Data Complexity and Success Probability for Various Cryptanalyses

Céline Blondeau, Benoît Gérard and Jean Pierre Tillich

INRIA project-team SECRET, France

Outline

Outline

Context

We focus on *symmetric cryptography*.

- **Block cipher**
  Block of ciphertext are function of block of plaintext and key.
- **Stream cipher**
  Bits of ciphertext are obtained by a XOR between bits of plaintext and bits of pseudo random keystream.
- **Hash function**
  Arbitrary block of data are reduced to a fixed-size bit string.

Statistical Attacks

Some statistical attacks against *stream ciphers* and *hash functions*.

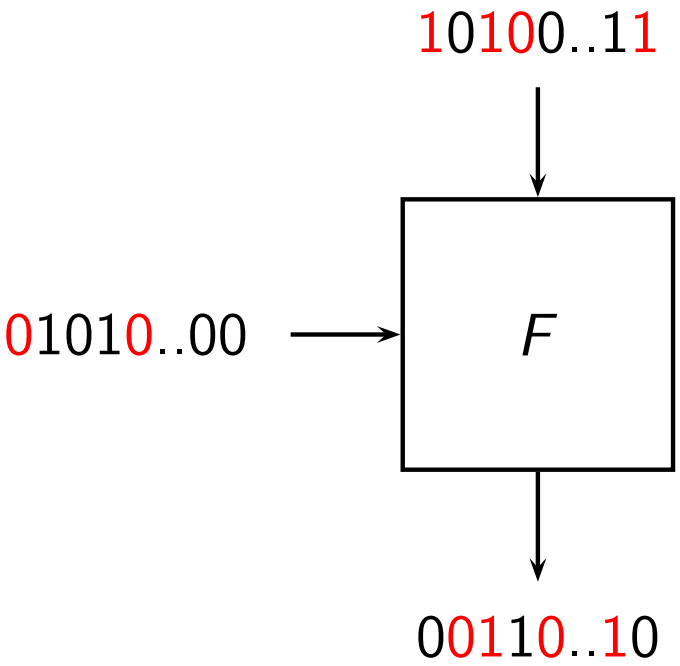
Some known statistical cryptanalyses against *block ciphers*:

- linear cryptanalysis [Matsui 93];
- differential cryptanalysis [Biham Shamir 91];
- truncated differential cryptanalysis [Knudsen 94];
- higher order differential cryptanalysis [Knudsen 94];
- impossible differential cryptanalysis [Biham Biryukov Shamir 99];
- ...

Statistical Attacks

More generally a statistical attack aims at distinguishing two probability distributions to obtain information on the key of the cipher.

Linear Cryptanalysis



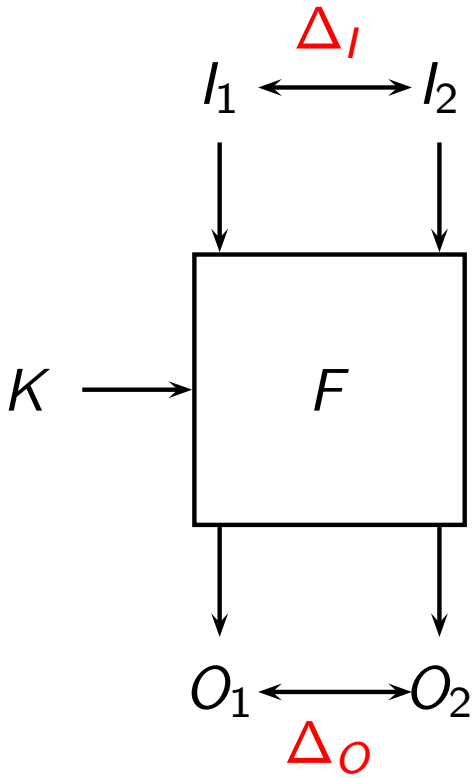
Approximation of the form:

$$\pi_I(I) \oplus \pi_K(K) = \pi_O(O)$$

Probability of a linear approximation:

$$\frac{1}{2} + \varepsilon \text{ with } \varepsilon \text{ small.}$$

Differential Cryptanalysis



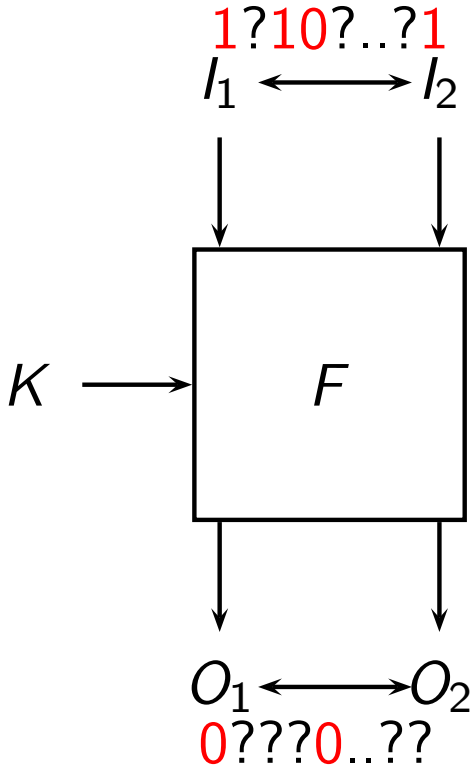
Equation of the form

$$F(x) + F(x + \Delta_I) = \Delta_O$$

Probability of a differential equation:

$$\frac{c}{2^s} \text{ with } c \text{ large.}$$

Truncated Differential Cryptanalysis



Equation of the form

$$\forall a \text{ such that } \pi_I(a) = A$$
$$\pi_O(F(x) + F(x+a)) = b$$

Probability of a truncated differential equation

$$\frac{c}{2^t} \text{ with } c \text{ small.}$$

Notation

We have **N** samples: $\mathcal{S}_1, \cdots, \mathcal{S}_N$. The composition of the sample depends on the type of cryptanalysis:

- One couple of known plaintext ciphertext (linear cryptanalysis).
- Two couples of chosen plaintext ciphertext (differential cryptanalysis).
- ...

- **Subkey:** Studied bits of the key.
- **Candidate:** A possible subkey.
- **n**: Number of candidates $k_0 ... k_{n-1}$.
- $\mathbf{k}_0$: The good candidate.

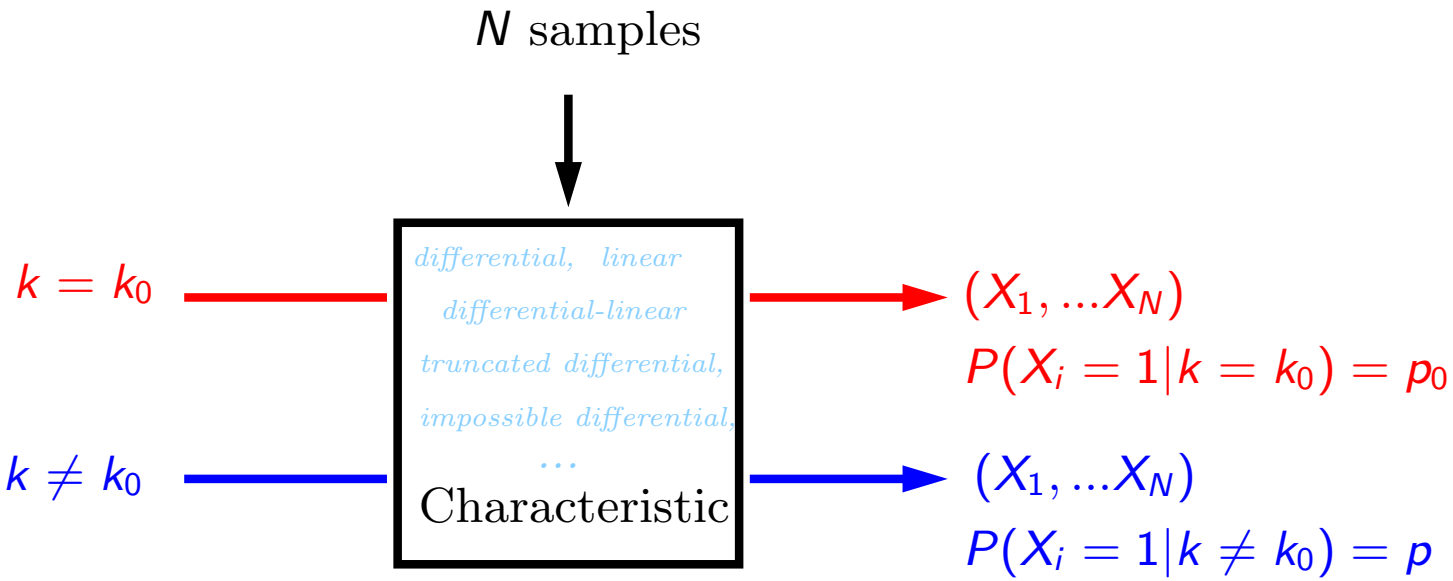
Steps of Statistical Attacks

1. **Distillation phase:** some statistic Σ is extracted from the available data.
2. **Analysis phase:** from Σ, the likelihood of each possible subkey is computed and a list $\mathcal{L}$ of the likeliest keys is suggested.
3. **Search phase:** for each subkey in $\mathcal{L}$, all the possible corresponding master keys are exhaustively tried until the good one is found.

Using a Characteristic

Let χ be some characteristic on a given cipher.

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



- $S_{k_0} = \sum_{i=1}^{N} X_i$ follows a binomial law of parameters (N, p_0).
- $k \neq k_0, \quad S_k = \sum_{i=1}^{N} X_i$ follows a binomial law of parameters (N, p).

Problematic

Important quantities in statistical cryptanalysis:

- N: Data Complexity
- P_S: Success Probability
- $\ell = \#\mathcal{L}$: Size of the list of kept keys

Aim

Evaluate the data complexity, the success probability and relate both quantities.

Paradigm

Fixed threshold (T)

$$\mathcal{L} = \{k_i, S_{k_i} > T\}.$$

- $\alpha \stackrel{\text{def}}{=} \Pr[k_0 \notin \mathcal{L}]$.
- $\beta \stackrel{\text{def}}{=} \Pr[k \neq k_0, k \in \mathcal{L}]$.

