Number Theory and Cryptography

Chapter 4

With Question/Answer Animations

Chapter Summary

4.1 Divisibility and Modular Arithmetic

4.2 Integer Representations and Algorithms

4.3 Primes and Greatest Common Divisors

4.4 Solving Congruences

4.5 Applications of Congruences

4.6 Cryptography

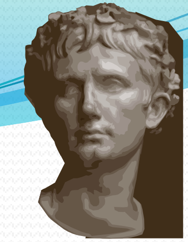
Cryptography

Section 4.6

Section Summary

- Classical Cryptography
- Cryptosystems
- Public Key Cryptography
- RSA Cryptosystem
- Crytographic Protocols

Caesar Cipher



The process of making a message secret is *encryption*.

Julius Caesar created secret messages by shifting each letter three letters forward (sending the last three letters to the first three letters.)

- For example,
  - B is replaced by E
  - X is replaced by A.

Here is how the encryption process works:

- Replace each letter by an integer from $\mathbf{Z}_{26}$, that is an integer from 0 to 25
  - representing one less than its position in the alphabet.
- The encryption function is $f(p) = (p + 3)$ **mod** 26.
  - It replaces each integer p in the set $\{0,1,2,\ldots,25\}$ by $f(p)$ in the set $\{0,1,2,\ldots,25\}$.
- Replace each integer p by the letter with the position $p + 1$ in the alphabet.

Example: Encrypt "MEET YOU IN THE PARK" using the Caesar cipher.

Solution: 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.

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

Translating the numbers back to letters produces the encrypted message

"PHHW BRX LQ WKH SDUN."

Caesar Cipher

The process of recovering the original message is *decryption*.

- To recover the original message, use $f^{-1}(p) = (p-3)$ **mod** 26.
  - So, each letter in the coded message is shifted back three letters (with the first three letters sent to the last three letters)
- The Caesar cipher is one of a family of *shift ciphers*.
- Letters are shifted by an integer k, with 3 just one possibility.
- The encryption function is $f(p) = (p + k)$ **mod** 26
- The decryption function is $f^{-1}(p) = (p-k)$ **mod** 26
- k is the *key*.

Shift Cipher

Example 1: Encrypt "STOP GLOBAL WARMING" using the shift cipher with $k = 11$.

Solution: Replace each letter with an element of $\mathbf{Z}_{26}$.

18 19 14 15 6 11 14 1 0 11 22 0 17 12 8 13 6.

Apply the shift $f(p) = (p + 11)$ **mod** 26, yielding

3 4 25 0 17 22 25 12 11 22 7 11 2 23 19 24 17.

Translating #s back to letters produces the ciphertext

"DEZA RWZMLW HLCXTYR."

Shift Cipher

Example 2: Decrypt "LEWLYPLUJL PZ H NYLHA ALHJOLY" that was encrypted using shift cipher with $k = 7$.

Solution: Replace each letter with an element of $\mathbf{Z}_{26}$:

11 4 22 11 24 15 11 20 9 11 15 25 7 13 24 11 7 0 0 11 7 9 14 11 24.

Shift each of #s by $-k = -7$ modulo 26, yielding

4 23 15 4 17 8 4 13 2 4 8 18 0 6 17 4 0 19 19 4 0 2 7 4 17

Translating the #s back to letters produces

"EXPERIENCE IS A GREAT TEACHER."

Affine Ciphers

Shift ciphers are special case of *affine ciphers,* whose encryption function is

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

where a and b are integers chosen so that f is a bijection (*i.e.*, $\gcd(a, 26) = 1$)

Example: What letter replaces the letter K when the function $f(p) = (7p + 3)$ **mod** 26 is used for encryption.

Solution: Since 10 represents K, $f(10) = (7\cdot10 + 3)$ **mod** 26 $=21$, which is then replaced by V.

