Attacks on search-RLWE

Hao Chen

Microsoft Research End-of-Internship Presentation

Mentor: Kristin Lauter
Joint work with: Katherine Stange

September 10, 2015

Overview

1. Background
   - Number fields and canonical embeddings
   - Definitions related to RLWE
2. The chi-square attack
   - Recap the [ELOS] attack
   - The new chi-square attack
3. Galois RLWE
   - Properties of Galois RLWE
   - Vulnerable instances
4. Cyclotomics
   - Unramified primes (are secure)
   - Ramified prime (is vulnerable)

Part 1/4: Background

Minkowski embedding and the embedded lattice

Let K be a number field of degree n with ring of integers R and let $\sigma_1, \cdots, \sigma_n$ be the embeddings of K into $\mathbb{C}$. Assume $\sigma_1, \cdots, \sigma_{r_1}$ are the real embeddings.

Minkowski embedding and the embedded lattice

Let K be a number field of degree n with ring of integers R and let $\sigma_1, \cdots, \sigma_n$ be the embeddings of K into $\mathbb{C}$. Assume $\sigma_1, \cdots, \sigma_{r_1}$ are the real embeddings.

Definition

The *Minkowski embedding* of K is

$$\begin{aligned}\iota : K &\to \mathbb{R}^n \\ x &\mapsto (\sigma_1(x), \cdots, \sigma_{r_1}(x), \mathrm{Re}(\sigma_{r_1+1})(x), \mathrm{Im}(\sigma_{r_1+1})(x), \cdots, \\ &\quad \mathrm{Re}(\sigma_{r_1+r_2})(x), \mathrm{Im}(\sigma_{r_1+r_2})(x))\end{aligned}$$

Minkowski embedding and the embedded lattice

Let K be a number field of degree n with ring of integers R and let $\sigma_1, \cdots, \sigma_n$ be the embeddings of K into $\mathbb{C}$. Assume $\sigma_1, \cdots, \sigma_{r_1}$ are the real embeddings.

Definition

The *Minkowski embedding* of K is

$$\begin{aligned}\iota : K &\to \mathbb{R}^n \\ x &\mapsto (\sigma_1(x), \cdots, \sigma_{r_1}(x), \mathrm{Re}(\sigma_{r_1+1})(x), \mathrm{Im}(\sigma_{r_1+1})(x), \cdots, \\ &\quad \mathrm{Re}(\sigma_{r_1+r_2})(x), \mathrm{Im}(\sigma_{r_1+r_2})(x))\end{aligned}$$

It turns out that

$$\Lambda_R := \iota(R)$$

is a lattice in $\mathbb{R}^n$, we call it the *embedded lattice of K*.

Discrete Gaussian distribution on lattices

For $\sigma > 0$, define the Gaussian function ρ_σ as

$$\rho_\sigma(x) = e^{-||x||^2/2\sigma^2}.$$

Discrete Gaussian distribution on lattices

For $\sigma > 0$, define the Gaussian function ρ_σ as

$$\rho_\sigma(x) = e^{-||x||^2/2\sigma^2}.$$

Definition

For a lattiace $\Lambda \subset \mathbb{R}^n$ and $\sigma > 0$, the *discrete Gaussian distribution on* Λ *with parameter* σ is

$$D_{\Lambda,\sigma}(x) = \frac{\rho_\sigma(x)}{\sum_{y\in\Lambda} \rho_\sigma(y)}, \ \forall x \in \Lambda.$$

Equivalently, the probability of sampling any lattice point x is proportional to $\rho_\sigma(x)$.

RLWE instance

Definition

An *RLWE instance* is a tuple $\mathcal{R} = (K, q, \sigma, s)$, where K is a number field, q is a prime, $\sigma > 0$, and s is an element of R/qR (s is the *secret*).

RLWE instance

Definition

An *RLWE instance* is a tuple $\mathcal{R} = (K, q, \sigma, s)$, where K is a number field, q is a prime, $\sigma > 0$, and s is an element of R/qR (s is the *secret*).

