Exploring relationships between pseudorandomness properties of sequences and cryptographic properties of Boolean functions

Konstantinos Limniotis

Hellenic Data Protection Authority,
Kifissias 1-3,
11523 Athens, Greece
Email: klimniotis@dpa.gr

Dept. of Informatics & Telecommunications,
National and Kapodistrian University of Athens,
15784 Athens, Greece
Email: klimn@di.uoa.gr

Athens Cryptography Day 2019 -
National Technical University of Athens

January 8th, 2019, Athens, Greece

Talk Outline

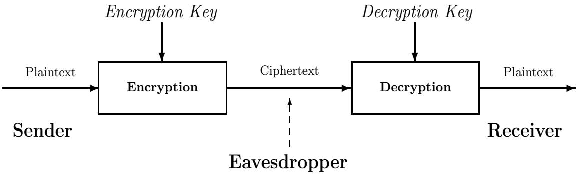
1. Introduction
   - Cryptographic properties of Boolean functions
   - Error linear complexity spectrum of sequences
     - The Games-Chan algorithm
     - The Lauder-Paterson algorithm
2. Investigating relationships

   Joint work with N. Kolokotronis (submitted - under review)

   - Bijection between 2^n-periodic binary sequences and Boolean functions on n variables
   - Properties of the error linear complexity spectrum provides information on how well a function can be approximated by a simpler function
     - with fewer number of variables
     - with lower degree
3. Conclusions

Symmetric ciphers

A typical cryptosystem



Symmetric cryptography

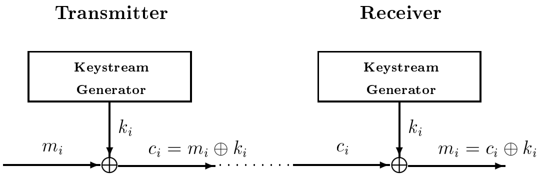
- Encryption Key = Decryption Key
- The key is only shared between the two parties
  - The security rests with the secrecy of the key (Kerchoffs principle)
  - Post-quantum resistant (for appropriate key sizes)

Two types of symmetric ciphers

- Stream ciphers
- Block ciphers

Stream ciphers

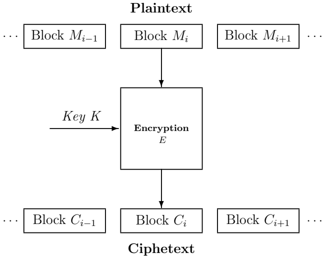
Simplest Case: Binary additive stream cipher



- Suitable in environments characterized by a limited computing power or memory, and the need to encrypt at high speed
- The seed of the keystream generators constitutes the secret key
- Security depends on
  - Pseudorandomness of the keystram k_i
  - Properties of the underlying functions that form the keystream generator

Block ciphers

Simplest Case: Electronic Codebook Mode of operation (ECB)



- Encryption on a per-block basis (typical block size: 128 bits)
- Several drawbacks of the ECB - Other modes of operation are being used in practice (CTR, GCM etc.)
  - Some modes resemble the operation of stream ciphers - the encryption function E stands as a keystream generator
- Current research trend: Authenticated cipher (CAESAR)

A common approach for block and stream ciphers

- Despite their differences, a common study is needed for their building blocks (multi-output and single-output Boolean functions)
- The attacks in block ciphers are, in general, different from the attacks in stream ciphers and vice versa. However:
  - For both cases, almost the same cryptographic criteria of functions should be in place
- Challenges:
  - There are tradeoffs between several cryptographic criteria
  - The relationships between several criteria are still unknown
  - How to construct functions that are mathematically bound to satisfy all the main criteria
  - New attacks $\Rightarrow$ New criteria

Boolean Functions

A Boolean function f on n variables ($f \in \mathbb{B}_n$) is a mapping from $\mathbb{F}_2^n$ onto $\mathbb{F}_2$

- The vector $f = \big(f(0,0,\ldots,0), f(1,0,\ldots,0),\ldots,f(1,1,\ldots,1)\big)$ of length 2^n is the truth table of f
- The Hamming weight of f is denoted by $\mathrm{wt}(f)$
  - f is balanced if and only if $\mathrm{wt}(f) = 2^{n-1}$
- The support $\mathrm{supp}(f)$ of f is the set $\{\boldsymbol{b} \in \mathbb{F}_2^n : f(\boldsymbol{b}) = 1\}$

