Proofs of Security

TUT n°4 — Asymmetric revisions

M2-SIF, M2-CRYPTO

Léo Ackermann

Answer all ▸ questions before looking at ⋆ questions.

Hash-and-Sign signature from lattices

The *Hash-and-Sign* paradigm is a generic method for deriving rather efficient signatures from trapdoor one-way functions[*]. The signature scheme is derived as follows:

- KeyGen select a trapdoor function f together with its trapdoor τ_f. The public key of the scheme is f, the secret key is τ_f.
- Sign(sk, m) first hashes the message m to some point $y = \mathcal{H}(m)$ within $f's$ range. Then, it computes $\sigma \in f^{-1}(m)$ using the trapdoor τ_f.
- Verif(pk, m, σ) simply checks that $\mathcal{H}(m) = f(\sigma)$.

This exercise focuses on a classical instanciation of the Hash-and-Sign paradigm in the lattice world. We briefly introduce lattices and some important results that can be used as blackboxes in this exercise[†].

A (full-rank) lattice Λ is a discrete subgroup of $\mathbb{R}^n$, and as such it always admits generators $\mathbf{b}_1, \cdots, \mathbf{b}_n \in \mathbb{R}^n$ (often compiled within a matrix $\mathbf{B}$) whose integer linear combinations yield the lattice $\Lambda = \Lambda(\mathbf{B})$. The security of the scheme we focus on relies on hypotheses related to the (Inhomogeneous) Short Integer problems, defined as follows.

Definition 1 ((Inhomogeneous) Short Integer Solution) *Given an integer q, a matrix $\boldsymbol{A} \in \mathbb{Z}_q^{m \times n}$, a real β, and a syndrome $\boldsymbol{u} \in \mathbb{Z}_q^n$.*

- ***SIS.*** *The problem is to find a nonzero vector $\boldsymbol{e} \in \mathbb{Z}_q^m$ such that $\mathbf{e}^\top \boldsymbol{A} = \mathbf{0}$ and $\|\mathbf{e}\| \leqslant \beta$.*
- ***ISIS.*** *The problem is to find a nonzero vector $\boldsymbol{e} \in \mathbb{Z}_q^m$ such that $\mathbf{e}^\top \boldsymbol{A} = \boldsymbol{u}$ and $\|\mathbf{e}\| \leqslant \beta$.*

The following results regarding lattice-based cryptography will be of importance for the rest of the tutorial.

- **Gaussian sampling.** There is a PPT algorithm that, given a basis $\boldsymbol{B}$ of an n-dimensional lattice $\Lambda = \Lambda(\boldsymbol{B})$, a parameter $s \geqslant \|\widetilde{\boldsymbol{B}}^{\ddagger}\| \cdot \omega(\sqrt{\log n})$, and a center $\boldsymbol{c} \in \mathbb{R}^n$, outputs a sample from a distribution that is statistically close to $G_{\Lambda,s,\boldsymbol{c}}$[§].
- **Lattice generation with trapdoor.** For any prime $q = \mathrm{poly}(n)$, and any $m \geqslant 5n \log q$, there is a PPT algorithm that outputs a matrix $\boldsymbol{A} \in \mathbb{Z}_q^{m \times n}$ together with a basis $\boldsymbol{T_A}$ of $\Lambda^\perp(\boldsymbol{A})$ such that: the distribution of $\boldsymbol{A}$ is statistically close to the uniform over $\mathbb{Z}_q^{m \times n}$ and $\|\widetilde{\boldsymbol{T_A}}\| \leqslant L \coloneqq m^2\sqrt{m}$.
- **Conditional distribution of syndrome.** Let $\boldsymbol{u} \in \mathbb{Z}_q^n$ and $\boldsymbol{t} \in \mathbb{Z}^m$ be an arbitrary solution to $\boldsymbol{t}^\top \boldsymbol{A} = \boldsymbol{u} \mod q$. Then the conditional distribution of $\boldsymbol{e} \longleftarrow G_{\mathbb{Z}^m,s}$ given $\boldsymbol{e}^\top \boldsymbol{A} = \boldsymbol{u} \mod q$ is exactly $\boldsymbol{t} + G_{\Lambda^\perp,s,-\boldsymbol{t}}$.
- **Distribution of syndrome.** Let n, q be integers with q prime, and let $m \geqslant 2n \log q$. Then for all but a $2q^{-n}$ fraction of all $\boldsymbol{A} \in \mathbb{Z}_q^{m \times n}$, and for any $s \geqslant \omega(\sqrt{\log n})$, the distribution of the syndrome $\boldsymbol{u} = \boldsymbol{e}^\top \boldsymbol{A} \mod q$ is statistically close to uniform over $\mathbb{Z}_q^n$, where $\boldsymbol{e} \longleftarrow G_{\mathbb{Z}^m,s}$.
- **Recheable syndromes.** Let $m \geqslant 2n \log q$. Then, for all but an atmost q^{-n} fraction of $\boldsymbol{A} \in \mathbb{Z}_q^{m \times n}$, for any syndrome $\boldsymbol{u} \in \mathbb{Z}_q^n$, there is a $\boldsymbol{e} \in \{0,1\}^m$ such that $\boldsymbol{e}^\top \boldsymbol{A} = \boldsymbol{u} \mod q$.

