CRYPTOGRAPHY

Dr. Shubhangi Sapkal

Cryptography-Introduction

- Cryptography is a technique of securing information and communications through use of codes so that only those person for whom the information is intended can understand it and process it. Thus preventing unauthorized access to information. The prefix “crypt” means “hidden” and suffix graphy means “writing”.
- In Cryptography the techniques which are used to protect information are obtained from mathematical concepts and a set of rule based calculations known as algorithms to convert messages in ways that make it hard to decode it.
- These algorithms are used for cryptographic key generation, digital signing, verification to protect data privacy, web browsing on internet and to protect confidential transactions such as credit card and debit card transactions.

Cryptography-Introduction

Modern cryptography concerns itself with the following four objectives:

- **Confidentiality**: the information cannot be understood by anyone for whom it was unintended
- **Integrity:** the information cannot be altered in storage or transit between sender and intended receiver without the alteration being detected
- **Non-repudiation**: the creator/sender of the information cannot deny at a later stage his or her intentions in the creation or transmission of the information
- **Authentication**: the sender and receiver can confirm each other's identity and the origin/destination of the information
- Procedures and protocols that meet some or all of the above criteria are known as cryptosystems.

Cryptography

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.

- **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.

Features Of Cryptography

- **Confidentiality:**
  Information can only be accessed by the person for whom it is intended and no other person except him can access it.
- **Integrity:**
  Information cannot be modified in storage or transition between sender and intended receiver without any addition to information being detected.
- **Non-repudiation:**
  The creator/sender of information cannot deny his or her intention to send information at later stage.
- **Authentication:**
  The identities of sender and receiver are confirmed. As well as destination/origin of information is confirmed.

Types Of Cryptography

In general there are three types Of cryptography:

- **Symmetric Key Cryptography:**
  It is an encryption system where the sender and receiver of message use a single common key to encrypt and decrypt messages. Symmetric Key Systems are faster and simpler but the problem is that sender and receiver have to somehow exchange key in a secure manner. The most popular symmetric key cryptography system is Data Encryption System(DES).
- **Hash Functions:**
  There is no usage of any key in this algorithm. A hash value with fixed length is calculated as per the plain text which makes it impossible for contents of plain text to be recovered. Many operating systems use hash functions to encrypt passwords.
- **Asymmetric Key Cryptography:**
  Under this system a pair of keys is used to encrypt and decrypt information. A public key is used for encryption and a private key is used for decryption. Public key and Private Key are different. Even if the public key is known by everyone the intended receiver can only decode it because he alone knows the private key.

Cryptographic algorithms

- Cryptosystems use a set of procedures known as cryptographic algorithms, or ciphers, to encrypt and decrypt messages to secure communications among computer systems, devices such as smartphones, and applications.
- A cipher suite uses one algorithm for encryption, another algorithm for message authentication, and another for key exchange.
- This process, embedded in protocols and written in software that runs on operating systems and networked computer systems, involves public and private key generation for data encryption/decryption, digital signing and verification for message authentication, and key exchange.

Types of cryptography

- **Single key or symmetric key encryption** algorithms create a fixed length of bits known as a block cipher with a secret key that the creator/sender uses to encipher data (encryption) and the receiver uses to decipher it. Types of symmetric-key cryptography include the Advanced Encryption Satndard (AES), a specification established in November 2001 by the National Institute of Standards and Technology as a Federal Information Processing Standard (FIPS 197), to protect sensitive information. The standard is mandated by the U.S. government and widely used in the private sector.

