RSA Background Theory and Algorithms

Scot Anderson, Ph.D.

March 8, 2009

1 Introductions

This lecture gives an introduction to the theory used in the RSA algorithm. We assume a knowlege of prime number concepts and factoring in general, however we do not assume a background in abstract algebra.

2 Background Theory

Notation 1 *The following notation will be helpful if you have not seen it before:*

- $x \in \mathbb{Q}$ *means that the element* x *is a member of the set* $\mathbb{Q}$. *There are a number of standard sets:*
  - $\mathbb{Z}$ *Integers*
  - $\mathbb{Z}_p$ *Integers* mod *an integer* p *(not necessarily prime)*
  - $\mathbb{Q}$ *Rationals*
  - $\mathbb{R}$ *Reals*
  - $\mathbb{N}$ *Natural numbers also sometimes listed as* $\mathbb{Z}^+$
  - $\mathbb{C}$ *Complex Numbers*

We start with a consideration of inverses to numbers in the rational number systems. Since grade school most of us have been able to tell you the multiplicative inverse of a $x \in \mathbb{Q}$. We can write $x = \frac{a}{b}$, where $a, b \neq 0$ and posit that $x^{-1} = \frac{b}{a}$. Proving this for any number x is one of the simplest proofs we present.

Theorem 2 *Let* $x \in \mathbb{Q}$ *where* $x \neq 0$. *Then the multiplicative inverse is of* x *is given by* $x^{-1} = \frac{1}{x}$.

Proof. Let $a, b \in \mathbb{Z}$, and $a, b \neq 0$, then we can write $x = \frac{a}{b}$ for some a, b. Now $\frac{1}{x} = \frac{b}{a}$ is defined and we have.

$$x \cdot x^{-1} = \frac{a}{b} \cdot \frac{b}{a} \quad (1)$$
$$= 1 \quad (2)$$

Hence we have the multiplicative inverse of $x = \frac{1}{x}$ for rational numbers. ■

The above will work easily for real numbers as well, but what if we want to work with a limited set of numbers such as the the set of integers. We can see quickly that multiplicative inverses do not exist for all integers. Given 2, find a multiplicative inverse in the integers. You might guess $\frac{1}{2}$, but $\frac{1}{2} \notin \mathbb{Z}$ so you can't use it. It should be clear without further discussion that only 1 and -1 have multiplicative inverses in $\mathbb{Z}$ and that they are 1 and -1 respectively.

Suppose we restrict the set of numbers to

$$\mathbb{Z}_p = \{0, 1, ..., p-1\} \quad (3)$$

by redefining addition and multiplication to be modulus p. That is consider a number system that has elements $0, ..., p-1$ and defines $\times$ as

$$a \times_p b \equiv (a \times b) \bmod p \quad (4)$$

and addition as

$$a +_p b \equiv (a + b) \bmod p. \tag{5}$$

Normally we don't use the subscripts to define + and ×. It is understood that addition and subtraction in $\mathbb{Z}_p$ is $\bmod p$.

But what about inverses for addition and multiplication.

Example 3 *First consider the additive inverse of $5 \in \mathbb{Z}_7$. We want*

$$5 + x = 0$$

If we try a couple of numbers from $\mathbb{Z}_7$, we can find the answer:

$$\begin{aligned}(5+1) \bmod 7 &= 6\\(5+2) \bmod 7 &= 0\end{aligned}$$

From this we can easily theorize the following.

Theorem 4 *For any $a \in \mathbb{Z}_p$ the additive inverse of a is given by $b = p - a$.*

Proof. Let $a \in \mathbb{Z}_p$ and $b = p - a$. Since $0 \le a < p$, we have that $p - a > 0$ and $p - a \le p$. If $p - a = p$ we will explicitely perform the subtraction modulus p and we obtain a value $b \in \mathbb{Z}_p$.

$$\begin{aligned}(a+b) \bmod p &= (a + p - a) \bmod p\\&= p \bmod p\\&= 0\end{aligned}$$

Hence we have that b is the additive inverse a for any $a \in \mathbb{Z}_p$. ■

(Ok, take a breath! Maybe that was a little more involved than you thought it would be but it is still quite simple.)

