Side Channel Analysis of an Embedded/Hardware Crypto Device

Sam Mitchell and Nathanael Weidler
Deptartment of Electrical and Computer Engineering
Utah Stat University
Logan, Utah 84322
e-mail: samuel.alan.mitchell@gmail.com, NWeidler@gmail.com

***Abstract*—This paper describes a side channel analysis attack on a microprocessor. The microprocessor is running a Data Encryption Standard (DES) algorithm. The goal of the attack is to recover the secret key from the device. This is done by capturing the power usage of the device and using differential power analysis to analyze the data.**

***Index Terms*—Data Encryption Standard, Differential Power Analysis, Correlation Power Analysis, Security.**

I. Introduction

Side channel analysis is one way to attack even the most computationally complex cryptographic devices. Through side channel analysis, information (such as secret keys) can be obtained from devices without using more traditional or brute force attacks. In order to carry out a side channel analysis attack, the attacker must have access to the device. The attacker then obtains power information using an oscilloscope connected to a resistor put in series with the device's power input.

In this paper a side channel analysis will be carried out on an Tiva-C microprocessor to obtain the secret key from a Data Encryption Standard (DES) algorithm. First, a typical DES implementation was written and loaded into the flash of the Tiva-C microprocessor. Next, the side channel analysis was carried out. After the power information, or traces, were obtained, a technique called differential power analysis (DPA) was used to analyze the data. Another analysis using correlation power analysis was also carried out. [1] [2]

A. Structure of paper

The organization is as follows: in Section II, the development of the DES implementation is presented. In Section III the experimental test setup for the capturing of the traces is described. Section IV contains the analysis of the data. Conclusions are detailed in Section V.

II. DES Implementation

As a part of this work an implementation of DES was written for the Tiva-C microprocessor. "The DES Algorith Illustrated" was a valuable resource during this process. All of the explicit details of DES will be left to other work, however, a brief overview will be presented in order for the reader to grasp the complexity of the algorithm. [1]

DES is a 16 round encryption algorithm. It works on 64-bit blocks of plaintext and outputs encrypted data of the same size. The output of the algorithm is the cipher text. The first step in DES is to divide the plaintext block in half, the left 32-bits and the right 32-bits. The heart of DES is the 56-bit master key which is used to create 16 distinct 48-bit round keys, one for each round of the algorithm. (The intent of this attack is to recover the key from the 16th round, from which all other keys can be easily obtained.) To create each round key the 56-bit key is permuted and divided into two halves. Then for each round each of those halves is shifted, recombined and permuted yielding the 48-bit round keys.

One half of the plaintext (32-bits) is expanded and permuted to create a 48-bit word which is xored with the round key. The result is divided into 8 6-bit parts which go through 8 distinct s-boxes. An s-box is method to substitute a 6-bit input with a 4-bit output. These 4-bit outputs are put back together and permuted to create a 32-bit word which is xored with the other half of the plaintext. This process is repeated 16 times until finally the two halves are put back together and permuted one last time to create the cipher text. An illustration of one round is shown in figure 1.

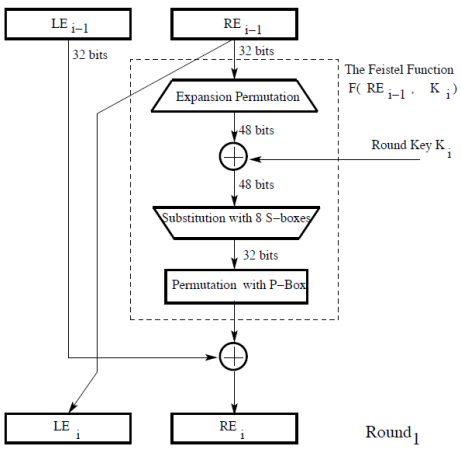

Figure 1. One round of DES.

