Exact Maximum Expected Differential and Linear Probability for 2-Round Advanced Encryption Standard (AES)

Liam Keliher and Jiayuan Sui

Department of Mathematics and Computer Science
Mount Allison University
Sackville, New Brunswick, Canada, E4L 1E6
{lkeliher,js}@mta.ca

Abstract. Provable security of a block cipher against differential / linear cryptanalysis is based on the *maximum expected differential / linear probability* (MEDP / MELP) over $T \geq 2$ core rounds. Over the past few years, several results have provided increasingly tight upper and lower bounds in the case $T = 2$ for the Advanced Encryption Standard (AES). We show that the *exact* value of the 2-round MEDP / MELP for the AES is equal to the best known lower bound: $53/2^{34} \approx 1.656 \times 2^{-29}$ / $109,953,193/2^{54} \approx 1.638 \times 2^{-28}$. This immediately yields an improved upper bound on the AES MEDP / MELP for $T \geq 4$, namely $\left(53/2^{34}\right)^4 \approx 1.881 \times 2^{-114}$ / $\left(109,953,193/2^{54}\right)^4 \approx 1.802 \times 2^{-110}$.

Keywords: AES, Rijndael, SPN, provable security, differential cryptanalysis, linear cryptanalysis, maximum expected differential probability, maximum expected linear probability

1 Introduction

Several recent papers have dealt with provable security against differential and linear cryptanalysis for block ciphers based on the *substitution-permutation network* (SPN) structure [2, 4–8, 11–13]. Most of these results apply directly to the Advanced Encryption Standard (AES) [3] (originally named *Rijndael*). Demonstrating provable security against differential / linear cryptanalysis involves proving that the *maximum expected differential / linear probability* (MEDP / MELP) is sufficiently small over T core rounds—this is because the *data complexity* of the attack (the number of plaintext-ciphertext pairs required) is proportional to the inverse of the MEDP / MELP.

Since in general it is difficult to compute the MEDP / MELP exactly, researchers have focused on bounds. A series of progressively smaller upper bounds has been obtained for the AES; the best of these is 1.161×2^{-111} (MEDP) / 1.064×2^{-106} (MELP) for $T \geq 4$ [12].[1] Many such bounds are based on careful

---

[1] The upper bounds as stated in [12] (and cited in [6]) are 1.144×2^{-111} (MEDP) and 1.075×2^{-106} (MELP). The difference here is due to rounding; the values in the current paper are more accurate.

examination of the case $T = 2$. Prior to this paper, the 2-round AES MEDP was known to lie between $53/2^{34}$ and $79/2^{34}$, and the 2-round AES MELP was known to lie between $109,953,193/2^{54}$ and $192,773,764/2^{54}$ [2, 6, 12]; in both cases, the upper bound had been shown not to be tight [6]. In this paper, we show that the 2-round AES MEDP / MELP is in fact equal to the known lower bound. This immediately yields an improved upper bound for the AES for $T \geq 4$, namely $\left(53/2^{34}\right)^4 \approx 1.881 \times 2^{-114}$ (MEDP) / $\left(109,953,193/2^{54}\right)^4 \approx 1.802 \times 2^{-110}$ (MELP).

There is a well-known duality between differential cryptanalysis and linear cryptanalysis that often allows results for one attack to be translated into corresponding results for the other [1]. Since this is applicable to what follows, we focus on differential cryptanalysis; the modifications relevant to linear cryptanalysis are outlined in Section 5.

2 Background Concepts

Let N denote the cipher block size. An SPN consists of a sequence of rounds, each of which involves: (a) XOR with an N-bit subkey (*key-mixing stage*), (b) parallel application of M bijective $n \times n$ s-boxes ($M = N/n$) (*substitution stage*), (c) processing through a linear transformation $\mathcal{L} : \{0,1\}^N \to \{0,1\}^N$ (*linear transformation stage*). For the purpose of analysis, we assume that the subkeys are chosen uniformly and independently from $\{0,1\}^N$. We number the s-boxes in any substitution stage $1 \dots M$, left to right.

