

# Secure implementations: side channel attacks and countermeasures

**Ingrid Verbauwhede**

ingrid.verbauwhede-at-esat.kuleuven.be

K.U.Leuven, ESAT- SCD - COSIC  
Computer Security and Industrial Cryptography  
[www.esat.kuleuven.be/cosic](http://www.esat.kuleuven.be/cosic)



Acknowledgements:

**Benedict Gierlic, Elke Demulder, Patrick Schaumont**

**Funding: EU, FWO, IUAP, K.U.Leuven, NSF, SRC**

## Outline & Goal

- Part I: Introduction and overview of side-channel attacks
  - Implementation of security and secure implementations
  - Basic concepts
  - Overview side-channel attacks
- Part II: Countermeasures

## Security for Embedded Systems

“Researcher has a new attack for embedded devices  
Vulnerability lies in ARM and XScale microprocessors”  
Computerworld – security  
April 4, 2007  
How: Use JTAG interface

“Secustick gives false sense of security”  
April 12, 2007  
<http://tweakers.net/reviews/683>  
Security completely broken



KUL - COSIC

Indocrypt 2007 – part I - 3

Chennai, December 2007

## Security for Embedded Systems

“Devices That Tell On You: The Nike+iPod Sport Kit”  
T. Saponas, J. Lester, C. Hartung, T. Kohno  
<http://www.cs.washington.edu/research/systems/privacy.html>  
Dec. 2006  
-Tracks up to 60 feet = 20 meter  
-No privacy measures included



[[www.apple.com](http://www.apple.com): nike+ipod]



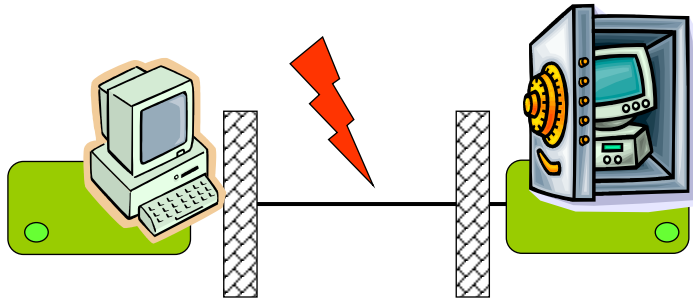
Tracking a Nike+iPod Sport Kit user

KUL - COSIC

Indocrypt 2007 – part I - 4

Chennai, December 2007

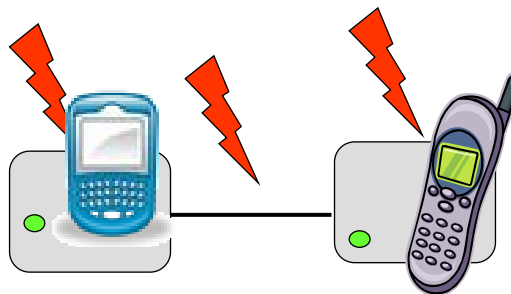
## Embedded Security



### Old Model (simplified view):

- Attack on channel *between* communicating parties
- Encryption and cryptographic operations in *black* boxes
- Protection by strong mathematic algorithms and protocols
- Computationally secure

## Embedded Security



### New Model (also simplified view):

- Attack channel *and* endpoints
- Encryption and cryptographic operations in *gray* boxes
- Protection by strong mathematic algorithms and protocols
- Protection by secure implementation

**Need secure *implementations* not only algorithms**

## Embedded Security

### NEED BOTH

- Efficient Implementation

- Within power, area, timing budgets
- Public key: 1024 bits RSA on 8 bit  $\mu\text{C}$  and 100  $\mu\text{W}$
- Public key on a passive RFID tag



- Trustworthy implementation

- Resistant to attacks
- Active attacks: probing, power glitches, JTAG scan chain
- Passive attacks: monitor electromagnetic radiation



## Why a hard engineering problem?



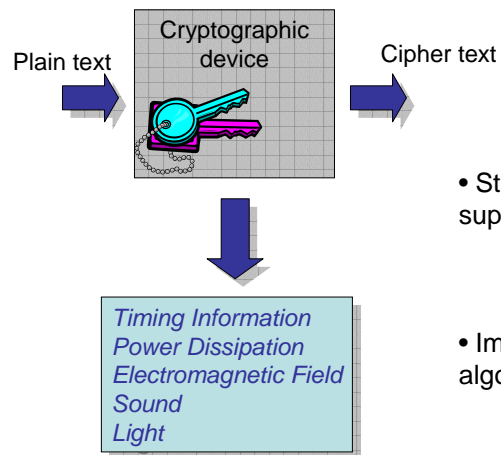
- More difficult to guarantee that something will not happen (attacks) than that something will happen.
- Engineers are trained to make something happen.

## Security as a design dimension

- Security consumes resources!
  - extra area, extra power, extra design time
  - E.g. communication – computation trade-off
- *Similar* to power or area optimization
  - Perfect security does not exist (zero-power design doesn't exist either)
  - Low-risk security does exist (low-power design does exist)
- *Different*: attacker will go for the easiest entry point:
  - If strong crypto algorithm: try side channel attack
  - Look at the JTAG interface
  - Monitor power consumption
  - Introduce glitches (= fault attacks)
  - Guess the password: Paris Hilton's dog
  - Bribe the security guard

## Basic concept side channel attack

## Concept: Side channel



- Standardized algorithms are supposed to be secure
  - *theoretically or computationally*
- Implementations of the algorithms should also be secure
  - *real-life*

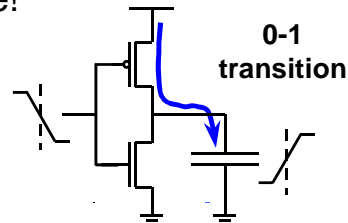
## Concept: Origin

- Due to the instruction which is executed
- Due to the data which is processed
- Due to some physical effects which are often not well understood, often called noise

## Concept: Intro to Static CMOS

- Consumes power when output makes a 0 to 1 transition
- Most popular circuit style!

