A First-Order DPA Attack Against AES in Counter Mode with Unknown Initial Counter

Josh Jaffe

Cryptography Research, Inc.
575 Market Street, suite 2150, San Francisco, CA 94105, USA.
josh@cryptography.com

Abstract. Previous first-order differential power analysis (DPA) attacks have depended on knowledge of the target algorithm's input or output. This paper describes a first-order DPA attack against AES in counter mode, in which the initial counter and output values are all unknown.

Keywords: power analysis, SPA, DPA, HO-DPA, AES, counter mode.

1 Introduction

Previous first-order differential power analysis (DPA) attacks have depended on knowledge of the target algorithm's input or output [1][2]. This paper describes a first-order DPA attack against the Advanced Encryption Standard (AES) [3] in counter mode, in which the initial counter, input values, and output values are all unknown.

The attack proceeds as follows. Suppose the input data to an algorithm is unknown, but can be expressed as single secret constant summed with known, variable data. The known, variable part of the data is used to mount a DPA attack, and the secret constant is treated as part of the key to be recovered. The "key" recovered by the DPA attack is then a function of the actual key and the secret constant. The known input values are then combined with the recovered "key" to compute the actual intermediate values produced by the algorithm. The recovered intermediates are then used to carry the attack forward into later rounds, enabling additional DPA attacks to recover the real key.

The attack also addresses the challenges to DPA presented by block ciphers used in counter mode [4]. DPA attacks target secrets when they are mixed with known *variable* quantities. In counter mode only the low-order bits of the input change with each encryption. Hence there are few variable intermediates to target in the first round of a typical block cipher. We demonstrate a method for propagating the attack into later rounds in which more known, variable data is available.

Although counter mode presents additional challenges to DPA attacks, in certain respects it also makes the attack easier. Unlike most first-order DPA attacks, the sequential nature of the counter enables the attack to succeed with

only knowledge of the power measurements. Knowledge of input, output, and initial counter values are not required to implement the attack.

1.1 Related Work

Simple power analysis (SPA) attacks have been used to extract portions of keys directly from power traces without requiring knowledge of input messages. Fahn and Pearson used inferential power analysis (IPA), an attack that exploits binary SPA leaks [5]. Mayer-Sommer presented attacks exploiting SPA leaks in high-amplitude power variations [6]. Mangard presented an SPA attack against the AES key expansion step [7]. Messerges et al described SPA attacks on Hamming weight and transition count leaks [8].

Side channel collision attacks were introduced by Dobbertin, and have traditionally targeted SPA leaks using chosen ciphertext [9] [10] [11]. Side channel collision attacks can be adapted to the case in which inputs are known to be successive values of a counter.

High-order differential power analysis (HO-DPA) [12] attacks target a hypothesized key-dependent relationship between data parameters in a computation. Previous work has noted that HO-DPA attacks can be applied to situations in which cipher input values are not known [13].

Fouque and Valette presented the "doubling attack" [14] which exploits the relationship between inputs in successive RSA decryptions to recover the exponent. The attack succeeds despite the fact that the input to the modular exponentiation step is masked by a blinding factor. Messerges presented a second-order DPA attack [15] that defeated a data whitening scheme.

Chari et al [16] and Akkar et al [17] also presented DPA attacks on block ciphers with a "whitening" step.

2 Preliminaries

2.1 Notation

Suppose X and Y are used to denote input and output data of a transformation. (Letters other than X or Y will also be used.) If the transformation is implemented as a sequence of rounds, the input and output of the i^{th} round are denoted by X_i and Y_i.

Within a round, data may be partitioned into bytes for processing. $X_{i,j}$ and $Y_{i,j}$ denote the j^{th} bytes of round data X_i and V_i.

K is used to denote input keys, K_i denotes the i^{th} round key derived from K, and $K_{i,j}$ denotes the j^{th} byte of round key K_i.

Symbols

The symbol '$\oplus$' denotes the bitwise XOR of two n-bit vectors.
The symbol '$+$' denotes the ordinary addition of two numbers.
The symbol '$\circ$' denotes multiplication between two elements of $GF(2^8)$.
The symbol '||' denotes the concatenation of two vectors.

