CSE 291: Lattices in Cryptography and Cryptanalysis Nov. 18, 1999

Lecture 9

Lecturer: Daniele Micciancio *Scribe: Anand Desai*

In the last lecture, we saw how lattices could be used to find small solutions to univariate polynomial equations. In this lecture we will present a general solution to finding small solutions to *multivariate linear equations*.

The problem is motivated by its relation to the subset sum problem, which is stated below for reference.

Definition 1 (Subset Sum Problem) *Given a set $\{a_1, a_2, \cdots, a_n\}$ of positive integers, a positive integer b and a modulus m, determine the $x_i \in \{0,1\}, 1 \leq i \leq n$, such that $\sum_{i=1}^{n} a_i x_i = b \bmod m$, provided that such x_i exist.*

The subset sum problem is **NP**-hard. The average-case difficulty of this problem has been the basis of the (presumed) security of a number of cryptographic primitives. The algorithm we describe in the next section can be used to solve subset sum problems of low "density". The density of a set such as $A = \{a_1, a_2, \cdots, a_n\}$ provides a measure of the size of the elements in the set A. For a modulus m of l bits, the subset sum density δ is $\frac{n}{l}$.

1 Finding small solutions to multivariate linear equations

We want to solve the following problem.

Problem 1 *Given a system of linear equations:*

$$A\vec{x} = \vec{b} \bmod \vec{m}$$

where $A = (a_{ij})_{k \times n}$ and $\vec{b} = \{b_1, b_2, \cdots, b_k\}$ and $\vec{m} = \{m_1, m_2, \cdots, m_k\}$ and for $1 \leq i \leq k : |b_i| = l$ find a small solution in $\vec{x}$ such that $\|\vec{x}\| \leq \sqrt[n]{2^{k \cdot l}}$.

Note that it is possible that $k \leq n$, in which case there may not be a unique solution and in general we could expect exponentially many solutions. Also we cannot hope to use Gaussian Elimination since the system of equations could be in different modulii. As before, we first show how to find solutions when they are smaller than some bound X to be determined. Then we will show that $X \approx \sqrt[n]{2^{k \cdot l}}$.
This problem could be solved using the LLL-algorithm-based technique we saw in the last lecture. However we will take a different approach which may be viewed as a dual to the above. The key is that we view solutions as vectors.

The main points of the algorithm are the following:

- Define a suitable basis B and a corresponding target vector $\vec{t}$.
- Look for a lattice vector $\vec{v} \in \mathcal{L}(B)$ close to $\vec{t}$ (using Babai's approximation algorithm for the closest vector problem to be defined).

Define the following basis:

$$B = \begin{bmatrix} \alpha M & \alpha A \\ \vec{0} & I_n \end{bmatrix}$$

where

- M is a $k \times k$ matrix (e_{ij}) with $e_{ii} = m_i$ and $e_{ij} = 0$ for $i \neq j$
- $A = (a_{ij})$ is the given $k \times n$ matrix
- $\vec{0}$ is a $n \times k$ zero matrix
- I_n is an n-dimensional identity matrix
- α is a scalar multiplier whose value will become apparent later.

It is easy to see that $\exists \vec{r}$:

$$B \begin{bmatrix} \vec{r} \\ \vec{x} \end{bmatrix} = \begin{bmatrix} \alpha(A\vec{x} + M\vec{r}) \\ \vec{x} \end{bmatrix} = \begin{bmatrix} \alpha\vec{b} \\ \vec{x} \end{bmatrix}$$

Note that

$$\mathcal{L}(B) = \left\{ B \begin{bmatrix} \vec{r} \\ \vec{x} \end{bmatrix} : \forall r, x \in \mathbb{Z} \right\}$$

Now consider the vector $\vec{t} = \begin{bmatrix} \alpha\vec{b} \\ \vec{0} \end{bmatrix}$.

We know that there are lattice vectors in $\mathcal{L}(B)$ at a distance X from $\vec{t}$. Before we explain how to find these we need to introduce a new problem.