- In June 2003, AES was approved by the U.S. government for classified information. It is a royalty-free specification implemented in software and hardware worldwide. AES is the successor of the Data Encryption Standard (DES)_ and DES3. It uses longer key lengths (128-bit, 192-bit, 256-bit) to prevent brute force and other attacks.
- Public-key or symmetric-key encryption algorithms use a pair of keys, a public key associated with the creator/sender for encrypting messages and a private key that only the originator knows (unless it is exposed or they decide to share it) for decrypting that information.
- The types of public-key cryptography include RSA, used widely on the internet; Elliptic Curve Digital Signature Algorithm (ECDSA) used by Bitcoin; Digital Signature Algorithm (DSA) adopted as a Federal Information Processing Standard for digital signatures by NIST in FIPS 86; and Diffie-Hellman key exchange.
- To maintain data integrity in cryptography, hash functions, which return a deterministic output from an input value, are used to map data to a fixed data size. Types of cryptographic hash functions include SHA-1 (Secure Hash Algorithm 1), SHA-2 and SHA-3.

symmetric Cipher Model

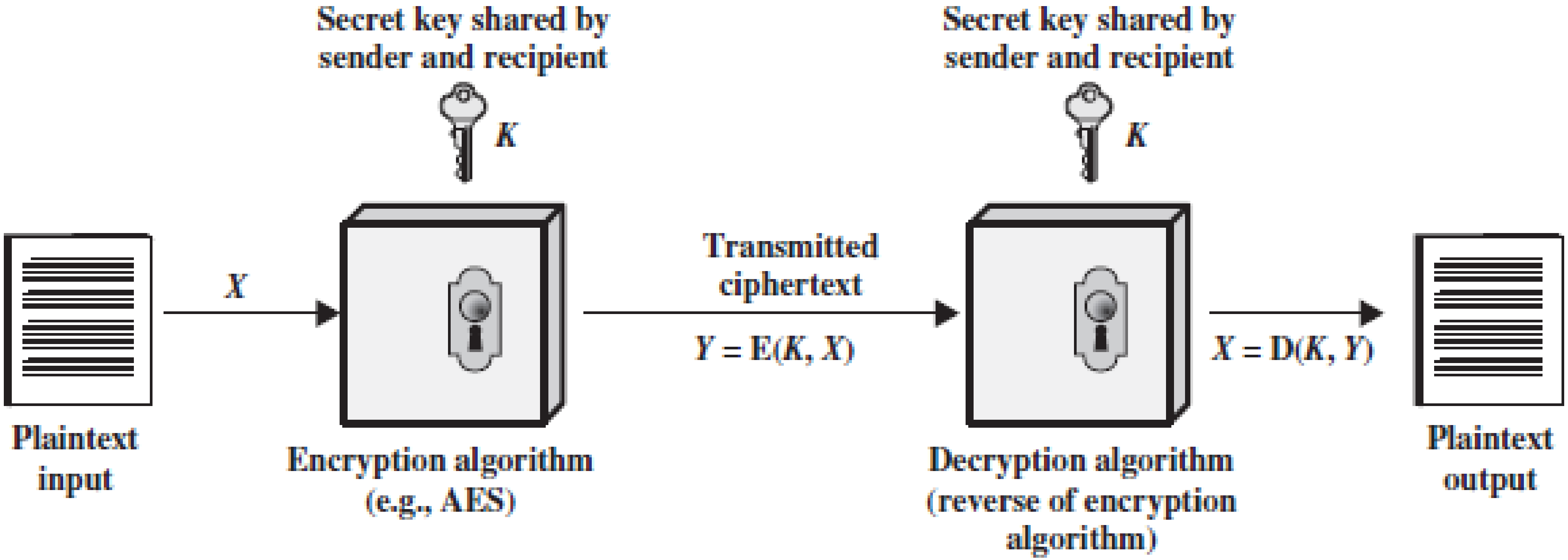

Secret key shared by sender and recipient
K
Secret key shared by sender and recipient
K
X
Transmitted ciphertext
$Y = \mathrm{E}(K, X)$
$X = \mathrm{D}(K, Y)$
Plaintext input
Encryption algorithm (e.g., AES)
Decryption algorithm (reverse of encryption algorithm)
Plaintext output

Figure 2.1 Simplified Model of Symmetric Encryption

Symmetric Encryption

A symmetric encryption scheme has five ingredients

- **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.

Symmetric Encryption

- **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.

