Université Paris 13 Année 2021–2022

M2 Mathématiques des données

Codes correcteurs – Feuille de TD 4

10 décembre 2021

Exercice 1. Attaque de Sidelnikov-Shestakov.

Dans cet exercice, on propose d'étudier l'attaque de Sidelnikov et Shestakov [1] sur une variation du cryptosystème de McEliece employant des codes de Reed-Solomon généralisés.

Originellement, un code de Reed-Solomon généralisé peut être introduit comme suit. On considère $\boldsymbol{x} = (x_1, \dots, x_n)$ des éléments deux-à-deux distincts de $\mathbb{F}_q$, et $\boldsymbol{y} = (y_1, \dots, y_n)$ des éléments non-nuls de $\mathbb{F}_q$. Ensuite, on définit

$$\mathrm{GRS}_k(\boldsymbol{x}, \boldsymbol{y}) = \{(y_1 f(x_1), \dots, y_n f(x_n)) \mid f \in \mathbb{F}_q[X], \deg f < k\} \subseteq \mathbb{F}_q^n .$$

Par la suite, on notera $\mathrm{ev}_{\boldsymbol{x},\boldsymbol{y}}(f) \coloneqq (y_1 f(x_1), \dots, y_n f(x_n))$.

Les codes de Reed-Solomon généralisés peuvent également être présentés comme des codes géométriques. On note $\mathbb{P}^1(\mathbb{F}_q)$ l'ensemble des points de la droite projective. Étant donné un ensemble de points $\mathcal{P} = \{P_1, \dots, P_n\} \subset \mathbb{P}^1(\mathbb{F}_q)$ et un diviseur D de $\mathbb{P}^1$, le code de Reed-Solomon généralisé est :

$$\mathcal{C}_{\mathbb{P}^1}(\mathcal{P}, D) = \{\mathrm{ev}_{\mathcal{P}}(f) \mid f \in L(D)\} .$$

On note $\mathcal{H}$ le groupe des homographies, autrement dit des automorphismes de la droite projective :

$$\mathcal{H} = \left\{ \sigma : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} \;\middle|\; ad - bc \neq 0 \right\}$$

Question 1.– Soit $\sigma \in \mathcal{H}$. Démontrer que $\mathcal{C}_{\mathbb{P}^1}(\mathcal{P}, D) = \mathcal{C}_{\mathbb{P}^1}(\sigma(\mathcal{P}), \sigma(D))$.

Dans une variation du schéma de McEliece employant des codes de Reed–Solomon, Alice choisit aléatoirement $\boldsymbol{x}$ et $\boldsymbol{y}$, et publie $\boldsymbol{G} \in \mathbb{F}_q^{k \times n}$, une matrice génératrice de $\mathcal{C} = \mathrm{GRS}_k(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{C}_{\mathbb{P}^1}(\mathcal{P}, D)$. Les vecteurs $\boldsymbol{x}$ et $\boldsymbol{y}$ sont gardés secrets par Alice.

On rappelle que la connaissance de $(\boldsymbol{x}', \boldsymbol{y}')$ tels que $\mathcal{C} = \mathrm{GRS}_k(\boldsymbol{x}', \boldsymbol{y}')$ permet de décoder n'importe quel message chiffré à l'aide de la clé publique $\boldsymbol{G}$. Une attaque sur la clé correspond donc à retrouver, à partir de $\boldsymbol{G}$ uniquement, des vecteurs $\boldsymbol{x}'$, $\boldsymbol{y}'$ (ou de manière équivalente, $\mathcal{P}'$ et D'), tels que

$$\mathrm{GRS}_k(\boldsymbol{x}', \boldsymbol{y}') = \mathcal{C} .$$

Notons qu'il n'est pas nécessaire retrouver exactement le couple $(\boldsymbol{x}, \boldsymbol{y})$ engendré pae Alice.

Question 2.– Déduire de la **Question 1** que l'on peut supposer que l'on connaît déjà trois des points de $\boldsymbol{x}$ choisis par Alice.