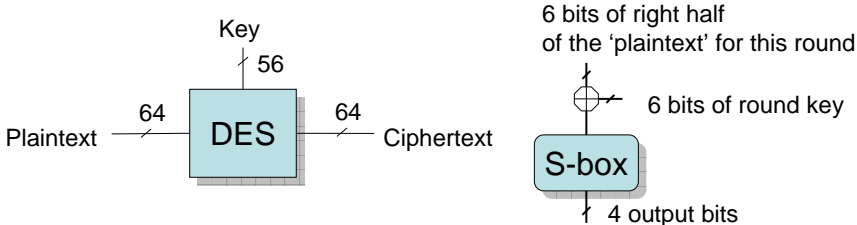
IN	OUT
0→0	0
0→1	discharge
1→0	charge
1→1	0



## Concept: Intermediate values (1)

- For secure algorithms it is 'impossible' to find the plaintext from the cipher text if the key is not known
- If key = n bytes long, exhaustive key search would mean trying  $2^n$  keys, because all cipher text bits depend on all key bits and all plain text bits
- Intermediate variables often depend on only a few key bits and plain text bits, if the intermediate variables are known, an 'exhaustive' key search is more feasible
- Design principle of ciphers: strong because a sequence of many small weaker steps

## Concept: Intermediate values (2)



- Iterated block cipher
- Only  $2^6$  sub-keys possible: exhaustive search feasible
- Strong cipher composed of weak components

*Divide and conquer (divide et impera)*

## Classification of attacks

## Classifications of Attacks

- Active versus passive
  - Active: power glitches or laser pulses
  - Passive: electromagnetic radiation
- Invasive versus non-invasive
  - Invasive: bus probing
  - Non-invasive: power measurements
- Side channel : passive and non-invasive
  - Very difficult to detect
  - Often cheap to set-up
  - Mostly: need lots of measurements
- Analysis capabilities
  - “Simple” attacks: one measurement – visual inspection
  - “Differential” and “higher”: multiple measurements – signal processing

## Active attacks

*Probing data: read out data from bus or individual cells directly*

- micro-probing
  - probe the bus with a very thin needle
  - Several needles concurrently
- SEM (Scanning electron microscope)
  - e-beam probing

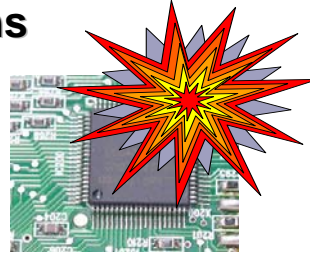
*Modify circuits*

- Connect or disconnect security mechanism
  - disconnect security sensors
  - RNG stuck at a fixed value
  - reconstruct blown fuses
- Cut or Paste tracks with laser or focused ion beam
- Add probe pads on buried layers

## Active Attacks: fault attacks

### Apply combinations of strange environmental conditions

- Vcc
- Glitch
- Clock
- Temperature
- UV
- Light
- X-Rays
- ...



input



error

and bypass or infer secrets

[courtesy: H. Handshuh]

## Focus: side-channel attacks

- Passive: observation of device
- Non-invasive:
  - Very difficult to detect
  - Often cheap to set-up
  - Mostly: need lots of measurements
- Mostly based on:
  - Timing attacks: e.g. cache attacks
  - Power variations
  - Electromagnetic radiation
- Also based on:
  - Acoustic: keyboard clicking

## Set-up

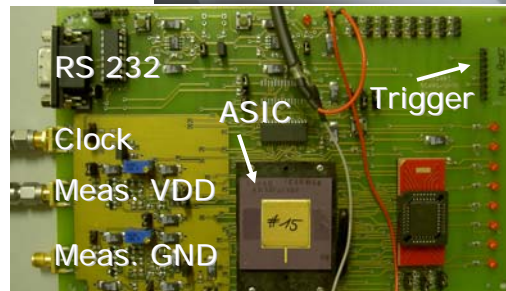
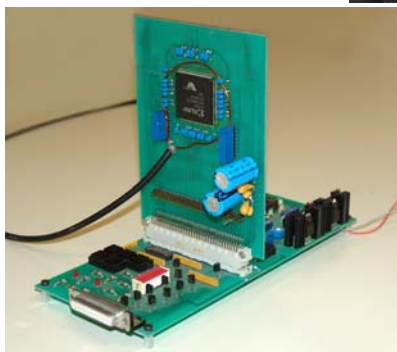
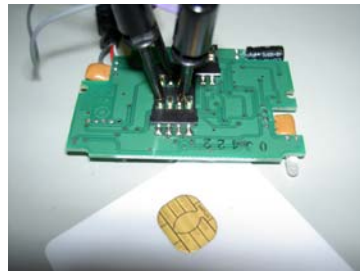
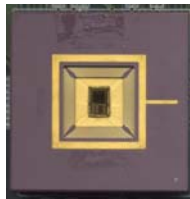
KUL - COSIC

Indocrypt 2007 – part I - 21

Chennai, December 2007

## Devices under attack

- SmartCard
- FPGA, ASIC
- Etc.



