Théorie de Shannon

Introduction au Concept de Sécurité

Yann Rotella

UVSQ - Université Paris-Saclay

22 janvier 2026

Plan du Cours

Quelques rappels de probabilité

L'information

L'Entropie

- Liens Information et Entropie
- Propriétés de l'Entropie

Sécurité au sens de l'information

- Sécurité Inconditionnelle
- Le théorème de Shannon
- Le chiffrement de Vernam

Vers une sécurité pratique

- Modèles d'attaque
- Sécurité Calculatoire

RAPPELS DE PROBABILITÉ (SIMPLIFIÉ)

- Ω l'univers
- X une variable aléatoire v.a à valeurs dans $\mathcal{X}$, un ensemble (fini)
- $X : \Omega \to \mathcal{X}$
- À chaque expérience $x \in \mathcal{X}$, on associe une mesure de probabilité

$$\Pr[X = x] = p_X(x) = p(x)$$

- La loi de probabilité de X est la donnée de

$$\begin{aligned} p : \mathcal{X} &\to [0,1] \\ x &\mapsto \Pr_X[X = x] \end{aligned}$$

Exemple (jet de dé modulo 3) :

$$\mathcal{X} = \{0,1,2\}$$

$\forall x \in \mathcal{X}$,

$$\Pr[X = x] = p_X(x) = \frac{1}{3}$$

Probabilité mutuelle et conditionnelle

Soient X et Y deux v.a. à valeurs dans $\mathcal{X}$ et $\mathcal{Y}$.

Définition (Probabilité mutuelle)

$$\Pr_{X,Y}(x,y) = \Pr[X = x, Y = y].$$

C'est la probabilité que l'événement $X = x$ ***et*** *se réalise que l'événement $Y = y$ se réalise.*

Définition (Probabilité conditionnelle)

$$\Pr_{X,Y}(x|y) = \Pr[X = x|Y = y].$$

C'est la probabilité que l'événement $X = x$ ***sachant*** *que l'événement $Y = y$ est réalisé.*

ESPACE PROBABILISÉ JOINT

On considère X et Y deux variables aléatoires. Nous avons donc une loi de probabilité sur l'espace $\mathcal{X} \times \mathcal{Y}$. Celle-ci est notée

$$\Pr[X = x, Y = y] = p_{XY}(x, y)\,.$$

De là, nous définissons les **lois marginales** sur $\mathcal{X}$ et $\mathcal{Y}$:

$$\Pr[X = x] = p_X(x) = \sum_{y \in \mathcal{Y}} p_{XY}(x, y)$$

$$\Pr[Y = y] = p_Y(y) = \sum_{x \in \mathcal{X}} p_{XY}(x, y)$$

Les probabilités conditionnelles peuvent alors être exprimées comme suit :

$$\Pr[Y = y | X = x] = \frac{p_{XY}(x, y)}{p_X(x)}$$

$$\Pr[X = x | Y = y] = \frac{p_{XY}(x, y)}{p_Y(y)}$$

Variables aléatoires indépendantes

Définition (Variables indépendantes)

Les variables X et Y sont dites ***indépendantes*** *si pour tout* $(x,y) \in \mathcal{X} \times \mathcal{Y}$,

$$p_{X,Y}(x,y) = p_X(x)p_Y(y)\,.$$

Exemple (jet de dé modulo 3 et les faces 1, 2 et 3 sont jaunes, les faces 4, 5 et 6 sont orange) :

$$\mathcal{X} = \{0,1,2\}, \mathcal{Y} = \{jaune, orange\}$$

- ✍ Donner les lois marginales de X et Y
- ✍ Montrer que les variables sont indépendantes
- ✍ Modifier la couleur des dés pour qu'elles ne le soient plus.

Notion d'Information

On considère une v.a. X et une v.a. Y et on s'intéresse à la **quantité d'information** apportée par la **réalisation** d'un événement x. On cherche :

- ✍ Une fonction **décroissante** de $p(x)$.
- ✍ Une fonction **additive** si les événements $X = x$ et $Y = y$ sont indépendants.
- ✍ Que vaut l'information apportée par un événement certain ?
- ✍ Que vaut l'information apportée par un événement n'arrivant jamais ?

INFORMATION

DÉFINITION (INFORMATION PROPRE)

*Soit X une v.a. et x un événement, **l'information propre** de $X = x$ est donnée par*

$$I(x) = -\log_2(p(x))$$

DÉFINITION (INFORMATION MUTUELLE)

*L'**information mutuelle** de $X = x$ et $Y = y$ est donnée par*

$$I(x;y) = \log_2\left(\frac{p(x|y)}{p(x)}\right) = \log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

- ✍ C'est une quantité en **bit d'information**.
- ✍ $I(x;y)$ est positive négative ou nulle.

