On the construction of generalized approximations for one filter generator key recovery method

Alekseev E.K.[1], Kushchinskaya L.A.[2]

[1]CryptoPro LCC, [2]Lomonosov Moscow State University

June 7, 2017

Figure 1: Filter generator

Filter generator is a construction built on the basis of linear mapping $A : V_n \to V_n$ and the Boolean function $f \in \mathcal{F}_n$ (see fig. 1a).

- Correlation method proposed by T. Siegenthaler
- Fast correlation attack proposed by W. Meier and O. Staffelbach
- Algebraic attacks proposed by N. Courtois and W. Meier

Definition

For a filter generator, the *trajectory* will refer to the three values $\mathrm{Traj} = \langle m, \mathbb{L}, \mathbb{T}\rangle$, where $m \in \mathbb{N}$ is the length of the trajectory,
$\mathbb{L} = \{L_i - \text{is a plane in } V_n | i = \overline{1,m}\}$,
$\mathbb{T} = \{t_i | t_i \in \mathbb{Z}, i = \overline{1,m};\ \ t_1 = 0\}$, such that

$$L_i = A^{t_i - t_{i-1}}(L_{i-1}),\ \ t_i, t_{i-1} \in \mathbb{T}, i = \overline{2,m}.$$

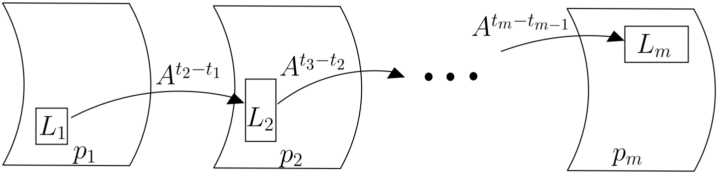
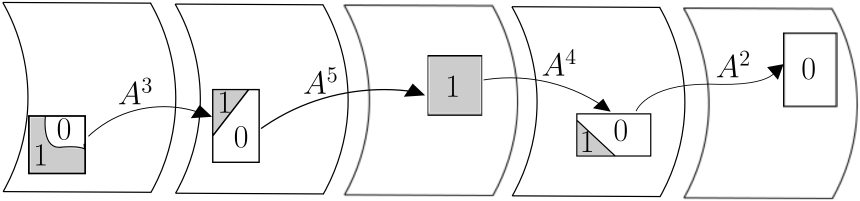
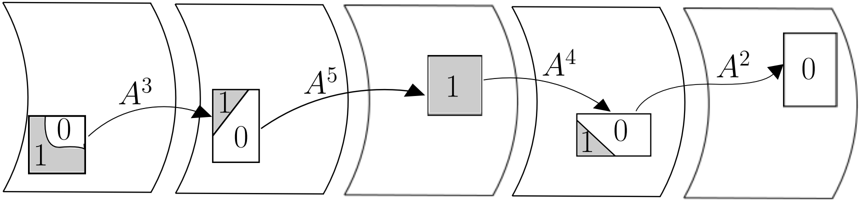


Figure 2: Trajectory

Trajectory

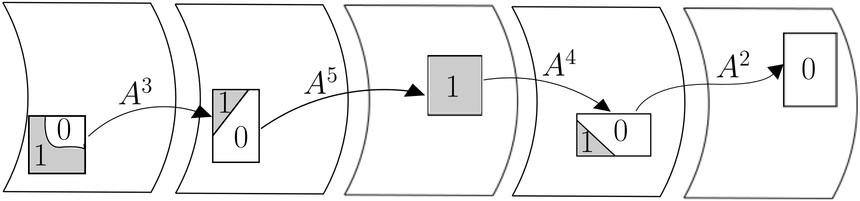


Trajectory



- $z = 0010110001011 0000\ldots$ - reject

Trajectory



- $z = 0010110001011 0000\ldots$ - reject
- $z = 1010110001010 0000\ldots$ - accept

Definition

The characteristic of trajectory $\mathrm{Traj} = \langle m, \mathbb{L}, \mathbb{T} \rangle$ is a pair of sets $(\mathbb{P}, C)$, where $\mathbb{P} = \{p_i | p_i \in \left(\frac{1}{2}; 1\right], i = \overline{1, m}\}$, $C = \{c_i | c_i \in \mathbb{F}_2, i = \overline{1, m}\}$, p_i is the probability that the value of the filter function f is the same as constant c_i in plane $L_i, i = \overline{1, m}$, provided that vector $v \in L_i$ is picked randomly with each value having the same probability of being selected.

Definition

The set of all the trajectories $\{\mathrm{Traj}^{(i)}\}$ will then be referred to as *the generalized approximation* of filter function $f \in \mathfrak{F}_n$ in the generator with linear mapping A.

