

Lec. 6

8- One-Time Pad

*Use a random key that is as long as the message, so that the key need not be repeated.

*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.

* one-time pad, is unbreakable.

((It produces random output that bears no statistical relationship to the plaintext)).

* Because the ciphertext provides no information about the original message to a cryptanalyst, there is simply no way to break the code .

ENCRYPTION

EXAMPLE:

H	E	L	L	O	↔	plaintext
7(h)	4(e)	11(l)	11(l)	14(o)	↔	
+ 23(X)	12(M)	2(C)	10(K)	11(L)	↔	key
= 30	16	13	21	25	↔	plaintext + key
= 4(E)	16(Q)	13(N)	21(V)	25(Z)	↔	(p+ key)mod26
E	Q	N	V	Z	↔	cipher text

DECRYPTION

EXAMPLE:

E	Q	N	V	Z	↔	cipher text
4(E)	16(Q)	13(N)	21(V)	25(Z)	↔	
- 23(X)	12(M)	2(C)	10(K)	11(L)	↔	key
= -19 = -19 + 26	4	11	11	14	↔	ciphertext-key
7					↔	(c-k)mod26
h	e	l	l	o	↔	plain text

Encrypt the following text “otp” using otp with logic operation key=xyz

Plaintext= otp

Key = xyz

$o = \text{ASCII code for letter "o"} = 111 \text{ as decimal number so } 111_{10} = (01101111)_2$
 $x = \text{ASCII code for letter "x"} = 120 \text{ as decimal number so } 120_{10} = (01111000)_2$

o XOR x =

$o = 111_{10} = (01101111)_2$
 $x = 120_{10} = (01111000)_2$

$$\begin{array}{r} 01101111 \\ \text{XOR} \\ 01111000 \\ \hline (00010111)_2 = 23_{10} \mod 26 = 23 = x \quad o \text{ xor } x = X = 23 + 65 = 88 = X \end{array}$$

t XOR y =

$t = 116_{10} = (01110100)_2$
 $y = 121_{10} = (01111001)_2$

$$\begin{array}{r} 01110100 \\ \text{XOR} \\ 01111001 \\ \hline (00001101)_2 = 13_{10} \mod 26 = 13 = n \quad t \text{ xor } y = N = 13 + 65 = 78 \end{array}$$

p XOR z =

$p = 112_{10} = (01110000)_2$
 $z = 122_{10} = (01111010)_2$

$$\begin{array}{r} 01110000 \\ \text{XOR} \\ 01111010 \\ \hline (00001010)_2 = 10_{10} \mod 26 = 10 = k \quad p \text{ xor } z = K = 10 + 65 = 75 \end{array}$$

Q2: Decrypt the following cipher text “XNK” using otp with logic operation key=xyz

Cipher text= XNK

Key = xyz

X=ASCII code for letter “X” =88 as decimal number so $88_{10} = 23_{10} = 00010111_2$

x= ASCII code for letter “x”= 120 as decimal number so $120_{10} = (01111000)_2$

X XOR x =

$X=88_{10} = 88 - \text{ascii for capital letter (65)} = 23 = (00010111)_2$

$x=120_{10} = (01111000)_2$

$$\begin{array}{r}
 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1 \\
 \text{XOR} \\
 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0 \\
 \hline
 (0\ 1\ 1\ 0\ 1\ 1\ 1\ 1)_2 = 111_{10} = \text{ascii code } 111 = \text{letter o}
 \end{array}$$

N XOR y =

$78-65=13$

$n=13_{10} = (0\ 0\ 0\ 0\ 1\ 1\ 0\ 1)_2$

$y=121_{10} = (01111001)_2$

$$\begin{array}{r}
 0\ 0\ 0\ 0\ 1\ 1\ 0\ 1 \\
 \text{XOR} \\
 0\ 1\ 1\ 1\ 1\ 0\ 0\ 1 \\
 \hline
 (0\ 1\ 1\ 1\ 0\ 1\ 0\ 0)_2 = 116_{10} = t \quad n \text{ XOR } y = t
 \end{array}$$

k XOR z =

$K=75-65=10$

$k=10_{10} = (00001010)_2$

$z=122_{10} = (01111010)_2$

$$\begin{array}{r}
 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0 \\
 \text{XOR} \\
 0\ 1\ 1\ 1\ 1\ 0\ 1\ 0 \\
 \hline
 (0\ 1\ 1\ 1\ 0\ 0\ 0\ 0)_2 = 112_{10} = p \quad K \text{ XOR } z = p
 \end{array}$$

(XNK) XOR (xyz)= (otp)

Cryptanalysis

In this section, we discuss some techniques of cryptanalysis. The general assumption that is usually made is that the opponent, Oscar, knows the cryptosystem being used. This is usually referred to as *Kerckhoff's principle*. Of course, if Oscar does not know the cryptosystem being used, that will make his task more difficult. But we do not want to base the security of a cryptosystem on the (possibly shaky) premise that Oscar does not know what system is being employed. Hence, our goal in designing a cryptosystem will be to obtain security under Kerckhoff's principle.

