Optimal Key Ranking Procedures in a Statistical Cryptanalysis

Pascal Junod and Serge Vaudenay

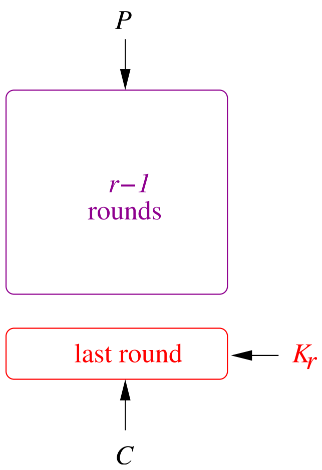
Security and Cryptography Laboratory (LASEC)
Swiss Federal Institute of Technology, Lausanne
{pascal.junod, serge.vaudenay}@epfl.ch

Contents

Introduction

⋆ A typical problem for a cryptanalyst: try to find something "deviant" in a cryptographic primitive.

Another typical problem: try to distinguish *efficiently* the (sub-) key(s) which makes deviate the primitive the most.



Introduction (2)

- ⋆ In this talk: we are interested in certain settings of the second problem.

- ⋆ One can view this problem in a more general way than the cryptographic one.

Introduction (3)

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

- ⋆ Result: one can prove optimality results.

- ⋆ Interestingly, this has practical applications !

Statistical Tests

⋆ D_0 and D_1, two different probability distributions defined on the same finite set $\mathcal{X}$.

⋆ Given an element $x \in \mathcal{X}$ (modeled by a random variable denoted X) drawn according either to D_0 or to D_1, one has to decide which is the case.

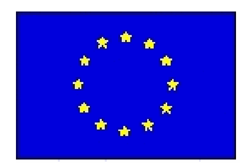
Statistical Tests (2)

- One uses a decision rule
$$\delta : \mathcal{X} \to \{0, 1\}$$
taking a sample of X as input and defining what should be the guess for each possible $x \in \mathcal{X}$.

- Two different types of error probabilities:
$$\alpha \triangleq \Pr_{X_0}[\delta(X) = 1]$$
$$\beta \triangleq \Pr_{X_1}[\delta(X) = 0]$$

Statistical Tests (3)

A Swiss instance of the problem: in 1992, Swiss people had to vote whether they wanted to become European or not.



Statistical Tests (4)

* It was possible to separate the Swiss (voting) population in two categories according to a simple criterion.

  1. In one part of the voters, a big majority was in favour of becoming European.
  2. In the other part of the voters, a big majority was in favour of not becoming European.

* Question: given a random Swiss citizen, what is the best way to decide whether (s)he voted YES or NO become an European ?

Statistical Tests (5)

⋆ In statistics, one calls this type of decision a binary hypothesis test (or *simple hypothesis test*).

⋆ In fact, each of these hypotheses *completely* specifies the probability distributions.

⋆ An hypothesis test which is not simple is called composite hypothesis test. For instance, a χ^2-test is a composite test.

Statistical Tests (6)

- The decision rule δ defines a partition of $\mathcal{X}$ in two disjoint subsets $\mathcal{A}$ and $\overline{\mathcal{A}}$.

- The optimal decision rule is given by the Neyman-Pearson Lemma based on the *likelihood-ratio*:

$$\mathcal{A} \triangleq \left\{ x \in \mathcal{X} : \frac{\Pr_{X \leftarrow \mathsf{D}_0}[x]}{\Pr_{X \leftarrow \mathsf{D}_1}[x]} \geq \tau \right\} \qquad (1)$$

Statistical Tests (7)

Definition 1 (Optimal Binary Hypothesis Test)
To test $X \leftarrow \mathsf{D}_0$ *against* $X \leftarrow \mathsf{D}_1$*, choose a constant* $\tau > 0$ *depending on* α *and* β *and define the likelihood ratio*

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

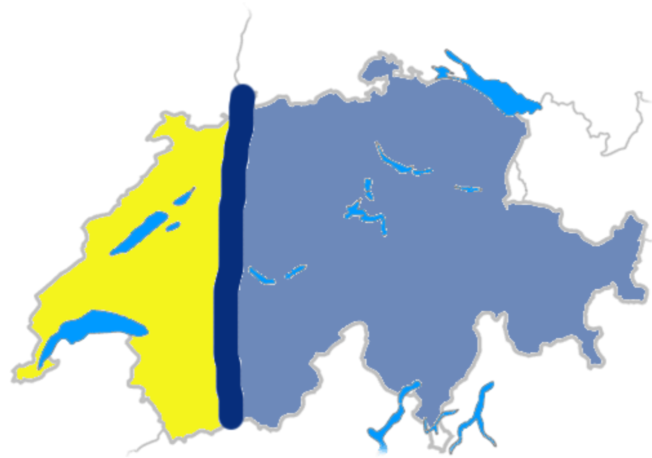
The optimal decision function is then defined by

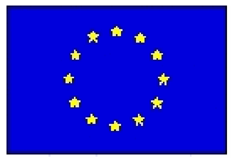
$$\delta_{\mathsf{opt}} \triangleq \begin{cases} 0 & (i.e \text{ accept } X \leftarrow \mathsf{D}_0) & \text{if } \mathsf{lr}(x) \geq \tau \\ 1 & (i.e. \text{ accept } X \leftarrow \mathsf{D}_1) & \text{if } \mathsf{lr}(x) < \tau \end{cases}$$

Statistical Tests (8)

Back to the Swiss instance of the problem: let us assume that our first hypothesis is "voted YES"; a likelihood-ratio decision rule could have been "*Is your mothertongue French ?*".

- $\alpha \equiv$ probability that a french-speaking Swiss citizen voted NO.
- $\beta \equiv$ probability that a german-speaking, italian-speaking or rumantsch-speaking Swiss citizen voted YES.





