# HAI709I - Fondements cryptographiques pour la sécurité

Cours 1 - Cryptographie classique et sécurité inconditionnelle

Rocco Mora - rocco.mora@umontpellier.fr

8 September, 2025

Université de Montpellier – Faculté des Sciences M1 informatique, parcours Algo, IASD, Imagine

# Cryptographie [ Cryptography]

### Étymologie (grec ancien) :

- kruptos (κρυπτός) "caché"
- graphein (γράφειν) "écrire"



« L'art de l'écriture et le déchiffrement de codes »

# Cryptographie [ Cryptography]

### Étymologie (grec ancien) :

- kruptos (κρυπτός) "caché"
- graphein (γράφειν) "écrire"



« L'art de l'écriture et le déchiffrement de codes »

Bonne définition pour la cryptographie classique mais pas pour la cryptographie moderne

# 

### Étymologie (grec ancien) :

- kruptos (κρυπτός) "caché"
- graphein (γράφειν) "écrire"



« L'art de l'écriture et le déchiffrement de codes »

Bonne définition pour la cryptographie classique mais pas pour la cryptographie moderne

#### Cryptographie classique

- très créative
- presque aucune théorie derrière

# Cryptographie moderne (années 1970 - aujourd'hui)

- rigoureuse
- science avec de solides bases mathématiques

# Cryptographie [ Cryptography]

### Étymologie (grec ancien) :

- kruptos (κρυπτός) "caché"
- graphein (γράφειν) "écrire"



« L'art de l'écriture et le déchiffrement de codes »

Bonne définition pour la cryptographie classique mais pas pour la cryptographie moderne

### Cryptographie classique

- permettre des communications sûrs (confidentialité)
- ⇒ Applications militaires/gouvernementales

#### Cryptographie moderne

- intégrité
- échange de clés
- authentification
- vote électronique
- cryptomonnaie, ...
- ⇒ Vie quotidienne

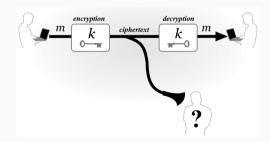
# Scénarios classiques en cryptographie

Cryptographie : consiste à concevoir et à utiliser des *codes* (ou chiffres) permettant à deux parties de

- 1. d'envoyer des messages,
- 2. de garder ces messages cachés à un espion/intercepteur [ eavesdropper] qui peut surveiller la communication.

**Scénario 1 : deux** parties communicantes séparées dans l'espace

⚠ Hypothèse : les deux parties ont partagé une clé à l'avance (par exemple, elles se sont rencontrées physiquement dans un lieu sécurisé)

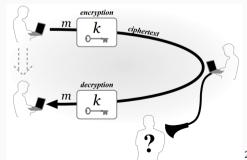


# Scénarios classiques en cryptographie

Cryptographie : consiste à concevoir et à utiliser des *codes* (ou chiffres) permettant à deux parties de

- 1. d'envoyer des messages,
- 2. de garder ces messages cachés à un espion/intercepteur [ eavesdropper] qui peut surveiller la communication.

Scénario 2 : même partie communiquant avec elle-même au fil du temps (par exemple, chiffrement de disque)



# Scénarios classiques en cryptographie

Cryptographie : consiste à concevoir et à utiliser des *codes* (ou chiffres) permettant à deux parties de

- 1. d'envoyer des messages,
- 2. de garder ces messages cachés à un espion/intercepteur [ eavesdropper] qui peut surveiller la communication.

Cryptanalyse [ Cryptanalysis]: a pour objectif de trouver des faiblesses ou des vulnérabilités dans un système cryptographique.

 ${\color{blue} {\sf Cryptologie}} = {\color{blue} {\sf Cryptographie}} + {\color{blue} {\sf Cryptanalyse}}$ 

# Syntaxe du chiffrement

Schéma de chiffrement (terme moderne pour codes) [ $\blacksquare$  Encryption scheme] : Espaces des messages/textes en clair [ $\blacksquare$  messages/plaintexts]  $\mathcal{M}$ , des textes chiffrés ciphertexts]  $\mathcal{C}$ , des clés [ $\blacksquare$  keys]  $\mathcal{K}$ . Il y a 3 algorithmes :