Let $B : \{0,1\}^d \to \{0,1\}^d$, let $\Delta\mathbf{x}, \Delta\mathbf{y} \in \{0,1\}^d$ be fixed, and let $\mathbf{X} \in \{0,1\}^d$ be a uniformly distributed random variable. The *differential probability* $DP(\Delta\mathbf{x}, \Delta\mathbf{y})$ is defined as

$$\mathrm{Prob}_{\mathbf{X}} \left\{ B(\mathbf{X}) \oplus B(\mathbf{X} \oplus \Delta\mathbf{x}) = \Delta\mathbf{y} \right\}.$$

We refer to $\Delta\mathbf{x}$ / $\Delta\mathbf{y}$ as input/output *differences*. It is natural to view the DP values as entries in a $2^d \times 2^d$ table.

If B is parameterized by a key, $\mathbf{k}$, we write $DP(\Delta\mathbf{x}, \Delta\mathbf{y}; \mathbf{k})$, and the *expected differential probability* $EDP(\Delta\mathbf{x}, \Delta\mathbf{y})$ is $E_{\mathbf{K}} [DP(\Delta\mathbf{x}, \Delta\mathbf{y}; \mathbf{K})]$, where $E[\ ]$ denotes expectation and $\mathbf{K}$ is uniformly distributed over the space of keys.

For T core cipher rounds, the *maximum EDP* (MEDP) is given by

$$\max_{\Delta\mathbf{x}, \Delta\mathbf{y} \in \{0,1\}^N \setminus \mathbf{0}} EDP(\Delta\mathbf{x}, \Delta\mathbf{y}).$$

An R-round block cipher is *provably secure* against differential cryptanalysis if, for certain values of $T \leq R$, the MEDP is sufficiently small that the corresponding data complexity is prohibitive (for SPNs, we often use $T = R - 2$). A particularly useful relationship exists for the AES and related SPNs: if μ is an upper bound on the 2-round MEDP (or MELP), then μ^4 is an upper bound on the MEDP (MELP) for $T \geq 4$ [12, 13].

Hereafter, all references to rounds are relative to $T \geq 2$ core rounds under consideration; often T will be implicit in the notation that is used. A *differential characteristic* is a vector $\Omega = \langle \Delta\mathbf{x}^1, \Delta\mathbf{x}^2, \dots, \Delta\mathbf{x}^{T+1} \rangle$, where $\Delta\mathbf{x}^t$ and

$\Delta\mathbf{x}^{t+1}$ are input/output differences for round t ($1 \leq t \leq T$). It follows that $\Delta\mathbf{x}^t$ and $\Delta\mathbf{y}^t = \mathcal{L}^{-1}(\Delta\mathbf{x}^{t+1})$ are input/output differences for the *substitution stage* of round t, yielding input/output differences for each s-box S_m^t in round t ($1 \leq m \leq M$), denoted $\Delta\mathbf{x}_m^t$ / $\Delta\mathbf{y}_m^t$. If $\Delta\mathbf{x}_m^t$ and $\Delta\mathbf{y}_m^t$ are both zero or both nonzero for any s-box, Ω is called *consistent* [14]; it suffices to limit consideration to consistent characteristics. For a given characteristic, Ω, an s-box with nonzero input/output differences is called *active*. The minimum number of active s-boxes in two consecutive rounds for any characteristic (excluding the all-zero characteristic) is the *differential branch number*, $\mathcal{B}_d$—this is determined by $\mathcal{L}$. The *expected differential characteristic probability* $EDCP(\Omega)$ is defined as

$$\prod_{t=1}^{T} \prod_{m=1}^{M} DP^{S_m^t}(\Delta\mathbf{x}_m^t, \Delta\mathbf{y}_m^t),$$

where $DP^{S_m^t}(\cdot, \cdot)$ is a DP value for s-box S_m^t.

