RSA CRYPTO-SYSTEM

Preliminaries.

Theorem (Fermat's Little Theorem)**.** *If p is prime, and a an integer, then*

$$a^p \equiv a \,(mod\, p)$$

Proof. (Use induction on a.)

If $a = 1$, then

$$a^p = 1^p = 1 \equiv a \,(\text{mod}\, p).$$

Suppose $k^p \equiv k \,(\text{mod}\, p)$. (Inductive hypothesis)

Then

$$(k+1)^p = k^p + \binom{p}{1} k^{p-1} + \ldots + \binom{p}{p-1} k + 1.$$

Now if p is prime and $0 < t < p$, then $\binom{p}{t}$ has a factor p. So

$$(k+1)^p \equiv k^p + 1 \equiv k + 1 \,(\text{mod}\, p).$$

□

Illustration. For $a = 3$ and $p = 7$,

$$\begin{aligned} 3^7 &= 3^{2\cdot 3+1} \\ &= \left(3^2\right)^3 \cdot 3^1 \\ &= 9^3 \cdot 3 \;\equiv\; 2^3 \cdot 3 \\ &= 8 \cdot 3 \;\equiv\; 1 \cdot 3 \;\equiv\; 3 \,(\text{mod}\, 7). \end{aligned}$$

Consequences. In the case that p is prime and $p \nmid a$,

- $a^{p-1} \equiv 1 \,(\text{mod}\, p)$ (by Theorem 2, Section 2.6)
- $a^{-1} \equiv a^{p-2} \,(\text{mod}\, p)$
- If $\gcd(a, p) = 1$ and $a^{p-1} \not\equiv 1 \,(\text{mod}\, p)$, then p ***is not*** prime.

Illustration. If $a = 2$ and $n = 91$, first observe that $2^{12} \equiv 1 \,(\text{mod}\, 91)$.

$$\begin{aligned} 2^{12} &= 4096 \;=\; 4000 + 96 \;=\; 40(100) + 96 \\ &\equiv 40(9) + 5 \;=\; 4 * (90) + 5 \;\equiv\; 4(-1) + 5 \\ &\equiv 1 \,(\text{mod}\, 91). \end{aligned}$$

So now,

$$\begin{aligned} 2^{91} &= 2^{7(12)+7} \\ &= (2^{12})^7 \cdot 2^7 \\ &\equiv 1^7 \cdot 128 \;\equiv\; 37 \,(\text{mod}\, 91). \end{aligned}$$

The Euler φ-function.

Definition (Euler φ-function)**.** *For any $n \in \mathbb{N}$, the Euler φ-function, also known as Euler's totient function, is defined to be the number of $m \in \mathbb{N}$ satisfying $m \leq n$ and $\gcd(m, n) = 1$.*

Properties of $\varphi(n)$.

- If p is prime, then $\varphi(p) = p - 1$.
- If p is prime, then $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha(1 - 1/p)$.

  Count Them. First, the multiples of p up to and including p^α are

  $$p, 2p, 3p, \ldots, (p-1)p, p^2, p^2 + p, p^2 + 2p, \ldots, p^\alpha$$

  and there are $p^{\alpha-1}$ of them.
  So,

  $$\begin{aligned} \varphi(p^\alpha) &= (\text{number of } n \leq p^\alpha) \\ &\quad -(\text{number of multiples of } p \leq p^\alpha) \\ &= p^\alpha - p^{\alpha-1} \end{aligned}$$

  $\square$

- If $\gcd(a, b) = 1$, then $\varphi(ab) = \varphi(a)\varphi(b)$.
- In general, for any integer $n > 1$, if the distinct prime numbers dividing n are $p_1, p_1, \ldots, p_k$—that is, $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$—then $\varphi(n) = n(1 - 1/p_1)(1 - 1/p_2) \cdots (1 - 1/p_k)$.

Theorem (Euler)**.** *If $\gcd(a, n) = 1$, then $a^{\varphi(n)} \equiv 1 \,(\mathit{mod}\, n)$.*

Observe.

- If n is prime, then $\varphi(n) = n - 1$ and we have

  $$a^{\varphi(n)} = a^{n-1} \equiv 1 \pmod{n}$$

  which is Fermat's Little Theorem.
- In our earlier computation involving $2^{91} \bmod 91$, we can see $a^{\varphi(n)} \equiv 1 \pmod{n}$ occurring as a subproblem by observing
  - $91 = 7 \cdot 13$ so

    $$\varphi(91) = \varphi(7)\varphi(13) = 6 \cdot 12 = 72, \qquad \text{and}$$

  - $2^{12} \equiv 1 \pmod{91}$.

