Code-Based Cryptography – Exercises

Summer School on Post-Quantum Cryptography 2017

TU Eindhoven

1. Decoding *vs* Syndrome Decoding [4]. Let $H \in \mathbf{F}_q^{(n-k)\times n}$ be a (full rank) matrix. Let $\mathcal{C} = \langle H \rangle^{\perp}$ be the linear $[n, k]$ code $\mathcal{C}$ over $\mathbf{F}_q$ of parity check matrix H. Let d be the minimum distance of $\mathcal{C}$ and let $t < d/2$. A t-bounded decoder for $\mathcal{C}$ is a mapping $\phi : \mathbf{F}_q^n \to \mathcal{C}$ such that for all $y \in \mathbf{F}_q^n$ and all $x \in \mathcal{C}$

$$|x - y| \leq t \Rightarrow \phi(y) = x$$

A t-bounded H-syndrome decoder is a mapping $\psi : \mathbf{F}_q^{n-k} \to \mathbf{F}_q^n$ such that for all $e \in \mathbf{F}_q^n$

$$|e| \leq t \Rightarrow \psi(eH^T) = e$$

1.a) Prove that there exists a polynomial time t-bounded decoder for $\mathcal{C} = \langle H \rangle^{\perp}$ if and only if there exists a polynomial time t-bounded H-syndrome decoder.

1.b) Prove that for any code $\mathcal{C}$, the McEliece and the Niederreiter public key encryption schemes using $\mathcal{C}$ as public code are equally secure. First restate the question in terms of adversary, success probability, and running time.

2. Resend Attack [1]. We consider an instance of McEliece using as public key the generator matrix $G \in \mathbf{F}_2^{k\times n}$ and errors of Hamming weight t. The same cleartext x is encrypted twice

$$\begin{aligned} x &\longmapsto y_1 = xG + e_1, |e_1| = t \\ x &\longmapsto y_2 = xG + e_2, |e_2| = t \end{aligned}$$

We denote $e_1 * e_2$ the component wise product of e_1 and e_2.

2.a) Assuming e_1 and e_2 are drawn uniformly and independently, what are the expected values of $|e_1 + e_2|$ and $|e_1 * e_2|$?

Let $\tilde{\cdot}$ denote the operation of removing the coordinates indexed by the non-zero positions of $e_1 + e_2$ (applied to a matrix or a vector).

2.b) Show that, given y_1 and y_2, recovering x can be achieved by decoding $t' = |e_1 * e_2|$ errors in a binary linear code of length $n - |e_1 + e_2|$ and dimension k. *Hint: remark that* $\widetilde{e_1 * e_2} = \tilde{e}_1 = \tilde{e}_2$.

2.c) Give an estimate for the cost for decrypting a resent message in the average case. How does it compare with the "normal" cost.
In this exercise, we will assume that the cost for generic decoding of t *errors in a binary* $[n, k]$ *code is* $2^{t \log_2 \frac{n}{n-k}}$.
Numerical application: $(n, k, t) \in \{(1024, 524, 50), (2048, 1608, 40), (4096, 3496, 50)\}$.

3. Reaction Attack [2]. We consider an instance of McEliece using as public key the generator matrix $G \in \mathbf{F}_2^{k \times n}$ and errors of Hammnig weight t. We have access (for a cost 1) to the following oracle, defined for all $y \in \mathbf{F}_2^n$,

$$\Theta_G(y) = \begin{cases} \text{TRUE} & \text{if } \operatorname{dist}(y, \langle G \rangle) \leq t, \\ \text{FALSE} & \text{else.} \end{cases}$$

This oracle correspond to a real life scenario in which the adversary is allowed to modify a ciphertext and to check whether or not the decryption device consider it as valid.

3.a) Build a polynomial time t-bounded decoder for $\langle G \rangle$ from Θ_G. How many calls to Θ_G are needed?

