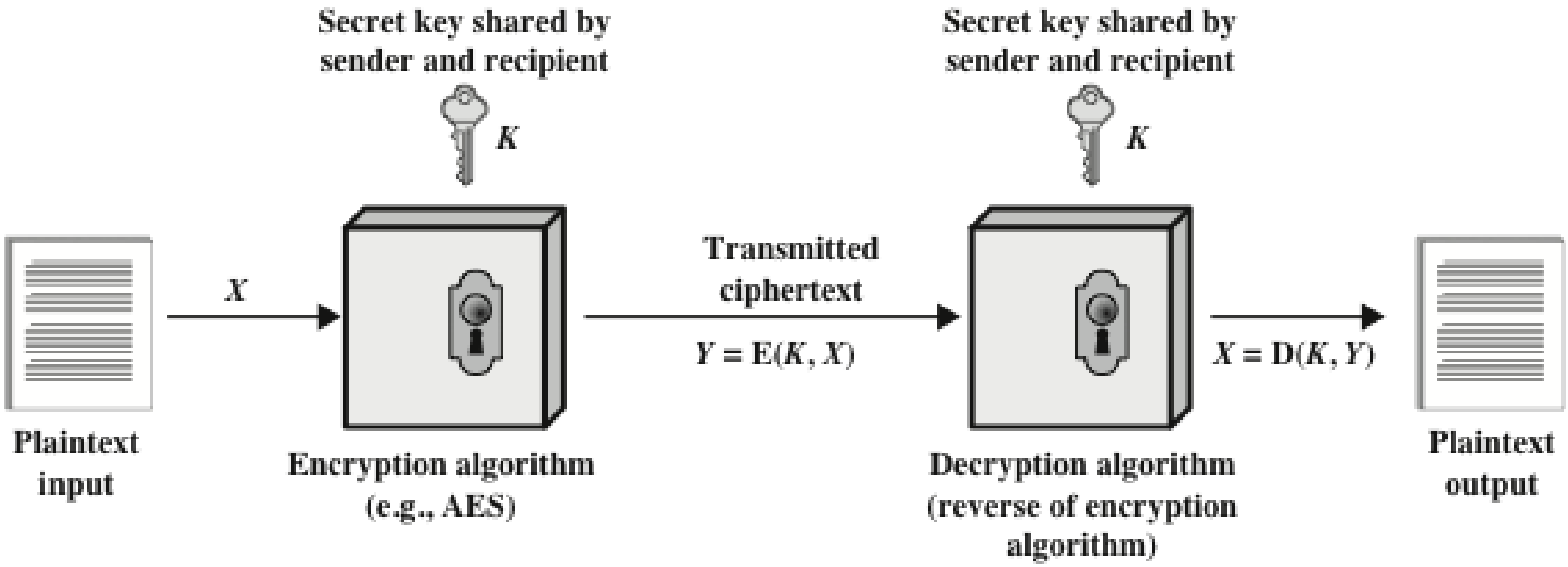
Definitions (1 of 2)

- **Plaintext**
  - An original message
- **Ciphertext**
  - The coded message
- **Enciphering/encryption**
  - The process of converting from plaintext to ciphertext
- **Deciphering/decryption**
  - Restoring the plaintext from the ciphertext

Definitions (2 of 2)

- **Cryptography**
  - The area of study of the many schemes used for encryption
- **Cryptographic system/cipher**
  - A scheme
- **Cryptanalysis**
  - Techniques used for deciphering a message without any knowledge of the enciphering details
- **Cryptology**
  - The areas of cryptography and cryptanalysis

