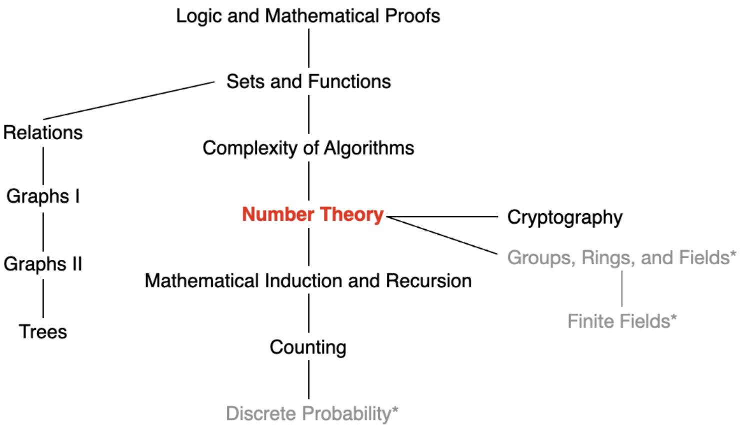
Discrete Mathematics for Computer Science

Lecture 9 + Lecture 10: Cryptography

Dr. Ming Tang

Department of Computer Science and Engineering
Southern University of Science and Technology (SUSTech)
Email: tangm3@sustech.edu.cn

This Lecture



Number Theory: divisibility and modular arithmetic, integer representations, primes, greatest common divisors, linear congruences, application, ...

Modular Arithmetic in CS

Modular arithmetic and congruencies are used in CS:

- Pseudorandom number generators
- Hash functions
- Cryptography

Pseudorandom Number Generators

Linear congruential method

We choose four numbers:

- the modulus m
- multiplier a
- increment c
- seed x_0

We generate a sequence of numbers $x_1, x_2, \ldots, x_n, \ldots$ with $0 \leq x_i < m$ by using the congruence

$$x_{n+1} = (ax_n + c) \textbf{ mod } m$$

Pseudorandom Number Generators

Linear congruential method

$$x_{n+1} = (ax_n + c) \textbf{ mod } m$$

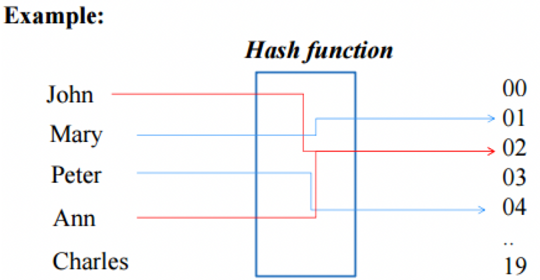
$m = 9$, $a = 7$, $c = 4$, and $x_0 = 3$:

$$\begin{aligned}
x_1 &= 7x_0 + 4 \textbf{ mod } 9 = 7 \cdot 3 + 4 \textbf{ mod } 9 = 25 \textbf{ mod } 9 = 7,\\
x_2 &= 7x_1 + 4 \textbf{ mod } 9 = 7 \cdot 7 + 4 \textbf{ mod } 9 = 53 \textbf{ mod } 9 = 8,\\
x_3 &= 7x_2 + 4 \textbf{ mod } 9 = 7 \cdot 8 + 4 \textbf{ mod } 9 = 60 \textbf{ mod } 9 = 6,\\
x_4 &= 7x_3 + 4 \textbf{ mod } 9 = 7 \cdot 6 + 4 \textbf{ mod } 9 = 46 \textbf{ mod } 9 = 1,\\
x_5 &= 7x_4 + 4 \textbf{ mod } 9 = 7 \cdot 1 + 4 \textbf{ mod } 9 = 11 \textbf{ mod } 9 = 2,\\
x_6 &= 7x_5 + 4 \textbf{ mod } 9 = 7 \cdot 2 + 4 \textbf{ mod } 9 = 18 \textbf{ mod } 9 = 0,\\
x_7 &= 7x_6 + 4 \textbf{ mod } 9 = 7 \cdot 0 + 4 \textbf{ mod } 9 = 4 \textbf{ mod } 9 = 4,\\
x_8 &= 7x_7 + 4 \textbf{ mod } 9 = 7 \cdot 4 + 4 \textbf{ mod } 9 = 32 \textbf{ mod } 9 = 5,\\
x_9 &= 7x_8 + 4 \textbf{ mod } 9 = 7 \cdot 5 + 4 \textbf{ mod } 9 = 39 \textbf{ mod } 9 = 3.
\end{aligned}$$

