NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

pace Physical Analysis and Cryptographic Engineering

JUST ONE FAULT

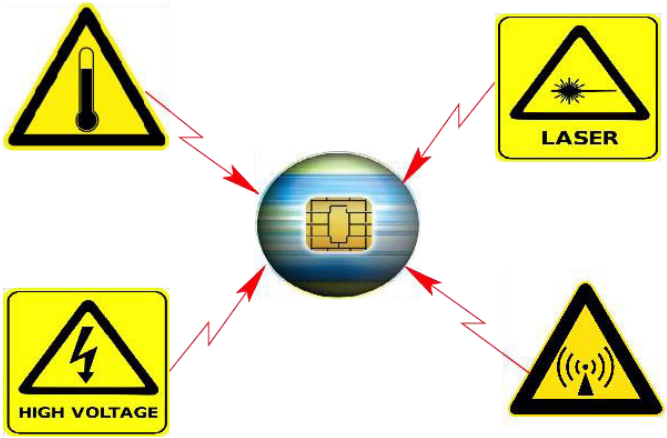
Persistent Fault Analysis on Block Ciphers

Shivam Bhasin
Temasek Labs @ NTU

ASK 2018, Kolkata, India
15 Nov 2018

Table of Contents

Fault Injection Attacks (FIA)



What is FIA?

- Physical Attacks
- Actively disturbs functioning of the target
- Exploits erroneous behavior

Injection Methods

- Global/Low-Cost/Low-Precision
  - Clock/Voltage glitch, temperature
- Local/High-Cost/High-Precision
  - Laser, Electromagnetic, Ion Beam

Impacts

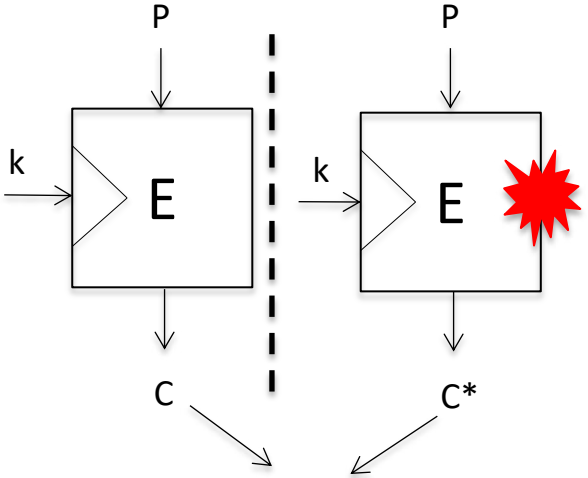
- Duration
  - Transient or Harmonic
- Effects
  - Data or Flow Modification
- Objectives
  - Corrupt computation, bypass security checks

Fault Models

- Single/multiple bit-flip – a target variable was altered either by single or multiple bit flip.
- Random byte fault – Some bits of a byte are flipped. No-precise multi-bit flip.
- Instruction skip – One or several instructions were not executed (for software)
- Stuck-at fault – target variable stuck at-0/1

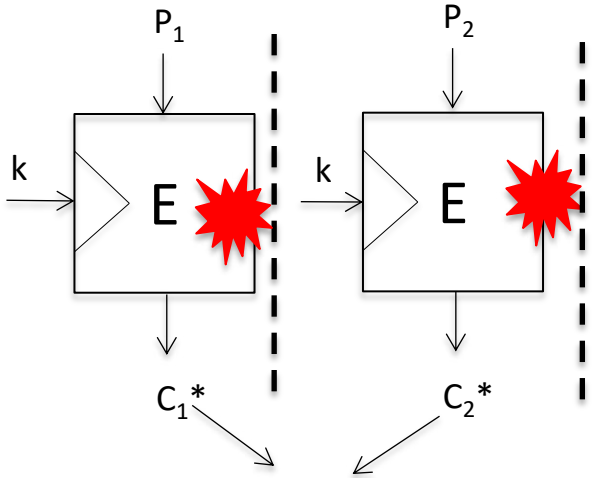
Fault Analysis

- Differential Fault Analysis (DFA)
- Usually few ciphertext pair
- Control over plaintext needed



Analysis K=f(C,C*)

- Statistical Fault Analysis (SFA)
- Need several ciphertext
- Several variants exist



Analysis K=f(C_1*,C_2*, ...)