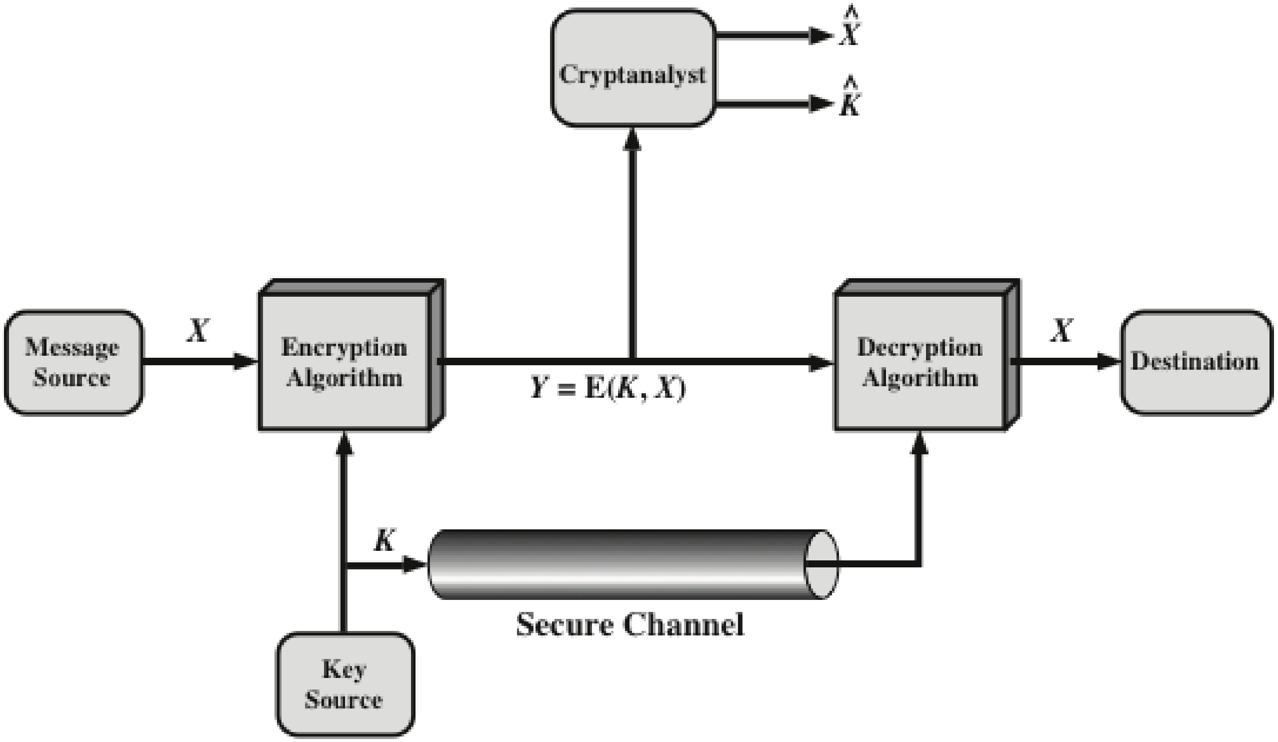
Figure 3-1 Simplified Model of Symmetric Encryption


Secret key shared by sender and recipient
K
Secret key shared by sender and recipient
K
X
Transmitted ciphertext
Y = E(K, X)
X = D(K, Y)
Plaintext input
Encryption algorithm (e.g., AES)
Decryption algorithm (reverse of encryption algorithm)
Plaintext output

Symmetric Cipher Model

- There are two requirements for secure use of conventional encryption:
  - A strong encryption algorithm
  - Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the key secure

Figure 3-2 Model of Symmetric Cryptosystem



Symmetric Key Cryptography

- **Basic Encryption Techniques:**
  - Substitution: Letters of plaintext are replaced by other letters or numbers or symbols
  - Transposition: Rearranging the letter order in the plaintext without altering the actual letters
  - Product: A substitution followed by a transposition
- **Basic Types:**
  - Stream Cipher: Plaintext digits are combined with a pseudorandom cipher digit stream (keystream) one bit at a time. E.g. RC4, A5/1, A5/2
  - Block Cipher: A symmetric key cipher operating on fixed-length groups of bits E.g. AES, DES
- **Possible Attacks:**
  - Cryptanalytic attack:
    - Carefully analyzing the nature of encryption algorithm and some characteristics of plaintext and/or ciphertext
  - Brute-force attack:
    - Trying every possible key until an intelligible translation of plaintext is obtained. On average, 50% of total keys must be tried to achieve success.

