Classical Encryption Techniques

Dr. Md. Mahbubur Rahman

Outline

1. To define the terms and the concepts of symmetric key ciphers
2. To emphasize the two categories of traditional ciphers: substitution and transposition ciphers
3. To describe the categories of cryptanalysis used to break the symmetric ciphers
4. To introduce the concepts of the stream ciphers and block ciphers
5. To discuss some very dominant ciphers used in the past, such as the Enigma machine

Cryptographic Algorithms

- Symmetric encryption: Used to conceal the contents of blocks or streams of data of any size, including messages, files, encryption keys and passwords.

- Asymmetric encryption: Used to conceal small blocks of data, such as encryption keys and hash function values, which are used in digital signatures.

- Data integrity algorithms: Used to protect blocks of data, such as messages, from alteration.

- Authentication protocols: These are schemes based on the use of cryptographic algorithms designed to authenticate the identity of entities.

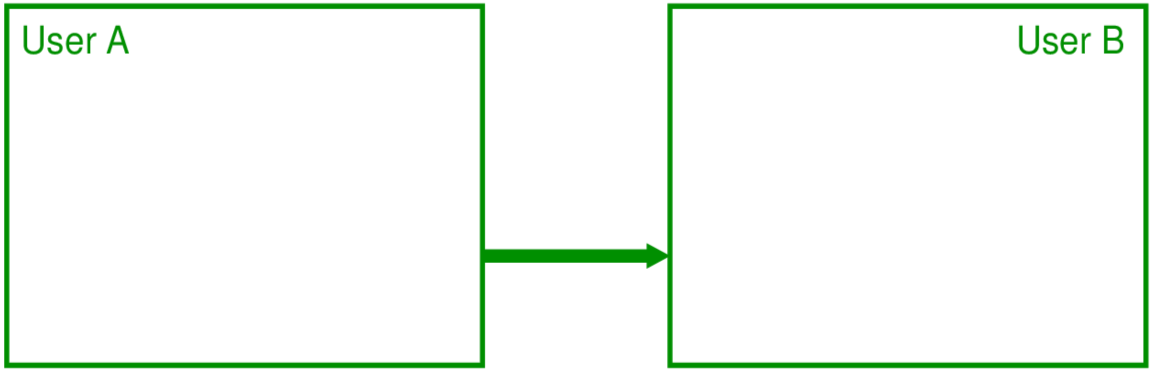
Encryption for Confidentiality

- Aim: assure confidential information not made available to unauthorized individuals (data confidentiality)

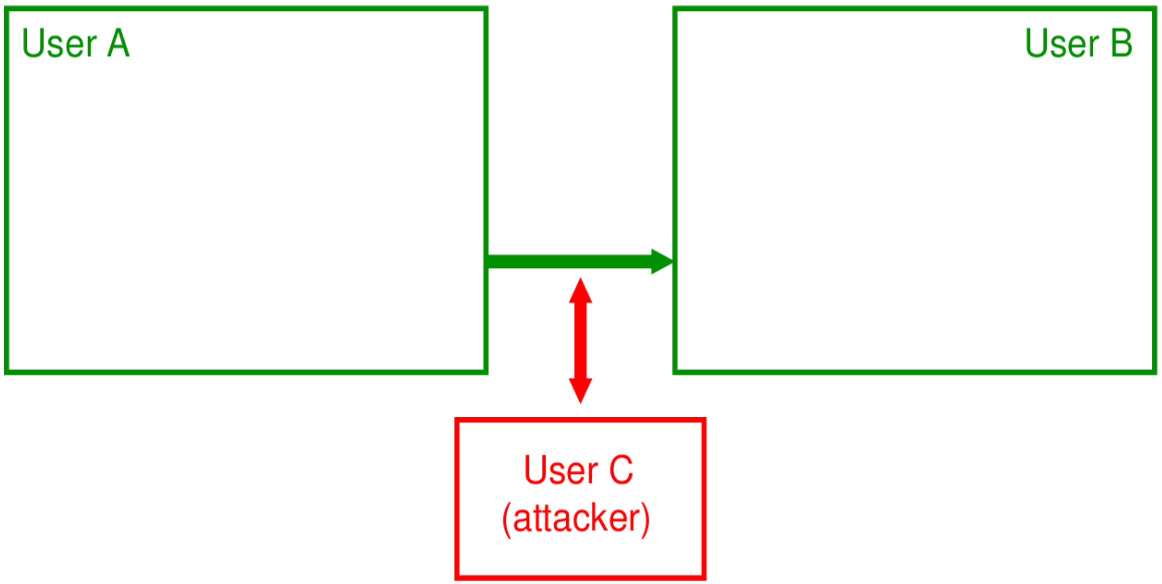
- How: encrypt the original data; anyone can see the encrypted data, but only authorized individuals can decrypt to see the original data

- Used for both sending data across network and storing data on a computer system

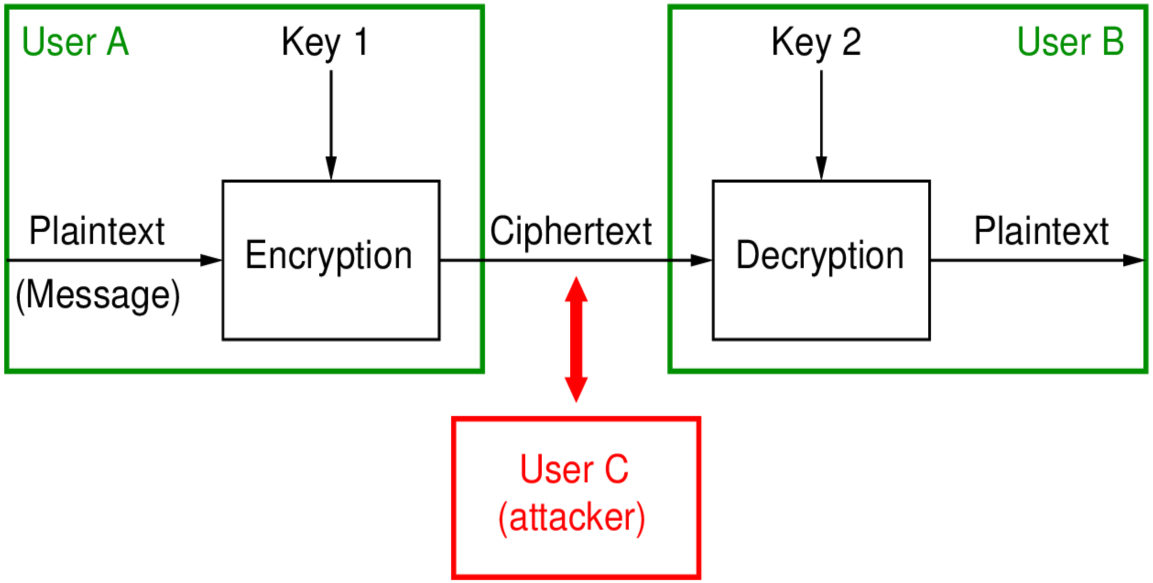
Model of Encryption for Confidentiality



Model of Encryption for Confidentiality



Model of Encryption for Confidentiality



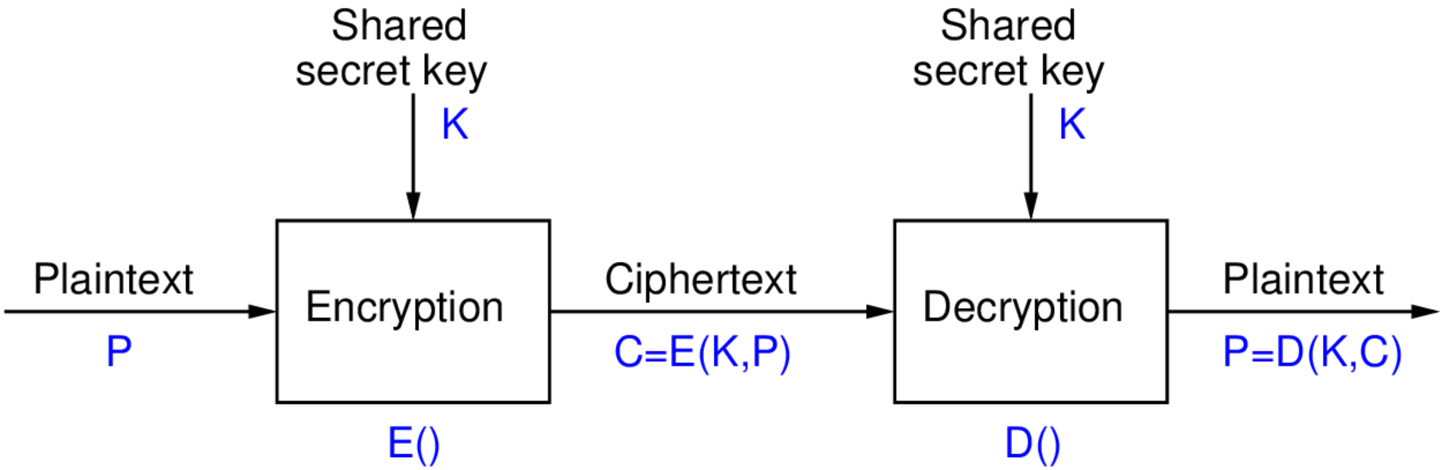
Terminology

Plaintext	original message
Ciphertext	encrypted or coded message
Encryption	convert from plaintext to ciphertext (enciphering)
Decryption	restore the plaintext from ciphertext (deciphering)
Key	information used in cipher known only to sender/receiver
Cipher	a particular algorithm (cryptographic system)
Cryptography	study of algorithms used for encryption
Cryptanalysis	study of techniques for decryption without knowledge of plaintext
Cryptology	areas of cryptography and cryptanalysis

Symmetric Encryption

- or conventional / private-key / single-key
- sender and recipient share a common key
- all classical encryption algorithms are private-key
- was only type prior to invention of public-key in 1970's
- and by far most widely used

Symmetric Cipher Model



Requirements and assumptions

- two requirements for secure use of symmetric encryption:
  - a strong encryption algorithm (cannot decrypt and know key)
  - a secret key known only to sender / receiver

- mathematically have:

  C = E(K, P)

  P = D(K, C)

- Assumptions
  - encryption algorithm is known
  - a secure channel to distribute key

Kerckhoff’s principle

- Although it may appear that a cipher would be more secure if we hide both the encryption/decryption algorithm and the secret key, this is not recommended.

- Based on Kerckhoff’s principle, one should always assume that the adversary , Eve, knows the encryption/decryption algorithm. The resistance of the cipher to attack must be based only on the secrecy of key.

- In other words, guessing the key should be so difficult that there is no need to hide the encryption/decryption algorithm.

Characterizing Cryptographic Systems

Operations used for encryption:

Substitution	replace one element in plaintext with another
Transposition	re-arrange elements
Product systems	multiple stages of substitutions and transpositions

Number of keys used:

Symmetric	sender/receiver use same key (single-key, secret-key, shared-key, conventional)
Public-key	sender/receiver use dierent keys (asymmetric)

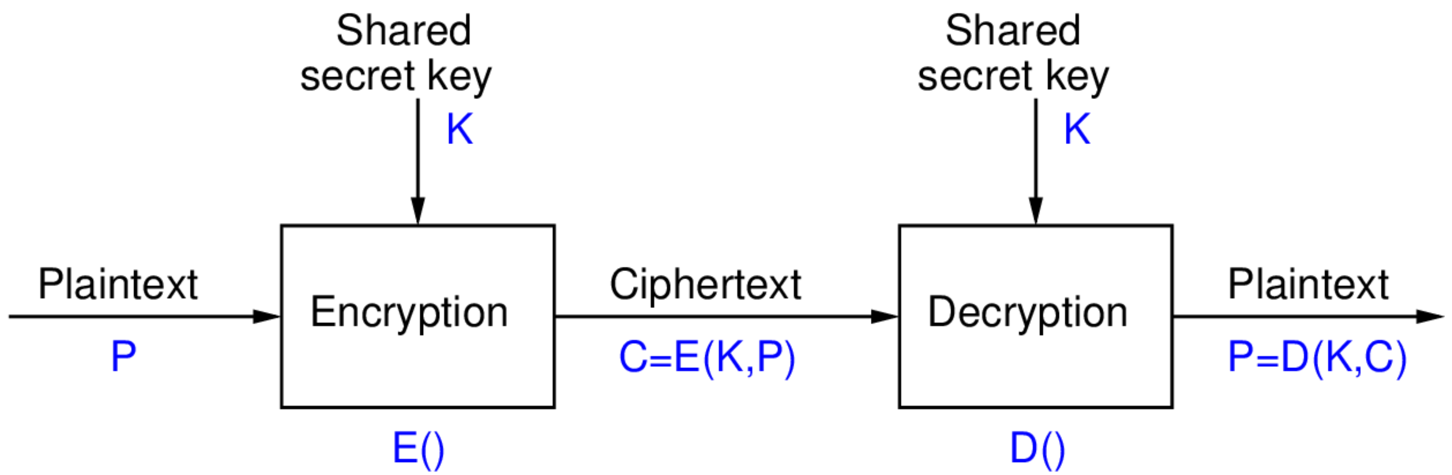
Processing of plaintext:

Block cipher	process one block of elements at a time
Stream cipher	process input elements continuously

Cryptography Classification

- By type of encryption operations used
  - Substitution: Meet Me ⇒ Offu Of
  - Transposition: Meet Me ⇒ Me etM
  - Product
- By number of keys used
  - Single-key or Secret Key
  - Two-key or Public Key
- By the way in which plaintext is processed
  - Block: ABCD EFGH IJKL
  - Stream: ABCDEFGHIJKL

Symmetric Key Encryption for Confidentiality



Requirements

- Strong encryption algorithm: given algorithm, ciphertext and known pairs of (plaintext, ciphertext), attacker should be unable to find plaintext or key
- Shared secret keys: sender and receiver both have shared a secret key; no-one else knows the key

Attacks

Goal of the Attacker

- Discover the plaintext (good)
- Discover the key (better)

Assumed Attacker Knowledge

- Ciphertext (want to decrypt)
- Algorithm (nature of the algorithm) or general idea of the type of plaintext
- Other pairs of (plaintext, ciphertext) using same key (not the plaintext in question)

Attack Methods

Brute-force attack	Try every possible key on ciphertext
Cryptanalysis	Exploit characteristics of algorithm to deduce plaintext or key

Assumption: attacker can recognize correct plaintext

Attacks on Block Ciphers

- Brute Force Attack

  Approach: try all keys in key space

  Metric: number of operations (time)

  k bit key requires 2^k operations

  Depends on key length and computer speed

- Cryptanalysis

  Approach: Find weaknesses in algorithms

  Methods: Linear cryptanalysis, differential cryptanalysis, meet-in-the-middle attack, side-channel attacks

  Metrics: Number of operations

  Amount of memory

  Number of known plaintexts/ciphertexts

If either succeed all key usages are compromised

Cryptanalysis and Brute-Force Attack

- Cryptanalysis : Cryptanalytic attacks rely on the nature of the algorithm plus perhaps some knowledge of the general haracteristics of the plaintext or even some sample plaintext–ciphertext pairs.

- Brute-Force Attack :The attacker tries every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained.

Brute Force Attack

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

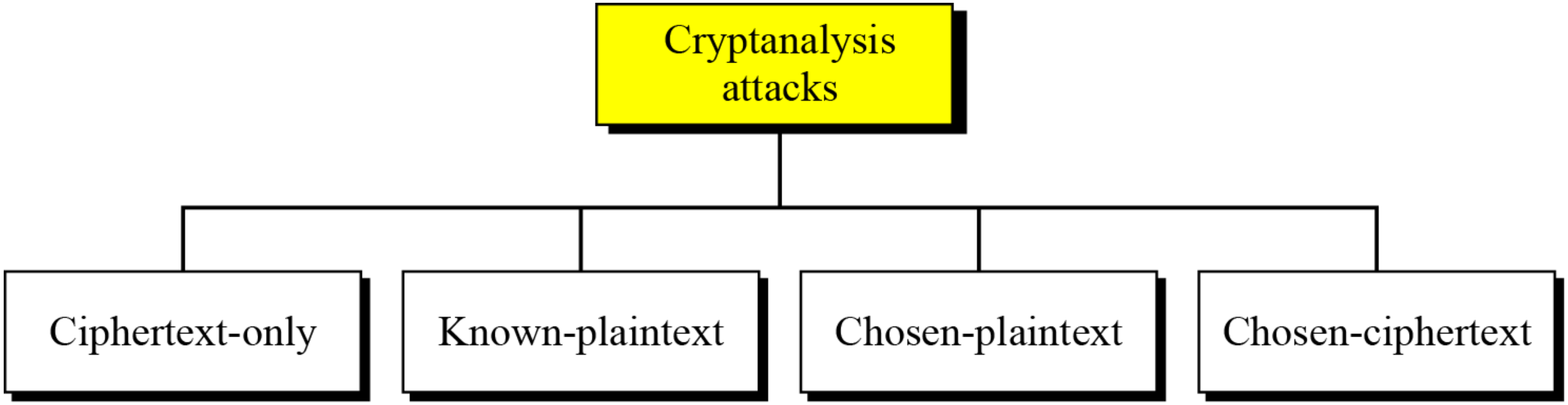
Key length	Key space	Worst case time at speed: 10^9/sec	10^{12}/sec	10^{15}/sec
32	2^{32}	4 sec	4 ms	4 us
56	2^{56}	833 days	20 hrs	72 sec
64	2^{64}	584 yrs	213 days	5 sec
128	2^{128}	10^{22} yrs	10^{19} yrs	10^{16} yrs
192	2^{192}	10^{41} yrs	10^{38} yrs	10^{35} yrs
256	2^{256}	10^{60} yrs	10^{57} yrs	10^{54} yrs
26!	2^{88}	10^{10} yrs	10^{7} yrs	10^{4} yrs

Age of Earth: 4×10^9 years

Age of Universe: 1.3×10^{10} years

Cryptanalysis

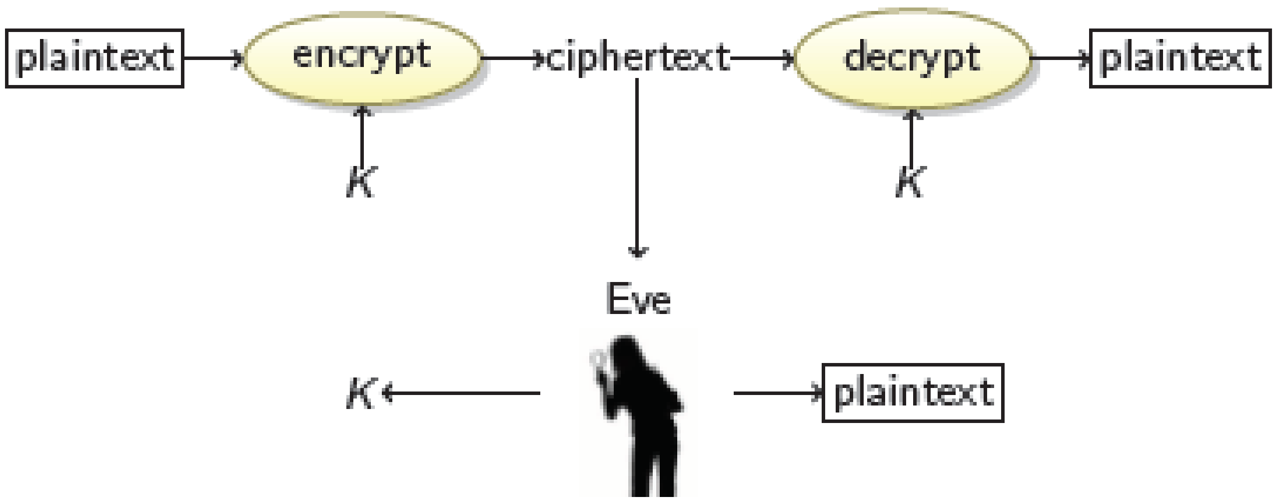
As cryptography is the science and art of creating secret codes, cryptanalysis is the science and art of breaking those codes.



