Homomorphic Encryption and Lattices, Spring 2011 Instructor: Shai Halevi

Collision-Resistant Hashing from Ideal Lattices

May 12, 2011 Scribe: Aviad Rubinstein

1 Algebraic Background (Reminders)

Definition 1. A *commutative ring with unity* $(R, +, \times)$ satisfies the following properties:

- $(R, +)$ is an abelian group;
- $(R, \times)$ is associative, commutative and has unity;
- $\times$ distributes over $+$

(Similar to a field, except not every non-zero element has an inverse)

Examples: $\mathbb{Z}, \mathbb{Z}[x]$

Definition 2. An *ideal* $I \subseteq R$ is a subset that satisfies:

- $(I, +)$ is a subgroup of $(R, +)$;
- I is closed under multiplication by R ($i \times r \in I\, \forall i \in I$, $r \in R$).

Definition 3. We say that an ideal is *finitely generated* if there is a finite set of generators $\{i_1, \ldots, i_t\}$ s.t. $I = \{\sum r_j \times i_j\} = \langle i_1, \ldots, i_t \rangle$.

Definition 4. For a ring R and an ideal I we can define the *quotient ring* to be the set $Q = R/I = \{[r]_I : r \in R\}$ where $[r]_I = \{r + i \colon i \in I\}$.

Examples:
$\mathbb{Z}_5 = \mathbb{Z}/5\mathbb{Z}$, $\mathbb{Z}[x] / \langle x^4 + 1 \rangle$, $\mathbb{Z}_5[x] / \langle x^4 + 1 \rangle = \mathbb{Z}[x] / \langle 5, x^4 + 1 \rangle$.

We will mostly be interested in rings of the forms $\mathbb{Z}[x] / \langle f \rangle$ and $\mathbb{Z}_p[x] / \langle f \rangle$ for some monic polynomial f. Note that if f is not irreducible, then $\mathbb{Z}[x] / \langle f \rangle$ has zero divisors: $a \times b = f$, $0 < \deg a, b < \deg f$, therefore $a, b \in \mathbb{Z}[x] / \langle f \rangle$. Thus we will mostly be interested in quotient rings over irreducible polynomials.

We can represent an element of $R = \mathbb{Z}[x] / \langle f \rangle$ by the vector of its coefficients:

$$\begin{aligned} (\deg \leq n-1)\text{-polynomial } g \in R \quad &\Longleftrightarrow \quad \text{vector } \vec{g} \in \mathbb{Z}^n \\ \text{additive subgroup of } R \quad &\Longleftrightarrow \quad \text{lattice in } \mathbb{Z}^n \end{aligned}$$

Other representations are also possible, but we do not deal with them here.

Definition 5. A lattice $\Lambda \in \mathbb{Z}^n$ is an *ideal lattice* if there exist a ring $R = \mathbb{Z}[x] / \langle f \rangle$ and an ideal $I \subseteq R$ s.t. Λ is associated with I.

For example:

$$\Lambda = \left\{\text{coeff. vector of } g() \mid \exists h() \text{ s.t. } g(x) = h(x) \times \left(x^2 + 5\right) \pmod{\left(x^4 + 1, 8\right)}\right\} \subset \mathbb{Z}^4$$

is associated with the ideal

$$\left\langle x^2 + 5 \right\rangle \subset \mathbb{Z}_8 / \left\langle x^4 + 1 \right\rangle$$

Lemma 1. $\mathbb{Z}[x]$ *is a Unique Factorization Domain (UFD)*

Proof. Follows since (a) $\mathbb{Z}$ is a UFD; and (b) if a ring R is a UFD then so is $R[x]$. $\square$

Corollary 1. *If $f \in \mathbb{Z}[x]$ is irreducible and $f \mid g \times h$ then $f \mid g$ or $f \mid h$.*

Lemma 2. *If f is irreducible and $I \subseteq \mathbb{Z}[x] / \langle f \rangle$ is a non-zero ideal then Λ_I is a full-rank lattice in $\mathbb{Z}^n$.*

Proof. Let $g \in I$ be a non-zero element. We will show that $\left\{g, x \times g, \ldots x^{n-1} \times g\right\}$ are linearly independent. Assume $\sum h_i x^i \times g(x) = 0$. Since $g(x) \in \mathbb{Z}[x] \subset \mathbb{Q}[x]$, we can assume wlog that $h_i \in \mathbb{Z}$: If $\left\{x^i \times g\right\}$ are linearly dependent, then, using Gauss elimination, we can find rational $\overline{h_i}$'s s.t. $\sum \overline{h_i} x^i \times g(x) = 0$. Multiply by the least common denominator of the $\overline{h_i}$'s we get integral h_i's.

