Post-Quantum Cryptography Definitions
Winter 2024
University of Waterloo
Instructor: Sam Jaques

January 9, 2024

1 General Cryptography

Negligible: We say a function $f(\lambda)$ is *negligible* if $f(\lambda) \in O(\frac{1}{p(\lambda)})$ for all polynomials $p(x)$.

1.1 Public Key Encryption

A public key encryption scheme is a set of 3 algorithms, $\mathsf{KeyGen}() \to (\mathbf{PK}, \mathbf{SK})$, $\mathsf{Enc}(\mathbf{PK}, m) \to c$, and $\mathsf{Dec}(\mathbf{SK}, c) \to m$. Intuitively, $\mathbf{PK}$ is the public key, $\mathbf{SK}$ is the secret key, m is a plaintext message, and c is a ciphertext.

Implicitly, all algorithms are parameterized by a security parameter λ.

Such a scheme should be correct: For all outputs $(\mathbf{PK}, \mathbf{SK})$, the probability that

$$\mathsf{Dec}(\mathbf{SK}, \mathsf{Enc}(\mathbf{PK}, m)) \neq m \tag{1}$$

is negligible in λ.

IND-CPA Security: For the IND-CPA game (indistinguishability against chosen plaintext attack), let $\mathcal{A}$ be an algorithm whose runtime is polynomial in λ. In the IND-CPA game:

1. A challenger generates a keypair: $\mathsf{KeyGen}() \to (\mathbf{PK}, \mathbf{SK})$.

2. $\mathcal{A}$ receives $\mathbf{PK}$, and can make a polynomial number of queries to an encryption oracle, which outputs $\mathsf{Enc}(\mathbf{PK}, \cdot)$.

3. $\mathcal{A}$ outputs two messages m_0 and m_1.

4. The challenger selects a uniformly random bit $b \in \{0, 1\}$, and returns $c_b = \mathsf{Enc}(\mathbf{PK}, m_b)$ to $\mathcal{A}$.

5. $\mathcal{A}$ can make another polynomial number of queries to an encryption oracle, which outputs $\mathsf{Enc}(\mathbf{PK}, \cdot)$.

6. $\mathcal{A}$ outputs a bit b'.

We say that $\mathcal{A}$ "wins" the IND-CPA game if $b' = b$.

An encryption scheme is IND-CPA secure if, for any polynomial time algorithm $\mathcal{A}$, the probability of winning is at most $\frac{1}{2} + \epsilon(\lambda)$ where $\epsilon(\lambda)$ is negligible.

IND-CCA Security: For the IND-CCA game (indistinguishability against chosen ciphertext attack), let $\mathcal{A}$ be an algorithm whose runtime is polynomial in λ. In the IND-CCA game:

1. A challenger generates a keypair: $\mathsf{KeyGen}() \to (\mathbf{PK}, \mathbf{SK})$.

2. $\mathcal{A}$ receives $\mathbf{PK}$, and can make a polynomial number of queries to:

   - an encryption oracle, which outputs $\mathsf{Enc}(\mathbf{PK}, m)$ on input m
   - a decryption oracle, which outputs $\mathsf{Dec}(\mathbf{SK}, c)$ on input c

3. $\mathcal{A}$ outputs two messages m_0 and m_1.

4. The challenger selects a uniformly random bit $b \in \{0, 1\}$, and returns $c_b = \mathsf{Enc}(\mathbf{PK}, m_b)$ to $\mathcal{A}$.

5. $\mathcal{A}$ can make another polynomial number of queries:

   - an encryption oracle, which outputs $\mathsf{Enc}(\mathbf{PK}, m)$, on input m.
   - a restricted decryption oracle, which outputs $\mathsf{Dec}(\mathbf{SK}, c)$ on input c if $c \neq c_b$, and outputs a fixed symbol (say, $\perp$) if $c = c_b$.

6. $\mathcal{A}$ outputs a bit b'.

We say that $\mathcal{A}$ "wins" the IND-CCA game if $b' = b$.

An encryption scheme is IND-CCA secure if, for any polynomial time algorithm $\mathcal{A}$, the probability of winning is at most $\frac{1}{2} + \epsilon(\lambda)$ where $\epsilon(\lambda)$ is negligible.

1.2 Digital Signatures

A digital signature scheme is a tuple of algorithms:

- $\mathbf{KeyGen}() \to (\mathbf{PK}, \mathbf{SK})$
- $\mathbf{Sign}(\mathbf{SK}, m) \to s$
- $\mathbf{Ver}(\mathbf{PK}, s, m) \to b \in \{0, 1\}$

A digital signature scheme is correct/complete if, for any keypair $(\mathbf{PK}, \mathbf{SK})$ generated by KeyGen, the probability is negligible in λ that

$$\mathbf{Ver}(\mathbf{PK}, \mathbf{Sign}(\mathbf{SK}, m), m) \neq 1 \tag{2}$$

Security: Security definitions are complicated; see here for a taxonomy: `https://crypto.stackexchange.com/questions/44188/what-do-the-signature-security-a`

Here I will give the definition of strong existential forgery under chosen-message attack. The game is as follows

1. A challenger generates $(\mathbf{PK}, \mathbf{SK}) \leftarrow \mathsf{KeyGen}()$ and initializes a set $\mathcal{M}$.
2. An adversary $\mathcal{A}$ runs for polynomial time and is allowed polynomial queries to a signing oracle, which does the following:
   - Computes $s \leftarrow \mathsf{Sign}(\mathbf{SK}, m)$
   - Adds (m, σ) to $\mathcal{M}$.
   - Returns σ to $\mathcal{A}$
3. The adversary $\mathcal{A}$ outputs (m^*, s^*).