- Algorithme de génération de clé Gen
   [ Key generation algorithm] (probabiliste)
  - sortie :  $k \in \mathcal{K}$  selon une certaine distribution (généralement uniforme dans  $\mathcal{K}$ )
- Algorithme de chiffrement Enc<sub>k</sub>(m)
   Encryption algorithm
  - ullet entrée : clé  $k\in\mathcal{K}$  et texte en clair  $m\in\mathcal{M}$
  - sortie : texte chiffré c
- Algorithme de déchiffrement  $Dec_k(c)$ 
  - $[ \begin{tabular}{ll} \blacksquare \begin{tabular}{ll$ 
    - entrée : clé  $k \in \mathcal{K}$  et texte chiffré c
    - sortie : texte en clair  $m \in \mathcal{M}$

Propriété de correction

[Section Content in C

pour tout  $k \in \mathcal{K}$ ,

 $\operatorname{Dec}_k(\operatorname{Enc}_k(m))=m.$ 

# Principe de Kerchoff

La sortie k de Gen est secrète, mais Dec ne l'est pas. Pourquoi?

Sécurité par l'obscurité : garder les algorithmes secrets améliore la sécurité

# Principe de Kerchoff

La sortie k de Gen est secrète, mais Dec ne l'est pas. Pourquoi?

Sécurité par obscurité : garder les algorithmes secrets améliore la sécurité

Principe de Kerchoff ["La cryptographie militaire", 1883]

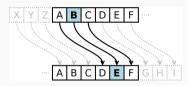
Le système de chiffrement ne doit pas nécessairement être secret et doit pouvoir tomber entre les mains de l'ennemi sans inconvénient.

#### Raisons:

- 1. Il est plus facile de garder secrète une clé courte qu'un algorithme
  - dans une organisation, tous les employés connaissent le système utilisé par les autres
  - risque de fuite
  - ingénierie inverse
- 2. plus facile de remplacer une clé en cas d'exposition
- 3. encourager l'examen public et standardizer/normalizer les systèmes

### Le chiffrement de César ["De Vita Caesarum, Divus Iulius", 121 apr. J.-C.]

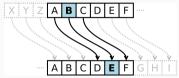
Chiffrement : Décaler les lettres de l'alphabet de 3 places vers l'avant (de manière cyclique)



Quel est le problème avec ce système?

### Le chiffrement de César ["De Vita Caesarum, Divus Iulius", 121 apr. J.-C.]

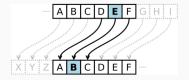
Chiffrement : Décaler les lettres de l'alphabet de 3 places vers l'avant (de manière cyclique)



Quel est le problème avec ce système?

Réponse : Il n'y a pas de clé

Déchiffrement : Décaler les lettres de l'alphabet de 3 places vers l'arrière



# 

Ajoutons donc une clé!  $k \in \mathcal{K} = \{0, 1, \dots, 25\} = \mathcal{M} = \mathcal{C}$ .

Nous devons associer les lettres à des chiffres :  $a \to 0$ ,  $b \to 1$ , ...,  $z \to 25$ .

Alors

$$\operatorname{Enc}_k(m_1,\ldots,m_\ell)=c_1,\ldots,c_\ell, \text{ où } c_i=m_i+k \mod 26$$

et

$$\operatorname{Dec}_k(c_1,\ldots,c_\ell)=m_1,\ldots,m_\ell, \text{ où } m_i=c_i-k \mod 26.$$

Prouvons la propriété de correction :

$$\begin{aligned} \operatorname{Dec}_k(\operatorname{Enc}_k(m_1, \dots, m_\ell)) = & \operatorname{Dec}_k(m_1 + k \mod 26, \dots, m_\ell + k \mod 26) \\ = & (m_1 + k \mod 26) - k \mod 26, \dots, (m_\ell + k \mod 26) - k \mod 26 \\ = & m_1 + k - k \mod 26, \dots, m_\ell + k - k \mod 26 \\ = & m_1 \mod 26, \dots, m_\ell \mod 26 \\ = & m_1, \dots, m_\ell \end{aligned}$$

# 

Ajoutons donc une clé!  $k \in \mathcal{K} = \{0, 1, \dots, 25\} = \mathcal{M} = \mathcal{C}$ .

Nous devons associer les lettres à des chiffres :  $a \to 0$ ,  $b \to 1$ , ...,  $z \to 25$ .