Denote $h(x) = \sum h_i x^i$. Then $h \times g = 0$ in the ring, (which is the same as $f \mid h \times g$, since we are working modulo f). Thus by the previous corollary, and since $g \neq 0$, we have that $h = 0$. $\square$

Definition 6. For any polynomial f we define:

$$\theta(f) = \max_{a \in \mathbb{Z}[x]/\langle f \rangle,\, i < \deg f} \frac{\left\| x^i \times a \bmod f \right\|_\infty}{\|a\|_\infty}$$

For any two polynomials $a, b \in \mathbb{Z}[x] / \langle f \rangle$:

$$\|a \times b \bmod f\|_\infty \leq \|a\|_\infty \cdot \|b\|_\infty \cdot n\theta(f)$$

Lemma 3. *If f is an n-degree irreducible polynomial and $I \subseteq \mathbb{Z}[x] / \langle f \rangle$ is a non-zero ideal then*

$$\lambda_n^\infty(\Lambda_I) \leq \lambda_1^\infty(\Lambda_I) \cdot \theta(f)$$

Proof. Let $\vec{g}$ be the shortest vector in Λ, and $g(x)$ the corresponding polynomial. Then the vectors that correspond to $\left\{g_i(x) = x^i \times g(x)\right\}$ are linearly independent. Clearly,

$$\lambda_n^\infty(\Lambda_I) \;\leq\; \max_{i=0}^{n-1} \|\vec{g}_i\|_\infty \;\leq\; \|\vec{g}\|_\infty \cdot \theta(f) \;=\; \lambda_1^\infty(\Lambda_I) \cdot \theta(f)$$

$\square$

2 The Shortest Vector Problem in Ideal Lattices

Just as in any lattice, we can ask how easy / hard it is to find a (good approximaiton of) the shortest vector in the lattice. Note that if $\theta(f)$ is small, then by the previous lemma estimating the *size* of the smallest vector is easy:

$$\begin{aligned}
\frac{1}{\sqrt{n}}\lambda_1(\Lambda) &\leq \det(\Lambda)^{1/n} \\
&\leq \lambda_n(\Lambda) \\
&\leq \sqrt{n}\lambda_n^\infty(\Lambda) \\
&\leq \sqrt{n}\lambda_1^\infty(\Lambda) \cdot \theta(f) \\
&\leq \sqrt{n}\lambda_1(\Lambda) \cdot \theta(f)
\end{aligned}$$

Still, finding the shortest vectors themselves seems hard. In particular, we don't know of methods that do much better on ideal lattices than on regular lattices. (Sometimes we can do slightly better, for example in [GS02] they are able to reduce an ideal lattice problem in dimension $2n$ to a non-ideal lattice problem in dimension n.)

2.1 The f-SVP$_\gamma$ Problem

For a family of polynomials $f = \{f_n\}$ (with $\deg f_n = n$): Given a lattice Λ_I corresponding to ideal $I \subseteq \mathbb{Z}[x] / \langle f_n \rangle$, find a non-zero vector $\vec{v} \in \Lambda_I$ s.t. $\|\vec{v}\| \leq \gamma \lambda_1(n)$. (Can also be stated with l_∞ or any other l_p norm).

