Baby Kyber - teoria

Małgorzata Zajęcka

3 stycznia 2024

Oznaczenia

Definicja

Dla liczby naturalnej m *m-tym wielomianem cyklotomicznym* nazywamy nierozkładalny (nad $\mathbb{Z}[X]$) wielomian o współczynnikach całkowitych, który dzieli $x^m - 1$, ale nie dzieli $x^k - 1$ dla $k < m$.

Innymi słowy m-ty wielomian cyklotomiczny to wielomian, którego pierwiastki to m-te pierwiastki z jedynki.

Niech $f \in \mathbb{Z}[X]$ będzie m-tym wielomianem cyklotomicznym postaci $f(X) = X^d + 1$, gdzie d jest potęgą dwójki ($d = 2^k$), niech q będzie liczbą pierwszą (różną od 2). Definiujemy pierścień R jako pierścień reszt $\mathbb{Z}[X]/f(X)$ oraz pierścień $R_q = R/qR$
Przez χ oznaczymy rozkład prawdopodobieństwa nad R.

M-LWE

Definicja (Problem Search M-LWE)

Mając dane $\boldsymbol{a}_1, ..., \boldsymbol{a}_m \in R_q^n$ (wyznaczone losowo z rozkładem jednostajnym) oraz $\boldsymbol{b} = (b_1, ..., b_m) \in R_q$ znaleźć $s \in R_q^n$ taki, że

$$b_i = \langle \boldsymbol{s}, \boldsymbol{a_i} \rangle + e_i \bmod qR, \;\; i = 1, ..., m$$

gdzie "błędy" e_i są wyznaczone z rozkładu χ.

M-LWE

Definicja (Problem Decision M-LWE)

Mając dane $\boldsymbol{a}_1, ..., \boldsymbol{a}_m \in \mathbb{R}_q^n$ oraz $\boldsymbol{b} = (b_1, ..., b_m) \in R_q$ rozstrzygnąć, czy

$$b_i = \langle \boldsymbol{s}, \boldsymbol{a_i} \rangle + e_i \bmod qR, \;\; i = 1, ..., m$$

dla pewnego $s \in R_q^n$, czy też $\boldsymbol{b}$ jest wyznaczony losowo z rozkładem normalnym i w żaden sposób nie zależny od $\boldsymbol{a}$.

CRYSTALS-Kyber

CRYSTALS-Kyber oparty jest na problemie M-LWE z następującymi parametrami:

1. $q = 3329 = 13 \cdot 2^8 + 1$
2. $n = 256$
3. $R = \mathbb{Z}[X]/(X^n + 1)$,
4. $R_q = R/qR = \mathbb{Z}_q[X]/(X^n + 1)$,

Generowanie kluczy

- Macierz $A \in R_q^{k \times k}$ dana jest losowo z rozkładu jednostajnego
- Współczynniki sekretu $\boldsymbol{s} \in R_q^k$ wybierane są z wycentrowanego rozkładu dwumianowego B_{η_1}
- Współczynniki wektora błędu $\boldsymbol{e} \in R_q^k$ są wybierane również z B_{η_1}
- Klucz publiczny: $\boldsymbol{t} = A\boldsymbol{s} + \boldsymbol{e}$
- Klucz prywatny: $\boldsymbol{s}$

Szyfrowanie

- Chcemy zaszyfrować wielomian m o współczynnikach z $\{0, 1\}$
- Próbkujemy $\boldsymbol{r} \in R_q^k$ z B_{η_1}
- Próbkujemy $\boldsymbol{e}_1 \in R_q^k$ z B_{η_2}
- Próbkujemy $e_2 \in R$ z B_{η_2}
- Obliczamy $\boldsymbol{u} = A^T \boldsymbol{r} + \boldsymbol{e}_1$
- Obliczamy $v = \boldsymbol{t}^T \boldsymbol{r} + e_2 + \lceil (q/2)m \rfloor$
- $c = (\boldsymbol{u}, v)$

Deszyfrowanie

- Obliczamy $m_n = v - \boldsymbol{s}^T \boldsymbol{u}$

Otrzymany wynik jest "zaszumiony", ponieważ obliczenia w rzeczywistości nie dają oryginalnej wiadomości m. Jednak dzięki zastosowanemu przeskalowaniu współczynniki m_n są albo bliskie $\lfloor q/2 \rceil$, (a zatem pierwotny współczynnik m wynosił 1) albo są bliskie zera, co oznacza, że oryginalny współczynnik w m wynosił 0.