Cryptanalysis attacks

Cryptanalysis (Cont.)

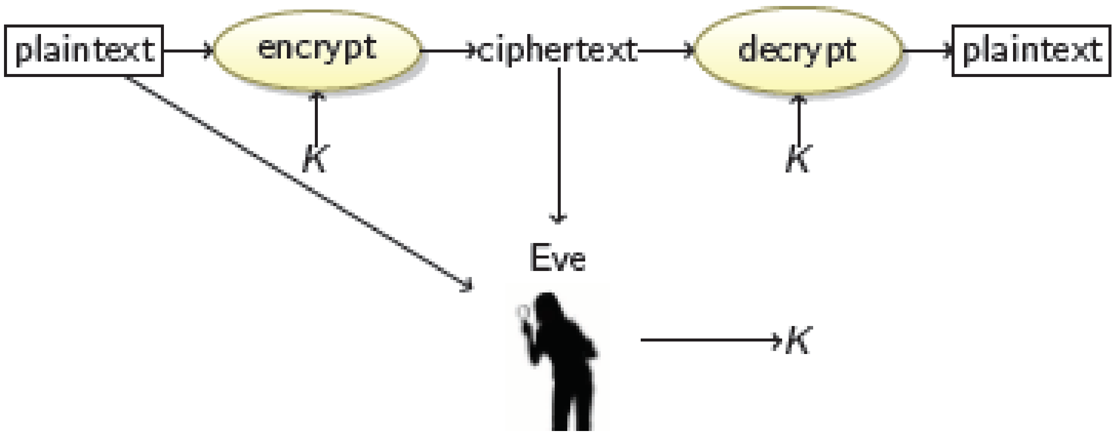
Ciphertext-only attack



We have: the ciphertext of several messages that have been encrypted with the same key, K.
We recover: the plaintexts, or K.

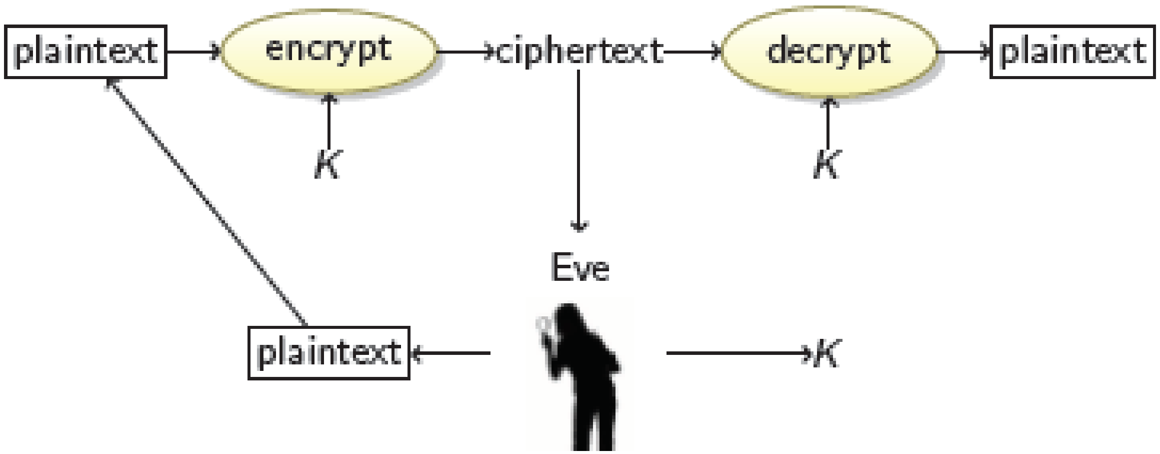
Cryptanalysis (Cont.)

Known-Plaintext Attack



We have: the ciphertexts and corresponding plaintexts of several messages, all encrypted with the same key K.
We recover: the key K.

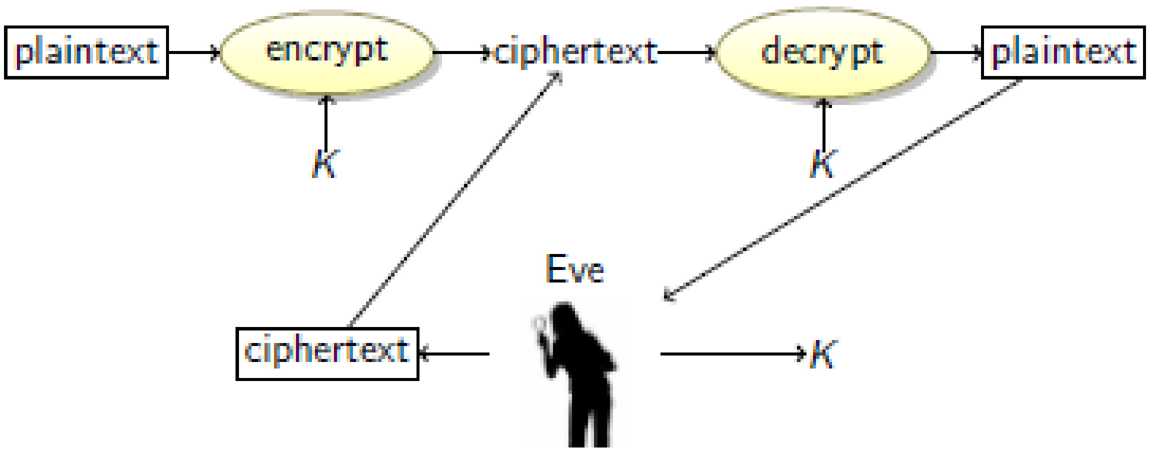
Chosen-Plaintext Attack



We have: the ciphertext of several messages that have been encrypted with the same key K, such that we get to choose the plaintexts.
We recover: the key K.

Cryptanalysis (Cont.)

Chosen-Ciphertext Attack



We have: the plaintext of several messages that have been encrypted with the same key K, such that we get to choose the ciphertexts.
We recover: the key K.

Cryptanalysis: Summary

Table 2.1 Types of Attacks on Encrypted Messages

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

Exercise

What type of attack is Eve employing here:

1. Eve tricks Alice into decrypting a bunch of ciphertexts that Alice encrypted last month.
2. Eve picks Alice's encrypted cell phone conversations.
3. Eve has given a bunch of messages to Alice for her to sign using the RSA signature scheme, which Alice does without looking at the messages and without using a one-way hash function. In fact, these messages are ciphertexts that Eve constructed to help her figure out Alice's RSA private key.
4. Eve has bet Bob that she can figure out the AES secret key he shares with Alice if he will simply encrypt 20 messages for Eve using that key. Bob agrees. Eve gives him 20 messages, which he then encrypts and emails back to Eve.

Cryptography

- Cryptographic systems are characterized along **three** independent dimensions:
- 1. The type of operations used for transforming plaintext to ciphertext.—substitution and transposition (or product)
- **2.** The number of keys used**.** (sym. and asym.)
- **3** The way in which the plaintext is processed. (block and stream)

Measures of Security

Unconditionally Secure

- Ciphertext does not contained enough information to derive plaintext or key
- One-time pad is only unconditionally secure cipher (but not very practical)

Computationally Secure

If either:

- Cost of breaking cipher exceeds value of encrypted information

-Time required to break cipher exceeds useful lifetime of encrypted information

- Hard to estimate value/lifetime of some information
- Hard to estimate how much effort needed to break cipher

Motivation for cryptanalysts

- All forms of cryptanalysis for symmetric encryption schemes are designed to exploit the fact that traces of structure or pattern in the plaintext may survive encryption and be discernible in the ciphertext.

Substitution ciphers

A substitution cipher replaces one symbol with another. Substitution ciphers can be categorized as either monoalphabetic ciphers or polyalphabetic ciphers.

A substitution cipher replaces one symbol with another.

Topics:

Monoalphabetic Ciphres
Polyalphabetic Ciphers

Monoalphabetic Ciphers

In monoalphabetic substitution, the relationship between a symbol in the plaintext to a symbol in the ciphertext is always one-to-one.

Encoding

- In these simple ciphers we typically
  1. convert all letters to upper case;
  2. remove spaces;
  3. remove punctuation;
  4. break into blocks of the same size (typically 5 letters);
  5. add some unusual letter (like Z) to the last block, if necessary.
- Example:

```
It wAs A DArk and sTormY NighT ...
```

  turns into

```
ITWAS ADARK ANDST ORMYN IGHTZ
```

- Knowing word boundaries can help with cryptanalysis.

Continued

Example

The following shows a plaintext and its corresponding ciphertext. The cipher is probably monoalphabetic because both l's (els) are encrypted as O's.

Plaintext: hello **Ciphertext:** KHOOR

Example

The following shows a plaintext and its corresponding ciphertext. The cipher is not monoalphabetic because each l (el) is encrypted by a different character.

Plaintext: hello **Ciphertext:** KHOLR

Additive Cipher

The simplest monoalphabetic cipher is the additive cipher.

This cipher is sometimes called a shift cipher and sometimes a Caesar cipher, but the term additive cipher better reveals its mathematical nature.

Plaintext and ciphertext in Z_{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	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
Value →	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Additive Cipher

- Replace each letter by the letter three positions along in alphabet

```
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
```

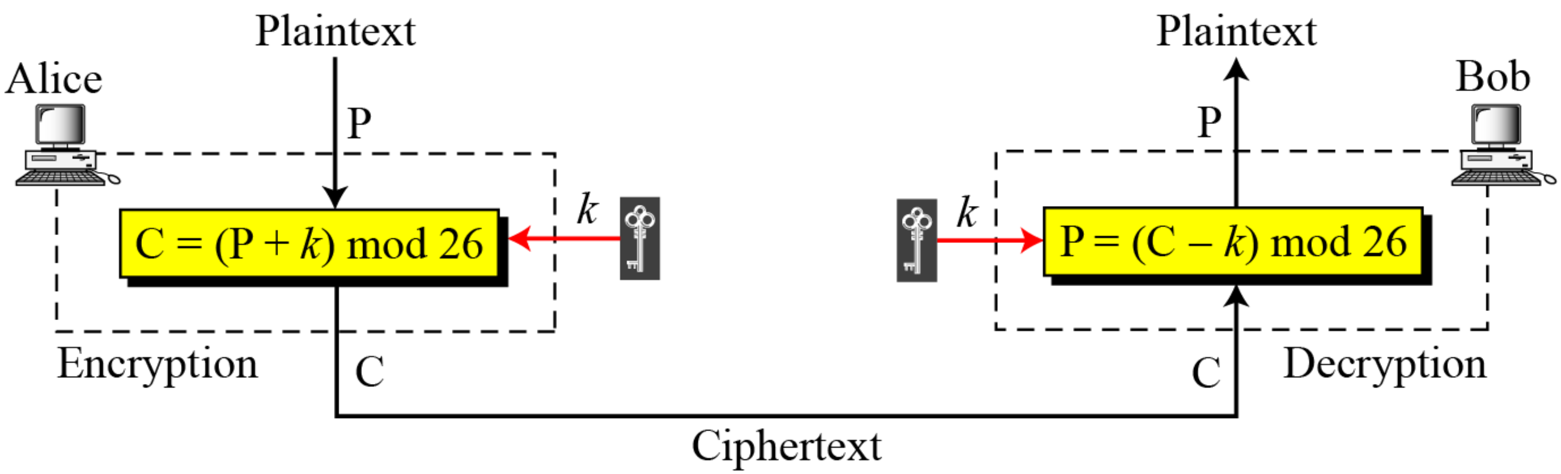
Generalized Caesar Cipher

- Allow shift by k positions
- Assume each letter assigned number (a = 0, b = 1, . . .)

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

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

Additive Cipher



When the cipher is additive, the plaintext, ciphertext, and key are integers in Z_{26}.

Use the additive cipher with key = 15 to encrypt the message "hello".

Solution

We apply the encryption algorithm to the plaintext, character by character:

Plaintext: h $\rightarrow$ 07	Encryption: (07 + 15) mod 26	Ciphertext: 22 $\rightarrow$ W
Plaintext: e $\rightarrow$ 04	Encryption: (04 + 15) mod 26	Ciphertext: 19 $\rightarrow$ T
Plaintext: l $\rightarrow$ 11	Encryption: (11 + 15) mod 26	Ciphertext: 00 $\rightarrow$ A
Plaintext: l $\rightarrow$ 11	Encryption: (11 + 15) mod 26	Ciphertext: 00 $\rightarrow$ A
Plaintext: o $\rightarrow$ 14	Encryption: (14 + 15) mod 26	Ciphertext: 03 $\rightarrow$ D

Shift Cipher and Caesar Cipher

Hitorically, additive ciphers are called shift ciphers. Julius Caesar used an additive cipher to communicate with his officers. For this reason, additive ciphers are sometimes referred to as the Caesar cipher. Caesar used a key of 3 for his communications.

Additive ciphers are sometimes referred to as shift ciphers or Caesar cipher.

Continued

Eve has intercepted the ciphertext “UVACLYFZLJBYL”. Show how she can use a brute-force attack to break the cipher.

Solution

Eve tries keys from 1 to 7. With a key of 7, the plaintext is “not very secure”, which makes sense.

Ciphertext: UVACLYFZLJBYL

K = 1 → **Plaintext:** tuzbkxeykiaxk
K = 2 → **Plaintext:** styajwdxjhzwj
K = 3 → **Plaintext:** rsxzivcwigyvi
K = 4 → **Plaintext:** qrwyhubvhfxuh
K = 5 → **Plaintext:** pqvxgtaugewtg
K = 6 → **Plaintext:** opuwfsztfdvsf
K = 7 → **Plaintext:** notverysecure

Breaking the Caesar Cipher

- Brute force attack

    -The encryption and decryption algorithms are known.

        -Try all 25 keys, e.g. k = 1, k = 2, . . .

        -Plaintext should be recognised

- Recognising plaintext in brute force attacks

        -Need to know “structure" of plaintext

        -Language? Compression?

- How to improve against brute force?

        -Hide the encryption/decryption algorithm: Not practical

        -Compress, use different language: Limited options

        -Increase the number of keys

Substitution in other forms: Monoalphabetic Ciphers

With only 25 possible keys, the Caesar cipher is far from secure. A dramatic increase in the key space can be achieved by allowing an arbitrary substitution (Random substitution).

Permutation: A **permutation** of a finite set of elements *S*
is an ordered sequence of all the elements of *S*, with each element appearing exactly once.

For example, if *S* = {a, b, c}, there are six permutations of *S*:
abc, acb, bac, bca, cab, cba
For n elementts, n! permutations.