◆ A primer on game-base proofs

▸ **Question 1.** *Let $\mathcal{G}_0$ and $\mathcal{G}_1$ be two security games, and $\mathcal{A}_0, \mathcal{A}_1$ be adversaries against them. Let $\mathcal{D}$ be an adversary trying to determine whether she is interacting with $\mathcal{G}_0$ or $\mathcal{G}_1$. Relate $\mathrm{Adv}(\mathcal{D})$, $\Pr(\mathcal{G}_0(\mathcal{A}_0) \to \top)$ and $\Pr(\mathcal{G}_1(\mathcal{A}_1) \to \top)$.*

▸ **Question 2.** *Let $\mathcal{G}_0, \cdots, \mathcal{G}_n$ be a sequence of game, and $(\mathcal{D}_i)_{0 \leqslant i < n}$ be distinguishers between games $\mathcal{G}_i$ and $\mathcal{G}_{i+1}$. Assuming that $Pr(\mathcal{G}_n \to \top) = p$, what can be said about $\Pr(\mathcal{G}_0 \to \top)$?*

[*] Informally, a family of trapdoor function is a family of efficiently computable functions that are hard to invert, except if one is given the associate trapdoor.

[†] The point here is really to understand how those results branch together and what operations they allow to perform.

[‡] This denote the Gram-Schmidt orthogonalization of the matrix $\mathbf{B}$.

[§] The discrete Gaussian distribution over a lattice is really the (scaled) continuous Gaussian distribution restricted to the lattice points.

◆ Preimage sampleable functions from (I)SIS hardness

In this part, we design preimage-sampleable function that are one-way and collision-resistant under the SIS and ISIS hypothesis.

> **Definition 2 (Collision-resistant preimage sampleable functions)** *A collection of collision-resistant preimage sampleable functions is a tuple of PPT algorithms* TrapGen*,* SDom *and* SPre*) such that:*
>
> - TrapGen *outputs a couple* (a, τ_a) *where* a *describe an efficiently computable function* $f_a : D \to R$*, and* τ_a *some trapdoor information.*
> - SDom(a) *samples* $x \in D$ *such that the distribution of* $f_a(x)$ *is (statistically close from) uniform.*
> - SPre$(\tau_a, y \in R)$ *samples from (a distribution statistically close from) the conditional distribution of* $x \leftarrow$ SDom(a)*, given* $f_a(x) = y$*. Additionally, it requieres that the later distribution has min-entropy at least* $\omega(\log n)$*.*
>
> *Moreover, the advantage of any adversary for producing a preimage of* $y \hookleftarrow \mathcal{U}(R)$ *by* f_a *is negligible if it was not handed* τ_a *(this is one-wayness), and the advantage of any adversary for producing* $x \neq x'$ *such that* $f_a(x) = f_a(x')$ *is negligible if it was not handed* τ_a *(this is collision-resistance).*

