Code-Based Cryptography with the Subspace Metric

Anna-Lena Horlemann

University of St. Gallen, Switzerland

SIAM AG21, August 17th, 2021

Main idea of code-based cryptosystems

- Decoding a random linear code is a hard problem.
- **Public key/information**: the parity check matrix of a random (looking) linear code, and a syndrome
- **Secret**: the solution to the corresponding syndrome decoding problem: usually a low-weight error vector (and/or the corresponding message/codeword)

$$\underbrace{\mathbf{s}}_{\text{syndrome}} = \underbrace{\mathbf{e}}_{\text{error vector}} \cdot \underbrace{\mathbf{H}^\top}_{\text{PC matrix}}$$

Main idea of code-based cryptosystems

- Decoding a random linear code is a hard problem.
- **Public key/information**: the parity check matrix of a random (looking) linear code, and a syndrome
- **Secret**: the solution to the corresponding syndrome decoding problem: usually a low-weight error vector (and/or the corresponding message/codeword)

$$\underbrace{\mathbf{s}}_{\text{syndrome}} = \underbrace{\mathbf{e}}_{\text{error vector}} \cdot \underbrace{\mathbf{H}^{\top}}_{\text{PC matrix}}$$

Various weights:

- Hamming weight
- rank weight
- Lee weight
- etc. (homogeneous weight, sum rank weight)

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

1. an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
2. in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

1. an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
2. in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

$\Longrightarrow$ we can relax or generalize both of the above

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

1. an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
2. in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

$\implies$ we can relax or generalize both of the above

But what do we really need?

Reformulation and generalization

The syndrome decoding problem asks for a vector that is

1. an element of the coset of the subspace $\ker(\mathbf{H})$ given by $\mathbf{s}$,
2. in the sphere $\{\mathbf{x} \mid \mathrm{wt}(\mathbf{x}) = w\} = \{\mathbf{x} \mid d_H(\mathbf{x}, \mathbf{0}) = w\}$.

$\Longrightarrow$ we can relax or generalize both of the above

But what do we really need?

1. We need an efficient representation of the code (e.g. by linearity).
2. If we do not think about weights/distances any more, it is not code-based crypto.
3. For PKE we need an efficient decoding algorithm.
4. For identification schemes decoding is not necessary. But we need transitive "linear" maps on the spheres (in the existing schemes), and identifiers of the cosets (e.g. syndromes).

How could we use the subspace metric?

Quick reminder

Definition

Denote by $\mathcal{P}_q(n)$ the set of all subspaces of $\mathbb{F}_q^n$ and by $\mathcal{G}_q(k,n)$ the set of all k-dimensional subspaces of $\mathbb{F}_q^n$ ("Grassmannian").

1. A subset $\mathcal{C} \subseteq \mathcal{P}_q(n)$ is called a **subspace code**. If $\mathcal{C} \subseteq \mathcal{G}_q(k,n)$, then it is also called a **constant-dimension code**.
2. The **subspace distance** on $\mathcal{P}_q(n)$ is defined as

$$d_S(\mathcal{U},\mathcal{V}) := \dim(\mathcal{U}) + \dim(\mathcal{V}) - 2\dim(\mathcal{U}\cap\mathcal{V}).$$

Problems to solve

1. Efficient representation of the public key (”linearity”):

2. Distance measure for the correct solution (”weight”):

[1] including multi-level construction and spread codes

Problems to solve

1. Efficient representation of the public key ("linearity"):
   - lifted rank-metric codes[1] – generator matrix of the rank-metric code
2. Distance measure for the correct solution ("weight"):
   - lifted rank-metric codes – rank weight ($\cong$ subsp. dist. to $\mathrm{rs}[I \mid 0]$)

[1] including multi-level construction and spread codes

Problems to solve

1. Efficient representation of the public key ("linearity"):
   - lifted rank-metric codes[1] – generator matrix of the rank-metric code
   - orbit codes – generators of the group in GL_n defining the orbit code
2. Distance measure for the correct solution ("weight"):
   - lifted rank-metric codes – rank weight ($\cong$ subsp. dist. to $\mathrm{rs}[I \mid 0]$)
   - orbit codes – subspace distance to a prescribed "zero" codeword

[1] including multi-level construction and spread codes

Problems to solve

1. Efficient representation of the public key ("linearity"):
   - lifted rank-metric codes[1] – generator matrix of the rank-metric code
   - orbit codes – generators of the group in GL_n defining the orbit code
2. Distance measure for the correct solution ("weight"):
   - lifted rank-metric codes – rank weight ($\cong$ subsp. dist. to $\mathrm{rs}[I \mid 0]$)
   - orbit codes – subspace distance to a prescribed "zero" codeword
3. For McEliece/Niederreiter type systems we also need an efficient decoding algorithm:
   - lifted rank-metric codes – Gabidulin code decoders
   - orbit codes – ???

