Attacks against Filter Generators Exploiting Monomial Mappings

Anne Canteaut & Yann Rotella

GT BaC, 20 October 2017

Inria - SECRET, Paris, France

Summary

Introduction : Stream ciphers

Linear Feedback Shift Registers

Monomial equivalence between filtered LFSR

Univariate correlation attacks

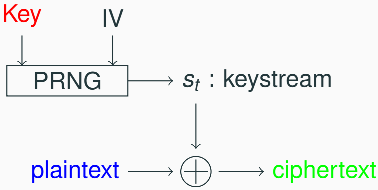
Impact on Boolean functions

Conclusions

Stream ciphers

Stream ciphers

- Symetric cryptography, $\neq$ block ciphers
- Based on Vernam cipher (one-time pad)
- PRNG



Stream ciphers

- Block cipher modes of operations (OFB, Counter)
- Specific design (LFSR, NLFSR)
- Internal state
- Large period
- A5/1 - A5/2, SNOW

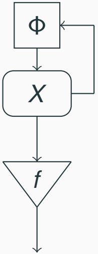
Stream ciphers

- Block cipher modes of operations (OFB, Counter)
- Specific design (LFSR, NLFSR)
- Internal state
- Large period
- A5/1 - A5/2, SNOW

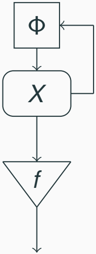
Interests

- Small latency
- No padding
- No error propagation
- Cheap

Generic attacks

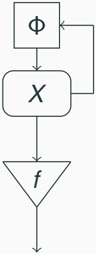


- Key recovering



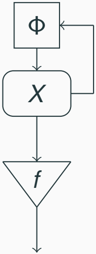
- Key recovering
- Initial state recovering

Generic attacks



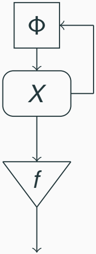
- Key recovering
- Initial state recovering
- Next-bit prediction

Generic attacks



- Key recovering
- Initial state recovering
- Next-bit prediction
- distinguishing s_t from a random sequence

Generic attacks



- Key recovering
- Initial state recovering
- Next-bit prediction
- distinguishing s_t from a random sequence

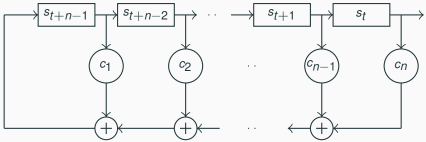
Always take an internal state twice bigger as the security level (i.e. key size)

LFSR

Linear feedback shift Register (LFSR)

Definition

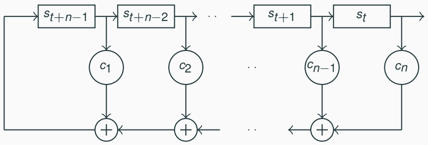
Fibonacci representation



Linear feedback shift Register (LFSR)

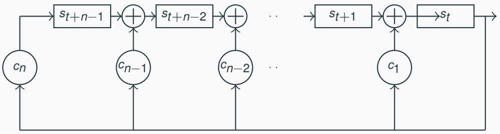
Definition

Fibonacci representation



Definition

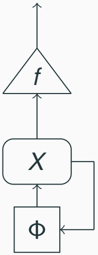
Gallois representation

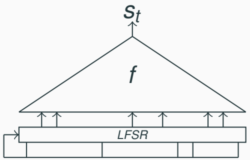


Classical properties of LFSR

- Nice statistical properties
- Linear
- $s_{t+L} = \sum_{i=1}^{n} c_i s_{t+n-i}, \forall t \leq 0$
- $P(X) = 1 - \sum_{i=1}^{n} c_i X^i$
- $P^*(X) = X^n P(1/X)$
- We wil take P primitive