Fault Countermeasures

- Two principle approches
- Detection
  - Incremental
  - Sensors to detect physical condition
  - Redundancy to detect data modification
- Prevention
  - Provable
  - Infect/Correct fault

Limitations of SoA

- Very tight time synchronization on the round calculation and the injection timing
- Very complicated analysis due to the random value and the fault propagation
- May not work if there are countermeasures against fault attacks

Revisiting Fault types

- Transient: Affect one encryption
- Permanent: Always present

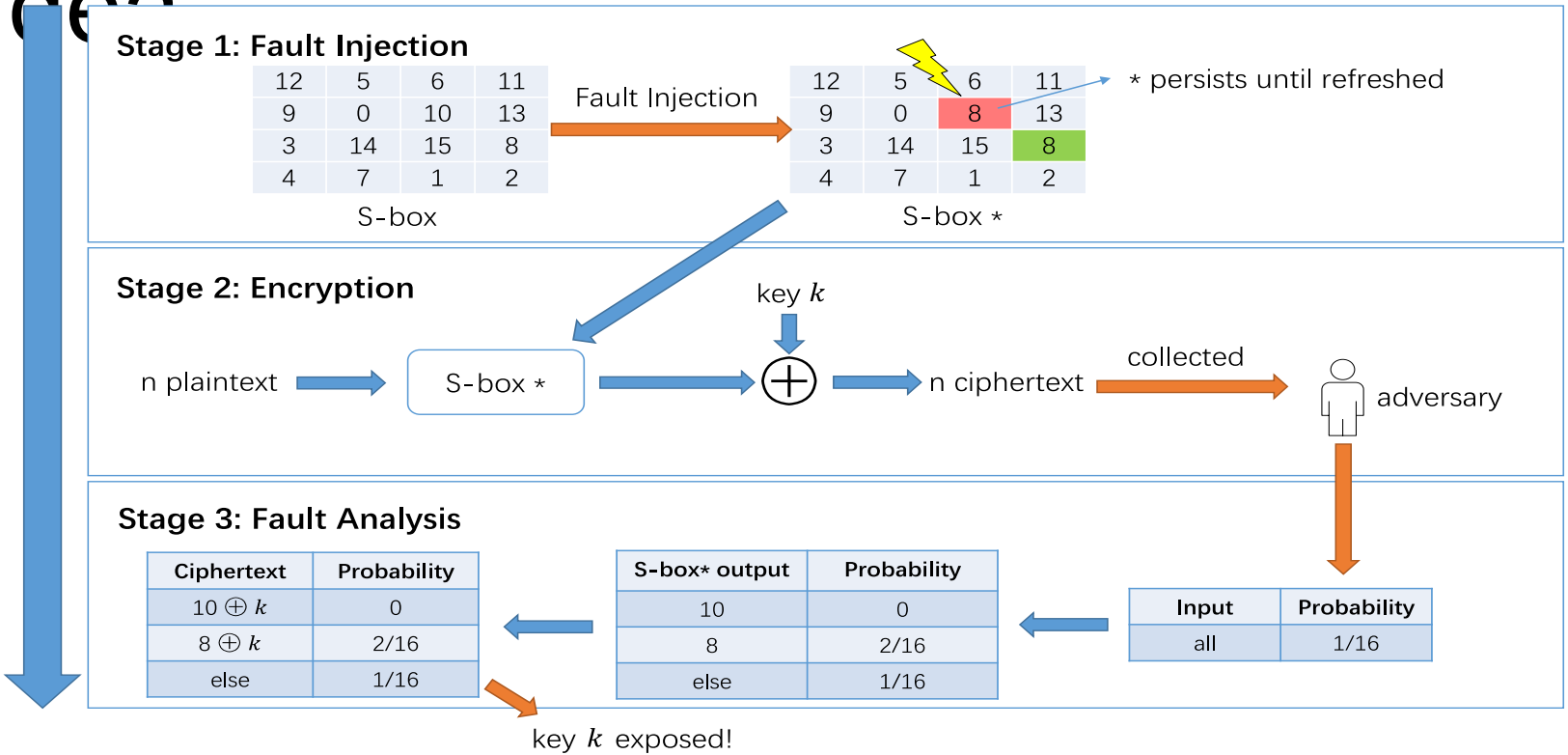
- **Persistent[1]:** Hybrid model between transient and permanent. Persist over several encryptions but disappears on reboot. Typically targets stored constants (ex. Sbox in memory)