Below we will typically use $f_n(x) = x^n + 1$, where n is a power of 2. Thus $f_n(x)$ is irreducible. Also, $\theta(f) = 1$ because:

- For any $g(x)$ with coefficient vector $(g_1, g_2, \ldots, g_n)$, we have that the coefficient vector of $x \times g(x)$ is $(-g_n, g_1, \ldots, g_{n-1})$.
- This means that lattices over $\mathbb{Z}[x] / \langle f_n \rangle$ are "almost circular": If $(v_i) \in \Lambda_I$, then also $\left(\text{sign}(i-k-1) \cdot v_{i-k \pmod n}\right) \in \Lambda_I$.

3 The Collision-Resistant Construction of [PR06], [LM06]

Parameters:

- n - security parameter
- d - size of entries in input (e.g. $d = \sqrt{n}$)
- p - modulus (e.g. $p = n^3$)
- m - another parameter (e.g. $m = 8$)

We need $m > \frac{\log p}{\log 2d}$ to make the function contracting, and $p > 4dmn^{1.5} \log n \cdot \theta^2(f)$ for the reduction to work. We get a hash function with keys of size $mn\lceil \log p \rceil$ bits, hashing $mn\lceil \log 2d \rceil$ inputs into $n\lceil \log p \rceil$-bits output.

- The construction is over the ring $R_n = \mathbb{Z}_p[x] / \langle f_n \rangle$.
- Keys: m-vector of ring elements $\vec{a} = (a_i)_{i=1}^m \in R_n^m \cong \mathbb{Z}_p^{mn}$
- Inputs: m-vector of small ring elements $\vec{x} = (a_i)_{i=1}^m$ with $\|x_i\|_\infty < d$ (so $\vec{x} \in \mathbb{Z}_{2d}^{mn}$).
- Hash function: Inner product over R_n:

$$h_{\vec{a}}(\vec{x}) = \sum_{i=1}^m \vec{a_i} \times \vec{x_i} \in R_n$$

Example: $m = 2$, $n = 4$, $p = 13$, and $\vec{a} = ((1,2,3,4), (5,6,7,8))$.

$$h_{\vec{a}}((x_1, x_2, x_3, x_4), (x_5, x_6, x_7 x_8)) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -4 & 1 & 2 & 3 \\ -3 & -4 & 1 & 2 \\ -2 & -3 & -4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} 5 & 6 & 7 & 8 \\ -8 & 5 & 6 & 7 \\ -7 & -8 & 5 & 6 \\ -6 & -7 & -8 & 5 \end{pmatrix} \begin{pmatrix} x_5 \\ x_6 \\ x_7 \\ x_8 \end{pmatrix} \pmod{13}$$

Theorem 1. *If there is an efficient algorithm that finds collisions with noticable probability on a random hash key $\vec{a} \in R_n^m$, then there is an efficient algorithm that given* <u>any ideal lattice</u> *Λ_I for $I \subseteq R_n$ finds a vector $\vec{v} \in R_n$ s.t.*

$$\|\vec{v}\|_\infty = \lambda_1^\infty(\Lambda) \cdot \underbrace{16dmn\log^2 n \cdot \theta^2(f)}_{\gamma}$$

- Note: this is a reduction from an average case collision finder to a worst case SVP-approximation - but only worst case for ideal lattices over R_n. This result can be extended to a single construction that will hold for SVP-approximation over ideal lattices in **any ring** $\mathbb{Z}[x]/\langle f\rangle$ provided that f is irreducible and has small $\theta(f)$.

- Note also the similarity to SIS-based hash function: we can write this hash function as $h_{\vec{a}}(\vec{x}) = A\vec{x} \pmod{p}$, where $A \in \mathbb{Z}_p^{mn\times mn}$ and $\vec{x} \in \mathbb{Z}_p^{mn}$, $\|\vec{x}\|_\infty < d$, so it really is a special case of SIS, but A is very far from what we had in the SIS case. Still, the proof shares many simliarities with Ajtai's proof. (Also, note that p here is of polynomial size, so we use Ajtai's iterative reduction, rather than the all-at-once we did in class).

