Cryptographie Post-Quantique (PQC)

De la théorie vers une normalisation du NIST

Boly SECK

Rencontres des jeunes chercheurs africains en France (RJCAF)
Cinquième Édition
Institut Henri Poincaré (Paris)

5 et 6 Décembre 2022

Équipe encadrante de la cotutelle

- Pierre-Louis Cayrel (Directeur de thèse, Maître de conférences HDR, UJM, FRANCE).

Équipe encadrante de la cotutelle

- Pierre-Louis Cayrel (Directeur de thèse, Maître de conférences HDR, UJM, FRANCE).

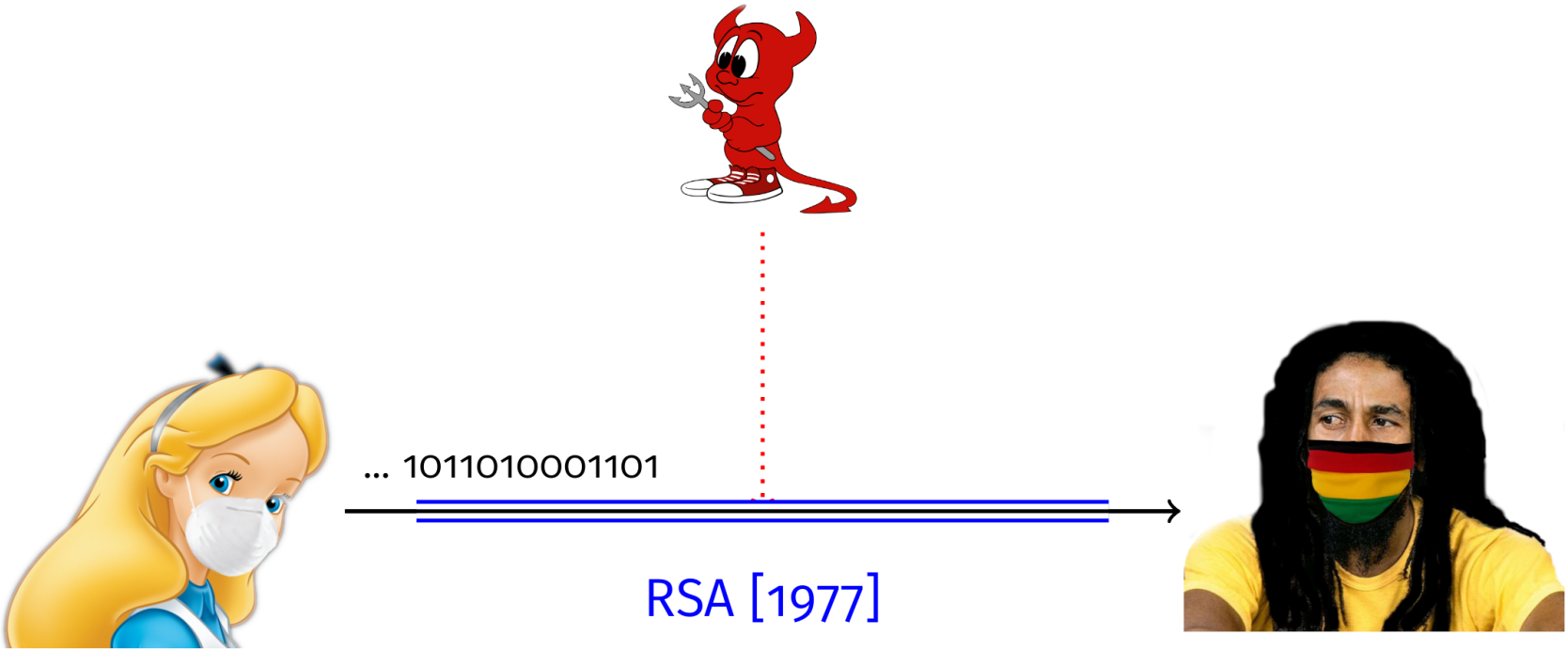
- Idy Diop (Co-directeur de thèse, Professeur, ESP, SENEGAL).

