

# Links Between Theoretical and Effective Differential Probabilities: Experiments on PRESENT

C. Blondeau and B. Gérard

INRIA project-team SECRET, France  
{celine.blondeau, benoit.gerard}@inria.fr

**Abstract.** Recent iterated ciphers have been designed to be resistant to differential cryptanalysis. This implies that cryptanalysts have to deal with differentials having so small probabilities that, for a fixed key, the whole codebook may not be sufficient to detect it. The question is then, do these theoretically computed small probabilities have any sense? We propose here a deep study of differential and differential trail probabilities supported by experimental results obtained on a reduced version of PRESENT.

**Keywords :** differential cryptanalysis, differential probability, iterated block cipher, PRESENT.

## 1 Introduction

Differential cryptanalysis has first been applied to the *Data Encryption Standard* (DES) in the early 90's by E. Biham and A. Shamir [BS91,BS92]. Since then, many ciphers have been cryptanalyzed using differential cryptanalysis or one of the large family of variants (truncated differential [Knu94], higher order differential [Knu94], impossible differential [BBS99], ...).

The basic differential cryptanalysis is based on a differential over  $r$  rounds of the cipher.

**Definition 1.** *A  $r$ -rounds differential*

*is a couple  $(\delta_0, \delta_r) \in \mathbb{F}_2^m \times \mathbb{F}_2^m$ . The probability of a differential  $(\delta_0, \delta_r)$  is*

$$p_* \stackrel{\text{def}}{=} \Pr_{\mathbf{X}, \mathbf{K}} [F_K^r(X) \oplus F_K^r(X \oplus \delta_0) = \delta_r],$$

*where  $m$  is the input/output size of the cipher and  $F$  the round function.*

Then,  $r$  rounds of the cipher can be distinguished from a random permutation using this differential. To break  $r + 1$  rounds of the block cipher, we look for differentials on  $r$  rounds. Then, for all the possible subkeys for the last round, the attacker does a partial decryption of the ciphertext pairs and count the number of time  $\delta_r$  appears. For a wrong candidate, the probability that  $\delta_r$  appears is around  $2^{-m}$  and for the correct subkey, this probability is around  $p_* + 2^{-m}$ . It is widely accepted that the number of pairs needed to distinguish those two probabilities is of order  $p_*^{-1}$  if the so called *signal-to-noise ratio* is large enough ( $S_N = \frac{p_*}{2^{-m}}$ ).

Recent ciphers, the *Advanced Encryption Standard* (AES) for instance, have been designed to be resistant to the basic differential cryptanalysis. Nevertheless, when a new cipher is proposed, cryptanalysts try to mount the best possible linear and differential attacks. In the case of PRESENT[BKL<sup>+</sup>07], the cipher we used for the experiments, the actual best published attack is the one of Wang [Wan08]. But actually, there is still lacks in the data complexity estimates of those differential attacks.

The first one is the use of Gaussian or Poisson distributions to estimate what actually is a binomial distribution. Since we are interested in differential cryptanalysis, Gaussian distribution is known to be worse than Poisson [Sel08] but such an approximation is used to estimate the success probability of most of the recent differential cryptanalyses. Nevertheless, Poisson distribution might not be tight if the differential probability is close to the uniform probability  $2^{-m}$ . Work has been done to give good estimates of the data complexity and the success probability of a statistical cryptanalysis for any setting [BGT10]. That is the reason why we chose to directly deal with binomial distributions without making any approximation.

The second point is the estimation of a differential probability. It is well known that a differential is composed of many trails and that the probability of one trail may not be a good estimate of the whole differential probability [NK92]. Again, in most of the recent differential cryptanalysis papers, the differential probability is estimated computing the probability of the main trail. We recall in Subsection 3.4 how to efficiently find many trails.

Last but not least, a widely used assumption is made in statistical cryptanalysis that is the assumption of *fixed key equivalence* or *stochastic equivalence* that is assuming that the probability of a differential that is computed over all the possible keys is roughly the same that the probability of a differential for some fixed key [LMM91].

