Towards constructions of Boolean Functions with good cryptographic properties and efficient implementation

Enes Pasalic

UP FAMNIT & IAM

BalkanCryptSec 2015

Overview

Why stream ciphers ...

- ... when we have so many good block ciphers ?
- IDEA, KASUMI, FEAL, DES, AES, ...
- Lightweight representatives : PRESENT, KATAN, KLEIN, SPECK ... (1000-2000 gate equivalents (GE))
- Block ciphers are : well understood and analyzed, standardized, and can work in stream cipher mode.

"Stream ciphers - Dead or Alive"

Asiacrypt 2004, invited talk by Adi Shamir

Why stream ciphers ...

- ... when we have so many good block ciphers ?
- IDEA, KASUMI, FEAL, DES, AES, ...
- Lightweight representatives : PRESENT, KATAN, KLEIN, SPECK ... (1000-2000 gate equivalents (GE))
- Block ciphers are : well understood and analyzed, standardized, and can work in stream cipher mode.

"Stream ciphers - Dead or Alive"

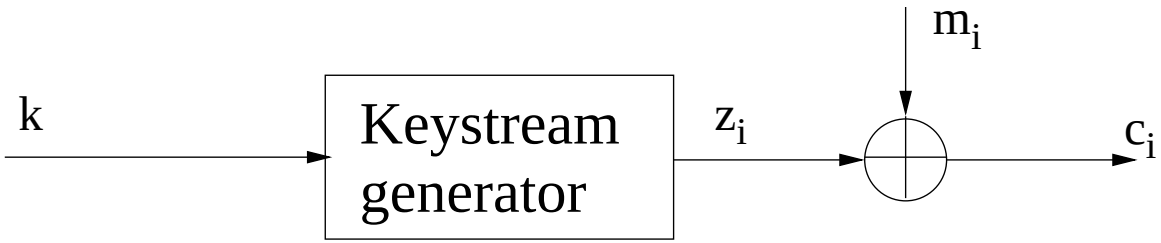
Asiacrypt 2004, invited talk by Adi Shamir

If not hardware efficient then :

"Why nonlinear combiners and filtering generators ? "

BCS 2015, tutorial talk by Enes Pasalic

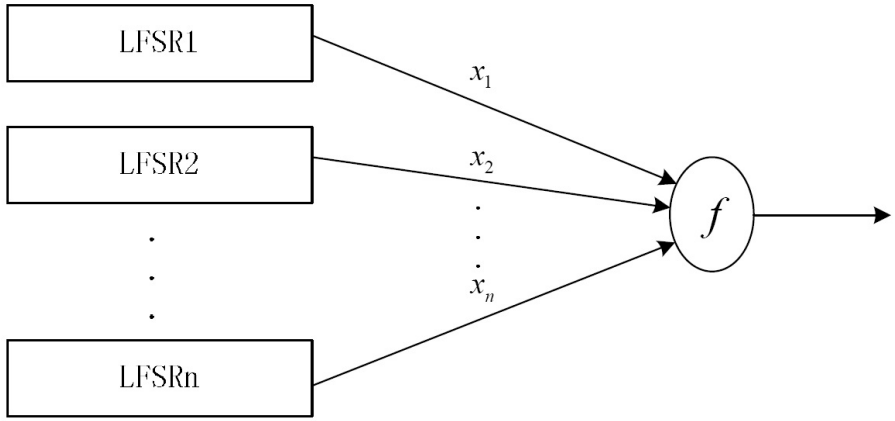
Additive stream ciphers



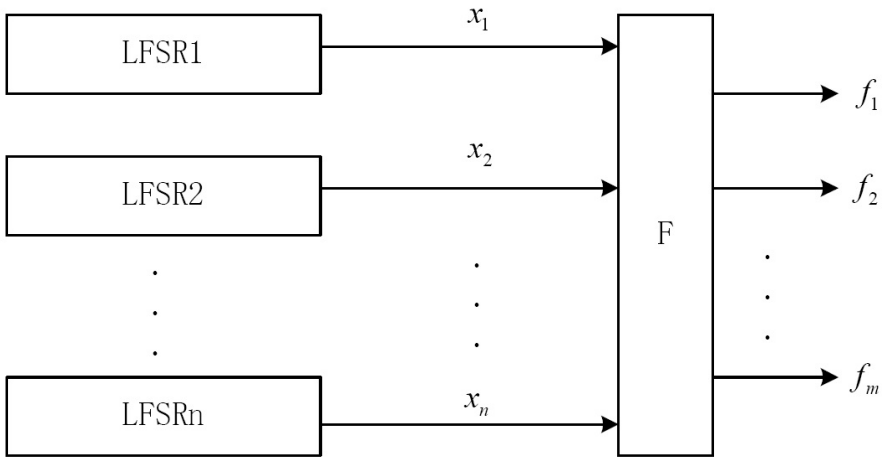
General model of a binary additive stream cipher

