

Introduction to Power Analysis

Benedikt Gierlichs

KU Leuven – COSIC, Belgium

benedikt.gierlichs@esat.kuleuven.be



Summer School on
Design and Security of Cryptographic
Functions, Algorithms and Devices

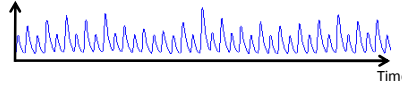
Albena, Bulgaria, 4 July 2013

Agenda

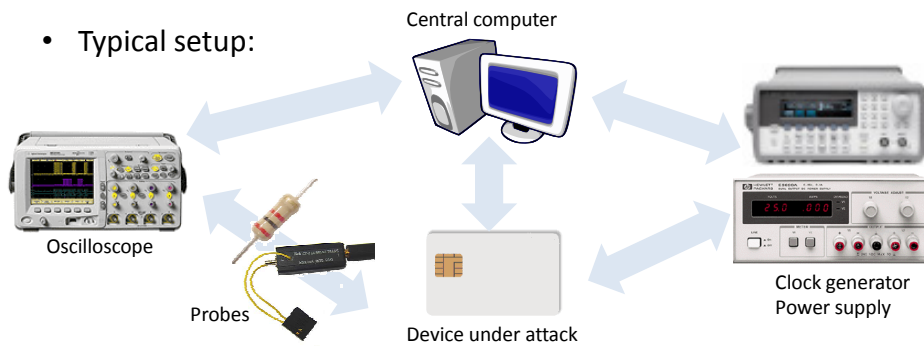
- Measuring power consumption
- Power analysis (exploration of power traces)
- Power analysis attacks (revealing secrets)
- Differential power analysis attacks: overview
- Practical problems
- Summary

Measuring power consumption

- Not average power over time, not peak power
- Instantaneous power over time
 - Trace or curve, many samples

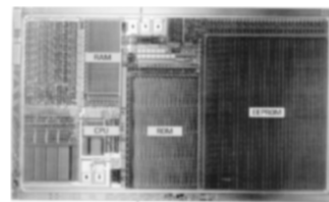


- Typical setup:

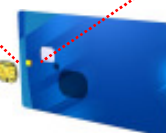
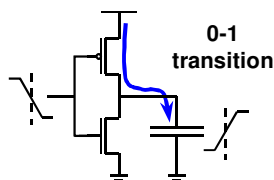


Measuring power consumption (2)

- Logic: constant supply voltage, supply current varies
- Predominant technology: CMOS
 - Low static power consumption
 - Relatively high dynamic power consumption
 - Power consumption depends on input
- CMOS inverter:



Input	Output	Current
0 0	1 1	Low
0 1	1 0	Discharge
1 0	0 1	Charge
1 1	0 0	Low



Measuring power consumption (3)

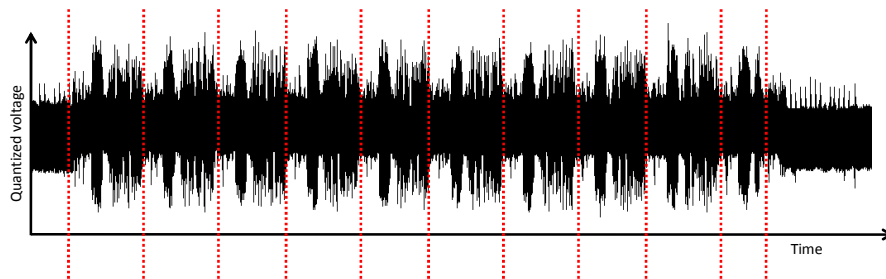
- Oscilloscope can only measure voltage
 - Generate voltage signal, proportional to current
- Measure in VDD or GND line
 - Resistor (Ohm's law: $U = R \times i$), measure U over resistor
 - Current probe: current field voltage
 - Dedicated measurement circuits
- Measure 'global' E or H field of the device
 - Field intensity proportional to power consumption
 - Field orientation depends on current direction