  So $2^{\varphi(91)} = 2^{12 \cdot 6} = (2^{12})^6 \equiv 1^6 \equiv 1 \pmod{91}$.

Description of RSA.

RSA encryption starts with a numerical plaintext P and converts it into a numerical ciphertext C by

$$C = P^e \bmod n.$$

Upon receipt, C is decrypted in a similar manner using the same modulus n and a different exponent d. That is

$$P = C^d \bmod n.$$

The values of n, e, and d are constructed as follows.

Key Generation.

- Randomly select two primes p & q.

  To keep the factoring of n from defaulting to something that might be "easy", p & q should be roughly the same size. In real world implementations, they are about 150 digits long. This corresponds to "1024-bit encryption", the 1024 bits referring to the size of n.
- Compute $n = pq$ and $\varphi(n) = (p-1)(q-1)$.
- Select a random integer e with $1 < e < \varphi(n)$ and $\gcd(e, \varphi(n)) = 1$.
- Compute the unique integer d, $1 < d < \varphi(n)$ such that

$$ed \equiv 1 \pmod{\varphi(n)}.$$

The Public Key and Encryption.

- Make public n and e.
- Encipher plaintext P by

$$C = P^e \bmod n.$$

The Private Key and Decryption.

- Keep private p, q, $\varphi(n)$, and d
- Decipher ciphertext C by

$$P = C^d \bmod n.$$

A Small Example.

Select two primes:

$p = 11$ and $q = 13$.
So $n = pq = 143$.
Now $\varphi(n) = (p-1)(q-1) = 10 \cdot 12 = 120$.

Choose e coprime with $\varphi(n)$:

Choose $e = 37$.

Find d:

We need $e \cdot d \equiv 1 \pmod{120}$.
Compute $37^{-1} \bmod 120$.
Now solve $37d \equiv 1 \pmod{120}$; that is, solve $37d + 120q = 1$ for d.

$$\begin{aligned} 120 &= 3 \cdot 37 + 9 \\ 37 &= 4 \cdot 9 + 1, \end{aligned}$$

so

$$\begin{aligned} 1 &= 37 - 4 \cdot 9 \\ &= 37 - 4(120 - 3 \cdot 37) \\ &= 13 \cdot 37 - 4 \cdot 120. \end{aligned}$$

Therefore $d = 13$.

Alternatively, we could compute $37^{\varphi(120)-1} \mod 120$:

$$\begin{aligned} \varphi(120) &= \varphi(12 \cdot 10) = \varphi(2^2 \cdot 3 \cdot 2 \cdot 5) = \varphi(2^3 \cdot 3 \cdot 5) \\ &= \varphi(2^3)\varphi(3)\varphi(5) = (2^3 - 2^2)(3-1)(5-1) = (8-4)(2)(4) = 4 \cdot 2 \cdot 4 = 32. \end{aligned}$$

So $\varphi(120) = 32$, and $\varphi(120) - 1 = 31$.

Now, reducing by mod 120,

$$\begin{aligned} 37^{\varphi(120)-1} &= 37^{31} = 37^{1+30} = 37^{1+2\cdot 15} = 37 \cdot (37^2)^{15} \\ &\equiv 37 \cdot 49^{15} = 37 \cdot 49^{1+2\cdot 7} = 37 \cdot 49 \cdot (49^2)^7 \\ &\equiv 37 \cdot 49 \cdot 1^7 = 31 \cdot 49 \\ &\equiv 13 \pmod{120}. \end{aligned}$$

Note: With this base (120), $e = 19, 29$, and 31 are all their own inverses! So these would be bad choices for e.

The Public Key:

$n = 143$, $e = 37$

The Private Key:

$n = 143$, $d = 13$

Encipher a Message: Let's encipher "Hi."

- Begin by converting our plaintext into a number or series of numbers. Using the ASCII values, we find that

  H $\longleftrightarrow$ 72
  i $\longleftrightarrow$ 105
  . $\longleftrightarrow$ 46

- Raise each to the power $e = 37$ and reduce mod 143.

$$\begin{aligned} 72^{37} &= 72^{1+2\cdot 18} = 72 \cdot 72^{2\cdot 18} = 72 \cdot (72^2)^{18} \\ &\equiv 72 \cdot 36^{18} = 72 \cdot 36^{2\cdot 9} = 72 \cdot (36^2)^9 \\ &\equiv 72 \cdot 9^9 = 72 \cdot 9^{3\cdot 3} = 72 \cdot (9^3)^3 \\ &\equiv 72 \cdot 14^3 \equiv 72 \cdot 27 \\ &\equiv 85. \end{aligned}$$

85 is the enciphered letter "H".

