On the Optimality of Linear, Differential and Sequential Distinguishers

Pascal Junod

Security and Cryptography Laboratory (LASEC)

Swiss Federal Institute of Technology, Lausanne

<pascal.junod@epfl.ch>

Introduction

- $\star$ One of the central problems for a cryptanalyst is to find a *statistically deviant* property in a block cipher.

- $\star$ Another problem: try to distinguish *efficiently* the deviant property from a "normal" behaviour.

- $\star$ Efficient $\equiv$ in terms of error probability *and* oracle queries.

Introduction (2)

- ⋆ Short survey of the litterature about cryptanalysis of block ciphers for the past 5 years (Eurocrypt-Crypto-Asiacrypt-SAC-FSE) : a big majority of the papers focuses on finding deviant properties.

- ⋆ In this paper, we are interested in the *efficiency problem*.

Goals

⋆ Goal: apply statistical concepts to well-known cryptanalytic techniques.

⟶ one can prove optimality results.

⟶ this can shed a new light on well-known cryptographic statistical procedures.

⟶ interestingly, one can derive practical applications !

Cryptanalysis and Statistics

- Old Cryptanalysis Era: statistics are widely used to break "old-school" ciphers.

- Modern Cryptanalysis Era

  - Davies (1987): attack against DES.

  - Biham and Shamir (1990): differential cryptanalysis.

  - Matsui (1993): linear cryptanalysis.

  - Vaudenay (1995): χ^2 cryptanalysis.

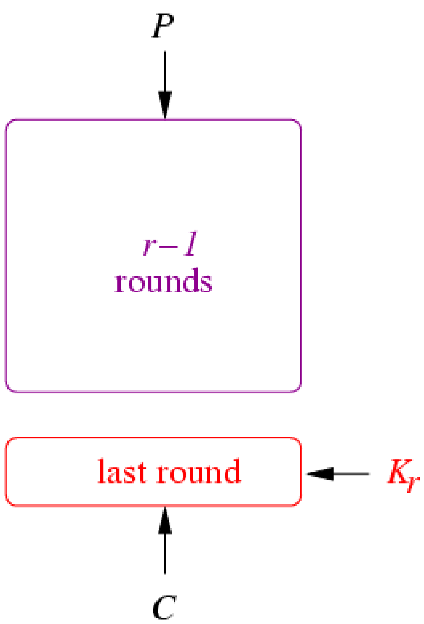
Cryptanalysis and Statistics (2)

- ⋆ Statistics give tools to break ciphers...

- ⋆ ... and (not frequently used in the crypto community) results about the performances and the behaviour of these tools !

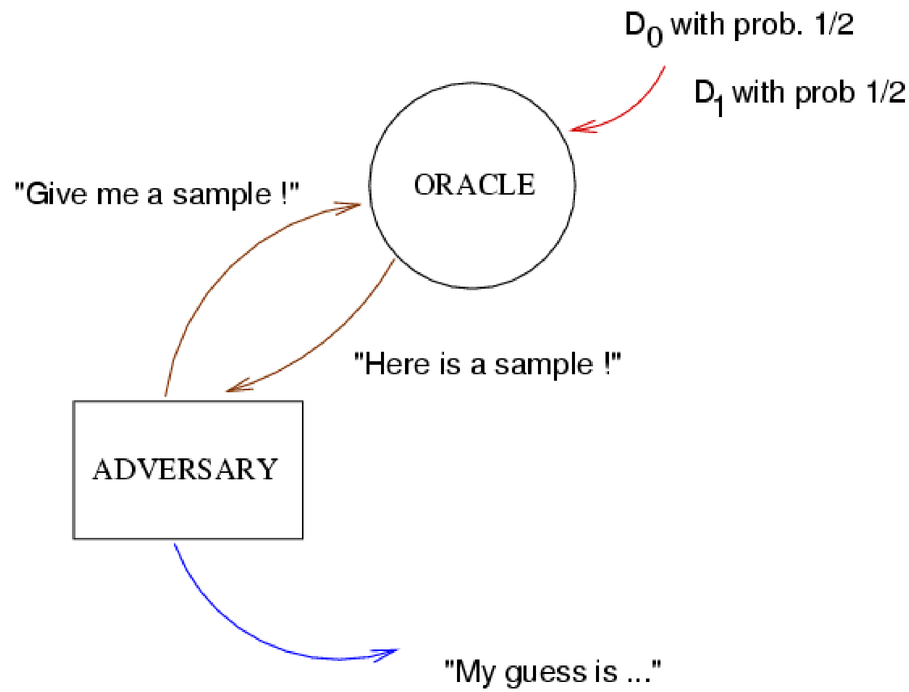
Last-Round Attack

- ⋆ A typical situation: a *1R-attack* (Biham-Shamir).
- ⋆ Under reasonable assumptions, we are in a *simple hypothesis test* situation.