Filtering Boolean functions in nonlinear combiners

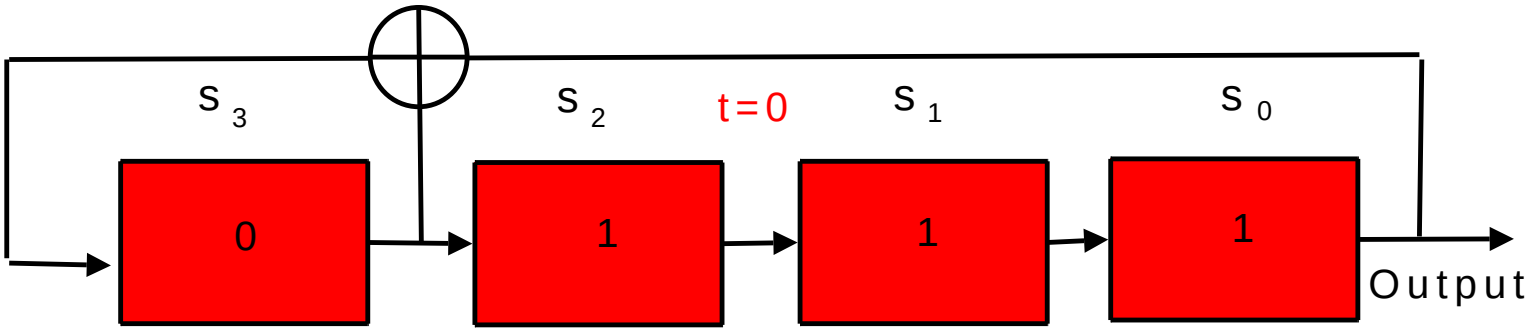
Boolean function in nonlinear combiner



Vectorial Boolean functions (S-boxes) for increased throughput



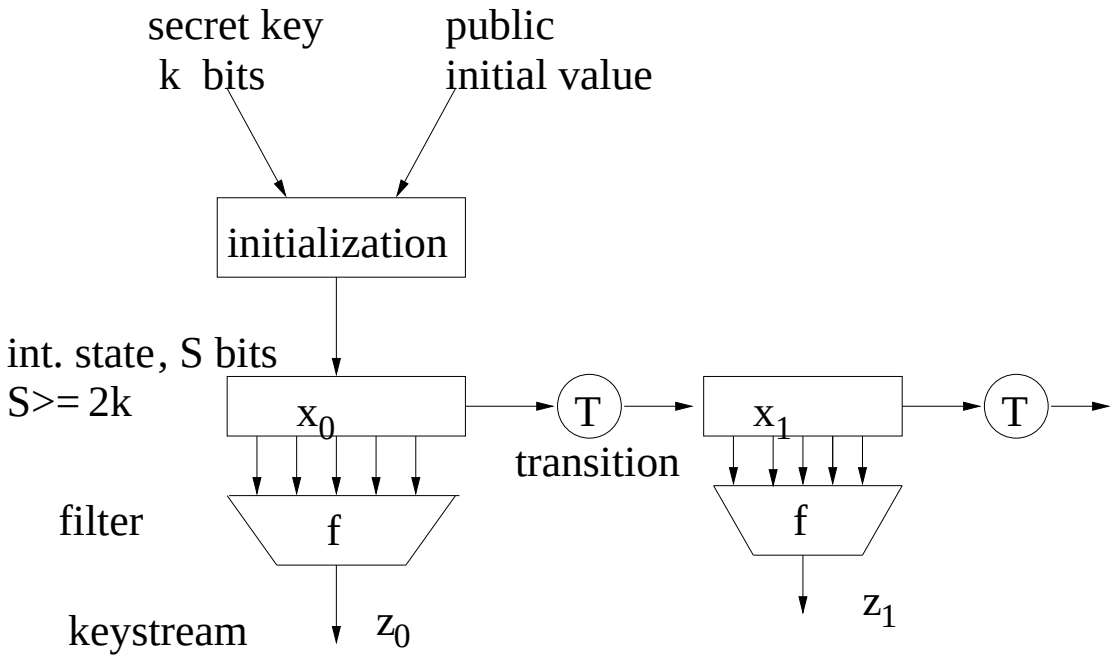
Linear Feedback Shift Registers (LFSR)



1	0	1	1	t=1
0	1	0	1	t=2
.				
.				
.				
0	1	1	1	t=15

- The recurrence is $s_{t+4} = s_{t+3} + s_t, \quad t \geq 0.$

LFSR filters - example



- Initial state $x_0 = s_0, \ldots, s_{S-1}$. Then $x_t = L^t(s_0, \ldots, s_{S-1})$.
- **Problem:** Recover $s_0, \ldots, s_{S-1}$ from $z_0, \ldots, z_{N-1}$,

$$z_t = f(x_t) = f \circ L^t(s_0, \ldots, s_{S-1}),\ 0 \leq t \leq N-1.$$

Cryptographic criteria for filtering Boolean functions

- High algebraic degree (linear complexity, alg. attacks)
- High nonlinearity (affine approximation attacks)
- Resiliency (correlation attacks)
- Optimal algebraic immunity (alg. attacks)
- Good resistance to (fast) algebraic attacks
- Hardware efficient (e.g. algebraic thickness Carlet 2003)
- ...

Cryptographic criteria for filtering Boolean functions

- High algebraic degree (linear complexity, alg. attacks)
- High nonlinearity (affine approximation attacks)
- Resiliency (correlation attacks)
- Optimal algebraic immunity (alg. attacks)
- Good resistance to (fast) algebraic attacks
- Hardware efficient (e.g. algebraic thickness Carlet 2003)
- ...

MAIN PROBLEM : Satisfy all requirements and do not forget implementation issues !

LFSR has (relatively) efficient hardware implementation, for $K = 80$ bits (at least) $|LFSR| = 160$ requiring c.a. 640 GE (gate equivalents). For Boolean function it remains c.a. 400-1000 GE for efficient implement. !

Boolean functions

A Boolean function $f(x_1, \ldots, x_n) \in \mathcal{B}_n$ maps from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ - generally represented by its **algebraic normal form** (ANF):

$$f(x_1, \ldots, x_n) = \bigoplus_{b \in \mathbb{F}_2^n} \lambda_b \left(\prod_{i=1}^{n} x_i^{b_i} \right), \tag{1}$$

where $\lambda_b \in \mathbb{F}_2$, $b = (b_1, \ldots, b_n) \in \mathbb{F}_2^n$.

The **algebraic degree** $deg(f)$, is defined as $\max_{b \in \mathbb{F}_2^n} \{wt(b) \mid \lambda_b \neq 0\}$, where $wt(b)$ denotes the Hamming weight of b.

- f is called an **affine function** when $deg(f) = 1$. If constant term is equal to zero f is called a **linear function**. Any linear function on $\mathbb{F}_2^n$ is denoted by:

$$\omega \cdot X_n = \omega_1 x_1 \oplus \cdots \oplus \omega_n x_n, \tag{2}$$

where $\omega = (\omega_1, \ldots, \omega_n) \in \mathbb{F}_2^n$ and $X_n = (x_1, \ldots, x_n)$.

Example: $f \in \mathcal{B}_3$

$$ANF: \quad f(x_1, x_2, x_3) = a_0 \oplus a_1x_1 \oplus a_2x_2 \oplus a_3x_3 \oplus a_4x_1x_2 \oplus a_5x_1x_3 \oplus a_6x_2x_3 \oplus a_7x_1x_2x_3$$

If $(a_0, a_1, \cdots, a_7) = (01110010)$, then

$$f(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3 \oplus x_2x_3.$$

We have $deg(f) = 2$.

Linear functions:

$$\ell(x_1, x_2, x_3) = \quad a_1x_1 \oplus a_2x_2 \oplus a_3x_3$$

Affine functions:

$$\rho(x_1, x_2, x_3) = a_0 \oplus a_1x_1 \oplus a_2x_2 \oplus a_3x_3$$

Table : $\mathbb{F}_2^3 \to \mathbb{F}_2$

x_1, x_2, x_3	f
000	0
001	1
010	1
011	0
100	1
101	0
110	1
111	0

Truth table and ANF correspondence

- From ANF to truth table is easy. In other direction it can be verified that,

$$f(x) = \sum_{\alpha | f(\alpha)=1} \prod_{i=1}^{n} (1 + x_i + \alpha_i), \quad \alpha \in \mathbb{F}_2^n.$$

- For the previous example we have

$$f(\alpha) = 1 \Leftrightarrow (\alpha_1, \alpha_2, \alpha_3) \in \{(0,0,1), (0,1,0), (1,0,0), (1,1,0)\}$$

- Then

$$\begin{aligned} f(x) &= (1 + x_1)(1 + x_2)x_3 + (1 + x_1)x_2(1 + x_3) + x_1(1 + x_2)(1 + x_3) \\ &\quad + x_1x_2(1 + x_3) = \ldots = x_1 + x_2 + x_3 + x_2x_3. \end{aligned}$$

IDEA : Select odd number of vectors in support (say $2^{n-1} - 1$) and you get maximal degree !