Power analysis

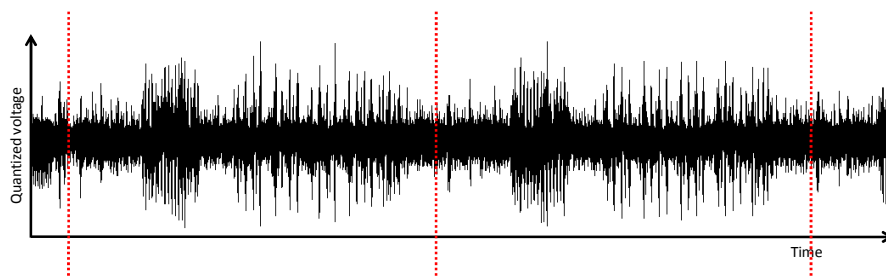
- What can we see looking at a curve?
- Information in:
 - Repetitive patterns: typically coarse, structure of algorithm and implementation (e.g. loops)
 - Time: what happens when, program flow
 - Amplitude: what happens at a given moment in time, data flow
 - the same operation, executed with different operand values, consumes more or less power
- Examples: trace inspection

Power analysis (2)



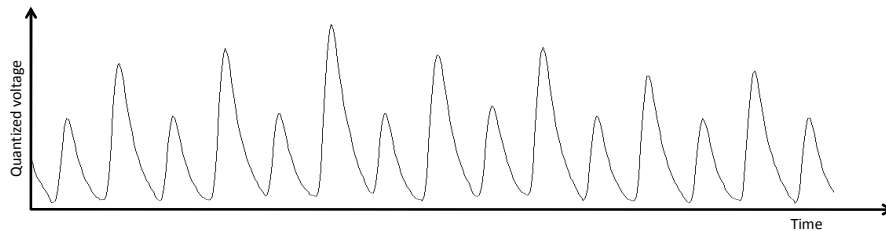
- Unprotected software implementation of AES-128 on 8-bit μC
 - Ten rounds, last round shorter, without MixColumns

Power analysis (3)



- Unprotected software implementation of AES-128 on 8-bit μC
 - Two rounds, four AES building blocks look different

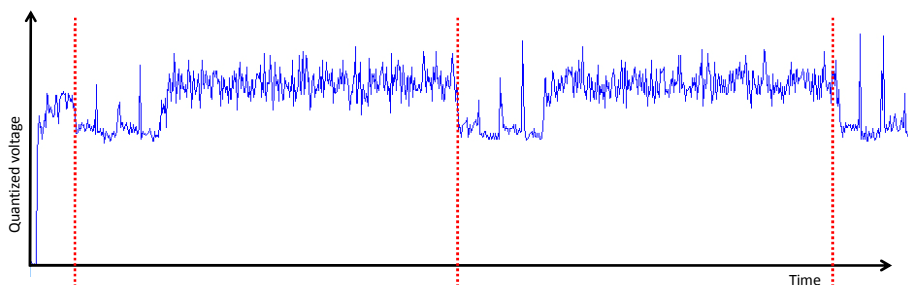
Power analysis (4)



- Few clock cycles on 8-bit μC
 - Capacitive charge and discharge effect visible in every clock cycle
 - Loading and unloading capacitors in the circuit
 - wires, input/output capacitances, parasitic capacitances, etc.
 - Amplitude depends on operation and operand value(s)

Power analysis (5)

- RSA signature generation with CRT



From power analysis to power analysis attacks

- If sequence of patterns, timing or amplitude depends on secret values, power analysis attacks can possibly reveal the secrets [JO05]
- Taxonomy: attacks categorized according to approach, requirements, adversarial power, etc.
- Categories and criteria not 100% clear, definitions vary, transitions are smooth

Power analysis attacks

[KJJ99]

- Simple power analysis (SPA) attacks
- Internal collision attacks
- Differential power analysis (DPA) attacks
- Orthogonal: ad-hoc (non-profiled) versus profiled
 - Non-profiled: little prior knowledge about how the device leaks and noise distribution, relies on assumptions
 - Profiled: more or less precise profiling of the leakage behaviour and noise distribution, typically training of a classifier (curve key-related information)

Simple power analysis attacks

- Anything but simple (except in examples :-))
- Visual inspection of few traces, worst/best case: single shot
- Often exploitation of direct key dependencies, input and output data need not be known (but they are useful for verification)
- Require: expertise, experience, detailed knowledge about target device and implementation
- Examples: patterns, amplitude, timing

Simple power analysis attacks (2)

- Patterns (many-cycle sequences) show, e.g.:
 - Symmetric crypto algorithms:
 - Number of rounds (resp. key length), loops
 - Memory accesses (sometimes higher power consumption)
 - Asymmetric crypto algorithms:
 - Key (if badly implemented, e.g. RSA / ECC)
 - Key length
 - Implementation details (e.g. RSA with CRT)
- Search for repetitive patterns