For Caesar cipher:

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

For mono-alphabetic substitution:

If, instead, the "cipher" line can be any permutation of the 26 alphabetic characters,

then there are 26! or greater than $4 * 10^{26}$ possible keys

Example

We can use the key in Figure in previous slide to encrypt the message

this message is easy to encrypt but hard to find the key

The ciphertext is

ICFVQRVVNEFVRNVSIYRGAHSLIOJICNHTIYBFGTICRXRS

Attacks on Mono-alphabetic Ciphers

- Exploit the regularities of the language
  - -Frequency of letters, digrams, trigrams
  - -Expected words
- Fundamental problem with mono-alphabetic ciphers
  - -Ciphertext reflects the frequency data of original plaintext
  - -Solution 1: encrypt multiple letters of plaintext
  - -Solution 2: use multiple cipher alphabets

Language Redundancy and Cryptanalysis

- Human languages are **redundant**

  e.g., "th lrd s m shphrd shll nt wnt"
- 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
- Have tables of single, double & triple letter frequencies for various languages

Continued

Frequency of characters in English

Letter	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*
E	12.7	H	6.1	W	2.3	K	0.08
T	9.1	R	6.0	F	2.2	J	0.02
A	8.2	D	4.3	G	2.0	Q	0.01
O	7.5	L	4.0	Y	2.0	X	0.01
I	7.0	C	2.8	P	1.9	Z	0.01
N	6.7	U	2.8	B	1.5		
S	6.3	M	2.4	V	1.0		

Frequency of diagrams and trigrams

Digram	TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI, OF
Trigram	THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH

Relative Frequency of Letters in English Text

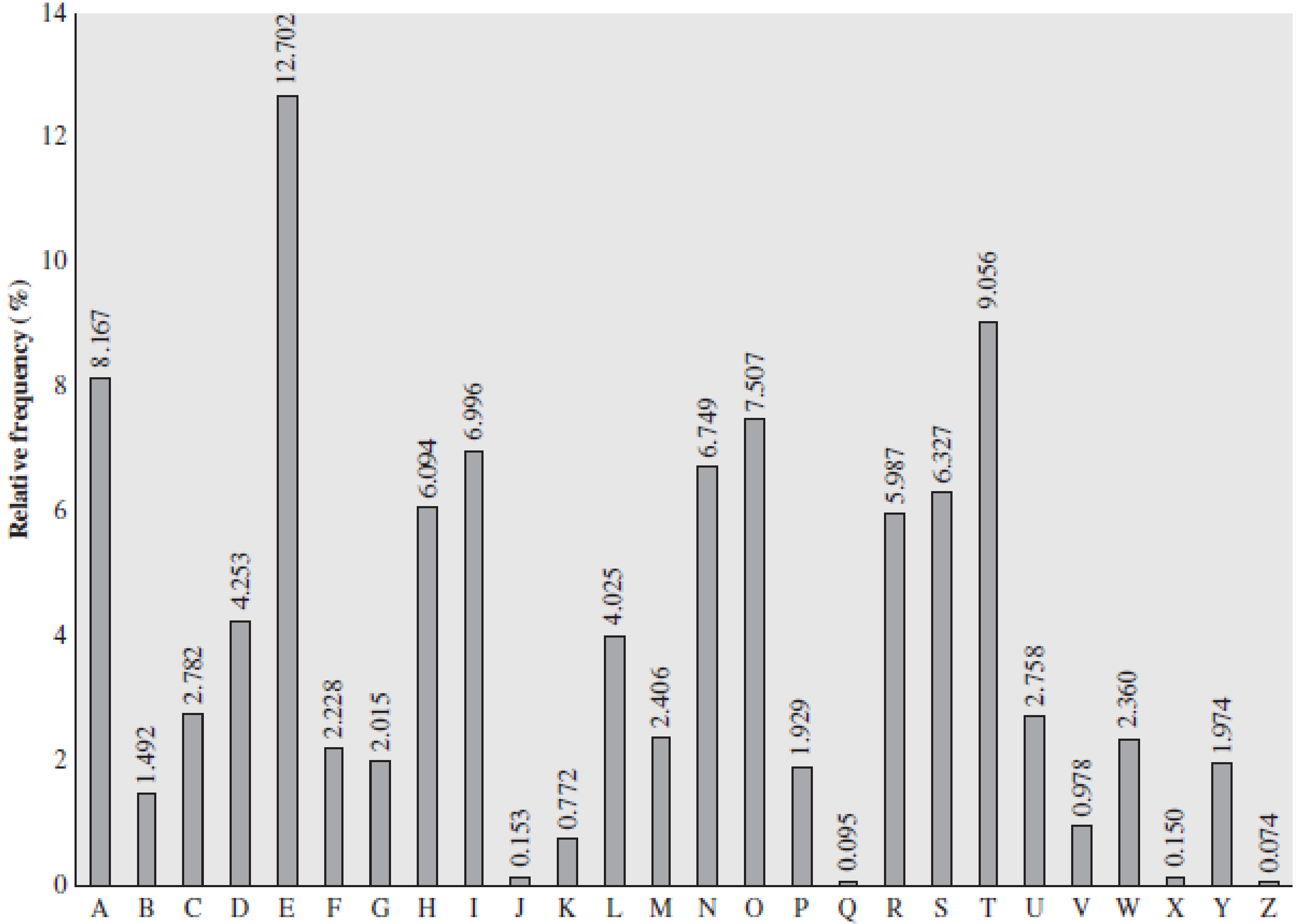
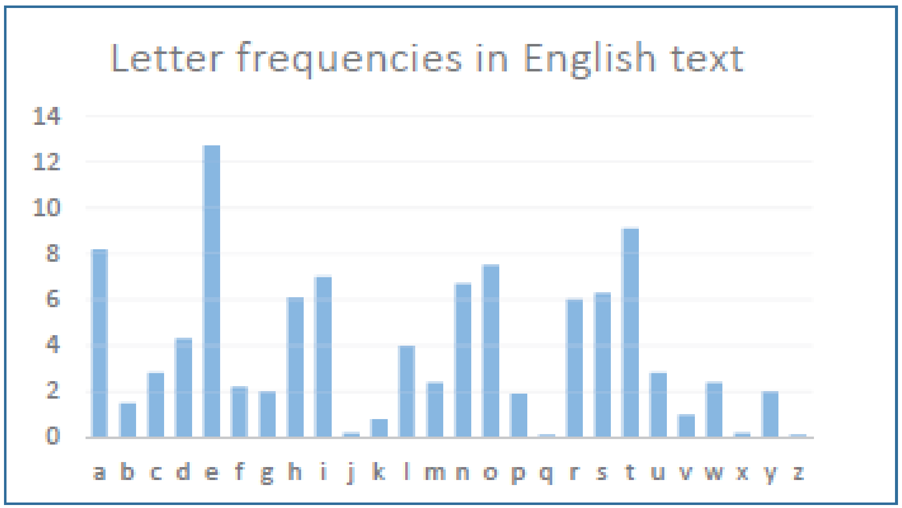
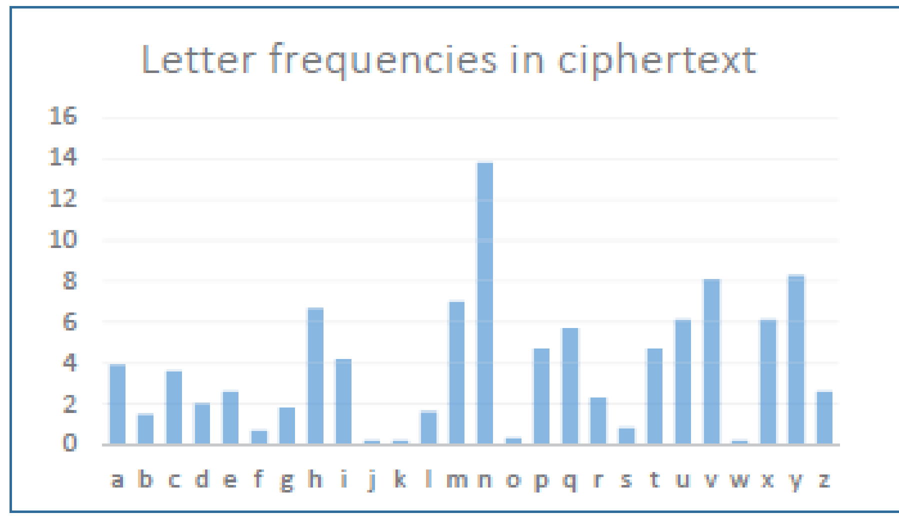


Figure 2.5 Relative Frequency of Letters in English Text

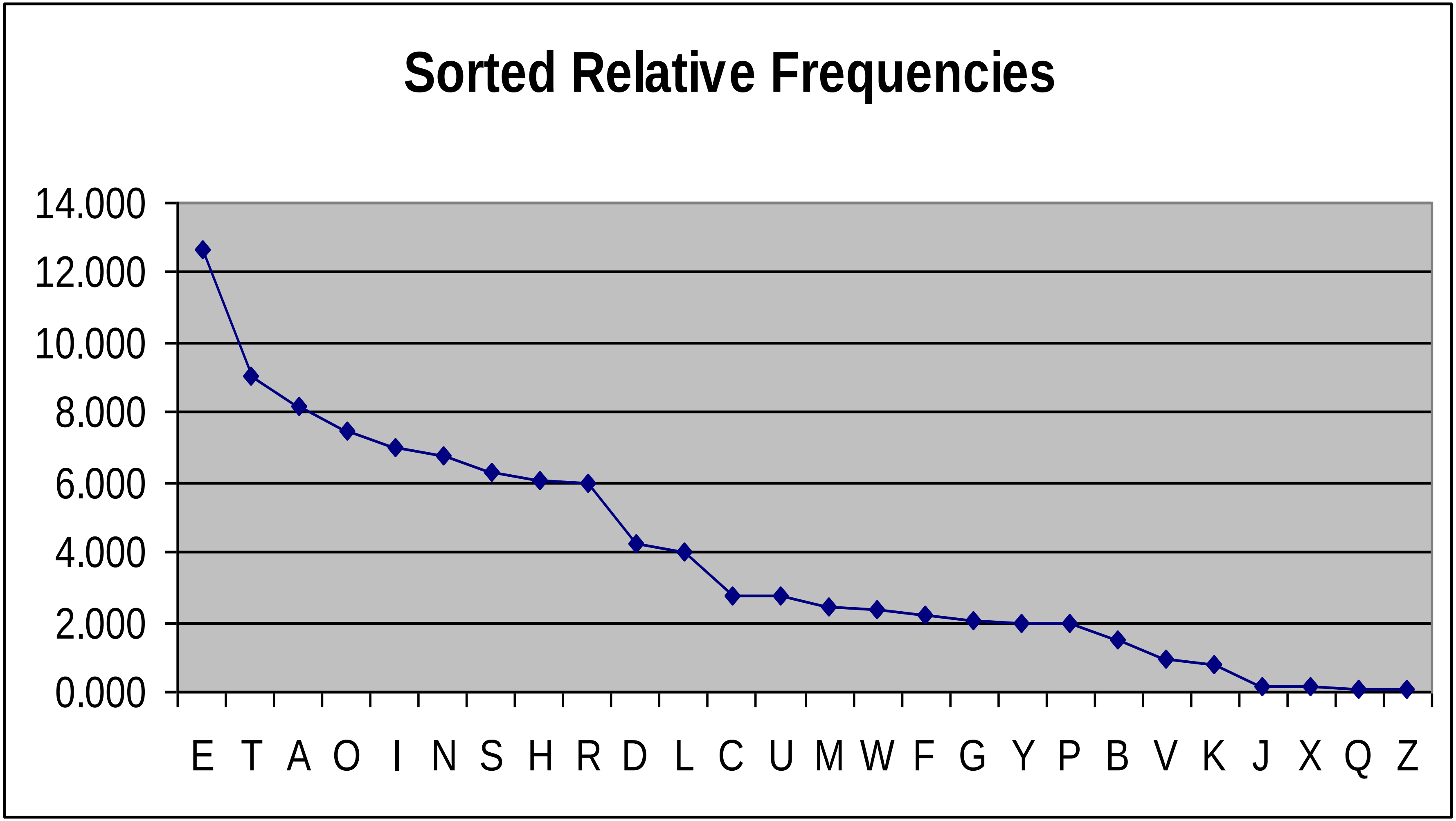
Breaking Monoalphabetic Substitution Cipher

- **Letter** Frequency Analysis results:





English Letter Frequencies



Example cryptanalysis

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

frequency

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								

As a first step, the relative frequency of the letters can be determined and compared to a standard frequency distribution for English

So far, then, we have

```
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
```

Example Cryptanalysis

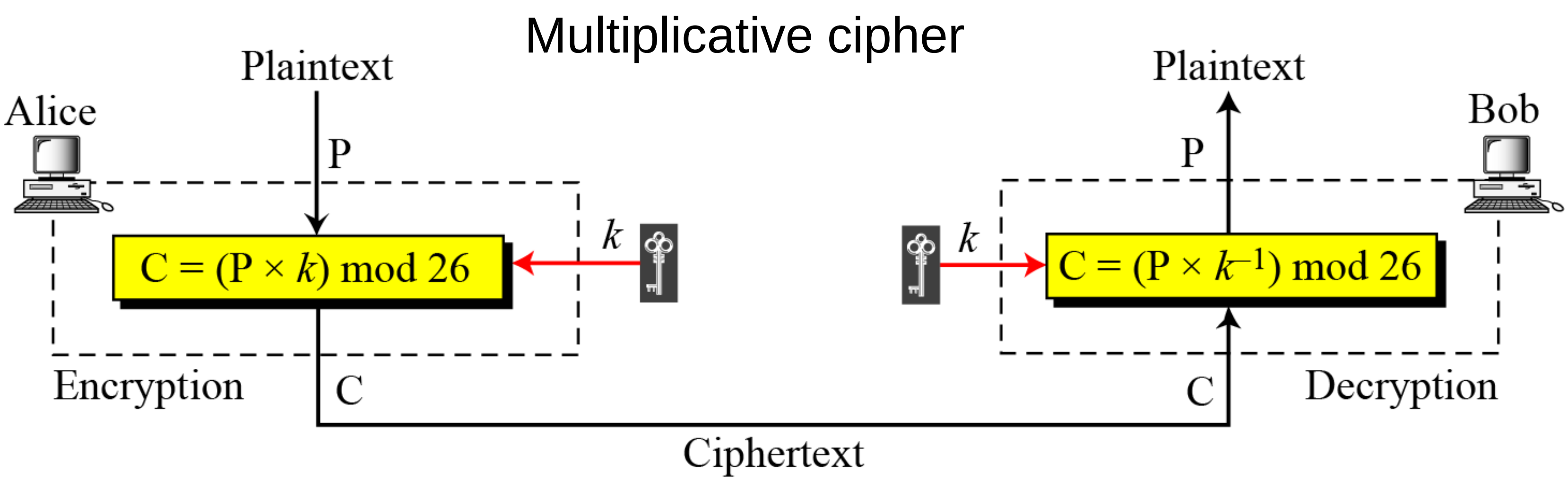
- given ciphertext

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- guess P & Z are e and t
- guess ZW is th and hence ZWP is “the”
- proceeding with trial and error finally get:

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

Multiplicative Ciphers



In a multiplicative cipher, the plaintext and ciphertext are integers in Z_{26}; the key is an integer in Z_{26}^*.

Continued.

Example

What is the key domain for any multiplicative cipher?

Solution

The key needs to be in $Z_{26}*$. This set has only 12 members: 1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25.

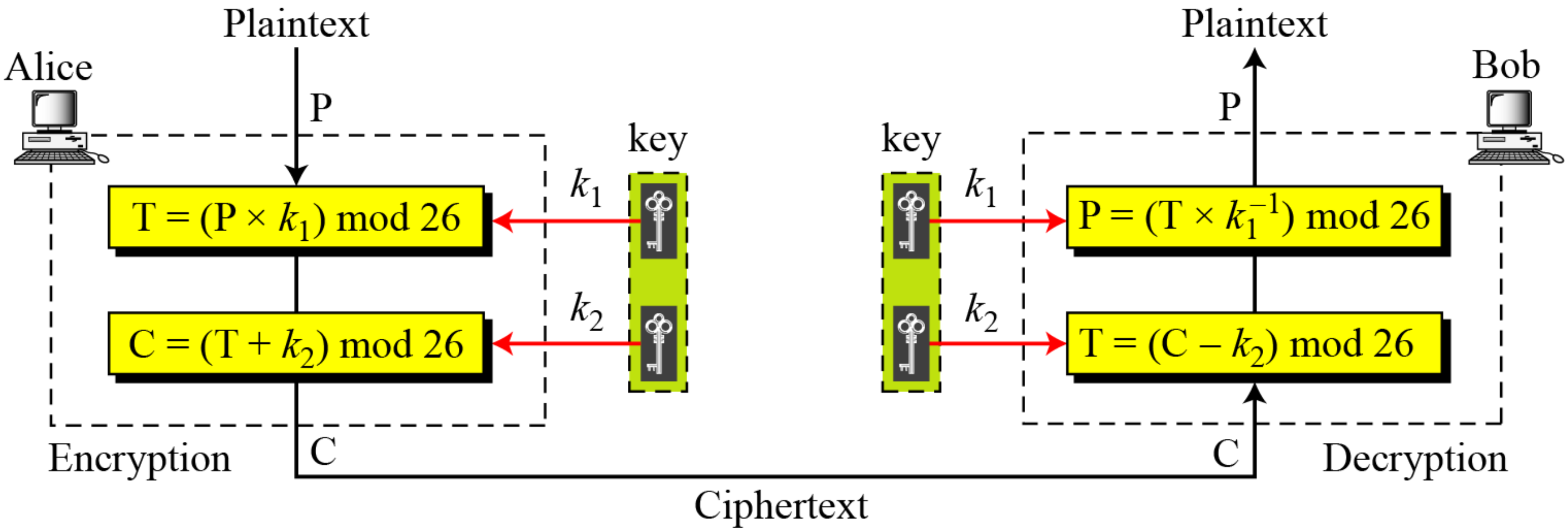
Example

We use a multiplicative cipher to encrypt the message "hello" with a key of 7. The ciphertext is "XCZZU".

Plaintext: h $\rightarrow$ 07	Encryption: (07×07) mod 26	ciphertext: 23 $\rightarrow$ X
Plaintext: e $\rightarrow$ 04	Encryption: (04×07) mod 26	ciphertext: 02 $\rightarrow$ C
Plaintext: l $\rightarrow$ 11	Encryption: (11×07) mod 26	ciphertext: 25 $\rightarrow$ Z
Plaintext: l $\rightarrow$ 11	Encryption: (11×07) mod 26	ciphertext: 25 $\rightarrow$ Z
Plaintext: o $\rightarrow$ 14	Encryption: (14×07) mod 26	ciphertext: 20 $\rightarrow$ U

Affine Ciphers

Combining additive and multiplicative cipher



$$C = (P \times k_1 + k_2) \bmod 26 \qquad P = ((C - k_2) \times k_1^{-1}) \bmod 26$$

where k_1^{-1} is the multiplicative inverse of k_1 and $-k_2$ is the additive inverse of k_2

Affine Ciphers

Example

The affine cipher uses a pair of keys in which the first key is from Z_{26}* and the second is from Z_{26}. The size of the key domain is $26 \times 12 = 312$.

Example

Use an affine cipher to encrypt the message "hello" with the key pair (7, 2).

P: h $\rightarrow$ 07	Encryption: $(07 \times 7 + 2)$ mod 26	C: 25 $\rightarrow$ Z
P: e $\rightarrow$ 04	Encryption: $(04 \times 7 + 2)$ mod 26	C: 04 $\rightarrow$ E
P: l $\rightarrow$ 11	Encryption: $(11 \times 7 + 2)$ mod 26	C: 01 $\rightarrow$ B
P: l $\rightarrow$ 11	Encryption: $(11 \times 7 + 2)$ mod 26	C: 01 $\rightarrow$ B
P: o $\rightarrow$ 14	Encryption: $(14 \times 7 + 2)$ mod 26	C: 22 $\rightarrow$ W

Affine Ciphers

Example

Use the affine cipher to decrypt the message "ZEBBW" with the key pair (7, 2) in modulus 26.

Solution

C: Z $\rightarrow$ 25	Decryption: $((25 - 2) \times 7^{-1}) \bmod 26$	P:07 $\rightarrow$ h
C: E $\rightarrow$ 04	Decryption: $((04 - 2) \times 7^{-1}) \bmod 26$	P:04 $\rightarrow$ e
C: B $\rightarrow$ 01	Decryption: $((01 - 2) \times 7^{-1}) \bmod 26$	P:11 $\rightarrow$ l
C: B $\rightarrow$ 01	Decryption: $((01 - 2) \times 7^{-1}) \bmod 26$	P:11 $\rightarrow$ l
C: W $\rightarrow$ 22	Decryption: $((22 - 2) \times 7^{-1}) \bmod 26$	P:14 $\rightarrow$ o

Example

The additive cipher is a special case of an affine cipher in which $k_1 = 1$. The multiplicative cipher is a special case of affine cipher in which $k_2 = 0$.