Question 3.– Si $n - k \leq 3$, donner une borne sur la capacité de correction du code $\mathcal{C}$. En déduire que le cryptosystème de McEliece n'est pas sûr dans ce cas.

Question 4.– Si $k = 1$, démontrer que l'on peut retrouver aisément $(\boldsymbol{x}', \boldsymbol{y}')$ tels que $\mathrm{GRS}_k(\boldsymbol{x}', \boldsymbol{y}') = \mathcal{C}$.

Dorénavant, on suppose donc que $k \geq 2$, $n - k \geq 4$ et $x'_1 = 0$, $x'_2 = 1$ et $x'_3 = \infty$. On rappelle également que pour tout code MDS, tout ensemble $I \subset [1, n]$ de cardinal k est un ensemble d'information du code.

On peut maintenant passer à l'attaque. Pour cela, on rappelle que l'on suppose que l'on connaît uniquement $\boldsymbol{G}$.

Question 5.– Comment peut-on calculer deux mots non-collinéaires $\boldsymbol{u}$ et $\boldsymbol{v}$ de $\mathcal{C}$ qui partagent $k - 2$ zéros en commun sur leurs $k - 2$ dernières coordonnées ? Démontrer que l'on peut également imposer $u_{n-k+1} = 0$ et $v_{n-k+2} = 0$.

Question 6.– Soient f et g tels que $\boldsymbol{u} = \mathrm{ev}_{\boldsymbol{x}',\boldsymbol{y}'}(f)$ et $\boldsymbol{v} = \mathrm{ev}_{\boldsymbol{x}',\boldsymbol{y}'}(g)$ où $\boldsymbol{u}$ et $\boldsymbol{v}$ sont déterminés à la question précédente. Que dire du degré de la fraction rationnelle f/g ? Connaît-on ses évaluations sur certains points x'_i ?

Question 7.– Démontrer que l'on peut déterminer exactement f/g par la résolution d'un système linéaire de trois équations à quatre inconnues, dont on extrait une solution non-nulle.

Question 8.– En déduire que l'on peut maintenant retrouver des valeurs de x'_i sur $n - k - 3$ nouveaux points que l'on déterminera. En déduire que l'on peut retrouver tous les points $\boldsymbol{x}'$.

Question 9.– On suppose maintenant qu'un vecteur $\boldsymbol{x}'$ a été déterminé. Comment peut-on trouver $\boldsymbol{y}'$ tel que $\mathrm{GRS}_k(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{GRS}_k(\boldsymbol{x}', \boldsymbol{y}')$?

Question 10.– Conclure avec un algorithme qui attaque la variante du cryptosystème de McEliece employant des codes de Reed-Solomon. On donnera sa complexité.

Exercice 2. Dual de codes de Reed–Solomon généralisés.

Soient $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$ deux à deux distincts, et $\boldsymbol{y} = (y_1, \ldots, y_n) \in (\mathbb{F}_q^\times)^n$. Le code de Reed–Solomon généralisé, de dimension $k \in \{0, \ldots, n\}$, de points d'évaluation $\boldsymbol{x}$ et de multiplieurs $\boldsymbol{y}$ est défini comme :

$$\mathrm{GRS}_k(\boldsymbol{x}, \boldsymbol{y}) = \{(y_1 f(x_1), \ldots, y_n f(x_n)) \mid f \in \mathbb{F}_q[X], \deg f \leq k - 1\}.$$

Question 1.– Démontrer que $\mathrm{GRS}_k(\boldsymbol{x}, \boldsymbol{y})$ est un code MDS.

Dans cet exercice, on souhaite démontrer que le dual d'un code GRS est également un code GRS (avec d'autres paramètres).

Pour cela, on commence par étudier le cas où $\boldsymbol{y} = \mathbf{1} \coloneqq (1, \ldots, 1) \in \mathbb{F}_q^n$.