Fixed size of list (ℓ)

$$\ell \stackrel{\text{def}}{=} \#\mathcal{L},$$
$$k \in \mathcal{L}, k' \notin \mathcal{L} \Rightarrow S_k > S_{k'}$$

- $P_S \stackrel{\text{def}}{=} \Pr[k_0 \in \mathcal{L}]$.
- ℓ/n: ratio of kept keys.

Outline

Gaussian Approximation of Binomial Tails

$$P\left[S_k \leq T\right] \simeq \int_{-\infty}^{T/N} \frac{1}{\sqrt{2\pi Np(1-p)}} \cdot e^{-\frac{N(x-p)^2}{2p(1-p)}} dx$$

Classically used in linear cryptanalysis:

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

But...

... not accurate for any $N \cdot p$. For instance, when $N \cdot p$ is too small as in differential cryptanalysis [Selçuk 08].

Poisson Approximation of Binomial Tails

$$P\left[S_k \leq T\right] \simeq \sum_{i=0}^{T} e^{-Np} \cdot \frac{(Np)^i}{i!}$$

Implicitly used in differential cryptanalysis:

- [Biham Shamir 91,93];
- [Gilbert 97];
- [Selçuk 08]
- ...

But ...

... not accurate for any $N \cdot p$. For instance, when $N \cdot p$ is too big as in linear cryptanalysis.

Good Approximation of Binomial Tails

Binomial tail:

$$P\left[S_k \leq N\tau\right] = \sum_{i=0}^{\lfloor N\tau \rfloor} \binom{N}{i} p^i (1-p)^{N-i}$$

Theorem

$$P(S_k \leq N\tau) \underset{N\to\infty}{\sim} \frac{p\sqrt{1-\tau}}{(p-\tau)\sqrt{2\pi N\tau}} \cdot 2^{-N\cdot D(\tau||p)}.$$

Where the *Kullback-Leibler divergence* is defined by:

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

Comparison

		Exact	Poisson	Gaussian	Ours
Lin Crypt: $p = 0.5$, $p_0 = 0.5 + 2^{-10}$	β	$\mathbf{8.12 \cdot 10^{-5}}$	$3.84 \cdot 10^{-3}$	$\mathbf{8.12 \cdot 10^{-5}}$	$8.62 \cdot 10^{-5}$
	α	$\mathbf{2.97 \cdot 10^{-2}}$	$9.14 \cdot 10^{-2}$	$\mathbf{2.97 \cdot 10^{-2}}$	$3.58 \cdot 10^{-2}$
Diff Crypt: $p = 2^{-27}$, $p_0 = 2^{-20}$	β	$\mathbf{2.03 \cdot 10^{-3}}$	$\mathbf{2.03 \cdot 10^{-3}}$	$8.84 \cdot 10^{-5}$	$1.97 \cdot 10^{-3}$
	α	$\mathbf{3.27 \cdot 10^{-3}}$	$\mathbf{3.27 \cdot 10^{-3}}$	$6.66 \cdot 10^{-3}$	$3.33 \cdot 10^{-3}$
Trunc Diff(1): $p = 2^{-4}$, $p_0 = 1.01 \cdot 2^{-4}$	β	$\mathbf{9.29 \cdot 10^{-5}}$	$1.46 \cdot 10^{-4}$	$\mathbf{9.23 \cdot 10^{-5}}$	$9.90 \cdot 10^{-5}$
	α	$\mathbf{9.80 \cdot 10^{-5}}$	$1.55 \cdot 10^{-4}$	$\mathbf{9.89 \cdot 10^{-5}}$	$1.04 \cdot 10^{-4}$
Trunc Diff(2): $p = 2^{-15}$, $p_0 = 1.5 \cdot 2^{-15}$	β	$\mathbf{5.05 \cdot 10^{-5}}$	$\mathbf{5.06 \cdot 10^{-5}}$	$3.17 \cdot 10^{-5}$	$5.34 \cdot 10^{-5}$
	α	$\mathbf{4.37 \cdot 10^{-4}}$	$\mathbf{4.38 \cdot 10^{-4}}$	$5.45 \cdot 10^{-4}$	$4.67 \cdot 10^{-4}$

Outline

Fixed Threshold Method

Neyman-Pearson (optimal) test:

Accept a candidate k if

$$\frac{P(X_1, X_2, \ldots, X_N | k = k_0)}{P(X_1, X_2, \ldots, X_N | k \neq k_0)} > t.$$

This (likelihood) ratio only depends on $S_k = \sum_{i=1}^{N} X_i$, p_0 and p and is increasing in S_k. Thus, the acceptance condition becomes, for some threshold $0 < T < N$,

$$\text{If } S_k \geq T \text{ then } k \in \mathcal{L} \text{ else } k \notin \mathcal{L}$$

Error Probabilities

- **Non-detection error probability**

$$P(k_0 \notin \mathcal{L}) = P(S_{k_0} < T) \underset{N\to\infty}{\sim} \frac{p_0\sqrt{1-\tau}}{(p_0-\tau)\sqrt{2\pi N\tau}} 2^{-N\ D(\tau||p_0)};$$

- **False alarm error probability**

$$k \neq k_0 \quad P(k \in \mathcal{L}) = P(S_k \geq T) \underset{N\to\infty}{\sim} \frac{(1-p)\sqrt{\tau}}{(\tau-p)\sqrt{2\pi N(1-\tau)}} 2^{-N\ D(\tau||p)}.$$

$$\tau \stackrel{\text{def}}{=} T/N\text{: Relative threshold}$$

Aim

Finding N minimal such that it exists T such that $P(S_{k_0} < T) \leq \alpha$ and $P(S_k \geq T,\ k \neq k_0) \leq \beta$ for given values of α and β.