This sequence contains nine different numbers before repeating.

Hash Functions

A hash function is an algorithm that maps data of arbitrary length to data of a fixed length. The values returned by a hash function are called hash values or hash codes.



Hash Functions

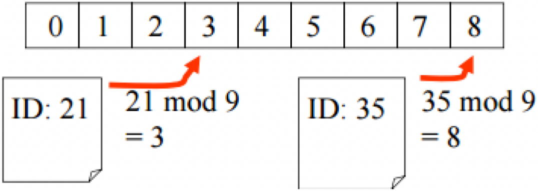
Problem: Given a large collection of records, how can we store and find a record quickly?

Solution: Use a hash function, calculate the location of the record based on the record's ID.

A common function is

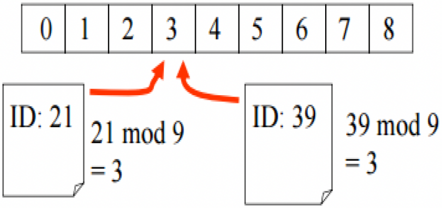
$$h(k) = k \ \mathbf{mod} \ m,$$

where m is the number of available storage locations.



Hash Functions

Two records mapped to the same location



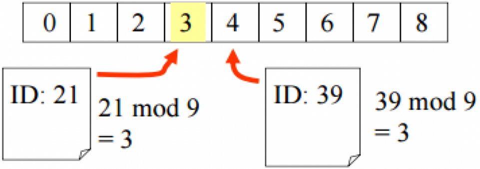
How to address this?

Hash Functions

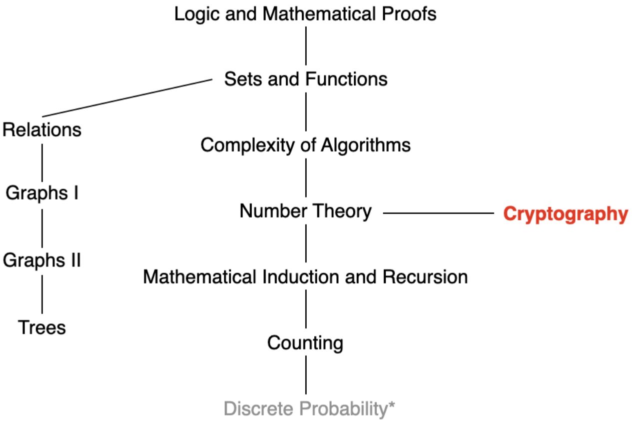
One way is to assign the first free location following the occupied memory location assigned by the hashing function.

try

$$h_0(k) = k \textbf{ mod } n$$
$$h_1(k) = (k+1) \textbf{ mod } n$$
$$\ldots$$
$$h_m(k) = (k+m) \textbf{ mod } n$$



This Lecture



Cryptograhy: classical cryptography, RAS cryptosystem, ...

Cryptography

History of almost 4000 years (from 1900 B.C.)

