Power Analysis – an overview

Benedikt Gierlichs
KU Leuven – COSIC, Belgium
benedikt.gierlichs@esat.kuleuven.be

COSIC

KU LEUVEN

Summer School on
Design and security of cryptographic algorithms
and devices for real-world applications

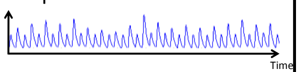
Šibenik, Croatia, 5 June 2014

Agenda

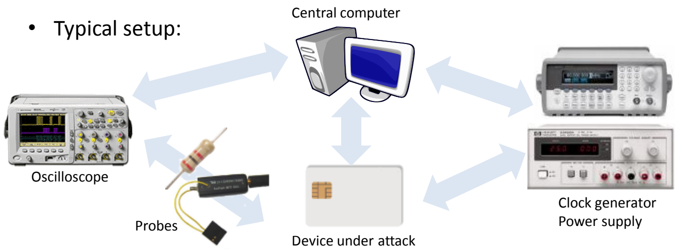
- Measurements
- Analysis
- Pre-processing
- Attacks
- Evaluation
- Countermeasures

Measuring power consumption

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

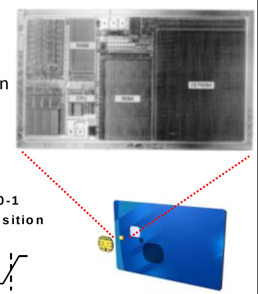


- Typical setup:

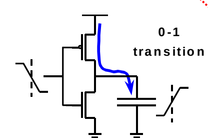


Measuring power consumption (2)

- 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:

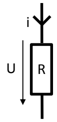


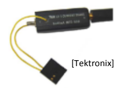
Input	Output	Current
0 → 0	1 → 1	Low
0 → 1	1 → 0	Discharge
1 → 0	0 → 1	Charge
1 → 1	0 → 0	Low



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 x 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

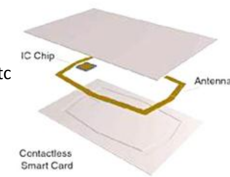



[Tektronix]


[Rohde&Schwarz]

Measuring power consumption (4)

- Contactless (passive RFID)
  - Public transport ticket, access control, etc.
  - Electronic passport, contactless credit card, etc
- Harvest energy from field supplied by reader
  - No immediate access to power lines
    - Would require "opening" the device, tamper evidence
- Measure how much power RFID took from field
  - Best with analogue processing



[gemsecuritysystem.com]

[KOP09, KOP11, OP11]

Measuring power consumption (5)

- What matters?
- Noise: will typically increase number of measurements required (see countermeasures later)
  - Intrinsic, ambient, quantization, countermeasures, etc.
- Bandwidth
  - How much is enough? Is sampling rate limiting factor? Probes etc.
- Sampling rate
- Trigger point
  - Stable trigger point simplifies many attacks

Power analysis attacks

- If power consumption "patterns" depend on secret values, power analysis attacks can possibly reveal the secrets [JO05]
- Simple power analysis (SPA) attacks [KJJ99]
- Differential power analysis (DPA) attacks
- Internal collision attacks
- Algebraic side channel attacks [RSV09]
- Orthogonal: non-profiled (ad-hoc) versus profiled
  - Non-profiled: little prior knowledge about how the device leaks and noise distribution, relies on assumptions
  - Profiled: profiling of the leakage behaviour and noise distribution, typically training of a classifier; machine learning; feature selection

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
- Example: patterns

Simple power analysis attacks (2)

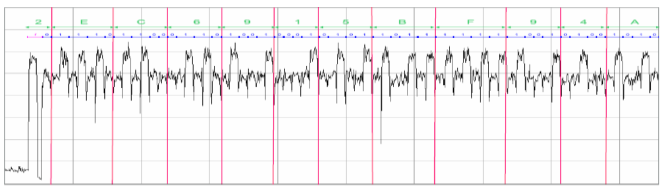
- Patterns (many-cycle sequences) show, e.g.:
  - Symmetric crypto algorithms:
    - Number of rounds (resp. key length), loops
    - Memory accesses (sometimes higher power consumption)
  - Asymmetric crypto algorithms:
    - Key (if badly implemented, e.g. RSA / ECC)
    - Key length
    - Implementation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