Table : The truth tables of 3-variable affine functions

x_1, x_2, x_3	**f**	l_0	l_1	l_2	l_3	l_4	l_5	l_6	l_7	l'_0	l'_1	l'_2	l'_3	l'_4	l'_5	l'_6	l'_7
000	**0**	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
001	**1**	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0
010	**1**	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0
011	**0**	0	1	1	0	0	1	1	0	1	0	0	1	1	0	0	1
100	**1**	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
101	**0**	0	1	0	1	1	0	1	0	1	0	1	0	0	1	0	1
110	**1**	0	0	1	1	1	1	0	0	1	1	0	0	0	0	1	1
111	**0**	0	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0
$N_f = 2$		4	6	4	2	4	2	4	2	4	2	4	6	4	6	4	6

The **nonlinearity** of a Boolean function $f \in \mathcal{B}_n$, denoted by N_f, is defined as the distance to the set of all affine functions,

$$N_f = \min_{\rho \in A(n)} \#\{X_n \in \mathbb{F}_2^n : f(X_n) \neq \rho(X_n)\}, \tag{3}$$

where $A(n)$ is the set of all affine functions on $\mathbb{F}_2^n$.

Theoretical research versus cryptographically practical

- For instance bent (highest nonlinearity) and semi-bent functions deserve special attention.
- Classification, counting, explicit algebraic expressions for bent functions (theoretical but also practical)

- Cryptographic drawbacks (at least in LFSR based schemes) are :
  - Nonbalanced (outputting $2^{n/2-1}$ more zeros or ones for 2^n output values)
  - Algebraic degree at most $n/2$

SOLUTION : Modify bent functions so that they are of optimal degree and very high nonlinearity.

Shannon's attack for linear transition ciphers

- Set up the enciphering equations:

$$\begin{aligned} z_0 &= f(s_0, s_1, \ldots, s_{S-1}) \\ z_1 &= f \circ L(s_0, s_1, \ldots, s_{S-1}) \\ &\vdots \\ z_t &= f \circ L^t(s_0, s_1, \ldots, s_{S-1}). \end{aligned}$$

- System of equations in S state variables of **degree** $d = deg(f)$. The number of terms is

$$\leq \sum_{i=0}^{d} \binom{S}{i} \approx \frac{S^d}{d!}$$

- Observe more than $\frac{S^d}{d!}$ bits and solve system using **linearization** (turn nonlinear system to linear) in **complexity** $\left(\frac{S^d}{d!}\right)^3$.

Algebraic attacks preliminaries

- Can we decrease the degree of the system ?
- If we can set up a true system of lower degree $r < d$ the complexity becomes smaller,

$$\left(\frac{S^r}{r!}\right)^3 \leftarrow \left(\frac{S^d}{d!}\right)^3$$

- How do we decrease the degree of the system ?
- What ciphers are vulnerable to this attack ?

Algebraic Immunity: Given $f \in \mathcal{B}_n$, define

$$AN(f) = \{g \in \mathcal{B}_n \mid f \cdot g = 0\}.$$

Any function $g \in AN(f)$ is called an *annihilator* of f. The *algebraic immunity*, denoted by $AI(f)$, of function f is the minimum degree of all non-zero annihilators of f and $f \oplus 1$.

Annihilators of Boolean function

- Let $f(x_3, x_2, x_1) = x_1x_2 + x_2x_3$.

x_3	x_2	x_1	$f(x)$	$g(x)$
0	0	0	0	$*$
0	0	1	0	$*$
0	1	0	0	$*$
0	1	1	1	0
1	0	0	0	$*$
1	0	1	0	$*$
1	1	0	1	0
1	1	1	0	$*$

- Assign "$*$" to get **annihilator** g, $f(x)g(x) = 0$, of low degree !
- For instance $g(x) = 1 + x_2$ gives

$$f(x)g(x) = [x_2(x_1 + x_3)][1 + x_2] = x_2(x_1 + x_3) + x_2(x_1 + x_3) = 0$$

Algebraic attacks- decreasing the degree of f

- **Idea of attack:** Find annihilator g of degree less than $deg(f)$. Observing $z^t = 1$,

$$f(x^t) = 1 \Rightarrow \underbrace{f(x^t)g(x^t)}_{=0} = g(x^t) \Rightarrow g \circ L^t(s_0, s_1, \ldots, s_{S-1}) = 0.$$

- Similarly for $h \in AN(1+f)$ and $f(x^t) = z^t = 0$,

$$h(x^t)(1 + f(x^t)) = 0 \Rightarrow h \circ L^t(s_0, s_1, \ldots, s_{S-1}) = 0.$$

- Solve a system of equations of degree $deg(g) = deg(h) < deg(f)$.

Fast algebraic attacks - Toyocrypt

- **Toyocrypt** uses LFSR of length 128 to generate $z^t = f(\mathbf{s}^t)$,

$$f(s_0, \ldots, s_{127}) = s_{127} + \sum_{i=0}^{62} s_i s_{\alpha_i} + s_{10}s_{23}s_{32}s_{42}$$

$$+ \quad s_1 s_2 s_9 s_{12} s_{18} s_{20} s_{23} s_{25} s_{26} s_{28} s_{33} s_{41} s_{42} s_{51} s_{53} s_{59} + \prod_{i=0}^{62} s_i .$$

- Now $T \approx \binom{128}{63} \approx 2^{124}$ which gives attack $Compl = 2^{124^3} = 2^{372}$.

Fast algebraic attacks - Toyocrypt

- **Toyocrypt** uses LFSR of length 128 to generate $z^t = f(\mathbf{s}^t)$,

$$f(s_0, \ldots, s_{127}) = s_{127} + \sum_{i=0}^{62} s_i s_{\alpha_i} + s_{10}s_{23}s_{32}s_{42}$$

$$+ \quad s_1 s_2 s_9 s_{12} s_{18} s_{20} s_{23} s_{25} s_{26} s_{28} s_{33} s_{41} s_{42} s_{51} s_{53} s_{59} + \prod_{i=0}^{62} s_i .$$

- Now $T \approx \binom{128}{63} \approx 2^{124}$ which gives attack $Compl = 2^{124^3} = 2^{372}$.

- But $f(\mathbf{s})(1 + s_{23})$ is of degree **3** ! System $f(\mathbf{s}^t)(1 + s_{23}) = z^t(1 + s_{23})$.

  Then $T \approx \binom{128}{3} = 2^{18}$, and attack complexity $\mathbf{2^{54}}$.

- **CRUCIAL :** Existence of low degree g, h s.t. $fg = h$ (here $\deg(g) = 1$, $\deg(h) = 3$)!

