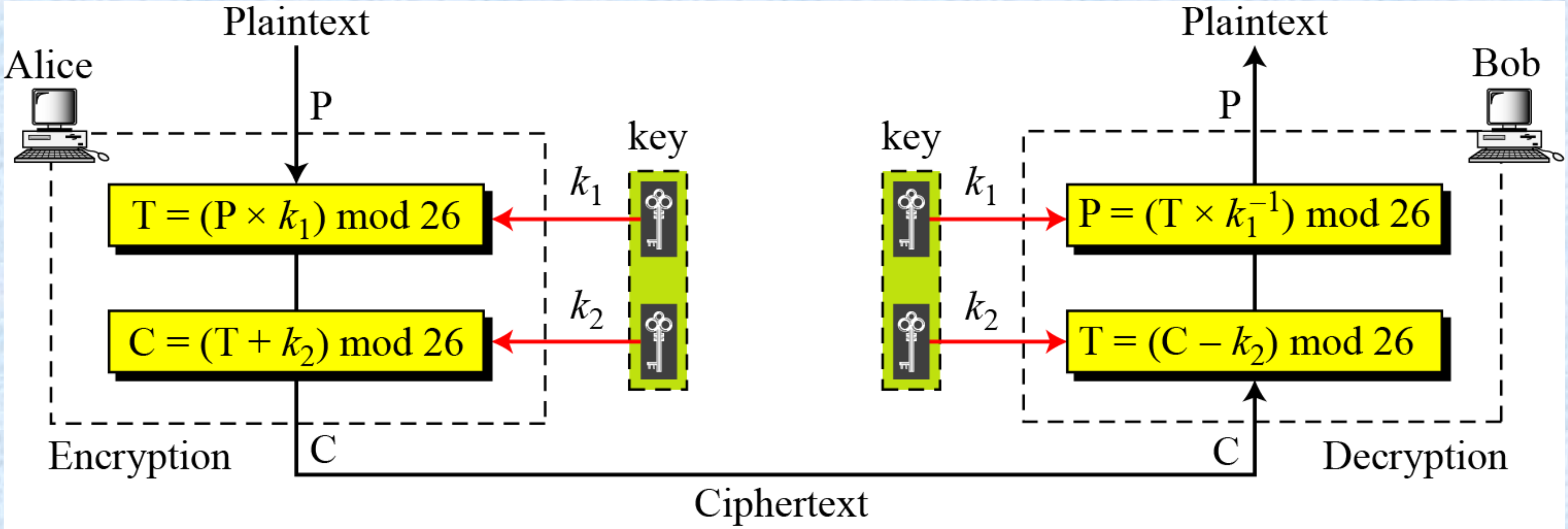
Cryptanalysis

By

Alaaddin Abbas Abdulhasen

Affine Ciphers

Figure 1 ***Affine cipher***



$$C = (P \times k_1 + k_2) \bmod 26 \qquad P = ((C - k_2) \times k_1^{-1}) \bmod 26$$

where k_1^{-1} is the multiplicative inverse of k_1 and $-k_2$ is the additive inverse of k_2

The affine cipher uses a pair of keys in which the first key is from Z_{26}* and the second is from Z_{26}. The size of the key domain is 25 × 11 = 275.

Example 1

Use an affine cipher to encrypt the message "hello" with the key pair (7, 2).

P: h → 07	Encryption: $(07 \times 7 + 2) \bmod 26$	C: 25 → Z
P: e → 04	Encryption: $(04 \times 7 + 2) \bmod 26$	C: 04 → E
P: l → 11	Encryption: $(11 \times 7 + 2) \bmod 26$	C: 01 → B
P: l → 11	Encryption: $(11 \times 7 + 2) \bmod 26$	C: 01 → B
P: o → 14	Encryption: $(14 \times 7 + 2) \bmod 26$	C: 22 → W

Example 2

Use the affine cipher to decrypt the message "ZEBBW" with the key pair (7, 2) in modulus 26.