ENTROPIE

DÉFINITION (ENTROPIE)

*L'**entropie** H d'une v.a. X est définie par*

$$H(X) = \sum_{x \in \mathcal{X}} -p(x) \log_2(p(x)$$

- ✍ Donne la quantité de désordre de la v.a.
- ✍ L'entropie est toujours positive

AUTRES GRANDEURS

DÉFINITION (ENTROPIE CONDITIONNELLE)

L' entropie conditionnelle de X sachant Y est définie par

$$H(X|Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} -p(x,y) \log_2 \left(p(x|y) \right)$$

DÉFINITION (INFORMATION MUTUELLE MOYENNE)

L'information mutuelle moyenne de X et de Y est définie par

$$I(X;Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log_2 \left(\frac{p(x,y)}{p(x)p(y)} \right)$$

PROPRIÉTÉS DE L'ENTROPIE

THÉORÈME (INFORMATION MUTUELLE MOYENNE ET ENTROPIE)

Pour toute v.a. X et Y, on a

$$I(X;Y) = H(X) - H(X|Y)$$

THÉORÈME (BORNES SUR L'ENTROPIE)

Pour toute v.a. X, on a

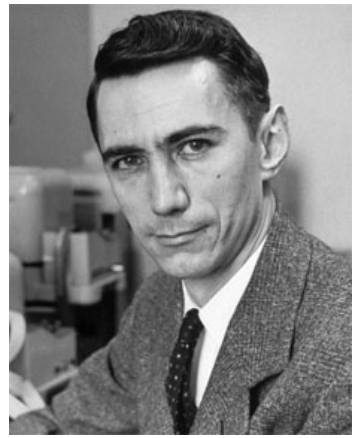
$$0 \leq H(X) \leq \log_2(|\mathcal{X}|)$$

THÉORÈME (INFORMATION MUTUELLE MOYENNE)

Pour toute v.a. X et Y, on a

$$I(X;Y) \geq 0$$

POURQUOI ON FAIT TOUT ÇA ?



« Communication Theory of Secrecy Systems », Claude Shannon, 1949

- Sécurité **Inconditionnelle** des systèmes cryptographiques.
- Diffusion (prochain cours)
- Confusion (prochain cours)

Sécurité Inconditionnelle

Pour simplifier l'étude on se place dans un contexte où l'attaquant voit des textes chiffrés (C)

- ✍ Quel événement l'attaquant peut observer ?
- ✍ Qu'est-ce qu'on ne veut pas ?
- ✍ Comment formaliser un cryptosystème inconditionnellement sûr dans ce contexte ?

FORMALISME AVEC LA THÉORIE DE L'INFORMATION

Un cryptosystème symétrique est décrit par :

- $\mathcal{M}$: l'ensemble des message (clairs) possibles, peut être noté $\mathcal{P}$ pour « plaintext »
- $\mathcal{C}$: l'ensemble des chiffrés possibles
- $\mathcal{K}$: l'ensemble des clefs possibles
- $E_k(m)$: la fonction de chiffrement qui prend en entrée une clef $k \in \mathcal{K}$ et un message $m \in \mathcal{M}$ et renvoie un chiffré $c \in \mathcal{C}$
- $D_k(c)$: la fonction de déchiffrement.

✍ Rappeler la propriété que l'on souhaite pour les fonctions de chiffrement et de déchiffrement.

Formalisme avec la théorie de l'information

Du point de vue de l'attaquant, celui-ci est capable d'observer des messages $c \in \mathcal{C}$, et cherche à obtenir de **l'information** sur le message clair m transmis (ou bien sur la clef secrète k utilisée).
Comme différents messages peuvent être envoyés, ceux-ci suivent donc une loi de probabilité $\Pr[M = m]$.
Il en va de même pour la clef $\Pr[K = k]$.
Et comme nous avons une fonction de chiffrement qui, à chaque clef et message associe un chiffré, ceux-ci peuvent être modéliser par une v.a. C.

Définition (Sécurité Inconditionnelle)

Un cryptosystème est dit inconditionnellement sûr si la connaissance d'un chiffré n'apporte ***aucune*** *information sur le message clair.*

SÉCURITÉ INCONDITIONNELLE ET ENTROPIE

DÉFINITION (SÉCURITÉ INCONDITIONNELLE)

Un cryptosystème est dit inconditionnellement sûr si la connaissance d'un chiffré n'apporte ***aucune*** *information sur le message clair.*

- Ceci se formalise mathématiquement en

$$H(M|C) = H(M)$$

- On se souvient que

$$H(M) - H(M|C) = I(M;C)$$

Indépendance statistique entre les clairs et les chiffrés

Théorème (Indépendance statistique et information mutuelle moyenne)

Soient X et Y deux v.a., alors

$$I(X;Y) \geq 0$$

avec égalité si et seulement si les v.a. sont indépendantes.

- ✍ Rappeler la formule de l'information mutuelle moyenne.
- ✍ **Indication :** $\log_2(x) \geq \log_2(e)(x-1)$ avec égalité si et seulement si x = 1.

IMPLICATIONS

Soient K, M et C les trois v.a. définies précédemment.
Soit k la clef utilisée. Les v.a. M et C sont liées par la relation $C = E_k(M)$.

