Cryptography and Security

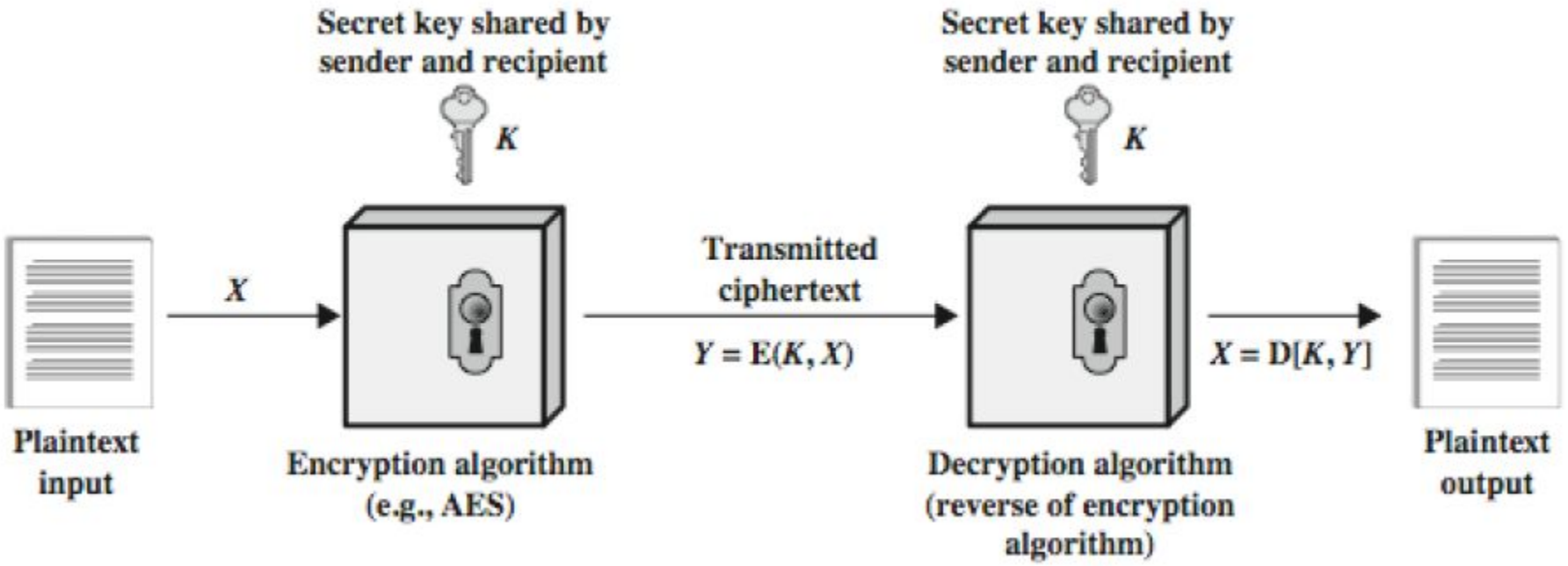
Lecture 2

Classical Encryption Techniques

Some Basic Terminology

- **plaintext**-original message
- **ciphertext**-coded message
- **cipher**-algorithm for transforming plaintext to ciphertext
- **key**-info used in cipher known only to sender/receiver
- **encipher (encrypt)**-converting plaintext to ciphertext
- **decipher (decrypt)**-recovering ciphertext from plaintext
- **cryptography**-study of encryption principles/methods
- **cryptanalysis (codebreaking)**-study of principles/ methods of deciphering ciphertext *without* knowing key
- **cryptology**-field of both cryptography and cryptanalysis