Solution

C: Z $\rightarrow$ 25	Decryption: $((25 - 2) \times 7^{-1}) \bmod 26$	P:07 $\rightarrow$ h
C: E $\rightarrow$ 04	Decryption: $((04 - 2) \times 7^{-1}) \bmod 26$	P:04 $\rightarrow$ e
C: B $\rightarrow$ 01	Decryption: $((01 - 2) \times 7^{-1}) \bmod 26$	P:11 $\rightarrow$ l
C: B $\rightarrow$ 01	Decryption: $((01 - 2) \times 7^{-1}) \bmod 26$	P:11 $\rightarrow$ l
C: W $\rightarrow$ 22	Decryption: $((22 - 2) \times 7^{-1}) \bmod 26$	P:14 $\rightarrow$ o

The additive cipher is a special case of an affine cipher in which $k_1 = 1$. The multiplicative cipher is a special case of affine cipher in which $k_2 = 0$.

Monoalphabetic Substitution Cipher

Because additive, multiplicative, and affine ciphers have small key domains, they are very vulnerable to brute-force attack.

A better solution is to create a mapping between each plaintext character and the corresponding ciphertext character. Alice and Bob can agree on a table showing the mapping for each character.

Figure 2 ***An example key for monoalphabetic substitution cipher***

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

Example 4

We can use the key in Figure 1 to encrypt the message

this message is easy to encrypt but hard to find the key

The ciphertext is

ICFVQRVVNEFVRNVSIYRGAHSLIOJICNHTIYBFGTICRXRS

Polyalphabetic Ciphers

In polyalphabetic substitution, each occurrence of a character may have a different substitute. The relationship between a character in the plaintext to a character in the ciphertext is one-to-many.

$$P = P_1 P_2 P_3 \ldots \qquad C = C_1 C_2 C_3 \ldots \qquad k = (k_1, P_1, P_2, \ldots)$$

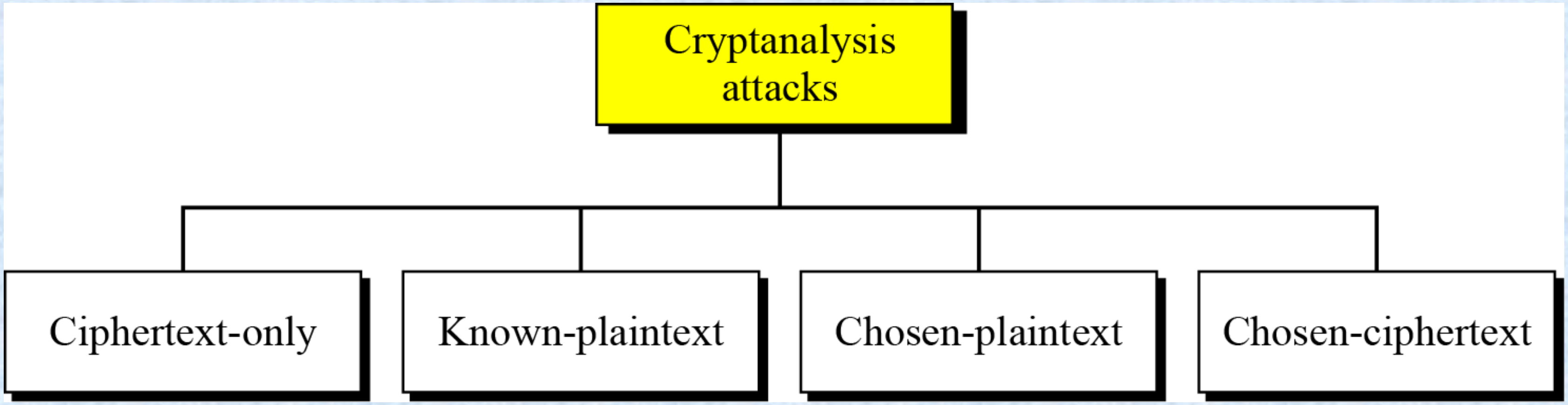
$$\text{Encryption: } C_i = (P_i + k_i) \bmod 26 \qquad \text{Decryption: } P_i = (C_i - k_i) \bmod 26$$