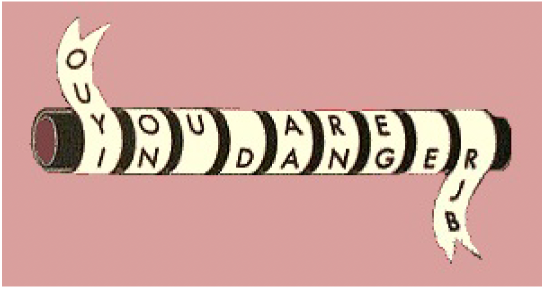
Cryptography = kryptos + graphos

- kryptos: secret
- graphos: writing

One-sentence definition: "Cryptography is the practice and study of techniques for secure communication in the presence of third parties called adversaries." – Ronald L. Rivest

Some Examples

In 405 B.C., the Greek general LYSANDER OF SPARTA was sent a coded message written on the inside of a servant's belt.



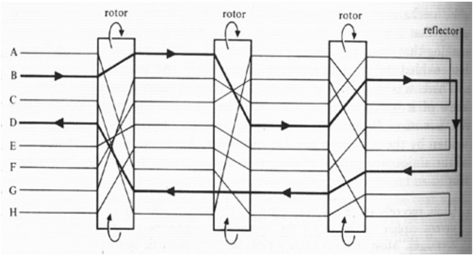
Some Examples

The Greeks also invented a cipher which changed letters to numbers. A form of this code was still being used during World War I.

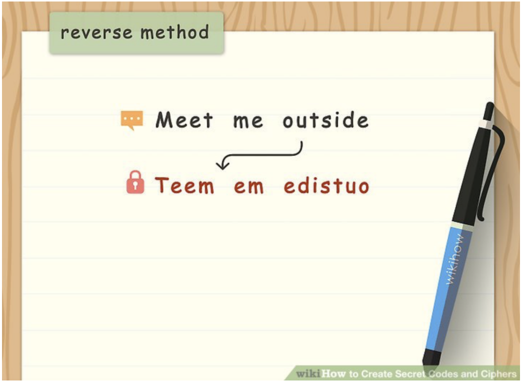
	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

Some Examples

Enigma, Germany coding machine in World War II.



Some Examples



Shift Ciphers

Shift Ciphers: Make messages secret by shifting each letter several letters forward in the alphabet.

Encryption:

- Assign each letter an integer $p \in \mathbf{Z}_{26} = \{0, 1, ..., 25\}$ based on the location of the letter in the alphabet.
- Replace p with $f(p)$:

$$f(p) = (p + k) \textbf{ mod } 26.$$

- Maps $f(p)$ back to the alphabet.

Shift Ciphers

Example: What is the secret message produced from the message "MEET YOU IN THE PARK" using the Shift cipher with $k = 3$?

Solution: First replace the letters in the message with numbers. This produces

12 4 4 19   24 14 20   8 13   19 7 4   15 0 17 10.

Now replace each of these numbers p by $f(p) = (p + 3)$ **mod** 26. This gives

15 7 7 22   1 17 23   11 16   22 10 7   18 3 20 13.

Translating this back to letters produces the encrypted message "PHHW BRX LQ WKH SDUN."

Shift Ciphers

Shift Ciphers: Make messages secret by shifting each letter several letters forward in the alphabet.

Decryption:

- Assign each letter an integer $p \in \mathbf{Z}_{26} = \{0, 1, ..., 25\}$ based on the location of the letter in the alphabet.
- Replace p with $f^{-1}(p)$:

$$f^{-1}(p) = (p - k) \textbf{ mod } 26.$$

- Maps $f^{-1}(p)$ back to the alphabet.

Shift Ciphers

We can generalize shift ciphers further to slightly enhance security by using a function of the form

$$f(p) = (ap + b) \textbf{ mod } 26.$$

How about the decryption? Suppose $\gcd(a, 26) = 1$.

Suppose that $c = (ap + b) \textbf{ mod } 26$ with $\gcd(a, 26) = 1$. To decrypt, we need to show how to express p in terms of c. That is, we solve the congruence for p:

$$c \equiv ap + b \pmod{26}.$$

Subtract b from both sides, we have $ap \equiv c - b \pmod{26}$. Since $\gcd(a, 26) = 1$, we know that there is an inverse $\bar{a}$ of a modulo 26:

$$p \equiv \bar{a}(c - b) \pmod{26}.$$

Cryptanalysis

The process of recovering plaintext from ciphertext without knowledge of both the encryption method and the key is known as cryptanalysis or breaking codes.

