Message Passing, RSA Encryption

Robert Y. Lewis

CS 0220 2024

March 11, 2024

Overview

Sending messages

Motivations:

- Why do we send each other messages? Communication is a pretty human activity. Coordination is a practical application.
- Why would we want them to be secret? Competition. Gossiping. Surprise. Private information.
- Why might the message need to be encrypted? Message can be intercepted, stolen/broadcast by a third party, accidentally revealed like by being left on screen connected to projector. Communication channel is open.

Encoding messages

We’ll assume all messages are fixed-length bit strings.

Is that sufficiently general? Can we encode arbitrary messages this way?

The Alphabet

00000	space	01000	H	10000	P	11000	X
00001	A	01001	I	10001	Q	11001	Y
00010	B	01010	J	10010	R	11010	Z
00011	C	01011	K	10011	S	11011	.
00100	D	01100	L	10100	T	11100	!
00101	E	01101	M	10101	U	11101	?
00110	F	01110	N	10110	V	11110	,
00111	G	01111	O	10111	W	11111	@

One-time pad: Encryption

Alice and Bob share 60 random bits (the “one-time pad”) in advance:

10110 00011 00000 00110 10010 00010
00000 11000 00110 01001 11111 11110

Alice: Wants to send a private message to Bob. She turns it into a sequence of 60 bits. She then computes the bitwise “xor” of her message and the one-time pad and transmits it:

00101 10111 00101 00001 11101 10001
00001 01101 10100 11100 01000 11110

One-time pad: Decryption

Bob: Wants to read Alice's message.

00101 10111 00101 00001 11101 10001
00001 01101 10100 11100 01000 11110

How can he recover it? Bitwise "xor" with the one-time pad will undo the encryption operation.

encrypted line 1:	00101 10111 00101 00001 11101 10001
pad line 1:	10110 00011 00000 00110 10010 00010
xor line 1:	10011 10100 00101 00111 01111 10011
text line 1:	S T E G O S
encrypted line 2:	00000 11000 00110 01001 11111 11110
pad line 2:	00001 01101 10100 11100 01000 11110
xor line 2:	00001 10101 10010 10101 10111 00000
text line 2:	A U R U S _

One-time pad: Cracking

Eve: Sees the encrypted message and wants to understand it. She doesn't have the one-time pad.

The encrypted message gives *no information* about the unencrypted message. All possible messages are *equally likely*.

Although, if one-time pad is reused, information is leaked.

Plus, The one-time pad is essential to the one-time pad scheme. How can Alice and Bob agree on the one-time pad if Eve is listening?

Definition

Definition: Integers that have no prime factor in common are called *relatively prime*.

In other words, they have no common divisor greater than 1. Or $\gcd(a, b) = 1$. Also, called “co-prime”.

Example: 9 and 14 are relatively prime. Neither are prime. But, they have no prime factors in common. Here, “relative” refers to “relative to each other”, not “kinda”.

What’s *not* relatively prime to 17? 34, sure. But, in general? Precisely the multiples of 17. True of any prime *p*.

Multiplicative inverse

Lemma: Let n be a positive integer. If k is relatively prime to n, then there exists an integer k^{-1} such that:

$$k \cdot k^{-1} \equiv 1 \pmod{n}.$$

Proof: Since n and k are relatively prime, gcdcombo$(n, k) = (s, t, 1)$. t must be the multiplicative inverse of k (mod n):

$$s \cdot n + t \cdot k = 1 \text{ implies } t \cdot k \equiv 1 \pmod{n}.$$

Corollary: Suppose n is a positive integer and k is relatively prime to n. Then,

$$ak \equiv bk \pmod{n} \quad \text{implies} \quad a \equiv b \pmod{n}$$

Proof: Multiply both sides by k^{-1} and simplify.

Relatively prime permutations

Lemma: Suppose n is a positive integer and k is relatively prime to n. Let $k_1, k_2, \cdots, k_r$ be all the integers in the interval $[1, n)$ that are relatively prime to n. Then, the sequence of remainders on division by n of

$$k_1 \cdot k, k_2 \cdot k, \ldots, k_r \cdot k$$

is a permutation of the sequence $k_1, k_2, \cdots, k_r$.

Example: $n = 18, k = 5$.

j	1	2	3	4	5	6	$= r$
k_j	1	5	7	11	13	17	
$k_j \cdot k$	5	25	35	55	65	85	
$k_j \cdot k \bmod n$	5	7	17	1	11	13	

Relatively Prime Permutation Proof (for reference)

Proof: We will show that the remainders in the first sequence are all distinct and are equal to some member of the sequence of k_js. Since the two sequences have the same length, the first must be a permutation of the second. (Kind of a bijection argument.)

If $k \cdot k_j \equiv k \cdot k_{j'} \pmod{n}$, then $k_j \equiv k_{j'} \pmod{n}$ by the Cancelation rule. Thus, the remainders are all distinct.

Next, we show that each remainder in the first sequence equals one of the k_js. By assumption, k_i and k are relatively prime to n, and therefore so is $k_i k$ by the "you can't split a prime" property. So, $\gcd(k \cdot k_i, n) = 1$. But, by the derivation of Euclid's algorithm, $\gcd(k \cdot k_i, n) = \gcd(n, \text{rem}(k \cdot k_i, n))$. Thus, $\text{rem}(k \cdot k_i, n)$ has a GCD of 1 with n, so it's on the list of relatively prime integers to n. QED.

Fermat's little theorem (reminder)

Theorem: Suppose p is prime and k is not a multiple of p. Then:

$$k^{p-1} \equiv 1 \pmod{p}.$$

Since $k^{p-1} \equiv 1 \pmod{p}$ and $k^{p-1} = k \cdot k^{p-2}$, that tells us that k^{p-2} is the multiplicative inverse for k.