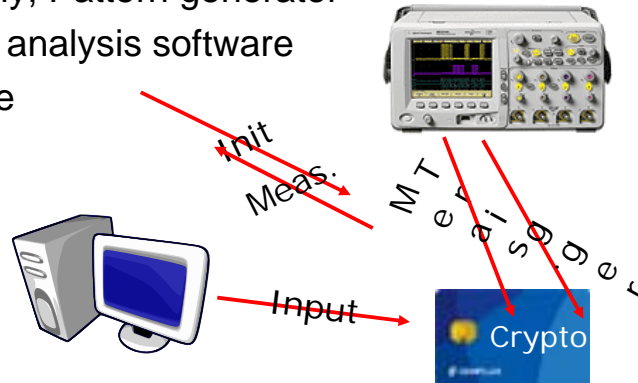
KUL - COSIC

Indocrypt 2007 – part I - 22

Chennai, December 2007

## The lab – measurement setup

- Cryptographic device under attack
- Probe, measurement circuit
- Power supply, Pattern generator
- Control and analysis software
- Oscilloscope
- PC

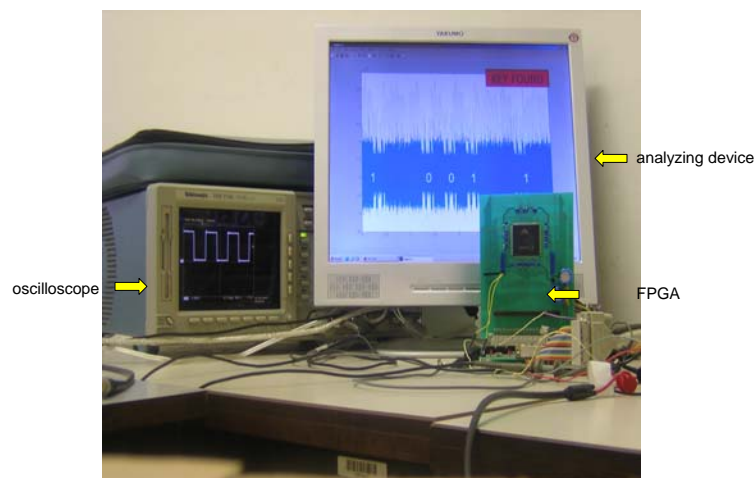


KUL - COSIC

Indocrypt 2007 – part I - 23

Chennai, December 2007

## Power Analysis: Measurement setup (1)

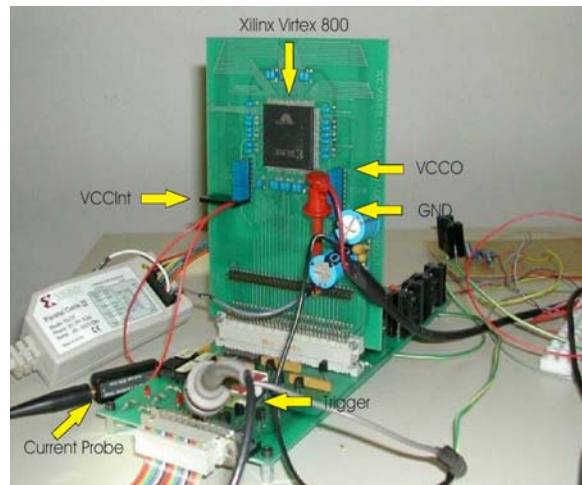


KUL - COSIC

Indocrypt 2007 – part I - 24

Chennai, December 2007

## Power Analysis: Measurement setup (2)



KUL - COSIC

Indocrypt 2007 – part I - 25

Chennai, December 2007

## Probe / Measurement circuit

- An oscilloscope can only measure voltage
  - Current flow needs to be transformed into a proportional voltage signal
- Simple resistor in series (Ohm's law:  $U = R \times I$ )
  - Measure voltage drop over the resistor
- Current probe (Current flow  $\rightarrow$  electric field)



- Dedicated measurement circuit in the design

KUL - COSIC

Indocrypt 2007 – part I - 26

Chennai, December 2007

## Practical Issues

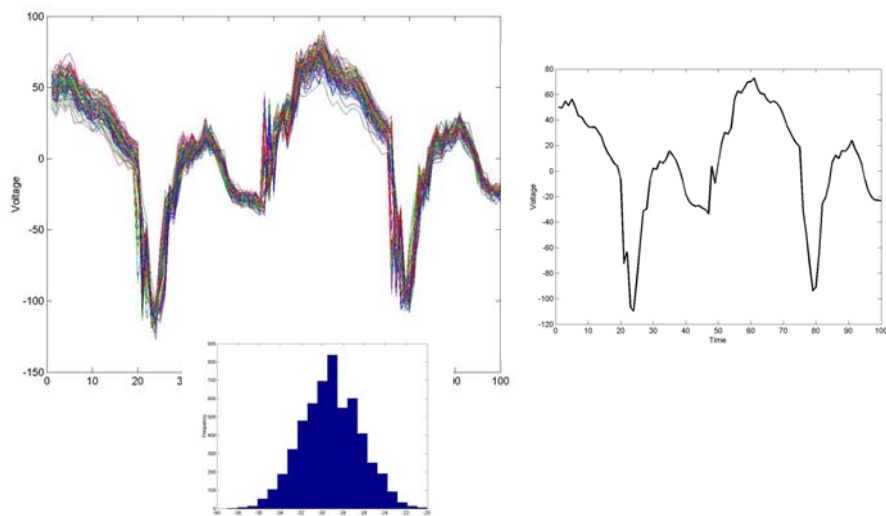
- Quality of measurements
  - Noise poses a problem
    - Instead of a deterministic value, we observe Gaussian distr.
    - Algorithmic Noise  $SNR : \sqrt{(r/m)}$   
 $r$  = number of bits under consideration  
 $m$  = total number of bits
    - External Noise (measurement setup)  $SNR : \sqrt{n}$   
 $n$  = number of measurements
    - 'Workaround': average multiple observations
  - De-synchronised measurements
    - 'Temporal noise' => re-synchronisation
  - Filter 'bad' measurements
- Amount of data
  - Processing duration, data compression