How to break messages that were encrypted using a shift cipher?

Solution 1: Try each 26 possible shifts.

Solution 2: Try different values of k based on the frequency of letters in the ciphertext. The nine most common letters in English text: E: 13%, T: 9%, A: 8%, O: 8%, I: 7%, N: 7%, S: 7%, H: 6%, and R: 6%.

Private Key Cryptosystem

In a **private key cryptosystem**, once you know an encryption key, you can quickly find the decryption key.

When a private key cryptosystem is used, two parties who wish to communicate in secret must share a secret key.

Any problems?

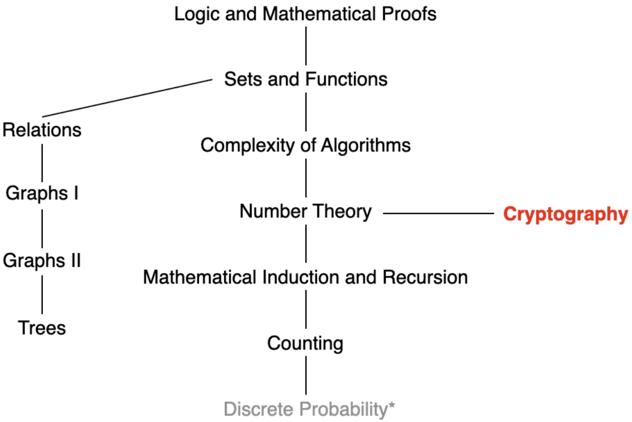
- Two people who want to communicate securely need to securely exchange this key.
- New key is used for each communication session between two parties.

Public Key Cryptosystem

In **public key cryptosystems**, knowing how to send an encrypted message does not help decrypt messages.

- Public key is known to the public.
- Private key is kept secret: only the intended recipient of a message can decrypt it.

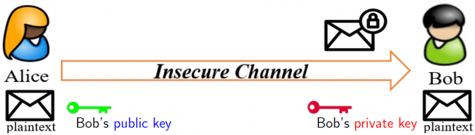
This Lecture



Cryptograhy: classical cryptography, RSA cryptosystem, ...

Overview

- RSA as Public Key System
  - Only target recipient can decrypt the message:



- RSA as Digital Signature
- Diffie-Hellman Key Exchange Protocol

RSA Cryptosystem

Rivest-Shamir-Adleman **2002 Turing Award**

2002	Ronald L. Rivest, Adi Shamir and Leonard M. Adleman	For their ingenious contribution for making **public-key cryptography** useful in practice.

Pick two large primes p and q. Let $n = pq$. **Encryption key** (n, e) and **decryption key** (n, d) are selected such that

- $\gcd(e, (p-1)(q-1)) = 1$
- $ed \equiv 1 \pmod{(p-1)(q-1)}$

RSA encryption: $C = M^e \bmod n$

RSA decryption: $M = C^d \bmod n$

RSA Encryption

1. Translate a plaintext message into integers, each with two digits, e.g, A is translated into 00, B into 01, . . . , and Z into 25.
2. Divide this string into equally sized blocks of $2N$ digits
   - $2N$ is the largest even number such that the number 2525...25 with $2N$ digits does not exceed n.
3. For each block, transform it into a ciphertext block:

$$C = M^e \textbf{ mod } n$$

RSA Encryption: Example

Encrypt the message "STOP" with key $(n = 2537, e = 13)$. Note that $2537 = 43 \cdot 59$, where $p = 43$ and $q = 59$ are primes, and $\gcd(e, (p-1)(q-1)) = 1$.

