SCAs against Embedded Crypto Devices

F.-X. Standaert

UCL Crypto Group, Université catholique de Louvain

Lecture 2 - Side-Channel Attacks (I)

Outline

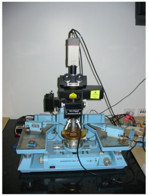
- Introduction
- Basics of Side-Channel Attacks
  - Origin of the leakages
  - Measurement setups
  - SPA, DPA
- Exemplary attack against the DES
- Improved attacks
- Countermeasures
- Key independence and asymptotic equivalences

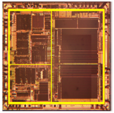
Cryptographic devices

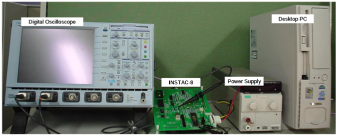
Attacks against cryptographic devices

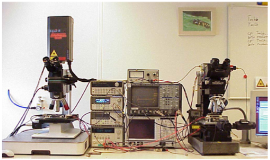
- Classical (or Black box) cryptanalysis: only uses the cryptographic primitives inputs and outputs, *e.g* the plaintexts, ciphertexts for block ciphers
- Physical attacks: additionally take advantage of physical specificities in the implementations
  - Probing attacks
  - Side-channel attacks
  - Fault insertion attacks
  - ...

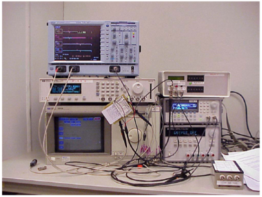
Physical attacks

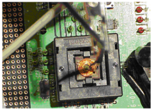






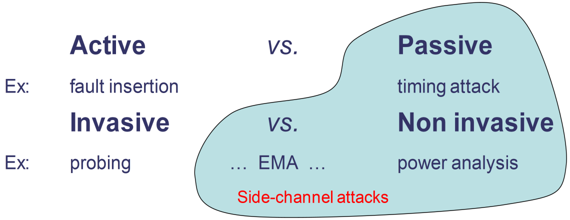






Classification of physical attacks

- According to the type of attack



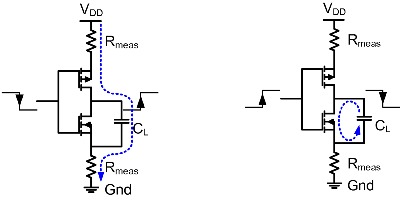
- According to the strength of the adversary: common criteria, FIPS 140-2, IBM taxonomy, . . .

Side-channel attacks

- Take advantage of physical leakages such as timing information (1996), power consumption (1998), electromagnetic radiation (2001), cache hits/misses (2005), branch predictions (2006), ...
- Continuous problem: there is a "certain" amount of information that is leaked ⇒ difficult to model
- By contrast probing and fault attacks are discrete problems: a wire can/cannot be read, a fault can/cannot be inserted ⇒ easier to model

Origin of the leakages

- *e.g.* Dynamic power consumption in CMOS devices



$$P_{dyn} \propto C_L \cdot V_{DD}^2 \cdot f_{op} \cdot P_{0\to 1}$$

- $P_{0\to 1} \Rightarrow$ data dependent physical leakage
- But $\not\Rightarrow$ P_{dyn} is the only source of information

Origin of the leakages

- *e.g.* EM radiation in CMOS devices

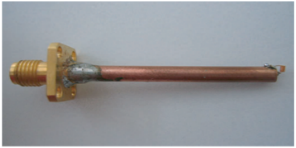
$$d\mathbf{B} = \frac{\mu I d\mathbf{l} \times \widehat{r}}{4\pi r^2}$$

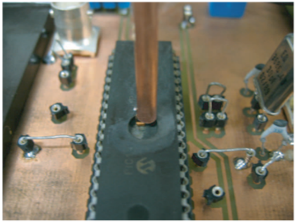
- Data dependent current intensity
    - As for the power consumption
- Field orientation depends on the current direction