Alors

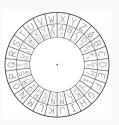
$$\operatorname{Enc}_k(m_1,\ldots,m_\ell)=c_1,\ldots,c_\ell, \ \operatorname{où}\ c_i=m_i+k \ \operatorname{mod}\ 26$$

et

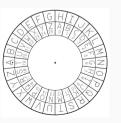
$$\operatorname{Dec}_k(c_1,\ldots,c_\ell)=m_1,\ldots,m_\ell, \text{ où } m_i=c_i-k \mod 26.$$

Le chiffrement par décalage est une généralisation du chiffrement de César où k=3 :





$$k = 19$$



Réponse : L'espace des clés est petit, ce qui permet d'effectuer une recherche exhaustive

Soit c = BQSHOFJEWHQFXYUUIJSEEB

Réponse : L'espace des clés est petit, ce qui permet d'effectuer une recherche exhaustive

Soit c = BQSHOFJEWHQFXYUUIJSEEB

Supposons que k=1 :  $\mathtt{Dec}_1(c)=\mathtt{APRGNEIDVGPEWXTTHIRDDA}$   $\pmb{\mathsf{X}}$ 

**Réponse** : L'espace des clés est petit, ce qui permet d'effectuer une recherche exhaustive

Soit c = BQSHOFJEWHQFXYUUIJSEEB

Supposons que k = 1:  $Dec_1(c) = APRGNEIDVGPEWXTTHIRDDA X$ 

Supposons que k=2 :  $\mathrm{Dec}_2(c)=\mathrm{ZOQFMDHCUFODVWSSGHQCCZ}$   $\color{red}{\mathsf{X}}$ 

Réponse : L'espace des clés est petit, ce qui permet d'effectuer une recherche exhaustive

Soit c = BQSHOFJEWHQFXYUUIJSEEB

Supposons que k=1 :  $\mathtt{Dec}_1(c)=\mathtt{APRGNEIDVGPEWXTTHIRDDA}$   $\pmb{\mathsf{X}}$ 

Supposons que k=2:  $\mathrm{Dec}_2(c)=\mathrm{ZOQFMDHCUFODVWSSGHQCCZ}$   $\color{red}{\mathsf{X}}$ 

Supposons que k = 3: Dec<sub>3</sub>(c) = YNPELCGBTENCUVRRFGPBBY X

Réponse : L'espace des clés est petit, ce qui permet d'effectuer une recherche exhaustive

```
Soit c = \operatorname{BQSHOFJEWHQFXYUUIJSEEB}
Supposons que k = 1 : \operatorname{Dec}_1(c) = \operatorname{APRGNEIDVGPEWXTTHIRDDA} X
Supposons que k = 2 : \operatorname{Dec}_2(c) = \operatorname{ZOQFMDHCUFODVWSSGHQCCZ} X
Supposons que k = 3 : \operatorname{Dec}_3(c) = \operatorname{YNPELCGBTENCUVRRFGPBBY} X
:
Supposons que k = 16 : \operatorname{Dec}_{16}(c) = \operatorname{LACRYPTOGRAPHIEESTCOOL} \checkmark
```

- l'espace de clés doit être suffisamment grand pour rendre la force brute impossible
- dépend des ressources disponibles (supercalculateurs, serveurs cloud, GPU, ...)
- ullet de nos jours,  $|\mathcal{K}| \geq 2^{80}$  au minimum, de préférence  $|\mathcal{K}| \geq 2^{128}$

# Chiffrement par substitution mono-alphabétique

Augmentons  $|\mathcal{K}|$ . Comment? **Idée**: au lieu d'utiliser des décalages, permutons les lettres.

# Exemple



En termes mathématiques, k est une permutation de  $\{0, 1, \dots, 25\}$ .

Combien y a-t-il de clés possibles?  $26 \cdot 25 \cdot \dots \cdot 1 = 26! \approx 2^{88}$ 

 $\rightarrow$  une attaque par force brute est irréalisable!

"26 factorielle"

Mais quelle est la faiblesse de ce chiffrement?