Definition

Let $\mathcal{R} = (K, q, \sigma, s)$ be an RLWE instance and let R be the ring of integers of K. The *error distribution* of $\mathcal{R}$, denote by $D_{\mathcal{R}}$, is the discrete Gaussian on the embedded lattice $\iota(R)$ with parameter σ:

$$D_{\mathcal{R}} = D_{\iota(R),\sigma}.$$

RLWE instance

Definition

An *RLWE instance* is a tuple $\mathcal{R} = (K, q, \sigma, s)$, where K is a number field, q is a prime, $\sigma > 0$, and s is an element of R/qR (s is the *secret*).

Definition

Let $\mathcal{R} = (K, q, \sigma, s)$ be an RLWE instance and let R be the ring of integers of K. The *error distribution* of $\mathcal{R}$, denote by $D_{\mathcal{R}}$, is the discrete Gaussian on the embedded lattice $\iota(R)$ with parameter σ:

$$D_{\mathcal{R}} = D_{\iota(R),\sigma}.$$

Remark: let V denote the covolume of the lattice $\iota(R)$. It is convenient to define a relative standard deviation parameter: $\sigma_0 = \frac{\sigma}{V^{\frac{1}{n}}}$.

RLWE samples

Definition (RLWE distribtuion)

Let $\mathcal{R} = (K, q, \sigma, s)$ be an RLWE instance with error distribution $D_{\mathcal{R}}$, and let R_q denote the quotient ring R/qR. Then a sample from the *RLWE distribtuion* of $\mathcal{R}$ is an ordered pair

$$(a, b = as + e \pmod{qR}) \in R_q \times R_q,$$

where the first coordiante a is chosen uniformly at random in R_q, and $e \leftarrow D_{\mathcal{R}}$.

Notation: $(a, b) \leftarrow \mathcal{R}$ means (a, b) is a sample from the RLWE distribution of $\mathcal{R}$.

RLWE problems

The RLWE problem has two variants: search and decision.

RLWE problems

The RLWE problem has two variants: search and decision.

Definition (Search)

Let $\mathcal{R}$ be an RLWE intance. The *search Ring-LWE* problem, denoted by SRLWE($\mathcal{R}$), is to discover s given access to arbitrarily many independent samples $(a, b) \leftarrow \mathcal{R}$.

RLWE problems

The RLWE problem has two variants: search and decision.

Definition (Search)

Let $\mathcal{R}$ be an RLWE intance. The *search Ring-LWE* problem, denoted by SRLWE($\mathcal{R}$), is to discover s given access to arbitrarily many independent samples $(a, b) \leftarrow \mathcal{R}$.

Definition (Decision)

Let $\mathcal{R}$ be an RLWE intance. The *decision Ring-LWE* problem, denoted by DRLWE($\mathcal{R}$), is to distinguish between the same number of independent samples in two distributions on $R_q \times R_q$. The first is the RLWE distribution of $\mathcal{R}$, and the second consists of uniformly random and independent samples from $R_q \times R_q$.

Part 2/4: The chi-square attack

Recap the [ELOS] attack

The [ELOS] attack picks a prime ideal $\mathfrak{q}$ above q and uses the reduction map

$$\pi : R/qR \to R/\mathfrak{q}R : \quad x \mapsto x \pmod{\mathfrak{q}}.$$

Recap the [ELOS] attack

The [ELOS] attack picks a prime ideal $\mathfrak{q}$ above q and uses the reduction map

$$\pi : R/qR \to R/\mathfrak{q}R : \quad x \mapsto x \pmod{\mathfrak{q}}.$$

It runs through possible guesses g of $\pi(s)$, and computes the "error"

$$\text{``}\pi(e)'' = \pi(b) - \pi(a) \cdot g.$$

Then it marks the correct guess based on the assumption that the distribution $\pi(D_{\Lambda_R,\sigma})$ is distinguishable from the uniform distribution over the finite field $F := R/\mathfrak{q}R$.

Recap the [ELOS] attack

The [ELOS] attack picks a prime ideal $\mathfrak{q}$ above q and uses the reduction map

$$\pi : R/qR \to R/\mathfrak{q}R : \quad x \mapsto x \pmod{\mathfrak{q}}.$$

It runs through possible guesses g of $\pi(s)$, and computes the "error"

$$\text{``}\pi(e)\text{''} = \pi(b) - \pi(a) \cdot g.$$

Then it marks the correct guess based on the assumption that the distribution $\pi(D_{\Lambda_R,\sigma})$ is distinguishable from the uniform distribution over the finite field $F := R/\mathfrak{q}R$.