- **Decryption algorithm:** This is essentially the encryption algorithm run in

reverse. It takes the ciphertext and the secret key and produces the original plaintext.

Requirements of Encryption

There are two requirements for secure use of conventional encryption:

1. We need a strong encryption algorithm. At a minimum, we would like the algorithm to be such that an opponent who knows the algorithm and has access to one or more ciphertexts would be unable to decipher the ciphertext or figure out the key. This requirement is usually stated in a stronger form: The opponent should be unable to decrypt ciphertext or discover the key even if he or she is in possession of a number of ciphertexts together with the plaintext that produced each ciphertext.

2. Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the key secure. If someone can discover the key and knows the algorithm, all communication using this key is readable.

Symmetric Encryption

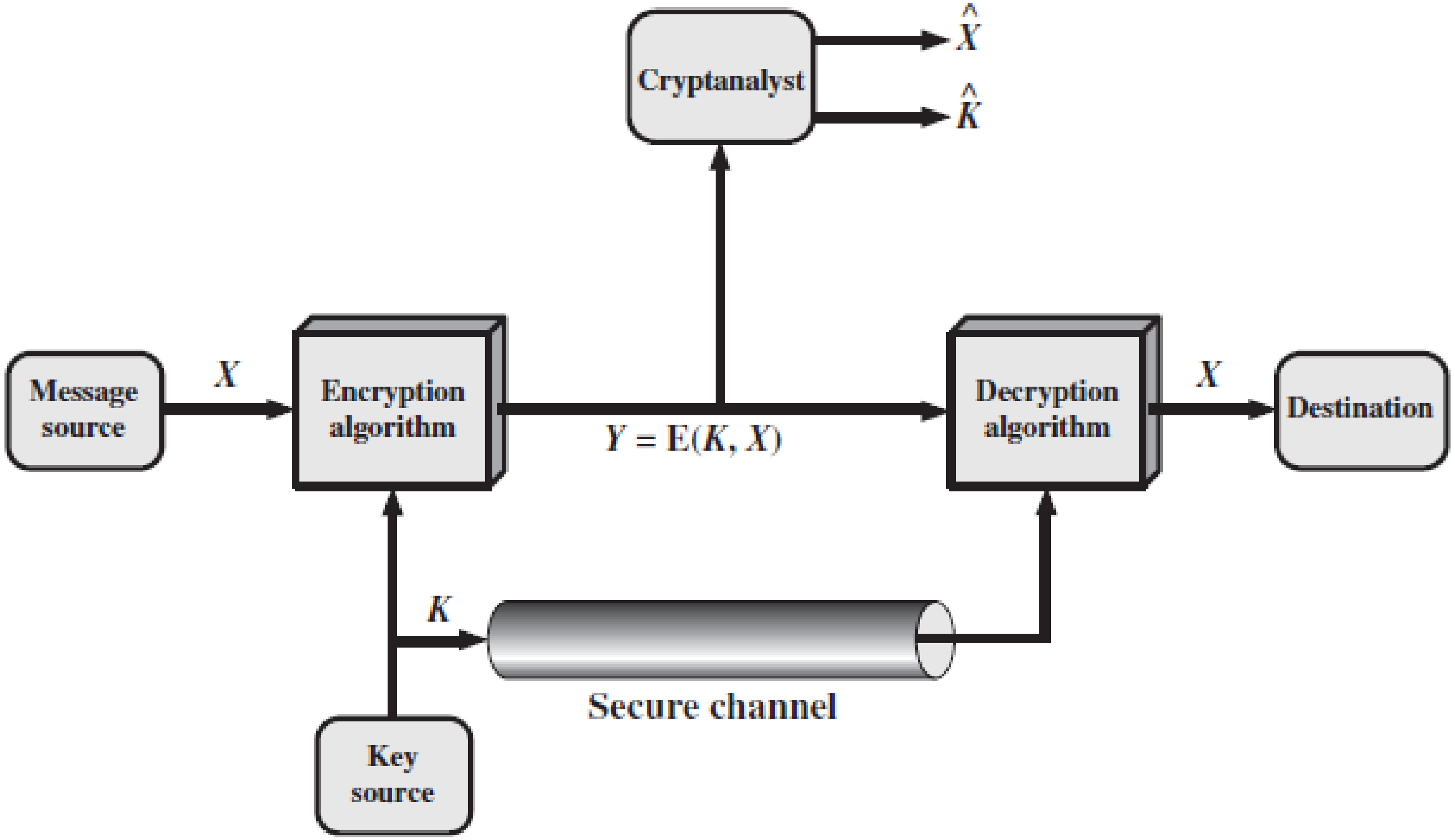


Figure 2.2 Model of Symmetric Cryptosystem

Cryptanalysis and Brute-Force Attack

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.

Cryptanalysis and Brute-Force Attack

An encryption scheme is said to be **computationally secure** if either of the two criteria are met.

1. The cost of breaking the cipher exceeds the value of the encrypted information.

2. The time required to break the cipher exceeds the useful lifetime of the information.