[1] including multi-level construction and spread codes

McEliece with lifted Gabidulin codes

- **Secret key:** Gabidulin code $\mathcal{C}_{Gab} \subseteq \mathbb{F}_{q^{n-k}}^{\kappa \times k}$
- **Public key:** Generator matrix G_{pub} of $\mathcal{C}_{pub} := \phi(\mathcal{C}_{Gab})$ [2]
- **Encryption (encoding plus random subspace errors):**

$$\mathcal{D}_\rho(\mathrm{rs}[I_k \mid \underbrace{mG_{pub}}_{\text{expanded over } \mathbb{F}_q}]) \oplus \mathcal{E}$$

  such that $\rho + \dim(\mathcal{E}) \leq t$ (error correction capability)
- **Decryption:** Use lifted Gabidulin decoder with application of ϕ^{-1}

[2] ϕ can be any valid rank-metric disguising function.

Decoding with transformation to secret code

By Silva-Kschischang (2009), decoding the ciphertext in the received word can be translated to

$$\underset{X \in \mathcal{C}_{pub}}{\operatorname{argmin}} \operatorname{rk} \begin{pmatrix} \hat{L} & X - R \\ 0 & \hat{E} \end{pmatrix} = \underset{X' \in \mathcal{C}_{Gab}}{\operatorname{argmin}} \operatorname{rk} \begin{pmatrix} \bar{L} & X' - \phi^{-1}(R) \\ 0 & \bar{E} \end{pmatrix}$$

which is in turn equivalent to a rank-metric decoding problem with row and column erasures.
($\hat{L}, \hat{E}, R$ are given by the structure of the cipher vector space; $\bar{L}, \bar{E}$ also depend on ϕ.)

Decoding with transformation to secret code

By Silva-Kschischang (2009), decoding the ciphertext in the received word can be translated to

$$\underset{X \in \mathcal{C}_{pub}}{\operatorname{argmin}} \operatorname{rk} \begin{pmatrix} \hat{L} & X - R \\ 0 & \hat{E} \end{pmatrix} = \underset{X' \in \mathcal{C}_{Gab}}{\operatorname{argmin}} \operatorname{rk} \begin{pmatrix} \bar{L} & X' - \phi^{-1}(R) \\ 0 & \bar{E} \end{pmatrix}$$

which is in turn equivalent to a rank-metric decoding problem with row and column erasures.
($\hat{L}, \hat{E}, R$ are given by the structure of the cipher vector space; $\bar{L}, \bar{E}$ also depend on ϕ.)

$\Longrightarrow$ For both the receiver and the attacker it is equivalent to a rank-metric decoding problem with row/column erasures.

Decoding with transformation to secret code

By Silva-Kschischang (2009), decoding the ciphertext in the received word can be translated to

$$\operatorname*{argmin}_{X \in \mathcal{C}_{pub}} \operatorname{rk} \begin{pmatrix} \hat{L} & X - R \\ 0 & \hat{E} \end{pmatrix} = \operatorname*{argmin}_{X' \in \mathcal{C}_{Gab}} \operatorname{rk} \begin{pmatrix} \bar{L} & X' - \phi^{-1}(R) \\ 0 & \bar{E} \end{pmatrix}$$

which is in turn equivalent to a rank-metric decoding problem with row and column erasures.
($\hat{L}, \hat{E}, R$ are given by the structure of the cipher vector space; $\bar{L}, \bar{E}$ also depend on ϕ.)

$\Longrightarrow$ For both the receiver and the attacker it is equivalent to a rank-metric decoding problem with row/column erasures.

$\Longrightarrow$ Subspace metric not necessary, can just use rank metric.

Multiplicative analog of linearity via orbit codes

- Group theoretic (multiplicative instead of additive) analog of linear codes: orbit codes in $\mathcal{G}_q(k,n)$

Definition

Let $G \leq \mathrm{GL}_n$ be a group and $\mathcal{U}_0 \in \mathcal{G}_q(k,n)$. Then $\mathcal{U}_0 G$ is an orbit code in $\mathcal{G}_q(k,n)$.

[3] except for the cases where the orbit code is also a lifted MRD or spread code

Multiplicative analog of linearity via orbit codes

- Group theoretic (multiplicative instead of additive) analog of linear codes: orbit codes in $\mathcal{G}_q(k,n)$

Definition

Let $G \leq \mathrm{GL}_n$ be a group and $\mathcal{U}_0 \in \mathcal{G}_q(k,n)$. Then $\mathcal{U}_0 G$ is an orbit code in $\mathcal{G}_q(k,n)$.

- No efficient decoders known (yet)[3] $\implies$ not usable for McEliece
- But for identification scheme?!