What about multiplicative inverses? Well, these are not quite so simple. Consider the following theorem:

Theorem 5 *For some $p \in \mathbb{Z}^+$ and some $a \in \mathbb{Z}_p$, a^{-1} does not exist.*

Proof. Let $p = 6$ and $a = 2$.

$$\begin{aligned}(2 \times 0) \bmod 6 &= 0\\(2 \times 1) \bmod 6 &= 2\\(2 \times 2) \bmod 6 &= 4\\(2 \times 3) \bmod 6 &= 0\\(2 \times 4) \bmod 6 &= 2\\(2 \times 5) \bmod 6 &= 4\end{aligned}$$

Here we have exausted all possiblilites for inverses of 2 in $\mathbb{Z}_6$ and $\forall x \in \mathbb{Z}_6$ we have that $(x \times 2) \bmod 6 \neq 1$. Thus we have that in some $\mathbb{Z}_p$ there exists elements that do not have multiplicative inverses. ■

That is surely a blow to doing certain types of arithmatic in just any $\mathbb{Z}_p$ but certainly there is some constraint that we can put on p that will guarantee inverses in $\mathbb{Z}_p$. Here we need Fermat's Little Theorm. First let's consider a little notation.

Notation 6 *If two numbers b and c have the property that their difference $b - c$ is integrally divisible by a number m (i.e. $(b - c)/m$ is an integer), then b and c are said to be "congruent modulo m. " The number m is called the modulus, and the statement "b is congruent to c (modulo m)" is written mathematically as*

$$b \equiv c \pmod m$$

or equivalenty as

$$b - c = m \cdot t$$

for some $t \in \mathbb{Z}$. Here b is called the ***base****, c is called the* ***residue*** *and m is called the modulus. If we require that $c \in \mathbb{Z}_m$, we can also say that*

$$b \bmod m = c$$

As a counter example where $c \notin \mathbb{Z}_m$ consider

$$10 \equiv 4 \pmod 3$$

gives

$$10 - 4 = 3t$$

for some integer t. Clearly $t = 2$ works

$$\begin{aligned} 10 - 4 &= 3(2) \\ 6 &= 6 \end{aligned}$$

However $10 \bmod 3 = 1$ and $1 \neq 4$. So be careful with your intuition.

Theorem 7 (Fermat's Little Theorem) *If p is prime and a is a positive integer not divisible by p, then*

$$a^{p-1} \equiv 1 (\bmod p) \tag{6}$$

Proof. Consider the set of possitive integers less than p: $W = \{1, 2, ..., p-1\}$. Multiply each element by a, modulo p, to get the set

$$X = \{a \bmod p, 2a \bmod p, ..., (p-1)\, a \bmod p\} \tag{7}$$

None of the elements of X is equal to zero because p does not divide a. Further no two elements of X are equal. To see this fact we will do a proof by contradiction (remember set theory or discreet math/structures?). *Assume* that two elements are equal.

$$ja \bmod p = ka \bmod p \tag{8}$$

or equivalently

$$ja \equiv ka \pmod p \tag{9}$$

Because a is relatively prime to p we can eliminate it from (9). This gives the contradition since $j \bmod p = j$ and $k \bmod p = k$ we have that $j = k$. Hence *no two of the $p-1$ elements of X are equal and therefore X* is identical to our first set W in some order. Multiplying both sets and taking the result mod p yields

$$\begin{aligned} a \times 2a \times ... \times (p-1)\, a &\equiv [1 \times 2 \times ... \times (p-1)] \pmod p \\ a^{p-1}\, (p-1)! &\equiv (p-1)! \pmod p \end{aligned} \tag{10}$$

Since $(p-1)!$ is relatively prime to p we can eliminate it giving the result

$$a^{p-1} \equiv 1 \pmod p \tag{11}$$

■

An alternative form to Theorem 7 states: If p is prime and a is a positive integer, then

$$a^p = a \pmod p \tag{12}$$

This result looks suspiciously like something that we might use for encryption/decryption. After all if you set $x < p$ and take a^x, all I need to know is $y = p/x$ and take the result $(a^x)^{p/x} = a^p \pmod p = a$. But taking a closer look, this doen't seem very secure. If you know x, (you must know p) then you can know y and vice versa. To overcome this problem we need to understand Euler's Totient function and then Euler's Theorem.