For the vulnerable instances found in [ELOS], one has $|\pi(D_{\Lambda_R,\sigma})| \ll |F|$, making distinguishing an easy task.

Recap the [ELOS] attack

The [ELOS] attack picks a prime ideal $\mathfrak{q}$ above q and uses the reduction map

$$\pi : R/qR \to R/\mathfrak{q}R : \quad x \mapsto x \pmod{\mathfrak{q}}.$$

It runs through possible guesses g of $\pi(s)$, and computes the "error"

$$``\pi(e)'' = \pi(b) - \pi(a) \cdot g.$$

Then it marks the correct guess based on the assumption that the distribution $\pi(D_{\Lambda_R,\sigma})$ is distinguishable from the uniform distribution over the finite field $F := R/\mathfrak{q}R$.

For the vulnerable instances found in [ELOS], one has $|\pi(D_{\Lambda_R,\sigma})| \ll |F|$, making distinguishing an easy task.

However, for Galois extensions, these examples are harder to find. So a new attack is needed.

Background on chi-sqaure test

Let S be a finite set partitioned into r subsets: $S = \sqcup_{j=1}^{r} S_j$. Given M samples $y_1, \cdots, y_M$ in S.
Null hypothesis: the samples are from taken the uniform distribution on S.
we compute the expected and the actual number of samples that lie in each subset. Then the χ^2 value is

$$\chi^2(S, y) = \sum_{j=1}^{r} \frac{(actual_j - expect_j)^2}{expect_j}.$$

Background on chi-sqaure test

Let S be a finite set partitioned into r subsets: $S = \sqcup_{j=1}^{r} S_j$. Given M samples $y_1, \cdots, y_M$ in S.
Null hypothesis: the samples are from taken the uniform distribution on S. we compute the expected and the actual number of samples that lie in each subset. Then the χ^2 value is

$$\chi^2(S, y) = \sum_{j=1}^{r} \frac{(actual_j - expect_j)^2}{expect_j}.$$

If the samples were drawn from the uniform distribution on S, then the χ^2 value follows the chi-square distribution with degree of freedom $d = r - 1$. Hence we may use this to test uniform distribution.

Idea of our attack

The goal of our chi-square attack is to recover $s \pmod{\mathfrak{q}}$ from a set of samples $(a, b) \leftarrow \mathcal{R}$.

Idea of our attack

The goal of our chi-square attack is to recover $s \pmod{\mathfrak{q}}$ from a set of samples $(a, b) \leftarrow \mathcal{R}$.

1. For each guess s' of $s \pmod{\mathfrak{q}}$:
   - compute the "errors" $e' = b \pmod{\mathfrak{q}} - a \pmod{\mathfrak{q}} s'$ for all samples (a, b).
   - run the chi-square uniform test on the "errors" e'.
   - accept s' as a good guess if the test rejects the uniform hypothesis.

Idea of our attack

The goal of our chi-square attack is to recover $s \pmod{\mathfrak{q}}$ from a set of samples $(a, b) \leftarrow \mathcal{R}$.

1. For each guess s' of $s \pmod{\mathfrak{q}}$:
   - compute the "errors" $e' = b \pmod{\mathfrak{q}} - a \pmod{\mathfrak{q}} s'$ for all samples (a, b).
   - run the chi-square uniform test on the "errors" e'.
   - accept s' as a good guess if the test rejects the uniform hypothesis.
2. Repeat (1) with more samples and the set of good guesses until there is only one good guess s_g left, and ouput s_g. (If there is no good guess left, output *fail*).

Idea of our attack

The goal of our chi-square attack is to recover $s \pmod{\mathfrak{q}}$ from a set of samples $(a, b) \leftarrow \mathcal{R}$.