A. DES Modifications

In order to facilitate the process by which traces are obtained, there was an addition to the traditional DES algorithm. This slight modification in no way compromises the integrity of the

algorithm or its outputs. The change was to hold a general purpose input/output (GPIO) high during the algorithm except during a single instruction in which the output of the Feistel function is xored with the other half of the plaintext in round 16 and the result is written back to a register. This allowed an oscilloscope to be triggered on the falling edge of the GPIO and frame this critical step which will be discussed later. There were also two no-operations (nops) added before this register write and one after to allow the GPIO to settle and facilitate a cleaner capture. The output of the assembly dump with comments added can be seen in figure 2. Note that the register where the xor is written to (R0) is set to zero before the output of the xor is written there.

```
MOVS        r0,#0x00       //setting R0 to 0
MOVS        r3,#0x00       //lowering GPIO
LDR         r4,[pc,#1016]  //lowering GPIO
STR         r3,[r4,#0x00]  //lowering GPIO
NOP                        //No Op
NOP                        //No Op
EOR         r0,r2,r1       //Single instruction to attack
NOP                        //No Op
MOVS        r3,#0x02       //raising GPIO
LDR         r4,[pc,#1004]  //raising GPIO
STR         r3,[r4,#0x3FC] //raising GPIO
```

Figure 2. The assembly showing the GPIO writes surrounding the xor and register write under attack.

III. Experimental Setup

This portion of the experiment was the most crucial to its success. The goal is to obtain information about the power usage of the microprocessor during a critical step of the DES algorithm. During a register write operation the power used will be different depending on the value of the word written. It is believed that by analyzing this information, the secret key to the algorithm can be obtained.

The oscilloscope used was a Techtronix DP0-2024 the sample rate was 1Gs per second and each trace contained about 270 samples. The Tiva-C used a 17 MHz clock so there 38.5 samples per clock.

The setup used involved the Tiva-C microprocessor powered by a bench top power supply through a breadboard. The ground connection between the microprocessor and the power supply had a 1-Ω resistor added in series. An Oscilloscope probe was placed on the resistor so that the voltage could be measured. Then another oscilloscope probe was placed on the GPIO pin discussed in the previous section to act as the trigger. The voltage data would be taken twenty times for each plain text value used in the DES algorithm and then averaged in an attempt to reduce noise. 91,840 captures with 4,592 unique plain text values would be obtained. Each trace takes about two seconds to capture so the time to obtain all traces is 183,680 seconds or about 51 hours. The voltage data acquired would be analyzed later to determine the secret key. The test setup is illustrated in figure 3

The test setup needed to be automated to facilitate the capturing of the traces. In order for this to be done a matlab file was written to interface with the oscilloscope. As trivial as this may sound it was not easy. It took many hours of learning the oscilloscope and the commands to control it through matlab. It also took time and patience to fine tune the DES implementation to create enough delay between each run of the algorithm to allow for the oscilloscope to transfer the data capture and re-arm itself.

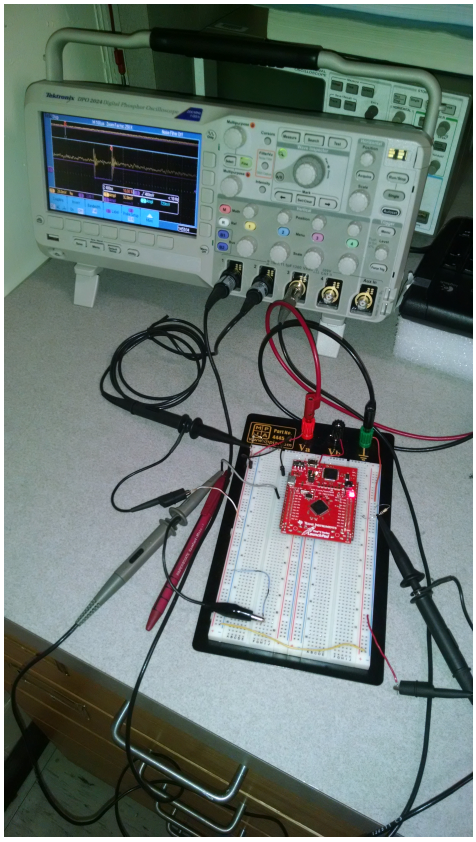

Figure 3. Test setup showing the oscilloscope connected to the Tiva-C board.

