On the Data Complexity of Statistical Attacks against Block Ciphers

Céline Blondeau and Benoît Gérard

INRIA project-team SECRET, France

WCC - May the 14th 2009
Outline

1. Introduction
2. Algorithm to find the data complexity
3. A problem: Approximation of the involved binomial tails
4. A formula to approximate the Data Complexity
5. Asymptotic behavior for some statistical attacks
Outline

1. Introduction

2. Algorithm to find the data complexity

3. A problem: Approximation of the involved binomial tails

4. A formula to approximate the Data Complexity

5. Asymptotic behavior for some statistical attacks
Some known statistical cryptanalyses:

- linear cryptanalysis [Matsui 93];
- differential cryptanalysis [Biham Shamir 91];
- higher order differential cryptanalysis [Knudsen 94];
- impossible differential cryptanalysis [Biham Biryukov Shamir 99];
- ...
Using a characteristic to distinguish from random

Let $\chi$ be some characteristic on a given cipher.

- If the sub-key guess is correct: $\chi$ occurs with probability $p_\star$.
- If the sub-key guess is not correct: $\chi$ occurs with probability $p$.

$$X_i = \begin{cases} 1 & \text{if } \chi \text{ occurs in sample } i, \\ 0 & \text{otherwise.} \end{cases}$$

$N$ samples

Characteristic

$K_{\text{good}}$

$K_{\text{wrong}}$

$(X_1, \ldots X_N)$

$P(X_i = 1|H_{\text{good}}) = p_\star$

$(X_1, \ldots X_N)$

$P(X_i = 1|H_{\text{wrong}}) = p$
Neyman-Pearson (optimal) test:
Accept a candidate $K$ if

$$\frac{P(X_1, X_2, \ldots, X_N|H_{\text{good}})}{P(X_1, X_2, \ldots, X_N|H_{\text{wrong}})} > t.$$ 

This (likelihood) ratio only depends on $S_N = \sum_{i=1}^{N} X_i$, $p_*$ and $p$ and is increasing in $S_N$.

Thus, the acceptance condition becomes, for some threshold $0 < T < N$,

$$S_N > T$$

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

Two kinds of errors can be made:

- **Non-detection error probability** \( P(S_{N,p*} < T) \);
- **False alarm error probability** \( P(S_{N,p} \geq T) \).

The **non-detection error probability** corresponds to the success probability of the cryptanalysis.

The **false alarm error probability** is the expected ratio of kept candidates and thus influences the time complexity of the cryptanalysis.

**Aim:** Finding \( N \) minimal and the corresponding \( T \) such that \( P(S_{N,p*} < T) \leq \alpha \) and \( P(S_{N,p} \geq T) \leq \beta \) for given values of \( \alpha \) and \( \beta \).
Motivation

Generalized Feistel Network
[Nyberg 96] with:
- 10 rounds;
- 4 S-boxes.

Truncated differential path: \( p^* = 1.18 \cdot 2^{-16} \) and \( p = 2^{-16} \)
Differential path: \( p^* = 1.53 \cdot 2^{-27} \) and \( p = 2^{-32} \)

Question:
Which couple of parameters gives the best cryptanalysis ???
Outline

1. Introduction

2. Algorithm to find the data complexity

3. A problem: Approximation of the involved binomial tails

4. A formula to approximate the Data Complexity

5. Asymptotic behavior for some statistical attacks
An algorithm for finding $N (1/2)$

- For a fixed $\tau = T/N$, error probabilities decrease when $N$ increases.
- For a fixed $N$, non-detection error increases with $\tau$.
- For a fixed $N$, false alarm error decreases when $\tau$ increases.

Problem

$N$ and $T$ are integers $\Rightarrow$ restrictions on $\tau$.

Solution

Approximations of probabilities ($G_{nd}$ and $G_{fa}$) that:

- allow non integers values for $N$ and $T$;
- have the expected properties.
An algorithm for finding $N$ (2/2)

**Input:** $(\alpha, \beta)$ and $(p_*, p)$  
**Output:** $N$ and $\tau$ the minimum number of samples and the corresponding relative threshold to reach error probabilities less than $(\alpha, \beta)$.