First, we want to differentiate between different levels of attacks on cryptosystems. The most common types are enumerated as follows.

Ciphertext-only

The opponent possesses a string of ciphertext, y .

Known plaintext

The opponent possesses a string of plaintext, x , and the corresponding ciphertext, y .

Chosen plaintext

The opponent has obtained temporary access to the encryption machinery. Hence he can choose a plaintext string, x , and construct the corresponding ciphertext string, y .

Chosen ciphertext

Chosen ciphertext

The opponent has obtained temporary access to the decryption machinery. Hence he can choose a ciphertext string, y , and construct the corresponding plaintext string, x .

In each case, the object is to determine the key that was used. We note that a chosen ciphertext attack is relevant to public-key cryptosystems, which we discuss in the later chapters.

We first consider the weakest type of attack, namely a ciphertext-only attack. We also assume that the plaintext string is ordinary English text, without punctuation or “spaces.” (This makes cryptanalysis more difficult than if punctuation and spaces were encrypted.)

Many techniques of cryptanalysis use statistical properties of the English language. Various people have estimated the relative frequencies of the 26 letters by compiling statistics from numerous novels, magazines, and newspapers. The estimates in Table 1.1 were obtained by Beker and Piper.

On the basis of the above probabilities, Beker and Piper partition the 26 letters into five groups as follows:

1. *E*, having probability about 0.120
2. *T, A, O, I, N, S, H, R*, each having probabilities between 0.06 and 0.09
3. *D, L*, each having probabilities around 0.04
4. *C, U, M, W, F, G, Y, P, B*, each having probabilities between 0.015 and 0.028

It may also be useful to consider sequences of two or three consecutive letters called *digrams* and *trigrams*, respectively. The 30 most common digrams are (in decreasing order) *TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI, and OF*. The twelve most common trigrams are (in decreasing order) *THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, and DTH*.

Table 1.1Probabilities of Occurrence of the 26 Letters

letter	probability	letter	probability
A	.082	N	.067
B	.015	O	.075
C	.028	P	.019
D	.043	Q	.001
E	.127	R	.060
F	.022	S	.063
G	.020	T	.091
H	.061	U	.028
I	.070	V	.010
J	.002	W	.023
K	.008	X	.001
L	.040	Y	.020
M	.024	Z	.001

Cryptanalysis Monoalphabetic Ciphers

Cryptanalysis is the science and study of methods of breaking ciphers. It is assumed that the ciphertext is sent over insecure communications lines and is available to the cryptanalyst. His aim is to recover the plaintext from the ciphertext without knowing the key parameters[17][11][18]. A cipher is breakable if it is possible to determine the plaintext or key parameters from the ciphertext. There are three basic methods of attack : ciphertext-only, known-plaintext, and chosen-plaintext. Under a ciphertext-only attack, a cryptanalyst must determine the key solely from intercepted ciphertext, though the method of encryption and certain probable words may be known. On the other hand, a known-plaintext attack requires a substantial amount of plaintext and ciphertext be known, and a chosen-plaintext attack is able to acquire the ciphertext corresponding to selected plaintext[21, 25].

2-5-1 Cryptanalysis a mono alphabetic substitution ciphers

If the cryptanalysis knows the nature of the plaintext (e.g., non compressed 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. The cipher text to be solved is:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

As a first step, the relative frequency of the letters can be determine and compare to a standard frequency distribution for English, such as is shown in fig. 3.3. If the message were long enough, this technique alone might be sufficient, but because this is a relatively short message, we cannot expect an exact match. In any case, the relative frequencies of the letters in the cipher text (in percentages) are as follows:-

P 13.33	H 5.83	F 3.33	B 1.67	C 0.000
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.000
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.000
U 8.33	V 4.17	T 2.50	I 0.83	N 0.000
O 7.50	X 4.17	A 1.67	J 0.83	R 0.000
M 6.67				

Comparing this breakdown with figure 3.3 , it seems likely that cipher letters P and Z are 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 The genetic algorithms are.