- To decrypt a message, solve $c \equiv ap + b \pmod{26}$ for p.
  - Subtract b from both sides to obtain $c - b \equiv ap \pmod{26}$.
  - Multiply both sides by inverse of a mod 26, which exists since $\gcd(a,26) = 1$.
  - $\bar{a}(c - b) \equiv \bar{a}ap \pmod{26}$, which simplifies to $\bar{a}(c - b) \equiv p \pmod{26}$.
  - $p \equiv \bar{a}(c - b) \pmod{26}$ is used to determine p in $\mathbf{Z}_{26}$.

Cryptanalysis of Affine Ciphers

The process of recovering plaintext from ciphertext without knowledge of the encryption method is known as *cryptanalysis*.

- An important tool for cryptanalyzing ciphertext produced with any a bijection of letters is the relative frequencies of letters.
- The 9 most common letters in the English texts are

  E 13%, T 9%, A 8%, O 8%, I 7%, N 7%, S 7%, H 6%, and R 6%.
- To analyze ciphertext where a shift cipher is suspected
  - Find the frequency of the letters in the ciphertext.
  - Hypothesize that the most frequent letter is produced by encrypting E.
  - If the value of the shift from E to the most frequent letter is k, shift the ciphertext by $-k$ and see if it makes sense.
  - If not, try T as a hypothesis and continue.
- **Example**: intercepted message "ZNK KGXRE HOXJ MKZY ZNK CUXS". Let's cryptanalyze.
- **Solution**: The most common letter in the ciphertext is K. So perhaps the letters were shifted by 6 since this would then map E to K. Shifting the entire message by -6 gives us "THE EARLY BIRD GETS THE WORM."

Block Ciphers

- Ciphers that replace each letter of the alphabet by another letter are *character* or *monoalphabetic* ciphers.
- They are vulnerable to cryptanalysis based on letter frequency.
- *Block ciphers* avoid this problem, by replacing blocks of letters with other blocks of letters.
- A simple block cipher is the *transposition cipher*.
  - The key is a *permutation* σ of the set $\{1,2,\ldots,m\}$, $m \in Z$

  (that is a one-to-one function from $\{1,2,\ldots,m\}$ to itself)
- To encrypt a message, split the letters into blocks of size *m*, adding additional letters to fill out the final block.
- We encrypt $p_1, p_2,\ldots, p_m$ as $c_1, c_2,\ldots, c_m = p_{\sigma(1)}, p_{\sigma(2)},\ldots, p_{\sigma(m)}$.
- To decrypt $c_1, c_2,\ldots, c_m$ apply the inverse permutation σ^{-1}.

Block Ciphers

Example: Using the permutation σ of {1,2,3,4} with

$$\sigma(1) = 3, \sigma(2) = 1, \sigma(3) = 4, \sigma(4) = 2,$$

a. Encrypt the plaintext PIRATE ATTACK
b. Decrypt the ciphertext message SWUE TRAEOEHS.

Solution:

a. Split into four blocks PIRA TEAT TACK.
   Apply the permutation σ giving IAPR ETTA AKTC.
b. σ^{-1}: $\sigma^{-1}(1) = 2, \sigma^{-1}(2) = 4, \sigma^{-1}(3) = 1, \ \sigma^{-1}(4) = 3.$
   Apply the permutation σ^{-1} giving USEW ATER HOSE.
   Split into words to obtain USE WATER HOSE.

Cryptosystems

Definition: A *cryptosystem* is a 5-tuple $(\mathcal{P},\mathcal{C},\mathcal{K},\mathcal{E},\mathcal{D})$, where

- $\mathcal{P}$ is the set of plaintext strings,
- $\mathcal{C}$ is the set of ciphertext strings,
- $\mathcal{K}$ is the *keyspace* (set of all possible keys),
- $\mathcal{E}$ is the set of encryption functions, and
- $\mathcal{D}$ is the set of decryption functions.