Kyber bez zębów - Baby Kyber

Przyjmijmy uproszczony model z parametrami:

- $k = 2$,
- $q = 17$ oraz $n = 4$, zatem $R = \mathbb{Z}[X]/(X^4 + 1)$, $R_q = \mathbb{Z}_{17}[X]/(X^4 + 1)$.

Generowanie klucza:

- macierz $A \in R_q^{k \times k}$ generujemy losowo, tzn. wybieramy cztery wielomiany o losowych współczynnikach z $\mathbb{Z}_{17}$, przykładowo

$$A = \begin{bmatrix} 6x^3 + 16x^2 + 16x + 11 & 9x^3 + 4x^2 + 6x + 3 \\ 5x^3 + 3x^2 + 10x + 1 & 6x^3 + x^2 + 9x + 15 \end{bmatrix}$$

- wybieramy dwa (bo $k = 2$) losowe wielomiany z R_q o współczynnikach ze zbioru $\{-1, 0, 1\}$, np.

$$\mathrm{s} = (-x^3 - x^2 + x, -x^3 - x)$$

- wybieramy losowo wektor błędu e składający się z dwóch wielomianów z R_q o współczynnikach ze zbioru $\{-1, 0, 1\}$ (jak w przypadku sekretu s), np. $\mathrm{e} = (x^2, x^2 - x)$
- obliczamy klucz publiczny $\mathrm{t} = A\mathrm{s} + \mathrm{e}$, gdzie wszystkie działania są wykonywane w arytmetyce modulo 17 (na współczynnikach) i modulo $x^4 + 1$ (na wielomianach).

$$\mathrm{t} = (16x^3 + 15x^2 + 7, 10x^3 + 12x^2 + 11x + 6)$$

Szyfrowanie

Powiedzmy, że chcemy zaszyfrować ciąg binarny 1011. Tworzymy odpowiadający mu wielomian

$$m = 1 \cdot x^3 + 0 \cdot x^2 + 1 \cdot x + 1 = x^3 + x + 1$$

- wybieramy losowo wektor błędu r z R_q o współczynnikach ze zbioru $\{-1, 0, 1\}$, np. $\mathrm{r} = (-x^3 + x^2, x^3 + x^2 - 1)$
- podobnie wybieramy losowo wektor błędu e_1, np. $\mathrm{e}_1 = (x^2 + x, x^2)$
- potrzebujemy jeszcze losowy wielomian e_2 stopnia co najwyżej 3 - dla uproszczenia również jego współczynniki wybieramy losowo z $\{-1, 0, 1\}$, przykładowo $e_2 = -x^3 - x^2$

- obliczamy $\mathrm{u} = A^T \mathrm{r} + \mathrm{e}_1$, gdzie znowu wszystkie działania są mod 17 i mod $x^4 + 1$.

$$\mathrm{u} = (11x^3 + 11x^2 + 10x + 3, 4x^3 + 4x^2 + 13x + 11)$$

- obliczamy $v = \mathrm{t}^T \mathrm{r} + e_2 + \lfloor \frac{q}{2} \rceil m$

$$v = 8x^3 + 6x^2 + 9x + 16$$

- otrzymujemy szyfrogram $\mathrm{c} = (\mathrm{u}, v)$

$$\mathrm{c} = ((11x^3+11x^2+10x+3, 4x^3+4x^2+13x+11), 8x^3+6x^2+9x+16)$$

Deszyfrowanie

- obliczamy $m_n = v - \mathrm{s}^T \mathrm{u}$

$$m_n = 8x^3 + 14x^2 + 8x + 6$$

Następnie zaokrąglamy otrzymane współczynniki do 0, jeżeli są bliżej 0 lub 17 niż 9 albo do 1, jeżeli są bliżej 9 niż 0 lub 17. Otrzymujemy $m_n = x^3 + 0 \cdot x^2 + x + 1$, a zatem zaszyfrowany wielomian to $m = x^3 + x + 1$ i odczytujemy odpowiadający mu ciąg binarny 1011.