Cryptanalysis

Cryptanalysis

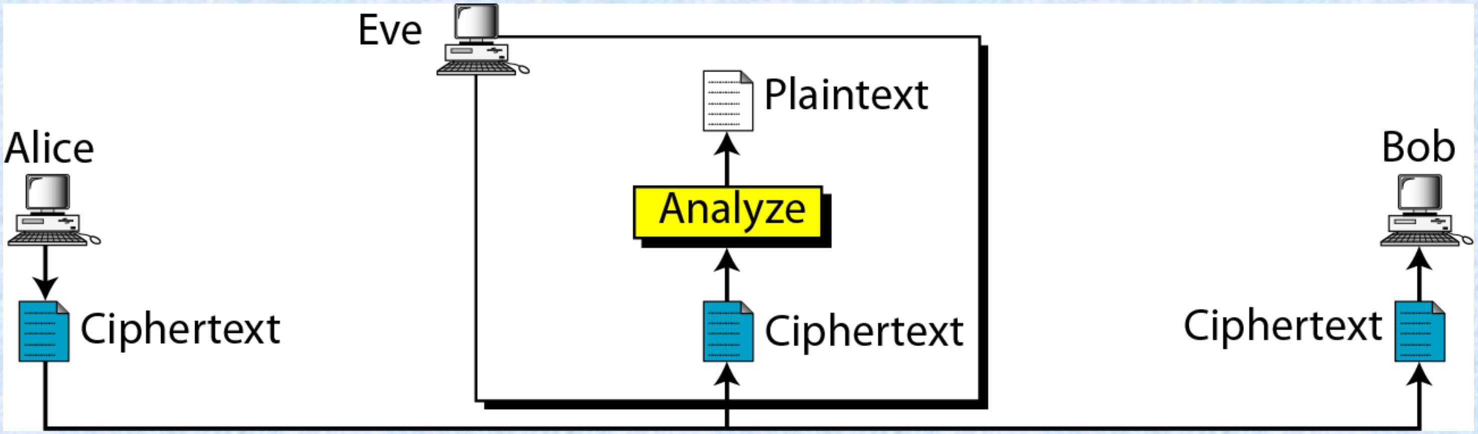
As cryptography is the science and art of creating secret codes, cryptanalysis is the science and art of breaking those codes.

