

# On the complexity of Matsui's attack against DES

Pascal Junod, [pascal.junod@epfl.ch](mailto:pascal.junod@epfl.ch)



Swiss Institute of Technology, Lausanne

*SAC'01, Toronto - Canada*

## Outline

Matsui's linear cryptanalysis against 16-rounds DES, as proposed at Crypto'94.

- Historical Overview
- Experimental Results
- Theoretical Analysis
- Conclusion

## Linear Cryptanalysis Performances: Historical Overview

- [Matsui, Eurocrypt'93, Crypto'94] Linear cryptanalysis, first experimental implementation
- [Blöcher-Dichtl, FSE'94] Some observations on the application of the piling-up lemma
- [Nyberg, Eurocrypt'94] Linear hull concept
- [Harpes-Kramer-Massey, Eurocrypt'95] Generalization of linear cryptanalysis

## Linear Cryptanalysis Performances: Historical Overview

- [Vaudenay, 1995] Statistical cryptanalysis concept
- [Kukorelly, 1999] Theoretical study on the piling-up lemma application
- [Selçuk, Indocrypt'00] Bias estimation in linear cryptanalysis

## Experiment Description

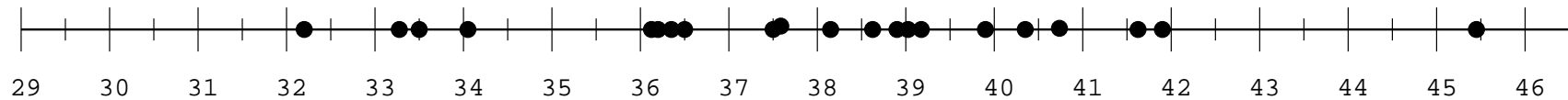
- Matsui attack has been implemented using today's technology
- Fast DES routine (bitsliced implementation on the Intel MMX architecture)
- Idle time of 12 - 18 CPUs
- 3-7 days to produce and analyze  $2^{43}$  known pairs
- The experiment has run 21 times

## Experimental Results (1)

- Widely accepted attack complexity: *Given  $2^{43}$  known pairs, it is possible to recover the key with a success probability of 85 % within  $C_{(0.85)}^{est} = 2^{43}$  DES computations.*

## Experimental Results (2)

- Real complexity  $C_{(0.85)}$  seems to be lower (logarithmic scale):



- Experimental results suggest: *Given  $2^{43}$  known pairs, it is possible to recover the key with a success probability of 85 % within  $C_{(0.85)} = 2^{41}$  DES computations.*

## Experimental Results (3)

Other experimental results:

- Given  $2^{43}$  known pairs,  $\mathcal{C}_{(0.5)} \approx 2^{38.5}$ .
- Given  $2^{42.5}$  known pairs,  $\mathcal{C}_{(0.5)} \approx 2^{42}$ .
- Given  $2^{40}$  known pairs,  $\mathcal{C}_{(0.5)} \approx 2^{51.5}$ .



