Autour de la cryptographie post-quantique

Journée commune NormaSTIC, Normandie Mathématiques

Magali Bardet, joint work with Vlad Dragoi, Ayoub Otmani, Jean-Gabriel Luque, Jean-Pierre Tillich, J. Chaulet

Laboratoire LITIS - Université de Rouen
Équipe C&A

20 mai 2016

Plan

Cryptographie asymétrique

Problèmes difficiles utilisés aujourd'hui

- les algorithmes basés sur la factorisation des nombres (RSA),
- les algorithmes basés sur le logarithme discret (El Gamal, DSA, DH, etc).

Tailles des clefs

- RSA : 2048 bits
- El Gamal sur courbes elliptiques : 256 bits.

Cryptographie asymétrique

Problèmes difficiles utilisés aujourd'hui

- les algorithmes basés sur la factorisation des nombres (RSA),
- les algorithmes basés sur le logarithme discret (El Gamal, DSA, DH, etc).

Tailles des clefs

- RSA : 2048 bits
- El Gamal sur courbes elliptiques : 256 bits.

Algorithme de Shor

Algorithme dans le modèle quantique qui résout ces problèmes efficacement.

Cryptographie post-quantique

Algorithmes cryptographique « sûrs » dans le modèle quantique

- Cryptographie basée sur les réseaux (ex : NTRU).
- Cryptographie multivariée (ex : UOV).
- Cryptographie basée sur les codes (ex : McEliece).

Cryptosystème de McEliece (1978)

Problème mathématique

- Clef privée : un code correcteur d'erreur linéaire avec un algorithme de décodage efficace (en temps polynomial) ;
- Clef publique : une base aléatoire de ce code.

Codes proposés par McEliece : les codes de Goppa.

Cryptosystème de McEliece naïf

Génération de la clef privée

- On choisit uniformément un code linéaire $\mathscr{C}$ sur $\mathbb{F}_q$, dans une famille de codes corrigeant t erreurs efficacement ;
- $\mathbf{G}$ matrice génératrice de $\mathscr{C}$ de taille $k \times n$,
  $\mathbf{P}$ matrice de permutation de taille n,
  $\mathbf{S}$ matrice inversible de taille k ;
- La clef privée est $(\mathbf{S}, \mathbf{G}, \mathbf{P})$ plus l'algorithme de décodage ;
- La clef publique est $(\mathbf{G}_{pub}, t)$ où $\mathbf{G}_{pub} = \mathbf{S} \times \mathbf{G} \times \mathbf{P}$.

Chiffrement/Déchiffrement

Chiffrement

Pour $\mathbf{m} \in \mathbb{F}_q^k$,

- tirer une erreur $\mathbf{e} \in \mathbb{F}_q^n$ de poids de Hamming t,
- chiffrer $\mathbf{c} = \mathbf{m}\mathbf{G}_{pub} + \mathbf{e}$.

Chiffrement/Déchiffrement

Chiffrement

Pour $\mathbf{m} \in \mathbb{F}_q^k$,

- tirer une erreur $\mathbf{e} \in \mathbb{F}_q^n$ de poids de Hamming t,
- chiffrer $\mathbf{c} = \mathbf{m}\mathbf{G}_{pub} + \mathbf{e}$.

Déchiffrement

- calculer $\mathbf{z} = \mathbf{c}\mathbf{P}^{-1}$,
- calculer $\mathbf{y} = Decode_{\mathbf{G}}(\mathbf{z})$,
- retourner $\mathbf{m}' = \mathbf{y}\mathbf{S}^{-1}$.

Taille des clefs

Taille de la clef publique pour une sécurité en 2^{128}

- code de Goppa : plus de 8.000.000 bits.
- code QC-MDPC : 65.000 bits.