Substitution Technique

- Is one in which the letters of plaintext are replaced by other letters or by numbers or symbols
- If the plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns

Popular Substitution Ciphers

- Simplest Substitution
  - Caesar Cipher
- Monoalphabetic Substitution (A single cipher alphabet substitution; 1→1 mapping)
  - Playfair Cipher
  - Hill Cipher
- Polyalphabetic Substitution (uses different monoalphabetic substitution as one proceeds through the plaintext message)
  - Vigenère Cipher
  - Autokey Cipher
  - One-Time Pad

Caesar Cipher

- Simplest and earliest known use of a substitution cipher
- Used by Julius Caesar
- Involves replacing each letter of the alphabet with the letter standing three places further down the alphabet
- Alphabet is wrapped around so that the letter following Z is A

plain: meet me after the toga party

cipher: PHHW PH DIWHU WKH WRJD SDUWB

Pearson

Caesar Cipher Algorithm (1 of 2)

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

$$c = E(3, p) = (p + 3)\mathrm{mod}(26)$$

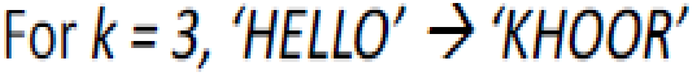
Caesar Cipher Algorithm (2 of 2)

- A shift may be of any amount, so that the general Caesar algorithm is:

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

- Where k takes on a value in the range 1 to 25; the decryption algorithm is simply:

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

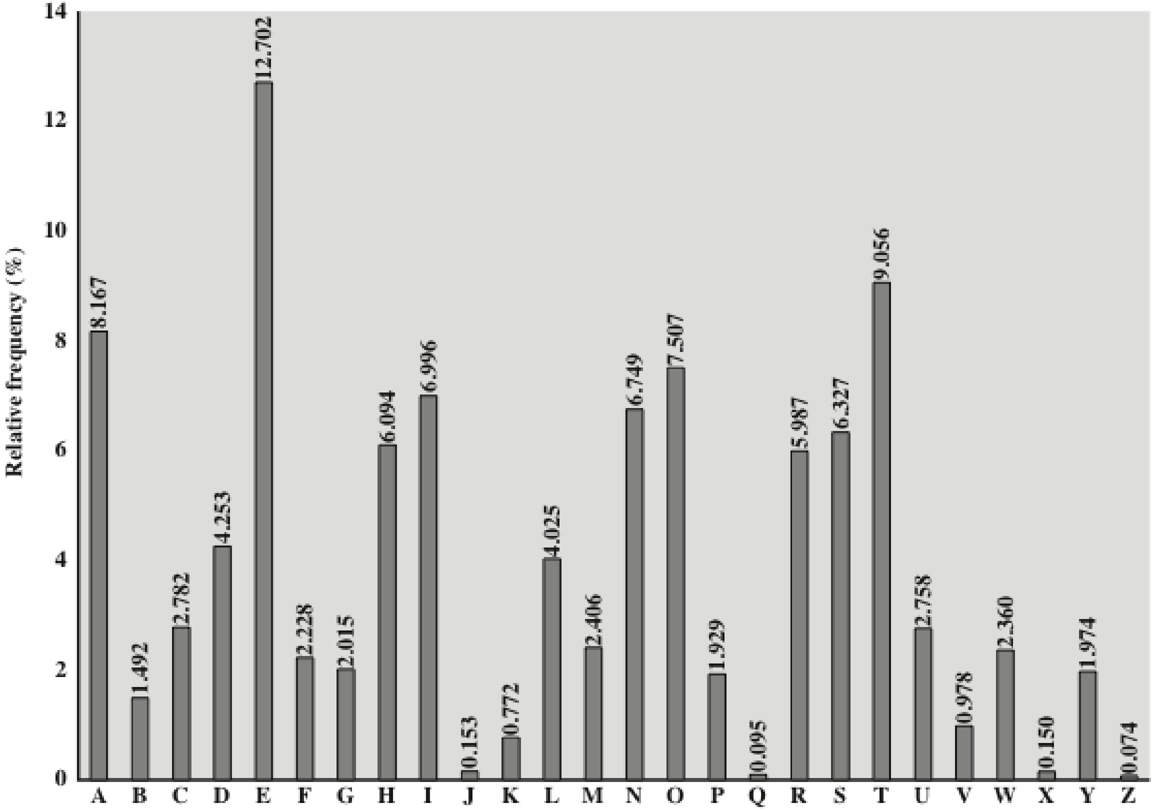


For $k = 3$, 'HELLO' → 'KHOOR'

