Computer Security

Lecture 3: Cryptography

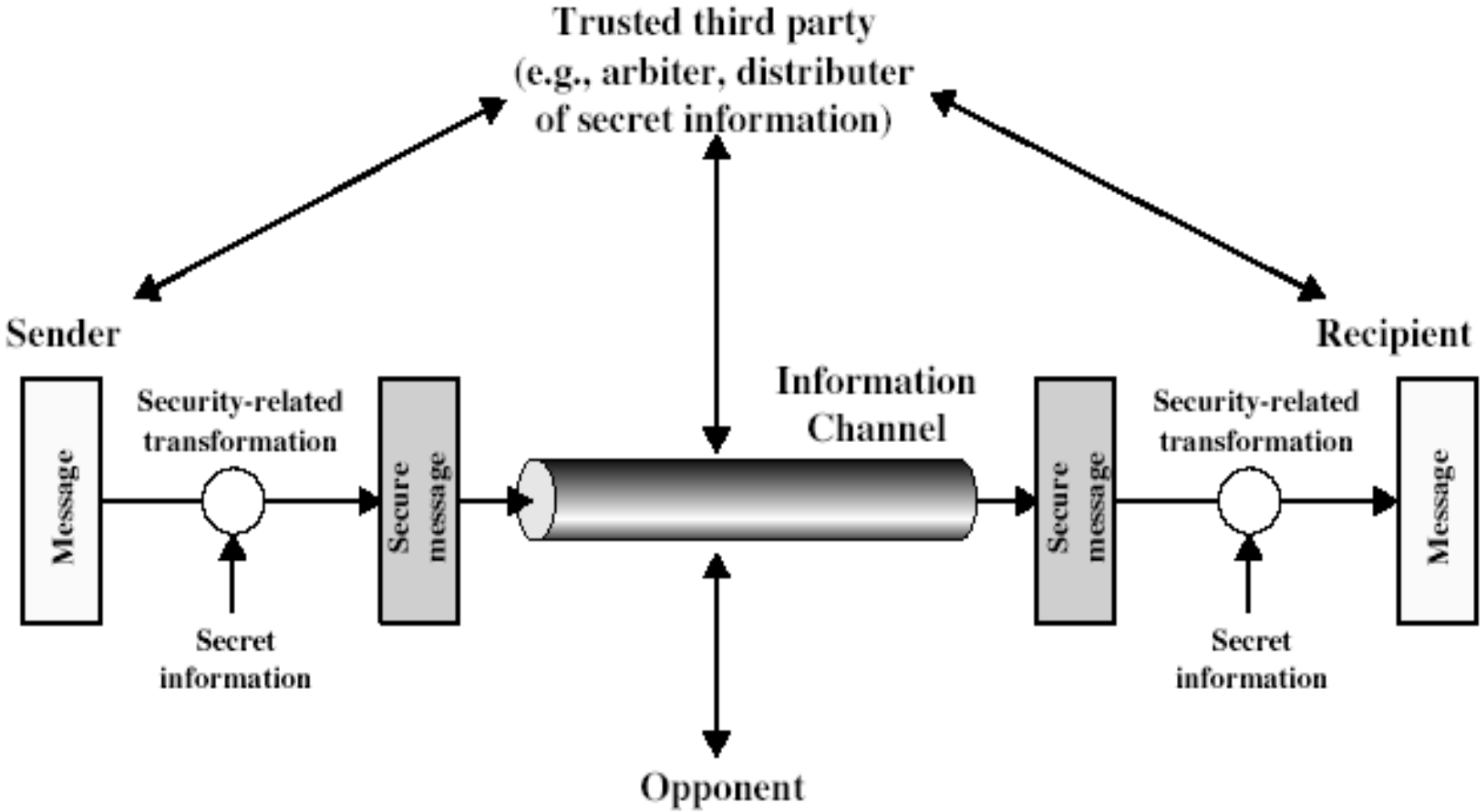
Prof. Dr. Sadeeq Jan

Department of Computer Systems Engineering

University of Engineering and Technology Peshawar

CRYPTOGRAPHY

Model for Network Security



Crypto Technology

- **Send a message that will be understandable only by the receiver**
- **Two ways to hide information**
  - **Stegnography**
    - Hiding the existence of message
    - Invisible ink, hide in picture
  - **Cryptology**
    - **Cryptogrphay**
      - Change plain text into unreadeable with the correct key
    - **Cryptanalysis**
      - Read the plain text without the correct key, cracking the cipher

Basic Terminology

- **plaintext** - the original message
- **ciphertext** - the 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 plaintext from ciphertext
- **cryptography** – art of achieving security by encoding messages to make then non-readable
- **cryptanalysis (codebreaking)** - the study of principles/ methods of deciphering ciphertext *without* knowing key
- **cryptology** - the field of both cryptography and cryptanalysis