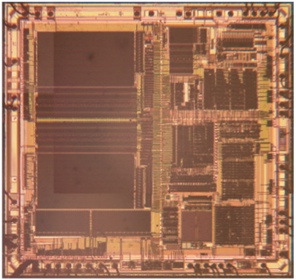
Measurement setups

- Target device: smart card ASIC, FPGA, . . .
- Measurement circuit: resistor inserted in supply circuit, small antenna (hand made coil), . . .
- Digital oscilloscope (1 Gsample/s)

Measurement setups

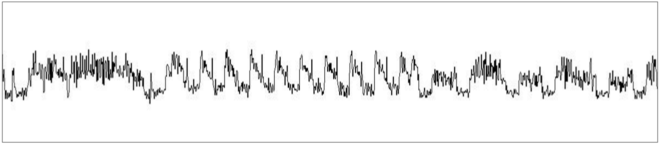






SPA

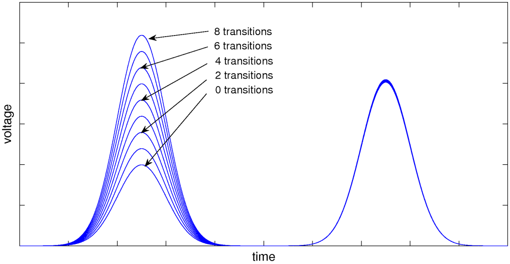
- Operation dependent leakage variations
- Example: AES encryption, 10 rounds



- Not an attack in itself for block ciphers
    - Preliminary step before other attacks
- May be very powerful (*e.g.* public key cryptography)

DPA

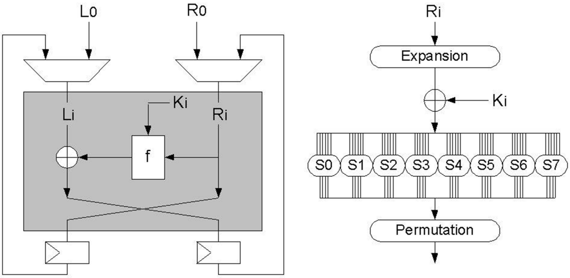
- Data dependent leakage variations



- *e.g.* CMOS: power consumption dependent on the number of bit switches within the target device

Exemplary attack against the DES

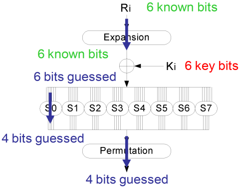
- The Data Encryption Standard
- FPGA implementation, loop architecture



(a) DES (b) f function

Exemplary attack against the DES

1. Input selection: random plaintexts
2. Internal values derivation
3. Leakage modeling (Hamming weights)



Key[0…5]

Ri		0	1	**2**	3
	0	5	12	**7**	2
	1	9	0	**12**	6
	2	14	4	**1**	13
	3	7	5	**5**	8
	4	3	10	**15**	1

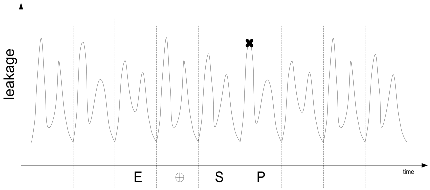
Key[0…5]

Ri		0	1	**2**	3
	0	2	2	**3**	1
	1	2	0	**2**	2
	2	3	1	**1**	3
	3	3	2	**2**	1
	4	2	2	**4**	1

- How to avoid any physical attack? {. . .}

Exemplary attack against the DES

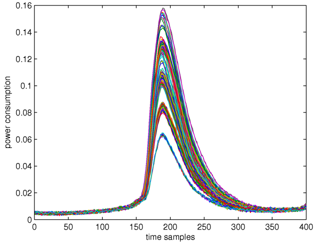
4. Leakage measurement
5. Leakage reduction (select representative samples)

