The $N = 1$ case of NTRU

Jeff Hoffstein

October 12, 2008

1 The $N = 1$ algorithm

In this section we will present a toy model of a real public key cryptosystem. It will turn out to have an unexpected connection with lattices of dimension 2, and hence a fatal vulnerability as the dimension is so low. However it is of some interest. For one thing it illustrates a tendency of lattices to be useful in cryptanalysis even though the problem involved may appear to have nothing to do with lattices. And for another thing, it provides a simple, low-dimensional example of NTRU, (another lattice-based cryptosystem).

In the following, Alice will set up a public-private key pair. She begins by fixing a large integer q. She then chooses smaller integers f, g satisfying $(f, q) = 1$ and

$$f, g < \sqrt{q/2}.$$

The integer g is also chosen to satisfy $g > \sqrt{q/4}$. Alice computes $h \equiv f^{-1}g \pmod{q}$ and posts the pair q, h as her public key. Alice does not reveal the pair f, g, her private key. Although f and g are small compared to q, i.e. they are $\mathcal{O}(\sqrt{q})$, the quantities $f^{-1} \pmod{q}$ and h will, with high probability, be $\mathcal{O}(q)$. Bob would now like to encrypt and send a secret message m to Alice. The message m is chosen so that $m < \sqrt{q/4}$. To accomplish the encryption, he chooses an integer r at random, $r < \sqrt{q/2}$ and computes $e \equiv rh + m \pmod{q}$. Bob sends the encrypted message e to Alice. To decrypt e, Alice first computes $fe \pmod{q}$ and chooses $a \equiv fe \pmod{q}$ with $0 \leq a < q$. Alice then computes $m' \equiv f^{-1}a \pmod{g}$ with $0 < m' < g$. But $m \equiv f^{-1}a \pmod{g}$ also, and as $0 < m < g$, it follows that $m = m'$ and Arthur has recreated m. This process is summarized in Table 1

Alice	Bob
Key Creation	
Chooses a large integer q. Chooses secret integers f, g, with $(f, q) = 1$, $f, g < \sqrt{q/2}$, and $g > \sqrt{q/4}$. Computes $h \equiv f^{-1}g \pmod{q}$. Publishes q and h.	
Encryption	
	Chooses plaintext m, with $m < \sqrt{q/4}$. Chooses random r, $r < \sqrt{q/2}$. Uses Alice's public key (q, h) to compute $e \equiv rh + m \pmod{q}$. Sends ciphtertext e to Alice.
Decryption	
Computes $a \equiv fe \pmod{q}$ with $0 \leq a < q$. Computes $m' \equiv f^{-1}a \pmod{g}$ with $0 \leq m' < \sqrt{q/4}$. Then m' equals the plaintext m.	

Table 1: N=1 case of NTRU: key creation, encryption, and decryption

We illustrate this cryptosystem with a small numerical example.

Key Creation

- Alice chooses $q = 122430513841$, $f = 231231 \approx 0.66\sqrt{q}$, $g = 195698 \approx 0.56\sqrt{q}$. Alice computes $f^{-1} \equiv 49194372303 \pmod{q}$ and $h \equiv f^{-1}g \equiv 39245579300 \pmod{q}$.
- Alice reveals her public key pair

$$q = 122430513841, h = 39245579300.$$

Encryption

- Bob converts his plaintext into an integer $m = 123456 \approx 0.35\sqrt{q}$.
- He chooses a random $r = 101010 \approx 0.29\sqrt{q}$.
- Bob uses Alice's public key $(q, h) = (122430513841, 39245579300)$ to compute

$$e \equiv rh + m \equiv 18357558717 \pmod{q}.$$

- Bob sends the ciphertext $e = 18357558717$ to Alice.

Decryption

- To decrypt e, Alice first computes

$$a \equiv fe \equiv 48314309316 \pmod{q}$$

and recovers the positive integer $a = 48314309316$ (which is *not* reduced modulo q.)

- She then computes first

$$f^{-1} \equiv 193495 \pmod{g},$$

and then

$$f^{-1}a \equiv 123456 \pmod{g},$$

which miraculously equals m.

