A new rank metric codes based cryptosystem

Pierre Loidreau

DGA MI and IRMAR, Université de Rennes 1

PQCrypto 2017, June 26th

Motivations

- Post-Quantum cryptography
  - Multivariate cryptography
  - Hash-based cryptography
  - Isogenies based cryptography
  - Decoding based cryptography
    - Lattices
    - Codes

$\Rightarrow$ Rank metric codes based

  - Smaller keys for a given security target
  - Another alternative to Hamming metric or Euclidian metric based primitives.

In theory

- Code-Based Encryption: solving general decoding problems in the metric is hard
- In Hamming metric: Dec-*Bounded Distance Decoding* is *NP*-complete, [BMvT78]
- Rank metric decoding related to two difficult problems:
  - *MinRank*, *NP*-complete
  - Dec-*Rank Syndrome Decoding* in *ZPP* $\Rightarrow$ *ZPP*=*NP*, [GZ15]

In practice

Consider a *random* $[n, Rn]$-code over $\mathbb{F}_{2^n}$,

- Decoding errors of rank δn, [GRS16]: $2^{c_{algo}(\delta)n^2+\Omega(\log(n))}$
- Decoding errors of Hamming weight δn: $2^{c_{algo}(\delta)n+o(1)}$

Dec. Complex.	Ham. Met. Gen. Mat.	Rank Met. Gen. Mat.
2^{128}	$[2400, 2006, 58]_2 \approx 100\ KB$	$[48, 39, 4]_{2^{48}} \approx 2.2\ KB$
2^{256}	$[4150, 3307, 132]_2 \approx 350\ KB$	$[70, 50, 5]_{2^{70}} \approx 8.7\ KB$

Table: Decoding complexity on classical computer, [CTS16]

$\Rightarrow$ Rank metric provides better security/size tradeoff
$\Rightarrow$ In *PQ*-world, exponential complexity is square-rooted, [GHT16]

Rank metric, [Gab85]

Definition

- $\gamma_1, \ldots, \gamma_m$, *a basis of* $\mathbb{F}_{2^m}/\mathbb{F}_2$,
- $\mathbf{e} = (e_1, \ldots, e_n) \in (\mathbb{F}_{2^m})^n$, $e_i \mapsto (e_{i1}, \ldots, e_{in})$,

$$\forall \mathbf{e} \in (\mathbb{F}_{2^m})^n, \quad \mathrm{Rk}(\mathbf{e}) \overset{def}{=} \mathrm{Rk}\begin{pmatrix} e_{11} & \cdots & e_{1n} \\ \vdots & \ddots & \vdots \\ e_{m1} & \cdots & e_{mn} \end{pmatrix}$$

- $[n, k, d]_r$ code: $\mathcal{C} \subset \mathbb{F}_{2^m}^n$, k-dimensional, $d = \min_{\mathbf{c} \neq \mathbf{0} \in \mathcal{C}} \mathrm{Rk}(\mathbf{c})$
- Singleton property $d - 1 \leq n - k$ (if $n \leq m$)
- $\mathrm{Rk}(\mathbf{e}) = t \Leftrightarrow \exists \mathcal{V} \subset \mathbb{F}_{2^m}$, s.t. $\dim_2(\mathcal{V}) = t$ and $e_i \in \mathcal{V}$, $\forall i$

Example

$$\mathbf{e} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

In $\mathbb{F}_{2^5}$ we have $\mathbf{e} = (\alpha, \beta, \alpha + \beta, \beta, \alpha + \beta)$

- Hamming weight: 5
- Rank: 2

Rank metric codes based encryption

Key generation

- Private-key
  - $\mathcal{C}$ a $[n, k, d]_r$ t-rank error decodable code over $\mathbb{F}_{2^m}$
  - $L : \mathbb{F}_{2^m}^n \mapsto \mathbb{F}_{2^m}^n$, s.t.
    - L is vector-space isomorphism
    - L is a rank isometry
- Public-key: $\mathcal{C}_{pub} = L^{-1}(\mathcal{C})$.