1. For each guess s' of $s \pmod{\mathfrak{q}}$:
   - compute the "errors" $e' = b \pmod{\mathfrak{q}} - a \pmod{\mathfrak{q}} s'$ for all samples (a, b).
   - run the chi-square uniform test on the "errors" e'.
   - accept s' as a good guess if the test rejects the uniform hypothesis.
2. Repeat (1) with more samples and the set of good guesses until there is only one good guess s_g left, and ouput s_g. (If there is no good guess left, output *fail*).

The complexity of our attack is $O(q^f)$ in time and $O(q^f)$ in space.

Algorithm 1 chi-square attack of $SRLWE(\mathcal{R}, \mathfrak{q})$

Require: $\mathcal{R} = (K, q, \sigma, s)$ – an RLWE instance. $\mathfrak{q}$ – a prime ideal in K above q. S – a collection of M ($M = \Omega(N)$) RLWE samples $(a, b) \sim \mathcal{R}$.

Ensure: a guess of the value $s \pmod{\mathfrak{q}}$, or **NON-RLWE**, or **INSUFFIICNET-SAMPLES**

$\alpha \leftarrow 1 - \frac{1}{10N}$, $\omega \leftarrow \Phi^{-1}((1+\alpha)/2)$, $G = \emptyset$.
for s in F **do**
    $E \leftarrow [b \pmod{\mathfrak{q}} - a \pmod{\mathfrak{q}} s$ for a, b in $S]$.
end for
Run χ^2 test on E with B bins and obtain the value $\chi^2(E)$.
if $|\chi^2(E) - (B-1)| > \omega\sqrt{2B-2}$ **then**
    add s to G
end if
if $G = \emptyset$ **then return NOT RLWE**
else if $G = \{g\}$ **then return** g
else return INSUFFIICNET-SAMPLES
end if

Part 3/4: Galois RLWE

Notations and properties

Recall: a number field K is Galois if $|Aut(K)| = [K : \mathbb{Q}]$.

Notations and properties

Recall: a number field K is Galois if $|Aut(K)| = [K : \mathbb{Q}]$.

RLWE instances with K Galois have nice properties.

Notations and properties

Recall: a number field K is Galois if $|Aut(K)| = [K : \mathbb{Q}]$.

RLWE instances with K Galois have nice properties.

Theorem (Search-to-Decision)

Let $\mathcal{R} = (K, q, \sigma, s)$ be an RLWE instance, where K is Galois of degree n and q is unramified in K with residue degree f. Suppose there is an algorithm A which recovers $s \pmod{\mathfrak{q}}$ for some prime $\mathfrak{q}$ above q using a set S of samples. Then the problem $\mathsf{SRLWE}(\mathcal{R})$ *can be solved using n/f calls to A and the same set S of samples.*

Notations and properties

Recall: a number field K is Galois if $|Aut(K)| = [K : \mathbb{Q}]$.

RLWE instances with K Galois have nice properties.

Theorem (Search-to-Decision)

Let $\mathcal{R} = (K, q, \sigma, s)$ be an RLWE instance, where K is Galois of degree n and q is unramified in K with residue degree f. Suppose there is an algorithm A which recovers $s \pmod{\mathfrak{q}}$ for some prime $\mathfrak{q}$ above q using a set S of samples. Then the problem $\mathrm{SRLWE}(\mathcal{R})$ can be solved using n/f calls to A and the same set S of samples.

Theorem (Independence of $\mathfrak{q}$)

Keeping the above assumptions. Then the error distribution $D(\mathcal{R}, \mathfrak{q}) := D_{\mathcal{R}} \pmod{\mathfrak{q}}$ is independent of the choice of prime ideal $\mathfrak{q}$ above q.

Vulnerable instances

We consider subfields of form $K_{m,H} = \mathbb{Q}(\zeta_m)^H$, where $H \leq (\mathbb{Z}/m\mathbb{Z})^*$.

Vulnerable instances

We consider subfields of form $K_{m,H} = \mathbb{Q}(\zeta_m)^H$, where $H \leq (\mathbb{Z}/m\mathbb{Z})^*$.

Notations: f – the residue degree of q; M – the number of samples; σ_0 – the relative standard deviation parameter.

