CSEN1001

Computer and Network Security

Mervat AbuElkheir
Mohamed Abdelrazik
Ahmad Helmy

Lecture (2)

Cryptographic Tools

Early History

- Egypt (Old Kingdom), ca. 1900BC. Use of non-standard hieroglyphs (probably not a serious attempt at secret communication)
- Mesopotamia, ca. 1500BC. Encrypted recipe for pottery glaze on clay tablet
- Hebrew, ca. 500BC. Monoalphabetic substitution cipher used by scholars

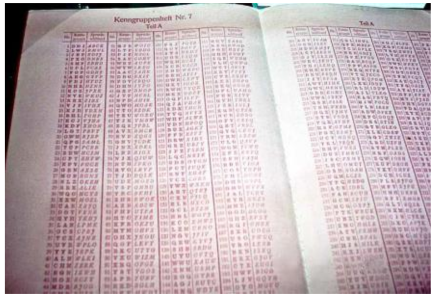
Not so Early History

- 800AD. Early description of frequency analysis for breaking substitution ciphers (credited to Arab mathematician Al-Kindi, 801–873). Works about cryptanalysis of single- and polyalphabetic ciphers
- Ahmad al-Qalqashandi 1355–1418 List of ciphers in his encyclopedia contains a cipher with multiple substitutions for each letter; frequency tables for letters to aid cryptanalysis
- 1467: Leon Battista Alberti ("father of Western cryptology") describes the polyalphabetic cipher

Recent History

- ❑WWII heavy use of rotor machines for complex polyalphabetic substitution ciphers

The “Enigma” machine:



Recent History

- The time of WWII brought massive advances in cryptography as well as cryptanalysis
- In the 20th century, mathematical cryptography was developed
  - Works by Claude Shannon. Any theoretically unbreakable cipher must have keys which are at least as long as the plaintext and used only once (one-time pad)
  - Publication of the Data Encryption Standard in 1970
- 1976 Ground-breaking paper: Whitfield Diffie and Martin Hellman. “New Directions in Cryptography” solves key exchange problem and sparks development of asymmetric key algorithms

