Introduction to Symmetric Cryptology

Eric Filiol

ESIEA - Laval
Laboratoire de cryptologie et de virologie opérationnelles
$(C+V)^O$
filiol@esiea.fr

2013 - 2014

Agenda

Agenda

Introduction

Also denoted *secret key cryptology*.

- Encryption $C = F(K, M)$
- Decryption $M = F(K, C)$

Key point

The secret key K is known to each communication actor only and must remain secret.

- Mathematical formalization :
  - Information theory - Claude Elwood Shannon - 1948.
  - Three essential theorems.

Introduction : Shannon Theory

A sends a message M to B through a channel $\mathcal{C}$. Three main issues are then to be considered :

- Source characterization.
  - Entropy - Source coding - Shannon's first theorem.
- Information channel modelling.
  - Mutual information - Error-correcting codes - Shannon's second theorem.
- Perfect secrecy
  - Equivocation - Shannon's third theorem.

Introduction : Source Coding

Aim : to find the most economic way to write the source symbols/messages.

- Morse code (alphabet $= \{., -, *\}$). The less frequent the letter is, the longer is the associated source word is (E est encoded by $.*$)

Source Coding Theorem

For any information source X with entropy $H(X)$, we always can find a (source) code whose average length is arbitrary close to $H(X)$ (as an upper bound).

- $H(X)$ is the best lower bound for any possible possible source encoding.

Introduction : Noisy Channels

Aim : to recover from noise over the channel as efficiently as possible.

Noisy Channel Theorem

For any channel, we can always found a code family whose residual error probability (after decoding) tends towards to 0, provided that the transmission rate does not exceed the channel capacity.

- The theory of error-detecting and error-correction codes deals primarily with the study of finding good families of codes.
  - Repetition codes, normalized spelling code (alpha for A...).
  - Cyclic, linear, geometric codes ...

Introduction : Perfect Secrecy

Let M be a message (plaintext), K a (secret) key and Γ the ciphertext, all three considered as random variables.

- **Key Equivocation** :

$$H(K|\Gamma) = H(K,\Gamma) - H(\Gamma)$$

It is the entropy (uncertainty) about the key once the ciphertext has been captured.

- **Plaintext Equivocation** :

$$H(M|\Gamma) = H(M,\Gamma) - H(\Gamma)$$

It is the entropy (uncertainty) about the plaintext once the ciphertext has been captured.

Perfect Secrecy (2)

Proposition

Key equivocation $\geq$ plaintext equivocation.

- This comes from the fact that once the ciphertext and the secret key are known, then the plaintext is totally and uniquely determined (unique deciphering condition) under the Kerckhoffs' laws assumptions.

Perfect Secrecy

$$H(\mathcal{M}|\Gamma) = H(\mathcal{M})$$

- In other words, the mutual information of the channel is void (the channel is totally hermetic).

Perfect Secrecy (3)

Perfect Secrecy Theorem

There is perfect secrecy if and only if

$$\forall M \in \mathcal{M}, \forall \Gamma \in \mathcal{G} \quad P[M|\Gamma] = P[M]$$

or if and only if

$$\forall M \in \mathcal{M}, \forall \Gamma \in \mathcal{G} \quad P[\Gamma|M] = P[\Gamma]$$

- If there perfect secrecy, then there are as many keys as plaintexts/ciphertexts.

$$|\mathcal{K}| \geq |\mathcal{M}|$$

- So contraposed, whenever $|\mathcal{K}| < |\mathcal{M}|$, there is no perfect secrecy.

Agenda

Introduction

- Plaintext (resp. ciphertext) symbols are enciphered (resp. deciphered) on the fly.
- This enables a very high encryption speed.
- Very large theoretical corpus of knowledge that enables a high level of security proof.
- A (pseudo-) random sequence is bitwise combined (XOR) to the plaintext/ciphertext.
- This sequence must be reused !
- The sequence is truly random : *Vernam encryption* (1917).
- The sequence is pseudo-random (simulation of physical randomness) : deterministic algorithms.

Error-resilience of Stream Ciphers

Operational Constraint

The ciphertext MUST NOT be significantly larger than the plaintext.

- Cost issue (bandwidth)
- Security issue.
- If we cannot add redundancy, how to revover from noise during the transmission with respect to encryptd data ?
- We use synchronous stream ciphers.

Synchronous Stream Ciphers

We call *synchronous cryptosystems* any system which can be described by the equations

$$\Gamma_t = f(K, M_t, t)$$

- We xor a (pseudo-) random sequence to the plaintext/ciphertext.
  - Encryption : $\Gamma_t = M_t \oplus K_t$
  - One error occurs : $\Gamma'_t = \Gamma_t \oplus E_t$
  - Decipherment : $M'_t = M_t \oplus E_t$
- The single error on the cryptogram bit has an effect on the deciphered bit M'_t only.
  - No propagation of errors !

Vernam Stream Ciphers

Let $(M_t)_{t\geq 0}$, $(K_t)_{t\geq 0}$ and $(\Gamma_t)_{t\geq 0}$ the sequences of plaintext, of key and ciphertext respectively.

- **Encryption** : $\Gamma_t = M_t \oplus K_t$
- **Decipherment** : $M_t = \Gamma_t \oplus K_t$

Truly Random Sequence

The encrypting sequence $(K_t)_{t\geq 0}$ is a truly random sequence (serie of independent, identically distributed variables of parameter $\frac{1}{2}$. These variables are not produced by a deterministic process.

- The one-time pad provides perfect secrecy (be careful not to reuse the random sequence).

Vernam Stream Ciphers (2)

Vernam stream ciphers provide :

- Theoretically unconditional security.
- Highest possible encryption speed.
- Critical issue regarding the random sequence generation.
  - The randomness quality must be as good as possible. However there is no absolute definition of what randonmess is !
  - Production by sampling thermal/electronic resistor + statistical tests (Thalès module GDA-2)
  - Production by NVidia GPU in parallel at the hardware level (MassiveRand products).
  - Critical issues of storage (random sequences must be as long as the plaintext/ciphertext).
- There are critical issues regarding key management as well.