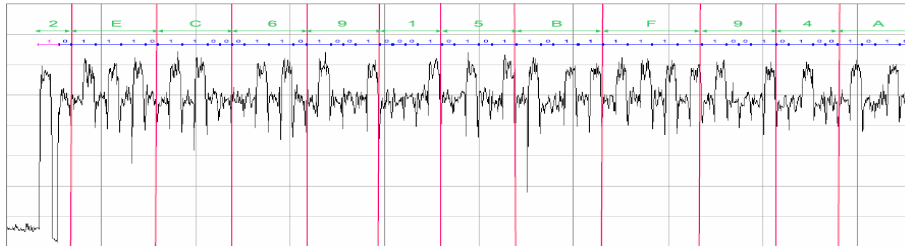
RSA decryption, $M = C^d \bmod N$
with $d = d_{n-1}d_{n-2}\dots d_0$

```
x = C
for j = n-2 to 0
  x = x2 mod N
  if dj == 1 then
    x = xC mod N
  end if
end for
return M = x
```

conditional operation

Simple power analysis attacks (3)

- Example: RSA exponentiation $M = C^d \bmod N$
- Crypto coprocessor optimized for squaring



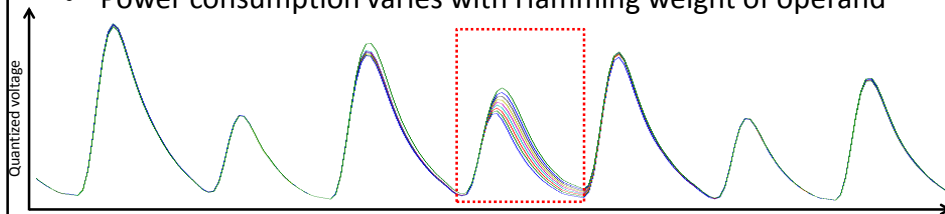
[courtesy: C. Clavier]

Simple power analysis attacks (4)

- Amplitude in a cycle can show:
 - Exact operand values (extreme case)
 - Often: Hamming weight or Hamming distance of operand(s)
 - Can greatly reduce key space
 - Operation being executed (software, microcontroller)
 - Reverse engineering of implementation details
 - Reverse engineering of e.g. proprietary algorithms (SCARE attacks)
- Typically requires device profiling to have a reference

Simple power analysis attacks (5)

- Example: a MOV instruction with different operand values
- Power consumption varies with Hamming weight of operand



- Suppose we have a 'dictionary' that translates power consumption values into Hamming weights
- Example: SPA attack on the AES key schedule [M02]
 - Extract HWs of round keys, generate list of suitable round keys
 - Requires 1 plaintext/ciphertext pair to check remaining candidate keys

Simple power analysis attacks (6)

- Timing, e.g. when/if an operation is executed, can show:
 - Data-dependent branches in software implementations
 - If branch condition does not only depend on key but on intermediate result, one also needs to know input (output)

- Example: a bad implementation of AES MixColumns [KQ99]

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

- Per output byte a couple of XORs and multiplication by 2 (in Rijndael's Galois field)

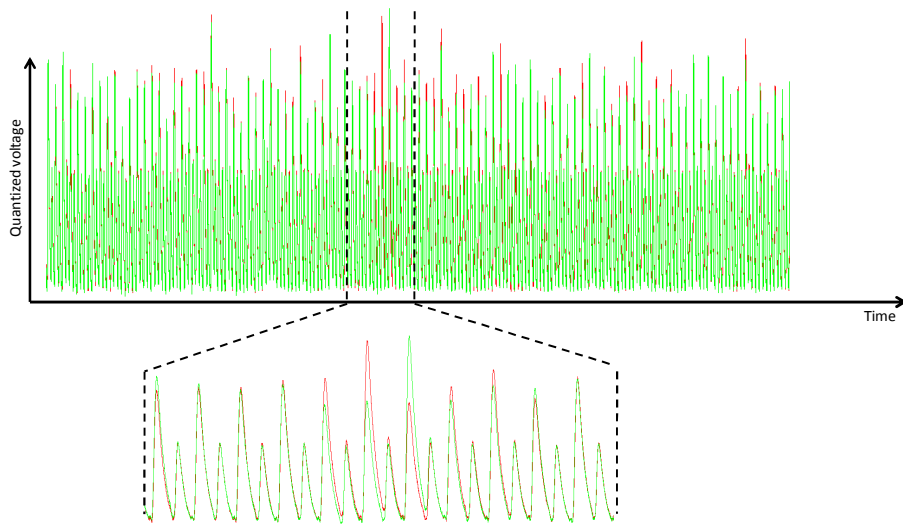
[DR98]

