ECE 547/647: Security Engineering

Side-channel Attacks

Assignment 3 Discuss

- Voting
- Power Grid

- Similarities
- Differences
- Threats
- Defenses
- Future Research

Announcements

- Start thinking about project ideas!
  - Read papers
  - Build and measure something
  - Presentation and Report
- Midterm Exam, Take-home (24 hours)
  - Open book and notes but no collaboration or help from other humans
  - Focus on slides (rather than more extensive content in book or readings)
  - Practice exam will be posted 1 week before exam

Side channels

- Side-channels can **leak** information from cryptographic circuits.
- This lesson will focus on how the power side-channel can allow cryptographic keys to be extracted through a combination of measurement and statistical analysis.

Objectives

- Analyze how to extract secret keys through information leaked through power consumption and other side channels
- Apply the Differential Power Analysis approach of retrieving secret keys from different cryptographic algorithm

Discussion

- What are examples of side-channels?
- What is SPA vs. DPA?

Power Consumption by an IC

- The power consumed by a circuit varies according to the activity of its individual transistors and other components.
- Measurements of the power can contain information about the operations being performed and the data being processed.
- DPA is one approach by which the information leaked through power consumption can be analyzed to extract secret keys.

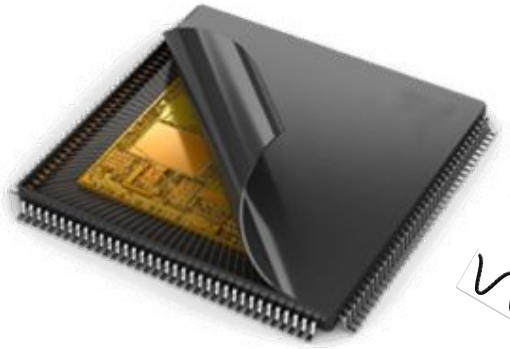
Introduction to DPA

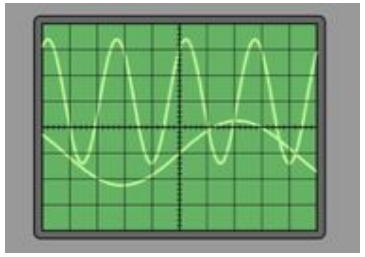
- Transistors activity depends on the data the circuit is processing.
  - For example: more transistors may switch when adding the hex value of A7 to B9 than when adding 01 to 00.
- Because the amount of power used by a device is influenced by the data being processed, power consumption measurements contain information about a circuit's calculations.
- When a device is processing cryptographic secrets, its data-dependent power usage can expose these secrets to attack.

