Introduction à la théorie de l'information et de la cryptographie

2005

INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON

FORMATION

IF

INFORMATIQUE

Auteur :
Gérard Beuchot

membre de UNIVERSITÉ DE LYON

INSA LYON

Introduction à la Théorie de l'Information et à la cryptographie

L’information, grandeur mesurable

Définition de l'information

- Une information est un couple constitué:

  d'une représentation matérielle, qui en constitue le formant

  et d'un ensemble d'interprétations, qui en constitue le formé

  dont la nature, événementielle, consiste en un changement d'état qui, par l'occurrence de cette représentation matérielle, provoque l'activation du champ interprétatif correspondant, selon les règles fixées par un code préétabli.

  Georges Ifrah - Histoire universelle des chiffres

  - Seule la composante matérielle (formant) d'une information fait l'objet d'une communication: ce n'est pas le sens (formé) que l'on transmet
  - L'information est la troisième dimension universelle après la matière et l'énergie. L'information n'est autre que la néguentropie (structure ordonnée)
  - Étymologie: informare= donner une forme....

- Quantité d'information

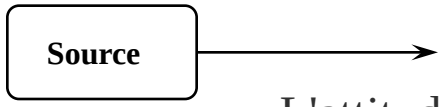
  Mesure quantitative de l'incertitude d'un message en fonction du degré de probabilité de chaque signal composant ce message

Messages et signaux

- Information (suite...)
  - Séquence de signaux, correspondant à des règles de combinaisons précises, transmise entre une source et un collecteur par l'intermédiaire d'un canal
  - Ecrit, fait, notion ou instruction susceptible d‘être traitée en tout ou partie par des moyens automatiques.
  - Renseignements obtenus de quelqu'un ou sur quelqu'un ou quelque chose, en particulier nouvelle communiquée par la presse, la radio,...
- Message
  - Lot d'information formant un tout intelligible ou exploitable et transmis en une seule fois
  - Séquence de signaux qui correspondent à des règles de combinaisons précises et qu'une source transmet à un collecteur par l'intermédiaire d'un canal
- Signal
  - Phénomène physique porteur d'une information et pouvant représenter des données
  - Variation d'une grandeur de nature quelconque grâce à laquelle, dans un équipement, un élément en influence un autre
  - Signe convenu pour avertir, annoncer, donner un ordre.

L'Information, Grandeur mesurable: Approche empirique

- On ne s'intéresse ici qu'à l'aspect QUANTITATIF
- L'aspect qualitatif de l'information, son "intérêt" est subjectif
- Mesure basée sur la probabilité d'occurrence d'un événement, du message



- L'attitude du collecteur de l'information est probabiliste
  - La communication n'a d'objet que si le contenu du message est inconnu
- Pour un événement X de probabilité P(X)
  - I(X) = f (1 / P(X)) pour que I(X) croît quand P(X) décroît
  - Il faut que I(X) soit toujours POSITIVE et ADDITIVE
    - on choisit donc f= log

$$I(X) = \log \frac{1}{P(X)} = -\log(P(X))$$

- si $\log_e$ unité : Nat
- si $\log_{10}$ unité : Hartley
- si $\log_2$ unité: bit

L'information, grandeur mesurable: Notations

- X — Symbole émis par une source
- Y — Symbole reçu par l'observateur au collecteur
- $P(x_k)$ — Probabilité que $X = x_k$
- $P(x_k / y_j)$ — Probabilité d'avoir émis $X = x_k$ si on a reçu $Y = y_j$
- $P(x_k ; y_j)$ — Probabilité émettre $X = x_k$ et de recevoir $Y = y_j$

$$I(x_k) = -\log_2 P(x_k) \text{ bits}$$

Définitions

- Alphabet: Ensemble fini de symboles appelés lettres
  - ex : a b c d e ...
  - alphabet binaire : 0 1
- Message: Suite finie de symboles ou lettres
  - ex : Beuchot, 0110100110
- Source de messages: Ensemble de TOUS les messages susceptibles d'être formés à partir d'un alphabet
  - Source discrète ou continue
  - ex : dictionnaire
  - alphabet N° 5 (AI5) à 7 bits (0000000 à 1111111)
- Extension d'une source
  - Soit une source codée par un alphabet de taille k *par exemple k = 2 [0,1]*

  Une séquence de longueur l de lettres de cet alphabet constitue une nouvelle source appelée l ème extension de k
  - ex : Le code binaire correspondant à l'alphabet AI5 est une extension de taille 7 de l'alphabet binaire.

Information moyenne: Approche axiomatique

- Source X de messages $x_1, x_2, \ldots x_i \ldots x_n$ de probabilité $p_1, p_2 \ldots p_i \ldots p_n$

$$p_i > 0 \quad \forall\, i \qquad \text{et} \qquad \sum_{i=1}^{n} p_i = 1$$

$$\text{Information moyenne} \quad H = \sum_{i=1}^{n} p_i I(x_i)$$

- Axiomes de Fadeev et Fenstein
  - si $p_1 = p_2 . = .. p_i .. = . p_n = 1/n$, H est fonction monotone croissante de n
  - Soit 2 ensembles $X(x_i)$ et $Y(y_j)$ sources de messages indépendantes
    et $p(x_i) = 1/n \quad \forall\, i$ et $p(y_j) = 1/m \quad \forall\, j$

    $$H(\ldots, 1/mn, \ldots) = H(\ldots, 1/n, \ldots) + H(\ldots, 1/m, \ldots)$$
  - (voir polycop) . On n'apporte pas plus d'information en fractionnant les expériences ....
  - Pour un alphabet binaire, H(p, 1-p) est une fonction continue de p sur [0,1]

Information moyenne: Approche axiomatique (suite)

- La seule fonction vérifiant ces 4 axiomes est :

$$H(p_1,\dots p_i,\dots p_n) = - C\sum_{i=1}^{n} p_i \log_a p_i$$

$$\text{Information moyenne } H = \sum_{i=1}^{n} p_i I(x_i)$$

  - C constante arbitraire > 0 et a base du logarithme

$$I(p_i) = - C \log_a p_i$$

  - I : information moyenne associée au résultat d'un épreuve
  - H : information moyenne, espérance mathématique de I
    - ENTROPIE (Shannon, 1948)

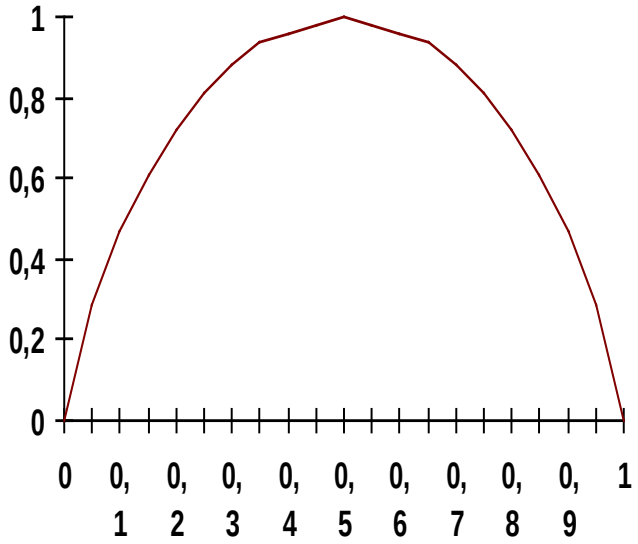
Exemples

- Alphabet + séparateur: 27 symboles
  - si équiprobables

$$H = \frac{1}{27}\sum_{i=1}^{27} \log_e \frac{1}{1/27} = \log_2 27 = 4{,}7549 \quad \text{bit par lettre}$$

  - en réalité, les lettres ne sont pas équiprobables et H = 3,98 bit par lettre
    - cette inégalité fait perdre 770 bits pour 1000 lettres
- Source binaire
  - quelconque
    - si p = 0,6 pour X_0 et p = 0,4 pour X_1
      - $I(X_0) = \log_2 0{,}6 = 0{,}73$
      - $I(X_1) = \log_2 0{,}4 = 1{,}32$ soit H = 0,97 bit
  - symétrique
    - p = 0,5 pour X_0 et X_1
      - $I(X_0) = \log_2 0{,}5 = 1 = I(X_1)$ soit H = 1 bit
- de équiprobable H = 2,585 bits

Propriétés de l'ENTROPIE



- l'entropie est maximale si les messages sont équiprobables

- Loi composée

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\log\left(p(x_{i,}y_j)\right)$$

  - généralisable à N sources

$$H(X,Y) \leq H(X) + H(Y)$$

    - avec égalité si indépendance des sources
    - généralisé par

$$H(X_1,..X_n) \leq \sum_{i=1}^{n} H(X_i)$$

Propriétés de l'ENTROPIE (suite)

$$p\left(Y = y_j / X = x_i\right)$$

$$H\left(Y / X = x_i\right) = -\sum_{j=1}^{k} p\left(y_j / x_i\right) \log\ p\left(y_j / x_i\right)$$

$$H\left(Y / X\right) = -\sum_{i=1}^{n} p(x_i) \sum_{j=1}^{m} p\left(y_j / x_i\right) \log\ p\left(y_j / x_i\right)$$

$$H\left(Y / X\right) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_{i,} y_j) \log\ p\left(y_j / x_i\right)$$

$$H(X,Y) = H(Y) + H\left(X / Y\right) = H(X) + H\left(Y / X\right)$$

$$H\left(Y / X\right) \leq H(Y)$$ égalité si indépendance

- Loi conditionnelle
  - L'entropie conditionnelle de Y relativement à $X = x_i$ est l'information moyenne apportée par Y si la source X a émis x_i
  - Si X et Y sont indépendantes, la connaissance de X n'apporte aucune information au sujet de Y. L'entropie conjointe H(X,Y) est alors la somme des entropies des sources indépendantes.
  - Si les sources sont liées, la connaissance de l'une apporte des information sur l'autre et l'entropie conjointe est inférieure à la somme des entropies de chaque source.