Polygraphic Substitution Ciphers

- In a polygram cipher blocks of characters in the plaintext are mapped to blocks of characters in the ciphertext:

$$\mathtt{ARF} \longrightarrow \mathtt{RTW}, \mathtt{ING} \longrightarrow \mathtt{PWQ}, \ldots$$

- We represent the cipher with a Substitution Box (S-Box):

	A	**B**	**C**	**D**	**E**	**F**
A	BA	CA	DC	DD	DE	FB
B	EA	AB	EC	BD	BE	AF
C	AA	BB	AC	ED	CE	BF
D	EB	DB	BC	CD	DF	FC
E	DA	CB	CC	AD	AE	FF
F	FA	CF	EE	FD	EF	FE

- Examples:

$$\begin{array}{lcl} \mathtt{AA} & \longrightarrow & \mathtt{BA} \\ \mathtt{AB} & \longrightarrow & \mathtt{CA} \\ \mathtt{EF} & \longrightarrow & \mathtt{FF} \end{array}$$

Substitution: Other forms (Cont)

- Use two-letter combinations: Playfair Cipher
- Use multiple letter combinations: Hill Cipher

Use multiple ciphers. Use a key to select which alphabet (code) is used for each letter of the message

Poly-alphabetic Ciphers

- Monoalphabetic ciphers are easy to break because they reflect the frequency data of the original alphabet. A countermeasure is to provide multiple substitutes known as homophones, for a single letter.

- For example, the letter e could be assigned a number of different cipher symbols, such as 16, 74, 35, and 21, with each homophone assigned to a letter in rotation or randomly.

- Use different mono-alphabetic substitutions as proceed through plaintext
  - Set of mono-alphabetic ciphers
  - Key determines which mono-alphabetic cipher to use for each plaintext letter
- Examples: Vigenere cipher, Vernam cipher, One time pad

Polyalphabetic Ciphers

In polyalphabetic substitution, each occurrence of a character may have a different substitute. The relationship between a character in the plaintext to a character in the ciphertext is one-to-many.

Autokey Cipher

$P = P_1P_2P_3 \ldots$ $\quad C = C_1C_2C_3 \ldots$ $\quad k = (k_1, P_1, P_2, \ldots)$

Encryption: $C_i = (P_i + k_i) \bmod 26$ $\quad$ Decryption: $P_i = (C_i - k_i) \bmod 26$

Autokey Cipher

Assume that Alice and Bob agreed to use an autokey cipher with initial key value k1 = 12. Now Alice wants to send Bob the message “Attack is today”. Enciphering is done character by character.

Plaintext:	a	t	t	a	c	k	i	s	t	o	d	a	y
P’s Values:	00	19	19	00	02	10	08	18	19	14	03	00	24
Key stream:	*12*	*00*	*19*	*19*	*00*	*02*	*10*	*08*	*18*	*19*	*14*	*03*	*00*
C’s Values:	12	19	12	19	02	12	18	00	11	7	17	03	24
Ciphertext:	**M**	**T**	**M**	**T**	**C**	**M**	**S**	**A**	**L**	**H**	**R**	**D**	**Y**

Autokey Cipher

- Vigenère proposed what is referred to as an autokey system**,** in which a keyword is concatenated with the plaintext itself to provide a running key. For our example,

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

Even this scheme is vulnerable to cryptanalysis. Because the key and the plaintext share the same frequency distribution of letters, a statistical technique can be applied.

Playfair Cipher

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

Playfair Cipher

- Initialization

  1. Create 5x5 matrix and write keyword (row by row)

  2. Fill out remainder with alphabet, not repeating any letters

  3. Special: Treat I and J as same letter

- Encryption

  1. Operate on pair of letters (digram) at a time

  2. Special: if digram with same letters, separate by special letter (e.g. x)

  3. Plaintext in same row: replace with letters to right

  4. Plaintext in same column: replace with letters below

  5. Else, replace by letter in same row as it and same column as other plaintext letter

Playfair Cipher

- Rules to encrypt the digraph $\alpha\beta$:
  1. If $\alpha = \beta$, add an **X**, encrypt the new pair.
  2. If one letter is left, add an **X**, encrypt the new pair.
  3. If α, β are in the same row:

$$\begin{array}{|c|c|c|c|c|} \hline * & * & * & * & * \\ \hline * & * & * & * & * \\ \hline \alpha & X & * & \beta & Y \\ \hline * & * & * & * & * \\ \hline * & * & * & * & * \\ \hline \end{array} \quad \Rightarrow \quad \alpha\beta \rightarrow XY$$

     If necessary, wrap around.
  4. If $\alpha\beta$ occur in the same column:

$$\begin{array}{|c|c|c|c|c|} \hline * & * & * & * & * \\ \hline * & * & \alpha & * & * \\ \hline * & * & X & * & * \\ \hline * & * & \beta & * & * \\ \hline * & * & Y & * & * \\ \hline \end{array} \quad \Rightarrow \quad \alpha\beta \rightarrow XY$$

Playfair Cipher

- And the final rule:
  5. If the letters are not on the same row or column:

X	$*$	$*$	α	$*$
$*$	$*$	$*$	$*$	$*$
$*$	$*$	$*$	$*$	$*$
β	$*$	Y	$*$	$*$
$*$	$*$	$*$	$*$	$*$

$$\Rightarrow \quad \alpha\beta \rightarrow XY$$

  Order matters: X is on the same row as α.

- To decrypt:
  1. Use the inverse of the last three rules.
  2. Drop any **X**s that don't make sense.

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

	balloon	**balxloxon**
ar		RM
mu		CM
hs		BP
ea		IM

In this case, the keyword is *monarchy.*

Plaintext is encrypted two letters at a time

An example of a secret key in the Playfair cipher

Secret Key =

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

Example

Let us encrypt the plaintext “hello” using the key in Figure

he $\rightarrow$ EC   lx $\rightarrow$ QZ   lo $\rightarrow$ BX

Plaintext: hello   Ciphertext: ECQZBX

An example of a secret key in the Playfair cipher

- Example plaintext:

IT WA SA DA RK AN DS TO RM YN IG HT

- IT→MP

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

- WA→XI

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

An example of a secret key in the Playfair cipher

- SA→XG

D	I	A	M	O
N	R	$\boxed{G}$	B	C
E	F	H	J	K
L	P	S	T	U
V	W	$\boxed{X}$	Y	Z

- DA→IM

D	$\boxed{I}$	A	$\boxed{M}$	O
N	R	G	B	C
E	F	H	J	K
L	P	S	T	U
V	W	X	Y	Z

Exercise

1. Construct a Playfair table using the key phrase BLINKENLIGHTS.
2. Encode the message Run, RAbbit, Run!
3. Encrypt the plaintext message from 2.
4. Decrypt the ciphertext message from 3.

Measuring Effectiveness of the Playfair and other ciphers

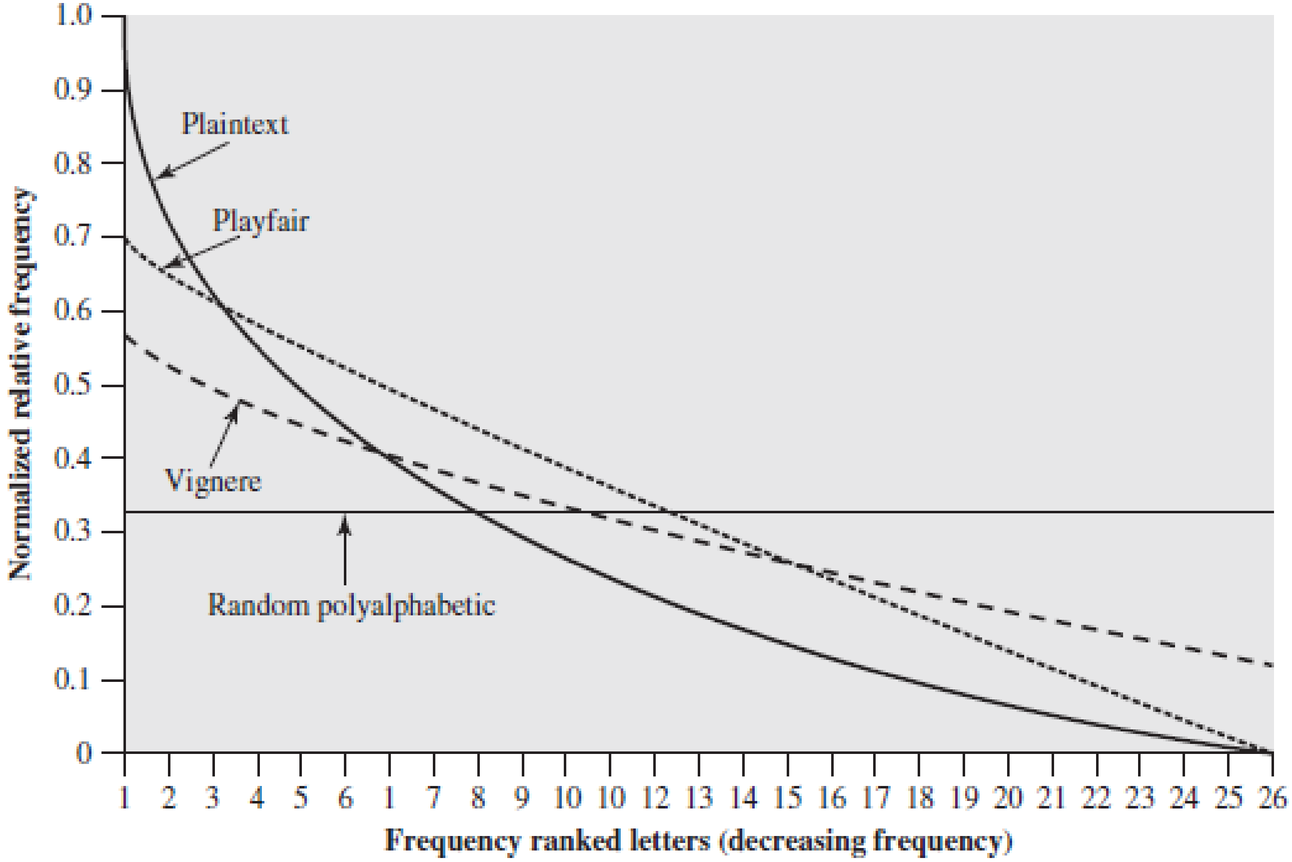
- The following plot is developed: (i) The number of occurrences of each letter in the text is counted and divided by the number of occurrences of the most frequently used letter. e is the most frequently used letter.

- To normalize the plot, the number of occurrences of each letter in the ciphertext was again divided by the number of ccurrences of e in the plaintext.

- If the frequency distribution information were totally concealed in the encryption process, the ciphertext plot

  of frequencies would be flat, and cryptanalysis using ciphertext only would be effectively impossible.

Relative Frequency of Occurrence of Letters



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 (versus 26 for a monoalphabetic) and correspondingly more ciphertext was widely used for many years
- E.g. by US & British military in WW1
- It can be broken, given a few hundred letters since still has much of plaintext structure

Playfair Cipher - Is it Breakable?

- Better than mono-alphabetic: relative frequency of digrams much less than of individual letters
- But relatively easy (digrams, trigrams, expected words)

Vigenere Cipher

- Set of 26 general Caesar ciphers

  26 Caesar ciphers with shifts of 0 through 25.
- Letter in key determines the Caesar cipher to use
  - Key must be as long as plaintext: repeat a keyword
  - Key: pqr
  - Plaintext: attack is today
- Example:

```
Plain:   a t t a c k i s t o d a y
Key:     p q r p q r p q r p q r p
Cipher:
```

Multiple ciphertext letters for each plaintext letter

$$P = p_0, p_1, p_2, \ldots, p_{n-1}$$

$$K = k_0, k_1, k_2, \ldots, k_{m-1}, \text{ where typically } m < n$$

$$C = C_0, C_1, C_2, \ldots, C_{n-1}$$

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

For the next *m letters of the plaintext, the key letters are repeated.*

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

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

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

Expressed numerically, we have the following result.

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

Vigenere Cipher

$$P = P_1P_2P_3 \ldots \qquad C = C_1C_2C_3 \ldots \qquad K = [(k_1, k_2, \ldots, k_m), (k_1, k_2, \ldots, k_m), \ldots]$$

$$\text{Encryption: } C_i = P_i + k_i \qquad \text{Decryption: } P_i = C_i - k_i$$

Vigenere Cipher

Example

We can encrypt the message "She is listening" using the 6-character keyword "PASCAL".

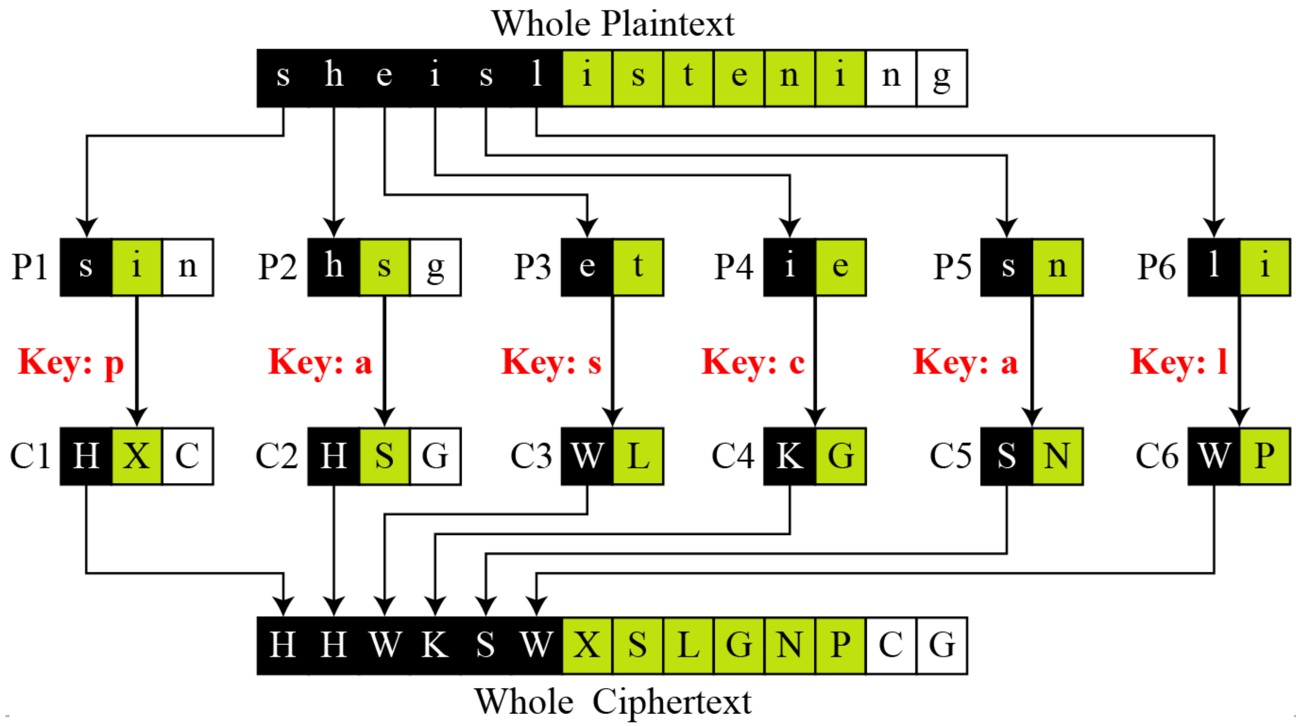
Let us see how we can encrypt the message "She is listening" using the 6-character keyword "PASCAL". The initial key stream is (15, 0, 18, 2, 0, 11). The key stream is the repetition of this initial key stream (as many times as needed).

Plaintext:	s	h	e	i	s	l	i	s	t	e	n	i	n	g
P's values:	18	07	04	08	18	11	08	18	19	04	13	08	13	06
Key stream:	***15***	***00***	***18***	***02***	***00***	***11***	***15***	***00***	***18***	***02***	***00***	***11***	***15***	***00***
C's values:	**07**	**07**	**22**	**10**	**18**	**22**	**23**	**18**	**11**	**6**	**13**	**19**	**02**	**06**
Ciphertext:	**H**	**H**	**W**	**K**	**S**	**W**	**X**	**S**	**L**	**G**	**N**	**T**	**C**	**G**

Vigenere Cipher

Vigenere cipher can be seen as combinations of m additive ciphers.

Figure A Vigenere cipher as a combination of m additive ciphers



Vigenere Cipher

Using Example , we can say that the additive cipher is a special case of Vigenere cipher in which m = 1.

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

Table
A Vigenere Tableau

Vigenere Cipher

(Cryptanalysis)

Example

Let us assume we have intercepted the following ciphertext:

LIOMWGFEGGDVWGHHCQUCRHRWAGWIOWQLKGZETKKMEVLWPCZVGTH-
VTSGXQOVGCSVETQLTJSUMVWVEUVLXEWSLGFZMVVWLGYHCUSWXQH-
KVGSHEEVFLCFDGVSUMPHKIRZDMPHHBVWVWJWIXGFWLTSHGJOUEEHH-
VUCFVGOWICQLTJSUXGLW

The Kasiski test for repetition of three-character segments yields the results shown in Table .

String	*First Index*	*Second Index*	*Difference*
JSU	68	168	100
SUM	69	117	48
VWV	72	132	60
MPH	119	127	8

Vigenere Cipher (Crypanalysis)

Example Cont.

Let us assume we have intercepted the following ciphertext:

LIOMWGFEGGDVWGHHCQUCRHRWAGWIOWQLKGZETKKMEVLWPCZVGTH-
VTSGXQOVGCSVETQLTJSUMVWVEUVLXEWSLGFZMVVWLGYHCUSWXQH-
KVGSHEEVFLCFDGVSUMPHKIRZDMPHHBVWVWJWIXGFWLTSHGJOUEEHH-
VUCFVGOWICQLTJSUXGLW

The Kasiski test for repetition of three-character segments in ciphertext yields the results shown in Table

String	*First Index*	*Second Index*	*Difference*
JSU	68	168	100
SUM	69	117	48
VWV	72	132	60
MPH	119	127	8

Example Cont.