- Multiplication by 2 in GF(2⁸):
 - Multiplication by 2 (left shift)
 - **Conditionally**: if carry is set, perform reduction (XOR with 0x1b)



- Execution time depends on MSB of a_i

SPA on a bad implementation of AES MixColumns



Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

19

Internal collision attacks

- Collision: a key-dependent intermediate result takes the same value for two different inputs: $f(\text{input1}, \text{key}) = f(\text{input2}, \text{key})$
- Detection:
 - Collision not visible in output, hence internal collision
 - If a collision occurs, the curves corresponding to the two inputs should be 'similar' at time/points where collision is expected
 - Statistical methods detect this, e.g. least-squares test, correlation
- Exploitation: relatively simple cryptanalysis
 - Exploit occurrence and absence of collisions
 - Possibly adaptively chosen inputs



[SWP03] (DES) and [SLFP04] (AES)

Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

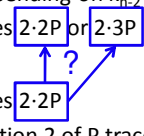
20

Internal collision attacks (2)

- Collision persists: for short up to long interval
 - Single intermediate result, long sequence of intermediate results
 - Typically: the longer, the easier to detect
 - One needs to know where to look for collision
- Extensions: collisions in two or more different intermediate results, one or multiple traces
 - $f_1(\text{input}_1, \text{key}) = f_2(\text{input}_1, \text{key})$ with $f_1 \neq f_2$
 - $f_1(\text{input}_1, \text{key}) = f_2(\text{input}_2, \text{key})$ with $\text{input}_1 \neq \text{input}_2$
 - ...
 - Requires shifting the traces before comparison

Internal collision attacks (3)

- Example for public-key crypto: ECC
 - ECC scalar multiplication kP usually works on the binary expansion of k ($k_{n-1}, k_{n-2}, \dots, k_1, k_0$)
 - A sequence of point doublings and point additions
- The doubling attack
 - To find out what happened in iteration i , test which values are computed in iteration $i+1$
 - First trace: input P
 - Iteration 1: P $2P$ or P $3P$ depending on k_{n-2}
 - Iteration 2: the doubling computes $2 \cdot 2P$ or $2 \cdot 3P$
 - Second trace: input $2P$
 - Iteration 1: the doubling computes $2 \cdot 2P$
 - Compare that to doubling in iteration 2 of P trace



[FV03]



Differential power analysis attacks

- Recall: divide and conquer principle
 - Block ciphers: strength from a sequence of many 'weak' steps
 - Intermediate results often depend only on a few key bits
 - Recover the secret in several small chunks
 - Problem: no access to weak intermediate results \perp
- Recall CMOS: power consumption of an operation varies with the operand value(s) intermediate results 'leak'
- Variation relatively small, not directly observable
 - Statistics detect weak signals

Differential power analysis attacks (2)

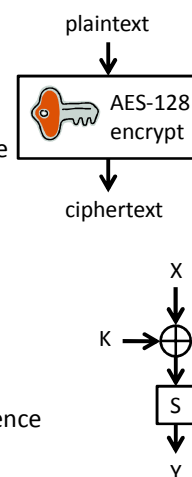
- Differential attacks use statistics to exploit the data-dependent variations of the power consumption
- ~50 to millions of traces
- Input or output of implementation need to be known (typically)
- Require little knowledge about target device and implementation (but extra knowledge helps!)
- Weak adversary + strong attack = highly relevant

Differential power analysis attacks (3)

- Three disciplines:
 - Cryptanalysis: target a sensitive intermediate result for which exhaustive key search is feasible
 - Engineering: access to good side channel measurements
 - Statistics: an "oracle" to verify key hypotheses
- Working principle:
 - Take a set of traces with varying inputs
 - Select sensitive intermediate variable
 - For each key hypothesis
 - Compute hypothetical values of intermediate, sort curves into subsets
 - Compute difference between the subsets
 - Intuition: wrong key guesses random subsets, no difference 
 correct key guess correct subsets, difference 

