

Chap. IV : La Théorie de Shannon - partie 1 : La confidentialité parfaite

Laurent Poinot

UMR 7030 - Université Paris 13 - Institut Galilée

Cours “ Sécrypt ”

En 1949, Claude Shannon publia un article intitulé “ *Communication Theory of Secrecy Systems* ” dans le *Bell Systems Technical Journal* qui eut une influence considérable sur l’étude de la cryptographie. Soixante ans plus tard, on estime que cet article a fondé les bases de la cryptographie moderne. Dans ce chapitre, on détaille plusieurs des idées de Shannon.

Sécurité calculatoire

Il y a deux approches fondamentales dans l'étude de la sécurité d'un système cryptographique : la notion de confidentialité parfaite de Shannon (que l'on étudie plus loin) et le concept de **sécurité calculatoire**.

La sécurité calculatoire mesure la quantité de calcul nécessaire pour casser un système. On dira qu'un procédé est **sûr au sens de la théorie de la complexité** si le meilleur algorithme pour le casser nécessite N opérations où N est un nombre beaucoup trop grand pour que cet algorithme soit applicable en pratique. Dans la pratique, on dit souvent qu'un système est *sûr* si la meilleure attaque connue ne peut se faire avec une quantité raisonnable de temps de calcul. Le problème de cette approche c'est qu'un cryptosystème peut être sûr pendant un moment puis ne plus l'être lorsque l'on découvre un algorithme plus efficace.

Sécurité inconditionnelle

Ceci mesure la sécurité du système sans borne sur la quantité de calcul que l'attaquant est capable de faire. Un procédé est **inconditionnellement sûr** s'il ne peut être cassé, même avec une puissance de calcul infinie.

Lorsque que l'on étudie la sécurité d'un système cryptographique, on doit également préciser le type d'attaque considéré. On verra au chapitre suivant que le chiffrement par substitution est peu sûr face à une attaque à texte chiffré connu (avec suffisamment de textes chiffrés). On étudie la sécurité inconditionnelle face à une attaque à texte chiffré connu.

La sécurité inconditionnelle ne peut évidemment pas s'envisager dans le cadre de la théorie de la complexité, puisque l'on permet une quantité infinie de calculs. Le cadre le mieux approprié est celui de la théorie des probabilités. On utilisera quelques notions élémentaires que l'on rappelle ici.

Rappels : Théorie élémentaire des probabilités

(1/3)

Soit $E = \{x_1, \dots, x_n\}$ un ensemble fini. Une **probabilité** sur E est une application

$$P : E \rightarrow [0, 1] \quad (1)$$

telle que $P(x_1) + P(x_2) + \dots + P(x_n) = 1$. Si $P(x) = 1$, on dit que x est un **événement sûr**, alors que si $P(x) = 0$, on dit que x est un **événement impossible**. Par exemple, on peut définir la **probabilité uniforme** sur E par

$$P(x) = \frac{1}{n} \quad (2)$$

quel que soit $x \in E$.

Rappels : Théorie élémentaire des probabilités (2/3)

Soient E et F deux ensembles finis (avec la possibilité $E = F$) munis des probabilités P_E et P_F respectivement. La **probabilité mutuelle** $P(x, y)$ est la probabilité que x et y soient réalisés simultanément. La **probabilité conditionnelle** $P(x|y)$ représente la probabilité de x sachant que y est réalisé. Les probabilités P_E et P_F sont **indépendantes** si $P(x, y) = P_E(x)P_F(y)$ pour tous x, y possibles. La probabilité mutuelle est liée à la probabilité conditionnelle par les formules

$$P(x, y) = P(x|y)P_F(y) = P(y|x)P_E(x) . \quad (3)$$

Rappels : Théorie élémentaire des probabilités

(3/3)

À partir des deux précédentes formules on obtient immédiatement le **théorème de Bayes** :

Si $P_F(y) > 0$, on a

$$P(x|y) = \frac{P_E(x)P(y|x)}{P_F(y)} . \quad (4)$$

À partir de maintenant, on suppose qu'une clef donnée est utilisée une et une seule fois dans le chiffrement. Supposons que le texte clair suive une probabilité particulière dans l'espace des textes clairs $\mathcal{P}$. On note $P_{\mathcal{P}}(x)$ la probabilité de tirer le texte clair x . On suppose également que la clef K a été choisie (par Alice et Bob) suivant la probabilité $P_{\mathcal{K}}$ de l'espace $\mathcal{K}$ des clefs. En rappelant que la clef est choisie avant de savoir quel message Alice doit transmettre, on peut raisonnablement supposer que la clef K et le texte clair x sont indépendants.