Let $p = \mathrm{poly}(n)$ prime, $m \geqslant 5n \log q$ and $s \geqslant L \cdot \omega(\sqrt{\log n})$. Under the hardness of $\mathsf{ISIS}_{q,m,s\sqrt{m}}$ and $\mathsf{SIS}_{q,m,2s\sqrt{m}}$, collision-resistant permutations exist over lattices. The trapdoor generation is performed as follows.

> TrapGen samples $(\boldsymbol{A}, \boldsymbol{T_A})$ such that $\boldsymbol{A}$ is statistically close from the uniform distribution, and $\boldsymbol{T_A}$ is a trapdoor for $\boldsymbol{A}$ – namely, a good basis of $\Lambda^\perp(\boldsymbol{A})$ – and the associated function is $f_{\boldsymbol{A}} : \boldsymbol{e} \mapsto \boldsymbol{e}^\top \boldsymbol{A} \mod q$, with domain $D = \{\boldsymbol{e} \in \mathbb{Z}^m \mid \|e\| \leqslant s\sqrt{m}\}$ and range $R = \mathbb{Z}_q^n$.

▸ **Question 3.** *Propose algorithms for* SDom *and* SPre *that fit this* TrapGen *algorithm.*

▸ **Question 4.** *Show that one-wayness and collision-resistance hold under the* (I)SIS *hypotheses for aforementioned parameters.*

◆ GPV signatures

We now present the scheme known as GPV's signature (in its stateful version), an instanciation of the Hash-and-Sign paradigm over lattices. Let $\mathcal{H}$ be a hash function modeled as a random oracle. We have:

> - KeyGen(1^λ) samples $(\boldsymbol{A}, \boldsymbol{T_A})$ by calling TrapGen(1^λ).
> - Sign$(\boldsymbol{T_A}, m)$ returns σ_m if the couple (m, σ_m) is in local storage[¶]. Otherwise, it let σ_m be a preimage of $\mathcal{H}(m)$ (found using SPre), stores the couple (m, σ_m) and outputs σ_m.
> - Verif$(\boldsymbol{A}, m, \sigma)$ computes $y = \mathcal{H}(m)$ and checks whether both $\sigma^\top \boldsymbol{A} = y$ and $\sigma \in D$.

▸ **Question 5.** *Write the* stateful-EUF-CMA *security game in a game-based style. You may defined auxiliary algorithm* $\mathcal{O}_{\mathsf{sign}}$ *and* $\mathcal{O}_{\mathsf{RO}}$*. This is the game* $\mathcal{G}_{\mathsf{real}}$*.*

▸ **Question 6.** *Write a game* $\mathcal{G}_{\mathsf{sim}}$ *that "looks like"* $\mathcal{G}_{\mathsf{real}}$ *in which the trapdoor* $\boldsymbol{T_A}$ *is no longer used. You may defined auxiliary algorithm* $\mathcal{S}_{\mathsf{sign}}$ *and* $\mathcal{S}_{\mathsf{RO}}$*. What is* $\Pr(\mathcal{G}_{\mathsf{sim}} \to \top)$*?*

▸ **Question 7.** *Write a game-based proof showing that the distance between the games is negligible. You may introduce an intermediate game ensuring* $\mathcal{H}(m)$ *has always be queried before signing the message* m*, and assume that* $\mathcal{A}$ *queried* $\mathcal{H}(m^*)$ *for the message* m^* *he produces a forgery on.*

▸ **Question 8.** *Deduce that stateful-GPV is* SEU-CMA *secure under the* (I)SIS *hypothesis in the* ROM*.*

▸ **Question 9.** *Propose non-stateful version of this signature scheme, and argue for its security.*

[¶] This is where the *stateful* adjective reflects.