

Experiments on the Multiple Linear Cryptanalysis of Reduced Round Serpent

B. Collard , F.-X. Standaert , J.-J. Quisquater

UCL Crypto Group
Microelectronics Laboratory
Catholic University of Louvain - UCL
Belgium

FSE 2008



Happy Birthday, Nathalie!

Outline

Various experimental attacks against reduced-round Serpent are presented.

We used the framework proposed by Biryukov *et al.* at crypto 2004 [2]

The purposes are the following :

- To confirm the relevance of their theoretical approach
- To show the practical improvements of multiple approximations
- To observe the consequences of linear dependancies in the approximations
- To compare the specificities of Matsui's *Algorithm 1* and *2*

Table of content

- 1 Linear cryptanalysis
- 2 Preliminary remarks
- 3 Experimental attacks with Algorithm 1
- 4 Experimental attacks with Algorithm 2
- 5 Conclusion and further work

1. Linear Cryptanalysis

- Initially proposed by Matsui [8] in 1993
- Exploits bias in the occurrence probability of a linear approximation
- Such expressions are obtained by linear approximations of the non-linear elements of the cipher

Linear Approximation

$$P[\chi_P] \oplus C[\chi_C] = K[\chi_K] \quad (1)$$

- P , C and K denote the plaintext, ciphertext and the secret key
- $A[\chi]$ stands for $A_{a_1} \oplus A_{a_2} \oplus \dots \oplus A_{a_n}$
- χ is usually denoted as a mask
- For a 'good' approximation, the equation holds with a probability significantly different than $1/2$

Given a r -round approximation $P[\chi_P] \oplus C[\chi_C] = K[\chi_K]$ with bias ϵ

Algorithm 1

Algorithm 1 attacks r -round cipher by simply evaluating $P[\chi_P] \oplus C[\chi_C]$ for a sufficiently large number of plaintext-ciphertext. The parity of $K[\chi_K]$ can then be guessed thanks to the probability of the left parity. This attack recovers one bit of key parity.

Algorithm 2

Algorithm 2 targets $(r+1)$ -rounds cipher by partially decrypting the last round with a key guess and then evaluates the experimental bias for each guess. Several bits can be recovered at the same time.

In both cases, the data complexity is proportional to $1/\epsilon^2$

Multiple linear cryptanalysis

- Improves cryptanalysis by using multiple approximations
- Introduced by Kalisky and Robshaw [5] in 1994
- Improved by Biryukov *et al.* [2] in 2004
- Defines capacity as $\bar{c}^2 = 4 \cdot \sum_{i=1}^n \epsilon_i^2$
⇒ Decreases the data complexity to $O(1/\bar{c}^2)$

Theoretical framework

Given m approximations on r rounds :

$$P[\chi_P^i] \oplus C[\chi_C^i] = K[\chi_K^i] \quad (1 \leq i \leq m), \quad (2)$$

We want to determine the value of the vector of parity :

$$\mathbf{Z} = (z_1, z_2, \dots, z_m) = (K[\chi_K^1], K[\chi_K^2], \dots, K[\chi_K^m]) \quad (3)$$

- Define a counter T_i for approximation i
- T_i is incremented when the approximation is verified for a P-C pair
- The experimental biases ϵ_i^* are evaluated as $(T_i - N/2)/N$
- A sorted list of the vector parity candidates is built according to the distance between theoretical and experimental biases
- The remaining unknown bits are guessed by exhaustive search.

Definition (Gain)

if an attack is used to recover an n -bit key and is expected to return the correct key after having checked M candidates in average , then the gain of the attack, expressed in bits, is defined as :

$$\gamma = -\log_2 \frac{2 \cdot M - 1}{2^n} \quad (4)$$

Intuitively, the gain is a measure of the remaining key candidates to test after a cryptanalysis has been performed. This gain is determined by the position of the correct vector of parity in the weighted list of candidates obtained during the analysis phase.