IV. Analysis of the Data

A. Differential Power Analysis

In order to analyze the data, differential power analysis (DPA) was employed. Round 16 was chosen for this because given the cipher text, the input to the Feistel function and the output from the second xor (seen in the bottom right of figure 1) can be determined. Then the round key is guessed 6 bits at a time. This greatly reduces the complexity of the key as there are only 64 possible combinations for a guess of 6 bits of the key. Using knowledge of the Feistel function and the 6-bit key guess, the output of the Feistel function can be guessed at 4 bits at a time. Remember, the s-boxes reduce the data from 48 bits to 32 bits. This output is xored with the D value, and we already know the output of this xor. So guessing one input and knowing the output determines the other input to the xor gate.

For each key guess 4-bits of the D word are deduced. These 4-bit groups are divided into 3 sets where set 0 contains all guesses for which the four bits guessed are "0000," set 1 contains all guess where the 4 bits are "1111," and set 2 contains all other guessed values. The average power for each sample point for sets 0 and 1 is calculated. The two average

powers are subtracted from each other. The correct key guess will show large spikes where the calculation was actually made on the trace. This is where the difference between the 2 sets of data is the greatest. This process is repeated 8 times, once for each s-box until the correct key is determined.

Figures 6 through 13 in Apendix A show the outputs of the DPA. Unfortunately, none of the key values were able to be recovered based on DPA. More discussion of this to follow.

B. Correlation Power Analysis

Correlation Power Analysis (CPA) was also used on the data captured to see if more clear results could be obtained than what DPA provided. In CPA the guessed at D word is sorted using the hamming distance of the change in the register value. I our case the register was always set to 0 prior to the calculation, so in this case the Hamming distance is the same as the Hamming weight of the result. Hamming weight is the number of 1s a data word contains. The Hamming distance is the number bit changes from one word to the next so you count the number of 0s that changed to 1s and add that number to the number of 1s that changed to 0s. The Hamming distance of anything from 0 is the Hamming weight of the data word. Each sample from every trace is entered into Pearson's sample correlation coefficient along with the Hamming weight of the guess. The result should show a high correlation for the correct key guess. The key guesses are the same 6-bit key guesses that were used in the DPA attack. The finest resolution the oscilloscope showed was $20mV$/div. The bulk of the traces had nominal values of $4mV$, with fluctuations of $\pm 5mV$. The oscilloscope has 8-bit resolution, which translates to a resolution of $0.1mV$. By using such low resolution, the representation of the significant portion of the signal (the xor operation) is reduced to 12 values or 4 bits. The error introduced by discretization could be one reason for inaccurate reading of the system. As an illustration of this, Figure 4 shows the measurement error between traces with matching plaintexts.

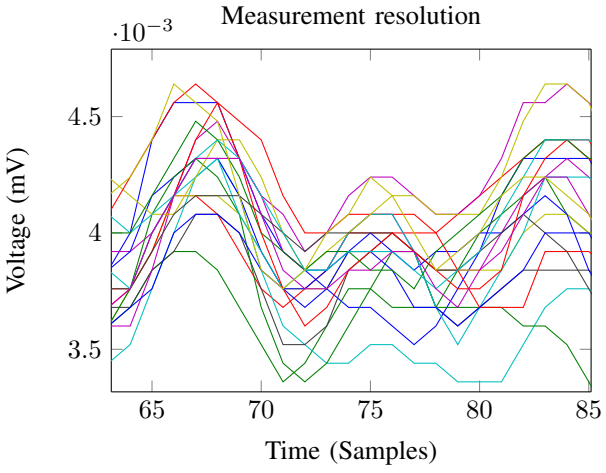

Figure 4. resoltion .

Figures 14 through 21 in Apendix A show the results of the CPA. Unfortunately, none of the key values were able to be recovered based on CPA. This will be discussed in more detail in the following section.