We say that $\mathcal{A}$ wins the game if:

- $(m^*, s^*) \notin \mathcal{M}$, and
- $\mathsf{Ver}(\mathbf{PK}, s^*, m^*) = 1$.

Notice that (m^*, s) could be in $\mathcal{M}$ and the adversary could still win, i.e., they could win by producing a new signature of a message that had already been signed.

A digital signature scheme is sEF-CMA-secure if, for any polynomial time $\mathcal{A}$, the probability of $\mathcal{A}$ winning this game is negligible.

2 Lattice Cryptography

2.1 General

We can define a norm on $\mathbb{Z}/q\mathbb{Z}$ by setting $|x|_q = |\overline{x}|$ where $\overline{x} \in [-q/2, q/2)$ and $\overline{x} \equiv x \mod q$. This can extend to a norm on $\mathbb{Z}/q\mathbb{Z}^n$ by setting $\|x\| = \sqrt{|x_1|_q + \cdots + |x_n|_q}$.

2.2 Learning With Errors

Learning With Errors (LWE): An LWE$(n, m, q, \chi_s, \chi_e)$ instance is formed by sampling a uniformly random $m \times n$ matrix A with entries in $\mathbb{Z}/q\mathbb{Z}$, a vector $s \in (\mathbb{Z}/q\mathbb{Z})^n$ from the distribution χ_s, and a vector $e \in (\mathbb{Z}/q\mathbb{Z})^m$ from the distribution χ_e, and outputting $(A, b := As + e \mod q)$.

The number m is sometimes referred to as the number of "samples".

The LWE$(n, m, q, \chi_s, \chi_e)$ *search* problem is, given (A, b) as sampled above, to recover s.

The LWE$(n, m, q, \chi_s, \chi_e)$ *decision* problem is: a bit $b' \in \{0, 1\}$ is drawn uniformly at random, and if $b' = 0$, then one is given an LWE sample (A, b) as above, and if $b' = 1$, then one is given (A, b) where A is a uniformly random $n \times m$ matrix and b is a uniformly random m-dimensional vector (both with entries in $\mathbb{Z}/q\mathbb{Z}$. The problem is to determine whether $b' = 0$ or $b' = 1$.

Normal form LWE sets $m = n$ and $\chi_s = \chi_e$.

A non-standard definition is that of "unique" LWE parameters, which is a set of parameters $(n, m, q, \chi_s, \chi_e)$ such that if s, s' are sampled from χ_s and e, e' are sampled from χ_e such that $As + e = As' + e'$, then with high probability $s = s'$ and $e = e'$. Generally the literature assumes this to be the case, but there are pathological parameter choices (e.g., χ_e uniformly random) where this does not hold.

Textbook LWE Encryption: This is a public key encryption scheme and thus consists of three algorithms. It is parameterized by $(n, m, q, \chi_s, \chi_e, \chi'_s, \chi'_e, \chi''_e)$ (though often $n = m$, $\chi'_s = \chi_s$, and $\chi'_e = \chi_e$).

- $\mathsf{KeyGen}() \to (\mathbf{PK}, \mathbf{SK})$: Sample a uniformly random matrix A with entries in $\mathbb{Z}/q\mathbb{Z}$, a vector s from the distribution χ_s, and a vector e from the distribution χ_e. Compute $b = As + e \mod q$, and set $\mathbf{SK} \leftarrow s$ and $\mathbf{PK} \leftarrow (A, b)$.
- $\mathsf{Enc}(\mathbf{PK}, m) \to c$: Sample a vector $s' \leftarrow \chi'_s$, $e' \leftarrow \chi'_e$, and $e'' \leftarrow \chi''_e$. Set $c_1 = s'^T A + e'^T \mod q$ and $c_2 = s'^T b + e'' + m \left\lfloor \frac{q}{2} \right\rceil \mod q$. Output $c = (c_1, c_2)$.
- $\mathsf{Dec}(\mathbf{SK}, c) \to m$. Compute $m' = c_2 s - c_1 \mod q$, where this is taken between $[-q/2, q/2)$. Round m' to $\left\lfloor \frac{q}{2} \right\rceil$, i.e, if $-\frac{q}{4} \leq m' \leq \frac{q}{4}$, set $m = 0$, otherwise set $m = 1$. Output m.

Never deploy this scheme, it is not IND-CCA secure.

Basic LWE Kyber: This not Kyber, but a toy version useful to explore parameters.