Proof. We will show a procedure that given a vector $0 \neq \vec{g} \in \Lambda_I$, uses the collision finder to find another vector $0 \neq \vec{h} \in \Lambda_I$ s.t. $\|\vec{h}\|_\infty < \|\vec{g}\|_\infty/2$. This procedure will succeed as long as $\|\vec{g}\|_\infty > \lambda_1^\infty(\Lambda)\cdot\gamma$. Repeating this procedure poly(n) times will get our short vector. Below, we denote by $B(\vec{g})$ the matrix with columns $\{\vec{g}, x\vec{g}, \ldots, x^{n-1}\vec{g}\}$. We also denote by $P(\vec{g})$ the parallelpiped that corresponds to $B(\vec{g})$.

For simplicity assume at first that $\vec{g}$ generates I (so $I = \langle\vec{g}\rangle$), so $B(\vec{g})$ is a basis for Λ_I. For this simplified case we get Algorithm 1. In the algorithm below, the × operation is defined over $\mathbb{R}[X]/\langle f_n\rangle$ (which includes the ring $R_n = \mathbb{Z}[X]/\langle f_n\rangle$).

```
Algorithm 1. Half-Size($\Lambda_I, \vec{g}$) with oracle access to the collision finder CF:
  1. For $i = 1$ to $m$
      (a) Choose at random $\vec{y_i} \leftarrow N(0, s^2)^n$ for $s = \frac{\|\vec{g}\|_\infty}{\gamma}\sqrt{n}\log n \cdot \theta(f)$.
          Note that $s > \lambda_n^\infty(\Lambda_I)\cdot\sqrt{n}\cdot\log n$ if $\|\vec{g}\|_\infty > \lambda_1^\infty(\Lambda)\cdot\gamma$.
      (b) Let $\vec{y_i}' = \vec{y_i} \bmod B(\vec{g})$. Note that $\vec{y'}_i = \vec{y_i} - \vec{k_i} \times \vec{g}$ for some $\vec{k_i} \in R_n$.
      (c) Let $\vec{w_i}$ be the representation of $\vec{y_i}$' in base $B(\vec{g})$, i.e., $\vec{y}' = \vec{w_i}B(\vec{g})$
          Note: $\vec{y_i}' = \vec{g} \times \sum_{j=0}^{n-1} w_{ij}x^i$. Also, $w_{ij} \in [0,1)$.
      (d) Let $a_{ij} = \lceil p w_{ij} \rfloor$.
  2. Run the CF on $(\vec{a_1}, \ldots, \vec{a_m}) \in R_n^m$: $((\vec{u_i}), (\vec{v_i})) \leftarrow CF(\vec{a_i})$
     Note: if CF succeeds, then $\sum_i \vec{a_i} \times (\vec{u_i} - \vec{v_i}) = 0$
  3. Let $\vec{z_i} = \vec{u_i} - \vec{v_i}$
  4. Output $h = \sum_{i=1}^m \left(\vec{g} \times \left(\vec{w_i} - \frac{\vec{a_i}}{p}\right) - \vec{y_i}\right) \times \vec{z_i}$
```

We need to show that with noticable probability we have $0 \neq \vec{h} \in \Lambda_I$ and $\|\vec{h}\|_\infty < \|\vec{g}\|_\infty/2$.

Claim 1. Ths distribution of the vectors $\{\vec{a_i}\}_{i=1}^m$ created by the algorithm above is statistically close to uniform over R_n^m.

Proof. Since $B(\vec{g})$ is a basis of Λ_I and $s > \lambda_n^\infty(\Lambda_I)\sqrt{n}\log n$ then choosing $\vec{y_i} \leftarrow N(0, s^2)^n$ and reducing mod $B(\vec{g})$ yields a distribution which is nearly uniform over $P(\vec{g})$. Hence $\vec{w_i}$ is nearly uniform over $[0,1)^n$, and $\vec{a_i} = \lceil p\vec{w_i} \rfloor$ is nearly uniform over $\mathbb{Z}_p^n$. □

Thus, the CF succeeds with noticeable probability. From now on we will consider only the case in which it succeeds.