Substitution Techniques

Caesar Cipher

- 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. For example,

  plain: meet me after the toga party

  cipher: PHHW PH DIWHU WKH WRJD SDUWB

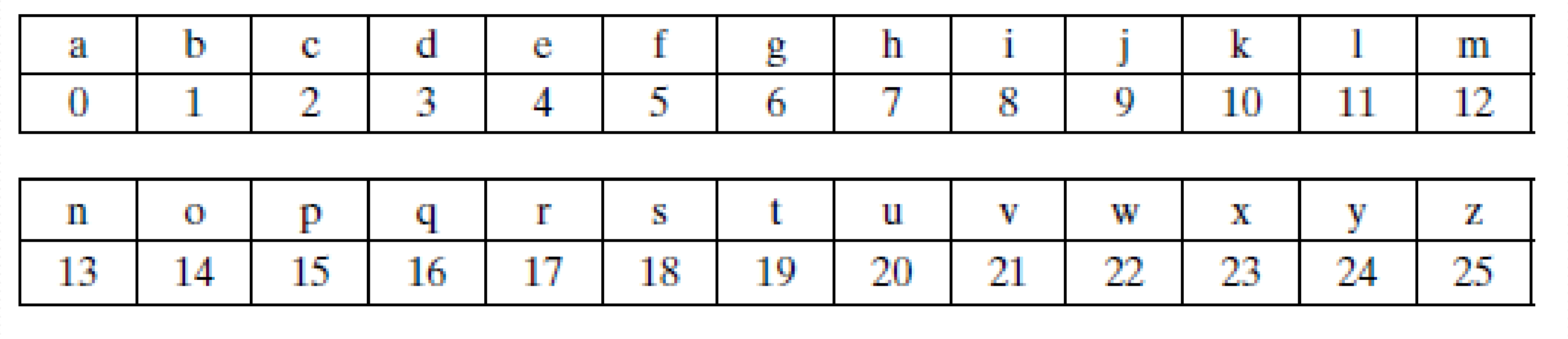
Caesar Cipher

We can define the transformation by listing all possibilities, as follows:

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

Let us assign a numerical equivalent to each letter:



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

Caesar Cipher

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

  $$C = \mathrm{E}(3, p) = (p + 3) \bmod 26$$

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

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

  The decryption algorithm is simply

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

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.
- 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.

Monoalphabetic Ciphers

- If, instead, the “cipher” line can be any permutation of the 26 alphabetic characters, then there are 26! or greater than 4 * 1026 possible keys.
- This is 10 orders of magnitude greater than the key space for DES and would seem to eliminate brute-force techniques for cryptanalysis. Such an approach is referred to as a **monoalphabetic substitution cipher**, because a single cipher alphabet (mapping from plain alphabet to cipher alphabet) is used per message.