$$\begin{aligned} 105^{37} &= 105^{1+2\cdot 18} = 105 \cdot 105^{2\cdot 18} = 105 \cdot (105^2)^{18} \\ &\equiv 105 \cdot 14^{18} = 105 \cdot 14^{2\cdot 9} = 105 \cdot (14^2)^9 \\ &\equiv 105 \cdot 53^9 = 105 \cdot 53^{3\cdot 3} = 105 \cdot (53^3)^3 \\ &\equiv 105 \cdot 14^3 \equiv 105 \cdot 27 \\ &\equiv 118. \end{aligned}$$

118 is the enciphered letter "i".

$$\begin{aligned} 46^{37} &= 46^{1+2\cdot 18} = 46 \cdot 46^{2\cdot 18} = 46 \cdot (46^2)^{18} \\ &\equiv 46 \cdot 114^{18} = 46 \cdot 114^{2\cdot 9} = 46 \cdot (114^2)^9 \\ &\equiv 46 \cdot 126^9 = 46 \cdot 126^{3\cdot 3} = 46 \cdot (126^3)^3 \\ &\equiv 46 \cdot 92^3 \equiv 46 \cdot 53 \\ &\equiv 7. \end{aligned}$$

7 is the enciphered letter ".".

The ciphertext C_t is 85 105 7.

Decipher a Message: Let's decipher the ciphertext we just received, $C_t = 851057$.

- Raise each number in the ciphertext to the power $d = 13$ and reduce mod 143. Then look up the letter in the ASCII table.

$$\begin{aligned}85^{13} &= 85^{1+2\cdot 2\cdot 3} = 85 \cdot 85^{2\cdot 2\cdot 3} = 85 \cdot ((85^2)^2)^3 \\ &\equiv 85 \cdot (75^2)^3 \equiv 85 \cdot 48^3 \equiv 85 \cdot 53 \\ &\equiv 72.\end{aligned}$$

72 is the ASCII value of "H".

$$\begin{aligned}118^{13} &= 118^{1+2\cdot 2\cdot 3} = 118 \cdot 118^{2\cdot 2\cdot 3} = 118 \cdot ((118^2)^2)^3 \\ &\equiv 118 \cdot (53^2)^3 \equiv 118 \cdot 92^3 \equiv 118 \cdot 53 \\ &\equiv 105.\end{aligned}$$

105 is the ASCII value of "i".

$$\begin{aligned}7^{13} &= 7^{1+2\cdot 2\cdot 3} = 7 \cdot 7^{2\cdot 2\cdot 3} = 7 \cdot ((7^2)^2)^3 \\ &\equiv 7 \cdot (49^2)^3 \equiv 7 \cdot 113^3 \equiv 7 \cdot 27 \\ &\equiv 46.\end{aligned}$$

46 is the ASCII value of ".".

The plaintext P_t is "Hi.".

Why It Works. In order to decode ciphertext C into the original plaintext P, we need

$$P = C^d = (P^e \bmod n)^d = P^{e \cdot d} \bmod n.$$

The requirement that $ed \equiv 1 \pmod{\varphi(n)}$, means that ed can be written as

$$ed = 1 + k \cdot \varphi(n)$$

for some integer k. Therefore

$$\begin{aligned} P^{d \cdot e} &= P^{1+k\varphi(n)} \\ &= P^1 \cdot P^{\varphi(n) \cdot k} \\ &= P \cdot \left(P^{\varphi(n)}\right)^k \\ &\equiv P \cdot 1 \equiv P \pmod{n}. \end{aligned}$$

Protocols.

The Context.

- Bob creates an RSA crypto-system with public key (n_B, e_B) and private key (n_B, d_B).
- Alice creates an RSA crypto-system with public key (n_A, e_A) and private key (n_A, d_A)

Implementations.

Alice sends a message P to Bob:

- Alice wants her message to Bob to be read only by him.
  1. Alice encrypts P into C using Bob's public key (n_B, e_B) and sends C to Bob.
  2. Bob uses his private key (n_B, d_B) to decipher C back into P.

Alice sends a signed message P to Bob:

- Alice wants her message to Bob to be read only by him.
- Bob wants assurance that it was Alice who sent him the message, and that Alice cannot deny that she sent it.
  1. Alice signs P by encrypting it into S using her private key (n_A, d_A), then she enciphers S into C using Bob's public key (n_B, e_B) and sends C to Bob.
  2. Bob deciphers C into S using his private key (n_B, d_B), then he "unsigns" S into P using Alice's public key (n_A, e_A). Since only Alice had the inverse of her decryption, the message had to come from Alice.

Practical Matters.

The implementation has several practical matters.