[3] except for the cases where the orbit code is also a lifted MRD or spread code

Theoretical setup for McEliece with orbit codes

- **Secret key:** Generators of orbit code $\mathcal{C}_{orb} = \mathcal{U}_0 G \subseteq \mathcal{G}_q(k,n)$
- **Public key:** Generators of disguised code $\mathcal{C}_{pub} := \phi(\mathcal{C}_{orb})$
- **Encryption (encoding plus random subspace errors):**

$$\mathcal{D}_\rho(\mathrm{rs}[I_k \mid \underbrace{mG_{pub}}_{\text{expanded over } \mathbb{F}_q}]) \oplus \mathcal{E}$$

  such that $\rho + \dim(\mathcal{E}) \leq t$ (error correction capability)
- **Decryption:** Use orbit decoder with application of ϕ^{-1}

Theoretical setup for McEliece with orbit codes

- **Secret key:** Generators of orbit code $\mathcal{C}_{orb} = \mathcal{U}_0 G \subseteq \mathcal{G}_q(k,n)$
- **Public key:** Generators of disguised code $\mathcal{C}_{pub} := \phi(\mathcal{C}_{orb})$
- **Encryption (encoding plus random subspace errors):**

$$\mathcal{D}_\rho(\mathrm{rs}[I_k \mid \underbrace{mG_{pub}}_{\text{expanded over } \mathbb{F}_q}]) \oplus \mathcal{E}$$

  such that $\rho + \dim(\mathcal{E}) \leq t$ (error correction capability)

- **Decryption:** Use orbit decoder with application of ϕ^{-1}

Questions:

1. What could ϕ be?
   It should keep the orbit structure (for representability), but hide the structure of the secret code.
2. Do we find orbit codes with an efficient decoder?

Idea for a subspace metric ZK-ID scheme

- **Secret:** coset leader $\mathcal{V} \in \mathcal{G}_q(k,n)$, s.t.
  - $\underset{B \in G}{\operatorname{argmin}}\, d_S(\mathcal{U}_0, \mathcal{V}B) = I_n$
  - $d_S(\mathcal{U}_0, \mathcal{V}) = t$
- **Public information:**
  - group $G \leq \mathrm{GL}_n(q)$ and $\mathcal{U}_0 \in \mathcal{G}_q(k,n)$ (orbit code $\mathcal{C} := \mathcal{U}_0 G$)
  - an identifier S of the orbit $\mathcal{V}G$
  - distance t
- **Interactive protocol:** Prove to the verifier one of the two per round:
  - secret is on the orbit $\mathcal{V}G$
  - secret has subspace distance t to $\mathcal{U}_0$

Idea for a subspace metric ZK-ID scheme

- **Secret:** coset leader $\mathcal{V} \in \mathcal{G}_q(k,n)$, s.t.
  - $\underset{B \in G}{\operatorname{argmin}}\, d_S(\mathcal{U}_0, \mathcal{V}B) = I_n$
  - $d_S(\mathcal{U}_0, \mathcal{V}) = t$
- **Public information:**
  - group $G \leq \mathrm{GL}_n(q)$ and $\mathcal{U}_0 \in \mathcal{G}_q(k,n)$ (orbit code $\mathcal{C} := \mathcal{U}_0 G$)
  - an identifier S of the orbit $\mathcal{V}G$
  - distance t
- **Interactive protocol:** Prove to the verifier one of the two per round:
  - secret is on the orbit $\mathcal{V}G$
  - secret has subspace distance t to $\mathcal{U}_0$

Questions:

- How to implement the interactive protocol (computationally)?
- What could the orbit identifier ($\cong$ syndrome) be?
- How difficult is the general coset leader decoding problem for orbit codes ($\implies$ security)?

Ideas and problems with the interactive protocol

Let $E \in \mathrm{GL}_n$ such that $\mathcal{V} = \mathcal{U}_0 E$.

Hamming metric codes	orbit codes
sample lin. isometry τ and $u \in \mathbb{F}_q^n$ reveal $y = \tau(u+e)$ and hashes of $u, \tau(u), uH^\top$	sample σ and $U \in \mathrm{GL}_n$ reveal $Y = \sigma(UE)$ and hashes of $U, \sigma(U)$, identifier of $\mathcal{U}_0 UG$

Ideas and problems with the interactive protocol

Let $E \in \mathrm{GL}_n$ such that $\mathcal{V} = \mathcal{U}_0 E$.