Figure 3 ***Cryptanalysis attacks***



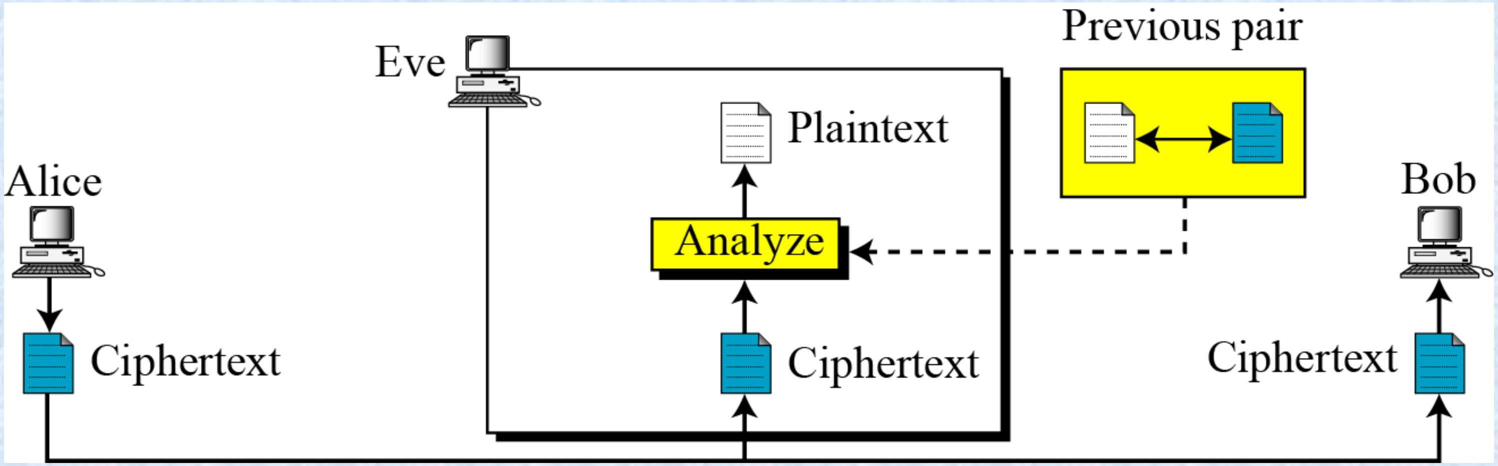
Ciphertext-Only Attack

Figure 4 ***Ciphertext-only attack***



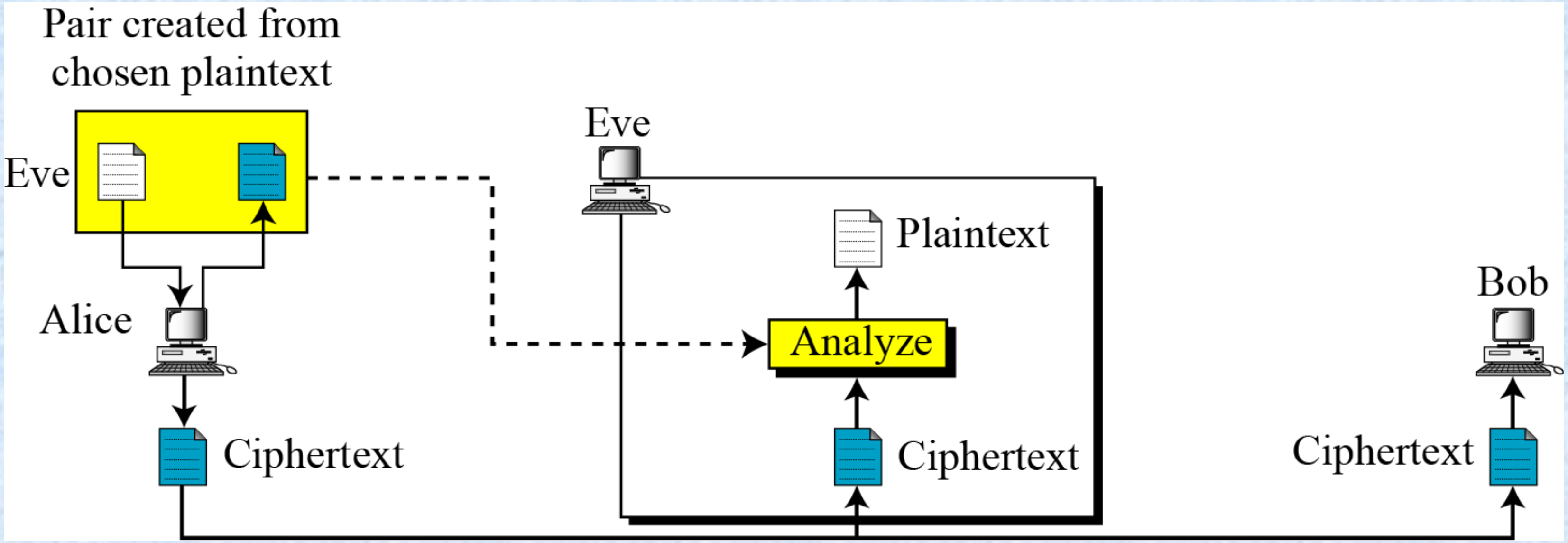
Known-Plaintext Attack

Figure 5 ***Known-plaintext attack***



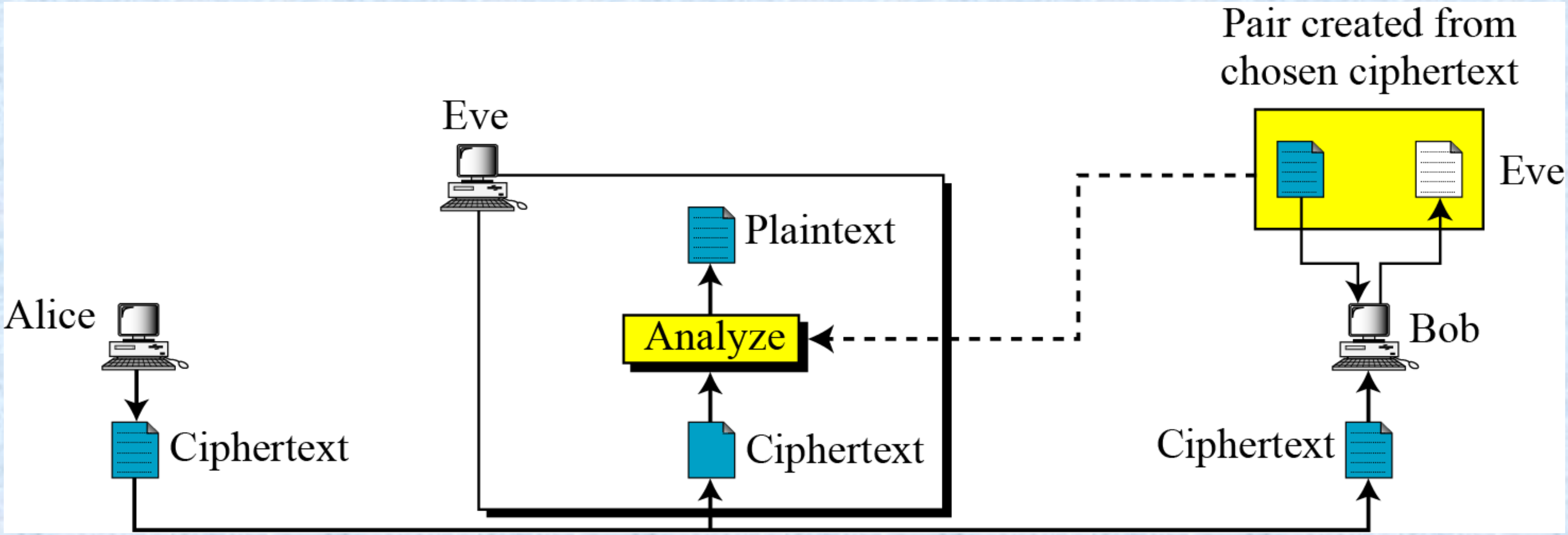
Chosen-Plaintext Attack

Figure 6 ***Chosen-plaintext attack***



Chosen-Cipher text Attack

Figure 7 ***Chosen-ciphertext attack***



Example 5 brute-force attack

Eve has intercepted the ciphertext "UVACLYFZLJBYL". Show how she can use a brute-force attack to break the cipher.

Solution

Eve tries keys from 1 to 7. With a key of 7, the plaintext is "not very secure", which makes sense.

Ciphertext: UVACLYFZLJBYL

K = 1	→	**Plaintext:** tuzbkxeykiaxk
K = 2	→	**Plaintext:** styajwdxjhzwj
K = 3	→	**Plaintext:** rsxzivcwigyvi
K = 4	→	**Plaintext:** qrwyhubvhfxuh
K = 5	→	**Plaintext:** pqvxgtaugewtg
K = 6	→	**Plaintext:** opuwfsztfdvsf
K = 7	→	**Plaintext:** notverysecure

Frequency attack

Table 1 *Frequency of characters in English*

Letter	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*
E	12.7	H	6.1	W	2.3	K	0.08
T	9.1	R	6.0	F	2.2	J	0.02
A	8.2	D	4.3	G	2.0	Q	0.01
O	7.5	L	4.0	Y	2.0	X	0.01
I	7.0	C	2.8	P	1.9	Z	0.01
N	6.7	U	2.8	B	1.5		
S	6.3	M	2.4	V	1.0		

Table 2 *Frequency of diagrams and trigrams*

Digram	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, OF
Trigram	THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH

Example 6

Eve has intercepted the following ciphertext. Using a statistical attack, find the plaintext.

