# Report on implementation of Difference of Mean Power Attack

Group No: 8

Pranjal Kumar Srivastava Roll No: 22111046

Gunj Hundiwala Roll No: 22111024

Kapilkumar Kathiriya Roll No: 22111028

Amit Kumar Roll No: 22111008

Instructor : Prof. Urbi Chatterjee

Indian Institute of Technology ,Kanpur

## Abstract

This report demonstrates the fundamental technique of power analysis, differential power analysis (DPA) . The DPA attack we implement is referred to as the Difference of Means attack. The cryptographic algorithm we have chosen to attack is AES-128. In particular,

We have written AES code in python and we have  power traces stored in CSV file. We have simulated this DOM power attack in python and we demonstrate how the full 16-byte cipher key can be deduced using the technique by monitoring the power consumption of the device during cryptographic operations. We were assigned the task of finding 4$^{th}$ and 5$^{th}$ byte of the secret key which we have successfully recovered.

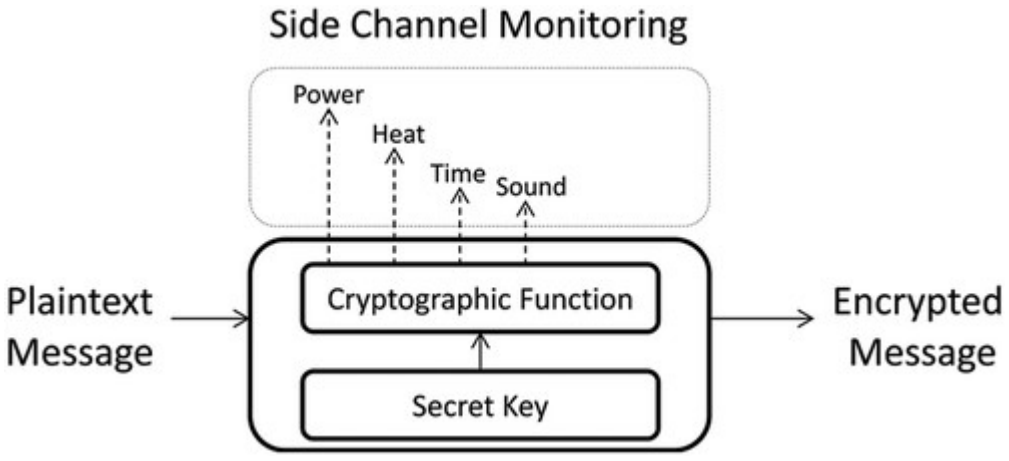
Our Performance in Assignment:

# 1. Background

## 1.1. Overview of Side Channel Attacks

A SCA is an attack on the information of a cryptographic device. The concept of 'information' in the context SCA generally refers to a secret key of a cryptographic algorithm. A SCA is carried out by monitoring the physical outputs of a device (e.g. power consumption, time taken to carry out an operation, emission of heat, light and sound). The hypothesis made in such an attack is that the physical outputs of a cryptographic device demonstrate correlation with the internal state of the device when conducting cryptographic operations. Figure 1 depicts an abstraction of this concept.

Figure 1. Side channel monitoring.



Rather than attempt to break the core implementation of a cryptographic device, SCA aims to monitor ‘leaked’ information produced by a device during its normal operation and attempt to deduce the secret key leakage gathered. Some of the most common attacks include timing attacks , fault attacks  and power analysis attacks. For the scope of our assignment we focus on power analysis attack, which is focused on monitoring the power consumption of a device in order to deduce information leakage.

## 1.2 Power analysis techniques

Power analysis attacks are carried out by monitoring the power consumption on a cryptographic device. A common piece hardware which can be used to monitor the device is an oscilloscope. In this type of attack, one must first assume that there is correlation between the level of power consumption and cryptographic operations of the device. Originally, there were two main categories of power analysis attacks including simple power analysis (SPA) and DPA. As research progressed in the area of SCA, a third category was found in literature which is formally referred to CPA.

An example of a more power attack is DPA. This attack makes use of statistical techniques to identify *differences* in power traces, thus revealing data leakage which may result in the correct secret key being guessed. CPA, on the other hand, aims to reveal data leakage by finding relationships between characteristics of power traces and a hypothesised power model. If correlation is found between these two variables, one has the capability to predict the correct secret key if enough power traces are gathered.

We focus on experiments to study one of the first original DPA attacks which is referred to as Difference of Means .

## 2. DPA: Difference of Means

In DPA, the hypothesis is that small variations in power level may be observed in a trace based on the output of an encryption algorithm. An example of this is to observe the least significant bit (LSB) of an output. The LSB is the unit bit of a binary number. For example, the LSB of binary value 11001101 is 1. If the LSB output produced by an encryption algorithm is 1 then, in theory, the device under test should consume more

power in comparison with a 0 value. If the hypothesis holds true, we may exploit this fact to deduce the cipher key during cryptographic operations.

