May-Ozerov Algorithm for Nearest Neighbor Problem over $\mathbb{F}_q$ and Its Application to Information Set Decoding

Shoichi Hirose

University of Fukui

Hot Topic Session
PQCrypto 2016 (2016/02/24-26, Fukuoka)

Introduction

Information set decoding

- Algorithm for decoding random linear codes
- Generic attack on code-based cryptography

For parity check matrix $\boldsymbol{H}$ and vector $\boldsymbol{x}$, $\boldsymbol{s} = \boldsymbol{H}\boldsymbol{x} = \boldsymbol{H}\boldsymbol{e}$ and $w_{\mathrm{H}}(\boldsymbol{e}) = w$

1. (Permutation step)
   1. Randomly permute the columns of $\boldsymbol{H}$
   2. Transform the permuted $\boldsymbol{H}$ with Gaussian elimination into

$$\begin{array}{cc} \overbrace{\hphantom{\quad\boldsymbol{Q}\quad}}^{k} \; \overbrace{\hphantom{1\;1\;\cdots\;1}}^{n-k} & \\ \left(\begin{array}{c|cccc} & 1 & & & 0 \\ & & 1 & & \\ \boldsymbol{Q} & & & \ddots & \\ & 0 & & & 1 \end{array}\right) & \Big\updownarrow n-k \end{array}$$

   $\boldsymbol{s}$ is also transformed into $\tilde{\boldsymbol{s}}$ accordingly

2. (Search step) For some fixed p, search a linear combination of p columns of $\boldsymbol{Q}$ whose Hamming distance to $\tilde{\boldsymbol{s}}$ is $(w - p)$

Nearest-neighbor algorithm is used for search step

Our Contribution

1. Generalize May-Ozerov algorithm for NN problem over $\mathbb{F}_q$
2. Apply May-Ozerov NN algorithm to Stern ISD algorithm
   - Stern-MO is more efficient than Stern only if $q = 2$

Definition ((m, γ, λ)-Nearest-Neighbor problem over $\mathbb{F}_q$)

- $m \in \mathbb{N}$
- $0 < \gamma < 1/2$
- $0 < \lambda < 1$

Input $\mathcal{U}$, $\mathcal{V}$ and γ, where $\mathcal{U} \subset \mathbb{F}_q^m$, $\mathcal{V} \subset \mathbb{F}_q^m$ and $|\mathcal{U}| = |\mathcal{V}| = q^{\lambda m}$

Output $\mathcal{C} \subset \mathcal{U} \times \mathcal{V}$ which have $(\boldsymbol{u}^*, \boldsymbol{v}^*)$ s.t. $w_{\mathrm{H}}(\boldsymbol{u}^* - \boldsymbol{v}^*) = \gamma m$

May-Ozerov Algorithm over $\mathbb{F}_q$: Overview

Repeat $m^{O(q^3)}$ times:

1. (Randomize & filter) Select a random permutation matrix $\boldsymbol{P}$ and a random balanced vector $\boldsymbol{r}$, and compute

$$\tilde{\mathcal{U}} \leftarrow \{\tilde{\boldsymbol{u}} \,|\, \tilde{\boldsymbol{u}} \in \boldsymbol{P}\mathcal{U} + \boldsymbol{r} \wedge (\tilde{\boldsymbol{u}} \text{ is balanced})\}$$
$$\tilde{\mathcal{V}} \leftarrow \{\tilde{\boldsymbol{v}} \,|\, \tilde{\boldsymbol{v}} \in \boldsymbol{P}\mathcal{V} + \boldsymbol{r} \wedge (\tilde{\boldsymbol{v}} \text{ is balanced})\}$$

2. (Create pairs of lists by filtering) Repeat $q^{O(m)}$ times:
   1. Select a random set $A \subset \{1, 2, \ldots, m\}$ such that $|A| = \beta m$
   2. Compute

$$\mathcal{U}' \leftarrow \{\boldsymbol{u} \,|\, \boldsymbol{u} \in \tilde{\mathcal{U}} \wedge \text{ (the number of } x \in \mathbb{F}_q \text{ in } \tilde{\boldsymbol{u}} \text{ on } A \text{ is } h_x \beta m)\}$$
$$\mathcal{V}' \leftarrow \{\boldsymbol{v} \,|\, \boldsymbol{v} \in \tilde{\mathcal{V}} \wedge \text{ (the number of } x \in \mathbb{F}_q \text{ in } \tilde{\boldsymbol{v}} \text{ on } A \text{ is } h_x \beta m)\}$$

