Institut
Mines-Télécom

Defining Perceived Information based on Shannon's Communication Theory

Cryptarchi 2016
June 21-24, 2016
La Grande Motte, France

Eloi de Chérisey, Sylvain Guilley, & Olivier Rioul

Télécom ParisTech, Université Paris-Saclay, France.

Contents

Contents

Motivation

- Consolidate the state of the art about **Perceived Information** (PI) metrics;
- Continue the work of Annelie Heuser presented last year at CryptArchi;
- Establish clear and coherent definitions for PI based on **optimal distinguishers** and **Shannon's theory**;

Motivation

- Consolidate the state of the art about **Perceived Information** (PI) metrics;
- Continue the work of Annelie Heuser presented last year at CryptArchi;
- Establish clear and coherent definitions for PI based on **optimal distinguishers** and **Shannon's theory**;
- Deduce **tests** in order to evaluate the success of an attack;

Motivation

- Consolidate the state of the art about **Perceived Information** (PI) metrics;
- Continue the work of Annelie Heuser presented last year at CryptArchi;
- Establish clear and coherent definitions for PI based on **optimal distinguishers** and **Shannon's theory**;
- Deduce **tests** in order to evaluate the success of an attack;
- Introduce **communication channels** in Side-Channel Analysis (SCA).
- Is Shannon's **channel capacity** useful in SCA?

Assumptions and Notations

What is an attack?

- Two phases: *profiling* phase & *attacking* phase.

Assumptions and Notations

What is an attack?

- Two phases: *profiling* phase & *attacking* phase.
- **Profiling phase**: secret key $\hat{k}$ is known. A vector of $\hat{q}$ textbytes $\hat{\mathbf{t}}$ is given and $\hat{q}$ traces $\hat{\mathbf{x}}$ are measured;
- **Attacking phase**: secret key $\tilde{k}$ is unknown. A vector of $\tilde{q}$ textbytes $\tilde{\mathbf{t}}$ is given and $\tilde{q}$ traces $\tilde{\mathbf{x}}$ are measured;

Assumptions and Notations

What is an attack?

- Two phases: *profiling* phase & *attacking* phase.
- **Profiling phase**: secret key $\hat{k}$ is known. A vector of $\hat{q}$ textbytes $\hat{\mathrm{t}}$ is given and $\hat{q}$ traces $\hat{\mathrm{x}}$ are measured;
- **Attacking phase**: secret key $\tilde{k}$ is unknown. A vector of $\tilde{q}$ textbytes $\tilde{\mathrm{t}}$ is given and $\tilde{q}$ traces $\tilde{\mathrm{x}}$ are measured;
- The leakages follow some **unknown** distribution $\mathbb{P}$;
- **Estimate** $\mathbb{P}$ based on either $\hat{\mathrm{x}}, \hat{\mathrm{t}}$ or $\tilde{\mathrm{x}}, \tilde{\mathrm{t}}$.

Assumptions and Notations (Cont'd)

Consider the following sets and variables.

- $\hat{\mathcal{X}}$ and $\tilde{\mathcal{X}}$ for $\hat{x}$ and $\tilde{x}$.
- $\hat{\mathcal{T}}$ and $\tilde{\mathcal{T}}$ for $\hat{t}$ and $\tilde{t}$.

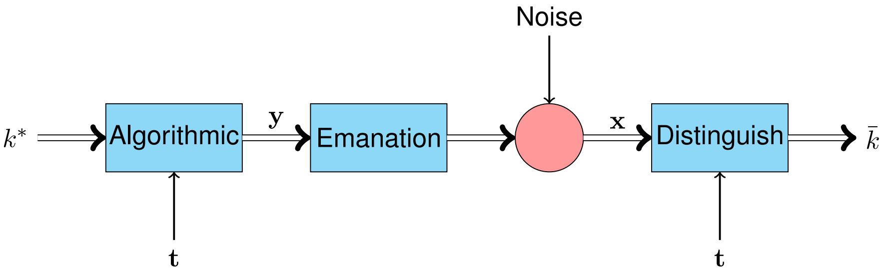
TELECOM ParisTech

Assumptions and Notations (Cont'd)

Consider the following sets and variables.

