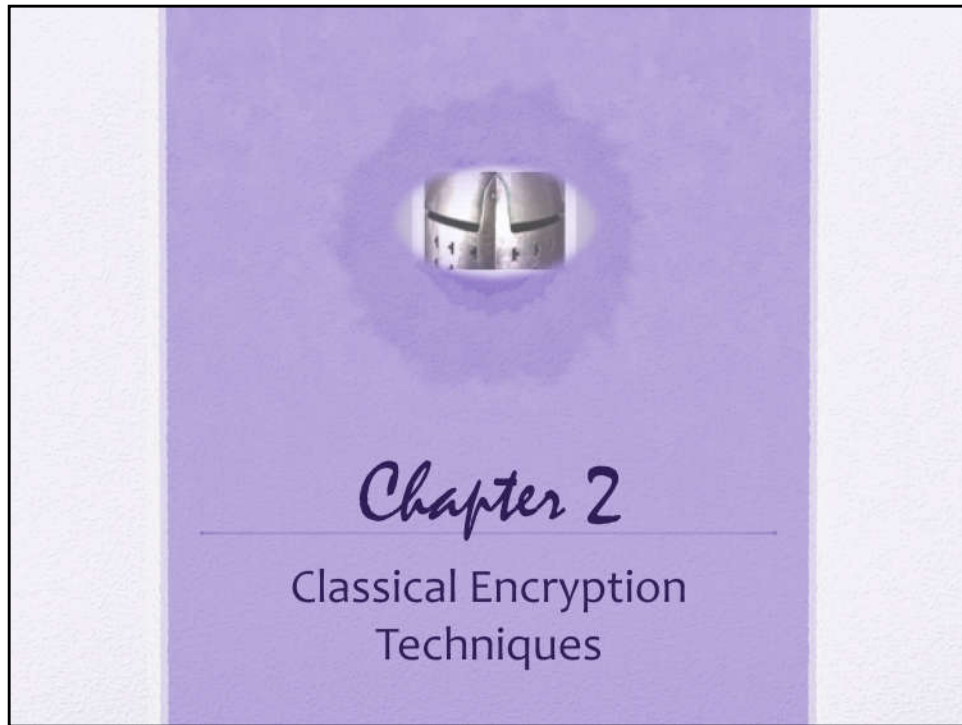


Lecture slides prepared for “Cryptography and Network Security”, 6/e, by William Stallings, Chapter 2 – “Classical Encryption Techniques”.



Classical Encryption Techniques

*"I am fairly familiar with all the forms of secret writings, and am myself the author of a trifling monograph upon the subject, in which I analyze one hundred and sixty separate ciphers," said Holmes.*

**—The Adventure of the Dancing Men,  
Sir Arthur Conan Doyle**

Opening quote.

# Symmetric Encryption

- Also referred to as conventional encryption or single-key encryption
- Was the only type of encryption in use prior to the development of public-key encryption in the 1970s
- Remains by far the most widely used of the two types of encryption



Symmetric encryption, also referred to as conventional encryption or single-key encryption,

was the only type of encryption in use prior to the development of public key

encryption in the 1970s. It remains by far the most widely used of the two types

of encryption. Part One examines a number of symmetric ciphers. In this chapter, we

begin with a look at a general model for the symmetric encryption process; this will

enable us to understand the context within which the algorithms are used.

Next, we

examine a variety of algorithms in use before the computer era. Finally, we look briefly

at a different approach known as steganography. Chapters 3 and 5 introduce the two

most widely used symmetric cipher: DES and AES.

# Basic Terminology

- Plaintext
  - The original message
- Ciphertext
  - The coded message
- Enciphering or encryption
  - Process of converting from plaintext to ciphertext
- Deciphering or decryption
  - Restoring the plaintext from the ciphertext
- Cryptography
  - Study of encryption
- Cryptographic system or cipher
  - Schemes used for encryption
- Cryptanalysis
  - Techniques used for deciphering a message without any knowledge of the enciphering details
- Cryptology
  - Areas of cryptography and cryptanalysis together

Before beginning, we define some terms. An original message is known as the

plaintext, while the coded message is called the ciphertext. The process of converting

from plaintext to ciphertext is known as enciphering or encryption; restoring the

plaintext from the ciphertext is deciphering or decryption. The many schemes used

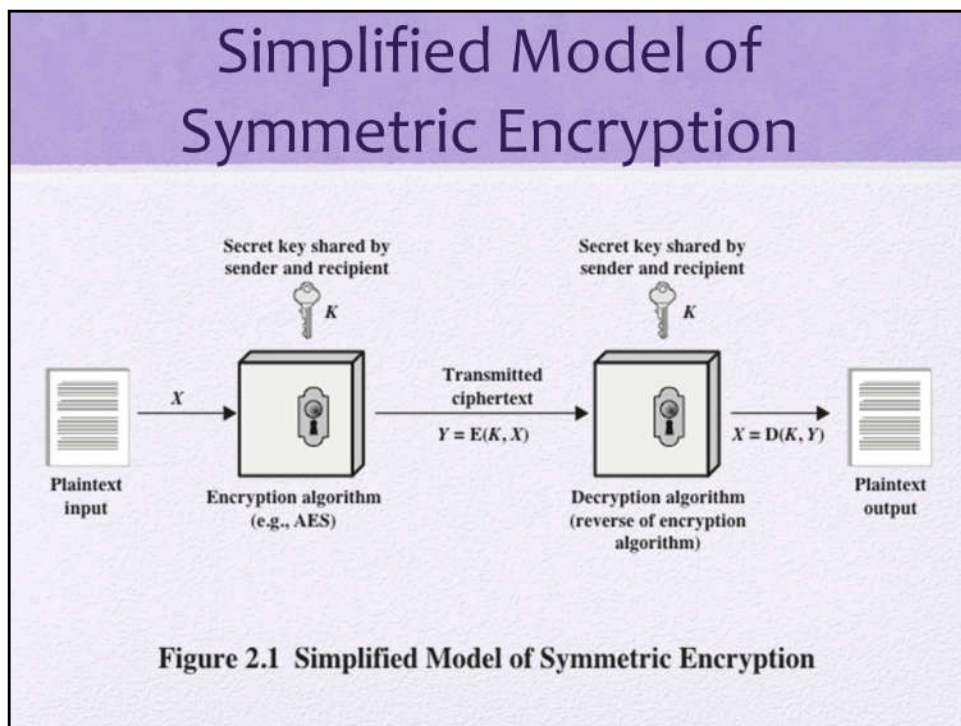
for encryption constitute the area of study known as cryptography. Such a scheme