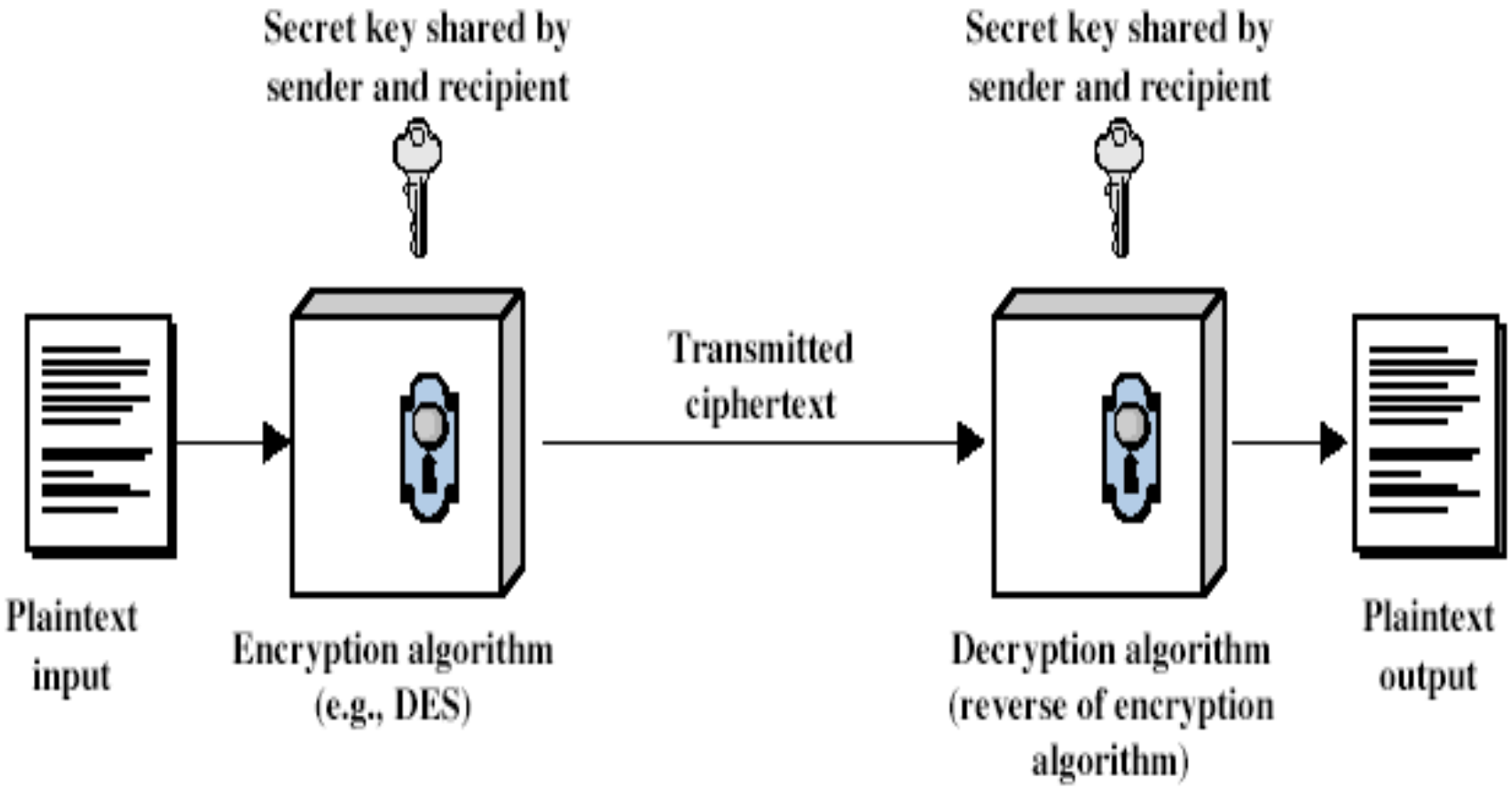
Authentication, Integirty and NonRepudiation

- Authentication
  - It should be possible for the receiver of a message to verify its origin; an intruder should not be able to masquerade as someone else.
- Integrity
  - It should be possible for the receiver of a message to verify that it has not been modified in transit; an intruder should not be able to substitute a false message for a legitimate one.
- Non-repudiation
  - A sender should not be able to falsely deny later that he sent a message.

