Chap 2. Basic Encryption and Decryption

H. Lee Kwang

Department of Electrical Engineering

& Computer Science,

KAIST

Objectives

- Concepts of encryption
- Cryptanalysis: how encryption systems are “broken”

2.1 Terminology and Background

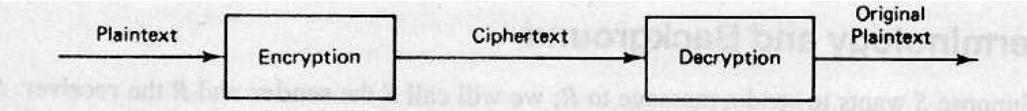
- Notations
  - S: sender
  - R: receiver
  - T: transmission medium
  - O: outsider, interceptor, intruder, attacker, or, adversary
- S wants to send a message to R
  - S entrusts the message to T who will deliver it to R
  - Possible actions of O
    - block(interrupt), intercept, modify, fabricate
    - Chapter 1

2.1.1 Terminology

- Encryption and Decryption
  - encryption: a process of encoding a message so that its meaning is not obvious
  - decryption: the reverse process
- encode(encipher) vs. decode(decipher)
  - encoding: the process of translating entire words or phrases to other words or phrases
  - enciphering: translating letters or symbols individually
  - encryption: the group term that covers both encoding and enciphering

2.1.1 Terminology

- Plaintext vs. Ciphertext
  - P(plaintext): the original form of a message
  - C(ciphertext): the encrypted form
- Basic operations
  - plaintext to ciphertext: encryption: C = E(P)
  - ciphertext to plaintext: decryption: P = D(C)
  - requirement: P = D(E(P))



2.1.1 Terminology

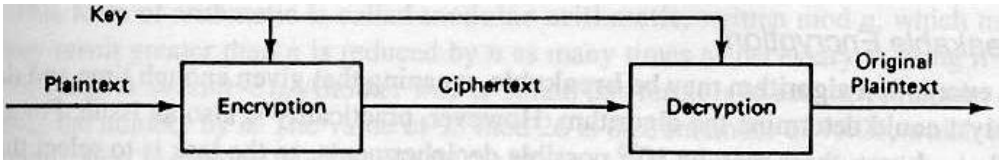
- Encryption with key
  - encryption key: K_E
  - decryption key: K_D
  - $C = E(K_E, P)$
  - $P = D(K_D, E(K_E, P))$

If the encryption algorithm should fall into the interceptor's hands, future messages can still be kept secret because the interceptor will not know the key value

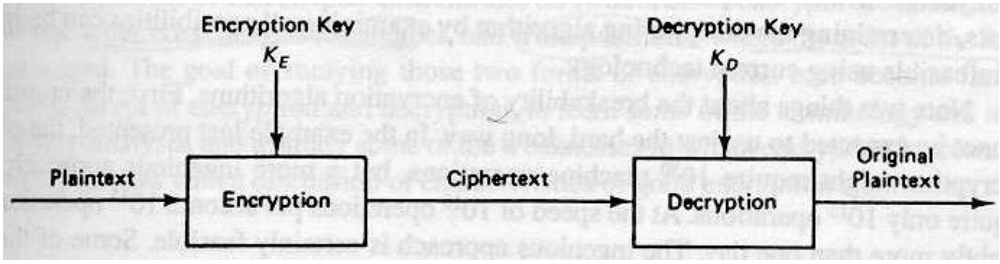
- Keyless Cipher
  - a cipher that does not require the use of a key
    - key cannot be changed

2.1.1 Terminology

- Symmetric Cryptosystem: $K_E = K_D$



- Asymmetric Cryptosystem: $K_E \neq K_D$



2.1.1 Terminology

- Cryptography
  - cryptography means hidden writing, the practice of using encryption to conceal text
- Cryptanalysis
  - cryptanalyst studies encryption and encrypted message, with the goal of finding the hidden meaning of the messages
- Cryptology
  - includes both cryptography and cryptanalysis

2.1.1 Terminology

- Cryptanalysis
  - break an encryption
  - cryptanalyst can do any or all of three different things:
    - attempt to break a single message
    - attempt to recognize patterns in encrypted messages, in order to be able to break subsequent ones by applying a straightforward decryption algorithm
    - attempt to find general weakness in an encryption algorithm, without necessarily having intercepted any messages

2.1.1 Terminology

- Breakable encryption
  - An encryption algorithm may be breakable, meaning that given enough time and data, an analyst could determine the algorithm
  - practicality is an issue
    - for a given cipher scheme, there may be 10^{30} possible decipherments, so the task is to select the right one out of the 10^{30}
    - cryptanalyst cannot be expected to try just the hard, long way
      - another efficient algorithm may exist
    - estimates of breakability are based on current technology
      - budget dependent