## Analysis (1)

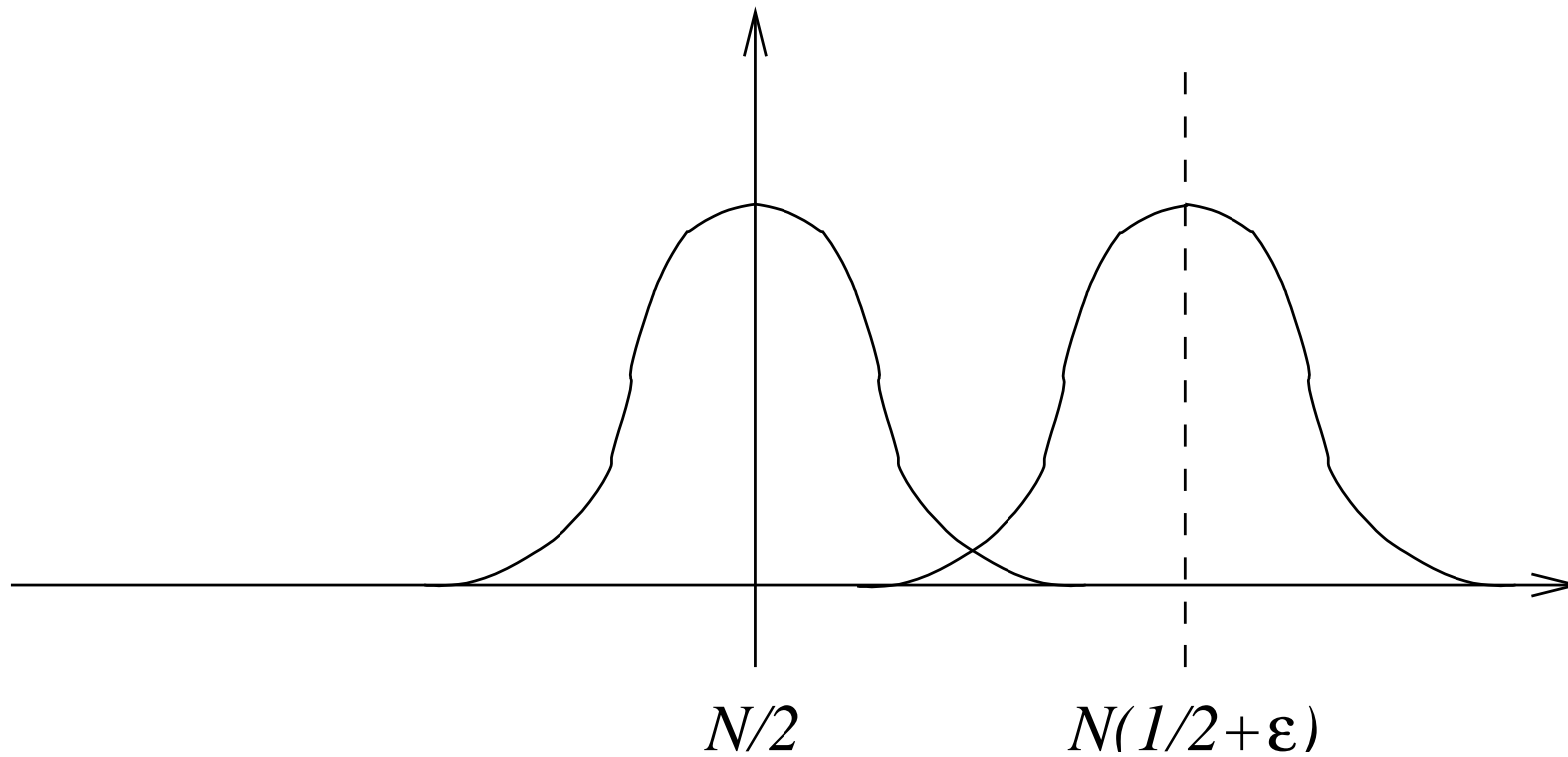
- Linear expression :  $P_{[i_1, \dots, i_r]} \oplus C_{[j_1, \dots, j_s]} = K_{[k_1, \dots, k_t]}$
- The expression must be biased in order to be useful:  
 $\Pr[\text{Expression holds}] = \frac{1}{2} + \epsilon, |\epsilon| > 0.$
- Wrong-key randomization hypothesis:

$$\frac{\left| \Pr[\text{Expression holds} \mid \text{right key}] - \frac{1}{2} \right|}{\left| \Pr[\text{Expression holds} \mid \text{wrong key}] - \frac{1}{2} \right|} \gg 1$$

## Analysis (2)

- *Assumption 1*: Bias produced by a wrong key is independent of the key
- *Assumption 2*: Bias produced by the right key is independent of the ones produced by wrong keys
- *Assumption 3*: The distribution of the biases is well approximated by a normal law

### Analysis (3)



*SAC'01, Toronto - Canada*

## Analysis (4)

- Counting / Analysis / Sorting / Searching phases
- Success Probability : key bits sum guessing, success within a given complexity
- Complexity is function of the right subkey rank  $\psi$  in the candidates list
- $n - 1$  wrong candidates follow a probability density  $f_W$ , the right one follows  $f_R$ .