Symmetric Encryption

- conventional /private-key /single-key encryption
- sender and recipient share a common key
- all classical encryption algorithms are private-key
- was only type prior to invention of public-key (Asymmetric key encryption) in 1970's

Symmetric Cipher Model

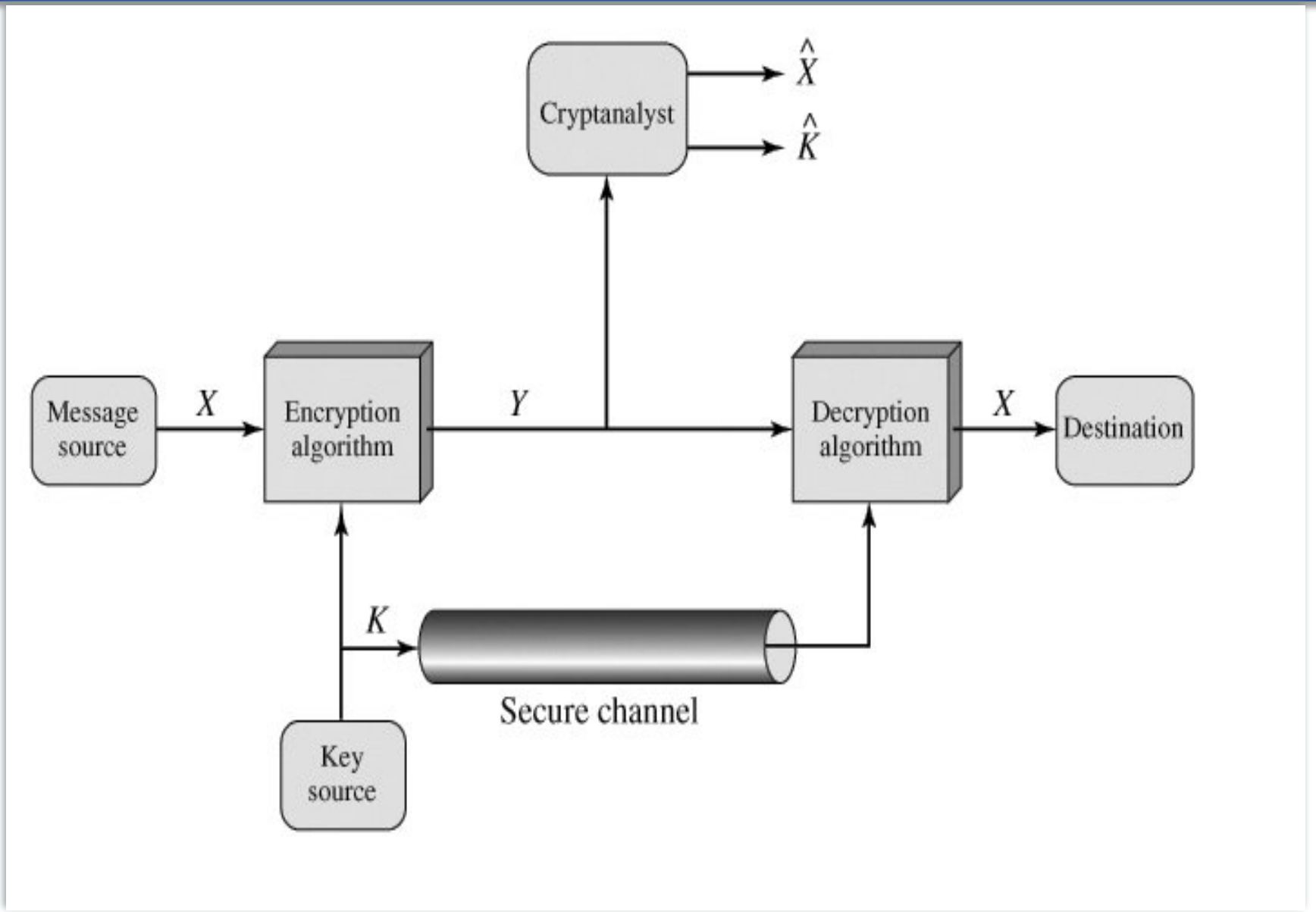


Requirements

- two requirements for secure use of symmetric encryption:
  - a strong encryption algorithm
  - a secret key known only to sender / receiver

    $E_K(M)=C$

    $D_K(C)=M$
- assume encryption algorithm is known
- implies a secure channel to distribute key


Cryptanalyst
$\hat{X}$
$\hat{K}$
Message source
X
Encryption algorithm
Y
Decryption algorithm
X
Destination
K
Secure channel
Key source

Cryptography

- can characterize by:
  - type of encryption operations used
    - substitution / transposition / product
    - substitution, in which each element in the plaintext (bit, letter, group of bits or letters) is mapped into another element, and transposition, in which elements in the plaintext are rearranged
  - number of keys used
    - single-key or private / two-key or public
  - way in which plaintext is processed
    - Block ciphers / stream cipher
    - A *block cipher* processes the input one block of elements at a time, producing an output block for each input block.
    - A *stream cipher* processes the input elements continuously, producing output one element at a time, as it goes along.

Types of Cryptanalytic Attacks

- **ciphertext only**
  - The opponent possesses a string of ciphertext y.
- **known plaintext**
  - The opponent possesses a string of plaintext x, and the corresponding ciphertext string y.
- **chosen plaintext**
  - The opponent has obtained temporary access to the encryption machinery. He can choose a plaintext x and encrypt it to get the corresponding output y.
- **chosen ciphertext**
  - The opponent has obtained a temporary access to the decryption machinery. He can choose a ciphertext y and construct the corresponding plaintext x.

