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

NIST First Post-Quantum Cryptography Standardization Conference

Carlos AguilarMelchor\textsuperscript{2} Nicolas Aragon\textsuperscript{1} Slim Bettaieb\textsuperscript{5}
Loïc Bidoux\textsuperscript{5} Olivier Blazy\textsuperscript{1} Jean-Christophe Deneuville\textsuperscript{1,4}
Philippe Gaborit\textsuperscript{1} Adrien Hauteville\textsuperscript{1} Gilles Zémor\textsuperscript{3}

\textsuperscript{1}University of Limoges, XLIM-DMI, France; \textsuperscript{2}ISAE-SUPAERO, Toulouse, France
\textsuperscript{3}IMB, University of Bordeaux; \textsuperscript{4}INSA-CVL, Bourges, France; \textsuperscript{5}Worldline, France.
OUROBOROS-R, an IND-CPA KEM based on Rank Metric
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 $x \in \mathbb{F}_{q^m}^n$ we can associate a matrix $M_x$

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

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

**Definition**

$$d_R(x, y) = \text{Rank}(M_x - M_y) \text{ and } |x|_r = \text{Rank } M_x.$$
## Support of a Word

### Definition

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

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

### Number of supports of weight $w$:

<table>
<thead>
<tr>
<th>Rank</th>
<th>Hamming</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\left[ \begin{array}{c} m \ w \end{array} \right]_q \approx q^{w(m-w)}$</td>
<td>$\binom{n}{w} \leq 2^n$</td>
</tr>
</tbody>
</table>

### Complexity in the worst case:
- quadratically exponential for Rank Metric
- simply exponential for Hamming Metric
**LRPC Codes**

**Definition**

Let $H \in \mathbb{F}_{q^m}^{(n-k) \times n}$ be a full-rank matrix such that the dimension $d$ of $\langle h_{ij} \rangle_{\mathbb{F}_q}$ is small.

By definition, $H$ is a parity-check matrix of an $[n, k]_{q^m}$ LRPC code. We say that $d$ is the weight of the matrix $H$.

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

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

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

- $He^T = s^T$
- $|e|_r = r$

Probabilistic reduction to the NP-Complete SD problem [Gaborit-Zémor, IEEE-IT 2016].
1. Presentation of the rank metric

2. Description of the scheme

3. Security and parameters
OUROBOROS-R scheme

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

<table>
<thead>
<tr>
<th>Alice</th>
<th>Bob</th>
</tr>
</thead>
<tbody>
<tr>
<td>seed$<em>h$ $\leftarrow {0, 1}^\lambda$, $h$ $\leftarrow \mathbb{F}</em>{q^m}^n$</td>
<td>$h,s$ $\xrightarrow{}$</td>
</tr>
<tr>
<td>$(x, y) \leftarrow S_1,w(\mathbb{F}_{q^m}^n)$, $s \leftarrow x + hy$</td>
<td>$(r_1, r_2, e_r) \leftarrow S_{w_r}(\mathbb{F}_{q^m}^n)$</td>
</tr>
<tr>
<td>$F \leftarrow \text{Supp}(x, y)$</td>
<td>$E \leftarrow \text{Supp}(r_1, r_2, e_r)$</td>
</tr>
<tr>
<td>$e_c \leftarrow s_e - y s_r$</td>
<td>$s_r \leftarrow r_1 + hr_2$, $s_e \leftarrow s r_2 + e_r$</td>
</tr>
<tr>
<td>$E \leftarrow \text{QCRS-Recover}(F, e_c, w_r)$</td>
<td></td>
</tr>
<tr>
<td>Hash$(E)$</td>
<td>Shared Secret</td>
</tr>
</tbody>
</table>

**Figure 1:** Informal description of OUROBOROS-R. $h$ and $s$ constitute the public key. $h$ can be recovered by publishing only the $\lambda$ bits of the seed (instead of the $n$ coordinates of $h$).
Why does it work?

\[ \mathbf{e}_c = s_\mathbf{e} - y\mathbf{s}_r = s\mathbf{r}_2 + \mathbf{e}_r - y(r_1 + h\mathbf{r}_2) \]
\[ = (x + hy)r_2 + \mathbf{e}_r - y(r_1 + h\mathbf{r}_2) = xr_2 - yr_1 + \mathbf{e}_r \]

1 \in \mathbb{F}, coordinates of \( \mathbf{e}_c \) generate a subspace of 
\( \text{Supp}(r_1, r_2, \mathbf{e}_r) \times \text{Supp}(x, y) \) on which one can apply the 
QCRS-Recover algorithm to recover \( E \) (LRPC decoder).

In other words: \( \mathbf{e}_c \) 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!

OUROBOROS-R, an IND-CPA KEM based on Rank Metric
1. Presentation of the rank metric
2. Description of the scheme
3. Security and parameters
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.

<table>
<thead>
<tr>
<th>Security</th>
<th>Key Size (bits)</th>
<th>Ciphertext Size (bits)</th>
<th>KeyGen Time (ms)</th>
<th>Encap Time (ms)</th>
<th>Decap Time (ms)</th>
<th>Probability of failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>5,408</td>
<td>10,816</td>
<td>0.18</td>
<td>0.29</td>
<td>0.53</td>
<td>$&lt; 2^{-36}$</td>
</tr>
<tr>
<td>192</td>
<td>6,456</td>
<td>12,912</td>
<td>0.19</td>
<td>0.33</td>
<td>0.97</td>
<td>$&lt; 2^{-36}$</td>
</tr>
<tr>
<td>256</td>
<td>8,896</td>
<td>17,792</td>
<td>0.24</td>
<td>0.40</td>
<td>1.38</td>
<td>$&lt; 2^{-42}$</td>
</tr>
</tbody>
</table>
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 ($\times 2$).
- Slightly larger parameters because of security reduction compared to LRPC.
- RSD problem studied since 27 years.
Questions!