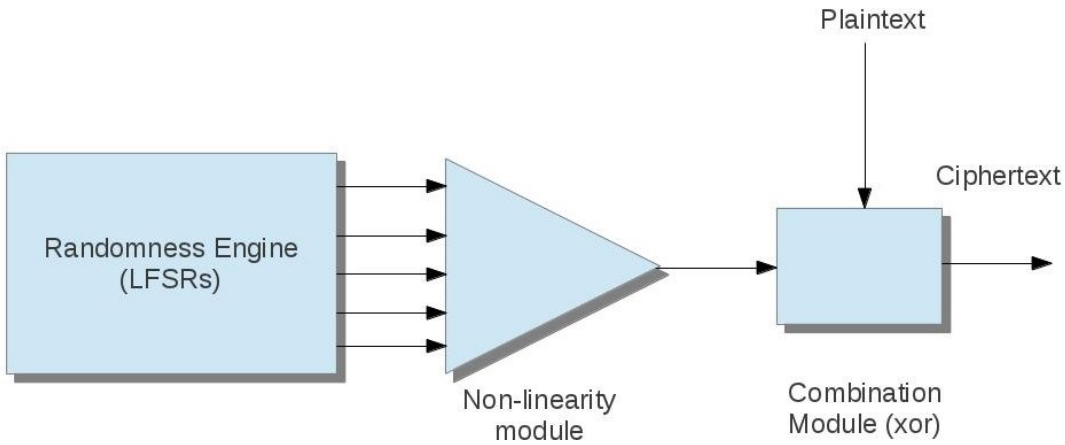
Pseudo-random Stream Ciphers

Pseudo-random Stream Ciphers

These systems are secure enough approximations of one-time pad system by means of deterministic (hence reproducible) automata.

- Use of a short K (a few tens of bits) from which the deterministic algorithm expands a pseudo-random sequence $E(K)$.
- We must use a message key to avoid parallel messages (reuse of the key).
- Made of *linear feedback shift registers* and Boolean functions.
- Pros : very high encryption speed, very cheap technology, large theoretical corpus enabling to achieve a high level of security.

Pseudo-random Stream Ciphers : General Structure



- The secret key initializes the register at time instant $t = 0$.
- There must be practically impossible to recover the secret key from the output sequence.
- There must be practically impossible to predict a part of the sequence from the knowledge of another part of the sequence.
- Most of those algorithms are generally proprietary (secret) algorithms.
  - Reconstruction techniques of those algorithms (Filiol, 2000).

Pseudo-random Stream Ciphers : Cryptanalysis

- All cryptanalysis techniques exploit the existence of correlation between inputs and outputs of the non-linearity module.
- Correlation attack : T. Siegenthaler (1985).
- Fast correlation attack : W. Meier - O. Staffelbach (1989).
- Various other variants and improvements.
- Algebraic attacks.
- Ciphertext-only attacks are interesting and make sense.
- The complexity of the attacks are still very high and they cannot be used in practice.

Pseudo-random Stream Ciphers : Examples

- Only a very few algorithms are public.
- A5/1 Algorithm (GSM) : 64-bit key - Broken by Shamir/Biryukov (2000).
- RC4 Algorithm (D. Rivest) (non officially published) : key size up to 2048 bits.
  - For export versions, key size is limited to 128 bits (Acrobat).
- E0 Algorithm (Bluetooth) : 132-bit key.
- Main manufacturers : Thalès, Sagem, Crypto-AG & Hagelin (CH), Racal (GB), Siemens (D),...
- Stork-Nessie Project (`http://www.stork.eu.org/index.html`).

Agenda

Introduction

The message M is split into n-bit blocks (DES 64 bits - AES 128 bits).

- Block B_i is encrypted as $C_i = f(K, B_i)$.
- The algorithm uses subkeys produced from the initial (base) key.
- To decipher, the same algorithm is used, with the subkeys in the reverse order.

Two main families

- Feistel schemes (1975) - Data Encryption Standard.
- Substitution/permutation networks - Advanced Encryption Standard.
- In both cases, a set of Boolean functions (or round) is iterated a number of times.

Block Cipher Security Principles

Shannon's rules.

Diffusion

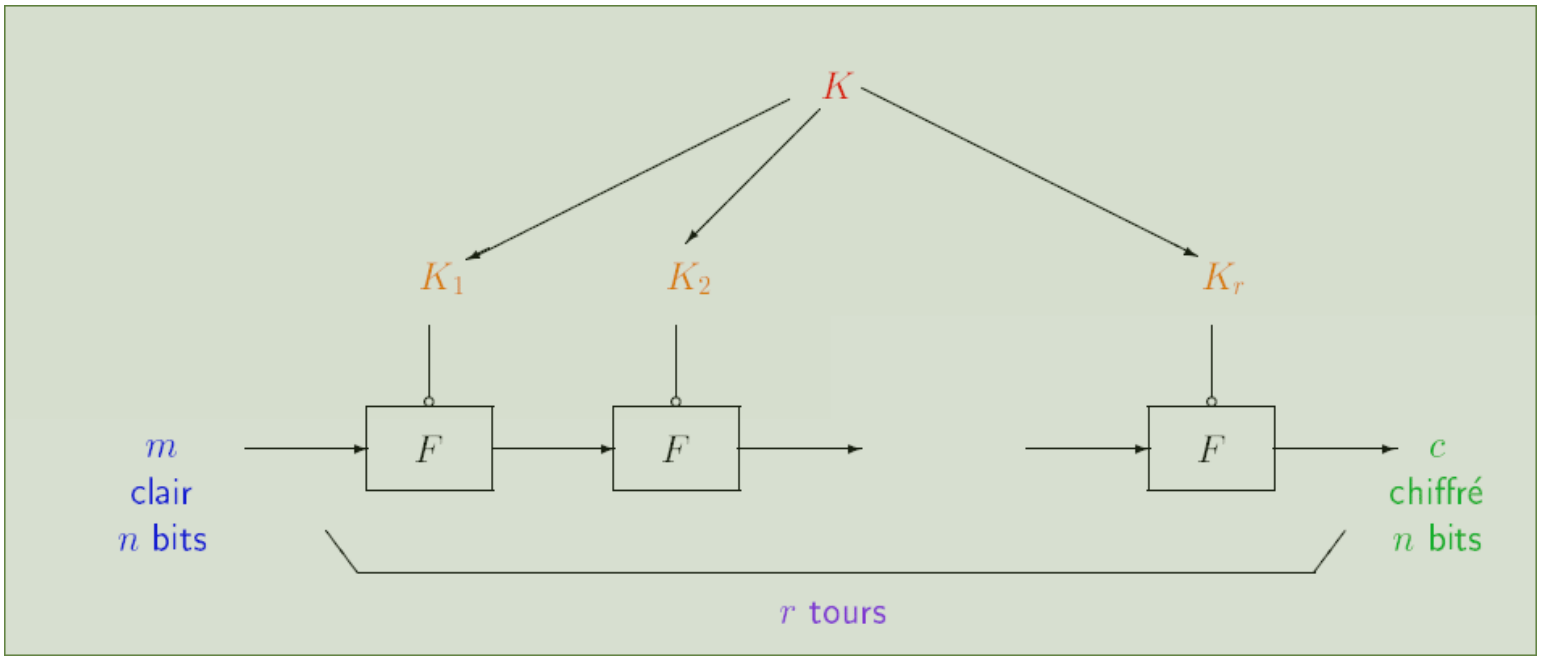
"In the method of diffusion the statistical structure of the plaintext which leads to its redundancy is dissipated into long range statistics."