Only for prime p!

Remainder reminder: Solving equation mod prime

$$
\begin{aligned}
3x + 9 &\equiv 2 \pmod{11} & & \\
3x &\equiv -7 \pmod{11} & & \text{additive shift} \\
3x &\equiv 4 \pmod{11} & & \text{pre-mod} \\
3^{11-2} \cdot 3x &\equiv 3^{11-2} \cdot 4 \pmod{11} & & \text{multiply both sides} \\
x &\equiv 3^{11-2} \cdot 4 \pmod{11} & & \text{Fermat's little theorem} \\
x &\equiv 4 \times 4 = 16 \pmod{11} & & \text{Maybe some repeated squaring} \\
x &\equiv 5 \pmod{11} & & \text{modding}
\end{aligned}
$$

Double check: $3 \times 5 + 9 = 15 + 9 = 24 = 2 \pmod{11}$.

Key step: $3^{-1} = 4 \pmod{11}$.

Via Fermat's little theorem: $3^{11-2} \bmod 11 = 19683 \bmod 11 = 4$.

But what if we wanted to work mod not-a-prime?

Counting relatively prime numbers

$\phi(n)$: The number of integers in $[0, n)$ that are relatively prime to n.

Examples:

- $\phi(7) = |\{1, 2, 3, 4, 5, 6\}| = 6$.
- $\phi(18) = |\{1, 5, 7, 11, 13, 17\}| = 6$.
- $\phi(20) = |\{1, 3, 7, 9, 11, 13, 17, 19\}| = 8$.
- $\phi(p) = p - 1$ if p is prime. Everybody below p is relatively prime to prime p !

Called Euler's ϕ or totient function.

Euler's Theorem

Theorem: Suppose n is a positive integer and k is relatively prime to n. Then,

$$k^{\phi(n)} \equiv 1 \pmod{n}.$$

Proof: Let $k_1, k_2, \ldots, k_r$ denote all integers relatively prime to n where $k_i \in [0, n)$. So, $\phi(n) = r$.

$$\begin{aligned} &= \text{rem}(k_1 \cdot k, n) \cdot \text{rem}(k_2 \cdot k, n) \cdot \cdots \cdot \text{rem}(k_r \cdot k, n) && \text{rel. prime perm.} \\ k_1 \cdot k_2 \cdot \cdots \cdot k_r &= (k_1 \cdot k) \cdot (k_2 \cdot k) \cdot \cdots \cdot (k_r \cdot k) \pmod{n} && \text{pre-mod} \\ &= k_1 \cdot k_2 \cdot \cdots \cdot k_r \cdot k^r \pmod{n} && \text{regroup} \end{aligned}$$

Applying the Cancellation lemma, the claim is proven. QED.

If we could compute $\phi(n)$, we could use it to compute multiplicative inverses. Can we?

Phi of product of two distinct primes

Lemma: For distinct primes p and q, $\phi(pq) = (p-1)(q-1)$.

Proof: Since p and q are prime, any number that is not relatively prime to pq must be a multiple of p or a multiple of q. Among the pq numbers in $[0, pq)$, there are precisely q multiples of p and p multiples of q. Since p and q are relatively prime, the only number in $[0, pq)$ that is a multiple of both p and q is 0. Hence, there are $p+q-1$ numbers in $[0, pq)$ that are not relatively prime to pq. So, $\phi(pq) = pq - (p+q-1) = (p-1)(q-1)$ as claimed. QED.

Phi for arbitrary numbers

Theorem: If p is prime, then $\phi(p^k) = p^k - p^{k-1}$ for $k \geq 1$. If a and b are relatively prime, $\phi(ab) = \phi(a)\phi(b)$.

Example:

$$
\begin{aligned}
&\phi(750) \\
&= \phi(2 \times 3 \times 5^3) \\
&= \phi(2) \times \phi(3) \times \phi(5^3) \\
&= (2-1) \times (3-1) \times (5^3 - 5^2) \\
&= 2 \times (125 - 25) \\
&= 200.
\end{aligned}
$$

Double check that this rule correctly generalizes the rules we already discussed for $\phi(p)$ and $\phi(pq)$.

Note: Practical if factorization is known. Otherwise, not so much.

An asymmetry

"Practical if factorization is known. Otherwise, not so much." This points to an asymmetry in difficulty: *if someone knows the factors of n, it is much easier for them to compute inverses mod n.*

We can exploit this asymmetry of difficulty!

Public key cryptography

Bob wants to be able to receive secret messages. Bob creates a private key, which must remain secret. Bob also creates a public key, which is made public. Anyone, Alice say, who wants to send a secret message to Bob can encrypt it with Bob's public key. Only Bob can decrypt such messages (using his private key).

No other communication or agreements or secret emails are needed.

Beforehand

1. Bob generates two distinct (hundreds of digits long) primes, p and q and keeps them hidden.
2. Bob sets $n ::= pq$.
3. Bob selects an integer $e \in [1, n)$ such that $\gcd(e, (p-1)(q-1)) = 1$. The public key is the pair (e, n).
4. Compute $d \in [1, n)$ such that $de \equiv 1 \pmod{(p-1)(q-1)}$. The private key is the pair (d, n).

Time to send and receive

Encryption: Alice wants to send unencryted message m. She computes and then sends the encrypted message $m^* = \text{rem}(m^e, n)$. (Uses e and n, which are public.)

Decryption: Bob receives m^*. He decrypts by computing $m' = \text{rem}((m^*)^d, n)$. (Uses d and n, where d is private.)

Is this *correct*? Is this *secure*?