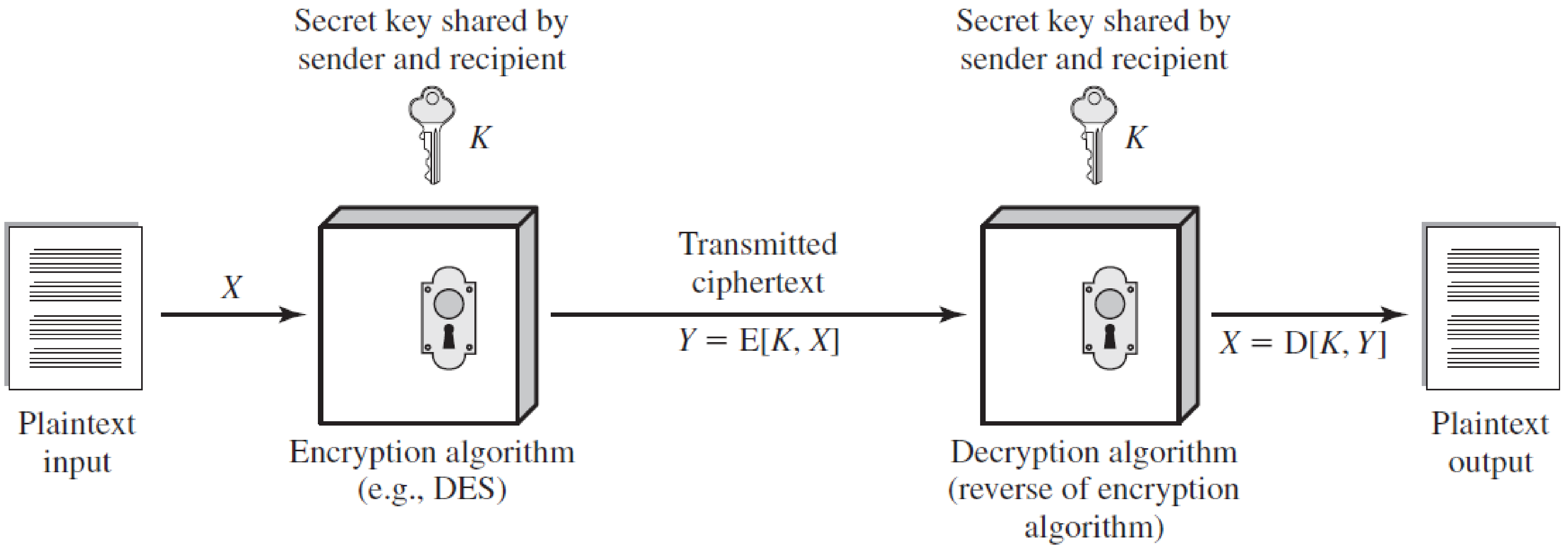
Basic Concepts

- Plaintext → original message
- Ciphertext → coded message
- Cryptography → collection of encryption schemes
- Cipher → an encryption scheme/algorithm
- Key → value used to code a message via an algorithm
- Encryption → process of coding a message
- Decryption → opposite process!
- Cryptanalysis → deciphering a coded message without knowledge of encryption technique/key

Cryptography Elements

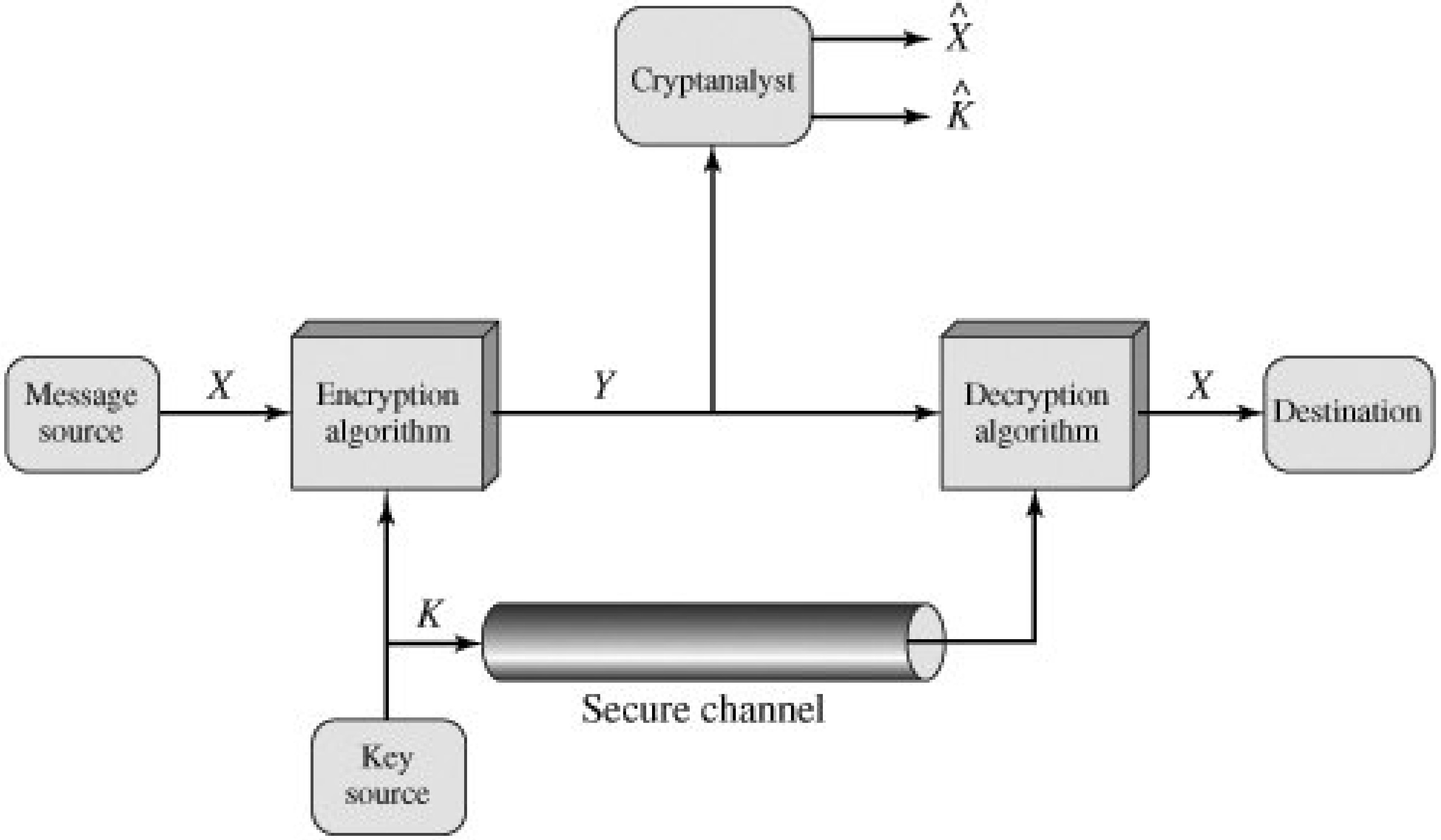
- Transformation
  - substitution, transposition
- # Keys
  - same key (symmetric, single, secret), different keys (asymmetric, two-key, public)
- Plaintext Processing
  - Block cipher, stream cipher