INFORMATION TRANSMISE - Equivocation

$$I(X,Y) = H(X) - H(X/Y)$$

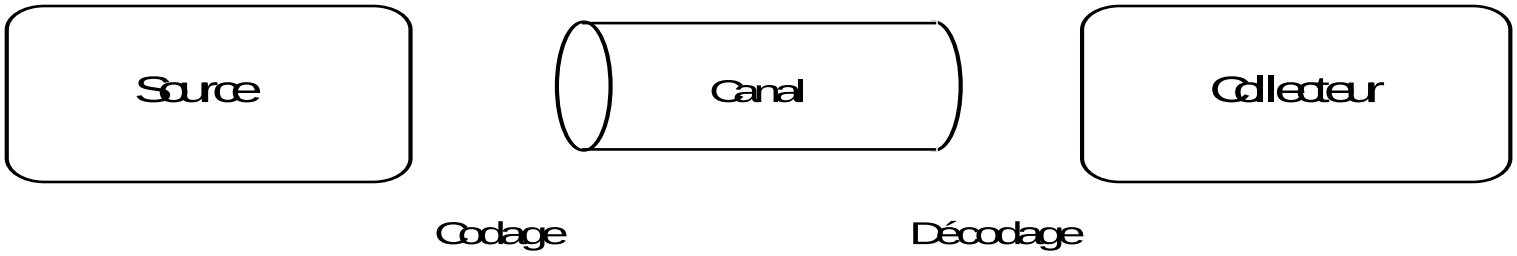
$$= H(Y) - H(Y/X)$$

$$I(X,Y) = I(Y,X)$$

– Notations
  - H(X) = entropie de la source
  - H(Y) = entropie du collecteur
  - H(X/Y) = partie de l'information source récupérée au vu de la sortie ou équivocation
  - H(Y/X) = bruit ou erreur sur le canal

– I(X,Y) est l'incertitude à priori sur X moins l'incertitude sur X lorsque Y a été réalisée
  - Pour un canal sans bruit I(X;Y) = H(X) = H(Y)
  - L'incertitude sur X décroît par la connaissance de Y
  - I(X,Y) est l'apport d'information de Y au sujet de X
  - $0 \leq I(X,Y)$. I(X,Y) = 0 si les sources sont indépendantes; Y ne dit rien de X

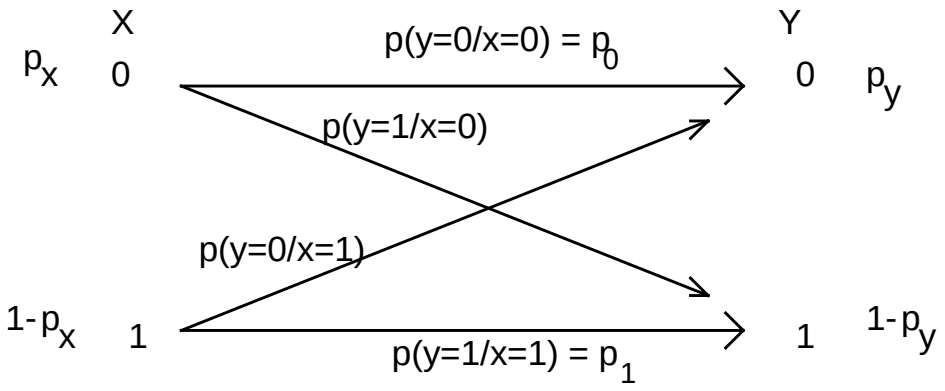
Application: CANAL de COMMUNICATION



- Au collecteur on observe Y qui est une conséquence de X
  - Stockage
  - Transmission
- Dans le canal l'information peut être perturbée, modifiée ....

Exemple: CANAL BINAIRE SYMETRIQUE

- $p_y = p_0 p_x + (1-p_1)(1-p_x)$
  - si $p_x = 0,5$ $p_y = 0,5\ (1 + p_0 - p_1)$
- Le canal est symétrique si $p_0 = p_1$ alors $p_y = 0,5$



$$H(Y/X) = -\sum_{i=1}^{2} p(x_i) \sum_{j=1}^{2} p(y_j/x_i) \log p(y_j/x_i)$$

$$H(Y/X) = -\left[\begin{array}{l} p(x=0)\left(p(0/0)\log p(0/0) + p(1/0)\log p(1/0)\right) + \\ p(x=1)\left(p(0/1)\log p(0/1) + p(1/1)\log p(1/1)\right) \end{array} \right]$$

si $p(x) = 0,5$ et $p(0/0) = p(1/1) = p$

$H(Y/X) = p \log(p) + (1-p) \log(1-p)$

p	H (Y / X)	I (X,Y)
1	0	1
0,9999	0,001473	0,9985
0,999	0,011408	0,988592
0,99	0,0808	0,9192
0,9	0,4690	0,5310
0,5	1	0

Le Canal de Transmission

- Entre source et collecteur caractérisé par :
  - alphabets d'entrée et de sortie
  - probabilités de transition
  - ensemble d'états

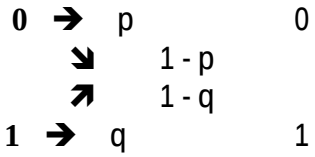
- si le canal est SANS MEMOIRE les probabilités de transition sont indépendantes de l'état du canal
  - Nous ne traiterons que ce cas
  - Une "mémoire" est un canal sans mémoire

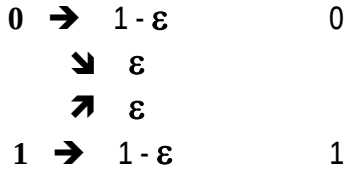
Canal sans mémoire

- sans perte (sortie implique l'entrée)
- déterministe (entrée détermine la sortie)
- sans erreur (déterministe et sans perte)

symétrique (matrice de transition symétrique)

- exemple : canal binaire symétrique
  - (B.S.C. binary symetric channel)





Extension du Canal

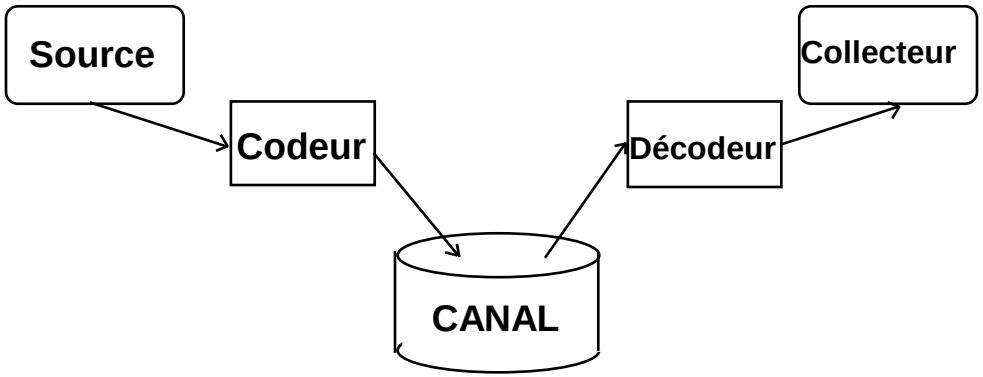
- Entrée prise dans l'extension de l'alphabet
  - ex : $p(00/00) = p^2$, $p(10/00) = qp$
  - Codage multiniveaux
- exemple:
  - matrice de probabilité de transition pour BSC

	00	01	10	11
00	$(1-e)^2$	$e(1-e)$	$(1-e)$	e^2
01	$e(1-e)$	$(1-e)^2$	e^2	$e(1-e)$
10	$e(1-e)$	e^2	$(1-e)^2$	$e(1-e)$
11	e^2	$e(1-e)$	$e(1-e)$	$(1-e)^2$

- CANAL AVEC MEMOIRE
  - La matrice de transition dépend de l'état du canal, donc des transitions antérieures
  - Application : paquets d'erreurs

Capacité du canal sans mémoire

- I(X,Y) = H(Y) - H(Y/X) = H(X) - H(X/Y)
  - Cette information dépend
    - de la source
    - de la matrice de transition fixée par le canal
  - cette matrice de transition est imposée.
- On recherche le maximum d'information transmise. Pour cela il faut jouer sur la source par un codage approprié.
  - analogie: adaptation d'impédance
  - On utilisera donc un codeur et un décodeur comme interfaces avec le canal



Capacité du Canal (suite)

$$H(Y) = - \sum_j p_{y_j} \log p_{y_j}$$

$$H(Y) = -\sum_j \left[\sum_i p_{x_i} \log p_{y_j/x_i} \right] \log \left(\sum_i p_{x_i} \log p_{y_j/x_i} \right)$$

$$H(Y/X) = - \sum_i \sum_j p_{x_i} p_{y_j/x_i} \log p_{y_j/x_i}$$

- $C = \underset{p(x_i)}{\text{Max}}[X,Y]$

$$I(X,Y) = - \sum_i \sum_j p_{x_i} p_{y_j/x_i} \left[\log \sum_i p_{x_i} p_{y_j/x_i} - \log p_{y_j/x_i} \right]$$

$$I(X,Y) = \sum_i \sum_j p_{x_i} p_{y_j/x_i} \log \frac{p_{y_j/x_i}}{\sum_i p_{x_i} p_{y_j/x_i}}$$

Capacité du Canal Symétrique

- Si L est la taille de l'alphabet B du collecteur

$$C = \text{Max}\,(H(Y) - H(Y/X))$$

  - H(Y) est maximal pour une source équiprobable

$$p(x_i) = 1 / L \quad \text{et} \quad \sum_{i=1}^{L} p_{x_i} = 1$$

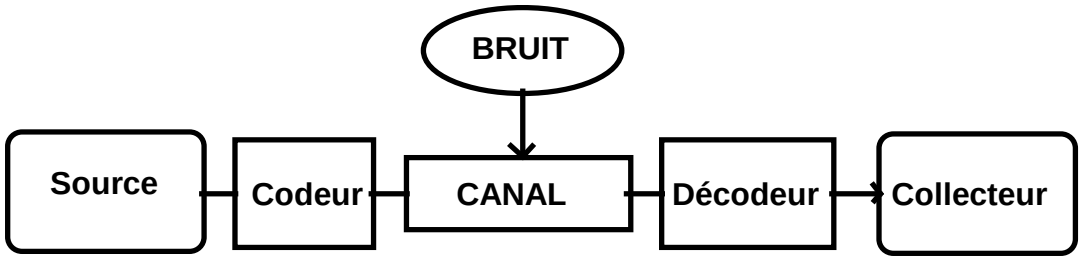
$$C = \log L + \sum_{j=1}^{L} p\left(\frac{y_j}{x_i}\right) \log p\left(\frac{y_j}{x_i}\right)$$

- Pour un canal binaire symétrique (BSC)

$$C = 1 + p \log p + (1-p) \log (1-p) \quad 1$$

  - si p = 0,5 log p = log (1-p) = log 0,5 = - 1 C = 0 bit
  - si p = 0,99 log p = -0,014495
    1-p = 0,01 log(1-p) -6,643856 C = 0,9192 bit

Canal Bruité: probabilité d'erreurs



- La matrice de transition du canal est liée à des perturbations induites par l'ajout d'une information parasite "le bruit" dans le canal.
- Exemple : Canal binaire
  - Source 1 1 0 1 1 0 0 1 0
  - bruit 0 0 0 1 0 1 0 0 0
  - Collecteur 1 1 0 0 1 1 0 1 0
  - erreur + +
- Nous observons 2 erreurs de décodage. Nous allons rechercher la probabilité d'erreur p_{ei} sur un mot X_i et la probabilité moyenne d'erreurs p_e.

Canal Bruité: probabilité d'erreurs (suite)

$$P_{e_i} = \sum_{Y_j} p(Y_j / X_i)$$

$$P_e = \sum_{i=1}^{M} p(X_i)\, p_{e_i}$$

- A l'entrée du canal sont placés M mots X_i de longueur N. A chaque mot reçu Y_i on assigne un mot émis X_i. Il y a erreur si on reçoit Y_j .
- Le décodeur, placé à la sortie du canal, doit minimiser cette probabilité d'erreurs en recherchant :
  - le maximum de la loi de probabilité à posteriori
  - ou le maximum de vraisemblance
- Pour une loi d'émission uniforme ces règles se confondent.
- Le calcul est en général très difficile. On se contente d'un majorant.