Confusion

"The method of confusion is to make the relation between the simple statistics of the ciphertext and the simple description of the key a very complex and involved one."

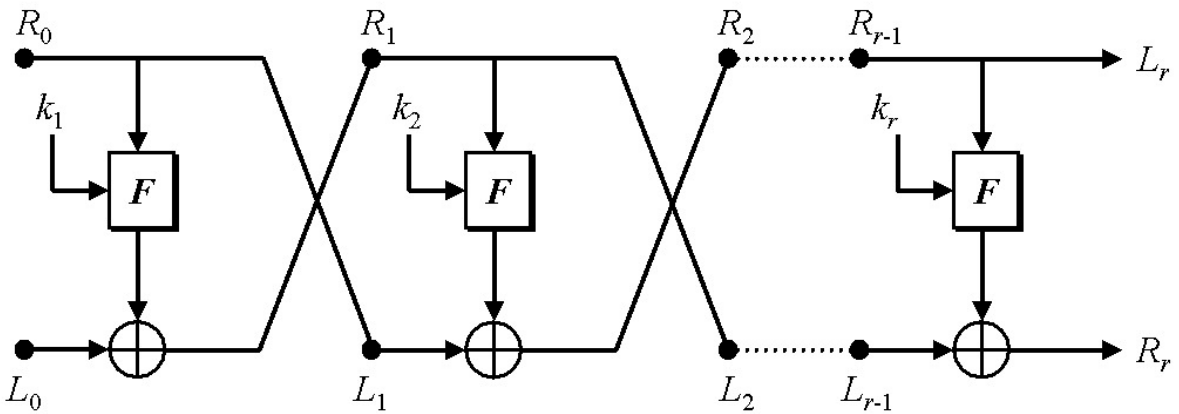
Only a very few mathematical formalizations of those two general principles (Filiol, 2002). The link between plaintext, ciphertext and key bits must be complex enough to be exploited by the cryptanalyst.

Block Iterated Encryption



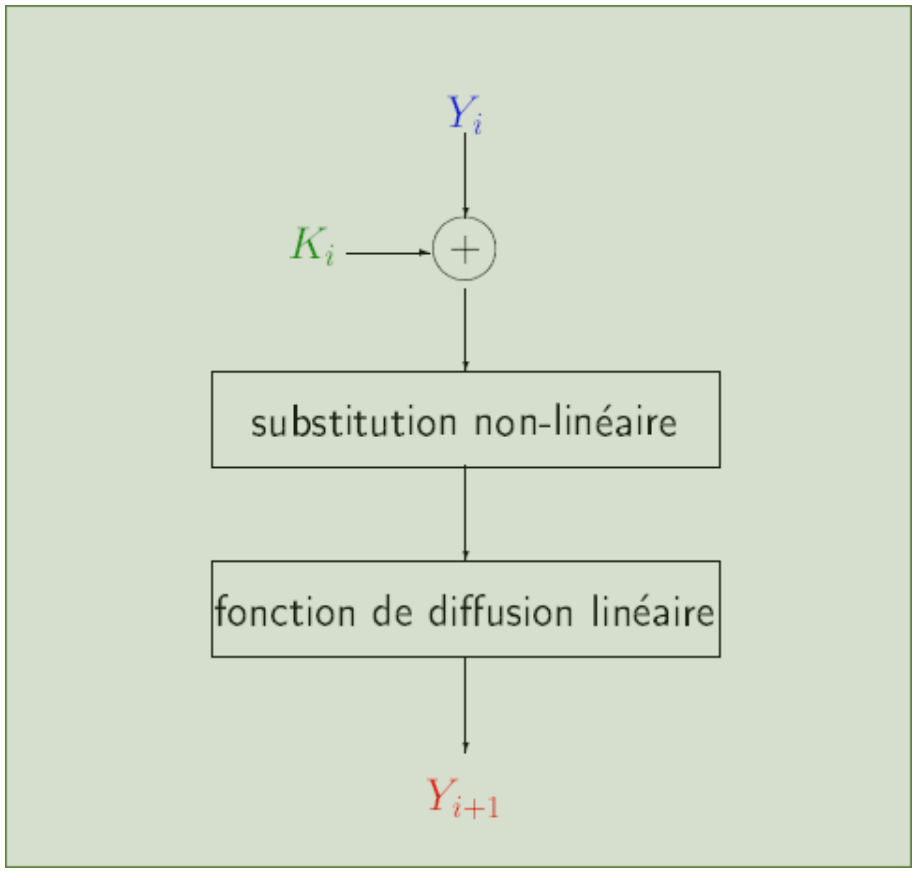
- F operates as a permutation over the set of n-bit words.
- DES : 16 rounds, $K = 56$ bits and $K_i = 48$ bits.
- AES-128 : 10 tours, $K = K_i = 128$ bits.