- The encryption function in $\mathcal{E}$ corresponding to the key k is denoted by E_k and the decryption function in $\mathcal{D}$ that decrypts cipher text encrypted using E_k is denoted by D_k. Therefore:

$$D_k(E_k(p)) = p, \text{ for all plaintext strings } p.$$

Cryptosystems

Example: Describe shift ciphers as a cryptosystem.

Solution:

- $\mathcal{P}$ is the set of strings of elements in $\mathbf{Z}_{26}$,
- $\mathcal{C}$ is the set of strings of elements in $\mathbf{Z}_{26}$,
- $\mathcal{K} = \mathbf{Z}_{26}$,
- $\mathcal{E}$ consists of functions $E_k(p) = (p + k)\ \mathbf{mod}\ 26$,
- $\mathcal{D}$ is the same as $\mathcal{E}$ where $D_k(p) = (p - k)\ \mathbf{mod}\ 26$.

Public Key Cryptography

- All classical ciphers are *private key cryptosystems*.
  - Knowing encryption key allows one to quickly determine decryption key.
- All parties who wish to communicate using private key cryptosystem must share the key and keep it a secret.
- In public key cryptosystems, invented in the 1970s, knowing how to encrypt does not help one to decrypt.
- Therefore, everyone can have a publicly known encryption key.
- Only the decryption key needs to be kept secret.

Clifford Cocks (Born 1950)

The RSA Cryptosystem

- A public key cryptosystem, now known as the RSA system was introduced in 1976 by three researchers at MIT.

Ronald Rivest (Born 1948)

Adi Shamir (Born 1952)

Leonard Adelman (Born 1945)

- It is now known that the method was discovered earlier by Clifford Cocks, working secretly for the UK government.

The public encryption key is (n, e):

- $n = pq$ (the modulus) is the product of two large (~200 digit) primes p, q,
- e is an exponent that is relatively prime to $(p-1)(q-1)$.
- The two large primes can be found using probabilistic primality tests.
- But $n = pq$, with approximately 400 digits, cannot be factored in a reasonable length of time.

RSA Encryption

To encrypt a message using RSA using a key (n,e) :

i. Translate the plaintext message M into sequences of two digit integers representing the letters. Use 00 for A, 01 for B, etc.

ii. Concatenate the two digit integers into strings of digits.

iii. Divide this string into equally sized blocks of $2N$ digits where $2N$ is the largest even number such that $\underbrace{2525\ldots25}_{2N}$ does not exceed n.

iv. The plaintext message M is now a sequence of integers $m_1, m_2, \ldots, m_k$.

v. Each block (an integer) is encrypted using the function $C = M^e \textbf{ mod } n$.

Example: Encrypt STOP using the RSA cryptosystem with key(2537,13).

- $2537 = 43 \cdot 59$,
- $p = 43$ and $q = 59$ are primes and $\gcd(e,(p-1)(q-1)) = \gcd(13, 42 \cdot 58) = 1$.

Solution: Translate letters in STOP to their numerical equivalents 18 19 14 15.

- Divide into blocks of four digits (because $2525 < 2537 < 252525$) to obtain 1819 1415.
- Encrypt each block using the mapping $C = M^{13} \textbf{ mod } 2537$.
- Since $1819^{13} \text{ mod } 2537 = 2081$ and $1415^{13} \text{ mod } 2537 = 2182$, the encrypted message is 2081 2182.

RSA Decryption

To decrypt a RSA ciphertext, the decryption key d, an inverse of e mod $(p-1)(q-1)$ is needed. The inverse exists since $\gcd(e,(p-1)(q-1)) = \gcd(13, 42\cdot 58) = 1$.

- With the decryption key d, we can decrypt each block with the computation

  $M = C^d$ **mod** $p{\cdot}q$. (*see text for full derivation*)

- RSA works as a public key system since the only known method of finding d is based on a factorization of n into primes.
- There is currently no known feasible method for factoring large #s into primes.