Les deux probabilités sur $\mathcal{P}$ et $\mathcal{K}$ induisent une probabilité sur l'espace des textes chiffrés $\mathcal{C}$. On peut facilement calculer la probabilité $P_{\mathcal{C}}(y)$ que le texte chiffré y soit transmis. Pour une clef $k \in \mathcal{K}$,

$$C(k) := \{E_k(x) : x \in \mathcal{P}\} . \quad (5)$$

Ainsi, $C(K)$ représente l'ensemble des textes chiffrés possibles avec la clef k . Pour tout $y \in \mathcal{C}$, on définit $C_y := \{k \in \mathcal{K} : y \in C(k)\}$, c'est-à-dire l'ensemble de toutes les clefs possibles qui permettent d'obtenir le chiffré y . On a alors

$$P_{\mathcal{C}}(y) = \sum_{k \in C_y} P_{\mathcal{K}}(k) P_{\mathcal{P}}(D_k(y)) . \quad (6)$$

L'événement “ le texte chiffré est y ” n'est possible que si on a une clef $k \in C_y$ et si le message clair x se chiffre (par k) en y , soit encore $E_k(x) = y$ ou $x = D_k(y)$. La probabilité d'obtenir le “ texte chiffré y ” est la somme de toutes les probabilités mutuelles $P(x, k)$. Les événements “ le message clair est x ” et “ la clef est k ” sont supposés indépendants, cette probabilité est égale à $P_{\mathcal{P}}(x)P_{\mathcal{K}}(k)$, d'où le résultat final.

On note également que pour tout $y \in \mathcal{C}$ et $x \in \mathcal{P}$, on peut calculer la probabilité conditionnelle $P(y|x)$ (c'est-à-dire la probabilité que y soit le texte chiffré en sachant que x est le texte clair) par

$$P(y|x) = \sum_{k \in D_{x,y}} P_{\mathcal{K}}(k) \quad (7)$$

où $D_{x,y} := \{k \in \mathcal{K} : x = D_k(y)\}$.

Avec tout cela, on peut maintenant calculer la probabilité conditionnelle $P(x|y)$ (c'est-à-dire la probabilité que x soit le texte clair en connaissant le chiffré y) en utilisant le théorème de Bayes :

$$P(x|y) = \frac{P_{\mathcal{P}}(x) \sum_{k \in D_{x,y}} P_{\mathcal{K}}(k)}{\sum_{k \in C_y} P_{\mathcal{K}}(k) P_{\mathcal{P}}(D_k(y))} \quad (8)$$

Exemple (1/5)

Voici un petit exemple pour illustrer ce calcul.

Soit $\mathcal{P} = \{a, b\}$ avec $\mathbf{P}_{\mathcal{P}}(a) = \frac{1}{4}$ et $\mathbf{P}_{\mathcal{P}}(b) = \frac{3}{4}$. Soit

$\mathcal{K} = \{K_1, K_2, K_3\}$ avec $\mathbf{P}_{\mathcal{K}}(K_1) = \frac{1}{2}$ et $\mathbf{P}_{\mathcal{K}}(K_i) = \frac{1}{4}$ pour $i = 2, 3$.

Soit $\mathcal{C} = \{1, 2, 3, 4\}$ et supposons que les règles de chiffrement soient définies par $E_{K_1}(a) = 1, E_{K_1}(b) = 2; E_{K_2}(a) = 2, E_{K_2}(b) = 3$ et $E_{K_3}(a) = 3, E_{K_3}(b) = 4$. Avec ces données on peut calculer la probabilité $\mathbf{P}_{\mathcal{C}}$.

Calculons $\mathbf{P}_{\mathcal{C}}(1)$. On a alors

$$\mathcal{C}_1 = \{k \in \mathcal{K} : \exists x \in \mathcal{P}, E_k(x) = 1\} = \{K_1\}.$$

Il en résulte que

$$\mathbf{P}_{\mathcal{C}}(1) = \mathbf{P}_{\mathcal{K}}(K_1)\mathbf{P}_{\mathcal{P}}(D_{K_1}(1)) = \mathbf{P}_{\mathcal{K}}(K_1)\mathbf{P}_{\mathcal{P}}(a) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8}.$$

Exemple (2/5)

Calculons $P_C(2)$. On a $C_2 = \{K_1, K_2\}$.

