Aalto University
School of Science

Linear Cryptanalysis of Long-Key Iterated Cipher with Applications to Permutation-Based Ciphers

Kaisa Nyberg

Aalto University School of Science
kaisa.nyberg@aalto.fi

ESC 2017
Luxemburg January 2017

Outline

Outline

Introduction

- Linear cryptanalysis on block ciphers based on iterated structures
- Theory for long-key block ciphers well developed [Nyberg 1994, Daemen 1994, Daemen-Rijmen 2006, Blondeau-Nyberg 2015]
- Statistics estimates based on signal-noise model [Bogdanov-Tischhauser 2013]
- Key-schedulings that imitate long-key properties [Leander 2016, Blondeau-Nyberg 2015]
- Focus of this presentation: extreme key-schedulings such as known-key ciphers
- We discuss how to apply known properties of long-key ciphers to permutation based ciphers such as EM ciphers

Permutation-Based Cipher

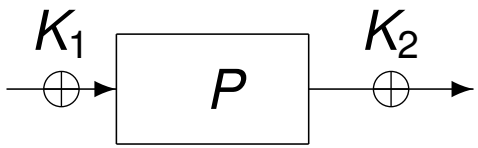


Figure : 1-EM cipher with permutation P and encryption key (K_1, K_2)

Key-recovery attacks use some property over P.

Problem. How to determine and evaluate a property over P?

We present an approach to answer this problem for linear cryptanalysis and a permutation based on an iterated permutation.

We make use of the noise-based statistical model of correlation.

Outline

Iterated Long-Key Cipher

$$E'(x; k_0, k_2 \ldots k_{r'}) = g_{r'}(\cdots(g_2(g_1(x + k_0) + k_1) + k_2)\cdots) + k_{r'}$$

where $x,\ k_0, k_1, \ldots, k_{r'} \in \mathbb{F}_2^n$ and $g_i : \mathbb{F}_2^n \to \mathbb{F}_2^n$

Iteration over r' rounds

Key Scheduling

Any r'-round key-alternating cipher in $\mathbb{F}_2^n$ can be seen as an application of such E using key-scheduling

$$\mathsf{KS}' : \mathbb{F}_2^L \to \mathbb{F}_2^{(r'+1)n}$$

which is an injective function that maps the initial key $K \in \mathbb{F}_2^L$ to $(k_0, k_1, \ldots, k_{r'})$.
Then the cipher BC' can be presented as

$$\mathsf{BC}'(X, K) \mapsto E'(x; \mathsf{KS}'(K))$$

where X is the plaintext.

Permutation by Iteration

Any r'-round permutation in $\mathbb{F}_2^n$

$$P(x) = g_{r'}(\cdots(g_2(g_1(x)))\cdots)$$

can also been seen as an application of a long-key cipher

$$E'(X; k_0, k_2, \ldots, k_{r'}) = g_{r'}(\cdots(g_2(g_1(X + k_0) + k_1) + k_2)\cdots) + k_{r'}$$

by setting $k_0 = k_1 = \ldots = k_{r'} = 0$

Examples

- Typical key-alternating ciphers: (DES), AES, PRESENT, Simon,...
- Permutation-based ciphers: EM constructions, in practice, based on
  - dedicated large permutations, or
  - cipher with known fixed key.

Example. Key-schedule of single-key 1-EM where permutation based on iterated long-key cipher

$$\mathsf{KS}'(K) = (K_1, 0, 0, \ldots, 0, K_2)$$

Outline

Key-Recovery Setting

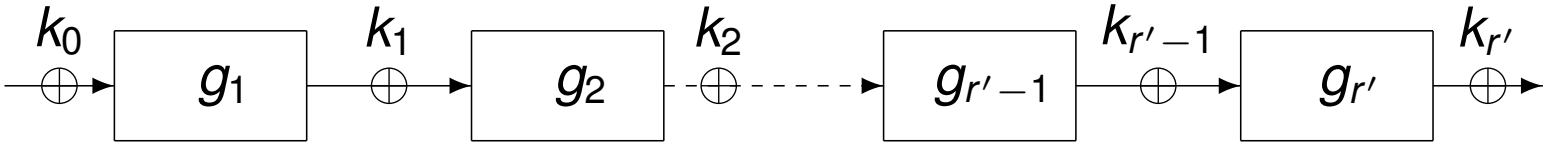


Figure : Key-alternating block cipher of r' rounds with round functions g_i and expanded encryption key $(k_0, k_1, \ldots, k_{r'})$

Key guesses over some first and last rounds. Then the long-key cipher E' is reduced to r rounds; denote it by E.