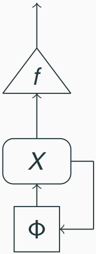
Filtered LFSR

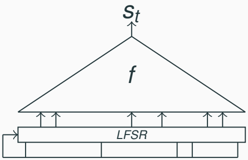




$$s_t = f(u_{t+\gamma_1}, \cdots, u_{t+\gamma_n})$$

Filtered LFSR





$$s_t = f(u_{t+\gamma_1}, \cdots, u_{t+\gamma_n})$$

Algebraic Normal Form

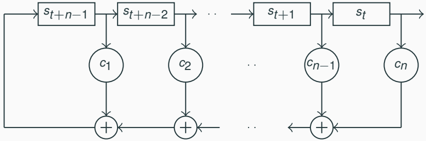
$$f(x_1, x_2, \cdots, x_n) = \sum_{u \in \mathbb{F}_2^n} a_u \prod_{i=1}^{n} x_i^{u_i}$$

$$= a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_3 x_1 x_2 + \cdots + a_{2^n-1} x_1 \cdots x_n$$

Monomial equivalence

LFSR over a Finite Field

- α : root of the primitive characteristic polynomial in $\mathbb{F}_{2^n}$
- Identify the n-bit words with elements of $\mathbb{F}_{2^n}$ with the dual basis of $\{1, \alpha, \alpha^2, \cdots, \alpha^{n-1}\}$

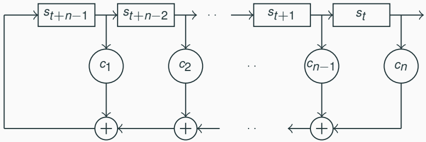


Proposition

The state of the LFSR at time $(t+1)$ is the state of the LFSR at time t multiplied by α.

LFSR over a Finite Field

- α : root of the primitive characteristic polynomial in $\mathbb{F}_{2^n}$
- Identify the n-bit words with elements of $\mathbb{F}_{2^n}$ with the dual basis of $\{1, \alpha, \alpha^2, \cdots, \alpha^{n-1}\}$



Proposition

The state of the LFSR at time $(t+1)$ is the state of the LFSR at time t multiplied by α.

For all t, $X_t = X_0\alpha^t$

Boolean functions

Proposition (Univariate representation)

$$F(X) = \sum_{i=0}^{2^n-1} A_i X^i$$

with $A_i \in \mathbb{F}_{2^n}$ *given by the discrete Fourier Transform of* F

Boolean functions

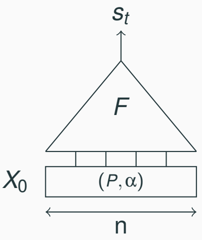
Proposition (Univariate representation)

$$F(X) = \sum_{i=0}^{2^n-1} A_i X^i$$

with $A_i \in \mathbb{F}_{2^n}$ *given by the discrete Fourier Transform of* F

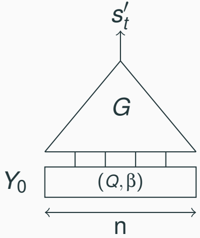
For all t**,** $s_t = F(X_0 \alpha^t)$

Monomial equivalence [Rønjom - Cid 2010]



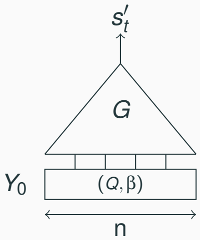
For all t**,** $s_t = F(X_0\alpha^t)$

Monomial equivalence [Rønjom - Cid 2010]



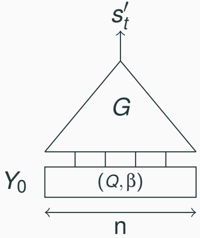
$$\beta = \alpha^k \text{ with } \gcd(k, 2^n - 1) = 1$$

Monomial equivalence [Rønjom - Cid 2010]



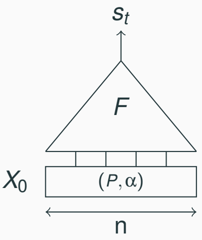
$$\beta = \alpha^k \text{ with } \gcd(k, 2^n - 1) = 1$$
$$s'_t = G(Y_0 \beta^t) = G(Y_0 \alpha^{kt})$$

