Classical Encryption Techniques

Ref: Cryptography and Network Security by William Stallings

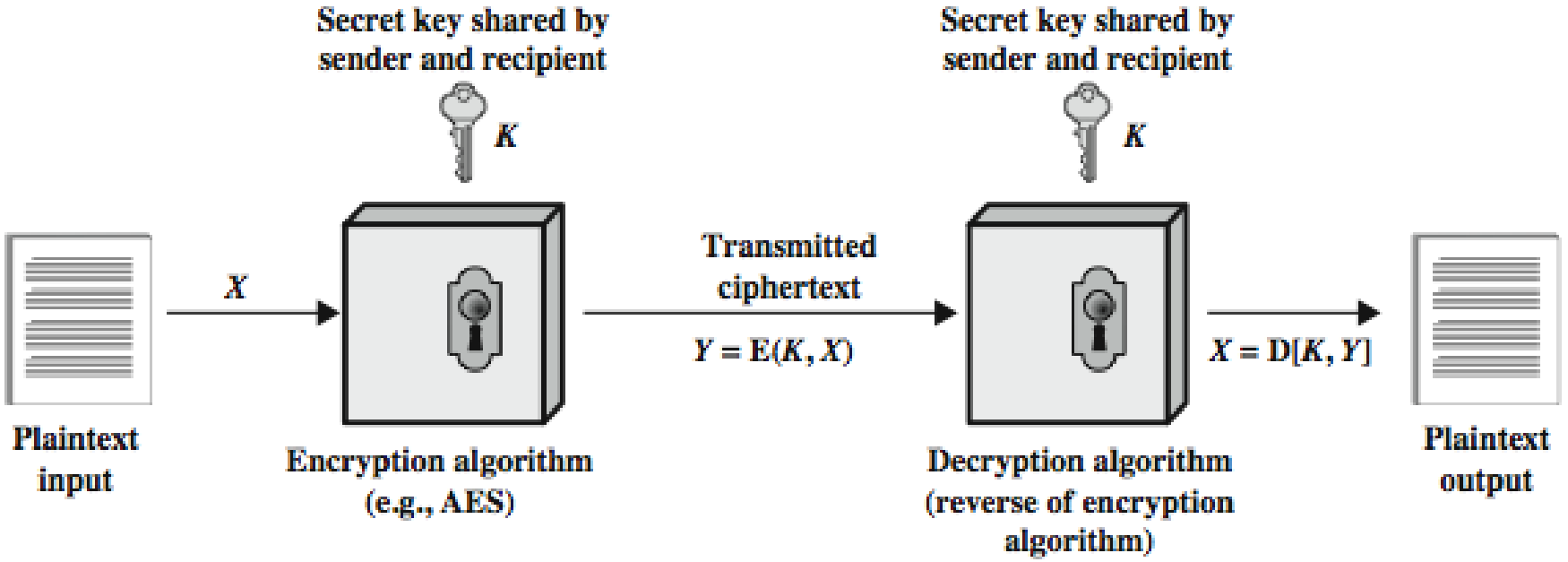
Symmetric Encryption

- Conventional / private-key / single-key
- Sender and recipient share a common key
- All classical encryption algorithms are private-key
- Only type prior to invention of public-key in 1970's

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 methods
- **Cryptanalysis (codebreaking)** - study of principles/ methods of deciphering ciphertext *without* knowing key
- **Cryptology** - field of both cryptography and cryptanalysis

Symmetric Cipher Model



Symmetric Cipher Model

- Plaintext - original message
- Encryption algorithm – performs substitutions/ transformations on plaintext
- Secret key – control exact substitutions/transformations used in encryption algorithm
- Ciphertext - scrambled message
- Decryption algorithm – inverse of encryption algorithm

Requirements

- Requirements for secure use of symmetric encryption:
  - a strong encryption algorithm
  - a secret key known only to sender / receiver
- It can be defined mathematically as:

  $Y = E(K, X)$

  $X = D(K, Y)$
- It is assume that encryption algorithm is known
- A secure channel is needed to distribute key