2.2 Description of AES

Although most readers are no doubt familiar with AES, this section gives a brief review of its design. The round transformations are grouped differently than in the AES standard to facilitate presentation of the attack, but the algorithm described here is equivalent to AES. The review will also familiarize the reader with the notation and concepts used in this paper.

AES is a block cipher that operates on 16-byte blocks of data. It is designed as a sequence of 10, 12, or 14 rounds, depending on whether the key K is 16, 24, or 32 bytes in length. The key is expanded by the AES key schedule into 16-byte round keys K_i.

The round structure of AES encryption. The following transformations are performed during each round of an AES encryption:

1. AddRoundKey
2. SubBytes
3. ShiftRows
4. MixColumns[1]

These operations are described below, using the following notation for intermediate round states:

X_i denotes the input to round i and the AddRoundKey transformation.
Y_i denotes the output of the AddRoundKey transformation and the input to the SubBytes transformation.
Z_i denotes the output of the SubBytes transformation and the input to the ShiftRows transformation.
U_i denotes the output of the ShiftRows transformation and the input to the MixColumns transformation.
V_i denotes the output of the MixColumns transformation and the input to the next round: $V_i = X_{i+1}$.

AddRoundKey Each byte of $Y_{i,j}$ is produced by computing the exclusive or (XOR) of a byte of incoming data $X_{i,j}$ with the corresponding byte of round key $K_{i,j}$:

$$Y_{i,j} = X_{i,j} \oplus K_{i,j} \ . \tag{1}$$

SubBytes Each byte of input data is transformed via an invertible non-linear 8-bit lookup table S:

$$Z_{i,j} = S[Y_{i,j}] = S[X_{i,j} \oplus K_{i,j}] \ . \tag{2}$$

[1] The MixColumns operation is not performed in the final round, and an additional AddRoundKey operation is performed after the final round.

ShiftRows ShiftRows permutes the bytes within the data vector:

$$U_i = \boxed{Z_{i,0}}\boxed{Z_{i,5}}\boxed{Z_{i,10}}\boxed{Z_{i,15}}\boxed{Z_{i,4}}\boxed{Z_{i,9}}\boxed{Z_{i,14}}\boxed{Z_{i,3}}\boxed{Z_{i,8}}\boxed{Z_{i,13}}\boxed{Z_{i,2}}\boxed{Z_{i,7}}\boxed{Z_{i,12}}\boxed{Z_{i,1}}\boxed{Z_{i,6}}\boxed{Z_{i,11}}$$

MixColumns The j^{th} column of U_i is defined to be the four bytes

$$\{U_{i,4j}, U_{i,4j+1}, U_{i,4j+2}, U_{i,4j+3}\} .$$

MixColumns is an invertible linear transformation over $GF(2^8)$ performed on the columns of U_i. The j^{th} column of output V_i is defined to be:

$$\begin{aligned}
V_{i,4j} &= (\{02\} \circ U_{i,4j}) \oplus (\{03\} \circ U_{i,4j+1}) \oplus (\{01\} \circ U_{i,4j+2}) \oplus (\{01\} \circ U_{i,4j+3}) \\
V_{i,4j+1} &= (\{01\} \circ U_{i,4j}) \oplus (\{02\} \circ U_{i,4j+1}) \oplus (\{03\} \circ U_{i,4j+2}) \oplus (\{01\} \circ U_{i,4j+3}) \\
V_{i,4j+2} &= (\{01\} \circ U_{i,4j}) \oplus (\{01\} \circ U_{i,4j+1}) \oplus (\{02\} \circ U_{i,4j+2}) \oplus (\{03\} \circ U_{i,4j+3}) \\
V_{i,4j+3} &= (\{03\} \circ U_{i,4j}) \oplus (\{01\} \circ U_{i,4j+1}) \oplus (\{01\} \circ U_{i,4j+2}) \oplus (\{02\} \circ U_{i,4j+3})
\end{aligned}$$

where $\{01\}$, $\{02\}$, $\{03\}$, and $U_{i,4j}$, $U_{i,4j+1}$, $U_{i,4j+2}$, $U_{i,4j+3}$ are considered 8-bit vectors representing elements in $GF(2^8)$.