XLILSYWIMWRSAJSVWEPIJSVJSYVQMPPMSRHSPPEVWMXMWASVX-LQSVILY-VVCFIJSVIXLIWIPPIVVIGIMZIWQSVISJJIVW

Solution

When Eve tabulates the frequency of letters in this ciphertext, she gets: I =14, V =13, S =12, and so on. The most common character is I with 14 occurrences. This means key = 4.

the house is now for sale for four million dollars it is worth more hurry before the seller receives more offers

Cryptanalysis

- As cryptanalysts develop techniques for breaking ciphers, cryptographers must develop new ciphers which are more difficult to break
- This has been an ongoing process for over 2000 years
- Current cryptographic techniques are highly mathematical in nature
- Government Communications Headquarters (GCHQ) and the National Security Agency (NSA) currently undertake such work in the UK and USA respectively

Caesar Substitution Cipher

- One of the earliest recorded uses of a cipher is by Julius Caesar
- This (now simple) type of cipher is commonly known as the Caesar Substitution Cipher
- Each letter of the alphabet is substituted by another letter, according to the cipher algorithm

Ciphertext

PCQ VMJYPD LBYK LYSO KBXBJXWXV BXV ZCJPO EYPD KBXBJYUXJ LBJOO KCPK. CP LBO LBCMKXPV XPV IYJKL PYDBL, QBOP KBO BXV OPVOV LBO LXRO CI SX'XJMI, KBO JCKO XPV EYKKOV LBO DJCMPV ZOICJO BYS, KXUYPD: 'DJOXL EYPD, ICJ X LBCMKXPV CPO PYDBLK Y BXNO ZOOP JOACMPLYPD LC UCM LBO IXZROK CI FXKL XDOK XPV LBO RODOPVK CI XPAYOPL EYPDK. SXU Y SXEO KC ZCRV XK LC AJXNO X IXNCMJ CI UCMJ SXGOKLU?'

