De Cifris incontra Torino

Politecnico di Torino – 14 Ottobre 2019

Boolean cryptographic functions and related combinatorial constructions

Alberto Leporati

Università degli Studi di Milano – Bicocca

Dip. di Informatica, Sistemistica e Comunicazione (DISCo)

Viale Sarca 336/14 – Milano - Italy

About me

- Associate Professor at the Department of Informatics, Systems and Communication (DISCo) of the University of Milan – Bicocca
- Founder and current director of Bicocca Security Lab
  - interests also in Cybersecurity
  - inside the lab, Luca Mariot and me have competencies on Cryptography
- Teacher of a course on Information Theory and Cryptography for the Master Degree on Computer Science, since 2008
- Supervisor of many bachelor (100+) and master (30+) theses
- Supervisor of two Ph.D. theses on Cryptography
- Supervisor of a post-doc research project on Cryptography
- Member of CINI Cybersecurity Lab (Milan – Bicocca node)

Research in Cryptography

- Theoretical foundations of cryptographic primitives
- Search for Boolean functions with good cryptographic properties: *k*-resiliency, nonlinearity, balancedness
- Relations with Secret Sharing Schemes, Orthogonal Arrays, combinatorial designs, linear codes
- Relations with parallel models of computation, mainly Boolean circuits and Cellular Automata

Boolean functions

We search for Boolean functions:

$$f: \{0,1\}^n \to \{0,1\}$$

(and we extend them to multi-output Boolean functions: $F: \{0,1\}^n \to \{0,1\}^m$)

with good cryptographic properties:

- balancedness: $f(x) = 0$ for half of the inputs $x \in \{0,1\}^n$
- nonlinearity: high distance from affine functions
- correlation immunity of order *k*: every subset of at most *k* variables is statistically independent of the value of $f(x)$
- *k*-resiliency: both balanced and correlation immune of order *k*

Boolean functions

- Functions that do not have these properties lead to known attacks
- All these properties can be expressed in terms of the Walsh transform

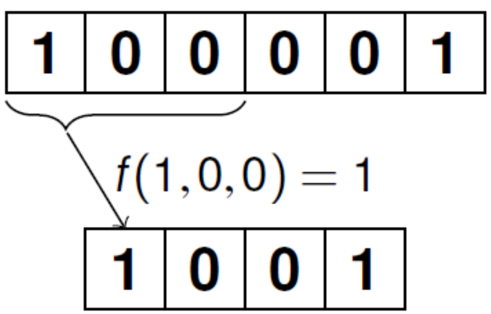
$$W_f(\omega) = \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus \omega \cdot x} \text{ , where } \omega \cdot x = \bigoplus_{i=1}^{n} \omega_i \cdot x_i$$

(in practice, it computes the projection of the truth table vector of *f* wrt the basis composed of the XORs of all possible subsets of the input variables)

- Some upper bounds for the properties are known (for example, on nonlinearity), and also relationships/constraints between them

Cellular Automata (CA)

- One-dimensional Cellular Automaton (CA): a discrete parallel computation model composed of a finite array of n cells
- Each cell updates its state $s \in \{0,1\}$ by applying a local rule $f: \{0,1\}^d \to \{0,1\}$ of diameter d to itself and to the $d-1$ neighboring cells to its right
- Example: $n = 6$, $d = 3$, $f(s_i, s_{i+1}, s_{i+2}) = s_i \oplus s_{i+1} \oplus s_{i+2}$