Example: Truth table of balanced f with $n = 3$

x_1	0	1	0	1	0	1	0	1
x_2	0	0	1	1	0	0	1	1
x_3	0	0	0	0	1	1	1	1
$f(x_1,x_2,x_3)$	0	1	0	0	0	1	1	1

A vectorial Boolean function F is a mapping from $\mathbb{F}_2^n$ onto $\mathbb{F}_2^m$, $m > 1$

Algebraic Normal Form and degree of functions

- Algebraic Normal Form (ANF) of f:

$$f(x) = \sum_{\boldsymbol{v} \in \mathbb{F}_2^n} a_{\boldsymbol{v}} x^{\boldsymbol{v}}, \quad \text{where } x^{\boldsymbol{v}} = \prod_{i=1}^{n} x_i^{v_i}$$

.

  - The sum is performed over $\mathbb{F}_2$ (XOR addition)
- The degree $\deg(f)$ of f is the highest number of variables that appear in a product term in its ANF.
- If $\deg(f) = 1$, then f is called affine function
  - If, in addition, the constant term is zero, then the function is called linear
- In the previous example: $f(x_1, x_2, x_3) = x_1 x_2 \oplus x_2 x_3 \oplus x_1$.
- $\deg(f) = 2$

Univariate representation of Boolean functions

- $\mathbb{F}_2^n$ is isomorphic to the finite field $\mathbb{F}_{2^n}$,
- $\Rightarrow$ Any function $f \in \mathbb{B}_n$ can also be represented by a univariate polynomial, mapping $\mathbb{F}_{2^n}$ onto $\mathbb{F}_2$, as follows

$$f(x) = \sum_{i=0}^{2^n-1} \beta_i x^i$$

where $\beta_0, \beta_{2^n-1} \in \mathbb{F}_2$ and $\beta_{2i} = \beta_i^2 \in \mathbb{F}_{2^n}$ for $1 \leq i \leq 2^n - 2$

- The coefficients of the polynomial determine the Discrete Fourier Transform of f
- The degree of f can be directly deduced by the univariate representation
- The univariate representation is more convenient in several cases

Walsh transform

Definition

The Walsh transform $\widehat{\chi}_f(\boldsymbol{a})$ at $\boldsymbol{a} \in \mathbb{F}_2^n$ of $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is

$$\widehat{\chi}_f(\boldsymbol{a}) = \sum_{\boldsymbol{x} \in \mathbb{F}_2^n} (-1)^{f(\boldsymbol{x}) \oplus \boldsymbol{a}\boldsymbol{x}^T} = 2^n - 2\,\mathrm{wt}(f \oplus \phi_{\boldsymbol{a}})$$

where $\phi_{\boldsymbol{a}}(\boldsymbol{x}) = \boldsymbol{a}\boldsymbol{x}^T = a_1 x_1 \oplus \cdots \oplus a_n x_n$

- Computational complexity: $\mathcal{O}(n2^n)$ (via fast Walsh transform)
- Parseval's theorem: $\sum_{a \in \mathbb{F}_2^n} \widehat{\chi}_f(a)^2 = 2^{2n}$

Cryptographic properties

Apart from the balancedness and the high algebraic degree, other important cryptographic criteria are the following:

- Correlation immunity
- Existence of linear structures
- Nonlinearity
  - Higher-order nonlinearity
- Minimum Hamming distance from a function with fewer number of variables
- (Fast) algebraic immunity

More recently, the structure of specific ciphers (e.g. the FLIP stream cipher) necessitates the study of appropriate modifications of (some of) the above criteria (Carlet, 2017).

Correlation immunity

- If the output of a Boolean function f is correlated to at least one of its inputs, then it is vulnerable to correlation attacks (Siegenthaler, 1984).
- The $f \in \mathbb{B}_n$ is t-th correlation immune if it is not correlated with any t-subset of $\{x_1, \ldots, x_n\}$; namely if

$$Pr(f(\boldsymbol{x}) = 0 | x_{i_1} = b_{i_1}, \ldots, x_{i_t} = b_{it}) = Pr(f(\boldsymbol{x}) = 0)$$

  for any t positions $x_{i_1}, \ldots, x_{i_t}$ and any $b_{i_1}, \ldots, b_{i_t} \in \mathbb{F}_2$
- If a t-th order correlation immune function is also balanced, then it is called t-th order resilient.