### Contributions of this work.

A deep study of this hypothesis has been done in [DR05] and this paper aims at providing evidences to confirm this theory by the way of practical experiments on a toy version of PRESENT.

We first present the cipher we used for experiments Section 2. Then, in Section 3, we focus on differential trails that is sets of intermediate differences taken by a pair that matches a differential. The classical way of estimating a trail probability relies on some hypotheses that are not true. Nevertheless, experiments show that this theoretical probability makes sense as an average of the probability over the keys. In Section 4 we recall that a differential probability is the sum of the corresponding trail probabilities. Then, we present some experiments about the key dependency of this differential probability that corroborate the results in [DR05]. Finally, we conclude in Section 5 and sum-up the results as well as the problematics left as open questions.

## 2 PRESENT

Experiments are made on a lightweight cipher presented in 2007 at *CHES* conference: PRESENT [BKL<sup>+</sup>07]. This cipher is a *Substitution Permutation Network* and thus is easy to describe.

### 2.1 Reduced version of PRESENT: SMALLPRESENT-[s]

For the experiments to be meaningful, we need to be able to exhaustively compute the ciphertexts corresponding to all possible plaintexts for all possible keys. That is the reason why we chose to work on a reduced version of PRESENT named SMALLPRESENT-[s] [Lea10]. The family of SMALLPRESENT-[s] has been designed for such experiments. The value of  $s$  indicate the number of Sboxes of the cipher. These Sboxes are all the same which is defined on  $\mathbb{F}_2^4$ . This substitution is described in Table 1. The size of the message is then  $4s$ . In this paper, we make experiments on SMALLPRESENT-[4] that is the version with 4 Sboxes (the full version of PRESENT has 16 Sboxes). One round of SMALLPRESENT-[4] and PRESENT are respectively depicted in Figure 8, Figure 9 (Appendix B).

### 2.2 Different key schedules for SMALLPRESENT-[4]

The problem with the reduced cipher presented in [Lea10] is the key schedule. Actually, in the whole PRESENT, most of the bits of a subkey are directly used in the subkey of the next round. Since, for SMALLPRESENT-[s], the number of key bits is always 80 but the state size is only  $4s$ , this

is not true anymore for a small  $s$ . We decided to introduce two additional key schedules for our experiments.

1. *Same key*: The cipher has a master key that has the same size as the state and each subkey is equal to this master key. Therefore SMALLPRESENT-[4] is parameterized by a 16 bits master key.
2. *80-bits*: This key schedule is the one used in the full version of PRESENT and proposed in [Lea10].
3. *20-bits*: a homemade key schedule used with SMALLPRESENT-[4] similar to the one of the full version.

The master key is represented as  $K = k_{19}k_{18} \dots k_0$ . At round  $i$  the 16-bits round key  $K_i = k_{19}k_{18} \dots k_4$  consists in the 16 left-most bits of the current content of register  $K$ . After extracting the round key  $K_i$ , the key register is updated as follows:

- (a)  $[k_{19}k_{18} \dots k_1k_0] = [k_6k_5 \dots k_8k_7]$
- (b)  $[k_{19}k_{18}k_{17}k_{16}] = S[k_{19}k_{18}k_{17}k_{16}]$
- (c)  $[k_7k_6k_5k_4k_3] = [k_7k_6k_5k_4k_3] \oplus \text{roundcounter}$

The key is rotated by 13 bit positions to the left, the left most four bits are passed through the PRESENT Sbox, and the *roundcounter* value is exclusive-ored with bits  $k_8k_7k_6k_5k_4$ . We keep the 5-bits counter version. But we only study less than 7 rounds of SMALLPRESENT-[4] so the counter can be represented in 3 bits.

### 3 Differential trail probability

#### 3.1 Notation

Let us denote by  $K$  the master key. The round subkeys derived from  $K$  are denoted by  $K_1, K_2, \dots, K_r$ . Let  $F_{K_i} : \mathbb{F}_2^m \mapsto \mathbb{F}_2^m$  be a round function of a block cipher. We will denote by  $F_K^r$  the application of  $r$  rounds of the block cipher.

$$F_K^r = F_{K_r} \circ F_{K_{r-1}} \circ \dots \circ F_{K_1}.$$

Generally, there is not one but many ways to go from the input difference to the output difference of a differential. Since the term if *differential characteristic* seems to be ambiguous, we use the linear cryptanalysis notation and call such a way a *differential trail*.