- $\hat{\mathcal{X}}$ and $\tilde{\mathcal{X}}$ for $\hat{x}$ and $\tilde{x}$.
- $\hat{\mathcal{T}}$ and $\tilde{\mathcal{T}}$ for $\hat{t}$ and $\tilde{t}$.
- Random variable $\hat{X}$, $\tilde{X}$, $\hat{T}$ and $\tilde{T}$.
- Random vectors $\hat{\mathbf{X}}$, $\tilde{\mathbf{X}}$, $\hat{\mathbf{T}}$ and $\tilde{\mathbf{T}}$.
- Generic notation $\mathbf{x}$ (either profiling or attacking)

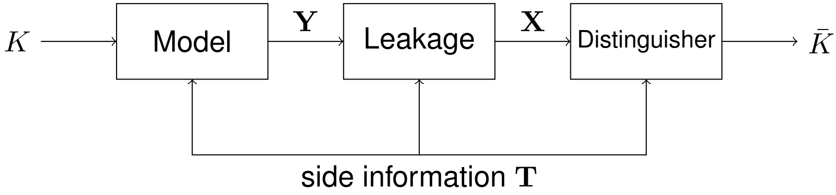
Leakage Model



Recall our notational conventions:

– profiling phase with a hat $\hat{\bullet}$.
– attacking phase with a tilde $\tilde{\bullet}$.

Leakage Equivalent Flow-Graph



Markov Chain

We have the following Markov Chain given $\mathbf{T}$:

$$K \longrightarrow \mathbf{Y} \longrightarrow \mathbf{X} \longrightarrow \bar{K}$$

The attacker receives $\mathbf{X}$.

Estimations of the Probability Distribution $\mathbb{P}$

Definition (Profiled Estimation: OffLine)

$$\forall x, t \quad \hat{\mathbb{P}}(x,t) = \frac{1}{\hat{q}} \sum_{i=1}^{\hat{q}} \mathbb{1}_{\hat{x}_i = x, \hat{t}_i = t} \tag{1}$$

TELECOM ParisTech

Estimations of the Probability Distribution $\mathbb{P}$

Definition (Profiled Estimation: OffLine)

$$\forall x, t \quad \hat{\mathbb{P}}(x,t) = \frac{1}{\hat{q}} \sum_{i=1}^{\hat{q}} \mathbb{1}_{\hat{x}_i = x, \hat{t}_i = t} \tag{1}$$

Definition (On-the-fly Estimation: OnLine)

$$\forall x, t \quad \tilde{\mathbb{P}}(x,t) = \frac{1}{\tilde{q}} \sum_{i=1}^{\tilde{q}} \mathbb{1}_{\tilde{x}_i = x, \tilde{t}_i = t} \tag{2}$$

Optimal Distinguisher

Theorem (Optimal Distinguisher)

The optimal distinguisher [2] is the maximum a posteriori (MAP) distinguisher defined by

$$\mathcal{D}_{\mathsf{Opt}}(\tilde{\mathbf{x}}, \tilde{\mathbf{t}}) = \arg\max \mathbb{P}(k|\tilde{\mathbf{x}}, \tilde{\mathbf{t}}) \tag{3}$$

Optimal Distinguisher

Theorem (Optimal Distinguisher)

The optimal distinguisher [2] is the maximum a posteriori (MAP) distinguisher defined by

$$\mathcal{D}_{\mathsf{Opt}}(\tilde{\mathbf{x}}, \tilde{\mathbf{t}}) = \arg\max \mathbb{P}(k|\tilde{\mathbf{x}}, \tilde{\mathbf{t}}) \tag{3}$$

As $\mathbb{P}$ is unknown, we may replace it by $\hat{\mathbb{P}}$ in the distinguisher :

$$\mathcal{D}(\tilde{\mathbf{x}}, \tilde{\mathbf{t}}) = \arg\max \hat{\mathbb{P}}(k|\tilde{\mathbf{x}}, \tilde{\mathbf{t}}) \tag{4}$$

Contents

SCA Seen as a Markov Chain

Theorem (SCA as a Markov Chain)

The following is a Markov Chain:

$$(K, \mathbf{T}) \longrightarrow (\mathbf{Y}, \mathbf{T}) \longrightarrow (\mathbf{X}, \mathbf{T}) \longrightarrow (\bar{K}, \mathbf{T})$$

In other words: as $\mathbf{T}$ is known everywhere we can put it at every stage. Therefore, Mutual Information $I(K, \mathbf{T}; \mathbf{X}, \mathbf{T})$ is a relevant quantity.