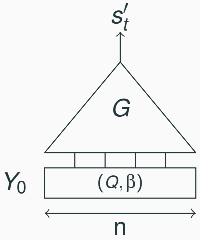
Monomial equivalence [Rønjom - Cid 2010]



$$\beta = \alpha^k \text{ with } \gcd(k, 2^n - 1) = 1$$
$$s'_t = G(Y_0\beta^t) = G(Y_0\alpha^{kt})$$
$$\textbf{If } G(x) = F(x^r)$$
$$\textbf{with } rk \equiv 1 \mod (2^n - 1)$$
$$\text{Then } s'_t = F(Y_0^r\alpha^t)$$

Monomial equivalence [Rønjom - Cid 2010]





For all t, $s_t = F(X_0 \alpha^t)$

$\beta = \alpha^k$ with $\gcd(k, 2^n - 1) = 1$

$s'_t = G(Y_0 \beta^t) = G(Y_0 \alpha^{kt})$

If $G(x) = F(x^r)$

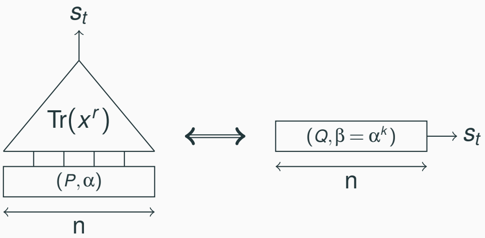
with $rk \equiv 1 \mod (2^n - 1)$

Then $s'_t = F(Y_0^r \alpha^t)$

For all t, $s'_t = s_t$ if $Y_0 = X_0^k$

Example

$F(x) = \mathrm{Tr}(x^r)$, with $\gcd(r, 2^n - 1) = 1$:
Let k be such that $rk \equiv 1 \mod (2^n - 1)$.



$\Longrightarrow$ The initial generator is equivalent to a plain LFSR of the same size.

Consequence

The security level of a filtered LFSR is the minimal security level for a generator of its equivalence class.

Consequence
The security level of a filtered LFSR is the minimal security level for a generator of its equivalence class.

- Algebraic attacks
- Correlation attacks

Algebraic attacks

Λ : Linear complexity

Proposition (Massey-Serconek 94)

Let an LFSR of size n filtered by a Boolean function F:

$$F(X) = \sum_{i=0}^{2^n-1} A_i X^i$$

Then

$$\Lambda = \#\{0 \leq i \leq 2^n - 2 : A_i \neq 0\}$$

Algebraic attacks

Λ : Linear complexity

Proposition (Massey-Serconek 94)

Let an LFSR of size n filtered by a Boolean function F:

$$F(X) = \sum_{i=0}^{2^n-1} A_i X^i$$

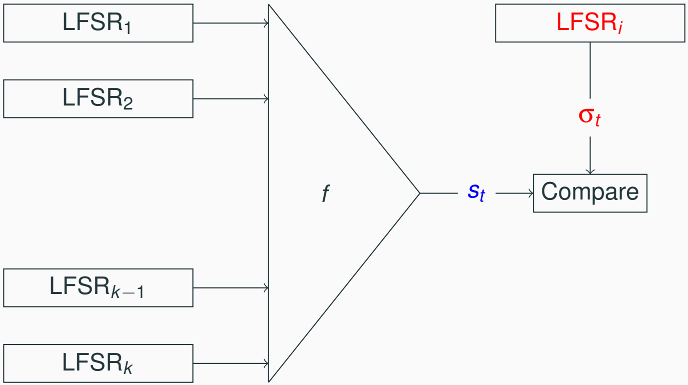
Then

$$\Lambda = \#\{0 \leq i \leq 2^n - 2 : A_i \neq 0\}$$

The monomial equivalence does not affect the complexity of algebraic attacks [Gong et al. 11]

Univariate correlation attacks

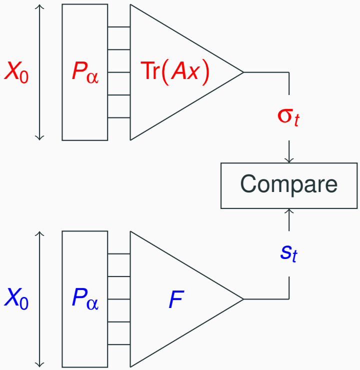
Correlation attack [Siegenthaler 85]



Criterion

The criterion besides the correlation attack is the **resiliency**.