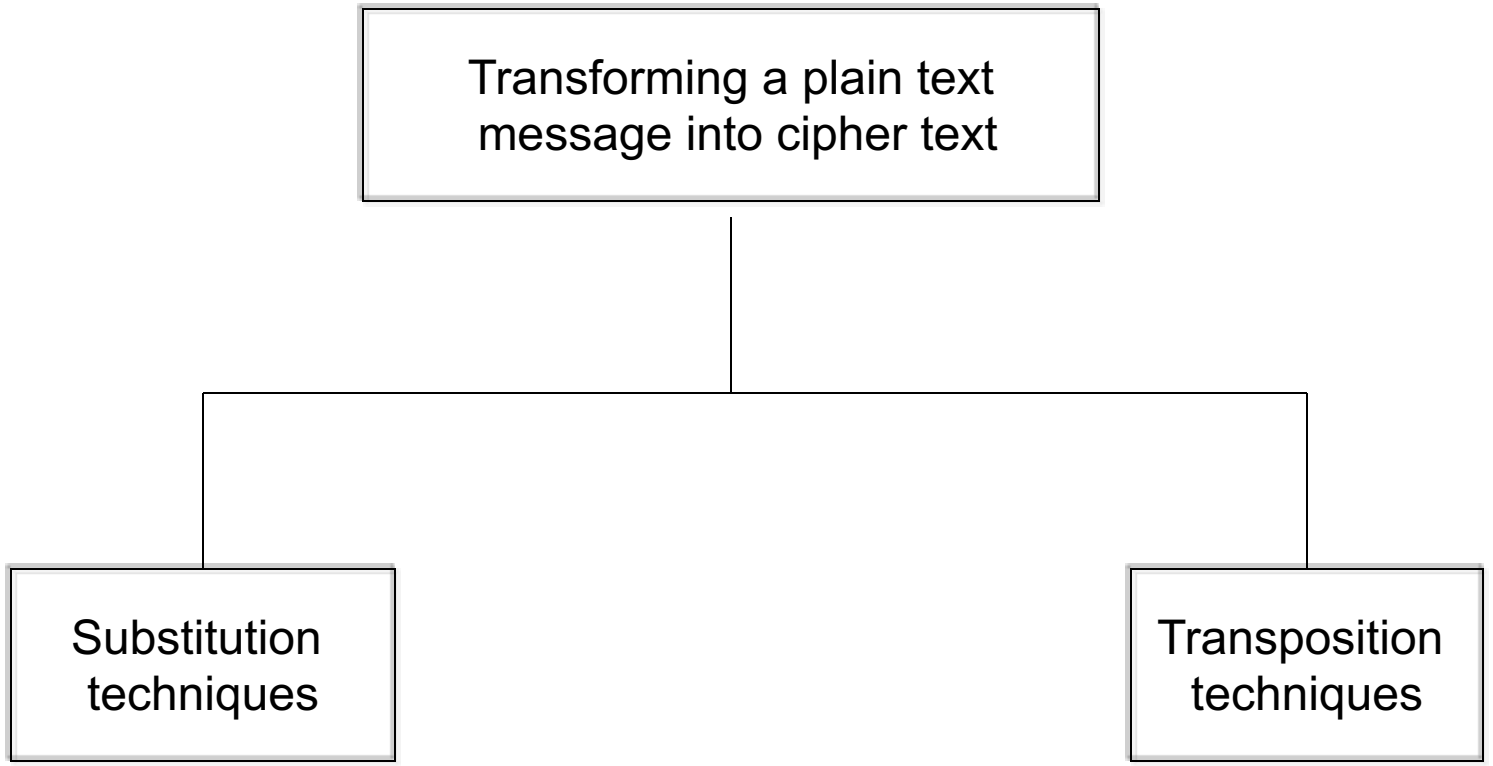
Brute Force Search

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

Key Size (bits)	Number of Alternative Keys	Time required at 1 encryption/μs	Time required at 10^6 encryptions/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

More Definitions

- **unconditional security**
  - no matter how much computer power is available, the cipher cannot be broken since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext
- **computational security**
  - given limited computing resources (eg time needed for calculations is greater than the useful lifetime of the information), the cipher cannot be broken


Transforming a plain text message into cipher text
Substitution techniques
Transposition techniques

Simple Substitution

- 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
- Key(alphabet) is arbitrary chosen. Must be written down.
  - A b c d e f g h I j k l m n o ….
  - A C B

Caesar Cipher

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

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

- Message = D a n g e r
- Ciphertext =

- Message = D a n g e r
- Ciphertext = G D Q J H N
- P+k = C
- K=8
- 8+20 = 28
- I = ??

