Rejection Sampling

Giovanni Tognolini

University of Trento

2023

In a Nutshell

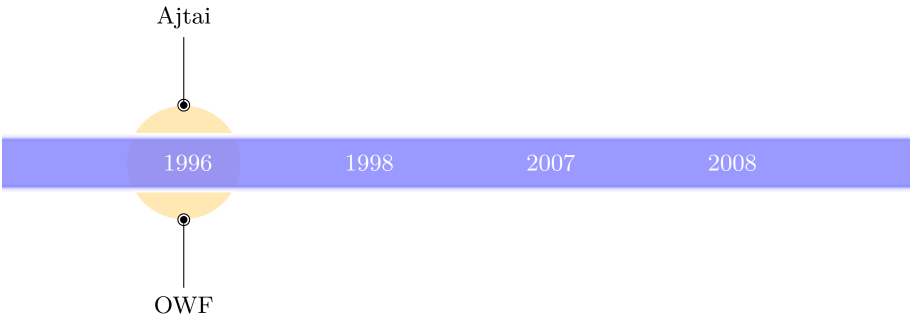
What we will see?

1. Small recap on lattices and cryptography:
2. The idea of Lyubashevsky.
   - Original proposal;
   - First variation;
   - Second variation;

Small Recap

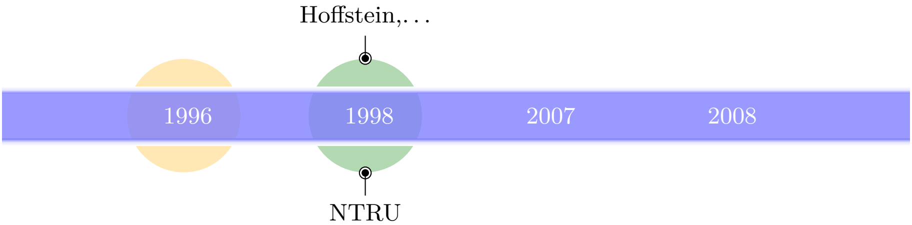
1996	1998	2007	2008

Small Recap



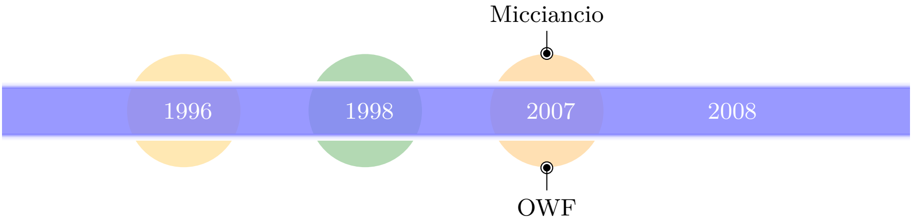
- Inefficient;
- New ideas are needed to make lattice encryption a valid alternative to the t.d.n based one.

Small Recap



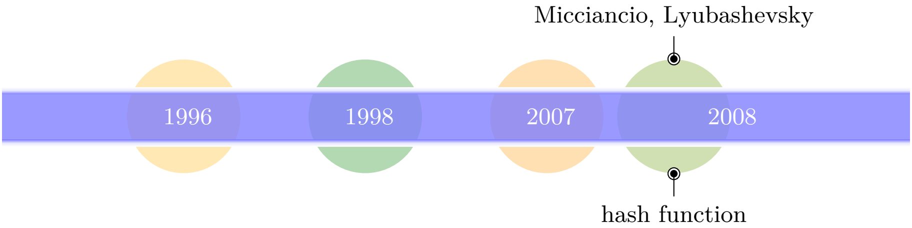
- Efficient;
- No security proof.

Small Recap



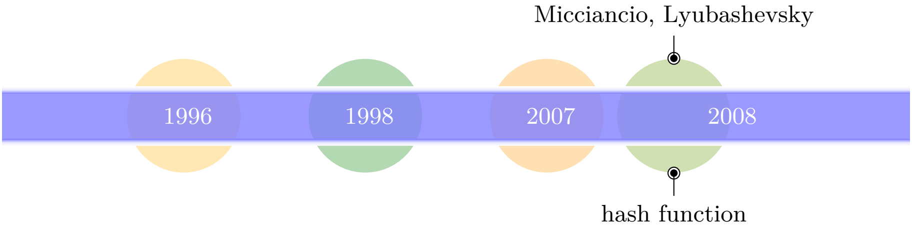
- Security based on solving some problems on structured lattices.