Codes proposés

- Codes de Goppa (1978-).
- Codes GRS (1986-2014).
- sous-code d'un GRS (2005-2010).
- Codes Reed-Muller (1994-2007).
- Codes de Goppa Géométriques (1996-2014).
- Codes LDPC (2000-2008), (2008-).
- Codes de Goppa sauvages (2010-2014).
- Codes MDPC (2012-)
- Codes Polaires (2014-)

Plan

QC-MDPC codes (MTSB12)

Matrices génératrice et de parité

Code linéaire
$\mathscr{C} = \{c \in \mathbb{F}_2^n \mid \exists m \in \mathbb{F}_2^k,\ c = m \times \mathbf{G}\} = \{c \in \mathbb{F}_2^n \,|\mathbf{H}^T c = 0\}.$

Quasi-Cyclic Moderate Density Parity Check codes

- chaque ligne de la matrice de parité a un poids constant w ($w = O(\sqrt{n \log n})$) ;
- décodage par l'algorithme « bit flipping » de Gallager ;
- quasi-cyclic : la matrice de parité est cyclique par blocs.

Attaque sur la clef privée

Soit $\mathscr{C}$ un code QC-MDPC sur $\mathbb{F}_2$ de paramètres $(2p, p, w)$.

Génération de la clef

- Une matrice de parité pour $\mathscr{C}$ peut être entièrement décrite par
$$(h_1, h_2) \in (\mathbb{F}_2[x]/(x^p - 1))^2,$$
où $\|h_1\| + \|h_2\| = w$ et h_2 inversible. La clef privée est (h_1, h_2).
- La clef publique est $f = \frac{h_1}{h_2} \in \mathbb{F}_2[x]/(x^p - 1)$.

p est premier.

Attaque sur la clef privée

Problème de Reconstruction Rationnelle

Étant donné $f \in \mathbb{F}_2[x]$ avec $\deg(f) < p$, trouver $(\varphi, \psi) \in \mathbb{F}_2[x]^2$ tels que

$$f = \frac{\varphi}{\psi} \mod x^p - 1, \quad \deg(\psi) < r, \quad \deg(\varphi) \leqslant p - r.$$

Algorithme d'Euclide Étendu appliqué à $x^p - 1$ et f. Complexité quadratique, sous-quadratique (Knuth, Schönhage 1971).

Attaque sur la clef privée

Problème de Reconstruction Rationnelle

Étant donné $f \in \mathbb{F}_2[x]$ avec $\deg(f) < p$, trouver $(\varphi, \psi) \in \mathbb{F}_2[x]^2$ tels que

$$f = \frac{\varphi}{\psi} \mod x^p - 1, \quad \deg(\psi) < r, \quad \deg(\varphi) \leqslant p - r.$$

Algorithme d'Euclide Étendu appliqué à $x^p - 1$ et f. Complexité quadratique, sous-quadratique (Knuth, Schönhage 1971).
Combien de clefs sont attaquables par cet algorithme ?

Comptage des Clefs faibles BDLO '16

Soit $\mathscr{C}$ un code QC-MDPC de paramètres $(2p, p, w)$ où $w = w_1 + w_2$ (w_i impair).

Notations

- $\mathscr{P}_{\omega_1,\omega_2} = \left\{(h_1, h_2) \in (\mathbb{K}[x]/(x^p - 1))^2 \mid \|h_i\| = \omega_i \text{ impairs}\right\}$.
- $\mathscr{P}_\omega = \bigcup_{\omega_1+\omega_2=\omega} \mathscr{P}_{\omega_1,\omega_2}$.
- $\mathscr{W}_\omega = \{(h_1, h_2) \in \mathscr{P}_\omega : \deg(h_1) + \deg(h_2) < p\}$.
- $\mathscr{W}_{\omega_1,\omega_2} = \mathscr{W}_\omega \cap \mathscr{P}_{\omega_1,\omega_2}$.

Comptage des Clefs faibles BDLO '16

Soit $\mathscr{C}$ un code QC-MDPC de paramètres $(2p, p, w)$ où $w = w_1 + w_2$ (w_i impair).