Properties of correlation immunity

- Siegenthaler, 1984: A known trade-off: If f is k-th order resilient for $1 \leq k \leq n-2$, then $\deg(f) \leq n-k-1$.
- Xiao-Massey, 1988: A function $f \in \mathbb{B}_n$ is t-th order correlation immune iff its Walsh transform satisfies

$$\widehat{\chi}_f(a) = 0, \forall\ 1 \leq \mathrm{wt}(a) \leq t$$

  - Note that f is balanced iff $\widehat{\chi}_f(\mathbf{0}) = 0$.
  - $\Rightarrow$ A function $f \in \mathbb{B}_n$ is t-th order resilient iff its Walsh transform satisfies $\widehat{\chi}_f(a) = 0, \forall\ 0 \leq \mathrm{wt}(a) \leq t$
- Siegenthaler also proposed a recursive procedure to construct m-th order resilient Boolean functions, for any desired m, with the maximum possible degree
- Several other constructions are currently known

Linear structures

- The derivative of f in the direction of the vector $\boldsymbol{a} \in \mathbb{F}_2^n$ is given by

$$D_{\boldsymbol{a}}(f(\boldsymbol{x})) = f(\boldsymbol{x}) \oplus f(\boldsymbol{x} \oplus \boldsymbol{a}).$$

- A vector $\boldsymbol{a} \in \mathbb{F}_2^n$ is called a linear structure of f if the derivative $D_{\boldsymbol{a}}(f)$ is constant.
- Boolean functions used in symmetric ciphers should avoid nonzero linear structures.
    - To thwart, e.g. differential cryptanalysis

The linear kernel of f

- The set of linear structures of f constitutes the so-called linear kernel of f, being a subspace of $\mathbb{F}_2^n$.
- A Boolean function admits a nonzero linear structure if and only if it is linear equivalent to a function of the form $f(x_1, \ldots, x_n) = g(x_1, \ldots, x_{n-1}) \oplus \epsilon x_n$.
- More generally, its linear kernel has dimension at least k if and only if it is linearly equivalent to a function of the form:

$$f(x_1, \ldots, x_n) = g(x_1, \ldots, x_{n-k}) \oplus \epsilon_{n-k+1} x_{n-k+1} \oplus \ldots \oplus \epsilon_n x_n \ ,$$

$$\epsilon_{n-k+1}, \ldots, \epsilon_n \in \mathbb{F}_2.$$

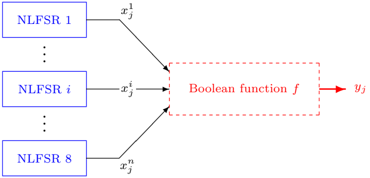
Linear approximation attacks

- The maximum possible degree of a balanced Boolean function with n variables is $n-1$
- High degree though is not adequate to prevent linear cryptanalysis (in block ciphers - Matsui, 1992) or best affine approximation attacks (in stream ciphers - Ding et. al., 1991)
- A function should not be well approximated by a linear/affine function
- Any function of degree 1 that best approximates f is a best affine/linear approximation of f
- An equivalent notion of describing the Hamming distance between two Boolean functions f, g is the so-called *bias* ϵ:

$$\epsilon = |p(f(\boldsymbol{x}) = g(\boldsymbol{x})) - 1/2|$$

Example of approximation attacks

The Achterbahn cipher [Gammel-Göttfert-Kniffler,2005] (candidate in eSTREAM project)



- Lengths of nonlinear FSRs: 22-31
- $f(x_1, \ldots, x_8) = \sum_{i=1}^{4} x_i \oplus x_5x_7 \oplus x_6x_7 \oplus x_6x_8 \oplus x_5x_6x_7 \oplus x_6x_7x_8$
- Johansson-Meier-Muller, 2006: cryptanalysis via the linear approximation $g(x_1, \ldots, x_8) = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_6$, satisfying $\mathrm{wt}(f \oplus g) = 64$ $(p(f = g) = 3/4,\ \epsilon = 0.25)$

The notion of nonlinearity

- The minimum distance between f and all affine functions is the nonlinearity of f:

$$\mathsf{nl}(f) = \min_{l \in \mathbb{B}_n : \deg(l)=1} \mathrm{wt}(f \oplus l)$$

- Relathionship with Walsh transform

$$\mathsf{nl}(f) = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_2^n} |\widehat{\chi}_f(a)|$$

