The Conditional Correlation Attack:

A Practical Attack on Bluetooth Encryption

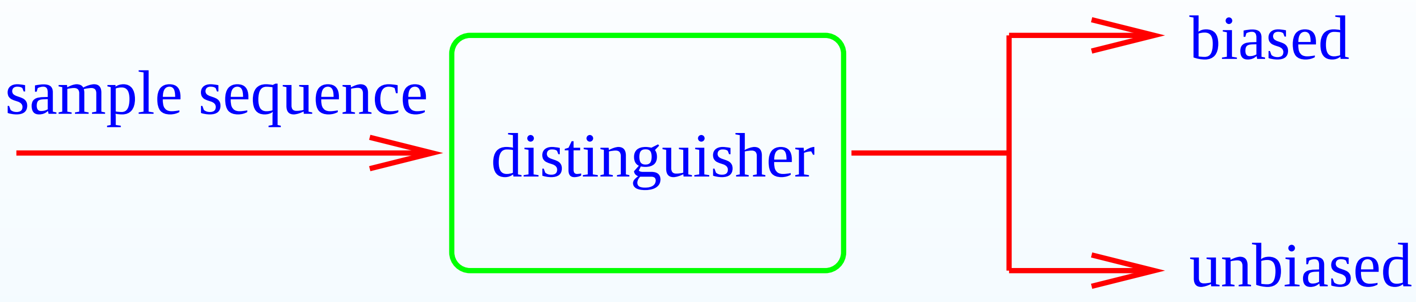
Yi Lu — EPFL

Willi Meier — FH Aargau

Serge Vaudenay — EPFL

Background

Simple Distinguisher



Sample length=1 bit:

$$\min \#\text{samples} = O\left(\frac{1}{\epsilon^2}\right),$$

where the bias $\epsilon \stackrel{\text{def}}{=} \Pr(\text{sample} = 1) - \Pr(\text{sample} = 0)$.

Simple Distinguisher

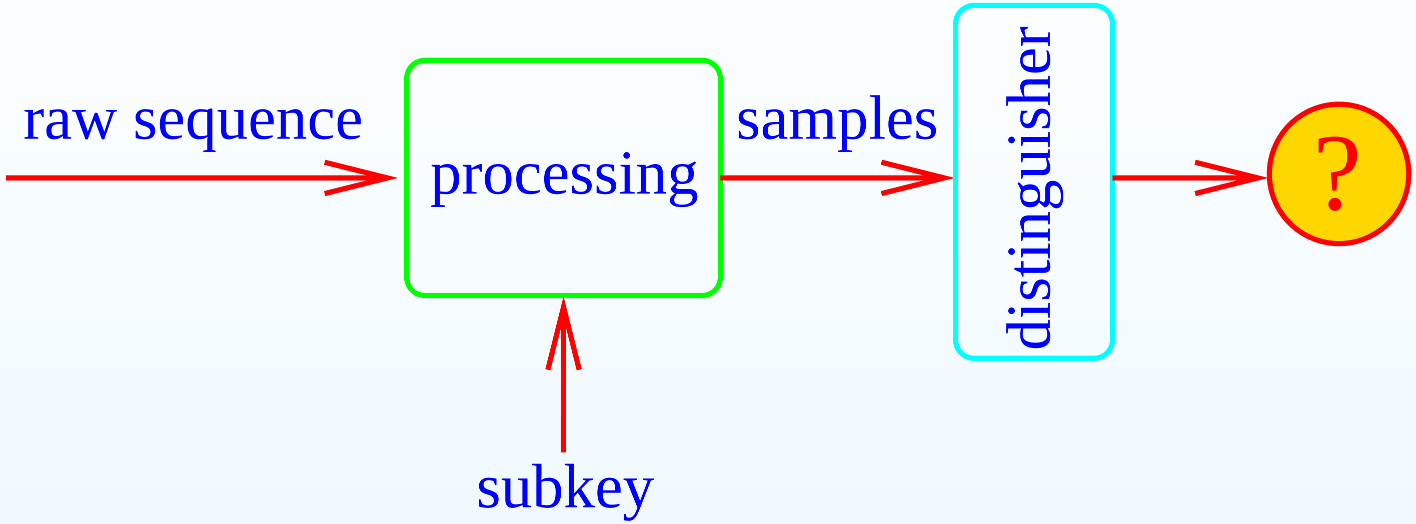
Sample length=r bits [BJV'04]:

$$\min \#\text{samples} = O\left(\frac{1}{\Delta(D)}\right),$$

where the Squared Euclidean Imbalance of the sample distribution D is defined by

$$\Delta(D) = 2^r \sum_a \left(D(a) - 2^{-r}\right)^2.$$

Distinguisher with Key-recovery



[BJV'04]:

- Assume the right key (resp. wrong key) transforms the raw sequences into biased (resp. unbiased) samples;
- to successfully recover L-bit key deterministically,

$$\min \#\text{samples} = \frac{4L \ln 2}{\Delta(D)}.$$

Distinguisher & Correlation Attack

- In correlation attacks,
  - raw sequence: output of LFSR-based keystream generators
  - correlation: biased relation between keystream and certain LFSR output sequence(s)
  - subkey: state(s) of a subset of involved LFSR(s)
  - subkey processing: linear transformation
- The distinguisher is used to solve the MLD problem.
- The distinguisher can be either
  - (often) probabilistic (eg, in fast correlation attacks), or
  - (rarely) deterministic

  depending on the key size L.

Conditional Correlation Attack

Related Work

Prior to our work,

- R. Anderson (FSE'94) initiated the work of conditional correlation attacks on the nonlinear filter generator.
- The notion of conditional correlation was formalized by Lee et al. (ASIACRYPT'96):

$$\text{given } X_0, Y_0 \Longrightarrow \left|\Pr(X \cdot X_0 = 0 | f(X) = Y_0) - 0.5\right| .$$

- Löhlein'03 extended conditional correlations and studied efficient attacks.

  However, the basic concept of conditional correlations remains the same: the linear correlation of the inputs conditioned on a given output pattern of a nonlinear function.

Our Problem

- We studied the correlation of the output of a function conditioned on the unknown (partial) input which is uniformly distributed.

- Given
  - a function $f(\mathcal{B}, X)$
  - n i.i.d. samples of pairs $(f_{\mathcal{B}}(X), \mathcal{B})$

  Q: What is the minimum n to spot above sequence from truly random sequences of equal length?

- Application: $\mathcal{B}$ is the key-related material, our problem is interesting in related-key attacks.

Smart Distinguisher

- 2^L sample sequences: $(Z_i^K, \mathcal{B}_i^K)$ for $i \in [1, n]$ and L-bit K.
- Related results:
  - [GBM'02] (conditional correlations): for $|Z| = 1$ bit,

$$n = \frac{2L}{\mathsf{E}[\Delta(f_{\mathcal{B}})]}.$$

  - [BJV'04] (unconditional correlations): for $|Z| \geq 1$ bit and sample sequences do not include $\mathcal{B}$'s,

$$n = \frac{4L \ln 2}{\Delta(f)}.$$

- Our theoretical result: based on [BJV'04], the deterministic smart distinguisher that maximizes $\prod_{i=1}^{n} D_{f_{\mathcal{B}_i^K}}(Z_i^K)$ solves our problem with time $O(n \cdot 2^L)$ and

$$n = \frac{4L \ln 2}{\mathsf{E}[\Delta(f_{\mathcal{B}})]}.$$

Optimal Smart Distinguisher

If $\mathcal{B}_i^K$'s and Z_i^K's exhibit special structures:

- computing $\prod_{i=1}^{n} D_{f_{\mathcal{B}_i^K}}(Z_i^K)$ reduces to computing convolution;
- thanks to Fast Walsh Transform, an optimal smart distinguisher is achieved within time

$$O(n + L \cdot 2^{L+1}),$$

after one-time precomputation $O(L \cdot 2^L)$.