Small Recap



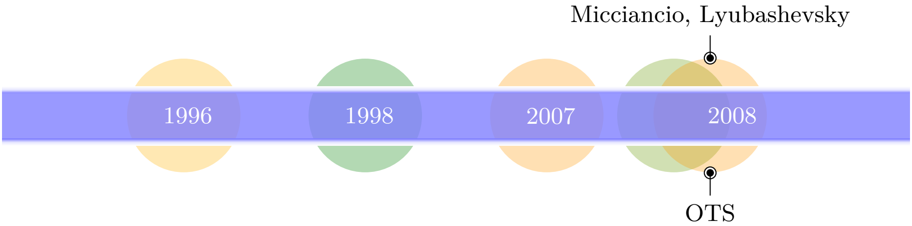
- Hash function based on structured lattices;
- Efficient

Small Recap



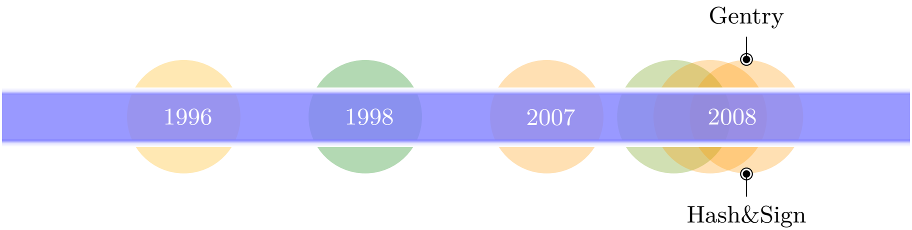
Hash functions and signature/ID schemes are very closely related.
Are there *efficient* lattice-based signature schemes?

Small Recap

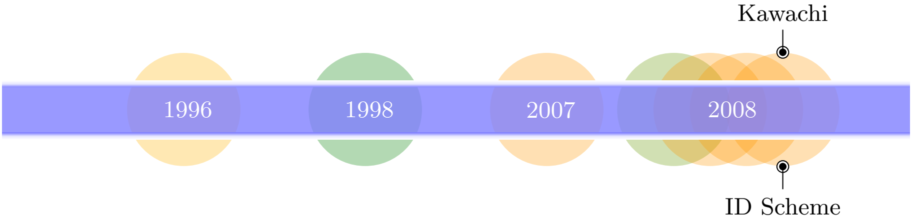


- One time signature;
- With standard techniques the signature becomes full-fledged.

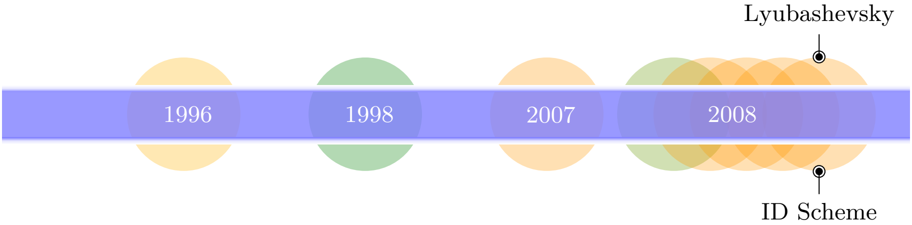
Small Recap



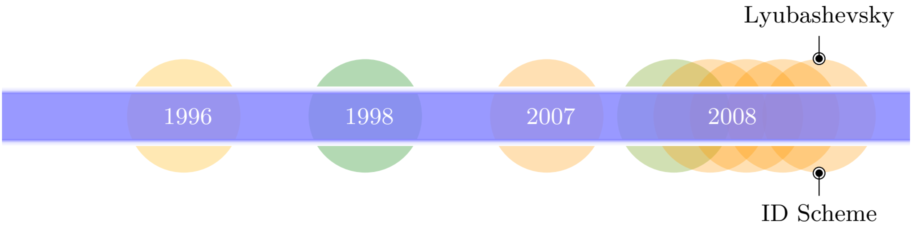
Small Recap



Small Recap



Small Recap



Why the scheme proposed by Lyubashevsky is so important?

The Idea of Lyubashevsky

ID Schemes

PROVER		*VERIFIER*
(sk, pk)		(pk)
	$\xrightarrow{commit}$	
	$\xleftarrow{challenge}$	
	$\xrightarrow{response}$	$f(pk, com, ch, res)$

- We would like to instantiate this framework with lattices;
- We need a problem to base the security on.

