Improved Collision-Correlation Power Analysis on First Order Protected AES

Christophe Clavier[1] Benoit Feix[1,2] Georges Gagnerot[1,2]
Mylène Roussellet[2] Vincent Verneuil[2,3]

[1]XLIM-Université de Limoges, France

[2]INSIDE Secure Aix-en-Provence, France

[3]Institut de Mathématiques de Bordeaux, France

CHES 2011 - September 29, 2011

Outline

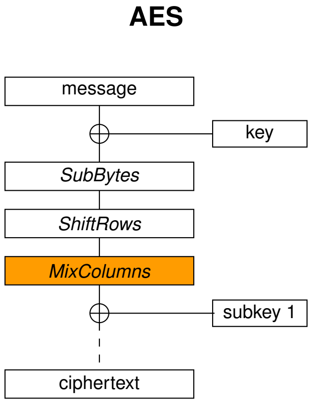
Outline

Collision attack

- K. Schramm, T. J. Wollinger and C. Parr. *A New Class of Collision Attacks and Its Application to DES.* FSE 2003.

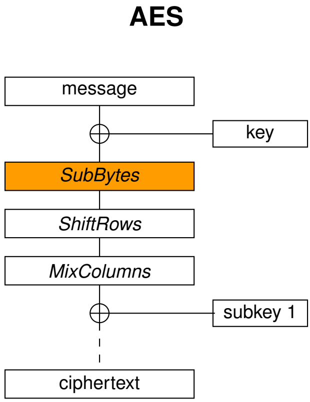
Collision attack

- K. Schramm, T. J. Wollinger and C. Parr. *A New Class of Collision Attacks and Its Application to DES.* FSE 2003.
- K. Schramm, G. Leander, P. Felke and C. Paar. *A Collision-Attack on AES: Combining Side Channel- and Differential-Attack.* CHES 2004.



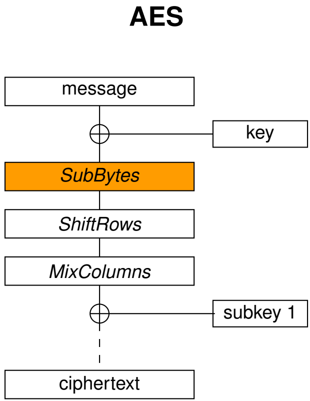
Collision attack

- K. Schramm, T. J. Wollinger and C. Parr. *A New Class of Collision Attacks and Its Application to DES.* FSE 2003.
- K. Schramm, G. Leander, P. Felke and C. Paar. *A Collision-Attack on AES: Combining Side Channel- and Differential-Attack.* CHES 2004.
- A. Bogdanov. *Improved Side-Channel Collision Attacks on AES.* SAC 2007



Collision attack

- K. Schramm, T. J. Wollinger and C. Parr. *A New Class of Collision Attacks and Its Application to DES.* FSE 2003.
- K. Schramm, G. Leander, P. Felke and C. Paar. *A Collision-Attack on AES: Combining Side Channel- and Differential-Attack.* CHES 2004.
- A. Bogdanov. *Improved Side-Channel Collision Attacks on AES.* SAC 2007
- A. Moradi, O. Mischke and T. Eisenbarth. *Correlation-Enhanced Power Analysis Collision Attack.* CHES 2010.
- . . .



Our Contribution

- Target first-order protected AES implementations
- Use correlation to detect internal collision
- Practical results on RISC 16-bit implementations
- Attacks validated using simulated and real curves
- Comparison with second-order techniques

Outline

Outline

AES Implementations

- We focus on AES-128 but our results can be applied to AES-192 and AES-256
  - message $M = (m_0\, m_1\, \dots\, m_{15})$
  - key $K = (k_0\, k_1\, \dots\, k_{15})$
  - ciphertext $C = (c_0\, c_1\, \dots\, c_{15})$
  - for $i \in [0, 15]$ we denote $x_i = m_i \oplus k_i$
- Attack on *SubBytes* function in first round
- Two protections against first-order attacks are considered:
  1. substitution table masking: $S'(x_i \oplus u) = S(x_i) \oplus v$
     same masks u and v for all bytes
  2. masked pseudo-inversion in $GF(2^8)$ using inversion in subfield $GF(2^4)$ (and $GF(2^2)$): $I'(x_i \oplus u_i) = I(x_i) \oplus u_i$
     16 different masks but same input and output masks

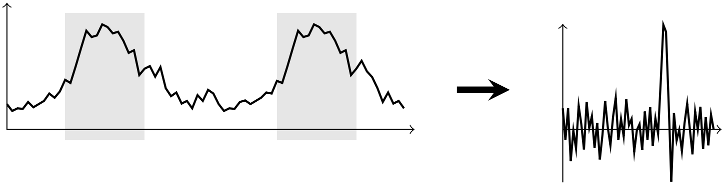
Outline

Principle

Attack Principle

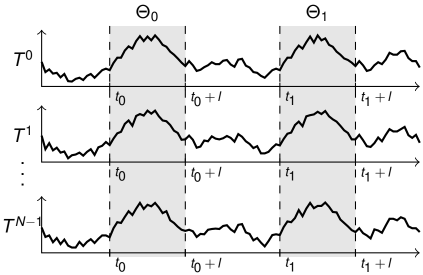
Detect internal collisions between data processed in blinded S-Boxes on the first AES round.