Équipe encadrante de la cotutelle

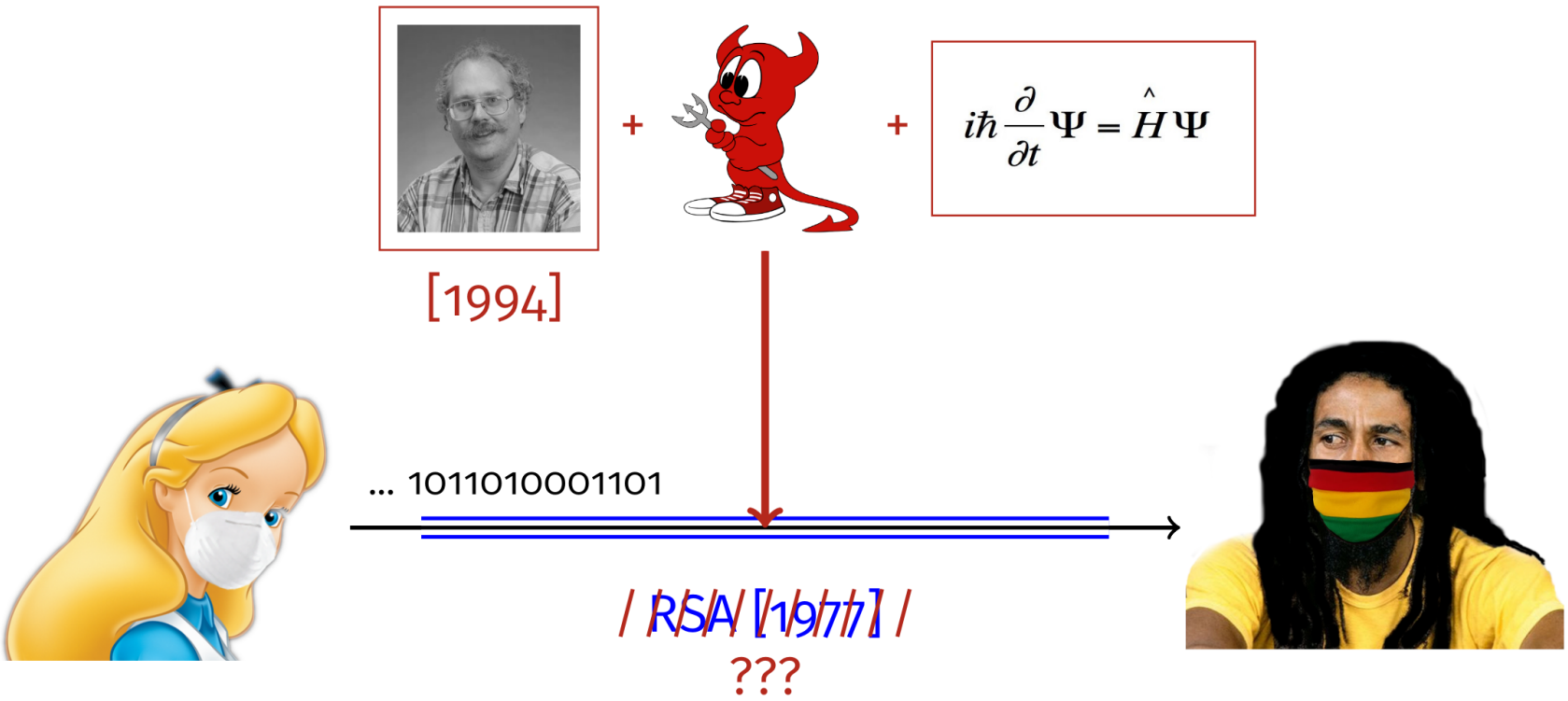
- Pierre-Louis Cayrel (Directeur de thèse, Maître de conférences HDR, UJM, FRANCE).
- Idy Diop (Co-directeur de thèse, Professeur, ESP, SENEGAL).
- Morgan Barbier (Encadrant, Maître de conférences, ENSICAEN, FRANCE)

Context : cryptologie

Context : cryptologie



Context : cryptologie



Context : post-quantique

Les standards de la cryptographie à clé publique utilisés aujourd'hui sont basés sur les problèmes difficiles :

Context : post-quantique

Les standards de la cryptographie à clé publique utilisés aujourd'hui sont basés sur les problèmes difficiles :

- **Factorisation d'entiers** (RSA) ;

Context : post-quantique

Les standards de la cryptographie à clé publique utilisés aujourd'hui sont basés sur les problèmes difficiles :

- **Factorisation d'entiers** (RSA) ;
- **Logarithme discret** (DH et DSA).

Context : post-quantique

Les standards de la cryptographie à clé publique utilisés aujourd'hui sont basés sur les problèmes difficiles :

- **Factorisation d'entiers** (RSA) ;
- **Logarithme discret** (DH et DSA).

Problème ?

Context : post-quantique

Les standards de la cryptographie à clé publique utilisés aujourd'hui sont basés sur les problèmes difficiles :

- **Factorisation d'entiers** (RSA) ;
- **Logarithme discret** (DH et DSA).

Problème ? = *Il existe des algorithmes quantiques qui permettent de résoudre ces problèmes*.

Context : post-quantique

Les standards de la cryptographie à clé publique utilisés aujourd'hui sont basés sur les problèmes difficiles :

- **Factorisation d'entiers** (RSA) ;
- **Logarithme discret** (DH et DSA).

Problème ? = *Il existe des algorithmes quantiques qui permettent de résoudre ces problèmes*.

Solution ?