- ✍ Pour que le système soit inconditionnellement sûr, il faut que M et C soient **indépendantes**
- ✍ La probabilité qu'un message clair m soit transmis sachant que l'on a observé un chiffré c est la même que la probabilité du message clair m soit transmis, sans avoir observé le chiffré c.

EXEMPLE

$\mathcal{M} = \{0,1,2\} = \mathcal{C}$, $\mathcal{K} = \{0,1\}$

$$E_k(m) = m + k \mod 3$$

- ✍ En supposant des lois uniformes pour K et M, donner la loi de C. Que peut-on dire de ce chiffrement ?
- ✍ Modifier le chiffrement pour avoir un chiffrement inconditionnellement sûr

Théorème de Shannon

Théorème (Théorème de Shannon)

Un système cryptographique tel que $|\mathcal{K}| = |\mathcal{C}| = |\mathcal{M}|$ *assure une confidentialité parfaite* **si et seulement si**

- $\Pr[K = k] = \frac{1}{|\mathcal{K}|}$, *pour tout* $k \in \mathcal{K}$
- *Pour tout* $m \in \mathcal{M}$, $c \in \mathcal{C}$, *il existe une clef unique* $k \in \mathcal{K}$ *telle que* $E_k(m) = c$.

✍ Preuve en TD.

UN PROBLÈME DE TAILLE

THÉORÈME

Si $|\mathcal{M}| > |\mathcal{K}|$ *alors aucun système n'offre une confidentialité parfaite.*

✍ Preuve en TD.

Le chiffrement de Vernam - One-Time-Pad

Définition (Chiffrement de Vernam)

Soit $n \geq 1$ et $\mathcal{M}, \mathcal{C}, \mathcal{K} = \{0,1\}^n$. Le chiffrement de Vernam est défini par la fonction de chiffrement suivante :

$$E_k(m) : ((k_0, \ldots, k_{n-1}), (m_0, \ldots, m_{n-1})) \rightarrow (m_0 \oplus k_0, \ldots, m_{n-1} \oplus k_{n-1})$$

- ✍ Quelle est la fonction de déchiffrement ?
- ✍ Quelles conditions pour que ce chiffrement soit inconditionnellement sûr ?
- ✍ Comment l'adapter si on travaille sur un plus grand alphabet

Problèmes

Pour avoir une sécurité parfaite, il faut que

- toutes les clefs soient tirées de manière uniforme
- chaque clef doit être aussi longue que le message transmis
- chaque clef n'est utilisée qu'une seule fois

✍ Donner des arguments expliquant pourquoi on ne peut pas utiliser un chiffrement parfait en pratique.

Conclusion

- On peut quantifier l'information avec une théorie mathématique
- Cela permet de définir ce que l'on cherche à construire
- Ce n'est pas pratique **du tout**

- Comment définir une sécurité **pratique** ?
- Quelles propriétés nous voulons ?
- Quels types d'attaque ?

Ce que l'attaquant peut connaître

$$E(k, m) = c$$

✍ Donner et justifier tout ce que l'attaquant peut connaître.

Modèles d'attaque

- Chiffré connu
- Clair connu
- Chiffré choisi
- Clair choisi

Définition un peu plus précise

Soit E un algorithme de chiffrement et k une clef secrète.

- Connaissant $c = E_k(m)$, puis-je retrouver m, k ?
- Connaissant $E_k(m_1)$, $E_k(m_2)$, $E_k(m_3)$, $E_k(m_i)$, puis-je retrouver un des m_i ?
- Connaissant m **et** $E_k(m) = c$, puis-je retrouver k ?
- Connaissant $E_k(m_1)$, $E_k(m_2)$, $E_k(m_3)$, $E_k(m_i)$ **et** $m_1, m_2, \ldots, m_i$, puis-je retrouver ... ?
- ✍ Prolonger la phrase suivante.
- ✍ Remplacer « connaître »par « choisir ».
- ✍ Ces contextes sont-ils pertinents en pratique ? Pourquoi ?

Sécurité Calculatoire

- Les cerveaux des attaquants
- Un ordinateur
- Plusieurs ordinateurs très performants
- Et dans 10 ans ? Dans 20 ans ? Dans 200 ans ?
- ✍ Nombre d'opérations par secondes d'un CPU ?
- Aujourd'hui : Nvidia, Bitcoin mining et GPU
- Frontier : ExaFLOPS : 10^{18} opérations par secondes
- ✍ Nombre d'atomes dans l'univers : 10^{80} : En supposant qu'un algorithme cryptographique peut coûter 100 opérations environ, donner, en puissance de deux le nombre d'opérations réalisables par la planète en une année.

La recherche exhaustive

Soient E une fonction de chiffrement et supposons que l'attaquant connaît un couple clair chiffré (m, c).

- ✍ Donner le pseudo-code correspondant à la recherche de la clef en les testant toutes une à une
- ✍ Quel est le coût **exact** de cet algorithme ?
- ✍ Pourquoi la recherche exhaustive ne fonctionne (a priori) pas si on a un chiffré seul ?
- ✍ Décrire cela en terme d'information.