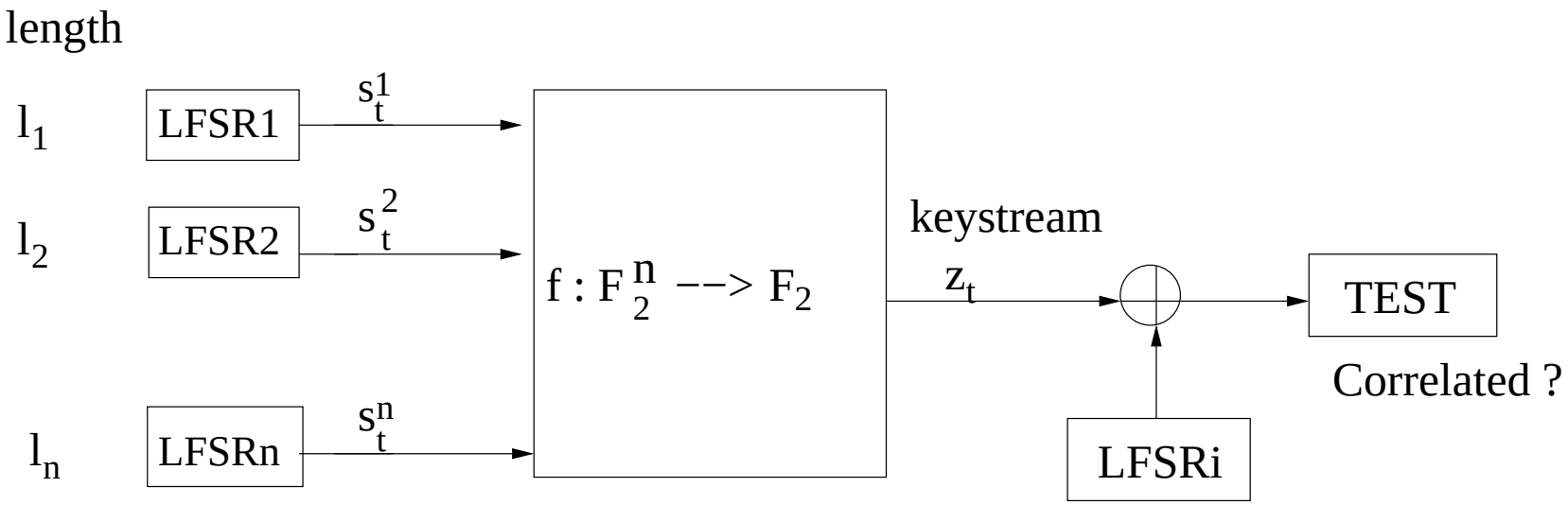
Introduction to correlation immunity

- Balanced Boolean functions in n variables are of degree $\leq n-1$.
- We might be interested in computing $Pb(f(x) = x_i)$! Consider the function $f(x) = x_3 + x_1x_2 + x_2x_3$.

x_3	x_2	x_1	$f(x)$	$f(x) + x_1$
0	0	0	0	0
0	0	1	0	1
0	1	0	0	0
0	1	1	1	1
1	0	0	1	0
1	0	1	1	0
1	1	0	0	0
1	1	1	1	0

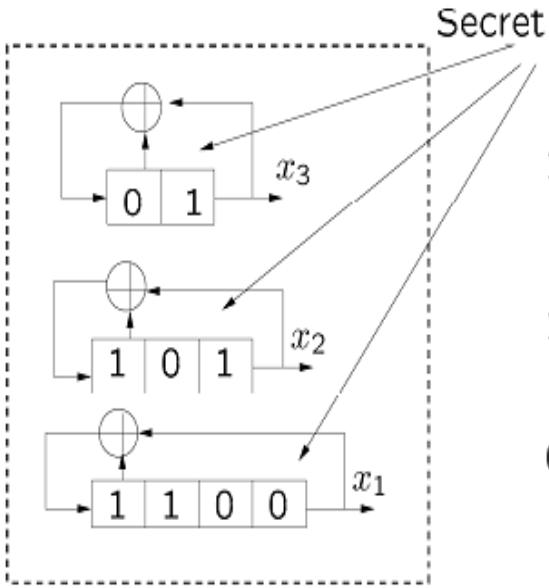
- Same situation, unbalancedness, for $f(x) + x_2$ and $f(x) + x_3$.

Correlation attacks



- Attack is performs by checking all states of LFSR1:
  - Guess not correct : We get a random sequence
  - Guess correct: Then $z_t \oplus x_1$ is biased, more zeros than ones .
- In previous example $Pb\{f(x) = x_i\} = 3/4$, thus possible to run test.

Example of correlation attacks



$z^t = f(x^t) = x_1 \oplus x_2 x_3$

1 0 1 1 0 1 1 0 1 1 0 1 1 0 1 1 0 1 1 0 1 1 0 1 1 0

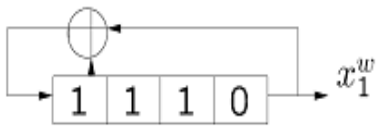
1 0 1 0 0 1 1 1 0 1 0 0 1 1 1 0 1 0 0 1 1 1 0 1 0 0

0 0 1 1 1 1 0 1 0 1 1 0 0 1 0 0 0 1 1 1 1 0 1 0 1 1

1 0 0 1 1 0 1 1 0 0 1 0 1 1 1 0 0 1 1 1 0 1 1 1 1 1

$f(x) \oplus x_1$

1 0 1 0 0 1 1 0 0 1 0 0 1 0 1 0 0 0 0 0 1 1 0 1 0 0 ← 16 zeros



wrong key

0 1 1 1 1 0 1 0 1 1 0 0 1 0 0 0 1 1 1 1 0 1 0 1 1 0

$f(x) \oplus x_1^w$

1 1 1 0 0 0 0 1 1 1 1 0 0 1 1 0 1 0 0 0 0 0 1 0 0 1

Extended affine (EA) equivalence

- The most important parameters (nonlinearity, degree and algebraic properties) invariant under transformation

$$g(x) = f(xL + b) + cx + d,$$

where L is invertible binary $n \times n$ matrix; $b, c \in \mathbb{F}_2^n$ and $d \in \mathbb{F}_2$.

- **Algebraic thickness** - the sparsest ANF of a given function !

Extended affine (EA) equivalence

- The most important parameters (nonlinearity, degree and algebraic properties) invariant under transformation

$$g(x) = f(xL + b) + cx + d,$$

where L is invertible binary $n \times n$ matrix; $b, c \in \mathbb{F}_2^n$ and $d \in \mathbb{F}_2$.

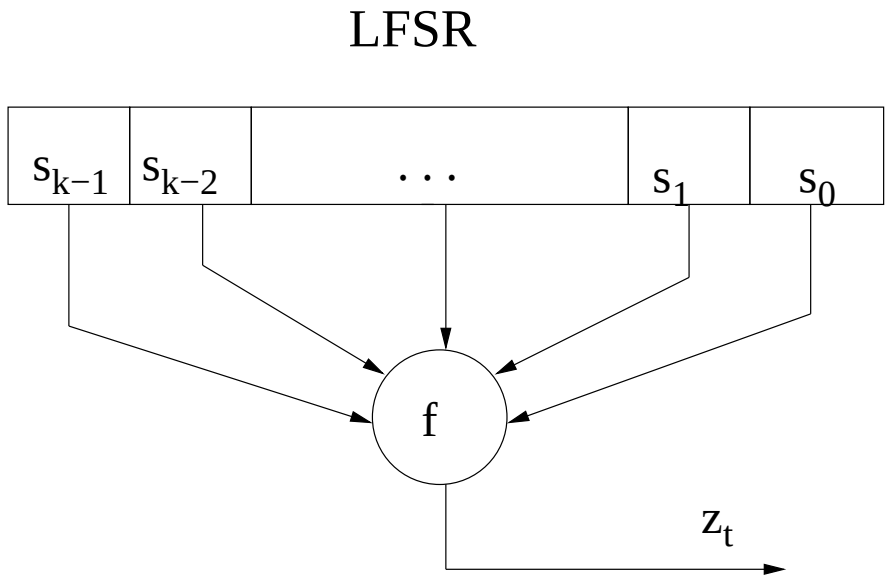
- **Algebraic thickness** - the sparsest ANF of a given function !