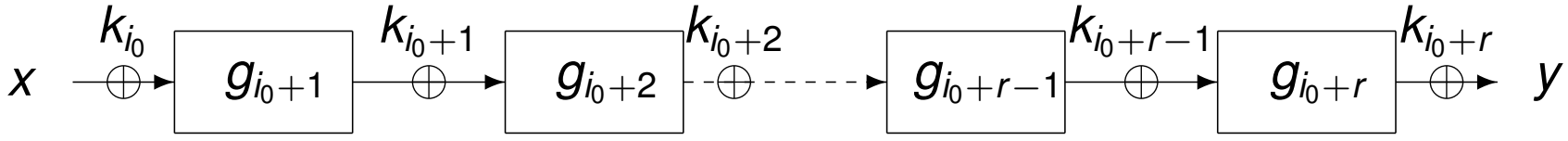


Figure : Property of key-alternating block cipher over r rounds with round functions g_i and expanded encryption key $(k_{i_0}, k_{i_0+1}, \ldots, k_{i_0+r})$

Linear Property

u	n-bit mask on x
v	n-bit mask on y
BC	r rounds of BC′
KS	r rounds of KS′

Known linear property $u \cdot x + v \cdot y$ with correlation $c_{\mathrm{BC}}(K)$ where

$$c_{\mathrm{BC}}(K) = \#\{\, x \mid u \cdot x + v \cdot \mathrm{BC}(x, K) = 0 \,\}$$

Observed Correlation

Given D, a data sample of size N of pairs (x, y), we call

$$\hat{c}(D,K) = \frac{2}{N}\#\{(x,y)) \in D \,|\, u \cdot x + v \cdot y = 0\} - 1 = \frac{2}{N} Z(D,K) - 1$$

the observed correlation where we denoted

$$Z(D,K) = \#\{(x,y)) \in D \,|\, u \cdot x + v \cdot y = 0\}$$

For any fixed key K,

$$Z(D,K) \sim \mathcal{B}(N, p(K)),$$

where $p(K)$ is some apriori probability. By the normal approximation of the binomial distribution, we obtain that for any fixed K

$$\hat{c}(D,K) \sim \mathcal{N}(c(K), \frac{1}{N}(1 - c(K)^2)) \approx \mathcal{N}(c(K), \frac{1}{N})$$

where $c(K) = 2p(K) - 1$.

Cipher and Random

1. $y = \mathrm{BC}(x, K) = E(x; \mathrm{KS}(K))$ (cipher): Then $c(K) = c_{\mathrm{BC}}(K)$ The parameters $\mathrm{Exp}_K c_{\mathrm{BC}}(K)$ and $\mathrm{Var}_K c_{\mathrm{BC}}(K)$ must be determined from the cipher in offline analysis.
2. $y \neq E(x; \mathrm{KS}(K))$ (random):

   Wrong-key randomization hypothesis: for each K, the bits $u \cdot x + v \cdot y$ are computed from a random linear approximation, that is, see [Daemen-Rijmen 2006],

$$\begin{aligned} \mathrm{Exp}_K c(K) &= 0 \\ \mathrm{Var}_K c(K) &= 2^{-n}. \end{aligned}$$

Statistics of Observed Correlation

$$\hat{c}(D,K) - c(K) \sim \mathcal{N}\left(0, \frac{1}{N}\right), \text{ for any fixed } K.$$

In the right key (cipher) case, the distribution of the observed correlation has parameters

$$\begin{aligned} \mathrm{Exp}_{D,K}\hat{c}(D,K) &= \mathrm{Exp}_K c_{\mathrm{BC}}(K) \\ \mathrm{Var}_{D,K}\hat{c}(D,K) &= \frac{1}{N} + \mathrm{Var}_K c_{\mathrm{BC}}(K) \end{aligned}$$