Side-Channel Analysis

Side Channel Leakage

Timing

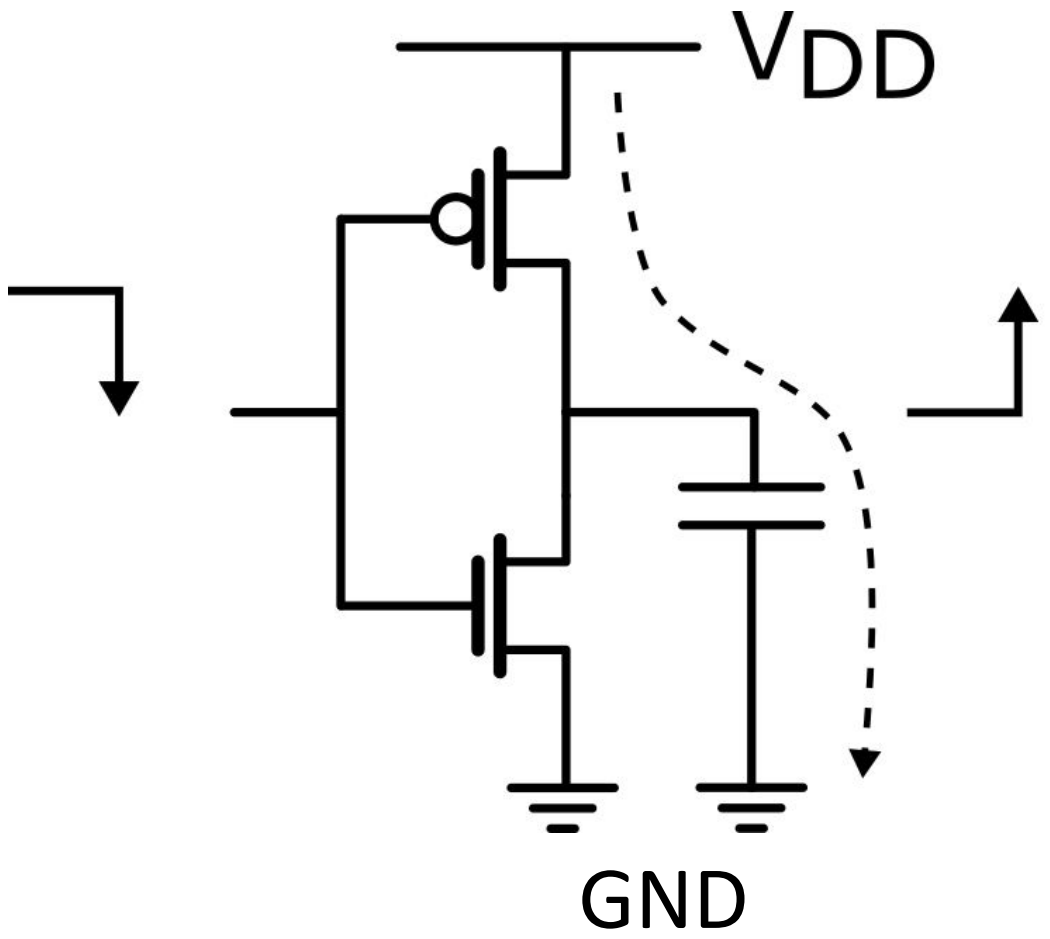