KUL - COSIC

Indocrypt 2007 – part I - 27

Chennai, December 2007

## Practical Issues



KUL - COSIC

Indocrypt 2007 – part I - 28

Chennai, December 2007

## Power Analysis

- What's the setting?
  - Power of the adversary, Device, Scenario, etc.
- Direct attacks (extract)
  - Simple Power Analysis (1999)
  - Single- / Multi-Bit Differential Power Analysis (1999/2002)
  - Correlation Power Analysis (2004)
  - Collision Attacks (2003)
- Two-step attacks (extract)
  - Template Attacks (2002)
  - Stochastic Model (2005)
  - Inferential Power Analysis (1999)

## Simple Attacks

Simple Timing  
Simple Power  
Simple Electromagnetic

## Simple Power Analysis

- Based on one or few measurements
- Discovery of data-(in)dependent properties
  - Symmetric:
    - Number of rounds (resp. key length)
    - Memory accesses (usually higher power consumption)
  - Asymmetric:
    - The key (if badly implemented, e.g. RSA / ECC) conditional operation
    - Keylength
    - Mode: for example RSA with ???
- Search for repetitive patterns

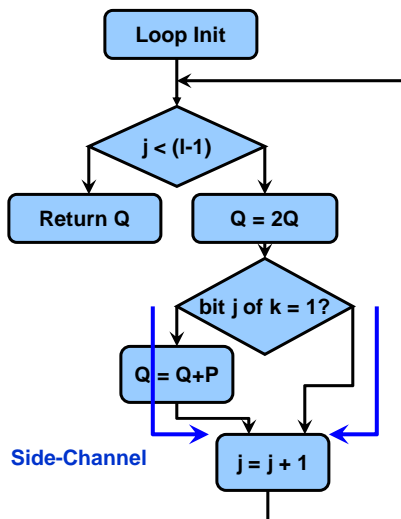
## Insecure ECC implementation

ECC point multiplication  
 In: point  $P$ , key  $k$  ( $l$  bits)  
 Output:  $Q = k.P$

```

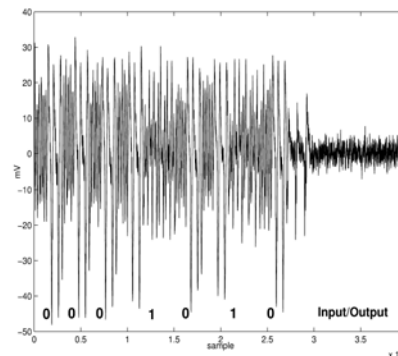
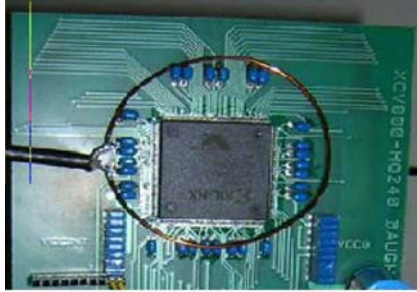
for j = 0 to l - 1
  Q = 2.Q /* double */

  if (bit j of k) is 1
  then
    Q = Q + P /* add */
Return Q
  
```



## Insecure ECC implementation

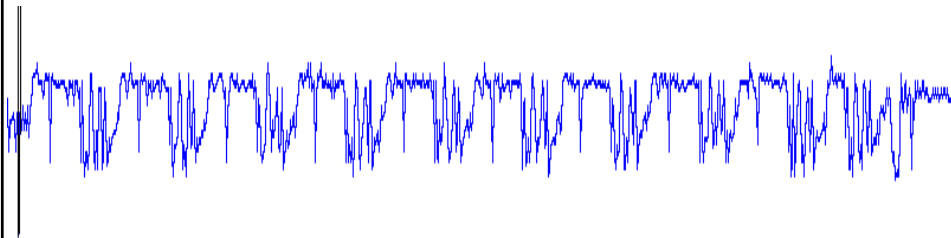
Electro Magnetic Attack:



[E. Demulder EUROCON 2005]

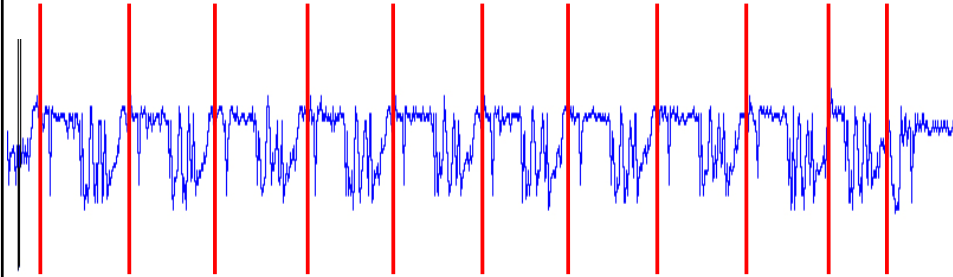
## Simple Power Analysis (AES)

- What is the keylength of this AES implementation?