is known as a cryptographic system or a cipher. Techniques used for deciphering

a message without any knowledge of the enciphering details fall into the area of

cryptanalysis. Cryptanalysis is what the layperson calls “breaking the code.” The areas

of cryptography and cryptanalysis together are called cryptology.



A symmetric encryption scheme has five ingredients (Figure 2.1):

- **Plaintext:** This is the original intelligible message or data that is fed into the algorithm as input.
- **Encryption algorithm:** The encryption algorithm performs various substitutions and transformations on the plaintext.
- **Secret key:** The secret key is also input to the encryption algorithm. The key is a value independent of the plaintext and of the algorithm. The algorithm will produce a different output depending on the specific key being used at the time. The exact substitutions and transformations performed by the algorithm depend on the key.
- **Ciphertext:** This is the scrambled message produced as output. It depends on the plaintext and the secret key. For a given message, two different keys will produce two different ciphertexts. The ciphertext is an apparently random stream of data and, as it stands, is unintelligible.



# Model of Symmetric Cryptosystem

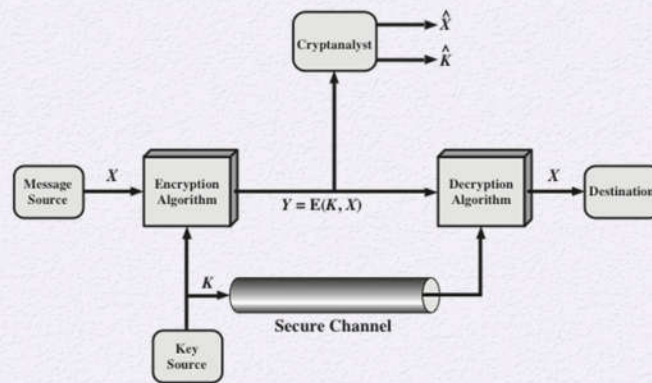
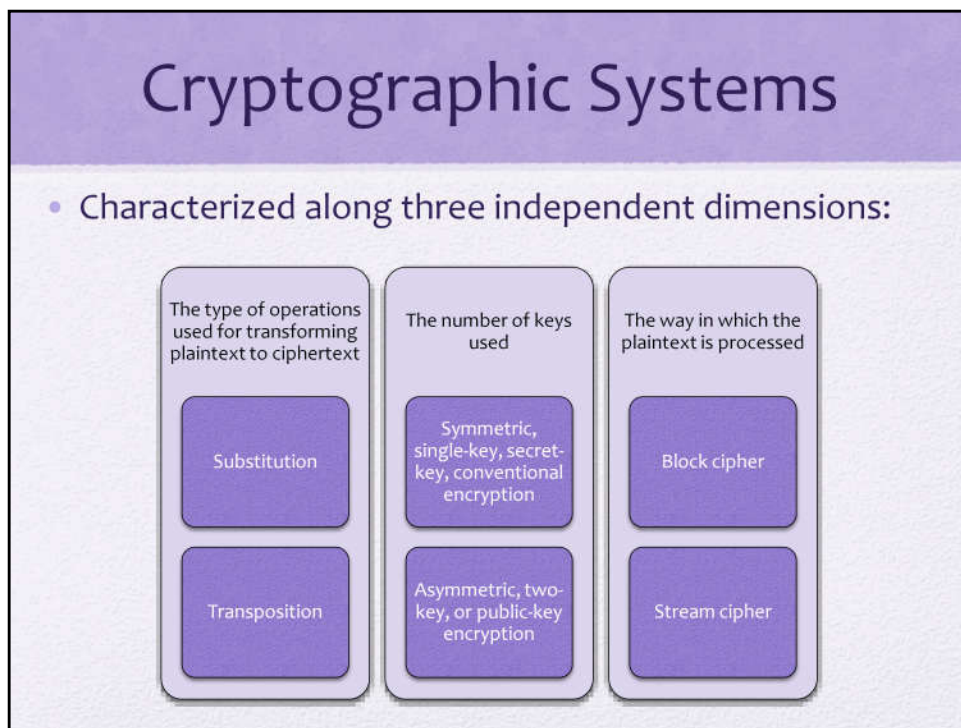