Canal Binaire Symétrique Bruité: probabilité d'erreurs

- Pour un canal B.S.C.
  - avec des probabilités de transition e et 1-e
  - des mots de taille N

$$P_e \leq \left(2\sqrt{\varepsilon(1-\varepsilon)}\right)^N$$

en réalité

$$P_e \approx \sqrt{\frac{2}{\pi N}}\left(2\sqrt{\varepsilon(1-\varepsilon)}\right)^N$$

Théorème Fondamental de SHANNON

- Soit une source émettant des mots de taille N
  - Son taux est R = H / N bit par lettre
  - La probabilité d'erreur p_e î si N ì mais R î si N ì
- En fait pour avoir une probabilité d'erreur $p_e = 0$ il faut N = ¥ soit R = 0 !
- On cherche à minimiser p_e
  - en fonction de R, N et de la capacité C du canal
  - Shannon a montré que si R £ C , il existe une suite de codes de longueur de mot N qui minimise la probabilité d'erreur.
  - La taille du code est *Partie entière* [exp (NR)]
  - Si R > C il n'existe aucune méthode de codage qui permette de transmettre de l'information avec un taux d'erreur négligeable.
  - Ce théorème fournit une condition limite qui permet d'orienter la recherche des codes les meilleurs ou de les comparer à cette valeur limite.

Le Canal Continu

$$H(X) = -\int_{-\infty}^{+\infty} f_1(x) \log f_1(x)\, dx$$

$$H(Y) = -\int_{-\infty}^{+\infty} f_2(y) \log f_2(y)\, dy$$

$$H(X/Y) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y) \log \frac{f(x,y)}{f_2(y)} dxdy$$

$$I(X,Y) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y) \log \frac{f(x,y)}{f_{1(x).} f_2(y)} dxdy$$

Capacité du canal continu

$$C = \mathrm{Max}_{f_{1(x)}}\left[I(X,Y)\right]$$

- Les définitions de l'information, de l'entropie, etc. sont étendues à des sources ou des collecteurs ayant des fonctions densité de probabilité continues (et non plus discrètes) pour utiliser pleinement les canaux de communication analogiques
- Cette quantité d'information est liée à un symbole unique

Débit du Canal Continu

- Si T est l'intervalle de temps entre 2 symboles ("Moment")

- le DEBIT d'information vaut

$$C_T = C / T \quad \text{bit/s}$$

- Le taux de la source est $R_T = H(X) / T$

- D'après le second théorème de Shannon,
- si $R_T \leq C_T$ on peut avoir un taux d'erreurs aussi petit que l'on veut.

Théorème de Shannon-Hartley-Tuller

$$H(Z) = -\int_{-\infty}^{+\infty} q(z) \log q(z) dz$$

avec, si B est la puissance du bruit

$$q(z) = \frac{1}{\sqrt{2\pi B}} \exp\left(-\frac{z^2}{2B}\right)$$

et si S est la puissance du signal

$$f_1(x) = \frac{1}{\sqrt{2\pi S}} \exp\left(-\frac{x^2}{2S}\right)$$

- On suppose que le canal est additif et qu'il est perturbé par un "bruit" gaussien.
- Soit H(Z) l'entropie de cette" source de bruit" de densité de probabilité q(z)
- On suppose aussi que le signal utile issu de la source possède une densité de probabilité gaussienne $f_1(x)$

Théorème de Shannon-Hartley-Tuller (suite)

- H (X) = log (2peS)
- H (Z) = log (2peB)

- Y = X + Z donc H (Y) = log (2pe(S+B))

- I(X,Y) = H(Y) - H (Z)

  = log (2pe(S+B)) - log (2peB)

I (X,Y) =

$$\frac{1}{2} \log_2 \left(1 + \frac{S}{B} \right)$$

- Si B est minimal et S maximal C = $\frac{1}{2} \log_2 \left(1 + \frac{S}{B} \right)$
  - Nota : Pour avoir le débit maximal théorique d'information nous devrons multiplier cette capacité par le nombre de symboles transmis par unité de temps.

Exemple 1 de capacité du canal continu

- Bruit blanc maximum sur canal téléphonique :
  48 db au dessous du signal
  - nota : décibel = $10 \log_{10} \frac{P_2}{P_1}$
- soit $10 \log_{10}$ (S/B) = 48 ou S/B » 63000
- C = $\log_2$ (1 + S/B) = $\frac{1}{2} \frac{\log_{10}(1 + \frac{S}{B})}{\log_{10} 2}$
- C = 3,32 $\log_{10}$ (1 + S/B)
  - = 3,32 * 4,8 » 7,970 bits par symbole

Exemple 2 de capacité du canal continu

- Signal émis 1mW
  - nota : sur une ligne téléphonique
    - puissance crête 0 dbm = 1 mw
    - puissance moyenne -10 dbm = 0,1 mw
  - Atténuation 30 db soit un facteur 1000
- Signal reçu 1W
- Bruit à la réception 10 nW
  - S/B = 100 (20 db)
  - $\log_2 (1 + S/B) = \log_2 (101) = 6,658$

- C = 3,329 bit par symbole
  - En pratique on pourra espérer transmettre 3 bit par symbole

Codage

Définitions

- SOURCES
  - Indépendantes
    - x_i sont des var. aléatoires indépendantes de même loi , Si M = $x_1x_2x_3$ $p(M) = p_1p_2p_3$
  - Stationnaires
    - x_i sont liées mais les probabilités pour une suite $x_1x_{2...}x_i$ ne dépendent pas de l'instant d'émission
  - de Markov
    - Source à mémoire finie
    - loi à priori $p(x_i)$ + suite de lois conditionnelles $p(x_{1/}x_2)$, ...,$p(x_{1/}x_n)$

      $p(x_{1/}x_2x_3)$

      $p(x_{1/}x_1x_2x_{3...}x_n)$
      - On utilise un diagramme d'état pour indiquer les probabilités à prendre en compte à un instant donné
    - exemple : langue naturelle (...ait, ...ment, ..)
    - Codage prédictif: images Vidéo

Définitions (suite)

- LONGUEUR MOYENNE
  - n_i nombre de caractères du mot m_i codant un message x_i
  - Longueur moyenne des messages : $\bar{n} = E(n_i) = \sum_{i=1}^{N} n_i p(x_i)$

  Cette grandeur est aussi appelé le coût moyen par message
- TYPES DE CODES
  - A longueur de mot fixe
    - tous les mots sont de même longueur
    - exemple : AI5 (7 bit/lettre) , AI2 (5 bit/lette)
  - A longueur de mots variable
    - exemple : code Morse, Huffman, ...

Autres définitions

- DEBITS, ETC.
  - Débit littéral : nombre de caractères par unité de temps D lettres/ s
  - Taux d'émission : $\frac{H(X)}{\bar{n}}$ bit/lettre
  - Débit d'information : $D\frac{H(X)}{\bar{n}}$ bit/s
  - Efficacité : $\eta = \frac{\bar{n}_{min}}{\bar{n}}$ $\bar{n}_{min}$ longueur moyenne pour le code le plus court
  - Redondance : $r = 1 - h$

Codes à longueur fixe

- Source codée par un alphabet de taille K : $a_1, a_2 \ldots a_k$
  - exemple [0,1] K = 2
- Soit une séquence de longueur l de lettres de cet alphabet
  - $x_1, x_2, \ldots\ldots x_l = \vec{x}_l$
  - On peut construire K^l exemple 2^l
  - Ces K^l séquences constituent une nouvelle source appelée l ème extension de K
  - exemples : AI2 l = 5 2^5 = 32 "lettres" de 00000 à 11111

    AI5 l = 7 2^7 = 128 "lettres de 0000000 à 1111111
- Pour coder ces suites par des mots de taille N à partir d'un alphabet de taille D,

on doit avoir $\frac{N}{l} \geq \frac{\log K}{\log D}$

- Exemple : Décimal codé binaire
  - D = 2 K = 10 chiffre de l = 1 lettre (0 à 9)
  - soit N = 4

$$N \geq 1 \frac{\log 10}{\log 2} = 3{,}32$$

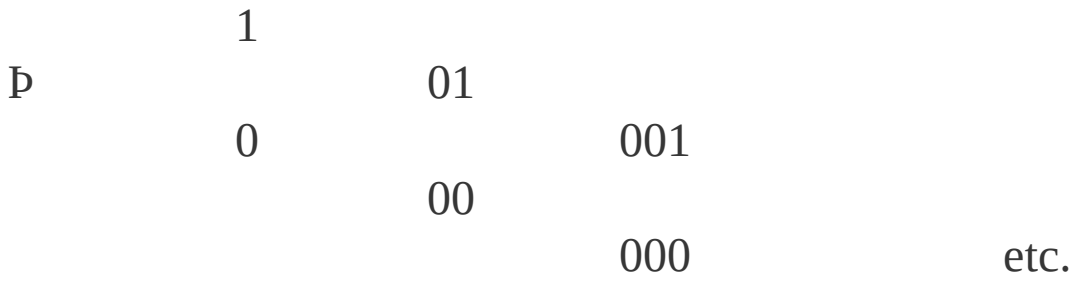
Codes à longueur variable

- Code caractérisé par :
  - longueur moyenne des mots
  - non ambiguïté de la lecture
- Code : partie d'un ensemble de suites finies de caractères issus d'un alphabet
  - Un tel code doit être REGULIER et DECHIFFRABLE
  - exemples :

	$P(x_i)$	code 1	code 2	code 3	code 4
x_1	0,5	1	0	1	1
x_2	0,25	1	1	10	01
x_3	0,15	0	11	100	001
x_4	0,1	00	01	1000	000
		non régulier pas injectif	non déchiffrable	non ambigu préfixé	irréductible

Classification des codes de longueur variable

- Un code est déchiffrable (conditions suffisantes) si
  - il a une longueur fixe
  - ou il est préfixé
- Code
  - singulier
  - régulier non déchiffrable
    déchiffrable
    - réductible
    - irréductible (ou instantané)
- un code irréductible est construit à l'aide d'un arbre



Longueur moyenne d'un code

- Soit un alphabet de taille D (nombre de mots code)
- Code irréductible

$$\bar{n} \leq \frac{H(X)}{\log D} + 1$$

- Code déchiffrable

$$\bar{n} \geq \frac{H(X)}{\log D}$$ Si égalité : code optimal

- Pour toute source indépendante, il existe un code irréductible de longueur moyenne aussi proche que l'on veut de
  - 1er Théorème de Shannon

$$\frac{H(X)}{\log D}$$

- Pour un code binaire $\min_{\bar{n}} = H(X)$ (log D =1)

Capacité de codage

- Soit k(N) le nombre de textes codés avec N lettres

$$k(N) \leq D^N$$

$$\lim_{N\to\infty} \frac{\log k(N)}{N} = C \leq \log D$$

$$\text{avec } C \text{ solution de } \sum_{i=1}^{N} 2^{-Cn_i} = 1$$

- C est la capacité de codage du code
- Exemple : code binaire [0,1] D = 2
  - longueur optimale

$$n_i = - \frac{\log p_i}{\log 2}$$

  - Si on classe $p_1 \geq p_2 \geq \ldots\ldots \geq p_{i\ldots\ldots} \geq p_n$ alors $n_1 \leq n_2 \leq \ldots\ldots \leq n_{i\ \ldots} \leq n_n$
    - On cherche les plus petits entiers tels que $p_i \geq 2^{-n_i}$

Code de Shannon-Fano

- A chaque étape on regroupe les messages en 2 sous-ensembles de probabilité la plus voisine possible. On continue jusqu'à obtenir 2 messages que l'on code.