Conditional Correlation & Regular Correlation

Property 1 *We have*

$$\mathsf{E}[\Delta(f_{\mathcal{B}})] \geq \Delta(f),$$

where equality holds iff $D_{f_{\mathcal{B}}}$ is independent of $\mathcal{B}$.

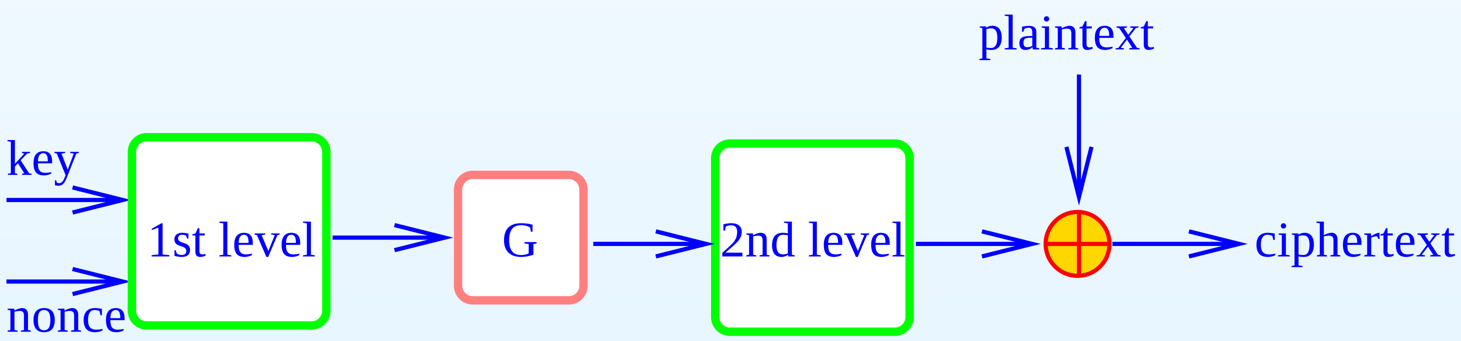
Comments:

- The conditional correlation is no smaller than the unconditional correlation.
- In particular, even if the traditional distinguisher fails with $\Delta(f) = 0$, the smart distinguisher would still work as long as $D_{f_{\mathcal{B}}}$ is dependent on $\mathcal{B}$ (i.e. $\mathsf{E}[\Delta(f_{\mathcal{B}})] > 0$).

Application to Attacking Bluetooth Encryption

About Bluetooth Encryption

- Encryption key size is a multiple of 8 and ranges over $\{8, 16, 24, \ldots, 128\}$.
- The keystream length is limited up to 2745 bits per frame.
- Uses a two-level reinitialization scheme.
- One secret key can be reinitialized for up to 2^{26} frames.



Known Attacks

- guess & determine:

  [Saarinen'00], [FL'01], [Fluhrer'02]

- algebraic attack:

  [Krause'02], [AK'03], [Courtois'03], [ALP'04]

- correlation attack:

  [HN'99], [GBM'02], [LV'04a], [LV'04b]

Known Correlations: Preliminaries

- For any ℓ,

$$f: \quad \underset{\text{LFSR input weights,}}{\underset{\Uparrow}{B = B_1 B_2 \cdots B_\ell,}} \quad \underset{\text{FSM state}}{\underset{\Uparrow}{X}} \quad \mapsto \quad \underset{\text{FSM outputs}}{\underset{\Uparrow}{Z = c_0^0 \cdots c_{\ell+1}^0}}$$

- For any $(\ell + 2)$-bit binary vector α,

$$f^\alpha(B, X) \stackrel{\text{def}}{=} \alpha \cdot f(B, X),$$

and B is considered to be partial input.