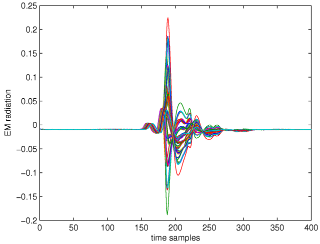


Ri	R(*L*)
0	1.675
1	1.432
2	1.221
3	1.498
4	1.937

Exemplary attack against the DES

- In practice, power consumption *vs.* EM radiation



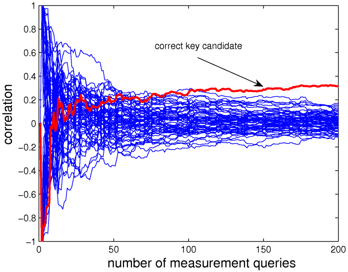


Exemplary attack against the DES

6. Statistical test
   - e.g. correlation coefficient

Key[0...5]	0	1	**2**	3
corr	-0.09	0.05	**0.3**	-0.11

$$\mathrm{corr}(M, L) = \frac{\sum_{m\in\mathcal{M}, l\in\mathcal{L}} \left(m - \overline{M}\right) \cdot \left(l - \overline{L}\right)}{\sqrt{\sum_{m\in\mathcal{M}} \left(m - \overline{M}\right)^2 \cdot \sum_{l\in\mathcal{L}} \left(l - \overline{L}\right)^2}}$$



- How to recover other bits of the master key? {. . .}

Example

- $\{\ldots\}$

Improved attacks

Improved attacks

- Improved measurement setups
  - Or combine different channels (*e.g.* power, EM)
- Adaptive selection of the inputs
- Pre-processing of the traces (*e.g.* averaging, filtering)
- Improved leakage models by profiling, characterization
- Exploitation of multiple samples, multivariate statistics
  - Higher-order attacks
  - Template attacks
- Different statistical tests
  - Difference of mean
  - Correlation analysis
  - Bayesian classification

Improved attacks

- Example: univariate template attack
  - Optimal statistical test
  - Profiled leakage model
  - Most powerful type of attack
  - (specially when extended to the multivariate case)
- Mainly identical to the previous attack
  - Only 3 steps vary...

Improved attacks

0. Preparation of the leakage model
   - Assume Gaussian noise:

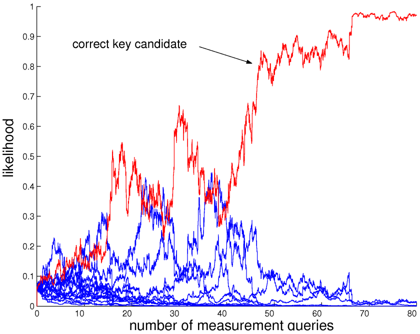
$$\mathcal{N}(\mathrm{R}(l_i)|\mu_{v_i}, \sigma_{v_i}) = \frac{1}{\sigma_{v_i}\sqrt{2\pi}} \exp\frac{-(\mathrm{R}(l_i) - \mu_{v_i})^2}{2{\sigma_{v_i}}^2}$$

   - Estimate the means μ_{v_i}'s and variances σ_{v_i}'s for each intermediate value v_i from N_t leakage traces

3. Leakage modeling: $\hat{\Pr}[\mathrm{R}(l_i)|v_i] = \mathcal{N}(\mathrm{R}(l_i)|\hat{\mu}_{v_i}, \hat{\sigma}_{v_i})$
   - In place of Hamming weights

Improved attacks

6. Statistical test: $\tilde{k} = \underset{k^*}{argmax} \prod_{i=1}^{q} \hat{\Pr}[R(l_i)|x_i, k^*]$



Countermeasures

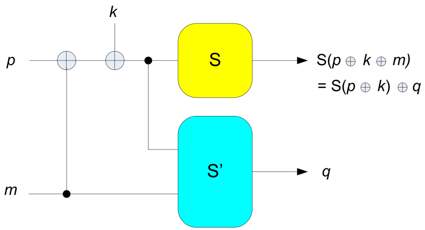
Countermeasures

- Never perfect (only make the attack harder)
- Can be implemented at different abstraction levels:
  - Physical (e.g. noise addition, decoupling C)
  - Technological (e.g. dual-rail logic styles)
  - HW/SW design (e.g. time/data randomization)
  - Algorithmic/protocol (e.g. key updates)

