University of Milano-Bicocca
Department of Informatics, Systems and Communications

Cryptographic Criteria of Boolean Functions and S-Boxes

Luca Mariot

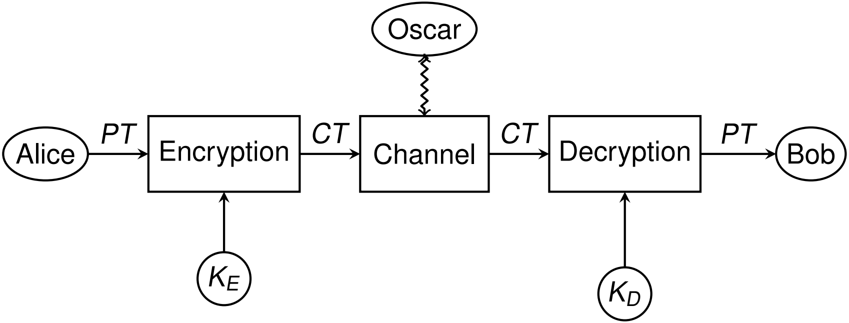
luca.mariot@unimib.it

Guest Lecture for Digital Communication

Durham – March 18, 2019

Cryptography

Basic Goal of Cryptography: Enable two parties (Alice and Bob, A and B) to securely communicate over an insecure channel, even in presence of an opponent (Oscar, O)



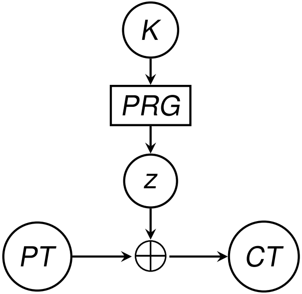
- PT: plaintext
- CT: ciphertext
- K_E: encryption key
- K_D: decryption key

Symmetric cryptosystems

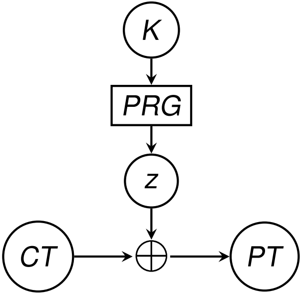
Symmetric cryptosystems ($K_E = K_D = K$) can be classified as:

- *Stream ciphers*: each symbol of *PT* is combined with a symbol of a *keystream*, computed from *K*
  - Grain
  - Trivium
  - ...
- *Block ciphers*: *PT* is divided in *blocks* combined with *round keys* derived from *K* through a *round function*
  - DES
  - Rijndael (AES)
  - ...

Vernam Stream Cipher



(a) Encryption

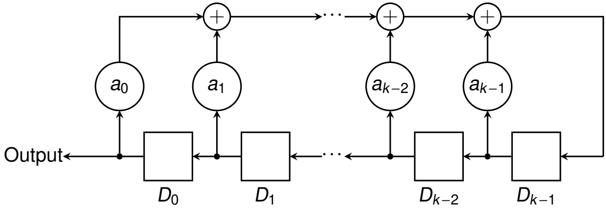


(b) Decryption

- *K*: secret key
- *PRG*: Pseudorandom Generator
- *z*: keystream
- $\bigoplus$: bitwise XOR
- *PT*: Plaintext
- *CT*: Ciphertext

- Device computing the binary linear recurring sequence

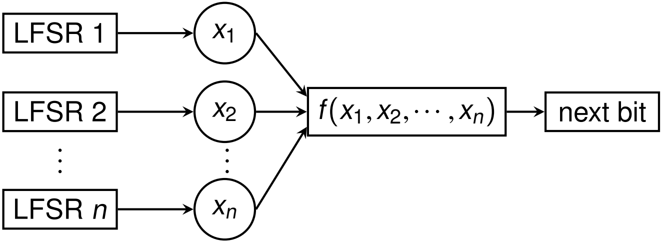
$$s_{n+k} = a + a_0 s_n + a_1 s_{n+1} + \cdots + a_{k-1} s_{n+k-1}$$



- Too weak as a PRG: $2k$ consecutive bits of keystream are enough to recover the LFSR initialization via the Berlekamp-Massey algorithm