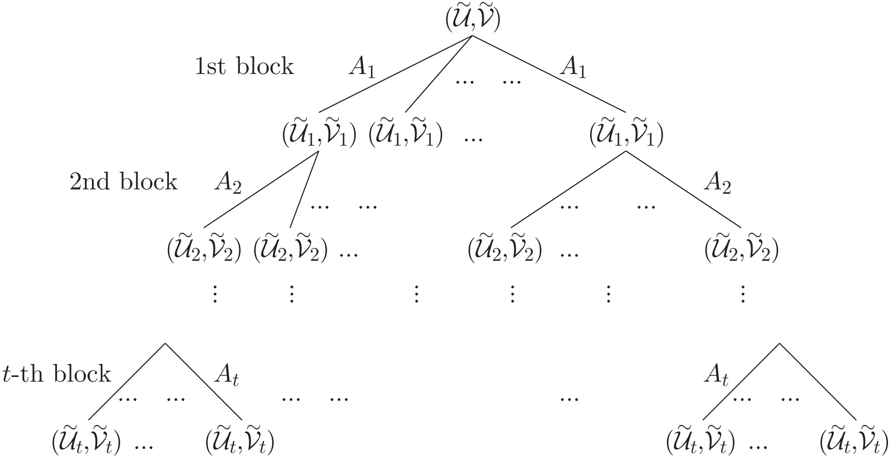
   If $(\boldsymbol{u}^*, \boldsymbol{v}^*) \in \mathcal{U}' \times \mathcal{V}'$ s.t. $w_{\mathrm{H}}(\boldsymbol{u}^* - \boldsymbol{v}^*) = \gamma m$, then success

Intuitive idea: Since $w_{\mathrm{H}}(\boldsymbol{u}^* - \boldsymbol{v}^*)$ is small, for $\exists A$,

$$(\boldsymbol{u}^* \text{ has a bias on } A) \Rightarrow (\boldsymbol{v}^* \text{ has the same bias on } A)$$

May-Ozerov Algorithm over $\mathbb{F}_q$: Some More Detail

- The vectors are divided into t blocks $\tilde{\boldsymbol{u}} = (\tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_t)$
- The second step of the algorithm is applied recursively



May-Ozerov Algorithm over $\mathbb{F}_q$: Time Complexity

Theorem

The May-Ozerov algorithm solves the (m, γ, λ)-NN problem over $\mathbb{F}_q$ in time $\tilde{O}(q^{(y+\varepsilon)m})$ with overwhelming probability, where

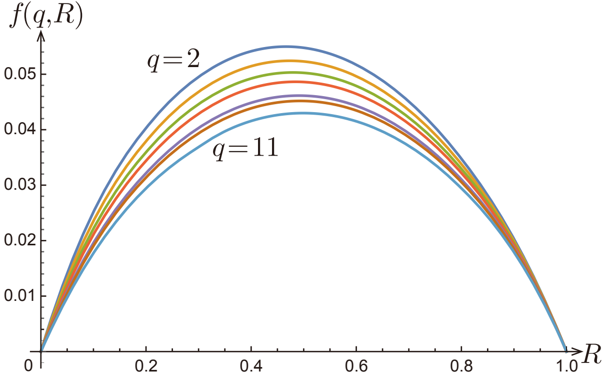
- $y = (1-\gamma)\left(H_q(\beta) - \frac{1}{q}\sum_{x \in \mathbb{F}_q} H_q\left(\frac{qh_x - \gamma}{1-\gamma}\beta\right)\right)$
- $\varepsilon > 0$ *is any real*

The time complexity is optimized under

- $0 < \beta < 1$
- $\frac{\gamma}{q} \le h_x \le \frac{\gamma}{q} + \frac{1-\gamma}{q\beta}$ *for* $\forall x \in \mathbb{F}_q$*, and* $\sum_{x \in \mathbb{F}_q} h_x = 1$
- $\lambda < H_q(\beta) - \frac{1}{q}\sum_{x \in \mathbb{F}_q} H_q(qh_x\beta)$

Asymptotic time complexity of Stern ISD with May-Ozerov NN for bounded distance decoding over $\mathbb{F}_q$

- Time complexity is $\tilde{O}(q^{f(q,R)n})$, where $R = k/n$.
- $q = 2, 3, 4, 5, 7, 8, 11$ in the decreasing order.



Asymptotic time complexity of worst cases for bounded distance decoding

- Time complexity is $\tilde{O}(q^{f(q,R)n})$, where $R = k/n$.
- All but one of h_x's are equal to h.

q	$f(q,R)$	R	p/n	β	h
2	.05498	.4663	.003848	.4998	.3981
3	.05242	.4736	.002979	.1792	.2322
4	.05032	.4796	.002201	.0932	.1644
5	.04864	.4843	.001704	.0593	.1279
7	.04614	.4909	.001164	.0326	.0893
8	.04519	.4933	.001006	.0263	.0778
11	.04299	.4989	.000727	.0166	.0563

Numerical Analysis: Stern-MO vs. Stern

Asymptotic time complexity of worst cases for bounded distance decoding

- Time complexity is $\tilde{O}(q^{f(q,R)n})$, where $R = k/n$.
- $\Delta = f(q,R) - f_{\mathrm{S}}(q,R')$

	Stern-MO		Stern		
q	$f(q,R)$	R	$f_{\mathrm{S}}(q,R')$	R'	Δ
2	.05498	.4663	.05563	.4655	−.00065
3	.05242	.4736	.05217	.4742	.00025
4	.05032	.4796	.04987	.4801	.00045
5	.04864	.4843	.04815	.4844	.00049
7	.04614	.4909	.04571	.4907	.00043
8	.04519	.4933	.04478	.4931	.00041
11	.04299	.4989	.04266	.4985	.00033

Stern-MO is more efficient than Stern only if $q = 2$.

Future Work

- Apply May-Ozerov NN algorithm to BJMM ISD algorithm