Ring-SIS

$$R := \mathbb{Z}_q[x]/(x^n + 1)$$
$$a_1, a_2, a_3 \in R \text{ random}$$
$$\text{Find "short" } r_1, r_2, r_3 \in R \text{ s.t. } a_1 r_1 + a_2 r_2 + a_3 r_3 = 0.$$

Observation

Sometimes the problem is rewritten with just the request to find short r_1, r_2, r_3 s.t. $a_1 r_1 + a_2 r_2 + r_3 = 0$.

Lyubashevsky ID Scheme

$$sk : s_1, s_2 \in R \text{ of "small norm"}$$
$$pk : a \in R, t = as_1 + s_2$$

PROVER (sk, pk)		*VERIFIER* (pk)
Pick $y_1, y_2 \sim D$		
	$\xrightarrow{w = ay_1 + y_2}$	
		Pick sparse $c \in R, \|\|c\|\|_1$ "small"
	$\xleftarrow{c}$	
$z_1 = y_1 + s_1 c$ $z_2 = y_2 + s_2 c$		
	$\xrightarrow{z_1, z_2}$	Check that

- $\|\|z_1\|\|, \|\|z_2\|\|$ are "small";
- $az_1 + z_2 - tc = w$.

A Couple of Observations

The scheme, as described above, is very vague, indeed:

- We did not specify the distribution for D;
- We didn't specify how small c should be;
- We didn't specified how small z_1, z_2 should be;
- We didn't said why it should be difficult for an opponent to break this scheme.

Security of the Scheme

Supp. there is an opponent capable of breaking the scheme just by looking at the transcripts. Let's see how to exploit this adversary to break SIS.

1. Supp. to have a_1, a_2 as input. We are asked to find r_1, r_2, r_3 "small" s.t. $a_1 r_1 + a_2 r_2 + r_3 = 0$.
2. We instantiate the Lyubashevsky ID scheme with public key

$$a := a_1, t := a_2$$

3. Can we create valid transcripts (or rather, indistinguishable from real transcripts) and show them to the attacker?

4. Supp. that the adversary, after having seen some of the transcript, tells us that he is able to break the signature. What happen?

So...

(under the assumption that the scheme is correct and does not leak informations)
If the parameters are chosen correctly then the security scheme can be reduced to ring-SIS.

Considerations on z_1, z_2

$$\begin{cases} z_1 = y_1 + s_1\, c \\ z_2 = y_2 + s_2\, c \end{cases}$$

Ideally we would like z_1, z_2 not to depend on the private key.

Observation (Possible ideas)

- y_1, y_2 big norm.
- (Idea of Lyubashevsky):
  - Choose a *distrib.* for y_1, y_2 so that they have a small norm;
  - Do *rejection sampling* on z_1, z_2 so that they don't depend on the private key.

What is the rejection sampling?

It is a technique for sampling from a distribution F, when we only know how to sample from a distribution G.

Intuition

Rejection Sampling in two Lines

- Sample $x \leftarrow G$;
- Accept x w.p. $\frac{F(x)}{M \cdot G(x)}$.

Observation

The number of steps we need to accept x is M.

Rejection Sampling and Lyubashevsky

Why do we care?

- Chiave privata: s
- Challenge: c
- Response: $z = y + sc$.

We would like the distribution of z to be independent of s.

We are only able to create objects that follow a distribution dependent on s.

Understanding the dependence

To understand how to eliminate it

Example in one dimension

$y \in [-10, \ldots, 10]$
$s \in [-1, 0, 1]$
$z = y + s$

Understanding the dependence

To understand how to eliminate it

Schema di Lyubashevsky

$y \in [-nb, \dots, nb]^n$, with $b := \max_{s,c} ||sc||_\infty$
$s \in [-1, 0, 1]^n$
$z = y + sc$

How likely is a good sample?

That is: how many times do we have to repeat rejection sampling before we find a good z?

Conclusion

If we do rejection sampling in this way we obtain vectors z such that:

- Indipendent from s.
- Uniformly distributed in $[-nb + b, nb - b]$.
- They have norm (mean) $||z||_2 \approx n^{1.5}$.