[1]Persistent Fault Analysis. CHES 2018
Joint work with Zhang, Fan, Xiaoxuan Lou, Xinjie Zhao, Wei He, Ruyi Ding, Samiya Qureshi, and Kui Ren

Adversary Model

- Block cipher with serial implementation
- Common Sbox as look-up table
- Persistent fault injected in one Sbox element
- Victim encrypts n plaintext with faulty Sbox
- Adversary can observe the n ciphertext
- No control on plaintext, except varying plaintext

Persistent Fault Analysis: Main Idea

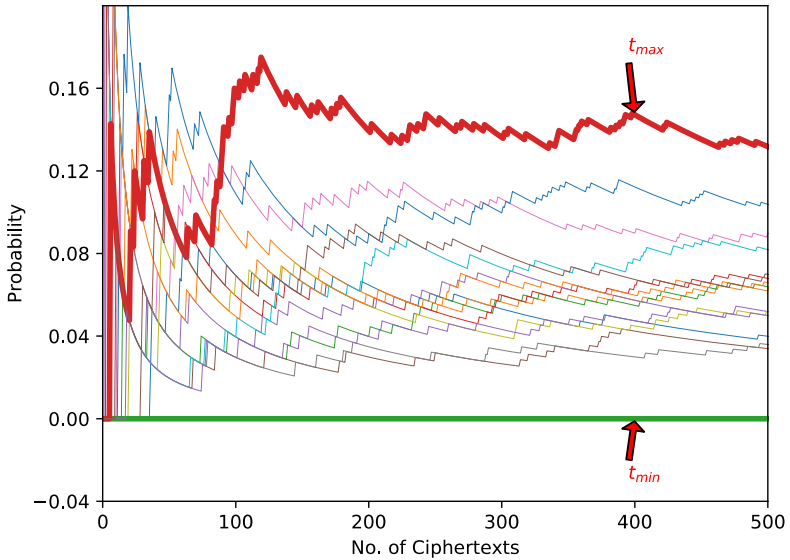


PFA: Modus Operandi

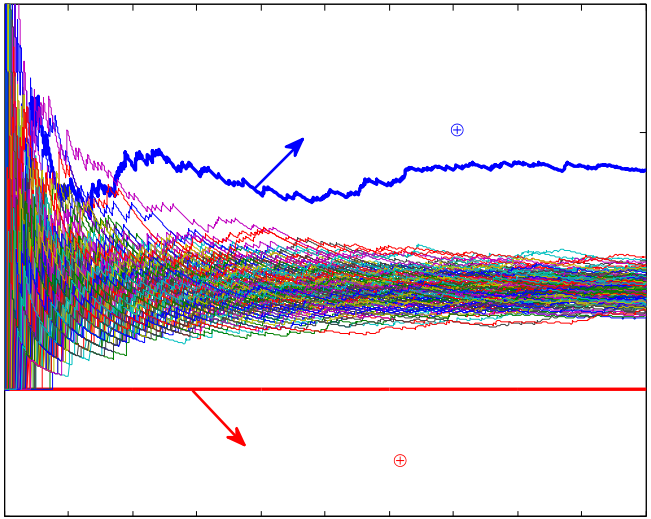
- Statistical analysis on last round with ciphertext only
- Fault changes one element $x \rightarrow x^*$ in Sbox (lets say 4X4 Sbox)
- Expectation $E(x)= 0$, $E(x^*)=2/16$, $E(y \neq (x,x^*))=1/16$
- Three analysis startegies:
  - t_{min}: find the missing value in Sbox table (x). Then $k = t_{min} \oplus x$;
  - $t \neq t_{min}$: find values t where $t \neq t_{min}$ and eliminate candidates for k;
  - t_{max}: find the value with max probability (x'). Then $k = t_{max} \oplus x^*$
- No. of ciphertext n can be determined by coupon collector's problem
- x, x* can be brute-forced if not known