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

```
x = 1
for j = n-1 to 0
    x = x² mod N
    if d_j == 1 then      [conditional operation]
        x = xM mod N
    end if
end for
return S = x
```


Simple power analysis attacks (3)

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



[courtesy: C. Clavier]

Internal collision attacks

- Collision: a key-dependent intermediate result takes the same value for two different inputs: f(input1,key) = f(input2,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)

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
  - $f_1(input_1,key) = f_2(input_1,key)$ with $f_1 \neq f_2$
  - $f_1(input_1,key) = f_2(input_2,key)$ with $input_1 \neq input_2$
  - ...
  - 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 $(k_{n-1}, k_{n-2},\ldots,k_1,k_0)$
  - 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 2·2P or 2·3P
  - Second trace: input 2P
    - Iteration 1: the doubling computes 2·2P ?
    - 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 ☹
- 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

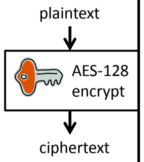
Differential power analysis attacks (2)

- Differential attacks use statistics to exploit the data-dependent variations of the power consumption
- ~50 to millions of measurements
- 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

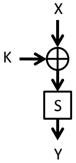
Differential power analysis attacks (3)

- 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)



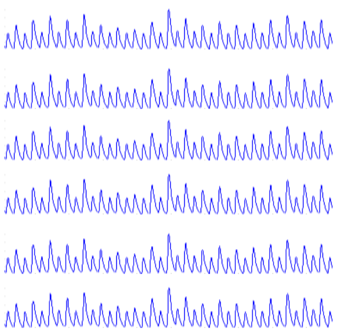
- Example: classical 1-bit DPA on AES-128 encryption
- Select Y = f(X,K) in implementation
  - Until first MixColumns, each byte of state depends on one plaintext byte and one key byte
  - Target S-boxes, recover key byte-by-byte
  - Here sensitive intermediate variable: LSB(Y)
- For each possible value of K, here [0..255]
  - Compute Y for each input and check if 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

0EFCAFE0....
A12F726A....
9115793B....
37A75474....
6E1AEDF8....
965A9F0F....



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

Differential power analysis attacks (6)

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

Y = AB, LSB = 1

Y = 32, LSB = 0

Y = 81, LSB =1

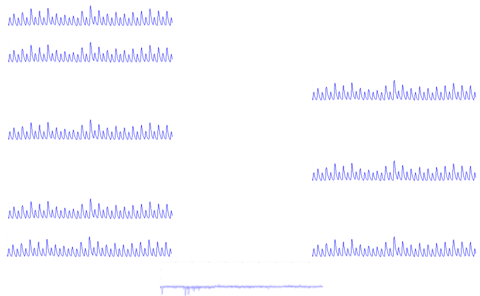
Y = 9A, LSB = 0

Y = 9F, LSB = 1

Y = 90, LSB = 0

AVERAGE

DIFFERENCE



Differential power analysis attacks (7)