Cryptography

- It can be characterized by:
  - type of encryption operations used
    - substitution
    - transposition
    - hybrid
  - number of keys used
    - single-key or private
    - two-key or public
  - way in which plaintext is processed
    - block
    - stream

Cryptanalysis

- Objective is to recover the key in use rather then simply to recover the plaintext
- General approaches:
  - **Cryptanalytic attack:** relies on the nature of the algorithm plus perhaps some knowledge of the general characteristics of the plaintext or even some sample plaintext- ciphertext pairs
  - **Brute-force attack:** try every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained
- If either type of attack succeeds in deducing the key, the effect is catastrophic

Classical Ciphers

- The two basic building blocks of all encryption technique are substitution and transposition.
- **Substitution technique:** letters of plaintext are replaced by other letters or by numbers or symbols
  - The core idea is to replace one basic unit (letter/byte) with another.
- **Transposition Technique:** Performing some sort of permutation on the plaintext letters.

Classical Substitution Ciphers

Caesar Cipher

- The earliest known, and the simplest, use of a substitution cipher was by Julius Caesar.
- First attested use in military affairs
- It involves replacing each letter of the alphabet with the letter standing three places further down the alphabet
- example:

  meet me after the toga party

  PHHW PH DIWHU WKH WRJD SDUWB

Caesar Cipher

- Message is transformed 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

- Mathematically Caesar cipher can be defined as:

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

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

Cryptanalysis of Caesar Cipher

- Only have 26 possible ciphers
  - A maps to A,B,..Z
- **brute force search:** simply try each in turn
- given ciphertext, just try all shifts of letters
- eg. break ciphertext "GCUA VQ DTGCM"

Monoalphabetic Cipher

- Rather than just shifting the alphabet, each plaintext letter maps to a different random ciphertext letter.
- It allows arbitrary substitution, where the translation alphabet can be any permutation of the 26 alphabetic characters.
- Key is 26 letters long