OFYRCDMO, LXROK IJCS LBO LBCMKXPV XPV CPO PYDBLK

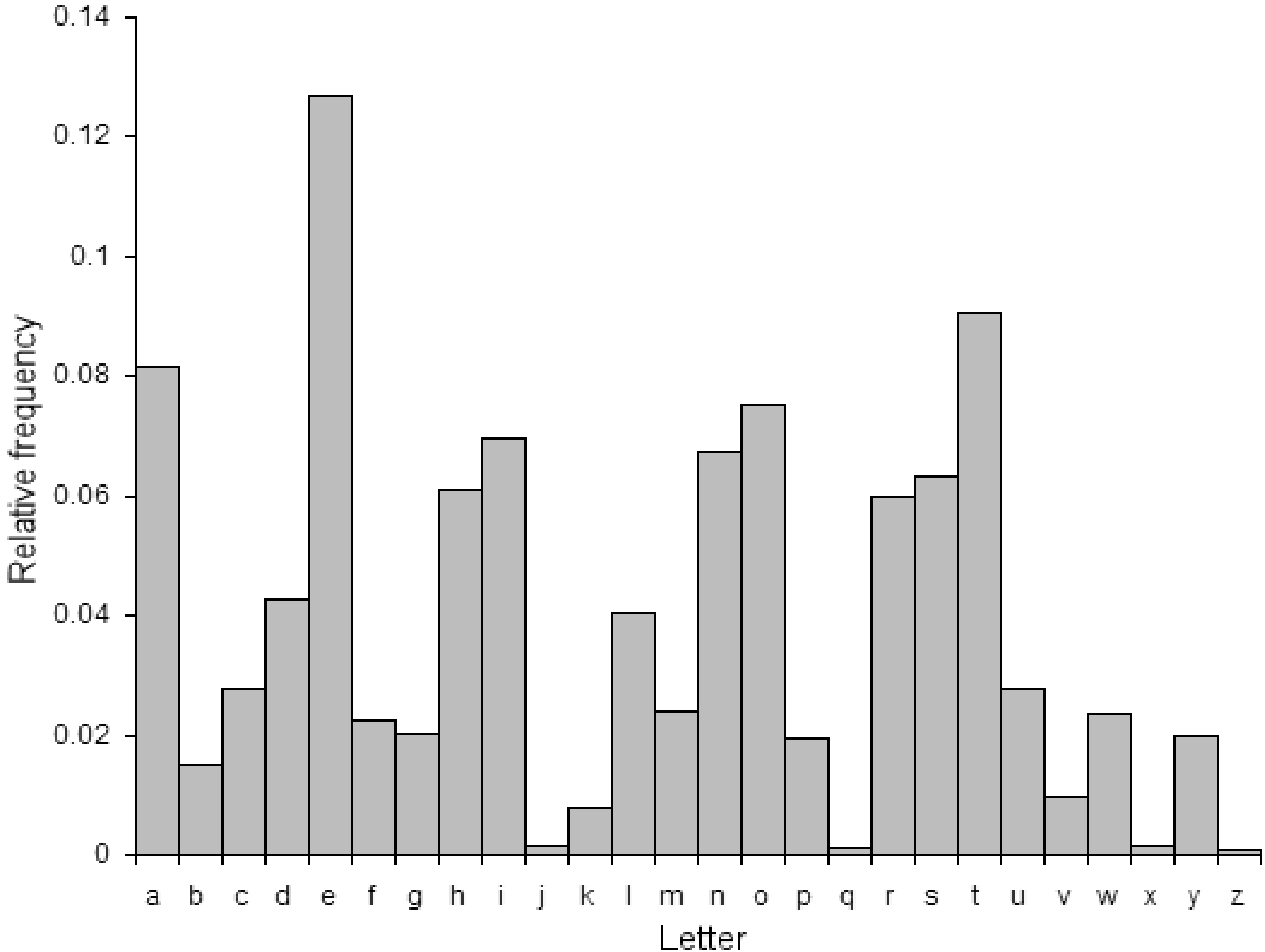
Breaking the Cipher

- Simple substitution cipher
- Plain text in English
- We can use a technique known as frequency analysis to begin with
  - In any given language, on average, each letter makes up a specific percentage of that written language
  - Dependant on type of text
  - Not effective for short messages
- Ciphertext is normally shown in capitals, whilst plaintext uses lower case

Frequency analysis of letters

- Frequency analysis for passages taken from English newspapers and novels (approx. 100,000 characters)

Letter	Percent	Letter	Percent	Letter	Percent	Letter	Percent
a	8.2	h	6.1	o	7.5	v	1.0
b	1.5	i	7.0	p	1.9	w	2.4
c	2.8	j	0.2	q	0.1	x	0.2
d	4.3	k	0.8	r	6.0	y	2.0
e	12.7	l	4.0	s	6.3	z	0.1
f	2.2	m	2.4	t	9.1		
g	2.0	n	6.7	u	2.8		



0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Relative frequency
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
Letter

Analysis of the Encrypted Message

- Frequency analysis for enciphered message

Letter	Percent	Letter	Percent	Letter	Percent	Letter	Percent
a	0.9	h	0.0	o	11.2	v	5.3
b	7.4	i	3.3	p	9.2	w	0.3
c	8.0	j	5.3	q	0.6	x	10.1
d	4.1	k	7.7	r	1.8	y	5.6
e	1.5	l	7.4	s	2.1	z	1.5
f	0.6	m	3.3	t	0.0		
g	0.3	n	0.9	u	1.8		

Analysis of the Encrypted Message

- The three most common letters in the ciphertext are `O`, `P` and `X`
- Therefore it seems likely that these represent `e`, `t` or `a` in plaintext
- Next, lets see which letters `O`, `P` and `X` are adjacent to in the ciphertext

Analysis of the Encrypted Message

- Number of occurrences of letters adjacent to `O`, `X` and `P` in the ciphertext

