Introduction à la Cryptographie

Notion de sécurité sémantique

Objectif d'un chiffrement symétrique

Objectif d'un chiffrement symétrique

Dans le TD "Chiffrements Ancestraux", nous avons vu de vieux cryptosystèmes et des notions de cryptanalyse.

Dans ce cours, nous allons voir deux notions de "sécurité" plus formelle : **Sécurité selon Shannon** et **Sécurité Parfaite**.

Objectif d'un chiffrement symétrique

Un **chiffrement symétrique** est défini par 3 algorithmes :

- un **générateur de clé** *KeyGen* qui retourne une clé "aléatoire" **k** ;
- un **algorithme de chiffrement** *Enc* ;
- un **algorithme de déchiffrement** *Dec*.

Remarque

Puisque la clé de chiffrement et la clé de déchiffrement peuvent être déduites l'une de l'autre, on peut supposer que la première opération de *Dec* consiste à calculer la clé de déchiffrement à partir de la clé de chiffrement. Ainsi, on suppose que la clé de chiffrement et de déchiffrement sont la même clé **k**.

Objectif d'un chiffrement symétrique

Notations :

- **k** : la clé secrète
- **m** : un texte clair
- **c** := Enc(**k, m**) : le chiffré associé au texte clair **m**.

On doit avoir Dec(**k, c**) = Dec(**k**, Enc(**k, m**)) = **m**
et il doit être difficile de trouver **m** à partir de **c** si on ne connait pas **k**.

Objectif d'un chiffrement symétrique

1. **Cacher la clé secrète**
   Cependant, même si la clé secrète est parfaitement cachée, cela ne garantit pas que le chiffré ne peut pas être décrypté...
   Exemple. Avec un chiffrement par substitution, si le chiffré ne contient pas toutes les lettres de l'alphabet, alors une analyse fréquentielle ne va pas retrouver toute la clé secrète mais seulement une partie.
2. **Cacher le texte clair**
   En effet, un attaquant peut vouloir décrypter des messages chiffrés particuliers sans spécialement retrouver la clé secrète.

Noter que l'objectif 2 implique nécessairement d'atteindre l'objectif 1

Objectif d'un chiffrement symétrique

Que signifie "cacher le texte clair" ?

- rendre impossible le fait de trouver le texte clair intégral
- rendre impossible le fait de trouver le texte clair partiellement
- rendre impossible le fait de trouver n'importe quelle information sur le texte clair

Exemple. Pour toute fonction f , il doit être impossible de calculer $f(\mathbf{m})$ à partir de la seule connaissance de $\mathbf{c}$. Par exemple, si $\mathbf{c}$ permet de trouver la fréquence des lettres triée du message clair $\mathbf{m}$ associé, alors la règle n'est pas respectée.

Objectif d'un chiffrement symétrique

Mais peut-on tout cacher du message clair ? On considère que la réponse est NON. En pratique, des éléments du message clair sont souvent supposé connu de l'attaquant (cf. modèle de l'attaquant).

Exemple. Dans un message en français, on pourrait supposer que le premier mot est "Bonjour" ou bien que le message contient la date ou un autre mot particulièrement fréquent. En particulier, c'est ce qui a permis aux Alliés de casser la machine Enigma durant la seconde guerre mondiale : les messages chiffrés Nazi contenaient généralement l'expression "Hi Hitler" et/ou la date du jour.

Objectif d'un chiffrement symétrique

3. **Cacher tout ce qui n'est pas déjà connu**
   En effet, on ne peut pas cacher quelque chose qui est déjà connu a priori. Mais le message chiffré doit cacher tout le reste...

De manière équivalente, un attaquant ne doit pas pouvoir apprendre quelque chose de nouveau (qu'il ne connait pas déjà) à propos du texte clair après avoir vu et analysé le texte chiffré.

Sécurité selon Shannon

Sécurité selon Shannon

On note :

- $\mathcal{M}$: l'espace des messages clairs
- $\mathcal{C}$: l'espace des messages chiffrés
- $\mathcal{K}$: l'espace des clé secrètes
- K : la variable aléatoire sur $\mathcal{K}$ correspondant à la sortie de *KeyGen*.

Sécurité selon Shannon

Soit M une variable aléatoire sur $\mathcal{M}$.

$\hookrightarrow$ M capture toutes les informations a priori connues par l'attaquant. Par exemple, si le message clair commence par "Bonjour", alors M prendra cette information en considération.

L'attaquant sait seulement que :

- **k** est tiré selon K ;
- **m** est tiré selon M ;
- Enc peut utiliser sa propre source d'aléa. Mais d'après le principe de Kerckhoffs, l'attaquant connait la distribution de la variable aléatoire $C := Enc(K, M)$ définie sur $\mathcal{C}$.

Sécurité selon Shannon

Finalement :

- Toute information sur le message clair **m** qui peut être connue **avant** l'observation du texte chiffré **c** est capturée par la variable aléatoire M.
- Toute information sur le message clair **m** qui peut être connue **après** l'observation du texte chiffré **c** est capturée par la variable aléatoire $M|C$.

$\Rightarrow$ **Sécurité selon Shannon :**

$$M \sim M|C$$

La connaissance du message chiffré n'apporte aucune information sur le message clair.

Sécurité selon Shannon

Définition : Sécurité selon Shannon

Un Chiffrement Symétrique $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, Enc, Dec)$ est **Shannon Sûr** par rapport à une variable aléatoire M sur $\mathcal{M}$ si pour tout $\mathbf{m} \in \mathcal{M}$ et tout $\mathbf{c} \in \mathcal{C}$:

$$\mathbb{P}(M = \mathbf{m}) = \mathbb{P}(M = \mathbf{m} | Enc(K, M) = \mathbf{c})$$

On dira qu'il est **Shannon Sûr** si il est **Shannon Sûr** par rapport à toutes les variables aléatoire $\mathcal{M}$.

