Crypanalysis of some Lattice-based Assumptions

Dipayan Das (joining NTT Japan in December)
Lattice-based cryptography

• Post-quantum candidate.
• Worst-case to average-case reductions (in asymptotic sense).
• Advanced cryptographic primitives (like FHE).

NIST standardized lattice-based algorithms for quantum-resistant cryptography (July, 2022).

More details, please visit:


The availability of a quantum computer is altogether a different question 😊
Lattice-based assumptions

Cryptography relies on the assumptions of computationally hard problems.

**Lattice-based assumptions**: The best known way to solve it is by lattice methods through a transformation to a lattice problem.

**Talk overview**: This doesn’t always guarantee hardness (by counterexamples).

Lattice methods might not be the optimal strategy to approach it.
A full rank matrix $B \in \mathbb{Z}^{n \times n}$ generates a **Lattice** $L = L(B) = \{Bz : z \in \mathbb{Z}^n\}$

- This lattice has $\text{dim} = n$ and $\text{Vol} = |\text{det}(B)|$
Algorithmic problem related to lattices

• Shortest (non-zero) vector problem (SVP)
• Minkowski’s theorem: Let $\nu$ be the SVP solution, then
  \[ ||\nu|| \leq \sqrt{nVol^n} \]
• In practice, we use lattice reduction algorithms to find approximate solutions.

\textbf{LLL}: Finds a lattice vector of norm $\leq 2^{\frac{n}{2}} Vol^{\frac{1}{n}}$ in polynomial time in the size of its input.

\textbf{BKZ with block size $\beta$}: Finds a lattice vector of norm $\leq \beta^\beta Vol^{\frac{1}{n}}$ in time $2^{O(\beta)}$. 
Cryptanalysis of the Finite Field Isomorphism problem

Based on the work: D. Das, A. Joux. On the Hardness of the Finite Field Isomorphism Problem. EUROCRYPT’23
Reminders from Finite field theory

• Finite field with $q$ elements: $F_q$, where $q$ is prime.

• Finite field with $q^n$ elements ($n$ degree extension of $F_q$): $F_{q^n}$

• Isomorphic representations of $F_{q^n}$ using irreducible polynomials of degree $n$ over $F_q$

$$F_q[x]/f(x) \approx F_q[y]/F(y) \approx ...$$

• To find an explicit isomorphism, it is enough to know the roots of one polynomial in $F_{q^n}$ in terms of the other representation
Finite Field Isomorphism (FFI) Distribution

<table>
<thead>
<tr>
<th>Private:</th>
<th>Public:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform Sparse ternary minimal polynomial of $x$: $f(x) = x^n + g(x)$, $\deg(g) \leq \frac{n}{2}$</td>
<td>Uniform minimal polynomial of $y$: $F(y)$</td>
</tr>
<tr>
<td>Pick an Isomorphism: $\phi$</td>
<td></td>
</tr>
<tr>
<td>Sample $\beta$-bounded linear combinations of powers of $x$: $a_i(x)$</td>
<td>$A_i(y) = \phi(a_i(x))$</td>
</tr>
</tbody>
</table>

Good Representation in polynomial $x$—basis

Bad Representation in polynomial $y$—basis
FFI problem [DHP+'18,HSWZ’20]

Given $q, F(y), A_1(y), A_2(y), ..., A_k(y)$ decide if $A_i(y)$ is from the FFI distribution or the uniform distribution.

This is the Decisional FFI (DFFI) problem.


Toy example

\[ n=16 \]
\[ q=32771 \]
\[ f(x)=x^{16} + x^7 + x^5 - x^3 - x^2 - x + 1 \]
\[ F(y)=y^{16} + 4152y^{15} + 2594y^{14} + 26843y^{13} + 27498y^{12} + 31444y^{11} + 15956y^{10} + 7616y^9 + 30326y^8 + 26759y^7 + 8558y^6 + 4785y^5 + 27721y^4 + 1190y^3 + 14992y^2 + 14544y + 11277 \]
\[ \phi(x)=28228y^{15} + 13643y^{14} + 21168y^{13} + 4909y^{12} + 25475y^{11} + 21646y^{10} + 23297y^9 + 19665y^8 + 5019y^7 + 1677y^6 + 6823y^5 + 15399y^4 + 23062y^3 + 242y^2 + 18578y + 31824 \]