  Plain: abcdefghijklmnopqrstuvwxyz

  Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

  Plaintext: ifwewishtoreplaceletters

  Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA

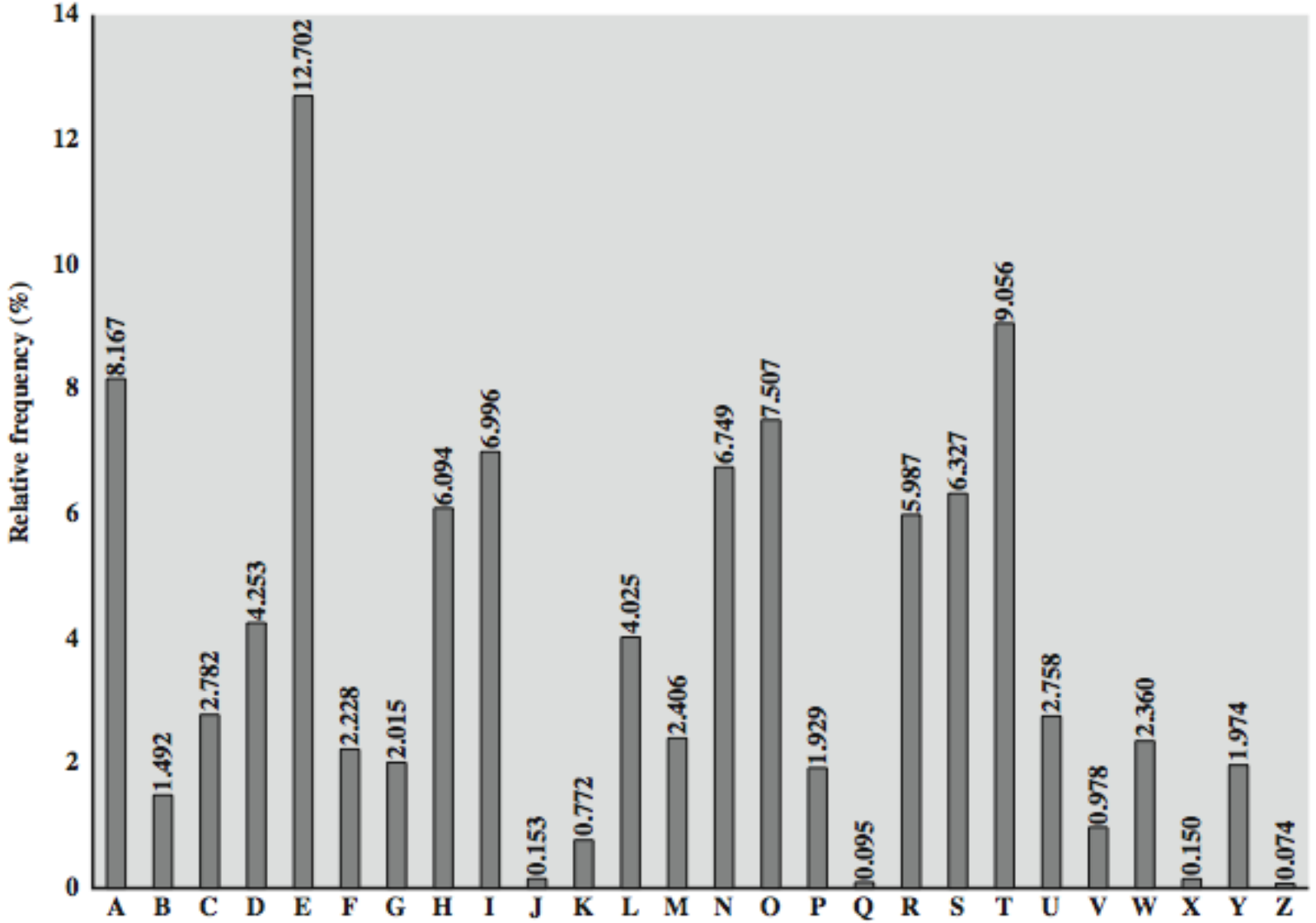
Monoalphabetic Cipher Security

- In this techniques we 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

- Letters are not equally commonly used
- in English E is by far the most common letter
  - followed by T,R,N,I,O,A,S
- other letters like Z,J,K,Q,X are fairly rare
- We have tables of single, double & triple letter frequencies for various languages

English Letter Frequencies



Use in Cryptanalysis

- Key concept - monoalphabetic substitution ciphers do not change relative letter frequencies
- Al-Kindi's book entitled *"Manuscript on Deciphering Cryptographic Messages"* published in the 9th century but only rediscovered in 1987, gave rise to the birth of cryptanalysis,
- Calculate letter frequencies for ciphertext
- Compare counts/plots against known values
  - tables of common double/triple letters help

Example Cryptanalysis

- given ciphertext:
  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSXEPYEPO
  PDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
- Count relative letter frequencies

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				

- Guess P & Z are e and t
- ZW appears three times
- Guess ZW is th and hence ZWP is the

```
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
```

- The final plain text:

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

Playfair Cipher

- Even the large number of keys in a monoalphabetic cipher do not provides security
- One approach to improving security was to encrypt multiple letters
- **Playfair Cipher** is an example
- invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair

Playfair Key Matrix

- Consider a 5X5 matrix of letters based on a keyword
- Fill in letters of keyword
- 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

Encrypting and Decrypting

- Plaintext is encrypted two letters at a time
  1. If a pair is a repeated letter, insert filler like ‘X’, eg. "balloon" encrypts as "ba lx lo on"
  2. If both letters fall in the same row, replace each with letter to right (wrapping back to start from end), eg. “ar" encrypts as "RM"
  3. If both letters fall in the same column, replace each with the letter below it (wrapping to top from bottom), eg. “mu" encrypts to "CM"
  4. Otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair, eg. “hs" encrypts to "BP", and “ea" to "IM" or "JM" .
- Decrypting of course works exactly in reverse. Can see this by working the example pairs shown, backwards.