## Analysis (5)

### Theorem 1

$$\Pr[\Psi \leq \psi] = \int_{-\infty}^{+\infty} B_{n+1-\psi, \psi}(F_W(x)) f_R(x) dx$$

and

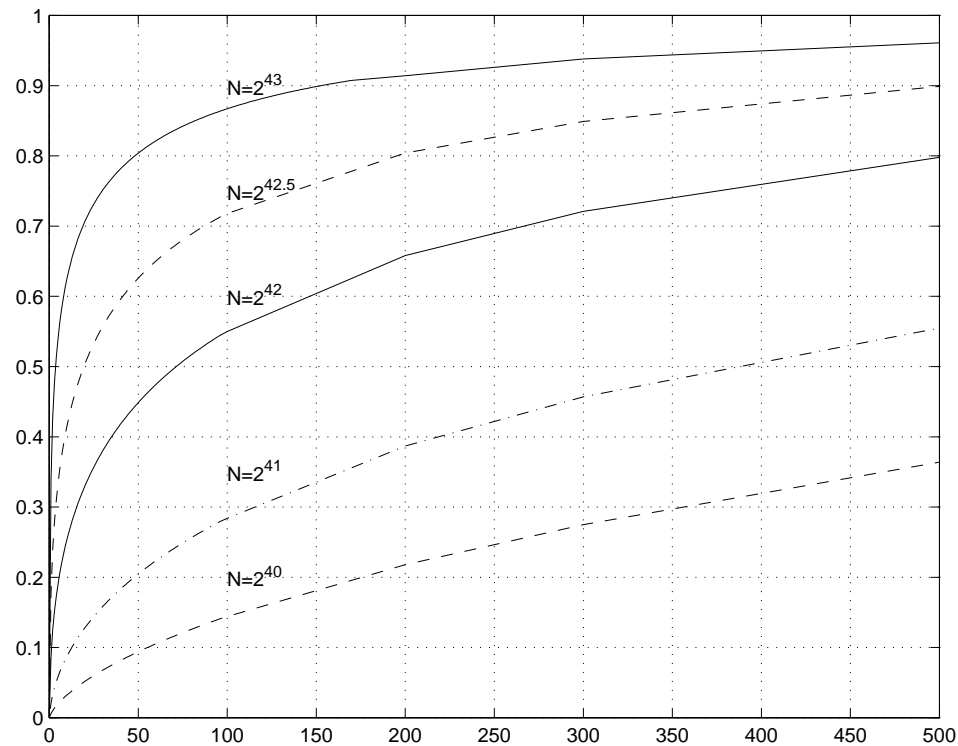
$$E[\Psi] = 1 + n \left( 1 - \int_{-\infty}^{+\infty} f_R(x) F_W(x) dx \right)$$

where

$$B_{a,b}(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1} (1-t)^{b-1} dt$$

is the incomplete beta function of order  $(a, b)$ .

## Analysis (6)



Theoretical rank distribution ( $\epsilon_w = 0$  and  $\epsilon_R =$  piling-up approximation) for various amounts of known pairs.