PFA on PRESENT and AES

PRESENT: n ≥50

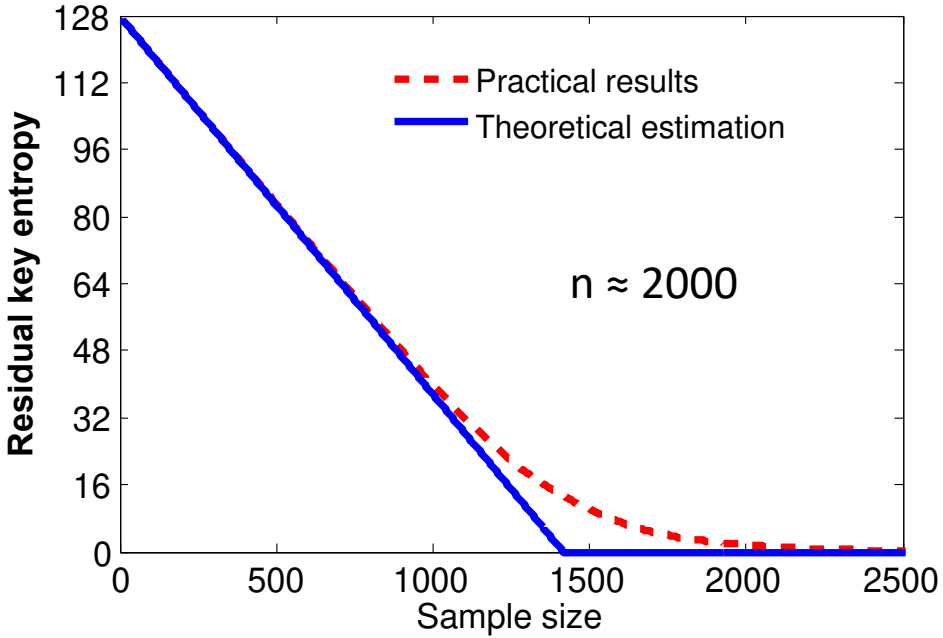


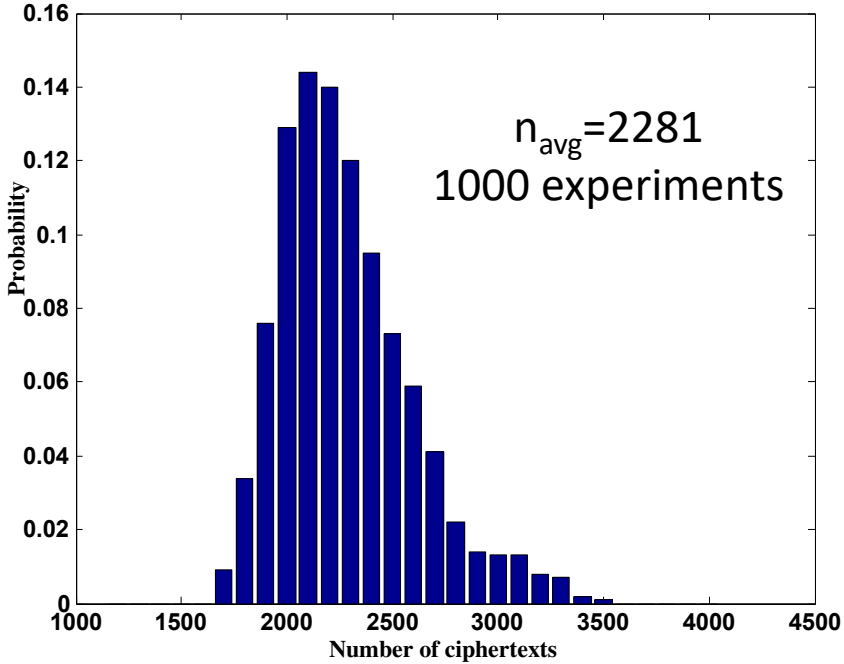
AES: n ≥ 1560



n= Minimum no of ciphertext needed by coupon collector's problem

Practical PFA on AES





Comparison vs Other Fault Attacks

(1) The attack is not differential in nature and thus the control over the plaintext is not required.