The linearity of the AES MixColumns transformation will be exploited during the attack. Suppose that input data can be selected such that in round i, one or more input bytes to the MixColumns operation are unknown, but are known to remain constant across multiple invocations of the AES algorithm. Then the contribution of these constant bytes to V_i is equivalent to XORing with fixed constants.

For example, suppose bytes $U_{1,4j+1}$, $U_{1,4j+2}$, and $U_{1,4j+3}$ are constant (but unknown) across multiple invocations of AES. Then the values

$$\begin{aligned}
E_{1,4j} &= (\{03\} \circ U_{1,4j+1}) \oplus (\{01\} \circ U_{1,4j+2}) \oplus (\{01\} \circ U_{1,4j+3}) \\
E_{1,4j+1} &= (\{02\} \circ U_{1,4j+1}) \oplus (\{03\} \circ U_{1,4j+2}) \oplus (\{01\} \circ U_{1,4j+3}) \\
E_{1,4j+2} &= (\{01\} \circ U_{1,4j+1}) \oplus (\{02\} \circ U_{1,4j+2}) \oplus (\{03\} \circ U_{1,4j+3}) \\
E_{1,4j+3} &= (\{01\} \circ U_{1,4j+1}) \oplus (\{01\} \circ U_{1,4j+2}) \oplus (\{02\} \circ U_{1,4j+3})
\end{aligned}$$

will be constant, and the MixColumns output can be expressed as

$$\begin{aligned}
V_{1,4j} &= (\{02\} \circ U_{1,4j}) \oplus E_{1,4j} \\
V_{1,4j+1} &= (\{01\} \circ U_{1,4j}) \oplus E_{1,4j+1} \\
V_{1,4j+2} &= (\{01\} \circ U_{1,4j}) \oplus E_{1,4j+2} \\
V_{1,4j+3} &= (\{03\} \circ U_{1,4j}) \oplus E_{1,4j+3} .
\end{aligned} \tag{3}$$

As will be shown in Section 3, the constant, unknown terms E can then be incorporated into the round key of the next round, and effectively ignored.

2.3 Counter Mode

Counter mode is a standard mode of operation for block ciphers in which ciphertext is produced by encrypting a counter and XORing the result with the plaintext block. Let B be a block cipher using key K, C the initial counter value, and X_T the T^{th} block of plaintext to be encrypted. Then the T^{th} block of ciphertext Y_T is given by

$$Y_T = X_T \oplus B_{enc}(C + T, K) \ .$$

Ciphertext is decrypted by XORing it with same encrypted counter value:

$$X_T = Y_T \oplus B_{enc}(C + T, K) \ .$$

Since counter values are inputs to the first round only, C_j and T_j will be used to denote the j^{th} bytes of C and T respectively, and not their values at round j. See [4] for more information on counter mode.

Galois counter mode Galois counter mode (GCM) [18] is a draft counter mode protocol currently being studied by NIST. In GCM, the initial counter value is derived from a variable-sized initialization vector (IV). If the length of the IV is not exactly 96 bits, then the initial counter value C is derived from the IV using a secret key. In protocols where the IV is exactly 96 bits long, at least part of the initial counter value may be secret. For example, in RFC 4106 [19] the first four bytes of the IV are derived with the AES key and may remain secret. The attack described in this paper assumes that the entire initial counter value C is unknown.

3 The Attack on AES in Counter Mode

This section will present a first-order DPA attack against AES in counter mode with unknown initial counter value C.

To keep the index notation from getting too cumbersome, the symbol "T" is omitted from subscripts. When data is described as constant or variable, however, it means that the data is constant or variable with respect to T. For example, when we say that an attack recovers a variable such as $Z_{1,15}$, it means that it recovers each value the variable took for each value of T.

3.1 Overview

The main stages of the attack are as follows:

1. Perform data collection.
2. Use DPA against the first round to recover $Z_{1,15}$ and $Z_{1,14}$.