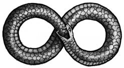
OUROBOROS-R, an IND-CPA KEM based on Rank Metric

NIST First Post-Quantum Cryptography Standardization Conference



Carlos AguilarMelchor[2] Nicolas Aragon[1] Slim Bettaieb[5]
Loic Bidoux[5] Olivier Blazy[1] Jean-Christophe Deneuville[1,4]
Philippe Gaborit[1] Adrien Hauteville[1] Gilles Zémor[3]

[1]University of Limoges, XLIM-DMI, France ; [2]ISAE-SUPAERO, Toulouse, France

[3]IMB, University of Bordeaux; [4]INSA-CVL, Bourges, France ; [5]Worldline, France.

Rank Metric

We only consider codes with coefficients in $\mathbb{F}_{q^m}$.
Let $\beta_1, \ldots, \beta_m$ be a basis of $\mathbb{F}_{q^m}/\mathbb{F}_q$. To each vector $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ we can associate a matrix $\boldsymbol{M_x}$

$$\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{F}_{q^m}^n \leftrightarrow \boldsymbol{M_x} = \begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \ldots & x_{mn} \end{pmatrix} \in \mathbb{F}_q^{m \times n}$$

such that $x_j = \sum_{i=1}^m x_{ij}\beta_i$ for each $j \in [1..n]$.

Definition

$d_R(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{Rank}(\boldsymbol{M_x} - \boldsymbol{M_y})$ and $|\boldsymbol{x}|_r = \mathrm{Rank}\, \boldsymbol{M_x}$.

Support of a Word

Definition

The support of a word is the $\mathbb{F}_q$-subspace generated by its coordinates:

$$\mathrm{Supp}(\boldsymbol{x}) = \langle x_1, \ldots, x_n \rangle_{\mathbb{F}_q}$$

Number of supports of weight w:

Rank	Hamming
$\begin{bmatrix} m \\ w \end{bmatrix}_q \approx q^{w(m-w)}$	$\binom{n}{w} \leqslant 2^n$

Complexity in the worst case:

- quadratically exponential for Rank Metric
- simply exponential for Hamming Metric

LRPC Codes

Definition

Let $\boldsymbol{H} \in \mathbb{F}_{q^m}^{(n-k)\times n}$ a full-rank matrix such that the dimension d of $\langle h_{ij} \rangle_{\mathbb{F}_q}$ is small.
By definition, $\boldsymbol{H}$ is a parity-check matrix of an $[n, k]_{q^m}$ LRPC code.
We say that d is the weight of the matrix $\boldsymbol{H}$.

A LRPC code can decode errors (recover support) of weight $r \leqslant \frac{n-k}{d}$ in polynomial time with a probability of failure

$$p_f < \max\left(q^{-(n-k-2(r+d)+5)}, q^{-2(n-k-rd+2)}\right)$$

$\rightarrow$ matrices based on random small weight codewords with same support can be turned into a decoding algorithm !

Difficult problems in rank metric

Problem (Rank Syndrome Decoding problem)

Given $\boldsymbol{H} \in \mathbb{F}_{q^m}^{(n-k)\times n}$, $\boldsymbol{s} \in \mathbb{F}_{q^m}^{n-k}$ *and an integer* r, *find* $\boldsymbol{e} \in \mathbb{F}_{q^m}^{n}$ *such that:*

- $\boldsymbol{He}^T = \boldsymbol{s}^T$
- $|\boldsymbol{e}|_r = r$

Probabilistic reduction to the NP-Complete SD problem [Gaborit-Zémor, IEEE-IT 2016].

OUROBOROS-R scheme

Vectors x of $\mathbb{F}_{q^m}^n$ seen as elements of $\mathbb{F}_{q^m}[X]/(P)$ for some polynomial P.