Process

- Encryption: $\mathbf{y} = \mathbf{c} \in \mathcal{C}_{pub} + \mathbf{e}$, where $\mathrm{Rk}(\mathbf{e}) \leq t$
- Decryption: $L(\mathbf{y}) = L(\mathbf{c}) \in \mathcal{C} + L(\mathbf{e}) \overset{Decode}{\Rightarrow} \mathbf{c}$

Gabidulin codes, [Gab85]

Definition (Gabidulin codes)

Let $\mathbf{g} = (g_1, \ldots, g_n) \in (\mathbb{F}_{2^m})^n$, $\mathbb{F}_2$-*l.i.*, $[i] \overset{def}{=} 2^i$

$$Gab_k(\mathbf{g}) = \langle \mathbf{G} \rangle, \ \textit{where } \mathbf{G} = \begin{pmatrix} g_1 & \cdots & g_n \\ \vdots & \ddots & \vdots \\ g_1^{[k-1]} & \cdots & g_n^{[k-1]} \end{pmatrix}$$

- Properties of $\text{Gab}_k(\mathbf{g})$
  - Optimal $[n, k, d]_r$ codes for rank metric: $n - k = d - 1$
  - P-time quadratic decoding up to $t = \lfloor (n-k)/2 \rfloor$, [Gab85]
- *Sufficiently* scrambled $\Rightarrow$ McEliece-like cryptosystems.

Rise and fall of GPT encryption - [GPT91, Ksh07, RGH10, OKN16]

- Linear rank preserving isometries of $\mathbb{F}_{2^m}^n$: $\mathbf{P} \in M_n(\mathbb{F}_2)$
- Since $\mathrm{Gab}_k(\mathbf{g})\mathbf{P} = \mathrm{Gab}_k(\mathbf{gP}) \Rightarrow$ Necessity of scrambling
- But
  1. For any published reparation, always possible to write

$$\mathbf{G}_{pub} = \mathbf{S}_1(\mathbf{X}_1 \mid \underbrace{\mathbf{G}_1}_{Gab_k(\mathbf{g}_1)})\mathbf{P}^*, \ \mathbf{P}^* \in M_n(\mathbb{F}_2)$$

  2. $\Rightarrow$ Stability through $g \mapsto g^{[i]}$,

$$(\mathbf{G}_{pub})^{[i]} = \mathbf{S}_1^{[i]}\left(\mathbf{X}_1^{[i]} \mid \mathbf{G}_1^{[i]}\right)\mathbf{P}^*$$

  3. $\Rightarrow$ Apply Overbeck's like attacks

How to strengthen ?

- Find less structured codes for rank metric
  - Use of subfield subcodes ? Not sufficient !,[GL08]
- Find a new way to mask the structure
  - Simple
  - Efficient
  - Convincing

A novel idea: LRPC codes, [GMRZ13]

- Let $\mathcal{V} \subset \mathbb{F}_{2^m}$ a λ dimensional $\mathbb{F}_2$-subspace
- Let $\mathcal{L} \subset \mathbb{F}_{2^m}^n$, $[n, k, d]_r$-code with parity-check $\mathbf{H}$ of low rank:

$$\mathbf{H} \in \mathcal{V}^{(n-k)\times n} \subset \mathbb{F}_{2^m}^{(n-k)\times n}$$