For random

$$\hat{c}(D,K) \sim \mathcal{N}\left(0, \frac{1}{N} + 2^{-n}\right)$$

Outline

Linear Approximations and Correlations

u	n-bit mask on x
v	n-bit mask on y
τ	L-bit mask on the key K

Given r rounds of a cipher

$$(x, K) \mapsto E(x; \mathrm{KS}(K))$$

we define

$$c(u, \tau, v) = \mathrm{cor}_{x,K}(u \cdot x + v \cdot E(x; \mathrm{KS}(K)) + \tau \cdot K).$$

This is the "trail correlation" for trail τ. Then

$$c(u, \tau, v) = 2^{-L} \sum_{K} (-1)^{\tau \cdot K} \mathrm{cor}_{x}(u \cdot x + v \cdot E(x; \mathrm{KS}(K))).$$

Correlation via Trail Correlations

We have

$$c_{\mathrm{BC}}(K) = \mathrm{cor}_x(u \cdot x + v \cdot E(x; \mathrm{KS}(K))$$

Taking the inverse Fourier transform we get

$$c_{\mathrm{BC}}(K) = \sum_{\tau} (-1)^{\tau \cdot K} c(u, \tau, v)$$

where the trail correlations $c(u, \tau, v)$ are independent of K, but hard to evaluate for a general iterated block cipher.

But trail correlations can be evaluated for the corresponding long-key cipher.

Trail Correlations from Long-Key Cipher

$$c(u,\tau,v) = \mathrm{cor}_{x,K}(u\cdot x + v\cdot E(x;\mathrm{KS}(K)) + \tau\cdot K) =$$

$$\sum_W \mathrm{cor}_K(\tau\cdot K + W\cdot \mathrm{KS}(K))\mathrm{cor}_{x,k_{i_0}\ldots k_{i_0+r}}(u\cdot x + v\cdot E(x;k_{i_0}\ldots k_{i_0+r}) + W\cdot(k_{i_0}\ldots k_{i_0+r}$$

$$= \sum_W \mathrm{cor}_K(\tau\cdot K + W\cdot \mathrm{KS}(K))c_E(u,W,v),$$

where

$$c_E(u,W,v) = \mathrm{cor}_{x,k_{i_0}\ldots k_{i_0+r}}(u\cdot x + v\cdot E(x;k_{i_0}\ldots k_{i_0+r}) + W\cdot(k_{i_0}\ldots k_{i_0+r}))$$

are the trail correlations of the iterated "long-key" cipher E over r rounds with $(r+1)n$-bit masks W.

Correlation for Cipher BC

This gives

$$
\begin{aligned}
c_{\mathrm{BC}}(K) &= \sum_{\tau} (-1)^{\tau \cdot K} c(u, \tau, v) \\
&= \sum_{\tau, W} (-1)^{\tau \cdot K} \mathrm{cor}_Z(\tau \cdot Z + W \cdot \mathrm{KS}(Z)) c_E(u, W, v) \\
&= \sum_{W} (-1)^{W \cdot \mathrm{KS}(K)} c_E(u, W, v)
\end{aligned}
$$

Outline

Trails over Long-Key Cipher E

$$c_{\mathrm{BC}}(K) = \sum_{W} (-1)^{W \cdot \mathrm{KS}(K)} c_E(u, W, v)$$

where

$$c_E(u, W, v) = \prod_{j=1}^{r} \mathrm{cor}_z \left(w_{i_0+j-1} \cdot z + w_{i_0+j} \cdot g_{i_0+j}(z) \right)$$

$W = (w_{i_0}, w_{i_0+1}, \ldots, w_{i_0+r})$, and and $u = w_{i_0}$ and $v = w_{i_0+r}$

Use Matsui's algorithm to search for such trails over E that have correlation $c_E(u, W, v)$ of high absolute value.