$data1 \oplus mask = data2 \oplus mask$



Collision-Correlation Analysis (1)

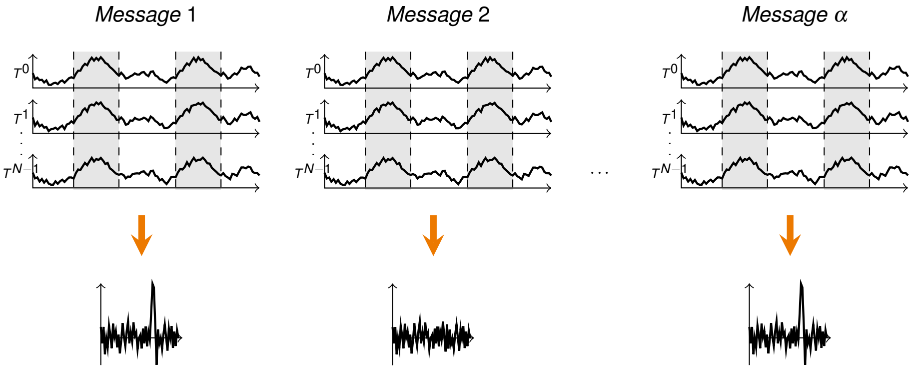
- Encrypt N times the same message M
- Collect the power traces T^n, $0 \leq n \leq N-1$
- Consider two instructions whose processing starts at times t_0 and t_1
  l points are acquired per instruction processing
- Construct the two series $\Theta_0 = (T^n_{t_0})_n$ and $\Theta_1 = (T^n_{t_1})_n$ of power consumptions segments



- Apply a statistical treatment to (Θ_0, Θ_1) to identify if same data was involved in $T^n_{t_0}$ and $T^n_{t_1}$
- We choose the Pearson correlation factor $\hat{\rho}_{\Theta_0,\Theta_1}(t) = \dfrac{\mathrm{Cov}(\Theta_0(t), \Theta_1(t))}{\sigma_{\Theta_0(t)}\sigma_{\Theta_1(t)}}$

Collision-Correlation Analysis (2)

Repeat with other messages until having enough information on key bytes



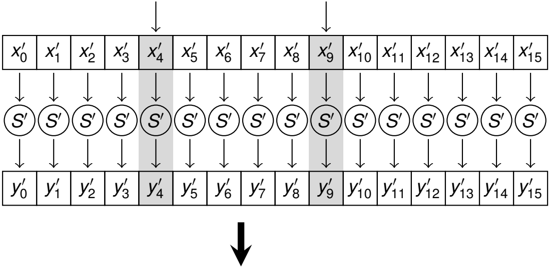
Outline

Outline

First Attack Description (1)

Principle = detect when two `SubBytes` inputs (and outputs) are equal in first AES round

$$m_4 \oplus k_4 \oplus u \quad = \quad m_9 \oplus k_9 \oplus u$$



$$k_4 \oplus k_9 \quad = \quad m_4 \oplus m_9$$

Result = provide a relation between two key bytes

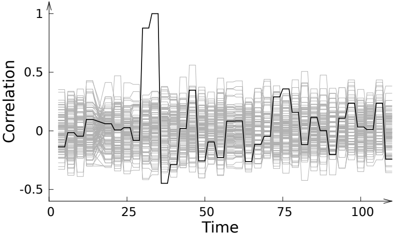
First Attack Description (2)

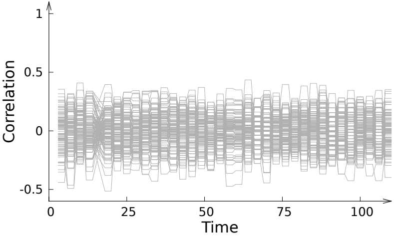
- Repeat for several random messages M until enough relations have been found
  - Encrypt N times the same message M and collect the N traces of first AES round
  - Construct the 16 series Θ_i corresponding to the computation of $S'(x_i \oplus u)$
  - For the 120 possible pairs (i_1, i_2) compute $\hat{\rho}_{\Theta_{i_1}, \Theta_{i_2}}(t)$
  - When a correlation peak appears a relation between k_{i_1} and k_{i_2} has been found

$\Rightarrow$ On average 59 messages are needed
Total number of curves $= 59 \times N$

Results on simulated curves

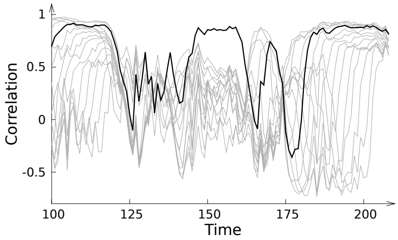
Correlation traces obtained on simulated curves for $N = 16$