Il en résulte que

$$\begin{aligned} P_C(2) &= P_K(K_1)P_{\mathcal{P}}(D_{K_1}(2)) + P_K(K_2)P_{\mathcal{P}}(D_{K_2}(2)) = \\ P_K(K_1)P_{\mathcal{P}}(b) + P_K(K_2)P_{\mathcal{P}}(a) &= \frac{1}{2} \times \frac{3}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{3}{8} + \frac{1}{16} = \frac{7}{16}. \end{aligned}$$

Exemple (3/5)

En raisonnant de façon identique, on trouve

$$P_c(3) = \frac{3}{16} + \frac{1}{16} = \frac{1}{4} \text{ et } P_c(4) = \frac{3}{16}.$$

Exemple (4/5)

On peut également calculer la probabilité conditionnelle du texte clair en connaissant le texte chiffré.

Regardons le cas $x = a$ et $y = 1$. On a

$D_{a,1} = \{k \in \mathcal{K} : a = D_k(1)\} = \{K_1\}$. Il nous suffit d'appliquer l'une des formules vues précédemment (car on connaît P_C).

On obtient donc par le calcul, $P(a|1) = 1$.

Exemple (5/5)

De façon identique on obtient

- $P(a|2) = \frac{1}{7}$;
- $P(a|3) = \frac{1}{4}$;
- $P(a|4) = 0$;
- $P(b|1) = 0$;
- $P(b|2) = \frac{6}{7}$;
- $P(b|3) = \frac{3}{4}$;
- $P(b|4) = 1$.

Confidentialité parfaite : approche intuitive

On peut maintenant définir la notion de **confidentialité parfaite**. Intuitivement, il y a confidentialité parfaite si un adversaire n'obtient aucune information sur le texte clair en observant seulement le texte chiffré. Cette idée se formalise en termes mathématiques à l'aide des probabilités.

Confidentialité parfaite : définition mathématique

Définition

Un système cryptographique assure une **confidentialité parfaite** si l'on a $P(x|y) = P_{\mathcal{P}}(x)$ pour tout texte clair x et tout texte chiffré y , c'est-à-dire si la probabilité que le texte clair soit x sachant que le chiffré est y est égale à la probabilité que le texte clair soit x .

Dans l'exemple précédent, la confidentialité parfaite n'est assurée que si le texte chiffré est 3.

Chiffrement par décalage et confidentialité parfaite (1/5)

Si les vingt-six lettres du chiffrement par décalage sont utilisées suivant la probabilité uniforme, c'est-à-dire avec la même probabilité $\frac{1}{26}$, alors quelle que soit la probabilité des textes clairs, on a une confidentialité parfaite.

Chiffrement par décalage et confidentialité parfaite (2/5)

Démontrons ce résultat. On rappelle que

$\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_{26} = \{0, 1, \dots, 25\}$, $E_K(x) = x + K \pmod{26}$, et $D_K(y) = y - K \pmod{26}$.

Calculons tout d'abord la probabilité $P_{\mathcal{C}}$. Soit $y \in \mathcal{C}$. On a

$$\begin{aligned} P_{\mathcal{C}}(y) &= \sum_{K \in \mathbb{Z}_{26}} P_{\mathcal{K}}(K) P_{\mathcal{P}}(D_K(y)) \\ &= \sum_{K \in \mathbb{Z}_{26}} \frac{1}{26} P_{\mathcal{P}}(y - K) \\ &= \frac{1}{26} \sum_{K \in \mathbb{Z}_{26}} P_{\mathcal{P}}(y - K) . \end{aligned} \tag{9}$$

Chiffrement par décalage et confidentialité parfaite (3/5)

Lorsque y parcourt $\mathbb{Z}_{26}$, la valeur $y - K$ également (l'application $y \mapsto y - K$ est une bijection). Comme $P_{\mathcal{P}}$ est une probabilité, on a

$$\sum_{K \in \mathbb{Z}_{26}} P_{\mathcal{P}}(y - K) = \sum_{K \in \mathbb{Z}_{26}} P_{\mathcal{P}}(y) = 1. \quad (10)$$

Par conséquent,

$$P_{\mathcal{C}}(y) = \frac{1}{26} \quad (11)$$

quel que soit le message chiffré y .

