Blind Side Channel Analysis using joint distributions

Christophe Clavier, Léo Reynaud, Antoine Wurcker

Université de Limoges - XLIM
Eshard
Table of contents

1. Introduction
2. Joint distributions
3. Passive attack with joint distributions
4. Higher order
5. Quadrivariate joint distributions
Cryptology

Cryptography

- Private communication
- Authentication
- Integrity
- Non repudiation

Cryptanalysis

Tires to break cryptography

Cryptology

Both cryptography and cryptanalysis
## Cryptanalysis

<table>
<thead>
<tr>
<th>Mathematical</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Targets algorithm itself</td>
</tr>
<tr>
<td>● Exploits mathematical properties between inputs/outputs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical attacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Targets physical implementation</td>
</tr>
<tr>
<td>● Three kinds:</td>
</tr>
<tr>
<td>● Invasives</td>
</tr>
<tr>
<td>● Semi-invasives</td>
</tr>
<tr>
<td>● Non-invasives / Passives</td>
</tr>
</tbody>
</table>
Side Channel attacks

Figure 1: Non exhaustive side channels attacks
Side Channel attacks

Figure 1: Non exaustive side channels attacks
Side Channel attacks

Common non profiled side channel attacks

- DPA [KJJ99]
- CPA [BCO04]
- MIA [GBTP08]

Figure 2: Internal states variables

Needs

- Leakage on some internal state
- Knowledge and variability of plain/ciphertext
Side Channel attacks: CPA

**Figure 2:** Internal states variables

**Figure 3:** CPA principle
The need of blind attacks

**Figure 4:** EMV session key derivation

**Figure 5:** Aes early/final rounds protected
The need of blind attacks

- $m$ is unknown
  (does not vary much)
- $y$ is unpredictable

Figure 6: Passive Joint Distribution attack principle

$\rightarrow$ Let’s observe both
Joint Distributions

- \( m, x, k \in GF(2) \)
- \( x = m \oplus k \)

\[
\begin{array}{c|cc}
\text{HW}(m) & 0 & 1 \\
\hline
0 & 0 & 1/2 \\
1 & 1 & 1/2 \\
\end{array}
\]

\[
\begin{array}{c|cc}
\text{HW}(x) & 0 & 1 \\
\hline
0 & 1/2 & 1/2 \\
1 & 0 & 1/2 \\
\end{array}
\]

**Figure 7**: Joint distribution \( k = 0 \)

**Figure 8**: Joint distribution \( k = 1 \)

→ Distribution related to \( k \)
HWs Joint Distributions [LDL13]

- \( m, y, k \in GF(2^8) \)
- \( y = S(m \oplus k) \)
- HWs joint distributions of \( m, y \):

**Figure 9:** Joint distribution \( k = 39 \)

**Figure 10:** Joint distribution \( k = 126 \)
HWs Joint Distributions

(a) HW(k) = 0
(b) HW(k) = 1
(c) HW(k) = 2
(d) HW(k) = 3
(e) HW(k) = 4
(f) HW(k) = 5
(g) HW(k) = 6
(h) HW(k) = 7
(i) HW(k) = 8

Figure 11: HWs joint distributions of m and x
The attack: Pros/Cons

Cons

- Needs to know the points of interest (PoI)
- Works on HWs not consumptions

Pros

- Works without plain/ciphertext
- Works with little variability of inputs
- Any round can be attacked
The attack: Steps

- **Step 1: Processing of the traces**
  Locate the PoIs where the considered variables leak

- **Step 2: Reverse the consumption model**
  Infer HWs from the observed leakages at the Pol

- **Step 3: Joint distributions**
  Build the joint distribution for each key (can be preprocessed)

- **Step 4: Distinguisher**
  Select the key whose distribution best fits the observations
Step 1: Test Vector Leakage Assessment [GJJR11]

- Uses Welch’s t-test
- Finds differences between distributions of samples
- Fixed vs Random: Non specific

\[ t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{N_x} + \frac{s_y^2}{N_y}}} \]

- \( \bar{x} \): mean of \( x \)
- \( s_x^2 \): variance of \( x \)
- \( N_x \): sample size of \( x \)
Step 2: Slices