Key Length

- 1996

Type of Attacker	Budget	Tool	Time and cost per key recovered 40bits	56bits	Length Needed for protection in Late 1995
Pedestrian Hacker					
	tiny	scavenged computer time	1 week	infeasible	45
	$400	FPGA	5 hours ($0.08)	38 years ($5,000)	50
Small Business					
	$10,000	FPGA	12 minutes ($0.08)	18 months ($5,000)	55
Corporate Department					
	$300K	FPGA or ASIC	24 seconds ($0.08) .18 seconds ($0.001)	19 days ($5,000) 3 hours ($38)	60
Big Company					
	$10M	FPGA or ASIC	.7 seconds ($0.08) .005 seconds ($0.001)	13 hours ($5,000) 6 minutes ($38)	70
Intelligence Agency					
	$300M	ASIC	.0002 seconds ($0.001)	12 seconds ($38)	75

2.1.2 Representation of Characters

- Encoding

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M
Code	0	1	2	3	4	5	6	7	8	9	10	11	12
Letter	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Code	13	14	15	16	17	18	19	20	21	22	23	24	25

  - modular arithmetic
    - $Y + 3 = B(24 + 3 = 27 \equiv 1 \bmod 26)$
- Two forms of encryption
  - substitution: one letter is exchanged for another
  - transposition: the order of the letters is rearranged

2.2 Monoalphabetic Ciphers(Substitution)

- Simple substitution
  - use a correspondence table
    - substitute each character by another character or symbol
  - monoalphabetic cipher
    - one-by-one

2.2.1 The Caesar Cipher

- Named for Julious Caesar
  - Caesar used a shift of 3

$$c_i = E(p_i) = (p_i + 3) \bmod 26$$

  - translation chart

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	d	e	f	g	h	I	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c

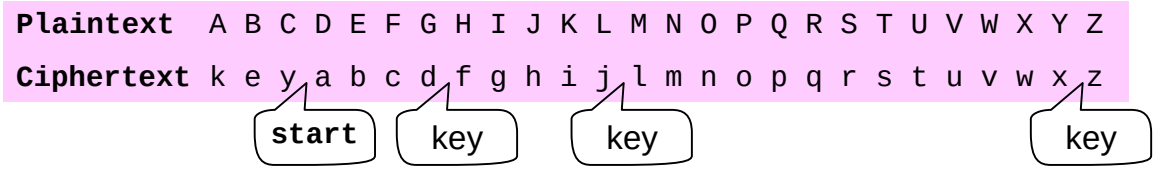
  - E(TREATY IMPOSSIBLE) = wuhdwb lpsrvvleoh
    - E(T) = w, E(R)=u, etc.

2.2.1 The Caesar Cipher

- Advantages and Disadvantages of the Caesar Cipher
  - advantage
    - easy to use
  - disadvantage
    - simple structure
    - easy to break

2.2.2 Other Monoalphabetic Substitutions

- Permutation based
  - generalization of the Caesar cipher
  - permutation $p : Z_{26} \rightarrow Z_{26}$
    - 1-1
    - example: $p(l) = (l + 3) \bmod 26$
  - use more complex rule
    - use a **key**, a word that controls the enciphering



2.2.2 Other Monoalphabetic Substitutions

- Another variation $p(l) = (3 \times l) \bmod 26$
  - $p(A) = a$ $p(0) = 0 \bmod 26$
  - $p(B) = d$ $p(1) = 3 \bmod 26$

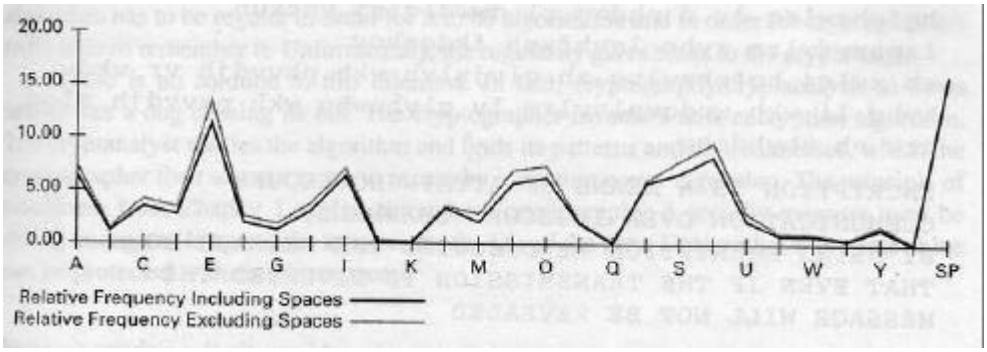
Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	a	d	g	j	m	p	s	v	y	b	e	h	k	n	q	t	w	z	c	f	i	l	o	r	u	x

2.2.3 Cryptanalysis of Monoalphabetic Ciphers

- Short words, words with repeated patterns, and common initial and final letters all give clues for guessing the permutation
- Cryptanalysis of Monoalphabetic Ciphers
  - a lot like working crossword puzzle
    - try a guess
    - continue to work to substantiate that guess until you have all the words in place, or until you reach a contradiction
  - we can use more advanced technique
    - frequency distribution