Feistel Scheme Principle



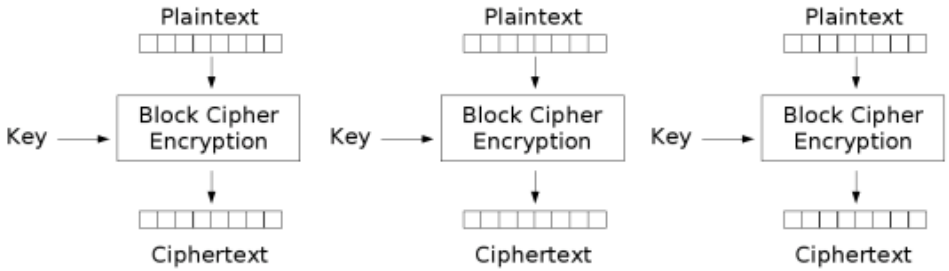
- The iterated function is an involution even if the core function F is not a bijection.

Substitution/Permutation Network (SPN) Principle



Modes of Operation : ECB Mode

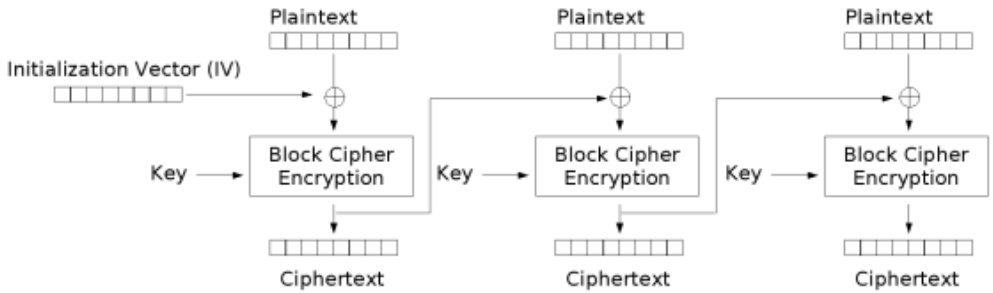
Block ciphers must use special modes to recover from noise and synchronization issues (due to avalanch effets).



Electronic Codebook (ECB) mode encryption

- *Electronic CodeBook* mode.
- Two identical plaintext blocks will result in the same ciphertext block for any key.
- Any single error on a ciphertext block cannot be recovered (avalanch effect : $\frac{n}{2}$ wrong bits).
- Any synchronization error cannot be recovered.

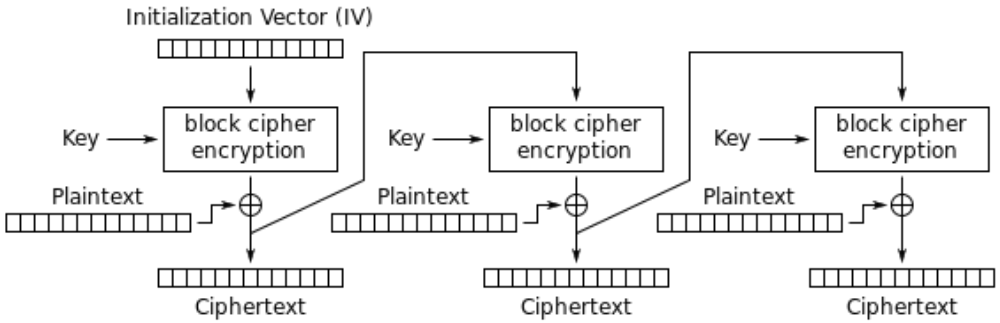
Modes of Operation : CBC Mode



Cipher Block Chaining (CBC) mode encryption

- *Cipher Block Chaining* mode (ciphertext feedback).
- Two identical plaintext blocks will result in two different ciphertext blocks for any key.
- Any single error on a ciphertext block will propagate on the next block only (self-correcting capability).
- Any synchronization error cannot be recovered.

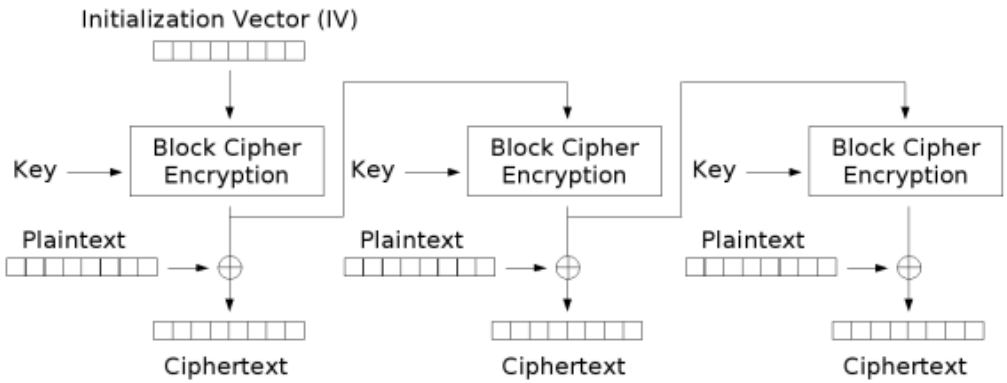
Modes of Operation : CFB Mode



Cipher Feedback (CFB) mode encryption

- *Cipher Block Chaining* mode (stream cipher emulation).
- Any single error on a ciphertext block will propagate on the next block only.
- Self-synchronizing cipher (recovers from any synchronization error).

Modes of Operation : OFB Mode



Output Feedback (OFB) mode encryption

- *Output Feedback Block* mode (stream cipher emulation).
- Any single error on a ciphertext block will result in a single error on the plaintext.
- Any synchronization error cannot be recovered.