Notations

- $\mathscr{P}_{\omega_1,\omega_2} = \left\{(h_1, h_2) \in (\mathbb{K}[x]/(x^p - 1))^2 \mid \|h_i\| = \omega_i \text{ impairs}\right\}$.
- $\mathscr{P}_\omega = \bigcup_{\omega_1+\omega_2=\omega} \mathscr{P}_{\omega_1,\omega_2}$.
- $\mathscr{W}_\omega = \{(h_1, h_2) \in \mathscr{P}_\omega : \deg(h_1) + \deg(h_2) < p\}$.
- $\mathscr{W}_{\omega_1,\omega_2} = \mathscr{W}_\omega \cap \mathscr{P}_{\omega_1,\omega_2}$.

Comptage naïf et asymptotique quand $n \to \infty$

$\#\mathscr{W}_{\omega_1,\omega_2} = \binom{p+1}{\omega}$ et $\#\mathscr{P}_{\omega_1,\omega_2} = \binom{p}{w_1}\binom{p}{w_2}$.

$\#\mathscr{W}_\omega / \#\mathscr{P}_\omega = \frac{w}{2^w p^{c/2}}(1 + o(1))$ si $\frac{w^2}{2p} = c \log p + O(\sqrt{\log p / p})$.

Comptage des Clefs faibles BDLO '16

Si $f = \frac{h_1}{h_2} \mod x^p - 1$ **alors**

$x^{i-j} f = \frac{x^i h_1}{x^j h_2} \mod x^p - 1$ pour tous i, j.

Comptage des Clefs faibles BDLO '16

Si $f = \frac{h_1}{h_2} \mod x^p - 1$ alors

$x^{i-j} f = \frac{x^i h_1}{x^j h_2} \mod x^p - 1$ pour tous i, j.

Exemple

$p = 7$, $f = \frac{x + x^5}{x^2 + x^4 + x^5}$ et $x^4 f = \frac{x^2 h_1}{x^5 h_2} = \frac{1 + x^3}{1 + x^2 + x^3}$.

Comptage des Clefs faibles BDLO '16

Si $f = \frac{h_1}{h_2} \mod x^p - 1$ alors

$x^{i-j} f = \frac{x^i h_1}{x^j h_2} \mod x^p - 1$ pour tous i, j.

Exemple

$p = 7$, $f = \frac{x+x^5}{x^2+x^4+x^5}$ et $x^4 f = \frac{x^2 h_1}{x^5 h_2} = \frac{1+x^3}{1+x^2+x^3}$.

Une clef est attaquable si l'un de ses shifts l'est, où si le shift correspondant au *mot de Lyndon* associé l'est.

Combinatoire

Compte du nombre de clefs faibles

Compter $L^k(p, w)$, le nombre de mots de Lyndon de longueur p, de poids w et de plus grande plage de 0 de taille k ?

Biblio

- 1961, Gilbert et Riordan comptent les mots de Lyndon de longueur p et poids w.
- Approche probabiliste de Feller et Schilling, Gordon, Waterman en 1986 sur des mots (non de Lyndon) ;
- Approche combinatoire de Bassino, Clément, Nicaud en 2005 mais sans fixer le poids.

Distribution

BDLO16, p premier

$$L^{\leqslant k}(p,w) = \frac{1}{w}\binom{w}{p-w}_k$$

où $\binom{i}{j}_k = [x^i](1+x+\cdots+x^k)^j$ coefficient de Pascal-De Moivre.