\[
\tau_{\text{min}} \leftarrow p \quad \text{and} \quad \tau_{\text{max}} \leftarrow p_*. 
\]

repeat
\[
\tau \leftarrow \frac{\tau_{\text{min}} + \tau_{\text{max}}}{2}. 
\]
Compute $N_{\text{nd}}$ such that $\forall N > N_{\text{nd}}, \ G_{\text{nd}}(N, \tau) \leq \alpha$.
Compute $N_{\text{fa}}$ such that $\forall N > N_{\text{fa}}, \ G_{\text{fa}}(N, \tau) \leq \beta$.

if $N_{\text{nd}} > N_{\text{fa}}$ then $\tau_{\text{max}} = \tau$ else $\tau_{\text{min}} = \tau$

until $N_{\text{nd}} = N_{\text{fa}}$.

return $N$ and $\tau$. 
Number of required samples $N$ for differential and truncated-differential cryptanalyses

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\log(N)$ (differential)</th>
<th>$\log(N)$ (truncated differential)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.001</td>
<td>27.35</td>
<td>24.31</td>
</tr>
<tr>
<td>0.5</td>
<td>$10^{-10}$</td>
<td>29.25</td>
<td>26.37</td>
</tr>
<tr>
<td>0.01</td>
<td>0.001</td>
<td>29.43</td>
<td>25.94</td>
</tr>
<tr>
<td>0.01</td>
<td>$10^{-10}$</td>
<td>30.54</td>
<td>27.29</td>
</tr>
</tbody>
</table>

**Answer to the question:**

In that case, truncated differential is better than differential.

Differential: $p_* = 1.53 \cdot 2^{-27}$ and $p = 2^{-32}$

Truncated differential: $p_* = 1.18 \cdot 2^{-16}$ and $p = 2^{-16}$
Outline

1 Introduction

2 Algorithm to find the data complexity

3 A problem: Approximation of the involved binomial tails

4 A formula to approximate the Data Complexity

5 Asymptotic behavior for some statistical attacks
Gaussian approximation of the binomial tail

\[ P [S_{N,p} \leq N\tau] \simeq \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi Np(1-p)}} \cdot e^{-\frac{N(x-p)^2}{2p(1-p)}} \, dx \]

Classically used in linear cryptanalysis:

- [Matsui 93,94];
- [Gilbert 97];
- [Junod 01,03,05];
- [Selçuk 08]
- ...

But...

... not valid everywhere. For instance, when \( N \cdot p \) is too small as in differential cryptanalysis [Selçuk 08].
Poisson approximation of the binomial tail

\[ P \left[ S_{N,p} \leq N\tau \right] \simeq \sum_{k=0}^{\lfloor N\tau \rfloor} e^{-Np} \cdot \frac{(Np)^k}{k!} \]

Implicitly used in differential cryptanalysis:
- [Biham Shamir 91, 93];
- [Gilbert 97];
- [Selçuk 08]
- ...

But...

... not valid everywhere. For instance, when \( N \cdot p \) is too big as in linear cryptanalysis.
A good approximation of the binomial tail

We recall the binomial tail:

\[
P [S_{N,p} \leq N\tau] = \sum_{k=0}^{\lfloor N\tau \rfloor} \binom{n}{k} p^k (1 - p)^{n-k}
\]

Approximation found, for instance, in [Arriata, Gordon 89]:

\[
P(S_{N,p*} \leq N\tau) \sim \frac{p_* \sqrt{1 - \tau}}{(p_* - \tau) \sqrt{2\pi N\tau}} \cdot 2^{-N \cdot D(\tau||p*)}.
\]

Where the Kullback-Leibler divergence is defined by:

\[
D(p||q) = p \log_2 \left( \frac{p}{q} \right) + (1 - p) \log_2 \left( \frac{1 - p}{1 - q} \right).
\]
## Experimental results