Monoalphabetic Ciphers

- 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.
- The ciphertext to be solved is

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

Monoalphabetic Ciphers

- 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 following Figure

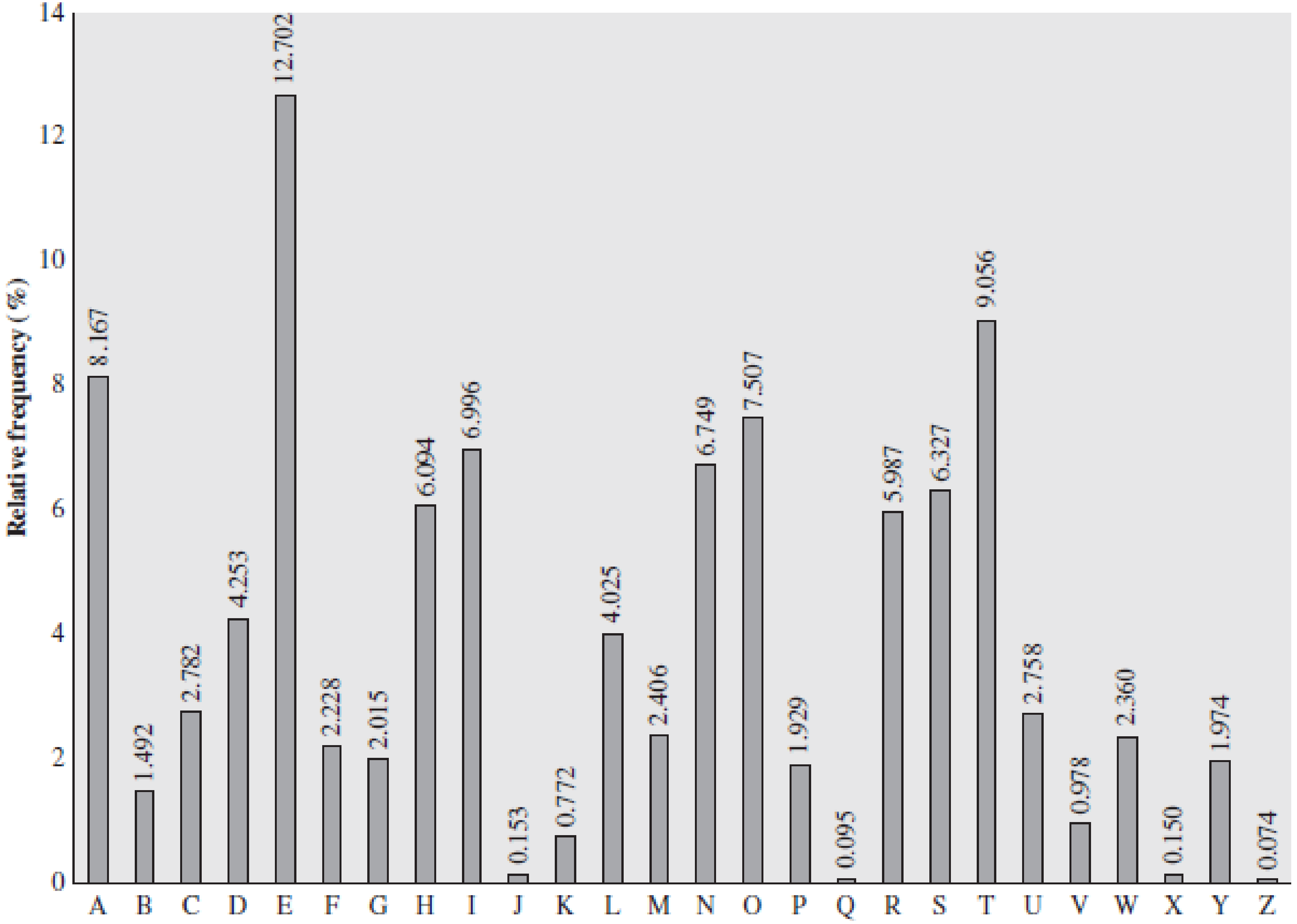
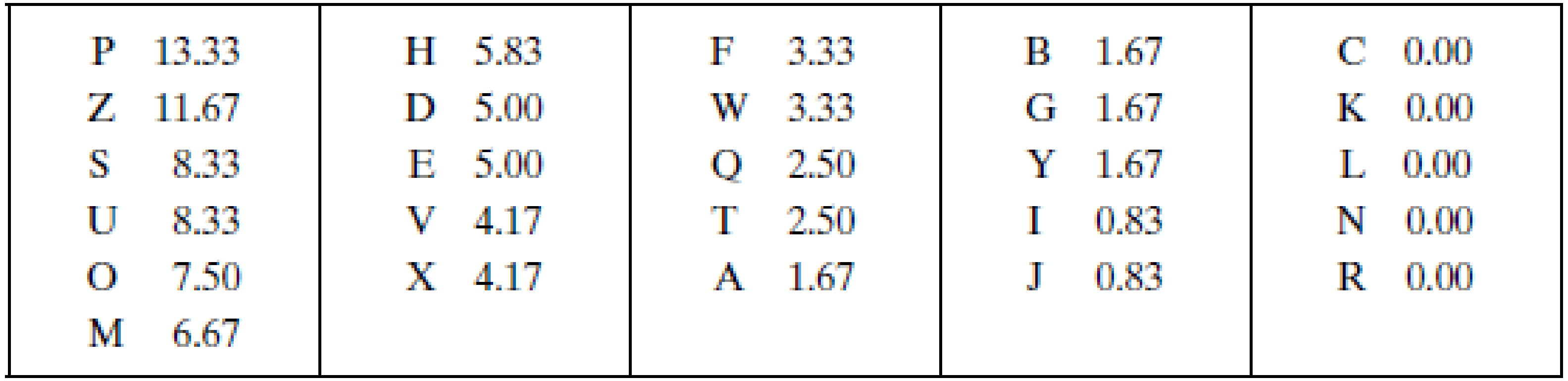