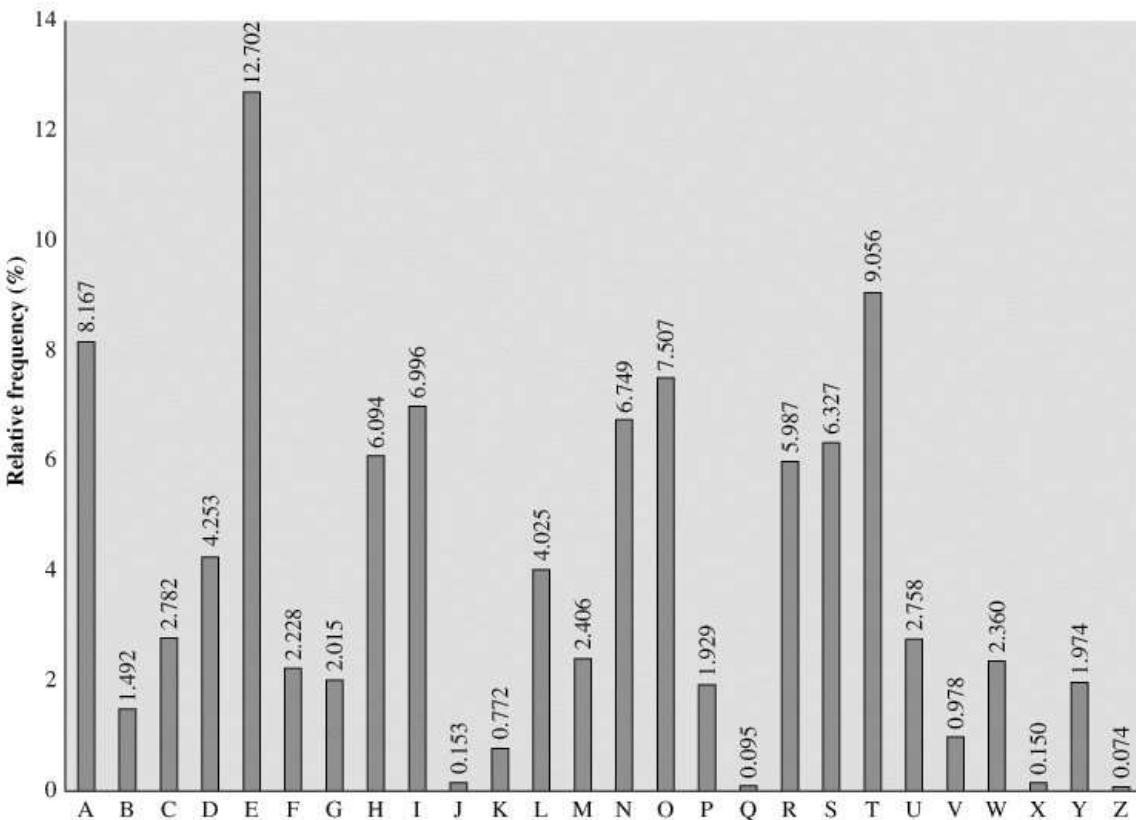


Figure.3.3: Letter Frequency for English Language

There are a number of ways to proceed at this point. We could make some tentative assignments and start to fill in the plaintext to see if it looks like a reasonable "skeleton" of a message. A more systematic approach is to look for other regularities. For example, certain words may be known to be in the text. Or we could look for repeating sequences of chapter letters and try to deduce their plaintext equivalents.

A powerful tool is to look at the frequency of two-letter combinations, known as digrams. A table similar to figure 3.3 could 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".

U	Z	Q	S	O	V	U	O	H	X	M	O	P	V	G	P	O	Z	P	E	V	S	G	Z	W	S	Z	O	P	F
	t		a								e				e		t	e			a		t	h	a	t		e	
P	E	S	X	U	D	B	M	E	T	S	X	A	I	Z	V	U	E	P	H	Z	H	M	D	Z	S	H	Z	O	W
e		a								a							e		t					t	a		t		h
S	F	P	A	P	P	D	T	S	V	P	Q	U	Z	W	Y	M	X	U	Z	U	H	S	X	E	P	Y	E	P	O
a		e		e	e			a		e			t	h					t			a			e			e	
P	D	Z	S	Z	U	F	P	O	M	B	Z	W	P	F	U	P	Z	H	M	D	J	U	D	T	M	O	H	M	Q
e		t	a	t			e				t	h	e					t											

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**

Mono alphabetic ciphers are easy to break because they reflect the frequency data of the original alphabet

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

Table 1: Types of attacks on encrypted messages

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