Definition 2 (Closest Vector Problem (CVP)) *Given a basis B and a target vector $\vec{t}$ find the lattice vector $\vec{v} \in \mathcal{L}(B)$ closest to $\vec{t}$.*

CVP is **NP**-hard. However there is a polynomial time approximation algorithm for CVP due to Babai. We will study this algorithm in the next lecture. For now we just assume that there exists a polynomial time algorithm that solves the following problem. Let $d(\mathcal{L}(B), \vec{t})$ be the distance of the closest vector in $\mathcal{L}(B)$ to $\vec{t}$.

Definition 3 (Approximate CVP) *Given a basis $B \in \mathbb{Z}^{m \times n}$ and a target vector $\vec{t} \in \mathbb{Z}^m$ find the lattice vector $\vec{v} \in \mathcal{L}(B)$ such that $\|\vec{v} - \vec{t}\| \leq 2^n \cdot d(\mathcal{L}(B), \vec{t})$.*

We will look at Babai's algorithm in the next lecture. Continuing with our problem,

$$d(\mathcal{L}(B), \vec{t}) \leq \left\| \begin{bmatrix} \alpha\vec{b} \\ \vec{x} \end{bmatrix} - \begin{bmatrix} \alpha\vec{b} \\ \vec{0} \end{bmatrix} \right\| = \left\| \begin{bmatrix} \vec{0} \\ \vec{x} \end{bmatrix} \right\| \leq X$$

Applying Babai's algorithm we get some $\begin{bmatrix} \vec{s} \\ \vec{y} \end{bmatrix} \in \mathbb{Z}^{n+k}$ such that

$$\left\| B \begin{bmatrix} \vec{s} \\ \vec{y} \end{bmatrix} - \vec{t} \right\| = \begin{bmatrix} \alpha(A\vec{y} + M\vec{s} - \vec{b}) \\ \vec{y} \end{bmatrix} . < 2^{n+k} \cdot X.$$

Let $\alpha = 2^{n+k} \cdot X$. Since the first entry in the above vector is an integer multiple of α, it must be 0 and $A\vec{y} + M\vec{s} - \vec{b} = \vec{0}$, i.e., $A\vec{y} = \vec{b} \bmod M$. Also

$$\left\| \begin{bmatrix} \vec{0} \\ \vec{y} \end{bmatrix} \right\| = \|\vec{y}\| \leq 2^{n+k} \cdot X.$$

So, we found a solution to the system of equations of size at most $\|\vec{y}\| < 2^{n+k}X$. We claim that for most A we have $\vec{x} = \vec{y}$ and therefore $\|\vec{y}\| < X$ (in general this isn't true).
We will show that there exists a unique solution $\vec{y}$ such that $\|\vec{y}\| < \alpha$.
Let $\vec{z} = \vec{x} - \vec{y}$. Then $A\vec{z} = \vec{0} \bmod m$.
Consider the set $H = \{\vec{z} : A\vec{z} = \vec{0} \bmod m\}$ of all possible solutions. Notice that H is a lattice because it is a set of integer vectors closed under addition.
We want to show that length of the shortest vector in H, $\lambda(H) > (2^{k+n+1} + 1)X$.
Then by the triangle inequality ($\|\vec{x} - \vec{y}\| < X + X \cdot 2^{k+n} < \lambda(H)$) we get that $\vec{x}$ and $\vec{y}$ must be the same vector.
It remains to pick the bound X. We analyze the problem when all the m_i are the same l bit prime. If the modulii are different square free numbers, a similar analysis applies.
Then $\Pr[\exists \vec{z} \neq \vec{0} \text{ s.t. } \|z\| < (2^{n+k+1})X,\ A\vec{z} = \vec{0} \bmod m]$

$$\begin{aligned} &\leq (2 \cdot 2^{n+k+1} \cdot X)^n \cdot \Pr[A\vec{z} = \vec{0} \bmod m \mid z] \\ &= (2 \cdot 2^{n+k+1} \cdot X)^n \cdot \left(\frac{1}{m}\right)^k \end{aligned}$$

We pick $X \leq \frac{m^{\frac{k}{n}}}{2^{n+k+3}} \leq 2^{\frac{kl}{n} - n - k - 3}$.