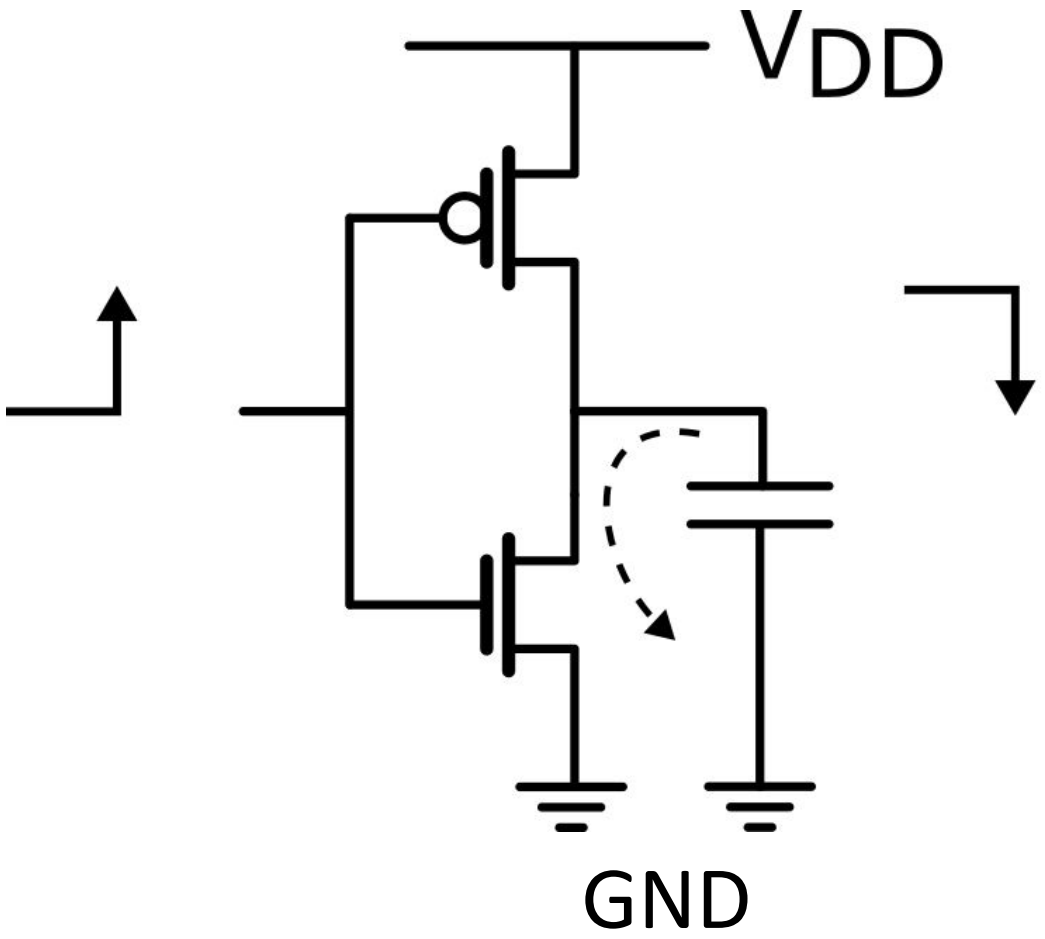
Current, Voltage

EM Emissions

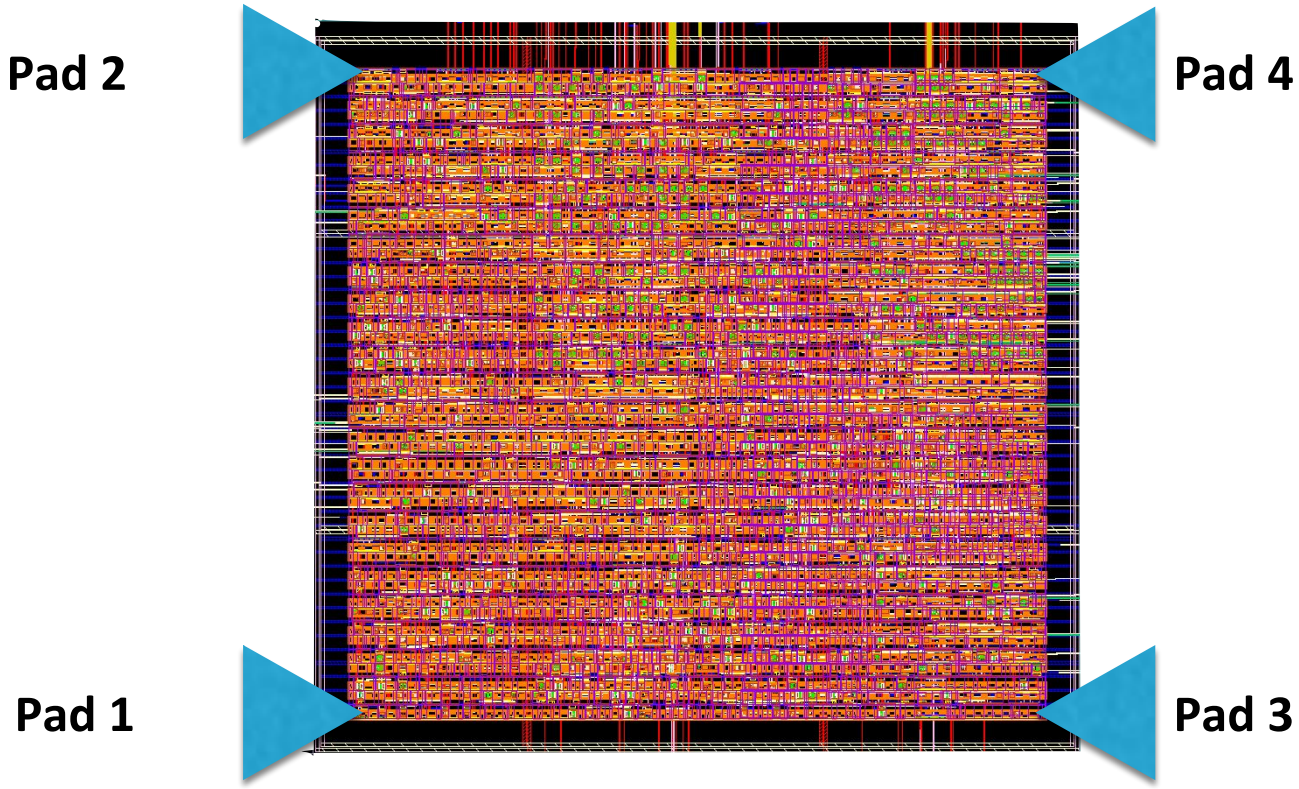
Side-Channel Analysis