Question 2.– Calculer $\sum_{x \in \mathbb{F}_q} x^i$ pour tout $i \in \{0, \ldots, q - 1\}$.

Question 3.– En déduire que si $\boldsymbol{x}'$ est de longueur $n = q$ (c'est-à-dire, s'il est constitué de tous les éléments de $\mathbb{F}_q$, alors $\mathrm{GRS}_k(\boldsymbol{x}', \mathbf{1}) = \mathrm{GRS}_{n-k}(\boldsymbol{x}', \mathbf{1})$.

Si $\boldsymbol{x}$ n'est pas de longueur q, alors le code $\mathrm{GRS}_k(\boldsymbol{x}, \mathbf{1}) \subseteq \mathbb{F}_q^n$ peut être vu comme le poinçonnement de $\mathrm{GRS}_k(\boldsymbol{x}', \mathbf{1}) \subseteq \mathbb{F}_q^q$, où $\boldsymbol{x}'$ est constitué de tous les éléments de $\mathbb{F}_q$.

On rappelle également que pour tout code $\mathcal{C} \subseteq \mathbb{F}_q^n$ et pour tout $I \subset [1, n]$, on a $\mathrm{Punct}_I(\mathcal{C})^\perp = \mathrm{Short}_I(\mathcal{C}^\perp)$.

Soit $\mathcal{A}(\boldsymbol{x}) \coloneqq \{a \in \mathbb{F}_q \mid \forall i \in [1, n], x_i \neq a\}$.

Question 4.– Démontrer que $\mathrm{GRS}_k(\boldsymbol{x}, \mathbf{1})^\perp = \mathrm{Short}_I(\mathrm{GRS}_{q-k}(\boldsymbol{x}', \mathbf{1}))$ où $I = \{i \in [1, q] \mid x'_i \in \mathcal{A}(\boldsymbol{x})\}$.

Question 5.– Démontrer que pour tout $f \in \mathbb{F}_q[X]$ de degré $\leq q - 1$, on a :

$$(f(x_1), \ldots, f(x_n)) \in \mathrm{GRS}_k(\boldsymbol{x}, \mathbf{1})^\perp \quad \iff \quad \prod_{a \in \mathcal{A}(\boldsymbol{x})} (X - a) \quad \text{divise } f(X).$$

Question 6.– En déduire que $\mathrm{GRS}_k(\boldsymbol{x}, \mathbf{1})^\perp = \mathrm{GRS}_{n-k}(\boldsymbol{x}, \boldsymbol{z})$ où

$$\boldsymbol{z} = (g(x_1), \ldots, g(x_n)) \quad \text{et} \quad g(X) = \prod_{a \in \mathcal{A}(\boldsymbol{x})} (X - a)\,.$$

On revient au cas général, où $\boldsymbol{y} \in (\mathbb{F}_q^\times)^n$ est potentiellement différent de $\mathbf{1}$.

Question 7.– Soit $\mathcal{C} \subseteq \mathbb{F}_q^n$ un code. Pour $\boldsymbol{y} \in (\mathbb{F}_q^\times)^n$, on note $\boldsymbol{y} \star \mathcal{C} \coloneqq \{(y_1 c_1, \ldots, y_n c_n) \mid \boldsymbol{c} \in \mathcal{C}\} \subseteq \mathbb{F}_q^n$.
Démontrer que $(\boldsymbol{y} \star \mathcal{C})^\perp = \boldsymbol{y}^{-1} \star \mathcal{C}^\perp$ où $\boldsymbol{y}^{-1} \coloneqq (y_1^{-1}, \ldots, y_n^{-1})$.

Question 8.– En déduire une expression de $\mathrm{GRS}_k(\boldsymbol{x}, \boldsymbol{y})^\perp$ comme un code GRS.

Références

[1] V.M. Sidelnikov and S.O. Shestakov. On cryptosystems based on generalized Reed-Solomon codes. *Discrete Mathematics*, 4(3) :57–63, 1991.