EXAMPLE : $f(x_1, \ldots, x_n) = (x_1 + 1)(x_2 + 2) \cdots (x_n + 1)$ has 2^n terms in its ANF. Its (cryptographically useless) equivalent is $g(x) = x_1 x_2 \cdots x_n$ has one term !!

PROBLEM Given f find its best affine equivalent with sparsest ANF !!!

Extended affine equivalence II

- Since **resiliency** of degree t (balancedness of $f(x) + \ell(x)$ with $wt(\ell(x)) \leq t$) is **not invariant** under EA transformation easier to design f for :



Nonlinear filtering generator

- No resiliency needed just balancedness of f !!

Cryptographic functions with non-efficient implementation

- **CARLET-FENG 2008** : Let $n \geq 2$ and α a primitive element of $\mathbb{F}_{2^n}$. Define f on $\mathbb{F}_{2^n}$ whose (balanced) support is

$$S = \{0, 1, \alpha, \ldots, \alpha^{2^{n-1}} - 2\}, \quad f(x) = 1 \leftrightarrow x \in S.$$

- f satisfies all of the cryptographic criteria (high degree and nonlinearity, good algebraic properties) ...

- But Carlet estimated that for $n = 18$ c.a. 40 000 transistors needed for implementation !!

- Need much more efficient implementation otherwise no practical use of these functions (e.g. CARLET-FENG) , use compact block ciphers instead !

Efficient hardware implementation

- Hard to expect that EA-equivalence would help here !
- To resist algebraic attacks number of variables $n \geq 16$ approx. !
- A random Boolean function has in average 2^{n-1} terms in its ANF !!
- Need some special classes with efficient implementation such as Maiorana-McFarland class (concatenation of affine/linear functions)

Basic ideas behind MM class

- Any linear function $\ell_n(x)$ in n variables can be seen as "linear" concatenation of some fixed affine couple $\ell_p(x)$ and $1+\ell_p(x)$ in p variables.
- Consider any linear function

$$\ell_n(x_1,\ldots,x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n.$$

$$\ell_n(x_1,\ldots,x_n) = \begin{cases} \ell_{n-1}(x_1,\ldots,x_{n-1})||\ell_{n-1}(x_1,\ldots,x_{n-1}); & a_n = 0, \\ \ell_{n-1}(x_1,\ldots,x_{n-1})||1+\ell_{n-1}(x_1,\ldots,x_{n-1}); & a_n = 1, \end{cases}$$

where $\ell_{n-1}(x_1,\ldots,x_{n-1}) = a_1x_1 + \ldots + a_{n-1}x_{n-1}$. Proceed with induction.

- **CONCLUSION:** We can measure the distance of $f(x)$ to $l_n(x)$ by looking at subfunctions in smaller dimension.

Maiorana-McFarland Construction

For any positive integers p, q such that $n = p + q$, an MM function $f \in \mathcal{B}_n$ is defined by

$$f(Y_q, X_p) = \phi(Y_q) \cdot X_p \oplus g(Y_q), \ X_p \in \mathbb{F}_2^p, \ Y_q \in \mathbb{F}_2^q, \tag{4}$$

where ϕ is any mapping from $\mathbb{F}_2^q$ to $\mathbb{F}_2^p$, $g \in \mathcal{B}_q$.

Maiorana-McFarland Construction

For any positive integers p, q such that $n = p + q$, an MM function $f \in \mathcal{B}_n$ is defined by

$$f(Y_q, X_p) = \phi(Y_q) \cdot X_p \oplus g(Y_q), \ X_p \in \mathbb{F}_2^p, \ Y_q \in \mathbb{F}_2^q, \tag{4}$$

where ϕ is any mapping from $\mathbb{F}_2^q$ to $\mathbb{F}_2^p$, $g \in \mathcal{B}_q$.

- For any fixed Y_q the function f is affine/linear (depending on g)

Maiorana-McFarland Construction

For any positive integers p, q such that $n = p + q$, an MM function $f \in \mathcal{B}_n$ is defined by

$$f(Y_q, X_p) = \phi(Y_q) \cdot X_p \oplus g(Y_q), \ X_p \in \mathbb{F}_2^p, \ Y_q \in \mathbb{F}_2^q, \tag{4}$$

where ϕ is any mapping from $\mathbb{F}_2^q$ to $\mathbb{F}_2^p$, $g \in \mathcal{B}_q$.

- For any fixed Y_q the function f is affine/linear (depending on g)
- If ϕ is **injective** then $p \geq q$ and affine subfunctions in p variables are distinct !

Maiorana-McFarland Construction

For any positive integers p, q such that $n = p + q$, an *MM* function $f \in \mathcal{B}_n$ is defined by

$$f(Y_q, X_p) = \phi(Y_q) \cdot X_p \oplus g(Y_q),\ X_p \in \mathbb{F}_2^p,\ Y_q \in \mathbb{F}_2^q, \tag{4}$$

where ϕ is any mapping from $\mathbb{F}_2^q$ to $\mathbb{F}_2^p$, $g \in \mathcal{B}_q$.

- For any fixed Y_q the function f is affine/linear (depending on g)
- If ϕ is **injective** then $p \geq q$ and affine subfunctions in p variables are distinct !
- Calculation of Hamming distance of f to any ℓ_n becomes easy (injectivity implies either a single or no match with linear functions in p variables)
- For a single match (if no match $d_H(f, \ell_n) = 2^{n-1}$)

$$d_H(f, \ell_n) = (2^q - 1)2^{p-1} + 0 = 2^{n-1} - 2^{p-1}.$$

Special case - bent functions

- Defined for $n = 2k$ even, bent functions achieve maximum nonlinearity.

- Essentially, $k = p = q$ and we concatenate all 2^k linear functions in k variables,