Symmetric Encryption Model



- Conventional / private-key / single-key
- Sender and recipient share a common key
- All classical encryption algorithms are **private-key**
- Was the only type prior to invention of public-key in 1970's, and by far most widely used

Cryptosystem Model



Classical Substitution Ciphers

- Letters of plaintext are replaced by other letters or by numbers or symbols
- Or if plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns

Caesar Cipher

- Earliest known substitution cipher (invented by Julius Caesar)
- First attested use in military affairs
- Replaces each letter by 3rd letter down in the alphabet. Example:

  Meet me after the toga party

  PHHW PH DIWHU WKH WRJD SDUWB

- 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

- If you 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

- Then you have Caesar cipher as:

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

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

Breaking Caesar Cipher

- Only have 26 possible ciphers for each letter
  - A maps to A,B,..Z
- Could simply try each in turn
- a **brute force search**
- Given ciphertext, just try all shifts of letters
- Do need to recognize when have plaintext
- Compression reduces chance of breaking

```
˜+Wµ"— Ω–0)≤4{∞‡¸ ë~Ω%ràu·¯Í ◊¯Z-
Ú≠2Ò#Åæ∂ œ«q7¸Ωn·®3N◊Ú Œz'Y–ƒ∞Í[±Û_ èΩ,<NO¬±«ˇxã  Åäƒèü3Å
x}ö§kºÂ
_yÍ ^ΔÉ] ¸¤ J/˚iTê&ı 'c<uΩ-
 ÄD(G WÄC~y_ïõÄW PÔı«ÎÜ†ç],¤¡ˇÌ^üÑπˇ≈ˇLˇ9Ogfl0ˇ&Œ≤ ¬≤ ØÔ§˝:
 ˇŒ!SGqèvo^ ú\¸S›h<-*6ø‡%x´˝|ﬁÓ#≈˜my‰ˇ≥ñP‹,ƒi Áj  Å◊¿˝Zù-
Ω¨Õ¯6Œÿ{% „ΩÊó ¸ï π÷ÁÎ˚úO2çSÿ´O–
2Äflßi /@^"∏Kºª PŒπ¸úé^´3∑ˇöˇÔZÌ"Y¬ŸΩœY› Ω+eô/˙‹Kƒ¿*÷˜"≤û˜
B ZøK˜Qßÿüƒ¸!ÒflÎzsS/]›ÈQ ü
```