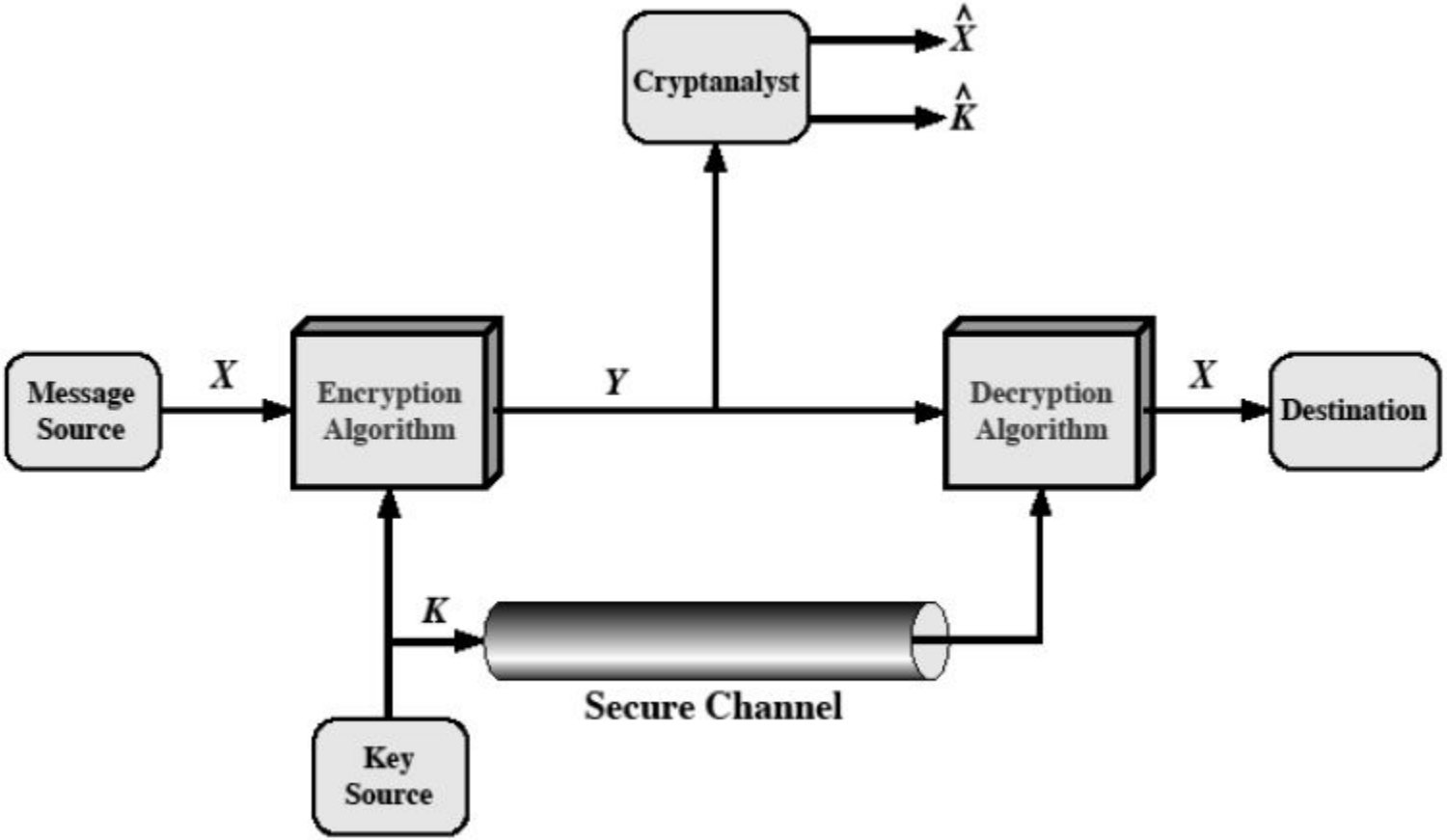
Symmetric Cipher Model



Requirements for Secure Use of Symmetric Key Encryption

1. **Need strong encryption algorithm.**
    - An opponent cannot decrypt a ciphertext if s/he has knowledge of the algorithm, access to some ciphertexts with associated plaintexts without explicit knowledge of key.
    - Algorithms are known.
2. **Maintain the secrecy of the key.**
    - Need secure channel to distribute key.

Model of Symmetric Cryptosystem



Characteristics of Cryptographic System

1. Type of encryption operations used
    - substitution
    - transposition
    - product
2. Number of keys used
    - single-key/ secret-key/symmetric key/ conventional encryption.
    - two-key/ public-key/asymmetric key encryption
3. way in which plaintext is processed
    - block
    - stream

Attacks on Conventional Encryption Scheme

- **Cryptanalysis**
  - Exploits the nature of algorithm based on some knowledge of plaintext or some plaintext-ciphertext pair.
- **Brute-force Attack**
  - Tries every possible keys on a piece of ciphertext.
  - On average, half of all possible keys must be tried.