- To balance with implementation cost!
- Next: two typical examples

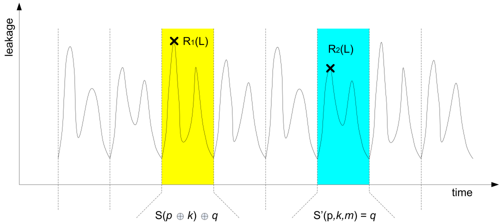
Countermeasure 1: masking

- Goal: have data-independent leakage
- How: by "randomizing" the computation
- *e.g.* block cipher S-box



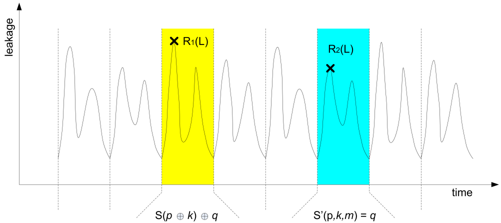
Countermeasure 1: masking

- $\mathrm{R}_1(L) \perp\!\!\!\perp k$, $\mathrm{R}_2(L) \perp\!\!\!\perp k$



Countermeasure 1: masking

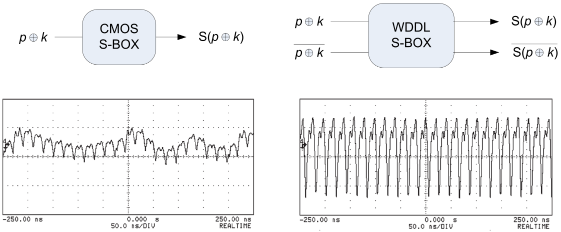
- $\mathrm{R}_1(L) \perp\!\!\!\perp k$, $\mathrm{R}_2(L) \perp\!\!\!\perp k$



- But $\exists f$ such that $f(\mathrm{R}_1(L), \mathrm{R}_2(L)) \propto k$
  - Univariate $\rightarrow$ bivariate
  - The rest of the attack remains unchanged

Countermeasure 2: dual-rails

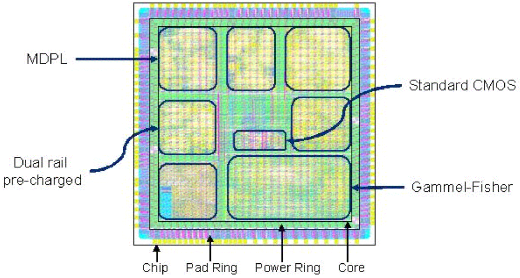
- Goal: have data-independent leakage
- How: by forcing constant leakage
- *e.g.* WDDL logic style



Countermeasure 2: dual-rails

- Hamming weight/distance models seem meaningless
- But $\exists$ data dependent leakage variations
- $\exists f$ such that $\mathrm{R}(L) \propto f(p, k)$
- An efficient attack may require to
  - Change the leakage model
    - But possibly involves a $\neq$ adversarial context
  - Use device-independent attacks

Countermeasures: cost



Key independence

- $\{\ldots\}$
- Under the assumptions that:

Key independence

- $\{\ldots\}$
- Under the assumptions that:
  - Plaintexts are uniformly distributed
  - $\mathsf{L}_t(x_i, k) = \mathsf{f}(x_i \oplus k) \neq \mathsf{f}(x_i, k)$

Asymptotic equivalences

- $\{\ldots\}$
- Under the additional assumption that:

Asymptotic equivalences

- $\{\ldots\}$
- Under the additional assumption that:
  - $\mathsf{L}_t(x_i, k) = \delta(x_i, k) + n$,
  - with n normally distributed, identical $\forall t$'s and independent of the data manipulated
  - The same models are used by all distinguishers

Summary

- Practical attacks (against real world devices)
- Device specific $\Rightarrow$ less generic but usually more powerful than black box attacks
- $\exists$ a wide variety of statistical tools, leakage models, ...
- Key independence can make evaluations easier
- Distinguishers can asymptotically equivalent in certain contexts (e.g. "standard univariate DPA")
- Attacks can be sophisticated, combined with other (computational) cryptanalytic techniques

Thanks