  - $\Rightarrow$ Nonlinearity is computed via the Fast Walsh Transform

- High nonlinearity is prerequisite for thwarting attacks based on affine (linear) approximations

Known results on nonlinearity of Boolean functions

- For even n, the maximum possible nonlinearity is $2^{n-1} - 2^{n/2-1}$, achieved by the so-called bent functions
  - Many constructions are known (not fully classified yet)
  - But bent functions are never balanced!
- For odd n, the maximum possible nonlinearity is still unknown
  - By concatenating bent functions, we can get nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$. Can we impove this?
    - For $n \leq 7$, the answer is no
    - For $n \geq 15$, the answer is yes (Patterson-Wiedemann, 1983 - Dobbertin, 1995 - Maitra-Sarkar, 2002)
    - For $n = 9, 11, 13$, such functions have been found (Kavut, 2006)
- Several constructions of balanced functions with high nonlinearity exist (e.g. Dobbertin, 1995). However:
  - Finding the highest possible nonlinearity of balanced Boolean functions is still an open problem

Higher-order nonlinearity

- Approximating a function by a low-order function (not necessarily linear) may also lead to cryptanalysis (Non–linear cryptanalysis - Knudsen-1996, low-order approximation attacks - Kurosawa et. al. - 2002)
- The rth order nonlinearity of a Boolean function $f \in \mathbb{B}_n$ is given by

$$\mathsf{nl}_r(f) = \min_{g \in \mathbb{B}_n : \deg(g) \leq r} \mathrm{wt}(f \oplus g)$$

- The rth order nonlinearity remains unknown for $r > 1$
  - Recursive lower bounds on $\mathsf{nl}_r(f)$ (Carlet, 2008)
  - Specific lower and upper bounds for $\mathsf{nl}_2(f)$ (Cohen, 1992 - Carlet, 2007)
  - More recent lower bounds for 2-nd order nonlinearity: Gangopadhyay et. al. - 2010, Garg et. al. - 2011, Singh - 2011, Singh et. al. - 2013

Computing best low order approximations

- Computing even the best 2-nd order approximations is a difficult task
  - Efficient solution for specific class of 3-rd degree functions (Kolokotronis-Limniotis-Kalouptsidis, 2009)
  - For the Achterbahn's combiner function: $q(x) = x_5x_7 \oplus x_6x_8 \oplus x_1 \oplus x_2 \oplus x_3 \oplus x_4$ is a best 2-nd approximation ((Limniotis, 2007))
  - $\mathrm{wt}(f+q) = 32$ ($p(f=q) = 7/8 > 3/4$, $\epsilon = 0.375$)
- No much is known regarding constructions of functions with high r-th nonlinearity, for $r \geq 2$
  - Even if a high lower bound on the nonlinearity is proved, best r-th order approximations cannot be computed
  - A class of highly nonlinear 3-rd degree functions satisfying $\mathsf{nl}_2(f) = \mathsf{nl}(f)$ (Kolokotronis-Limniotis, 2012)

Approximation by a function depending on fewer variables

- Exploiting an approximation of a cryptographic Boolean function by a function of fewer variables may result in specific attacks, such as divide-and-conquer attacks (Canteaut et. al., 2002)
- If $f \in \mathbb{B}_n$ depends only on $k < n$ variables, then we say that $f \in \mathbb{B}_n(k)$
  - Linearly equivalent to a function g depending on $x_1, x_2, \ldots, x_k$
  - The linear kernel of f has dimension $n-k$ (if $g \in \mathbb{B}_k$ has no linear structures).
- A function with high nonlinearity cannot be efficiently approximated by other function depending on a small subset of its input variables (Canteaut et. al., 2002)
- If $f \in \mathbb{B}_n$ is a t-resilient function, then:

$$d_H(f, \mathbb{B}_n(k)) \geq 2^{n-1} - \frac{\max_{\boldsymbol{a} \in \mathbb{F}_2^n} |\widehat{\chi}_f(\boldsymbol{a})|}{2} \left(\sum_{i=t+1}^{k} \binom{k}{i} \right)^{1/2}$$

Annihilators and algebraic immunity

Definition

Given $f \in \mathbb{B}_n$, we say that $g \in \mathbb{B}_n$ is an annihilator of f if and only if g lies in the set

$$\mathcal{AN}(f) = \{g \in \mathbb{B}_n : f * g = 0\}$$

Definition

The algebraic immunity $\mathsf{AI}_n(f)$ of $f \in \mathbb{B}_n$ is defined by

$$\mathsf{AI}_n(f) = \min_{g \neq 0}\{\deg(g) : \ g \in \mathcal{AN}(f) \cup \mathcal{AN}(f \oplus 1)\}$$