**Definition 2.** *Differential trail*

A differential trail of a cipher is a  $(r+1)$ -tuple  $(\beta_0, \beta_1, \dots, \beta_r) \in (\mathbb{F}_2^m)^{r+1}$  of intermediate differences at each round. The probability of a differential trail  $\beta = (\beta_0, \beta_1, \dots, \beta_r)$  is

$$\Pr_{\mathbf{X}, \mathbf{K}} [\forall i \ F_K^i(X) \oplus F_K^i(X \oplus \beta_0) = \beta_i].$$

Computing the exact value of a trail probability is not possible for real ciphers since it needs to encipher the whole codebook for all possible keys. The classical way of estimating a trail probability is to chain trails on 1 round. This approach is based on a formalism introduced by Lai, Massey and Murphy [LMM91].

*Markov cipher*

A Markov cipher is a cipher for which, the probability

$$\Pr_{\mathbf{X}, \mathbf{K}} [F_K^r(X) \oplus F_K^r(X \oplus \delta_0) = \delta_r | X = x]$$

does not depend on  $x$  if the subkeys  $\mathbf{K}_i$  are uniformly distributed.

In the case of Markov ciphers where the subkeys are xored to the state, the theoretical probability of a trail  $\beta = (\beta_0, \beta_1, \dots, \beta_r)$  is computed as follow

$$p_\beta^t = \prod_{i=1}^r \Pr_{\mathbf{X}} [F(X) \oplus F(X \oplus \beta_{i-1}) = \beta_i].$$

Notice that we did not use the notation  $F_{K_i}$  because when the subkeys are xored to the state, the probability  $\Pr_{\mathbf{X}} [F(X) \oplus F(X \oplus \beta_{i-1}) = \beta_i]$  does not depend on the value of the subkey.

### 3.2 Key dependency of a trail

The probability of a differential trail can be influenced by the choice of the master key used to encipher samples. This remark is the main motivation of the work in [DR05]. In order to take into account this fact, let us introduce some notation. For a  $r$ -round differential trail  $\beta = (\beta_0, \beta_1, \dots, \beta_r)$ , let us define

$$\begin{aligned} T_K &\stackrel{\text{def}}{=} \frac{1}{2} \#\{X \in \mathbb{F}_2^m | F_K^i(X) \oplus F_K^i(X + \beta_0) = \beta_i \quad \forall 1 \leq i \leq r\}, \\ T[j] &\stackrel{\text{def}}{=} \#\{K | T_K = j\}. \end{aligned} \tag{1}$$

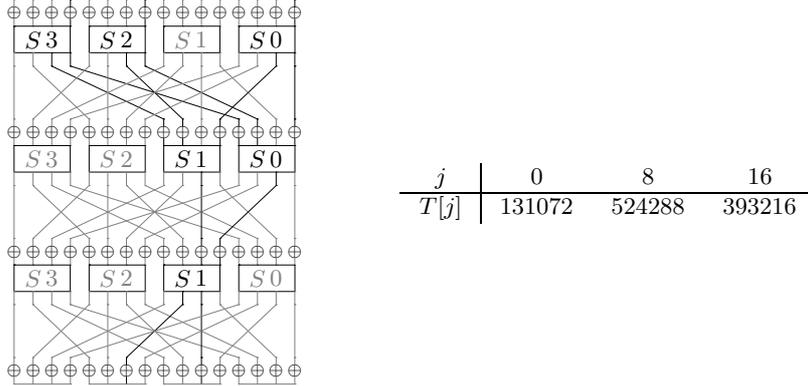
Let  $n_k$  be the number of bits of the master key. The *real* or *effective value* of the trail probability is

$$p_\beta = 2^{-m-1} \sum_{K \in \mathbb{F}_2^{n_k}} T_K = 2^{-m-1-n_k} \sum_j T[j] \cdot j.$$

To motivate these new notation, we give an example where the key dependency is obvious.