3.b) Same questions with the oracle

$$\Theta'_G(y) = \begin{cases} \text{TRUE} & \text{if } \operatorname{dist}(y, \langle G \rangle) = t, \\ \text{FALSE} & \text{else.} \end{cases}$$

4. Codeword Finding. We consider Lee & Brickell [3] variant of Prange [6] algorithm. Let $G \in \mathbf{F}_2^{k \times n}$ be a generator matrix of a code $\mathcal{C}$. Given an integer $w > 0$, the problem is to find $c \in \mathcal{C}$ such that $|c| = w$. The algorithm uses an integer parameter $p > 0$.

repeat

1: pick a $n \times n$ permutation matrix P and compute

$$G' = UGP = \left(\begin{array}{ccc|c} 1 & & & \\ & \ddots & & R \\ & & 1 & \end{array} \right)$$

2: for all $x \in \mathbf{F}_2^k$ of Hamming weight p

if $|xR| = w - p$

return xUG

4.a) Prove that if the algorithm stops it solves the target problem. What is its average cost, denoted $\mathrm{WF}_{\mathrm{LB}}(n, k, w)$, assuming there is a single word of weight w in the code? You may use an unspecified polynomial factor (in n), but give an indication on its degree (plus or minus 1).

Binomial coefficient can be accurately estimated. For all λ, $0 < \lambda < 1$, and sufficiently large n, we have

$$\frac{1}{\sqrt{8}} \frac{2^{nh(\lambda)}}{\sqrt{\lambda(1-\lambda)n}} \leq \binom{n}{\lambda n} \leq \frac{1}{\sqrt{2\pi}} \frac{2^{nh(\lambda)}}{\sqrt{\lambda(1-\lambda)n}},$$

where $h(x) = -x \log_2 x - (1 - x) \log_2(1 - x)$ is the binary entropy function.

4.b) *Asymptotic Analysis.* We assume a choice of p such that

$$\mathrm{WF}_{\mathrm{LB}}(n, k, w) = Q(n) \frac{\binom{n}{w}}{\binom{n-k}{w}}, \deg(Q) \text{ is 2 or 3.}$$

Let $\Omega = w/n$ and $R = k/n$, using $h(\cdot)$, give an expression for $c(\Omega, R)$ such that $\mathrm{WF}_{\mathrm{LB}}(n, k, w)$ equals $2^{c(\Omega,R)w}$ up to a polynomial factor. Prove that

$$\lim_{\Omega \to 0} c(\Omega, R) = \log_2 \frac{1}{1 - R}$$

5. Finding Codewords in a Cyclosymmetric Code [5]. We keep the context and notations of the previous question. We assume the existence of a word $c = (c_0, \ldots, c_{n-1}) \in \mathcal{C}$ of small weight w such that $c_{2i} = c_{2i+1}$, $0 \leq i < n/2$.

5.a) Adapt the Lee & Brickell procedure to search the codeword c of weight w with duplicate coordinates as defined above. Estimate the cost. You should be able to essentially divide the exponent by 2.
Hint: only allow particular permutations in step **1:**.

5.b) A binary $[n, k]$ code is cyclosymmetric if it admits a generator matrix formed of $r \times r$ blocks that both circulant and symmetric. If $r \geq 2$ and if a cyclosymmetric generator matrix of the code admits a row of weight w, show that this row can be recovered for a cost which do not exceeds $\sqrt{\mathrm{WF}_{\mathrm{LB}}(n, k, w)}$ up to a polynomial factor.

References

[1] T. Berson. Failure of the McEliece public-key cryptosystem under message-resend and related-message attack. In B. Kalisky, editor, *Advances in Cryptology - CRYPTO '97*, volume 1294 of *LNCS*, pages 213–220. Springer, 1997.

[2] K. Kobara and H. Imai. Countermeasure against reaction attacks (in japanese). In *The 2000 Symposium on Cryptography and Information Security : A12*, January 2000.

[3] P.J. Lee and E.F. Brickell. An observation on the security of McEliece's public-key cryptosystem. In C.G. Günther, editor, *Advances in Cryptology - EUROCRYPT '88*, volume 330 of *LNCS*, pages 275–280. Springer, 1988.

[4] Y.X. Li, R.H. Deng, and X.M. Wang. On the equivalence of McEliece's and Niederreiter's public-key cryptosystems. *IEEE Transactions on Information Theory*, 40(1):271–273, January 1994.

[5] Ray A. Perlner. Optimizing information set decoding algorithms to attack cyclosymmetric MDPC codes. In *PQCrypto 2014*, volume 8772 of *LNCS*, pages 220–228. Springer, 2014.

[6] E. Prange. The use of information sets in decoding cyclic codes. *IRE Transactions*, IT-8:S5–S9, 1962.