Consumption model

\[ \ell = \alpha \text{HW}(v) + \beta + \omega \]

Consumptions

- Sort \( N \) consumptions
- \( \frac{N^* C^2}{2^8} \) first \( \rightarrow \) HW = 0
- \( \frac{N^* C^1}{2^8} \) next \( \rightarrow \) HW = 1

Figure 12: Slice method to infer HWs
Step 1/2 : Variance [CR17]

Goal : infer HW of a variable $v$

$\rightarrow$ Infer parameters $\alpha$, $\beta$ of consumption model

\[
\begin{align*}
\text{Var}(\ell) &= \text{Var}(\alpha \text{HW}(v) + \beta) + \text{Var}(\omega) \\
&= \alpha^2 \text{Var}(\text{HW}(v)) + \text{Var}(\omega) \\
&= 2\alpha^2 + \text{Var}(\omega)
\end{align*}
\]

\[
\begin{align*}
\alpha &= \pm \sqrt{\frac{\text{Var}(\ell) - \text{Var}(\omega)}{2}} \\
\beta &= \mathbb{E}(\ell) - \alpha \mathbb{E}(\text{HW}(v))
\end{align*}
\]

\[
\text{HW}(v) = \frac{\ell - \beta}{\alpha}
\]

**Figure 13:** Standard deviation trace
Step 4 : Distances [LDL13]

- Observe HWs
- Build histograms
- Apply distances between experimental and theoretical :
  - Inner product
  - $\chi^2$
  - ...
- Select $k$ such that distance is minimum
### Step 4: Maximum likelihood [LB14]

<table>
<thead>
<tr>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_m^<em>, h_y^</em>$: correct HWs (integers)</td>
</tr>
<tr>
<td>$\omega_m, \omega_y$: noise</td>
</tr>
</tbody>
</table>
| $h_m = h_m^* + \omega_m$  
$h_y = h_y^* + \omega_y$ |

<table>
<thead>
<tr>
<th>Bayes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Pr(k</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Law of total probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Pr((h_m, h_y)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Noise probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Pr((h_m, h_y)</td>
</tr>
</tbody>
</table>
Fault/Templates

- Fault attacks [Kor16]
- Templates [HTM09]

→ Faults and templates in step 2
Improvements: More PoIs [CR17]

<table>
<thead>
<tr>
<th>x</th>
<th>000</th>
<th>001</th>
<th>010</th>
<th>011</th>
<th>100</th>
<th>101</th>
<th>110</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>S(x)</td>
<td>010</td>
<td>110</td>
<td>011</td>
<td>101</td>
<td>001</td>
<td>111</td>
<td>100</td>
<td>000</td>
</tr>
</tbody>
</table>

$k = 011$

<table>
<thead>
<tr>
<th>m</th>
<th>x</th>
<th>y</th>
<th>HW(m)</th>
<th>HW(x)</th>
<th>HW(y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>000</td>
<td>011</td>
<td>101</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>001</td>
<td>010</td>
<td>011</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>010</td>
<td>001</td>
<td>110</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>011</td>
<td>000</td>
<td>010</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>100</td>
<td>111</td>
<td>000</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>101</td>
<td>110</td>
<td>100</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>110</td>
<td>101</td>
<td>111</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>111</td>
<td>100</td>
<td>001</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

- More $\neq$ between distributions $\rightarrow$ More efficient
- Wrong PoI is catastrophic

Figure 14: Three PoIs
A secret $k$ is vulnerable if:

- We can observe at least 2 variables $a$ and $b$ such as $b = \varphi_k(a)$
- The joint distribution of the HWs of $a$ and $b$ is not identical for all $k$
Generalization: First order masking

A secret $k$ is vulnerable in the case of boolean masking if:

- We can observe at least 2 variables $a$ and $b$ such as $b = \varphi_k(a)$
- The joint distribution of the HWs of $a$ and $b$ is not identical for all $k$
- $a$ and $b$ are masked with the same mask:
  - $a' = a \oplus r$
  - $b' = b \oplus r = \varphi_k(a) \oplus r$

→ Distributions take into account all couples $(a' = a \oplus r, b' = b \oplus r)$
HWs Joint Distributions first order masking