Alice		Bob
$\text{seed}_{\mathbf{h}} \leftarrow \{0,1\}^\lambda$, $\mathbf{h} \overset{\text{seed}_{\mathbf{h}}}{\longleftarrow} \mathbb{F}_{q^m}^n$ $(\mathbf{x}, \mathbf{y}) \leftarrow \mathcal{S}_{1,w}^{2n}(\mathbb{F}_{q^m})$, $\mathbf{s} \leftarrow \mathbf{x} + \mathbf{hy}$ $\mathbf{F} \leftarrow \text{Supp}\,(\mathbf{x}, \mathbf{y})$	$\xrightarrow{\mathbf{h},\mathbf{s}}$	$(\mathbf{r}_1, \mathbf{r}_2, \mathbf{e}_\mathbf{r}) \leftarrow \mathcal{S}_{w_r}^{3n}(\mathbb{F}_{q^m})$ $\mathbf{E} \leftarrow \text{Supp}\,(\mathbf{r}_1, \mathbf{r}_2, \mathbf{e}_\mathbf{r})$
$\mathbf{e}_\mathbf{c} \leftarrow \mathbf{s}_\mathbf{e} - \mathbf{y}\mathbf{s}_\mathbf{r}$ $\mathbf{E} \leftarrow \text{QCRS-Recover}(\mathbf{F}, \mathbf{e}_\mathbf{c}, w_r)$	$\xleftarrow{\mathbf{s}_\mathbf{r},\mathbf{s}_\mathbf{e}}$	$\mathbf{s}_\mathbf{r} \leftarrow \mathbf{r}_1 + \mathbf{h}\mathbf{r}_2$, $\mathbf{s}_\mathbf{e} \leftarrow \mathbf{s}\mathbf{r}_2 + \mathbf{e}_\mathbf{r}$
$\boxed{\text{Hash}\,(\mathbf{E})}$	Shared Secret	$\boxed{\text{Hash}\,(\mathbf{E})}$

Figure 1: Informal description of OUROBOROS-R. **h** and **s** constitute the public key. **h** can be recovered by publishing only the λ bits of the seed (instead of the n coordinates of **h**).

Why does it work ?

$$\mathbf{e_c} = \mathbf{s_e} - \mathbf{y}\mathbf{s_r} = \mathbf{s}\mathbf{r}_2 + \mathbf{e_r} - \mathbf{y}(\mathbf{r}_1 + \mathbf{h}\mathbf{r}_2)$$
$$= (\mathbf{x} + \mathbf{h}\mathbf{y})\mathbf{r}_2 + \mathbf{e_r} - \mathbf{y}(\mathbf{r}_1 + \mathbf{h}\mathbf{r}_2) = \mathbf{x}\mathbf{r}_2 - \mathbf{y}\mathbf{r}_1 + \mathbf{e_r}$$

$1 \in \mathbf{F}$, coordinates of $\mathbf{e_c}$ generate a subspace of $\mathrm{Supp}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{e_r}) \times \mathrm{Supp}(\mathbf{x}, \mathbf{y})$ on which one can apply the QCRS-Recover algorithm to recover E (LRPC decoder).

In other words: e_c seen as syndrome associated to an LRPC code based on the secret key (x, y)
$\rightarrow$ a reasonable decoding algorithm is used to decode a SMALL weight error !

Semantic Security

Theorem

Under the assumption of the hardness of the $[2n, n]$-Decisional-QCRSD and $[3n, n]$-Decisional-QCRSD problems, OUROBOROS-R is IND-CPA in the Random Oracle Model.

Best Known Attacks

- Combinatorial attacks: try to guess the support of the error or of the codeword. The best algorithm is GRS+(Aragon et al. ISIT 2018). On average:

$$\mathcal{O}\left((nm)^3 q^{r\left\lceil \frac{km}{n} \right\rceil - m}\right)$$

- Quantum Speed Up : Grover's algorithm directly applies to GRS+ $\implies$ exponent divided by 2.

Examples of parameters

All the times are given in **ms**, performed on an Intel Core i7-4700HQ CPU running at 3.40GHz.

Security	Key Size (bits)	Ciphertext Size (bits)	KeyGen Time(ms)	Encap Time(ms)	Decap Time(ms)	Probability of failure
128	5,408	10,816	0.18	0.29	0.53	$< 2^{-36}$
192	6,456	12,912	0.19	0.33	0.97	$< 2^{-36}$
256	8,896	17,792	0.24	0.40	1.38	$< 2^{-42}$

Advantages and Limitations

Advantages:

- Small key size
- Very fast encryption/decryption time
- **Reduction to decoding a random (QC) code**.
- Well understood decryption failure probability

Limitations:

- Longer ciphertext (compared to LRPC) because of reconciliation ($\times 2$).
- Slighlty larger parameters because of security reduction compared to LRPC.
- RSD problem studied since 27 years.

Questions !