Differential power analysis attacks (4)

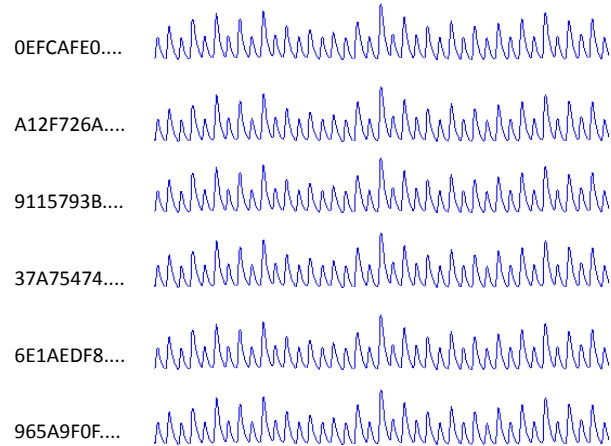
- Example: classical 1-bit DPA on AES-128 encryption
- Select $Y = f(X, K)$ in implementation
 - Until first MixColumns, each byte of state depends on one plaintext byte and one key byte
 - Target S-boxes, recover key byte-by-byte
 - Here sensitive intermediate variable: $\text{LSB}(Y)$
- For each possible value of K , here $[0..255]$
 - Compute Y for each input and check if $\text{LSB}(Y) = 0$ or $= 1$
 - Group curves in two subsets
 - Compute mean curves for both subsets, then their difference
- Analyse the differential curves
 - For correct guess of K , differential curve shows peaks at point(s) in time when selected bit is manipulated



Differential power analysis attacks (5)

Plaintexts

Traces



Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

27

Note

- Usually not mentioned but important for beginners
- The adversary typically does not know **when** the targeted intermediate value is computed
- Analyze all time samples (typically separately) in the same way
- Search over time samples and possible key values
- Some advanced attacks analyze multiple time samples jointly

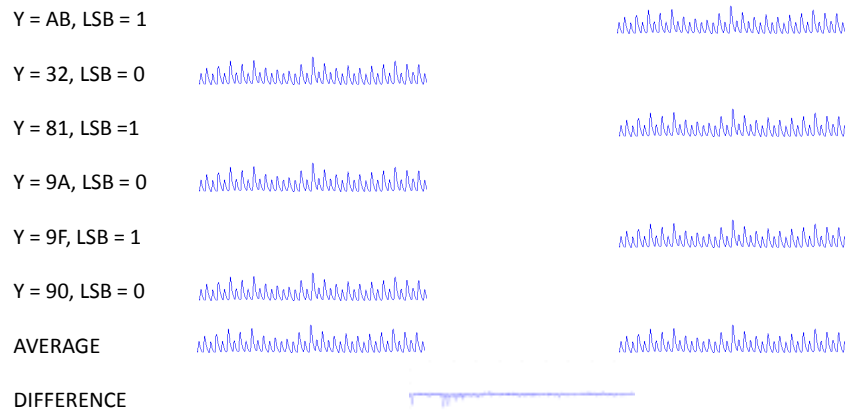
Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

28

Differential power analysis attacks (6)

- Attack on first key byte in round 1 of AES-128
- If $K = 00$



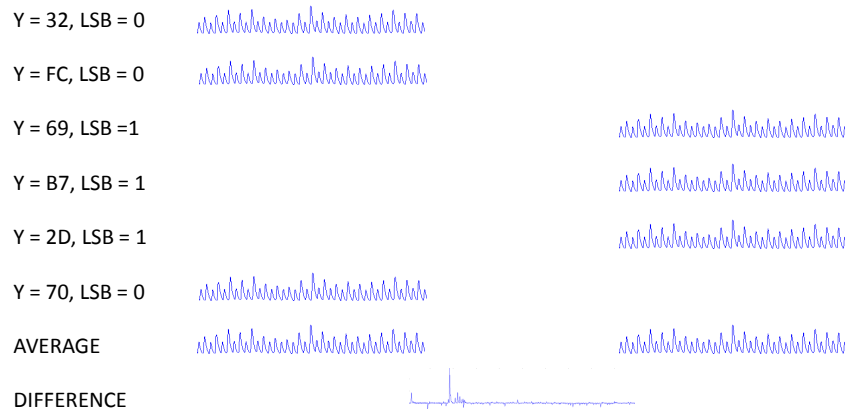
Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