C. Actual Key Values Analysed

This experiment is unique from an actual side channel analysis attack because we knew what the secret key was. This meant that we could run the correct values through DPA and CPA as was done previously with guesses.

Figure 5 shows the Hamming distance model. This figure shows the actual key run through Pearson's sample correlation coefficient with the captured traces. A good correlation would be close to 1, however as can be seen in figure 5 the highest correlation we saw was close to 0.03. This shows that the data wasn't closely correlated with the actual output of the s-boxes anywhere. This is a problem because if we know the key and we know we should see a correlation somewhere but we do not, there is no way we could use guesses to find out the key. The actual key doesn't correlate to any value on our traces. There will be more discussion of this in the conclusion.

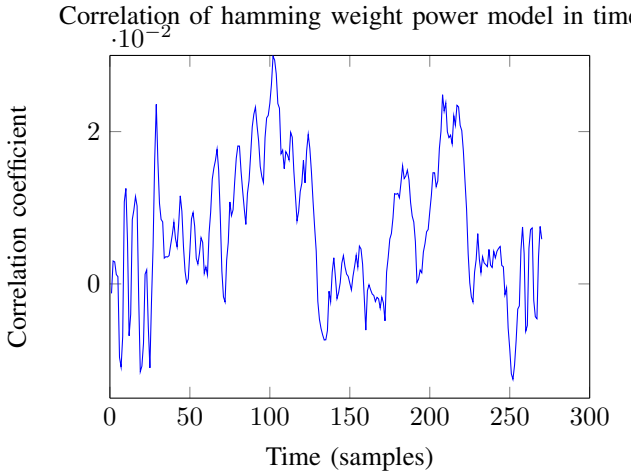

Figure 5. Measured Hamming Distance.

V. Conclusion

Both DPA and CPA have previously been shown to work at recovering secret keys from DES. We were not able to reproduce such promising results. We still believe that DPA and CPA should work to recover secret keys from DES but there were some potential flaws in our methods.

One aspect to this project which could have been done better would have been to take more traces for a single plain text value. We took 20 traces for each plain text, but we believe that if we had taken many more than this the noise would have averaged out more nicely and the results may have been better. Perhaps we should have taken an order of magnitude more traces than what we did. We believe that this was the main contributing factor to the side channel attacks not producing the expected results.

We know there is a problem with our data because no correlation is seen to the actual key. We think the most likely

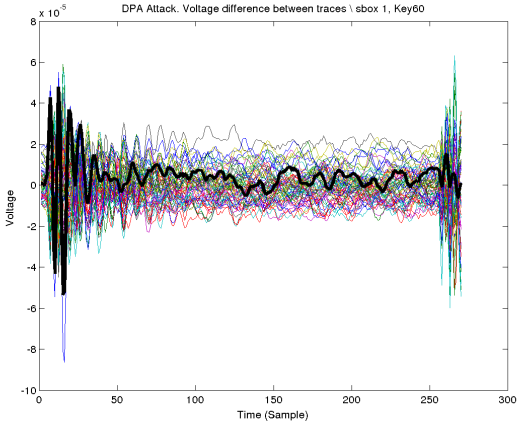

Figure 6. SBox 1 using DPA.

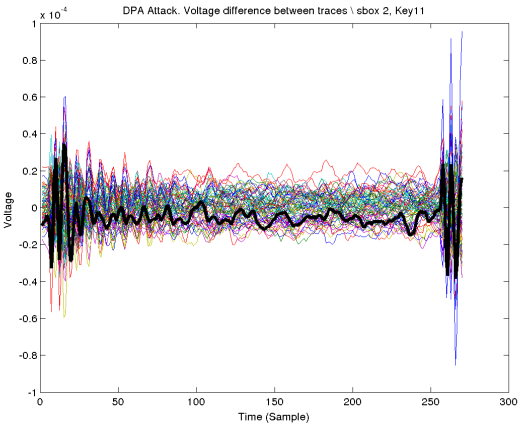

Figure 7. SBox 2 using DPA.

reason is the noise present in the system. Given more time we might be able to come up with a new experimental setup and method of capturing data that would yield better results. The lesson learned is that your analysis can only be as good as your data is. If the data is not good you can't analyze it.