**Example of a trails with experimental probability not equal to the theoretical one**

We illustrate this phenomena by a differential trail over 3 rounds on SMALLPRESENT - [4]:  $\beta = (0x1101, 0xdd, 0x30, 0x220)$  (see Figure 1).



**Fig. 1.** Trail  $\beta = (0x1101, 0xdd, 0x30, 0x220)$  and the corresponding  $T[j]$ 's

First, we are going to compute the theoretical probability of this trail. We suppose that SMALLPRESENT-[4] is a Markov cipher and that uses independent subkeys.

- *Round 1* 3 S-boxes with input difference  $0x1$  and output difference  $0x3$ .
- *Round 2* 2 S-boxes with input difference  $0xd$  and output difference  $0x2$ .
- *Round 3* 1 S-box with input difference  $0x3$  and output difference  $0x6$ .

We have 6-Sboxes with transition probability  $2^{-2}$  therefore  $p_{\beta}^t = 2^{-12}$ . This means that the number of plaintext such that  $(X, X \oplus 0x1101)$  follows this trail for a fixed key should be  $2^{16-1} \cdot 2^{-12} = 2^3$ . We made experiments to check this assumption. For a fixed key we computed the number of plaintexts that follows this trail. In Figure 1 are given the values taken by  $T[j]$  for all keys in  $\mathbb{F}_2^{20}$  using the 20-bit key schedule. We can see that there are three kinds of key leading to three different values of  $T[j]$ .

Experiments on this trail show that for a fixed key, the theoretical probability of a differential trail do not always match with the value of

this trail probability for a fixed key. Experiments also show that for some keys the trail can be impossible. This can be of real significance because such phenomenon is also existing on 3 rounds of PRESENT. That means that, maybe, some differential trails used in a differential cryptanalysis may not have the expected probability for most of the keys.

Averaging the probabilities over the keys, we see that the effective probability of this differential trail is  $2^{-11.6}$  (the theoretical one is  $2^{-12}$ ). This difference between real and theoretical probabilities may weaken (or strengthen) some attacks. Does a lot of trails have such a difference between their theoretical value and the average probability? In the next subsection we will show that, most of the times, the theoretical value of a differential trail is close to the effective one (averaged over the keys).

### 3.3 Theoretical probability and average probability of a trail over the keys

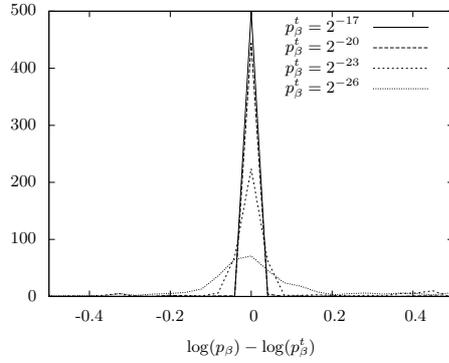
We observed that the theoretical probability is likely to be the average of the trail probabilities over all the possible keys. We made experiments on SMALLPRESENT-[4] with different key schedules. Let us recall that the theoretical probability of the differential trail  $\beta$  is denoted by  $p_\beta^t$  and the effective one (averaged over the keys) is denoted by  $p_\beta$ . In Figure 4, Figure 3 and Figure 2, we have computed the difference between  $\log(p_\beta^t)$  and  $\log(p_\beta)$  for 500 random trails.

- In *Figure 4* we assume that the round subkeys are derived from the 20-bits key schedule. We average the probabilities over the whole set of  $2^{20}$  keys to obtain the value  $p_\beta$ .
- In *Figure 3* we assume that all the round subkeys are the same. We average the probabilities over the whole set of  $2^{16}$  keys to obtain the value  $p_\beta$ .
- In *Figure 2* we assume that the round subkeys are derived from the 80-bits key schedule. Since we cannot average probabilities over the  $2^{80}$  possible master keys, the computed value is obtained averaging over  $2^{20}$  keys.