29

Differential power analysis attacks (7)

- Attack on first key byte in round 1 of AES-128
- If $K = 2B$



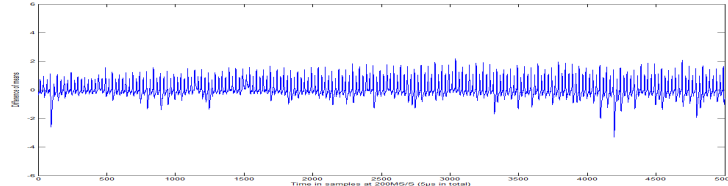
Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

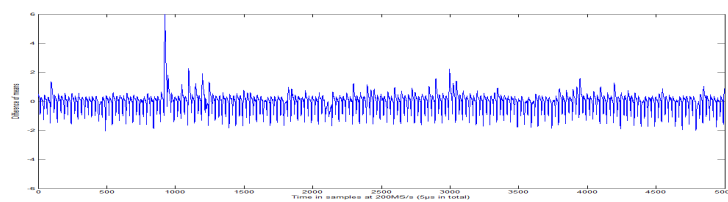
30

Differential power analysis attacks (8)

- Differential trace for a wrong hypothesis on K

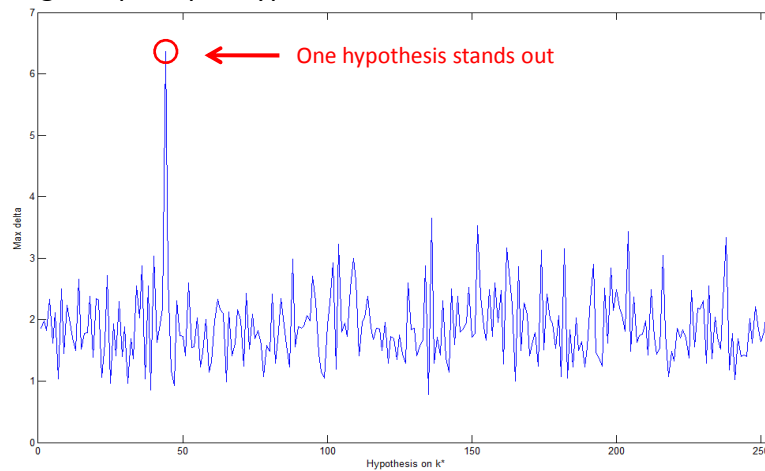


- Differential trace for correct hypothesis on K



Differential power analysis attacks (9)

- Highest peak per hypotheses on K



Differential power analysis attacks (10)

- x_i input for encryption i
- $p_i(t)$ power trace i
- k' denotes the targeted sub-key
- $f_{k'}(x_i)$ denotes the chosen intermediate result
- d_i is one bit of $f_{k'}(x_i)$, e.g. the LSB

$$\Delta_{k'} = \frac{\frac{\sum_{i=1}^n (1 - d_i) \cdot p_i(t)}{\sum_{i=1}^n (1 - d_i)}}{\frac{\sum_{i=1}^n d_i \cdot p_i(t)}{\sum_{i=1}^n d_i}}$$

Average power $d_i = 0$ Average power $d_i = 1$

Differential power analysis attacks (11)

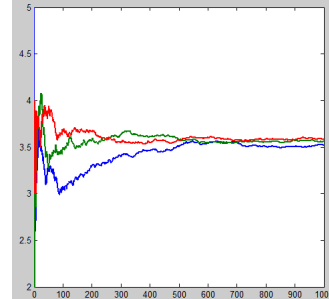
- For the correct key hypothesis the predicted d_i are correct
 - The groups for $d_i=0$ and $d_i=1$ are different
 - Averaging reveals the footprint of the target bit in the power traces
- For a wrong key hypothesis the subsets are random
 - The groups for $d_i=0$ and $d_i=1$ both represent $d_i=0.5$

$$\Delta_{k'} = \frac{\frac{\sum_{i=1}^n (1 - d_i) \cdot p_i(t)}{\sum_{i=1}^n (1 - d_i)}}{\frac{\sum_{i=1}^n d_i \cdot p_i(t)}{\sum_{i=1}^n d_i}}$$

Average power $d_i = 0$ Average power $d_i = 1$

Differential power analysis attacks (12)