<table>
<thead>
<tr>
<th>Method</th>
<th>Exact</th>
<th>Poisson</th>
<th>Gaussian</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lin Crypt:</strong></td>
<td>( p = 0.5 )</td>
<td>( \beta )</td>
<td>( 8.12 \cdot 10^{-5} )</td>
<td>( 8.12 \cdot 10^{-5} )</td>
</tr>
<tr>
<td>( p = 0.5 + 2^{-10} )</td>
<td>( \alpha )</td>
<td>( 2.97 \cdot 10^{-2} )</td>
<td>( 3.84 \cdot 10^{-3} )</td>
<td>( 2.97 \cdot 10^{-2} )</td>
</tr>
<tr>
<td><strong>Diff Crypt:</strong></td>
<td>( p = 2^{-27} )</td>
<td>( \beta )</td>
<td>( 2.03 \cdot 10^{-3} )</td>
<td>( 2.03 \cdot 10^{-3} )</td>
</tr>
<tr>
<td>( p = 2^{-20} )</td>
<td>( \alpha )</td>
<td>( 3.27 \cdot 10^{-3} )</td>
<td>( 3.27 \cdot 10^{-3} )</td>
<td>( 6.66 \cdot 10^{-3} )</td>
</tr>
<tr>
<td><strong>Trunc Diff(1):</strong></td>
<td>( p = 2^{-4} )</td>
<td>( \beta )</td>
<td>( 9.29 \cdot 10^{-5} )</td>
<td>( 1.46 \cdot 10^{-4} )</td>
</tr>
<tr>
<td>( p = 1.01 \cdot 2^{-4} )</td>
<td>( \alpha )</td>
<td>( 9.80 \cdot 10^{-5} )</td>
<td>( 1.55 \cdot 10^{-4} )</td>
<td>( 9.89 \cdot 10^{-5} )</td>
</tr>
<tr>
<td><strong>Trunc Diff(2):</strong></td>
<td>( p = 2^{-15} )</td>
<td>( \beta )</td>
<td>( 5.05 \cdot 10^{-5} )</td>
<td>( 5.06 \cdot 10^{-5} )</td>
</tr>
<tr>
<td>( p = 1.5 \cdot 2^{-15} )</td>
<td>( \alpha )</td>
<td>( 4.37 \cdot 10^{-4} )</td>
<td>( 4.38 \cdot 10^{-4} )</td>
<td>( 5.45 \cdot 10^{-4} )</td>
</tr>
</tbody>
</table>

These values are given for \( N = 2^{23} \) and \( \tau = \frac{p_* + p}{2} \).
Outline

1. Introduction
2. Algorithm to find the data complexity
3. A problem: Approximation of the involved binomial tails
4. A formula to approximate the Data Complexity
5. Asymptotic behavior for some statistical attacks
**Aim:** Finding a simple formula to estimate the data complexity.

Fixing $T$ simplifies the problem.

So we take $T = Np_*$ what implies $\alpha \simeq 50 \%$. 
Approximation of the data complexity (2)

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

where \( \lambda = \frac{(p_* - p)\sqrt{2\pi (1-p_*)}}{(1-p)\sqrt{p_*}}. \)

\[ N' \leq N_\infty \leq N' \left[ 1 + \frac{(\theta - 1) \log(\theta)}{\log(N')} \right], \]

with \( \theta = \left[ 1 + \frac{1}{2\log(\lambda \beta)} \log \left( -\frac{\log(\lambda \beta)}{D(p_*\|p)} \right) \right]^{-1}. \)

This is a good approximation of \( N \) when \( \beta \) tends to 0.
Experimental results (1)

Differential cryptanalysis of DES

\[ p_\ast = 1.87 \cdot 2^{-56}, \quad p = 2^{-64} \]

Linear cryptanalysis of DES

\[ p_\ast = 0.5 + 1.19 \cdot 2^{-21}, \quad p = 0.5 \]
Truncated differential (1)

\[ p_\ast = 1.01 \cdot 2^{-4}, \quad p = 2^{-4} \]

Truncated differential (2)

\[ p_\ast = 1.5 \cdot 2^{-15}, \quad p = 2^{-15} \]
Outline

1. Introduction

2. Algorithm to find the data complexity

3. A problem: Approximation of the involved binomial tails

4. A formula to approximate the Data Complexity

5. Asymptotic behavior for some statistical attacks
Simplified formula for the data complexity

Recall that:

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

Using Taylor series, \( \log(2\sqrt{\pi D(p^*||p)}) \) is a good estimate of \( \log(\lambda) \).

\[ N'' = -\frac{\log(2\sqrt{\pi \beta})}{D(p^*||p)}. \]

So comparing the data complexity of two statistical cryptanalyses boils down to comparing the Kullback Leibler divergences of those cryptanalyses.
Behavior of the data complexity for some statistical attacks

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

We were interested in the Data Complexity of statistical attacks. This work provides:

- an algorithm to accurately compute the DC;
- an asymptotic formula of the DC;
- the asymptotic behavior of the DC.

Perspectives:

- Key ranking.
- Using an approximation that catches the lattice behavior of the considered random variables.
- Generalizing this work to other distributions than Bernoulli.