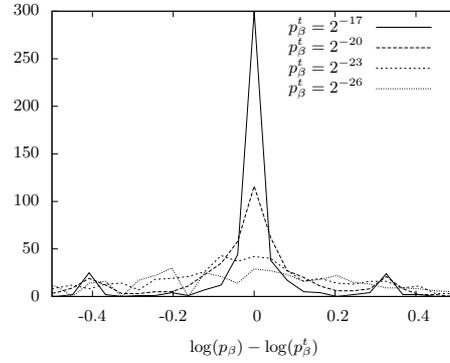
**Remark:**

We can see that the phenomenon is not the same depending on the key schedule. Indeed, in Figure 3 when the same key is taken over the 5 rounds the dependency of the round key is more important than in Figure 2 where the 80-bits key schedule is used implying that all the key

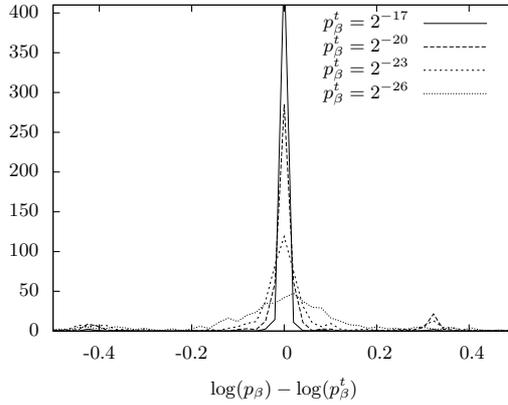
**Number of trails as a function of  $\log(\mathbf{p}_\beta) - \log(\mathbf{p}_\beta^t)$  for a sample of 500 trails on 5 rounds using different key schedules**



**Fig. 2.** 80-bits key schedule.



**Fig. 3.** same subkey for all rounds.



**Fig. 4.** 20-bits key schedule.

bits are used only once (on average). This remark has motivated the 20-bit key schedule we use in the following experiments that seems to be the most appropriate.

Experiments show that the average proportion of pairs satisfying a differential trail is close to the theoretical probability. We can observe that this behavior is getting worse as the probability is decreasing. Nevertheless, it seems to be some symmetry what leads to the idea that taking enough trails into account will correct this and give better results.

### 3.4 Automatic search of differential trails

In order to find the best trails, we use a Branch and Bound algorithm (the one used in linear cryptanalysis). This one is explained in Appendix A.

## 4 Differential probability

### 4.1 Differential probability and trail probabilities

The first thing to say here is that the probability of a differential is the sum of the probabilities of the corresponding differential trails.

**Lemma 1.** *Let  $(\delta_0, \delta_r)$  be a  $r$ -round differential. Then the probability  $p_*$  of this differential is*

$$p_* = \sum_{\beta=(\delta_0, \beta_1, \dots, \beta_{r-1}, \delta_r)} p_\beta.$$

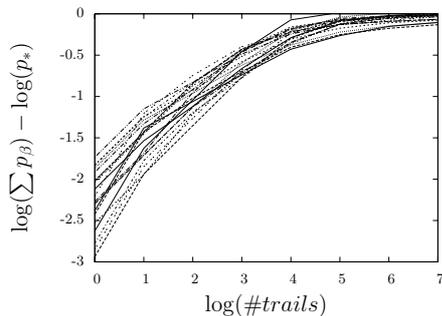
*Proof.* A pair that matches a trail cannot match any other (they are disjoint events) and thus  $\Pr[\cup_i A_i] = \sum_i \Pr[A_i]$ .

For a large number of rounds, it is impossible to compute the probability of a differential  $(\delta_0, \delta_r)$  because there is too much differential trails that go from  $\delta_0$  to  $\delta_r$ . Actually, in differential cryptanalysis, one uses a lower bound on the probability of the differential  $(\delta_0, \delta_r)$  by considering the sum of the likeliest trail probabilities.

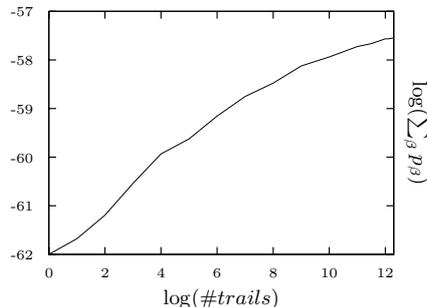
In Section 3, we saw that the effective trail probability may not match with the theoretical one. Nevertheless, it seems to be some symmetry what leads to the idea that the sum of theoretical trail probabilities may give a good estimate of a effective differential probability.