- exemple 1
  - mes. $p(x_i)$ code
  - x_1 0,51 0,51 0 0
  - x_2 0,29 0,29 10 10
  - x_3 0,08 0,49 1 0,20 11 0,08 110 110
  - x_4 0,12 0,12 111 111

  - longueur moyenne : 0,51 + 0,29 * 2 + (0.08 +0,12) * 3 = 1,69

Code de Shannon-Fano (suite)

- Exemple 2

mes.	$p(x_i)$								
code									
x_1	0,4			0,4	10				
10									
x_2	0,3	0,5	0	0,3	00				
00									
x_3	0,2			0,2	01				
01									
x_4	0,05	0,5	1			0,05	110		
110									
x_5	0,03			0,1	11	0,05	111	0,03	1110
1110									
x_6	0,02							0,02	1111
1111									

– longueur moyenne : $\overline{n}_{min} = (0,4 + 0,3 + 0,2) * 2 + 0,05 * 3 + (0,03 + 0,02) * 4 = 2,15$

– $H(X) = 1,994987$

- $\log D = 1$ $1,995$
- $h = 0,9279$ $r = 0,072$

Code de Huffman

- Propriétés
  - si $p_i > p_j$ alors n_i n_j
  - les 2 mots les moins probables ont même longueur
  - parmi les mots de longueur maximale n_m il y en a au moins 2 ayant les mêmes n_{m-1} premières lettres
  - On classe les messages et on les regroupe 2 par 2
- exemple 1:

mes.	$p(x_i)$						code
x_1	0,51	0,51	0,51	0			0
x_2	0,29	0,29	0,49	1	10		10
x_3	0,12	0,20			11	110	110
x_4	0,08					111	111

  - Pour cet exemple même longueur moyenne que Shannon-Fano

Code de Huffman (suite)

- Exemple 2

mes.	$p(x_i)$										code
x_1	0,4	0,4	0,4	0,4	0,4	1					1
x_2	0,3	0,3	0,3	0,3	0,6	0	00				00
x_3	0,2	0,2	0,2	0,3			01	010			010
x_4	0,05	0,05	0,1					011	0110		0110
x_5	0,03	0,05						0111	01110	01110	
x_6	0,02									01111	01111

  - longueur moyenne : 0,4 + 0,3 * 2 + 0,2 * 3 + 0,05 *4 + (0,03+0,02) * 5 = 2,05
  - H(X) = 1,994987
    - log D = 1

$\overline{n}_{min} = \eta = 0,9732$ 1,995 $\rho = 0,027$

  - Ce code est optimal

Applications

- COMPRESSION DE DONNEES

Si des données sont codées avec des code de longueur fixe on peut les comprimer

- en enlevant les caractères inutiles
- en évitant les répétitions : Code RLE (Run Lenght Encoding)
  - xxxxxxxxxxx Rx11
- en utilisant un code de taille variable
  - Huffman statique ou adaptatif
  - Huffman -Shannon- Fano autosynchronisant ...
  - Ziv et Lempel (codage par dictionnaire)
    - certaines séquences de caractères sont considérés comme des messages valides : extension de l'alphabet
    - On applique un code d'Huffman sur cette extension

Applications (suite)

- On peut aussi utiliser
  - des codes Arithmétiques
  - des transformations mathématiques etc.
- nota : code JPEG (Joint Photographic Expert Group)
  - CCITT T.83 ou ISO 10917-1
- code MPEG (Motion Picture Expert Group)
  - CCITT ?? , ISO ??
    - MPEG2 Codage vidéo : prédictif + Code d'Huffman
    - CCITT H.261

- QUESTIONNAIRES DICHOTOMIQUES
  - Questionnaire à réponse par oui ou non à nombre de questions minimal

Codage des images vidéo : MPEG

- MPEG : Moving Picture Expert Group
  - MPEG-2 : ISO/IEC 13818
  - IUT : H26x (H262) , H32x (visioconférence sur ATM)
- Différents formats d'image
  - Niveau (résolution) (pel = picture element)
    - Low (SIF) : 352*288 pels
    - Main (625l) : 720*576 pels (aspect 4/3)
    - High1440 : 1440*1152 pels (aspect 4/3)
    - High (TVHD): 1920*1152 pels (aspect 16/9)
  - Profil (format) : décomposition des macroblocs
    - Simple : 4:2:0 (Main)
    - Main : 4:2:0 (tous niveaux)
    - SNR : 4:2:0 (Low, Main)
    - Spatial : 4:2:0 (High-1440)
    - High : 4:2:0, 4:2:2 (Main, High1440, TVHD)

Structure des images

- Séquence vidéo
- Images décomposées en trames (entrelacées ou non)
  - I : Intra Indépendante de toute autre (Non Prédite)
  - P : Prédites à partir des images I ou P Précédentes)
  - B : Bidirectionnelle (prédites par rapport à I ou P voisine)
- trame : ensembles de tranches (slice)
- tranche : un ou plusieurs macroblocs
- macrobloc :
  - 16*16pels (picture element)
  - luminance (jaune)+ Chrominance (2 couleurs : rouge +bleu)
  - 6, 8 ou 12 blocs ➔ 4:2:0 , 4:2:2 , 4:4:4
- bloc :
  - 8*8 pels luminance et 0 ou 8*8 chrominance

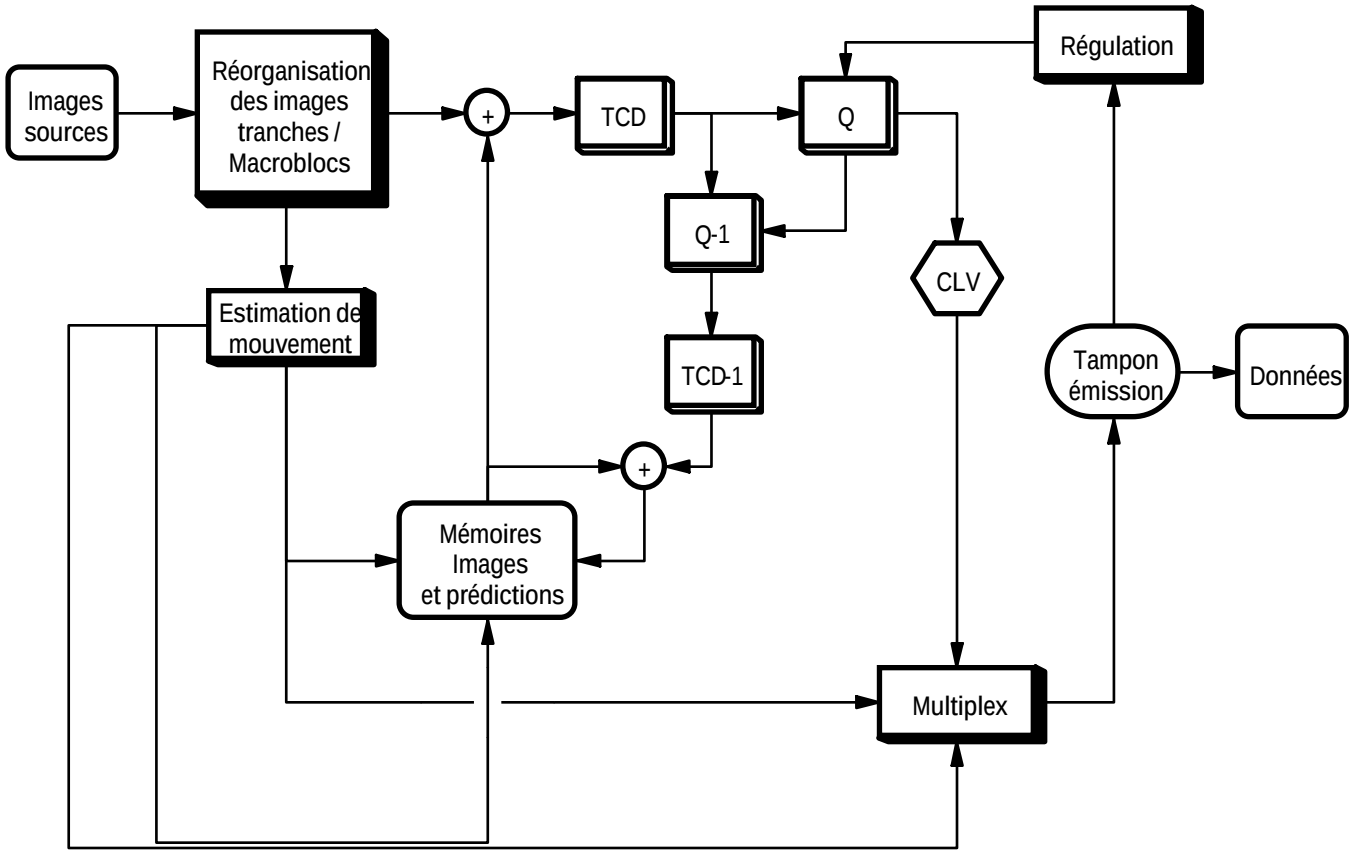
Codage et transfert des images

- Seuls les blocs différents d'une trame à la suivante sont transmis
  - L'adresse des blocs est codée par un code à longueur variable

- Seules les différences entre les blocs sont transmises
  - La représentation des blocs est transformée par DCT : Transformation en Cosinus Discrète
  - Les coefficients de la transformée sont codés sur 12 bits (Signe + 11 bits)
  - La différence entre les coefficients, pour deux trames successives, est codée par un code à longueur variable (3 à 17 bits)

- Les types de macroblocs et d'image sont codés par des codes à longueur variable

Codage MPEG pour images vidéo: principe



TCD:

Transformée en Cosinus Discrète

Q: Quantification

CLV:

Codage à longueur variable (Huffmann)

Codes détecteurs et correcteurs d'erreurs

Distance de Hamming

- Pour éviter les erreurs de décodage (dues à des perturbations dans le canal) les messages doivent le plus "différent" possible les uns des autres.
- On appelle DISTANCE DE HAMMING de deux mots codes et le nombre de chiffres par lequel ils diffèrent.
  - $\vec{x}_1$ = 00111010 $\vec{x}_2$ = 10110010 $d(\vec{x}_1, \vec{x}_2) = 2$
- Pour des mots de longueur N

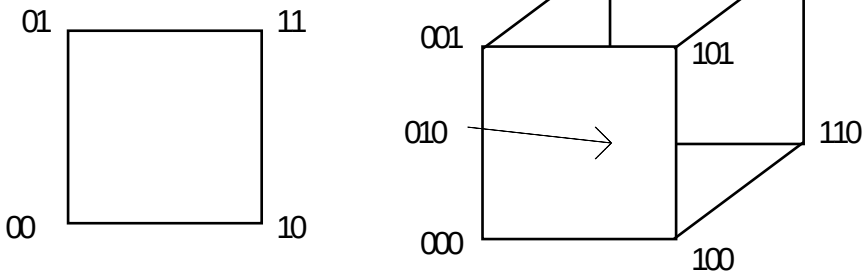
$$d(\vec{x}_1, \vec{x}_2) = \sum_{i=1}^{N} x_{1_i} \oplus x_{2_i}$$

- La DISTANCE DE HAMMING d'un code est la plus petite distance observée entre les mots du code (pris 2 à 2).