KEY	PHHW PH DIWHU WKH WRJD SDUWB
1	oggv og chvgt vjg vqic rctva
2	nffu nf bgufs uif uphb qbsuz
3	meet me after the toga party
4	ldds ld zesdq sgd snfz ozqsx
5	kccr kc ydrcp rfc rmey nyprw
6	jbbq jb xcqbo qeb qldx mxoqv
7	iaap ia wbpan pda pkcw lwnpu
8	hzzo hz vaozm ocz ojbv kvmot
9	gyyn gy uznyl nby niau julns
10	fxxm fx tymxk max mhzt itkmr
11	ewwl ew sxlwj lzw lgys hsjlq
12	dvvk dv rwkvi kyv kfxr grikp
13	cuuj cu qvjuh jxu jewq fqhjo
14	btti bt puitg iwt idvp epgin
15	assh as othsf hvs hcuo dofhm
16	zrrg zr nsgre gur gbtn cnegl
17	yqqf yq mrfqd ftq fasm bmdfk
18	xppe xp lqepc esp ezrl alcej
19	wood wo kpdob dro dyqk zkbdi
20	vnnc vn jocna cqn cxpj yjach
21	ummb um inbmz bpm bwoi xizbg
22	tlla tl hmaly aol avnh whyaf
23	skkz sk glzkx znk zumg vgxze
24	rjjy rj fkyjw ymj ytlf ufwyd
25	qiix qi ejxiv xli xske tevxc

Monoalphabetic Cipher

- Rather than just shifting the alphabet, shuffle (jumble) the letters **arbitrarily**
- Each plaintext letter maps to a different random ciphertext letter
- Hence key is 26 letters long

```
Plain:  abcdefghijklmnopqrstuvwxyz
Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext:  ifwewishtoreplaceletters
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
```

- Now have a total of $26! = 4 \times 10^{26}$ keys
- With so many keys, might think is secure
- But would be **!!!WRONG!!!**

Breaking the Monoalphabetic Cipher

- Given ciphertext:

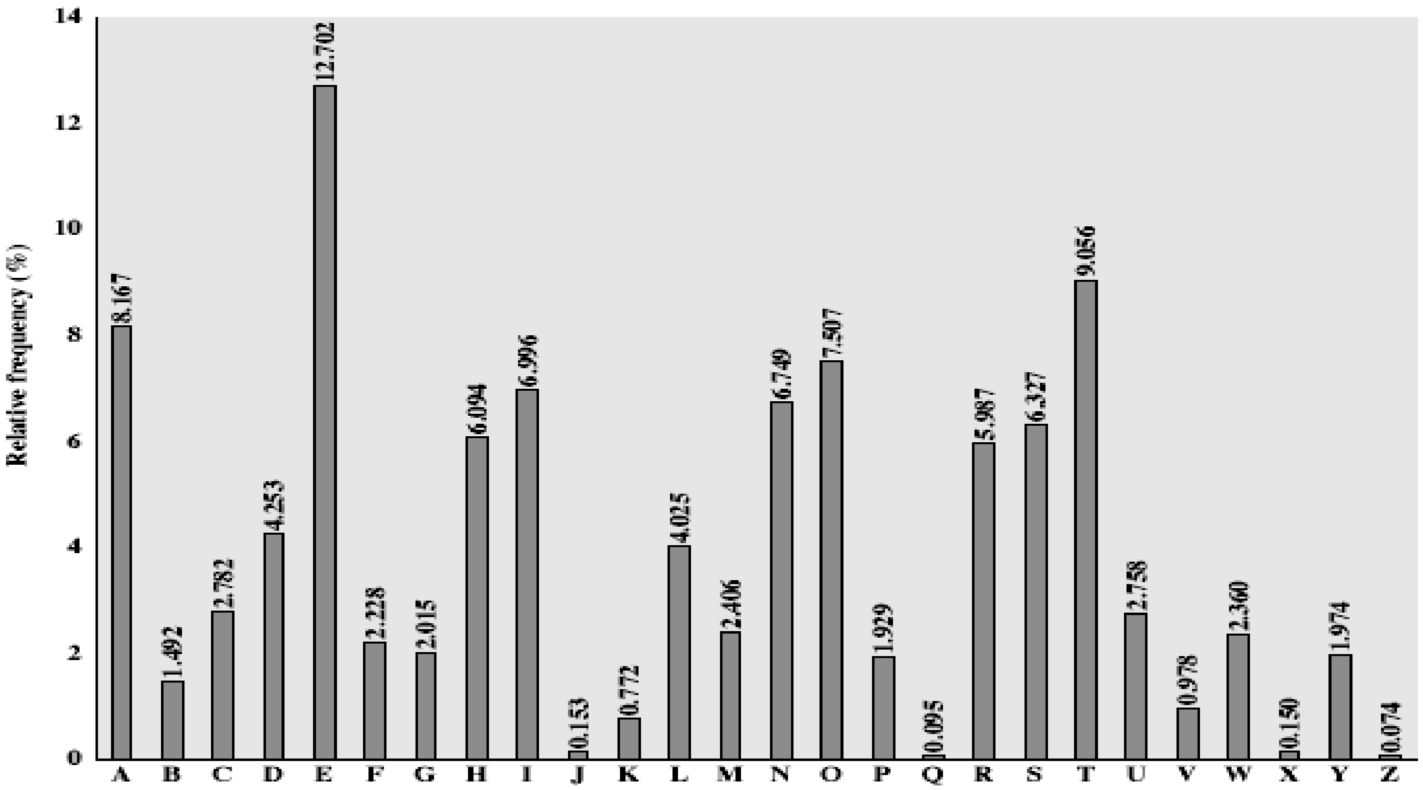
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

Breaking the Monoalphabetic Cipher

- Human languages are **redundant**, e.g. “secrty s awsm"
- letters are not equally commonly used
- In English E is by far the most common letter
  - followed by T,R,N,I,O,A,S
- Other letters like Z,J,K,Q,X are fairly rare
- Have tables of single, double & triple letter frequencies for various languages



Breaking the Monoalphabetic Cipher

- Key concept - monoalphabetic substitution ciphers do not change relative letter frequencies
- Discovered by Arabian scientists in 9^{th} century
- Calculate letter frequencies for ciphertext
- Compare counts/plots against known values
- If Caesar cipher look for common peaks/troughs
  - peaks at: A-E-I triple, NO pair, RST triple
  - troughs at: JK, X-Z
- For monoalphabetic must identify each letter
  - tables of common double/triple letters help

Breaking the Monoalphabetic Cipher

- Given ciphertext:
  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
- Count relative letter frequencies

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				

- Guess P & Z are e and t
- Guess ZW is “th” and hence ZWP is “the”. Frequency of two-letter combinations is known as digrams
- Proceeding with trial and error finally get:
  it was disclosed yesterday that several informal but
  direct contacts have been made with political
  representatives of the viet cong in moscow

Problem 1

The following ciphertext was produced using a Caesar cipher. Break the encryption to obtain the plaintext. Note that the most commonly occurring letter in the English language is “e” and the second most commonly occurring letter is “t”.

MAX YTNEM, WXTK UKNMNL, EBXL GHM BG HNK LMTKL UNM BG HNKLXEOXL

Problem 1

The following ciphertext was produced using a Caesar cipher. Break the encryption to obtain the plaintext. Note that the most commonly occurring letter in the English language is “e” and the second most commonly occurring letter is “t”.