We made some experiments on 5 rounds of SMALLPRESENT-[4] with the 20 bits key schedule to see how many trails are required to get a good estimate of a differential probability. We computed the sum of the theoretical probabilities of many trails corresponding to the same differential. Since the cipher is small we also have computed the effective value of the differential by averaging over all plaintexts and all keys. In Figure 5 we have plotted the difference between both values for 20 differentials. We can see that taking many trails give a better estimation of the differential probability.

Looking at the results in Figure 5 we can wonder whether it is possible to determine the number of trails to consider for estimating a differential probability. In this example we see that taking  $2^7$  trails seems to be sufficient but when we look at the whole cipher it is obviously not enough (see the following paragraph).



**Fig. 5.** Convergence of the sum of trails probabilities to the real value of the differential probability.



**Fig. 6.** Probability of the differential presented in [Wan08] as a function of the logarithm of the number of trails.

### Remark on Wang’s paper [Wan08]

In [Wan08], the target is the whole PRESENT (64 bits). One of the differentials used is

$$(d_0, d_{14}) = (0x0700000000000700, 0x0000000900000009)$$

on 14 rounds obtained by iterating 3 times a differential trail on 4 rounds and by adding one more round at the beginning and at the end.

- This trail on 4 rounds is not one of the best since it has a probability equal to  $2^{-18}$  and we found a lot of differential trails with probability  $2^{-12}$ . Nevertheless, it is the best iterative differential on 4 rounds.
- There exists lots of differential trails on 14 rounds with probability  $2^{-62}$  and using algorithm given in Subsection 3.4,  $2^{-62}$  seems to be the best trail probability over 14 rounds.
- We have theoretically computed all differential trails with input difference  $d_0$ , output difference  $d_{14}$  and probability greater than  $2^{-73}$ . Summing the probabilities of the  $2^{12}$  best trails, we observe that the probability of the differential  $(d_0, d_{14})$  is greater than  $2^{-57.53}$  and that it does not seem to converge yet (see Figure 6).

### 4.2 Key dependency of a differential probability

We now consider a differential  $(\delta_0, \delta_r)$  that is to be used in a differential cryptanalysis. The attacker will get some samples enciphered with a fixed master key. Depending on this key, the real probability of the differential will be smaller/equal/larger than the theoretically computed value.

For a fixed key  $K$ , let us denote by  $D_K$  the number of pairs of plaintexts with input difference  $\delta_0$  that lead to an output difference  $\delta_r$ . Since we do not want to count a pair twice, we introduce a  $\frac{1}{2}$  coefficient.

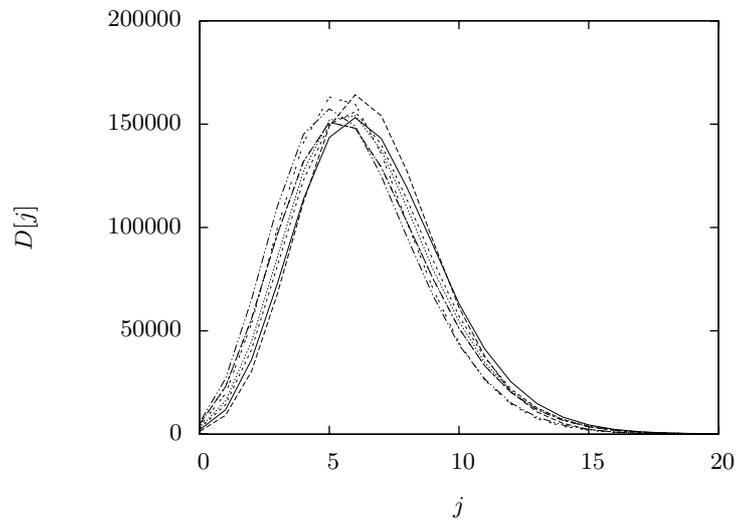
$$D_K \stackrel{\text{def}}{=} \frac{1}{2} \#\{X | F_K^r(X) + F_K^r(X + \delta_0) = \delta_r\}.$$

We are going to study the distribution of  $D_K$ 's.

$$D[j] \stackrel{\text{def}}{=} \#\{K | D_K = j\}.$$

It is proven in [DR05] that  $D_K$  follows a hypergeometric distribution that, in cryptography setting, is tightly approximated by a binomial distribution of parameters  $(2^{m-1}, p_*)$ .