Claim 2. $\vec{h} \in I$

Proof.

$$\begin{aligned}
\vec{h} &= \sum_i \left(\vec{g} \times \left(\vec{w_i} - \frac{\vec{a_i}}{p}\right) - \vec{y_i}\right) \times \vec{z_i} \;=\; \sum_i \left(\vec{g} \times \left(\vec{w_i} - \frac{\vec{a_i}}{p}\right) \underbrace{- \vec{y_i'} - \vec{k_i} \times \vec{g}}_{\vec{y_i} = \vec{y_i'} + \vec{k_i} \times \vec{g},\ \vec{k_i} \in R_n} \right) \times \vec{z_i} \\
&= \sum_i \left(\underbrace{\vec{g} \times \vec{w_i} - \vec{y_i'}}_{=0}\right) \times \vec{z_i} \;-\; \vec{g} \times \sum_i \vec{k_i} \times \vec{z_i} \;+\; \vec{g} \times \left(\frac{1}{p} \cdot \underbrace{\sum_i \vec{a_i} \times \vec{z_i}}_{\equiv 0 \pmod p}\right) \\
&= \vec{g} \times \text{some integer polynomial} \;\in\; \langle \vec{g} \rangle \;=\; I
\end{aligned}$$

□

Claim 3. w.h.p.:

$$\left\| \vec{h} \right\|_\infty < \left\| \vec{g} \right\|_\infty / 2$$

Proof. We are guaranteed that:

$$\begin{aligned}
\vec{h} &= \sum \left(\vec{g} \times \left(\vec{w_i} - \frac{\vec{a_i}}{p}\right) - \vec{y_i}\right) \times \vec{z_i} \\
&= \sum \vec{g} \times \left(\vec{w_i} - \frac{\vec{a_i}}{p}\right) \times \vec{z_i} + \sum \vec{y_i} \times \vec{z_i} \\
\Rightarrow \left\| \vec{h} \right\|_\infty &\leq \sum \left\| \vec{g} \times \left(\vec{w_i} - \frac{\vec{a_i}}{p}\right) \times \vec{z_i} \right\|_\infty + \sum \left\| \vec{y_i} \times \vec{z_i} \right\|_\infty
\end{aligned}$$

Now we have:

$$\begin{aligned}
\left\| \vec{w_i} - \frac{\vec{a_i}}{p} \right\|_\infty &\leq 1/p \\
\left\| \vec{z_i} \right\|_\infty &\leq 2d \\
\left\| \vec{y_i} \right\|_\infty &< s \cdot \log n \cdot \sqrt{n} \text{ w.h.p.}
\end{aligned}$$

From definition of $\theta(f)$:

$$\begin{aligned}
\left\| \vec{g} \times \left(\vec{w_i} - \frac{\vec{a_i}}{p}\right) z_i \right\|_\infty &\leq n^2 \theta^2(f) \left\| \vec{g} \right\|_\infty \left\| \vec{w_i} - \frac{\vec{a_i}}{p} \right\|_\infty \left\| \vec{z_i} \right\|_\infty \\
&\leq n^2 \cdot \theta^2(f) \cdot \left\| \vec{g} \right\|_\infty \cdot \frac{1}{p} \cdot 2d \\
\left\| \vec{y_i} \times \vec{z_i} \right\|_\infty &\leq n\theta(f) \left\| \vec{y_i} \right\|_\infty \cdot \left\| \vec{z_i} \right\|_\infty \\
&\leq n\theta(f)\, s \cdot \log n \cdot \sqrt{n} \cdot 2d \text{ w.h.p.}
\end{aligned}$$

This is not quite small enough for the parameters that we stated before. We can either use this bound as is and get slightly worse parameters, or use independence to improve the bound (w.h.p.).

(More details can be found in [Lyu08].) The resulting bounds are:

$$\begin{aligned}\left\|\vec{g}\times\left(\vec{w_i}-\frac{\vec{a_i}}{p}\right)z_i\right\|_\infty &\leq n^{1.5}\log n\cdot\theta^2\left(f\right)\cdot\left\|\vec{g}\right\|_\infty\cdot\frac{1}{p}\cdot d\\ &\leq\frac{\left\|\vec{g}\right\|_\infty}{4m}\\ \left\|\vec{y_i}\times\vec{z_i}\right\|_\infty &\leq 4\theta\left(f\right)s\cdot\log n\cdot\sqrt{n}\cdot d\\ &\leq\frac{\left\|\vec{g}\right\|_\infty}{4m}\end{aligned}$$