Réponse : dans les langues réelles, les lettres apparaissent avec des probabilités différentes

Par exemple, en français



Lettre																										
Fréq. (%)	7.6	0.9	3.3	3.7	14.7	1.1	1.0	0.7	7.5	0.6	0.05	5.5	2.7	7.1	5.2	3.0	1.3	6.5	7.9	7.0	6.4	1.8	0.04	0.4	0.3	0.1

Analyse de fréquence : comparer les fréquences théoriques avec celles du texte chiffré.

- Le texte doit être suffisamment long.
- Certaines suppositions peuvent être erronées, mais les informations sont tout de même suffisantes pour récupérer le message.
- On peut exploiter les propriétés spécifiques d'une langue.

# Le chiffrement de Vigenère (ou chiffrement polyalphabétique)

Le chiffrement lettre par lettre n'est pas sécurisé! Appliquez la clé à un bloc de caractères. [Giovan Battista Bellaso, 1553]

En termes mathématiques,  $k=(k_1,\ldots,k_t)$  est un vecteur dans  $\mathcal{K}=\{0,1,\ldots,25\}^t$ .

L'entier t est la période de la clé.

$$\operatorname{Enc}_k(m_1, \dots, m_\ell) = m_1 + k_1 \mod 26, \qquad m_2 + k_2 \mod 26, \dots, \qquad m_t + k_t \mod 26, \\ m_{t+1} + k_1 \mod 26, \qquad m_{t+2} + k_2 \mod 26, \dots, \qquad m_{2t} + k_t \mod 26, \\ \vdots$$

$$\operatorname{Dec}_k(c_1,\ldots,c_\ell) = c_1 - k_1 \mod 26, \qquad c_2 - k_2 \mod 26,\ldots, \quad c_t - k_t \mod 26,$$
  $c_{t+1} - k_1 \mod 26, \quad c_{t+2} - k_2 \mod 26,\ldots, \quad c_{2t} - k_t \mod 26,$  :

# Le chiffrement de Vigenère (ou chiffrement polyalphabétique)

Le chiffrement lettre par lettre n'est pas sécurisé! Appliquez la clé à un bloc de caractères. [Giovan Battista Bellaso, 1553]

Exemple : $k = INFO$																						
Texte en clair	L	Α	С	R	Y	Р	Т	О	G	R	Α	Р	Н	1	Ε	Е	S	Т	С	О	0	L
Clé	Ι	Ν	F	0	I	Ν	F	0	Ι	Ν	F	0	1	Ν	F	0	1	Ν	F	0	Ι	N
Texte chiffré	Т	N	Н	F	G	С	Υ	С	0	Е	F	D	Р	٧	J	S	Α	G	Н	С	W	Υ

Maintenant, certains caractères du texte chiffré sont identiques, mais les caractères correspondants dans le texte en clair ne le sont pas.

→ l'attaque statistique/l'analyse de fréquence ne s'applique pas directement

Mais alors, quel est le problème avec ce chiffrement?

Pour tous les j, le j-ième flux

$$c_j, c_{j+t}, c_{j+2t}, \ldots$$

a été chiffré avec la même valeur  $k_j$ .

Devinez t et effectuez une analyse de fréquence sur chaque flux séparé

 $\Rightarrow$  la complexité devient 26t au lieu de 26t

- toujours plus difficile que le chiffrement par décalage, car il est impossible de vérifier si le flux "a du sens" dans la même langue
- améliorations possibles de l'attaque :
  - La méthode de Kasiski examine les apparences et les motifs pour déterminer la période
  - La méthode de l'indice de coïncidence peut aider à automatiser l'attaque
- Le chiffrement de Vigenère n'a été cassé qu'après plusieurs siècles

# Qu'avons-nous appris de la cryptographie classique?

#### De quoi avons-nous besoin:

- Des définitions formelles des garanties de sécurité à atteindre, par exemple qu'un attaquant ne soit pas en mesure de récupérer
  - la clé
  - le texte en clair à partir du texte chiffré
  - n'importe quel caractère du texte en clair à partir du texte chiffré
  - n'importe quelle information sur le texte en clair à partir du texte chiffré
- Définitions formelles des attaquants (modèle de menace), par exemple l'attaquant peut
  - observer les textes chiffrés
  - apprendre certaines paires texte clair/texte chiffré, sans choisir
  - obtenir certaines paires texte clair/texte chiffré, en choisissant les textes clairs
  - obtenir certaines paires texte clair/texte chiffré, en choisissant les textes chiffrés
- Preuves de sécurité basées sur des hypothèses mathématiques

# Rappels sur la théorie des probabilités

Nous avons vu des chiffrements historiques faibles. Nous allons maintenant nous intéresser à l'autre extrême.

