:

$$I(X, Y | T) \leq q \cdot I(X; Y | T)$$

Proof.
Memoryless channel assumption.

However, the same $K$ is used $q$ times (huge repetition !)
First Upper Bound on $I(X; Y \mid T)$

Linear Bound
For $q$ queries:

$$I(X, Y \mid T) \leq q \cdot I(X; Y \mid T)$$

Proof.
Memoryless channel assumption.

- However, the same $K$ is used $q$ times (huge repetition !)
- Therefore, $I(X, Y \mid T) \leq H(Y \mid T) \leq H(K) = n$ should be bounded by $n$ bits as $q \rightarrow +\infty$. 
Second Upper Bound on $I(X; Y \mid T)$

Divergence Bound

$$I(X; Y \mid T) \leq -\mathbb{E}_T \mathbb{E}_K \log \mathbb{E}_{K'} \exp \left[ -D(\mathbb{P}_{X|K,T} || \mathbb{P}_{X|K',T}) \right]$$

where $K'$ is an independent copy of $K$.

Proof.
Apply the (equivalent) inequalities

$$-\mathbb{E}_Y \log \mathbb{E}_X [\exp(f(X, Y))] \leq - \log \mathbb{E}_X [\exp(\mathbb{E}_Y f(X, Y))].$$

$$\exp \mathbb{E}_Y \log \mathbb{E}_X [g(X, Y)] \geq \mathbb{E}_X [\exp(\mathbb{E}_Y \log g(X, Y))]$$

This upper bound is bounded by $n$ bits as $q \to \infty$. 

E.}
Figure: Mutual information $I(X; Y \mid T = t)$, where $t$ is a fixed balanced vector. Comparison for $n = 8$, assuming Hamming weight leakage model in AES, AWGN with $\sigma = 4$. 
(Scalar) mutual info does not exceed Shannon channel’s capacity:

\[ I(X; Y \mid T) \leq \frac{1}{2} \log_2(1 + \text{SNR}). \]

Theorem (Lower bound for AWGN in terms of SNR)

To reach success \( P_s \), q should be at least

\[ q \geq n + (P_s - 1) \log_2(2^n - 1) - H_2(P_s) \frac{1}{2} \log_2(1 + \text{SNR}). \]
The number of traces $q$ needed to recover the key reliably is lower-bounded by:

$$
\lim_{P_s \to 1} q \geq \frac{n}{\frac{1}{2} \log_2(1 + \text{SNR})}
$$

where $\text{SNR}$ can be measured on the fly (for balanced text $T$):

$$
\text{SNR} = \frac{\text{Var}(\mathbb{E}[X \mid T])}{\text{Var}(X) - \text{Var}(\mathbb{E}[X \mid T])}.
$$

No more leakage if $\text{SNR} \to 0$. 
In the AWGN model, \( \mathbb{P}_{X|K_i,T} \) follows a multivariate normal distribution \( \mathcal{N}(y(K_i, T), \sigma^2 I_q) \).

\[
D(\mathbb{P}_{X|K,T} \mid \mathbb{P}_{X|K',T}) = \frac{\|y(K, T) - y(K', T)\|_2^2}{2\sigma^2}.
\]

Besides, for balanced \( T \):

\[
\frac{1}{q} \frac{1}{2} \frac{\|y(k, t) - y(k', t)\|_2^2}{2} \xrightarrow{q \to \infty} \kappa(k, k'), \quad \text{// LLN}
\]

where

\[
\kappa(k, k') = \frac{1}{2^n} \sum_{t=0}^{2^n-1} \left( \frac{y(k, t) - y(k', t)}{2} \right)^2 \quad \text{(confusion coefficient)}
\]
Implicit bound:

\[ H_2(\mathbb{P}_s) + (1 - \mathbb{P}_s) \log_2(2^n - 1) \geq \frac{n_{\min}}{2^n} \exp \left( -\frac{q \min_{k \neq k'} \kappa(k, k')}{8 \sigma^2} \right). \]

where \( n_{\min} \) is the number of ex aequo key pairs \((k, k')\) such that \( \kappa(k, k') \) is minimum.
Comparison with Duc et al. [DFS15]
Making Masking Security Proofs Concrete (EC 2015, Duc, Faust, Standaert)

(Duc et al. use Pinsker’s inequality)
Simulation for Monobit Leakage

Monobit leakage model:  \( Y(T, K) = \text{LSB}(S_{\text{box}}(T \oplus K)) \)
where \( S_{\text{box}} = \text{AES substitution box} \) and \( \text{LSB} = \text{least significant bit} \).
Simulation for Hamming Weight Leakage

AES SubBytes based on bytes: \[ Y(T, K) = w_H(S_{\text{box}}(T \oplus K)) \]
where \( S_{\text{box}} \) = AES substitution box and \( w_H \) is the Hamming weight.
Conclusion

We obtained *universal* bounds to the success probability in terms of mutual information, in the sense that they are independent of the channel and leakage models;

Our results were presented within the specific framework of “power-line attacks” (e.g., monobit leakage or Hamming weight leakage);

The resulting bounds were found to be empirically tight.
Announcements

Secure-IC recruits:

- R&D team director, based in Paris (10 people in Paris, Rennes, Singapour and Tokyo)
- Tokyo “Security Science Factory” laboratory manager

TELECOM-Paris recruits (Palaiseau, France):

- PhD candidate in IT-powered SCA
- Researcher in embedded security, in Jean-Luc Danger’s team

Good Is Not Good Enough - Deriving Optimal Distinguishers from Communication Theory.

[NIS01] NIST/ITL/CSD.
Best Information is Most Successful
CHES, Atlanta, GA, USA,
Aug 27, 2019