It must be emphasised that this variation in power consumption is very small and almost impossible to detect by the naked eye. Thus, in DPA, one must gather a large number of traces, sort them into two subsets and calculate the average of each subset. This technique is referred to as the Difference of Means attack. The difference in means between each subset allows one to deduce whether there is any significance in the proposed hypothesis. In the case of significance, the difference in averages between the two subsets will highlight the variation in power consumption when a LSB of 0 is compared against an LSB of 1 (thus proving the hypothesis). However, if the subsets are not sorted correctly, no significance should be found.

Sorting of traces is formally referred to as the selection function but, ultimately, it comes down to an educated guess when deciding which subset a trace belongs to. Referring back to the concept of LSB, the selection function criterion for this example would be to sort traces into a subset where it is believed that all traces produce the LSB of 0 and a second subset where all traces produce the LSB of 1.

Once gathering and sorting of traces into two separate subsets is complete, the average of each subset is calculated in a point-by-point basis. The difference in averages can then be calculated by simply subtracting the points of the first subset against the points of the second subset. If the subsets are sorted correctly, a line plot of this difference of averages will result in unique peaks and nadirs easily discernible to the naked eye. On the other hand, if the subsets were sorted incorrectly, the process averaging and subtracting the difference in averages will cancel out each subset and results in a line plot close to 0 (although in practice, the values will never be exactly 0 due to noise and other interferences).

To mount such an attack in practice, one must have some form known variable whether it is an input or an output. An example of an input may be known plaintext values while an example of a output could be the ciphertext values produced by the cryptographic algorithm. By having a known input or output, the selection function can be implemented to sort power traces into the two subcategories based on hypothesis of how the known input will affect variations in power level (e.g. will LSB 1 or will LSB 0 be produced?) during key stages of a cryptographic operation which involve or relate back to the cipher key.

The Difference of Mean attack can be expressed as Equation 1. In this equation, $T$ is each individual power consumption trace while $T_i$ is the $i$th trace. The variable $j$ represents the power consumption value at time $j$th offset. $C$ represents the known inputs or outputs during an attack with $C_i$ being relative to the $i$th trace. The selection function itself is

represented as $D(C_i, K_n)$ which takes in the parameters of our known input or output $C_i$ along with the key guess $K_n$ where $n$ is relative to one of the cipher keys we are attempting to deduce ($n$ would be 16 in the case of AES-128). The top fraction of this equation represents the first subcategory of traces (as sorted by the selection function) while the bottom represents the second category of traces. The significant difference $\Delta_D$ for each point $j$ is calculated by subtracting the average of the first subcategory against the second subcategory.

$$\Delta_D[j] = \frac{\sum_{i=1}^{m} D(C_i, K_n) T_i[j]}{\sum_{i=1}^{m} D(C_i, K_n)} - \frac{\sum_{i=1}^{m} (1 - D(C_i, K_n) T_i[j])}{\sum_{i=1}^{m} (1 - D(C_i, K_n))}$$

## Our Work in Assignment:

We were provided the code of AES algorithm in python programming language, which was successfully implementing various functionalities like converting 128 bit plaintext into 4*4 matrix form which has each element in byte form, Substitution Box, Shift Rows, Mix Columns, Add Round key etc. Basically we are repeating this process for 10 rounds.

We were also provided the code to predict the 0$^{th}$ byte of the secret key where the DOM procedure we have applied as mentioned in the theory part of the report.

We are basically taking all the 256 values for the 4$^{th}$ byte and 5$^{th}$ byte of the key and for each of the 256 values i.e. from 0 to 255 we will calculate Difference of Mean (DOM) values for the 2 bins we have created by segregating each of the power traces based on the LSB value whether it is 0 or 1.

The DOM value will be maximum for which value of the 4$^{th}$ and 5$^{th}$ byte of the secret key will be the secret key. For the power trace file our group was assigned we have calculated the values of the **4$^{th}$ byte as 178 and of the 5$^{th}$ byte as 196**.

The whole code needs to run two times, one time for 4$^{th}$ byte and second time for 5$^{th}$ byte. We have used iteration variable in for loop for that purpose.

We have also plotted a graph corresponding to the DOM values of the 4$^{th}$ byte and the 5$^{th}$ byte , where the red spike is showing the highest value of the DOM.