Figure 2.5 Relative Frequency of Letters in English Text

Monoalphabetic Ciphers

- the relative frequencies of the letters in the ciphertext (in percentages) are as follows:



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								

Monoalphabetic Ciphers

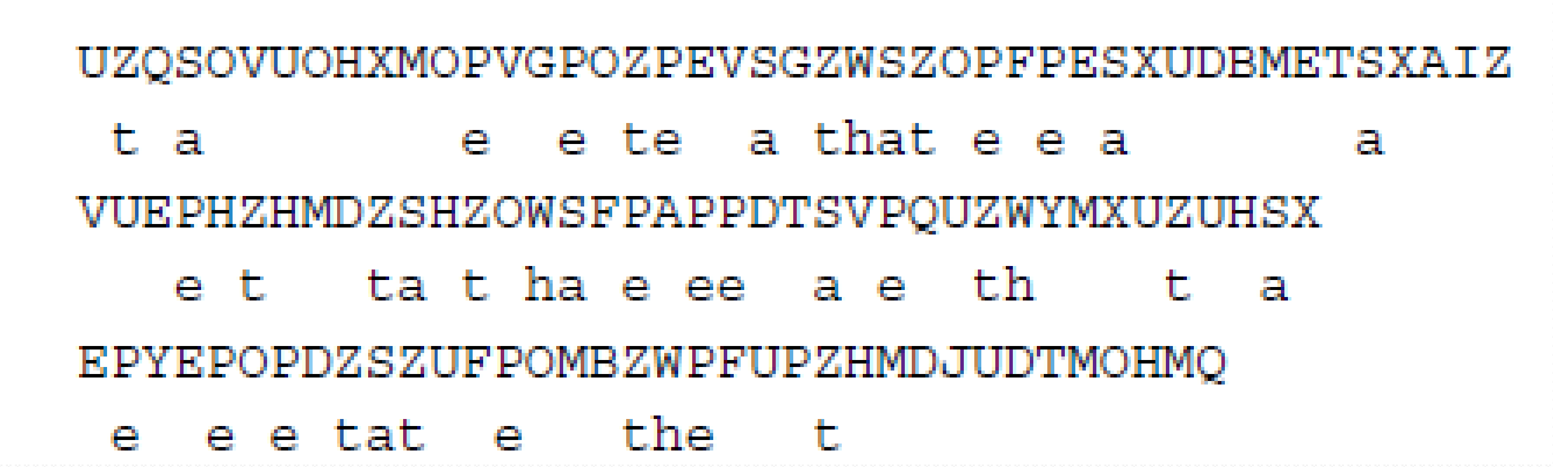
- Comparing this breakdown with Figure 2.5, it seems likely that cipher letters P and Z are the equivalents of plain letters e and t, but it is not certain which is which.
- The letters S, U, O, M, and H are all of relatively high frequency and probably correspond to plain letters from the set {a, h, i, n, o, r, s}.
- The letters with the lowest frequencies (namely, A, B, G, Y, I, J) are likely included in the set {b, j, k, q, v, x, z}.