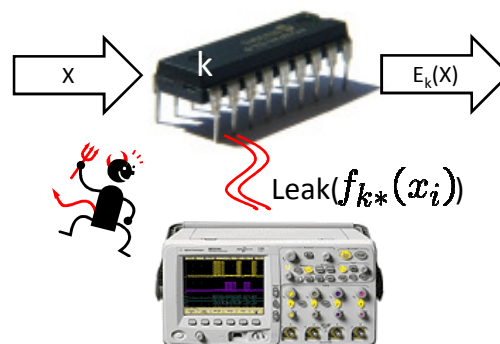
- Noise:
 - Electrical, quantization, ambient, etc.
 - The more traces (larger n), the better
 - Law of large numbers: estimation error of the averages decreases by \sqrt{n}



Six-sided dice, three experiments: average approaches 3.5

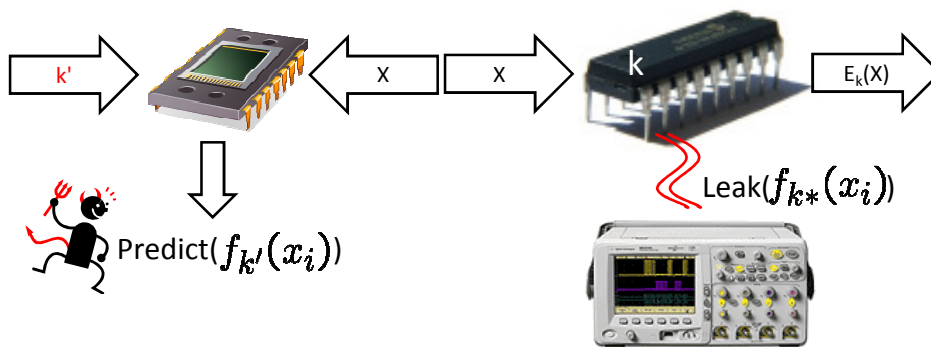
- This approach to DPA is easy to explain, but often not the best
 - Ignores the parallel activity of other bits (algorithmic noise)
 - Simple statistics: not error prone, but large n required

Modern view of differential attacks



- Observe power consumption of targeted intermediate value $f_{k*}(x_i)$, multiple executions on varying input x_i

Modern view of differential attacks



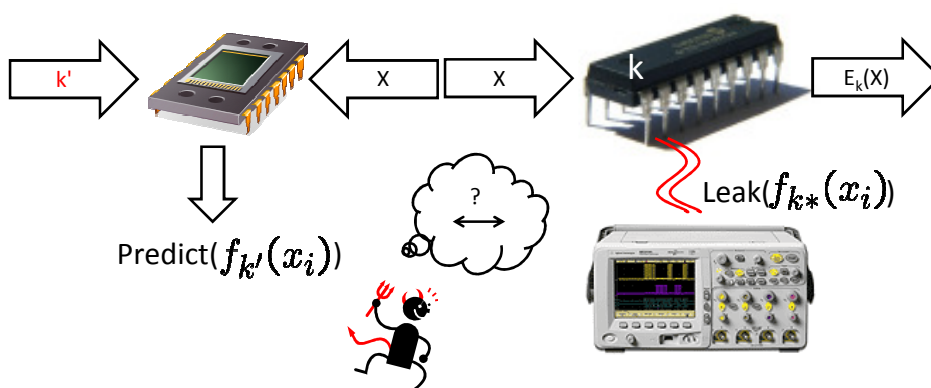
- Build a model to predict 'power consumption'
 $\text{Predict}(f_{k'}(x_i))$ parameterized by guess k' on the secret k^*

Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

37

Modern view of differential attacks



- For each k' , evaluate statistical dependence between $\text{Predict}(f_{k'}(x_i))$ and $\text{Leak}(f_{k^*}(x_i))$ with some distinguisher
- Correct guess $k' = k^*$ should yield strongest dependency

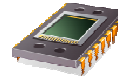
Albena, 04.07.2013

Summer School on Design and Security - Benedikt Gierlichs

38

Differential attacks: overview

- Power models: better model → more powerful attack
 - More precise model requires to know or assume more details
 - Bad model → unsuccessful attack (≠ device is secure)
 - Often: Hamming weight or distance of operand value(s), single bits
- Distinguishers: close link to power models
 - Should focus on and exploit properties of power model
 - Should tolerate some errors in power model
 - Often: Difference of means, Pearson correlation
- Trade-off: efficiency (# traces) versus generality
 - Recently: generic attacks, e.g. using mutual information (MIA)