- A high algebraic immunity is prerequisite for preventing algebraic attacks (Meier-Pasalic-Carlet, 2004)
- Well-known upper bound: $\mathsf{AI}_n(f) \leq \lceil \frac{n}{2} \rceil$

Fast algebraic attacks

- An extension of the conventional algebraic attacks
- Maximum AI does not imply resistance to fast algebraic attacks

Definition

The fast algebraic immunity $\mathsf{FAI}_n(f)$ of $f \in \mathbb{B}_n$ is defined by

$$\mathsf{FAI}_n(f) = \min_{1 \leq \deg(g) \leq \mathsf{AI}_n(f)} \{2\,\mathsf{AI}_n(f), \deg(g) + \deg(f * g)\}$$

- Upper bound: $\mathsf{FAI}_n(f) \leq n$
- If $\mathsf{FAI}_n(f) = n$, then f is a perfect algebraic immune function

The Carlet-Feng construction

- Carlet-Feng, 2008: $\mathrm{supp}(f) = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{2^{n-1}-2}\}$, where α a primitive element of the finite field $\mathbb{F}_{2^n}$.
  - Degree $n-1$ (i.e. the maximum possible)
  - High (first-order) nonlinearity is ensured
    - Lower bound (Tang et. al., 2013:)

$$\mathsf{nl}(f) \geq 2^{n-1} - \left(\frac{n\ln(2)}{\pi} + 0.74\right)2^{n/2} - 1$$

    - Experiments show that the actual values of nonlinearities may be higher enough
  - Optimal against fast algebraic attacks, as subsequently shown (Liu-Zhang-Lin, 2012)
- Several generalizations of the Carlet-Feng construction
  - The most recent is based on exploiting properties of punctured Reed-Muller codes (Limniotis-Kolokotronis, 2018)

Predictability of sequences: Linear complexity

Several criteria to measure pseudorandomness of a sequences s

- Widely studied:
  - Linear complexity $\mathrm{c}(s)$ of a sequence s (the length of the shortest Linear Feedback Shift Register that generates s)
    - Berlekamp-Massey algorithm
    - Games-Chan algorithm (for 2^n-periodic binary sequences)
  - Linear complexity profile (how linear complexity increases as the sequence length grows)
- Generalized complexity measures:
  - k-error linear complexity $\mathrm{c}_k(s)$: $\min_{\mathrm{wt}(e)\leq k} \mathrm{c}(s+e)$ (how the linear complexity can be reduced if at most k errors are introduced)
  - k-error linear complexity spectrum (how linear complexity decreases as the error weight k increases)

The Games-Chan algorithm

A recursive algorithm

- $s = [L\ \ R]$
- $B(s) = L \oplus\ R$ (of period 2^{n-1})
- Is $B(s)$ different from the all-zeroes sequence?
    - If yes, then $\mathrm{c}(s) = 2^{n-1} + \mathrm{c}(B(s))$;
    - otherwise, $\mathrm{c}(s) = \mathrm{c}(L)$

Example

- $s = 01000111$
- $B(s) = 0011,\ \mathrm{c}(s) = 4 + \mathrm{c}(B(s))$
- $B(B(s)) = 11,\ \mathrm{c}(B(s)) = 2 + \mathrm{c}(B(B(s))) = 2 + 1 = 3$
- $\mathrm{c}(s) = 4 + 3 = 7$

Critical Error Linear Complexity Spectrum

2^n-periodic binary sequences attracted great attention, due to special properties implied by the Games-Chan algorihm

- Critical Error Linear Complexity Spectrum (CELCS): the ordered set of points $(k, \mathrm{c}_k(s))$ satisfying $\mathrm{c}_k(s) > \mathrm{c}_{k'}(s)$, for $k' > k$.
- Each point in CELCS is called critical point (CP)

Milestones

- Stamp-Martin, 1993: an algorithm for computing $c_k(s)$,
- Kurosawa et. al., 2000: the minimum number of bits that should be altered in order to reduce the complexity: $2^{\mathrm{wt}(2^n - \mathrm{c}(s))}$,
- Lauder-Paterson, 2003: generalization of the Stamp-Martin algorithm, to compute the entire CELCS
- Etzion-Kalouptsidis-Kolokokotronis-Limniotis-Paterson, 2009: Detailed study on the properties of the CELCS