An Example of PRG: The Combiner Model

- a Boolean function $f : \{0,1\}^n \to \{0,1\}$ combines the outputs of n LFSR [2]



- Security of the combiner ⇔ cryptographic properties of f

Boolean Functions - Basic Definitions

Boolean function: a mapping $f : \mathbb{F}_2^n \to \mathbb{F}_2$, where $\mathbb{F}_2 = \{0, 1\}$

- Truth table: vector Ω_f specifying $f(x)$ for all $x \in \mathbb{F}_2$

(x_1, x_2, x_3)	000	100	010	110	001	101	011	111
Ω_f	0	1	1	1	1	0	0	0

- Algebraic Normal Form (ANF): Sum (XOR) of products (AND) over the finite field $\mathbb{F}_2$

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

- Walsh Transform: correlation with the *linear* functions defined as $\omega \cdot x = \omega_1 x_1 \oplus \cdots \oplus \omega_n x_n$

$$\hat{F}(\omega) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus \omega \cdot x}$$

Cryptographic Properties: Balancedness

- Hamming weight $w_H(f)$: number of 1s in Ω_f
- A function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is balanced if $w_H(f) = 2^{n-1}$
- Walsh characterization: f balanced $\Leftrightarrow \hat{F}(0) = 0$

(x_1, x_2, x_3)	000	100	010	110	001	101	011	111
Ω_f	0	1	1	1	1	0	0	0

$$\Downarrow$$

f is balanced

- Unbalanced functions present a statistical bias that can be exploited in attacks

Cryptographic Properties: Algebraic Degree

- Algebraic degree d: the degree of the multivariate polynomial representing the ANF of f

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

$$\Downarrow$$

$$f \text{ has degree } d = 2$$

- *Linear* functions $\omega \cdot x = \omega_1 x_1 \oplus \cdots \oplus \omega_n x_n$ have degree $d = 1$
- Boolean functions of high degree make the attack based on Berlekamp-Massey algorithm less effective

Cryptographic Properties: Nonlinearity

- Nonlinearity $nl(f)$: Hamming distance of f from linear functions
- Walsh characterization:

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{\omega \in \mathbb{F}_2^n} \left\{ \left| \hat{F}(\omega) \right| \right\}$$

(x_1, x_2, x_3)	000	100	010	110	001	101	011	111
Ω_f	0	1	1	1	1	0	0	0
$\hat{F}(\omega)$	0	0	0	0	-4	4	4	4

$$\Downarrow$$

$$nl(f) = 2^{3-1} - \frac{1}{2} \cdot 4 = 2$$

- Functions with high nonlinearity resist fast-correlation attacks

Cryptographic Properties: Resiliency

- t-Resiliency: when fixing any t variables, the restriction of f stays balanced
- Walsh characterization:

$$\hat{F}(\omega) = 0 \ \forall \omega : w_H(\omega) \leq t$$

(x_1, x_2, x_3)	000	100	010	110	001	101	011	111
Ω_f	0	1	1	1	1	0	0	0
$\hat{F}(\omega)$	0	0	0	0	−4	4	4	4

$$\Downarrow$$

$$F(001) = -4 \Rightarrow f \text{ is NOT 1-resilient}$$

- Resilient functions of high order t resist to correlation attacks

Bounds and Trade-offs

In summary, $f : \mathbb{F}_2^n \to \mathbb{F}_2$ should:

- be balanced
- be resilient of high order m
- have high algebraic degree d
- have high nonlinearity nl

But most of these properties cannot be satisfied simultaneously!

- *Covering Radius bound*: $nl \leq 2^{n-1} - 2^{\frac{n}{2}-1}$
- *Siegenthaler's bound*: $d \leq n - t - 1$
- *Tarannikov's bound*: $nl \leq 2^{n-1} - 2^{t+1}$

Constructions of good Boolean Functions

- Number of Boolean functions of n variables: 2^{2^n}
- $\Rightarrow$ too huge for exhaustive search when $n > 5$!
- Functions used in the combiner model have $n \geq 13$ variables

In practice, one usually resorts to:

- Algebraic constructions [2]
  - *Maiorana-McFarland construction*
  - *Rothaus' construction*
  - ...
- Heuristic techniques
  - *Simulated Annealing* [3]
  - *Evolutionary Algorithms* [6]
  - ...