(2) The adversary does not necessarily need live synchronization

(3) The fault model remains relaxed (no biased faults needed)

(4) PFA can also be applied in multiple fault setting

(5) PFA can bypass some redundancy based countermeasures

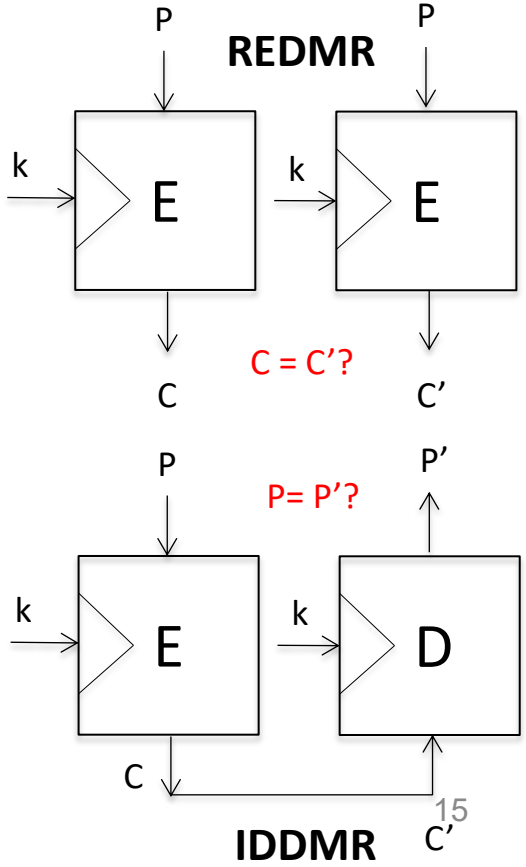
(6) An adversary can always inject the persistent fault before the victim is switched to the sensitive mode

(1) It needs higher number of ciphertexts as compared to DFA

(2) Persistent faults can be detected by some built-in health test mechanism or fault counters.

Dual Modular Redundancy (DMR) Countermeasure

- Compute twice and compare (REDMR)
- Compute forward-inverse and compare (IDDMR)
- If ≠
  - NCO: No Ciphertext output
  - ZVO: Zero Value output
  - RCO: Random Ciphertext output
- Provably secure against single fault
- Adversary can either target the encryption or comparison but not both
- REDMR broken by design if same S-box is used
- Lets target IDDMR, more difficult of the two