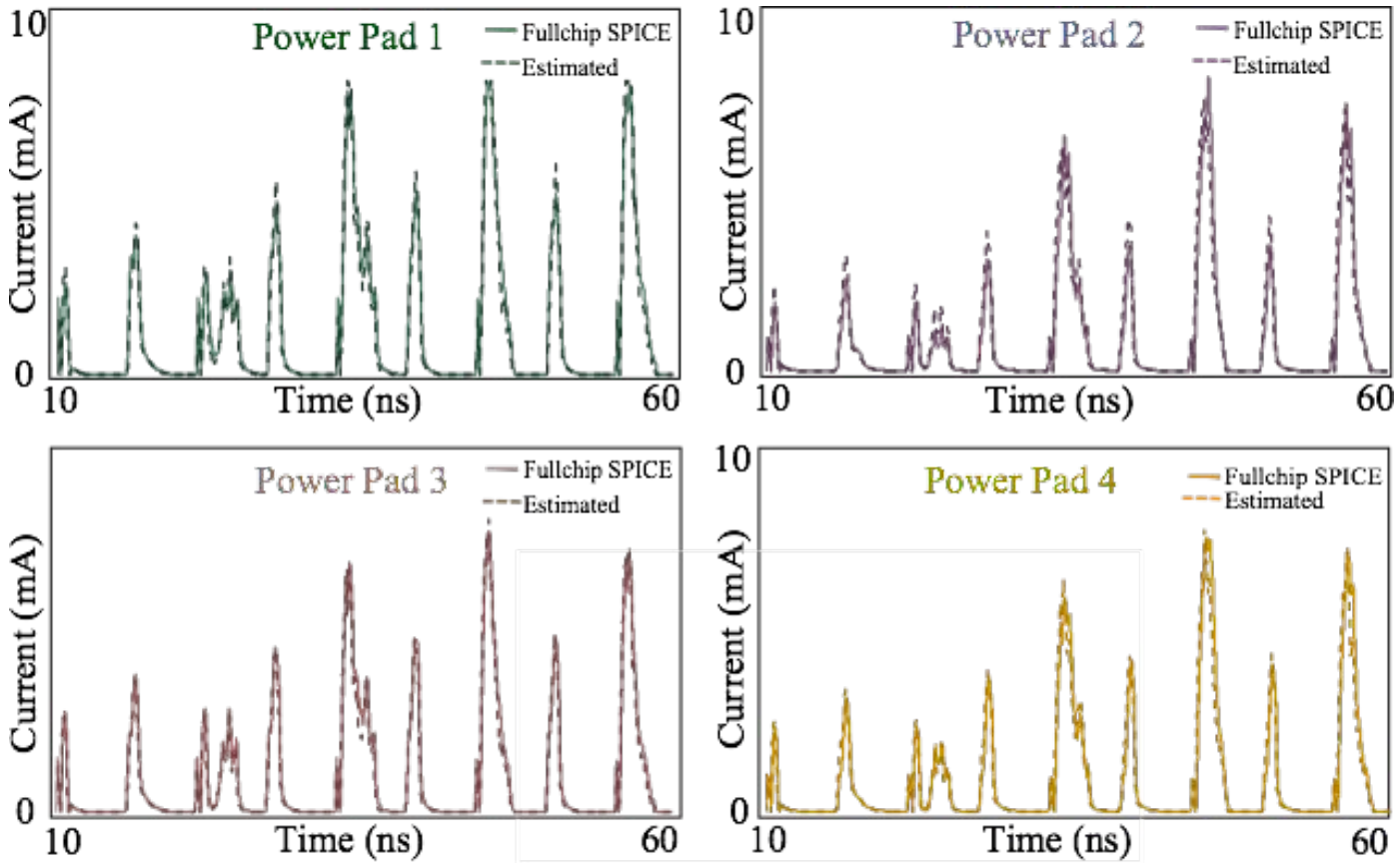
Side-Channel Analysis



Layout of Circuit under Test (DES)