Types of Cryptanalytic Attacks

Type of Attack	Known to Cryptanalyst
Ciphertext Only	• Encryption algorithm • Ciphertext
Known Plaintext	• Encryption algorithm • Ciphertext • One or more plaintext–ciphertext pairs formed with the secret key
Chosen Plaintext	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen Ciphertext	• Encryption algorithm • Ciphertext • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen Text	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

More Definitions

- **unconditional security**

– no matter how much computer power or time is available, the cipher cannot be broken as the ciphertext provides ***insufficient information*** to uniquely determine the corresponding plaintext.

- **computational security**

- given limited computing resources the cipher cannot be broken.

Brute-Force Attack

- always possible to simply try every key
- most basic attack, proportional to key size
- assume able to know / recognise plaintext

Table 2.2 Average Time Required for Exhaustive Key Search

Key Size (bits)	Number of Alternative Keys	Time Required at 1 Decryption/μs	Time Required at 10^6 Decryptions/μs
32	$2^{32} = 4.3 \times 10^9$	$2^{31}\mu s = 35.8$ minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	$2^{55}\mu s = 1142$ years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	$2^{127}\mu s = 5.4 \times 10^{24}$ years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	$2^{167}\mu s = 5.9 \times 10^{36}$ years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	$2 \times 10^{26}\mu s = 6.4 \times 10^{12}$ years	6.4×10^6 years

Substitution Techniques

- where letters of plaintext are replaced by other letters or by numbers or symbols
- or if plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns

Caesar Cipher

- earliest known substitution cipher
- first attested use in military affairs
- replaces each letter by 3rd letter
- example:

```
plain:  meet me after the toga party
cipher: PHHW PH DIWHU WKH WRJD SDUWB
```

Caesar Cipher

Can define transformation as:

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
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

- mathematically give each letter a number

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	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Caesar cipher is defined as:

$$c = E(k, p) = (p + k) \bmod (26)$$

$$p = D(k, c) = (c - k) \bmod (26)$$

Brute-Force Attack on Caesar Cipher

- Brute-force attack on Caesar cipher is possible:
  - The encryption and decryption algorithm are known.
  - For each ciphertext component, only 25 keys to try.
  - The language of plaintext is known and easily recognizable.