**x-basis representation**

\[ x^{14} + x^{12} + x^{10} + x^9 + x^8 - x^7 - x^5 - x^4 - x^3 - x \]

\[ 28795y^{15} + 787y^{14} + 4649y^{13} + 30560y^{12} + 21773y^{11} + 19702y^{10} + 14924y^9 + 22468y^8 + 39755y^7 + 7215y^6 + 5478y^5 + 4488y^4 + 9598y^3 + 3290y^2 + 1955y + 25737 \]

\[ x^{13} - x^{12} + x^{10} - x^9 + x^5 - x^4 + x^3 - x^2 - x + 1 \]

\[ 22173y^{15} + 15726y^{14} + 3731y^{13} + 2685y^{12} + 29516y^{11} + 30642y^{10} + 8601y^9 + 12332y^8 + 6722y^7 + 3340y^6 + 28353y^5 + 9853y^4 + 32035y^3 + 25337y^2 + 19076y + 29241 \]

\[ -x^{15} + x^{12} - x^{11} - x^{10} + x^9 - x^6 - x^5 - x^3 - x^2 - x - 1 \]

\[ 25606y^{15} + 24744y^{14} + 20203y^{13} + 1563y^{12} + 10690y^{11} + 20596y^{10} + 22744y^9 + 30883y^8 + 16058y^7 + 10331y^6 + 30479y^5 + 27544y^4 + 19820y^3 + 3869y^2 + 6933y + 2377 \]
Previous attack on Decisional FFI problem [DHP+’18, HSWZ’20]

Lattice attack

Find unusually short lattice vectors of the lattice $L \subseteq \mathbb{Z}^k$ spanned by the columns $A_i(y)$.
FHE from FFI problem (oversimplified) [DHP+’18]

- Let $p = 2$
- $m_a, m_b \in \{0,1\}$
- $\text{Enc}(m_a) = C_a = pC(y) + m_a$, $\text{Enc}(m_b) = C_b = pC'(y) + m_b$
- $\text{Dec}(C_a) = (pc(x) + m_a) \mod p = m_a$
- $\text{Dec}(C_a + C_b) = (pc(x) + pc'(x) + m_a + m_b) \mod p = m_a + m_b$
- $\text{Dec}(C_a \cdot C_b) = (p^2c(x)c'(x) + pc(x)m_b + pc'(x)m_a + m_a \cdot m_b) \mod p = m_a \cdot m_b$
- Correctness: Choose $q$ sufficiently large to avoid modular reductions in $x$-basis representations
- When $q = 2^{n \delta}$, $\delta \in (0,1)$, the Encryption scheme is FHE [DHP+18]
Trace of finite field

- Let $\alpha \in F_{q^n}$, trace is defined by
  \[ \text{Tr}(\alpha) = \alpha + \alpha^q + \cdots + \alpha^{q^{n-1}} \in F_q \]
- Trace is linear.
- Trace computation is polynomial time.
- Trace is invariant under basis representations.
Symmetric polynomials

- Roots of $f(x)$ in $F_{q^n}$ (in terms of polynomial $x$-basis):
  \[
  \{\alpha_0 = x, \alpha_1 = x^q, \ldots, \alpha_{n-1} = x^{q^{n-1}}\}
  \]

- Define Symmetric polynomials
  \[
  \sigma_1(\alpha_i) = -\sum \alpha_i, \sigma_2(\alpha_i) = \sum \sum \alpha_i \alpha_j, \ldots, \sigma_n(\alpha_i) = (-1)^n \prod \alpha_i
  \]
Trace of polynomial $x$-basis

\[ f(x) = x^n + \sigma_1 x^{n-1} + \cdots + \sigma_n \] where $\sigma_d = 0$ for $1 \leq d \leq \frac{n}{2} - 1$

\[ \sigma_d \in \{0, \pm 1\} \text{ for } \frac{n}{2} \leq d \leq n \]