Statistical Tests

- $\star$ D_0 and D_1, two *known*, *different* probability distributions defined on the same finite set $\mathcal{X}$.
- $\star$ Given an observation $x \in \mathcal{X}$ drawn according either to D_0 or to D_1, one has to decide which is the case.



Statistical Tests (2)

- ⋆ A decision rule $\delta : \mathcal{X} \rightarrow \{0, 1\}$ takes a sample x as input and defines what should be the guess for each possible $x \in \mathcal{X}$.

- ⋆ The optimal decision rule (in terms of error probabilities) is given by the Neyman-Pearson Lemma. It is based on the likelihood-ratio (denoted $\mathsf{lr}(x)$) concept:

$$\mathsf{lr}(x) \triangleq \frac{\Pr_{\mathsf{D}_0}[X = x]}{\Pr_{\mathsf{D}_1}[X = x]}$$

Statistical Tests (3)

In [BiSha90], describing the differential cryptanalysis, Biham and Shamir wrote:

"*We observed experimentally that when the signal-to-noise ratio is about 1-2, about 20-40 occurences of right pairs are sufficient. When the signal-to-noise ratio is much higher, even 3-4 pairs are usually enough. When the signal-to-noise ratio is much smaller, the identification of the right value of the subkey bits requires an unreasonably large number of pairs.*"

Linear Cryptanalysis

- ⋆ Linear Cryptanalysis: generic technique invented by Matsui in 1993 in an application to DES.

- ⋆ Principles: Find $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ such that

$$\mathbf{a} \cdot X \oplus \mathbf{b} \cdot C(X) = \mathbf{c} \cdot K$$

is probabilistically biased.

Optimality of a Linear Distinguisher

Vaudenay's modelization of a linear distinguisher δ_{lin}

Initialize a counter u to 0.
for $i = 1 \ldots n$ **do**
    Pick uniformly at random x and query $C(x)$ to the oracle Ω.
    if $\mathbf{a} \cdot x = \mathbf{b} \cdot C(x)$ **then**
        Increment u
    end if
end for
if $u \in \mathcal{A}^{(n)}$ **then**
    Output 0
else
    Output 1
end if

Optimality of a Linear Distinguisher (2)

- Optimality in terms of advantage

$$\left| \Pr[\delta_{\text{lin}} = 0 | \Omega = 0] - \Pr[\Pr[\delta_{\text{lin}} = 0 | \Omega = 1] \right| \qquad (1)$$

$\longrightarrow$ Neyman-Pearson is the solution!

- Optimality in terms of number of oracle queries: please wait a few slides !

Optimality of a Linear Distinguisher (3)

Theorem 1
The optimal acceptance region for δ_{lin} is

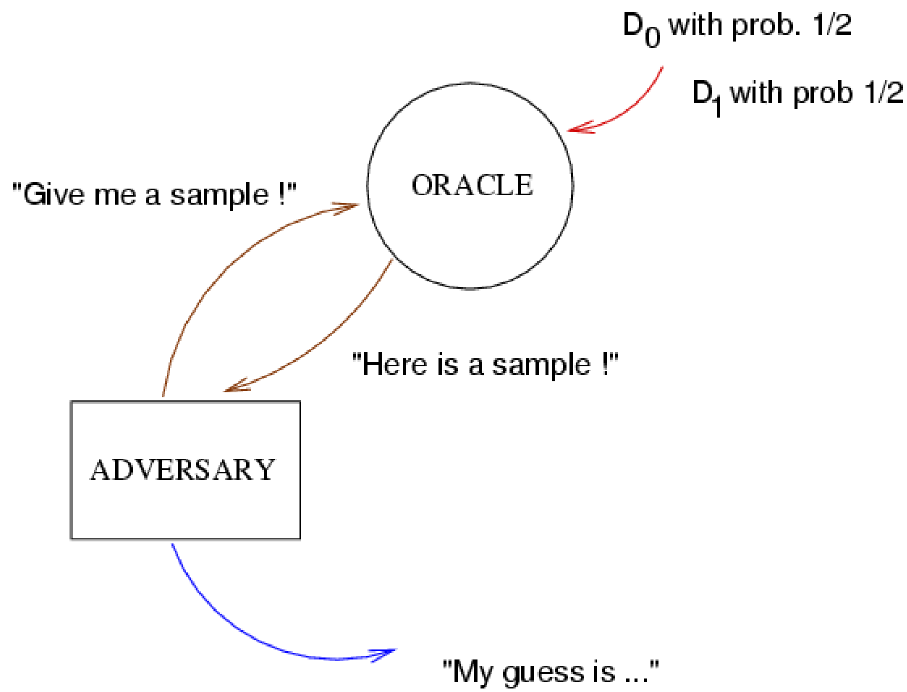
$$\mathcal{A}_{\text{opt}}^{(n)} = \left\{ u \in \{0, \ldots, n\} : u \geq n \cdot \frac{\log_2(1 - 2\epsilon)}{\log_2(1 - 2\epsilon) - \log_2(1 + 2\epsilon)} \right\}$$