Vulnerable instances

We consider subfields of form $K_{m,H} = \mathbb{Q}(\zeta_m)^H$, where $H \leq (\mathbb{Z}/m\mathbb{Z})^*$.

Notations: f – the residue degree of q; M – the number of samples; σ_0 – the relative standard deviation parameter.

Table: Attacked sub-cyclotomic RLWE instances

m	gens of H	n	q	f	σ_0	M	runtime (in hours)
2805	[1684, 1618]	40	67	2	1	22445	3.49
15015	[12286, 2003, 11936]	60	43	2	1	11094	1.05
90321	[90320, 18514, 43405]	80	67	2	1	26934	4.81
255255	[97943, 162436, 253826, 248711, 44318])	90	2003	2	1.25	15000	1114.44 (estimated)
285285	[181156, 210926, 87361]	96	521	2	1.1	5000	75.41 (estimated)
1468005	[198892, 978671, 431521, 1083139]	144	139	2	1	4000	5.72

Why are higher degree primes vulnerable?

Imagine the following (unlikely) scenario that Λ_R has an orthogonal basis $v_1, \cdots, v_n$ such that $||v_i||_2 \ll ||v_{i+1}||_2$ for all i.

Why are higher degree primes vulnerable?

Imagine the following (unlikely) scenario that Λ_R has an orthogonal basis $v_1, \cdots, v_n$ such that $||v_i||_2 \ll ||v_{i+1}||_2$ for all i.

Q: What happens if for some prime q of degree 2, the first 90 percent of $v_i \pmod{\mathfrak{q}}$ lie in the prime subfield F_q instead of F_{q^2}?

Why are higher degree primes vulnerable?

Imagine the following (unlikely) scenario that Λ_R has an orthogonal basis $v_1, \cdots, v_n$ such that $||v_i||_2 \ll ||v_{i+1}||_2$ for all i.

Q: What happens if for some prime q of degree 2, the first 90 percent of $v_i \pmod{\mathfrak{q}}$ lie in the prime subfield F_q instead of F_{q^2}?

A: $D_{\Lambda_R,\sigma} \pmod{\mathfrak{q}}$ will take on values in F_q with significantly higher probability. Hence we can distinguish it from uniform $U(F_{q^2})$!

Why are higher degree primes vulnerable?

Imagine the following (unlikely) scenario that Λ_R has an orthogonal basis $v_1, \cdots, v_n$ such that $||v_i||_2 \ll ||v_{i+1}||_2$ for all i.

Q: What happens if for some prime q of degree 2, the first 90 percent of $v_i \pmod{\mathfrak{q}}$ lie in the prime subfield F_q instead of F_{q^2}?

A: $D_{\Lambda_R,\sigma} \pmod{\mathfrak{q}}$ will take on values in F_q with significantly higher probability. Hence we can distinguish it from uniform $U(F_{q^2})$!

The real situation is similar to the hypothetical one above.

Why are higher degree primes vulnerable?

Imagine the following (unlikely) scenario that Λ_R has an orthogonal basis $v_1, \cdots, v_n$ such that $||v_i||_2 \ll ||v_{i+1}||_2$ for all i.

Q: What happens if for some prime q of degree 2, the first 90 percent of $v_i \pmod{\mathfrak{q}}$ lie in the prime subfield F_q instead of F_{q^2}?

A: $D_{\Lambda_R,\sigma} \pmod{\mathfrak{q}}$ will take on values in F_q with significantly higher probability. Hence we can distinguish it from uniform $U(F_{q^2})$!

The real situation is similar to the hypothetical one above.
One could optimise the attack based on this observation, reducing the space complexity to $O(q)$.

A detailed example

We demonstrate search-to-decision and the "degree 2" phenomenon with an example:

A detailed example

We demonstrate search-to-decision and the "degree 2" phenomenon with an example:

- $m = 3003$, $H = \langle 2276, 2729, 1123 \rangle$, $n = 30$, $q = 131$, $f = 2$, $\sigma_0 = 1$.

A detailed example

We demonstrate search-to-decision and the "degree 2" phenomenon with an example:

- $m = 3003$, $H = \langle 2276, 2729, 1123 \rangle$, $n = 30$, $q = 131$, $f = 2$, $\sigma_0 = 1$.
- There are $g = n/f = 15$ prime ideals $\mathfrak{q}_1, \cdots, \mathfrak{q}_{15}$ above q.