Modelling the Correlation

Noise-based approach [Bogdanov-Tischhauser 2013, Vejre et al. 2016] for long-key cipher E:
There is a set $\mathcal{S}$ of identified (dominant) trails

$$\begin{aligned} & c_E(k_{i_0}, k_{i_0+1} \dots k_{i_0+r}) \\ &= \mathrm{cor}(u \cdot x + v \cdot E(x; k_{i_0}, k_{i_0+1} \dots k_{i_0+r}) \\ &= \sum_{w \in \mathcal{S}} (-1)^{w \cdot (k_{i_0}, k_{i_0+1} \dots k_{i_0+r})} c_E(u, w, v) + \mathcal{R}_E(k_{i_0}, k_{i_0+1} \dots k_{i_0+r}) \end{aligned}$$

where $\mathcal{R}_E(k_{i_0}, k_{i_0+1} \dots k_{i_0+r})$ is normally distributed with mean zero.

This approach has been tested in experiments for scaled PRESENT variants, and it seems to work.

That is, cryptanalyst can collect many trails such that the remainder $\mathcal{R}_E(k_{i_0}, k_{i_0+1}, \dots, k_{i_0+r})$ behaves like random and has variance $\approx 2^{-n}$.

Dominant Trails Over Cipher

Our approach: Use the same set $\mathcal{S}$ also for the cipher BC with key-scheduling KS to estimate

$$\begin{aligned} c_{\mathrm{BC}}(K) &= \sum_{W} (-1)^{W \cdot \mathrm{KS}(K)} c_E(u, W, v) \\ &= \sum_{W \in \mathcal{S}} (-1)^{W \cdot \mathrm{KS}(K)} c_E(u, W, v) + \mathcal{R}_{\mathrm{BC}}(K) \end{aligned}$$

What can we assume about the behaviour of $\mathcal{R}_{\mathrm{BC}}(K)$?

Typical Key-Alternating Cipher

... has a “strong” key-scheduling.

Note. Here “strong” means something which behaves like the long-key cipher.

Then one can use the estimated variance of

$$c_E(k_{i_0}, k_{i_0+1} \ldots k_{i_0+r})$$

to estimate the variance of $c_{\mathrm{BC}}(K)$

$$\mathrm{Var}_K c_{\mathrm{BC}}(K) \approx \sum_{w \in \mathcal{S}} c_E(u, w, v)^2 + 2^{-n}$$

This approach has been tested with scaled versions of SMALLPRESENT, see [Vejre et al 2016] and (with different key schedules) [Blondeau-Nyberg 2017]

Outline

Permutation-Based 1-EM Cipher

Assume the permutation is based on an iterated structure E'. Then the cipher has an can be represented like

$$\mathrm{BC}'(x, K) = E'(x; K_1, k_1 \ldots k_{r'-1}, K_2)$$

where K is the secret key and $k_1 \ldots k_{r'-1}$ are known constants (e.g. zero). Again we use the properties of E to examine the correlations of linear approximations of the cipher

$$\begin{aligned} c_{\mathrm{BC}}(K) &= \sum_{W} (-1)^{W \cdot (k_{i_0} \ldots k_{i_0+r})} c_E(u, W, v) \\ &= \sum_{W \in \mathcal{S}} (-1)^{W \cdot (k_{i_0} \ldots k_{i_0+r})} c_E(u, W, v) + \mathcal{R}_{\mathrm{BC}}(k_{i_0} \ldots k_{i_0+r}), \end{aligned}$$

Given $c_E(u, W, v)$, $W \in \mathcal{S}$, the sum over $\mathcal{S}$ has a fixed value, say c, which can be computed. The remainder

$$\mathcal{R}_{\mathrm{BC}}(k_{i_0} \ldots k_{i_0+r})$$

also has a fixed value, but cannot be computed.

Modelling the Uncomputable Remainder

We model the uncomputable remainder

$$\mathcal{R}_{\mathrm{BC}}(k_{i_0} \ldots k_{i_0+r})$$

according to

$$\mathcal{R}_{E}(k_{i_0} \ldots k_{i_0+r})$$

Then we get

$$\begin{aligned} \mathrm{Exp}_K c_{\mathrm{BC}}(K) &= c \\ \mathrm{Var}_K c_{\mathrm{BC}}(K) &= 2^{-n} \end{aligned}$$

This model should work when the permutation is based on a known-key PRESENT.