Definition

The starting set $\mathbb{L}_{start}$ of the generalized approximation is the collection of sets $\{L_1^{(i)}\}$ from each trajectory.

$$z_i = f(A^i u^*), \quad i = \overline{0, N-1}.$$

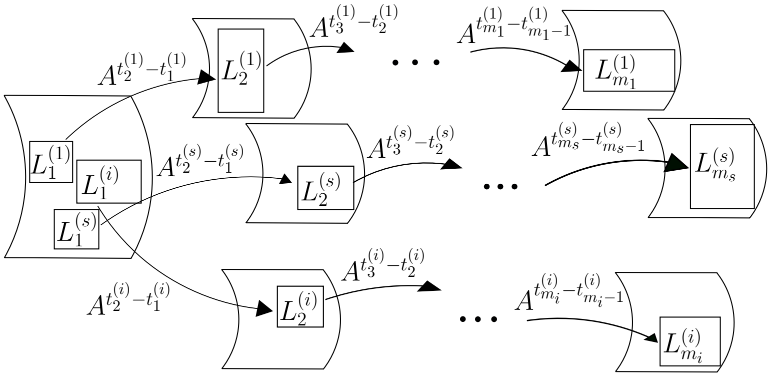


Figure 3: Generalized approximation

Lets build a vector

$$w = (c_1 \oplus \widetilde{z}_1, \dots, c_m \oplus \widetilde{z}_m), \;\; \widetilde{z}_i = z_{t_i}, \;\; i = \overline{1, m}.$$

for each trajecrory from the generalized approximation.
Lets assume that there is some deciding rule of the form $F(L) \geqslant 0$, that allows us to accept or reject the trajectory Traj (plane $L \in \mathbb{L}_{start}$).

Description of the algorithm

Let $\hat{L} = \mathbb{L}_{start}$, $M = V_n \setminus \left(\bigcup_{L \in \mathbb{L}_{start}} L\right)$.

1 Stage one (selecting the 'correct' trajectories). $\widetilde{L} := \emptyset$.

1.a) If $\hat{L} = \emptyset$, then go to stage two. Otherwise select a random element $L_1^{(i)}$ from the set $\hat{L}$; $\hat{L} := \hat{L} \setminus \{L_1^{(i)}\}$.

1.b) Build vector $w \in V_{m_i}$. If the inequality $F(L_1^{(i)}) \geqslant 0$ holds, then assume $\widetilde{L} := \widetilde{L} \cup \{L_1^{(i)}\}$. Go to step 1.a).

Description of the algorithm

2 Stage two (thorough testing of the 'correct' trajectories).

2.a) If $\widetilde{L} = \emptyset$, then go to stage three, else select Y from set $\widetilde{L}$; $\widetilde{L} := \widetilde{L} \setminus \{Y\}$.

2.b) If $Y = \emptyset$, then go to step 2.a). Otherwise select $v \in Y$; $Y := Y \setminus \{v\}$.

2.c) If $f(A^i v) = z_i$ for any $i = \overline{0, N-1}$, then return v and stop, otherwise go to step 2.b).

Description of the algorithm

3 Stage three (viewing set M).

3.a) If $M = \emptyset$, then quit without returning anything, otherwise select $u \in M$; $M := M \setminus \{u\}$.

3.b) If $f(A^i u) = z_i$ for any $i = \overline{0, N-1}$, then return u as the result and exit, otherwise go to 3.a).

The general characteristics of the method are proved in *E.K.Alekseev, L.A.Kushchinskaya.* Generalizing one method for recovering the key of a filter generator. Discrete Mathematics and Applications, 2017, forthcoming. (In Russian).

- α is the probability to accept "false" plane (the method laboriousness);
- β is the probability to reject "true" plane (the method reliability).

Theorem

Lets assume $Pr\left[u^* = v\right] = \frac{1}{2^n}, \ \ \forall v \in V_n$. *The method reliability satisfies the following inequality*

$$\pi \geqslant 1 - \frac{1}{2^n} \sum_{j=1}^{s} \beta_j \cdot |L_1^{(j)}|.$$

Planes from $\mathbb{L}_{start}$ do not intersect with each other.