MAX YTNEM, WXTK UKNMNL, EBXL GHM BG HNK LMTKL UNM BG HNKLXEOXL

Solution

By analyzing the ciphertext, the most commonly occurring letter is *M*, followed by *X*. If we assume that *M* is equivalent to *e* and *X* is equivalent to *t* we do not get a meaningful phrase.

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
												e											t		

8 shifts

30 shifts!

MAX YTNEM, WXTK UKNMNL, EBXL GHM BG HNK LMTKL UNM BG HNKLXEOXL

eAt YTNEe, WtTK UKNeNL, EBtL GHe BG HNK LeTKL UNe BG HNKLtEOtL

Problem 1

The following ciphertext was produced using a Caesar cipher. Break the encryption to obtain the plaintext. Note that the most commonly occurring letter in the English language is “e” and the second most commonly occurring letter is “t”.

MAX YTNEM, WXTK UKNMNL, EBXL GHM BG HNK LMTKL UNM BG HNKLXEOXL

Solution

By analyzing the ciphertext, the most commonly occurring letter is *M*, followed by *X*. If we assume that *M* is equivalent to *e* and *X* is equivalent to *t* we do not get a meaningful phrase.

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
												t											e		

19 shifts 19 shifts

MAX YTNEM, WXTK UKNMNL, EBXL GHM BG HNK LMTKL UNM BG HNKLXEOXL

tAe YTNEt, WeTK UKNtNL, EBeL GHt BG HNK LtTKL UNt BG HNKLeEOeL

Problem 1

The following ciphertext was produced using a Caesar cipher. Break the encryption to obtain the plaintext. Note that the most commonly occurring letter in the English language is "e" and the second most commonly occurring letter is "t".

MAX YTNEM, WXTK UKNMNL, EBXL GHM BG HNK LMTKL UNM BG HNKLXEOXL

Solution

By analyzing the ciphertext, the most commonly occurring letter is *M*, followed by *X*. If we assume that *M* is equivalent to *e* and *X* is equivalent to *t* we do not get a meaningful phrase. Thus, we can estimate that *M* = *t* and *X* = *e*. Analysis confirms that this is a shift of 19 letters in both substitutions. Using this same shift for all letters gives the 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
h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g

19 shifts 19 shifts

MAX YTNEM, WXTK UKNMNL, EBXL GHM BG HNK LMTKL UNM BG HNKLXEOXL

the faNEt, WeTK UKNtNL, EBeL GHt BG HNK LtTKL UNt BG HNKLeEOeL

The fault, dear Brutus lies not in our stars but in ourselves

Polyalphabetic Cipher

- Improve security using multiple cipher alphabets
- Make cryptanalysis harder with more alphabets to guess and flatter frequency distribution
- Use a key to select which alphabet is used for each letter of the message
- Use each alphabet in turn
- Repeat from start after end of key is reached

Vigenère Cipher

- Simplest polyalphabetic substitution cipher
- Effectively multiple Caesar ciphers
- Key is multiple letters long $K = k_1\ k_2\ \ldots\ k_d$, i^{th} letter specifies i^{th} alphabet to use
- Use each alphabet in turn
- Repeat from start after d letters in message
- Decryption simply works in reverse
- Write the plaintext out. Write the keyword repeated above it
- Use each key letter as a Caesar cipher key. Encrypt the corresponding plaintext letter
- e.g. using keyword *deceptive*
  - key:
  - plaintext:
  - ciphertext:

key:	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e
plaintext:	w	e	a	r	e	d	i	s	c	o	v	e	r	e	d	s	a	v	e	y	o	u	r	s	e	l	f
ciphertext:																											

Plaintext

Key

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

Vigenère Tableau (Tabula Recta)