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

Y = 32, LSB = 0

Y = FC, LSB = 0

Y = 69, LSB =1

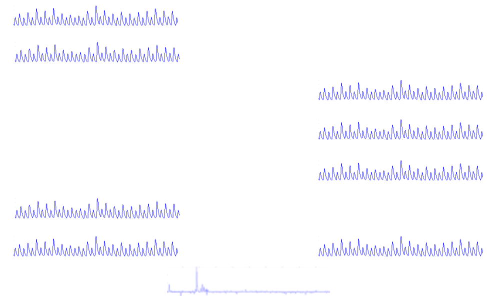
Y = B7, LSB = 1

Y = 2D, LSB = 1

Y = 70, LSB = 0

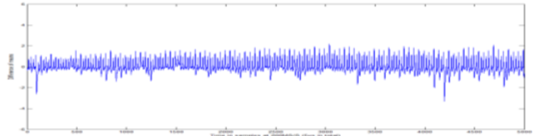
AVERAGE

DIFFERENCE

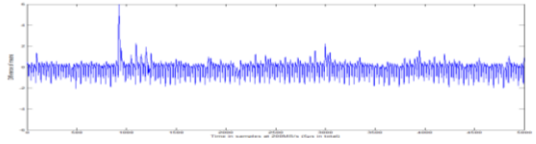


Differential power analysis attacks (8)

- Differential trace for a wrong hypothesis on K

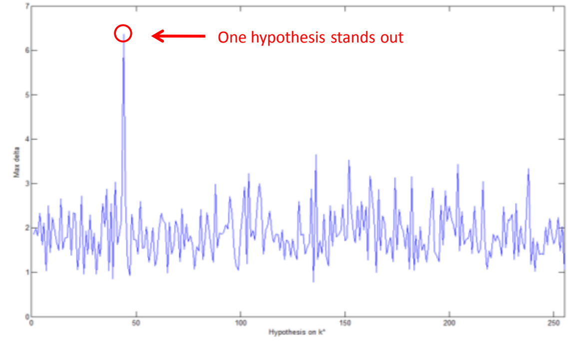


- Differential trace for correct hypothesis on K

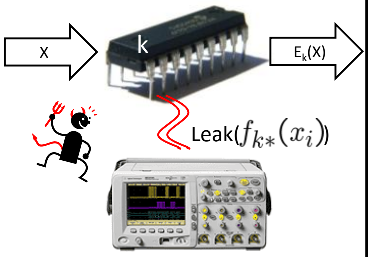


Differential power analysis attacks (9)

- Highest peak per hypotheses on K

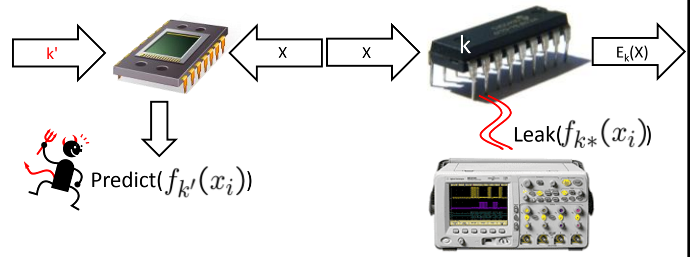


Modern view of differential attacks



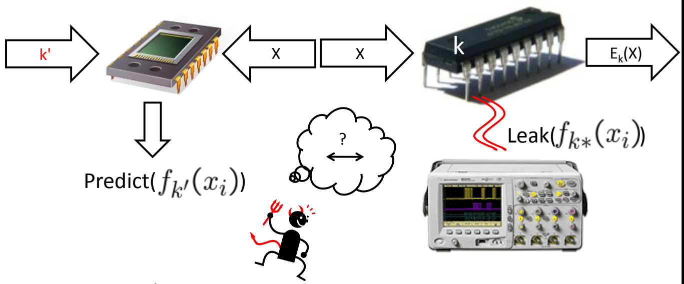
- Observe power consumption of targeted intermediate value $f_{k*}(x_i)$, multiple executions on varying input x_i

Modern view of differential attacks



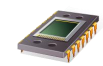
- Build a model to predict 'power consumption' Predict($f_{k'}(x_i)$) parameterized by guess k' on the secret k*

Modern view of differential attacks



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

Differential attacks: overview



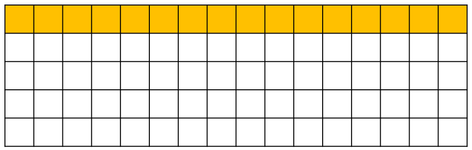
- Power models: better model → more powerful attack
  - More precise model requires to know or assume more details
  - Bad model → unsuccessful attack (≠ device is secure)
  - Often: Hamming weight or distance of operand value(s), single bits

- 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 [BCO04]
- Trade-off: efficiency (# traces) versus generality
  - Recently: generic attacks, e.g. using mutual information (MIA) [GBTP08]

After the fact

- Most power analysis attacks apply divide and conquer
  - Recover the secret in chunks, e.g. bits or bytes
  - For each chunk, hypotheses are ranked according to some score
  - What if the combination of the best hypotheses for each chunk does not yield the correct secret?