The *differential* $DIFF(\Delta\mathbf{x}, \Delta\mathbf{y})$ is the set of all characteristics whose first difference is $\Delta\mathbf{x}$ and whose last difference is $\Delta\mathbf{y}$. The following well-known equality is central to our analysis [9]:

$$EDP(\Delta\mathbf{x}, \Delta\mathbf{y}) = \sum_{\Omega \in DIFF(\Delta\mathbf{x}, \Delta\mathbf{y})} EDCP(\Omega). \tag{1}$$

Given an input or output difference, $\Delta\mathbf{z}$, for the substitution stage of round t, the corresponding pattern of active s-boxes is denoted $\gamma_{\Delta\mathbf{z}} = \gamma_1 \gamma_2 \cdots \gamma_M \in \{0,1\}^M$, where $\gamma_m = 1$ if S_m^t is active, and $\gamma_m = 0$ otherwise.

The following table of values, determined by $\mathcal{L}$, is useful. For $\gamma, \hat{\gamma} \in \{0,1\}^M$,

$$W_d[\gamma, \hat{\gamma}] \stackrel{\text{def}}{=} \#\left\{ \Delta\mathbf{x} \in \{0,1\}^N : \gamma_{\Delta\mathbf{x}} = \gamma, \gamma_{\mathcal{L}(\Delta\mathbf{x})} = \hat{\gamma} \right\}.$$

3 Analysis of 2-Round SPN MEDP

Consider two consecutive SPN rounds; without loss of generality, omit $\mathcal{L}$ from round 2. Let $\gamma, \hat{\gamma} \in \{0,1\}^M \setminus \mathbf{0}$, and choose any $\Delta\mathbf{x}, \Delta\mathbf{y} \in \{0,1\}^N \setminus \mathbf{0}$ satisfying $\gamma_{\Delta\mathbf{x}} = \gamma$, $\gamma_{\Delta\mathbf{y}} = \hat{\gamma}$. It follows that $W = W_d[\gamma, \hat{\gamma}]$ is the number of characteristics in $DIFF(\Delta\mathbf{x}, \Delta\mathbf{y})$. Enumerate the active s-boxes as $S_1, S_2, \ldots, S_A$, where $A = wt(\gamma) + wt(\hat{\gamma})$. For each $\Omega_w \in DIFF(\Delta\mathbf{x}, \Delta\mathbf{y})$ ($1 \leq w \leq W$) and for each S_a ($1 \leq a \leq A$), let ε_a be the "inner" difference for S_a (an inner difference is either an output difference for a round-1 s-box, or an input difference for a round-2 s-box), and define the vector $V_w = \langle \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_A \rangle$; note that each $\varepsilon_a \in \{0,1\}^n \setminus \mathbf{0}$. Clearly $\{V_w\}_{w=1}^W$ depends only on $\gamma, \hat{\gamma}$, not on the specific values of $\Delta\mathbf{x}, \Delta\mathbf{y}$.

Lemma 1 ([12]). *For* $\gamma, \hat{\gamma} \in \{0,1\}^M \setminus \mathbf{0}$, *let* $W = W_d[\gamma, \hat{\gamma}]$, *and form the set of vectors* $\{V_w\}_{w=1}^W$. *(**Case I**) If* $wt(\gamma) + wt(\hat{\gamma}) = \mathcal{B}_d$, *then all the values in any one vector position are distinct.* *(**Case II**) If* $wt(\gamma) + wt(\hat{\gamma}) > \mathcal{B}_d$, *isolate any*

$(wt(\gamma) + wt(\hat{\gamma}) - \mathcal{B}_d)$ *vector positions, and fix a value in $\{0,1\}^n \setminus \mathbf{0}$ for each such position. Form the subset $\mathcal{V} \subseteq \{V_w\}$ consisting of all vectors containing the fixed values in the specified positions. Then for each position whose value was not fixed, all the values in that position are distinct as we range over $\mathcal{V}$.*

Definition 1. *A $\mathcal{B}_d$-list is a set of vectors, each of length $\mathcal{B}_d$, that has been derived in one of two ways:*