2.2.3 Cryptanalysis of Monoalphabetic Ciphers

- Frequency Distribution
  - relative frequencies of characters in English text



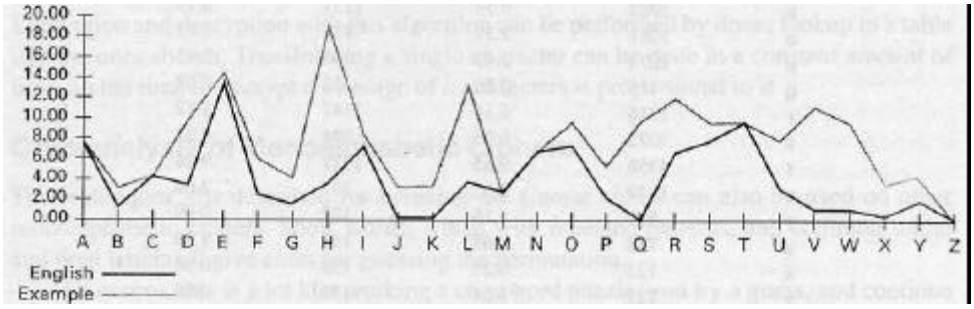
2.2.3 Cryptanalysis of Monoalphabetic Ciphers

- Ciphertext

```
hqfubswlrq lv d phdqv ri dwwdlqlqj vhfxuh
frpsxwdxlrq ryhu lqvhfxuh fkdqqhov
eb xvlqj hqfubswlrq zh glvjxlvh wkh phvvdjh vr wkdw
hyhq li wkh wudqvplvvlrq lv glyhuwhg whk phvvdjh
zloo qrw eh uhyhdohg
```

2.2.3 Cryptanalysis of Monoalphabetic Ciphers

- Cryptanalysis using frequency table



  - h may be one of e, a, i, o, etc.

2.2.3 Cryptanalysis of Monoalphabetic Ciphers

- Rule
  - h → e, d → a, l → i, r → o, etc.
- Results
  - `Ciphertext`
    - `hqfubswlrq lv d phdqv ri dwwdlqlqj vhfxuh…`
  - `Plaintext`
    - `ENCRYPTION IS A MEANS OF ATTAINING SECURE…`

2.2.3 Cryptanalysis of Monoalphabetic Ciphers

- Cryptographer's Dilemma
  - an encryption algorithm has to be regular in order for it to be algorithmic and in order for cryptographers to be able to remember it
  - the regularity gives clues to the cryptanalyst
  - there is no solution to this dilemma
- Cryptography : principle of timeliness

Principle of timeliness(Adequate Protection) :

A security measure must be strong enough to keep out the attacker for the life of the data. Data with a short time value can be protected with simple measures

2.3 Polyalphabetic Substitution Ciphers

A cipher is more cryptographically secure would display a rather flat distribution, which gives no information to a cryptanalyst

- Weakness of monoalphabetic ciphers
  - their frequency distribution reflects the distribution of the underlying alphabet
- Polyalphabetic substitution cipher
  - one way to flatten the distribution
  - to combine distributions that are high with ones that are low
    - T is sometimes enciphered as 'a' and sometimes as 'b'
    - frequency of 'a' is high and that of 'b' is low

2.3 Polyalphabetic Substitution Ciphers

- Simple example
  - odd position $p_1(l) = (3 \times l) \bmod 26$

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	a	d	g	j	m	p	s	v	y	b	e	h	k	n	q	t	w	z	c	f	i	l	o	r	u	x

  - even position $p_2(l) = ((5 \times l) + 13) \bmod 26$

Plaintext	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext	n	s	x	c	h	m	r	w	b	g	l	q	v	a	f	k	p	u	z	e	j	o	t	y	d	i

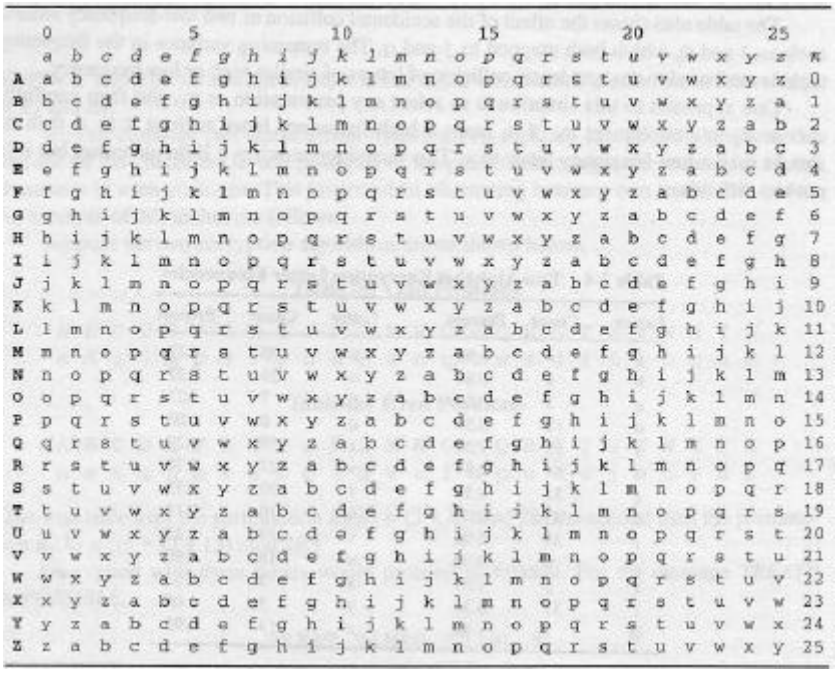
  - more complex but still simple
    - extend the number of permutations