Figure 2.2 Model of Symmetric Cryptosystem

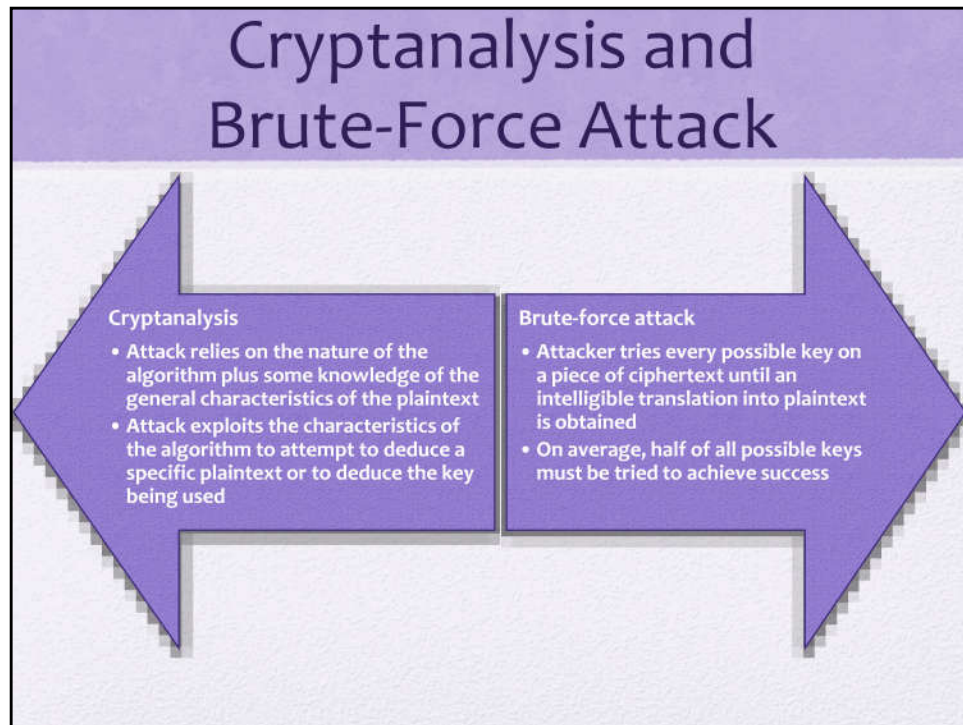
Let us take a closer look at the essential elements of a symmetric encryption scheme, using Figure 2.2.



Cryptographic systems are characterized along three independent dimensions:

1. The type of operations used for transforming plaintext to ciphertext. All encryption algorithms are based on two general principles: 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. The fundamental requirement is that no information be lost (i.e., that all operations are reversible). Most systems, referred to as product systems, involve multiple stages of substitutions and transpositions.
2. The number of keys used. If both sender and receiver use the same key, the system is referred to as symmetric, single-key, secret-key, or conventional encryption. If the sender and receiver use different keys, the system is referred to as asymmetric, two-key, or public-key encryption.
3. The way in which the plaintext is processed. 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.





Typically, the objective of attacking an encryption system is to recover the key in

use rather than simply to recover the plaintext of a single ciphertext. There are two

general approaches to attacking a conventional encryption scheme:

- **Cryptanalysis:** Cryptanalytic attacks rely on the nature of the algorithm plus perhaps some knowledge of the general characteristics of the plaintext or even some sample plaintext–ciphertext pairs. This type of attack exploits the characteristics of the algorithm to attempt to deduce a specific plaintext or to deduce the key being used.

- **Brute-force attack:** The 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.

If either type of attack succeeds in deducing the key, the effect is catastrophic: All future and past messages encrypted with that key are compromised.

Type of Attack	Known to Cryptanalyst
Ciphertext Only	<ul style="list-style-type: none"> <li>• Encryption algorithm</li> <li>• Ciphertext</li> </ul>
Known Plaintext	<ul style="list-style-type: none"> <li>• Encryption algorithm</li> <li>• Ciphertext</li> <li>• One or more plaintext-ciphertext pairs formed with the secret key</li> </ul>
Chosen Plaintext	<ul style="list-style-type: none"> <li>• Encryption algorithm</li> <li>• Ciphertext</li> <li>• Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key</li> </ul>
Chosen Ciphertext	<ul style="list-style-type: none"> <li>• Encryption algorithm</li> <li>• Ciphertext</li> <li>• Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key</li> </ul>
Chosen Text	<ul style="list-style-type: none"> <li>• Encryption algorithm</li> <li>• Ciphertext</li> <li>• Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key</li> <li>• Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key</li> </ul>

**Table 2.1**  
Types of  
Attacks  
on  
Encrypted  
Messages

Table 2.1 summarizes the various types of cryptanalytic attacks based on the amount of information known to the cryptanalyst. The most difficult problem is presented when all that is available is the ciphertext only . In some cases, not even

the encryption algorithm is known, but in general, we can assume that the opponent

does know the algorithm used for encryption. One possible attack under these circumstances is the brute-force approach of trying all possible keys. If the key space

is very large, this becomes impractical. Thus, the opponent must rely on an analysis

of the ciphertext itself, generally applying various statistical tests to it. To use this

approach, the opponent must have some general idea of the type of plaintext that

is concealed, such as English or French text, an EXE file, a Java source listing, an

accounting file, and so on.

The ciphertext-only attack is the easiest to defend against because the opponent has the least amount of information to work with. In many cases, however,

# Encryption Scheme Security

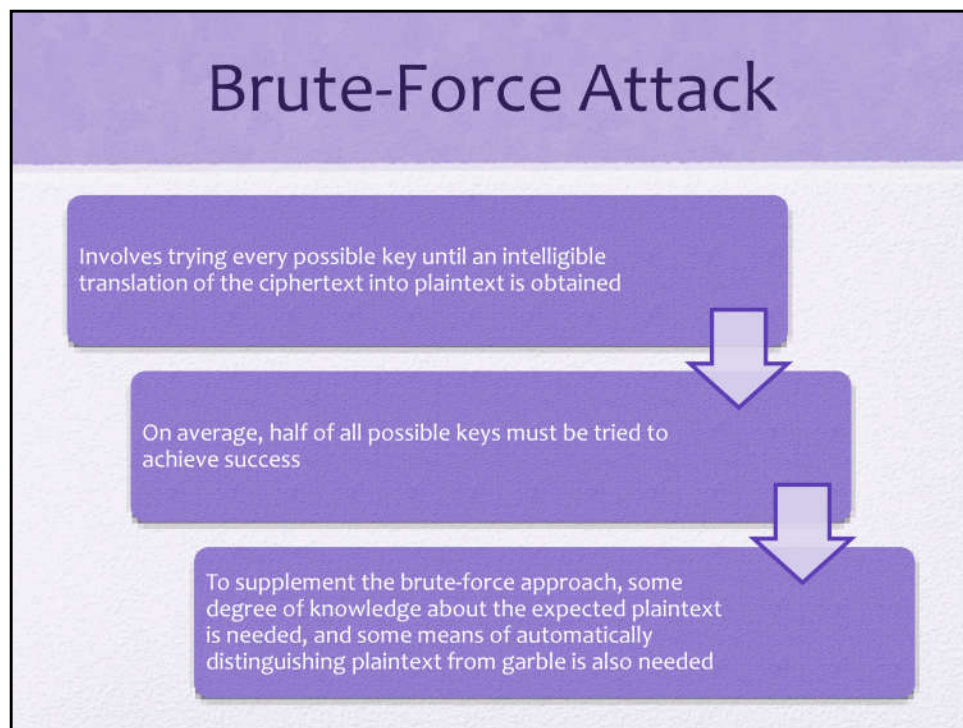
- Unconditionally secure
  - No matter how much time an opponent has, it is impossible for him or her to decrypt the ciphertext simply because the required information is not there
- Computationally secure
  - The cost of breaking the cipher exceeds the value of the encrypted information
  - The time required to break the cipher exceeds the useful lifetime of the information