Then

\[ |Tr(x^d)| = n \mod q \text{ for } d = 0 \]

\[ = 0 \mod q \text{ for } 1 \leq d \leq \frac{n}{2} - 1 \]

\[ = d \mod q \text{ for } \frac{n}{2} \leq d \leq n - 1 \text{ and } \sigma_d \neq 0 \]

\[ = 0 \mod q \hspace{1cm} \sigma_d = 0 \]
Trace of polynomial $x$-basis

\[ f(x) = x^n + \sigma_1 x^{n-1} + \cdots + \sigma_n \text{ where } \sigma_d = 0 \text{ for } 1 \leq d \leq \frac{n}{2} - 1 \]

\[ \sigma_d \in \{0, \pm 1\} \text{ for } \frac{n}{2} \leq d \leq n \]

- Then for $1 \leq d \leq \frac{n}{2} - 1$
  - $\sigma_d = 0$
  - $Tr(x^d) = 0 \ mod \ q$

Using Newton-Girard formula:

\[ Tr(x^d) = (-1)^d d \sum_{r \in \mathbb{N} : r_1 + 2r_2 + \cdots + dr_d = d} \frac{(r_1 + r_2 + \cdots + r_d - 1)!}{r_1! r_2! \cdots r_d!} \prod_{j=1}^{d} (-\sigma_j)^{r_j} \]
Trace of polynomial $x$-basis

$f(x) = x^n + \sigma_1 x^{n-1} + \cdots + \sigma_n$ where $\sigma_d = 0$ for $1 \leq d \leq \frac{n}{2} - 1$

\[ \sigma_d \in \{0, \pm 1\} \text{ for } \frac{n}{2} \leq d \leq n \]

• Then for $\frac{n}{2} \leq d \leq n - 1$

Only one solution for $r_i: r_1 + 2r_2 + \cdots + dr_d = d$ that contributes in the sum:

\[(r_1 = 0, r_2 = 0, \ldots, r_d = 1)\]

\[|Tr(x^d)| = d \text{ mod } q \text{ when } \sigma_d \neq 0\]

\[= 0 \text{ mod } q \text{ when } \sigma_d = 0\]

Using Newton-Girard formula:

\[Tr(x^d) = (-1)^d d \sum_{r_i \in \mathbb{N}; r_1 + 2r_2 + \cdots + dr_d = d} \frac{(r_1 + r_2 + \cdots + r_d - 1)!}{r_1! r_2! \cdots r_d!} \prod_{j=1}^{d} (-\sigma_j)^{r_j}\]
Trace of FFI samples

• Let $a_i(x)$ is a $\beta$-linear combinations of $x$-basis. Then $|Tr(a_i(x))| = |Tr(A_i(y))| \leq \beta n^2$
Polynomial-time attack on DFFI problem

• Let \( q > 4\beta n^2 \)
• Let \( A_1(y), A_2(y), \ldots, A_k(y) \) be the given samples.

Compute the trace of the samples.

If the absolute value of traces \( \leq \beta n^2 \), output FFI distribution.

Otherwise, output uniform distribution.

• Advantage: \( 1 - \frac{1}{2^k} \)

Trace is uniformly distributed in \( F_q \) for uniform samples.
Polynomial-time semantic attack on the FHE

• Let $p$ is not a divisor of $n$
• $C_a = pC(y) + m$, where $m \in \{0,1\}$
• $Tr(C_a) = pTr(c(x)) + Tr(m)$ is small.

\[
\begin{align*}
Tr(C_a) \mod p &= 0, \text{Return } m = 0 \\
&= 1, \text{Return } m = 1
\end{align*}
\]
Polynomial-time semantic attack on the FHE

• Let $p$ is a divisor of $n$
• $C_a = pC(y) + m$, where $m \in \{0,1\}$
• Pick any FFI sample $C^*$ such that $p$ is not a divisor of $Tr(C^*)$
• $Tr(C_a \cdot C^*) = pTr(c^*(x) \cdot c(x)) + m Tr(c^*(x))$ is still small.

The choice of $f(x)$ makes sure the coefficients of the product in $x$-basis are small.