	A	B	C	D	E	F	G	H	I	J	K	L	M
O	1	9	0	3	1	1	1	0	1	4	6	0	1
X	0	7	0	1	1	1	1	0	2	4	6	3	0
P	1	0	5	6	0	0	0	0	0	1	1	2	2

	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
O	2	2	8	0	4	1	0	0	3	0	1	1	2
X	3	1	9	0	2	4	0	3	3	2	0	0	1
P	0	8	0	0	0	0	0	0	11	0	9	9	0

Analysis of the Encrypted Message

- Both `O` and `X` are neighbors with the majority of letters
  - probably vowels `a` and `e` ?
- P avoids being neighbors with 15 letters
  - possibly a consonant `t` ?
- In the ciphertext `OO` appears twice
  - suggests `O = e` and therefore `X = a`

Analysis of the Encrypted Message

- `X` appears on its own in a single letter word
  - confirms `X = a`
- `Y` also appears on its own
  - suggests `Y = i`
- Also in the English language, `h` often appears before `e`, but rarely after `e`
  - suggests `B = h`

Add the lowercase ‘guessed’ letters

PCQ VMJiPD LhiK LiSe KhahJaWaV haV ZCJPe EiPD
KhahJiUaJ LhJee KCPK. CP Lhe LhCMKaPV aPV IiJKL PiDhL,
QheP Khe haV ePVeV Lhe LaRe CI Sa’aJMI, Khe JCKe aPV
EiKKeV Lhe DJCMPV ZeICJe hiS, KaUiPD: ‘DJeaL EiPD, ICJ a
LhCMKaPV CPe PiDhLK i haNe ZeeP JeACMPLiPD LC UCM
Lhe IaZReK CI FaKL aDeK aPV Lhe ReDePVK CI aPAiePL EiPDK.
SaU i SaEe KC ZCRV aK LC AJaNe a IaNCMJ CI UCMJ SaGeKLU?’

eFiRCDMe, LaReK IJCS Lhe LhCMKaPV aPV CPe PiDhLK

Look for common words

- The most common three letter words in the English language are `the` and `and`
- `Lhe` appears six times
  - suggests `L = t`
- `aPV` appears five times
  - suggests `P = n` and `V = d`

Write in the common words

nCQ dMJinD thiK tiSe KhahJaWad had ZCJne EinD KhahJiUaJ thJee KCnK. Cn the thCMKand and IiJKt niDht, Qhen Khe had ended the taRe CI Sa'aJMI, Khe JCKe and EiKKed the DJCMnd ZeICJe hiS, KaUinD: 'DJeat EinD, ICJ a thCMKand Cne niDhtK i haNe Zeen JeACMntinD tC UCM the IaZReK CI FaKt aDeK and the ReDendK CI anAient EinDK. SaU i SaEe KC ZCRd aK tC AJaNe a IaNCMJ CI UCMJ SaGeKtU?'

eFiRCDMe, taReK IJCS the thCMKand and Cne niDhtK

Keep Going

- We can continue the process and in this way end up with the original text

The Resultant Plaintext

Now during this time Shahrazad had borne King Shahriyar three sons. On the thousand and first night, when she had ended the tale of Ma'aruf, she rose and kissed the ground before him, saying: 'Great King, for a thousand and one nights I have been recounting to you the fables of past ages and the legends of ancient kings. May I make so bold as to crave a favour of your majesty?'

Epilogue, Tales from the Thousand and One Nights

One for you to try

- PSIIZWFPCMFSW UWTFWUUAG CAU MQU JUSJDU OQS JASVFYU MQU FWXACGMAZPMZAU XSA PSIIZWFPCMFSW KUMOUUW FWYFVFYZCDG CWY SATCWFGCMFSWG. MAN MS FICTFWU C OSADY FW OQFPQ CDD XSAIG SX PSIIZWFPCMFSW QCVU KUUW AUISVUY; WS MUDUVFGFSW, WS ACYFS, WS MUDUJQSWUG, WS FWMUAWUM, WS GCM-WCV.

- communication engineers are the people who provide the infrastructure for communication between individuals and organisations. try to imagine a world in which all forms of communication have been removed; no television, no radio, no telephones, no internet, no sat-nav.