Two more definitions are worthy of note. An encryption scheme is unconditionally secure if the ciphertext generated by the scheme does not contain enough information to determine uniquely the corresponding plaintext, no matter how much ciphertext is available. That is, no matter how much time an opponent has, it is impossible for him or her to decrypt the ciphertext simply because the required information is not there. With the exception of a scheme known as the one-time pad (described later in this chapter), there is no encryption algorithm that is unconditionally secure. Therefore, all that the users of an encryption algorithm can strive for is an algorithm that meets one or both of the following criteria:

- The cost of breaking the cipher exceeds the value of the encrypted information.
- The time required to break the cipher exceeds the useful lifetime of the information.

An encryption scheme is said to be computationally secure if either of the



A brute-force attack involves trying every possible key until an intelligible translation of the ciphertext into plaintext is obtained. On average, half of all possible keys must be tried to achieve success. That is, if there are  $X$  different keys, on average an attacker would discover the actual key after  $X/2$  tries. It is important to note that there is more to a brute-force attack than simply running through all possible keys. Unless known plaintext is provided, the analyst must be able to recognize plaintext as plaintext. If the message is just plain text in English, then the result pops out easily, although the task of recognizing English would have to be automated. If the text message has been compressed before encryption, then recognition is more difficult. And if the message is some more general type of data, such as a numerical file, and this has been compressed, the problem becomes even more difficult to automate. Thus, to supplement the brute-force approach, some degree of knowledge about the expected plaintext is needed, and some means of automatically

# 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



In this section and the next, we examine a sampling of what might be called classical encryption techniques. A study of these techniques enables us to illustrate the basic approaches to symmetric encryption used today and the types of cryptanalytic attacks that must be anticipated.

The two basic building blocks of all encryption techniques are substitution and transposition. We examine these in the next two sections. Finally, we discuss a system that combines both substitution and transposition.

A 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.





# 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

The earliest known, and the simplest, use of a substitution cipher was by Julius

Caesar. The Caesar cipher involves replacing each letter of the alphabet with the

letter standing three places further down the alphabet.

Note that the alphabet is wrapped around, so that the letter following Z is A.



# Caesar Cipher Algorithm

- Can define transformation as:

```
abcdefghijklmnopqrstuvwxyz  
DEFGHIJKLMNOPQRSTUVWXYZABC
```

- Mathematically give each letter a number

```
abcdefghijklmnopqrstuvwxyz  
012345678910111213141516171819202122232425
```

- Algorithm can be expressed as:

$$c = E(3, p) = (p + 3) \bmod 26$$

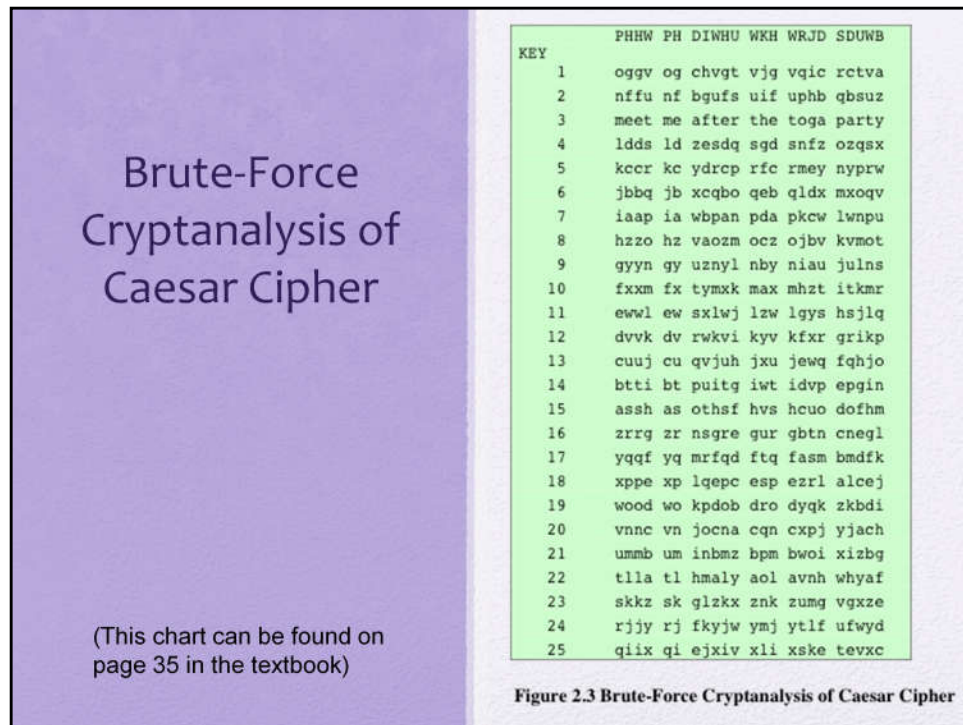
- A shift may be of any amount, so that the general Caesar algorithm is:

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

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

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

Then the algorithm can be expressed as follows. For each plaintext letter  $p$ , substitute the ciphertext letter  $C$



If it is known that a given ciphertext is a Caesar cipher, then a brute-force cryptanalysis is easily performed: simply try all the 25 possible keys. Figure 2.3

shows the results of applying this strategy to the example ciphertext. In this case, the plaintext leaps out as occupying the third line.