Optimal Key Ranking Procedures

⋆ Linear Cryptanalysis: generic technique invented by Matsui in 1993 in an application to DES. Refined and implemented in 1994.

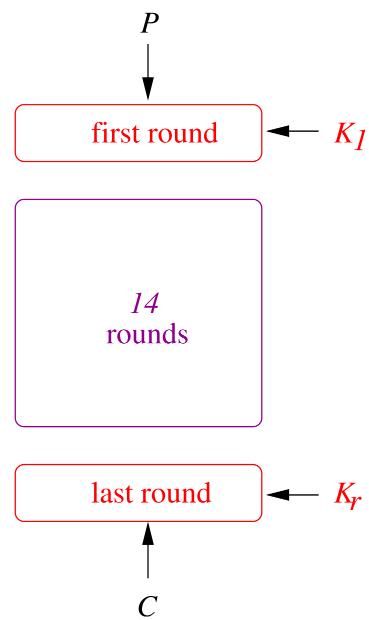
⋆ Principles: Find $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ such that

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

is probabilistically biased.

Optimal Key Ranking Procedures (2)

With full-DES (16 rounds), take the best 14-rounds linear characteristic, then decrypt the first and last rounds with subkey candidates.



Optimal Key Ranking Procedures (3)

⋆ For each subkey candidate, count the number of times that the linear approximation is equal to 0, given all the plaintext and ciphertext pairs ($N \approx 2^{43}$ for DES)

⋆ If there is enough plaintext-ciphertext pairs, the good subkey candidate should deviate the most from $\frac{N}{2}$.

⋆ Search exhaustively for the remaining missing key bits for the best candidate.

Optimal Key Ranking Procedures (4)

Prepare m counters $u_i, 1 \leq i \leq m$ and initialize them to 0.
for all Known plaintext-ciphertext pairs at disposal **do**
    for all Subkey candidates **do**
        Decrypt the first and last rounds and evaluate the linear expression.
        if It evaluates to 0 **then**
            Increment the corresponding counter
        end if
    end for
end for
Output the subkey candidate corresponding to the most biased counter as the right one.

Optimal Key Ranking Procedures (5)

⋆ Data complexity: the number N of needed known plaintext-ciphertext pairs.

⋆ Computational complexity: the number of DES evaluations during the exhaustive search part.

⋆ Key ranking was introduced in 1994 Matsui's paper; instead of taking the most biased, take the ℓ most biased and search them one after the other for the remaining unknown bits.

Optimal Key Ranking Procedures (6)

- ⋆ Ranking strategy ?

- ⋆ Intuitive way (the one in Matsui's paper): rank them from the highest to the smallest bias.

- ⋆ Is it optimal in terms of computational complexity ?

Optimal Key Ranking Procedures (7)

⋆ Neyman-Pearson Ranking Procedure: if probability distributions modelling the subkeys are available, one can rank the candidates by decreasing likelihood-ratio.

⋆ Under reasonable hypotheses, they are known in the case of a linear cryptanalysis [Jun01].

Optimal Key Ranking Procedures (8)

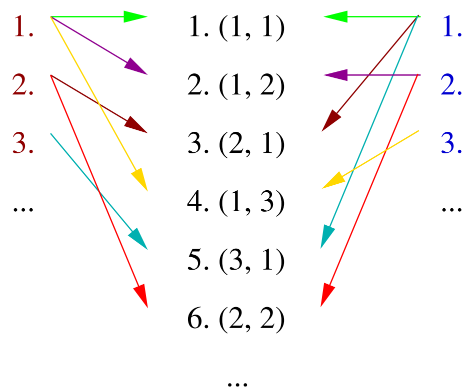
- ⋆ One can show that this ranking procedure is optimal in terms of computational complexity.

- ⋆ Matsui's ranking procedure is equivalent to a Neyman-Pearson Ranking Procedure (and thus optimal).

Optimal Key Ranking Procedures (9)

- More interesting problem: Matsui's refined attack (1994) uses *two* linear approximations involving *disjoint* key bits subsets.

- Matsui's proposition (based on intuition): rank them independantly following their bias, and then build a single list sorted by increasing product of ranks.



Optimal Key Ranking Procedures (10)

- ⋆ Interestingly, one can easily use a NP-Ranking Procedure.

- ⋆ Optimal in terms of computational complexity.

Optimal Key Ranking Procedures (11)

⋆ In the case of DES, the likelihood-ratio is given by

$$\mu_{(\ell_1,\ell_2)} = 2e^{-2n\epsilon^2} \cdot \cosh(4\epsilon\Sigma_{\ell_1}) \cdot \cosh(4\epsilon\Sigma_{\ell_2}) \tag{2}$$

⋆ Taylor approximation:

$$\mu_{(\ell_1,\ell_2)} \approx 2 + (16\Sigma_{\ell_1}^2 + 16\Sigma_{\ell_2}^2 - 4n)\epsilon^2 + O(\epsilon^4) \tag{3}$$

⋆ Simple to implement: sort by decreasing sum of the squares of the biases !

Optimal Key Ranking Procedures (12)

- ⋆ Experimental results on 21 linear cryptanalysis of DES: decrease of about 50 % of the computational complexity.

- ⋆ One can convert this gain in a decrease of N (about 31 %).

- ⋆ A possible tradeoff: given $2^{42.46}$ known plaintext-ciphertext pairs, it was possible to recover a complete DES key within $2^{44.46}$ DES evaluations with a success probability equal to 85 %.

Conclusion

- Situations of binary hypothesis tests occurs very frequently in cryptography.

- Using concepts of statistics, one can design optimal distinguishing procedures.

THANK YOU !