Recap

$$sk: s_1, s_2 \in R \text{ of "short norm"}$$
$$pk: a \in R, t = as_1 + s_2$$

PROVER (sk, pk)		*VERIFIER* (pk)
Pick $y_1, y_2 \sim D$		
	$\xrightarrow{w = ay_1 + y_2}$	
		Pick sparse $c \in R, \|\|c\|\|_1$ "small"
	$\xleftarrow{c}$	
$z_1 = y_1 + s_1 c$ $z_1 = y_2 + s_2 c$		
	$\xrightarrow{z_1, z_2}$	Check that

- $\|\|z_1\|\|, \|\|z_2\|\|$ are "small";
- $az_1 + z_2 - tc = w$.

Question

$||z_i||$ smaller $\longrightarrow$ Ring-SIS harder.
Is it possible to change the pair (distribution - rejection sampling)
to get a lower $||z_i||$?

Answer

It is possible with Gaussian distributions

Memo

Gaussian Distribution

Gaussian on n-dimension: $\rho_{\sigma,v}(x) := \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{||x-v||^2}{2\sigma^2}}$

Discretized normal: $D_{\sigma,v}(x) := \frac{\rho_{\sigma,v}(x)}{\sum_{y \in \mathbb{Z}^n} \rho_{\sigma,v}(y)}$

Claim

If $y_1, y_2 \sim D_{\sigma,v}$ then z_1, z_2 are smaller than the ones obtained with the previous rejection sampling method.

Let's Get to work

$z := y + sc$ will have a normal distribution centered in sc.
What do we have? $D_{\sigma,sc}$
What do we want? $D_{\sigma,0}$

Observation

We have already seen how to transform the first distribution into the second:

- Sample $x \leftarrow D_{\sigma,sc}$.
- Accept w.p. $\frac{D_{\sigma,0}}{M \cdot D_{\sigma,sc}}$.

How Do We Estimate the Earning?

The number of steps I need to accept x is M.

We know that $\frac{D_{\sigma,0}}{M \cdot D_{\sigma,sc}} \leq 1$ and we want that $M \approx e$

Let's impose $\frac{D_{\sigma,0}}{D_{\sigma,sc}} \leq e$

$\downarrow$

We get an estimate of σ

$\downarrow$

We get an estimate of $||z||$

Comparison

- Before: $||z|| \approx n^{1.5}$.
- After: $||z|| \approx 12n$.

Advantage

- Same prob. as before.
- Shorter vectors.

Can we do Better?

Observation (Memo: rejection sampling)

- Sample $x \leftarrow D_{\sigma,sc}$.
- Accept w.p. $\frac{D_{\sigma,0}}{M \cdot D_{\sigma,sc}}$.

What Did We Do Before??

Is There Any Distribution That Wraps $D_{\sigma,0}$ With Less Effort?

How Efficient is This Approach?

$$\frac{D_{\sigma,0}}{\frac{1}{2}\left(D_{\sigma,sc}+D_{\sigma,-sc}\right)} = \frac{e^{-\frac{||x||^2}{2\sigma^2}}}{\frac{1}{2}\left(e^{-\frac{||x+sc||^2}{2\sigma^2}}+e^{-\frac{||x-sc||^2}{2\sigma^2}}\right)}$$
$$= (\dots)$$
$$= \frac{e^{\frac{||sc||^2}{2\sigma^2}}}{\cosh\frac{\langle x,sc\rangle}{\sigma^2}}$$
$$\leq e^{\frac{||sc||^2}{2\sigma^2}}$$

This result is much better than the previous ones!

- Original proposal: $||z|| \approx n^{1.5}$.
- First variation: $||z|| \approx 12n$.
- Second variation: $||z|| \approx n/\sqrt{2}$.

Generate a Suitable Scheme

$sk: s_1, s_2 \in R$ with short norm
$pk: a \in R, t = as_1 + s_2$

PROVER (sk, pk)		*VERIFIER* (pk)
Pick $y_1, y_2 \sim D$		
	$\xrightarrow{w = ay_1 + y_2}$	
		Pick sparse $c \in R, \|\|c\|\|_1$ "small"
	$\xleftarrow{c, b \in \{0,1\}}$	
$z_1 = y_1 + b \cdot s_1 c$ $z_1 = y_2 + b \cdot s_2 c$		
	$\xrightarrow{z_1, z_2}$	Check that

- $||z_1||, ||z_2||$ are "small";
- $az_1 + z_2 - tc = w.$

Final Touches

The verifying phase is not working right now

Takeaway

Thanks