A detailed example

We demonstrate search-to-decision and the “degree 2” phenomenon with an example:

- $m = 3003$, $H = \langle 2276, 2729, 1123 \rangle$, $n = 30$, $q = 131$, $f = 2$, $\sigma_0 = 1$.
- There are $g = n/f = 15$ prime ideals $\mathfrak{q}_1, \cdots, \mathfrak{q}_{15}$ above q.
- We use LLL algorithm on a given basis and obtained a reducebasis $v_1, \cdots, v_n$ for R, ordered by length, out of which $v_1, \cdots, v_{n/2}$ (mod $\mathfrak{q}_i$) lie in the smaller field $\mathbb{F}_q$ (and the rest lie in $\mathbb{F}_{q^2} - \mathbb{F}_q$).

A detailed example

We demonstrate search-to-decision and the "degree 2" phenomenon with an example:

- $m = 3003$, $H = \langle 2276, 2729, 1123 \rangle$, $n = 30$, $q = 131$, $f = 2$, $\sigma_0 = 1$.
- There are $g = n/f = 15$ prime ideals $\mathfrak{q}_1, \cdots, \mathfrak{q}_{15}$ above q.
- We use LLL algorithm on a given basis and obtained a reducebasis $v_1, \cdots, v_n$ for R, ordered by length, out of which $v_1, \cdots, v_{n/2} \pmod{\mathfrak{q}_i}$ lie in the smaller field $\mathbb{F}_q$ (and the rest lie in $\mathbb{F}_{q^2} - \mathbb{F}_q$).
- We run the attack to recover $s \pmod{\mathfrak{q}_i}$ for each $1 \leq i \leq g$. Then we use Chinese remainder theorem to recover the whole key $s \in R/qR$.

A detailed example

We demonstrate search-to-decision and the "degree 2" phenomenon with an example:

- $m = 3003$, $H = \langle 2276, 2729, 1123 \rangle$, $n = 30$, $q = 131$, $f = 2$, $\sigma_0 = 1$.
- There are $g = n/f = 15$ prime ideals $\mathfrak{q}_1, \cdots, \mathfrak{q}_{15}$ above q.
- We use LLL algorithm on a given basis and obtained a reducebasis $v_1, \cdots, v_n$ for R, ordered by length, out of which $v_1, \cdots, v_{n/2}$ $\pmod{\mathfrak{q}_i}$ lie in the smaller field $\mathbb{F}_q$ (and the rest lie in $\mathbb{F}_{q^2} - \mathbb{F}_q$).
- We run the attack to recover $s \pmod{\mathfrak{q}_i}$ for each $1 \leq i \leq g$. Then we use Chinese remainder theorem to recover the whole key $s \in R/qR$.

Result: we used 1000 samples; the attack succeeded in 32.8 hours.

A detailed example

We demonstrate search-to-decision and the "degree 2" phenomenon with an example:

- $m = 3003$, $H = \langle 2276, 2729, 1123 \rangle$, $n = 30$, $q = 131$, $f = 2$, $\sigma_0 = 1$.
- There are $g = n/f = 15$ prime ideals $\mathfrak{q}_1, \cdots, \mathfrak{q}_{15}$ above q.
- We use LLL algorithm on a given basis and obtained a reducebasis $v_1, \cdots, v_n$ for R, ordered by length, out of which $v_1, \cdots, v_{n/2} \pmod{\mathfrak{q}_i}$ lie in the smaller field $\mathbb{F}_q$ (and the rest lie in $\mathbb{F}_{q^2} - \mathbb{F}_q$).
- We run the attack to recover $s \pmod{\mathfrak{q}_i}$ for each $1 \leq i \leq g$. Then we use Chinese remainder theorem to recover the whole key $s \in R/qR$.

Result: we used 1000 samples; the attack succeeded in 32.8 hours.

Remark: the last step is parallelizable.