Plaintext	0	1	2	3	4	5	6	7	8	9
Ciphertext	A	B	C	D	E	F	G	H	I	J
	D	E	F	G	H	I	J	K	L	M

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
0  1  2  3  4  5  6  7  8  9  10 11 12
n  o  p  q  r  s  t  u  v  w  x  y  Z
13 14 15 16 17 18 19 20 21 22 23 24 25
```

- then have Caesar cipher as:

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

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

Cryptanalysis of Caesar Cipher

- Knowing encryption/decryption algorithms
- only have 26 possible ciphers
  - A maps to A,B,..Z
- could simply try each in turn
- a **brute force search**
- given ciphertext, just try all shifts of letters
- do need to recognize when have plaintext
- eg. break ciphertext "GCUA VQ DTGCM"

Quiz 1

- Q1: Encrypt and Decrypt the following given Plain Text. Key = 6
  - Plain Text: "**Computer Security**"

- Q2: Break the ciphertext without the key.
  - "**GCUA VQ DTGCM**"

Quiz 1

- Q1: Break the ciphertext without the key.
  - "**Ymnx nx dtzw vzne**"

- Q2: Encrypt and Decrypt the following given Plain Text. Key =7
  - Plain Text: "**Computer Systems**"

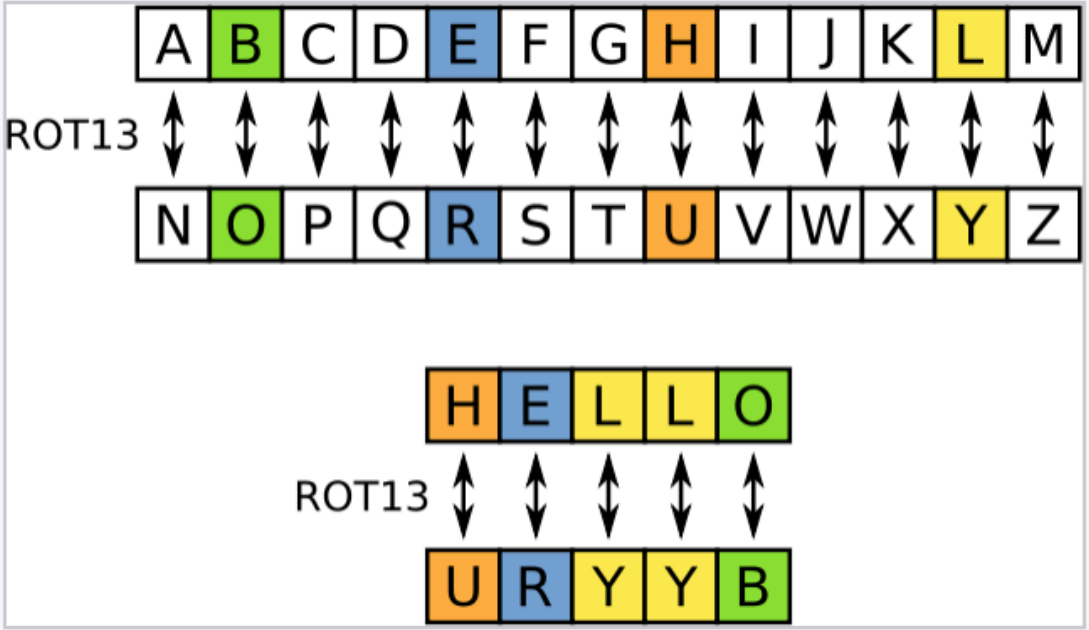
Figure 2.3. Brute-Force Cryptanalysis of Caesar Cipher

(This item is displayed on page 37 in the print version)

KEY	PHHW PH DIWHU WKH WRJD SDUWB
1	oggv og chvgt vjg vqic rctva
2	nffu nf bgufs uif uphb qbsuz
3	meet me after the toga party
4	ldds ld zesdq sgd snfz ozqsx
5	kccr kc ydrcp rfc rmey nyprw
6	jbbq jb xcqbo qeb qldx mxoqv
7	iaap ia wbpan pda pkcw lwnpu
8	hzzo hz vaozm ocz ojbv kvmot
9	gyyn gy uznyl nby niau julns
10	fxxm fx tymxk max mhzt itkmr
11	ewwl ew sxlwj lzw lgys hsjlq
12	dvvk dv rwkvi kyv kfxr grikp
13	cuuj cu qvjuh jxu jewq fqhjo
14	btti bt puitg iwt idvp epgin
15	assh as othsf hvs hcuo dofhm
16	zrrg zr nsgre gur gbtn cnegl
17	yqqf yq mrfqd ftq fasm bmdfk
18	xppe xp lqepc esp ezrl alcej
19	wood wo kpdob dro dyqk zkbdi
20	vnnc vn jocna cqn cxpj yjach
21	ummb um inbmz bpm bwoi xizbg
22	tlla tl hmaly aol avnh whyaf
23	skkz sk glzkx znk zumg vgxze
24	rjjy rj fkyjw ymj ytlf ufwyd
25	qiix qi ejxiv xli xske tevxc

Reciprocal Cipher

- applies the same transformation to decrypt a message as the one used to encrypt it.
- If x→y then y→x.
- E.g.,
  - XoR
  - Rot 13



ROT13 replaces each letter by its partner 13 characters further along the alphabet. For example, HELLO becomes URYYB (or, conversely, URYYB becomes HELLO again).

Monoalphabetic Cipher

- rather than just shifting the alphabet
- could shuffle (jumble) the letters arbitrarily
- each plaintext letter maps to a different random ciphertext letter
- hence key is 26 letters long

