On the Data Complexity of Statistical Attacks against Block Ciphers

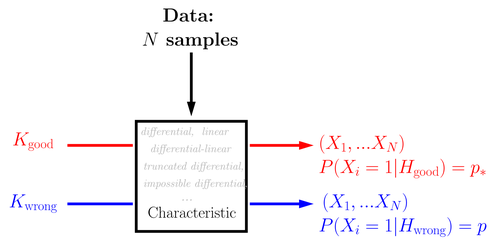
C. Blondeau and B. Gérard
INRIA, Paris Rocquencourt, France

Issue & Notation

Context: statistical attacks based on a distinguisher.

Two hypotheses:

- H_{good}: "correct subkey guess"
- H_{wrong}: "non correct subkey guess"



- $S_{N,p_*} = \sum_{i=1}^{N} X_i$ follows a binomial law of parameters (N, p_*)
- $S_{N,p} = \sum_{i=1}^{N} X_i$ follows a binomial law of parameters (N, p)

Two kinds of errors:

- $\boldsymbol{\alpha}$ Non detection error probability $\alpha = P(S_{N,p_*} < T)$
- $\boldsymbol{\beta}$ False alarm error probability $\beta = P(S_{N,p} \geq T)$

Aim: Finding the minimal number N of samples required to reach some given error probabilities.

An algorithm for finding N

Input: (α, β) and (p_*, p)
Output: N and τ the minimum number of samples and the corresponding relative threshold to reach error probabilities less than (α, β).

$\tau_{\min} \leftarrow p$ and $\tau_{\max} \leftarrow p_*$.
repeat
  $\tau \leftarrow \dfrac{\tau_{\min} + \tau_{\max}}{2}$.
  Compute N_{nd} such that $\forall N > N_{\text{nd}},\ P(S_{N,p_*} < \tau N) \leq \alpha$.
  Compute N_{fa} such that $\forall N > N_{\text{fa}},\ P(S_{N,p} \geq \tau N) \leq \beta$.
  if $N_{\text{nd}} > N_{\text{fa}}$ **then** $\tau_{\max} = \tau$ **else** $\tau_{\min} = \tau$
until $N_{\text{nd}} = N_{\text{fa}}$.
return N and τ.

Related work

[Junod 01,03,05]: Extensive study of linear cryptanalysis (Gaussian approximation).

[Baignères, Junod and Vaudenay 04]: General study of distinguishers. In particular the exponential behavior of binomial tail is exhibited .

[Selçuk 08]: Key ranking study for linear and differential cryptanalysis.
"It can reasonably be said that the normal approximation for the binomial counters may not be accurate when $Np_*(1-p_*) < 4$"

Some approximations

Approximations of the binomial law :

Linear cryptanalysis $\longrightarrow$ Gaussian approximation is valid.
Differential cryptanalysis $\longrightarrow$ Poisson approximation is valid.
Truncated-differential cryptanalysis $\longrightarrow$???

Parameters	Exact probabilities	Poisson	Gaussian	Kullback
L: $N = 2^{23}, \tau = 0.5 + 2^{-10.58}$ $p_* = 0.5 + 2^{-10}, p = 0.5$	$\boldsymbol{\beta = 8.12 \cdot 10^{-5}}$ $\boldsymbol{\alpha = 2.97 \cdot 10^{-2}}$	$\beta = 3.84 \cdot 10^{-3}$ $\alpha = 9.14 \cdot 10^{-2}$	$\boldsymbol{\beta = 8.12 \cdot 10^{-5}}$ $\boldsymbol{\alpha = 2.97 \cdot 10^{-2}}$	$\beta = 8.62 \cdot 10^{-5}$ $\alpha = 3.58 \cdot 10^{-2}$
D: $N = 2^{23}, \tau = 2^{-23}$ $p_* = 2^{-20}, p = 2^{-27}$	$\boldsymbol{\beta = 2.03 \cdot 10^{-3}}$ $\boldsymbol{\alpha = 3.27 \cdot 10^{-3}}$	$\boldsymbol{\beta = 2.03 \cdot 10^{-3}}$ $\boldsymbol{\alpha = 3.27 \cdot 10^{-3}}$	$\beta = 8.84 \cdot 10^{-5}$ $\alpha = 6.66 \cdot 10^{-3}$	$\beta = 1.97 \cdot 10^{-3}$ $\alpha = 3.33 \cdot 10^{-3}$
TD(1): $N = 2^{23}, \tau = 1.005 \cdot 2^{-4}$ $p_* = 1.01 \cdot 2^{-4}, p = 2^{-4}$	$\boldsymbol{\beta = 9.29 \cdot 10^{-5}}$ $\boldsymbol{\alpha = 9.80 \cdot 10^{-5}}$	$\beta = 1.46 \cdot 10^{-4}$ $\alpha = 1.55 \cdot 10^{-4}$	$\beta = 9.23 \cdot 10^{-5}$ $\alpha = 9.89 \cdot 10^{-5}$	$\beta = 9.90 \cdot 10^{-5}$ $\alpha = 1.04 \cdot 10^{-4}$
TD(2): $N = 2^{23}, \tau = 1.25 \cdot 2^{-15}$ $p_* = 1.5 \cdot 2^{-15}, p = 2^{-15}$	$\boldsymbol{\beta = 5.05 \cdot 10^{-5}}$ $\boldsymbol{\alpha = 4.37 \cdot 10^{-4}}$	$\boldsymbol{\beta = 5.06 \cdot 10^{-5}}$ $\boldsymbol{\alpha = 4.38 \cdot 10^{-4}}$	$\beta = 3.17 \cdot 10^{-5}$ $\alpha = 5.45 \cdot 10^{-4}$	$\beta = 5.34 \cdot 10^{-5}$ $\alpha = 4.67 \cdot 10^{-4}$