Mutual Information

Theorem (i.i.d. Channel)

For an i.i.d. channel, we have:

$$I(K, \mathbf{T}; \mathbf{X}, \mathbf{T}) = q \cdot I(K, T; X, T) \qquad (5)$$

The relevant quantity becomes $I(K, T; X, T)$*.*

Proof.

Using independence,

$$\begin{aligned} I(K, \mathbf{T}; \mathbf{X}, \mathbf{T}) &= H(\mathbf{X}, \mathbf{T}) - H(\mathbf{X}, \mathbf{T} | K, \mathbf{T}) \\ &= q \cdot H(X, T) - H(\mathbf{X} | K, \mathbf{T}) \\ &= q \cdot H(X, T) - qH(X | K, T) \\ &= q \cdot I(K, T; X, T) \end{aligned}$$

The Role of Perceived Information

Mutual Information $I(K,T;X,T)$ is important in order to evaluate the attack. We have:

$$I(K,T;X,T) = \underbrace{H(K,T)}_{=H(K)+H(T)} - \underbrace{H(K,T|X,T)}_{=H(K|X,T)} \tag{6}$$

The Role of Perceived Information

Mutual Information $I(K,T;X,T)$ is important in order to evaluate the attack. We have:

$$I(K,T;X,T) = \underbrace{H(K,T)}_{=H(K)+H(T)} - \underbrace{H(K,T|X,T)}_{=H(K|X,T)} \qquad (6)$$

giving

$$\boxed{I(K,T;X,T) = H(K) + H(T) - \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \mathbb{P}(x|k,t) \log \mathbb{P}(k|x,t)}. \qquad (7)$$

The Role of Perceived Information (Cont'd)

Issues

- $\mathbb{P}(k|x,t)$ is unknown!
- It has to be estimated: $\hat{\mathbb{P}}$ and $\tilde{\mathbb{P}}$.
- How to use $\hat{\mathbb{P}}$ and $\tilde{\mathbb{P}}$ in order to estimate the Mutual Information?

The Role of Perceived Information (Cont'd)

Issues

- $\mathbb{P}(k|x,t)$ is unknown!
- It has to be estimated: $\hat{\mathbb{P}}$ and $\tilde{\mathbb{P}}$.
- How to use $\hat{\mathbb{P}}$ and $\tilde{\mathbb{P}}$ in order to estimate the Mutual Information?

Answer

We define the **Perceived Information** as the estimation of Mutual Information using the MAP distinguisher.

Deriving the Perceived Information

The MAP distinguishing rule is given by

$$\begin{aligned}
\text{MAP} &= \arg\max \hat{\mathbb{P}}(k|\tilde{\mathbf{x}}, \tilde{\mathbf{t}}) \\
&= \arg\max \prod_{i=1}^{\tilde{q}} \hat{\mathbb{P}}(k|x_i, t_i) \\
&= \arg\max \prod_{x,t} \hat{\mathbb{P}}(k|x, t)^{\tilde{n}_{x,t}} \\
&= \arg\max \sum_{x,t} \tilde{\mathbb{P}}(x, t|k) \log \hat{\mathbb{P}}(k|x, t) \\
&= \arg\max \sum_{t} \tilde{\mathbb{P}}(t|k) \sum_{x} \tilde{\mathbb{P}}(x|k, t) \log \hat{\mathbb{P}}(k|x, t)
\end{aligned}$$

The Role of Perceived Information (Cont'd)

One obtains

$$\text{MAP} = \arg\max \sum_t \tilde{\mathbb{P}}(t|k) \sum_x \tilde{\mathbb{P}}(x|k,t) \log \hat{\mathbb{P}}(k|x,t) \tag{8}$$

Summming over $\mathbb{P}(k)$ and adding $H(K) + H(T)$ yields the form

$$H(K) + H(T) + \sum_k \mathbb{P}(k) \sum_t \tilde{\mathbb{P}}(t) \sum_x \tilde{\mathbb{P}}(x|k,t) \log \hat{\mathbb{P}}(k|x,t)$$

The Role of Perceived Information (Cont'd)

One obtains

$$\text{MAP} = \arg\max \sum_t \tilde{\mathbb{P}}(t|k) \sum_x \tilde{\mathbb{P}}(x|k,t) \log \hat{\mathbb{P}}(k|x,t) \tag{8}$$

Summing over $\mathbb{P}(k)$ and adding $H(K) + H(T)$ yields the form

$$H(K) + H(T) + \sum_k \mathbb{P}(k) \sum_t \tilde{\mathbb{P}}(t) \sum_x \tilde{\mathbb{P}}(x|k,t) \log \hat{\mathbb{P}}(k|x,t)$$

To be compared with MI:

$$H(K) + H(T) + \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \mathbb{P}(x|k,t) \log \mathbb{P}(k|x,t)$$

Definition of Perceived Information

This leads to the following definition.