```
Plain:  abcdefghijklmnopqrstuvwxyz
Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN
Plaintext:  ifwewishtoreplaceletters
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
```

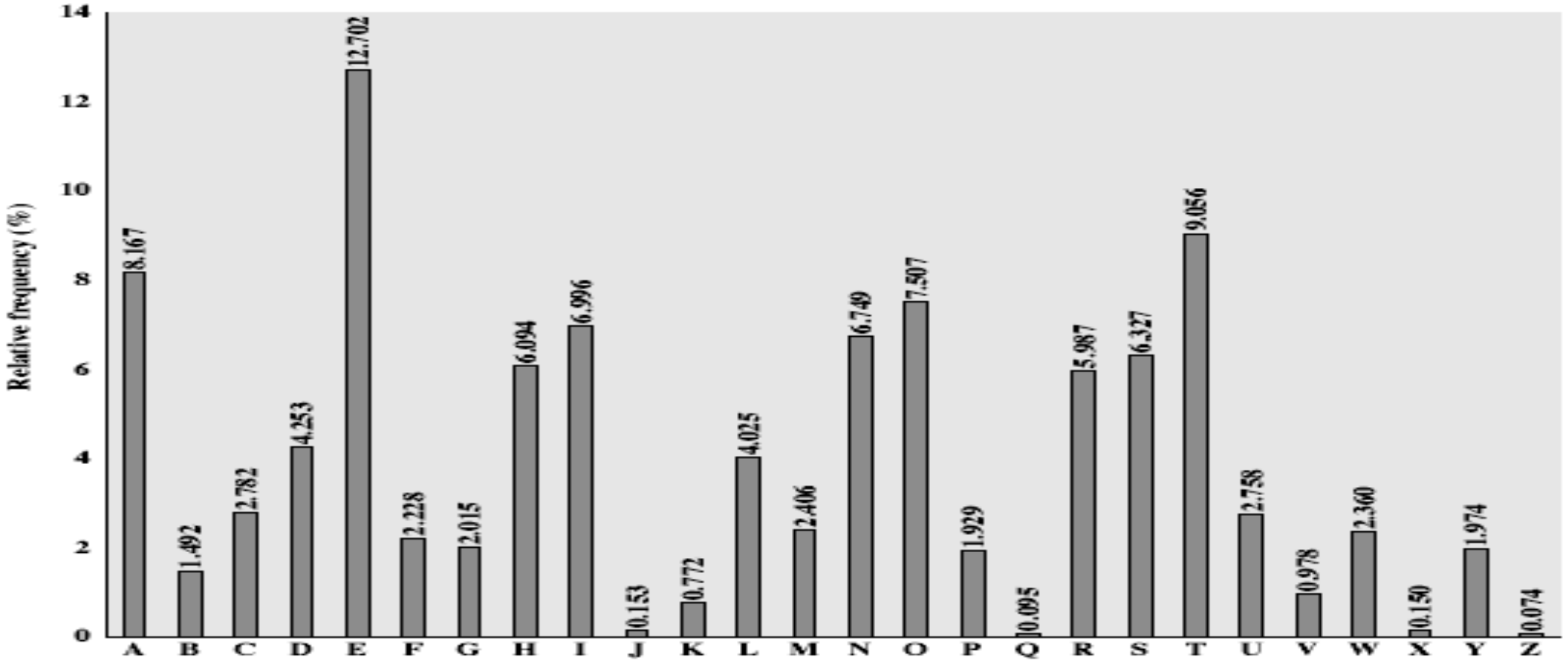
Monoalphabetic Cipher Security

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

Language Redundancy and Cryptanalysis

- human languages are **redundant**
- eg "th lrd s m shphrd shll nt wnt"
- letters are not equally commonly used
- in English **e** is by far the most common letter
- then T,R,N,I,O,A,S
- other letters are fairly rare
- cf. Z,J,K,Q,X
- have tables of single, double & triple letter frequencies

English Letter Frequencies



Character Frequency in English language

Letter	Frequency	Letter	Frequency
E	12.702	m	2.406
t	9.056	w	2.360
a	8.167	f	2.228
o	7.507	g	2.015
i	6.966	y	1.974
n	6.749	p	1.929
s	6.327	b	1.492
h	6.094	v	0.978
r	5.987	k	0.772
d	4.253	j	0.153
l	4.025	x	0.150
c	2.782	q	0.095
u	2.758	z	0.074

Characters in English Language

Char	Frequency	Most Common Bigram (in order)	Most Common Trigram (in order)
e	0.12702	**th**	the
t	0.09056	he	and
a	0.08167	in	tha
o	0.07507	an	ing
i	0.06966	nt	hat
n	0.06749	re	ion
s	0.06327	er	tio
h	0.06094	an	for
r	0.05987	ti	nde
d	0.04253	es	has
l	0.04025	on	nce
c	0.02782	at	edt
u	0.02758	is	tis
m	0.02406	nd	oft
w	0.02360	or	sth
f	0.02228	ar	men
g	0.02015	al	
y	0.01974	te	
p	0.01929	co	
b	0.01492	de	
v	0.00978	to	
k	0.00772	ra	
j	0.00153	et	
x	0.00150	ed	
q	0.00095	it	
z	0.00074	sa	

Use in Cryptanalysis

- key concept - monoalphabetic substitution ciphers do not change relative letter frequencies
- discovered by Arabian scientists in 9th century
- calculate letter frequencies for ciphertext
- compare counts/plots against known values
- if Caesar cipher look for common peaks/troughs
  - peaks at: A-E-I triple, NO pair, RST triple
  - Lowest at: JK, X-Z