Part 4/4: Cyclotomics

Background on Fourier analysis

Suppose f is a real-valued function on $\mathbb{F}_q$. The *Fourier transform* of f is defined as

$$\hat{f}(y) = \sum_{a \in \mathbb{F}_q} f(a) e^{-2\pi i a y / q}.$$

Background on Fourier analysis

Suppose f is a real-valued function on $\mathbb{F}_q$. The *Fourier transform* of f is defined as

$$\hat{f}(y) = \sum_{a \in \mathbb{F}_q} f(a) e^{-2\pi i a y / q}.$$

Let δ be the dirac delta function and $u \equiv 1/q$.

Fact (Properties of Fourier transform)

1. $\hat{\delta} = qu$, $\hat{u} = \delta$.
2. $\widehat{f * g} = \hat{f} \cdot \hat{g}$.
3. $f(a) = \frac{1}{q} \sum_{y \in \mathbb{F}_q} \hat{f}(y) e^{2\pi i a y / q}$.

A simplified error distribution

Definition

For any even integer $k \geq 2$, let $\mathcal{V}_k$ denote the distribution over $\mathbb{Z}$ such that

$$\mathrm{Prob}(\mathcal{V}_k = t) = \begin{cases} \frac{\binom{k}{t+\frac{k}{2}}}{2^k} & \text{if } |t| \leq \frac{k}{2} \\ 0 & \text{otherwise.} \end{cases}$$

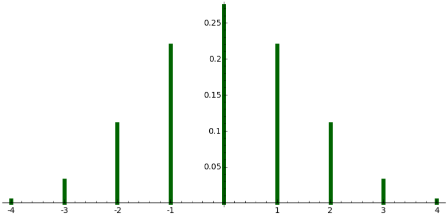

Figure: Probability density function of $\mathcal{V}_8$

Modified error distribution

Definition (Modified error distribtuion $MD_{m,q,k}$)

Let m, q be integers such that $q \equiv 1 \pmod{m}$ and let $k \geq 2$ be even. Then a sample from the *modified error distribtuion* $MD_{m,q,k}$ is

$$e = \sum_{i=0}^{n-1} e_i \zeta_m^i \pmod{qR},$$

where the coefficients e_i are sampled i.i.d. from $\mathcal{V}_k$.

Modified error distribution

Definition (Modified error distribtuion $MD_{m,q,k}$)

Let m, q be integers such that $q \equiv 1 \pmod{m}$ and let $k \geq 2$ be even. Then a sample from the *modified error distribtuion* $MD_{m,q,k}$ is

$$e = \sum_{i=0}^{n-1} e_i \zeta_m^i \pmod{qR},$$

where the coefficients e_i are sampled i.i.d. from $\mathcal{V}_k$.

Let α be a primitive m-th root of unity in $\mathbb{F}_q$, corresponding to a prime $\mathfrak{q}$. Then

$$\bar{e} = e \pmod{\mathfrak{q}} = \sum_i e_i \alpha^i.$$

Note that $\bar{e}$ is a random variable with value in $\mathbb{F}_q$. We abuse notations and let $\bar{e}$ denote its own probability density function.

Cyclotomics

Lemma

$$\widehat{e}(y) = \prod_{i=1}^{n} \cos\left(\frac{\alpha^i \pi y}{q}\right)^k .$$

Cyclotomics

Lemma

$$\widehat{\bar{e}}(y) = \prod_{i=1}^{n} \cos\left(\frac{\alpha^i \pi y}{q}\right)^k .$$

Theorem

For all $a \in \mathbb{F}_q$,

$$|\bar{e}(a) - 1/q| \leq \frac{1}{q} \sum_{y \in \mathbb{F}_q, y \neq 0} |\hat{\bar{e}}(y)|. \quad (4.1)$$

Cyclotomics

Lemma

$$\widehat{\bar{e}}(y) = \prod_{i=1}^{n} \cos\left(\frac{\alpha^i \pi y}{q}\right)^k .$$

Theorem

For all $a \in \mathbb{F}_q$,

$$|\bar{e}(a) - 1/q| \leq \frac{1}{q} \sum_{y \in \mathbb{F}_q, y \neq 0} |\hat{\bar{e}}(y)|. \tag{4.1}$$

Corollary

Let $u \equiv 1/q$ denote the p.d.f. for the uniform distribution, then

$$d(\bar{e}, u) \leq \frac{1}{2} \sum_{y \in \mathbb{F}_q, y \neq 0} |\hat{\bar{e}}(y)| =: \epsilon(m, q, k, \alpha).$$

A table of ϵ values

Let $\epsilon(m, q, k) = \max\{\epsilon(m, q, k, \alpha) : \alpha$ has order m in $\mathbb{F}_q\}$.