Definition (Perceived Information)

$$PI(K,T;X,T) = H(K) + H(T) + \sum_k \mathbb{P}(k) \sum_t \tilde{\mathbb{P}}(t) \sum_x \tilde{\mathbb{P}}(x|k,t) \log \hat{\mathbb{P}}(k|x,t) \quad (9)$$

Interpretation of PI

Interpretation

We defined PI under the prism of Mutual Information estimation, with the MAP distinguisher base for the estimated distributions.

PI has been first proposed by[1] in order to check if the estimated distribution of a chip is relevent or not.
They tested $\hat{\mathbb{P}}$ under $\mathbb{P} \rightarrow \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \mathbb{P}(x|k,t) \log \hat{\mathbb{P}}(k|x,t)$.
In our case, we test $\hat{\mathbb{P}}$ under $\tilde{\mathbb{P}}$ $\rightarrow$Eq. 9, meaning that we define PI as a way to check whether online and offline distributions are coherent.
We have chosen this particular Mutual Information $I(K,T;X,T)$ as it will be very usefull for the next computations.

Contents

A Lower Bound

Consider the Markov Chain defined earlier:

$$(K, \mathbf{T}) \longrightarrow (\mathbf{Y}, \mathbf{T}) \longrightarrow (\mathbf{X}, \mathbf{T}) \longrightarrow (\bar{K}, \mathbf{T})$$

Theorem (Minimum Number of Traces)

With such a Markov Chain, we have the universal inequality

$$q \geq \frac{n\mathbb{P}_s - H_2(\mathbb{P}_s)}{I(X;Y|T)} \tag{10}$$

This inequation is true whatever the attack and the leakage. In fact, it is a weak inequality, but is gives the minimum nuber of traces to have a chance to reach a certain success.

Sketch of Proof

By the Data Processing Inequality (DPI) in Information Theory:

$$I(K, \mathbf{T}; \bar{K}, \mathbf{T}) \leq I(\mathbf{Y}, \mathbf{T}; \mathbf{X}, \mathbf{T})$$

The l.h.s. in the DPI takes the form

$$\begin{aligned} I(K, \mathbf{T}; \bar{K}, \mathbf{T}) &= H(K, \mathbf{T}) - H(K, \mathbf{T} | \bar{K}, \mathbf{T}) \\ &= H(K) + q \cdot H(T) - H(K | \bar{K}, \mathbf{T}) \\ &\geq H(K) + q \cdot H(T) - H(K | \bar{K}) \end{aligned}$$

By the information -theoretic inequality of Fano, we get:

$$\boxed{I(K, \mathbf{T}; \bar{K}, \mathbf{T}) \geq H(K) + qH(T) - n(1 - \mathbb{P}_s) - H_2(\mathbb{P}_s)}$$

Where $\mathbb{P}_s$ is the probability of success : $\mathbb{P}_s = \mathbb{P}(K = \bar{K})$.

Sketch of Proof (Cont'd)

The r.h.s. in the DPI takes the form

$$\begin{aligned} I(\mathbf{Y},\mathbf{T};\mathbf{X},\mathbf{T}) &= q\cdot I(Y,T;X,T) \\ &= q\cdot (H(Y,T)-H(Y,T|X,T)) \\ &= q\cdot (H(T)+H(Y|T)-H(T|X,T)-H(Y|X,T)) \\ &= q\cdot (H(T)+I(X;Y|T)) \end{aligned}$$

Combining we obtain:

$$\boxed{H(K)+qH(T)-n(1-\mathbb{P}_s)-H_2(\mathbb{P}_s)\leq q(H(T)+I(X;Y|T))}$$

where $H(K)=n$ for equiprobable keys. This proves the theorem $\square$

AWGN Case

We consider an Additive White Gaussian Noise N such that $X = Y + N$.