## Simple Power Analysis (AES)

- 10 rounds => AES-128



KUL - COSIC

Indocrypt 2007 – part I - 35

Chennai, December 2007

## Simple Power Analysis (RSA)

- What is the private RSA exponent?

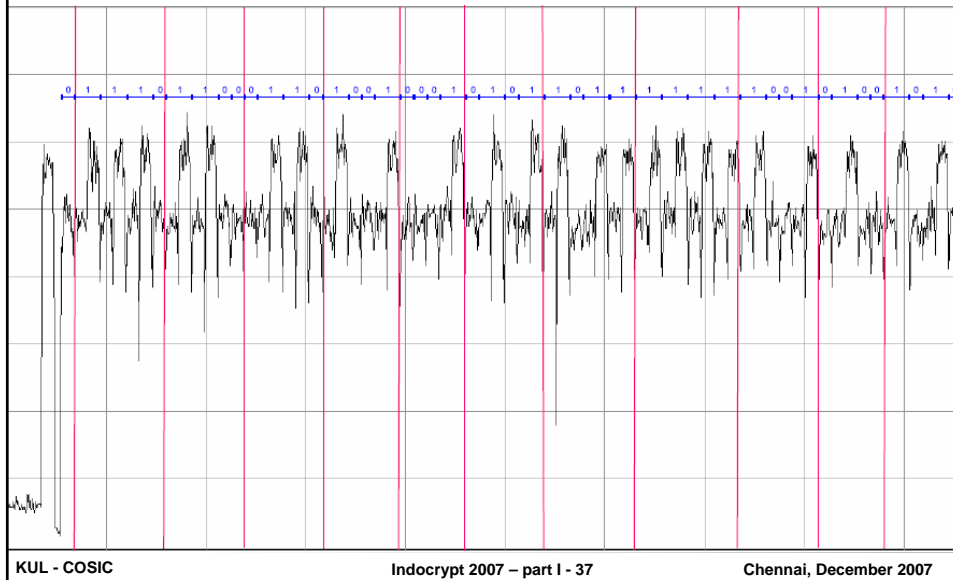


KUL - COSIC

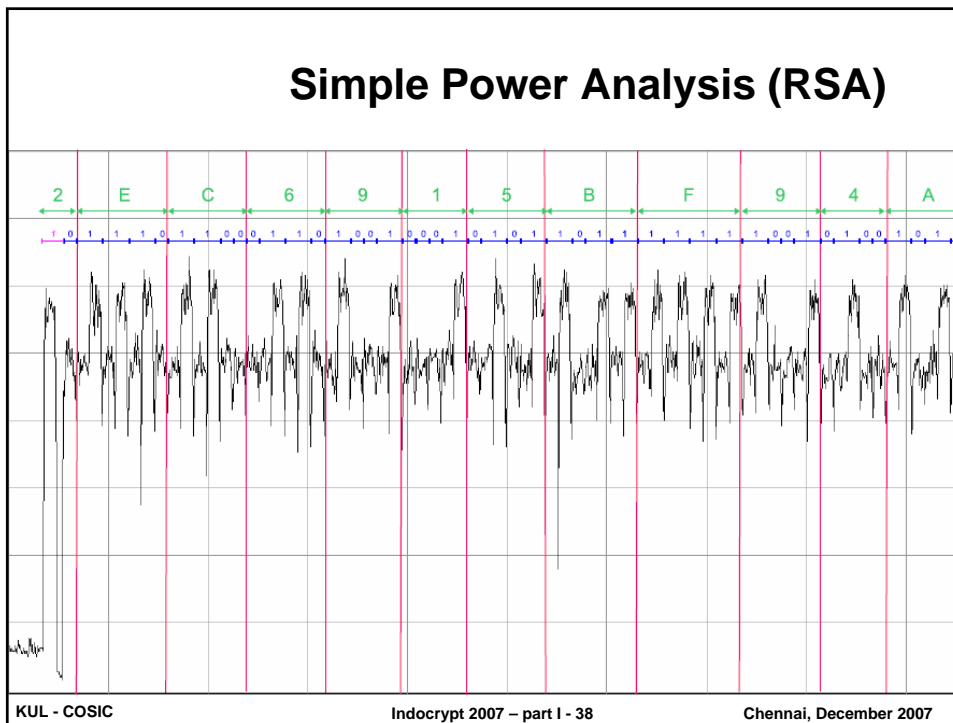
Indocrypt 2007 – part I - 36

Chennai, December 2007

## Simple Power Analysis (RSA)

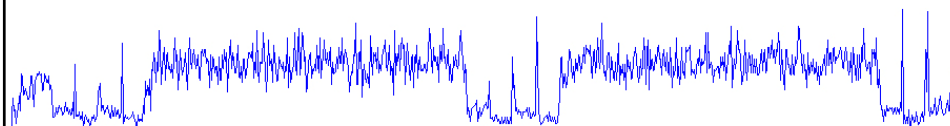


## Simple Power Analysis (RSA)



## Simple Power Analysis (RSA)

- RSA signature generation with...?