Fast correlation attack [Meier - Staffelbach 88]

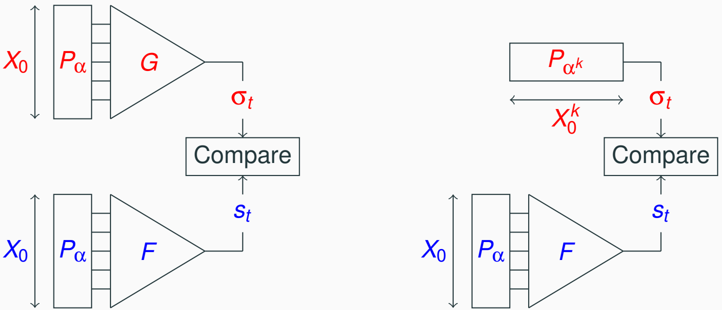


Criterion

The criterion besides the fast correlation attack is the **non-linearity**.

Generalized fast correlation attacks

$G(x) = \mathrm{Tr}(Ax^k)$



Generalized non-linearity [Gong & Youssef 01]

Relevant security criterion:

Generalized non-linearity

$$\mathrm{GNL}(f) = d(f, \{\mathrm{Tr}(\lambda x^k, \lambda \in \mathbb{F}_{2^n}, \gcd(k, 2^n - 1) = 1\})$$

Generalized non-linearity [Gong & Youssef 01]

Relevant security criterion:

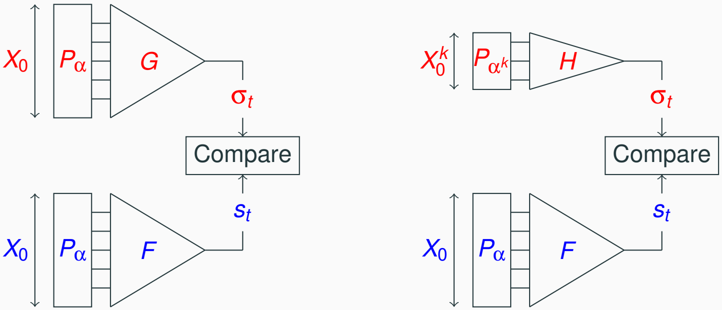
Generalized non-linearity

$$\mathrm{GNL}(f) = d(f, \{\mathrm{Tr}(\lambda x^k, \lambda \in \mathbb{F}_{2^n}, \gcd(k, 2^n - 1) = 1\})$$

And if k is not coprime to $2^n - 1$?

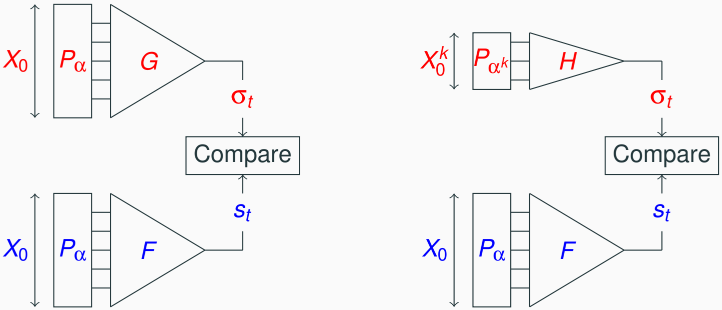
A more efficient correlation attack

When $\gcd(k, 2^n - 1) > 1$ and F correlated to $G(X) = H(X^k)$.