### Sécurité inconditionnelle (définition informelle)

L'adversaire dispose d'une puissance de calcul illimitée. L'observation du texte chiffré n'a aucun effet sur les connaissances de l'adversaire.

Afin de formaliser cette notion, on a besoin de la théorie des probabilités.

- Univers : ensemble  $\Omega$  de tous les résultats d'une expérience aléatoire.
- Espace d'événements : ensemble  $\mathcal{F}$  d'éléments  $A \subseteq \Omega$  (appelés événements).
- Mesure de probabilité : fonction  $Pr: \mathcal{F} \to \mathbb{R}$  telle que
  - $\forall A \in \mathcal{F}, \Pr(A) \geq 0$ ;
  - $Pr(\Omega) = 1$ ;
  - si  $A_1,A_2,\ldots$  sont disjoints (c. -à-d  $A_i\cap A_j=\emptyset$  pour  $i\neq j$ ), alors probabilité

$$\Pr(\bigcup_i A_i) = \sum_i \Pr(A_i).$$

Axiomes de

- Quelques propriétés :
  - $\Pr(A) \in [0,1]$ ;
  - $A \subseteq B \Rightarrow \Pr(A) \leq \Pr(B)$ ;
  - $\Pr(A \cap B) \leq \min(\Pr(A), \Pr(B))$ ;
  - $Pr(A \cup B) \leq Pr(A) + Pr(B)$ ;
  - $Pr(\Omega \setminus A) = 1 Pr(A)$
  - if  $A_1, \ldots, A_k$  sont disjoints et  $\bigcup_{i=1}^k A_i = \Omega$ , alors  $\sum_{i=1}^k \Pr(A_i) = 1$ .
- La probabilité conditionnelle d'un événement A étant donné un événement B avec une probabilité non nulle est

$$\Pr(A \mid B) \stackrel{\text{def}}{=} \frac{\Pr(A \cap B)}{\Pr(B)}.$$

- Deux événements sont indépendants si et seulement si  $\Pr(A \cap B) = \Pr(A)\Pr(B)$ , ou, de manière équivalente,  $\Pr(A \mid B) = \Pr(A)$ .
- Théorème de Bayes

$$\Pr(A \mid B) \stackrel{\mathsf{def}}{=} \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B)}.$$

**Variable aléatoire** : est une fonction  $X : \Omega \to E$ , avec E mesurable.

La probabilité que X prenne la valeur  $k \in E$  est

$$\Pr(X = k) = \Pr(\{\omega \in \Omega \mid X(\omega) = k\}).$$

De manière plus générale, étant donné un ensemble mesurable  $S\subseteq E$ ,

$$\Pr(X \in S) = \Pr(\{\omega \in \Omega \mid X(\omega) \in S\}).$$

- Nos univers sont les espaces des clés, des messages et des textes chiffrés  $\mathcal{K}, \mathcal{M}, \mathcal{C}$ .
- La distribution sur  $\mathcal{K}$  est donnée par Gen, et nous supposons qu'elle est uniforme, c'est-à-dire que toutes les clés ont la même probabilité d'être choisies.
- K, M, C sont les variables aléatoires désignant respectivement la valeur de la clé générée par Gen, le message, et le texte chiffré résultant.

### Chiffrement inconditionnellement sûr

### Sécurité inconditionnelle [ Perfect secrecy] [Shannon, 1949]

Un schéma de chiffrement (Gen, Enc, Dec) avec un espace de messages  $\mathcal{M}$  est inconditionnellement sûr si, pour chaque distribution de probabilité pour M, chaque message  $m \in \mathcal{M}$  et chaque texte chiffré  $c \in \mathcal{C}$  avec  $\Pr(\mathcal{C} = c) > 0$ , nous avons

$$\Pr(M = m \mid C = c) = \Pr(M = m).$$

#### **Proposition**

Un schéma de chiffrement (Gen, Enc, Dec) avec un espace de messages  $\mathcal{M}$  est inconditionnellement sûr si et seulement si, pour chaque  $m, m' \in \mathcal{M}$  et chaque  $c \in \mathcal{C}$ ,

$$\Pr(\operatorname{Enc}_K(m)=c)=\Pr(\operatorname{Enc}_K(m')=c).$$

Preuve : à faire dans le TD.