2.3.1 Vigenère Tableaux

- Goal: flat distribution
- Algorithm
  - use Vigenère Tableau(26×26 matrix)
  - low: letter
  - column: key(suppose key length is 6, i.e. key = (k_1, …k_6))
  - encryption
    - $c_j = (p_j, k_i)$ where $j \equiv i \bmod 6$
  - decryption
    - p_j =x s.t. $c_j = (x, k_i)$ where $j \equiv i \bmod 6$

2.3.1 Vigenère Tableaux

- table



	0				5					10					15					20					25	
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	A
A	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	0
B	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	1
C	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	2
D	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	3
E	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	4
F	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	5
G	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	6
H	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	7
I	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	8
J	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	9
K	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	10
L	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	11
M	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	12
N	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	13
O	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	14
P	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	15
Q	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	16
R	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	17
S	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	18
T	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	19
U	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	20
V	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	21
W	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	22
X	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	23
Y	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	24
Z	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	25

2.3.1 Vigenère Tableaux

- Example

```
Key         julie tjuli etjul ...
Plaintext   BUTSO FTWHA TLIGH ...
Ciphertext koeas ycqsi ...
```

  - k = (B, j), o = (U, u), e = (T, l), …

- Long keywords can be used, but a keyword of length three usually suffices s to smooth out the distribution

2.3.2 Cryptanalysis of Polyalphabetic Substitutions

- Polyalphabetic substitutions are apparently more secure than monoalphabetic substitutions
  - but still insecure
  - key length
- Two tools
  - Kasiski method for repeated patterns
  - index of coincidence

2.3.2 Cryptanalysis of Polyalphabetic Substitutions

- Method of Kasiski
  - find the candidates of the key length
  - rely on the regularity of English
    - high frequency
    - ending: -th, -ing, -ed, -ion, -tion, -ation, etc.
    - beginning: im-, in-, un-, re-, etc.
    - pattern: -eek-, -oot-, -our-, etc.
    - word: of, and, to, the, with, are, is, that, etc.

2.3.2 Cryptanalysis of Polyalphabetic Substitutions

- Steps(Method of Kasiski)
  1. identify repeated patterns of three or more characters
  2. for each pattern write down the position at which each instance of the pattern begins
  3. compute the difference between the starting points of successive instances
  4. determine all factors of each difference
  5. if a polyalphabetic substitution cipher wae used, the key length will be one of the factors that appears often in step 4.

2.3.2 Cryptanalysis of Polyalphabetic Substitutions

- Example(Method of Kasiski)

Starting Position	Distance from Previous	Factors
20		
83	63(83-20)	3, 7, 9, 21, 63
104	21(104-83)	3, 7, 21

  - key length is probably 3 or 7

2.3.2 Cryptanalysis of Polyalphabetic Substitutions

- Index of Coincidence
  - a measure of the variation between frequencies in a distribution
- Notation
  - Prob_λ : probability of an λ, where $\lambda \in \{a, \ldots, z\}$

$$\text{Prob}_a + \text{Prob}_b + \cdots + \text{Prob}_z = \sum_{\lambda=a}^{\lambda=z} \text{Prob}_\lambda = 1$$

  - Freq_λ : frequency of an λ among n ciphertext letters
  - RFreq_λ : relative frequency of λ
    - $\text{RFreq}_\lambda = \text{Freq}_\lambda / n$
    - $n \to \infty$, $\text{RFreq}_\lambda \to \text{Prob}_\lambda$

2.3.2 Cryptanalysis of Polyalphabetic Substitutions

- Index of Coincidence
  - measure of roughness(variance)

$$\text{var} = \sum_{\lambda=a}^{\lambda=z} (\text{Prob}_\lambda - \frac{1}{26})^2 = \{\sum_{\lambda=a}^{\lambda=z} \text{Prob}_\lambda{}^2\} - \frac{1}{26}$$

$$\text{IC} = \sum_{\lambda=a}^{\lambda=z} \frac{\text{Freq}_\lambda \times (\text{Freq}_\lambda - 1)}{n \times (n-1)} \approx \sum_{\lambda=a}^{\lambda=z} \text{Prob}_\lambda{}^2 = \text{var} + \frac{1}{26}$$