Thus,

$$\begin{aligned}\left\|\vec{h}\right\|_\infty &\leq\sum\left(\frac{\left\|\vec{g}\right\|_\infty}{4m}\right)+\sum\left(\frac{\left\|\vec{g}\right\|_\infty}{4m}\right)\text{ w.h.p.}\\ &\leq\frac{\left\|\vec{g}\right\|_\infty}{2}\end{aligned}$$

□

Claim 4. w.h.p.: $\vec{h}\neq 0$.

Proof. Note that $\vec{w_i}$'s and $\vec{z_i}$'s depend only on in which cosets of Λ_I $\vec{y_i}$'s fall. Hence we can view the process of randomly choosing $\vec{y_i}$'s as first choosing the cosets (and thus determining $\vec{w_i}$'s and $\vec{z_i}$'s), and then choosing $\vec{y_i}$'s from the induced distributions, $D_{coset_i,s}$. Since the $\vec{z_i}$'s are not all zero, we can assume w.l.o.g. that $\vec{z_1}\neq 0$. For any fixed choice of $\vec{w_i}$'s, $\vec{z_i}$'s, $\vec{a_i}$'s, and $\vec{y_i}\mid_{i>1}$, we have that $h=0$ iff:

$$\vec{y_1}\times\vec{z_1}=-\sum\vec{g}\times\left(\vec{w_i}-\frac{\vec{a_i}}{p}\right)\times\vec{z_i}-\sum_{i=2}^{m}\vec{y_i}\times\vec{z_i}$$

Since f_n is irreducible, there is at most (in the appropriate coset) that will make this equation true. Since s is sufficiently large (in particular, $s>\lambda_n\left(\Lambda_I\right)\sqrt{n}\log n$) then $D_{coset_i,s}$ is sufficiently close to uniform. Thus the probability of choosing $\vec{y_1}$ that will satisfy the equation is negligible, and thus also the probability that $\vec{h}=0$. □

□

3.1 Removing the Assumption $I=\langle\vec{g}\rangle$

Clearly $\langle\vec{g}\rangle\subseteq I$. The only place where we used equality was in the proof of Claim 1. We could set the parameter $s\geq\lambda_n^\infty\left(\Lambda_{\langle\vec{g}\rangle}\right)$, and then the proof would go as before, but we will no longer have a bound on $\frac{s}{\lambda_n^\infty(\Lambda_I)}$.

Instead, we modify the reduction. We choose a random coset of $I/\langle\vec{g}\rangle$: $\langle\vec{g}\rangle+t$. We now set $\vec{y_i}'=\vec{y_i}+\vec{t}(\bmod B\left(\vec{g}\right))$. Since $\vec{y_i}$ is nearly uniform mod $B\left(\Lambda_I\right)$, with the addition of a random coset we have that $\vec{y_i}'$ is nearly uniform in $B\left(\vec{g}\right)$, as needed for Claim 1.

Note that we still have $\vec{y_i}'-\vec{y_i}\in I$, as needed for Claim 2. The proofs of Claim 3 and Claim 4 remain unaffected.

References

[GS02] Craig Gentry and Michael Szydlo, *Cryptanalysis of the revised NTRU signature scheme*, Advances in Cryptology - EUROCRYPT'02, Lecture Notes in Computer Science, vol. 2332, Springer, 2002, pp. 299–320.

[LM06] Vadim Lyubashevsky and Daniele Micciancio, *Generalized compact knapsacks are collision resistant*, ICALP'06 (2), Lecture Notes in Computer Science, vol. 4052, Springer, 2006, pp. 144–155.

[Lyu08] Vadim Lyubashevsky, *Towards practical lattice-based cryptography*, Ph.D. thesis, University of California, San Diego, 2008.

[PR06] Chris Peikert and Alon Rosen, *Efficient collision-resistant hashing from worst-case assumptions on cyclic lattices*, TCC'06, Lecture Notes in Computer Science, vol. 3876, Springer, 2006, pp. 145–166.