Modes of Operation : PCBC Mode

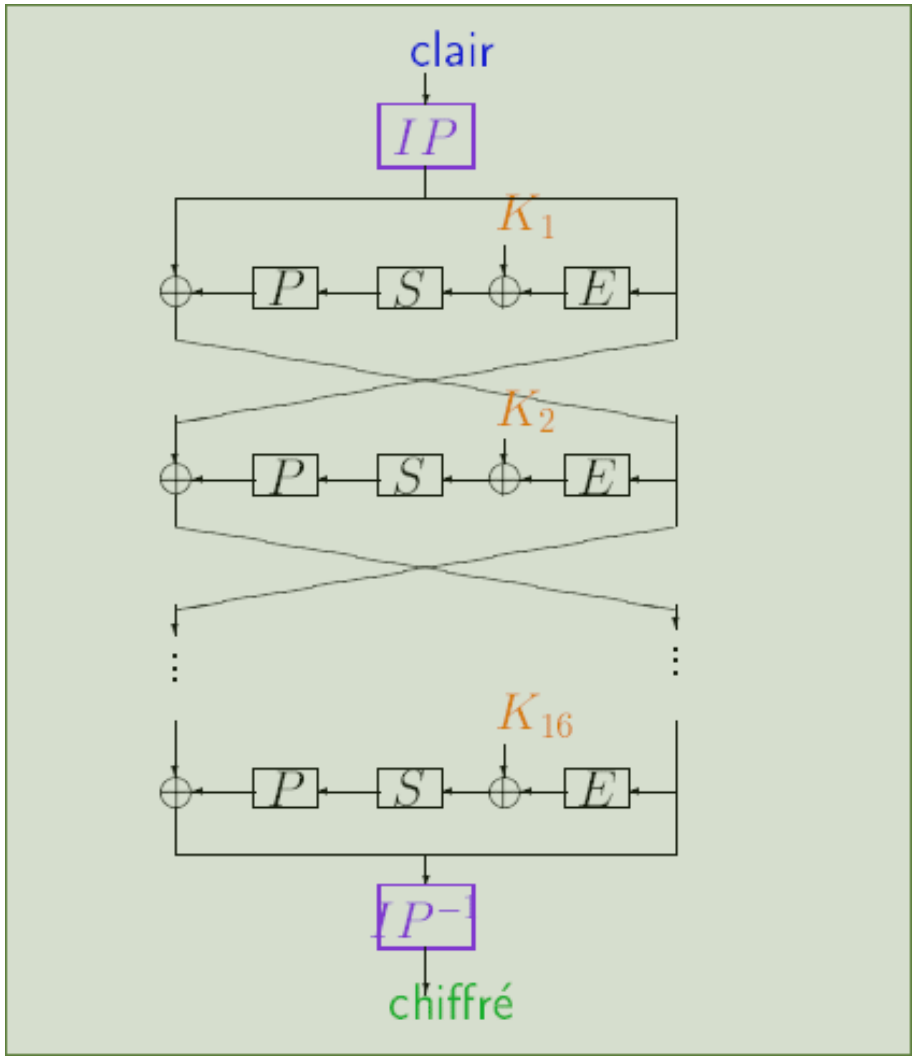
- *Propagating Cipher Block Chaining* mode (plaintext feedback)

$$C_i = E_K(P_i \oplus C_{i-1} \oplus P_{i-1})$$

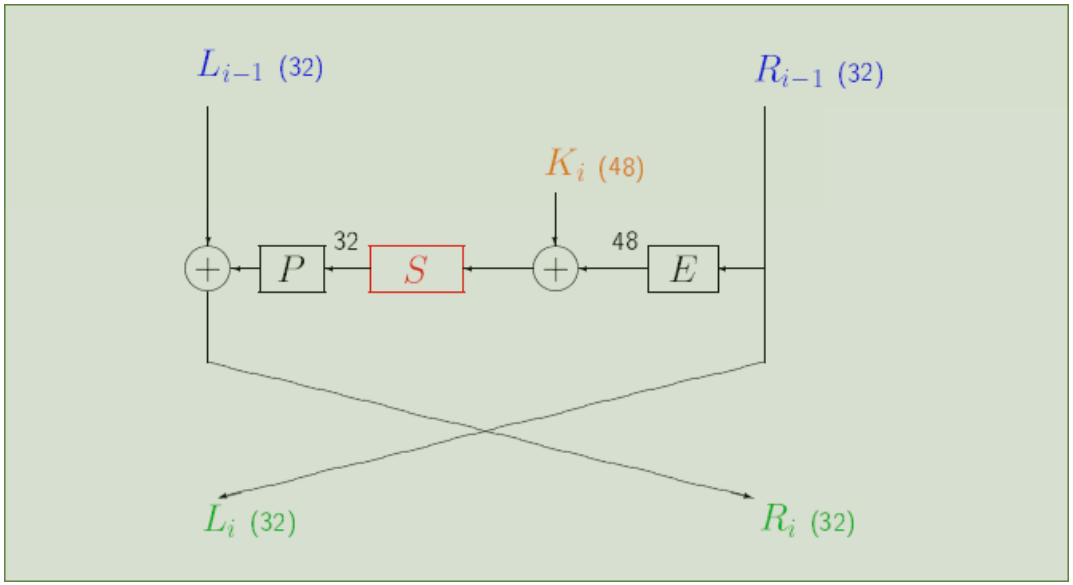
$$P_i = C_{i-1} \oplus P_{i-1} \oplus D_K(C_i)$$

- Provide encryption and integrity at the same time.
- Any single error on a ciphertext block will make decipherment impossible (endless error propagation).
- Used in the Kerberos protocol version 4 (authentication protocol with trusted third party in TCP/IP network)

The Data Encryption Standard (DES)



DES Iterated Function



- E expands a 32-bit block into a 48-bit block.
- S transforms a 48-bit block into a 32-bit block.
- P is a 32-bit permutation.