The greatest common divisor of differences is 4, which means that the key length is multiple of 4. First try m = 4.

```
C1: LWGWCRAOKTEPGTQCTJVUEGVGUQGECVPRPVJGTJEUGCJG
P1: jueuapymircneroarhtsthihytrahcieixsthcarrehe
C2: IGGGQHGWGKVCTSOSQSWVWFVYSHSVFSHZHWWFSOHCOQSL
P2: ussscsiswhofeaeceihcetesoecatnpntherhctecex
C3: OFDHURWQZKLZHGVVLUVLSZWHWKHFDUKDHVIWHUHFWLUW
P3: lcaerotnwhiwedssirsiirhketehretltiideatrairt
C4: MEVHCWILEMWVVXGETMEXLMLCXVELGMIMBWXLGEVVITX
P4: iardysehaisrrtcapiafpwtethecarhaesfterectpt
```

In this case, the plaintext makes sense.

Julius Caesar used a cryptosystem in his wars, which is now referred to as Caesar cipher. It is an additive cipher with the key set to three. Each character in the plaintext is shifted three characters to create ciphertext.

Vigenere Cipher - Is it Breakable?

- Yes
- Monoalphabetic or Vigenere cipher? Letter frequency analysis
- Determine length of keyword
- For keyword length m, Vigenere is m mono-alphabetic substitutions
- Break the mono-alphabetic ciphers separately

Weakness is repeating, structured keyword

Hill Cipher

Another interesting multiletter cipher is the Hill cipher, developed by the mathematician Lester Hill in 1929.

Plaintext are divided into equal size blocks. Each character in a block contributes to the encryption of the other characters in the block. (Block Cipher)

$$K = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1m} \\ k_{21} & k_{22} & \cdots & k_{2m} \\ \vdots & \vdots & & \vdots \\ k_{m1} & k_{m2} & \cdots & k_{mm} \end{bmatrix}$$

$$C_1 = P_1 k_{11} + P_2 k_{21} + \cdots + P_m k_{m1}$$
$$C_2 = P_1 k_{12} + P_2 k_{22} + \cdots + P_m k_{m2}$$
$$\cdots$$
$$C_m = P_1 k_{1m} + P_2 k_{2m} + \cdots + P_m k_{mm}$$

The key matrix in the Hill cipher needs to have a multiplicative inverse.

Examp Hilll1

For example, the plaintext "code is ready" can make a 3 × 4 matrix when adding extra bogus character "z" to the last block and removing the spaces. The ciphertext is "OHKNIHGKLISS".

Figure Example

$$\underset{C}{\begin{bmatrix} 14 & 07 & 10 & 13 \\ 08 & 07 & 06 & 11 \\ 11 & 08 & 18 & 18 \end{bmatrix}} = \underset{P}{\begin{bmatrix} 02 & 14 & 03 & 04 \\ 08 & 18 & 17 & 04 \\ 00 & 03 & 24 & 25 \end{bmatrix}} \underset{K}{\begin{bmatrix} 09 & 07 & 11 & 13 \\ 04 & 07 & 05 & 06 \\ 02 & 21 & 14 & 09 \\ 03 & 23 & 21 & 08 \end{bmatrix}}$$

a. Encryption

$$\underset{P}{\begin{bmatrix} 02 & 14 & 03 & 04 \\ 08 & 18 & 17 & 04 \\ 00 & 03 & 24 & 25 \end{bmatrix}} = \underset{C}{\begin{bmatrix} 14 & 07 & 10 & 13 \\ 08 & 07 & 06 & 11 \\ 11 & 08 & 18 & 18 \end{bmatrix}} \underset{K^{-1}}{\begin{bmatrix} 02 & 15 & 22 & 03 \\ 15 & 00 & 19 & 03 \\ 09 & 09 & 03 & 11 \\ 17 & 00 & 04 & 07 \end{bmatrix}}$$

b. Decryption

Example Hill2

Assume that Eve knows that m = 3. She has intercepted three plaintext/ciphertext pair blocks (not necessarily from the same message) as shown in Figure .

Figure

$$\begin{bmatrix} 05 & 07 & 10 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 03 & 06 & 00 \end{bmatrix}$$

$$\begin{bmatrix} 13 & 17 & 07 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 14 & 16 & 09 \end{bmatrix}$$

$$\begin{bmatrix} 00 & 05 & 04 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 03 & 17 & 11 \end{bmatrix}$$

P C

Example Hill2 cont.

She makes matrices P and C from these pairs. Because P is invertible, she inverts the P matrix and multiplies it by C to get the K matrix as shown in Figure.

Figure Example

$$\begin{bmatrix} 02 & 03 & 07 \\ 05 & 07 & 09 \\ 01 & 02 & 11 \end{bmatrix} = \begin{bmatrix} 21 & 14 & 01 \\ 00 & 08 & 25 \\ 13 & 03 & 08 \end{bmatrix} \begin{bmatrix} 03 & 06 & 00 \\ 14 & 16 & 09 \\ 03 & 17 & 11 \end{bmatrix}$$

$$\mathrm{K} \qquad \mathrm{P}^{-1} \qquad \mathrm{C}$$

Now she has the key and can break any ciphertext encrypted with that key.

- Concepts from Linear Algebra

We define the inverse $\mathbf{M}^{-1}$ of a square matrix $\mathbf{M}$ by the equation $\mathbf{M}(\mathbf{M}^{-1}) = \mathbf{M}^{-1}\mathbf{M} = \mathbf{I}$, where **I** is the identity matrix. **I** is a square matrix that is all zeros except for ones along the main diagonal from upper left to lower right.
(we are concerned with matrix arithmetic modulo 26).

$$\mathbf{A} = \begin{pmatrix} 5 & 8 \\ 17 & 3 \end{pmatrix} \qquad \mathbf{A}^{-1} \bmod 26 = \begin{pmatrix} 9 & 2 \\ 1 & 15 \end{pmatrix}$$

$$\mathbf{AA}^{-1} = \begin{pmatrix} (5 \times 9) + (8 \times 1) & (5 \times 2) + (8 \times 15) \\ (17 \times 9) + (3 \times 1) & (17 \times 2) + (3 \times 15) \end{pmatrix}$$

$$= \begin{pmatrix} 53 & 130 \\ 156 & 79 \end{pmatrix} \bmod 26 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Matrix Operations

- Matrix addition/subtraction
  - Matrices must be of same size.

- Matrix multiplication

$$\begin{array}{ccc} m \times n & n \times p & m \times p \\ \begin{bmatrix} a_{11} & a_{12} & . & a_{1n} \\ a_{21} & a_{22} & . & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & . & a_{mn} \end{bmatrix} & \begin{bmatrix} b_{11} & b_{12} & . & b_{1p} \\ b_{21} & b_{22} & . & b_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ b_{q1} & b_{q2} & . & b_{qp} \end{bmatrix} = & \begin{bmatrix} c_{11} & c_{12} & . & c_{1p} \\ c_{21} & c_{22} & . & c_{2p} \\ \cdots & \cdots & c_{ij} & \cdots \\ c_{m1} & c_{m2} & . & c_{mp} \end{bmatrix} \end{array}$$

Condition: n = q

$$AB \neq BA$$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Identity Matrix

$$AI = IA = A, \text{ where } I = \begin{bmatrix} 1 & 0 & . & 0 \\ 0 & 1 & . & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & . & 1 \end{bmatrix}$$

Matrix Transpose

$$A = \begin{bmatrix} a_{11} & a_{12} & . & a_{1n} \\ a_{21} & a_{22} & . & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & . & a_{mn} \end{bmatrix}, \; A^T = \begin{bmatrix} a_{11} & a_{21} & . & a_{m1} \\ a_{12} & a_{22} & . & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & . & a_{mn} \end{bmatrix}$$

Property: $(AB)^T = B^T A^T$

Symmetric Matrices

$$A = A^T \ (a_{ij} = a_{ji})$$

Example:

$$\begin{bmatrix} 4 & 5 & -3 \\ 5 & 7 & 2 \\ -3 & 2 & 10 \end{bmatrix}$$

Determinants

2 x 2

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

3 x 3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

n x n

$$det(A) = \sum_{j=1}^{m} (-1)^{j+k} a_{jk} det(A_{jk}), \text{ for any } k\text{: } 1 \le k \le m$$

Properties:

$$det(AB) = det(A)det(B)$$

$$det(A+B) \neq det(A) + det(B)$$

Matrix Inverse

- The inverse A^{-1} of a matrix A has the property:

$$AA^{-1}=A^{-1}A=I$$

- A^{-1} exists only if

$$det(A) \neq 0$$

- Terminology
  - **Singular matrix:** A^{-1} does not exist
  - **Ill-conditioned matrix:** A is "close" to being singular

Matrix Inverse (cont'd)

- Properties of the inverse:

$$det(A^{-1}) = \frac{1}{det(A)}$$

$$(AB)^{-1} = B^{-1} A^{-1}$$

$$(A^T)^{-1} = (A^{-1})^T$$

Inverse of a Matrix

Determinant

For any square matrix (m × m), the determinant equals the sum of all the products that can be formed by taking exactly one element from each row and exactly one element from each column, with certain of the product terms preceded by a minus sign

$$\begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix}$$

the determinant is $k_{11}k_{22} - k_{12}k_{21}$. For a 3×3 matrix, the value of the determinant is $k_{11}k_{22}k_{33} + k_{21}k_{32}k_{13} + k_{31}k_{12}k_{23} - k_{31}k_{22}k_{13} - k_{21}k_{12}k_{33} - k_{11}k_{32}k_{23}$. If a square matrix **A** has a nonzero determinant, then the inverse of the matrix is computed as $[\mathbf{A}^{-1}]_{ij} = (\det \mathbf{A})^{-1}(-1)^{i+j}(\mathrm{D}_{ji})$, where (D_{ji}) is the subdeterminant formed by deleting the jth row and the ith column of **A**, det(**A**) is the determinant of **A**, and $(\det \mathbf{A})^{-1}$ is the multiplicative inverse of (det **A**) mod 26.

$$\mathbf{A} = \begin{pmatrix} 5 & 8 \\ 17 & 3 \end{pmatrix} \qquad \mathbf{A}^{-1} \bmod 26 = \begin{pmatrix} 9 & 2 \\ 1 & 15 \end{pmatrix}$$

$$\mathbf{A}\mathbf{A}^{-1} = \begin{pmatrix} (5 \times 9) + (8 \times 1) & (5 \times 2) + (8 \times 15) \\ (17 \times 9) + (3 \times 1) & (17 \times 2) + (3 \times 15) \end{pmatrix}$$

$$= \begin{pmatrix} 53 & 130 \\ 156 & 79 \end{pmatrix} \bmod 26 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Det(A)= $\det\begin{pmatrix} 5 & 8 \\ 17 & 3 \end{pmatrix} = (5 \times 3) - (8 \times 17) = -121 \bmod 26 = 9$

$\text{Det(A)}^{-1} = 9^{-1} \bmod 26 = 3$ (not 1/9)

As, $9 \times 3 = 27 \bmod 26 = 1$

Note: $[\mathbf{A}^{-1}]_{ij} = (\det \mathbf{A})^{-1}(-1)^{i+j}(\mathbf{D}_{ji})$

Det(A)= $\det\begin{pmatrix} 5 & 8 \\ 17 & 3 \end{pmatrix} = (5 \times 3) - (8 \times 17) = -121 \bmod 26 = 9$

Det(A)$^{-1}$= 9^{-1} mod 26=3

As, $9 \times 3 = 27 \bmod 26 = 1$

$$\mathbf{A} = \begin{pmatrix} 5 & 8 \\ 17 & 3 \end{pmatrix}$$

$$\mathbf{A}^{-1} \bmod 26 = 3\begin{pmatrix} 3 & -8 \\ -17 & 5 \end{pmatrix} = 3\begin{pmatrix} 3 & 18 \\ 9 & 5 \end{pmatrix} = \begin{pmatrix} 9 & 54 \\ 27 & 15 \end{pmatrix} = \begin{pmatrix} 9 & 2 \\ 1 & 15 \end{pmatrix}$$

$$[\mathbf{A}^{-1}]_{ij} = (\det \mathbf{A})^{-1}(-1)^{i+j}(\mathbf{D}_{ji})$$

$$\det\begin{pmatrix}5 & 8\\ 17 & 3\end{pmatrix} = (5 \times 3) - (8 \times 17) = -121 \bmod 26 = 9$$

We can show that $9^{-1} \bmod 26 = 3$, because $9 \times 3 = 27 \bmod 26 = 1$ Therefore, we compute the inverse of **A** as

$$\mathbf{A} = \begin{pmatrix}5 & 8\\ 17 & 3\end{pmatrix}$$

$$\mathbf{A}^{-1} \bmod 26 = 3\begin{pmatrix}3 & -8\\ -17 & 5\end{pmatrix} = 3\begin{pmatrix}3 & 18\\ 9 & 5\end{pmatrix} = \begin{pmatrix}9 & 54\\ 27 & 15\end{pmatrix} = \begin{pmatrix}9 & 2\\ 1 & 15\end{pmatrix}$$

The Hill Algorithm

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, c, z = 25). For m = 3, the system can be described as

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

This can be expressed in terms of row vectors and matrices:[6]

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

or

$$\mathbf{C} = \mathbf{PK} \bmod 26$$

In general

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

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

$$\mathbf{K} = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

"paymoremoney"

The first three letters of the plaintext are represented by the vector (15 0 24). Then(15 0 24)$\mathbf{K}$ = (303 303 531) mod 26 = (17 17 11) = RRL. Continuing in this fashion, the ciphertext for the entire plaintext is RRLMWBKASPDH.

One-Time Pad

- One of the goals of cryptography is perfect secrecy.
- A study by Shannon has shown that perfect secrecy can be achieved if each plaintext symbol is encrypted with a key randomly chosen from a key domain.
- This idea is used in a cipher called one-time pad, invented by Vernam.

One Time Pad

- Similar to Vigenere, but use random key as long as plaintext
- Only known scheme that is unbreakable (unconditional security or perfect security)
  - Ciphertext has no statistical relationship with plaintext
  - Given two potential plaintext messages, attacker cannot identify the correct message

A cipher system has perfect secrecy if the ciphertext gives the cryptanalyst no information about the key. The one time pad achieves perfect secrecy.

Continued

- Mauborgne suggested using a random key that is as long as the message, so that the key need not be repeated.
- In addition, the key is to be used to encrypt and decrypt a single message, and then is discarded.
- Each new message requires a new key of the same length as the new message.
- Such a scheme, known as a **one-time pad,** is unbreakable.

- Two practical limitations:
  1. Difficult to provide large number of random keys
  2. Distributing unique long random keys is difficult
- Limited practical use

Ciphertext

```
ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
```

We now show two different decryptions using two different keys:

```
ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
key:        pxlmvmsydofuyrvzwc tnlebnecvgdupahfzzlmnyih
plaintext:  mr mustard with the candlestick in the hall

ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
key:        pftgpmiydgaxgoufhklllmhsqdqogtewbqfgyovuhwt
plaintext:  miss scarlet with the knife in the library
```

So, for using random key, the cryptanalyst will be confused.

Vernam Cipher

- The ultimate defense against such a cryptanalysis is to choose a keyword that is as long as the plaintext and has no statistical relationship to it.

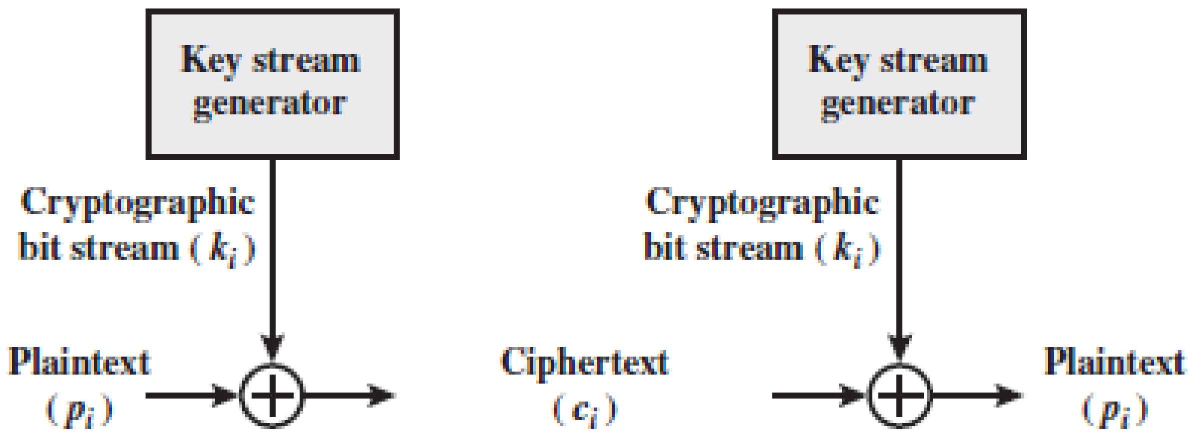


Figure 2.7 Vernam Cipher

His system works on binary data (bits) rather than letters. The system can be expressed succinctly as follows

$$c_i = p_i \oplus k_i$$

where

p_i = ith binary digit of plaintext
k_i = ith binary digit of key
c_i = ith binary digit of ciphertext
$\oplus$ = exclusive-or (XOR) operation

$$p_i = c_i \oplus k_i$$

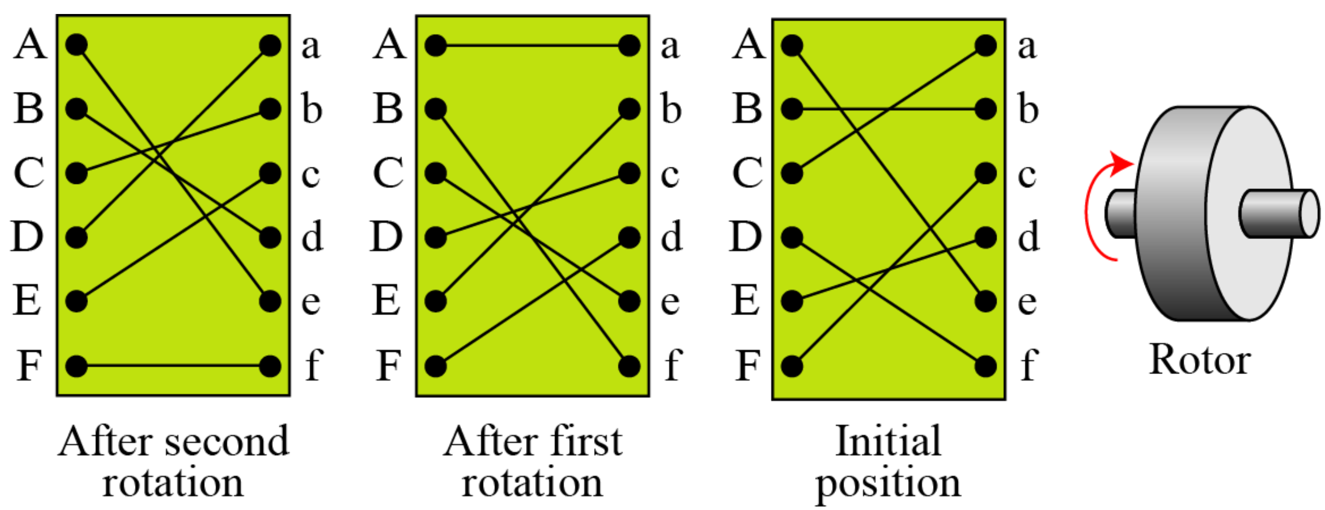
Vernam proposed the use of a running loop of tape that eventually repeated the key, so that in fact the system worked with a very long but repeating keyword.