Three important characteristics of this problem enabled us to use a brute-force cryptanalysis:

1. The encryption and decryption algorithms are known.
2. There are only 25 keys to try.
3. The language of the plaintext is known and easily recognizable.

## Sample of Compressed Text

```

~+Wu*- Ω-0)S4(==t. e-Ωêràü.-í ô~z-
û#20#Åæð æ=q7,Ωn-@3N0Û æz'Y-f=í[±0_ èΩ,<NO~t«~xã ÅæFè03Å
x)05k=Å
-yí "ΔÉ] .# J/'iTê&1 'c<uΩ-
AD(G WAC~y_10ÅW P0i<f0+c|,n;~i^uñπ~="L~90gñ0~&0S ~S 005":
"0!SGqèvo" ú\,S>h<-*6ø+8x'"|ñ0#="myk~zñP<,f1 Åj Å0ç"zù-
Ω~0~60y{8_Ω8ó .1 π+Å1'ú02ç8y'O-
2Åñ8i /0^~[[K~*P0π,úé^'3Σ~0~0Zi~Y~?ΩmY> Ω+e0/'<Kfç*+~*S0~
B ZøK~Q8y0f,!0ñIzsS/)»8Q ú

```

Figure 2.4 Sample of Compressed Text

In most networking situations, we can assume that the algorithms are known. What generally makes brute-force cryptanalysis impractical is the use of an algorithm

that employs a large number of keys. For example, the triple DES algorithm, examined in Chapter 6, makes use of a 168-bit key, giving a key space of  $2^{168}$  or

greater than  $3.7 \times 10^{50}$  possible keys.

The third characteristic is also significant. If the language of the plaintext is unknown, then plaintext output may not be recognizable. Furthermore, the input may be abbreviated or compressed in some fashion, again making recognition

difficult. For example, Figure 2.4 shows a portion of a text file compressed using an algorithm called ZIP. If this file is then encrypted with a simple substitution

cipher (expanded to include more than just 26 alphabetic characters), then the plaintext may not be recognized when it is uncovered in the brute-force cryptanalysis.

# 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  $4 \times 10^{26}$  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

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. Before proceeding,

we define the term 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

In general, there are  $n!$  permutations of a set of  $n$  elements, because the first element can be chosen in one of  $n$  ways, the second in  $n - 1$  ways, the third in  $n - 2$  ways, and so on.

If, instead, the “cipher” line can be any permutation of the 26 alphabetic characters, then there are  $26!$  or greater than  $4 \times 10^{26}$  possible keys. This is 10 orders of magnitude greater than the key space for DES and would seem to eliminate brute force

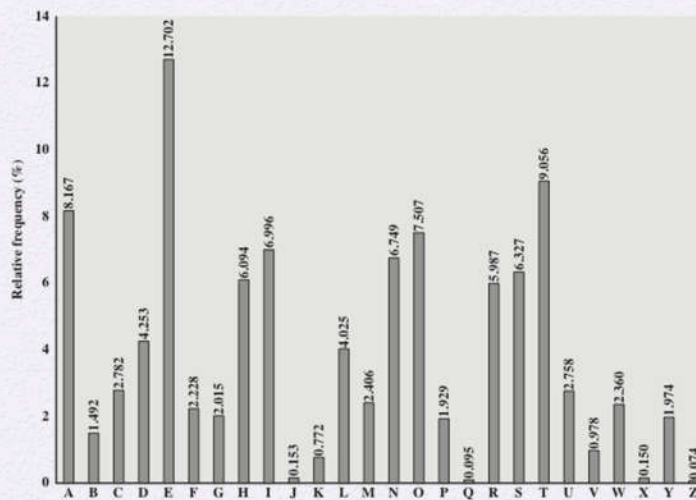


Figure 2.5 Relative Frequency of Letters in English Text

There is, however, another line of attack. If the cryptanalyst knows the nature of the plaintext (e.g., noncompressed English text), then the analyst can exploit the

regularities of the language. To see how such a cryptanalysis might proceed, we give

a partial example here that is adapted from one in [SINK09]. The ciphertext to be

solved is

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

VUEPHZHMZSHZOWSFPAPPDTSVPQUZWMXUZHUSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

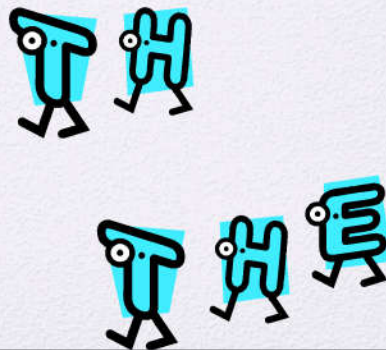
As a first step, the relative frequency of the letters can be determined and compared to a standard frequency distribution for English, such as is shown in Figure 2.5 (based on [LEWA00]). If the message were long enough, this technique

alone might be sufficient, but because this is a relatively short message, we cannot



# 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*



A powerful tool is to look at the frequency of two-letter combinations, known as digrams. A table similar to Figure 2.5 could be drawn up showing the relative frequency

of digrams. The most common such digram is *th*. In our ciphertext, the most common digram is *ZW*, which appears three times. So we make the correspondence

of *Z* with *t* and *W* with *h*. Then, by our earlier hypothesis, we can equate *P* with *e*.

Now notice that the sequence *ZWP* appears in the ciphertext, and we can translate

that sequence as “*the*.” This is the most frequent trigram (three-letter combination)

in English, which seems to indicate that we are on the right track.

Next, notice the sequence *ZWSZ* in the first line. We do not know that these four letters form a complete word, but if they do, it is of the form *th\_t*. If so, *S* equates with *a*.