Monoalphabetic Cipher

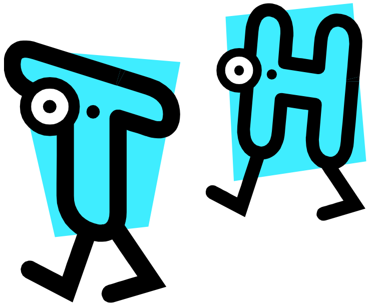
- Permutation
  - Of a finite set of elements *S* is an ordered sequence of all the elements of *S* , with each element appearing exactly once
- If the "cipher" line can be any permutation of the 26 alphabetic characters, then there are 26! or greater than 4×10^{26} possible keys
  - This is 10 orders of magnitude greater than the key space for DES
  - Approach is referred to as a **monoalphabetic substitution cipher** because a single cipher alphabet is used per message

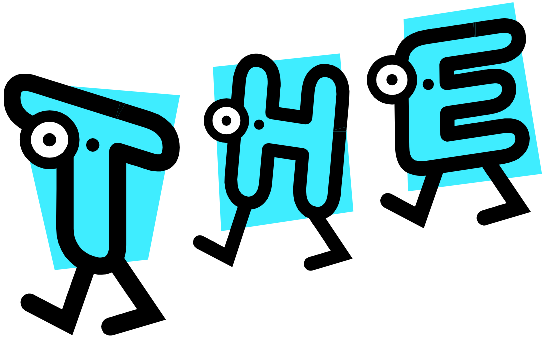
Figure 3-5 Relative Frequency of Letters in English Text



Monoalphabetic Ciphers

- Easy to break because they reflect the frequency data of the original alphabet
- Countermeasure is to provide multiple substitutes (homophones) for a single letter
- Digram
  - Two-letter combination
  - Most common is **th**
- Trigram
  - Three-letter combination
  - Most frequent is **the**





Playfair Cipher

- Invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair
- A 5X5 matrix of letters based on a keyword
- Fill in letters of keyword
- Fill rest of matrix with other letters
- E.g. 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

Choice of either I or J depends upon the wish of Encipherer OR one may omit a less frequent letter like Q to complete 5x5 matrix

Plaintext is encrypted two letters at a time

1. if a pair is a repeated letter, insert filler like 'X'
   - E.g. FLOOR → FL OX OR
2. if both letters fall in the same row, replace each with letter to right (wrapping back to start from end) E.g., “ar" encrypts as "RM"
3. if both letters fall in the same column, replace each with the letter below it (wrapping to top from bottom) E.g., “mu" encrypts to "CM"
4. otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair. E.g., “HS" encrypts to ”BP"

Pearson

Security of Playfair Cipher

- Better security than Monoalphabetic Cipher
- 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
- Widely used for many years
  - E.g. by US & British military in WORLD WAR 1
- It **can** be broken, given a few hundred letters
- since still has much of plaintext structure

Hill Cipher (by Lester S. Hill, 1929)

- A substitution cipher based on *Linear Algebra*
- Each letter is identified with a number

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

- Encryption
  - A block of **n** letters from plaintext is considered as a vector of **n** dimensions, and multiplied by an *invertible* **n×n** key matrix, modulo 26

$$\begin{pmatrix}6&24&1\\13&16&10\\20&17&15\end{pmatrix}\begin{pmatrix}0\\2\\19\end{pmatrix}\begin{array}{c}A\\C\\T\end{array}=\begin{pmatrix}67\\222\\319\end{pmatrix}\equiv\begin{pmatrix}15\\14\\7\end{pmatrix}\begin{array}{c}P\\O\\H\end{array}\pmod{26}$$

- Decryption
  - A block of **n** letters from ciphertext is considered as a vector of **n** dimensions, and multiplied by the inverse of key matrix, modulo 26

$$\begin{pmatrix}6&24&1\\13&16&10\\20&17&15\end{pmatrix}^{-1}\equiv\begin{pmatrix}8&5&10\\21&8&21\\21&12&8\end{pmatrix}\pmod{26}$$