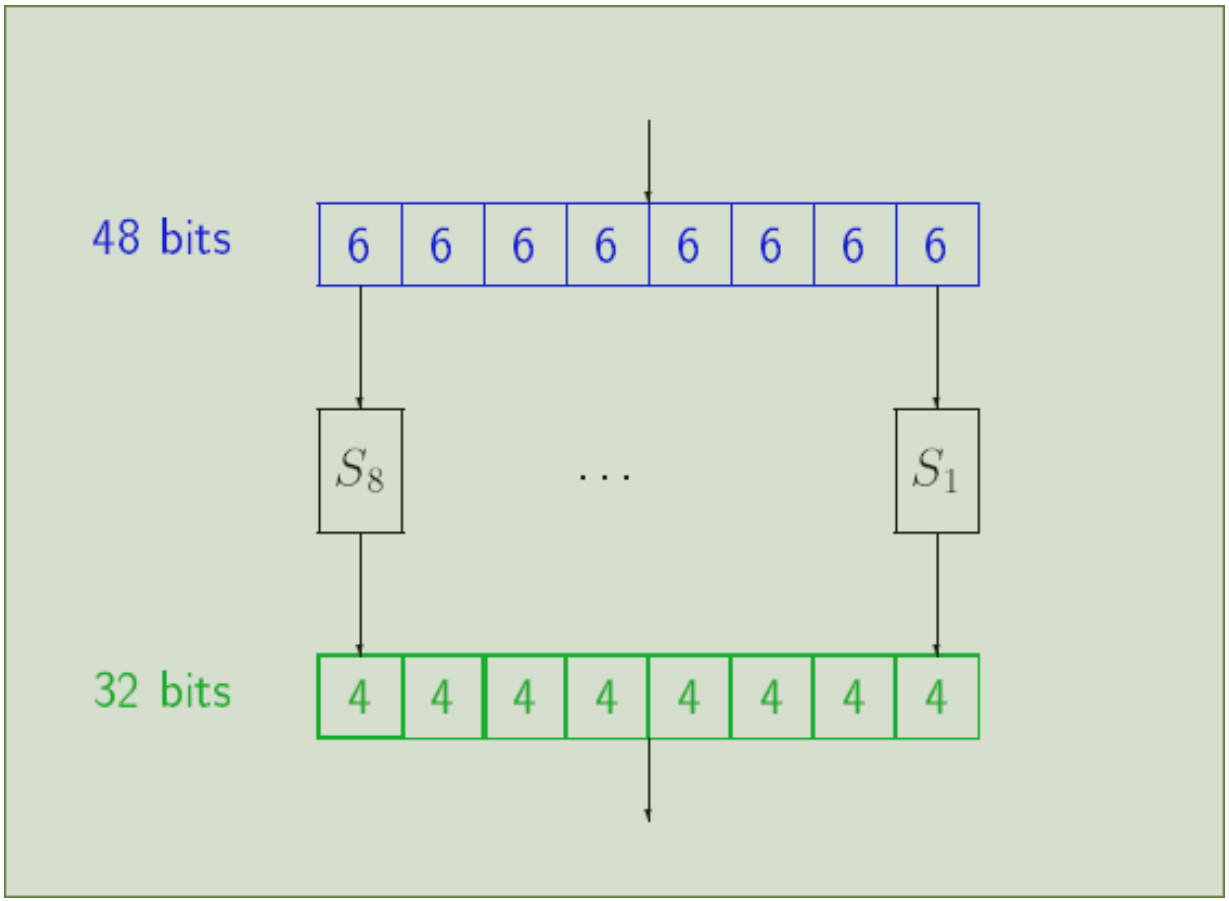
The Expansion Function

$$E : (x_1, x_2, \ldots, x_{32}) \longmapsto (y_1, y_2, \ldots, y_{48})$$

where bits of y correspond to bits of x taken into the following order :

E					
32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

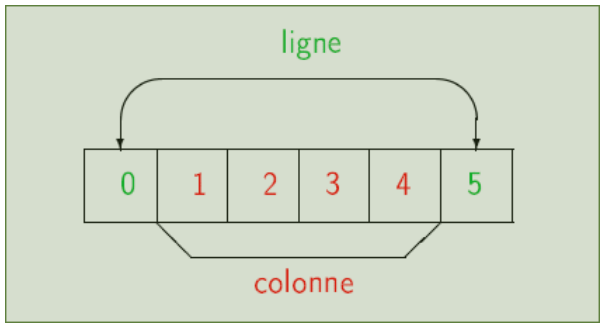
Data Encryption Standard S-boxes



Data Encryption Standard S-box 1

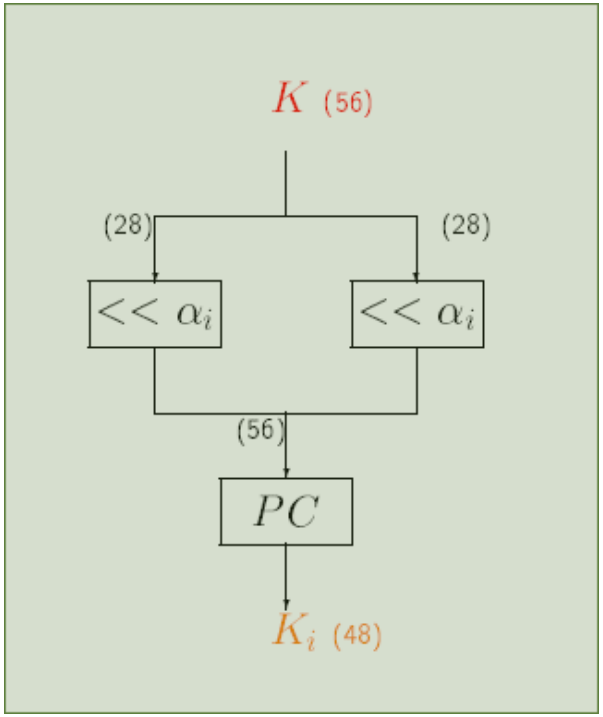
- Represented as a 4×16 array of nibbles (4 bits)