Hill Cipher

- This encryption algorithm takes m successive plaintext letters and substitutes for them m ciphertext letters,
- The substitution is determined by m linear equations in which each character is assigned a numerical value (a = 0, b = 1, .., z = 25).
- For $m = 3$, the system can be described as

$$c1 = (k11p1 + k21p2 + k31p3) \bmod 26$$

$$c2 = (k12p1 + k22p2 + k32p3) \bmod 26$$

$$c3 = (k13p1 + k23p2 + k33p3) \bmod 26$$

Hill Cipher

- This can be expressed in terms of row vectors and matrices:

$$(c_1\ c_2\ c_3) = (p_1\ p_2\ p_3)\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \bmod 26$$

or **C** = **PK** mod 26, where **C** and **P** are row vectors of length 3 representing the plaintext and ciphertext and **K** is a 3 * 3 matrix representing the encryption key.

Hill Cipher

- Decryption requires using the inverse of the matrix **K**.
- In general terms, the Hill system can be expressed as

$$\mathbf{C} = E(\mathbf{K}, \mathbf{P}) = \mathbf{PK} \bmod 26$$

$$\mathbf{P} = D(\mathbf{K}, \mathbf{C}) = \mathbf{CK}^{-1} \bmod 26 = \mathbf{PKK}^{-1} = \mathbf{P}$$

- Example: Suppose that the plaintext “hillcipher” is encrypted using a 2 * 2 Hill cipher using the K as follows:

$$\begin{pmatrix} 3 & 2 \\ 8 & 5 \end{pmatrix}$$

Polyalphabetic Ciphers

- **Polyalphabetic substitution ciphers:** Use of different monoalphabetic substitutions as one proceeds through the plaintext message.
- **Following rules are applied:**
  - A set of related monoalphabetic substitution rules is used.
  - A key determines which particular rule is chosen for a given transformation.
- Make cryptanalysis harder, since:
  - the frequency distribution is more complex,

Vigenère Cipher

- Simplest polyalphabetic substitution cipher
- Effectively multiple caesar ciphers
- Key is multiple letters long $K = k_1\ k_2\ ...\ k_d$
- i^{th} letter specifies i^{th} alphabet to use
- Repeat from start after d letters in message
- Decryption simply works in reverse

Vigenère Cipher

- Assume a sequence of plaintext letters $P = p_0, p_1, p_2, .., p_{n-1}$ and a key consisting of the sequence of letters $K = k_0, k_1, k_2, c, k_{m-1}$, where typically $m < n$.
- The sequence of ciphertext letters $C = C_0, C_1, C_2, .., C_{n-1}$ is calculated as follows:
  - $C = C_0, C_1, C_2, .., C_{n-1} = E(K, P) = E[(k_0, k_1, k_2, \ldots, k_{m-1}), (p_0, p_1, p_2, .., p_{n-1})] = (p_0 + k_0) \bmod 26, (p_1 + k_1) \bmod 26, \ldots, (p_{m-1} + k_{m-1}) \bmod 26, (p_m + k_0) \bmod 26, (p_{m+1} + k_1) \bmod 26, \ldots, (p_{2m-1} + k_{m-1}) \bmod 26, \ldots$

Example of Vigenère Cipher

- Write the plaintext out
- Write the keyword repeated above it
- Use each key letter as a caesar cipher key
- encrypt the corresponding plaintext letter
- eg using keyword *deceptive*

  key: deceptivedeceptivedeceptive

  plaintext: wearediscoveredsaveyourself

  ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ

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
- The system works on binary data (bits rather than letters)
- The system can be expressed concisely as follows: $c_i = p_i \; XOR \; k_i$
- Originally proposed using a very long but eventually repeating key

Transposition Ciphers

- **Transposition: Encryption is achieved by** performing some sort of permutation on the plaintext letters.
- These hide the message by rearranging the letter order Without altering the actual letters used
- This technique is referred to as a transposition cipher, and form the second basic building block of ciphers

Rail Fence cipher

- Write message letters out diagonally over a number of rows then read off cipher row by row
- eg. write message out as:

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

  e t e f e t e o a a t
- giving ciphertext

  MEMATRHTGPRYETEFETEOAAT

Row Transposition Ciphers

- 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: 4312567

  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

- The transposition cipher can be made significantly more secure by performing more than one stage of transposition