1
2
3
4
...



After the fact

- Most power analysis attacks apply divide and conquer
  - Recover the secret in chunks, e.g. bits or bytes
  - For each chunk, hypotheses are ranked according to some score
  - What if the combination of the best hypotheses for each chunk does not yield the correct secret?
- Enumeration
  - Guided exhaustive search
  - Problem: given a list of ranked hypotheses for chunks, generate list of ranked hypotheses for secret (in decreasing order of rank)
  - State of the art: 2^{32} hypotheses feasible

[VCGRS12]

Countermeasures

- Classified according to what they do
  - Hiding
  - Masking
  - Limits
- Classified according to how they can be implemented
  - Protocol
  - Non-crypto software
  - Algorithm implementation (how the algorithm is computed)
  - Digital logic
  - Analogue circuit

Countermeasures (2)

- Hiding
  - Increase noise (amplitude domain, time domain)
  - Decrease signal (filters, indistinguishable operations)
- Masking [CJRR99, S+10]
  - Compute function on randomized representation of the data
- Limits
  - Limit number of operations with the same key
    - Low frequency use, offline: counters (e.g. passport)
    - High frequency use, online: re-keying (e.g. pay TV) [MSGR10]

Countermeasures (3)

	Hiding	Masking	Limits
Protocol		X (Public key)	X
Non-crypto SW	X		X
Algo. implement.	X	X	
Digital logic	X	X	
Analogue	X		

- Examples
  - RSA signature generation
    - Blinded key prevents attacks requiring >1 measurement with same key
    - Regular sequence of operations prevents SPA
  - Digital Logic with almost data independent power consumption: Ingrid
  - Masked hardware implementations: Svetla

Pre-processing

- Reduce noise, increase or re-construct signal
  - Averaging, filtering, ...
  - Amplification (low-noise) before sampling, reduce quantization error
  - Alignment: synchronize time samples in measurements
    - Remove misalignment due to unstable trigger signal
    - Remove effect of countermeasures (random delays, unstable clock, ...)
  - Compression: reduce amount of data to process
    - After all, we often process many GB to extract a few bits
  - Transformation: alternative representation, e.g. in frequency domain
    - FFT, wavelets, ...: mix information in all time samples
  - Combination: join information in different time samples to create new traces, e.g. to break masking; trace length n has n(n-1)/2 pairs!
  - Normalize: usually zero mean and std dev 1

Pre-processing

- Reduce noise, increase or re-construct signal
  - Statistical moments: process measurements to expose certain statistical property that should contain information, e.g. to break (well-)masked implementations that process all shares in parallel

Evaluation

- Which attack is better, A or B?
  - Define "better" (often number of measurements)
  - Old days: if A works with n measurements and B does not, A is better than B
  - Today: sampling process, repeat attacks many times on independent data sets and calculate average scores (success rate, guessing entropy)
    - Keep all other parameters constant
    - Fully empirical, can be infeasible
  - Distinguishing margins: measure a distinguisher's ability to distinguish correct from incorrect key hypotheses [WO11]
    - Intuition: greater margin -> better distinguisher
    - But: 2 distinguishers with identical success rate can have different margins
    - Still informative, but interpret with care [RGV14]