References

[1] J. O. G. Grabbe, "The des algorithm illustrated," *Laissez Faire City Times*, vol. 2, no. 28.

[2] T. S. M. Messerges, E. A. D. Dabbish, and R. H. Sloan, "Investigations of power analysis attacks on smartcards," *USENIX Workshop on Smartcard Technology*, vol. 7, no. 49, pp. 14–16, May 1999.

VI. Apendix A

DPA and CPA figures :

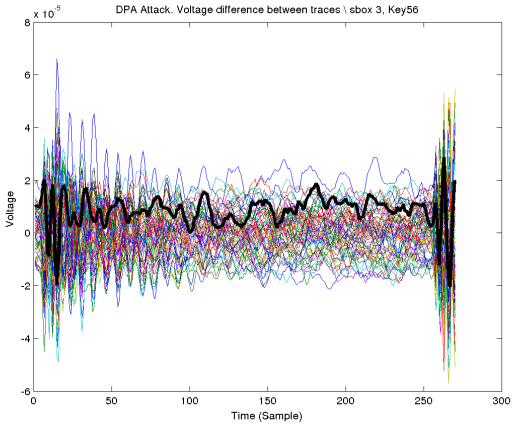

Figure 8. SBox 3 using DPA.

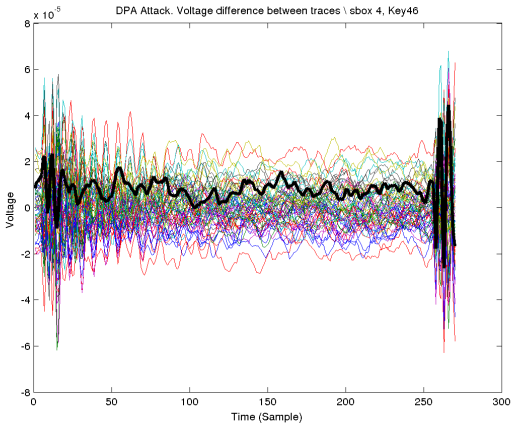

Figure 9. SBox 4 using DPA.

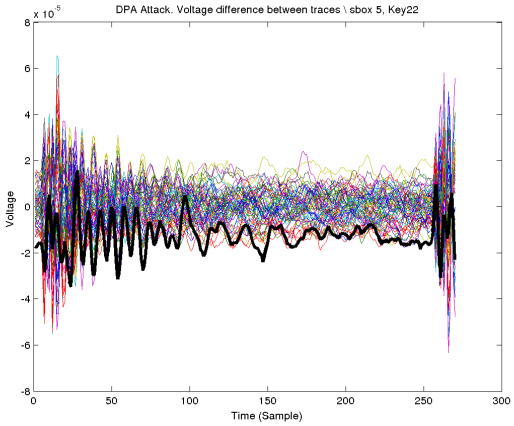

Figure 10. SBox 5 using DPA.

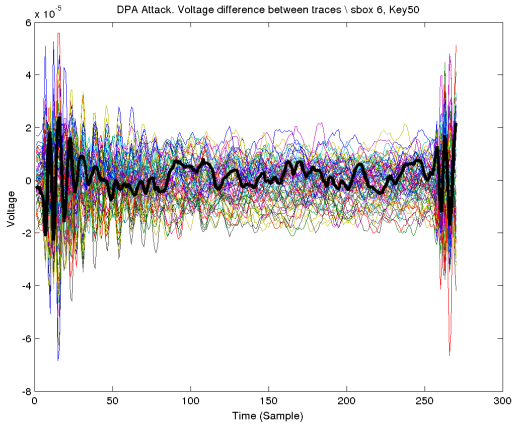

Figure 11. SBox 6 using DPA.