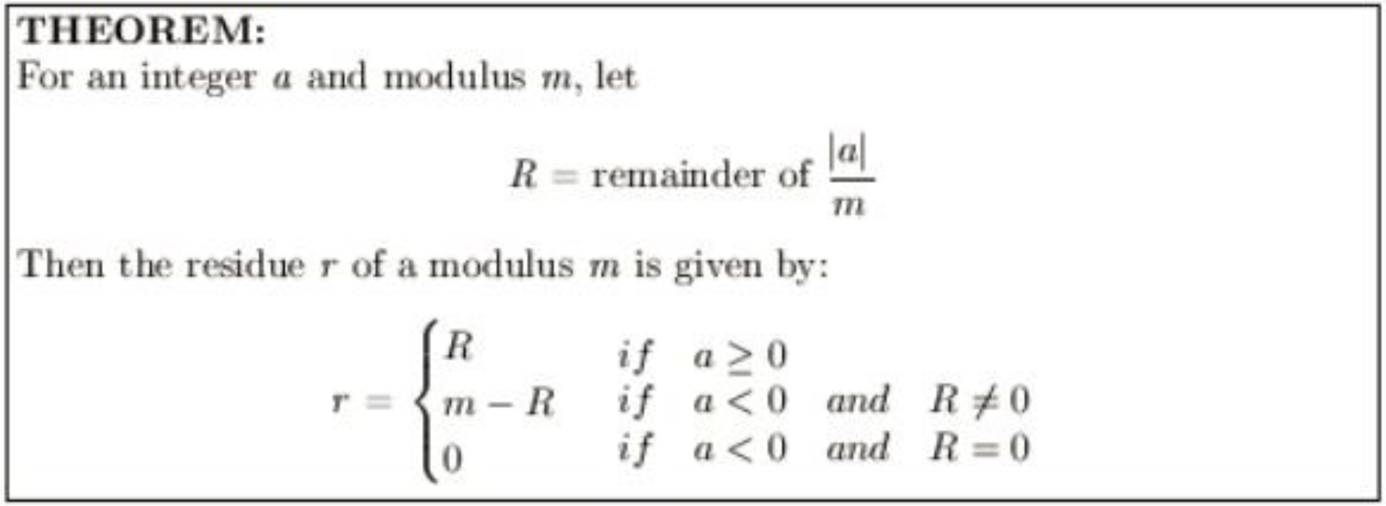
$$\begin{pmatrix}8&5&10\\21&8&21\\21&12&8\end{pmatrix}\begin{pmatrix}15\\14\\7\end{pmatrix}=\begin{pmatrix}260\\574\\539\end{pmatrix}\equiv\begin{pmatrix}0\\2\\19\end{pmatrix}\begin{array}{c}A\\C\\T\end{array}\pmod{26}$$

Be careful !!!
Not all the matrices are invertible.

Pearson

Finding the Modular Inverse Matrix

- Modulus Theorem



THEOREM:
For an integer a and modulus m, let

$$R = \text{remainder of } \frac{|a|}{m}$$

Then the residue r of a modulus m is given by:

$$r = \begin{cases} R & if \quad a \geq 0 \\ m - R & if \quad a < 0 \quad and \quad R \neq 0 \\ 0 & if \quad a < 0 \quad and \quad R = 0 \end{cases}$$

- Example

1. 80 Mod 26

$$\frac{|80|}{26} = 3 \text{ R } 2 \Rightarrow 80 \geq 0 \Rightarrow 80 \text{ Mod } 26 = 2$$

2. -80 Mod 26

$$\frac{|-80|}{26} = 3 \text{ R } 2 \Rightarrow -80 < 0 \text{ and } 2 \neq 0 \Rightarrow -80 \text{ Mod } 26 = 26 - 2 = 24$$

3. -52 Mod 26

$$\frac{|-52|}{26} = 2 \text{ R } 0 \Rightarrow -52 < 0 \text{ and } 0 = 0 \Rightarrow -52 \text{ Mod } 26 = 0$$

Finding the Modular Inverse Matrix

- Modular Inverse
  - Modular Inverse for **mod** ***m*:**

$$(a \cdot a^{-1}) \bmod m = 1$$

  - For Modular Inverses, *a* and *m* must NOT have any prime factors in common (Must be coprime, i.e. GCD=1)
- Example
  - Modular inverses of 4, 6, 8, 10, 12 … modulo 26 are not possible
  - Modular inverses of 2, 3, 5, 7, 9, modulo 26 are 9, 9, 21, 15, 3 respectively

Pearson

Finding the Modular Inverse Matrix

- Calculate determinant of the matrix
- Make sure that determinant has a modular inverse for mod 26
- Calculate the adjugate of A, adj A

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \operatorname{adj}(\mathbf{A}) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

- Multiply adj A by modular inverse of det A
- Calculate Mod 26 of the result
- Use the resulting matrix to decrypt

Finding the Modular Inverse Matrix

Let $A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$

det A = (2 x 4) - (1 x 3) = 8 - 3 = 5

modular inverse of 5 for Mod 26 = 21

$B = 21\begin{bmatrix} 4 & -1 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 84 & -21 \\ -63 & 42 \end{bmatrix}$