Why does this work? When Alice computes $a \equiv fe \pmod{q}$ she is actually computing

$$a \equiv f(rh + m) \equiv frf^{-1}g + fm \equiv rg + fm \pmod{q}$$

By the size restrictions on f, g, r, m, it follows that $0 \leq rg + fm < q$. Because of this, $a \equiv rg + fm \pmod{q}$ and $0 \leq a < q$ implies that $a = rg + fm$ as an integer, not simply modulo q. Reducing $rg + fm$ modulo g now gives $fm \pmod{g}$ and multiplying by the

inverse of f modulo g recovers m modulo g. But $m < \sqrt{q/4} < g$, (this is the reason for the lower bound on g), so recovering m modulo g recovers m over the integers.

How might Eve attack this system? First of all, there is the question of recovering the private key (f, g) from the public key (q, h). To search for the private key f an attacker would multiply h by a candidate $F = \mathcal{O}(\sqrt{q})$ and compute $G \equiv Fh \pmod{q}$. If $G = \mathcal{O}(\sqrt{q})$ then the chances are very good that $f = F$ and $g = G$. This is because as F varies, the corresponding G will be distributed randomly modulo q. The chance that a random G will be this small compared to q is $\mathcal{O}(1/\sqrt{q})$, which will be quite small if q is large. Interestingly, even if $(f, g) \neq (F, G)$, the chances are still good that if F, G are $\mathcal{O}(\sqrt{q})$, then F will function perfectly well as a private key. A brute force attack on the public key would then be expected to require $\mathcal{O}(\sqrt{q})$ operations. Similarly, brute force attacks can be launched on an encrypted message. The attacker would simply try a random r, with $r = \mathcal{O}(\sqrt{q})$ and check to see if $e - rh \pmod{q}$ were small. Thus it appears that to obtain security on the order of 2^{80} one should choose $q \approx 2^{160}$.

Unfortunately for Alice, there is another, completely different attack on this system. An attacker, given the public key (q, h), can simply compute the continued fraction expansion of h/q. This will create a sequence of approximations to h/q. In the example above, one of these is the fraction $\frac{74122}{231231}$. Thus f surprisingly appears in the denominator! "What in the world is going on here" one might well ask. In Section 2 we will see why the continued fraction expansion reveals the key, and the connection between this method and lattices.

2 Continued fractions and the $N = 1$ cryptanalysis

Recall that in Chapter **??** the extended Euclidian algorithm was described. This provided a method for finding an integral solution to the equation $ax + by = \pm 1$ when a, b are relatively prime integers. The following example was given in the case $a = 73, b = 25$..

$$\begin{aligned} 73 &= 2 \cdot 25 + 23 \\ 25 &= 1 \cdot 23 + 2 \\ 23 &= 11 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0. \end{aligned}$$

The sequence of numbers $\{2, 1, 11, 2\}$ are the coefficients in what is called the *continued fraction expansion* of the rational number $73/25$. This name arises because, as is easy to check,

$$\frac{73}{25} = 2 + \frac{1}{1 + \frac{1}{11 + \frac{1}{2}}}$$

The "magic box" method then described how to use this expansion to construct an integral solution to $73x + 25y = \pm 1$:

		2	1	11	2
0	1	2	3	35	73
1	0	1	1	12	25

In particular, $x = 12, y = -35$ and $73 \cdot 12 - 25 \cdot 35 = 1$. The general situation, for $a, b > 0$ and $(a, b) = 1$ is as follows. If $\{a_1, a_2, a_3, \ldots, a_t\}$ is the continued fraction expansion of a/b then the box has the form

		a_1	a_2	$\ldots$	a_{t-1}	a_t
0	1	p_1	p_2	$\ldots$	p_{t-1}	a
1	0	q_1	q_2	$\ldots$	q_{t-1}	b