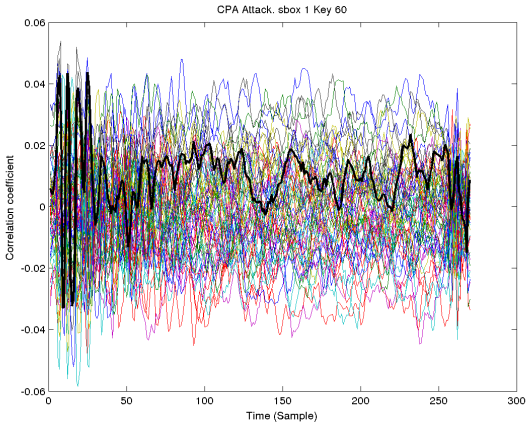

Figure 14. SBox 1 using CPA.

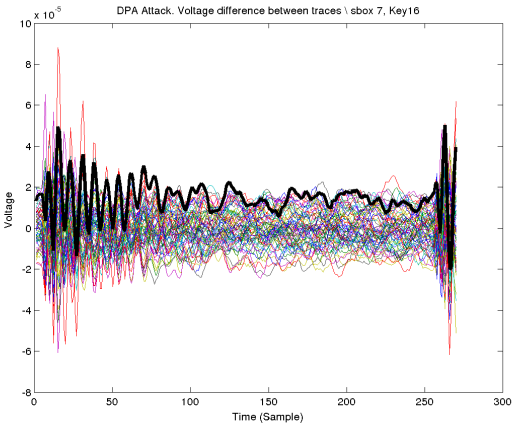

Figure 12. SBox 7 using DPA.

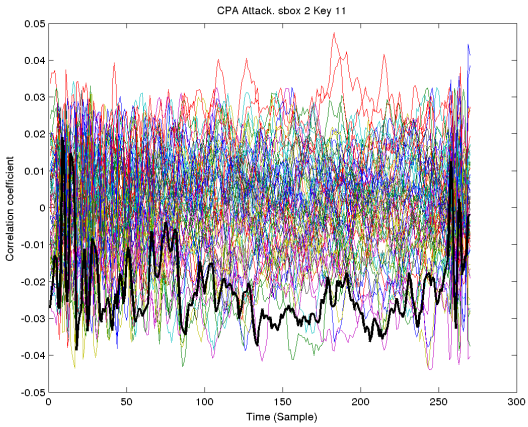

Figure 15. SBox 2 using CPA.

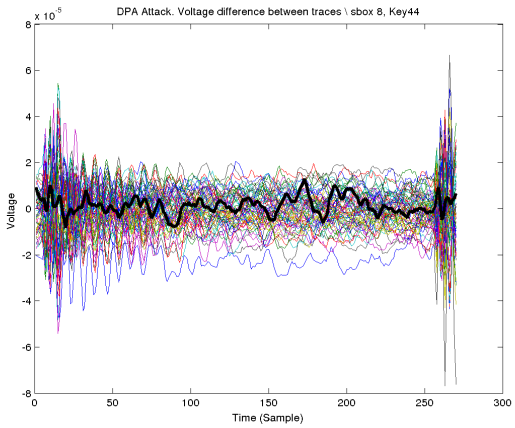

Figure 13. SBox 8 using DPA.

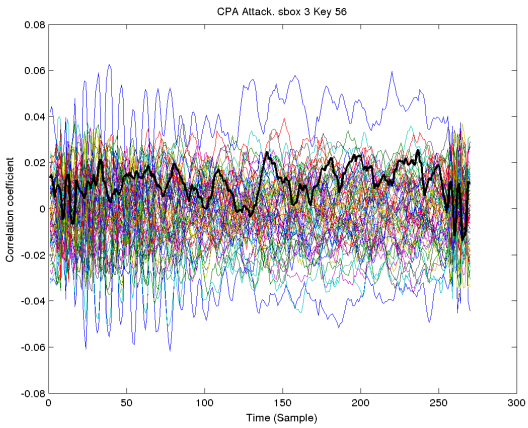

Figure 16. SBox 3 using CPA.