A more efficient correlation attack

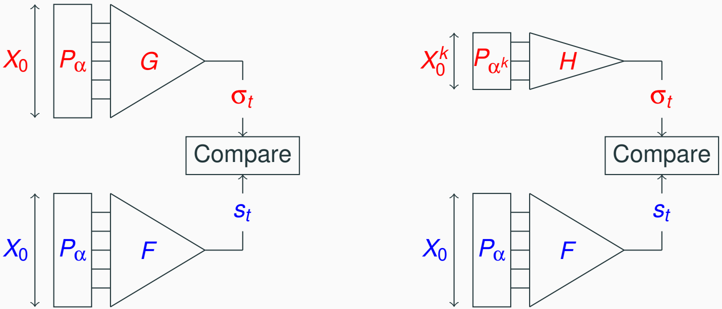
When $\gcd(k, 2^n - 1) > 1$ and F correlated to $G(X) = H(X^k)$.



- Number of states of the small generator: $\tau_k = \mathrm{ord}(\alpha^k)$.

A more efficient correlation attack

When $\gcd(k, 2^n - 1) > 1$ and F correlated to $G(X) = H(X^k)$.



- Number of states of the small generator: $\tau_k = \mathrm{ord}(\alpha^k)$.
- Exhaustive search on X_0^k: Time $= \frac{\tau_k \log(\tau_k)}{\varepsilon^2}$

Recovering the remaining bits of the initial state

Property

We get $\log_2(\tau_k)$ bits of information on X_0 where $\tau_k = \mathrm{ord}(\alpha^k)$:

Recovering the remaining bits of the initial state

Property

We get $\log_2(\tau_k)$ bits of information on X_0 where $\tau_k = \mathrm{ord}(\alpha^k)$:

If we perform two distinct correlation attacks with k_1 et k_2, then we get $\log_2(\mathrm{lcm}(\tau_{k_1}, \tau_{k_2}))$ bits of information.

First improvement

The complexity

$$\text{Time} = \frac{\tau_k \log(\tau_k)}{\varepsilon^2}$$

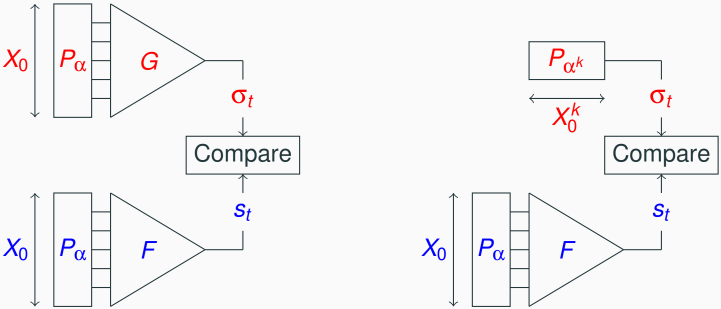
can be reduced to

$$\text{Time} = \tau_k \log \tau_k + \frac{2 \log(\tau_k)}{\varepsilon^2} \ .$$

with a fast Fourier transform [Canteaut - Naya-Plasencia 2012]

Second improvement

$G(X) = H(X^k)$ when H is linear:



- Size of the small LFSR: $L(k) = \mathrm{ord}(2) \mod \tau_k$.
- If $L(k) < n$ and H is linear $\longrightarrow$ fast correlation attack.

What we really do

- Split the state on the multiplicative subgroups
- recover independantly the information
- gather information

Impact on Boolean functions

New criterion

Definition (Multiplicative subgroup resiliency ?)

Let F be a Boolean function with n variables, let k dividing $2^n - 1$, and τ the multiplicative order of α^k and $d = \gcd(k, \tau)$, we say that F is k - MS resilient if and only if

$$\max_{G(x)=H(x^k)} \varepsilon(F(x), G(x)) = \frac{\tau}{d} 2^{-n}$$

Question

Is it possible to reach the value of τ/d for every possible τ ?

When H is linear

Question

What is the value of

$$\min_{f} \max_{G(x)=\mathrm{Tr}(\lambda x^k)} \varepsilon(F(x), G(x))$$

Conclusions

Conclusion and open questions

Conclusion

- Generalized criterion for f besides the generalized non-linearity.
- The attack does not apply when $(2^n - 1)$ is prime.

Open questions

- Find good filtering Boolean functions ?
- Compute efficiently a good approximation of the filtering function ?

Thank You for your attention !

Thank You for your attention !

Questions ?