We made some experiments on 5 rounds of SMALLPRESENT-[4] to check this. Using the 20-bits key schedule we computed the repartition of the  $D_K$ 's. In Figure 7, we see that the  $D_K$ 's seems to follow a binomial distribution.



**Fig. 7.** Distribution of  $D[j]$ 's for 8 differentials over 5 rounds of SMALLPRESENT-[4].

This observation should be taken into account when computing the success probability of an attack. Let us denote by  $p_*^t$  the theoretical probability of the differential used in a cryptanalysis. We recall that  $n_k$  is the number of key bits. For  $K \in \mathbb{F}_2^{n_k}$ , the effective probability of the differential is  $\frac{D_K}{2^{m-1}}$  where  $D_K$  is a random variable that follows a binomial distribution of parameters  $(2^{m-1}, p_*^t)$ . If we denote by  $P_S(p_*)$  the success probability of a differential cryptanalysis using a differential with **effective** probability  $p_*$  (see [Sel08,BGT10]). Then the success probability of a differential cryptanalysis using a differential with **theoretical** probability  $p_*^t$  is

$$P_{\text{success}} = \sum_{i=0}^{2^{m-1}} P_S\left(\frac{i}{2^{m-1}}\right) \cdot \left[ (p_*^t)^i (1 - p_*^t)^{2^{m-1}-i} \binom{2^{m-1}}{i} \right].$$

## 5 Conclusion

We have presented lots of experiments on differential cryptanalysis. The main teaching of this work is that claimed complexities of differential cryptanalyses on recent ciphers may be under/over-estimated.

The first point is the fact that estimating a differential probability with the probability of its main trail is really not suitable. To illustrate the first point, we estimated the probability of a differential used in [Wan08] to  $2^{-57.53}$  while the author only takes into account the best trail and provides an estimate of  $2^{-62}$ .

The second point is the key dependency of a differential probability. Experiments confirmed the theory exposed in [DR05] and thus we propose a formula for the success probability that takes this phenomenon into account.

This work give some elements for understanding differential cryptanalysis but it still remains some open questions. The two main problematics that seems to be of great interest are the following.

- The theoretical probability of a trail seems to be less meaningful as this probability decreases. How far does this theoretical value make sense?
- How can we get a good estimate of a differential probability without finding all the corresponding differential trails?

## References

- [BBS99] E. Biham, A. Biryukov, and A. Shamir. Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. In *EUROCRYPT '99*, volume 1592 of *LNCS*, pages 12–23, 1999.
- [BGT10] C. Blondeau, B. Gérard, and J.-P. Tillich. Accurate Estimates of the Data Complexity and Success Probability for Various Cryptanalyses. *DCC special issue on Coding and Cryptography*, 2010. To appear.
- [BKL<sup>+</sup>07] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw, Y. Seurin, and C. Vikkelsoe. PRESENT: An Ultra-Lightweight Block Cipher. In *CHES '07*, volume 4727 of *LNCS*, pages 450–466. SV, 2007.
- [BS91] E. Biham and A. Shamir. Differential Cryptanalysis of DES-like Cryptosystems. *Journal of Cryptology*, 4(1):3–72, 1991.
- [BS92] E. Biham and A. Shamir. Differential Cryptanalysis of the Full 16-round DES. In *CRYPTO'92*, volume 740 of *LNCS*, pages 487–496. Springer–Verlag, 1992.
- [DR05] J. Daemen and V. Rijmen. Probability distributions of Correlation and Differentials in Block Ciphers. Cryptology ePrint Archive, Report 2005/212, 2005. <http://eprint.iacr.org/>.
- [Knu94] L. R. Knudsen. Truncated and Higher Order Differentials. In *FSE '94*, volume 1008 of *LNCS*, pages 196–211. Springer–Verlag, 1994.
- [Lea10] G. Leander. Small Scale Variants Of The Block Cipher PRESENT. Cryptology ePrint Archive, Report 2010/143, 2010. <http://eprint.iacr.org/>.
- [LMM91] X. Lai, J. L. Massey, and S. Murphy. Markov Ciphers and Differential Cryptanalysis. In *EUROCRYPT '91*, volume 547, pages 17–38, 1991.
- [NK92] K. Nyberg and L.R. Knudsen. Provable Security Against Differential Cryptanalysis. In *CRYPTO'92*, volume 740 of *LNCS*, pages 566–574. Springer–Verlag, 1992.
- [Sel08] A. A. Selçuk. On Probability of Success in Linear and Differential Cryptanalysis. *Journal of Cryptology*, 21(1):131–147, 2008.
- [Wan08] M. Wang. Differential Cryptanalysis of Reduced-Round PRESENT. In *AFRICACRYPT '08*, volume 5023 of *LNCS*, pages 40–49. SV, 2008.