Context : post-quantique

Les standards de la cryptographie à clé publique utilisés aujourd'hui sont basés sur les problèmes difficiles :

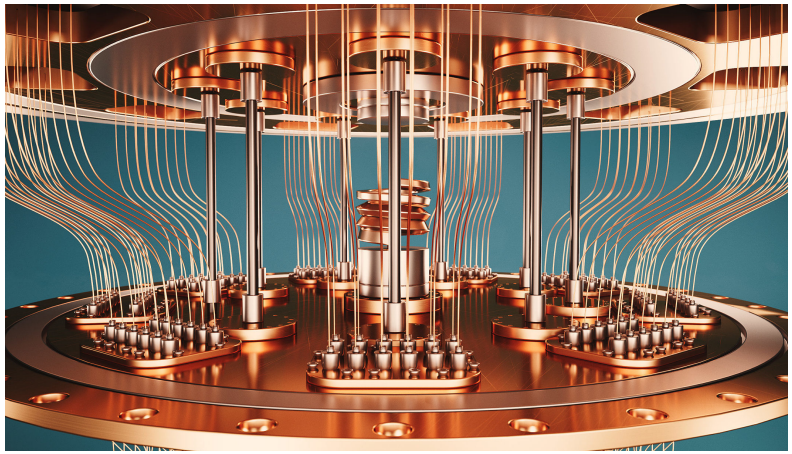
- **Factorisation d'entiers** (RSA) ;
- **Logarithme discret** (DH et DSA).

Problème ? = *Il existe des algorithmes quantiques qui permettent de résoudre ces problèmes*.

Solution ? = *Trouver des alternatives qui résistent aux attaques quantiques*.

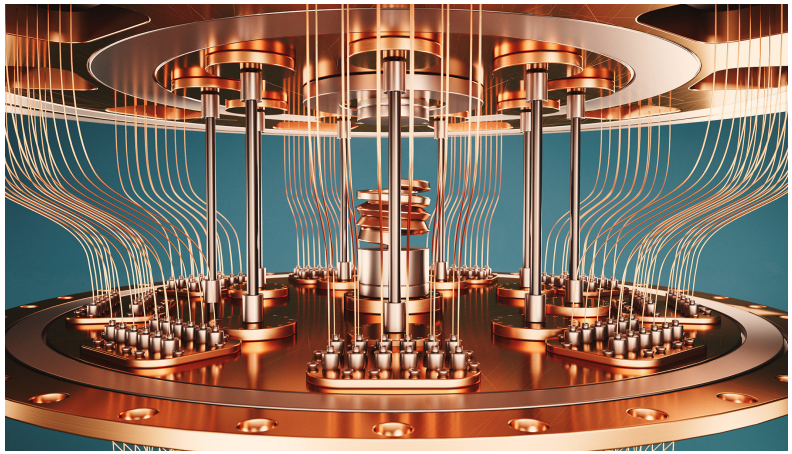
Context : post-quantique

- Hypothèse : l'attaquant a accès à un ordinateur quantique.



Context : post-quantique

- Hypothèse : l'attaquant a accès à un ordinateur quantique.



- L'agence américaine des standards à lancé depuis décembre 2016 un processus de normalisation de cryptosystème post-quantique pour 2024.

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Context : post-quantique

Cyptosystème post-quantique ?

Context : post-quantique

Cyptosystème post-quantique ? = Protocole cryptographique basé sur un problème mathématique prouvé difficile différend de la théorie des nombres.

Context : post-quantique

Cyptosystème post-quantique ? = Protocole cryptographique basé sur un problème mathématique prouvé difficile différend de la théorie des nombres.

Les approches les plus populaires sont :

1. Recherche de mots de petits poids pour les réseaux (lattices) ;

Context : post-quantique

Cyptosystème post-quantique ? = Protocole cryptographique basé sur un problème mathématique prouvé difficile différend de la théorie des nombres.

Les approches les plus populaires sont :

1. Recherche de mots de petits poids pour les réseaux (lattices) ;
2. Problème de décodage des codes aléatoires (codes) ;

Context : post-quantique

Cyptosystème post-quantique ? = Protocole cryptographique basé sur un problème mathématique prouvé difficile différend de la théorie des nombres.

Les approches les plus populaires sont :

1. Recherche de mots de petits poids pour les réseaux (lattices) ;
2. Problème de décodage des codes aléatoires (codes) ;
3. Résolution de systemes de polynômes à plusieurs variables (multivariate) ;

Context : post-quantique

Cyptosystème post-quantique ? = Protocole cryptographique basé sur un problème mathématique prouvé difficile différend de la théorie des nombres.