1. *by selecting any $\gamma, \hat{\gamma} \in \{0,1\}^M \setminus \mathbf{0}$ satisfying $wt(\gamma) + wt(\hat{\gamma}) = \mathcal{B}_d$, and forming the set $\{V_w\}$;*
2. *by selecting any not-yet-selected pair $\gamma, \hat{\gamma} \in \{0,1\}^M \setminus \mathbf{0}$ satisfying $wt(\gamma) + wt(\hat{\gamma}) > \mathcal{B}_d$, forming the set $\{V_w\}$, isolating $(wt(\gamma) + wt(\hat{\gamma}) - \mathcal{B}_d)$ vector positions, and then forming all possible subsets $\mathcal{V} \subseteq \{V_w\}$ in accordance with Case II of Lemma 1 (i.e., by using all possible choices of fixed values from $\{0,1\}^n \setminus \mathbf{0}$ for the isolated positions); each such $\mathcal{V}$ yields a $\mathcal{B}_d$-list by "shrinking" the vectors in $\mathcal{V}$ to length $\mathcal{B}_d$ via removal of the positions with fixed values.*

Let $\mathcal{B}_d\text{-LIST}^{(i)}$ be the set of all $\mathcal{B}_d$-lists formed by Option i above, for $i = 1, 2$, and let

$$\mathcal{B}_d\text{-LIST} = \mathcal{B}_d\text{-LIST}^{(1)} \cup \mathcal{B}_d\text{-LIST}^{(2)} .$$

Note that $\mathcal{B}_d\text{-LIST}^{(2)}$ is not uniquely defined.[2] For any $\mathcal{Z} \in \mathcal{B}_d$-LIST, let $\delta(\mathcal{Z})$ denote the number of vectors in $\mathcal{Z}$. Lemma 1 implies that $\delta(\mathcal{Z}) \leq (2^n - 1)$. For any vector $\mathbf{z} = \langle \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \ldots, \boldsymbol{\zeta}_{\mathcal{B}_d} \rangle$ in any $\mathcal{B}_d$-list, if $\boldsymbol{\zeta}_j$ is an output difference for a round-1 s-box, let $\boldsymbol{\alpha}_j$ be any *input* difference for the s-box, and let $DP^*(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j) = DP(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j)$. If $\boldsymbol{\zeta}_j$ is an input difference for a round-2 s-box, let $\boldsymbol{\alpha}_j$ be any *output* difference for the s-box, and let $DP^*(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j) = DP(\boldsymbol{\zeta}_j, \boldsymbol{\alpha}_j)$. (For simplicity, the specific s-box is implicit in the notation.)

Definition 2. *Let $\mathcal{Z} \in \mathcal{B}_d$-LIST. Define $\sigma(\mathcal{Z})$ as*

$$\max_{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{\mathcal{B}_d} \in \{0,1\}^n \setminus \mathbf{0}} \left(\sum_{\langle \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_{\mathcal{B}_d} \rangle \in \mathcal{Z}} \prod_{j=1}^{\mathcal{B}_d} DP^*(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j) \right) .$$

Theorem 1 ([6]). *The 2-round MEDP is lower bounded by*

$$\max \left\{ \sigma(\mathcal{Z}) : \mathcal{Z} \in \mathcal{B}_d\text{-LIST}^{(1)} \right\} .$$

Theorem 2 ([6]). *The 2-round MEDP is upper bounded by*

$$\max \left\{ \sigma(\mathcal{Z}) : \mathcal{Z} \in \mathcal{B}_d\text{-LIST} \right\} .$$

---

[2] This definition of $\mathcal{B}_d\text{-LIST}^{(2)}$ differs from [6]. Here, given $\gamma, \hat{\gamma} \in \{0,1\}^M \setminus \mathbf{0}$ in Option 2, each $\mathcal{B}_d$-list is formed for the *same* (arbitrary) choice of positions to be assigned fixed values; in [6], *all* such choices are used, but this is not necessary for our purposes (nor is it necessary for the results in [6]).