2. Preliminary remarks

Serpent

- AES candidate - rated second behind Rijndael
- Designed by Anderson, Biham and Knudsen [1]
- Conservative design

Architecture

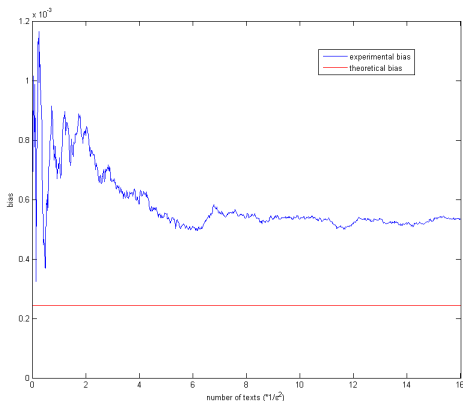
- Substitution-Permutation Network (SPN)
- Composed of 32 rounds
- For each round :
 - A subkey addition
 - A passage through S-boxes
 - A linear transformation

Best known attack

Linear-differential cryptanalysis on 11 rounds (Biham *et al.* [12]).

Evolution of the experimental biases according to the data complexity :

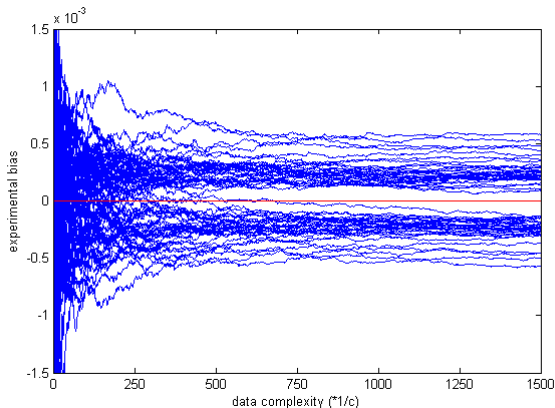
- We used a 4-round linear approximation with a bias of 2^{-12}
- We evaluated the experimental bias with up to $16 * 2^{24}$ texts



- The bias becomes stable after about $8/\epsilon^2$ texts.
- The underestimated theoretical bias suggests that the linear hull effect [4] is not negligible

Evolution of the experimental bias according to the data complexity :

- We used 64 4-round linear approximations with various biases
- We evaluated the experimental biases for up to $1500 * 2^{24}$ texts



- Approximations separate into 2 according to the sign of their bias
- Each approximation provides some information about the key

3. Experimental attacks with Algorithm 1

Linear approximation search

- Generation of the approximation is computationally demanding
- A branch-and-bound algorithm was proposed by Matsui [10]
- We used a modified heuristic [3]

Selection of the approximations

With *Algorithm 1*, an adversary recovers linear combination of subkey bits

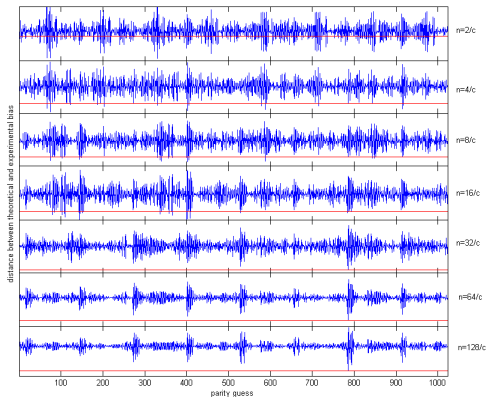
This drawback can be partially relaxed using multiple approximations :

- The best linear approximation found is selected
- Then only the input/output masks of the linear trail are modified
- Finally, by carefully choosing the linear dependancies, the adversary ends up with an exploitable information on the cipher key.

As the linear trail is the same for all the approximations except in the input/output, the adversary can easily recover first/last subkey bits.