    - flat distribution : IC = 0.0384
    - Number of Enciphering Alphabets vs. IC

Alphabets	1	2	3	4	5	10	Large
IC	.068	.052	.047	.044	.044	.041	.038

2.3.3 Concluding Remarks on Polyalphabetic Ciphers

- Steps in analyzing a polyalphabetic cipher
  1. Use the Kasiski method to predict likely numbers of enciphering alphabets. If no numbers emerge fairly regularly, the encryption is probably not simply a polyalphabetic substitution
  2. Compute the index of coincidence to validate the predictions from step 1
  3. Where steps 1 and 2 indicate a promising value, separate the ciphertext into appropriate subsets and independently compute the index of coincidence of each subset

2.3.4 The Perfect Substitution Cipher

- One-time pad
  - A large nonrepeating set of keys is written on sheet of paper, glued together into a pad
  - Algorithm
    - keys are 20 characters long
    - plain text is 300 characters long
    - encryption
      - sender encrypt the plaintext using the table, like Vigenère Tableau, with 15 pages of keys
    - decryption
      - use the same pad identical to the sender

2.3.4 The Perfect Substitution Cipher

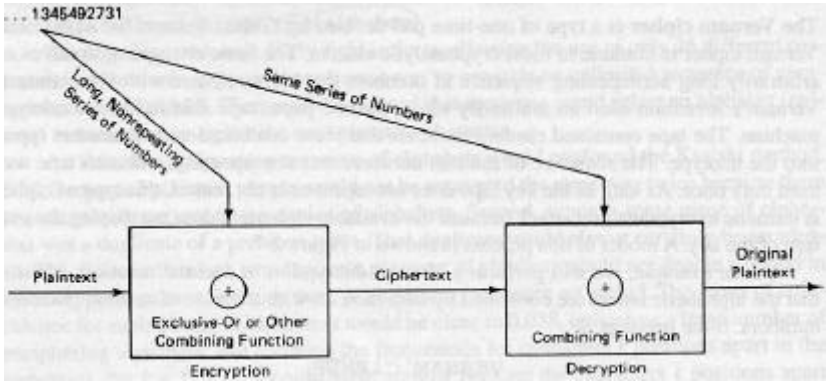
- Advantage of one-time pad
  - perfectly secure
  - ciphertext does not reveal any information of the corresponding plaintext
- Problems
  - the need for absolute synchronization between sender and receiver
  - the need for an unlimited number of keys

2.3.4 The Perfect Substitution Cipher

- Random Number Generator
  - A close approximation of a one-time pad for use on computers is a random number generator.
    - Computer random numbers are not random
    - they really form a sequence with a very long period.
    - A generator with a long period can be acceptable for a limited amount of time or plaintext

2.3.4 The Perfect Substitution Cipher

- The Vernam Cipher
  - a type of one-time pad
  - algorithm
    - $c_j = p_j + k_j \bmod 26$
- The Binary Vernam Cipher
  - algorithm
    - p_j and k_j are 1-bit
    - $c_j = p_j + k_j \bmod 2$
      $= p_j$ XOR k_j



2.3.4 The Perfect Substitution Cipher

- Linear congruential random number generator(LCRNG)
  - $r_{i+1} = (a \times r_i + b) \bmod n$
    - r_0 : seed
    - a, b, n: constant
  - properties
    - if r_0 and a are relatively prime to n, each number between *0*, and *n-1* will be generated before the sequence repeats

2.3.4 The Perfect Substitution Cipher

- Cracking LCRNG
  - Assumption
    - $c_j = p_j + r_j \bmod n$
    - $(p_0, p_1, p_2, p_3) = (M(12), E(4), M(12), O(14))$
  - Interceptor can find a, b, n, and r_0
    - solve the following equations

$$r_1 = a \times r_0 + b \bmod n$$
$$r_2 = a \times r_1 + b \bmod n$$
$$r_3 = a \times r_2 + b \bmod n$$

$$r_0 = c_0 - 12 \bmod n$$
$$r_1 = c_1 - 4 \bmod n$$
$$r_2 = c_2 - 12 \bmod n$$
$$r_3 = c_3 - 14 \bmod n$$

2.3.4 The Perfect Substitution Cipher

- Long Sequences form Books
  - Example: telephone book
    - the sender and the receiver agree to an identical telephone book
    - both of them agree to start at page 35
    - use two middle digits(ddd-DDdd) of each phone number mod 26 as a key letter for a polyalphabetic substitution cipher using preagreed form of Vigenère tableau
  - CAUTIONS
    - key text must be evenly distributed