Les approches les plus populaires sont :

1. Recherche de mots de petits poids pour les réseaux (lattices) ;
2. Problème de décodage des codes aléatoires (codes) ;
3. Résolution de systemes de polynômes à plusieurs variables (multivariate) ;
4. Isogénies (isogenies) ;

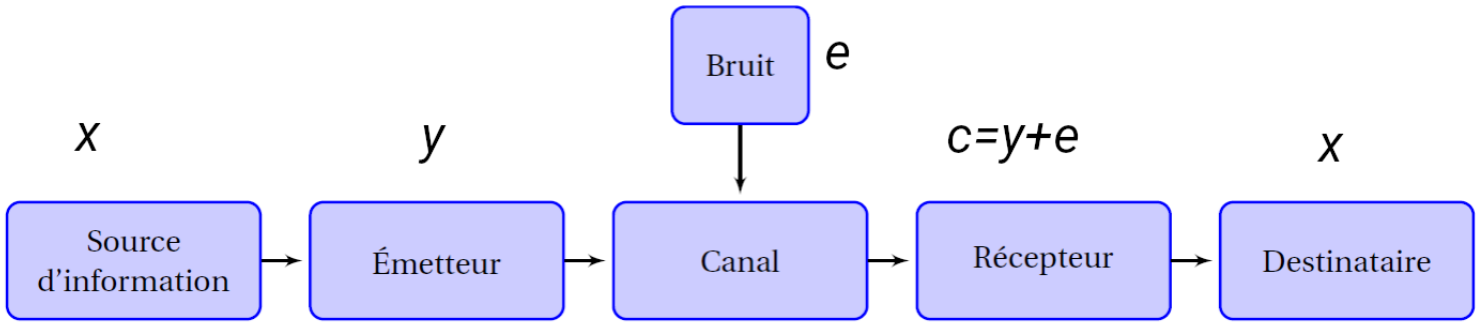
Context : post-quantique

Cyptosystème post-quantique ? = Protocole cryptographique basé sur un problème mathématique prouvé difficile différend de la théorie des nombres.

Les approches les plus populaires sont :

1. Recherche de mots de petits poids pour les réseaux (lattices) ;
2. Problème de décodage des codes aléatoires (codes) ;
3. Résolution de systemes de polynômes à plusieurs variables (multivariate) ;
4. Isogénies (isogenies) ;
5. Fonctions de hachage pour la signature (hash).

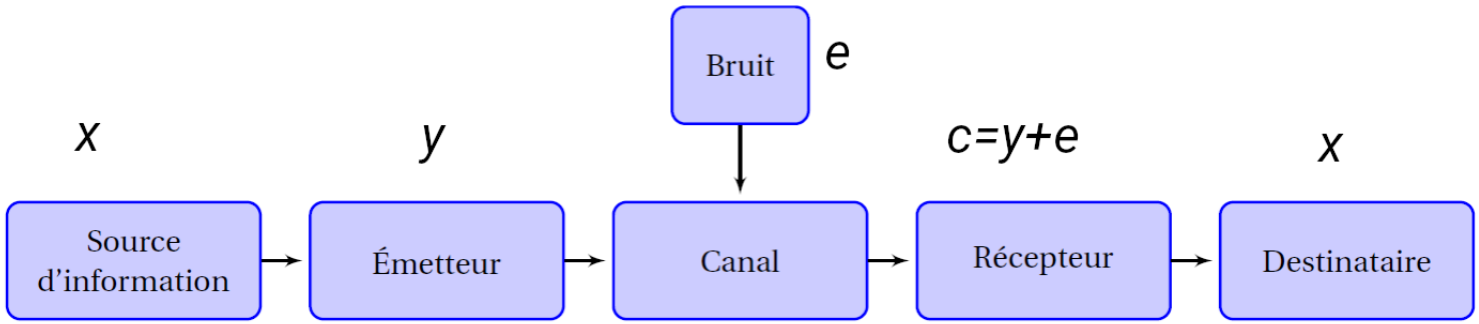
Codes correcteurs d'erreurs



Soit $\mathcal{C}$ un **code** de dimension k et de longueur n et $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_k)$ un message de longueur k.

1. $\boldsymbol{x}$ est **encodé** en un mot de **code** $\boldsymbol{y}$ de longueur n ;

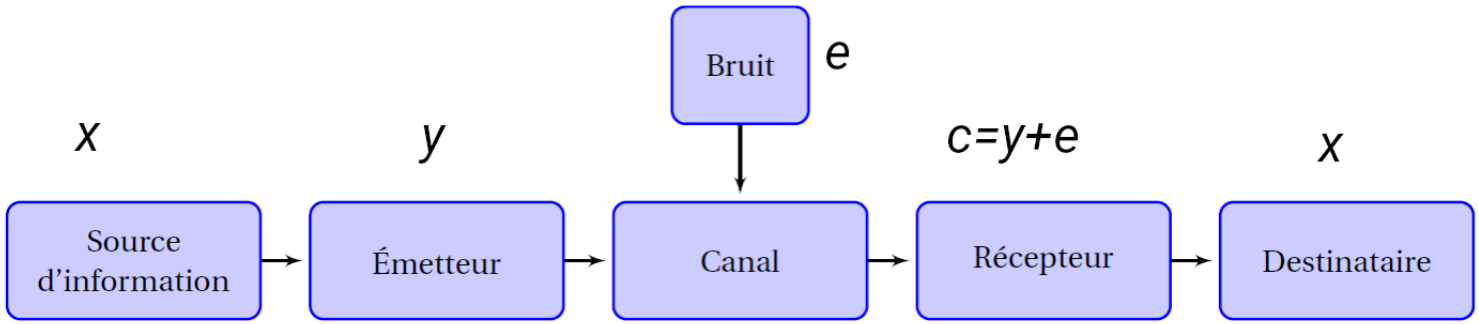
Codes correcteurs d'erreurs



Soit $\mathcal{C}$ un **code** de dimension k et de longueur n et $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_k)$ un message de longueur k.