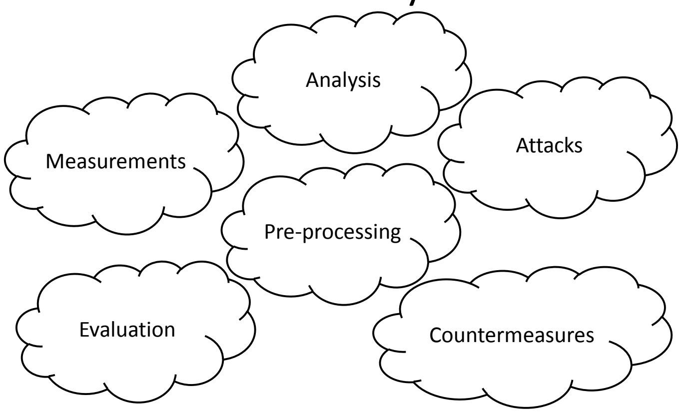
Evaluation

- Which countermeasure is better, A or B?
  - Define "better" (often number of measurements)
  - Old days: if A can be broken with n measurements and B cannot, B is better than A
  - Today: sampling process, repeat the attacks many times on independent data sets and calculate average scores (success rate, guessing entropy)
    - Keep all other parameters constant
    - Fully empirical, can be infeasible
  - Also: which attack is best? Try all?
  - Mutual information metric: how much information is leaked? [SMY09]
  - Leakage detection: does it leak? [GJJR11]

Power analysis – other uses

- IP protection
  - IP cores have distinct (unique?) power signature
  - Compare power signatures to detect IP fraud
  - Side-channel based watermarking [BKMP10]
- Hardware Trojan horse detection [ABKR07]
  - Record power signature of golden circuit
    - Verification that it is golden may require destructive reverse engineering
  - Compare power signatures to detect trojan

Summary



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
- [GBTP08] 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
- [KOP11] T. Kasper, D. Oswald, C. Paar: Side-Channel Analysis of Cryptographic RFIDs with Analog Demodulation. RFIDSec 2011: 61-77
- [OP11] D. Oswald, C. Paar: Breaking Mifare DESFire MF3ICD40: Power Analysis and Templates in the Real World. CHES 2011: 207-222

Bibliography

- [KOP09] T. Kasper, D. Oswald, C. Paar, EM Side-Channel Attacks on Commercial Contactless Smartcards Using Low-Cost Equipment, WISA 2009
- [VCGRS12] N. Veyrat-Charvillon, B. Gerard, M. Renauld, F.-X. Standaert, An optimal Key Enumeration Algorithm and its Application to Side-Channel Attacks, SAC 2012
- [WO11] C. Whitnall, E. Oswald, A Fair Evaluation Framework for Comparing Side-Channel Distinguishers, IACR eprint 2011:403
- [RGV14] O. Reparaz, B. Gierlichs, I. Verbauwhede, A note on the use of margins to compare distinguishers, COSADE 2014
- [SMY09] F.-X. Standaert, T.G. Malkin, M. Yung, A Unified Framework for the Analysis of Side-Channel Key Recovery Attacks, Eurocrypt 2009
- [GJJR11] G. Goodwill, B. Jun, J. Jae, P. Rohatgi. A testing methodology for side channel resistance validation. NIST NIAT workshop, 2011
- [BKMP10] G. Becker, M. Kasper, A. Moradi, C. Paar, Side-channel based Watermarks for Integrated Circuits, HOST 2010
- [ABKR07] D. Agrawal, S. Baktir, D. Karakoyunlu, P. Rohatgi, B. Sunar, Trojan Detection using IC Fingerprinting, IEEE S&P, 2007

Bibliography

- [RSV09] M. Renauld, F.-X. Standaert, N. Veyrat-Charvillon, Algebraic Attacks on the AES: Why Time also Matters in DPA, CHES 2009
- [MSGR10] M. Medwed, F.-X. Standaert, J. Großschädl, F. Regazzoni, Fresh ReKeying: Security against Side Channel and Fault Attacks for Low-Cost Devices, Africacrypt 2010