Handling Long Messages:

If the message is long, break it up into numbers P_t where

$$0 < P_t < n$$

and perform RSA on each P_t.

Randomly Selecting Primes p and q:

Security requires that p and q not be guessed easily, so they should have no special characteristics other than being prime. This is achieved using probabilistic methods.

(1) "Randomly" generate a string of digits of the appropriate length (ending in an odd digit other than 5).
(In a binary implementation, just require that the units bit be 1.)
This becomes a candidate for p (or q).

(2) Run a probabilistic test for primality k times. If it passes k times then the probability that it is prime is

$$1 - \frac{1}{b^k}$$

where b depends on the particular test. (E.g. $b = 2$ for the Solovay-Strassen test, $b = 4$ for the Miller-Rabin test.)

Preliminary Checks:

Before fixing values for p, q, and e, a good implementation will involve a computation of d to see that the choices yield no unfortunate surprises.

- If $p - q$ is small, then $p \approx \sqrt{n}$, in which case n could be factored efficiently merely by trial division of all odd numbers close to $\sqrt{n}$.
- A good implementation will involve a check that $d \neq e$. This is rare that $d = e$, but it is not impossible.
(If $p = 11$ and $q = 13$, then if $e = 19$, 29, or 31, then $d = e$.)

Raising Powers:

Because the size of P^e and C^d increase exponentially in their computation, *it is vital that the "square and multiply" algorithm be used and that modulo-n reduction be performed at each step.*

Square and Multiply

Compute $x^b \bmod n$ where $b = (b_t \dots b_1 b_0)_2$.

`Input:` x and b
$z := 1$
`for` $i := t$ `down to` 0 `do`
  $z := z^2 \bmod n$
  `if` $b_i = 1$ `then` $z = (z \cdot x) \bmod n$.

Example. Compute x^{11} : Note that

$$(11)_{10} = (1011)_2 = (b_3 b_2 b_1 b_0)_2.$$

$z := 1$
  $z := z \cdot x$ (I.e. $z = x$. This handles $b_3 = 1$)
 $z := z^2$ (I.e. $z = x^2$. This handles $b_2 = 0$)
 $z := z^2$ (I.e. $z = x^4$)
  $z := z \cdot x$ (I.e. $z = x^5$. This handles $b_1 = 1$)
 $z := z^2$ (I.e. $z = x^{10}$)
  $z := z \cdot x$ (I.e. $z = x^{11}$. This handles $b_0 = 1$)

Fair Warning:

Regardless of your choice of p, q, and e, there will always be plaintexts P for which $P^e \equiv P \pmod{n}$. (For example, $P = 0$, 1, and

$n-1$.) In fact, the number of such "unconcealed messages" is exactly

$$(1 + \gcd(e-1, p-1)) \cdot (1 + \gcd(e-1, q-1))$$

and since $e-1$, $p-1$, and $q-1$ are all even, there will always be at least 9 unconcealed messages.

Fortunately, if p and q are prime, and if e is randomly selected, then the proportion of messages left unconcealed by RSA is generally negligibly small.

Why $\varphi(n)$ Must Be Kept Secret.

If both n (i.e. $p \cdot q$) and $\varphi(n)$ (i.e. $(p-1)(q-1)$) are known, then the values of p and q can be computed using the following technique.

$$(p-1)(q-1) = pq - p - q + 1$$

so

$$\begin{aligned} \varphi(n) - n - 1 &= (pq - p - q + 1) - pq - 1 \\ &= -(p+q) \end{aligned}$$

Also,

$$x^2 - (a+b)x + ab = 0$$

has solutions a and b.

Now use the quadratic formula to find the zeros of

$$x^2 + \underbrace{(\varphi(n) - n - 1)}_{-(p+q)} x + \underbrace{n}_{pq} = 0$$

Example. Suppose $n = 253$ and $\varphi(n) = 220$.

Solve

$$\begin{aligned} x^2 + (220 - 253 - 1)x + 253 &= \\ x^2 - 34x + 253 &= 0 \end{aligned}$$

$$\begin{aligned} x &= \frac{-(-34) \pm \sqrt{(-34)^2 - 4 \cdot 253}}{2} \\ &= \frac{34 \pm \sqrt{1156 - 1012}}{2} \\ &= \frac{34 \pm \sqrt{144}}{2} \\ &= \frac{34 \pm 12}{2} \\ &= \frac{46}{2} \text{ or } \frac{22}{2} = 23 \text{ or } 11 \end{aligned}$$

Notice: 11 and 23 are primes, with

$$11 \cdot 23 = 253 = n$$

and

$$(11-1)(23-1) = 10 \cdot 22 = 220 = \varphi(n)$$