Definition 8 *Euler's Totient function $\phi(n)$ returns the number of positive integers less than n and relatively prime to n. By convention, $\phi(1) = 1$.*

Example 9 *Determine $\phi(37)$. Since 37 is prime (and we count the number 1) there are 36 positive numbers less than 37 that are relatively prime to 37.*

Now consider $\phi(35)$. The numbers relatively prime are

$$1, 2, 3, 4, 6, 8, 9, 11, 12, 13, 16, 17, 18, 19, 22, 23, 24, 26, 27, 29, 31, 32, 33, 34$$

There are 24 numbers in the list so $\phi(35) = 24$.

It should be clear by now that for a prime number p

$$\phi(p) = p - 1 \tag{13}$$

Now suppose that we have two prime numbers p and q such that $p \neq q$. Then we can show that for $n = pq$

$$\begin{aligned} \phi(n) &= \phi(p)\,\phi(q) && (14)\\ &= (p-1)(q-1) && (15) \end{aligned}$$

Consider the set

$$\{1, ..., pq - 1\}$$

. The integers in this set not relatively prime to n are the set

$$\{p, 2p, ..., (q-1)\,p\} \cup \{q, 2q, ..., (p-1)\,q\} \tag{16}$$

Hence we have

$$\begin{aligned} \phi(n) &= (pq-1) - [(q-1) + (p-1)]\\ &= pq - (p+q) + 1\\ &= (p-1)(q-1)\\ &= \phi(p)\,\phi(q) \end{aligned}$$

Now we are ready to use this fact.

Theorem 10 (Euler's Theorem) *For every a and n that are relatively prime:*

$$a^{\phi(n)} = 1 \pmod{n} \tag{17}$$

Proof. Equation (17) is true if n is prime, because in that case $\phi(n) = (n-1)$ and Theorem 7 holds. However, it also holds for any integer n. Consider the set of integers

$$R = \left\{x_1, x_2, ..., x_{\phi(n)}\right\} \tag{18}$$

That is, each element x_i of R is a unique positive integer less than n with $\gcd(x_i, n) = 1$. Now multiply each element by a, modulo n:

$$S = \left\{(ax_1 \bmod n), ..., \left(ax_{\phi(n)} \bmod n\right)\right\} \tag{19}$$

The set S is a permutation of R, by the following line of reasoning:

1. Because a is relatively prime to n and x_i is relatively prime to n, $a\, x_i$ must also be relatively prime to n. Thus, all member of S are integers that are less than n and relatively prime to n.

2. There are no duplicates in S. (similar to Theorem 7)

Therefore

$$\begin{aligned}\prod_{i=1}^{\phi(n)} (ax_i \bmod n) &= \prod_{i=1}^{\phi(n)} x_i \\ \prod_{i=1}^{\phi(n)} (ax_i) &\equiv \prod_{i=1}^{\phi(n)} x_i \pmod n \\ a^{\phi(n)} \times \prod_{i=1}^{\phi(n)} (x_i) &\equiv \prod_{i=1}^{\phi(n)} x_i \pmod n \\ a^{\phi(n)} &\equiv 1 \pmod n \qquad (20)\end{aligned}$$

However, if a and n are relatively prime then a^k and n are relatively prime and we have

$$\left(a^k\right)^{\phi(n)} \equiv 1 \pmod n \qquad (21)$$

■

Alternatively we can state the result as

$$a^{\phi(n)+1} = a \pmod n \qquad (22)$$

where a does not have to be relatively prime to n.

Notice that this is looks very similar to Theorem 7! As we will see, it also provides us with what we need to build the RSA algorithm.

3 The RSA Algorithm

First we have a Plaintext block $M < n$, that is the block size b must satisfy $b \leq \log_2 n$. In practice, the block size b is i bits, where $2^i < n \leq 2^{i+1}$. The cipher block C is given by

$$C = M^e \bmod n \qquad (23)$$

and decryption by

$$M = C^d \bmod n = (M^e)^d = M^{ed} \bmod n \qquad (24)$$

Both the sender and receiver must know n. The sender knows the value of e, and only the receiver knows the value of d.

$$\begin{aligned} PU_{key} &= \{e, n\} \qquad (25) \\ PR_{key} &= \{d, n\} \qquad (26)\end{aligned}$$

For this algorithm to be satisfactory for PKI we must meet the following requirements:

1. It is possible to find values of e, d, n such that $M^{ed} \bmod n = M$ for all $M < n$.
2. It is relatively easy to calculate $M^e \bmod n$ and $C^d \bmod n$ for all values of $M < n$.
3. It is **infeasible** to determine d given e and n.

