Cryptography

2. Shannon and Perfect secrecy

Contact : kevin.carrier@cyu.fr

Plan

1. What does mean secure ?
2. Goal of a symmetric cipher
3. Shannon Secrecy
4. Perfect Secrecy
5. One-Time-Pad
6. Shannon Theorem

What does mean secure ?

What does mean secure ?

In the exercise sheet “ancestral ciphers”, we have seen some old ciphers and some cryptanalysis notions.

In this lecture, we will see a more formal notion of “security”: **Shannon Secrecy** and **Perfect Secrecy**.

What does mean secure?

Recall. A symmetric cipher is made of:

- a **key generator** *KeyGen* which returns a random key **k**
- an **encryption algorithm** *Enc*
- a **decryption algorithm** *Dec*

Remark

Since the encryption key and the decryption key can be deduced from each other, we suppose that **k** is the encryption key and the first operation of *Dec* consists in computing the decryption key from **k**. Thus, *Enc* and *Dec* are both parameterized by **k**.

Goal of a symmetric cipher

Notations

Let:

- **k**: a secret key
- **m**: a plaintext
- $\mathbf{c} := \mathit{Enc}(\mathbf{k}, \mathbf{m})$: the ciphertext associated to the plaintext **m**.

We must have $\mathit{Dec}(\mathbf{k}, \mathbf{c}) = \mathit{Dec}(\mathbf{k}, \mathit{Enc}(\mathbf{k}, \mathbf{m})) = \mathbf{m}$

and it must be hard to find **m** from **c** if we do not know **k**.

Goal of a symmetric cipher

1) **Hide the secret key**

   However, even if the secret key is perfectly hidden, that doesn't guarantee the ciphertext cannot be decrypted...

   Example. With a substitution encryption, if the ciphertext doesn't contains all the letters from the alphabet, then a frequency analysis will not recover the whole secret key but only a part.

2) **Hide the plaintext**

   Indeed, an attacker may want to recover some particular plaintexts without especially recover the secret key.

Note that the goal 2) necessarily implies to achieve the goal 1)

Goal of a symmetric cipher

What does mean “hide the plaintext”?

- make it impossible to find the entire plaintext
- make it impossible to partially find the message
- make it impossible to find any information about the plaintext

Example. For all function f, it must be impossible to compute $f(\mathbf{m})$ from the mere knowledge of $\mathbf{c}$. For instance, if $\mathbf{c}$ allows to find the permuted frequency letters of the original message $\mathbf{m}$, then the rule is not met.

Goal of a symmetric cipher

But can we hide everything from the plaintext? We consider the answer is NO. In practical, some elements from the plaintext can be assumed.

Example. In an English message, we can suppose that the first word is “Hello” or that the message contains the date or another particular word. For instance, an encrypted Nazi message during the WW2 usually contained “Hi Hitler”. This allowed the Allies to break the Enigma cipher.

Goal of a symmetric cipher

3) **Hide everything that is not already known**

Indeed, we cannot hide something that would be already known a priori. But the encrypted message must hide everything else...

Equivalently, an attacker should not be able to learn anything new (that he does not already know) about the plaintext after seeing and analyzing the ciphertext.

Shannon Secrecy

Shannon Secrecy

We denote:

- $\mathcal{M}$: the space of the plaintexts
- $\mathcal{C}$: the space of the ciphertexts
- $\mathcal{K}$: the space of the secret keys
- K: the random variable over $\mathcal{K}$ corresponding to the output of *KeyGen*.

Shannon Secrecy

Let M be a random variable over $\mathcal{M}$.

$\hookrightarrow$ M captures all the information a priori known by the attacker. For instance, if the plaintext start with "Hello", then M will take this information into account.

The attacker only knows that:

- **k** is drawn according to K ;
- **m** is drawn according to M ;
- *Enc* can use its own random source. But using Kerckhoffs principle, the attacker knows the distribution of the random variable $C := Enc(K, M)$ defined on $\mathcal{C}$.

Shannon Secrecy

Finally:

- all the information about the plaintext **m** that could be known **before** watching the ciphertext **c** is captured by the random variable M.
- all the information about the plaintext **m** that could be known **after** watching the ciphertext **c** is captured by the random variable $M|C$.

$\Rightarrow$ **Shannon Secrecy:**

$$M \sim M|C$$

Shannon Secrecy

Definition (Shannon Secrecy)

A symmetric encryption scheme $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, \mathit{Enc}, \mathit{Dec})$ is **Shannon Secure** according to a random variable M over $\mathcal{M}$ if for all $\mathbf{m} \in \mathcal{M}$ and all $\mathbf{c} \in \mathcal{C}$:

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

We say it is Shannon Secure if it is Shannon Secure according to all the random variable over $\mathcal{M}$.