Rotor Cipher

Although one-tme pad is not practical, one step toward more secured encipherment is rotor cipher.

It uses the idea behind monalphabetic substitution, but changes the mapping between plaintext and cipphertext characters for each plaintext character.

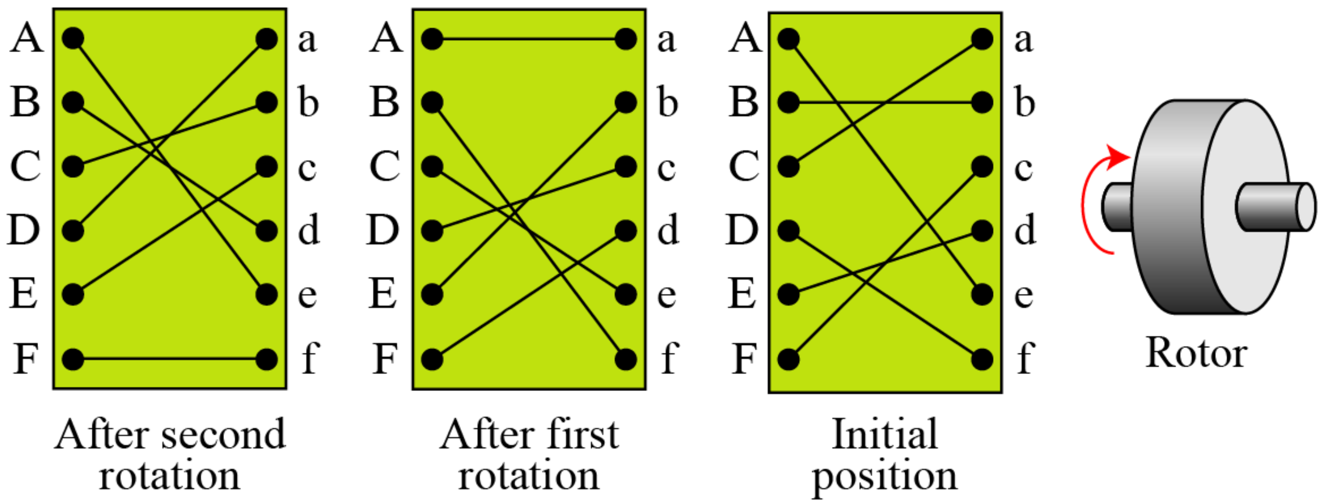
Figure A rotor cipher



Rotor Cipher

It uuses only 6 letter, but actual rotor uses 26 letters.
Initial setting is the secret key.

'bee' encrypts as "BAA" if rotor is stationery,but become "BCA" as rotates.



Enigma Machine

Originally degined by Sherbius, modified by Geerman army and extensively used in WWII.

‘.

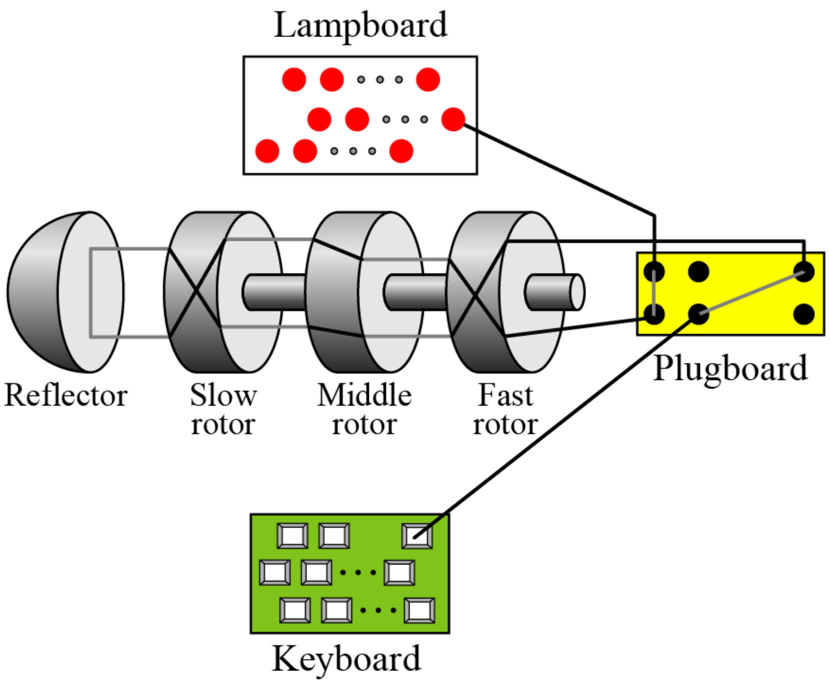


Figure A schematic of the Enigma machine

Rotor machines

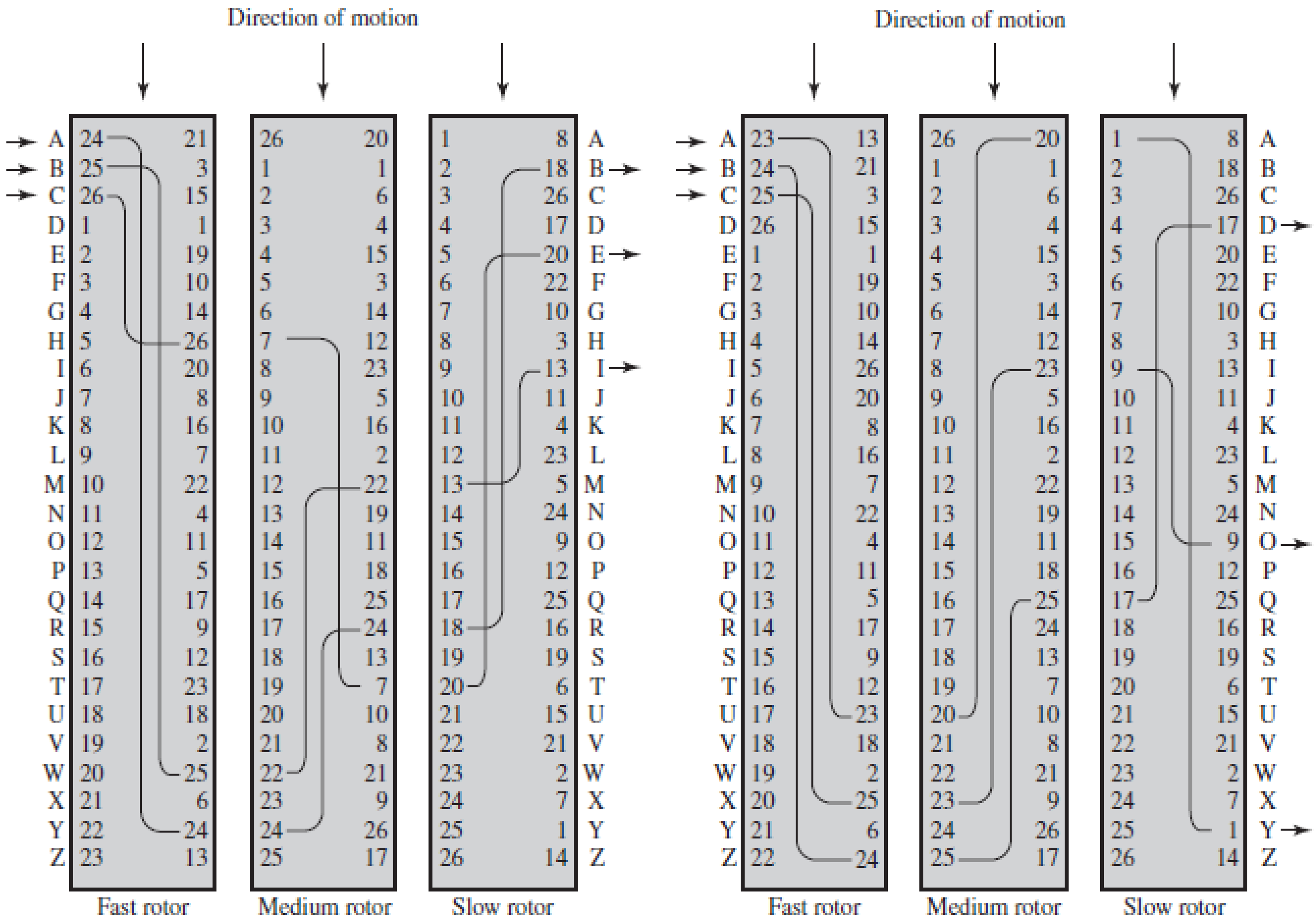
- The machine consists of a set of independently rotating cylinders through which electrical pulses can flow.

- Each cylinder has 26 input pins and 26 output pins, with internal wiring that connects each input pin to a unique output pin.

- After each input key is depressed, the cylinder rotates one position, so that the internal connections are shifted accordingly.

- After 26 letters of plaintext, the cylinder would be back to the initial position.

- For every complete rotation of the inner cylinder, the middle cylinder rotates one pin position. Finally, for every complete rotation of the middle cylinder, the outer cylinder rotates one pin position.
- Thus, a given setting of a 5-rotor machine is equivalent to a Vigenère cipher with a key length of 11,881,376.



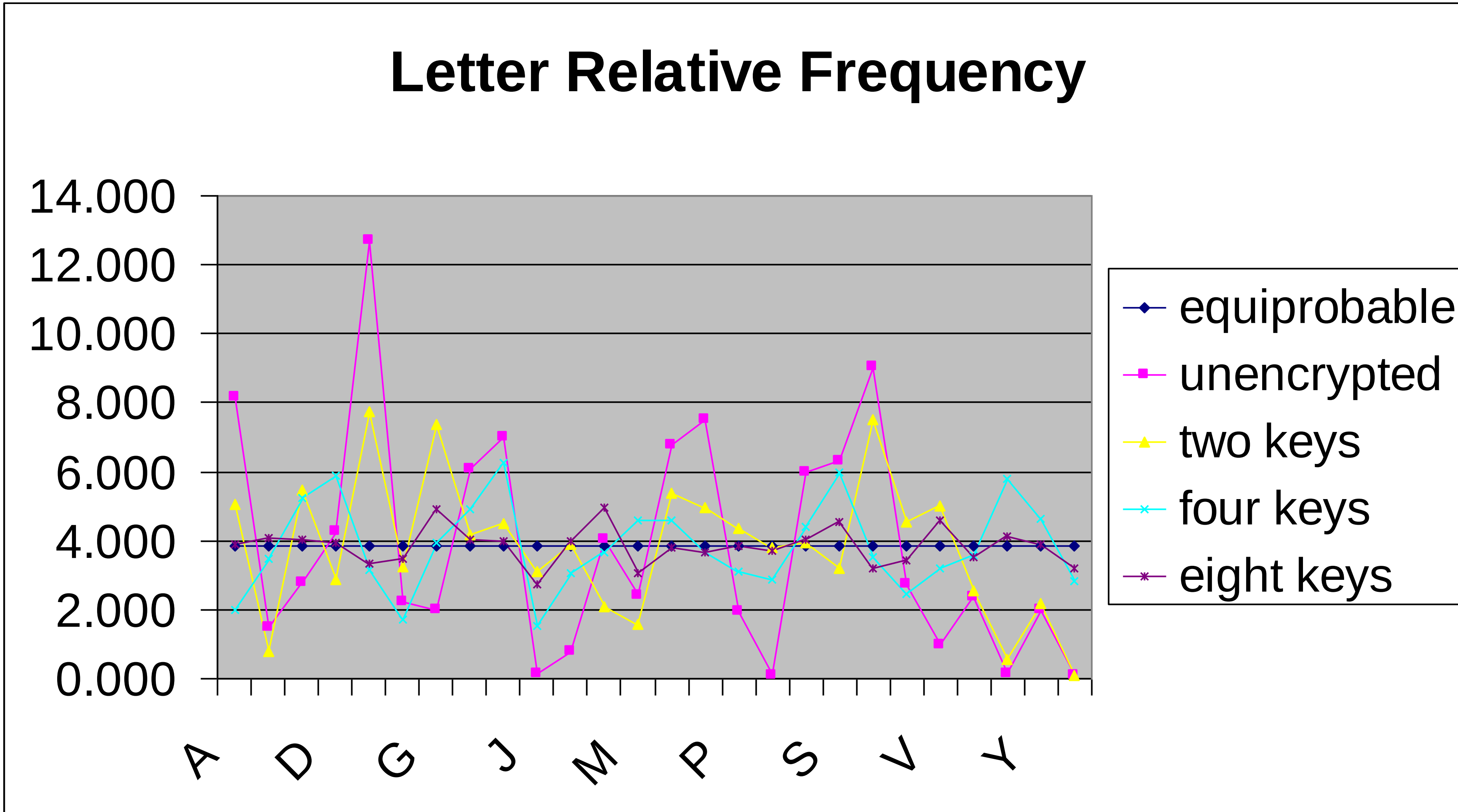
(a) Initial setting (b) Setting after one keystroke

Figure 2.8 Three-Rotor Machine with Wiring Represented by Numbered Contacts

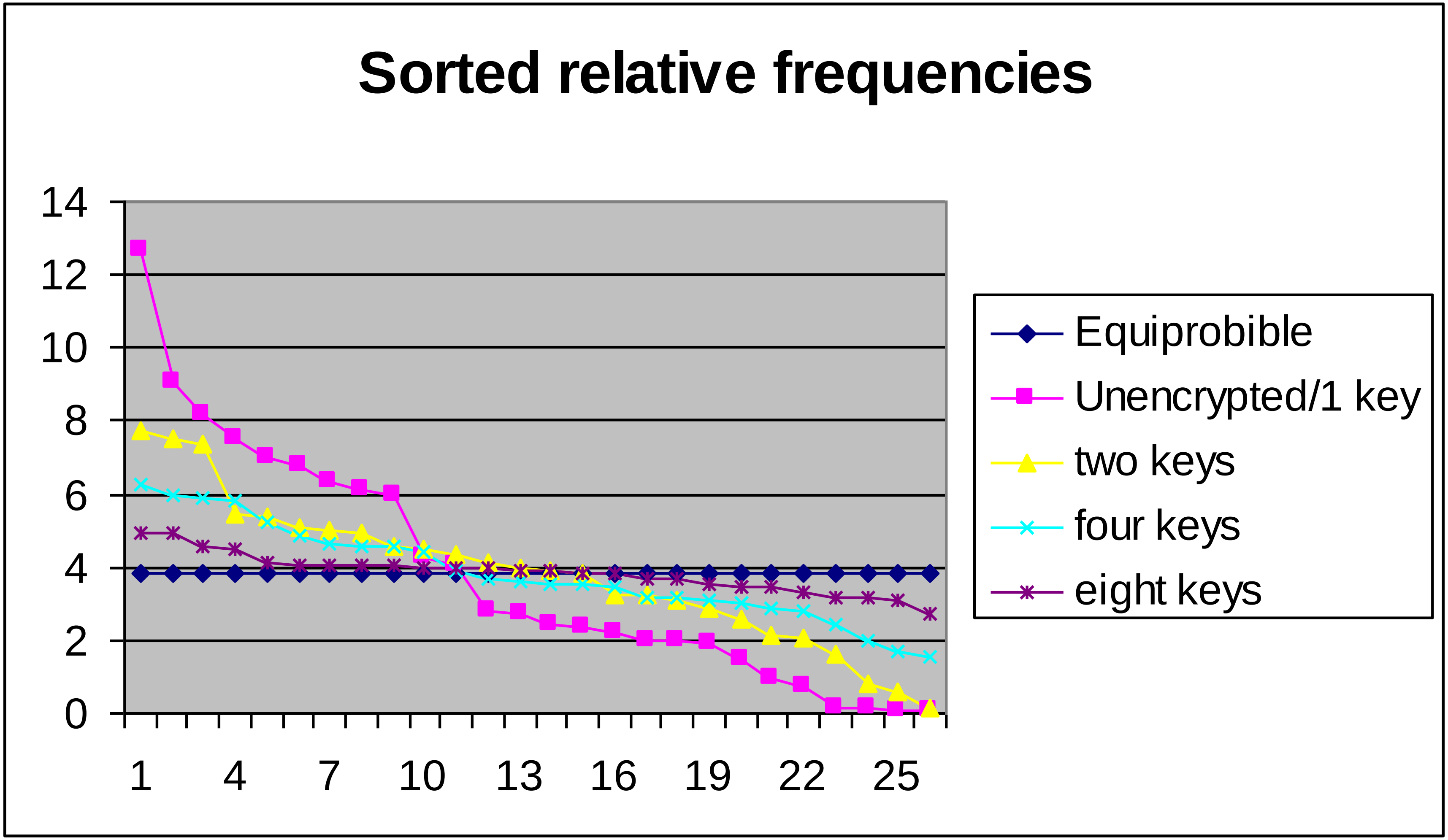
Poly-alphabetic Ciphers Summary

- **polyalphabetic substitution 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
- use each alphabet in turn
- repeat from start after end of key is reached

Frequencies After Polyalphabetic Encryption



Frequencies After Polyalphabetic Encryption



Transposition Ciphers

A transposition cipher does not substitute one symbol for another, instead it changes the location of the symbols.

A transposition cipher reorders symbols.

Topics:

Keyless Transposition Ciphers
Keyed Transposition Ciphers
Combining Two Approaches

Keyless Transposition Ciphers

Simple transposition ciphers, which were used in the past, are keyless.

Example

A good example of a keyless cipher using the first method is the rail fence cipher. The ciphertext is created reading the pattern row by row. For example, to send the message “Meet me at the park” to Bob, Alice writes

```
m       e       m       a       t       e       a       k
  ↘   ↗   ↘   ↗   ↘   ↗   ↘   ↗   ↘   ↗   ↘   ↗   ↘   ↗
    e       t       e       t       h       p       r
```

She then creates the ciphertext “MEMATEAKETETHPR”.