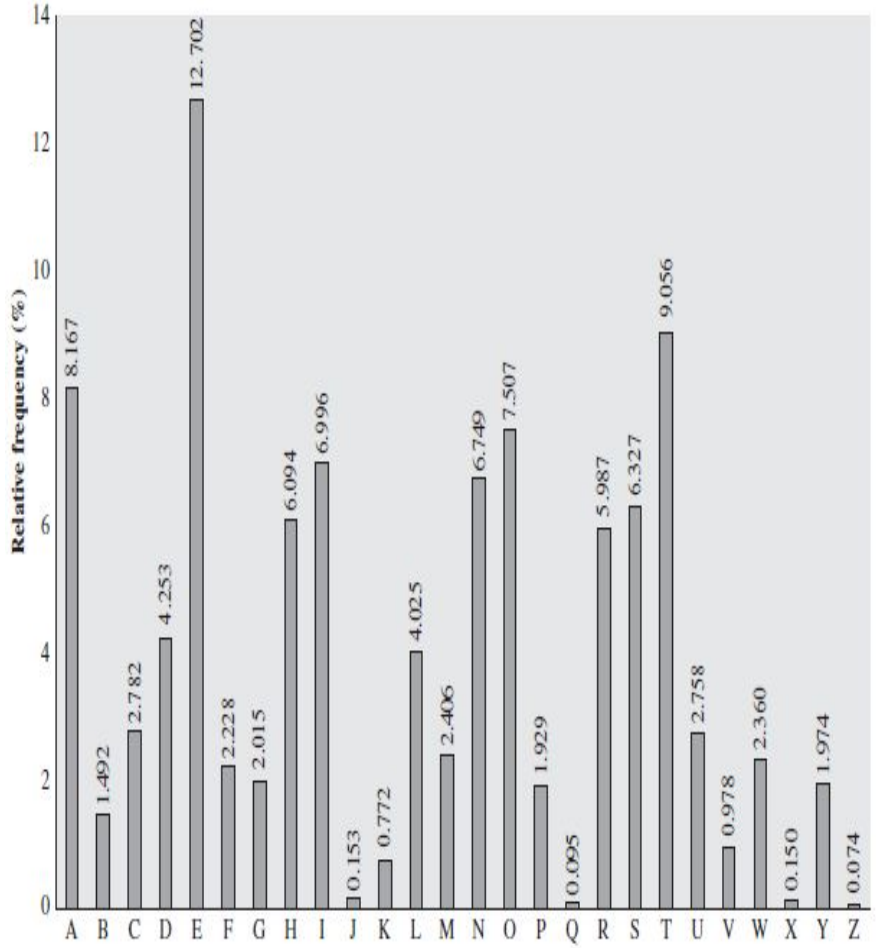
Monoalphabetic Cipher

- A single plain alphabet is mapped to an individual cipher alphabet (per message).
- rather than just shifting the alphabet could shuffle (jumble) the letters arbitrarily

```
plain:  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
cipher: 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
```

- now have a total of 26! = 4 x 10^{26} keys
- with so many keys, might think is secure
- but would be **!!WRONG!!**
- problem is language characteristics

Language Redundancy and Cryptanalysis



- key concept - monoalphabetic substitution ciphers do not change relative letter frequencies.

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

P 13.33	H 5.83	F 3.33	B 1.67	C 0.00
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.00
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.00
U 8.33	V 4.17	T 2.50	I 0.83	N 0.00
O 7.50	X 4.17	A 1.67	J 0.83	R 0.00
M 6.67				