- for monoalphabetic must identify each letter
  - tables of common double/triple letters help

Example Cryptanalysis

- given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- count relative letter frequencies (see text)

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				

Example Cryptanalysis

- given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- Plaintext..

guess P & Z are e and t

  - UtQSOVUOHXMOeVGeOteEVSGtWStOeFeESXUDBMETSXAItVUEeHtHMDtSHtOWSFe
    AeeDTSVeQUtWYMXUtUHSXEeYEeOeDtStUFeOMBtWeFUetHMDJUDTMOHMQ

Example Cryptanalysis

- given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

ZW is th
Hence ZWP is the

- Plaintext..
  - UtQSOVUOHXMOeVGeOteEVSGtWStOeFeESXUDBMETSXAItVUEeHtHMDtSHtOWSFe AeeDTSVeQUtWYMXUtUHSXEeYEeOeDtStUFeOMBtWeFUetHMDJUDTMOHMQ

Example Cryptanalysis

- given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

ZW is th
Hence ZWP is the

- Plaintext..
  - UtQSOVUOHXMOeVGeOteEVSGtWStOeFeESXUDBMETSXAItVUEeHtHMDtSHtOWSFe AeeDTSVeQUtWYMXUtUHSXEeYEeOeDtStUFeOMBtWeFUetHMDJUDTMOHMQ
  - UtQSOVUOHXMOeVGeOteEVSGthStOeFeESXUDBMETSXAItVUEeHtHMDtSHtOhSFeA eeDTSVeQUthYMXUtUHSXEeYEeOeDtStUFeOMBtheFUetHMDJUDTMOHMQ

Example Cryptanalysis

- given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

ZW is th
Hence ZWP is the

- Plaintext..
  - UtQSOVUOHXMOeVGeOteEVSGtWStOeFeESXUDBMETSXAItVUEeHtHMDtSHtOWSFe AeeDTSVeQUtWYMXUtUHSXEeYEeOeDtStUFeOMBtWeFUetHMDJUDTMOHMQ
  - UtQaOVUOHXMOeVGeOteEVaGthatOeFeEaXUDBMETaXAItVUEeHtHMDtaHtOhSFeA eeDTSVeQUthYMXUtUHaXEeYEeOeDtatUFeOMBtheFUetHMDJUDTMOHMQ

Example Cryptanalysis

- given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

ZW is th
Hence ZWP is the

- Plaintext..
  - UtQSOVUOHXMOeVGeOteEVSG thSt
    OeFeESXUDBMETSXAItVUEeHtHMDtSHtOhSFeAeeDTSVeQUthYMX
    UtUHSXEeYEeOeDtStUFeOMB the FUetHMDJUDTMOHMQ

Example Cryptanalysis

- 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
```