1. $\boldsymbol{x}$ est **encodé** en un mot de **code** $\boldsymbol{y}$ de longueur n ;
2. $\boldsymbol{y}$ est transmis sur un canal bruité et on obtient $\boldsymbol{c} = \boldsymbol{y} + \boldsymbol{e}$ ($\boldsymbol{e}$ représente les erreurs de transmission) ;

Codes correcteurs d'erreurs



Soit $\mathcal{C}$ un **code** de dimension k et de longueur n et $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_k)$ un message de longueur k.

1. $\boldsymbol{x}$ est **encodé** en un mot de **code** $\boldsymbol{y}$ de longueur n ;
2. $\boldsymbol{y}$ est transmis sur un canal bruité et on obtient $\boldsymbol{c} = \boldsymbol{y} + \boldsymbol{e}$ ($\boldsymbol{e}$ représente les erreurs de transmission) ;
3. Le destinataire **décode** $\boldsymbol{c}$ pour retrouver le message $\boldsymbol{x}$.

Définitions

Definition (Code linéaire)

Un **code** $[n, k]$ linéaire $\mathcal{C}$ sur $\mathbb{F}_q$ est un sous-espace vectoriel de $\mathbb{F}_q^n$ de dimension k et les éléments de $\mathcal{C}$ sont appelés mots de code.

Définitions

Definition (Code linéaire)

Un **code** $[n, k]$ linéaire $\mathcal{C}$ sur $\mathbb{F}_q$ est un sous-espace vectoriel de $\mathbb{F}_q^n$ de dimension k et les éléments de $\mathcal{C}$ sont appelés mots de code.

Definition (Décodage)

Un **décodeur** du code $\mathcal{C}$ est une fonction $\Phi_{\mathcal{C}}$ qui peut décoder jusqu'à t erreurs tel que :

$$\forall \boldsymbol{y} \in \mathcal{C}, \forall \boldsymbol{e} \in \mathbb{F}_q^n, w(\boldsymbol{e}) \leq t \Rightarrow \Phi_{\mathcal{C}}(\boldsymbol{y} + \boldsymbol{e}) = \boldsymbol{y}.$$

Définitions

Definition (Matrice génératrice)

Une **matrice génératrice** de $\mathcal{C}$ est une matrice $\boldsymbol{G} \in \mathbb{F}_q^{k \times n}$ telle que ses lignes forment une base de l'espace vectoriel $\mathcal{C}$:

$$\mathcal{C} = \{\boldsymbol{x}\boldsymbol{G} | \boldsymbol{x} \in \mathbb{F}_q^k\}.$$

Définitions

Definition (Matrice génératrice)

Une **matrice génératrice** de $\mathcal{C}$ est une matrice $\boldsymbol{G} \in \mathbb{F}_q^{k \times n}$ telle que ses lignes forment une base de l'espace vectoriel $\mathcal{C}$:

$$\mathcal{C} = \{\boldsymbol{x}\boldsymbol{G} | \boldsymbol{x} \in \mathbb{F}_q^k\}.$$

Definition (Matrice de parité)

Une **matrice de parité** de $\mathcal{C}$ est une matrice $\boldsymbol{H} \in \mathbb{F}_q^{(n-k) \times n}$ telle que :

$$\mathcal{C} = \{\boldsymbol{y} \in \mathbb{F}_q^n | \boldsymbol{H}\boldsymbol{y}^T = 0\}.$$

Problème de décodage par syndrome

Definition (Décodage par syndrome binaire)

Soient $\boldsymbol{H} \in \mathbb{F}_2^{(n-k)\times n}$ une matrice de parité, $\boldsymbol{s} \in \mathbb{F}_2^{n-k}$ un vecteur et $t \in \mathbb{N}^*$, est-il possible de trouver un vecteur $\boldsymbol{e} \in \mathbb{F}_2^n$ de poids de Hamming $wt(\boldsymbol{e}) \leq t$ tel que $\boldsymbol{H}\boldsymbol{e}^T = \boldsymbol{s}$.