Distribution

Notations

X_{p,w_1} variable aléatoire représentant la plus grande plage de 0 d'un mot de Lyndon choisit uniformément parmi les mots de taille p et de poids w et $Y_{p,w_1,w_2} = X_{p,w_1} + X_{p,w_2}$.

$$P(Y_{p,\omega_1,\omega_2} \geqslant p-1) \sim \omega_1\omega_2 \frac{\binom{p-1}{\omega-2}}{\binom{p-1}{\omega_1-1}\binom{p-1}{\omega_2-1}} \quad \textit{quand } p \to \infty \qquad (1)$$

Gain

$$P(Y_{p,\omega_1,\omega_2} \geqslant p-1) \sim \omega^2 \times \frac{\#\mathscr{W}_{\omega_1,\omega_2}}{\#\mathscr{P}_{\omega_1,\omega_2}}$$

Autre action

$$\alpha \cdot \left(\sum_{i=0}^{p-1} a_i x^i \right) = \sum_{i=0}^{p-1} a_i x^{\alpha i}.$$

Résultats numériques

Figure : $\omega_1 + \omega_2 = \omega$.

Security level	p	$\frac{\omega}{2}$	$\frac{\lvert\mathscr{W}_\omega\rvert}{\lvert\mathscr{P}_\omega\rvert}$ exact value	$P(Y_{p,\omega} \geqslant p-1)$ upper bound	$P([Y_{p,\omega}] \geqslant p-1)$ upper bound
80	4801	45	2^{-84}	2^{-71}	2^{-60}
	3593	51	2^{-96}	2^{-83}	2^{-72}
	3079	55	2^{-105}	2^{-91}	2^{-80}
128	9857	71	2^{-136}	2^{-121}	2^{-109}
	7433	81	2^{-156}	2^{-141}	2^{-129}
	6803	85	2^{-164}	2^{-149}	2^{-137}

Résultats numériques

Figure : $\omega_1 + \omega_2 = \omega$.

Security level	p	$\frac{\omega}{2}$	$\frac{\lvert\mathscr{W}_\omega\rvert}{\lvert\mathscr{P}_\omega\rvert}$ exact value	$P(Y_{p,\omega} \geqslant p-1)$ upper bound	$P([Y_{p,\omega}] \geqslant p-1)$ upper bound
80	4801	45	2^{-84}	2^{-71}	2^{-60}
	3593	51	2^{-96}	2^{-83}	2^{-72}
	3079	55	2^{-105}	2^{-91}	2^{-80}
128	9857	71	2^{-136}	2^{-121}	2^{-109}
	7433	81	2^{-156}	2^{-141}	2^{-129}
	6803	85	2^{-164}	2^{-149}	2^{-137}

Plan

Les codes polaires

- Très bonne capacité de correction.
- Algorithme de décodage efficace (Arikan 2009).
- Structure algébrique ?

Définition des codes polaires

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Définition des codes polaires

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

$$\mathbf{G}_m \stackrel{\text{def}}{=} \underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}_{m \text{ times}}.$$

Définition des codes polaires

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

$$\mathbf{G}_m \stackrel{\text{def}}{=} \underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}_{m \text{ times}}.$$

- le code polaire de longueur $n = 2^m$ et dimension k est un ensemble de k lignes déterminées de $\mathbf{G}_m$.
- Le code de Reed-Muller d'ordre r, $\mathscr{R}(r,m)$ est le code polaire où l'on choisit les lignes de poids $\geqslant 2^{m-r}$.

Les codes polaires comme codes monomiau

Les codes de Reed-Muller peuvent être vus comme des codes d'évaluation :

- $\mathbf{R}_m = \mathbb{F}_2[x_0, \ldots, x_{m-1}]/(x_0^2 - x_0, \ldots, x_{m-1}^2 - x_{m-1})$.
- Les mots de code sont les

$$ev(g) = (g(a_0, \ldots, a_{m-1}))_{(a_0,\ldots,a_{m-1}) \in \mathbf{R}_m^m} .$$

  pour $g \in \mathbf{R}_m$.
- Le code $\mathscr{R}(r, m)$ est engendré par les $\{ev(m) : m \in \mathscr{M}_m\}$.