Perfect Secrecy

Perfect Secrecy

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

Let C_1 and C_2 be the two following random variables:

$$\begin{aligned} C_1 &:= Enc(K, \mathbf{m}_1) \\ C_2 &:= Enc(K, \mathbf{m}_2) \end{aligned}$$

$\Rightarrow$ **Perfect Secrecy:**

$$\forall(\mathbf{m}_1, \mathbf{m}_2) \in \mathcal{M}^2,\ \textbf{we have } C_1 \sim C_2$$

Perfect Secrecy

Definition (Perfect Secrecy)

A symmetric encryption scheme $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, \mathit{Enc}, \mathit{Dec})$ is **Perfectly Secure** if for all $(\mathbf{m}_1, \mathbf{m}_2) \in \mathcal{M}^2$ and all $\mathbf{c} \in \mathcal{C}$:

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

This notion of security is much more simple to use than Shannon Secrecy.

Equivalence Theorem

Theorem (Equivalence Theorem)

A symmetric encryption scheme is Perfectly Secure if and only if it is Shannon Secure.

Proof on the board.

One-Time-Pad

One-Time-Pad (OTP)

Definition of the OTP cryptosystem:

- $\mathcal{M} := \{0,1\}^n$ $\longrightarrow$ binary string of length n
- $\mathcal{K} := \{0,1\}^n$ $\longrightarrow$ **the key is as long as the message to encrypt**
- K is the uniform distribution over $\mathcal{K}$
- $Enc(\mathbf{k}, \mathbf{m}) := \mathbf{k} \oplus \mathbf{m}$ where $\oplus$ stands for the bit by bit XOR
- $Dec(\mathbf{k}, \mathbf{m}) := Enc(\mathbf{k}, \mathbf{m})$

One-Time-Pad (OTP)

Definition of the OTP cryptosystem:

- $\mathcal{M} := \{0,1\}^n$ $\longrightarrow$ binary string of length n
- $\mathcal{K} := \{0,1\}^n$ $\longrightarrow$ **the key is as long as the message to encrypt**
- K is the uniform distribution over $\mathcal{K}$
- $Enc(\mathbf{k}, \mathbf{m}) := \mathbf{k} \oplus \mathbf{m}$ where $\oplus$ stands for the bit by bit *XOR*
- $Dec(\mathbf{k}, \mathbf{m}) := Enc(\mathbf{k}, \mathbf{m})$

Theorem (One-Time-Pad)

The One-Time-Pad cryptosystem is Perfectly Secure.

Proof on the board.

One-Time-Pad (OTP)

⚠ OTP is difficult to use in practice because:

- the entities must exchange a lot of keys because **a key can only be used once** ;
- and these **keys are incompressible** since they’ve been drawn uniformly at random.

One-Time-Pad (OTP)

⚠ OTP is difficult to use in practice because:

- the entities must exchange a lot of keys because **a key can only be used once** ;
- and these **keys are incompressible** since they've been drawn uniformly at random.

Remarks:

- We can replace the *XOR* by the addition/substraction modulo q if the alphabet is $\{0, \cdots, q-1\}$. For instance, in Vigenère or Caesar cipher, we use the addition modulo 26 to encrypt and the substraction modulo 26 to decrypt.
- Vigenère cipher where is also called Vernam cipher and it is Perfectly Secure

Shannon Theorem

Shannon Theorem

Theorem (Shannon)

*For all symmetric encryption scheme $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, Enc, Dec)$ which is **Perfectly Secure**, we necessarily have $\sharp\mathcal{K} \geq \sharp\mathcal{M}$.*

Proof on the board.

Remark. If $\mathcal{K}$ and $\mathcal{M}$ are respectively $\mathcal{A}^t$ and $\mathcal{A}^n$ with $\mathcal{A}$ an alphabet of size (cardinal) $\sharp\mathcal{A} = q$ then $\sharp\mathcal{K} = q^t$ and $\sharp\mathcal{M} = q^n$.

Thus, the Shannon theorem implies that the keys must be longer than the plaintexts which makes these cryptosystems unusable in practice!!!

Shannon Theorem

Theorem (Shannon)

*For all symmetric encryption scheme $(\mathcal{M}, \mathcal{C}, \mathcal{K}, K, Enc, Dec)$ which is **Perfectly Secure**, we necessarily have $\sharp\mathcal{K} \geq \sharp\mathcal{M}$.*

Proof on the board.

Remark. If $\mathcal{K}$ and $\mathcal{M}$ are respectively $\mathcal{A}^t$ and $\mathcal{A}^n$ with $\mathcal{A}$ an alphabet of size (cardinal) $\sharp\mathcal{A} = q$ then $\sharp\mathcal{K} = q^t$ and $\sharp\mathcal{M} = q^n$.

Thus, the Shannon theorem implies that the keys must be longer than the plaintexts which makes these cryptosystems unusable in practice!!!

Solution: Computational Security