  Example: The message 0981 0461 is received. What is the decrypted message if it was encrypted using the RSA cipher from the previous example.

  Solution: The message was encrypted with $n = 43\cdot 59$ and exponent 13. An inverse of 13 modulo $42\cdot 58 = 2436$ (*exercise* 2 *in Section* 4.4) is $d = 937$.

  - To decrypt a block C, $M = C^{937}$ **mod** 2537.
  - Since 0981^{937} **mod** 2537 = 0704 and 0461^{937} **mod** 2537 = 1115, the decrypted message is 0704 1115.
  - Translating back to English letters, the message is HELP.

Cryptographic Protocols: Key Exchange

Cryptographic protocols are exchanges of messages carried out by two or more parties to achieve a particular security goal.

- *Key exchange* is a protocol by which two parties can exchange a secret key over an insecure channel without having any past shared secret information. Here the *Diffie-Hellman key agreement protocol* is described by example.
  i. Suppose that Alice and Bob want to share a common key.
  ii. Alice and Bob agree to use a prime p and a primitive root a of p.
  iii. Alice chooses a secret integer k_1 and sends a^{k_1} **mod** p to Bob.
  iv. Bob chooses a secret integer k_2 and sends a^{k_2} **mod** p to Alice.
  v. Alice computes $(a^{k_2})^{k_1}$ **mod** p.
  vi. Bob computes $(a^{k_1})^{k_2}$ **mod** p.

  At the end of the protocol, Alice and Bob have their shared key

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

- To find the secret information from the public information would require the adversary to find k_1 and k_2 from a^{k_1} **mod** p and a^{k_2} **mod** p respectively.
- This is an instance of the discrete logarithm problem, considered to be computationally infeasible when p and a are sufficiently large.

Cryptographic Protocols: Digital Signatures

Adding a *digital signature* to a message is a way of ensuring the recipient that the message came from the purported sender.

- Suppose that Alice's RSA public key is (n,e) and her private key is d. Alice encrypts a plain text message x using $E_{(n,e)}(x) = x^d \mathbf{mod}\ n$. She decrypts a ciphertext message y using $D_{(n,e)}(y) = y^d \mathbf{mod}\ n$.
- Alice wants to send a message M so that everyone who receives the message knows that it came from her.
  1. She translates the message to numerical equivalents and splits into blocks, just as in RSA encryption.
  2. She then applies her decryption function $D_{(n,e)}$ to the blocks and sends the results to all intended recipients.
  3. The recipients apply Alice's encryption function and the result is the original plain text since $E_{(n,e)}(D_{(n,e)}(x)) = x$.

Everyone who receives the message can then be certain that it came from Alice.

Cryptographic Protocols: Digital Signatures

Example: Suppose Alice's RSA cryptosystem is same as in earlier example with key(2537,13), $2537 = 43\cdot 59$, $p = 43$, $q = 59$ and

$$\gcd(e,(p-1)(q-1)) = \gcd(13, 42\cdot 58) = 1.$$

Her decryption key is d = 937. She wants to send "MEET AT NOON" to her friends so that they can be certain that the message is from her.

Solution: Alice translates the message into blocks of digits

$$1204\ 0419\ 0019\ 1314\ 1413.$$

1. She then applies her decryption transformation $D_{(2537,13)}(x) = x^{937}$ **mod** 2537 to each block.
2. She finds (using her laptop, programming skills, and knowledge of discrete mathematics) that 1204^{937} **mod** 2537 = 817, 419^{937} **mod** 2537 = 555 , 19^{937} **mod** 2537 = 1310, 1314^{937} **mod** 2537 = 2173, and 1413^{937} **mod** 2537 = 1026.
3. She sends 0817 0555 1310 2173 1026.

When one of her friends receive the message, they apply Alice's encryption transformation $E_{(2537,13)}$ to each block. They then obtain the original message which they translate back to English letters.