Chiffrement par décalage et confidentialité parfaite (4/5)

Soit $x \in \mathcal{P}$ et $y \in \mathcal{C}$. Calculons $P(y|x)$. On a

$$P(y|x) = P_{\mathcal{K}}(y - x \pmod{26}) = \frac{1}{26} \quad (12)$$

car l'unique clef K telle que $E_K(x) = y$ est $K = y - x \pmod{26}$. En utilisant le théorème de Bayes, on obtient

$$\begin{aligned} P(x|y) &= \frac{P_{\mathcal{P}}(x)P(y|x)}{P_{\mathcal{C}}(y)} \\ &= \frac{P_{\mathcal{P}}(x)\frac{1}{26}}{\frac{1}{26}} \\ &= P_{\mathcal{P}}(x) \end{aligned} \quad (13)$$

et la confidentialité parfaite est assurée.

Chiffrement par décalage et confidentialité parfaite (5/5)

Le chiffrement par décalage est donc “ incassable ” si d’une part on se contente d’observer les messages chiffrés et si, d’autre part, chaque nouveau message clair est chiffré avec une nouvelle clef tirée aléatoirement.

Caractérisation de la confidentialité parfaite

Théorème (Shannon)

Soit un procédé de chiffrement tel que les nombres d'éléments dans $\mathcal{P}$, dans $\mathcal{C}$ et dans $\mathcal{K}$ soient tous égaux à un certain entier $n > 0$. Ce système assure une confidentialité parfaite si, et seulement si, chaque clef est utilisée avec une probabilité $\frac{1}{n}$, et pour chaque message clair $x \in \mathcal{P}$ et chaque message chiffré $y \in \mathcal{C}$, il existe une unique clef $K \in \mathcal{K}$ telle que $E_K(x) = y$ (et donc $D_K(y) = x$).

Chiffrement de Vernam (1/4)

Une réalisation célèbre de la confidentialité parfaite est le **chiffrement de Vernam**, également connu sous les noms **masque jetable** ou **one-time pad**. Il fut inventé par Gilbert Vernam en 1917 pour chiffrer et déchiffrer des messages télégraphiques. Il est intéressant de noter que le chiffrement de Vernam fut admis “incassable” pendant des années avant que Shannon ne le prouve trente ans plus tard grâce à la notion de confidentialité parfaite. Le transparent suivant contient la description de ce cryptosystème.

Chiffrement de Vernam (2/4)

Soit un entier $n \geq 1$ fixé. Soient $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_2^n$, c'est-à-dire que les messages clairs, les messages chiffrés et les clefs sont des blocs (des vecteurs) de n bits (puisque $\mathbb{Z}_2 = \{0, 1\}$). Pour $K = (K_1, K_2, \dots, K_n) \in \mathbb{Z}_2^n$ (donc K_i est un bit) et $x = (x_1, x_2, \dots, x_n)$ (x_i est un bit), on définit

$$E_K(x) = (x_1 \oplus K_1, x_2 \oplus K_2, \dots, x_n \oplus K_n) \quad (14)$$

où $\oplus$ est l'opération de **ou-exclusif** (ou **XOR**) donnée par la table suivante

$\oplus$	0	1
0	0	1
1	1	0

(15)

Le ou-exclusif n'est en fait rien d'autre que l'addition modulo deux.

Chiffrement de Vernam (3/4)

Le déchiffrement est identique au chiffrement. Si $y = (y_1, y_2, \dots, y_n)$, alors

$$D_K(y) = (y_1 \oplus K_1, y_2 \oplus K_2, \dots, y_n \oplus K_n) \quad (16)$$

car l'addition et la soustraction modulo sont **identiques** ($\oplus$).

En utilisant le théorème de Shannon, on voit facilement que le chiffrement de Vernam assure une confidentialité parfaite. Néanmoins ce procédé est vulnérable aux attaques à texte clair connu. En effet, si on connaît le texte chiffré $y = (y_1, \dots, y_n)$ et le texte clair correspondant $x = (x_1, \dots, x_n)$, alors on retrouve la clef K utilisée en calculant

$$(x_1 \oplus y_1, \dots, x_n \oplus y_n) = K. \quad (17)$$

Chiffrement de Vernam (4/4)

Ceci tient au fait à la propriété suivante du ou-exclusif : $b \oplus b = 0$ quelle que soit la valeur du bit b . Si $y = E_K(x)$, alors quel que soit $i = 1, \dots, n$, $x_i \oplus K_i = y_i$. Donc

$$x_i \oplus y_i = x_i \oplus (x_i \oplus K_i) = \underbrace{x_i \oplus x_i}_{=0} \oplus K_i = K_i. \text{ En résumé, la}$$

connaissance du i ème bit x_i de x et du i ème bit y_i de y permet de retrouver le i ème bit K_i de la clef K utilisée.