Algorithm for finding N (1/2)

Some properties:

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

Idea

Dichotomic search for τ.

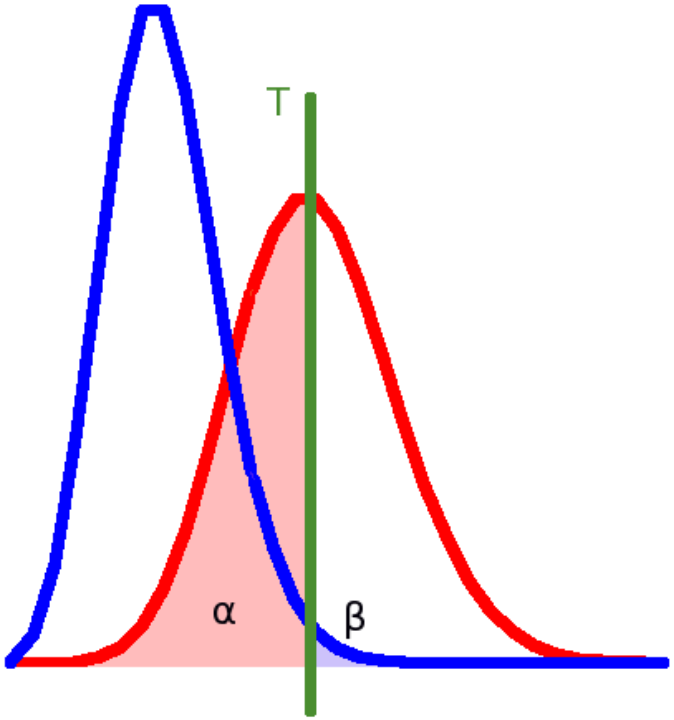
Algorithm for finding N (2/2)

Input: (α, β) and (p_0, 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 $\tau_{\max} \leftarrow p_0$.
repeat
  $\tau \leftarrow \dfrac{\tau_{\min} + \tau_{\max}}{2}$.
  Compute N_{nd} such that $\forall N > N_{\mathrm{nd}},\ P(S_{k_0} < N\tau) \leq \alpha$.
  Compute N_{fa} such that $\forall N > N_{\mathrm{fa}},\ P(S_k \geq N\tau\ ,\ k \neq k_0) \leq \beta$.
  if $N_{\mathrm{nd}} > N_{\mathrm{fa}}$ **then** $\tau_{\max} = \tau$ **else** $\tau_{\min} = \tau$
until $N_{\mathrm{nd}} = N_{\mathrm{fa}}$.
return N and τ.

Approximation of the Data Complexity (1)

Aim: Finding a simple formula to estimate the data complexity.



Fixing $\tau = p_0$ simplifies the problem.

Thus α is close to 50%.

Approximation of the Data Complexity (2)

$$\beta = \frac{(1-p)\sqrt{p_0}}{(p_0-p)\sqrt{2\pi N(1-p_0)}} 2^{-N\, D(p_0||p)}.$$

$$N = f(N) = -\frac{\log(\lambda\beta\sqrt{N})}{D\,(p_0||p)} \qquad \text{where } \lambda = \frac{(p_0-p)\sqrt{2\pi(1-p_0)}}{(1-p)\sqrt{p_0}}.$$

$$N_{i+1} = f(N_i),\ N_1 = 1$$

Theorem

$$N_3 = N' = -\frac{1}{D\,(p_0||p)}\left[\log\left(\frac{\lambda\beta}{\sqrt{D\,(p_0||p)}}\right) + 0.5\log\left(-\log(\lambda\beta)\right)\right]$$

Approximation of the Data Complexity (3)