Ordre monomial partiel

Définition

- On ordonne les monômes par

$$x_{i_1} \cdots x_{i_s} \preceq x_{j_1} \cdots x_{j_s}$$

  ssi pour tout l, $i_l \leqslant j_l$ (avec $i_1 < \cdots < i_s$ et $j_1 < \cdots < j_s$).
- On étend par divisibilité : $f \preceq g$ ssi il existe un diviseur g^* de g de même degré que f tel que $f \preceq g^*$.

Decreasing Monomial Code

1

Decreasing Monomial Code

$x_0 \longleftarrow 1$

Decreasing Monomial Code

$$x_1 \longleftarrow x_0 \longleftarrow 1$$

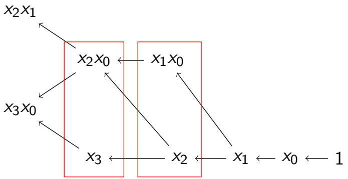
Decreasing Monomial Code

$$\begin{tikzcd}
 & x_1x_0 & & \\
x_2 & x_1 \arrow[l] \arrow[u] & x_0 \arrow[l] & 1 \arrow[l]
\end{tikzcd}$$

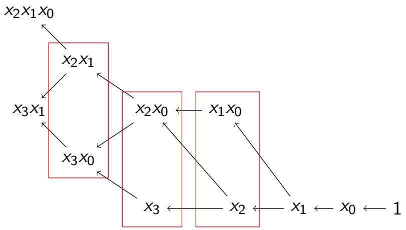
Decreasing Monomial Code

$$\begin{tikzcd}
x_2x_0 & x_1x_0 \arrow[l] & & & \\
x_3 & x_2 \arrow[l] \arrow[ul] & x_1 \arrow[l] \arrow[ul] & x_0 \arrow[l] & 1 \arrow[l]
\end{tikzcd}$$