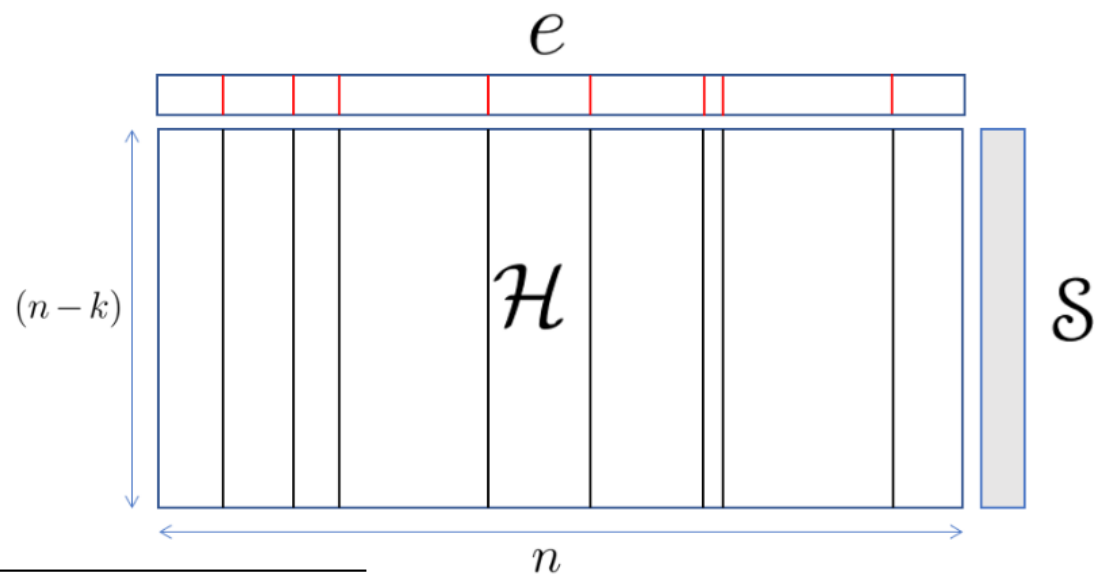
1. E. R. Berlekamp, R. J. McEliece, and H. C. A. van Tilborg. "On the inherent intractability of certain coding problems (Corresp.)". In : IEEE Transactions on Information Theory 24.3 (1978), pp. 384–386.

Problème de décodage par syndrome

Definition (Décodage par syndrome binaire)

Soient $\boldsymbol{H} \in \mathbb{F}_2^{(n-k)\times n}$ une matrice de parité, $\boldsymbol{s} \in \mathbb{F}_2^{n-k}$ un vecteur et $t \in \mathbb{N}^*$, est-il possible de trouver un vecteur $\boldsymbol{e} \in \mathbb{F}_2^n$ de poids de Hamming $wt(\boldsymbol{e}) \leq t$ tel que $\boldsymbol{H}\boldsymbol{e}^T = \boldsymbol{s}$.

Ce problème est $\mathcal{NP}$-complete [1].



1. E. R. Berlekamp, R. J. McEliece, and H. C. A. van Tilborg. "On the inherent intractability of certain coding problems (Corresp.)". In : IEEE Transactions on Information Theory 24.3 (1978), pp. 384–386.

Design de cryptosystème basé sur les codes

Étape 1

- Choisir une famille $\mathcal{F}$ de codes structurés ;
- Un décodeur $\Phi_{\mathcal{F}}$ qui peut corrigé jusqu'à t erreurs ;
- Un shaker !

Design de cryptosystème basé sur les codes

Étape 1

- Choisir une famille $\mathcal{F}$ de codes structurés ;
- Un décodeur $\Phi_{\mathcal{F}}$ qui peut corrigé jusqu'à t erreurs ;
- Un shaker !

Étape 2

Génération des clefs

$\boldsymbol{G} \leftarrow_{\$} \mathcal{F}$

$\boldsymbol{G}_{\text{pub}} \leftarrow$ Shake $(\boldsymbol{G})$

Chiffrement

$\boldsymbol{e} \leftarrow_{\$} \mathbb{F}_2^n$, s.t. $w(\boldsymbol{e}) = t$

$\boldsymbol{c} \leftarrow \boldsymbol{x}\boldsymbol{G}_{\text{pub}} + \boldsymbol{e}$

Déchiffrement

$\boldsymbol{x} \leftarrow \Phi_{\mathcal{F}}(\boldsymbol{G}, \boldsymbol{c})$

Design de cryptosystème basé sur les codes

Étape 1

- Choisir une famille $\mathcal{F}$ de codes structurés ;
- Un décodeur $\Phi_{\mathcal{F}}$ qui peut corrigé jusqu'à t erreurs ;
- Un shaker !