- Decoding $\mathbf{y} = \mathbf{c} + \mathbf{e}$, $\mathbf{e} \in \mathcal{E}^n$ where $\dim_2(\mathcal{E}) \leq t$
  1. Since $\mathbf{e} \in \mathcal{E}^n \Rightarrow \mathbf{yH}^t = \mathbf{eH}^t \subset (\mathcal{E} \cdot \mathcal{V})^{n-k}$
  2. $(\mathcal{E} \cdot \mathcal{V}) \stackrel{def}{=} \langle \alpha\beta,\ \alpha \in \mathcal{E},\ \beta \in \mathcal{V} \rangle \Rightarrow \dim_2(\mathcal{E} \cdot \mathcal{V}) \leq t\lambda$
  3. If $t\lambda \leq n - k$, knowing $\mathcal{V} \Rightarrow$ recovers $\mathcal{E}$ from $(\mathcal{E} \cdot \mathcal{V})$

$\Rightarrow$ LRPC based encryption was designed

Mixing the ideas

Weaknesses and strengths

- Gabidulin codes:
  - Advantages: efficient deterministic decoding
  - Drawbacks: too much structured
- LRPC codes:
  - Advantages: not structured
  - Drawbacks: probabilistic decoding with failure $2^{-(n-k-\lambda t)}$
  - Questions about attacks on MDPC's.

$\Rightarrow$ use rank multiplication to scramble structure of Gabidulin codes

The new encryption scheme

Proposition

Let $\mathcal{V} \subset \mathbb{F}_{2^m}$ *with* $\dim_2(\mathcal{V}) = \lambda$*, and let* $\mathbf{P} \in M_n(\mathcal{V})$*, then*

$$\forall \mathbf{x} \in \mathbb{F}_{2^m}^n, \ \mathrm{Rk}(\mathbf{xP}) \leq \lambda \, \mathrm{Rk}(\mathbf{x})$$

- Private-key:
  - $\mathrm{Gab}_k(\mathbf{g})$
  - $\mathcal{V} = \langle \alpha_1, \ldots, \alpha_\lambda \rangle_2$, λ-dimensional
  - $\mathbf{P} \in M_n(\mathcal{V})$
- Public-key: $\mathcal{C}_{pub} = \mathrm{Gab}_k(\mathbf{g})\mathbf{P}^{-1}$
- Encryption: $\mathbf{y} = \mathbf{c} \in \mathcal{C}_{pub} + \mathbf{e}$, where $\mathrm{Rk}(\mathbf{e}) \leq \lfloor (n-k)/(2\lambda) \rfloor$
- Decryption: $\mathbf{yP} = \mathbf{cP} \in \mathcal{C} + \mathbf{eP}$, where $\mathrm{Rk}(\mathbf{eP}) \leq \lfloor (n-k)/2 \rfloor$

Security arguments

- Indistinguishability of the public-code:
  - $\mathcal{V}$ not 2-stable $\Rightarrow \mathrm{Gab}_k(\mathbf{g})\mathbf{P}^{-1} \neq \mathrm{Gab}_k(\mathbf{g}\mathbf{P}^{-1})$:
  - $\mathcal{C}_{pub}$ and $\mathcal{C}_{pub}^{[i]}$, behave independently
  - Complexity evaluation: Harder than enumerating $\lambda - 1$ dimensional subspaces in $\mathbb{F}_{2^m}$:

$$> 2^{m(\lambda-1)-(\lambda-1)^2}$$

$\Rightarrow$ it is a *OWE* (One-Way Encryption)

- If $\lambda(n-k) \geq n$, (Couvreur, Coggia)

How to choose the parameters

- For a given security parameter s:
  - Choose $m, \lambda \geq 2$, s.t. $2^{m(\lambda-1)-(\lambda-1)^2} > 2^s$
  - Choose k, n s.t. solving $BDR(\lfloor (n-k)/2\lambda \rfloor) > 2^s$
  - Check that $\lambda(n-k) \geq n$