Simple Power Analysis



Traces and Frequency Distribution

- A trace is a sequence of measurements taken across a cryptographic operation or sequence of operations.
- The trace data are captured by placing a resistor in series with the device's ground line, then using an oscilloscope to measure the voltage at the ground input.

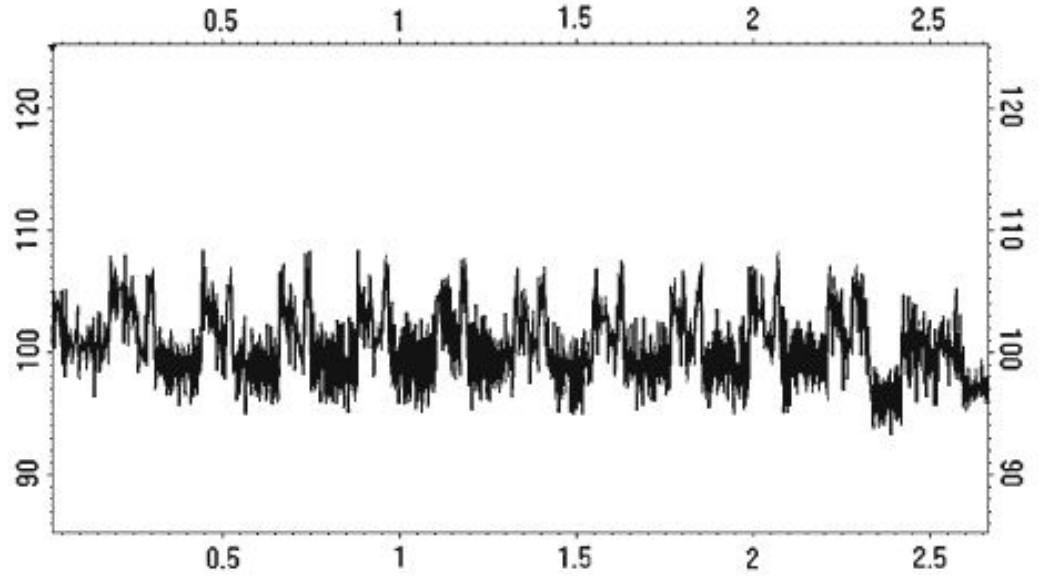


Fig. 1 Power trace from a smart card performing an AES-128 encryption, with the ten rounds clearly visible

Group Discussion and Report Back (Hand Raise): Power Traces

- How practical is it to capture the power trace?
  - For a smart card, for other implementations?

- How practical is it to capture thousands of power traces?

Traces and Frequency Distribution

- The Figure below shows the region of the power trace from Fig. 1 during the first round of the AES-128 operation.
- The moment in time when the card computes the output of the first S-box is marked by a vertical line.

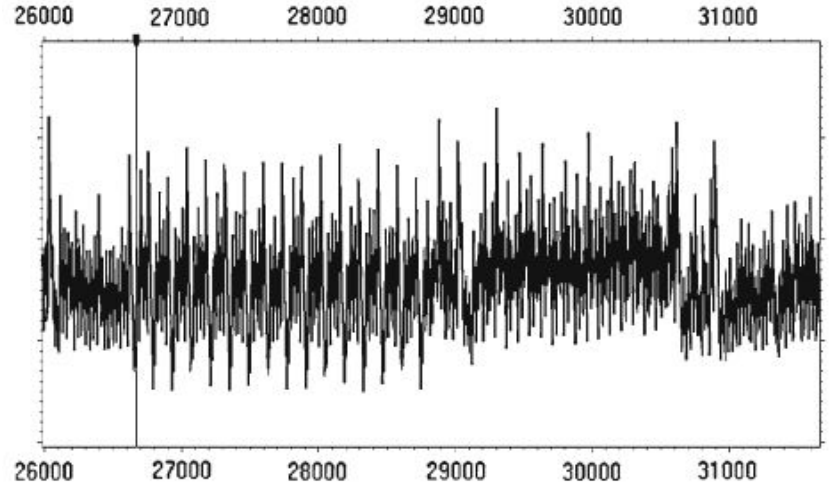


Fig. 3 Power trace segment showing the first round of AES-128 encryption on a smart card. A vertical line marks the location of first S-box lookup