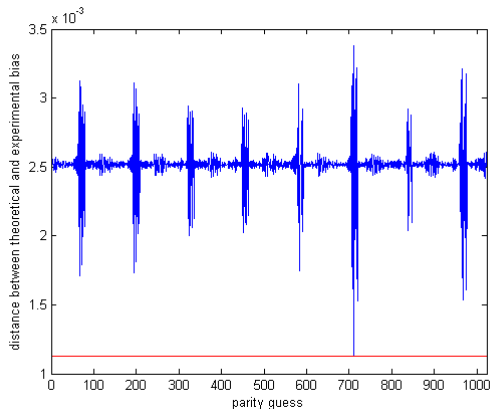
Evolution of the distance between theoretical and experimental biases :

- We used 64 4-round linear approximations with various biases
- Between $2/c^2$ and $128/c^2$ texts were used



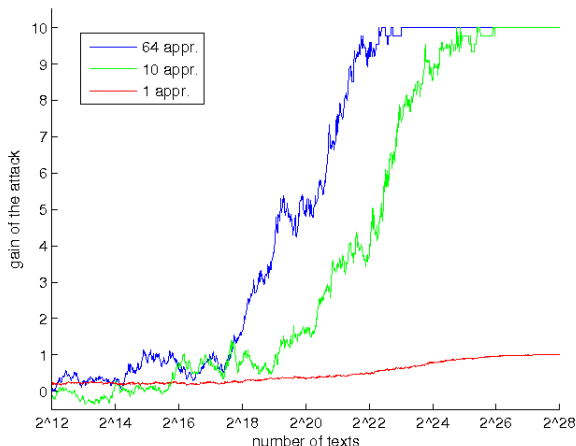
- Attack results improve with the number of texts
- A regular structure underlines the impact of the Hamming distance.

Same experiment using $4096/c^2$ texts :



- 10 parity bits $K[\chi'_k]$ have to be guessed
- The regular structure is even more remarkable

Gain of three attacks with respectively 1, 10 and 64 approximations :



- Only 10 linearly independent approximations
- Gain with 64 approx. increases $\simeq 8$ times faster than with 10 approx.
- The graph shows no influence of the linear dependencies

Definition (success rate)

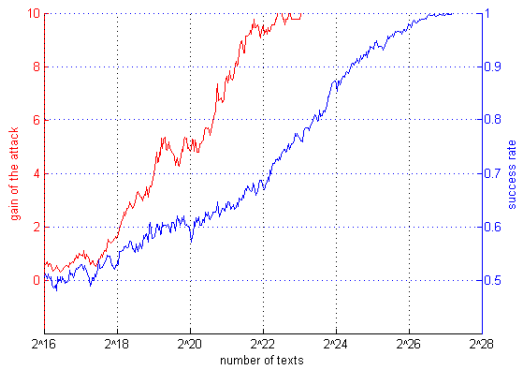
The success rate of an attack using n approximations is the percentage of parity bits guessed correctly among the n parities when they are chosen so as to minimize the distance between experimental and theoretical biases.

Rationale

- Unlike the gain, it doesn't take the linear dependencies into account
- Comparison allows to determine the advantage of multiple approximations.

Error Correcting code effect :

- Using 64 approximations, only 10 linearly independent



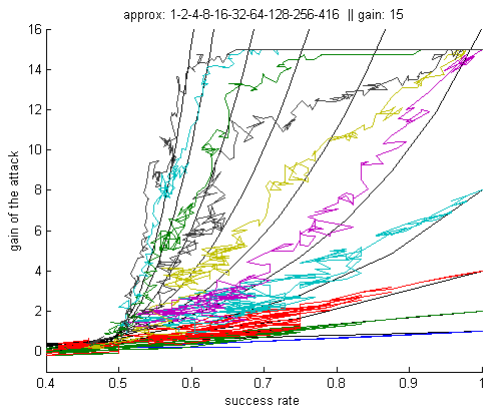
- The gain increases much faster than the success rate
 - Consequence of linear dependencies in the approximations
 - The correct vector of parity must respect these dependencies
 - This gives an efficient way to check a parity candidate
- Some parity candidates can be rejected a-priori.