Solution:

1. Translate into integers: 18191415
2. Divide this into blocks of 4 digits (because $2525 < 2537 < 252525$): 1819 1415
3. Encrypt each block using the mapping

$$C = M^{13} \textbf{ mod } 2537.$$

We have $1819^{13} \textbf{ mod } 2537 = 2081$ and $1415^{13} \textbf{ mod } 2537 = 2182$. The encrypted message is 2081 2182.

RSA Decryption

For each block, transform the ciphertext into plaintext message:

$$M = C^d \textbf{ mod } n$$

Example: What is the decrypted message of 0981 0461 with $e = 13$, $p = 43$, $q = 59$?

Solution: Recall that $ed \equiv 1 \ (\textbf{mod } (p-1)(q-1))$. Thus, $d = 937$ is an inverse of 13 modulo $42 \cdot 58 = 2436$.

For each block, transform it into plaintext message:

$$M = C^{937} \textbf{ mod } 2537.$$

Since 0981^{937} **mod** $2537 = 0704$ and 0461^{937} **mod** $2537 = 1115$, the plaintext message is 0704 1115, which is "HELP".

RAS Cryptosystem

Pick two large primes p and q. Let $n = pq$. **Encryption key** (n, e) and **decryption key** (n, d) are selected such that

(1) $\gcd(e, (p-1)(q-1)) = 1$

(2) $ed \equiv 1 \ (\mathbf{mod}\ (p-1)(q-1))$

RSA encryption: $C = M^e \ \mathbf{mod}\ n$;
RSA decryption: $M = C^d \ \mathbf{mod}\ n$. Why?

According to (1), the inverse d exists. According to (2), there exists an integer k such that

$$de = 1 + k(p-1)(q-1).$$

It follows that $C^d \equiv (M^e)^d = M^{de} = M^{1+k(p-1)(q-1)} (\mathbf{mod}\ n)$.

Assuming that $\gcd(M, p) = \gcd(M, q) = 1$, we have $M^{p-1} \equiv 1 \ (\mathbf{mod}\ p)$ and $M^{q-1} \equiv 1 (\mathbf{mod}\ q)$.

RAS Cryptosystem

According to (1), the inverse d exists. According to (2), there exists an integer k such that

$$de = 1 + k(p-1)(q-1).$$

It follows that $C^d \equiv (M^e)^d = M^{de} = M^{1+k(p-1)(q-1)} (\textbf{mod } n)$.

Assuming that $\gcd(M, p) = \gcd(M, q) = 1$, we have $M^{p-1} \equiv 1 \ (\textbf{mod } p)$ and $M^{q-1} \equiv 1 (\textbf{mod } q)$.

$$C^d \equiv M \cdot (M^{p-1})^{k(q-1)} \equiv M \cdot 1 = M \ (\textbf{mod } p)$$
$$C^d \equiv M \cdot (M^{q-1})^{k(p-1)} \equiv M \cdot 1 = M \ (\textbf{mod } q).$$

Because $\gcd(p, q) = 1$, we have

$$C^d \equiv M \ (\textbf{mod } pq).$$

This basically implies that

$$M = C^d \textbf{ mod } n$$

RSA as Public Key System

Pick two large primes p and q. Let $n = pq$. **Encryption key** (n, e) and **decryption key** (n, d) are selected such that

(1) $\gcd(e, (p-1)(q-1)) = 1$

(2) $ed \equiv 1 \pmod{(p-1)(q-1)}$

RSA encryption: $C = M^e \textbf{ mod } n$;
RSA decryption: $M = C^d \textbf{ mod } n$.

RSA as a Public Key System

- Public key: (n, e)
- Private key: d
- p, q must be kept secret!

Why is the RSA cryptosystem suitable for public key cryptography?

RSA as Public Key System

RSA as a Public Key System

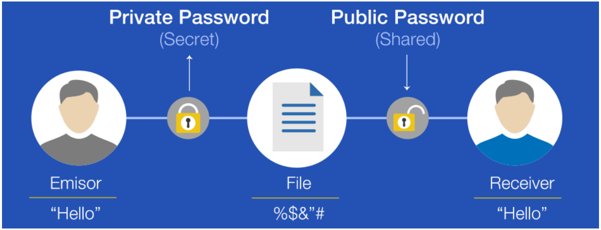
- Public key: (n, e); Private key: (n, d)
- p, q must be kept secret!