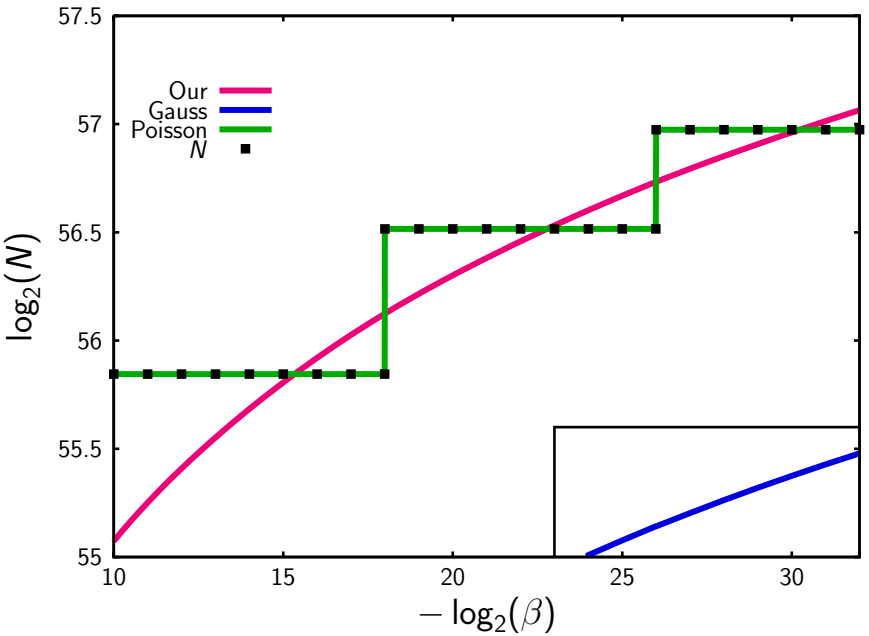
Theorem

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

with $\theta = \left[1 + \frac{1}{2\log(\lambda\beta)} \log\left(-\frac{\log(\lambda\beta)}{D(p_0||p)}\right)\right]^{-1}$.

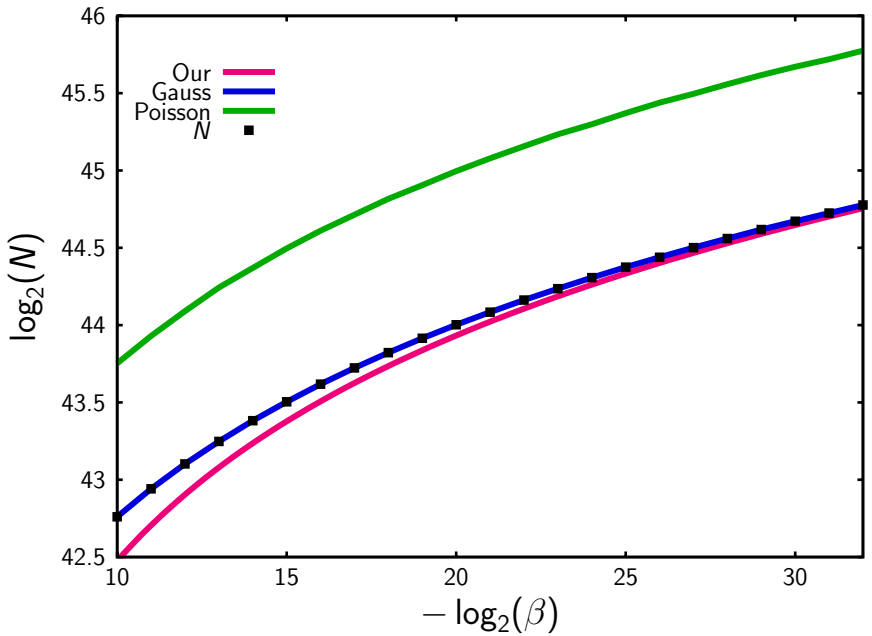
This is a good approximation of N when β tends to 0.

Experimental Results (1)



Differential cryptanalysis of DES

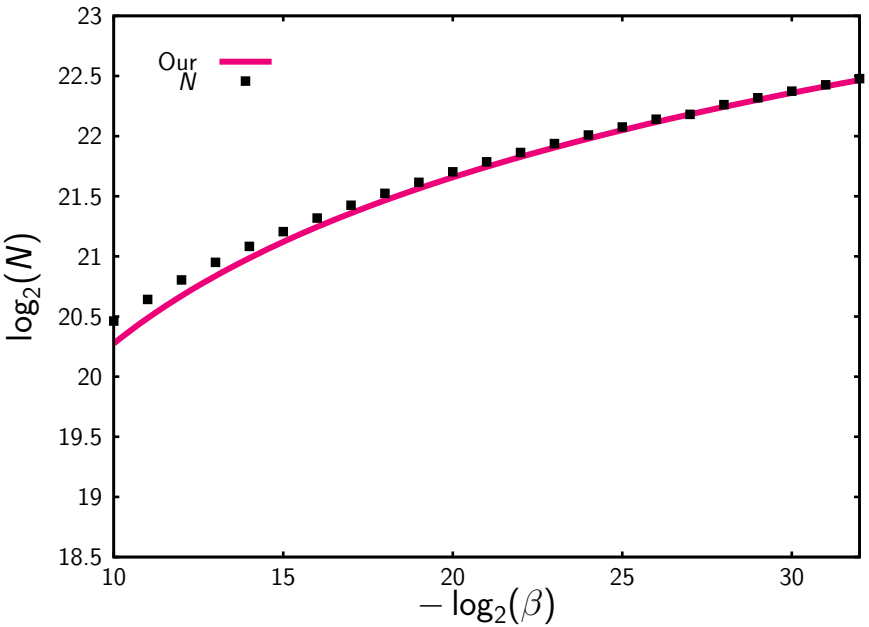
$p_0 = 1.87 \cdot 2^{-56}$, $p = 2^{-64}$