Traces and Frequency Distribution

- The figure shows two distributions of power measurements.
- These two distributions are significantly different, demonstrating that the power consumption is statistically correlated to the LSB of the S-box output.
- Given a large set of measurements, these distributions can be reliably distinguished.

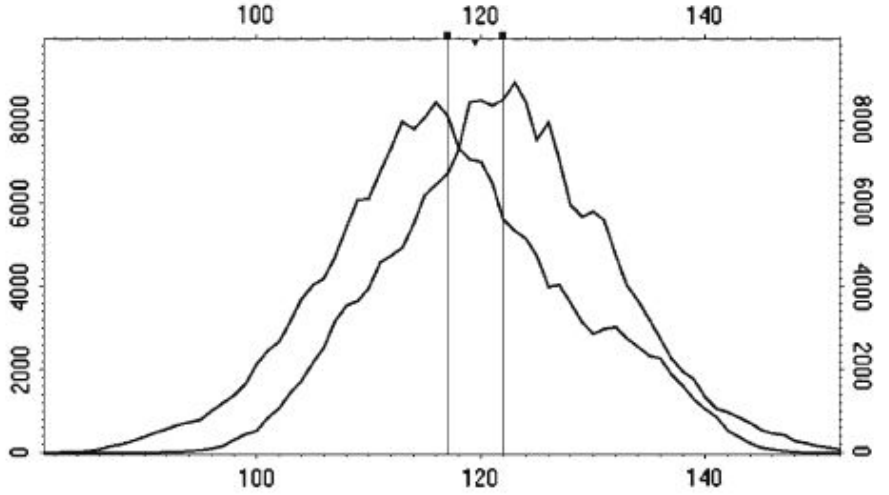


Fig. 5 Distributions of power consumption measurements for traces with the LSB of the output of the first S-box being 1 (left) and 0 (right)

Orchestrated Discussion (Hand Raise): Statistics Background

- Who in this class has taken probability and statistics?
- What is a Gaussian distribution?
- Why is this distribution Gaussian?

Differential Power Analysis (DPA)

- A statistical method for analyzing sets of measurements to identify data-dependent correlations.
- The basic method involves partitioning a set of traces into subsets, then computing the difference of the averages of these subsets.
  - If the choice of which trace is assigned to each subset is uncorrelated to the measurements contained in the traces, the difference in the subset's averages will approach zero as the number of traces increases.
  - Otherwise, if the partitioning into subsets is correlated to the trace measurements, the averages will approach a nonzero value.