Distance de Hamming : Propriétés

- $d(\vec{X}_1, \vec{X}_2) = 0$ si et seulement si $= \vec{X}_1 \ \vec{X}_2$
- $d(\vec{X}_1, \vec{X}_2) = d(\vec{X}_2, \vec{X}_1) > 0$ si $\vec{X}_2 \neq \vec{X}_1$
- $d(\vec{X}_1, \vec{X}_3) + d(\vec{X}_2, \vec{X}_3) \geq d(\vec{X}_1, \vec{X}_2)$

  - Exemples "géométriques"



  - Si pour l'espace à 3 dimensions (N=3) on ne garde que des mots codes sur des sommets opposés (000 et 111 ou 011 et 100 par exemple) leur distance est d = 3

Règles de Décodage

- Pour un canal B.S.C. décoder selon la règle du maximum de vraisemblance revient à prendre le mot le plus proche, au sens de la distance de Hamming, d'un mot code.
  - Par exemple si le code choisi est [001,110]

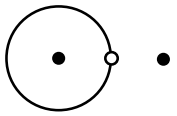
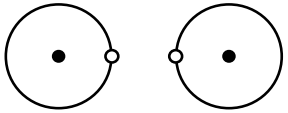
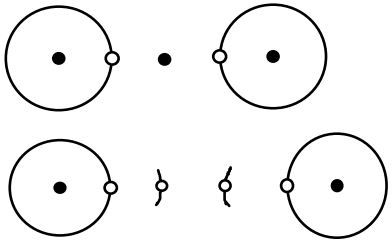
    et si on observe 000 à la sortie du canal,

    on décide que 001 a été émis à la source

- Ceci nous conduit à la notion de boule de décodage dans l'espace à N dimensions
  - Dans une boule de rayon r il y a $\sum_{i=0}^{N-1} C_N^i$ mots code
  - On pourra corriger une erreur simple ou multiple si le mot code erroné observé reste dans la boule de rayon r.
  - On pourra détecter une erreur simple ou multiple si le mot code observé n'est pas un mot du code.

Théorèmes de Hamming

- Avec un code de distance de Hamming d on peut
  - détecter p erreurs si $d \geq p+1$
  - corriger q erreurs si $d \geq 2q+1$
  - corriger q et détecter p erreurs si $d \geq q+p+1$

d	Puissance
1	erreur non détectable
2	détection d'une erreur simple
3 	correction d'une erreur simple
4 	correction d'une erreur simple et détection d'une erreur double
5 	correction d'une erreur double

Typologie des codes correcteurs d'erreurs

- Codes à contrôle de parité
  - Application d'un espace à 2^k éléments dans un espace à 2^n éléments
  - Dans ces codes un ou plusieurs bits de redondance sont ajoutés à une ensemble de k bits d'information
  - notation : code (n,k)
- Codes de blocs
  - longueur de bloc (ou de code) n dont k bits d'information
  - codes linéaires
    - Les bits de redondance sont placés sur des sites particuliers du bloc de taille n
    - ex : Code de Hamming
      - correction d'erreur simple
      - n-k bits de "parité"
      - $n \leq 2^{n-k} - 1$
      - code optimal si $n = 2^{n-k} - 1$

Typologie des codes correcteurs d'erreurs (suite)

- codes cycliques
  - En général ("codage par division") les bits de redondance sont placés après les k bits d'information
  - ex : Code de Hamming
    Code B.C.H. (Bose-Chaudhuri-Hocquenghem)

- Codes d'arbre ou codes continus
  - codes convolutionnels
    - des bits de redondance sont placés au fil de l'eau, régulièrement, dans la séquence de bits d'information et portent sur les m bits précédents (qui peuvent être des bits de redondance ...)

Utilisation d'une notation matricielle

- Codes systématiques
  - Matrice génératrice G k lignes , n colonnes
  - Un mot code est obtenu à partir d'une séquence d'information par $\vec{x} = G\,\vec{u}$
  - si $\vec{u}$ = 00.. 010..0 (1 en ième position) $\vec{x}$ est la ième ligne de G
  - Toute combinaison linéaire de lignes de cette matrice est un mot du code

$$G = \begin{bmatrix} 1 & 0 & . & .0 & a_{1,k+1} & . & a_{1,n} \\ 0 & 1 & 0 & .0 & a_{2,k+1} & . & a_{2,n} \\ 0 & . & . & .0 & a_{.,k+1} & . & a_{.,n} \\ 0 & 0 & 0 & .1 & a_{k,k+1} & . & a_{k,n} \end{bmatrix}$$

$$H = \begin{bmatrix} a_{1,k+1} & a_{1,k+2} & . & a_{1,n} \\ a_{2,k+1} & . & . & a_{2,n} \\ a_{.,k+1} & . & . & a_{.,n} \\ a_{k,k+1} & a_{k,k+2} & . & a_{k,n} \end{bmatrix}$$

- matrice de parité: k lignes n-k colonnes
- Permet de contrôler la réception par $\vec{x}$ $H = 0$ si aucune erreur

Codes cycliques: Bases

- Ensemble [0,1] muni
  - de l'addition (+ est le "ou exclusif")
  - de la multiplication

forme le corps de Galois GF(2)

- si $a_i \in GF(2)$ ($a_i = 0$ ou1)

$$a_n x^n + a_{n-1} x^{n-1} + \quad + a_0 = P(x)$$

  - P(x) est un polynôme de degré n sur GF(2)
  - L'ensemble de ces polynômes à une structure d'anneau
    - pas un groupe pour la multiplication

Division de polynômes sur GF(2); Classes résiduelles

- Soit 2 polynômes $P_1(x)$ et $P_2(x)$ sur GF(2)

  \$ deux polynômes $Q(x)$ et $R(x)$ tels que $P_1(x) = Q(x)\ P_2(x) + R(x)$

- 2 polynômes son équivalents modulo $Q(x)$ si les restes de leur division par $Q(x)$ sont identiques.

- Ceci permet de définir une classe d'équivalence. Ces polynômes peuvent être représenté par leur classe résiduelle de plus bas degré, soit ce reste $R(x)$

- exemple : $Q(x) = x^2$ Å 1

  classes résiduelles :

  0,1,x,x+1

- Pour un polynôme de degré N, il y a 2^N classes résiduelles

Irréductibilité

- Un polynôme P(x) est réductible sur un corps s'il existe deux polynômes G(x) et H(x) tels que

$$P(x) = G(x)\,H(x)$$

- Si ce n'est pas le cas P(x) est irréductible

- Si P(x) est irréductible les racines de P(x) = 0 sont des éléments d'une extension de GF(2)

- soit $P(x) = \prod_i (x + \alpha_i)$

Codes cycliques: Définitions

- Classe particulière des codes de groupe obtenu par permutation circulaire des chiffres
  - si $a_1a_2 \ldots a_n$ C alors $a_2a_3 \ldots a_na_1$ C, $a_3a_4 \ldots a_2$ Î C, etc
  - On utilise en général la notation polynomiale définie ci-dessus
  - si U(x) représente un mot du code, $x^iU(x)$ modulo $(x^n + 1)$ est aussi un mot du code
    - Ceci correspond à la ième permutation de U(x)
  - Toute combinaison linéaire de mots du code est encore un mot du code
  - Il existe un polynôme de plus bas degré qui divise $x^n + 1$
  - Ce polynôme $U_0(x)$, de degré m, est appelé POLYNOME GENERATEUR du code C
  - Il existe 2^{n-m} polynômes quotients de $x^n + 1$ par $U_0(x)$
    - Ce code contient donc k = n-m mots
    - code (n, n-m)

Matrice Génératrice et Matrice de Parité d'un Code cyclique

- MATRICE GENERATRICE
  - U_0 et ses n-m-1 permutés forment une base du code, notée par la matrice G

$$G = \begin{bmatrix} a_m & a_{m-1} & . & . & a_0 & 0 & . & 0 \\ 0 & a_m & a_{m-1} & . & . & a_0 & . & 0 \\ 0 & . & . & . & 0 & . & . & 0 \\ 0 & 0 & 0 & a_m & . & . & . & a_0 \end{bmatrix}$$

Exemple : Code (7,4) engendré par $X^3 + x + 1$

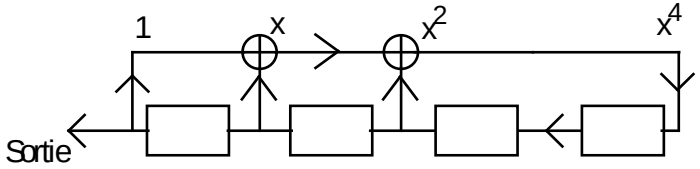
$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

- MATRICE de PARITE

$$H = \begin{bmatrix} \overline{a}_0 & 0 & . & 0 \\ \overline{a}_1 & \overline{a}_0 & . & 0 \\ . & . & . & . \\ 0 & . & . & \overline{a}_{n-m} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

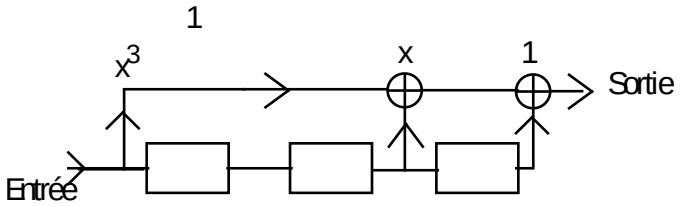
Codage par matériel

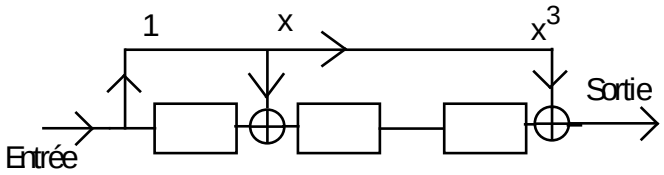


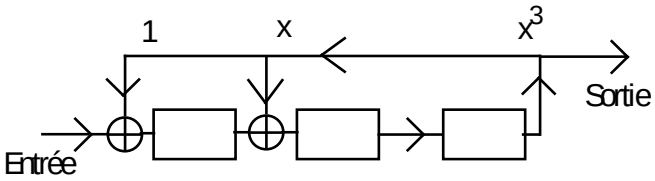
- On utilise des registres à décalage et des "ou exclusif" (addition)
- La multiplication
  - par 1 est notée par une connexion
  - par 0 par l'absence de connexion

- CODEUR A k = n-m ETAGES (ENTREE PARALLELE)
  - Exemple : codeur (7,4)
  - Polynôme orthogonal à $x^3 + x + 1$ soit $x^4 + x^2 + x + 1$ (voir matrice de parité)
- CODEUR A m ETAGES (ENTREE SERIE)
  - Le registre est initialisé à 0. L'information est introduite poids forts en tête

Codage par matériel (suite)





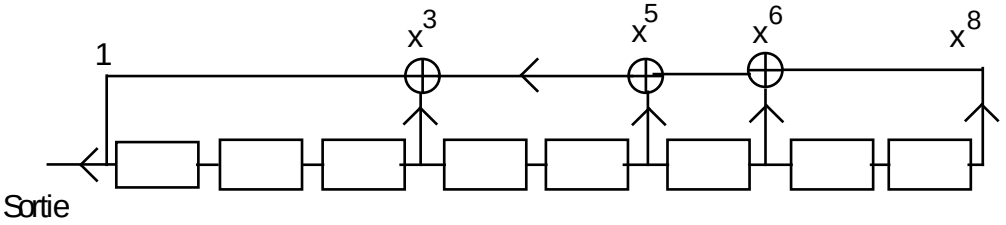


- Codeurs par multiplication.
  - Exemples: Codeurs (7,4)

- Codeur par division
  - Les m premiers décalages donnent 0 en sortie et ne sont pas utilisés.
  - Exemple: Codeur (7,4)

Application: Générateur Pseudo-aléatoire

- Une séquence pseudo-aléatoire de taille $2^{m-1} - 1$ est générée par un polynôme de degré m : G(x) diviseur de $x^{2^m - 1}$
  - Exemple : m = 8, n = 255
  - $G(x) = x^8 + x^6 + x^5 + x^3 + 1$



- La séquence de 255 bits comporte 128 bits à 1 et 127 bit à 0 répartis aléatoirement
- Si le circuit génère des bits en continu, la séquence est répétée périodiquement
- Technique utilisée pour
  - les générateurs de bruit
  - les séquences de test

Détection et Correction d'erreurs par code cyclique

- Soit un polynôme irréductible G(x) de degré m, diviseur ou produit de diviseurs de $x^{2^m - 1} + 1$
- et une séquence à protéger M(x) et E(x) un syndrome d'erreurs
  - sans erreur E(x) est une séquence nulle 000000....