2.3.5 Summary of Substitutions

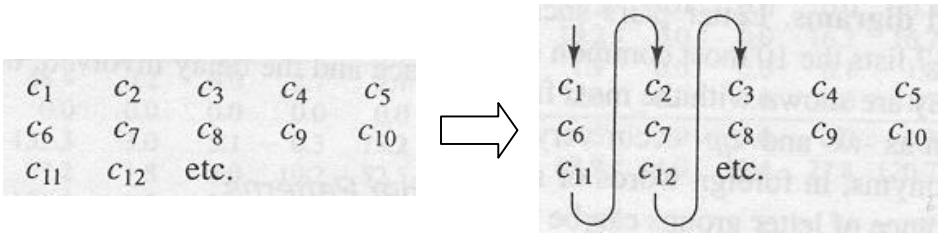
- Substitutions are effective cryptographic devices
  - the basis of many cryptographic algorithms used for diplomatic communication through the first half of the previous century
- Cryptanalystic tools
  - frequency distribution
  - index of coincidence
  - consideration of highly likely letters and probable words
  - repeated pattern analysis and the Kasiski approach
  - persistence, organization, ingenuity, and lock

2.4 Transpositions(Permutations)

- Transposition
  - an encryption in which the letters of the message are rearranged
  - permutation
    - a transposition is a rearrangement of the symbols of a message
- Substitution vs. Transposition
  - the goal of a substitution: confusion
  - the goal of a transposition: diffusion

2.4.1 Columnar Transpositions

- Columnar Transposition
  - a rearrangement of the characters of the plaintext into columns



2.4.1 Columnar Transpositions

- Example

2.4.1 Columnar Transpositions

- Encipherment/Decipherment Complexity
  - Substitution vs. Transposition

	Substitution	Transposition
Time Complexity	O(n)	O(n)
Space Complexity	O(c)	O(n)
Delay	O(c)	O(n)

  - Transposition is not especially appropriate for long messages

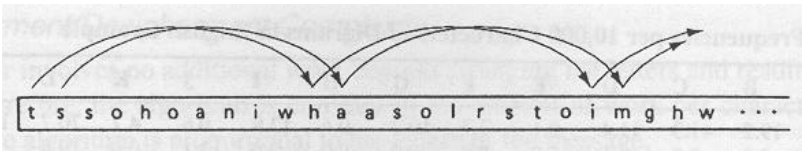
2.4.1 Columnar Transpositions

- Diagrams, Trigrams, and Other Patterns
  - there are characteristic patterns of adjacent letters
  - most common Digrams and Trigrams

Digrams	Trigrams
EN	ENT
RE	ION
ER	AND
NT	ING
TH	IVE
ON	TIO
IN	FOR
TF	OUR
AN	THI
OR	ONE

2.4.1 Columnar Transpositions

- Cryptanalysis by Digram Analysis
  - Positions of adjacent letters in Ciphertext



2.4.1 Columnar Transpositions

- Double Transposition Algorithm
  - Product Cipher
    - $C=E_2(E_1(P))$
  - the first transposition displaces adjacent letters
  - the second transposition breaks up the adjacency of short series of letters that happened to appear in adjacent columns of the first transposition

T	H	(I	S)	I
S	A	(M	E)	S
S	A	(G	E)	T
O	S	(H	O)	W
H	O	(W	A)	C
O	L	(U	M)	N
A	R	(T	R)	A
N	S	(P	O)	S
I	T	(I	O)	N
W	O	(R	K)	S

T	S	S	O	H	O	A
N	I	W	H	A	A	S
O	L	R	S	T	O	(I
(M	(G	(H	(W	(U	(T	(P
(I	(R	S)	E)	E)	O)	A)
M)	R)	O)	O)	K)	I	S
T	W	C	N	A	S	N
S	X	X	X	X	X	X

2.4.1 Columnar Transpositions

- Cryptanalysis
  - There is a functional relationship between plaintext and ciphertext character positions
    - 10×5 $E(i) = 10 * ([(i-1) \bmod 5] + (i-1)/5 + 1)$

      $c_{10*([(i-1) \bmod 5]+(i-1)/5+1)} = p_i$
    - 8×7 $E(i) = 8 * ([(i-1) \bmod 7] + (i-1)/7 + 1)$

      $c_{8*([(i-1) \bmod 7]+(i-1)/7+1)} = p_i$

2.4.2 Generalized Transpositions

- Transpositions are somewhat difficult to analyze because they move plaintext to unpredictable places in the ciphertext
- In fact, any permuting algorithm can be used, as long as it is reversible.

2.5 Fractionated Morse

- Morse Code
  - a means of representing letters as sequences of dots and dashed
  - used with telegraphs, flashing fights, and semaphore flags

A	.-	H	....	O	---	U	..-
B	-...	I	..	P	.--.	V	...-
C	-.-.	J	.---	Q	--.-	W	.--
D	-..	K	-.-	R	.-.	X	-..-
E	.	L	.-..	S	...	Y	-.--
F	..-.	M	--	T	-	Z	--..
G	--.	N	-.				