Known (Unconditional) Correlations

- One-level E0 ([HN'99], [EJ'00], [GBM'02], [LV'04a]): notable biases up to 26 bits are

α	1,1,0,1	1,0,1,1	1,1,1,1,1	1,0,0,0,0,1
$\lvert \text{bias}(f^\alpha) \rvert$	0	2^{-4}	$\approx 2^{-3.3}$	$\approx 2^{-3.3}$

- Two-level E0 [LV'04b]: at some specific positions of the header of the keystream,

$$\text{bias}(F^\alpha) = \text{bias}^4(f^\alpha) \cdot \text{bias}(f^{\bar{\alpha}}),$$

for any α of at most 8 bits, where $\bar{\alpha}$ is the vector in reverse order of α. Notable biases up to 8 bits are

α	1,1,0,1	1,0,1,1	1,1,1,1,1	1,0,0,0,0,1
$\lvert \text{bias}(F^\alpha) \rvert$	0	0	$\approx 2^{-3.3\times 5}$	$\approx 2^{-3.3\times 5}$

Conditional Correlations

- One-level E0:

α	1,1,0,1	1,0,1,1	1,1,1,1,1	1,0,0,0,0,1
$\Delta(f^\alpha)$	0	2^{-8}	$\approx 2^{-6.7}$	$\approx 2^{-6.7}$
$\mathsf{E}[\Delta(f^\alpha_B)]$	2^{-3}	2^{-4}	$\approx 2^{-2.9}$	$\approx 2^{-2.5}$

- Two-level E0: for any α of at most 8 bits,

$$\mathsf{E}[\Delta(F^\alpha_{\mathcal{B}})] = \mathsf{E}^4[\Delta(f^\alpha_B)] \cdot \Delta(f^{\bar{\alpha}}).$$

α	1,1,0,1	1,0,1,1	1,1,1,1,1	1,0,0,0,0,1
$\Delta(F^\alpha)$	0	0	$2^{-33.5}$	$2^{-33.5}$
$\mathsf{E}[\Delta(F^\alpha_{\mathcal{B}})]$	2^{-20}	0	$2^{-18.3}$	$2^{-16.7}$
$\log_2 \lvert\mathcal{B}\rvert$	33	33	49	65

Experiments

- With

$$\alpha = (1,1,0,1) \text{ and } n = \frac{4L\ln 2}{\mathsf{E}[\Delta(F_B^\alpha)]} \approx 2^{26} \text{ frames,}$$

- Experiments allow to discover: 256 33-bit subkeys always have the same rank (i.e. the 25-bit subkey). This can halve the run-time.

Table 1: Experiment Settings

CPU	RAM	HD	OS	Compiler
2.4G	2G	128G (32M/s)	LINUX	GCC

Table 2: Partial Key Recovery Attack Results

PreComp.	Run Time	#Tests	$\text{Prob}_{\text{Success}}$
37Hr	19Hr	30	100%

Full Attack

- In the same spirit of [LV'04a], more sophisticated techniques allow to use multi-biases to reduce data complexity to $2^{23.8}$ frames.

Table 3: Attack Comparison to Recover 128-bit Key

Attack	PreC.	Time	Frames	Data	Space
FL'01	-	2^{73}	-	2^{43}	2^{51}
F'02	2^{80}	2^{65}	2	$2^{12.4}$	2^{80}
GBM'02	2^{80}	2^{70}	45	2^{17}	2^{80}
LV'04b	-	2^{40}	2^{35}	$2^{39.6}$	2^{35}
Ours (A)	2^{38}	2^{38}	$2^{26.5}$	$2^{31.1}$	2^{33}
Ours (B)	2^{38}	2^{38}	$2^{23.8}$	$2^{28.4}$	2^{33}

Conclusion

- Based on conditional correlations ([Anderson'94], [Lee et al'96], [Löhlein'03]) and the generalized distinguisher [BJV'04], we have further generalized conditional correlations and studied a general statistical model for dedicated key-recovery distinguishers.
- The application leads to a practical known-plaintext attack on Bluetooth encryption.
- It remains to be a big challenge to investigate the redundancy in the header of each frame for a practical ciphertext-only attack on Bluetooth encryption.