Theorem (Highest Mutual Information)

We show that:

$$\max_{T-Y-X} I(X;Y|T) = \max_{Y} I(X;Y) = \frac{1}{2}\log_2(1+\mathsf{SNR}) \tag{11}$$

Therefore, according to Eq. 10, in order to reach a full success rate ($\mathbb{P}_s = 1$), the attacker needs to get at least $q \geq \frac{2n}{\log_2(1+\mathsf{SNR})}$ traces.

Link With Channel Capacity

Definition (Channel Capacity)

We can define the Channel Capacity by:

$$C = \max_{Y} I(X;Y) \tag{12}$$

As we saw earlier, in the case of an AWGN, the capacity of the channel is $C = \frac{1}{2}\log_2(1 + \text{SNR})$.

Protection Rule

In order to protect hardwares from leakages, according to Eq. 10, we have to ensure that C is as small as possible and therefore SNR **as small as possible.**

Link With Perceived Information

We now consider the worst possible case for the attacker: no model! Therefore, $Y = K, T$. The Mutual Information $I(X; Y|T)$ becomes $I(X; K, T|T)$.

Link With Perceived Information

We now consider the worst possible case for the attacker: no model! Therefore, $Y = K, T$. The Mutual Information $I(X;Y|T)$ becomes $I(X;K,T|T)$.

$$\begin{aligned} I(X;K,T|T) &= H(K,T|T) - H(K,T|X,T) \\ &= H(K) - H(K|X,T) \\ &= I(K,T;X,T) - H(T) \\ &= H(K) + \sum_k \mathbb{P}(k) \sum_t \mathbb{P}(t) \sum_x \mathbb{P}(x|k,t) \log \mathbb{P}(k|x,t) \end{aligned}$$

Including PI

Once again, $I(X;K,T|T)$ is unknown. We use the PI estimation defined in Eq. 9

Inequality With PI

Estimation of $I(X;Y|T)$

The estimation of $I(X;K,T|T)$ is:

$$H(K) + \sum_k \mathbb{P}(k) \sum_t \tilde{\mathbb{P}}(t) \sum_x \tilde{\mathbb{P}}(x|k,t) \log \hat{\mathbb{P}}(k|x,t) = PI(K,T;X,T) - H(T) \qquad (13)$$

Now, rewriting Eq. 10 with the estimation:

$$q_{\text{est}} \geq \frac{n\mathbb{P}_s - H_2(\mathbb{P}_s)}{PI(K,T;X,T) - H(T)}$$

If $PI(K,T;X,T) - H(T) \leq 0$, it means that PI is not a correct estimation of MI. Calculations are not relevant in this case.

Contents

Conclusion

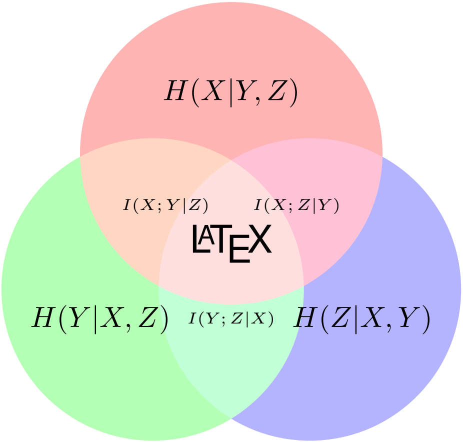
- A coherent definition of PI.
- SCA seen as a Markov Chain structure.
- Lower bounds of the number of traces - Shannon limit.
- Implication with PI.

Thank you!

Questions?

eloi.de-cherisey@mines-telecom.fr

Wenn Diagrams



References I

François Durvaux, François-Xavier Standaert, and Nicolas Veyrat-Charvillon.
How to Certify the Leakage of a Chip?
In Phong Q. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, volume 8441 of *Lecture Notes in Computer Science*, pages 459–476. Springer, 2014.

References II

Annelie Heuser, Olivier Rioul, and Sylvain Guilley.
Good Is Not Good Enough - Deriving Optimal Distinguishers from Communication Theory.
In Lejla Batina and Matthew Robshaw, editors, *Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings*, volume 8731 of *Lecture Notes in Computer Science*, pages 55–74. Springer, 2014.