S_1															
14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13



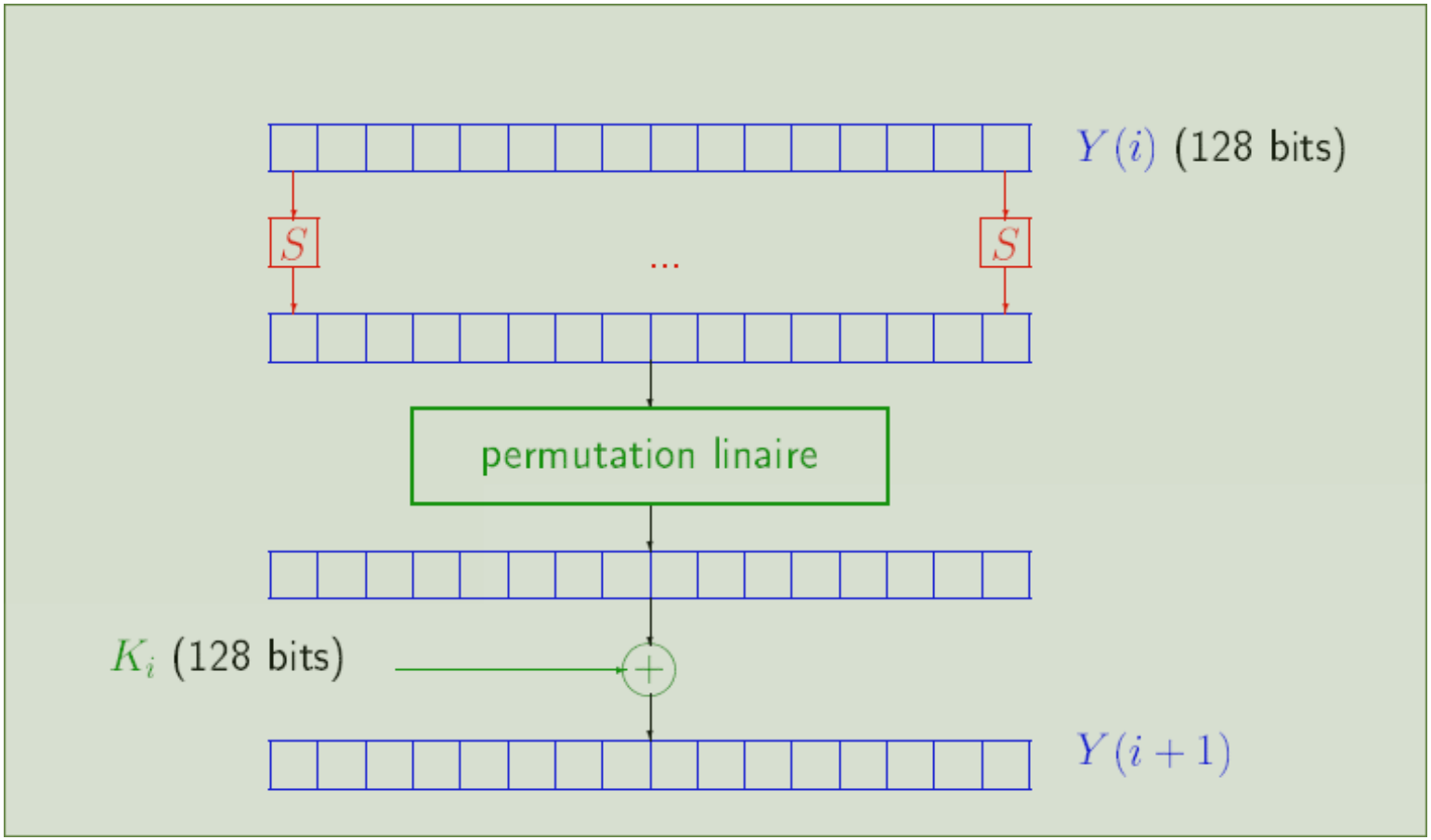
- Example : 1011 $\Rightarrow$ line 2 (10) & column 11 (1011) $\Rightarrow$ output = 0111 (7)

Data Encryption Standard Key Scheduling



tour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
α_i	1	1	2	2	2	2	2	2	1	2	2	2	2	2	2	1

Advanced Encryption Standard (AES)



Agenda

Introduction

Kerckhoffs' hypothesis (1883) : the attacker knows the algorithm. He tries to guess the secret key.

- **Ciphertext-only attack** : the attacker has only ciphertext blocks.
- **Known plaintext attack** : the attacker has pairs of {plaintext, ciphertext} blocks.
- **Chosen ciphertext attack** : the attacker has pairs of {plaintext, ciphertext} blocks corresponding to plaintext blocks of his choice.
- **Adaptative chosen ciphertext attack** : the attacker chooses plaintext block while adapting his choice to the corresponding ciphertext blocks he gets.

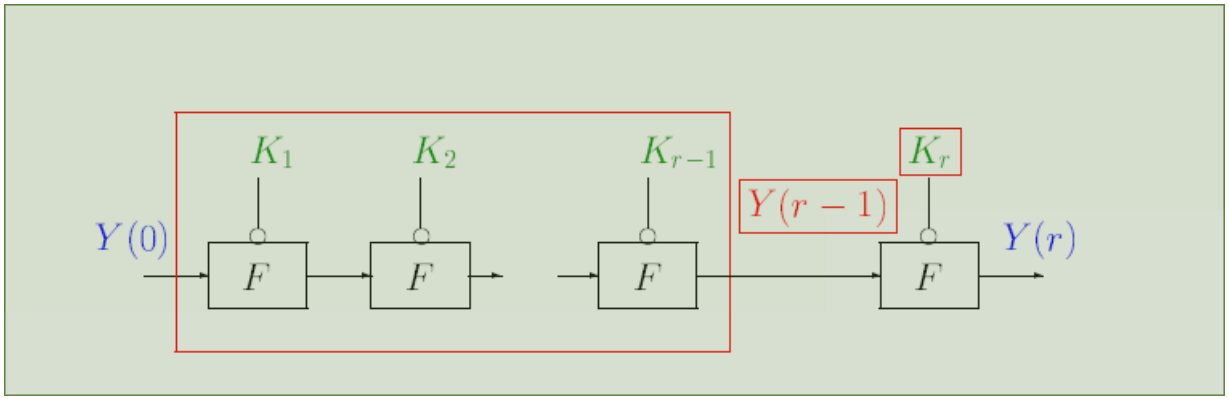
Exhaustive Key-search

- We need to perform 2^k encryption for a k-bit key.
- For $k = 56$ bits :
  - 41 days on 10,000 PC (february 1998).
  - 56 hours with a dedicated computer (EFF DES cracker).
  - 22 hours with EFF DES Cracker v2.0 and 100,000 PC (january 1999).
  - non-official record (2 hours with FPGA technology).
  - **Recommanded key-size** : 80 bits at least.
- Exhaustive key-search are no longer efficient for present day keys.