Language Redundancy and Cryptanalysis

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
 t a          e   e te   a that e e a          a
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
   e t    ta t ha e ee   a e   th     t   a
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
 e   e e tat   e    the    t
```

- Finally the decrypted message is

it was disclosed yesterday that several informal but direct contacts have been made with political representatives of the viet cong in moscow

Playfair Cipher

- not even the large number of keys in a Monoalphabetic cipher provides security.
- one approach to improve security was to encrypt multiple letters
- the Playfair Cipher is an example
- Invented by Charles Wheatstone in 1854.

Playfair Cipher

- a 5X5 matrix of letters based on a keyword
- fill in letters of keyword (no duplicates)
- fill rest of matrix with other letters
- eg. using the keyword MONARCHY

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

Playfair Cipher

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

plaintext is encrypted with two letters at a time

1. if a pair is a repeated letter, insert filler like 'X'
2. if both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
3. if both letters fall in the same column, replace each with the letter below it (wrapping to top from bottom)
4. otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair

Security of Playfair Cipher

- security much improved over monoalphabetic
- since have 26 x 26 = 676 digrams
- would need a 676 entry frequency table to analyse (verses 26 for a monoalphabetic)
- and correspondingly more ciphertext
- was widely used for many years
  - eg.. by US & British millitary in WW1
- it **can** be broken, given a few hundred letters
- since still has much of plaintext structure

Security of Ciphers

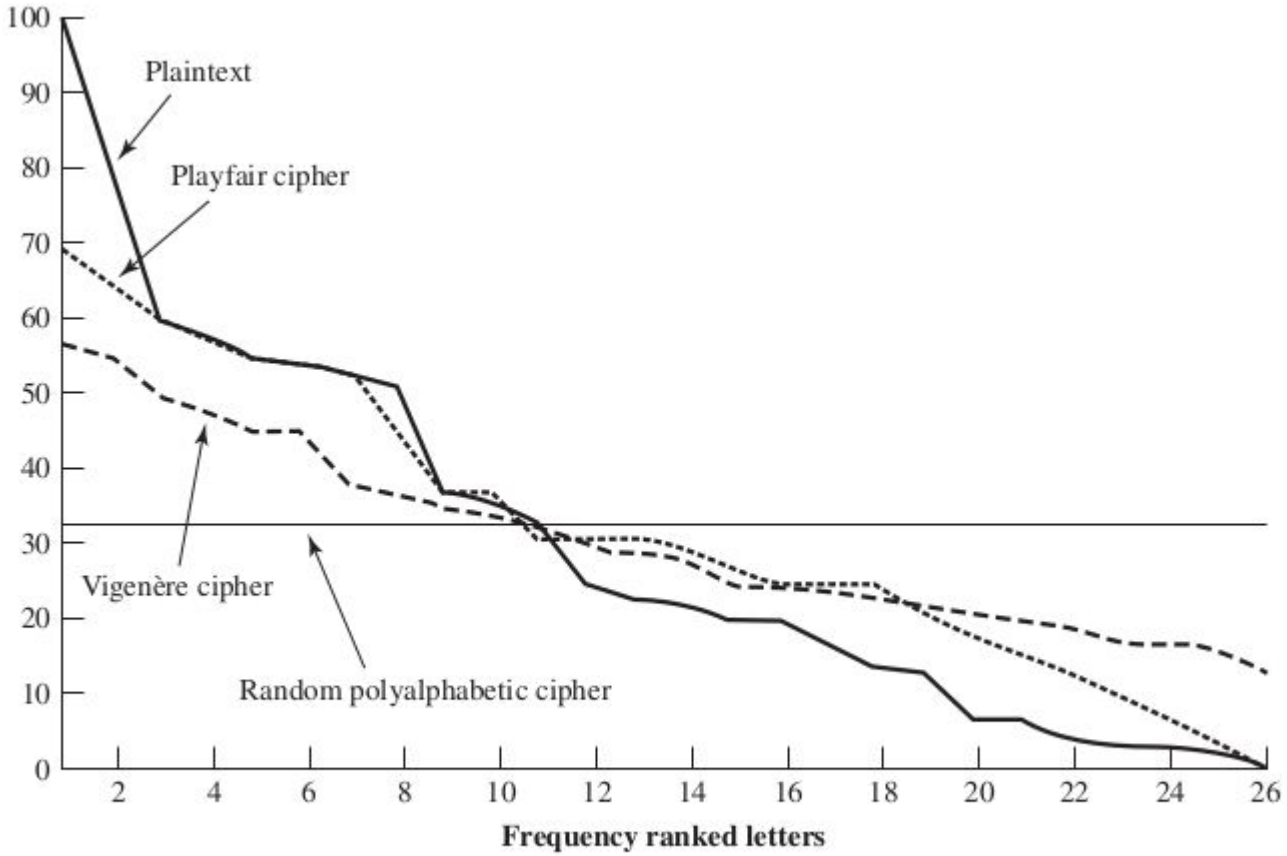


Figure 2.6 Relative Frequency of Occurrence of Letters

Polyalphabetic Ciphers

- improve security using multiple cipher alphabets
- make cryptanalysis harder with more alphabets to guess and flatter frequency distribution
- use a key to select which alphabet is used for each letter of the message
- Common features:
  - A set of related monoalphabetic substitution rules is used.
  - A key determines which particular rule is used for a given transformation.

Vigenère Cipher

- Simplest form of polyalphabetic cipher.
- Plaintext $P=p_0p_1p_2\ldots\ldots p_{n-1}$

  Key $K = k_0k_1k_2\ldots\ldots k_{m-1}$, $m<n$

  Ciphertext $C =C_0C_1C_2\ldots\ldots C_{n-1}$ is expressed as follows:

$$C = C_0, C_1, C_2, \ldots, C_{n-1} = \mathrm{E}(K, P) = \mathrm{E}[(k_0, k_1, k_2, \ldots, k_{m-1}), (p_0, p_1, p_2, \ldots, p_{n-1})]$$
$$= (p_0 + k_0) \bmod 26, (p_1 + k_1) \bmod 26, \ldots, (p_{m-1} + k_{m-1}) \bmod 26,$$
$$(p_m + k_0) \bmod 26, (p_{m+1} + k_1) \bmod 26, \ldots, (p_{2m-1} + k_{m-1}) \bmod 26, \ldots$$

- The general equation for encryption

$$C_i = (p_i + k_{i \bmod m}) \bmod 26$$

- The general equation for decryption

$$p_i = (C_i - k_{i \bmod m}) \bmod 26$$

Vigenère Cipher

- Example:

```
key:        deceptivedeceptivedeceptive
plaintext:  wearediscoveredsaveyourself
ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

key	3	4	2	4	15	19	8	21	4	3	4	2	4	15
plaintext	22	4	0	17	4	3	8	18	2	14	21	4	17	4
ciphertext	25	8	2	21	19	22	16	13	6	17	25	6	21	19

key	19	8	21	4	3	4	2	4	15	19	8	21	4
plaintext	3	18	0	21	4	24	14	20	17	18	4	11	5
ciphertext	22	0	21	25	7	2	16	24	6	11	12	6	9

Security of Vigenère Ciphers

- have multiple ciphertext letters for each plaintext letter, obscuring letter frequency
- Considerable frequency information exists.