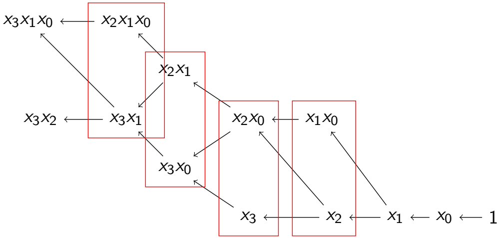
Decreasing Monomial Code



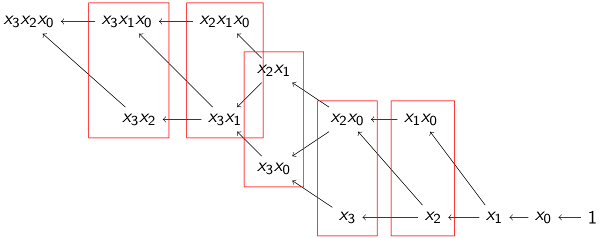
Decreasing Monomial Code



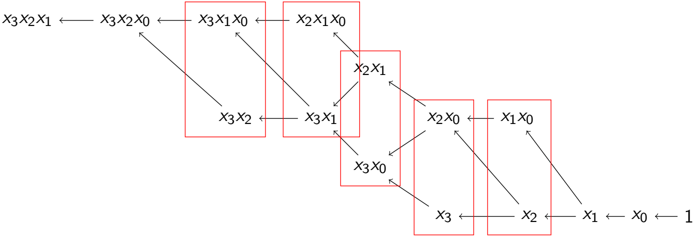
Decreasing Monomial Code



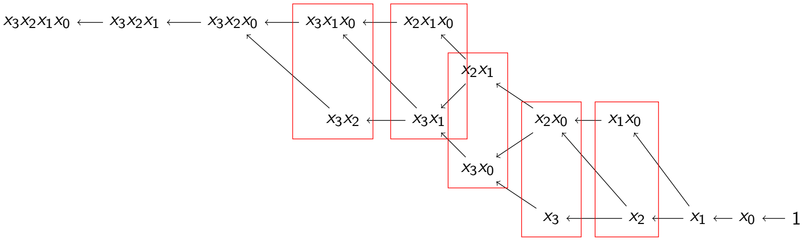
Decreasing Monomial Code



Decreasing Monomial Code



Decreasing Monomial Code



Codes Monomiaux Décroissants

Ensemble décroissant

Un ensemble $I \subseteq \mathscr{M}_m$ est décroissant si

$$f \in I \text{ et } g \preceq f \Longrightarrow g \in I.$$

Codes Monomiaux Décroissants

Ensemble décroissant

Un ensemble $I \subseteq \mathscr{M}_m$ est décroissant si

$$f \in I \text{ et } g \preceq f \Longrightarrow g \in I.$$

Code monomial décroissant

- Un code linéaire d'évaluation est définit par un ensemble I de polynômes et vérifie $\mathscr{C}(I) = Vect(\{ev(f) \mid f \in I\})$.
- Si $I \subseteq \mathscr{M}_m$, on dit que le code est monomial.
- Si de plus I est décroissant, on dit que le code est monomial décroissant.

Propriétés des codes monomiaux décroissants

Théorème (BDTO, PQcrypto 2016)

- Les codes polaires sont des codes monomiaux décroissants.
- Le dual d'un code monomial décroissant est un code monomial décroissant.
- Leur groupe de permutation contient $\mathrm{LTA}(m,2)$ l'ensemble des transformations affines de la forme $x \to Ax + b$ où A est une matrice triangulaire inférieure binaire avec des '1' sur la diagonale.
- On peut compter le nombre de mots de code de poids minimum.

Construction d'un distingueur

Définitions (codes poinçonnés et raccourcis)

- $\mathscr{P}_{\mathscr{I}}(\mathscr{C}) \stackrel{\text{def}}{=} \left\{ (c_i)_{i \notin \mathscr{I}} \mid \mathbf{c} \in \mathscr{C} \right\}$;
- $\mathscr{S}_{\mathscr{I}}(\mathscr{C}) \stackrel{\text{def}}{=} \left\{ (c_i)_{i \notin \mathscr{I}} \mid \exists \mathbf{c} = (c_i)_i \in \mathscr{C} \text{ tq } \forall i \in \mathscr{I}, \, c_i = 0 \right\}$.

Attaque par signature

- Si r est le degré maximum d'un monôme définisant $\mathscr{C}$, alors I contient $x_0 \dots x_{r-1} \in I$.
- Le mot de code associé est de poids minimum.
- On choisit $\mathbf{c}$ un mot de code de poids minimum dans $\mathscr{C}^\pi$.
- On calcule une signature : $(\dim(\mathscr{S}_{\mathrm{supp}(\mathbf{c})}(\mathscr{C})^\perp), W_{min}(\mathscr{S}_{\mathrm{supp}(\mathbf{c})}(\mathscr{C})^\perp)$ pour déterminer les mots de $\mathscr{C}^\pi$ correspondants à $x_0 \dots x_{r-1}$ et ses translatés.
- On trouve la permutation sur le support de $x_0 \dots x_{r-1}$.
- On poursuit par induction pour les monômes $x_0 \dots x_i$ avec $i \leqslant r$.

Conclusion

- Des outils mathématiques appliqués à la cryptographie.
- Codes MDPC : étendre l'attaque.
- Étude des sous-codes des codes polaires.
- Problème de l'équivalence de codes : point de vue algébrique.

Références

- M. Bardet, J. Chaulet, V. Dragoi, A. Otmani, and J.-P. Tillich, *Cryptanalysis of the McEliece Public Key Cryptosystem Based on Polar Codes*, PQCrypto 2016.
- M. Bardet, V. Dragoi, J.-G. Luque, A. Otmani, *Weak Keys for the Quasi-Cyclic MDPC Public Key Encryption Scheme*, Africacrypt 2016.
- M. Bardet, V. Dragoi, A. Otmani, and J.-P. Tillich, *Algebraic Properties of Polar Codes From a New Polynomial Formalism*, ISIT 2016.