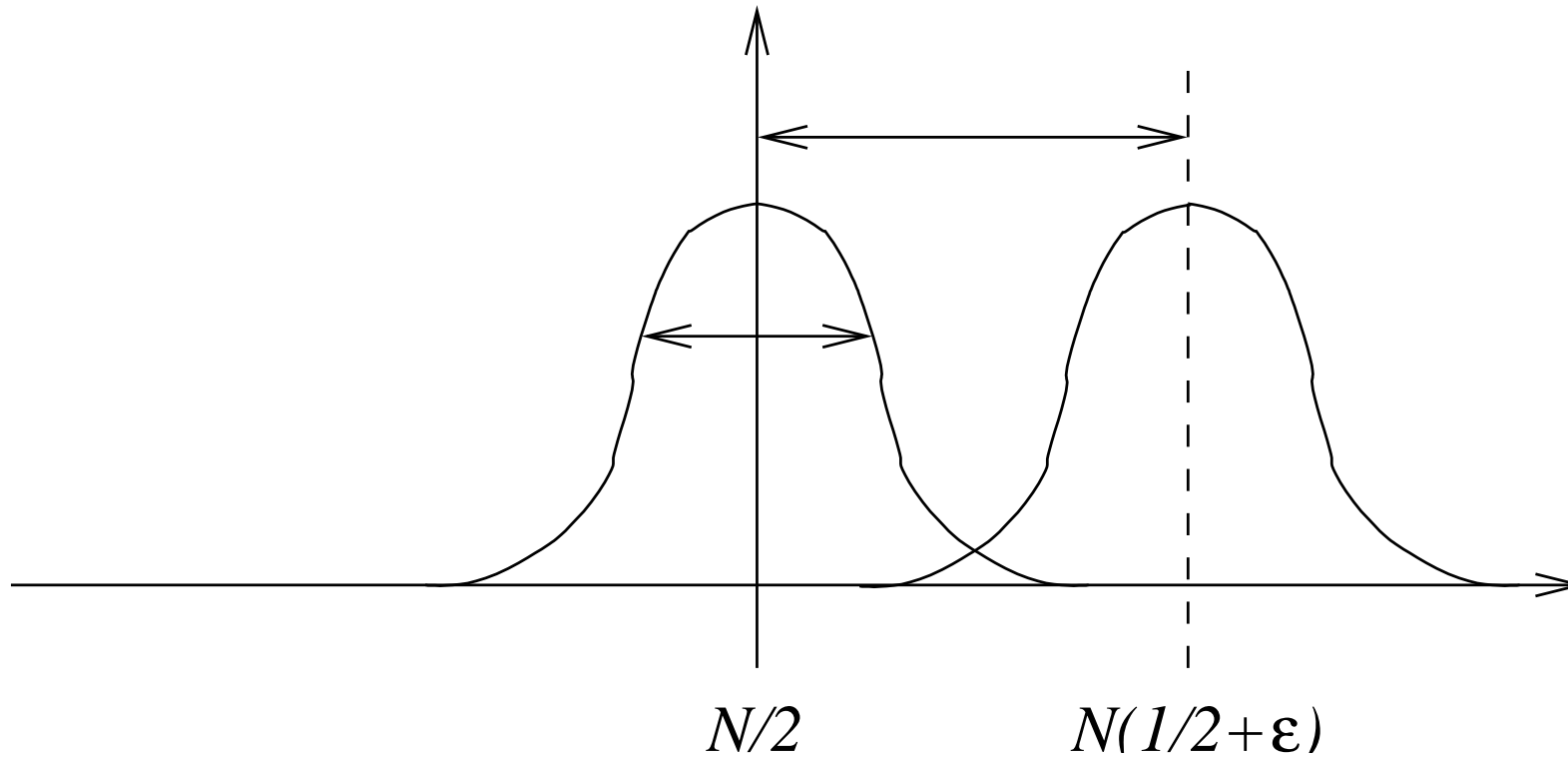
*SAC'01, Toronto - Canada*

## Analysis (7)

Some observations:

- Wrong-key randomization hypothesis holds well
- $\hat{\epsilon}_r - \epsilon_r$  is small (piling-up lemma approximation is OK, no linear hull effect)
- $\hat{\epsilon}_w \neq 0$ , but it doesn't matter a lot

### Analysis (8)



The experimental variances are smaller than the expected ones.



## Conclusion

- Experimental complexity analysis
- Theoretical analysis
- Partial inaccuracy of the model explained by experimental observations