Components of a Typical Successful DPA Result

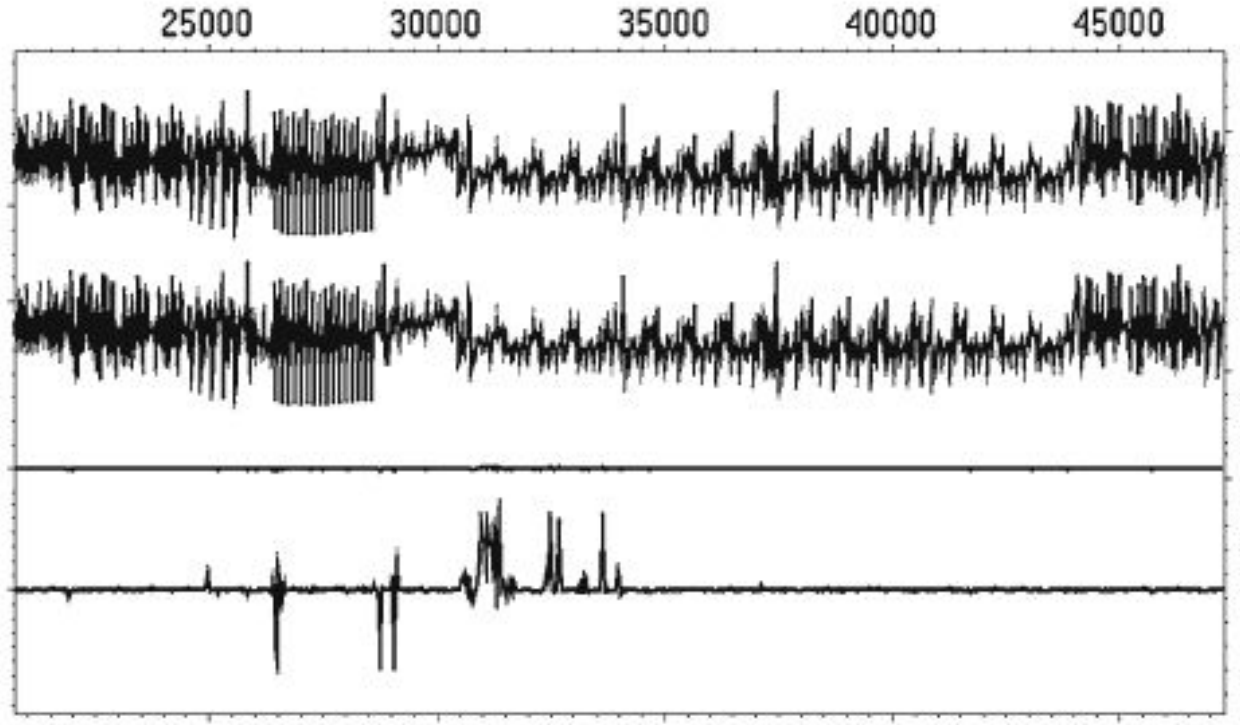


Fig. 6 Typical DPA result showing (from top to bottom) the average of the traces where the LSB of the output of first S-box in round 1 is 1, the average of traces where the LSB is 0, the difference between the top two traces, and the difference with the *Y* axis magnified by a factor of 15

Selection Function for DPA

- The information revealed by a DPA test is determined by the choice of selection function.
- A selection function is used to assign traces to subsets and is typically based on an educated guess as to a possible value for one or more intermediates within a cryptographic calculation.
  - If the final DPA trace shows significant spikes, the cryptanalyst knows the selection function output is correlated to (or equals) a value actually computed by the target device.
  - If no correlation is observed, then selection function output was not correlated.

DPA Attack on AES

The first round of AES-128 encryption consists of the following steps:

1. Initialization: The initial 16-byte state of the cipher, organized as a 4x4 byte matrix, is initialized to the 16 bytes of the plaintext.
2. AddRoundKey: The 16-byte secret key is exclusive-ORed with the 16 bytes of the plaintext state.
3. SubBytes: Each byte of the state is replaced by another using the S-box, which is an invertible lookup table.
4. ShiftRows: Bytes in each row of the state are shuffled.
5. MixColumns: Each column of bytes of the state is mixed using a linear operation.

DPA Attack on AES

- The DPA attack will target the output of AddRoundKey and SubBytes in AES.
- For each trace i, Let I_i denote the 16-byte intermediate state of the cipher just after the SubBytes step in round 1.
- Let the nth byte of this state (where $n \in \{0, \ldots, 15\}$) be denoted by $I_{i,n}$.
- $X_{i,n}$denote the nth byte of plaintext X_i used for the i_{th} trace.