2.5.2 Morse Code for Encryption

- Morse Code for Encryption
  - Morse code is a three-symbol coding scheme using the symbols “dash”, “dot”, and “separator”
  - there are 26 letters in English alphabet
  - $3^3=27=26+1$, English letter ≈ {-, ., |}^{3} - {|||}
  - if keyword is wovenflax
    - English letters associated with Morse Code symbols

w	...	a	.|-	i	--|	r	|.|
o	..-	x	.||	j	-|.	s	|-.
v	..|	b	-..	k	-|-	t	|--
e	.-.	c	-.-	m	-||	u	|-|
n	.--	d	-.|	p	|..	y	||.
f	.-|	g	--.	q	|.-	z	||-
l	.|.	h	---				

2.5.2 Morse Code for Encryption

- Algorithm
  - step 1. the English plaintext is converted to Morse code, using separator
  - step 2. the Morse code message is divided into blocks of three symbols
  - step 3. each block is encoded as the letter corresponding to that three symbol pattern

2.5.2 Morse Code for Encryption

- Example(key = wovenflax)
  - plaintext `FRACTIONATED MORSE`
  - step 1

```
..-.|.-.|.-|-.-.|-|..---|-.|.-|
-|.|-..||--|---|.-.|...|.
```

  - step 2

```
..- .|. -.| .-| -.- .|- |.. |-- -|- .|.
-|- |.| -.. ||- -|- --| .-. |.. .|.
```

  - step 3

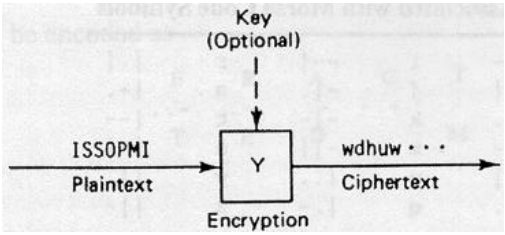
```
..-  .|.  -.|  .-|  -.-  .|-  |..  |--  -|-  .|.
 o    l    d    f    c    a    p    t    k    l

-|-  |.|  -..  ||-  -|-  --|  .-.  |..  .|.
 k    r    b    z    k    i    e    p    l
```

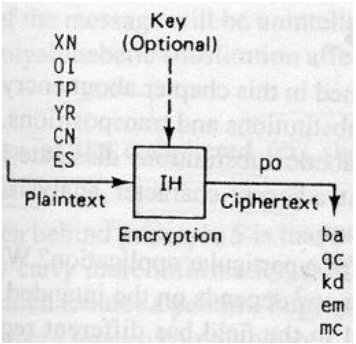
2.5.3 Cryptanalysis of Fractionated Morse

- Cryptanalysis of Fractionated Morse
  - This cipher has some of the same characteristics of the polyalphabetic cipher
    - a single letter is encoded in several different ways
    - unlike the polyalphabetic encodings, however, there is no necessary fixed repetition of the key between similar representations
  - This encryption is not very secure
    - if the basic algorithm is knows, but not the codeword, frequent ciphertext letters indicate frequent patterns

2.6 Stream and Block Ciphers

- Stream Cipher
  - convert one symbol of plaintext immediately into a symbol of ciphertext
  - figure
- Block Cipher
  - convert a group of plaintext symbols as one block
  - figure





2.6.1 Stream Ciphers

- Advantage
  - Speed of Transposition
  - Low error propagation
- Disadvantage
  - Low diffusion
    - subject to the tools such as frequency distribution, digram analysis, the index of coincidence, and the Kasiski method
  - Susceptibility to malicious insertions and modifications
    - integrity

2.6.2 Block Ciphers

- Example
  - transposition(columnar transposition and other transposition)
  - fractional Morse
- Disadvantages
  - the strengths of stream cipher
    - speed
    - error propagation

2.6.2 Block Ciphers

- Advantages
  - Diffusion
    - information from the plaintext if diffused into several ciphertext symbols
    - one ciphertext block may depend on several plaintext letters
  - Immunity to insertions: integrity
    - it is impossible to insert a single symbol into one block
    - the length of the block would then be incorrect, and the decipherment would quickly reveal the insertion
    - active interceptor cannot simply cur one ciphertext letter out of a message and paste a new one in to change an account, a time, a date, or a name of a message

2.7 Characteristics of Good Ciphers

- What does it mean for a cipher to be good?
  - The meaning of good depends on the intended use of the cipher
  - A cipher to be used by military personnel in the field has different requirements from one that will be used in a secure installation with substantial computer support

2.7.1 Shannon Characteristics

- Characteristics of Good Ciphers by Claude Shannon
  - The amount of secrecy needed should determine the amount of labor appropriate for the encryption and decryption
  - The set of keys and the enciphering algorithm should be free from complexity
  - The implementation of the process should be as simple as possible
  - Errors in ciphering should not propagate and cause corruption of further information in the message
  - The size of the enciphered text should be no larger then the text of the original message