Here we take the textbook LWE encryption above and set $n = m = 512$, $q = 3329$, and set $\chi_s = \chi'_s$ have each component be independently and identically distributed as a centered binomial distributions with parameters $(n = 6, p = \frac{1}{2})$, and χ_e, χ'_e and χ''_e to have each component independently and identically distributed as a centered binomial distribution with parameters $(n = 4, p = \frac{1}{2})$.

2.3 Distributions

Discrete Gaussian : A discrete Gaussian distribution on $\mathbb{Z}/q\mathbb{Z}$ with mean $\mu \in \mathbb{Z}/q\mathbb{Z}$ and standard deviation σ is defined by setting $\rho(x) = e^{-\frac{(x'-\mu)^2}{2\sigma^2}}$, where $x' \equiv x$ and $x' \in [\mu - \frac{q}{2}, \mu + \frac{q}{2})$. Then the probability of x in the discrete Gaussian distribution is proportional to $\rho(x)$, i.e.,

$$\Pr(x) = \frac{\rho(x)}{\sum_{y=0}^{q-1} \rho(y)} \tag{3}$$

Centered Binomial Distribution : This has parameters $n \in \mathbb{N}$ and $p \in [0, 1]$. Let $\mu = np$ and assume $\mu \in \mathbb{N}$. Then this is a distribution on $[-\mu, n - \mu]$ where

$$\Pr(k) = \Pr_{\mathrm{Bin}(n,p)}(k + \mu) = \binom{n}{k + \mu} p^{k+\mu}(1 - p)^{n-k-\mu} \tag{4}$$

where $\Pr_{\mathrm{Bin}(n,p)}(k)$ is the probability of x in the binomial distribution with parameters n and p.

If $[-\mu, n - \mu] \subseteq [-\frac{q}{2}, \frac{q}{2})$, then this distribution can be defined in $\mathbb{Z}/q\mathbb{Z}$ by using the equivalence class in $[-\frac{q}{2}, \frac{q}{2})$ and setting the probability to be 0 for all values outside of $[-\mu, n - \mu]$.

2.4 Lattices

A *lattice* is a discrete additive subgroup of $\mathbb{R}^n$. Equivalently, a lattice can be defined by a set $\mathcal{B}$ of lienarly independent vectors in $\mathbb{R}^n$ as

$$\mathcal{L}(\mathcal{B}) = \left\{ \sum_{i=1}^{m} a_i b_i \middle| a_i \in \mathbb{Z}, b_i \in \mathcal{B} \right\}. \tag{5}$$

The kth successive minima of a lattice, denoted $\lambda_k(\mathcal{L})$, is defined as:

$$\min \left\{ \max_{i=1}^{k} \{\|v_i\|\} \middle| \{v_1, \ldots, v_k\} \subseteq \mathcal{L} \text{ and is linearly independent in } \mathbb{R}^n \right\} \tag{6}$$

The value $\lambda_1(\mathcal{L})$ is of special importance: this is the length of the shortest non-zero vector in the lattice.

The *dual* of a lattice $\mathcal{L}$ is defined as

$$\mathcal{L}^\vee := \{v \in \text{real span of } \mathcal{L} | \langle v, w\rangle \subseteq \mathbb{Z}, \forall w \in \mathcal{L}\} \tag{7}$$

Lattice problems: The γ-shortest vector problem (γ-SVP): given a basis $\mathcal{B}$ for a lattice $\mathcal{L}(\mathcal{B})$, find a vector $v \in \mathcal{L}(\mathcal{B})$ such that $\|v\| \leq \gamma\lambda_1(\mathcal{L}(\mathcal{B}))$.

The γ, k-shortest independent vector problem (γ, k-SIVP): given a basis $\mathcal{B}$ for a lattice $\mathcal{L}(\mathcal{B})$, find k vectors $v_1, \ldots, v_k \in \mathcal{L}(\mathcal{B})$ which are linearly independent in $\mathbb{R}^n$ such that $\|v_i\| \leq \gamma\lambda_k(\mathcal{L}(\mathcal{B}))$ for all $1 \leq i \leq k$.

Given a vector $t \in \mathbb{R}^n$, we can define

$$\|t - \mathcal{L}\| = \min\{\|v - t\| | v \in \mathcal{L}\} \tag{8}$$

The γ-closest vector problem (γ-CVP): given a basis $\mathcal{B}$ for a lattice $\mathcal{L}(\mathcal{B})$ and a vector $t \in \mathbb{R}^n$, find a vector $v \in \mathcal{L}(\mathcal{B})$ such that $\|t - v\| \leq \gamma \|t - \mathcal{L}\|$.

The bounded distance decoding problem (BDD): Given β, a lattice $\mathcal{L}$, and a vector $t \in \mathbb{R}^n$, with the promise that $\|t - \mathcal{L}\| \leq \beta$, find v such that $\|t - v\| = \|t - \mathcal{L}\|$.

The β-short integer solutions problem (β-SIS): Given a matrix B, find an integer vector v such that $Bv \equiv 0 \mod q$ such that $\|v\| \leq \beta$.