$B = \begin{bmatrix} 84 & -21 \\ -63 & 42 \end{bmatrix}$ Mod 26 $= \begin{bmatrix} 6 & 5 \\ 15 & 16 \end{bmatrix}$

Therefore $\begin{bmatrix} 6 & 5 \\ 15 & 16 \end{bmatrix}$ is the modular inverse of $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$ for Mod 26. $\begin{bmatrix} 6 & 5 \\ 15 & 16 \end{bmatrix}$

Pearson

Security of Hill Cipher

- Unfortunately, the basic Hill cipher is vulnerable to a known-plaintext attack because it is completely linear.
- An opponent who intercepts n2 plaintext/ ciphertext character pairs can set up a linear system which can (usually) be easily solved;
- if it happens that this system is indeterminate, it is only necessary to add a few more plaintext/ ciphertext pairs.
- Calculating this solution by standard linear algebra algorithms then takes very little time.

Polyalphabetic Ciphers

- Polyalphabetic substitution cipher
  - Improves on the simple monoalphabetic technique by using different monoalphabetic substitutions as one proceeds through the plaintext message

All these techniques have the following features in common:

- A set of related monoalphabetic substitution rules is used
- A key determines which particular rule is chosen for a given transformation

Vigenère Cipher

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

- The letters are identified with the numbers 0, 1, …, 25
- The secret key is a short sequence of letters (e.g. any English word, or any sequence of letters).
- If the key is 'SESAME', encryption operates as follows:

T	H	I	S	I	S	A	T	E	S	T	M	E	S	S	A	G	E	–	plaintext
S	E	S	A	M	E	S	E	S	A	M	E	S	E	S	A	M	E	–	keystream
L	L	A	S	U	W	S	X	W	S	F	Q	W	W	K	A	S	I	–	ciphertext

- Same letter/number mapping as for the Caesar cipher:

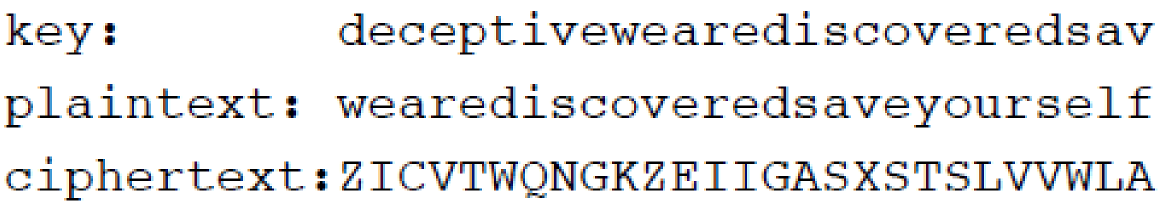
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Encryption involves adding ***k mod* 26** to the numerical equivalent of each letter.
- The Keystream is simply a keyword repeated as necessary
- The Vigenère & related polyalphabetic ciphers still do not completely obscure the underlying language characteristics.
- **Security:** Identifying the number of translation alphabets, and then attack each separately, can break the Vigenère Cipher

Pearson

Vigenère Autokey System

- Key is as long as the message
- Proposed by Vigenère
- Keyword is prefixed to message as key
- Knowing keyword can recover the first few letters which becomes the basis to guess rest of the message
- Still have frequency characteristics to attack
- E.g. given key ***deceptive***



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

One-Time Pad

- Improvement to Vernam cipher proposed by an Army Signal Corp officer, Joseph Mauborgne
- Use a random key that is as long as the message so that the key need not be repeated
- Key is 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
- Scheme is unbreakable
  - Produces random output that bears no statistical relationship to the plaintext
  - Because the ciphertext contains no information whatsoever about the plaintext, there is simply no way to break the code

Difficulties (1 of 2)

- The one-time pad offers complete security but, in practice, has two fundamental difficulties:
  - There is the practical problem of making large quantities of random keys
    - Any heavily used system might require millions of random characters on a regular basis
  - Mammoth key distribution problem
    - For every message to be sent, a key of equal length is needed by both sender and receiver

Difficulties (2 of 2)