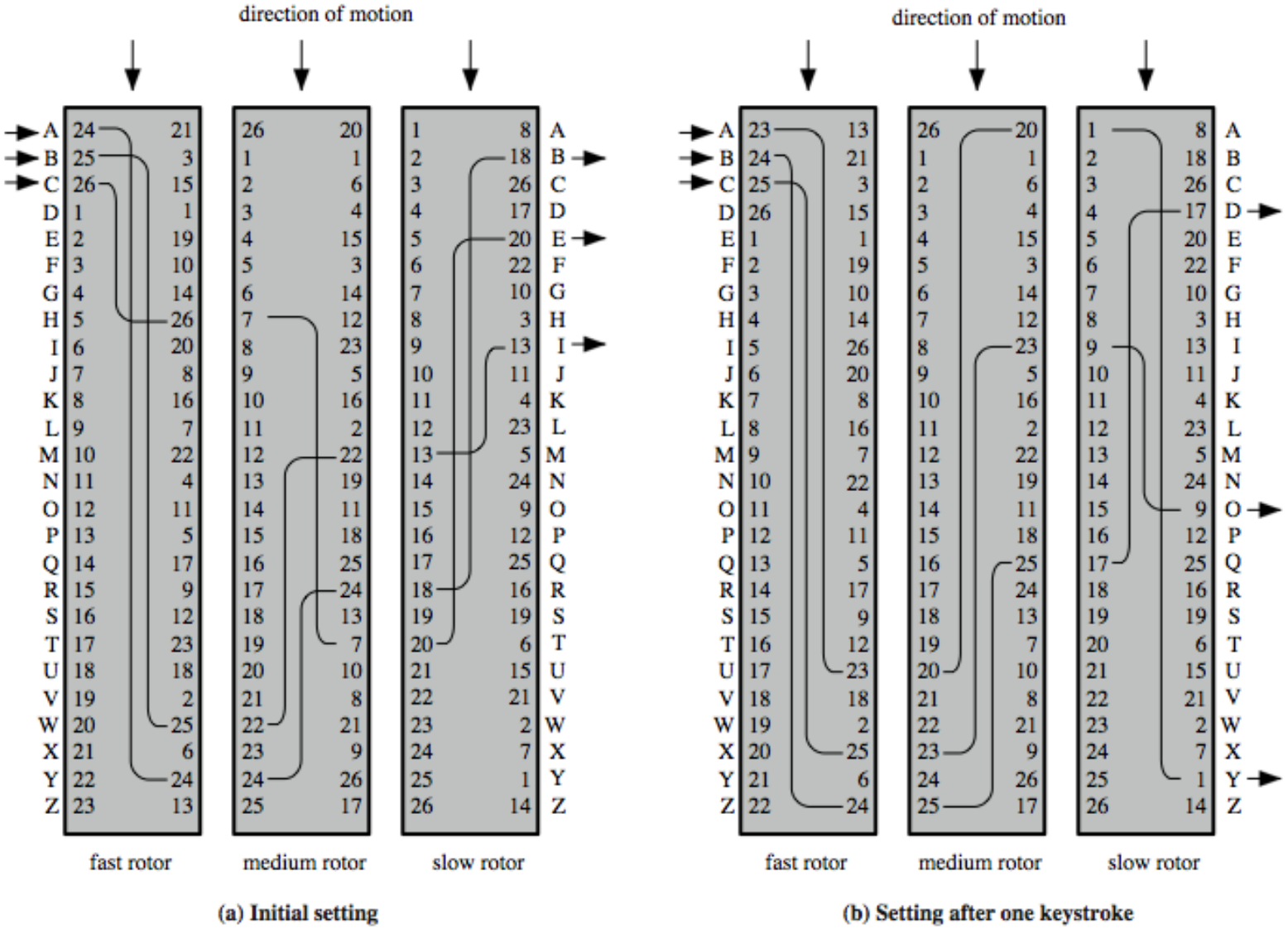
Product Ciphers

- 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
  - substitution followed by a transposition makes a new much harder cipher
- This is bridge from classical to modern ciphers