Example

Alice and Bob can agree on the number of columns and use the second method. Alice writes the same plaintext, row by row, in a table of four columns.

m	**e**	**e**	**t**
m	**e**	**a**	**t**
t	**h**	**e**	**p**
a	**r**	**k**	

She then creates the ciphertext "MMTAEEHREAEKTTP".

Example

The cipher in Example is actually a transposition cipher. The following shows the permutation of each character in the plaintext into the ciphertext based on the positions.

01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
01	05	09	13	02	06	10	13	03	07	11	15	04	08	12

The second character in the plaintext has moved to the fifth position in the ciphertext; the third character has moved to the ninth position; and so on. Although the characters are permuted,
there is a pattern in the permutation: (01, 05, 09, 13), (02, 06, 10, 13), (03, 07, 11, 15), and (08, 12). In each section, the difference between the two adjacent numbers is 4.

Keyed Transposition Ciphers

The keyless ciphers permute the characters by using writing plaintext in one way and reading it in another way The permutation is done on the whole plaintext to create the whole ciphertext.

Another method is to divide the plaintext into groups of predetermined size, called blocks, and then use a key to permute the characters in each block separately.

Example

Alice needs to send the message "Enemy attacks tonight" to Bob..

e	n	e	m	y	a	t	t	a	c	k	s	t	o	n	i	g	h	t	z

The key used for encryption and decryption is a permutation key, which shows how the character are permuted.

Encryption ↓

3	1	4	5	2
1	2	3	4	5

↑ Decryption

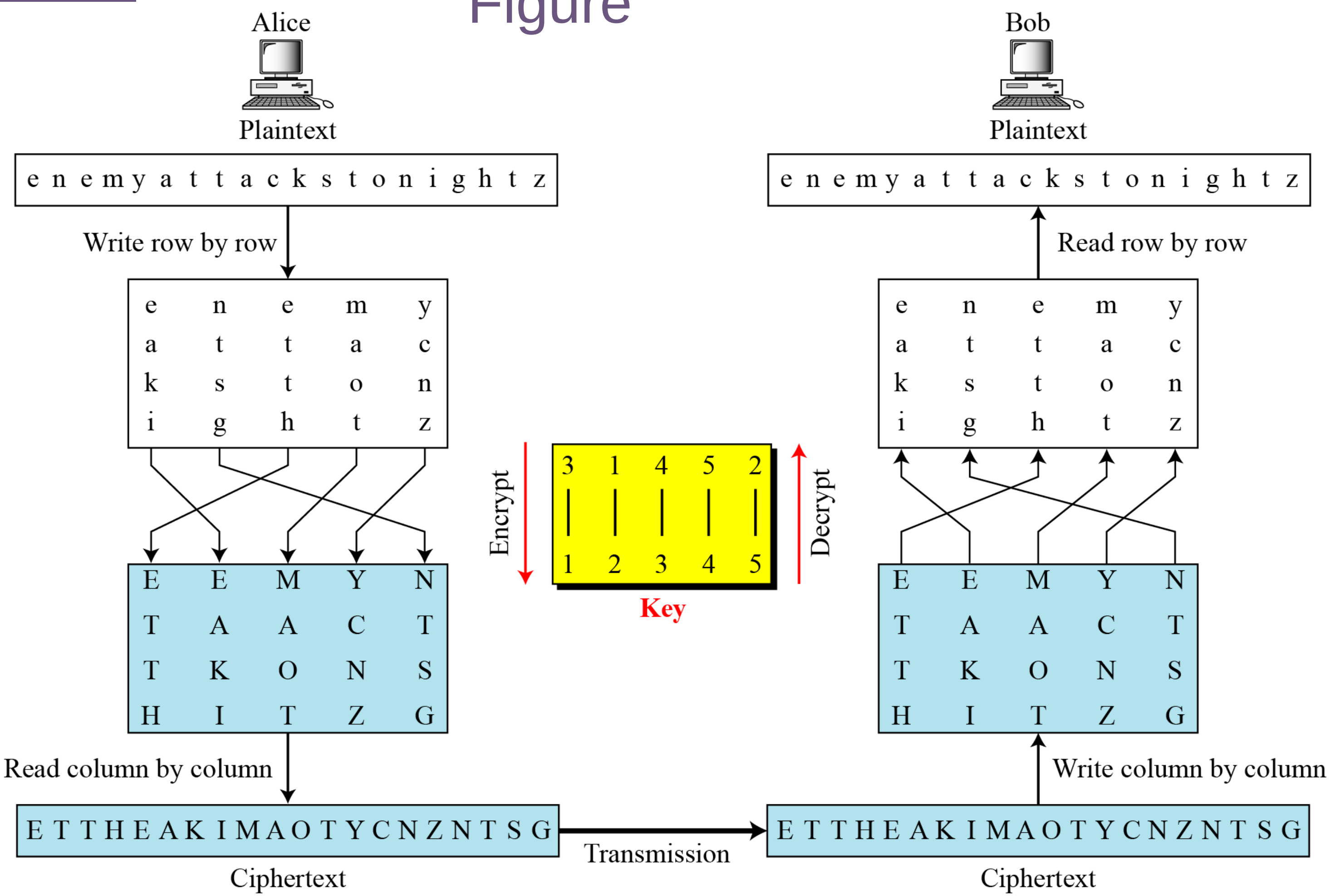
The permutation yields

E	**E**	**M**	**Y**	**N**	**T**	**A**	**A**	**C**	**T**	**T**	**K**	**O**	**N**	**S**	**H**	**I**	**T**	**Z**	**G**

Combining Two Approaches

Example

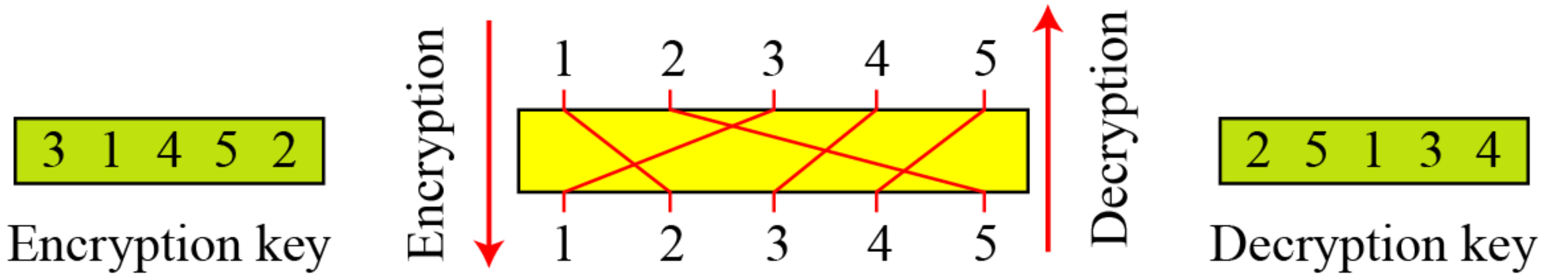
Figure



Keys

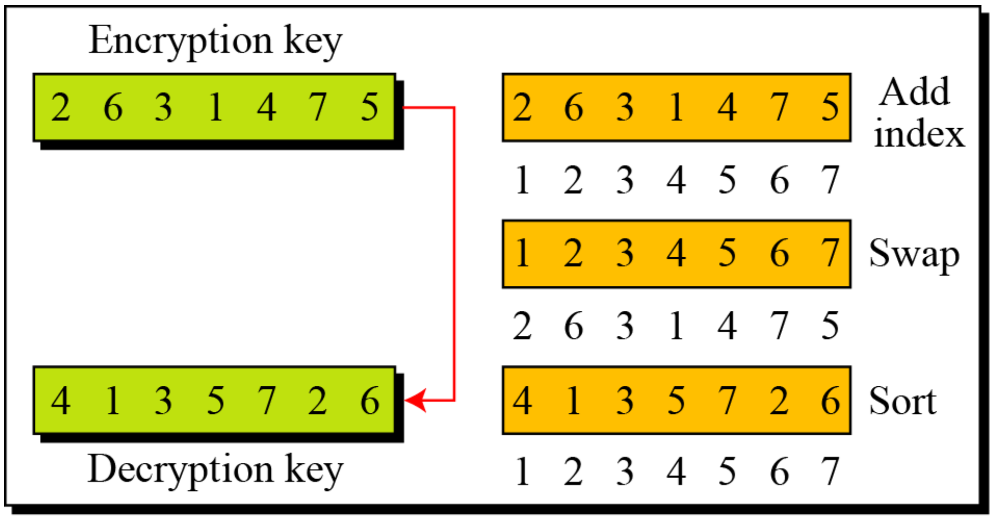
In Example , a single key was used in two directions for the column exchange: downward for encryption, upward for decryption. It is customary to create two keys.

Figure Encryption/decryption keys in transpositional ciphers



Continued

Figure Key inversion in a transposition cipher



a. Manual process

```
Given: EncKey [index]
index ← 1
while (index ≤ Column)
  {
    DecKey[EncKey[index]] ← index
    index ← index + 1
  }
Return : DecKey [index]
```

b. Algorithm

Continued

We can use matrices to show the encryption/decryption process for a transposition cipher.

Example

Figure Representation of the key as a matrix in the transposition cipher

$$\underset{\text{Plaintext}}{\begin{bmatrix} 04 & 13 & 04 & 12 & 24 \\ 00 & 19 & 19 & 00 & 02 \\ 10 & 18 & 19 & 14 & 13 \\ 08 & 06 & 07 & 19 & 25 \end{bmatrix}} \times \underset{\text{Encryption key}}{\overset{\begin{matrix} 3 & 1 & 4 & 5 & 2 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{matrix}}{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}}} = \underset{\text{Ciphertext}}{\begin{bmatrix} 04 & 04 & 12 & 24 & 13 \\ 19 & 00 & 00 & 02 & 19 \\ 19 & 10 & 14 & 13 & 18 \\ 07 & 08 & 19 & 25 & 06 \end{bmatrix}}$$

Continued

Figure shows the encryption process. Multiplying the 4 × 5 plaintext matrix by the 5 × 5 encryption key gives the 4 × 5 ciphertext matrix.

Figure Representation of the key as a matrix in the transposition cipher

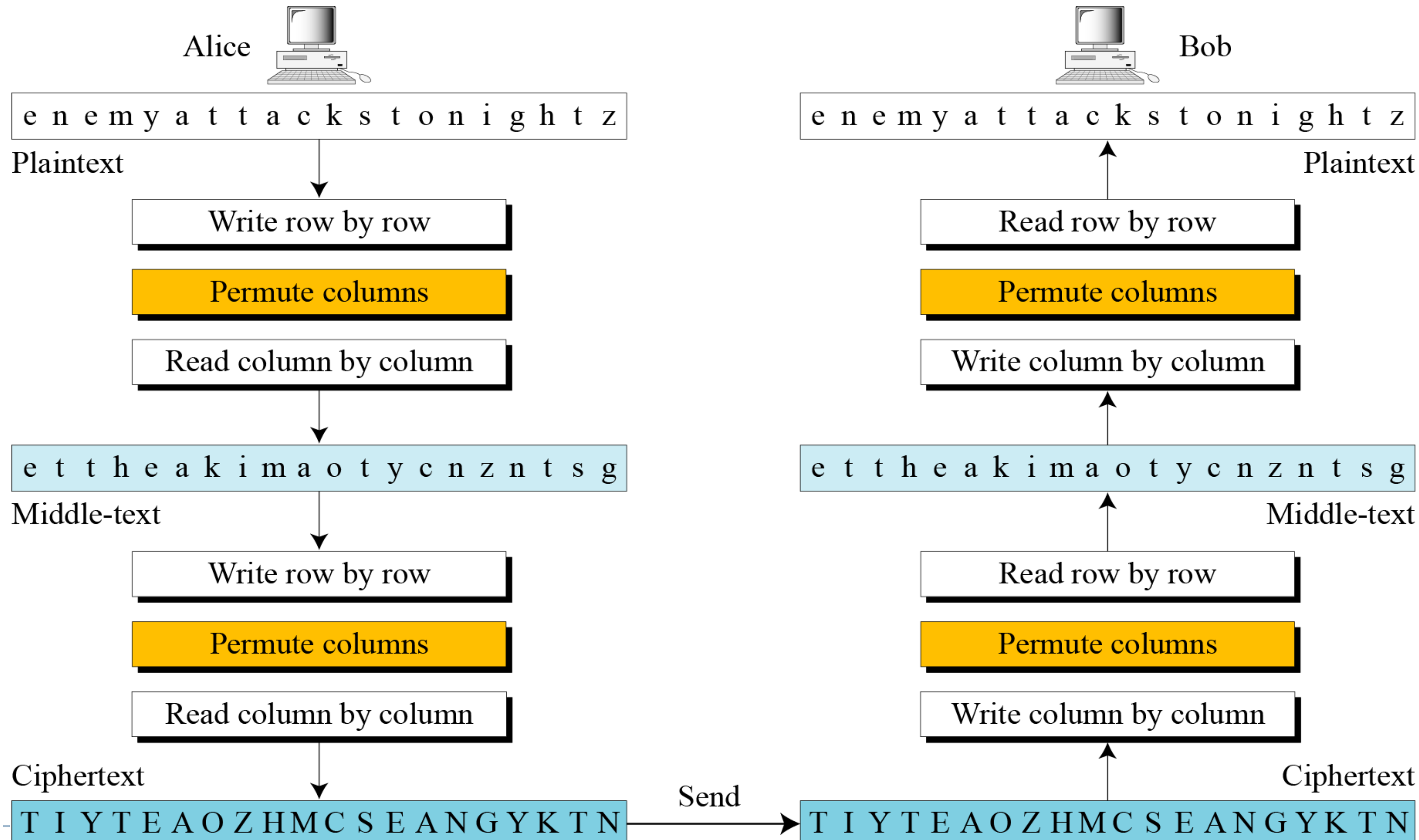
$$
\begin{array}{ccccc} 3 & 1 & 4 & 5 & 2 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{array}
$$

$$
\underset{\text{Plaintext}}{\begin{bmatrix} 04 & 13 & 04 & 12 & 24 \\ 00 & 19 & 19 & 00 & 02 \\ 10 & 18 & 19 & 14 & 13 \\ 08 & 06 & 07 & 19 & 25 \end{bmatrix}} \times \underset{\text{Encryption key}}{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}} = \underset{\text{Ciphertext}}{\begin{bmatrix} 04 & 04 & 12 & 24 & 13 \\ 19 & 00 & 00 & 02 & 19 \\ 19 & 10 & 14 & 13 & 18 \\ 07 & 08 & 19 & 25 & 06 \end{bmatrix}}
$$

Continued

Double Transposition Ciphers

Figure Double transposition cipher



Stream and Block Ciphers

The literature divides the symmetric ciphers into two broad categories: stream ciphers and block ciphers. Although the definitions are normally applied to modern ciphers, this categorization also applies to traditional ciphers.

Topics :

Stream Ciphers
Block Ciphers
Combination

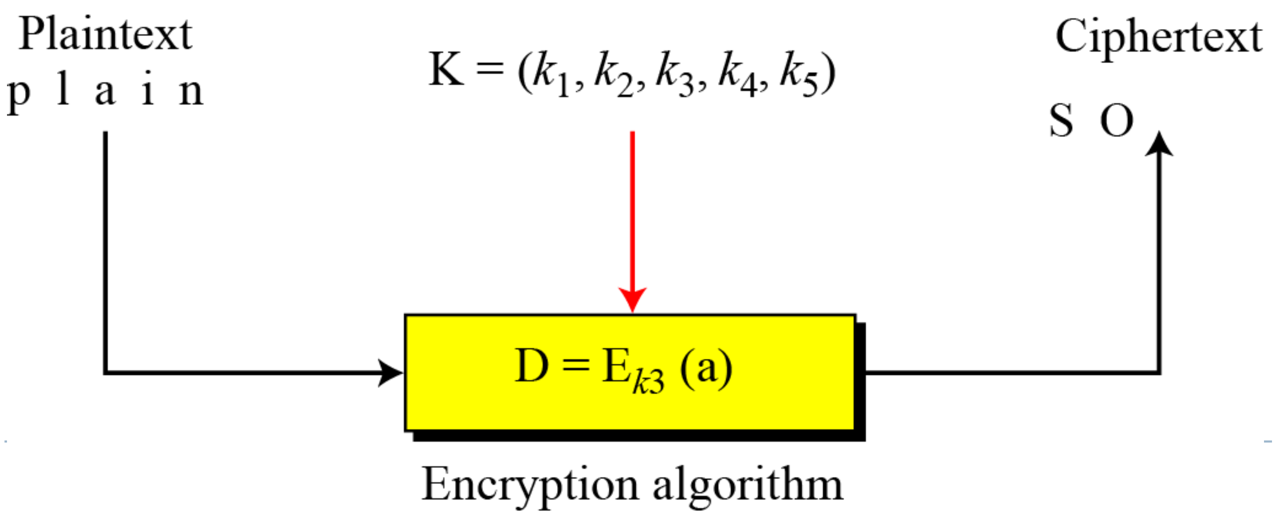
Stream Ciphers

Call the plaintext stream P, the ciphertext stream C, and the key stream K.

$$P = P_1P_2P_3, \ldots \qquad C = C_1C_2C_3, \ldots \qquad K = (k_1, k_2, k_3, \ldots)$$

$$C_1 = E_{k1}(P_1) \qquad C_2 = E_{k2}(P_2) \qquad C_3 = E_{k3}(P_3) \ \ldots$$

Figure Stream cipher



Continued

Example

Additive ciphers can be categorized as stream ciphers in which the key stream is the repeated value of the key. In other words, the key stream is considered as a predetermined stream of keys or K = (k, k, …, k). In this cipher, however, each character in the ciphertext depends only on the corresponding character in the plaintext, because the key stream is generated independently.

Example

The monoalphabetic substitution ciphers discussed in this chapter are also stream ciphers. However, each value of the key stream in this case is the mapping of the current plaintext character to the corresponding ciphertext character in the mapping table.

Continued

Example

Vigenere ciphers are also stream ciphers according to the definition. In this case, the key stream is a repetition of m values, where m is the size of the keyword. In other words,

$$K = (k_1, k_2, \ldots k_\mathrm{m}, k_1, k_2, \ldots k_m, \ldots)$$

Example

We can establish a criterion to divide stream ciphers based on their key streams. We can say that a stream cipher is a monoalphabetic cipher if the value of k_i does not depend on the position of the plaintext character in the plaintext stream; otherwise, the cipher is polyalphabetic.