Vigenère Cipher

- Simplest polyalphabetic substitution cipher
- Effectively multiple Caesar ciphers
- Key is multiple letters long $K = k_1\ k_2\ \ldots\ k_d$, i^{th} letter specifies i^{th} alphabet to use
- Use each alphabet in turn
- Repeat from start after d letters in message
- Decryption simply works in reverse
- Write the plaintext out. Write the keyword repeated above it
- Use each key letter as a Caesar cipher key. Encrypt the corresponding plaintext letter
- e.g. using keyword *deceptive*

key:	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e
plaintext:	w	e	a	r	e	d	i	s	c	o	v	e	r	e	d	s	a	v	e	y	o	u	r	s	e	l	f
ciphertext:	Z	I	C	V	T	W	Q	N	G	R	Z	G	V	T	W	A	V	Z	H	C	Q	Y	G	L	M	G	J

Breaking the Vigenère Cipher

- Have multiple ciphertext letters for each plaintext letter
- Hence letter frequencies are obscured, but not totally lost
- Start with letter frequencies
  - see if look like monoalphabetic or not
- If not, then need to determine number of alphabets, since then can attack each
- Kasiski method developed by Babbage/Kasiski is a way of breaking Vigenère
- Repetitions in ciphertext give clues to period, so find same plaintext an exact period apart which results in the same ciphertext
- Of course, could also be random fluke
- e.g. repeated “VTW” in previous example suggests size of 3 or 9
- Then attack each monoalphabetic cipher individually using same techniques as before

Kasiski Method

abcdefabcdefabcdefab**cdefab**cdefabc (key is 6 letters long)

crypto is short for cryptography

abcdeabcdeabcdeabcde**abcdea**bcdeabc (key is 5 letters long)

crypto is short for cryptography

1. Compute distances between all repeated groups in ciphertext.
2. Factor each of the distances. If any number is repeated in majority of factorizations, it is likely length of key.
3. If the keyword is N letters long, then every Nth letter must have been enciphered using the same letter of the keytext. Group every N letters of ciphertext together, and perform frequency analysis on one block of N letters.
4. Infer key from one block, and use to break other blocks of ciphertext.

One-time Pad

- A random key as long as the message is used to encrypt and decrypt a single message
- The key is then discarded, never to be used again
- The output bears no statistical relationship to the plaintext
- 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 which was the intended plaintext

```
ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
key:        pxlmvmsydofuyrvzwc tnlebnecvgdupahfzzlmnyih
plaintext:  mr mustard with the candlestick in the hall
ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
key:        mfugpmiydgaxgoufhklllmhsqdqogtewbqfgyovuhwt
plaintext:  miss scarlet with the knife in the library
```

- Large number of keys need to be used. Key distribution is a big problem

Transposition Ciphers

- Now consider classical **transposition** or **permutation** ciphers
- These hide the message by rearranging the letter order without altering the actual letters used
- Can recognise these since have the same frequency distribution as the original text
- Simplest transposition cipher is the **Rail Fence Cipher**
- Write message letters out vertically over two rows
- Then read off cipher row by row
- e.g. 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 Ciphers

- A more complex transposition
- Write letters of message out in rows over a specified number of columns, then reorder the columns according to some key before reading off the rows

Key:
Plaintext:

4	3	1	2	5	6	7
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

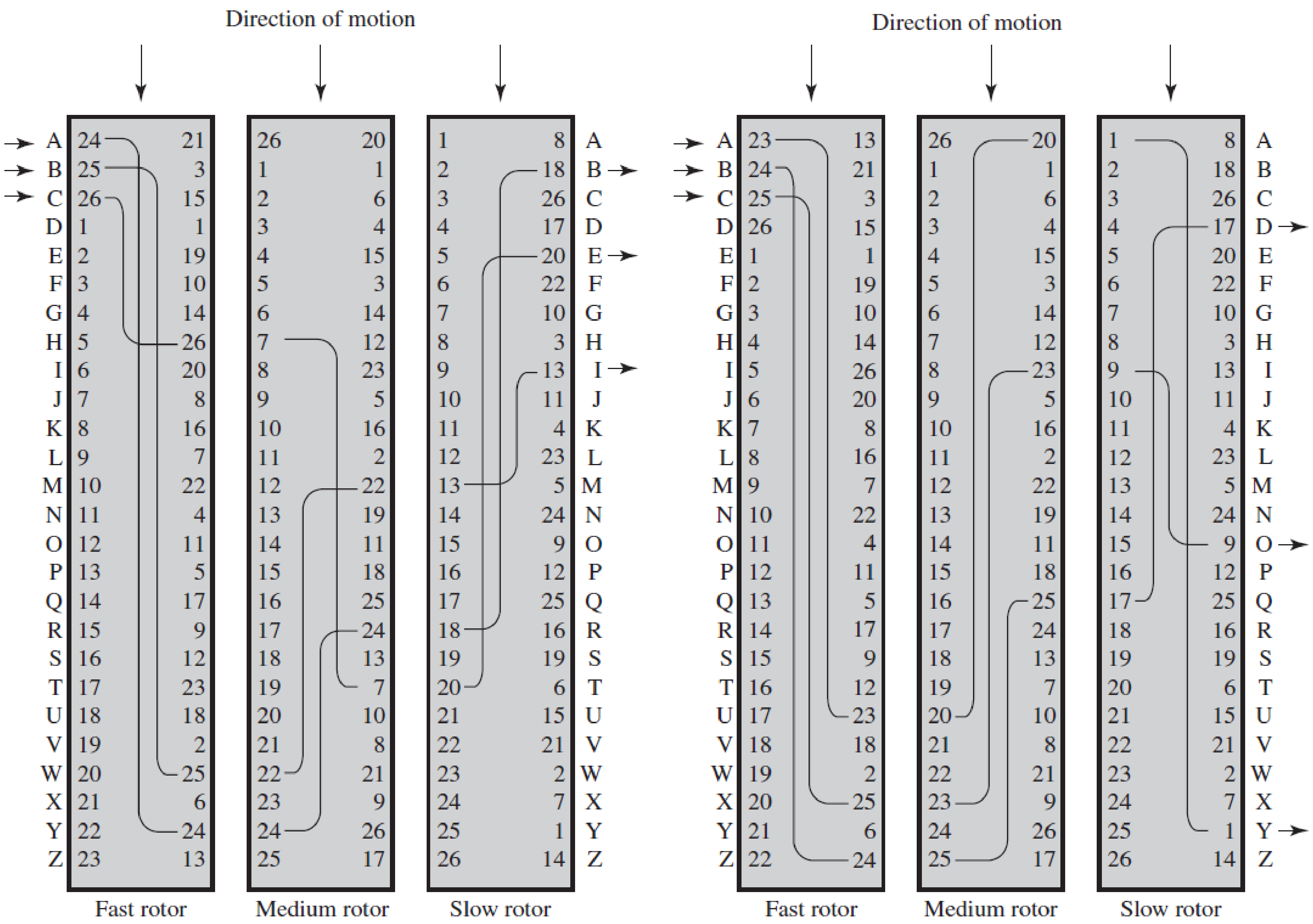
- Ciphers using substitutions or transpositions are not secure because of language characteristics
- Hence consider using several ciphers in succession to make harder, but:
  - two substitutions make a more complex substitution
  - two transpositions make more complex transposition
  - a substitution followed by a transposition makes a new much harder cipher

Rotor Machines

- Before modern ciphers, rotor machines were most common complex ciphers in use
- Widely used in WW2
  - German Enigma, Allied Hagelin, Japanese Purple
- Implemented a very complex, varying substitution cipher
- Used a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- With 3 cylinders have 26^3=17576 alphabets

- For a demo of the Enigma machine, check https://www.youtube.com/watch?v=mcX7iO_XCFA&t=401s

Rotor Machines



(a) Initial setting

(b) Setting after one keystroke

Note

The material in this lecture can be found in Chapter 2, Cryptography and network security.

Next Lecture

Block Ciphers and AES