Linear key-recovery attack with success probability $> 1/2$ and advantage > 1 possible if $c \neq 0$.

Permutation-Based 2-EM Cipher

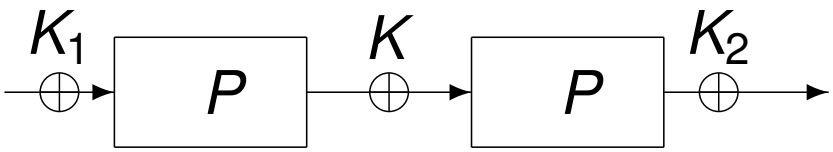


Figure : 2-EM block cipher with permutation P

Assume the two instances of permutation P are based on an iterated structure with long keys, as before.
After peeling off some first and last rounds, we consider linear properties over the following cipher

$$\mathrm{BC}(x, K) = P_2(P_1(x; RK_1) \oplus K; RK_2)$$

where RK_1 and RK_2 are known constants. Now the underlying long-key cipher E is

$$E(x; RK_1, K, RK_2) = P_2(P_1(x; RK_1) \oplus K; RK_2)$$

Correlations over 2-EM Cipher

Again we use the properties of E to examine the correlations of linear approximations of the cipher

$$c_{\mathrm{BC}}(K) \quad = \quad \sum_{W_1,w,W_2} (-1)^{W_1 \cdot RK_1 + w \cdot K + W_2 \cdot RK_2} c_E(u; W_1, w, W_2; v),$$

where

$$c_E(u; W_1, w, W_2; v) = c_{P_1}(u; W_1; w) c_{P_2}(w; W_2; v).$$

is independent of K. Hence

$$\mathrm{Exp}_K c_{\mathrm{BC}}(K) = 0$$

and

$$\mathrm{Var}_K c_{\mathrm{BC}}(K) = \mathrm{Exp}_K(c_{\mathrm{BC}}(K)^2)$$

Variance Estimation

Taking the noise-based approach, assume there exist sets $\mathcal{S}$, $\mathcal{S}_1$ $\mathcal{S}_2$ and a random and independent remainder $\mathcal{R}$ such that

$$c(RK_1, K, RK_2) = \sum_{W_1 \in \mathcal{S}_1, w \in \mathcal{S}, W_2 \in \mathcal{S}_2} (-1)^{W_1 \cdot RK_1 + w \cdot K + W_2 \cdot RK_2} c_{P_1}(u; W_1; w) c_{P_2}(w; W_2; v) + \mathcal{R}$$

$$= \sum_{w \in \mathcal{S}} (-1)^{w \cdot K} \sum_{W_1 \in \mathcal{S}_1} (-1)^{W_1 \cdot RK_1} c_{P_1}(u; W_1; w) \sum_{W_2 \in \mathcal{S}_2} (-1)^{W_2 \cdot RK_2} c_{P_2}(w; W_2; v) + \mathcal{R}$$

... to obtain

$$\mathrm{Var}_K c_{\mathrm{BC}}(K) = \sum_{w \in \mathcal{S}} \left(\sum_{W_1 \in \mathcal{S}_1} (-1)^{W_1 \cdot RK_1} c_{P_1}(u; W_1; w) \right)^2 \left(\sum_{W_2 \in \mathcal{S}_2} (-1)^{W_2 \cdot RK_2} c_{P_2}(w; W_2; v) \right)^2 + 2^{-n}$$