4 Exact 2-Round MEDP for the AES

The AES is an SPN with $N = 128$, $n = 8$, and all s-boxes identical [3]. The mapping $\mathcal{L}$ consists of a bytewise permutation followed by four identical 32-bit linear transformations applied in parallel. Consequently, analysis of the 2-round AES reduces to analysis of the simplified structure in Figure 1 for certain attacks—this is the case for differential (and linear) cryptanalysis. The branch number for the 32-bit linear transformation is $\mathcal{B}_d = 5$; hereafter we refer to 5-lists.
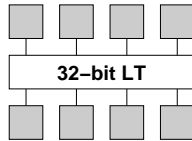

Fig. 1. Reduced 2-round AES

Our basic strategy for determining the exact value of the 2-round AES MEDP is to show that the lower bound of Theorem 1 and the upper bound of Theorem 2 are equal. Since computing $\sigma(\mathcal{Z})$ for a single 5-list $\mathcal{Z}$ involves a maximum over approximately 2^{40} terms, we use a pruning search to reduce complexity. (It is easy to show that 5-$LIST^{(1)}$ has size 56, which is manageable, but 5-$LIST^{(2)}$ has size approximately 2^{24}.)

We use the fact that all nontrivial rows and columns of the AES s-box DP table have the same distribution of values [12], given in the nonincreasing sequence $\langle d_1, d_2, \ldots, d_{256} \rangle$, where $d_1 = 2^{-6}$, $d_2, \ldots, d_{127} = 2^{-7}$, and $d_{128}, \ldots, d_{256} = 0$.

View any 5-list $\mathcal{Z}$ as a table of size $\delta(\mathcal{Z}) \times 5$ (each entry is a nonzero byte). Suppose we have selected values $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_J$ in Definition 2, with $1 \leq J \leq 5$. Let $\hat{\sigma}(\mathcal{Z}, J)$ be the largest value that can be contributed to the maximum $\sigma(\mathcal{Z})$ given the choice of $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_J$, i.e., if $1 \leq J < 5$, then

$$\hat{\sigma}(\mathcal{Z}, J) \;=\; \max_{\boldsymbol{\alpha}_{J+1}, \ldots, \boldsymbol{\alpha}_5 \in \{0,1\}^n \backslash \mathbf{0}} \left(\sum_{\langle \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_5 \rangle \in \mathcal{Z}} \; \prod_{j=1}^{5} DP^*(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j) \right),$$

and (trivially) if $J = 5$, then

$$\hat{\sigma}(\mathcal{Z}, J) \;=\; \sum_{\langle \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_5 \rangle \in \mathcal{Z}} \; \prod_{j=1}^{5} DP^*(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j).$$

Now form the sequence $\mathcal{S} = \langle s_1, s_2, \ldots, s_{\delta(\mathcal{Z})} \rangle$, where $s_i = \prod_{j=1}^{J} DP^*(\boldsymbol{\alpha}_j, \mathcal{Z}[i,j])$, and sort this sequence in *nonincreasing* order to obtain $\overline{\mathcal{S}} = \langle \bar{s}_1, \bar{s}_2, \ldots, \bar{s}_{\delta(\mathcal{Z})} \rangle$.

It follows from a generalized form of Lemma 5 in [7] that

$$\hat{\sigma}(\mathcal{Z}, J) \leq \Theta(\mathcal{S}, J) \overset{\text{def}}{=} \sum_{i=1}^{\delta(\mathcal{Z})} \bar{s}_i d_i^{(5-J)}, \tag{2}$$

and therefore $\Theta(\mathcal{S}, J)$ can be used as an easily computed "lookahead" value for pruning purposes. (Note that the *unsorted* $\mathcal{S}$ is passed to Θ.) Clearly equality holds in (2) when $J = 5$, since

$$\Theta(\mathcal{S}, 5) = \sum_{i=1}^{\delta(\mathcal{Z})} \bar{s}_i = \sum_{i=1}^{\delta(\mathcal{Z})} s_i = \hat{\sigma}(\mathcal{Z}, 5).$$