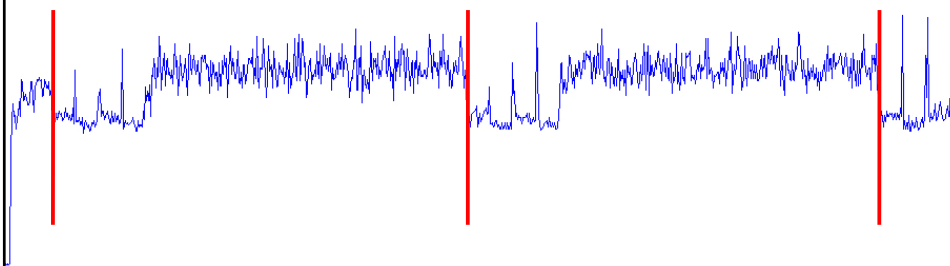
KUL - COSIC

Indocrypt 2007 – part I - 39

Chennai, December 2007

## Simple Power Analysis (RSA)

- RSA signature generation with CRT



KUL - COSIC

Indocrypt 2007 – part I - 40

Chennai, December 2007

## Differential Power Analysis

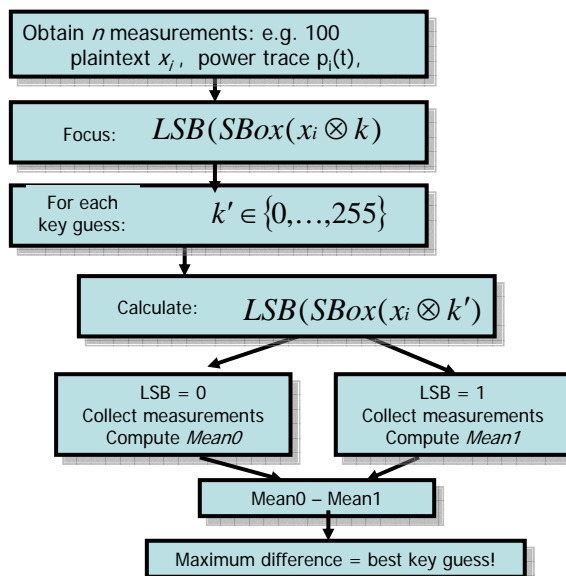
## Differential Power Analysis

- Recall: CMOS has data-dependent dynamic power dissipation (very small differences)
- Requires many measurements
  - Usually a few hundred for software implementations
  - Usually a few thousand for hardware implementations
  - Can go up to several 100k if implementation is protected
- Discovery of data-dependencies by statistical means (uni- and multivariate)
- Applies to **symmetric** and asymmetric schemes

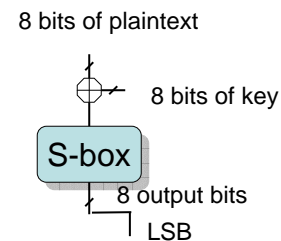
## DPA – how it works

- Obtain sufficient number ( $n$ ) of measurements
  - In general: uniform, random inputs; fixed, unknown key  $k$
- Choose an appropriate intermediate result
  - Preferably only a few bits involved (e.g. for AES the bytes are processed separately until the first MixCol operation)
  - Preferably high diffusion within these bits
  - Preferably after a non-linear transformation (e.g. Sbox)
- For each key hypothesis  $k'$ :
  - based on known plain-/ciphertext and key hypothesis  $k'$ , predict the intermediate result for each measurement
  - Apply a statistical test to reject/verify the key hypothesis
    - Here: difference of means

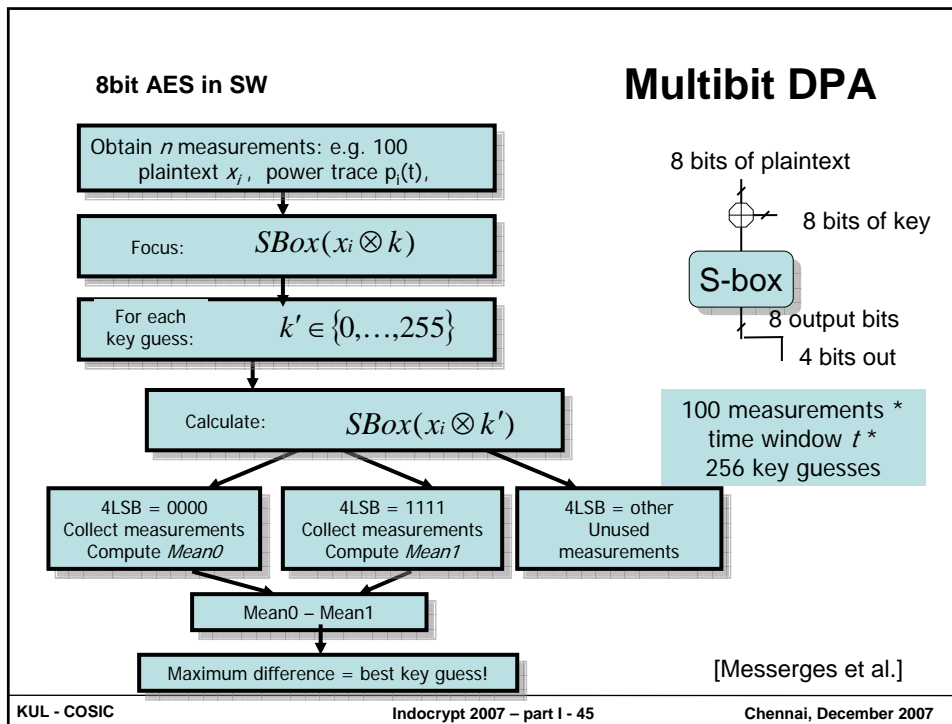
### 8bit AES in SW



### Classical 1-bit DPA



100 measurements \*  
time window  $t^*$   
256 key guesses



## Multi bit DPA

- 1 bit DPA: Bin0 = 0, Bin1 = 1, all measurements used
- 2 bit DPA, Bin0 = 00, Bin1 = 11, Bin2 = 01 or 10  
Half of measurements not used!
- 4 bit DPA, Bin0 = 0000, Bin1 = 1111, Bin2 = all the rest  
 $2^4 - 2 = 14$  out of 16 not used!
- Idea: improve bin model