Continued

- Additive ciphers are definitely monoalphabetic because k_i in the key stream is fixed; it does not depend on the position of the character in the plaintext.

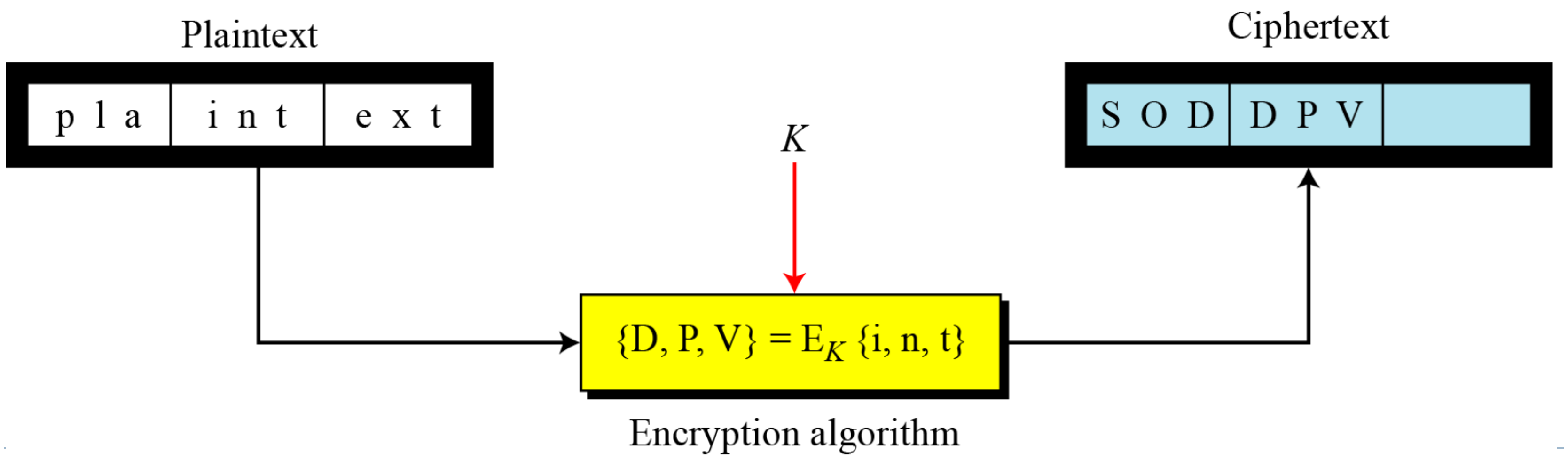
- Monoalphabetic substitution ciphers are monoalphabetic because k_i does not depend on the position of the corresponding character in the plaintext stream; it depends only on the value of the plaintext character.

- Vigenere ciphers are polyalphabetic ciphers because k_i definitely depends on the position of the plaintext character. However, the dependency is cyclic. The key is the same for two characters m positions apart.

Stream Ciphers

In a block cipher, a group of plaintext symbols of size m (m > 1) are encrypted together creating a group of ciphertext of the same size. A single key is used to encrypt the whole block even if the key is made of multiple values. Figure 3.27 shows the concept of a block cipher.

Figure 3.27 Block cipher



Continued

Example

Playfair ciphers are block ciphers. The size of the block is m = 2. Two characters are encrypted together.

Example

Hill ciphers are block ciphers. A block of plaintext, of size 2 or more is encrypted together using a single key (a matrix). In these ciphers, the value of each character in the ciphertext depends on
all the values of the characters in the plaintext. Although the key is made of $m \times m$ values, it is considered as a single key.

Example

From the definition of the block cipher, it is clear that every block cipher is a polyalphabetic cipher because each character in a ciphertext block depends on all characters in the plaintext block.

Combination

In practice, blocks of plaintext are encrypted individually, but they use a stream of keys to encrypt the whole message block by block. In other words, the cipher is a block cipher when looking at the individual blocks, but it is a stream cipher when looking at the whole message considering each block as a single unit.

Product Ciphers

- ciphers using substitutions or transpositions are not secure because of language characteristics
- hence consider using several ciphers in succession to make harder, but:
  - 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
- this is bridge from classical to modern ciphers

- The transposition cipher can be made significantly more secure by performing more than one stage of transposition. If we apply previous mapping again:

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

To visualize the result of this double transposition, designate the letters in the original plaintext message by the numbers designating their position.

```
01 02 03 04 05 06 07 08 09 10 11 12 13 14
15 16 17 18 19 20 21 22 23 24 25 26 27 28
```

- After the first transposition, we have

```
03 10 17 24 04 11 18 25 02 09 16 23 01 08
15 22 05 12 19 26 06 13 20 27 07 14 21 28
```

But after the second transposition, we have

```
17 09 05 27 24 16 12 07 10 02 22 20 03 25
15 13 04 23 19 14 11 01 26 21 18 08 06 28
```

- The example just given suggests that multiple stages of encryption can produce an algorithm that is significantly more difficult to cryptanalyze
- This is as true of substitution ciphers as it is of transposition ciphers.

S-box

- We can extend the substitution box idea to binary words.
- Here's a 4×4 S-box that maps 4 bits to 4 bits:

S	**00**	**01**	**10**	**11**
00	0011	1000	1111	0001
01	1010	0110	0101	1011
10	1110	1101	0100	0010
11	0111	0000	1001	1100

S	**0**	**1**	**2**	**3**
0	3	8	15	1
1	10	6	5	11
2	14	13	4	2
3	7	0	9	12

- Examples:
$$\begin{array}{ccc} 0000 & \longrightarrow & 0011 \\ 0001 & \longrightarrow & 0100 \\ 1010 & \longrightarrow & 0100 \end{array}$$

P-box

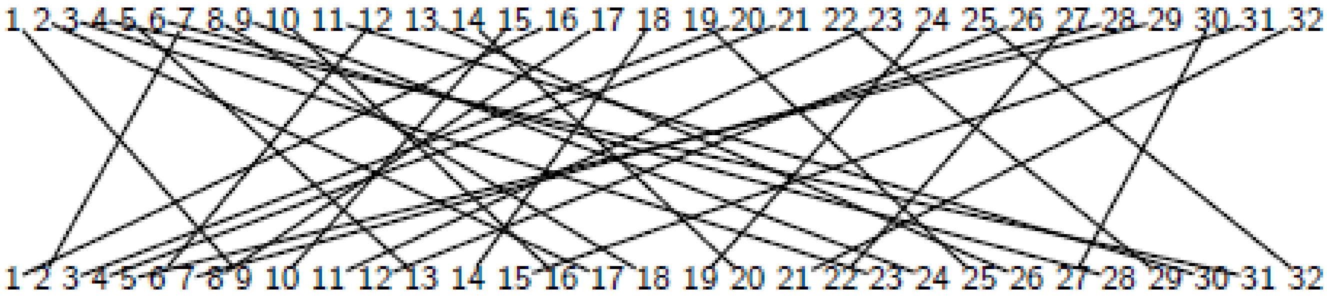
- We can extend the transposition cipher idea to binary words.
- Here's a 32-bit P-box that is used by the DES cipher:

P	moved to position							
1-8	9	17	23	31	13	28	2	18
9-16	24	16	30	6	26	20	10	1
17-24	8	14	25	3	4	29	11	19
25-32	32	12	22	7	5	27	15	21

P-box

P	moved to position							
1-8	9	17	23	31	13	28	2	18
9-16	24	16	30	6	26	20	10	1
17-24	8	14	25	3	4	29	11	19
25-32	32	12	22	7	5	27	15	21

⇕



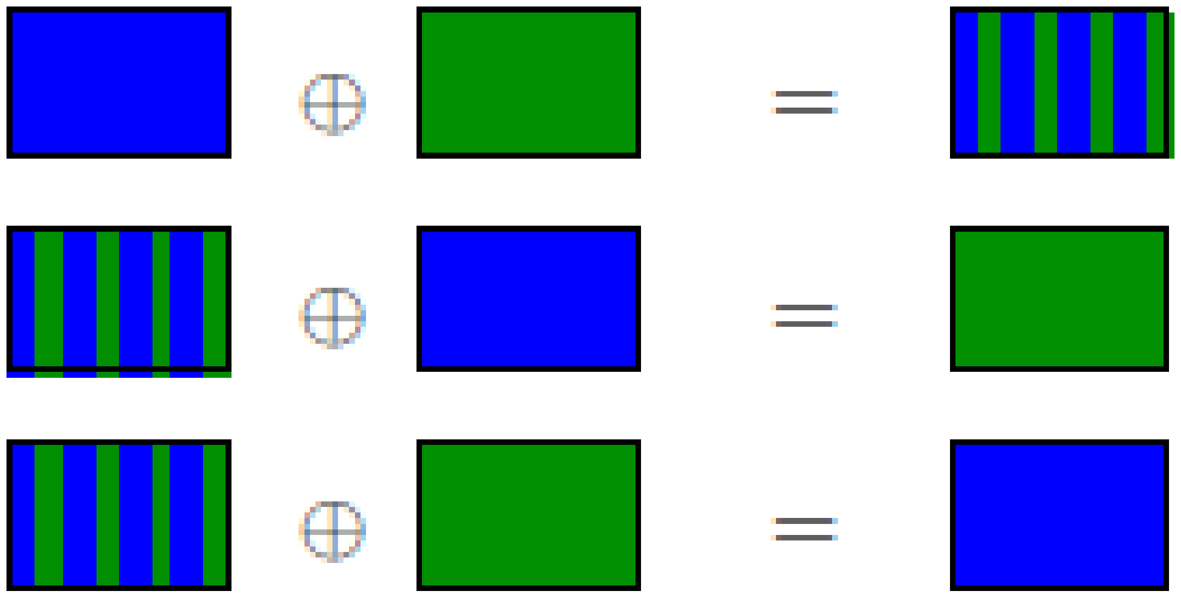
Exclusive-OR

$$
\begin{array}{ccccc|ccccccc}
0 & \oplus & 0 & = & 0 & a & \oplus & a & & & = & 0 \\
0 & \oplus & 1 & = & 1 & a & \oplus & b & \oplus & b & = & a \\
1 & \oplus & 0 & = & 1 & a & \oplus & a & \oplus & a & = & a \\
1 & \oplus & 1 & = & 0 & & & & & & &
\end{array}
$$

- Since xor-ing the same value twice gives us the original, we get a simple symmetric algorithm:

$$
\begin{aligned}
P \oplus K &= C \\
C \oplus K &= P
\end{aligned}
$$

Exclusive-OR



Steganography & Cryptography

- The methods of **steganography** **conceal the existence of the message, whereas the methods of** **cryptography** render the message unintelligible to outsiders by various transformations of the text

Steganography

- Hide a real message in a fake, but meaningful, message
- Assumes recipient knows the method of hiding
- Examples:
  - Selected letters in a document are marked to form the hidden message
  - Invisible ink (letters only become visible when exposed to a chemical or heat)
  - Using selected bits in images or videos to carry the message
- Advantages
  - Does not look like you are hiding anything
- Disadvantages
  - Once attacker knows your method, everything is lost
  - Can be ineffcient (need to send lot of information to carry small message)

Steganography Example

Dear George,

Greetings to all at Oxford. Many thanks for your
letter and for the Summer examination package.
All Entry Forms and Fee Forms should be ready
for final despatch to the Syndicate by Friday
20th or at the very latest, I'm told, by the 21st.
Admin has improved here, though there's room
for improvement still; just give us all two or three
more years and we'll really show you! Please
don't let these wretched 16+ proposals destroy
your basic O and A pattern. Certainly this
sort of change, if implemented immediately,
would bring chaos.

Sincerely yours.

Review

Probability distributions related to a cryptosystem

- Let us suppose that a cryptosystem is specified by $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$.
- Assume that it is possible to define a probability distribution on the plaintext space $\mathcal{P}$, the key space space $\mathcal{K}$.
- The probability distribution on $\mathcal{P}$ and $\mathcal{K}$ induces a probability distribution on $\mathcal{C}$.
- Let the random variables associated to plaintexts, keys and ciphertexts be X, Y and K respectively.
- The probability that $X = x$ is denoted by $\Pr[X = x]$.
- The probability that $K = k$ is denoted by $\Pr[K = k]$.
- The probability that $Y = y$ is denoted by $\Pr[Y = y]$.

Distribution of ciphertexts

- $C(k) = \{e_k(x) : x \in \mathcal{P}\}$ is the set of all possible ciphertexts.
- The probability distribution of ciphertexts is

$$\Pr[Y = y] = \sum_{\{k : y \in C(k)\}} \Pr[K = k] \Pr[x = d_k(y)]$$

- The probability distribution of ciphertexts given a plaintext is

$$\Pr[Y = y | X = x] = \sum_{\{k : x = d_k(y)\}} \Pr[K = k] .$$

Distribution of plaintexts given a ciphertext

- The probability distribution of plaintexts conditional to ciphertexts is

$$\Pr[X = x|Y = y] = \frac{\Pr[(X = x) \cap (Y = y)]}{\Pr[Y = y]}$$

$$= \frac{\Pr[X = x]\Pr[Y = y|X = x]}{\Pr[Y = y]}$$

$$= \frac{\Pr[X{=}x] \times \sum_{\{k:x=d_{k(y)}\}} \Pr[K{=}k]}{\sum_{\{k:y \in C(k)\}} \Pr[K{=}k]\Pr[X{=}d_k(y)]}$$

Computation of these probabilities

$$\Pr[Y=y] = \sum_{\{k\,:\,y\in C(k)\}} \Pr[(K=k)\cap(y=e_k(x))]$$

$$= \sum_{\{k\,:\,y\in C(k)\}} \Pr[(K=k)\cap(x=d_k(y))]$$

$$= \sum_{\{k\,:\,y\in C(k)\}} \Pr[K=k]\,\Pr[X=x=d_k(y)]$$

Computation of these probabilities

- Let $\mathcal{P} = \{a, b\}$ with $\Pr[X = a] = \frac{1}{4}, \Pr[X = b] = \frac{3}{4}$.
- Let $\mathcal{K} = \{k_1, k_2, k_3\}$ with $P[K = k_1] = \frac{1}{2}, \Pr[K = k_2] = \Pr[K = k_3] = \frac{1}{4}$.
- Let $\mathcal{C} = \{1,2,3,4\}$.
- The cryptosystem is represented by the following encryption matrix:

	a	b
k1	1	2
k2	2	3
k3	3	4

$\Pr[X = a] = \frac{1}{4}, \Pr[X = b] = \frac{3}{4}$

$P[K = k_1] = \frac{1}{2}$,
$\Pr[K = k_2] = \Pr[K = k_3] = \frac{1}{4}$.

$$\Pr[X = x|Y = y] = \frac{\Pr[(X = x) \cap (Y = y)]}{\Pr[Y = y]} = \frac{\Pr[X = x]\Pr[Y = y|X = x]}{\Pr[Y = y]}$$

Computation of these probabilities

	a	b
k1	1	2
k2	2	3
k3	3	4

$\Pr[X=a] = \frac{1}{4}, \Pr[X=b] = \frac{3}{4}$

$P[K=k_1] = \frac{1}{2},$
$\Pr[K=k_2] = \Pr[K=k_3] = \frac{1}{4}.$

$$\Pr[X=x|Y=y] = \frac{\Pr[(X=x)\cap(Y=y)]}{\Pr[Y=y]} = \frac{\Pr[X=x]\Pr[Y=y|X=x]}{\Pr[Y=y]}$$

$$\Pr[Y=1] = \Pr[K=k_1]\Pr[X=a] = \frac{1}{2}\cdot\frac{1}{4} = \frac{1}{8}$$

$$\Pr[Y=2] = \Pr[X=b]\Pr[K=k_1] + \Pr[X=a]\Pr[K=k_2] = \frac{3}{4}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{4} = \frac{3}{8} + \frac{1}{16}$$

Computation of these probabilities

$$\Pr[X=a \mid Y=2] = \frac{\Pr[X=a]\,\Pr[Y=2 \mid X=a]}{\Pr[Y=2]}$$

$$= \frac{\frac{1}{4} \cdot \frac{1}{4}}{\frac{7}{16}} = \frac{1}{7}.$$

Perfect Secrecy

- Perfect secrecy means that an adversary (Oscar) cannot get any information about the plaintext by observing the ciphertext.
- A precise formulation of this was given by Claude Elwood Shannon which is as follows:

  A cryptosystem has perfect secrecy if

  $$\mathbf{Pr}[X = x | Y = y] = \mathbf{Pr}[X = x]$$

  for all $x \in \mathcal{P}, y \in \mathcal{C}$.

Perfect secrecy and Shift Cipher

- Suppose that the 26 keys in the shift cipher are used with equal probability $\frac{1}{26}$. Then for any plaintext probability distribution, the Shift Cipher has perfect secrecy.
- $\Pr[Y = y] = \sum_{k \in \mathbb{Z}_{26}} \Pr[K = k] \Pr[X = d_k(y)]$
$$= \sum_{k \in \mathbb{Z}_{26}} \frac{1}{26} \Pr[X = y - k] = \frac{1}{26} \sum_{k \in \mathbb{Z}_{26}} \Pr[X = y - k] = \frac{1}{26}.$$
- $\Pr[Y = y | X = x] = \Pr[K = (y - x) \bmod 26] = \frac{1}{26}$.
- $\Pr[X = x | Y = y] = \frac{\Pr[X=x]\Pr[Y=y|X=x]}{\Pr[Y=y]} = \frac{\Pr[X=x]\frac{1}{26}}{\frac{1}{26}} = \Pr[X = x]$.

Computation of these probabilities

- $\Pr[Y = 1] = \frac{1}{8}$; $\Pr[Y = 2] = \frac{3}{8} + \frac{1}{16} = \frac{7}{16}$

  $\Pr[Y = 3] = \frac{3}{16} + \frac{1}{16} = \frac{1}{4}$; $\Pr[Y = 4] = \frac{3}{16}$.
- $\Pr[X = a|Y = 1] = 1$; $\Pr[X = a|Y = 2] = \frac{1}{7}$;

  $\Pr[X = a|Y = 3] = \frac{1}{4}$; $\Pr[X = a|Y = 4] = 0$.
- $\Pr[X = b|Y = 1] = 0$; $\Pr[X = b|Y = 2] = \frac{6}{7}$;

  $\Pr[X = b|Y = 3] = \frac{3}{4}$; $\Pr[X = b|Y = 4] = 1$.

Summary

- 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
  - steganography