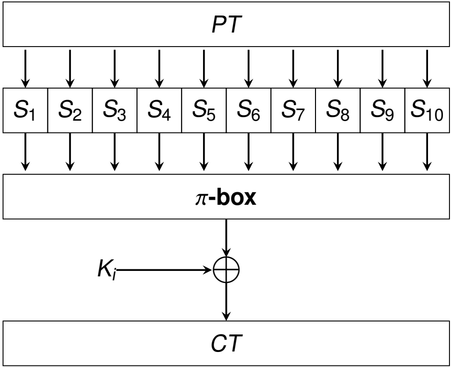
Special classes of functions

Special classes of functions:

- Bent functions: $\hat{F}(\omega) = \pm 2^{\frac{n}{2}}$ for all ω
  - Reach covering radius bound for n even (maximum nonlinearity)
  - Unfortunately, they are unbalanced: $\hat{F}(0) = \pm 2^{\frac{n}{2}}$
- Plateaued functions: $\hat{F}(\omega) \in \{-2^{\lambda}, 0, 2^{\lambda}\}$ for all ω
  - Can be balanced
  - Reach both Siegenthaler's and Tarannikov's bounds

Block Ciphers: Substitution-Permutation Network

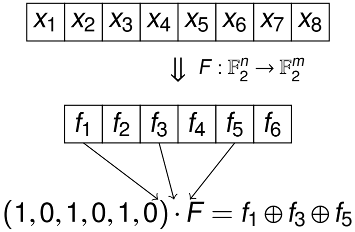
Round function of a SPN cipher:



- $S_i : \mathbb{F}_2^n \to \mathbb{F}_2^n$ are S-boxes providing confusion [8]
- Security of confusion layer $\Leftrightarrow$ cryptographic properties of S_i

S-Boxes: General definitions

- A Substitution Box (S-box) is a mapping $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ defined by m coordinate functions $f_i : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$
- The component functions $v \cdot F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ for $v \in \mathbb{F}_2^m$ of F are the linear combinations of the f_i



- In SPN ciphers, one uses S-boxes with $m = n$

Balancedness and Algebraic Degree

Balancedness:

- $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ balanced if $|F^{-1}(y)| = 2^{n-m}$ for all $y \in \mathbb{F}_2^m$
- F is balanced $\Leftrightarrow$ all its component functions $v \cdot F$ are balanced
- Balanced functions with $m = n$ are bijective S-boxes

Algebraic degree:

- Degree of the ANF of F over $\mathbb{F}_2^m$
- Equal to the maximum degree of all coordinate functions
- S-boxes of high degree thwart higher-order differential attacks

Nonlinearity

- Walsh transform for component $v \cdot F$:

$$\hat{F}(v,\omega) = \sum_{x \in \mathbb{F}_2^n} (-1)^{v \cdot F(x) \oplus \omega \cdot x}$$

- Nonlinearity for component $v \cdot F$:

$$nl(v \cdot F) = 2^{n-1} - \frac{1}{2} \max_{\omega \in \mathbb{F}_2^n} \left\{ \left| \hat{F}(v,\omega) \right| \right\}$$

- The nonlinearity of a S-box F is defined as the minimum nonlinearity among all its component functions
- S-boxes with high nonlinearity allow to resist to linear cryptanalysis attacks

Differential Uniformity

- delta difference table of F wrt a, b:

$$D_F(a,b) = \left\{ x \in \mathbb{F}_2^n : F(x) \oplus F(x \oplus a) = b \right\}.$$

- Given $\delta_F(a,b) = |D_F(a,b)|$, the differential uniformity of F is:

$$\delta_F = \max_{\substack{a \in \{0,1\}^{n*} \\ b \in \{0,1\}^m}} \delta_F(a,b).$$