Theorem

Let $C = \sum_{j=1}^{s} |L_1^{(j)}| \cdot \alpha_j$. *The method laboriousness is equal to*

$$s + C + \frac{|M|}{2^n}\left(\frac{|M|+1}{2} + \sum_{i=1}^{s} |L_1^{(i)}| \cdot \beta_i\right) + \frac{1}{2^n}\sum_{i=1}^{s} |L_1^{(i)}|^2 \cdot (1 - \alpha_i - \beta_i).$$

Planes from $\mathbb{L}_{start}$ do not intersect with each other and $\bigcup_{L \in \mathbb{L}_{start}} L = V_n$, while $dim(L) = k, \forall L \in \mathbb{L}_{start}$.
The method laboriousness is equal to

$$D = 2^{n-k} + 2^n \cdot \alpha + 2^k \cdot (1 - \alpha - \beta).$$

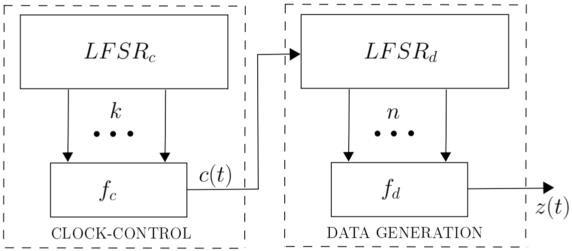


Figure 4: LILI-128

- DATA GENERATION: $2^{89} \to 2^{76}$
- CLOCK-CONTROL: $2^{39} \to 2^{23}$
- LILI-128: $2^{128} \to 2^{118}$

What do we need?

What do we need?

- Laboriousness Q

What do we need?

- Laboriousness Q
- Reliability π_0

What do we need?

- Laboriousness Q
- Reliability π_0
- The minimal possible amount of the generators output

Description of the method

- Parameter $k \in \{1, 2, \ldots, n-1\}$ is the number of dimensions for the plane in the trajectory, $N = 2^k$.

- Parameter $k \in \{1, 2, \ldots, n-1\}$ is the number of dimensions for the plane in the trajectory, $N = 2^k$.
- Parameter $\delta \in \{1, 2, \ldots, N\}$ will be responsible for the minimal predominance of some constant in the plane. Then $T_0 = \frac{N}{2} - \frac{\delta}{2}$ is the boundary for the number of zero values and $T_1 = \frac{N}{2} + \frac{\delta}{2}$ is the boundary for ones.

Select a random plane L_0 with k dimensions. For each $i = 0, 1, 2, \ldots$:

- If in plane L_i filter function f equals 1 a certain number of times different than $N/2$ by a large enough value then we add it to the trajectory were constructing.
- $L_{i+1} := A(L_i)$.
- Repeat the steps above until weve achieved the desired length of the trajectory.

- Lets assume that plane L_i is selected randomly and independently from the another.

- Lets assume that plane L_i is selected randomly and independently from the another.
- The predominance in accuracy equals $1/2 + \delta/2N$.

- Lets assume that plane L_i is selected randomly and independently from the another.
- The predominance in accuracy equals $1/2 + \delta/2N$.
- Let $p(\delta, k)$ be the probability that a random plane gets selected for the trajectory, let N_1 be the length of the trajectory, $N_2 = \frac{N_1}{p(\delta,k)}$.

$\underline{k \to 0}$:

method $\to$ *B-F* *on the first stage*

$\underline{k \to n}$:

method $\to$ *B-F* *on the second stage*

$\underline{\delta \to 0}$:

$p \uparrow, N_1 \uparrow$

$\underline{\delta \to N}$:

$p \downarrow, N_1 \downarrow$

Figure 5: Just B-F?

- L is a random set of vectors from V_n with a capacity of 2^k.
- $S_N = \sum\limits_{v \in L} f(v)$. $S_N \sim HG(2^{n-1}, 2^n, 2^k)$.
- $S_N \approx Bin(N, \frac{1}{2}) \Rightarrow S_N \approx N(\frac{N}{2}, \frac{N}{4})$.

The following expression holds

$$Pr\left[T_0 \leq S_N \leq T_1\right] = 2\Phi\left(\frac{\delta}{\sqrt{N}}\right) - 1,$$

where $\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{x^2}{2}} dx$ is a distribution function for a normally distributed random value. From this derive the following

$$p(\delta, k) = 1 - Pr[T_0 \leq S_N \leq T_1] = 2\left(1 - \Phi\left(\frac{\delta}{\sqrt{N}}\right)\right).$$

- $\beta = 1 - \pi_0$ is the probability of second type errors.
- $Q = 2^{n-k} + \alpha \cdot 2^{n-k} \cdot 2^k + (1-\beta) \cdot 2^k \Rightarrow$
  $\alpha = 2^{-n} \cdot \left(Q - 2^{n-k} - \pi_0 \cdot 2^k\right)$.
- $$N_1 \approx \frac{\left(u_\alpha \sqrt{q_0(1-q_0)} + u_\beta \sqrt{q_1(1-q_1)}\right)^2}{(q_1-q_0)^2},$$
  where u_α, u_β are the quantiles of a standard normal distribution, $q_0 = \frac{1}{2}$, $q_1 = \frac{1}{2} + \frac{\delta}{2N}$.