Polygram substitution

- Using two letters at a time
- If arbitrary chosen: (26^2)!=676!
  - Aa ab ac ad ae af ag ah ai
  - RL TC YB FR UU SN JA IL AP
- The key needs to be formulated via some means

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 Key Matrix

- 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

Encrypting and Decrypting

- plaintext encrypted two letters at a time:
  1. if a pair is a repeated letter, insert a 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 (again wrapping to top from bottom), eg. “mu" encrypts to "CM"
  4. otherwise each letter is replaced by the one in its row in the column of the other letter of the pair, eg. “hs" encrypts to "BP", and “ea" to "IM" or "JM" (as desired)

Security of the 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. US & British military in WW1)
- it **can** be broken, given a few hundred letters
- since still has much of plaintext structure

Polyalphabetic Ciphers

- another approach to improving security is to use multiple cipher alphabets
- called **polyalphabetic substitution ciphers**
- makes 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

Vigenère Cipher

- simplest polyalphabetic substitution cipher is the **Vigenère Cipher**
- effectively multiple caesar ciphers
- key is multiple letters long K = k1 k2 ... kd
- i^{th} letter specifies i^{th} alphabet to use
- use each alphabet in turn
- repeat from start after d letters in message
- decryption simply works in reverse

Example

- 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
```

Plain text

Key

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

Security of Vigenère Ciphers

- have multiple ciphertext letters for each plaintext letter
- hence letter frequencies are obscured
- but not totally lost
- start with letter frequencies
  - see if look monoalphabetic or not
- if not, then need to determine number of alphabets, since then can attach each

Kasiski Method

- method developed by Babbage / Kasiski
- repetitions in ciphertext give clues to period
- so find same plaintext an exact period apart
- which results in the same ciphertext
- eg repeated “VTW” in previous example
- suggests size of 3 or 9
- then attack each monoalphabetic cipher individually using same techniques as before

Autokey Cipher

- ideally want a key as long as the message
- Vigenère proposed the **autokey** cipher
- with keyword is prefixed to message as key
- knowing keyword can recover the first few letters
- use these in turn on the rest of the message
- but still have frequency characteristics to attack I.e. the key and message share the same frequency distribution.
- eg. given key *deceptive*

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

One-Time Pad

- if a truly random key as long as the message is used, the cipher will be secure
- called a One-Time pad
- is unbreakable since ciphertext bears no statistical relationship to the plaintext
- since for **any plaintext** & **any ciphertext** there exists a key mapping one to other
- can only use the key **once** though
- have problem of safe distribution of key

Transposition Ciphers

- now consider classical **transposition** or **permutation** ciphers
- these 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
- 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

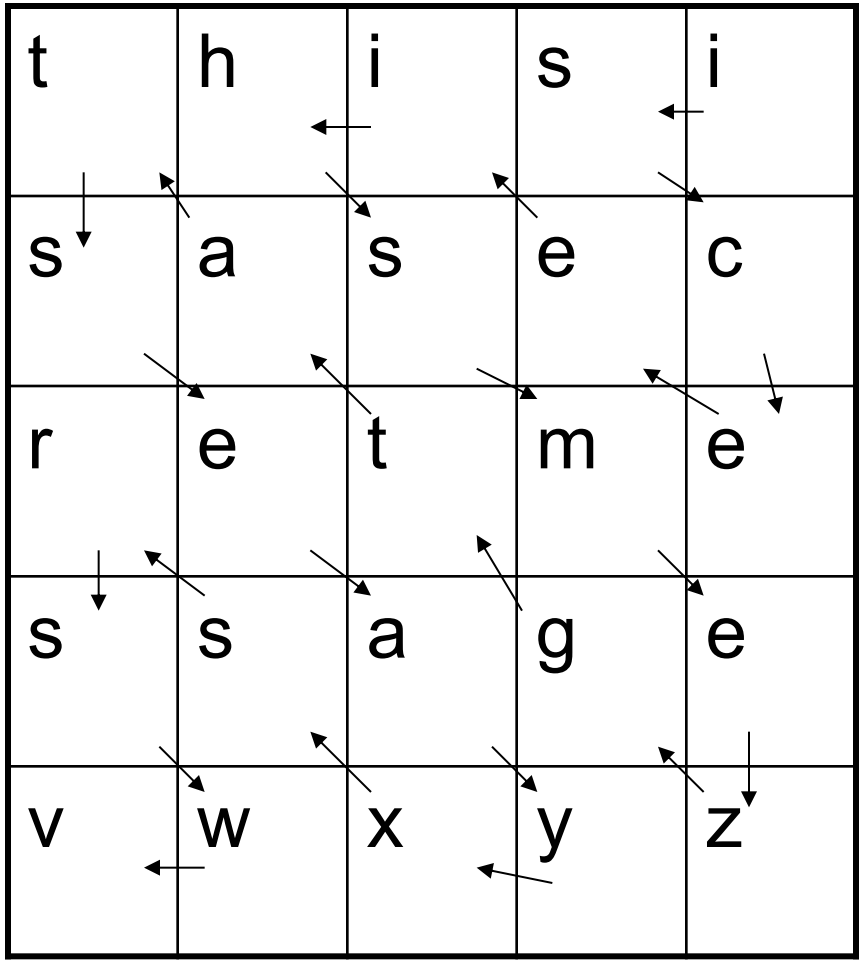
```
MEMATRHTGPRYETEFETEOAAT
```

Row Transposition Ciphers

- a more complex scheme
- 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
```

Route transposition



→ISCEE IHSME

ZGTAT SEAYX

SRSWV

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

Steganography

- an alternative to encryption
- hides existence of message
  - using only a subset of letters/words in a longer message marked in some way
  - using invisible ink
- has drawbacks
  - high overhead to hide relatively few info bits

Summary

- have considered:
  - classical cipher techniques and terminology
  - monoalphabetic substitution ciphers
  - cryptanalysis using letter frequencies
  - Playfair ciphers
  - polyalphabetic ciphers
  - transposition ciphers
  - product ciphers and rotor machines
  - stenography

END