$$= \sum_{w \in \mathcal{S}} c_1(w)^2 c_2(w)^2 + 2^{-n}$$

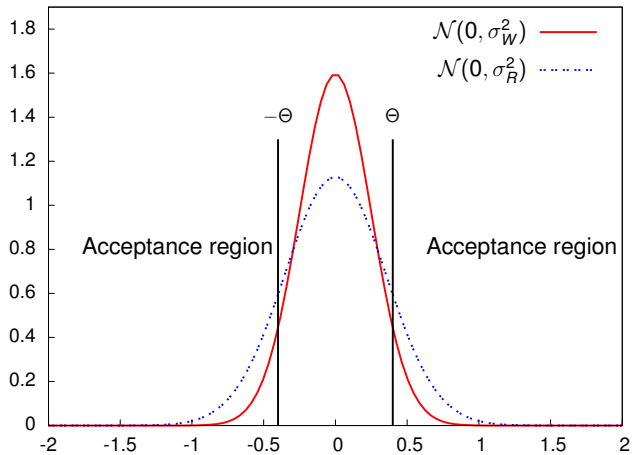
Attack Based on Variance

$$\mathrm{Var}_K c_{\mathrm{BC}}(K) = \sum_{w \in \mathcal{S}} c_1(w)^2 c_2(w)^2 + 2^{-n}$$

where $c_1(w)$ and $c_1(w)$ are the evaluated parts of the linear hull.

This gives

$$\begin{aligned} \mathrm{Exp}_{D,K} \hat{c}(D,K) &= 0 \\ \mathrm{Var}_{D,K} \hat{c}(D,K) &= \frac{1}{N} + \sum_{w \in \mathcal{S}} c_1(w)^2 c_2(w)^2 + 2^{-n} \end{aligned}$$



Example distribution of $\hat{c}(D,K)$ for random (red solid line) and right key (blue dotted line)

On the Choice of Trails

Question: When is $\mathcal{S}$ (or $\mathcal{S} \times \mathcal{S}_1 \times \mathcal{S}_2$) sufficiently large?

Possible answers:

- When $\mathrm{Var}_{k_{i_0} \ldots k_{i_0+r}} \left(R_E(k_{i_0}, \ldots, k_{i_0+r}) \right) = 2^{-n}$.
  How do we know if this has been reached?
- For permutation-based ciphers, when the value c is stable, that is, adding trails to $\mathcal{S}$ does not (essentially) change the value

$$\sum_{w \in \mathcal{S}} (-1)^{w \cdot (k_{i_0} \ldots k_{i_0+r})} c_E(u, w, v)$$

Security of 1-EM vs. 2-EM

- Assume 1-EM vs. 2-EM have the same total number r' of rounds
- 2-EM has a secret key K after round $r'/2$ while in 1-EM it is fixed to a known $k_{r'/2}$
- Assume the same identified set $\mathcal{S}_1 \times \mathcal{S} \times \mathcal{S}_2$ of trails over r rounds is used for both

2-EM

$$q_{2-\mathrm{EM}}(K) = \sum_{w \in \mathcal{S}} (-1)^{w \cdot K} c_1(w) c_2(w)$$

1-EM

$$q_{1-\mathrm{EM}} = \sum_{w \in \mathcal{S}} (-1)^{w \cdot k_{r'/2}} c_1(w) c_2(w) = q_{2-\mathrm{EM}}(k_{r'/2})$$

Security of 1-EM vs. 2-EM

It may be possible (?) to select $k_{r'/2}$ such that

$$q_{1-\mathrm{EM}} = 0$$

for all strong linear approximations (u, v).

If this cannot be done, it may happen that the attacker finds a (u, v) such that

$$q_{1-\mathrm{EM}} = q_{2-\mathrm{EM}}(k_{r'/2}) \neq 0$$

or even more,

$$\mathrm{Exp}_K(q_{2-\mathrm{EM}}(K)^2) < q_{2-\mathrm{EM}}(k_{r'/2})^2$$

For the c defined earlier, it holds

$$c = q_{1-\mathrm{EM}}$$

Further denote

$$\sigma = \sqrt{\mathrm{Exp}_K(q_{2-\mathrm{EM}}(K)^2)}$$

Linear Attack on 1-EM vs. 2-EM

- a advantage
- N number of known plaintext-ciphertext pairs
- Φ standard normal

Success probabilities:

1-EM

$$P_S = \Phi\left(|c|\sqrt{N} - \Phi^{-1}(1 - 2^{-a})\sqrt{1 + N2^{-n}}\right)$$

2-EM

$$P_S = 2\Phi\left(-\Phi^{-1}(1 - 2^{-a-1})\sqrt{\frac{1 + N2^{-n}}{1 + N2^{-n} + N\sigma^2}}\right)$$

For $|c| = \sigma$ the attack on 1-EM is much stronger.