Suppose n' out of the n approximations are guessed correctly :

- The success rate is n'/n .
- The gain is evaluated according to the position of the correct parity vector in the list of parity candidates :
 - Choose the first candidate so as to minimize the euclidian distance between theoretical and experimental biases.
 - Assume one guess is incorrect ; choose one parity bit and take its complement ; try the $\binom{n}{1}$ possible candidates ;
 - Assume two guesses are incorrect ; choose two parity bits and take their complements ; try the $\binom{n}{2}$ possible candidates ;
 - ...
 - Assume $n - n'$ guesses are incorrect ; choose $n - n'$ parity bits and take their complements ; try the $\binom{n}{n-n'}$ possible candidates ;
 - After $n - n'$ steps, we have necessarily found the correct candidate
 - Thus the gain of the attack equals :

$$\gamma = -\log_2\left(\frac{\sum_{i=0}^{n-n'} \binom{n}{i}}{2^n}\right) \quad (5)$$

Gain vs. Success rate (up to 416 approx. and 15 independent one) :



- Predictions (in black) assume independence of the approximations
- Observations fit well as long as the gains do not saturate
- For a given success rate, the gain increases with the number of approximations

4. Experimental attacks with Algorithm 2

Difference between Algorithm 1 and Algorithm 2

- With *Algorithm 1*, parity guesses are chosen so as to minimize :

$$\min_g \sum_{i=1}^m (\epsilon_i - (-1)^{g(i)} \cdot \epsilon_i^*)^2, \quad (6)$$

Algorithm 1 works even if the theoretical biases are underestimated.

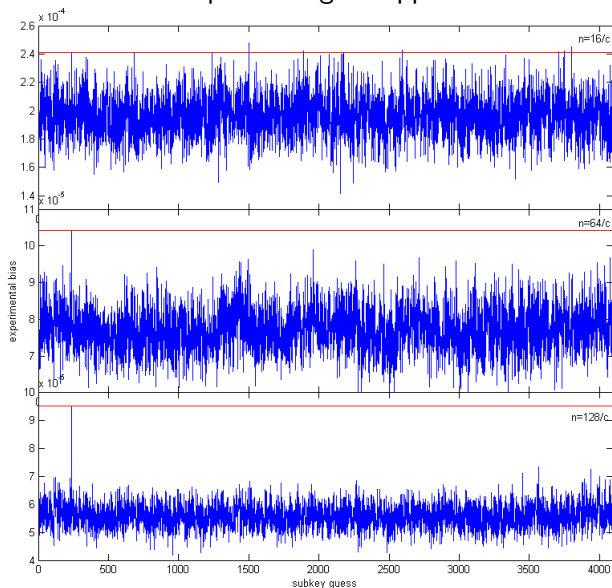
- With *Algorithm 2*, subkey and parity guesses are chosen to minimize :

$$\min_k \left(\min_g \sum_{i=1}^m (\epsilon_i - (-1)^{g(i)} \cdot \epsilon_{i,k}^*)^2 \right) \quad (7)$$

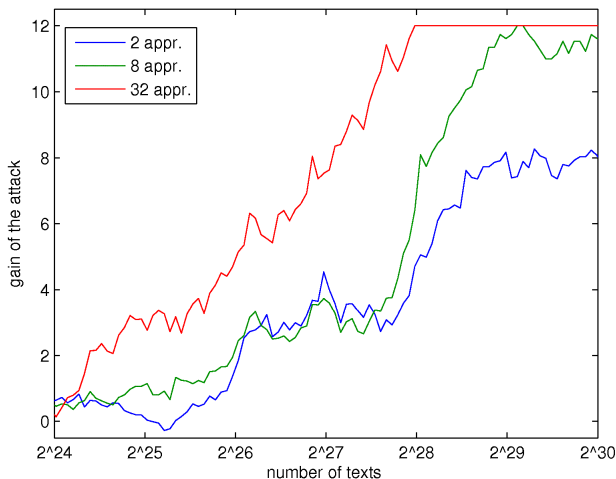
Algorithm 2 requires good theoretical estimations of experimental biases

- The framework of Biryukov cannot be directly applied in this context
- We look for the guess with the highest experimental bias instead

Attacks against 5-round Serpent using 32 approximations :



Gain of the attack :



- Multiple approximations allows to increase the gain of an attack
- Increasing the number of approximations does not involve reductions of the data complexity according to the capacity as for *Algorithm 1*.