Figure 15: First order

Figure 16: Second order
HWs Joint Distributions first order masking

Figure 17: First order

Figure 18: Second order
Masked schemes

Figure 19: Examples of Boolean masking
Quadrivariate joint distributions [CRW18]

Figure 20: Two consecutives masked bytes

Three masks vertically → Unable to attack

But usually

Same masks horizontally

$m_i \oplus u$

$x_i \oplus w$

$y_i \oplus v$

$k_i$

$m_j \oplus u$

$x_j \oplus w$

$y_j \oplus v$

$k_j$
**Quadrivariate joint distributions [CRW18]**

HWs joint distributions of $m'_i, m'_j, y'_i, y'_j$

\[
m'_i = m_i \oplus u \\
m'_j = m_j \oplus u \\
y'_i = y_i \oplus v \\
y'_j = y_j \oplus v
\]

→ Related to $k_i \oplus k_j$

HWs joint distributions of $m'_i, m'_j, x'_i, x'_j$

\[
m'_i = m_i \oplus u \\
m'_j = m_j \oplus u \\
x'_i = x_i \oplus v \\
x'_j = x_j \oplus v
\]

→ Related to $HW(k_i \oplus k_j)$
Quadrivariate joint distribution [CRW18]

Figure 21: Quadrivariate $m_y$

Figure 22: Quadrivariate $m_x$
Quadrivariate joint distribution: Recap

Cons

- Same issues as bivariate:
  - Need to locate the Pols
  - Infer HW from leakages
  - $m\ y$ not very efficient when masked
- Less efficient than classical joint distribution

New possibilities

A lot more masked schemes vulnerable:
  → Any two bytes sharing the same couple of masks
Quadrivariante $m \times$ : Key recovery on AES

Quadrivariante $m \times$ retrieves $\text{HW}(k_i \oplus k_j)$

Configurations considered

We will consider three configurations :

- Cfg1 : All bytes are masked the same way (1 set of masks)
- Cfg2 : Only bytes of a round are masked the same way (11 sets)
- Cfg3 : Only bytes of a same position in the state are masked the same way (16 sets)
**Quadrivariate** $m \times :$ Key recovery on AES

A guess-compute-backtrack approach between key bytes can be used

- Guess one or several extended key byte
- Compute some other related key bytes (key expansion)
- Backtrack in case of inconsistency
Full information

Number of key candidates

- Cfg1: Distances between all extended key bytes
- Cfg2: Distances between all bytes of a same round at every round
- Cfg3: Distances between all bytes of a same position for every position

→ Only 1 candidate
Local information

Number of key candidates

- Cfg1/Cfg2: Distances between all key bytes of a single round
  → Few millions

- Cfg 3: Distances between all key bytes of a single position
  → Millions (66 key bytes involved)

→ Adding another round/position ends up with one/few candidates

Figure 23: Propagation 1 subgroup cfg3
Thank you
I will be pleased to answer your questions
(Bibliography is next)
Eric Brier, Christophe Clavier, and Francis Olivier. 
**Correlation power analysis with a leakage model.**
Christophe Clavier and Léo Reynaud.

**Improved blind side-channel analysis by exploitation of joint distributions of leakages.**

Christophe Clavier, Léo Reynaud, and Antoine Wurcker. Quadrivariate improved blind side-channel analysis on boolean masked AES.
Benedikt Gierlichs, Lejla Batina, Pim Tuyls, and Bart Preneel. **Mutual information analysis.**

Gilbert Goodwill, Benjamin Jun, J. Jaffe, and Pankaj Rohatgi. **A testing methodology for side channel resistance.** 2011.

Roman Korkikian.  
**Side-channel and fault analysis in the presence of countermeasures : tools, theory, and practice.**  
Theses, PSL Research University, October 2016.

Hélène Le Bouder.  
**A FORMALISM FOR PHYSICAL ATTACKS ON CRYPTOGRAPHIC DEVICES AND ITS EXPLOITATION TO COMPARE AND RESEARCH NEWS ATTACKS.**  
Theses, Ecole Nationale Supérieure des Mines de Saint-Etienne, October 2014.