Étape 2

Génération des clefs

$\boldsymbol{G} \leftarrow_{\$} \mathcal{F}$

$\boldsymbol{G}_{\text{pub}} \leftarrow$ Shake $(\boldsymbol{G})$

Chiffrement

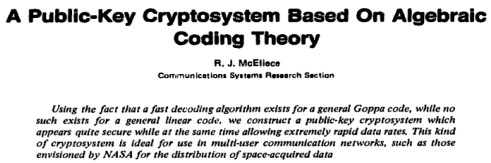
$\boldsymbol{e} \leftarrow_{\$} \mathbb{F}_2^n$, s.t. $w(\boldsymbol{e}) = t$

$\boldsymbol{c} \leftarrow \boldsymbol{x}\boldsymbol{G}_{\text{pub}} + \boldsymbol{e}$

Déchiffrement

$\boldsymbol{x} \leftarrow \Phi_{\mathcal{F}}(\boldsymbol{G}, \boldsymbol{c})$

Original McEliece : [McE78]



A Public-Key Cryptosystem Based On Algebraic Coding Theory

R. J. McEliece

Communications Systems Research Section

Using the fact that a fast decoding algorithm exists for a general Goppa code, while no such exists for a general linear code, we construct a public-key cryptosystem which appears quite secure while at the same time allowing extremely rapid data rates. This kind of cryptosystem is ideal for use in multi-user communication networks, such as those envisioned by NASA for the distribution of space-acquired data

Exemples

Choix de $\mathcal{F}$:

- Codes de Goppa [Original McEliece] ;
- Codes MDPC-QC [BIKE] ;
- Codes basés sur la métrique rang [ROLLO] ;
- Codes de Reed Solomon Généralisés(GRS)...

Exemples

Choix de $\mathcal{F}$:

- Codes de Goppa [Original McEliece] ;
- Codes MDPC-QC [BIKE] ;
- Codes basés sur la métrique rang [ROLLO] ;
- Codes de Reed Solomon Généralisés(GRS)...

Shakers :

- Mélange de lignes ;
- Permutation de colonnes ;
- Ajout de colonnes aléatoires ...

Cryptosystème de McEliece

Génération de clefs $(n, k, t) = (\mathrm{pk}, \mathrm{sk})$

$\boldsymbol{G} \in \mathbb{F}_2^{k \times n}$, une matrice génératrice du code de Goppa binaire ;

Une matrice de permutation $\boldsymbol{P} \in \mathbb{F}_2^{n \times n}$;

Une matrice inversible $\boldsymbol{S} \in \mathbb{F}_2^{k \times k}$;

Calcule $\boldsymbol{G}_{\mathrm{pub}} = \boldsymbol{SGP}$.

$\mathrm{pk} = (\boldsymbol{G}_{\mathrm{pub}}, t)$

$\mathrm{sk} = (\boldsymbol{S}, \boldsymbol{G}, \boldsymbol{P})$

Chiffrement $(\boldsymbol{x}, \mathrm{pk}) = \boldsymbol{c}$

Encode $\boldsymbol{x} \to \boldsymbol{y} = \boldsymbol{x}\boldsymbol{G}_{\mathrm{pub}}$;

$\boldsymbol{c} = \boldsymbol{y} + \boldsymbol{e}$, avec $\boldsymbol{e}$ un vecteur d'erreur de poids $\mathrm{wt}(\boldsymbol{e}) = t$.

Déchiffrement $(\boldsymbol{c}, \mathrm{sk}) = \boldsymbol{x}$

Compute $\boldsymbol{c}^* = \boldsymbol{c}\boldsymbol{P}^{-1}$;

$\boldsymbol{c}^* = \boldsymbol{x}\boldsymbol{S}\boldsymbol{G} + \boldsymbol{e}\boldsymbol{P}^{-1}$;

$\boldsymbol{x}^* = \Phi_{\mathcal{F}}(\boldsymbol{c}^*, \boldsymbol{G})$;

Retrouver $\boldsymbol{x}$ à partir de $\boldsymbol{x}^*\boldsymbol{S}^{-1}$.

Hypothèse de sécurité

+

+

$$i\hbar \frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

[1994]

... 1011010001101

McEliece

Hypothèse de sécurité

Comment Eve peut casser le schéma de McEliece ?

1. Reconstruire $\boldsymbol{G}_{\mathrm{sk}}$ à partir de $\boldsymbol{G}_{\mathrm{pk}}$ et utilise ensuite $\Phi_{\mathcal{C}}$ pour décoder.

Indistinguabilité

$\boldsymbol{G}_{\mathrm{pk}}$ est indistinguable d'une matrice aléatoire.

Hypothèse de sécurité

Comment Eve peut casser le schéma de McEliece ?

1. Reconstruire $\boldsymbol{G}_{\mathrm{sk}}$ à partir de $\boldsymbol{G}_{\mathrm{pk}}$ et utilise ensuite $\Phi_{\mathcal{C}}$ pour décoder.

Indistinguabilité

$\boldsymbol{G}_{\mathrm{pk}}$ est indistinguable d'une matrice aléatoire.

2. Décoder avec $\boldsymbol{G}_{\mathrm{pk}}$.

Problème de décodage par syndrome

Décoder t erreurs d'un code aléatoire est un problème $\mathcal{NP}$-complete [1].