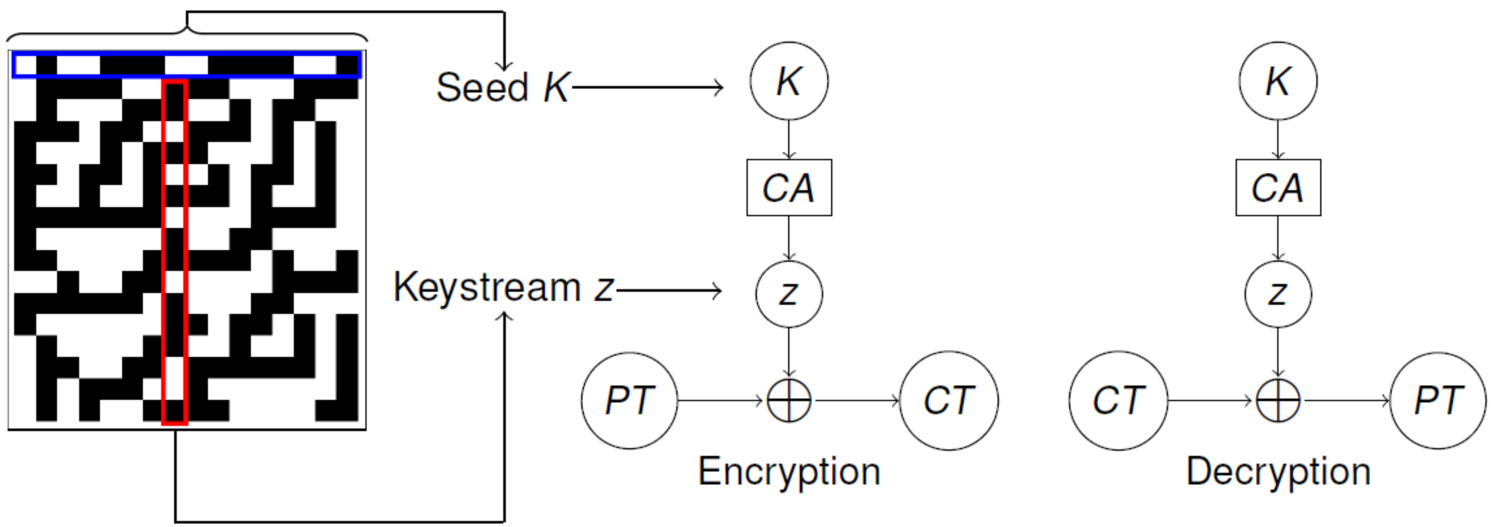


Note: No-boundary CA

It shrinks in size at each step

Use of CA in Cryptography

- Example: as a PRNG in stream ciphers



- Sometimes CA are used but not explicitly mentioned: see, for example, Keccak

Use of CA in Cryptography

- CA-based block cipher design:

$$\cdots \boxed{0}\underbrace{\boxed{1}\boxed{1}\boxed{0}}\boxed{0} \cdots$$

$$\downarrow f : \{0,1\}^d \to \{0,1\}$$

$$\boxed{0}$$

From local rules (Boolean functions)

to

global rules (multi-output Boolean functions)

  - global rules can be seen as S-boxes
  - goal: find S-boxes with high nonlinearity and with low differential uniformity

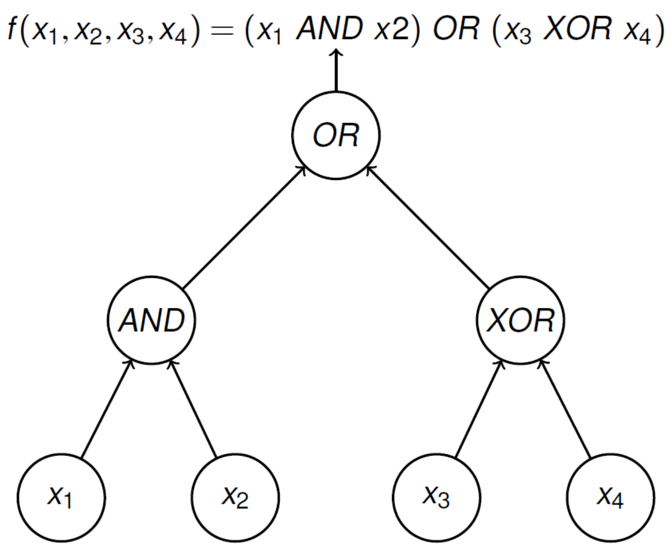
$$\boxed{1}\boxed{0}\boxed{0}\boxed{0}\boxed{0}\boxed{1}\boxed{0}\boxed{1}$$

$$\Downarrow F : \{0,1\}^n \to \{0,1\}^m$$

$$\boxed{1}\boxed{0}\boxed{0}\boxed{1}\boxed{1}\boxed{0}$$

In search of Boolean functions

- The number of Boolean functions grows in a double exponential way wrt to the number n of inputs: 2^{2^n}. Exhaustive search becomes impossible
- Evolutionary techniques: Genetic Algorithms, and Genetic Programming
- Search spaces:
  - truth tables of Boolean functions
  - Walsh spectra of pseudo-Boolean real functions
  - trees of Boolean operators
- Example of encoding in GP:



In search of Boolean functions

- Results obtained:
  - for $n = 4$ and $n = 5$, we obtained CA rules inducing S-boxes with optimal crypto properties, and with implementation cost similar to or slightly better than the state of the art in the literature
  - for $n > 5$, GP finds S-boxes with optimal cryptographic properties up to $n = 7$, but with too high implementation costs
- In general, Genetic Programming seems to work better than Genetic Algorithms, both with truth table and Walsh spectrum representations (Why???)