# One-Time pad (OTP)/Chiffre de Vernam/Masque jetable [Vernam, 1917]

(Utilisé pendant la guerre froide pour les communications entre les gouvernements américain et soviétique.)

Fixons  $\ell > 0$  et posons  $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0,1\}^{\ell}$ .

- Gen choisit  $k \in \mathcal{K}$  de manière uniforme;
- $c = \operatorname{Enc}_k(m) = m \oplus k$  (XOR des chaînes);

# One-Time pad (OTP)/Chiffre de Vernam/Masque jetable [Vernam, 1917]

(Utilisé pendant la guerre froide pour les communications entre les gouvernements américain et soviétique.)

Fixons  $\ell > 0$  et posons  $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0, 1\}^{\ell}$ .

- Gen choisit  $k \in \mathcal{K}$  de manière uniforme;
- $c = \operatorname{Enc}_k(m) = m \oplus k$  (XOR des chaînes);
- $\operatorname{Dec}_k(c) = c \oplus k$ .

#### **Théorème**

Le chiffre de Vernam est inconditionnellement sûr.

→ le chiffrement de Vernam a été proposé avant que la notion de sécurité inconditionnelle ne soit définie!

# Preuve de la sécurité inconditionnelle du OTP (1/2)

Pour tout  $c \in \mathcal{C}$ , et  $m \in \mathcal{M}$ , avec  $\Pr(M = m) > 0$ ,

$$Pr(C = c \mid M = m) = Pr(K \oplus M = c \mid M = m)$$
$$= Pr(K \oplus m = c \mid M = m)$$
$$= Pr(K = m \oplus c \mid M = m)$$
$$= 2^{-\ell}.$$

Fixons une distribution sur  $\mathcal{M}$ . Pour tout  $c \in \mathcal{C}$ ,

$$\Pr(C = c) = \sum_{m \in \mathcal{M}} \Pr(C = c \mid M = m) \cdot \Pr(M = m)$$
$$= 2^{-\ell} \cdot \sum_{m \in \mathcal{M}} \Pr(M = m)$$
$$= 2^{-\ell}.$$

# Preuve de la sécurité inconditionnelle du OTP (2/2)

Alors, d'après le théorème de Bayes,

$$\Pr(M = m \mid C = c) = \frac{\Pr(C = c \mid M - m) \cdot \Pr(M = m)}{\Pr(C = c)}$$
$$= \frac{2^{-\ell} \cdot \Pr(M = m)}{2^{-\ell}}$$
$$= \Pr(M = m).$$

Par définition, le OTP est inconditionnellement sûr.

Il ne s'agit certainement pas d'une propriété insignifiante. Par exemple

#### **Proposition**

Un chiffrement de Vigenère dont la clé est plus courte que le texte en clair n'est pas inconditionnellement sûr.

### Limites de la sécurité inconditionnelle

Le OTP est essentiellement un cas particulier du chiffrement de Vigenère où la clé est aussi longue que le texte en clair.

#### Théorème

Si (Gen, Enc, Dec) est un schéma de chiffrement inconditionnellement sûr, alors  $|\mathcal{K}| \geq |\mathcal{M}|$ .

**Preuve**: Prouvons-le par contradiction, c'est-à-dire supposons que  $|\mathcal{K}| < |\mathcal{M}|$ . Fixons  $c \in \mathcal{C}$ . Définissons

$$\mathcal{M}(c) \stackrel{\mathsf{def}}{=} \{ m \in \mathcal{M} \mid \exists k \in \mathcal{K} \; \mathsf{tel} \; \mathsf{que} \; m = \mathsf{Dec}_k(c) \}.$$

Nous avons  $|\mathcal{M}(c)| \leq |\mathcal{K}| < |\mathcal{M}|$ . Soit  $m' \in \mathcal{M} \setminus \mathcal{M}(c)$ . Alors

$$\Pr(M = m' \mid C = c) = 0 \neq \Pr(M = m').$$

Par conséquent, le schéma ne serait pas inconditionnellement sûr.

On veut des clés plus courtes ightarrow on a besoin de notions de sécurité plus faibles