$Tr(C_a C^*) \mod p = 0$, Return $m = 0$

$= 1$, Return $m = 1$

• The large $q$ makes sure there is no modular reduction!
Cryptanalysis of the Partial Vandermonde Knapsack Problem

Based on the work: D. Das, A. Joux. Key Recovery Attack on the Partial Vandermonde Knapsack Problem. In submission
Partial Vandermonde (PV) Knapsack Problem

Let $R_q = F_q[x]/g(x)$ be a quotient polynomial ring, where

- $g(x) = x^n - 1$ for prime $n$
  $\quad = x^n + 1$ for power of two $n$

- Prime $q$ such that $g(x)$ splits linearly over $F_q$

  When $n$ is prime, $q = 1 \mod n$
  When $n$ is power-of-two, $q = 1 \mod 2n$

$\Omega$: The set of all the primitive roots of $g(x)$ over $F_q$
PV Knapsack Problem
[HPSSW’14, HS’15, DHSS’20, LZA’18, BSS’22]

- $\Omega_t$: Uniformly random subset of $\Omega$ with $t$ distinct elements.
- $f(x) \in R_q$: Coefficients are sampled uniformly at random from the set $\{-1, 0, 1\}$.

PV Knapsack problem:
Given $R_q$, $\Omega_t$, and $f(\omega)$ for $\omega \in \Omega_t$ find $f(x)$ when $t \approx \frac{n}{2}$.

Initially PV Knapsack problem was called the partial Fourier recovery problem.


Previous attack (Direct primal attack) [HPSSW’14]

for \( \omega \in \Omega_t \)

\[
\begin{bmatrix}
1 & \omega & \omega^2 & \ldots & \omega^{n-1} & f(\omega)
\end{bmatrix} \mod q
\]
Previous attack (Direct primal attack)[HPSSW’14]

• PV Knapsack problem: Find the uSVP solution \((f, -1)\) on the Kernel lattice

\[
L^\perp = \{ x \in \mathbb{Z}^{n+1} : Vx = 0 \mod q \}
\]

With \( \text{Dim} = n + 1 \) \( \text{Vol} = q^t \)

• \( ||(f, -1)|| \approx \sqrt{\frac{2n}{3}} \) which is unusually short in the lattice \( L^\perp \).
Previous attack (Dual attack)[BGP’22]

• Distinguishing attack

• Doesn’t affect the hardness of recovering $f$.

“We note however that this does not fully invalidate the claim made in [LZA18], since the 128 bit-security is claimed against search attackers, and not distinguishing attackers.” [BGP’22]

• The attack exploits specific Ideal structure of the problem to map to an SVP instance of smaller dimension.

Attack on the PV Knapsack problem

• For any \( f(x) \in R_q \), we can interpret \( f \left( \frac{1}{x} \right) \in R_q \)

• \( \frac{1}{x} = x^{n-1} \in R_q \) when \( n \) is prime.

• \( \frac{1}{x} = -x^{n-1} \in R_q \) when \( n \) is power-of-two.
Attack on the PV Knapsack problem

- Consider $\Omega_{2t_1} = \{\omega \in \Omega_t: (\omega, \omega^{-1}) \in \Omega_t\} \subseteq \Omega_t$ with $0 \leq t_1 \leq \lfloor \frac{t}{2} \rfloor$
- We know the evaluations $f(\omega)$ and $f(\omega^{-1})$
- We can compute $f(\omega) \pm f(\omega^{-1})$ for $\omega \in \Omega_{2t_1}$

This gives $t_1$ evaluations of $\psi_{\pm}(x) = f(x) \pm f\left(\frac{1}{x}\right)$ at $\omega \in \Omega_{2t_1}$

Idea: Find $\psi_{\pm}(x)$ using lattice of smaller dimensions and do linear algebra to recover $f(x)$. Finding each of $\psi_{\pm}(x)$ can be performed in parallel.
Attack on the PV Knapsack problem

- The mapping
  \[ x^i \rightarrow x^i + 1/x^i \text{ for } 0 \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor \]
  is well defined.