Latin squares and SSS

- A Latin square (LS) is a $N \times N$ matrix where each row and each column permutes $[N] = \{1, \dots, N\}$
- Example, with $N = 4$:

1	4	2	3
3	2	4	1
4	1	3	2
2	3	4	1

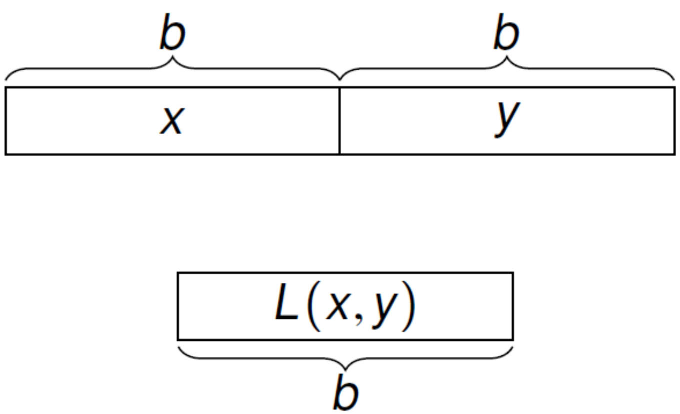
- Latin squares are examples of combinatorial designs
- They can be used as perfect (2,2)-threshold Secret Sharing Schemes (SSS)
  - secret: $s \in [N]$
  - shares: number of row, number of column

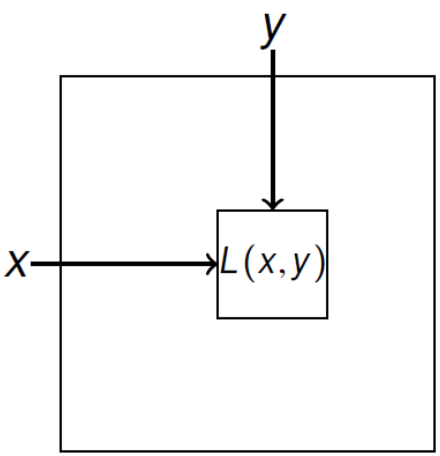
Latin squares and CA

- Any CA with a bipermutive rule:

$$f(x_1, \dots, x_d) = x_1 \oplus \underbrace{\varphi(x_2, \dots, x_{d-1})}_{\text{(generating function)}} \oplus x_d$$

of diameter $d = b + 1$ can be used to generate a LS of order $N = 2^b$





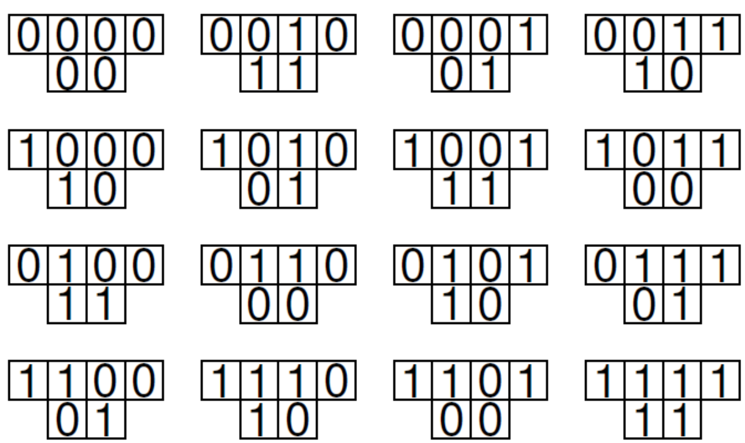
Latin squares and CA

- Example, with the so-called rule 150:

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

with the encoding:

$00 \to 1 \qquad 10 \to 2 \qquad 01 \to 3 \qquad 11 \to 4$



(a) Rule 150 on 4 bits

1	4	3	2
2	3	4	1
4	1	2	3
3	2	1	4

(b) Latin square L_{150}

Orthogonal Latin squares and SSS

- The LS $L_1, \ldots, L_n$ are mutually orthogonal (n-MOLS) if their pairwise superposition yields all the pairs $(x, y) \in [N] \times [N]$
- Example, with $n = 2$ and $N = 4$:

1	3	4	2
4	2	1	3
2	4	3	1
3	1	2	4

(a) L_1

1	4	2	3
3	2	4	1
4	1	3	2
2	3	4	1

(b) L_2

1,1	3,4	4,2	2,3
4,3	2,2	1,4	3,1
2,4	4,1	3,3	1,2
3,2	1,3	2,1	4,4

(c) (L_1, L_2)