- CODAGE PAR MULTIPLICATION
  - C(x) = M(x) • G(x) est le mot du code transmis
  - On reçoit $C^*(x) = C(x) + E(x)$
  - On décode $Q(x) = C^*(x) / G(x) + R(x)$
    - si R(x) = 0 M(x) = Q(x)
    - sinon erreur détectée
  - Correction
    - On peut initialiser une table donnant E(x) en fonction de R(x)
    - On peut alors corriger les erreurs par $C(x) = C^*(x) + E(x)$

Détection et Correction d'erreurs par code cyclique (suite)

- CODAGE PAR DIVISION
  - On calcule
    - $I(x) = x^m \bullet M(x)$
      - ajout en fin de séquence de m bits à 0
    - $r(x) = I(x) - G(x)\ Q(x)$
      - reste de la division de I(x) par G(x)
    - $C(x) = I(x) + r(x)$
      - remplacer les m bits à 0 de fin par r(x)
  - On décode
    - $Q(x) = C^*(x) / G(x) + R(x)$
    - Si $R(x) = 0$ $C(x) = C^*(x)$
      - $M(x) = C^*(x) / x^m$ troncature de C(x)

      Sinon erreur détectée
  - ou correction par $C(x) = C^*(x) + E(x)$

    et utilisation d'une table de correction E (R(x))

Tables de Peterson

- Considérons le polynôme $x^{n-1} + 1$
  - Ses racines $_i$ sont solutions de $x^{n-1} + 1 = 0$
  - Alors $x^{n-1} + 1 = \prod_{i=1}^{n-1}(x - \alpha_i)$
- On peut montrer que $x^{n-1} + 1$ peut être décomposé en un produit de polynômes irréductibles $f_j(x)$
- $$x^{n-1} + 1 = \prod_{j=1}^{L} f_j(x)$$
  - $f_j(x)$ a pour racines des racines $_i$ de $x^{n-1} + 1$ soit $f_j(x) = \prod_i (x - \alpha_i)$
- Soit P(x) un polynôme de degré n -1
- Si α est racine de p(x) alors $\alpha^2, \alpha^4, \alpha^8, \ldots,$ (2^k modulo n-1) sont racines de p(x)
- Les polynômes $f_j(x)$ sont difficiles à calculer . Plutôt que de donner des règles, fonctions du degré du polynôme permettant de les déterminer, Peterson en 1961 a fourni des tables qui les donnent tous jusqu'au degré 16 en partiellement jusqu'au degré 34.

Tables de Peterson: Exemple

- On lit dans la table
  - degré 5 1 45E 3 75G 5 67H
  - Ceci indique 3 polynômes $f_j(x)$ correspondant aux racines de base α^1 , α^3 et α^5
  - La lettre donne les propriétés du polynôme $f_j(x)$
    - A,B,C,D non primitif
    - E,F,G,H primitif
    - A,B,E,F racines dépendantes
    - C,D,G,H racines indépendantes
    - A,C,E,G racines du réciproque dépendantes
    - B,D,F,H racines du réciproque indépendantes
  - La valeur donne le polynôme codé en octal (base 8)

Exemple d'utilisation des tables de Peterson

- On décompose $x^{31} + 1$
  - 31 est de la forme $2^5 - 1$
  - Il sera décomposé en 6 polynômes de degré 5 et le polynôme (x+1)
- On recherche ces polynômes dans la table de Peterson
  - Les degrés des racines sont modulo 31
    - par exemple $\alpha^{48} = \alpha^{17}$
  - 45 100101 X^5+x^2+1 racines $\alpha, \alpha^2, \alpha^4, \alpha^8, \alpha^{16}$
  - 75 111101 $X^5+x^4+x^3+x^2+1$ racines $\alpha^3, \alpha^6, \alpha^{12}, \alpha^{24}, \alpha^{17}$
  - 67 110111 $X^5+x^4+x^2+x+1$ racines $\alpha^5, \alpha^{10}, \alpha^{20}, \alpha^9, \alpha^{18}$
- Les polynômes réciproques sont obtenus en prenant les bits dans l'ordre inverse
  - soit 100101 101001 = 51 racines $\alpha^{30}, \alpha^{29}, \alpha^{27}, \alpha^{23}, \alpha^{15}$
  - 111101 101111 = 57 racines $\alpha^{28}, \alpha^{25}, \alpha^{19}, \alpha^7, \alpha^{14}$
  - 110111 111011 = 73 racines $\alpha^{26}, \alpha^{21}, \alpha^{11}, \alpha^{22}, \alpha^{13}$
- Ils correspondent à la racine α^{30}, α^{28} et α^{26}. Les puissances sont complémentaires de 31
- On peut vérifier que $x^{31} + 1 = (x+1)(X^5 + x^2 + 1)(X^5 + x^4 + x^3 + x^2 + 1)$
  $(X^5 + x^4 + x^2 + x + 1)(X^5 + x^3 + 1)(X^5 + x^3 + x^2 + x + 1)(X^5 + x^4 + x^3 + x + 1)$

Codes de Hamming et Codes B.C.H.

- CODES DE HAMMING
  - Codes de distance 3, correcteurs d'une erreur
  - Ces codes sont optimaux
  - Construits à partir d'un seul polynôme irréductible G(x) de degré m
    - génèrent des mots de longueur $n = 2^m - 1$
    - exemple pour m = 5 , n = 31 code (31,26)
  - G(x) est un polynôme irréductible (normal ou réciproque) quelconque pris dans la table correspondante de Peterson
    - par exemple $X^5 + x^2 + 1$ ou $X^5 + x^3 + x^2 + x + 1$

- CODES B.C.H.
  - Hocquenghem et Bose et Chaudhuri ont donné une condition suffisante pour qu'un code ai une distance de Hamming donnée :
  - Leur polynôme générateur doit être un produit de polynômes irréductibles ayant au moins $s = d - 1$ racines a^i consécutives

Codes de Hamming et Codes B.C.H. (suite)

- Code BCH: Exemple
  - $(X^5 + x^2 + 1)(x^5 + x^4 + x^3 + x^2 + 1)$
  - a 4 racines consécutives α , α^2, α^3, α^4
- s = 4 , d = 5 . Il est correcteur de 2 erreurs (d = 2q +1)
- taille n = 31, m =10 code (31,21)

- CODES TRONQUES
  - On peut n'avoir que l < n bits à protéger (par exemple l = 16)
  - On allonge fictivement le mot par n - l zéros et on utilise de code de Hamming ou B.C.H. correspondant
    - (31,26) $\Rightarrow$ (21,16) Hamming
  - Mémoires protégées (22,16) avec $(x+1)(X^5 + x^2 + 1)$
    - (31,21) $\Rightarrow$ (26,16) B.C.H.
      $\Rightarrow$ (27,16) en ajoutant (x+1) à G(x)

Codes convolutionnels

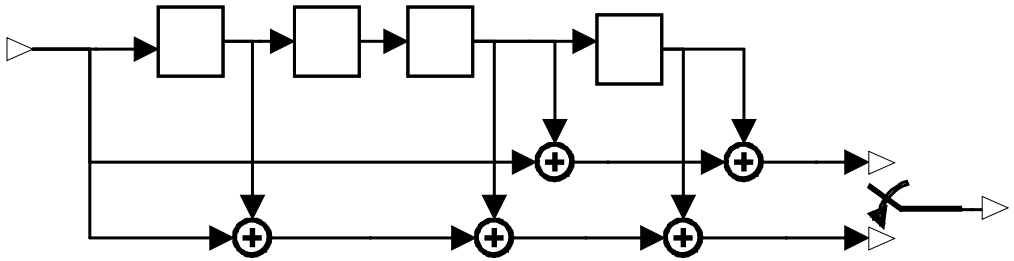
- Codes continus, au fil de l'eau…
  - Ajout de bits de parité en fonctions des X bits précédents
  - taux de code: k/n
    - k bits d 'information
    - n bits au total
    - en pratique k/n = 1/2 ou 4/5 (Hagelbarger)
- Codes correcteurs d'erreurs
  - Peuvent être très performants en théorie mais pas de codes connus avec un très bon rendement
  - nécessité d'une séquence assez longue au décodage pour retrouver l'information
    - si blocs, traînée nécessaire pour que tous les bits utiles soient protégés
  - Distance libre= distance de Hamming minimale
    - d_{free}= capacité de correction
    - si taux = 1/2 , d_{free} = 7 correction de 3 erreurs

Codes convolutionnels: codage

- Pour un code 1/2:
  - Entrée U_i, sortie $C_i = (C_i^{'}, C_i^{''})$
  - Polynômes $G'(X)$ et $G''(X)$

$$C'(X) = U(X)G'(X)$$
$$C''(X) = U(X)G''(X)$$

  - Décodage $S_i = C_i^{'} + C_i^{''}$



Utilisation dans le GSM

- Codes cycliques

nature	Canal logique	Taille k	Taille CRC reXonXance	polynômes
Parole Classe I.a	TCH/FS	50	3	X^3+X+1
Signalisation Contrôle	SACCH, FACCH, BCCH, PCH, …	184	40	$(X^{23}+1)(X^{17}+X^3+1)$
Accès	RACH	8	6	$X^6+ X^5+ X^3+ X^2 +X+1$
Synchronis.	SCH	25	10	$(X^5+1)(X^7+1)/(X+1)^2$

- Codes convolutionnels
  - code principal utilisé pour les données: taux 1/2
  - polynômes : $X^4+ X^3+ 1$ et $X^4+ X^3+X+1$

Introduction à la Cryptographie

Gérard Beuchot

beuchot@if.insa-lyon.fr

http://ictt.insa-lyon.fr/beuchot/

Quelques définitions ……

- Cryptographie
  - L'art d'écrire en caractères secrets...
- Chiffre
  - Méthode d'écriture secrète qui remplace un message clair par un message chiffré (cryptogramme)
- Chiffrement ou chiffrage
  - action de rédiger un texte en chiffre (régulier ou irrégulier)
- Chiffrer
  - transformer un langage clair en langage chiffré
- Déchiffrer
  - reconvertir en clair un message chiffré, en utilisant la clé que l 'on possède de droit.
- Décrypter
  - déchiffrer un cryptogramme alors qu'on en possède pas la clé ou après avoir reconstitué celle-ci

Quelques autres définitions ……

- Clair
  - message avant qu'il ai été chiffré ou codé
- Clé
  - mot, locution, phrase, nombre utilisé pour chiffrer ou déchiffrer un message
- Code
  - système cryptographique selon lequel des groupes de lettres sont substitués à des éléments d'un message clair
  - Moyen secret de communication autre que le chiffre, grâce auquel deux personnes peuvent échanger secrètement des informations
- Transposition
  - chiffre dans lequel chaque lettre du clair est reprise dans le cryptogramme mais placée à un autre emplacement
- Substitution
  - chiffre dans lequel chaque lettre ou groupe de lettres du clair est remplacée par une autre lettre, groupe de lettres, figure, symbole ….
- Cryptanalyse
  - Science de décrypter les messages secrets par analyse et déductions

Quelques cryptographes anciens….