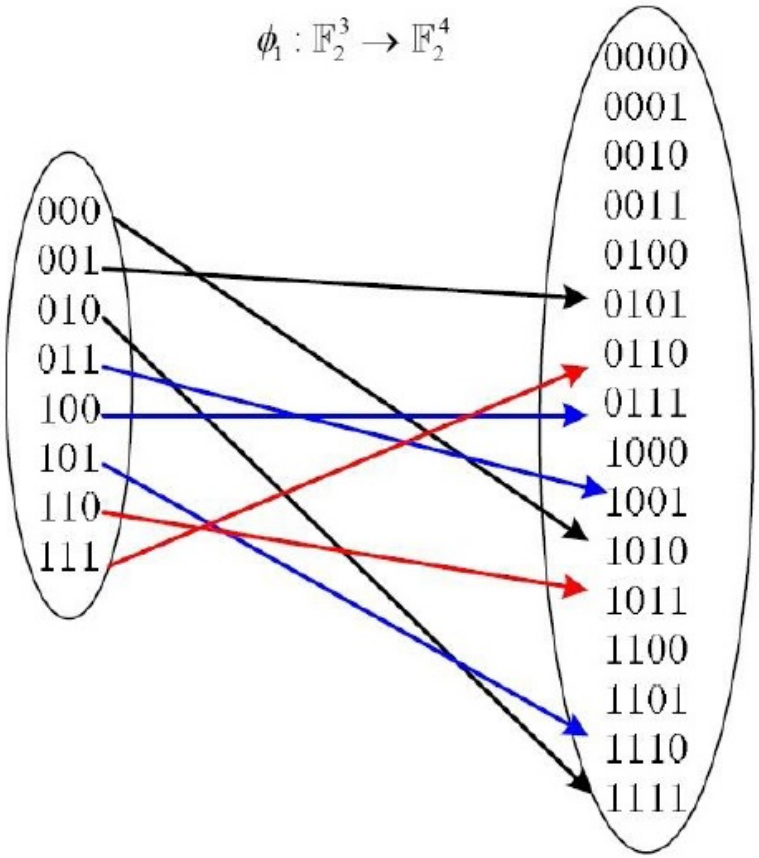
$$b(x) = \ell^{(0)}(X_k)||\ell^{(1)}(X_k)||\cdots||\ell^{(2^k-1)}(X_k),$$

where $\ell^{(i)}(X_k) = c^{(i)} \cdot X_k$ and $c^{(i)}$ is a binary k-bit representation of integer $i = 0, 1, \ldots, 2^k - 1$.

- In this case

$$d_H(f, \ell_n) = 2^{n-1} \pm 2^{n/2-1}.$$

Example of construction



$$
\begin{aligned}
f_1(Y_3, X_4) = {} & \overline{y_1}\,\overline{y_2}\,\overline{y_3}\,(x_1 \oplus x_3) \\
& \oplus \overline{y_1}\,\overline{y_2} y_3\,(x_2 \oplus x_4) \\
& \oplus \overline{y_1} y_2\,\overline{y_3}\,(x_1 \oplus x_2 \oplus x_3 \oplus x_4) \\
& \oplus \overline{y_1} y_2 y_3\,(x_1 \oplus x_4) \\
& \oplus y_1\,\overline{y_2}\,\overline{y_3}\,(x_2 \oplus x_3 \oplus x_4) \\
& \oplus y_1\,\overline{y_2} y_3\,(x_1 \oplus x_2 \oplus x_3) \\
& \oplus y_1 y_2\,\overline{y_3}\,(x_1 \oplus x_3 \oplus x_4) \\
& \oplus y_1 y_2 y_3\,(x_2 \oplus x_3)
\end{aligned}
$$

- Easy to get high nonlinearity; moderate alg. degree $\leq q + 1$
- Easy to construct t resilient functions - take linear functions with at least $t + 1$ terms.

Properties of the construction

Main drawbacks :

- Linearity of subfunctions on relatively large variable space may give rise to other attacks !
- Relatively bad algebraic properties (resistance to (fast) algebraic attacks)

Properties of the construction

Main drawbacks :

- Linearity of subfunctions on relatively large variable space may give rise to other attacks !
- Relatively bad algebraic properties (resistance to (fast) algebraic attacks)

Main advantages : Good nonlinearity and efficient hardware implementation. We implement linear functions in p variables, only smart addressing of these needed.

Properties of the construction

Main drawbacks :

- Linearity of subfunctions on relatively large variable space may give rise to other attacks !
- Relatively bad algebraic properties (resistance to (fast) algebraic attacks)

Main advantages : Good nonlinearity and efficient hardware implementation. We implement linear functions in p variables, only smart addressing of these needed.

EXAMPLE : For bent functions we implement $2^{n/2}$ linear functions using one-to-one direct addressing.

MM class uses direct product $\mathbb{F}_2^q \times \mathbb{F}_2^p = \mathbb{F}_2^n$ to decompose the space $\mathbb{F}_2^n$. But we can decompose the space in many different ways !!!

Generalized Maiorana-McFarland (GMM) Construction

Let for $1 \leq i \leq n-1$, $E_i \subseteq \mathbb{F}_2^i$ and $E_i' = E_i \times \mathbb{F}_2^{n-i}$ such that

$$\bigcup_{i=1}^{n-1} E_i' = \mathbb{F}_2^n, \tag{5}$$

and $E_{i_1}' \cap E_{i_2}' = \emptyset, \quad 1 \leq i_1 < i_2 \leq n-1$.

- Let $X_i' = (x_1, \ldots, x_i) \in \mathbb{F}_2^i$ and $X_{n-i}'' = (x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-i}$.
- Let ϕ_i be a mapping from E_i to $\mathbb{F}_2^{n-i}$ and $g_i \in \mathcal{B}_i$ arbitrary.

Generalized Maiorana-McFarland (GMM) Construction

Let for $1 \leq i \leq n-1$, $E_i \subseteq \mathbb{F}_2^i$ and $E_i' = E_i \times \mathbb{F}_2^{n-i}$ such that

$$\bigcup_{i=1}^{n-1} E_i' = \mathbb{F}_2^n, \tag{5}$$

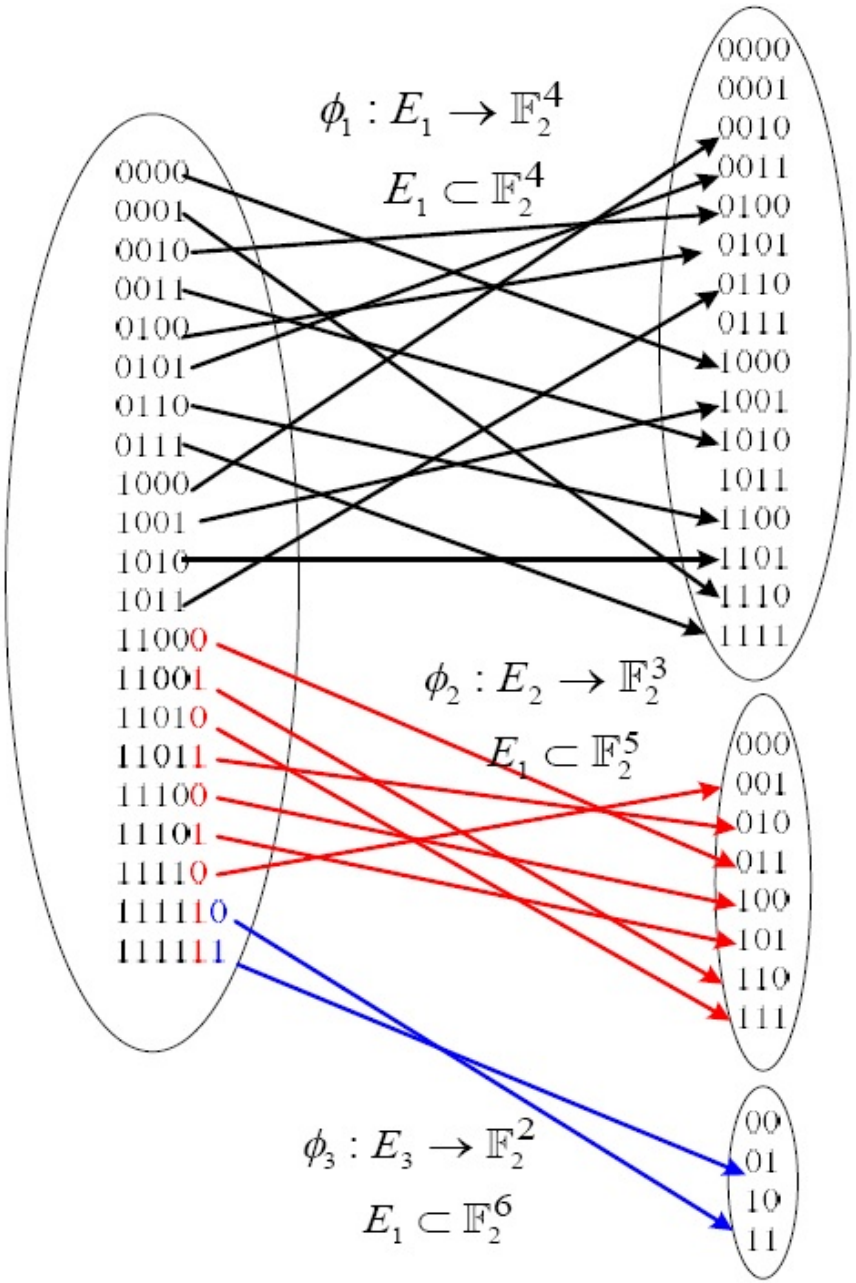
and $E_{i_1}' \cap E_{i_2}' = \emptyset, \quad 1 \leq i_1 < i_2 \leq n-1$.