Requirement (2) can be easily satisfied using ordinary arithmetic modulo n. (3) relies on the difficulty in factoring large primes as we will see. That leaves Relationship (1).

We need to find a relationship of the form

$$M^{ed} \pmod n = M \quad \text{or} \quad M^{ed} \equiv M \pmod n \qquad (27)$$

If

$$\begin{aligned} ed - 1 &= k\ \phi(n) \quad (28)\\ \Longleftrightarrow ed &= k\phi(n) + 1 \quad (29)\\ \Longleftrightarrow ed &\equiv 1 \pmod{\phi(n)} \quad (30)\end{aligned}$$

Then By Theorem 10 (see Equation 21)

$$M^{ed-1} \equiv 1 \pmod{n} \quad (31)$$

is known to hold. The alternate form of Theorem 10 gives.

$$M^{ed} \equiv M \pmod{n}$$

From Equation 30 we must have that e and d are inverses of each other modulo $\phi(n)$. That is,

$$\begin{aligned} ed &\equiv 1 \bmod \phi(n) \quad (32)\\ d &\equiv e^{-1} \bmod \phi(n) \quad (33)\end{aligned}$$

This gives the method of calculating d or e. Also note that, according to the rules of modular arithmetic, this is true only if d (and therefore e) is relatively prime to $\phi(n)$. Equivalently, $\gcd(\phi(n), d) = 1$. We can check the gcd and find the inverse using Euclid's Extended algorithm.

Table 1 gives the values needed for the RSA scheme. Notice that $\phi(n)$ is never divulged in the public or private keys. Generating a public key from the private (or vice versa) requires knowledge of $\phi(n)$. No problem you say, I'll just factor n. But here is the rub: factoring large primes is difficult and thus requirement (3) from above is met.

p, q, two prime numbers	(private, chosen)
$n = pq$	(public, calculated)
e, with $\gcd(\phi(n), e) = 1$; $1 < e < \phi(n)$	(public, calculated)
$d \equiv e^{-1} \pmod{\phi(n)}$	(private, calculated)

Table 1: RSA Values

This leads to the key generation algorithm given in Table 2.

Key Generation		
Select p, q		p, q both prime $p \neq q$
Calculate $n = p \times q$		
Calculate $\phi(n) = (p-1)(q-1)$		
Select integer e		$\gcd(\phi(n), e) = 1$; $1 < e < \phi(n)$
Calculate d		$d = e^{-1} \bmod (\phi(n))$
Return Public Key		$PU_{key} = \{e, n\}$
Return Private Key		$PR_{key} = \{d, n\}$

Table 2: Key Generation Algorithm

Example 11 (Simple RSA)

1. *Let* $p = 17$ *and* $q = 11$.
2. *Then* $n = 187$
3. *and* $\phi(n) = 160$.
4. *Select* e *such that* $\gcd(e, \phi(n)) = 1$ *(relatively prime) and* $e < \phi(n)$. *Let* $e = 7$.
5. *Determine* $d = 23$ *using Euclid's extended algorithm.*

6. Return the public and private keys $PU_{key} = \{7, 187\}$, and $PR_{key} = \{23, 187\}$.

Suppose we have $M = 88$. Encrypting this with PU_{key} and exploiting the properties of modular arithmetic gives:

$$\begin{aligned}
88^7 &= \left(88^4 \bmod 187\right)\left(88^2 \bmod 187\right)\left(88^1 \bmod 187\right) \\
88^1 \bmod 187 &= 88 \\
88^2 \bmod 187 &= 77 \\
88^4 \bmod 187 &= 77^2 \bmod 187 = 132 \\
88^7 \bmod 187 &= 132 \times 77 \times 88 \bmod 187 = 11
\end{aligned}$$

So $C = 88^7 \bmod 187 = 11$.

4 Homework

1. Using the example above, decrypt $C = 11$.
2. Program the RSA algorithm in jave to generate key pairs and encrypt/decrypt 32-bit blocks of data.