“Power model” in correlation DPA

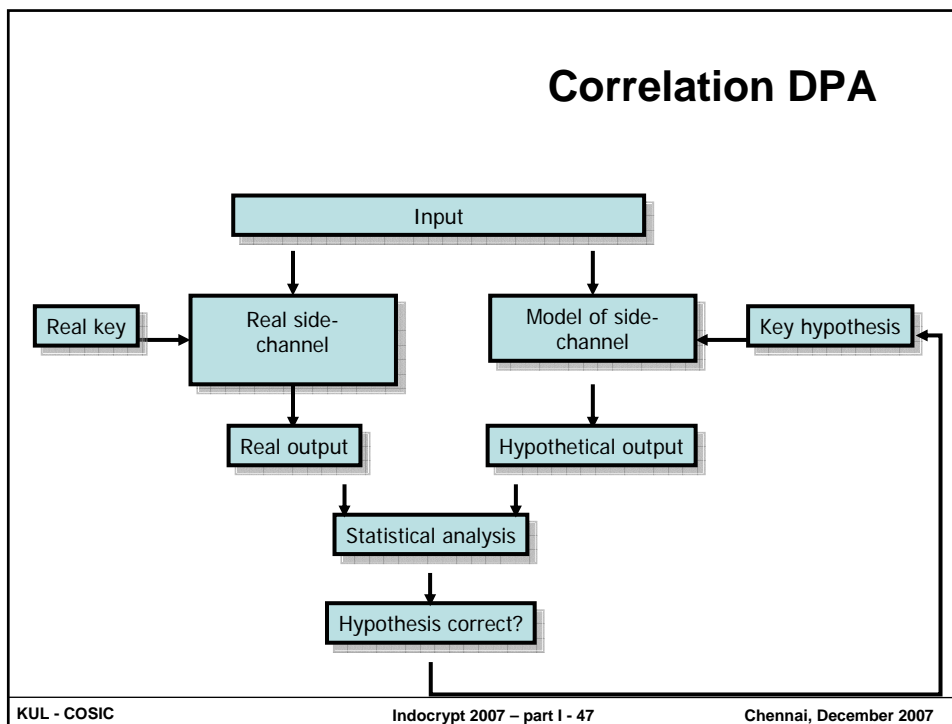
Based on an intermediate result, predict the power consumption for each measurement and correlate prediction and observation

KUL - COSIC

Indocrypt 2007 – part I - 46

Chennai, December 2007

## Correlation DPA



## Typical Power Models

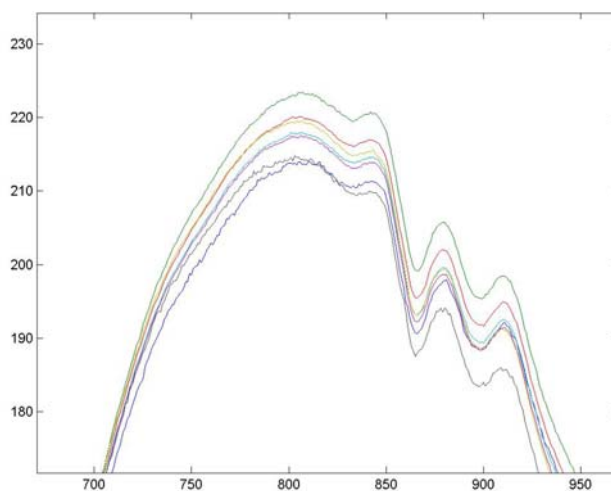
- Bit model: two bins or multiple bins
- Hamming weight model
  - Typically for pre-charged busses
- Hamming distance model
  - Typically for register outputs in ASIC's
- Weighted Hamming weight/distance model
- Dedicated models for combinational circuits
  - E.g. the ripple effect of adders or multipliers

## Example: power model bus

- 8 bit bus on a smart card, pre-charged
- (relatively) large capacitance
- Hamming weight model  
= numbers of bits set to 1

Side-note: on a pre-charged bus which is pre-set to 1, maximum power consumption is for data all zero.