- S-boxes with low differential uniformity are able to resist differential cryptanalysis attacks

Bounds and Special Classes

For nonlinearity:

- *Covering Radius* Bound ($m < n$): $nl(F) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$
  - Bent functions reach this bound (n even)
- *Sidelnikov-Chabaud-Vaudenay* Bound ($m = n$): $nl(F) \leq 2^{n-1} - 2^{\frac{n-1}{2}}$
  - Almost Bent functions (AB) reach this bound (n odd)

Bounds for differential uniformity:

- For $m < n$: $\delta_F \geq 2^{n-m}$
  - Bent functions reach this bound (n even)
- For $m = n$: $\delta_F \geq 2$
  - Almost Perfect Nonlinear functions (APN) reach this bound (AB $\Rightarrow$ APN)
  - Exist for even and odd n

Rijndael S-box [5]

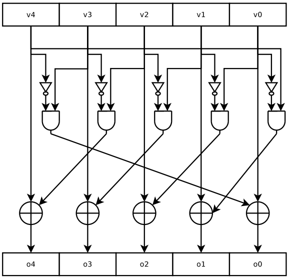
- Size 8×8 (works on bytes)
- Composition of an affine transformation and a nonlinear transformation
- Nonlinear transformation: Inversion in $\mathbb{F}_{2^8}$

$$F(x) = \begin{cases} x^{-1} & \text{, if } x \neq 0 \\ 0 & \text{, if } x = 0 \end{cases}$$

- Nonlinearity: 112, Differential uniformity: 4

Keccak χ S-box

- Cellular Automaton invertible for every odd size n [4]
- : Local rule: $\chi(x_i, x_{i+1}, x_{i+2}) = x_i \oplus (1 \oplus (x_{i+1} \cdot x_{i+2}))$



- Used as a 5×5 S-box in the Keccak specification of SHA-3 standard [1]
- Nonlinearity: 32, Differential uniformity: 8
- Other CA S-boxes with optimal properties found in [7]

Conclusions

- Boolean functions and S-boxes play a fundamental role in the design of symmetric ciphers
- The design of Boolean functions and S-boxes with good properties is a hard optimization problem
- Several other topics not covered here (see [2]:
  - Affine equivalence relation
  - Other properties (algebraic immunity, ...)
  - Relationship with error-correcting codes (Reed-Muller codes)

References I

G. Bertoni, J. Daemen, M. Peeters, and G. V. Assche.
Keccak.
In *Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings*, pages 313–314, 2013.

C. Carlet.
Boolean Functions for Cryptography and Error Correcting Codes.
In Y. Crama and P. L. Hammer, editors, *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, pages 257–397. Cambridge University Press, 2010.

J. A. Clark, J. L. Jacob, S. Stepney, S. Maitra, and W. Millan.
Evolving boolean functions satisfying multiple criteria.
In *Progress in Cryptology - INDOCRYPT 2002, Third International Conference on Cryptology in India, Hyderabad, India, December 16-18, 2002*, pages 246–259, 2002.

J. Daemen, R. Govaerts, and J. Vandewalle.
An efficient nonlinear shift-invariant transformation.
In *Proceedings of the 15th Symposium on Information Theory in the Benelux, B. Macq, Ed., Werkgemeenschap voor Informatie-en Communicatietheorie*, pages 108–115, 1994.

References II

J. Daemen and V. Rijmen.
The Design of Rijndael: AES - The Advanced Encryption Standard.
Information Security and Cryptography. Springer, 2002.

L. Mariot and A. Leporati.
A genetic algorithm for evolving plateaued cryptographic boolean functions.
In *Theory and Practice of Natural Computing - Fourth International Conference, TPNC 2015, Mieres, Spain, December 15-16, 2015. Proceedings*, pages 33–45, 2015.

L. Mariot, S. Picek, A. Leporati, and D. Jakobovic.
Cellular automata based s-boxes.
Cryptography and Communications, 11(1):41–62, 2019.

C. E. Shannon.
Communication theory of secrecy systems.
Bell system technical journal, 28(4):656–715, 1949.