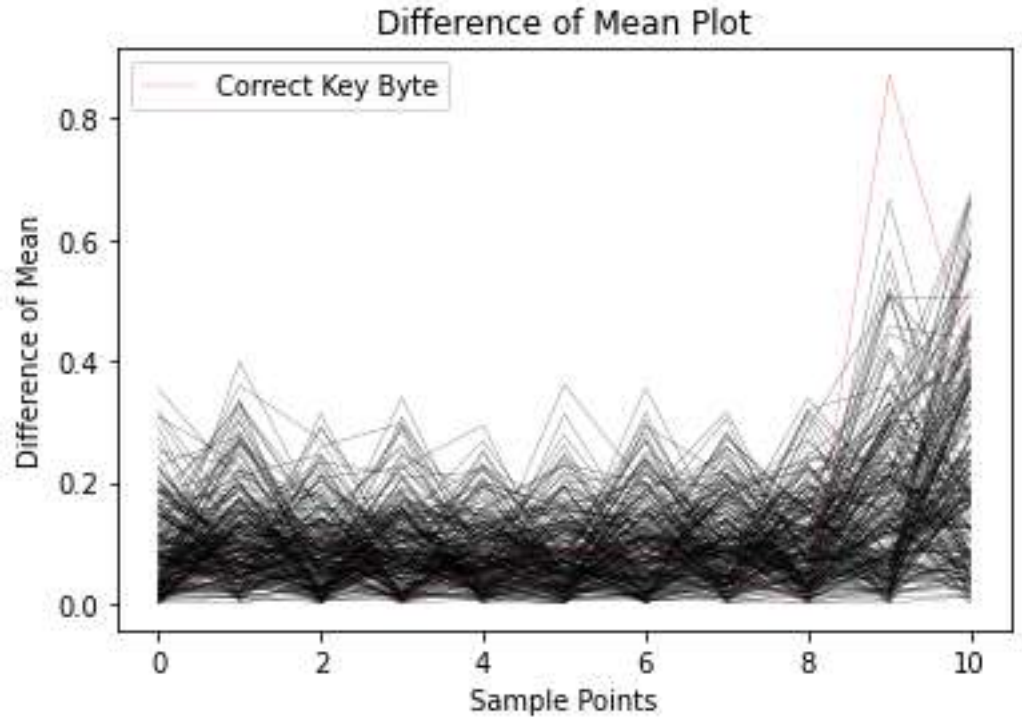



**Figure:** Graph for 4th Key Byte Output

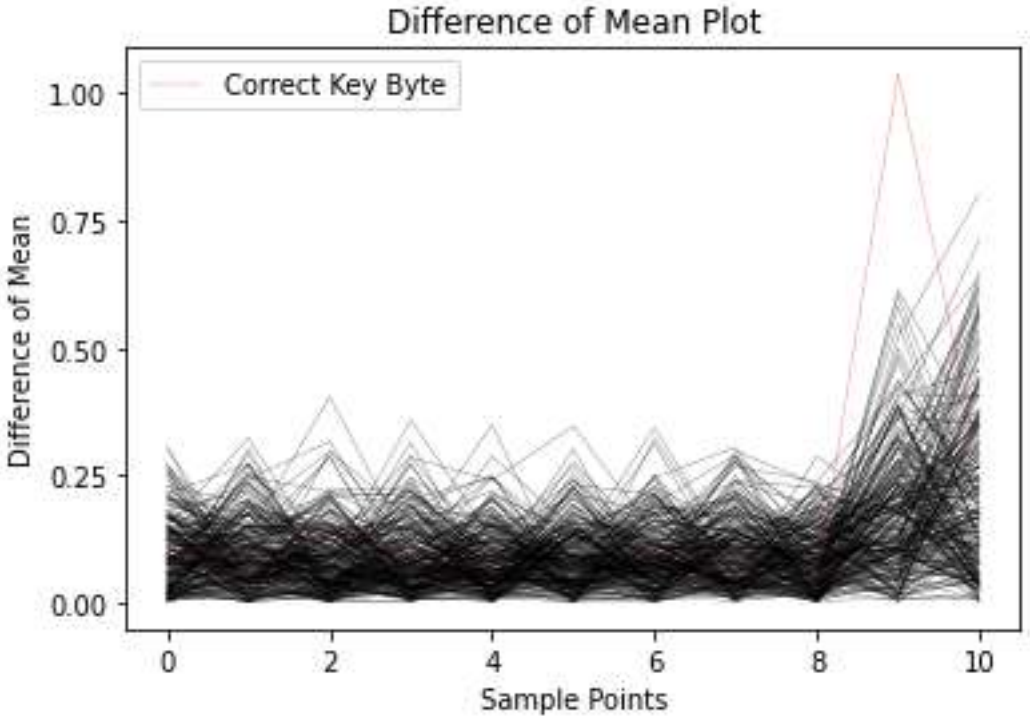



**Figure:** Graph for 5th Key Byte Output