[BCO04]



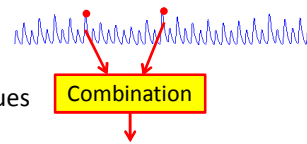
[GBTPO8]

Differential attacks: overview (2)

- Profiled attacks: related to machine learning
 - Profiling: training of a classifier (e.g. for $k=0, k=1, \dots, k=255$)
 - Typically using an 'open' clone device
 - Online attack: classify traces from target device
 - Rank key candidates, e.g. assign probabilities
 - Can exploit multiple points per trace, multivariate statistics
 - Typically require selection of points of interest (feature selection)
- Examples:
 - Template Attack [CRR02]: typically precise density estimation
 - Stochastic Model [SLP05]: approximation
 - Online attack: both maximum likelihood
 - Stochastic approach less precise, but more efficient profiling (# traces)

Differential attacks: overview (3)

- Second-order attacks: break masked implementations [M00, CJRR99]
 - Two intermediate results protected by the same mask
 - Each individually secure, but they leak jointly
 - Analyze pairs of points in each trace
 - Combination function: maps pairs to single values
 - Perform usual attacks on pre-processed traces
- Combination function:
 - Strong impact on success of attack
 - No generic combination function (?), loss of information [S+10]
 - Way out: no combination function, multivariate processing
 - E.g. multivariate MIA, profiled attacks
- Special case: both intermediate results processed in parallel
 - Traces contain already combined information, univariate but second-order



Practical problems in power analysis attacks

- Measurement quality
 - Noise, bandwidth, sampling frequency, amplitude resolution, etc.
 - Improve measurement setup
- Temporal de-synchronization
 - Unstable trigger points, etc.
 - Trace alignment
- Amount of measurement data
 - Many and long curves, processing time
 - Trace compression
- In the real world: secure devices have countermeasures!

Summary

- Attacks begin with measurements
 - Measurement quality is important
- Power traces are a rich source of information
 - Repetitive patterns, timing, amplitude
- Simple power analysis attacks
- Internal collision attacks
- Differential power analysis attacks
- Orthogonal: ad-hoc versus profiled attacks

Thank you for your attention!



Bibliography

- [JO05] M. Joye, F. Olivier: Side-channel analysis, Encyclopedia of Cryptography and Security, 2005
- [KJJ99] P. Kocher, J. Jaffe, B. Jun: Differential power analysis, CRYPTO 1999
- [M02] S. Mangard: A Simple Power-Analysis (SPA) Attack on Implementations of the AES Key Expansion, ICISC, 2002
- [DR98] J. Daemen, V. Rijmen: AES proposal Rijndael, 1998
- [KQ99] F. Koeune and J.-J. Quisquater: A timing attack against Rijndael, UCL Crypto Group technical report CG-1999/1, 1999
- [SLFP04] K. Schramm, G. Leander, P. Felke, C. Paar: A Collision-Attack on AES Combining Side Channel- and Differential-Attack, CHES, 2004
- [FV03] P.-A. Fouque, F. Valette: The Doubling Attack - Why Upwards Is Better than Downwards, CHES, 2003
- [BCO04] E. Brier, C. Clavier, F. Olivier: Correlation power analysis with a leakage model, CHES, 2004
- [GBTPO8] B. Gierlichs, L. Batina, P. Tuyls, B. Preneel: Mutual information analysis, CHES, 2008

Bibliography

- [CRR02] S. Chari, J.R. Rao, P. Rohatgi: Template Attacks, CHES, 2002
- [SLP05] W. Schindler, K. Lemke, C. Paar: A Stochastic Model for Differential Side Channel Cryptanalysis, CHES 2005
- [M00] T.S. Messerges: Using second-order power analysis to attack DPA resistant software, CHES, 2000
- [CJRR99] S. Chari, C.S. Jutla, J.R. Rao, P. Rohatgi: Towards sound approaches to counteract power-analysis attacks, CRYPTO, 1999
- [S+10] F.-X. Standaert, N. Veyrat-Charvillon, E. Oswald, B. Gierlichs, M. Medwed, M. Kasper, S. Mangard: The world is not enough: Another look on second-order DPA, ASIACRYPT, 2010
- [SWP03] K. Schramm, T. Wollinger, C. Paar: A New Class of Collision Attacks and Its Application to DES, FSE 2003