So far, then, we have

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ



# Playfair Cipher

- Best-known multiple-letter encryption cipher
- Treats digrams in the plaintext as single units and translates these units into ciphertext digrams
- Based on the use of a 5 x 5 matrix of letters constructed using a keyword
- Invented by British scientist Sir Charles Wheatstone in 1854
- Used as the standard field system by the British Army in World War I and the U.S. Army and other Allied forces during World War II

The best-known multiple-letter encryption cipher is the Playfair, which treats digrams in the plaintext as single units and translates these units into ciphertext

Digrams.

The Playfair algorithm is based on the use of a 5 \* 5 matrix of letters constructed using a keyword.

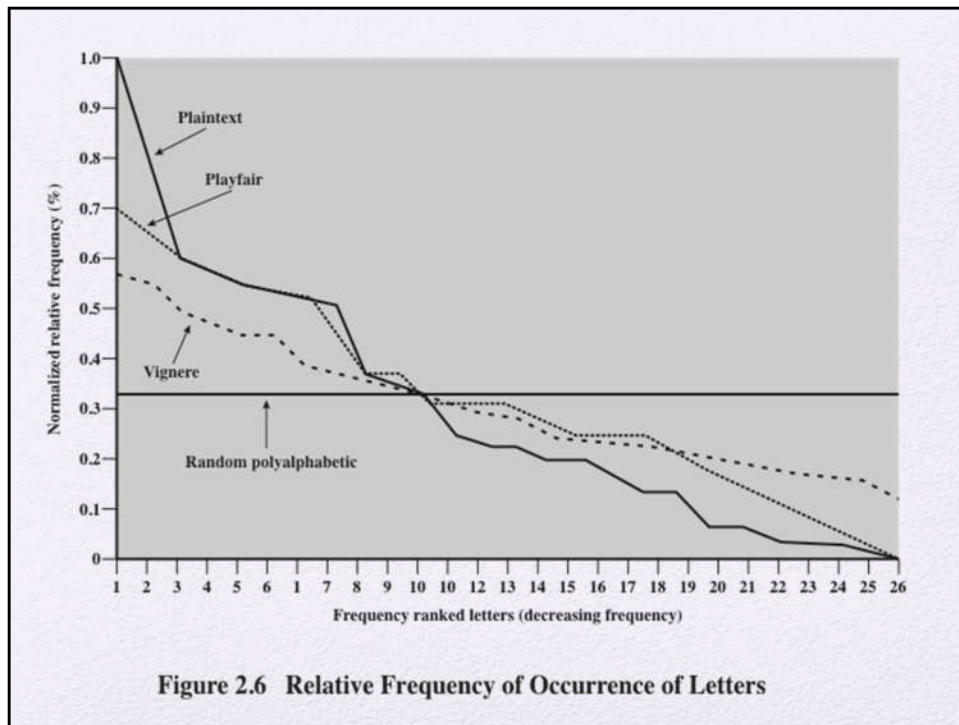
## Playfair Key Matrix

- Fill in letters of keyword (minus duplicates) from left to right and from top to bottom, then fill in the remainder of the matrix with the remaining letters in alphabetic order
- 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

In this case, the keyword is monarchy . The matrix is constructed by filling in the letters of the keyword (minus duplicates) from left to right and from top to bottom, and then filling in the remainder of the matrix with the remaining letters in alphabetic order. The letters I and J count as one letter. Plaintext is encrypted two letters at a time, according to the following rules:

1. Repeating plaintext letters that are in the same pair are separated with a filler letter, such as x, so that balloon would be treated as ba lx lo on.
2. Two plaintext letters that fall in the same row of the matrix are each replaced by the letter to the right, with the first element of the row circularly following the last. For example, ar is encrypted as RM.
3. Two plaintext letters that fall in the same column are each replaced by the letter beneath, with the top element of the column circularly following the last. For example, mu is encrypted as CM.



Despite this level of confidence in its security, the Playfair cipher is relatively easy to break, because it still leaves much of the structure of the plaintext language intact. A few hundred letters of ciphertext are generally sufficient.

One way of revealing the effectiveness of the Playfair and other ciphers is shown in Figure 2.6. The line labeled plaintext plots a typical frequency distribution of the 26 alphabetic characters (no distinction between upper and lower case) in ordinary text. This is also the frequency distribution of any monoalphabetic substitution cipher, because the frequency values for individual letters are the same, just with different letters substituted for the original letters. The plot is developed in the following way: 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. Using the results of Figure 2.5, we see that e is the most frequently used letter. As a result, e has a relative frequency of 1, t of  $9.056/12.702 = 0.72$ , and so on. The points on the horizontal axis correspond to the letters in order of decreasing frequency.

Figure 2.6 also shows the frequency distribution that results when the text is encrypted using the Playfair cipher. To normalize the plot, the number of occurrences of each letter in the ciphertext was again divided by the number of occurrences of e in the plaintext. The resulting plot therefore shows the extent to which the frequency distribution of letters, which makes it trivial to solve substitution ciphers, is masked by encryption. 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. As the figure shows, the Playfair cipher has a flatter distribution than does plaintext, but nevertheless, it reveals plenty of structure for a cryptanalyst to work with. The plot also shows the Vigenère cipher, discussed subsequently. The Hill and Vigenère curves on the plot are based on results reported in [SIMM93].

# Hill Cipher

- Developed by the mathematician Lester Hill in 1929
- Strength is that it completely hides single-letter frequencies
  - The use of a larger matrix hides more frequency information
  - A 3 x 3 Hill cipher hides not only single-letter but also two-letter frequency information
- Strong against a ciphertext-only attack but easily broken with a known plaintext attack

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

Before describing the Hill cipher, let us briefly review some terminology from linear algebra. In this discussion, we are concerned with matrix arithmetic modulo 26. For the reader who needs a refresher on matrix multiplication and inversion, see Appendix E.