Characteristics of the method

- $$N_2 = \frac{N_1}{p(\delta,k)} \approx \frac{\left(u_\alpha + u_\beta \cdot \sqrt{1-\left(\frac{\delta}{N}\right)^2}\right)^2 \cdot \left(\frac{N}{\delta}\right)^2}{2\left(1-\Phi\left(\frac{\delta}{\sqrt{N}}\right)\right)}$$

- $N_2 = \frac{N_1}{p(\delta,k)} \approx \frac{\left(u_\alpha + u_\beta \cdot \sqrt{1 - \left(\frac{\delta}{N}\right)^2}\right)^2 \cdot \left(\frac{N}{\delta}\right)^2}{2\left(1 - \Phi\left(\frac{\delta}{\sqrt{N}}\right)\right)}$
- Let $t = \delta/\sqrt{N}, t \in (0; \sqrt{N}]$:

$$N_2 \approx \frac{N}{2} \cdot \left(u_\alpha + u_\beta \sqrt{1 - \frac{t^2}{N}}\right)^2 \cdot \frac{1}{t^2(1 - \Phi(t))}.$$

- $Q \ll 2^n$ and $\alpha \ll \beta$ ($u_\alpha \gg u_\beta$). Then $N_2 \approx \frac{N}{2} \cdot u_\alpha^2 \cdot \frac{1}{t^2(1-\Phi(t))}$.

- $Q \ll 2^n$ and $\alpha \ll \beta$ $(u_\alpha \gg u_\beta)$. Then $N_2 \approx \frac{N}{2} \cdot u_\alpha^2 \cdot \frac{1}{t^2(1-\Phi(t))}$.
- $f(t) = t^2(1-\Phi(t)) \to$ maximum, $t \in (0; \sqrt{N}]$. $f'(t) = 0$:

$$2(1-\Phi(t)) = t \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}},$$

which is equivalent to

$$\frac{t}{2} = \frac{1-\Phi(t)}{\varphi(t)},$$

where $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$.

- $R(t) = \frac{1-\Phi(t)}{\varphi(t)}$ is known as the Mills ratio.
- The equation $\frac{t}{2} = \frac{1-\Phi(t)}{\varphi(t)}$ has just one solution
- $t_0 = 1.19061\ldots$
- Thus, $\delta \approx \lceil t_0 \cdot \sqrt{N} \rceil$. Value N_2 in the minimum:

$$N_2 \approx \frac{N}{2} \cdot u_\alpha^2 \cdot C_\Phi,$$

where $C_\Phi = \frac{1}{t_0^2(1-\Phi(t_0))} \approx 6.03442$.

- $u_\alpha \approx \sqrt{-ln(2\pi\alpha^2)}$ for small $\alpha \Rightarrow N_2$ reaches the minimum at the minimal possible k.
- $k \in \{1, 2, \ldots, n-1\}$: $\alpha = 2^{-n} \cdot \left(Q - 2^{n-k} - \pi_0 \cdot 2^k\right) > 0$
- Minimal k:

$$k = \left\lceil log_2 \left(\frac{Q - \sqrt{Q^2 - \pi_0 2^{n+2}}}{2\pi_0} \right) \right\rceil .$$

The values of the functions in the minimum are as follows:

- $N_1 = (u_\alpha)^2 \cdot \left(\frac{N}{\delta}\right)^2 = \left(\frac{u_\alpha}{t_0}\right)^2 \cdot N,$
- $N_2 = \left(\frac{u_\alpha}{t_0}\right)^2 \cdot \frac{N}{2} \cdot C_\Phi = N_1 \cdot \frac{C_\Phi}{2};$

- $n = 128$, $\pi_0 = 1/2$

Method characteristics

	k	δ	N_1	N_2
$Q = 2^{70}$	59	2^{30}	2^{65}	2^{67}
$Q = 2^{80}$	49	2^{25}	2^{54}	2^{56}
$Q = 2^{90}$	39	2^{20}	2^{25}	2^{27}

- $n = 128$, $\pi_0 = 1/2$

Method characteristics

	k	δ	N_1	N_2
$Q = 2^{70}$	59	2^{30}	2^{65}	2^{67}
$Q = 2^{80}$	49	2^{25}	2^{54}	2^{56}
$Q = 2^{90}$	39	2^{20}	2^{25}	2^{27}

- Experimental verification
  $n = 32$, $Q = 2^{24}$, $\pi_0 = 1/2$
  Expected: $N_2 = 12861$, obtained: $N_2 = 12910$.

Thank you for your attention.

alekseev@cryptopro.ru
lyudmila.kuschinskaja@yandex.ru