Hamming metric codes	orbit codes
sample lin. isometry τ and $u \in \mathbb{F}_q^n$ reveal $y = \tau(u+e)$ and hashes of $u, \tau(u), uH^\top$	sample σ and $U \in \mathrm{GL}_n$ reveal $Y = \sigma(UE)$ and hashes of $U, \sigma(U)$, identifier of $\mathcal{U}_0 UG$
1) *secret is solution to syndr. eq.* reveal τ, verify that $\mathrm{Hash}(\tau^{-1}(y)H^\top - s) = \mathrm{Hash}(uH^\top)$	1) *secret is on the orbit* $\mathcal{V}_0 G$ reveal σ, verify that (hashed) identifier of $\mathcal{U}_0 \sigma^{-1}(Y) G \odot S$ is equal to the one of $\mathcal{U}_0 UG$

Need operation $\odot S$, mapping identifier of $\mathcal{U}_0 UEG$ to the one of $\mathcal{U}_0 UG$, and σ with $d_S(\mathcal{U}_0 E, \mathcal{U}_0) = d_S(\mathcal{U}_0 \sigma(E), \mathcal{U}_0)$ and $\sigma(UE) = \sigma(U)\sigma(E)$.

Ideas and problems with the interactive protocol

Let $E \in \mathrm{GL}_n$ such that $\mathcal{V} = \mathcal{U}_0 E$.

Hamming metric codes	orbit codes
sample lin. isometry τ and $u \in \mathbb{F}_q^n$ reveal $y = \tau(u+e)$ and hashes of $u, \tau(u), uH^\top$	sample σ and $U \in \mathrm{GL}_n$ reveal $Y = \sigma(UE)$ and hashes of $U, \sigma(U)$, identifier of $\mathcal{U}_0 UG$
1) *secret is solution to syndr. eq.* reveal τ, verify that $\mathrm{Hash}(\tau^{-1}(y)H^\top - s) = \mathrm{Hash}(uH^\top)$	1) *secret is on the orbit* $\mathcal{V}_0 G$ reveal σ, verify that (hashed) identifier of $\mathcal{U}_0 \sigma^{-1}(Y) G \odot S$ is equal to the one of $\mathcal{U}_0 UG$
2) *secret has weight* t reveal $e' = \tau(e)$, verify that $\mathrm{wt}(e') = t$ and $\mathrm{Hash}(y - e') = \mathrm{Hash}(\tau(u))$	2) *secret has distance* t *to* $\mathcal{U}_0$ reveal $E' = \sigma(E)$, verify that $d_S(\mathcal{U}_0 E', \mathcal{U}_0) = t$ and $\mathrm{Hash}(Y(E')^{-1}) = \mathrm{Hash}(\sigma(U))$

Need operation $\odot S$, mapping identifier of $\mathcal{U}_0 UEG$ to the one of $\mathcal{U}_0 UG$, and σ with $d_S(\mathcal{U}_0 E, \mathcal{U}_0) = d_S(\mathcal{U}_0 \sigma(E), \mathcal{U}_0)$ and $\sigma(UE) = \sigma(U)\sigma(E)$.

Essential open problems

1. We need a complexity estimate for a generic orbit decoder in $\mathcal{G}_q(k,n) \implies$ security level
2. We need a syndrome-like identifier for the orbits, and a corresponding map $\odot$ such that we can recover the orbit $\mathcal{U}_0 UG$ from the orbit $\mathcal{U}_0 UEG$.
   (Non-commutativity makes this problem really hard.)
3. We need a "$\mathcal{U}_0$-isometry"σ with $d_S(\mathcal{U}_0 E, \mathcal{U}_0) = d_S(\mathcal{U}_0 \sigma(E), \mathcal{U}_0)$ and $\sigma(UE) = \sigma(U)\sigma(E)$.
4. The maps/operators need to come from large enough sets to make it cryptographically secure.

Summary and conclusions

- Using different metrics in code-based cryptography has shown to be beneficial – what about the subspace metric?
- We need efficient representation of the code.
  $\implies$ lifted rank-metric or orbit codes
- For lifted rank-metric codes the decoding problem is equivalent to rank-metric decoding with row and column erasures.
  $\implies$ no real advantage
- For orbit codes we have no efficient decoder.
  $\implies$ no McEliece/Niederreiter system
  But possibly a ZK-ID scheme... $\implies$ many open questions![4]

[4] You can do the same for any type of group code, with similar questions.

Summary and conclusions

- Using different metrics in code-based cryptography has shown to be beneficial – what about the subspace metric?
- We need efficient representation of the code.
  $\Longrightarrow$ lifted rank-metric or orbit codes
- For lifted rank-metric codes the decoding problem is equivalent to rank-metric decoding with row and column erasures.
  $\Longrightarrow$ no real advantage
- For orbit codes we have no efficient decoder.
  $\Longrightarrow$ no McEliece/Niederreiter system
  But possibly a ZK-ID scheme... $\Longrightarrow$ many open questions![4]

Thank you for your attention!
Questions? – Comments?

[4]You can do the same for any type of group code, with similar questions.