## A Algorithm for finding differential trails

Let  $\mathbf{B}$  be a lower bound on the probability of the trails we are looking for. We suppose that we are interested in differential trails over  $r$  rounds and that we already know the best trail probabilities for a smaller number of rounds.

We are going to traverse the tree defined as follow.

- Each node contains a difference.
- The root contains the input difference.
- The sons of a node correspond to all differences that are reachable after one round of the cipher from the input difference contained in the node.

- An edge has a weight that corresponds to the probability of transition from the father’s difference to the son’s one.

Then, we do a depth-first traversal of this tree and only consider leaves of the tree. The path from the root to the leaf is a differential trail. The probability of this trail is computed multiplying the weights of the path edges.

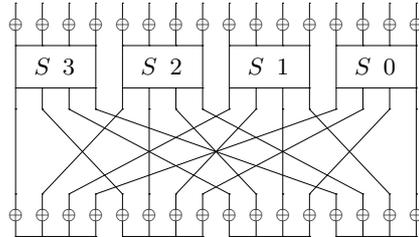
There is a simple criterion to avoid some useless branches. When going from a father to a son, we compute the path probability from the root to the son and multiply it to the best trail probability for the remaining rounds (that is the depth between the son and the leaves). If it is smaller than **B**, then no leaves under the son will leads to a trail with probability greater than **B**. Then we look at another son and so on... Notice that this criterion can be quickly checked because the probability of a trail is computed as one advances through the tree and thus when looking at a node, the probability of the path from the root to that node is already known (the cost is one multiplication when going one step deeper and one division when returning to the father node).

## B Characteristics of PRESENT

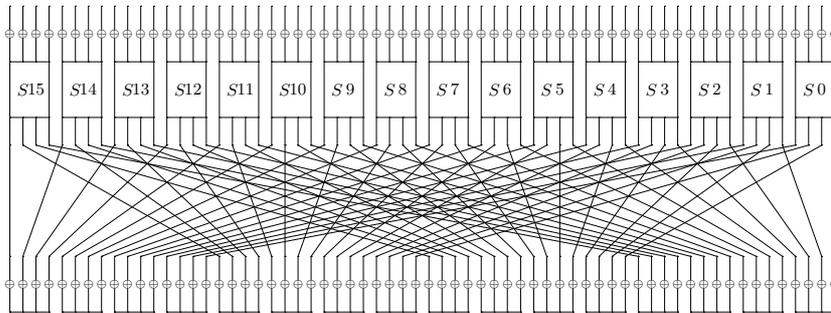
Here are the Sbox and the round function of both PRESENT and SMALL-PRESENT - [4].

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
S(x)	C	5	6	B	9	0	A	D	3	E	F	8	4	7	1	2

**Table 1.** The S-box of PRESENT



**Fig. 8.** 1 round of SMALLPRESENT-[4]



**Fig. 9.** 1 round of PRESENT