Monoalphabetic Ciphers

- 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.
- 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.

Monoalphabetic Ciphers

- So far, then, we have



```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFFPESXUDBMETSXAIZ
 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
```

Monoalphabetic Ciphers

- Only four letters have been identified, but already we have quite a bit of the message. Continued analysis of frequencies plus trial and error should easily yield a solution from this point.
- The complete plaintext, with spaces added between words, follows:

“it was disclosed yesterday that several informal but

direct contacts have been made with political

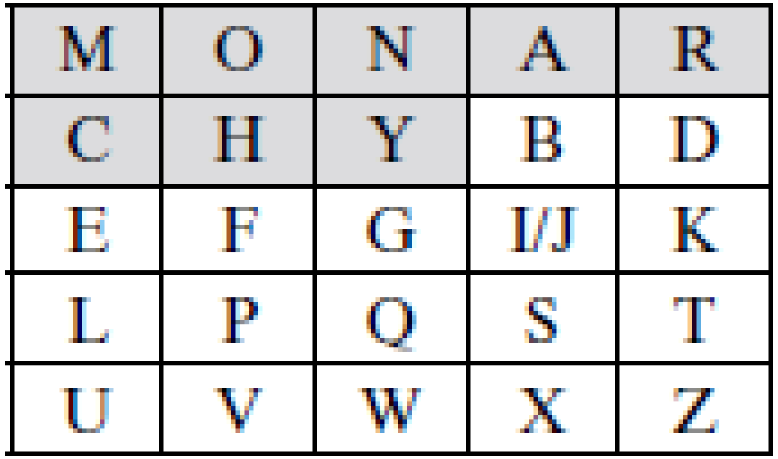
representatives of the viet cong in Moscow”

Monoalphabetic Ciphers

- Monoalphabetic ciphers are easy to break because they reflect the frequency data of the original alphabet.
- If the number of symbols assigned to each letter is proportional to the relative frequency of that letter, then single-letter frequency information is completely obliterated.

Playfair Cipher

- 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. Here is an example,



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

Playfair Cipher

- 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:

Playfair Cipher

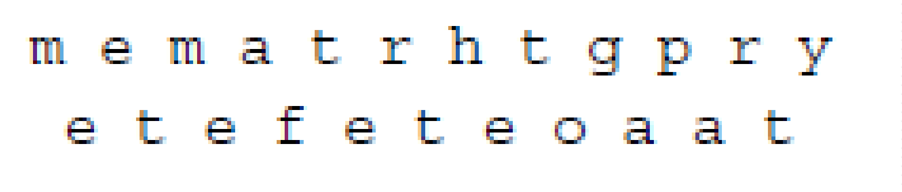
1. Repeating plaintext letters that are in the same pair are separated with a filler letter, such as x, so that balloon would be treated as ba lx lo on.
2. Two plaintext letters that fall in the same row of the matrix are each replaced by the letter to the right, with the first element of the row circularly 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.
4. Otherwise, each plaintext letter in a pair is replaced by the letter that lies in its own row and the column occupied by the other plaintext letter. Thus, hs becomes BP and ea becomes IM (or JM, as the encipherer wishes).