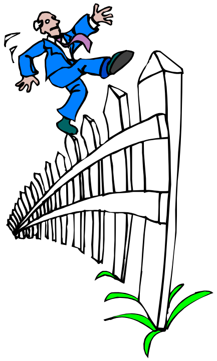
- Because of these difficulties, the one-time pad is of limited utility
  - Useful primarily for low-bandwidth channels requiring very high security
- The one-time pad is the only cryptosystem that exhibits **perfect secrecy** (see Appendix F)

Transposition Ciphers

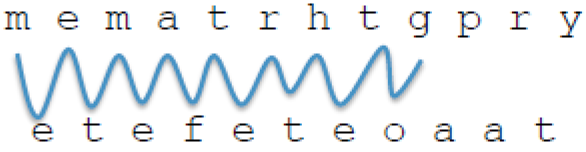
- Rail Fence cipher
- Row Transposition ciphers

Pearson

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:



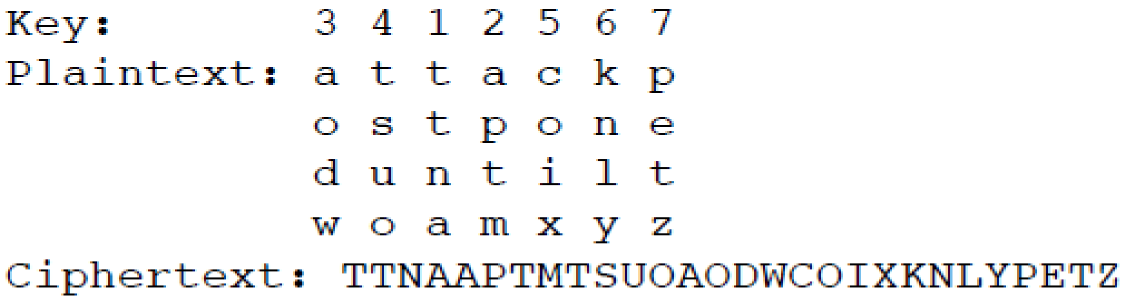
```
m e m a t r h t g p r y
 e t e f e t e o a a t
```

- giving ciphertext

MEMATRHTGPRYETEFETEOAAT

Row Transposition Cipher

- A more complex transposition
- Steps:
  1. Write letters of message out in rows over a specified number of columns
  2. then reorder the columns according to some key
  3. Write the columns (top to bottom) from start



```
Key:        3 4 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
```

Types of encryption security

- **Computational Security**
  - The cost of breaking the cipher exceeds the value of the encrypted information
  - Time required to break the cipher exceeds the useful lifetime of the information
- **Unconditional/Perfect/Information Theoretic Security**
  - Secure, even when the adversary has unlimited computing power.
  - There is not enough information to break the security.
  - Encryption scheme does not depend for its effectiveness on unproven assumptions about computational hardness
  - An algorithm which is not vulnerable to future developments in Computing, e.g. quantum computing.

Cryptanalysis and Brute-Force Attack

- **Cryptanalysis**
  - Attack relies on the nature of the algorithm plus some knowledge of the general characteristics of the plaintext
  - Attack exploits the characteristics of the algorithm to attempt to deduce a specific plaintext or to deduce the key being used
- **Brute-force attack**
  - Attacker tries every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained
  - On average, half of all possible keys must be tried to achieve success

Shannon’s Theorem

- Shannon’s Theorem is the most important theorem in the information theoretic study of cryptography.
- Suppose (P, C, K, E_k (·), D_k (·)) is a cryptosystem with #P = #C = #K.
- This cryptosystem provides **perfect secrecy** if and only if
  - every key is used with equal probability 1/#K and,
  - for each m ∈ P and c ∈ C, there is a unique key k such that $E_k(m)=c$.