The Lauder-Paterson algorithm

Example (*Cont.*)

- The sequence $s = 01000111$ has 3 CPs
  - $(0, 7)$
  - $(2, 2)$
    - s' =01010101
    - The sequence $e = 00010010$ such that $\mathrm{c}(s \oplus e) = \mathrm{c}_2(s)$ is a critical error sequence
  - $(4, 0)$
- For length $N = 2^n$, $\mathcal{O}(N \log(N)^2)$ bit operations
- The Lauder-Paterson algorithm computes all the CPs, but appropriately modified can also compute the critical error sequences
- For any 2^n-periodic binary sequence s, the minimum possible number of CPs is two:
  - $(0, \mathrm{c}(s)), (\mathrm{wt}(s), 0)$ (the *trivial* CPs)
- Etzion et. al., 2009: Full characterization of sequences with 2 CPs

A bijection between sequences and functions

Definition

If $\boldsymbol{s} = (s_0, s_1, \ldots, s_{2^n-1})$ is the vector corresponding to a periodic binary sequence s with period 2^n, then we define the corresponding n-variable Boolean function f, denoted by f_s, to be the function whose truth table equals $\boldsymbol{f_s} = (s_0, s_1, \ldots, s_{2^n-1})$

- We write $s \leftrightarrow f_s$.
- Conversely, for any function $f' \in \mathbb{B}_n$, there is a unique 2^n-periodic binary sequence s' such that $s' \leftrightarrow f'$.

Proposition

Let s be a 2^n-periodic binary sequence, with linear complexity $c(s)$. It holds $2^{n-\ell-1} \leq c(s) < 2^{n-\ell}$ for some $1 \leq \ell < n-1$ if and only if the ANF of $f_s(x_1, \ldots, x_n)$ depends only on $x_1, \ldots, x_{n-\ell}$.

"Linear complexity" of Boolean functions

- Due to the aforementioned bijection, the linear complexity of a sequence s reflects the number of variables that appear in the ANF of the corresponding Boolean function f_s
- Similarly, we may proceed with the CELCS of f_s

Theorem

- Let $(k, c_k(s))$ be a CP of s satisfying $2^{n-\ell-1} \leq c_k(s) < 2^{n-\ell}$ for some integer $\ell \geq 1$
  - Let k be the least integer with this property
- $f_s \leftrightarrow s$.
- Let e be a critical error sequence of s such that $\mathrm{wt}(e) = k$
- $\Rightarrow$ The function $h = f_s + f_e$ depends on the first $n - \ell$ variables and, moreover, there is no function $g \in \mathbb{B}_n$ with $\mathrm{wt}(g) < k$ such that $f_s + g$ depends on at most the first $n - \ell - 1$ variables.

The CELCS of a Boolean function

- The CELCS provides info on how well a function can be approximated by another function with fewer number of variables
- $\Rightarrow$ Use of the Lauder-Paterson algorithm for efficient computation

Example - The function f of the first version of the Achterbahn cipher

Use of the Lauder-Paterson algorithm for finding approximations of f depending on $k < 8$ variables

k	$distance$	Bias
7	32	0.375
6	64	0.25
5	96	0.125

- There exist functions depending on 7 and 6 variables that approximate $f \in \mathbb{B}_8$ with bias 0.375 (equal to the bias of the best 2nd-order approximation of f) and 0.25 (equal to the bias of the best affine approximation of f) respectively.

Other examples

- The Lauder-Paterson also provides useful results for the 2nd version of the Achterbahn, having a function with 13 variables
- For the 3rd-order resilient function $f \in \mathbb{B}_{10}$ of the LILI-128 cipher, we found out function depending on 4 variables, whose distance from f is very close to the relative lower bound proved in (Canteaut et. al., 2002)
  - The Carlet-Feng function $f_{CF} \in \mathbb{B}_9$ (perfect algrebraic immune)

k	distance	CP	Bias
8	$\mathbf{130}$	$(130, 97)$	0.2461
7	$\mathbf{162}$	$(162, 99)$	0.1836
6	$\mathbf{192}$	$(192, 57)$	0.1250
5	$\mathbf{220}$	$(220, 26)$	0.0703
4	$\mathbf{232}$	$(232, 9)$	0.0469
3	246	$(246, 5)$	0.0195

What if the number of CPs is only two?