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  $\mathbf{I}$  is the identity matrix.  $\mathbf{I}$  is a square matrix that is all zeros except for ones along the main diagonal from upper left to lower right. The inverse of a matrix does not always exist, but when it does, it satisfies the preceding equation.

To explain how the inverse of a matrix is computed, we begin with the concept of determinant. For any square matrix ( $m \times m$ ), the determinant equals the sum of all the products that can be formed by taking exactly one element from each

# 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

Another way to improve on the simple monoalphabetic technique is to use different monoalphabetic substitutions as one proceeds through the plaintext message. The general name for this approach is polyalphabetic substitution cipher . All these techniques have the following features in common:

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



# Vigenère Cipher

- Best known and one of the simplest polyalphabetic substitution ciphers
- In this scheme the set of related monoalphabetic substitution rules consists of the 26 Caesar ciphers with shifts of 0 through 25
- Each cipher is denoted by a key letter which is the ciphertext letter that substitutes for the plaintext letter a

The best known, and one of the simplest, polyalphabetic ciphers is the Vigenère cipher. In this scheme, the set of related monoalphabetic substitution rules consists of the 26 Caesar ciphers with shifts of 0 through 25. Each cipher is denoted by a key letter, which is the ciphertext letter that substitutes for the plaintext letter a. Thus, a Caesar cipher with a shift of 3 is denoted by the key value 3.



## Example of Vigenère Cipher

- To encrypt a message, a key is needed that is as long as the message
- Usually, the key is a repeating keyword
- For example, if the keyword is *deceptive*, the message “we are discovered save yourself” is encrypted as:

key:       deceptivedeceptivedeceptive  
plaintext: wearediscoveredsaveyourself  
ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

To encrypt a message, a key is needed that is as long as the message. Usually,

the key is a repeating keyword. For example, if the keyword is *deceptive*, the message “we are discovered save yourself” is encrypted as

key: deceptivedeceptivedeceptive

plaintext: wearediscoveredsaveyourself

ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

The strength of this cipher is that there are multiple ciphertext letters for each plaintext letter, one for each unique letter of the keyword. Thus, the letter frequency information is obscured. However, not all knowledge of the plaintext structure is lost. For example, Figure 2.6 shows the frequency distribution for a Vigenère cipher with a keyword of length 9. An improvement is achieved over the

Playfair cipher, but considerable frequency information remains.

# Vigenère Autokey System

- A keyword is concatenated with the plaintext itself to provide a running key
- Example:  
key:           deceptivewearediscoveredsav  
plaintext:   wearediscoveredsaveyourself  
ciphertext:  ZICVTWQNGKZEIIGASXSTSLVWLA
- 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

The periodic nature of the keyword can be eliminated by using a nonrepeating keyword that is as long as the message itself. 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:  ZICVTWQNGKZEIIGASXSTSLVWLA

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. For example, *e* enciphered by *e* , by Figure 2.5, can be expected to

occur with a frequency of  $(0.127)^2 = 0.016$ , whereas *t* enciphered by *t* would occur

only about half as often. These regularities can be exploited to achieve successful

cryptanalysis.

# Vernam Cipher

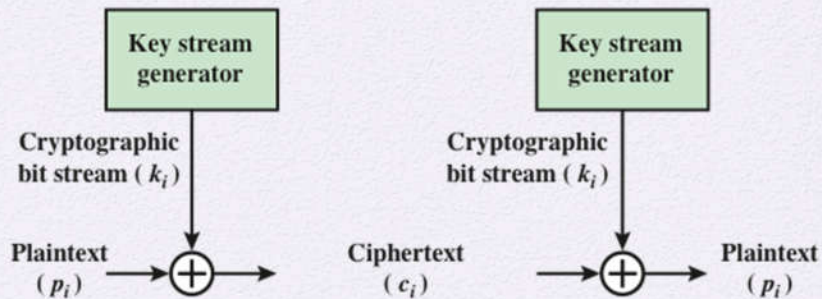


Figure 2.7 Vernam Cipher

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. Such

a system was introduced by an AT&T engineer named Gilbert Vernam in 1918. His system works on binary data (bits) rather than letters.

The essence of this technique is the means of construction of the key. 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.

Although

such a scheme, with a long key, presents formidable cryptanalytic difficulties, it can be broken with sufficient ciphertext, the use of known or probable plaintext sequences, or both.

# 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



An Army Signal Corp officer, Joseph Mauborgne, proposed an improvement to the

Vernam cipher that yields the ultimate in security. 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. It 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.

In fact, given any plaintext of equal length to the ciphertext, there is a key that produces that plaintext. Therefore, if you did an exhaustive search of all possible

keys, you would end up with many legible plaintexts, with no way of knowing



# Difficulties

- 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
- 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)

In theory, we need look no further for a cipher. The one-time pad offers complete

security but, in practice, has two fundamental difficulties:

1. 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. Supplying truly random characters in this volume is a significant task.
2. Even more daunting is the problem of key distribution and protection. For every message to be sent, a key of equal length is needed by both sender and receiver. Thus, a mammoth key distribution problem exists.

Because of these difficulties, the one-time pad is of limited utility and is useful primarily for low-bandwidth channels requiring very high security.

The one-time pad is the only cryptosystem that exhibits what is referred to as perfect secrecy . This concept is explored in Appendix F.

# Rail Fence Cipher

- Simplest transposition cipher
- Plaintext is written down as a sequence of diagonals and then read off as a sequence of rows
- To encipher the message “meet me after the toga party” with a rail fence of depth 2, we would write:

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

Encrypted message is:

MEMATRHTGPRYETEFETEOAAT



All the techniques examined so far involve the substitution of a ciphertext symbol for a plaintext symbol. A very different kind of mapping is achieved by performing some sort of permutation on the plaintext letters. This technique is referred to as a transposition cipher.

The simplest such cipher is the rail fence technique, in which the plaintext is written down as a sequence of diagonals and then read off as a sequence of rows.