The heart of our algorithm is the function F in Figure 2, which uses a global variable $\mathcal{E}$. For positive integer L, let $\mathbf{1}_L$ be the sequence $\langle 1, \ldots, 1 \rangle$ of length L.
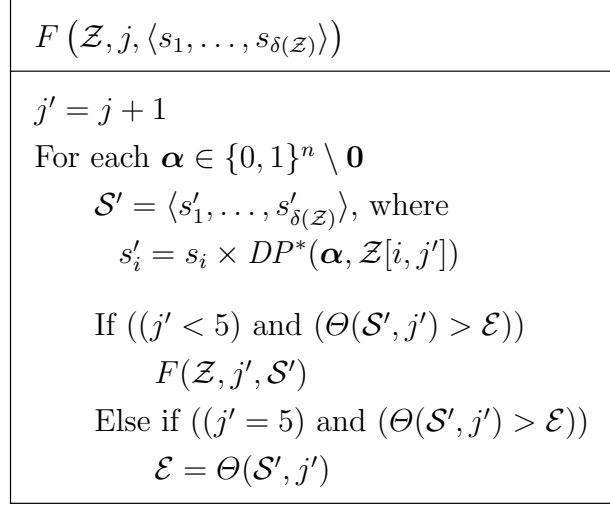
$$\boxed{\begin{array}{l} F\left(\mathcal{Z}, j, \langle s_1, \ldots, s_{\delta(\mathcal{Z})} \rangle\right) \\ \hline j' = j + 1 \\ \text{For each } \boldsymbol{\alpha} \in \{0,1\}^n \setminus \mathbf{0} \\ \qquad \mathcal{S}' = \langle s_1', \ldots, s_{\delta(\mathcal{Z})}' \rangle, \text{ where} \\ \qquad s_i' = s_i \times DP^*(\boldsymbol{\alpha}, \mathcal{Z}[i, j']) \\ \\ \qquad \text{If } ((j' < 5) \text{ and } (\Theta(\mathcal{S}', j') > \mathcal{E})) \\ \qquad\qquad F(\mathcal{Z}, j', \mathcal{S}') \\ \qquad \text{Else if } ((j' = 5) \text{ and } (\Theta(\mathcal{S}', j') > \mathcal{E})) \\ \qquad\qquad \mathcal{E} = \Theta(\mathcal{S}', j') \end{array}}$$

Fig. 2. Pruning search function F

Phase I. Initialize $\mathcal{E}$ to 0. For each $\mathcal{Z} \in 5\text{-}LIST^{(1)}$, call $F(\mathcal{Z}, 0, \mathbf{1}_{\delta(\mathcal{Z})})$. It is easy to see that if $\sigma(\mathcal{Z}) > \mathcal{E}$ prior to the call to F, then $\mathcal{E} = \sigma(\mathcal{Z})$ afterwards; otherwise, $\mathcal{E}$ is unchanged. It follows that when this phase is complete, $\mathcal{E}$ is equal to the lower bound of Theorem 1.

Phase II. Retain the value of $\mathcal{E}$ from Phase I. Call $F(\mathcal{Z}, 0, \mathbf{1}_{\delta(\mathcal{Z})})$ for each $\mathcal{Z} \in 5\text{-}LIST^{(2)}$. Then the final value of $\mathcal{E}$ is the upper bound of Theorem 2. If this upper bound is equal to the lower bound from Phase I, $\mathcal{E}$ is the *exact* 2-round MEDP.

4.1 Algorithm Results (MEDP)

Phase I of the above algorithm yields the lower bound $53/2^{34}$, a known result [2, 6]. What is significant is that *Phase II does not increase the value of* $\mathcal{E}$*,* and therefore the exact 2-round AES MEDP is equal to $53/2^{34} \approx 1.656 \times 2^{-29}$.

Further, making use of the fact that the 4th power of an upper bound on the 2-round AES MEDP is an upper bound for 4 or more rounds (as mentioned in Section 2), we obtain a new upper bound on the AES MEDP for $T \geq 4$, namely $\left(53/2^{34}\right)^4 \approx 1.881 \times 2^{-114}$.

5 Application to Linear Cryptanalysis

As stated above, the duality between differential cryptanalysis and linear cryptanalysis allows us to apply our approach, mutatis mutandis, to compute the exact 2-round AES maximum expected linear probability (MELP). The significant changes are as follows:

- Differential probability values are replaced by *linear probability* (LP) values (and EDP by ELP). For $B : \{0,1\}^d \to \{0,1\}^d$ and *masks* $\mathbf{a}, \mathbf{b} \in \{0,1\}^d$,

$$LP(\mathbf{a}, \mathbf{b}) \;=\; \left(2 \cdot \text{Prob}_{\mathbf{X}} \left\{\mathbf{a} \bullet \mathbf{X} = \mathbf{b} \bullet B(\mathbf{X})\right\} - 1\right)^2 \;,$$

  where $\bullet$ is the inner product over $\{0,1\}$.