Why is the RSA cryptosystem suitable for public key cryptography?

- It is possible to rapidly construct a public key by finding two large primes p and q, each with more than 200 digits.
- When we know p and q, we can quickly find an inverse d.
- However, no method is known to decrypt messages that is not based on finding a factorization of n.
  - Factorization is believed to be a difficult problem.
  - The most efficient factorization methods known (as of 2010) require billions of years to factor 400-digit integers.

Overview

- RSA as Public Key System
- RSA as Digital Signature
  - The recipient of the message knows that it came from the person they think it came from.



- Diffie-Hellman Key Exchange Protocol

RSA as Digital Signature

Alice's RSA public key is (n, e) and her private key is d.

Alice splits the plaintext message into blocks and applies her decryption function:

$$S = M^d \ \mathbf{mod}\ n \quad (\textbf{RSA signature})$$

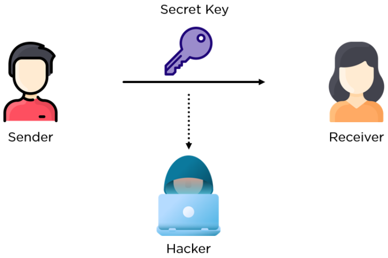
When a recipient receives her message, they apply Alice's encryption function:

$$M = S^e \ \mathbf{mod}\ n \quad \textbf{(RSA verification)}$$

Alice can send her message to as many people as she wants and by signing it in this way, every recipient can be sure it came from Alice.

Overview

- RSA as a Public Key System
- RSA as Digital Signature
- Diffie-Hellman Key Exchange Protocol
  - Exchange a secret key over an insecure communications channel



Diffie-Hellman Key Exchange Protocol

Two parties exchange a secret key over an insecure communications channel without having shared any information in the past.

Diffie-Hellman Key Exchange Protocol

Before introducing the protocol:

Definition: A primitive root modulo a prime p is an integer r in $\mathbf{Z}_p$ such that every nonzero element of $\mathbf{Z}_p$ is a power of r.

Example: Whether 2 is a primitive root modulo 11?

When we compute the powers of 2 in $\mathbf{Z}_{11}$, we obtain $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, $2^4 = 5$, $2^5 = 10$, $2^6 = 9$, $2^7 = 7$, $2^8 = 3$, $2^9 = 6$, $2^{10} = 1$. Because every nonzero element of $\mathbf{Z}_{11}$ is a power of 2, 2 is a primitive root of 11.

Diffie-Hellman Key Exchange Protocol

Suppose that Alice and Bob want to share a common key. Consider $\mathbf{Z}_p$.

(1) Alice and Bob agree to use a prime p and a primitive root a of p.
(2) Alice chooses a secret integer k_1 and sends $a^{k_1} \textbf{ mod } p$ to Bob.
(3) Bob chooses a secret integer k_2 and sends $a^{k_2} \textbf{ mod } p$ to Alice.
(4) Alice computes $(a^{k_2})^{k_1} \textbf{ mod } p$.
(5) Bob computes $(a^{k_1})^{k_2} \textbf{ mod } p$.

Alice and Bob have computed their shared key:

$$(a^{k_2})^{k_1} \textbf{ mod } p = (a^{k_1})^{k_2} \textbf{ mod } p.$$

- Public information: p, a, $a^{k_1} \textbf{ mod } p$, and $a^{k_2} \textbf{ mod } p$
- Secret: k_1, k_2, $(a^{k_2})^{k_1} \textbf{ mod } p = (a^{k_1})^{k_2} \textbf{ mod } p$

Note that it is very hard to determine k_1 with a, p, and $a^{k_1} \textbf{ mod } p$.

Next Lecture