Soumission : round 1 (2018)

Figure 1 : Géographie des soumissions (source : NIST)

Soumission : round 2 (2019)

Figure 2 : Géographie des soumissions (source : NIST)

Round 3 et premier résultat (2020-2022)

- En juillet 2020, le NIST a annoncé le round 3 du processus avec 7 finalistes et 8 candidats alternatifs.

Round 3 et premier résultat (2020-2022)

- En juillet 2020, le NIST a annoncé le round 3 du processus avec 7 finalistes et 8 candidats alternatifs.

- Le 05 juillet 2022, le NIST a publié les 4 premiers vainqueurs du processus qui feront partie de la norme cryptographique post-quantique.

Round 3 et premier résultat (2020-2022)

- En juillet 2020, le NIST a annoncé le round 3 du processus avec 7 finalistes et 8 candidats alternatifs.

- Le 05 juillet 2022, le NIST a publié les 4 premiers vainqueurs du processus qui feront partie de la norme cryptographique post-quantique.

  1. Pour le chiffrement à clef publique et les algorithmes d'établissement de clef : **CRYSTALS-KYBER** (lattice).

Round 3 et premier résultat (2020-2022)

- En juillet 2020, le NIST a annoncé le round 3 du processus avec 7 finalistes et 8 candidats alternatifs.

- Le 05 juillet 2022, le NIST a publié les 4 premiers vainqueurs du processus qui feront partie de la norme cryptographique post-quantique.

  1. Pour le chiffrement à clef publique et les algorithmes d'établissement de clef : **CRYSTALS-KYBER** (lattice).

  2. Pour la signature numérique (vérifier les identités lors d'une transaction numérique ou signer un document à distance) : **CRYSTALS-Dilithium** (lattice), **FALCON** (lattice) et **SPHINCS+** (hash).

Suite du processus (2023-2024)

- À coté de ces résultats, le NIST a annoncé un round 4 du processus pour 4 algorithmes d'établissement de clef :

Suite du processus (2023-2024)

- À coté de ces résultats, le NIST a annoncé un round 4 du processus pour 4 algorithmes d'établissement de clef :

  1. **Classic McEliece** (codes) ;

Suite du processus (2023-2024)

- À coté de ces résultats, le NIST a annoncé un round 4 du processus pour 4 algorithmes d'établissement de clef :

  1. **Classic McEliece** (codes) ;
  2. **HQC** (codes) ;

Suite du processus (2023-2024)

- À coté de ces résultats, le NIST a annoncé un round 4 du processus pour 4 algorithmes d'établissement de clef :

  1. **Classic McEliece** (codes) ;
  2. **HQC** (codes) ;
  3. **BIKE** (codes) ;

Suite du processus (2023-2024)

- À coté de ces résultats, le NIST a annoncé un round 4 du processus pour 4 algorithmes d'établissement de clef :

  1. **Classic McEliece** (codes) ;
  2. **HQC** (codes) ;
  3. **BIKE** (codes) ;
  4. **SIKE** (isogénies).

Suite du processus (2023-2024)

- À coté de ces résultats, le NIST a annoncé un round 4 du processus pour 4 algorithmes d'établissement de clef :
  1. **Classic McEliece** (codes) ;
  2. **HQC** (codes) ;
  3. **BIKE** (codes) ;
  4. **SIKE** (isogénies).
- Pour diversifier les normes pour la signature post-quantique, le NIST a lancé un appel a soumission avant le 1er juin 2023.

Conclusion

1. Dynamisme de la recherche en cryptographie post-quantique ;

Conclusion

1. Dynamisme de la recherche en cryptographie post-quantique ;
2. En France, le **Plan Quantique** prévoit 150 millions d'euros dédiés à la cryptographie post-quantique ;

Conclusion

1. Dynamisme de la recherche en cryptographie post-quantique ;
2. En France, le **Plan Quantique** prévoit 150 millions d'euros dédiés à la cryptographie post-quantique ;
3. Récemment, la France a transmis son premier télégramme diplomatique en cryptographie post-quantique avec son ambassade aux États-Unis ;

Conclusion

1. Dynamisme de la recherche en cryptographie post-quantique ;
2. En France, le **Plan Quantique** prévoit 150 millions d'euros dédiés à la cryptographie post-quantique ;
3. Récemment, la France a transmis son premier télégramme diplomatique en cryptographie post-quantique avec son ambassade aux États-Unis ;
4. La **suprématie quantique** est proche (Google) ;

Conclusion

1. Dynamisme de la recherche en cryptographie post-quantique ;
2. En France, le **Plan Quantique** prévoit 150 millions d'euros dédiés à la cryptographie post-quantique ;
3. Récemment, la France a transmis son premier télégramme diplomatique en cryptographie post-quantique avec son ambassade aux États-Unis ;
4. La **suprématie quantique** est proche (Google) ;
5. Il reste beaucoup à faire !

Merci pour votre attention !