- Given input/output masks for round t, $\mathbf{a}^t$ / $\mathbf{a}^{t+1}$, the output mask for the substitution stage is $\mathbf{b}^t = \mathcal{L}'(\mathbf{a}^{t+1})$, where $\mathcal{L}'$ is the matrix transpose of $\mathcal{L}$ when $\mathcal{L}$ is viewed as an $N \times N$ binary matrix (we use column vectors).

- Consistent differential characteristics are replaced by consistent *linear characteristics,* which are identically structured, but the constituent vectors from $\{0,1\}^N$ are interpreted as masks. EDCP is replaced by ELCP.

- The concept of *linearly active s-boxes* parallels that of differentially active s-boxes. Differential branch number is replaced by *linear branch number,* $\mathcal{B}_l$.

- Differentials $DIFF(\Delta\mathbf{x}, \Delta\mathbf{y})$ are replaced by *linear hulls* $ALH(\mathbf{a}, \mathbf{b})$, which consist of all linear characteristics (over T core rounds) having input mask $\mathbf{a}$ and output mask $\mathbf{b}$. The equation corresponding to (1) is given in [10]:

$$ELP(\mathbf{a}, \mathbf{b}) \;=\; \sum_{\Omega \in ALH(\mathbf{a}, \mathbf{b})} ELCP(\Omega) \,.$$

- An input or output mask, $\mathbf{z}$, for a substitution stage determines a pattern of active s-boxes, $\gamma_{\mathbf{z}} \in \{0,1\}^M$, just as in the differential setting. The table $W_d[\cdot, \cdot]$ is replaced by $W_l[\cdot, \cdot]$, where for $\gamma, \hat{\gamma} \in \{0,1\}^M$,

$$W_l[\gamma, \hat{\gamma}] \;=\; \#\left\{\mathbf{y} \in \{0,1\}^N : \gamma_{\mathbf{x}} = \gamma, \; \gamma_{\mathbf{y}} = \hat{\gamma}, \;\; \text{where } \mathbf{x} = \mathcal{L}'\mathbf{y}\right\} \,.$$

– All nontrivial rows and columns of the AES s-box LP table have the same distribution of values, given in Table 1 (ρ_i is a distinct value, and ϕ_i is the frequency with which it occurs) [7]. The sequence $\langle d_1, d_2, \ldots, d_{256} \rangle$ is modified accordingly.

i	1	2	3	4	5	6	7	8	9
ρ_i	$\left(\frac{8}{64}\right)^2$	$\left(\frac{7}{64}\right)^2$	$\left(\frac{6}{64}\right)^2$	$\left(\frac{5}{64}\right)^2$	$\left(\frac{4}{64}\right)^2$	$\left(\frac{3}{64}\right)^2$	$\left(\frac{2}{64}\right)^2$	$\left(\frac{1}{64}\right)^2$	0
ϕ_i	5	16	36	24	34	40	36	48	17

Table 1. Distribution of LP values for the AES s-box

5.1 Algorithm Results (MELP)

For the linear version of our algorithm, Phase I produced the known lower bound, $109,953,193/2^{54} \approx 1.638 \times 2^{-28}$ [2, 6]. And, as in the differential setting, Phase II did not increase this value, and therefore we conclude that this is the exact 2-round AES MELP.

In addition, we use the relationship stated in Section 2 to obtain a new upper bound on the AES MELP for $T \geq 4$, namely $\left(109,953,193/2^{54}\right)^4 \approx 1.802 \times 2^{-110}$.