Error probabilities obtained with different binomial approximations compared to real ones for:
L: linear cryptanalysis, **D:** differential cryptanalysis and **TD(1)-(2):** truncated differential cryptanalysis.

Approximating binomial tails

Aim : Finding a valid approximation for all parameters p_* and p.
Approximation from [Arriata, Gordon, 89].

$$P(S_{N,p} \geq \tau N) \underset{N \to \infty}{\sim} \frac{(1-p)\sqrt{\tau}}{(\tau - p)\sqrt{2\pi N(1-\tau)}}\ 2^{-N D(\tau||p)},$$

The Kullback-Leibler divergence:

$$D(p||q) = p \log_2\left(\frac{p}{q}\right) + (1-p)\log_2\left(\frac{1-p}{1-q}\right)$$

Rather good approximations

Approximation for N in the case where $\tau = p_*$ (implying $\alpha \approx 0.5$).

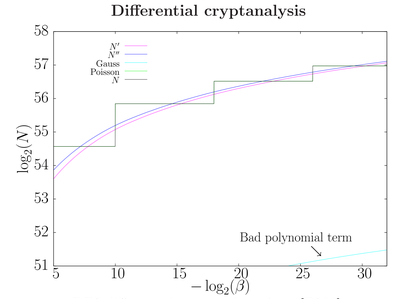
$$N' = -\frac{1}{D(p_*||p)}\left[\log_2\left(\frac{\lambda\beta}{\sqrt{D(p_*||p)}}\right) + 0.5\log_2\left(-\log_2(\lambda\beta)\right)\right],$$

$$\text{with} \quad \lambda = \frac{(p_* - p)\sqrt{2\pi(1-p_*)}}{(1-p)\sqrt{p_*}}.$$

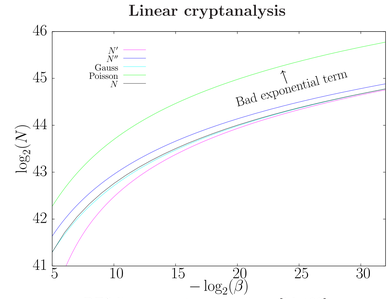
This formula can be simplified with a small loss of precision to give the estimate N''.

$$N'' = -\frac{\log_2(2\sqrt{\pi}\beta)}{D(p_*||p)}.$$

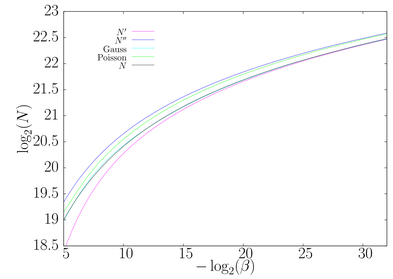
Experimental results

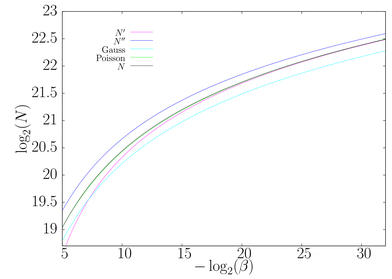


DES differential cryptanalysis from [BS93].



DES linear cryptanalysis from [Mat93].





Asymptotic data complexity

The estimate N'' shows that for fixed parameters α and β, the required number of requests depends essentially on the Kullback-Leibler divergence.

Attack	Parameters	Classical results	$\dfrac{1}{D(p_*\|\|p)}$
Linear	$p = 0.5$ $\quad p_* = 0.5 + \epsilon$	$\dfrac{1}{(p_* - p)^2}$	$\dfrac{1}{2(p_* - p)^2}$
Differential	$p_* \ll 1$ $\quad p_* \gg p$	$\dfrac{1}{p_*}$	$\dfrac{1}{p_* \log_2(p_*/p) - p_*}$
Differential-linear	$p = 0.5$ $\quad p_* - p \ll p$	$\dfrac{1}{(p_* - p)^2}$	$\dfrac{1}{2(p_* - p)^2}$
Truncated differential	$p_* \ll 1$ $\quad p_* - p \ll p$	unknown	$\dfrac{2p}{(p_* - p)^2}$
Impossible differential	$p_* = 0$ $\quad p \ll 1$	implicitly : $\dfrac{1}{p}$	$\dfrac{1}{p}$
k-th order differential	$p_* = 1$ $\quad p \ll 1$	1	$-\dfrac{1}{\log_2 p}$