Here $p_1 = a_1, q_1 = 1, p_2 = a_2 \cdot p_1 + 1, q_2 = a_2 \cdot q_1$ and for $i \geq 3$, $p_i = a_i \cdot p_{i-1} + p_{i-2}$ and $q_i = a_i \cdot q_{i-1} + q_{i-2}$. Also

$$a \cdot q_{t-1} - b \cdot p_{t-1} = (-1)^t.$$

Thus the computation of p_{t-1}, q_{t-1} gives x and y up to a minus sign. In general, the fractions p_i/q_i are called *convergents* to a/b. To see why this is so, refer back to the example above and note that the sequence of fractions

$$2, 2 + \frac{1}{1}, 2 + \frac{1}{1 + \frac{1}{11}}, 2 + \frac{1}{1 + \frac{1}{11 + \frac{1}{2}}}$$

reduces to the sequence of fractions

$$\frac{2}{1}, \frac{3}{1}, \frac{35}{12}, \frac{73}{25}.$$

In fact every positive real number x has a continued fraction expansion and this expansion is finite if and only if x is rational. The key useful point here is that each of the convergents p_i/q_i must be particularly close to the real number x. That is, they give very good approximations to x compared to other rational numbers with a similar denominator. This is quantified by the following Theorem: Let $x > 0$ be any real number. If p_i/q_i is a convergent in the continued fraction expansion of x then

$$\left| x - \frac{p_i}{q_i} \right| < \frac{1}{q_i^2},$$

If p/q is any positive rational number with the property that

$$\left| x - \frac{p}{q} \right| < \frac{1}{2q^2},$$

then p/q must appear as a convergent in the continued fraction expansion of x.

We refer the reader to the book of Hardy and Wright [] for a proof of this theorem and a complete presentation of the beautiful theory of continued fractions.

To return to the $N = 1$ case of NTRU, recall that the relevant information was: for some large public positive integer q, and some public positive integer h less than and relatively prime to q, there exist unknown integers f, g satisfying $f, g < \sqrt{q/2}$ and

$$f^{-1}g \equiv h \pmod{q}.$$

This means that for some integer k, $fh - g = kq$. This in turn means that

$$\frac{h}{q} - \frac{k}{f} = \frac{g}{fq}$$

and as $f, g < \sqrt{q/2}$,

$$\left|\frac{h}{q} - \frac{k}{f}\right| < \frac{1}{f\sqrt{2q}} < \frac{1}{2f^2}.$$

Because of this the fraction k/f must in fact be one of the convergents in the continued fraction expansion of h/q and can be found by a simple application of the extended Euclidean algorithm.

In the previous example the continued fraction expansion has coefficients

$$\{0, 3, 8, 2, 1, 3, 3, 8, 1, 4, 2, 1, 1, 2, 8, 1, 5, 3, 1, 4, 2, 1, 3, 1, 3, 3\}.$$

Expanding the intermediate convergent

$$\{0, 3, 8, 2, 1, 3, 3, 8, 1, 4, 2, 1, 1\}$$

one obtains the fraction

$$\frac{74122}{231231}.$$

There is another way of thinking about attacks on this system that is more clearly connected to lattices. Consider the 2 by 2 matrix formed with the public key:

$$\begin{pmatrix} 1 & h \\ 0 & q \end{pmatrix}.$$

The linear combinations of these rows with integer coefficients generate a two dimensional lattice L, with determinant q. As $h \equiv f^{-1}g \pmod{q}$ it follows that $fh = g + kq$ for some integer k. The linear combination $f(1, h) - k(0, q) = (f, g)$ will thus be contained in L. This vector has norm

$$|(f, g)| = \mathcal{O}(\sqrt{q}).$$

On the other hand, as the determinant equals q, the expected length of the shortest vector in L is $\approx 0.564\sqrt{q}$. The vector (f, g) has length somewhat larger than this, but it is

still small enough that (f, g) will, with high probability, be the smallest vector in L. An efficient method of searching for small vectors in a two dimensional lattice, like the method of Gauss, will certainly locate (f, g).

In the numerical example of the $N = 1$ case of NTRU we had as the public key the pair $q = 122430513841$ and $h = 39245579300$. The secret key was $f = 231231$ and $g = 195698$, where $f^{-1}g \equiv h \pmod{q}$. If Eve applies LLL to the matrix

$$\begin{pmatrix} 1 & 39245579300 \\ 0 & 122430513841 \end{pmatrix}$$

she recovers the short vector $(-231231, -195698)$ which does in fact reveal f and g.