- Jules César (Imperator ….)
  - inventeur d'un code à substitution élémentaire qu'il abandonna quand il ne fit plus confiance à Cicéron…
  - Tyro : scribe de Cicéron qui inventa une écriture secrète proche de la sténographie …
- Lysandre de Sparte 405 av. JC
  - Inventeur (ou utilisateur) du « scytale » qui fournissait une sorte de transposition
- Abbé Trithème 1499
  - auteur de « Polygraphia » premier traité de cryptographie
  - inventeur d 'un système chiffré : les Ave Maria (14 alphabets où une lettre est remplacée par un mot)
- Giovani Batista Belaso Della Porta 1563
  - porte le titre de père de la cryptographie moderne
- Vigenère (Blaise de) (
  - 1523-1596 - diplomate, alchimiste, kabbaliste, écrivain, historien né à St Pourçain sur Sioule …)
  - code de substitution basé sur une table de codage et un mot clé quelconque en usage jusqu'à la fin du 19ème siècle

Quelques cryptographes modernes

.....

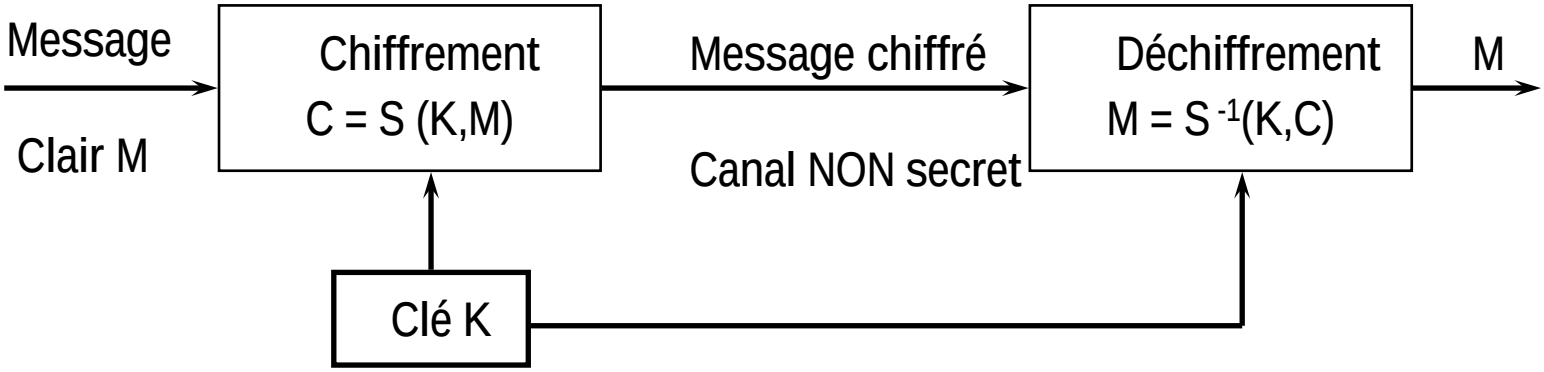
- Diffie et Hellman
- Rivest, Shamir et Adleman

Typologie des crypto-systèmes

- Méthodes de chiffrement
  - Sûr ou inconditionnellement sûr
  - Au vol ou Bloc par bloc
  - à clé PRIVEE (secrète) ou à clé PUBLIQUE
- Seul code inconditionnellement sûr : Vernam
  - Clé aléatoire utilisée une seule fois
  - $C = M \otimes K$ et $M = C \otimes K$
  - exemple
    - M = 1 0 1 1 0 0 0 1 1 1 0 1 0 1
    - K = 1 0 0 1 1 0 0 1 0 0 1 0 1 1
    - C = 0 0 1 0 1 0 0 0 1 1 1 1 1 0
    - K = 1 0 0 1 1 0 0 1 0 0 1 0 1 1
    - M = 1 0 1 1 0 0 0 1 1 1 0 1 0 1

CRYPTO-SYSTEME à CLE PRIVEE

- Utilise la même clé SECRETE pour chiffrement et déchiffrement
- Fragilité : partage de cette clé



La fonction Crytographique S est inversible

- exemples :
  - DES : Data Encryption Standard 2^{56} # 7,2 10^{6} clés possibles
  - IDEA: International Data Encryption Algorithm - clé sur 128 bits
- Pour casser DES (trouver la clé)
  il suffit théoriquement de 18 caractères des textes clair et chiffré

Un code manuel : Vigenère

- Exemple de table (volontairement simplifiée …)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
B	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a
C	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b
D	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c
E	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d
F	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e
G	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f
H	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g
I																										
R	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q
S	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r

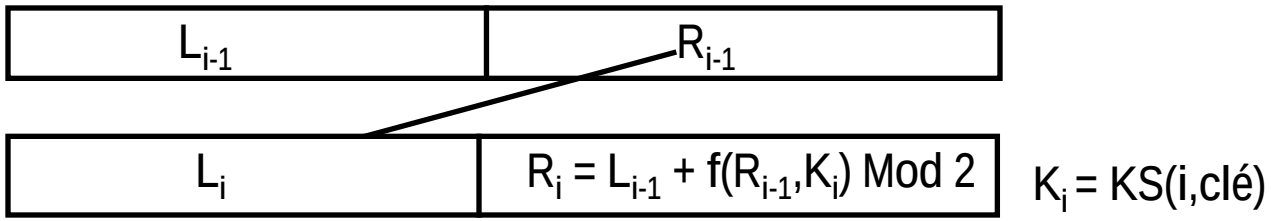
- Mot clé : CADRE
  - mot clé étendu : CADRECA DR ECAD RE CADRECADRE
- Message : exemple de code de Vigenère
- Codage
  - e est sous le C de CADRE : le code est à la ligne C et « g » est substitué à « e »
  - x est sous le A de CADRE : le code est à la ligne A et « x » est substitué à « x »
  - e est sous le D de CADRE : le code est à la ligne D et « h » est substitué à « e »
- Cryptogramme : gxhdtne gv gogv gi xijvrgrg

Crypto-systèmes à clé privé : Codes utilisés

- DES (1977)
  - Ancien standard . Voir ci-dessous
- DES-3
  - DES-EEE3 : 3 DES avec 3 clés différentes
  - DES-EDE3 : 3 Opérations en séquence (chiffrement-déchiffrement-chiffrement) avec 3 clés
  - DES-EEE2 ou DES-EDE2 : comme ci-dessus mais avec clés 1 et 3 identiques
- IDEA
  - Code de blocs proposé par Lai et Massey. Clé de 128 bits, blocs de 64 bits, 8 itérations
- RC2
  - Rivest (rfc2268). Code de bloc (64 bits) clé variable de 8 à 128 bits. Plus rapide que DES
- RC4
  - Code des flux d'octets (stream) par de permutations aléatoires (très rapide par logiciel)
- RC5
  - Code de blocs. Rapide. Clé de 8 à 2048 bits. Bloc de 32, 64 ou 128 bits. 1 à 255 itérations
- AES
  - remplaçant de DES : voir ci-dessous

Code DES

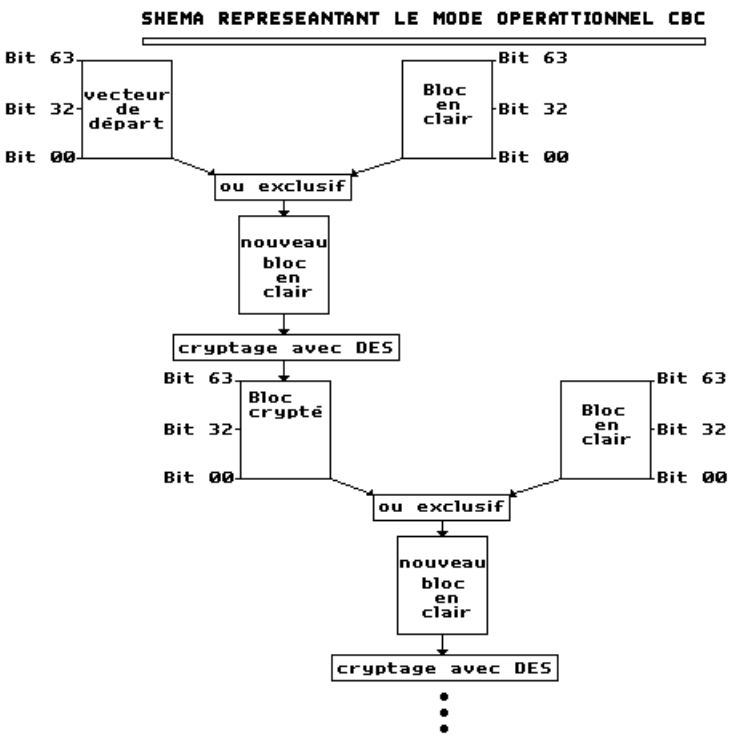
- Blocs d'information de 64 bits (8 octets)
- Clé à 56 bits (+8 de contrôle)
- Entre Transposition initiale et Transposition finale 16 itérations d'une Fonction mêlant Transposition et Substitution NON LINEAIRE (table)



R_{i-1} est étendu de 32 à 48 bits par duplication de 16 de ses bits
La clé K_i est ajoutée (bit à bit , ou exclusif)
Le champ est réduit à 32 bits grâce à une table (publique)
qui fait correspondre des champs de 4 bits aux champs de 6 bits trouvés.

- Toute la puissance du DES vient de cette table qui n'a pas de propriétés mathématiques
- Difficulté: QUALITE DE LA CLE (nombre réellement aléatoire ...)