Linear cryptanalysis of DES

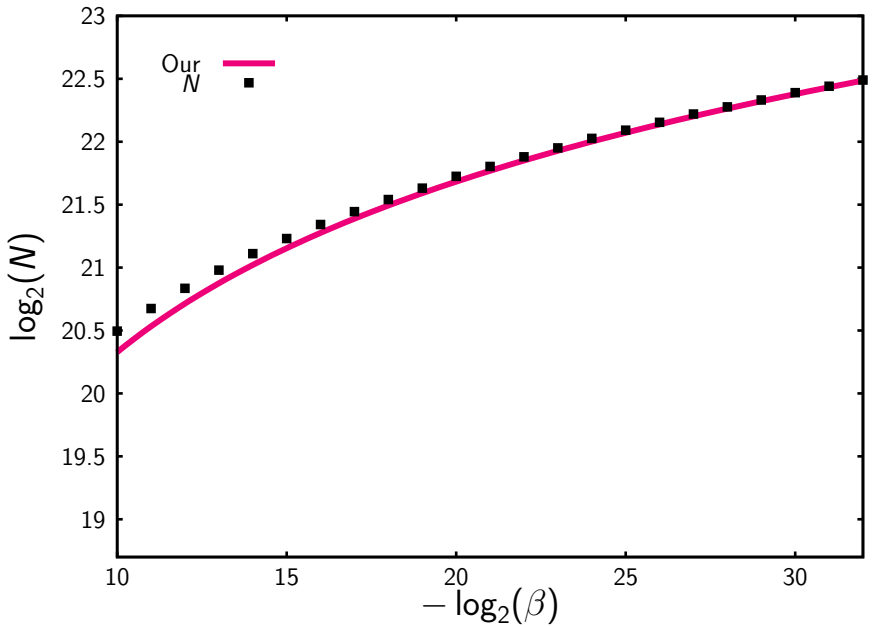
$p_0 = 0.5 + 1.19 \cdot 2^{-21}$, $p = 0.5$

Experimental Results (2)



Truncated differential (1)

$p_0 = 1.01 \cdot 2^{-4}$, $p = 2^{-4}$



Truncated differential (2)

$p_0 = 1.5 \cdot 2^{-15}$, $p = 2^{-15}$

Simplified Formula for the Data Complexity

Recall that:

$$N' = -\frac{1}{D\left(p_0||p\right)}\left[\log\left(\frac{\lambda\beta}{\sqrt{D\left(p_0||p\right)}}\right) + 0.5\log\left(-\log(\lambda\beta)\right)\right]$$

Using Taylor series, $\frac{1}{\sqrt{D\left(p_0||p\right)}} \approx \frac{2\sqrt{\pi}}{\lambda}$.

Theorem

$$N'' \approx N' \quad N'' = -\frac{\log(2\sqrt{\pi}\beta)}{D\left(p_0||p\right)}.$$

Rule

The best cryptanalysis is the one with the largest $D\left(p_0||p\right)$.

Behavior of the Data Complexity for Statistical Attacks (1)

Attack	Classical results	$\frac{1}{D(p_0 \| \| p)}$
Linear	$\frac{1}{(p_0 - p)^2}$	$\frac{1}{2(p_0 - p)^2}$
Differential	$\frac{1}{p_0}$	$\frac{1}{p_0 \log_2(p_0/p) - p_0}$
Differential-linear	$\frac{1}{(p_0 - p)^2}$	$\frac{1}{2(p_0 - p)^2}$

Behavior of the Data Complexity for Statistical Attacks (2)

Attack	Classical results	$\frac{1}{D(p_0 \| \| p)}$
Truncated differential	unknown	$\frac{2p}{(p_0 - p)^2}$
Impossible differential	implicitly: $\frac{1}{p}$	$\frac{1}{p}$
k-th order differential	1	$-\frac{1}{\log_2 p}$

Outline

Notation

We have n possible candidates:

- k_0, the correct one;
- $k_1, \ldots, k_{n-1}$.

For all random variables, $S_{k_j} \sim \begin{cases} \mathcal{B}(N, p_0) \text{ if } j = 0, \\ \mathcal{B}(N, p) \text{ if } j \neq 0. \end{cases}$

$$P_s = P(k_0 \in \mathcal{L})$$

Order Statistic Tools

We want to keep a list $\mathcal{L}$ of size ℓ which contains the most likely candidates.

$$\forall k \notin \mathcal{L}, \quad \forall k' \in \mathcal{L} \qquad S_k < S_{k'}$$

We sort $S_{k_1}, \ldots, S_{k_{n-1}}$ in decreasing order $\longrightarrow S^*_{k_1}, \ldots S^*_{k_{n-1}}$.