Transposition Techniques

- 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.

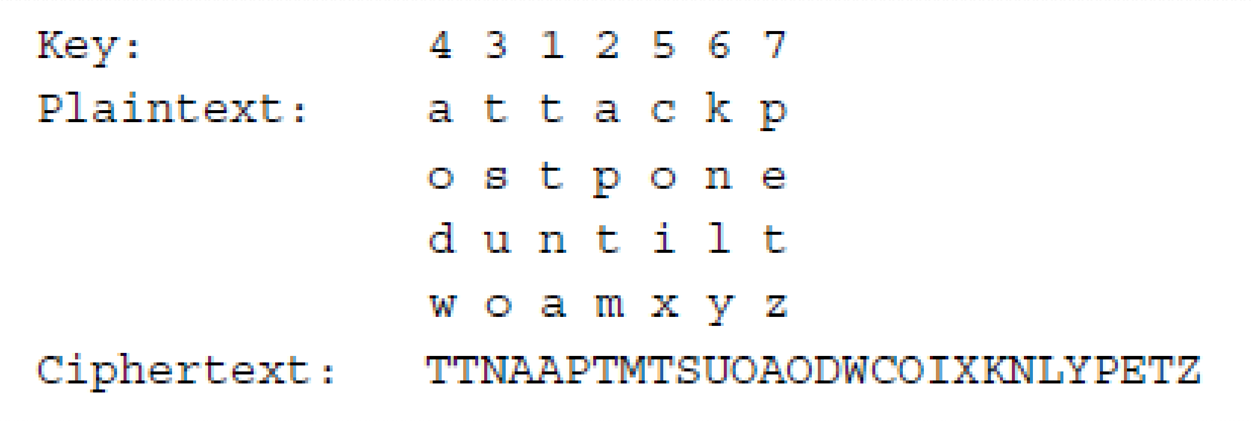
rail fence technique

- In this technique, 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
```

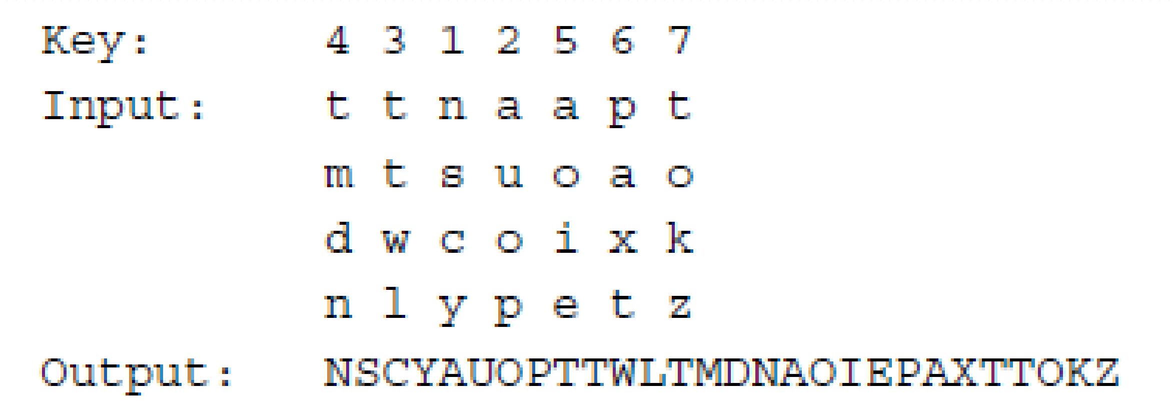
- The encrypted message is
  MEMATRHTGPRYETEFETEOAAT
- A more complex scheme is to write the message in a rectangle, row by row, and read the message off, column by column, but permute the order of the columns. The order of the columns then becomes the key to the algorithm. For example,



```
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 transposition cipher can be made significantly more secure by performing more than one stage of transposition. The result is a more complex permutation that is not easily reconstructed.
- Thus, if the foregoing message is reencrypted using the same algorithm,



```
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
```