Modes d ’utilisation du DES



- Mode EBC
  - Electronic Code Book ("catalogue électronique de codes").
  - Blocs de 64 bits indépendants les uns des autres
    - Problème si blocs identiques et même clé

Mode CBC

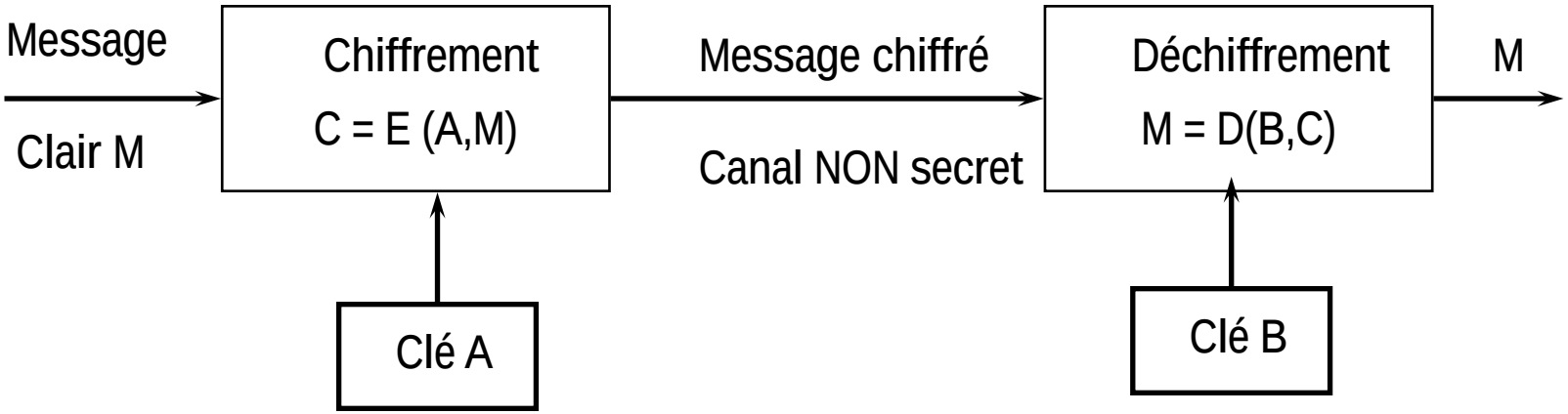
  - Chain Block Cipher ("Cryptogramme à blocs chaînés").
  - Chaque bloc clair est soumis avec bloc précédemment chiffré
    - en plus de la clé commune : vecteur initial

Code AES

- Nouveau standard en cours d'élaboration pour remplacer DES
- Cahier des charges
  - Très robuste
  - Blocs de 128 bits (autres tailles en option)
  - Clés symétriques (privées) de 128, 192 et 256 bits
  - Plus efficace et sécurisant que Triple DES
  - Élaboré et évalué publiquement - libre de droits
- 15 propositions dont DFC de ENS (Vaudanay) : 5 retenues
  - MARS d'IBM : même principe que DES - blocs de 4*32bits - clés 128 à 448 bits
    - très robuste - 8 itérations initiales et finales - S-Bloc de 512 mots de 32bits
  - RC6 de RSA : extension de RC5 - simple et rapide - clé jusqu'à 2040 bits
  - Rijndael de J. Daemen et V. Rijmen
  - Serpent de R. Anderson, E. Biham, L. Knudsen : robuste le + lent par soft
  - Twofish de B.Schneier, J.Kelsey, D.Whiting, D.Wagner, C.Hall, N.Ferguson
    - flexible - implication sur sécurité difficile à analyser

Codes à clé publique

- RSA (Rivest-Shamir-Adleman) est le plus connu …(1977)
- D'autres codes existent sur des principes voisins



  - L'algorithme E et la clé A sont publiques
  - D et B sont secrets et permettent de décoder le message chiffré C
- Basé sur la décomposition en facteurs premiers de nombres très grands nombres produits de 2 nombres premiers de plus de 100 chiffres ...

Codage RSA: Clés et exemple

Exemple:

p=31, q=47, n=1457
Φ(n)=30*46=1380

E= 889

889*1009modulo1380= 1
D=1009

B=« 11101010 »
B=234

C=234^{889}mod1457
C=892
C=« 1101111100 »

B=892^{1009}mod1457
B=234
B=« 11101010 »

- Clés
  - Chiffrement
  - Déchiffrement

pet q premiers très grands; n=pq
Φ(n)= Indicateur d'Euler de N
nombre aléatoire E, 3 <E<Φ(n)

D tel que E*D modulo Φ(n) = 1

- Chiffrement
  - Message découpé en blocs B_i
  - Codage
- Déchiffrement

$$C_i = B_i^E \text{ modulo n}$$

$$B_i = C_i^D \text{ modulo n}$$

Code Diffie-Hellman

- 1976
- Vulnérable à attaque par personne intermédiaire
- Blocs de taille n (assez grand), taille de la clé
- Code
  - B=[b_1,b_2, ,b_n] clé privée du destinataire b_i entier naturel aléatoire
  - A=[a_1,a_2, ,a_n] clé publique utilisée par la source a_i: entier naturel
    - $a_i = b_i * w$ modulo m
    - gâche $z = w^{-1}$ soiy z*w modulo m = 1
  - M=[x1,x2, ,xn] un bloc du message clair x_i : bit du message (0 ou 1)
- Chiffrement : $C=AM=a_1x_1+a_2x_2+ \ldots.+a_nx_n$
- Déchiffrement :
  - on calcule x tel que Bx = C*z modulo m = AM*z modulo m
  - puis algorithme d 'empilement pour i = n jusqu 'à 1
- Exemple
  - w=889 m=1457 B=[3,7,12,23,47,95,189,377] z=1398 tel que 889*1398 modulo 1457=1
  - $a_i=889b_i$ modulo1457 soit A=[1210,395,469,49,987,1406,466,43]

$$b_i > \sum_{j=0}^{i-1} b_j$$

$$\text{si } b_i > C - \sum_{j=i+1}^{n} x_j b_j \text{ alors } x_i = 0 \text{ sinon } x_i = 1$$

Code Diffie-Hellman : exemple de chiffrement et déchiffrement

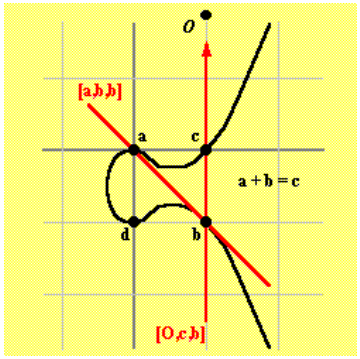
- M = 01001101
- Chiffrement
  - C=395+987+1406+43 =2831
- Déchiffrement
  - Bx=2831*1398 modulo 1457 = 526

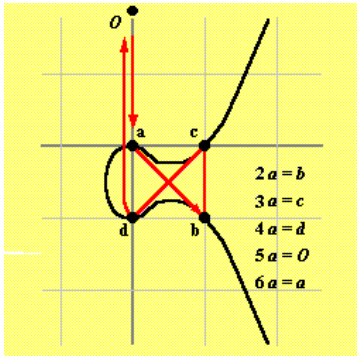
$$\text{soit } x_i = 0 \text{ si } b_i > Bx - \sum_{j=i+1}^{n} x_j b_j \text{ sinon } x_i = 1$$

  - 377 < 526 — $x_8 = 1$
  - 189 > 526-377 = 149 — $x_7 = 0$
  - 95 < 149 — $x_6 = 1$
  - 47 < 149 - 95 = 56 — $x_5 = 1$
  - 23 > 56 - 47 = 7 — $x_4 = 0$
  - 12 > 7 — $x_3 = 0$
  - 7 ≤ 7 — $x_2 = 1$
  - 3 > 7 -7 = 0 — $x_1 = 0$

Crypto-systèmes basés sur les courbes elliptiques

exemple :
$y^2 + y = x^3 - x^2$





- Plus rapide et clés plus courtes que RSA
  - si la fonction elliptique est bien choisie ……
  - donc plus facile à implanter sur carte à puce
- Fonction du type $y^2 + K\,y = x^3 + A\,x^2 + B$ (modulo p)
  - p premier très grand
- On s'intéresse aux points de coordonnées entières
  - certains de ces points forment un groupe
  - dans espace de p lignes et p colonnes cette courbe contient N points qui forment un groupe elliptique convenable où N est presque égal à p avec N=k*q où k est petit et q premier
  - exemple : $p = 2^{192} - 2^{64} - 1$
  - p = 6 277 101 735 386 680 763 835 789 423 207 666 416 083 908 700 390 324 961 279
  - N = 6 277 101 735 386 680 763 835 789 423 337 720 473 986 773 608 255 189 015 329
- le code est bâti à partir de couples de ces points ?

Aspects légaux

- En cours d'harmonisation au niveau européen
- pas encore tout à fait libre en France mais en voie de libéralisation complète
- loi 96-659 du 26/7/1996 et décret du 17/3/1999 + arrêté du 17/3/1999
- 2 cas
  - usage ou importation de services ou moyens cryptographiques
    - Voir ci-dessous
  - Fourniture de services et moyens cryptographiques
    - système d 'enregistrement
      - création, import (hors CE), export de fonctions d 'authentification ou de confidentialité à clé courte
    - système d 'autorisation préalable
      - tous les autres cas
      - sauf développement, test, démonstration : prévenir SCSSI 2 semaines à l 'avance

Aspects légaux (suite)

Usage ou importation de services ou moyens cryptographiques

- pour authentification : totalement libre

- pour confidentialité, libre si :
  - clé de moins de 40 bits (décret du 24/2/1998….)
  - clé de 40 à 128 bits pour usage individuel
  - clé de 40 à 128 bits pour usage collectif si enregistrée auprès d ’un Tiers de Confiance
  - le fournisseur a une autorisation générale
    - par exemple groupement bancaire ….

- SCSSI : service central de la sécurité des systèmes d’information
  - 18 rue du Dr Zamenhof 92131 Issy les Moulineaux
  - tel +33 1 4146 3700 Fax: +33 1 4146 3701
  - http://www.legifrance.gouv.fr/citoyen/officiels.ow

Attaques : selon informations disponibles…..

- Attaque : Rechercher le texte clair ou la clé
- Base de départ
  - Texte chiffré seulement
  - Texte clair connu (avec texte chiffré)
  - Texte clair défini
    - chiffrer et comparer à texte chiffrer
  - Texte chiffré défini
    - déchiffrer et comparer à texte clair
  - Clé choisie
    - procéder à de modifications de clés ou des comparaisons entre clés
  - Temps
    - mesure du temps de chiffrement pour avoir informations sur clé ou données
  - Analyse des défauts
    - défauts supposés du système de chiffrement
  - Man-in-the-Middle
    - s'introduire au centre du système par exemple système d 'authentification

Stratégies d'attaques

- Force brute ou « cassage »
  - cassage si résultat demande moins d'effort que la force brute : essai de toutes les clés
- Livre code
  - apprche classique du cassage de code: rechercher les transformations entre texte clair et texte chiffré
- Cryptoanalyse différentielle
  - corrélations statistiques
- Cryptoanalyse linéaire
  - approximation linéaire des S-Boxes (tables non linéaires) pour trouver clé
- Meet-in-the-middle
  - pour codage à deux niveaux
- plan de clés (Key schedule)
  - choisir clés qui produisent des effets connus à différentes itérations
- Date de naissance
  - paradoxe de la… : recherche de valeurs particulières

Stratégies d’attaques (suite)

- Codage formel
  - Propriétés algébriques
- Corrélation
  - Dans un chiffrement au vol, rechercher des corrélations entre séquences, des propriétés statistiques ,….
- Dictionnaire
  - Essayer les clés une par une à partir d’un dictionnaire de mots fréquemment utilisés
- Rejouer (Replay)
  - Enregistrer et sauver des messages ou des blocs chiffrés et renvoyer ces blocs quand besoin (exemple : mots de passe chiffrés)

- De nombreuses attaques essayent d ’isoler de petits composants ou des aspects qui peuvent être traités séparément

Bibliographie

- Historique
  - Histoire de la cryptologie : http://www.multimania.com/marief/
    - histoire et nombreux liens
  - Petit code des codes secrets J.Laffin Ed. Arts et voyages
- Cryptographie
  - http://www.scssi.gouv.fr/document/chiffre.html
  - http://www.cryptosoft.com/html/privacy.htm
  - http://www.rsa.com/rsalabs/faq/html/sections.html
  - http://www.reapertech.com.it/RSAEuro/RSAEuro/rsaann.html
    - téléchargement de programmes sources
- Aspects légaux
  - http://www.scssi.gouv.fr/
- Glossaires, acronymes
  - http://www.io.com/~ritter/GLOSSARY.HTM
  - http://www.cnet.com/Resources/Info/Glossary/num.html
  - http://www.csrstds.com/acro-a-d.html