- Let $X_i' = (x_1, \ldots, x_i) \in \mathbb{F}_2^i$ and $X_{n-i}'' = (x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-i}$.
- Let ϕ_i be a mapping from E_i to $\mathbb{F}_2^{n-i}$ and $g_i \in \mathcal{B}_i$ arbitrary.
- A Boolean function $f \in \mathcal{B}_n$ in GMM class can be constructed as follows:

$$f(X_n) = \phi_i(X_i') \cdot X_{n-i}'' \oplus g_i(X_i'), \text{ if } X_i' \in E_i, \ i = 1, \ldots, n-1. \tag{6}$$

- Wei-Guo Zhang, Enes Pasalic, Generalized Maiorana-McFarland construction of resilient Boolean functions with high nonlinearity and good algebraic properties, IEEE Transactions on Information Theory, vol. 60, no. 10, pp. 6681-6695, 2014.

Example: An 8-variable balanced GMM function



$$
\begin{aligned}
f_1(X_8) = {} & \overline{x_1}\,\overline{x_2}\,\overline{x_3}\,\overline{x_4}\,(x_5) \\
\oplus\ & \overline{x_1}\,\overline{x_2}\,\overline{x_3}x_4\,(x_5 \oplus x_6 \oplus x_7) \\
\oplus\ & \overline{x_1}\,\overline{x_2}x_3\,\overline{x_4}\,(x_6) \\
\oplus\ & \overline{x_1}\,\overline{x_2}x_3x_4\,(x_5 \oplus x_7) \\
\oplus\ & \overline{x_1}x_2\,\overline{x_3}\,\overline{x_4}\,(x_6 \oplus x_8) \\
\oplus\ & \overline{x_1}x_2\,\overline{x_3}x_4\,(x_7 \oplus x_8) \\
\oplus\ & \overline{x_1}x_2x_3\,\overline{x_4}\,(x_5 \oplus x_6) \\
\oplus\ & \overline{x_1}x_2x_3x_4\,(x_5 \oplus x_6 \oplus x_7 \oplus x_8) \\
\oplus\ & x_1\overline{x_2}\,\overline{x_3}\,\overline{x_4}\,(x_7) \\
\oplus\ & x_1\overline{x_2}\,\overline{x_3}x_4\,(x_5 \oplus x_8) \\
\oplus\ & x_1\overline{x_2}x_3\overline{x_4}\,(x_5 \oplus x_6 \oplus x_8) \\
\oplus\ & x_1\overline{x_2}x_3x_4\,(x_6 \oplus x_7) \\
\oplus\ & x_1x_2\,\overline{x_3}\,\overline{x_4}\,\overline{x_5}\,(x_7 \oplus x_8) \\
\oplus\ & x_1x_2\,\overline{x_3}\,\overline{x_4}\,x_5\,(x_6 \oplus x_7) \\
\oplus\ & x_1x_2\,\overline{x_3}\,x_4\,\overline{x_5}\,(x_6 \oplus x_7 \oplus x_8) \\
\oplus\ & x_1x_2\,\overline{x_3}\,x_4\,x_5\,(x_7) \\
\oplus\ & x_1x_2x_3\,\overline{x_4}\,\overline{x_5}\,(x_6) \\
\oplus\ & x_1x_2x_3\,\overline{x_4}\,x_5\,(x_6 \oplus x_8) \\
\oplus\ & x_1x_2x_3x_4\,\overline{x_5}\,(x_8) \\
\oplus\ & x_1x_2x_3x_4\,x_5\,\overline{x_6}\,(x_7 \oplus x_8) \\
\oplus\ & x_1x_2x_3x_4\,x_5\,x_6\,(x_8)
\end{aligned}
$$

Theorem: Let n be **even** and $E_i = \emptyset$, for $1 \leq i \leq n/2 - 1$. Let $0 \leq m \leq n/2 - 2$, and $(a_{n/2}, \ldots, a_{n-m-1}) \in \mathbb{F}_2^{n/2-m}$ be the binary vector such that $\sum_{i=n/2}^{n-m-1} a_i 2^i$ is maximal, satisfying at the same time,

$$\sum_{i=n/2}^{n-m-1} \left(a_i \cdot 2^{n-i} \sum_{j=m+1}^{n-i} \binom{n-i}{j} \right) \geq 2^n. \tag{7}$$

Let $e = \max\{i \mid a_i \neq 0,\ n/2 \leq i \leq n-m-1\}$. For $n/2 \leq i \leq e-1$, set

$$|E_i| = \begin{cases} 0, & \text{if } a_i = 0 \\ \sum_{j=m+1}^{n-i} \binom{n-i}{j}, & \text{if } a_i = 1. \end{cases} \tag{8}$$

Theorem: Let n be **even** and $E_i = \emptyset$, for $1 \leq i \leq n/2 - 1$. Let $0 \leq m \leq n/2 - 2$, and $(a_{n/2}, \ldots, a_{n-m-1}) \in \mathbb{F}_2^{n/2-m}$ be the binary vector such that $\sum_{i=n/2}^{n-m-1} a_i 2^i$ is maximal, satisfying at the same time,

$$\sum_{i=n/2}^{n-m-1} \left(a_i \cdot 2^{n-i} \sum_{j=m+1}^{n-i} \binom{n-i}{j} \right) \geq 2^n. \tag{7}$$

Let $e = \max\{i \mid a_i \neq 0,\ n/2 \leq i \leq n-m-1\}$. For $n/2 \leq i \leq e-1$, set

$$|E_i| = \begin{cases} 0, & \text{if } a_i = 0 \\ \sum_{j=m+1}^{n-i} \binom{n-i}{j}, & \text{if } a_i = 1. \end{cases} \tag{8}$$