Attacking Vigenère Ciphers

- start with letter frequencies
  - see if they look monoalphabetic or not
- if not, then check for Vigenere cipher.

Kasiski Method:

- method developed by Babbage / Kasiski
- repetitions in ciphertext give clues to period
- of course, could also be random fluke
- Example from previous slide.
- suggests size of 3 or 9
- then attack each monoalphabetic cipher individually

Autokey System

- ideally want a key as long as the message
- Vigenère proposed the **autokey** cipher where keyword is prefixed to message as key
- eg. given key *deceptive*

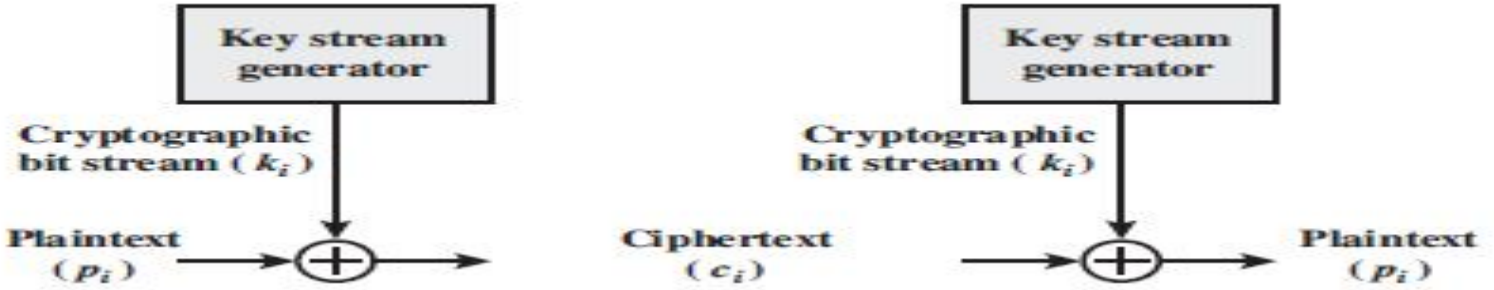
```
key:            deceptivewearediscoveredsav
plaintext:      wearediscoveredsaveyourself
ciphertext:     ZICVTWQNGKZEIIGASXSTSLVVWLA
```

- but can still attack frequency characteristics

Vernam Cipher

- ultimate defense is to use a key as long as the plaintext.
- with no statistical relationship to it
- invented by AT&T engineer Gilbert Vernam in 1918
- originally proposed using a very long but eventually repeating key

Vernam Cipher



- Equation for encryption:

$$c_i = p_i \oplus k_i$$

- Equation for decryption:

$$p_i = c_i \oplus k_i$$

One-Time Pad

- a truly random key as long as the message is used, that makes the cipher unbreakable.
- The key is used only once.
- for **any plaintext** & **any ciphertext** there exists a key mapping one to other
- is unbreakable since ciphertext bears no statistical relationship to the plaintext
- problems in generation & safe distribution of key
- One-time pad has ***perfect secrecy***.

Transposition Ciphers

- known as classical **transposition/permutation** cipher.
- hide the message by rearranging the letter order without altering the actual letters used
- can recognise these since have the same frequency distribution as the original text

Rail Fence cipher

- write message letters out diagonally over a number of rows and then read off cipher row by row.
- Example: Original message

  meet me after the toga party

  m e m a t r h t g p r y
  e t e f e t e o a a t

- The ciphertext is

  MEMATRHTGPRYETEFETEOAAT

Row Transposition Ciphers

- is a more complex transposition
- write letters of message out in rows over a specified number of columns
- then reorder the columns according to some key before reading off the rows

```
Key:          4 3 1 2 5 6 7
Plaintext:    a t t a c k p
              o s t p o n e
              d u n t i l t
              w o a m x y z
Ciphertext:   TTNAAPTMTSUOAODWCOIXKNLYPETZ
```