DES Cryptanalysis State-of-the-Art

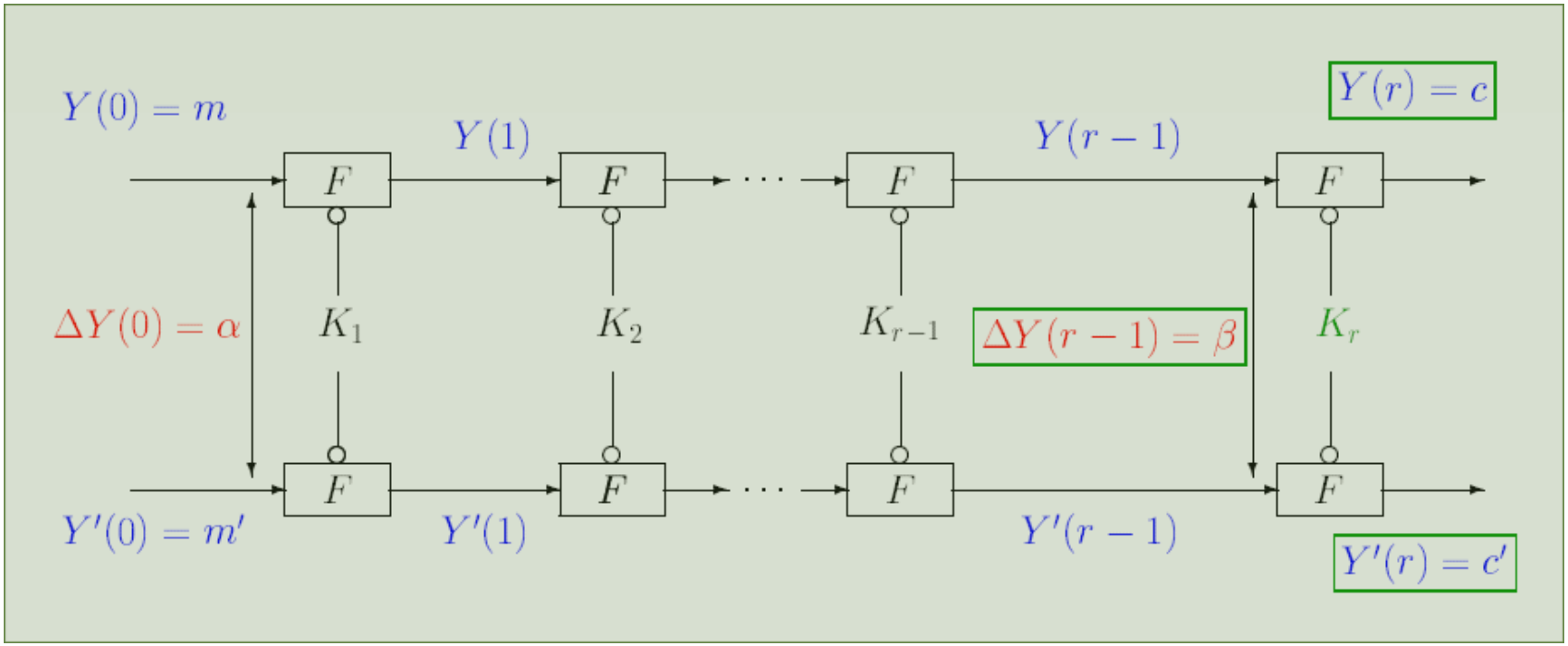
- **Differential cryptanalyse** : Biham - Shamir (1990).
  - For the DES, we need to have 2^{47} pairs of chosen {plaintext, ciphertext} blocks.
- **Linear cryptanalysis** : Matsui (1993).
  - For the DES, we need to have 2^{47} pairs of known {plaintext, ciphertext} blocks.
- A few other variants :
  - Quadratic approximation attacks (Knudsen/Robshaw - 1996).
  - Interpolation attacks (Jacobsen/Knudsen - 1997 & Jacobsen - 1998).
  - Algebraic attacks with quadratic equations (Courtois/Piepzryk - 2002).

Last Round Attacks : General Principle



- We identify a statistical bias in the distribution of $(Y(0), Y(r-1))$
- For any value of a subset of bits from K_r we compute $Y(r-1) = F^{-1}(Y(r), K_r)$.
- The comparison of the theoretical and observed distribution enables to guess the right value of the subset of bits from K_r.

Differential Cryptanalysis : General Principle



where $\Delta X = X \oplus X'$

Linear Cryptanalysis : General Principle

- Known plaintext attack.
- Let $F : (X, K) \mapsto F_K(X)$ the iterated permutation.
- We look for a linear approximation between some key bits, some plaintext bits and some ciphertext bits, which holds with a "high" probability :

$$< a \cdot X > + < b \cdot F_K(X) > + < c \cdot K > \approx \varepsilon$$

- By chaining such linear approximations, we get a final and general linear approximation regarding the last round $Y(r-1)$:

$$\alpha \cdot Y(0) + \beta \cdot Y(r-1) + \gamma \cdot (K_1, \ldots, K_{r-1}) + \varepsilon \approx \text{constant} \ .$$

Agenda

Conclusion

- Symmetric cryptology provides confidentiality.
- Based on Shannon's information theory.
- We can provide security proof.
- Stream ciphers should be preferred to block ciphers.