6 Conclusion

Numerous papers have tackled the problem of determining (or bounding) the values of the 2-round maximum expected differential probability (MEDP) and maximum expected linear probability (MELP) for the AES. In this paper, we present a pruning search algorithm that enables us to prove that these values are equal to the best existing lower bounds, $53/2^{34} \approx 1.656 \times 2^{-29}$ (MEDP) and $109,953,193/2^{54} \approx 1.638 \times 2^{-28}$ (MELP). This immediately gives improved upper bounds on the AES MEDP and MELP for 4 or more rounds, namely $\left(53/2^{34}\right)^4 \approx 1.881 \times 2^{-114}$ and $\left(109,953,193/2^{54}\right)^4 \approx 1.802 \times 2^{-110}$, respectively.

Acknowledgments

References

1. E. Biham, *On Matsui's linear cryptanalysis,* Advances in Cryptology—EUROCRYPT'94, in: Lecture Notes in Comput. Sci., Vol. 950, Springer, Berlin, 1995, pp. 341–355.

2. K. Chun, S. Kim, S. Lee, S.H. Sung, and S. Yoon, *Differential and linear cryptanalysis for 2-round SPNs,* Inform. Process. Lett. 87 (2003) 277–282.

3. J. Daemen and V. Rijmen, *The Design of Rijndael: AES—The Advanced Encryption Standard,* Springer, Berlin, 2002.

4. S. Hong, S. Lee, J. Lim, J. Sung, and D. Cheon, *Provable security against differential and linear cryptanalysis for the SPN structure,* Fast Software Encryption—FSE 2000, in: Lecture Notes in Comput. Sci., Vol. 1978, Springer, Berlin, pp. 273–283.

5. J.-S. Kang, S. Hong, S. Lee, O. Yi, C. Park, and J. Lim, *Practical and provable security against differential and linear cryptanalysis for substitution-permutation networks,* ETRI J. 23 (2001) 158–167.

6. L. Keliher, *Refined analysis of bounds related to linear and differential cryptanalysis for the AES,* Fourth Conference on the Advanced Encryption Standard—AES4, in: Lecture Notes in Comput. Sci., Vol. 3373, Springer, Berlin, 2005, pp.42–57.

7. L. Keliher, H. Meijer, and S. Tavares, *New method for upper bounding the maximum average linear hull probability for SPNs,* Advances in Cryptology—EUROCRYPT 2001, in: Lecture Notes in Comput. Sci., Vol. 2045, Springer, Berlin, 2001, pp. 420–436.

8. L. Keliher, H. Meijer, and S. Tavares, *Improving the upper bound on the maximum average linear hull probability for Rijndael,* Selected Areas in Cryptography—SAC 2001, in: Lecture Notes in Comput. Sci., Vol. 2259, Springer, Berlin, 2001, pp. 112–128.

9. X. Lai, J. Massey, and S. Murphy, *Markov ciphers and differential cryptanalysis,* Advances in Cryptology—EUROCRYPT'91, in: Lecture Notes in Comput. Sci., Vol. 547, Springer, Berlin, 1991, pp. 17–38.

10. K. Nyberg, *Linear approximation of block ciphers,* Advances in Cryptology—EUROCRYPT'94, in: Lecture Notes in Comput. Sci., Vol. 950, Springer, Berlin, 1995, pp. 439–444.

11. S. Park, S.H. Sung, S. Chee, E-J. Yoon, and J. Lim, *On the security of Rijndael-like structures against differential and linear cryptanalysis,* Advances in Cryptology—ASIACRYPT 2002, in: Lecture Notes in Comput. Sci., Vol 2501, Springer, Berlin, 2002, pp. 176–191.

12. S. Park, S.H. Sung, S. Lee, and J. Lim, *Improving the upper bound on the maximum differential and the maximum linear hull probability for SPN structures and AES,* Fast Software Encryption—FSE 2003, in: Lecture Notes in Comput. Sci., Vol. 2887, Springer, Berlin, 2003, pp. 247–260.

13. F. Sano, K. Ohkuma, H. Shimizu, and S. Kawamura, *On the security of nested SPN cipher against the differential and linear cryptanalysis,* IEICE Trans. Fund. Elec., Commun. and Comput. Sci. E86-A (1) (2003) 37–46.

14. S. Vaudenay, *On the security of CS-Cipher,* Fast Software Encryption—FSE'99, in: Lecture Notes in Comput. Sci., Vol. 1636, Springer, Berlin, 1999, pp. 260–274.