$$P_S = P[k_0 \in \mathcal{L}] = P[S^*_{k_\ell} < S_{k_0}].$$

[Selçuk 2008]'s Result

Using a normal approximation:

$$S_{k_0} \approx \widetilde{S}_{k_0} \overset{\text{def}}{\sim} \mathcal{N}(Np_0, Np_0(1-p_0)), \qquad \text{density } \varphi_0(t)$$

$$\forall k \neq k_0, \quad S_k \approx \widetilde{S}_k \overset{\text{def}}{\sim} \mathcal{N}(Np, Np(1-p)), \qquad \text{density } \varphi(t)$$

$$P_S \approx \int_{\Phi^{-1}(1-\ell/n)}^{\infty} \varphi_0(t)\, dt.$$

$$\Phi(u) \overset{\text{def}}{=} \int_{-\infty}^{u} \varphi(t)\mathrm{dt},$$

Our Result

Without approximation:

$$S_{k_0} \sim \mathcal{B}(N, p_0), \qquad k \neq k_0 \quad S_k \sim \mathcal{B}(N, p).$$

Theorem

$$P_S \approx \sum_{i=F^{-1}(1-(\ell-1)/(n-2))}^{N} P[S_{k_0} = i].$$

Where

$$F^{-1}(y) \stackrel{\text{def}}{=} \min_{x \in \mathbb{N}} \{P[S_{k \neq k_0} \leq x] \geq y\}.$$

Comparison

Using a normal approximation:

$$P_S \approx P(\widetilde{S}_{k_0} > \tau) \quad \text{where} \quad P(\widetilde{S}_k < \tau) = 1 - \frac{\ell}{n}.$$

Without approximation:

$$P_S \approx P(S_{k_0} > \tau) \quad \text{where} \quad P(S_k < \tau) = 1 - \frac{\ell - 1}{n - 2}.$$

Experimental Result and Comparison with Selçuk

Type of cryptanalysis	Parameters $N = 2^{48}$ $n = 2^{20}$	Experimental results	Ours	Selcuk
Linear	$\ell = 2^{15}$	**86.81**	86.81	86.81
Linear	$\ell = 2^{10}$	**45.33**	45.33	45.33
Differential	$\ell = 2^{15}$	**82.57**	82.47	90.50
Differential	$\ell = 2^{10}$	**82.50**	82.47	90.50

Linear: $p = 0.5$ and $p_0 = 0.5 + 1.49 \cdot 2^{-24}$

Differential: $p = 2^{-64}$ and $p_0 = 2^{-47.2}$

Sketch of Proof (1/2)

$S^*_{k_1}, \ldots S^*_{k_{n-1}}$: Order statistics

$$P_S = P[S^*_{k_\ell} < S_{k_0}].$$

For $j \neq 0$, $F(x) \stackrel{\text{def}}{=} P[S_{k_j} \leq x]$.
The function F is increasing:

$$P_S = P\left[F(S^*_{k_\ell}) < F(S_{k_0})\right].$$

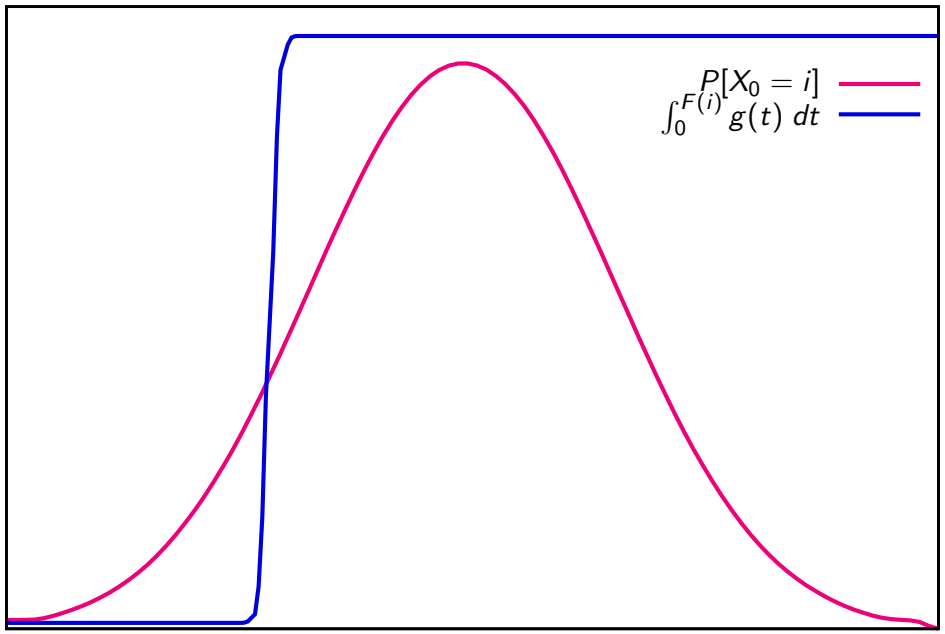
And,

$$F(S_{k_j}) \sim \mathcal{U}([0;1]) \Longrightarrow F(S^*_{k_\ell}) \sim \mathrm{Beta}(n-\ell-1, \ell-1).$$