- (n-MOLS) can be used as perfect $(2, n)$-threshold Secret Sharing Schemes
  - secret: s is the pair (row number, column number)
  - shares: the entries at that (row, column), for each LS

Orthogonal Latin squares and SSS

- Bipermutive linear rule: $f(x) = x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_{n-1} x_{n-1} \oplus x_n$
- Associated polynomial: $P_f(X) = a_1 + a_2 X + \cdots + a_n X^{n-1}$

> Theorem: Bipermutive linear rules $f, g: \{0,1\}^n \to \{0,1\}$ generate orthogonal Latin squares if and only if $P_f(X)$ and $P_g(X)$ are coprime

- Enumeration of OLS in the linear case → enumeration of pairs of coprime polynomials (but that's another story...)
- ... What about the nonlinear case?
- MOLS arising from nonlinear constructions have relevance in cheater-immune Secret Sharing Schemes
- **Goal:** Exhaustive enumeration of pairs of bipermutive rules of size *n* generating orthogonal Latin squares, classified by nonlinearity

Orthogonal Latin squares and SSS

- Pairs of bipermutive rules of n variables:

n	3	4	5	6	7
$\mathcal{B}_n$	16	256	65536	4294967296	$\approx 1.84 \cdot 10^{19}$

- Exhaustive enumeration possible up to $n = 6$
- By considering some symmetries, we can divide the search space size by 8
- f, g are pairwise balanced (PWB) if each pair (0,0), (0,1), (1,0), (1,1) occurs 2^{n-2} times in the superposition of the two truth tables
- Property: if f, g are bipermutive and generate OLS, then they are PWB
  - sufficient but not necessary condition!
  - counterexamples already available for $n = 4$
- We thus consider balanced quaternary strings of length 2^{n-2} $(= BalG_n)$

Orthogonal Latin squares and SSS

- We have:

n	$\#\mathcal{B}_n$	$\#Bal\mathcal{G}_n$	$\#Bal\mathcal{B}_n$
3	16	0	8
4	256	24	96
5	65536	2520	17920
6	4294967296	63006300	843448320
7	$\approx 1.84 \cdot 10^{19}$	$\approx 9.96 \cdot 10^{15}$	$\approx 2.58 \cdot 10^{18}$

- Even by focusing on $BalB_n$, we cannot exhaustively search beyond $n = 6$
- We used a 40-core machine to span $BalB_n$, and it took 22 hours to complete

Orthogonal Latin squares and SSS

- Classification results:

n	LS_size	#total	#linear	#nonlinear	$(NI(f), NI(g), \#pairs)$
3	4×4	1	1	0	–
4	8×8	9	5	4	$(4,4,4)$
5	16×16	213	21	192	$(4,4,96)$, $(8,8,96)$
6	32×32	66685	85	66600	$(4,4,512)$, $(12,12,17992)$, $(8,8,4020)$, $(16,16,28388)$, $(20,20,14384)$, $(4,12,8)$, $(8,16,160)$, $(12,20,128)$, $(16,24,88)$

Orthogonal Latin squares and SSS

- Summing up:
  - we considered the problem of exhaustively enumerating pairs of bipermutive CA generating orthogonal Latin squares, and classify them wrt nonlinearity
  - we proved that pairwise balancedness is a necessary condition for two rules to generate OLS
  - we used this condition to enumerate pairs up to size $n = 6$
- Future directions:
  - find sufficient conditions for two rules to generate OLS
  - combinatorial encoding to evolve pairs of PWB bipermutive rules through Genetic Algorithms (work in progress)

Blockchains

- Design of blockchain-based applications
  - supply chain management
  - definition of utility (crypto) tokens backed by tangible assets
  - development of smart contracts with Ethereum (Solidity) and Hyperledger (Node.js)

- Analysis of Ethereum smart contracts, for security and correcteness properties

Blockchains

Two examples of use cases:

- Anti-counterfeiting of luxury clothes and accessories, using a blockchain + RFIDs
  - each cloth / accessory has a unique RFID
  - every production / assembly / transportation / sell operation is written on the blockchain
  - it becomes incredibly difficult to sell counterfeit items!
- Storage of sensor data from (non-autonomous) vehicles
  - hashes of contents of the car's black box are regularly saved on the blockchain
  - when needed, the driver can prove that his/her data have not been altered

Thanks for your attention!

Alberto Leporati

alberto.leporati@unimib.it