5. Conclusion and further work







- We presented experimental results of multiple linear cryptanalysis against 4- and 5-round Serpent.
- In practice, our experiments confirmed the significant improvement of multiple linear cryptanalysis attacks compared to Matsui's original attack.
- As expected, Attack showed no influence of linear dependencies on the gain

By contrast with experiments against the DES, we observed a significant linear hull effect, with the following consequences :

- Optimal attacks using Matsui's *Algorithm 1* closely followed the data complexities predicted with the capacity value, even if the theoretical biases of the approximations were underestimated.
- Optimal attacks using Matsui's *Algorithm 2* did not lead to successful key recoveries because of the lack of good theoretical estimations of the biases. Modified heuristics allowed us to take advantage of multiple approximations. But the improvement is not following the predictions of the capacity values anymore.

Thanks for your attention !

-  R. Anderson, E. Biham, L. Knudsen, *Serpent : A Proposal for the Advanced Encryption Standard*, in the proceedings of the First Advanced Encryption Standard (AES) Conference, Ventura, CA, 1998.
-  A. Biryukov, C. De Cannière, M. Quisquater, *On Multiple Linear Approximations*, in the proceedings of CRYPTO 2004, Lecture Notes in Computer Science, vol. 3152, pp.1-22, Santa Barbara, California, USA, August 2004.
-  B. Collard, F.-X. Standaert, J.-J. Quisquater, *Improved and Multiple Linear Cryptanalysis of Reduced Round Serpent*, in the proceedings of InsCrypt 2007, LNCS, pp. 47-61, Xining, China, September 2007.
-  P. Junod, *On the Complexity of Matsui's Attack*, in the proceedings of SAC 2001, LNCS, vol. 2259, pp. 199-211, Toronto, Ontario, Canada, August 2001.
-  B.S. Kaliski, M.J.B. Robshaw, *Linear Cryptanalysis using Multiple Approximations*, in the proceedings of CRYPTO 1994, Lecture Notes in Computer Sciences, vol. 839, pp. 26-39, Santa Barbara, California, USA, August 1994.
-  L.R. Knudsen, *Practically Secure Feistel Ciphers*, in the proceedings of FSE 1993, LNCS, vol. 809, pp. 211-221, Cambridge, UK, December 1993.

-  S. Kumar, C. Paar, J. Pelzl, G. Pfeiffer, M. Schimmler, *Breaking Ciphers with COPACOBANA - A Cost-Optimized Parallel Code Breaker*, in the proceedings of Cryptographic Hardware and Embedded Systems - CHES 2006, Lecture Notes in Computer Science, vol. 4249, Springer, 2006.
-  M. Matsui, *Linear cryptanalysis method for DES cipher*, in the proceedings of Eurocrypt 1993, LNCS, vol. 765, pp. 386–397, Lofthus, Norway, May 1993.
-  K. Nyberg, *Linear Approximations of Block Ciphers*, in the proceedings of Eurocrypt 1994, LNCS, vol. 950, pp. 439-444, Perugia, Italy, May 1994.
-  M. Matsui, *On Correlation Between the Order of S-boxes and the Strength of DES*, in the proceedings of Eurocrypt 1994, Lecture Notes in Computer Science, vol. 950, pp. 366-375, Perugia, Italy, May 1994.
-  S. Murphy, *The Independence of Linear Approximations in Symmetric Cryptology*, IEEE Transactions on Information Theory, Vol. 52, pp. 5510-5518, 2006.
-  E. Biham, O. Dunkelman, N. Keller, *Differential-linear Cryptanalysis of Serpent*, in the Proceedings of Fast Software Encryption 2003, Lecture Notes in Computer Science, vol. 2887, pp. 9-21, Springer, 2004.