Sketch of Proof (2/2)

$$g(t) \stackrel{\text{def}}{=} (n-1) \cdot \binom{\ell-1}{n-2} t^{n-\ell-1}(1-t)^{\ell-1}.$$

$$P_S = \sum_{i=0}^{N} P[S_{k_0} = i] \int_0^{F(i)} g(t)\, dt.$$



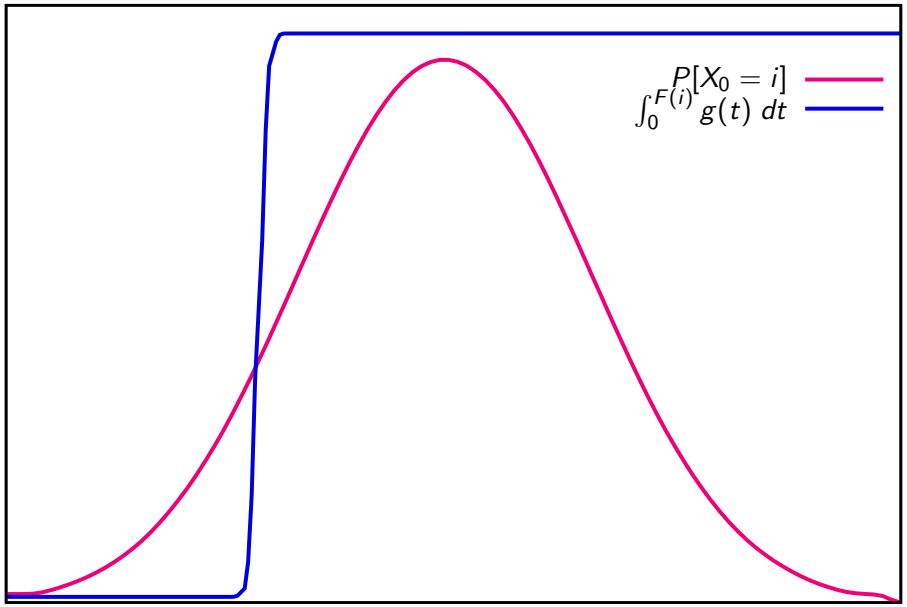
- The maximum of g is reached in

$$t_0 = 1 - \frac{\ell-1}{n-2}.$$

- Step in the figure at point

$$i = F^{-1}(t_0).$$

Error Term



Selçuk: no error term

$$\delta \stackrel{\text{def}}{=} \sum_{i=0}^{F^{-1}\left(1-\frac{\ell-1}{n-2}\right)-1} P[S_{k_0} = i]$$

Theorem

$$P_S = 1 - \delta + O\left(\delta\sqrt{\frac{\ln(\ell/\delta^2)}{\ell}} + \frac{1}{\ell^2} + \frac{1}{n}\right)$$

Outline

Data Complexity

With *fixed threshold* method, we obtain:

- An accurate algorithm
- A formula with an error estimate

$$N' = -\frac{1}{D\left(p_0||p\right)}\left[\log\left(\frac{\lambda\beta}{\sqrt{D\left(p_0||p\right)}}\right) + 0.5\log\left(-\log(\lambda\beta)\right)\right],$$

- A simpler formula

$$N'' = -\frac{\log(2\sqrt{\pi}\beta)}{D\left(p_0||p\right)}.$$

Success Probability

With *a list of fixed size* method, we obtain:

- a generalization of Selçuk's formula;
- an error term.

$$P_S = 1 - \delta + O\left(\delta\sqrt{\frac{\ln(\ell/\delta^2)}{\ell}} + \frac{1}{\ell^2} + \frac{1}{n}\right).$$

$$\delta \stackrel{\text{def}}{=} \sum_{i=0}^{F^{-1}\left(1-\frac{\ell-1}{n-2}\right)-1} P[S_{k_0} = i],$$

Link

- $\beta \longrightarrow$ probability to kept a bad key.
- $\ell/n \longrightarrow$ ratio of kept keys.

Idea

Taking N with $N = -c \cdot \dfrac{\log\left(2\sqrt{\pi}\,\dfrac{\ell}{n}\right)}{D\left(p_0||p\right)}$.

Link between P_S and N

We notice that using N of the previous form, gives us that P_S only depends on the constant c.

Experimental Results

	$c = 1$				$c = 1.5$			
	ℓ				ℓ			
	2^{10}	2^{14}	2^{18}	2^{22}	2^{10}	2^{14}	2^{18}	2^{22}
L1	60.11	60.38	60.58	60.47	**91.69**	89.61	86.70	82.20
L2	60.11	60.38	60.60	60.48	**91.69**	89.61	86.71	82.20
TD1	**59.81**	60.91	60.83	61.24	91.05	89.21	86.15	**82.01**
TD2	59.81	60.91	62.63	61.24	91.05	89.81	86.99	83.29
D1	68.16	61.92	73.79	67.49	88.75	90.21	82.65	86.39
D2	68.17	61.93	**73.80**	67.65	88.75	90.21	82.65	86.39

Success probability express in percent for:
Linear, **D**ifferential and **T**runcated **D**ifferential cryptanalysis

Conclusion

Removing the Gaussian or Poisson assumption, this work provides:

- An accurate algorithm for computing the data complexity;
- The asymptotic formula of the data complexity;
- The asymptotic behavior of the data complexity for some statistical attacks;
- A general formula expressing the success probability;
- A link between the data complexity and the success probability.

Perspectives:

- Generalizing this work to other distributions than Bernoulli;
- Studying stream ciphers and hash functions.