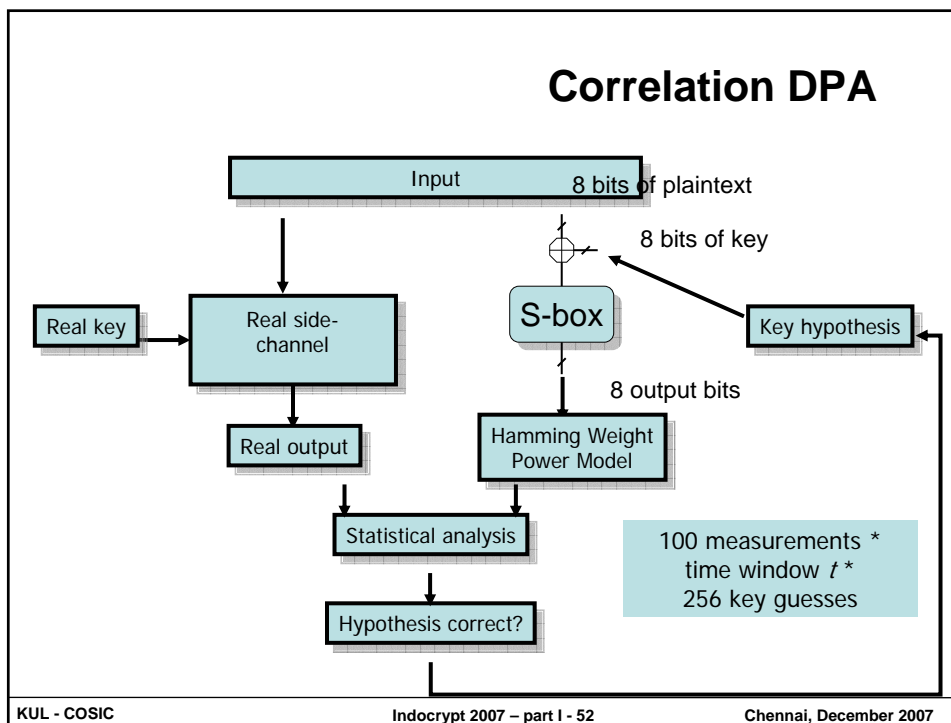
## CPA – the Hamming Weight model



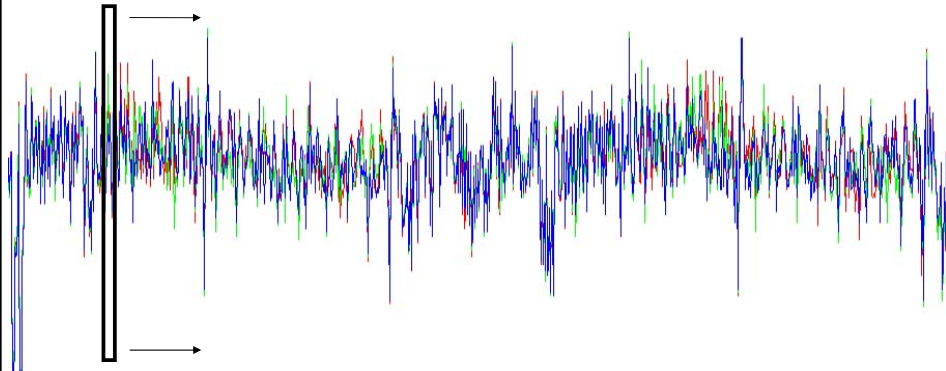
## Correlation Power Analysis

- R := reference state
  - Which bit pattern was previously present? E.g.
    - A pre-charged value
    - An opcode on the bus
    - A previously stored value in a register
- Hamming Weight := number of bits set to '1'
- Power model:  $a \times HW(SBox(x_i \otimes k') \otimes R) + b$ 
  - a,b are constant, linear model
- Statistical model: compute correlation

## Correlation DPA

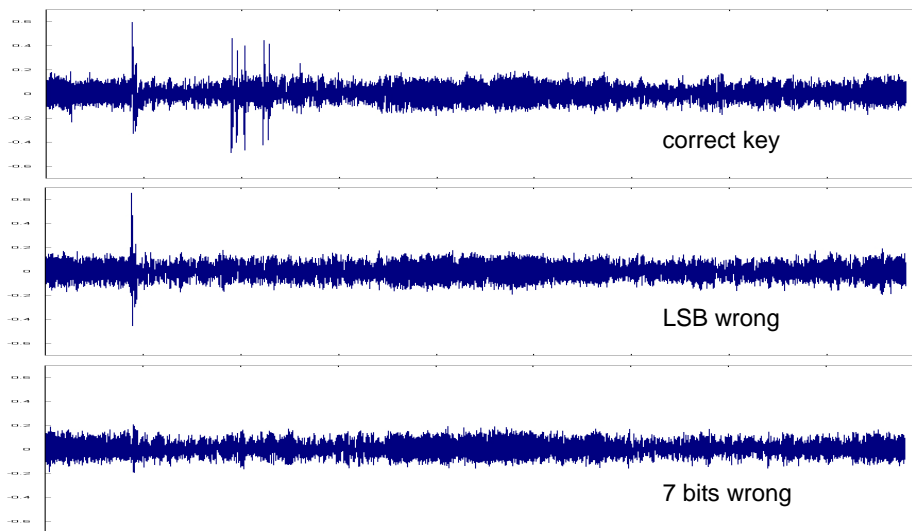


## CPA (Example: 8-bit AES, software)



- Slide the window, each position  $t$  yields one correlation coefficient

## CPA (Example: 8-bit AES, software)



## Advanced attacks

## Collision Attacks

- Usually: collision attacks on hash functions
  - Identical output for different inputs
- **Internal** collision attacks
  - Use Side Channel information to detect transient internal collisions
  - Schramm et al.: 2003 DES
  - Schramm et al.: 2004 AES
  - Novak: recover substitution tables, 2003 GSM A3/A8
  - Clavier: Scare, 2004 GSM A3/A8

## Profiling Attacks

- What can we do with a single measurement?
  - Template Attack
    - Profiling step (using an identical device)
      - Estimate a multivariate probability distribution for each possible value of the intermediate result of interest
    - Classification step (the real attack)
      - Check for which estimated probability distribution the observed measurement is most likely
    - High work load during profiling, success rate depends on profiling, can easily exceed 95%
  - Stochastic Model
    - Implements more engineering insights
    - Significantly less measurements, more computation
    - Doesn't reach precision of Templates, but very practical

**End of Part 1.**