2.7.2 Confusion and Diffusion

- Confusion
  - The interceptor should not be able to predict what changing one character in the plaintext will do to the ciphertext
  - An algorithm providing good confusion will have a complex functional relationship between the plaintext.key pair and the ciphertext
- Diffusion
  - The cipher should also spread the information from the plaintext over the entire ciphertext
  - Good diffusion means that the interceptor needs access to much ciphertext to infer the algorithm
  - The substitution and permutation ciphers do not provide any diffusion

2.7.3 Information Theoretic Tests

- Secure Cryptographic System
  - one where an interceptor cannot recover a plaintext message from its ciphertext
  - One-time pad(perfect transformation)
    - immune to cryptanalystic attack
    - fails if one assumes access to the pad through theft, collusion, coercion, or bribery between the interceptor and a sender or receiver

2.7.3 Information Theoretic Tests

- Notation
  - C = E(P)
  - h is the cryptanalyst;s suspectioed decipherment
  - h(C): a set of possible plaintext
    - h(C)={Poss1, Poss2,…}
- An encryption scheme is effectively secure
  - if the probability that h(C) is P is arbitrarily small; that is

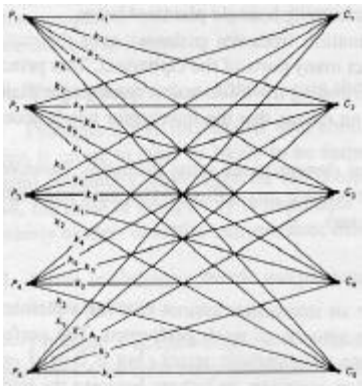
$$\text{Prob}(h(C) = P) < e$$

for arbitrarily small value ε

2.7.3 Information Theoretic Tests

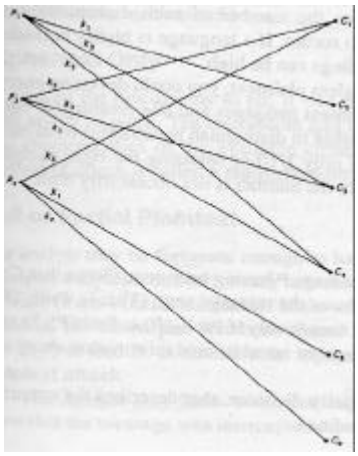
- Perfect Secrecy

$$\text{Prob}_{C_1}(h(C_1) = P) = \text{Prob}(P)$$
$$= \text{Prob}(h(C_1) = P)$$



- Imperfect Secrecy

$$\text{Prob}_{C_1}(h(C_1) = P_1) = 0$$



2.7.3 Information Theoretic Tests

- Redundancy
  - Languages are inherently redundant
  - absolute rate of the language with k letters in the alphabet

$$\lceil \log_2(k) \rceil$$

    - English: $\lceil \log_2(26) \rceil = 5$
    - the number of possible n-letter message: 2^{An}
  - The rate of language = R
    - the number of possible meaningful message: 2^{Rn}
  - Redundancy of the language = D
    - D = A - R

2.7.4 Unicity Distance

- Shannon's Information Theory
  - Unicity distance describes the amount of ciphertext needed in order to break a cipher

$$H_C(P) = \sum_P Prob_C(P) * \log_2\left(\frac{1}{Prob_C(P)}\right)$$

  - The unicity distance is the smallest message length n for which $H_C(P)$ is close to 0
    - unicity distance is the measure the amount of uncertainty in an encryption system

2.7.4 Unicity Distance

- Example
  - if encryption has $2^{H(P)}$ keys, the unicity distance is

$$N = \frac{H(P)}{D}$$

    - D is the redundancy of the language
  - The probability of getting a spurious decryption of a ciphertext is *p=(q-1)*, where

$$q = \frac{2^{Rn}}{2^{An}} = 2^{(R-A)n} = 2^{-Dn}$$

    - q is the likelihood of getting a string that represents nothing

2.8 What the Cryptanalyst Has to work With

- Four possible situations
  - Ciphertext only - ciphertext only attack
  - full or partial plaintext
    - known plaintext attack
    - probable plaintext attack
  - ciphertext of any plaintext - chosen plaintext attack
  - algorithm and ciphertext - chosen ciphertext attack

2.9 Summary

- Encipherment
  - substitution and transposition or permutation
- Cryptanalysis
  - frequency distribution, digram study, index of coincidence, searching for repeated patterns, study of probable letters
  - Five classic cryptanalytic attacks
    - ciphertext only attack, known plaintext attack, probable plaintext attack, chosen plaintext attack, and chosen ciphertext attack
- Formal material
  - redundancy and unicity distance

Further Reading

- History
  - David Kahn, “THE CODEBREAKERS: The Story of Secret Writing,” Macmillan publishing Co., Inc., 1967