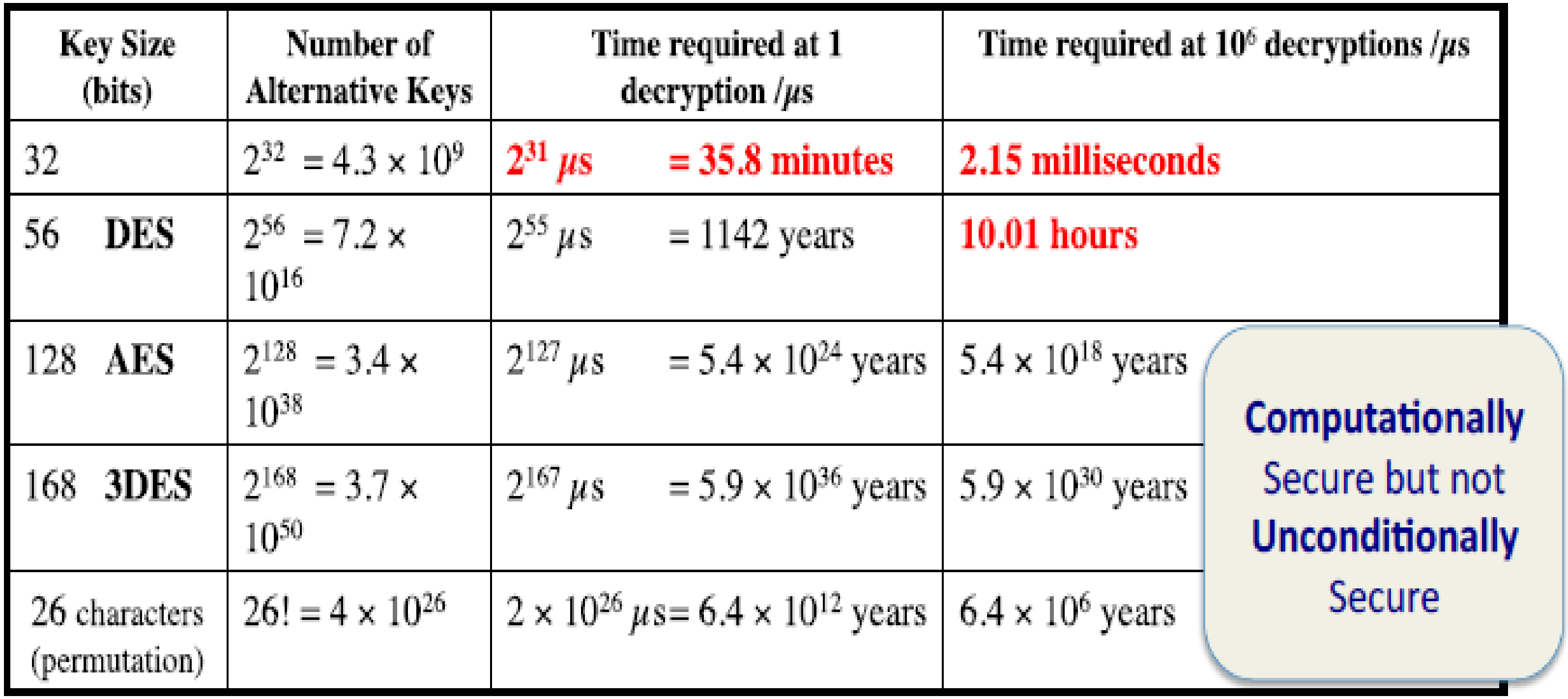
Unconditionally Security of Shift Cipher

- Consider the **Shift Cipher** Scheme
  - **P=K=C=Z_{26}**
  - $E_k(m) = m + k \bmod 26$
- According to Shannon's Theorem the Scheme will provide **perfect security**
  - if encryption is applied separately and independently on each letter (or one letter at a time)
  - The probability of using each key letter would be p(K=k)=1/$(26)^n$
- OTP is a special case of Shift Cipher with modulo 2 instead of modulo 26

Pearson

Brute-force attack vs Key Length

Also called exhaustive key search



Key Size (bits)	Number of Alternative Keys	Time required at 1 decryption $/\mu s$	Time required at 10^6 decryptions $/\mu s$
32	$2^{32} = 4.3 \times 10^9$	**2^{31} μs = 35.8 minutes**	**2.15 milliseconds**
56 **DES**	$2^{56} = 7.2 \times 10^{16}$	2^{55} μs = 1142 years	**10.01 hours**
128 **AES**	$2^{128} = 3.4 \times 10^{38}$	2^{127} μs = 5.4×10^{24} years	5.4×10^{18} years
168 **3DES**	$2^{168} = 3.7 \times 10^{50}$	2^{167} μs = 5.9×10^{36} years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	2×10^{26} μs= 6.4×10^{12} years	6.4×10^{6} years

Computationally Secure but not **Unconditionally** Secure

Table 3-1 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

Assignment / Homework

- Write programs to Encrypt/Decrypt using one of the Substitution Ciphers?
- Write programs to Encrypt/Decrypt using one of the Transformation Ciphers?
  - See course homepage for details!