$$I_{i,n} = S[X_{i,n} \oplus K_n]$$

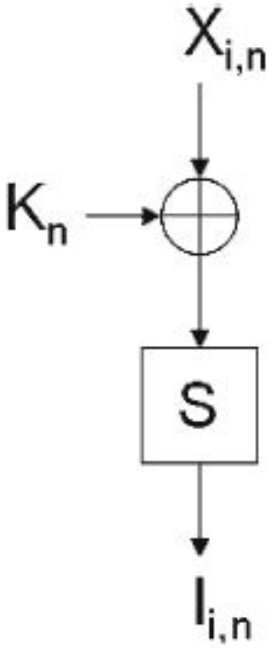


Fig. 7 AES S-box lookup during first round

DPA Attack on AES

- DPA provides a practical way to test if a candidate value of K_n is correct.
- The candidate K_n is used with below equation to derive the value of $I_{i,n}$ for each trace's $X_{i,n}$.

$$I_{i,n} = S[X_{i,n} \oplus K_n]$$

- In this example, bit 0 (the LSB) of $I_{i,n}$ was used as the selection function output.

- Each trace is assigned to one of two subsets, depending on whether the selection function result is 0 or 1 for the candidate K_n and the plaintext being encrypted when the trace was captured.

DPA Attack on AES

- The difference of the subset's averages is then examined.

- If the value of the S-box output bit predicted by the selection function has even a tiny correlation to the power traces, the DPA test will show spikes indicating that the candidate K_n is correct.

- For each wrong K_n, the predicted values of $I_{i,n}$ will be (largely) unrelated to any data being processed by the target device, and the DPA test will not be (or will be much less) statistically significant.

DPA Result

- Figure shows, from top to bottom, five traces for $K = 101, \ldots, 105$.
- The correct value for K is 103, as is obvious from the presence of large spikes in the $K = 103$.
- Traces for incorrect K values have much smaller spikes or are relatively flat.

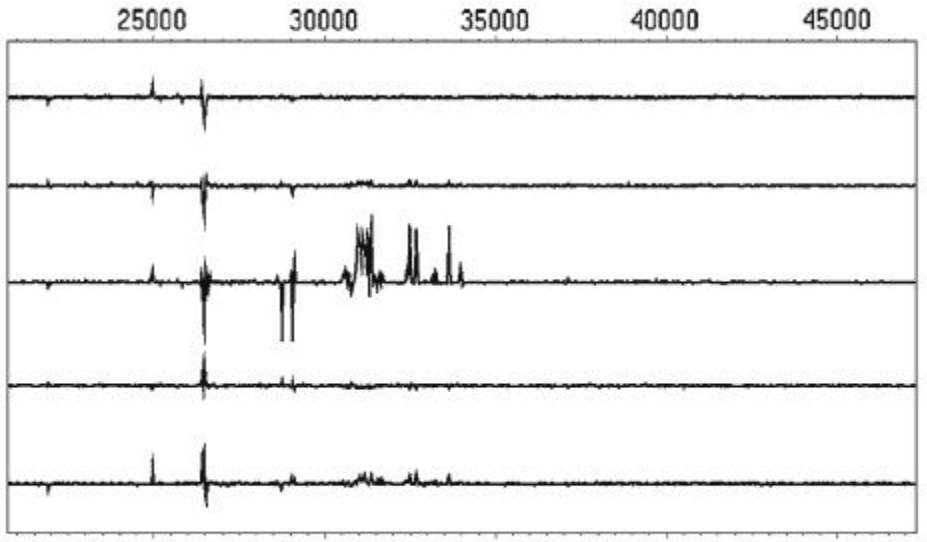


Fig. 8 Five differential traces for the DPA test predicting the LSB of $I_{i,0}$ for guesses $K_0 = 101, \ldots, 105$ from top to bottom, with the correct key $K_0 = 103$, corresponding to the third trace

DPA Result

- The same analysis can be repeated for all the 16 bytes of the state $(n = 0, \dots, 15)$ to recover the entire 128-bit AES secret key from the device.

- The same traces can be reused in finding each key byte; it is not necessary to collect separate data, since each test is checking for different correlations in the data set.

Group Discussion and Report Back (Short Answer): DPA Result

- Calculate file size for 20,000 traces at 100 MHz and 16-bit data.
  - Assume AES takes 16 clock cycles.

dpacontest.org

Post-work

- Work on finding a group and a presentation/project
- Read KeeLoq paper and 180nm AES chip paper.