where ϵ is the bias of the linear expression.

For $\epsilon > 0$ small, a good (and *intuitive*) approximation is given by

$$\mathcal{A}_{\text{opt}}^{(n)} \approx \left\{ u \in \{0, \ldots, n\} : u \geq n \cdot \left(\frac{1}{2} + \frac{\epsilon}{2} \right) \right\}$$

Asmptotic Behaviour of δ_{lin}



Asmptotic Behaviour of δ_{lin} (2)

Theorem 2
Let m be the block size of the involved permutations. For any distinguisher in the model described four slides ago,

$$1 - \frac{(n+1)}{2^{n\nu-1}} \leq \mathsf{BestAdv}^{n}_{\delta_{\mathsf{lin}}}(C, C^*) \leq 1 - \frac{1}{(n+1) \cdot 2^{n\nu-1}} \tag{2}$$

where $\nu = C(\mathsf{D}_0, \mathsf{D}_1)$ is the Chernoff information between D_0, a binary distribution having a bias equal to $\max\{\frac{1}{2^m-1}, \epsilon\}$ *such that* $\mathsf{ELP}^C(\mathbf{a}, \mathbf{b}) = 4\epsilon^2$ *and the uniform binary distribution D_1.*

Chernoff Information:

$$C(\mathsf{D}_0, \mathsf{D}_1) \triangleq - \min_{0 \leq \lambda \leq 1} \log \left(\sum_{x \in \mathcal{X}} \Pr_{X_0}[x]^{\lambda} \Pr_{X_1}[x]^{1-\lambda} \right)$$

Optimality of a Differential Distinguisher

Vaudenay's modelization of a differential distinguisher δ_{diff}

```
for i = 1 ... n do
    Pick uniformly at random x and query C(x) and C(x + a)
    to the oracle Ω.
    if C(x + a) = C(x) + b then
        Output 0 and stop.
    end if
end for
Output 1.
```

Sequential Distinguishers

- ⋆ Interesting point: in this modelization, the number N of queries is not constant, but merely a random variable !

- ⋆ It is a sequential distinguisher !

- ⋆ This kind of algorithm appeared explicitely in only two locations: in Davies and Murphy's paper (Journal of Cryptology, 1995), and in Murphy *et al*, an unpublished technical report (1995).

Sequential Distinguishers (2)

A sequential distinguisher is made of:

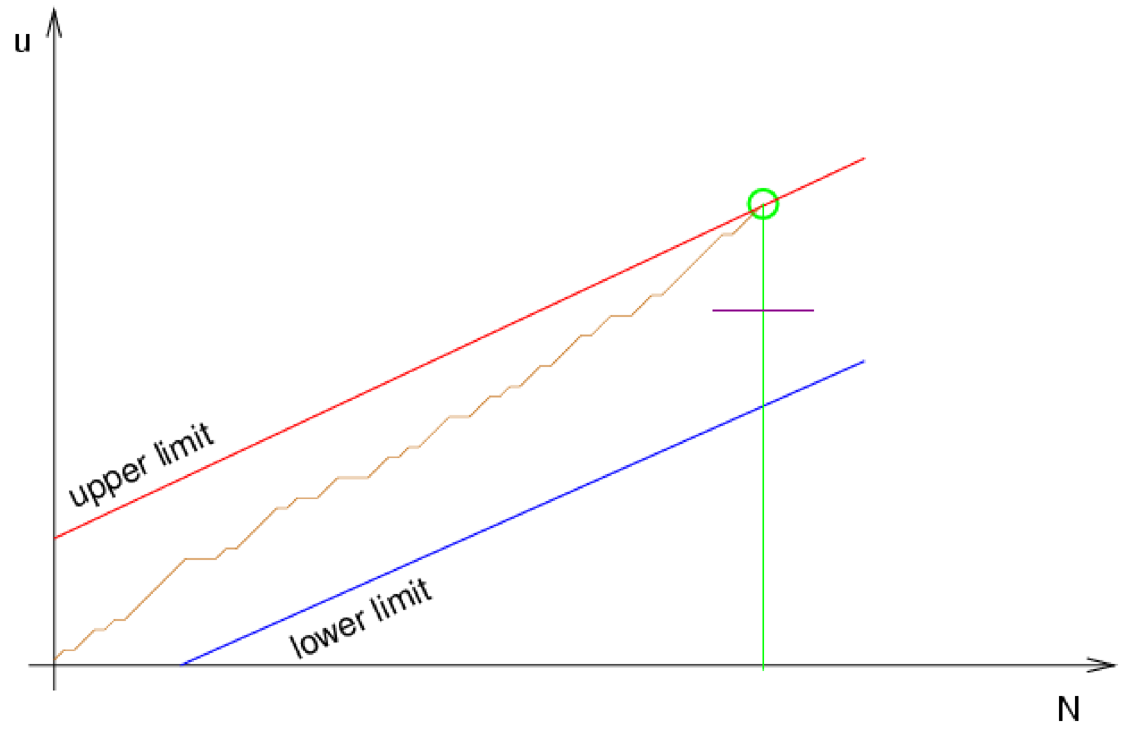
- ⋆ a *stopping rule* which decides either to take a decision or to query another sample,

- ⋆ a *decision rule* which specifies the guess to be taken.

Sequential Distinguishers (3)

Theorem 3 (Wald)
For testing a simple hypothesis against a simple alternative with independent, identically distributed observations, a sequential likelihood-ratio test *is optimal in the sense of* minimizing the expected sample size *among all tests having* no larger error probabilities*.*

Sequential Distinguishers (4)



Optimality of a Differential Distinguisher (bis)

for $i = 1 \ldots n$ **do**
    Pick uniformly at random x and query $C(x)$ and $C(x+a)$ to the oracle Ω.
    if $C(x+a) = C(x) + b$ **then**
        Output 0 and stop.
    end if
end for
Output 1.

$\rightarrow$ it is optimal in both aspects (sample size and advantage) if ϵ is *small*.

Use of Sequential Distinguishers

⋆ Linear cryptanalysis of 5-rounds DES.

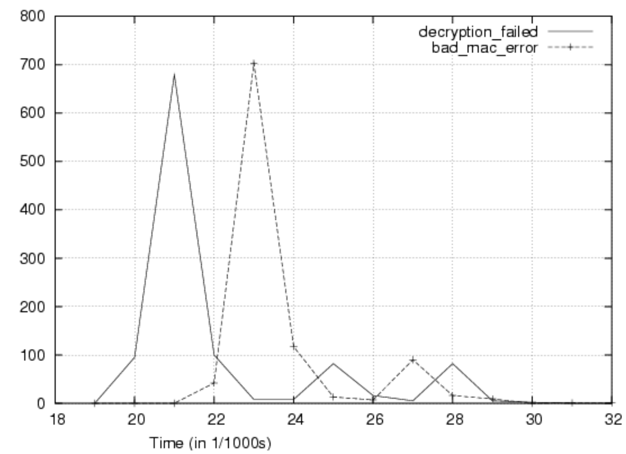
⋆ We try to guess the parity of the sum of involved key bits.

⋆ Using a (classical) static test, we need 2800 pairs in order to get a success probability equal to 97 %.

⋆ Result: we need an average number of queries equal to 1218 queries (instead of 2800) for the same success probability !

Use of Sequential Distinguishers (2)

⋆ LASEC's timing attack against SSL (see CRYPTO'03 paper of Canvel *et al.*): even under rough assumptions, a sequential distinguisher allows to decrease the number of queries to the attacked server by a factor of 5.



Use of Sequential Distinguishers (3)

⋆ This kind of distinguishers may be applied with success everytime when one has good approximations of the underlying probability distributions.

⋆ But please note that the costs of getting the information needed to compute the likelihood-ratio have to be taken into account !

Conclusion

- ⋆ We considered classical modelization of linear and differential cryptanalysis under a statistical point of view.

- ⋆ We provided results about the optimality and the asymptotic behaviour of these distinguishers.

- ⋆ We have "exhumed" the concept of sequential distinguisher.

THANK YOU !

The long version of this paper is available on

`http://eprint.iacr.org/2003/64`