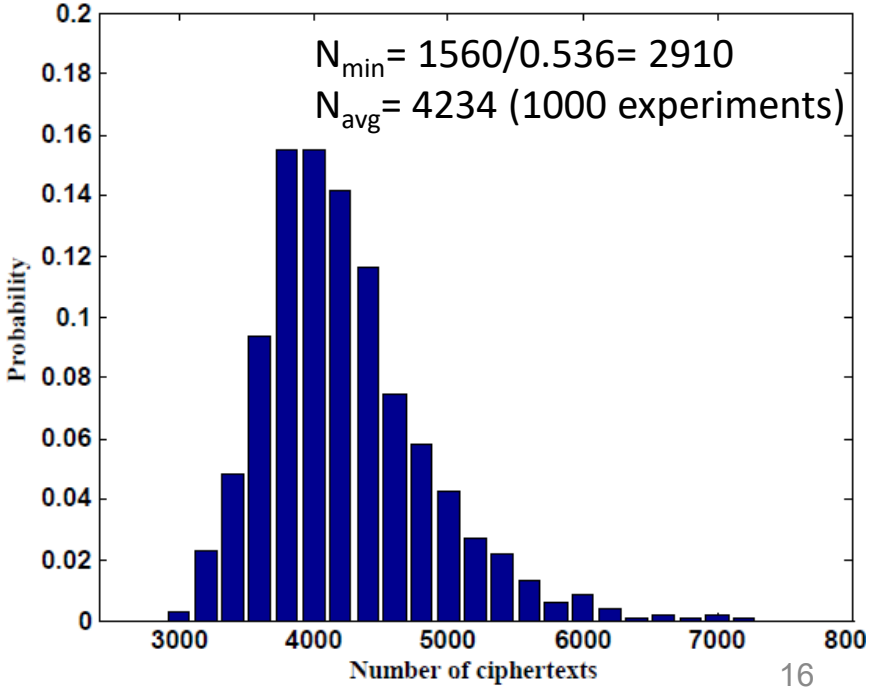


Attacking IDDMR with NCO/ZVO

- Faulty outputs are supressed
- Some output will be not affected by fault
- Probability p of correct output is f(x,k)
- *p* for AES

$$p = (1 - \frac{1}{256})^{160} = 0.5346$$

- Adversary roughly needs n/p ciphertext



Attacking IDDMR with RCO

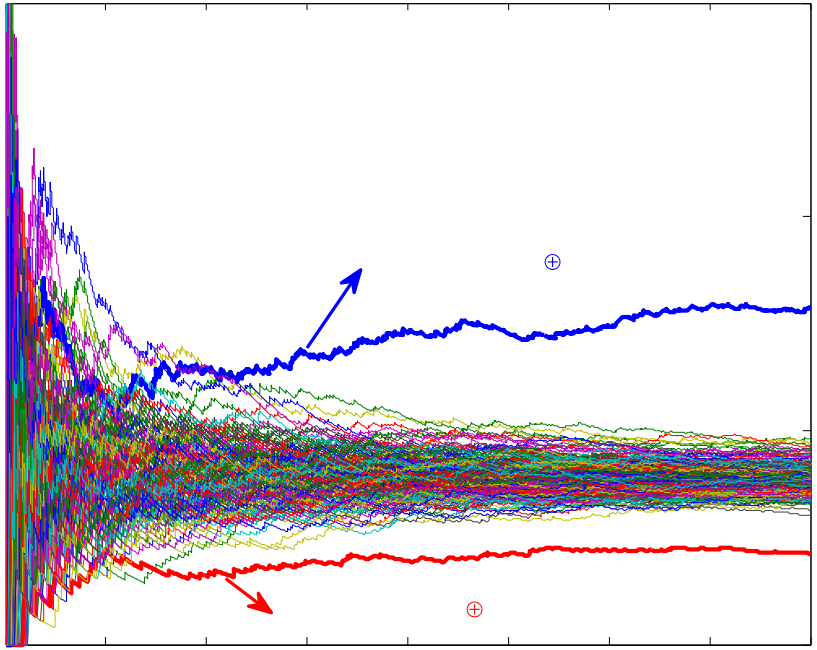
- Faulty output is replaced by uniformly random
- Slight difference in distribution of random output and correct ciphertext
- The bias can be detected with more ciphertext (n)

$$Pr(y = x) = 0 \times p + \frac{1}{256} \times (1 - p) = \frac{0.4654}{256}$$

$$Pr(y = x^*) = \frac{2}{256} \times p + \frac{1}{256} \times (1 - p) = \frac{1.5346}{256}$$

$$Pr(y \neq x \wedge y \neq x^*) = \frac{1}{256} \times p + \frac{1}{256} \times (1 - p) = \frac{1}{256}$$

- Roughly n≈10000 resulted in attack success



Conclusions

- Proposed Persistent Fault Analysis (PFA)
  - A novel attack on general block ciphers
  - Defeat popular fault countermeasures
  - Can work with multiple faults
  - One one fault injection required
- Validated with practical experiments
  - Used Rowhammer on Intel CPU to attack AES-128 in cryptographic library *Libgcrypt*

Fuure Works

- Attack higher-order masking
  - Accepted at DATE 2018
  - Target public implementation of higher-order masking with one fault
- Analyze combined countermeasure
- Develop countermeasures
- How to reverse key-scheduling?
- Application on PKC/PQC