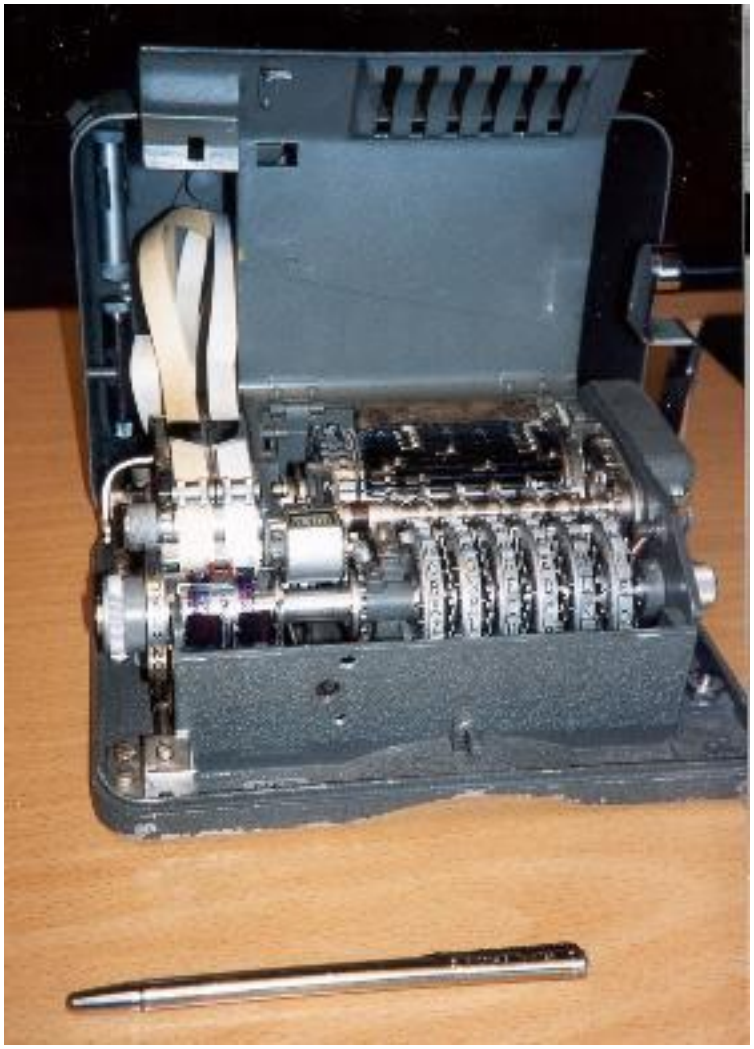
Rotor Machines

- 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

Rotor Machine Principles



Hagelin Rotor Machine



Rotor Machine

- The machine consists of a set of independently rotating cylinders through which electrical pulses can flow.
- Each cylinder has 26 input pins and 26 output pins, with internal wiring that connects each input pin to a unique output pin.
- A single cylinder defines a monoalphabetic substitution.
- If an operator depresses the key for the letter A, an electric signal is applied to the first pin of the first cylinder and flows through the internal connection to the twenty-fifth output pin.
- After each input key is depressed, the cylinder rotates one position, so that the internal connections are shifted accordingly.
- Thus, a different monoalphabetic substitution cipher is defined. After 26 letters of plaintext, the cylinder would be back to the initial position

Steganography

- An alternative to encryption
- **Character marking:** Selected letters of printed or typewritten text are overwritten. The marks are ordinarily not visible unless the paper is held at an angle to bright light.
- **Invisible ink:** A number of substances can be used for writing but leave no visible trace until heat or some chemical is applied to the paper.
- **Pin punctures:** Small pin punctures on selected letters are ordinarily not visible unless the paper is held up in front of a light.
- Concerns/issues:
  - High overhead to hide relatively few info bits

Summary

- We have considered:
  - Classical cipher techniques and terminology
  - Monoalphabetic substitution ciphers
  - Cryptanalysis using letter frequencies
  - Playfair cipher
  - Polyalphabetic ciphers
  - Transposition ciphers
  - Product ciphers and rotor machines
  - Stenography