Sécurité Parfaite

Sécurité Parfaite

Soient $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}$.

Soient C_1 et C_2 deux variables aléatoires définies comme suit :

$$C_1 := Enc(K, \mathbf{m}_1)$$

$$C_2 := Enc(K, \mathbf{m}_2)$$

$\Rightarrow$ **Sécurité Parfaite :**

$$\forall (\mathbf{m}_1, \mathbf{m}_2) \in \mathcal{M}^2, \text{ on a } C_1 \sim C_2$$

Sécurité Parfaite

Définition : Sécurité Parfaite

Un Chiffrement Symétrique $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, Enc, Dec)$ est **Parfaitement Sûr** si pour tout couple $(\mathbf{m}_1, \mathbf{m}_2) \in \mathcal{M}^2$ et tout $\mathbf{c} \in \mathcal{C}$:

$$\mathbb{P}(Enc(K, \mathbf{m}_1) = \mathbf{c}) = \mathbb{P}(Enc(K, \mathbf{m}_2) = \mathbf{c})$$

Cette notion de sécurité est beaucoup plus simple à utiliser que celle de Shannon.

Théorème d'équivalence

Théorème d'équivalence

Un Chiffrement Symétrique est Parfaitement Sûr si et seulement si il est Shannon Sûr.

One-Time-Pad

One-Time-Pad (OTP)

Définition du chiffrement OTP :

- $\mathcal{M} := \{0,1\}^n \longrightarrow$ chaine binaire de longueur n
- $\mathcal{K} := \{0,1\}^n \longrightarrow$ **la clé est aussi longue que le message à chiffrer**
- K est la distribution uniforme sur $\mathcal{K}$
- $Enc(\mathbf{k}, \mathbf{m}) := \mathbf{k} \oplus \mathbf{m}$ où $\oplus$ est le *XOR* bit à bit
- $Dec(\mathbf{k}, \mathbf{m}) := Enc(\mathbf{k}, \mathbf{m})$

One-Time-Pad (OTP)

Définition du chiffrement OTP :

- $\mathcal{M} := \{0,1\}^n \longrightarrow$ chaine binaire de longueur n
- $\mathcal{K} := \{0,1\}^n \longrightarrow$ **la clé est aussi longue que le message à chiffrer**
- K est la distribution uniforme sur $\mathcal{K}$
- $Enc(\mathbf{k}, \mathbf{m}) := \mathbf{k} \oplus \mathbf{m}$ où $\oplus$ est le *XOR* bit à bit
- $Dec(\mathbf{k}, \mathbf{m}) := Enc(\mathbf{k}, \mathbf{m})$

Théorème : *One-Time-Pad*

Le chiffrement *One-Time-Pad* est Parfaitement Sûr.

One-Time-Pad (OTP)

OTP est difficile à utiliser en pratique car :

- les entités doivent échanger de nombreuses clés car **une clé ne peut être utilisée qu'une fois** ;
- et ces **clés sont incompressibles** puisqu'elles ont été tirées uniformément.

One-Time-Pad (OTP)

OTP est difficile à utiliser en pratique car :

- les entités doivent échanger de nombreuses clés car **une clé ne peut être utilisée qu'une fois** ;
- et ces **clés sont incompressibles** puisqu'elles ont été tirées uniformément.

Remarque :

- On peut remplacer le *XOR* par une addition/soustraction modulo q si l'alphabet est $0, \ldots, q-1$. Par exemple, dans les chiffrements de Vigenère ou César, on utilise l'addition modulo 26 pour chiffrer et la soustraction modulo 26 pour déchiffrer.
- Le chiffrement de Vigenère est aussi appelé chiffrement de Vernam et est parfaitement sûr si la clé est de la même taille que le chiffré.

Théorème de Shannon

Théorème de Shannon

Théorème de Shannon

Si un Chiffrement Symétrique $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, Enc, Dec)$ est **Parfaitement Sûr**, alors nous avons nécessairement $\sharp\mathcal{K} \geq \sharp\mathcal{M}$.

Remarque. Si $\mathcal{K}$ et $\mathcal{M}$ sont respectivement $\mathcal{A}^t$ et $\mathcal{A}^n$ avec $\mathcal{A}$ un alphabet de taille (cardinal) $\sharp\mathcal{A} = q$ alors $\sharp\mathcal{K} = q^t$ et $\sharp\mathcal{M} = q^n$.

Ainsi, le théorème de Shannon implique que les clés doivent être plus longues que les messages clairs. Ce qui rend les cryptosystèmes parfaitement sûr inutilisables en pratique !

Théorème de Shannon

Théorème de Shannon

Si un Chiffrement Symétrique $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, Enc, Dec)$ est **Parfaitement Sûr**, alors nous avons nécessairement $\sharp\mathcal{K} \geq \sharp\mathcal{M}$.

Remarque. Si $\mathcal{K}$ et $\mathcal{M}$ sont respectivement $\mathcal{A}^t$ et $\mathcal{A}^n$ avec $\mathcal{A}$ un alphabet de taille (cardinal) $\sharp\mathcal{A} = q$ alors $\sharp\mathcal{K} = q^t$ et $\sharp\mathcal{M} = q^n$.

Ainsi, le théorème de Shannon implique que les clés doivent être plus longues que les messages clairs. Ce qui rend les cryptosystèmes parfaitement sûr inutilisables en pratique !

Solution : Sécurité Calculatoire