By linearity, \( \psi_+(x) = f(x) + f \left( \frac{1}{x} \right) \) can be generated by the basis (of order \( \lceil \frac{n}{2} \rceil \))

\[ \left\{ 2, \left( x + \frac{1}{x} \right), \left( x^2 + \frac{1}{x^2} \right), \ldots, \left( x^{\left\lfloor \frac{n}{2} \right\rfloor} + \frac{1}{x^{\left\lfloor \frac{n}{2} \right\rfloor+1}} \right) \right\} \]

Similarly, \( \psi_-(x) = f(x) - f \left( \frac{1}{x} \right) \) can be generated by the basis (of order \( \left\lfloor \frac{n}{2} \right\rfloor \))

\[ \left\{ \left( x - \frac{1}{x} \right), \left( x^2 - \frac{1}{x^2} \right), \ldots, \left( x^{\left\lfloor \frac{n}{2} \right\rfloor} - \frac{1}{x^{\left\lfloor \frac{n}{2} \right\rfloor+1}} \right) \right\} \]

- If \( f(x) \) has uniformly random coefficients in \( \{-1,0,1\} \), \( \psi_\pm(x) \) has coefficients in \( \{-2,-1,0,1,2\} \) and

\[ ||\psi_\pm|| \approx \sqrt{\frac{4\left\lfloor \frac{n}{2} \right\rfloor}{3}} \]

in the new basis representations.
Attack on the PV Knapsack problem
for $\omega \in \Omega_{2t_1}$

\[ W_+ = \begin{array}{cccc}
2 & \omega + \omega^{-1} & \cdots & \omega^{\left[\frac{n}{2}\right]} + \omega^{-\left(\left[\frac{n}{2}\right]+1\right)} \psi_+(\omega)
\end{array} \]

\[ \psi_+ \equiv 0 \mod q \]
New Primal Attack on the PV Knapsack problem

PV Knapsack problem reduced to finding the uSVP solution on the Kernel lattice

\[ L_{W_+}^1 = \{ x \in \mathbb{Z}_{\lceil n/2 \rceil + 1} : W_+ x = 0 \mod q \} \]

With \( \text{Dim} = \lceil \frac{n}{2} \rceil + 1 \) \( \text{Vol} = q^{t_1} \)

\[ \|(\psi_\pm, -1)\| \approx \sqrt{\frac{4\lceil n/2 \rceil}{3}} \] which is also unusually short in the lattice \( L_{W_\pm}^1 \).
Analysis of the attack

- uSVP cost depends on the root Hermite factor $\delta = \gamma^{1/dim}$, $\gamma = \frac{\lambda_2}{\lambda_1}$ is the uniqueness gap [GN’08].
- The attack gets faster as $t_1$ increases.

Probability distribution of the number of pairs $t_1$:

$$\pi_1(t_1) = \binom{n/2}{t_1} \binom{n/2}{t_2-t_1} 2^{t_2-t_1} \binom{2[n/2]}{t}$$


$$\pi_1(t_1)$$ for $n = 512, t = 256$
All the parameters from the literature contain a non-negligible fraction of weak keys, which are easily identified and extremely susceptible to our attack.

Example: We recovered the secret key of a parameter set from [LZA’18] for a fraction of

- \(2^{-15}\) of the public keys in about 117 hours (\(\approx 2^{50}\) bits operation)
- \(2^{-19}\) of the public keys in about 30 hours (\(\approx 2^{48}\) bits operation)
- \(2^{-23}\) of the public keys in about 10 hours (\(\approx 2^{46}\) bits operation)
- \(2^{-30}\) of the public keys in about 8 hours (\(\approx 2^{45}\) bits operation)

The direct primal attack provides 54-bits security using LWE estimator [APS’15].

It was initially claimed to have a 128-bit security against key recovery attack [LZA18], which was reduced to 87-bit security using the distinguishing attack from [BGP’22].

Conclusion

“40 years Advances in Cryptology: How will future judge Us?”

Crypto’20 Rump talk by Yvo Desmedt available at https://www.youtube.com/watch?v=MTafClFZOi8&list=PLeeS-3Ml-rppZMjRn2bNhb1FU-JOLMjRU&index=36&t=4650s

• Lattice-based assumptions are “relatively” NEW.

• CRYPTANALYSIS challenges our assumptions.