A table of ϵ values

Let $\epsilon(m, q, k) = \max\{\epsilon(m, q, k, \alpha) : \alpha$ has order m in $\mathbb{F}_q\}$.

Punchline: the value $\epsilon(m, q, k)$ is usually negligibly small. As a consequence, the reduced error distribution $\bar{e}$ is computationally indisinguishable from uniform distribution.

Hence these instances are secure agianst our chi-square attack.

A table of ϵ values

Let $\epsilon(m, q, k) = \max\{\epsilon(m, q, k, \alpha) : \alpha$ has order m in $\mathbb{F}_q\}$.

Punchline: the value $\epsilon(m, q, k)$ is usually negligibly small. As a consequence, the reduced error distribution $\bar{e}$ is computationally indisinguishable from uniform distribution.

Hence these instances are secure agianst our chi-square attack.

Table: Values of $\epsilon(m, q, 2)$

m	n	q	$-[\log_2(\epsilon(m, q, 2))]$
244	120	1709	230
101	100	1213	177
256	128	3329	194
256	128	14081	208
55	40	10891	44
197	196	3547	337
96	32	4513	35
160	64	20641	61
145	112	19163	176
512	256	10753	431
512	256	19457	414

Ramified prime (is vulnerable)

We consider $K = \mathbb{Q}(\zeta_p)$ and $q = p$. Then there is a unique prime ideal $\mathfrak{p} = (1 - \zeta_p)$ above p, and the reduction map $\pi : R/pR \to \mathbb{F}_p$ satisfies

$$\pi(\zeta_p^i) = 1, \quad \forall i.$$

We will exploit this property for our attack.

PLWE on power basis

The error is $e = \sum_{i=0}^{p-2} e_i \zeta_p^i$, where $e_i \sim D_{\mathbb{Z},\sigma}$ i.i.d.

PLWE on power basis

The error is $e = \sum_{i=0}^{p-2} e_i \zeta_p^i$, where $e_i \sim D_{\mathbb{Z},\sigma}$ i.i.d.

We have $e \pmod{\mathfrak{p}} = \sum_i e_i$, and when $p \gg 1$, by central limit theorem,

$$\sum_i e_i \to N(0, \sigma^2(p-1)).$$

PLWE on power basis

The error is $e = \sum_{i=0}^{p-2} e_i \zeta_p^i$, where $e_i \sim D_{\mathbb{Z},\sigma}$ i.i.d.

We have $e \pmod{\mathfrak{p}} = \sum_i e_i$, and when $p \gg 1$, by central limit theorem,

$$\sum_i e_i \to N(0, \sigma^2(p-1)).$$

Hence $|e \pmod{\mathfrak{p}}| \leq \sqrt{2\pi}\sigma\sqrt{p-1}$ with overwhelming probability, so [ELOS] attack will work.

PLWE on power basis

The error is $e = \sum_{i=0}^{p-2} e_i \zeta_p^i$, where $e_i \sim D_{\mathbb{Z},\sigma}$ i.i.d.

We have $e \pmod{\mathfrak{p}} = \sum_i e_i$, and when $p \gg 1$, by central limit theorem,

$$\sum_i e_i \to N(0, \sigma^2(p-1)).$$

Hence $|e \pmod{\mathfrak{p}}| \leq \sqrt{2\pi}\sigma\sqrt{p-1}$ with overwhelming probability, so [ELOS] attack will work.
Remark: we actually want to attack RLWE examples. So I tried chi-square attack and RLWE errors generated by the sampling method in [GPV].

Attacked ramified prime for prime cyclotomic RLWE

Examples:

- $p = 251$, $\sigma = 0.55$.
- (ongoing)

Thank you to everyone for a great summer!