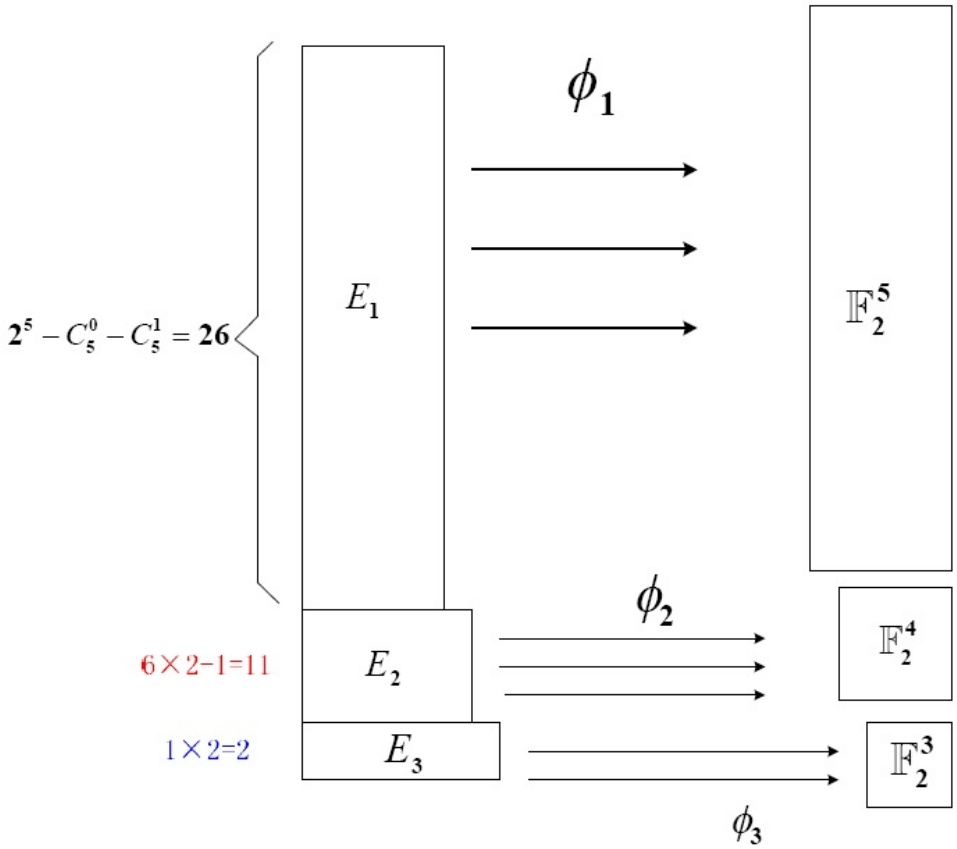
For $n/2 \leq i \leq e$ and $a_i = 1$, let $\phi_i : E_i \to T_i$ be **injective** mapping where

$$T_i = \{c \mid wt(c) \geq m+1,\ c \in \mathbb{F}_2^{n-i}\}. \tag{9}$$

Then the function $f \in \mathcal{B}_n$ is a SAO m-resilient function with nonlinearity

$$N_f \geq 2^{n-1} - 2^{n/2-1} - \sum_{i=n/2+1}^{e} a_i \cdot 2^{n-i-1} > 2^{n-1} - 2^{n/2}. \tag{10}$$

Construction of a GMM $(10, 1, 8, 484)$ function (SAO)



$$1024 \text{ bits } = 26 \times 32 + 11 \times 16 + 2 \times 8 \text{ bits } .$$

COMPARISON: For MM class injectivity of ϕ and resiliency imply that $q = 4$ and $p = 6$. Then, $N_f = 2^{n-1} - 2^{p-1} = 480$, and $\deg(f) \leq 5$!

Cryptographic properties of GMM class

Main advantages :

- Better nonlinearity and degree than MM class
- Much richer class - many decompositions of $\mathbb{F}_2^n$!
- Efficient implementation unless the decomposition of $\mathbb{F}_2^n$ is too complex

Main drawbacks :

- Relatively bad resistance against fast algebraic attacks (existence of low degree g, h s.t. $fg = h$.

Cryptographic properties of GMM class

Main advantages :

- Better nonlinearity and degree than MM class
- Much richer class - many decompositions of $\mathbb{F}_2^n$!
- Efficient implementation unless the decomposition of $\mathbb{F}_2^n$ is too complex

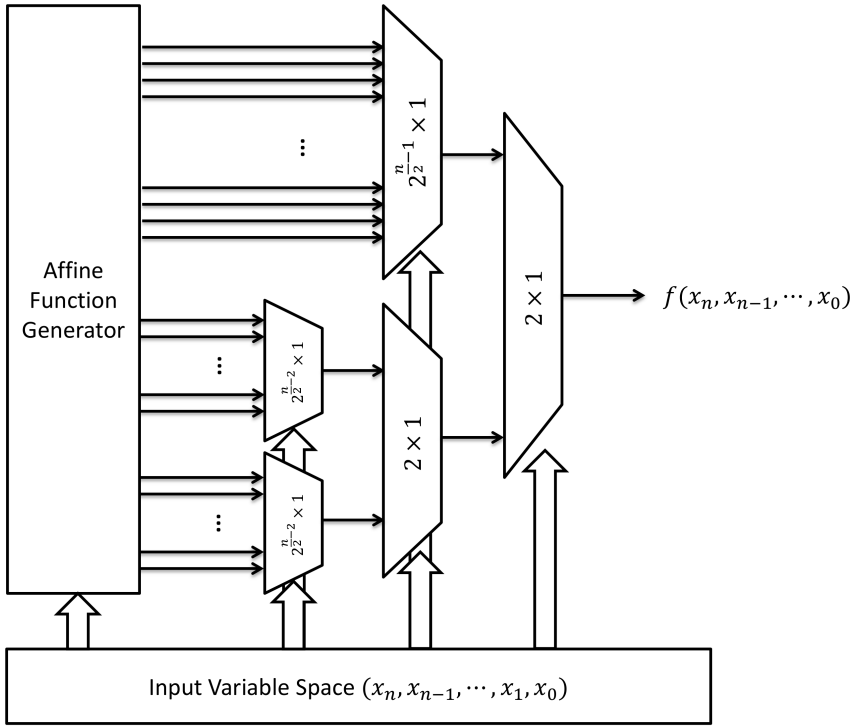
Main drawbacks :

- Relatively bad resistance against fast algebraic attacks (existence of low degree g, h s.t. $fg = h$.

Trade-off necessary - **DO NOT use too many** $n/2$**-variable linear functions** and get better resistance to (fast) algebraic attacks BUT nonlinearity decreases !!!

- Better understanding regarding fast algebraic attacks needed !

Implementation results



Design	Total Area (Gates)	Clock Period (ns)	Clock Frequency (MHz)
10-variable	167	2.11	473.93
12-variable	409	2.63	380.23
16-variable	1538	3.59	278.55

Table : GMM Construction 2 ASIC Synthesis Results

THANK YOU FOR

YOUR ATTENTION !