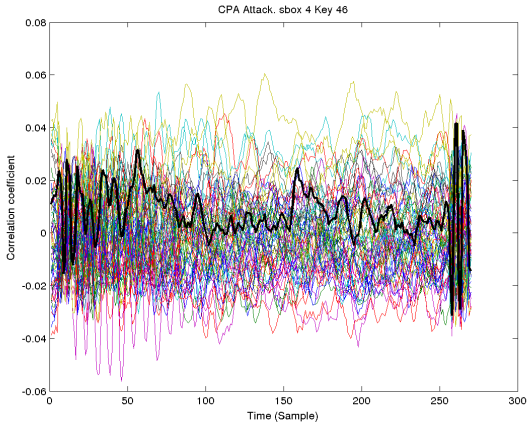

Figure 17. SBox 4 using CPA.

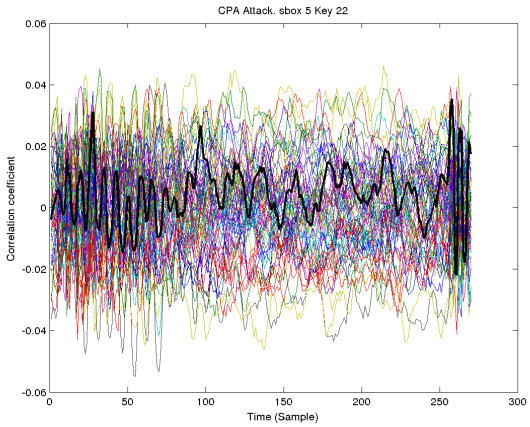

Figure 18. SBox 5 using CPA.

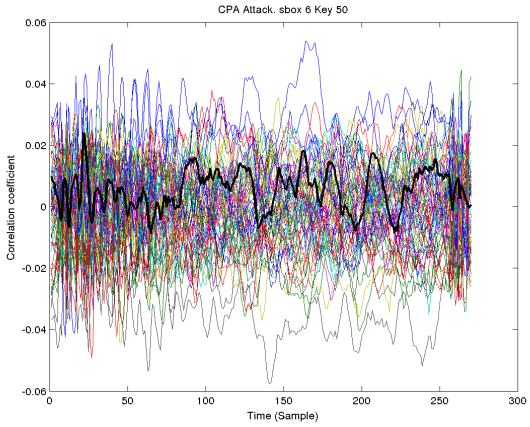

Figure 19. SBox 6 using CPA.

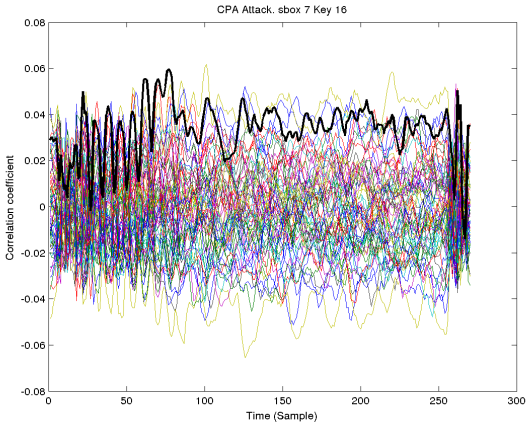

Figure 20. SBox 7 using CPA.

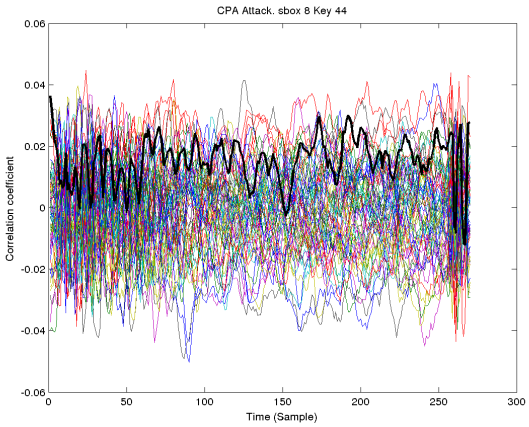

Figure 21. SBox 8 using CPA.