Proposition of parameters

Param.	Dec.	PQ	K. Rec.	Key
$m = n = 50,\ k = 32,\ \lambda = 3,\ t = 3$	$\approx 2^{81}$	$\approx 2^{49}$	$\approx 2^{96}$	$3.6\ KB$
$m = 96,\ n = 64,\ k = 40,\ \lambda = 3,\ t = 4$	$\approx 2^{139}$	$\approx 2^{80}$	$\approx 2^{188}$	$11.5\ KB$
$m = n = 112,\ k = 80,\ \lambda = 4,\ t = 4$	$\approx 2^{259}$	$\approx 2^{139}$	$\approx 2^{327}$	$36\ KB$

- Key-size for McEliece with Goppa codes: $\approx 850\ KB$ for 128 bits PQ-security, [AB315]
- Key-size factor gain: ≈ 23

Perspectives

- Reducing key-size by some structural property
- Thorough study of the security of the system
- Designing additional cryptographic services

Why do we believe in the design

- Choice of $\mathcal{V}$: Similar to subcodes in Hamming metric
  - For adequate parameters: Distinguishing the public-key is difficult.
- Versatility of the parameters

References I

Initial recommendations of long-term secure post-quantum systems.
Technical report, 2015.
http://pqcrypto.eu.org/docs/initial-recommendations.pdf.

E. Berlekamp, R. J. McEliece, and H. van Tilborg.
On the inherent intractability of certain coding problems.
IEEE Trans. Inform. Theory, 24(3):384–386, May 1978.

R. Canto-Torres and N. Sendrier.
Analysis of information set decoding for a sub-linear error weight.
In *Post-Quantum Cryptography 2016*, Lecture Notes in Comput. Sci., pages 144–161, Fukuoka, Japan, February 2016.

References II

E. M. Gabidulin.
Theory of codes with maximum rank distance.
Probl. Inf. Transm., 21(1):3–16, 1985.

P. Gaborit, A. Hauteville, and J.-P. Tillich.
Ranksynd a PRNG based on rank metric.
In *Post-Quantum Cryptography 2016*, pages 18–28, February 2016.

E. M. Gabidulin and P. Loidreau.
Properties of subspace subcodes of Gabidulin codes.
Adv. in Math. of Comm., 2(2):147–157, 2008.

References III

P. Gaborit, G. Murat, O. Ruatta, and G. Zémor.
Low rank parity check codes and their application to cryptography.
In *Proceedings of the Workshop on Coding and Cryptography WCC'2013*, Bergen, Norway, 2013.
Available on www.selmer.uib.no/WCC2013/pdfs/Gaborit.pdf.

E. M. Gabidulin, A. V. Paramonov, and O. V. Tretjakov.
Ideals over a non-commutative ring and their applications to cryptography.
In *Advances in Cryptology - EUROCRYPT'91*, number 547 in Lecture Notes in Comput. Sci., pages 482–489, Brighton, April 1991.

References IV

P. Gaborit, O. Ruatta, and J. Schrek.
On the complexity of the rank syndrome decoding problem.
IEEE Trans. Information Theory, 62(2):1006–1019, 2016.

P. Gaborit and G. Zémor.
On the hardness of the decoding and the minimum distance problems for rank codes.
IEEE Trans. Inform. Theory, 62(12):7245–7252, 2015.

A. Kshevetskiy.
Security of GPT-like public-key cryptosystems based on linear rank codes.
In *3rd International Workshop on Signal Design and Its Applications in Communications, IWSDA 2007*, 2007.

References V

A. May and I. Ozerov.
On Computing Nearest Neighbors with Applications to Decoding of Binary Linear Codes.
In E. Oswald and M. Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015*, volume 9056, pages 203–228, 2015.

A. Otmani, H. T. Kalashi, and S. Ndjeya.
Improved cryptanalysis of rank metric schemes based on Gabidulin codes.
http://arxiv.org/abs/1602.08549v1, 2016.

H. Rashwan, E. M. Gabidulin, and B. Honary.
A smart approach for GPT cryptosystem based on rank codes.
In *Proc. IEEE Int. Symposium Inf. Theory - ISIT*, pages 2463–2467, 2010.