For example, to encipher the message “meet me after the toga party” with a rail fence of depth 2, we write the following:

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

The encrypted message is  
MEMATRHTGPRYETEFETEOAAT



# Row Transposition Cipher

- Is a more complex transposition
- Write the message in a rectangle, row by row, and read the message off, column by column, but permute the order of the columns
  - The order of the columns then becomes the key to the algorithm

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

A more complex scheme is

to write the message in a rectangle, row by row, and read the message off, column

by column, but permute the order of the columns. The order of the columns then

becomes the key to the algorithm. For example,

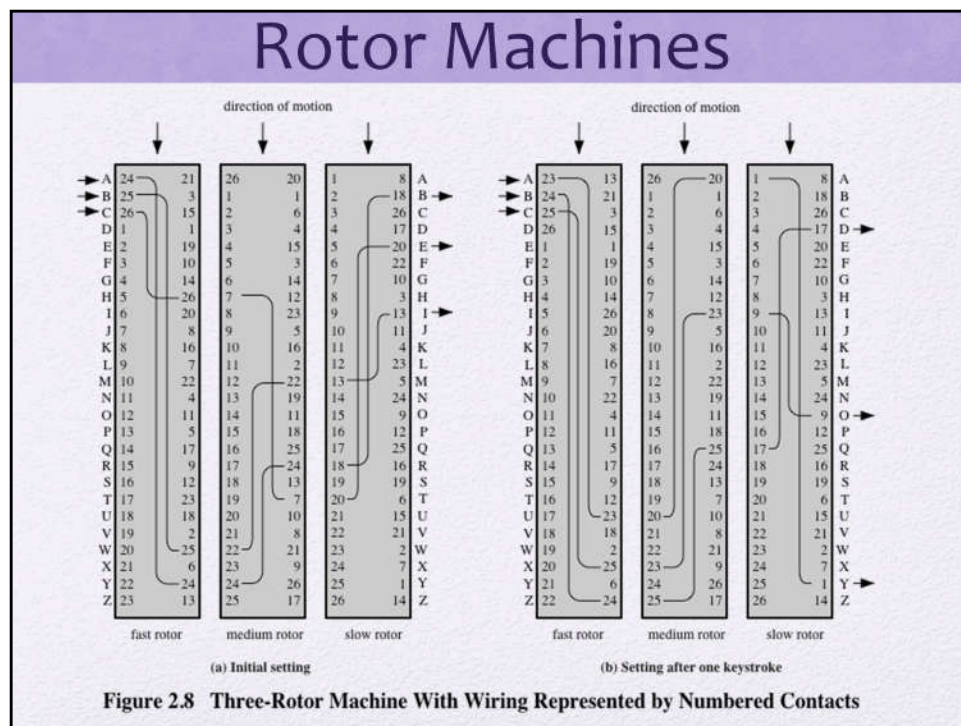
Key: 4 3 1 2 5 6 7

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

Thus, in this example, the key is 4312567. To encrypt, start with the column that is labeled 1, in this case column 3. Write down all the letters in that column.

Proceed to column 4, which is labeled 2, then column 2, then column 1, then columns 5, 6, and 7.



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. Before the introduction of DES, the most important application of the principle of multiple stages of encryption was a class of systems known as rotor machines.

The basic principle of the rotor machine is illustrated in Figure 2.8. 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. For simplicity, only three of the internal connections in each cylinder are shown.

If we associate each input and output pin with a letter of the alphabet, then a single cylinder defines a monoalphabetic substitution. For example, in Figure 2.8, if an operator depresses the key for the letter A, an electric signal is applied to

# Steganography

3rd March

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 26th. 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 those wretched 16 proposals destroy your basic O and A pattern. Certainly this sort of change, if implemented immediately, would bring chaos.

Sincerely yours,


Figure 2.9 A Puzzle for Inspector Morse  
(from *The Silent World of Nicholas Quinn*, by Colin Dexter)

We conclude with a discussion of a technique that (strictly speaking), is not encryption, namely, steganography .

A plaintext message may be hidden in one of two ways. 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.

A simple form of steganography, but one that is time-consuming to construct, is one in which an arrangement of words or letters within an apparently innocuous text spells out the real message. For example, the sequence of first letters of each word of the overall message spells out the hidden message. Figure 2.9 shows an example in which a subset of the words of the overall message is used to convey the hidden message. See if you can decipher this; it's not too hard.

## Other Steganography Techniques



- **Character marking**
  - Selected letters of printed or typewritten text are over-written in pencil
  - The marks are ordinarily not visible unless the paper is held at an angle to bright light
- **Invisible ink**
  - A number of substances can be used for writing but leave no visible trace until heat or some chemical is applied to the paper
- **Pin punctures**
  - Small pin punctures on selected letters are ordinarily not visible unless the paper is held up in front of a light
- **Typewriter correction ribbon**
  - Used between lines typed with a black ribbon, the results of typing with the correction tape are visible only under a strong light

Various other techniques have been used historically; some examples are the following [MYER91]:

- **Character marking:** Selected letters of printed or typewritten text are overwritten in pencil. The marks are ordinarily not visible unless the paper is held at an angle to bright light.
- **Invisible ink:** A number of substances can be used for writing but leave no visible trace until heat or some chemical is applied to the paper.
- **Pin punctures:** Small pin punctures on selected letters are ordinarily not visible unless the paper is held up in front of a light.
- **Typewriter correction ribbon:** Used between lines typed with a black ribbon, the results of typing with the correction tape are visible only under a strong light.

Although these techniques may seem archaic, they have contemporary equivalents.

OWAYN001 proposes hiding a message by using the least significant bits of

# Summary

- Symmetric Cipher Model
  - Cryptography
  - Cryptanalysis and Brute-Force Attack
- Transposition techniques
- Rotor machines
- Substitution techniques
  - Caesar cipher
  - Monoalphabetic ciphers
  - Playfair cipher
  - Hill cipher
  - Polyalphabetic ciphers
  - One-time pad
- Steganography



Chapter 2 summary.