Results on real curves

Correlation traces obtained on real curves for $N = 25$



Total number of acquisitions : $25 \times 59 \approx 1500$

First Attack Improvement

Remark: only collision events are exploited but they are not so frequent

Idea: exploit non-collision events as they are numerous

- For a given message only 0, 1 or 2 collisions most of the time among 120
- All other pairs (i_1, i_2) reveal impossible values for $k_{i_1} \oplus k_{i_2}$
  $\Rightarrow$ they are added to a blacklist
- Choose a message which have the maximum probability to generate a collision
- The penalty of a candidate message corresponds to the number of pairs (i_1, i_2) for which $m_{i_1} \oplus m_{i_2}$ is already blacklisted

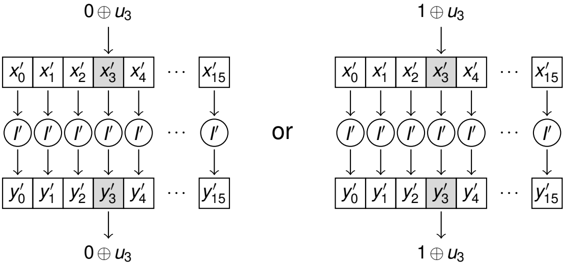
$\Rightarrow$ On average 27.5 messages are needed
Total number of curves $= 27.5 \times N$

On previous exemple we need $27.5 \times 25 \approx 700$ instead of 1500 curves.

Outline

Second Attack Description (1)

Previous attack cannot be applied to masked inversion as masks are different per bytes.



Collision between input and output reveals one key byte except one bit:

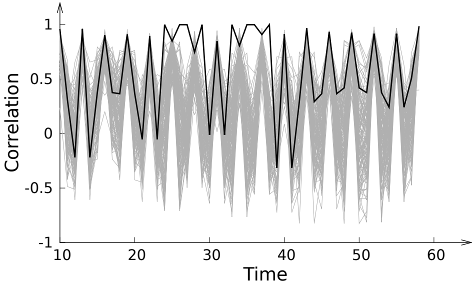
$$k_i = m_i \qquad \text{or} \qquad k_i = m_i \oplus 1$$

Second Attack Description (2)

- For each guess $g \in [0, 127]$
  - Encrypt N times message M s.t. $m_0 = g$ and collect traces $T^{n,g}$, $0 \leq n \leq N-1$
  - Construct series:
    Θ_0^g corresponding to the load of $x_0 \oplus u_0$ before inversion
    Θ_1^g corresponding to the store of $I(x_0) \oplus u_0$ after inversion
  - Compute $\hat{\rho}_{\Theta_0^g, \Theta_1^g}(t)$
- The highest correlation peak reveals k_0 except 1 bit

Practical Results

Correlation traces obtained on simulated curves for the pseudo-inversion of the first byte in $GF(2^8)$ with $N = 16$



Outline

Definitions

Target first implementation i.e. S-Box masking.

Consider three functions commonly used for second-order attacks:

- $f_1(x,y) = |x - y|$
- $f_2(x,y) = |x - y|^2$
- $f_3(x,y) = |x \times y|$

Use as distinguisher the Pearson correlation factor $\hat{\rho}$

Second Order Attack Modeling

- Construct the series of power consumptions of two S-Box outputs for N messages

$$\theta_0 = (HW_n(S(x_{i_1} \oplus u) \oplus v) + \omega_\sigma)_{0 \le n \le N-1}$$

$$\theta_1 = (HW_n(S(x_{i_2} \oplus u) \oplus v) + \omega_\sigma)_{0 \le n \le N-1}$$

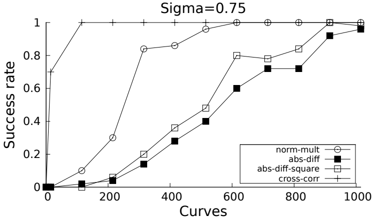
- Compute the series of estimated values of S-Box outputs for key guesses g_{i_1} and g_{i_2}

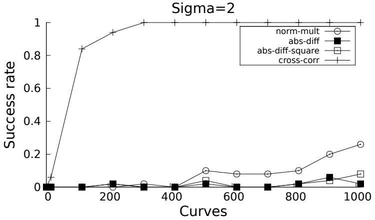
$$W_{g_{i_1}, g_{i_2}} = (HW_n(S(m_{i_1} \oplus g_{i_1}) \oplus S(m_{i_2} \oplus g_{i_2})))_{0 \le n \le N-1}$$

- The right key byte is obtained for the highest correlation value $\hat{\rho}(f_i(\theta_0, \theta_1), W_{g_{i_1}, g_{i_2}})$

Comparison

Compare the success rate of second-order power analysis methods with the collision-correlation one by simulating these attacks for different standard deviation σ of noise ω.





Outline

Conclusion

- Improved collision-correlation technique defeats some first-order protected implementations
- Need less than 1500 acquisitions
- More powerful than previous second-order power analyses
- No need to establish a consumption model for correlation

Thanks for your attention.