- If s has two CPs, then it seems that the Lauder-Paterson algorithm does non provide useful information - in terms of the previous analysis - on the Boolean function f_s
- However, in such a case, f_s is not of cryptographic strength

Lemma

- If s has two CPs, it is “highly probable” that the linear kernel of f_s has dimension at least 1
- Conversely, if $f_s(x_1, \ldots, x_n) = g(x_1, \ldots, x_{n-1}) \oplus \epsilon x_n$, $\epsilon \in \{0, 1\}$ then its linear kernel has dimension at least 1 and s has exactly two CPs.

An interesting observation

- Permuting the variables of f_s result in a linearly-equivalent function $f_{s'}$
  - Actualy, $f_{s'}$ is the same with f_s, having changed the names of the variables
- The CELCS of s' is generally different from the CELCS of s

Definition

Let $f \in \mathbb{B}_n$. Then, for any $0 \leq k \leq n$, the k-error linear complexity of f, denoted as $c_k(f)$ is defined as

$$c_k(f) = \min_{A \in P_n} \left\{ c_k(s) : s \leftrightarrow f(A\boldsymbol{x}) \right\}$$

where P_n is the set of all permutation matrices over $\mathbb{F}_2$ of order n.

The CELCS of f is similarly defined

The Lauder-Paterson algorithm for computing low-order approximations

- The Lauder-Paterson algorithm finds out critical error vectors
- If e is a critical error sequence of s, when it holds $\deg(f_{s\oplus e}) < \deg(f_s)$?

Proposition

Let $s = [L_1\ R_1]$, $s' = [L_2\ R_2]$ be two binary sequences of length 2^n. If $R_1 = R_2$ and $\deg(f_{B(s)}) < \deg(f_{B(s')})$, then it holds $\deg(f_s) \leq \deg(f_{s'})$.

- The proof of this Proposition illustrates that $\deg(f_s) < \deg(f_{s'})$ with high probability (i.e. equality is not expected to be common)

The Lauder-Paterson algorithm for computing low-order approximations *(Cont.)*

Proposition

Let s be a binary sequence with period 2^n such that

$$2^{n-2} < \mathrm{wt}(B(s)) < 2^{n-1}$$

. Then, there exists a non-trivial critical error sequence e of s such that $\deg(f_{s\oplus e}) \leq \deg(f_s)$.

- Hence, the Lauder-Paterson algorithm also finds out low-order approximations
- Experiments illustrate that, in some cases, best low-order approximations are obtained

Conclusions - Open problems

- Via defining a bijection between 2^n-periodic binary sequences and Boolean functions on n variables, information on pseudorandomness properties of sequences also reflect cryptographic properties of functions
- Known algorithms on sequences may be used for efficient computation of cryptographic properties of functions (known to be hard to be computed otherwise)
- The Lauder-Paterson algorithm for determining approximations:
  - depending on fewer number of variables
  - of lower degree

Open problems (not an exhaustive list...)

- When are these approximations the best?
- How to use these results for constructing cryptographically strong functions?

- T. Etzion, N. Kalouptsidis, N. Kolokotronis, K. Limniotis and K. G. Paterson, Properties of the error linear complexity spectrum, *IEEE Trans. Inform. Theory, vol. 55, pp. 4681-4686, Oct. 2009*.
- N. Kolokotronis, K. Limniotis and N. Kalouptsidis, Best affine and quadratic approximations of particular classes of Boolean functions', *IEEE Trans. Inform. Theory, vol. 55, pp. 5211-5222, Nov. 2009*.
- N. Kolokotronis and K. Limniotis, On the second-order nonlinearity of cubic Maiorana-McFarland Boolean functions, *Int. Symp. on Inform. Theory and its Applications (ISITA), pp. 596-600, 2012*.
- K. Limniotis and N. Kolokotronis, Boolean functions with maximum algebraic immunity: further extensions of the Carlet–Feng construction, *Designs, Codes and Cryptography, vol. 86, pp. 1685–1706, Springer, 2018*.
- K. Limniotis and N. Kolokotronis, The error linear complexity spectrum as a cryptographic criterion of Boolean Functions. *Submitted to IEEE Trans. Inform. Theory (under review)*.

Questions & Answers

Thank you for your attention!