Product Cipher

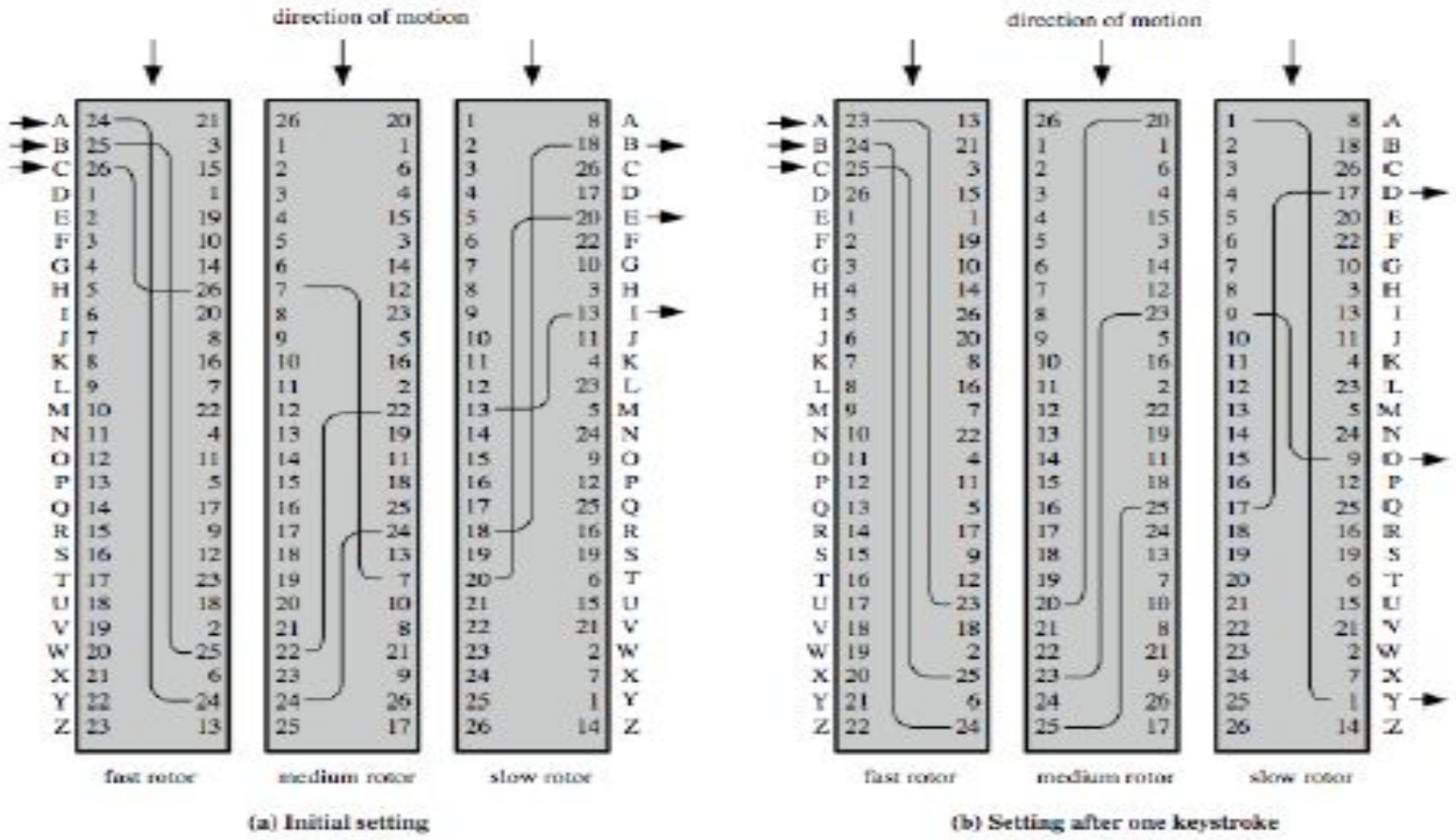
- ciphers using substitutions or transpositions are not secure because of language characteristics
- hence consider using several ciphers in succession to make harder:
  - two substitutions make a more complex substitution
  - two transpositions make more complex transposition
  - but a substitution followed by a transposition makes a new much harder cipher

Rotor Machine

- before modern ciphers, rotor machines were most common complex ciphers in use
- widely used in WW2
  - German Enigma, Allied Hagelin, Japanese Purple
- implemented a very complex, varying substitution cipher
- used a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- with 3 cylinders have 26^3=17576 alphabets

Hagelin Rotor Machine

Rotor Machine Principles



- **Classical Encryption Techniques from the book of William Stallings**