Common English Digrams and Trigrams

Digrams	Trigrams
EN	ENT
RE	ION
ER	AND
NT	ING
TH	IVE
ON	TIO
IN	FOR
TF	OUR
AN	THI
OR	ONE

monoalphabetic cryptanalysis

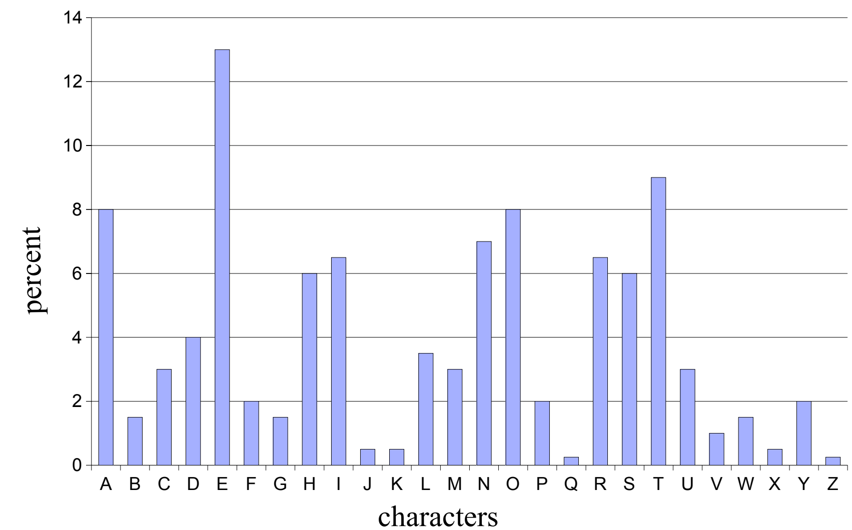
- See class example

Security in Computing

Chapter 2

Elementary Cryptography (part 2)

Character Frequencies (English)



Vigenere Cipher

- construct a table (a *Vigenere tableau*)
- each row in table is a different shift (alphabet)
- sender and receiver agree on sequence of rows
- helps to disguise patterns

Vigenere Tableau

0	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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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

cryptographer's counter-move

- cryptanalysts use properties of plaintext
- what can be cryptographers' counter-moves?

Hiding Patterns

- polyalphabetic ciphers
  - use multiple alphabets
- homophonic ciphers
  - multiple possible output characters for an input character
- polygram ciphers
  - encipher groups of letters at once

More on Vigenere Keys

- usually think of choice of rows as a keyword
- example: keyword "BASE"
- row order is b,a,s,e,b,a,s,e, ...

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
s	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

Don't Need to Construct the Table

- keyword = BRAKE
  - equivalent K={1, 17, 0, 10, 4}
- plaintext "I am sick of school"
- convert to numeric vals
- add to K mod 26

	I	A	M	S	I	C	K	O	F	S	C	H	O	O	L
	8	0	12	18	8	2	10	14	5	18	2	7	14	14	11
	B	R	A	K	E	B	R	A	K	E	B	R	A	K	E
+	1	17	0	10	4	1	17	0	10	4	1	17	0	10	4
	9	17	12	2	12	3	1	14	15	22	3	24	14	24	15

Vigenere Example

- suppose we agree on the key{1,5,9,16,21,22}
- encrypt:
  - char 1 with row 1
  - ' 2 ' ' 5
  - ' 3 ' ' 9
  - ...
  - char 6 with row 22
  - char 7 with row 1
  - char 8 with row 5
  - etc.

0	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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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

Vigenere Example

M = "*Chappelle*"

K = 1,5,9,16,21,22

$E_K(M)$ =
DMJFKAMQN

- *note:* our key is in ascending order. this isn't required

0	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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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

Homophonic Example

Plaintext	Homophones					
A	624	18	329	19	4	
B	5	333	511			
C	919	14	67	83		
D	414	309	238	71	15	6
E	8	13	12			
F	61	422				
G	413	2	16			
…	…	…	…	…	…	…

- So "*cabbage*" could be encrypted as:
  - 14 329 511 5 624 2 8

Homophonic Ciphers (cont'd)

- Are there disadvantages to this?

Don't Need to Construct the Table

- keyword = BRAKE
  - equivalent K={1, 17, 0, 10, 4}
- plaintext "I am sick of school"
- convert to numeric vals
- add to K mod 26

Encrypts to:
J RM CMDB OP WDYOYP

	I	A	M	S	I	C	K	O	F	S	C	H	O	O	L
	8	0	12	18	8	2	10	14	5	18	2	7	14	14	11
	B	R	A	K	E	B	R	A	K	E	B	R	A	K	E
+	1	17	0	10	4	1	17	0	10	4	1	17	0	10	4
	9	17	12	2	12	3	1	14	15	22	3	24	14	24	15

Homophonic Ciphers

- Try to hide plaintext patterns
- Map each plaintext character m to any of a set of ciphertext characters
- set of possible ciphertext characters that map to a single plaintext character m called **homophones**

Playfair Cipher

- 1850s. named after Playfair
- actually invented by his friend Wheatstone
- write keyword without dups. into 5x5 matrix
- treat *I* and *J* as the same character
- example:
  - keyword "MACARONI"

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

Playfair Encryption

- divide plaintext into pairs
- double characters separated by dummy character (*e.g.* x)
  - mi ss is si pp i becomes
  - mi sx si sx si px pi
- if plaintext has odd number of chars, append dummy char.
- encrypt plaintext pairs
- only 3 possibilities
  - same row
  - same column
  - different row and col.

Homophonic Ciphers (cont'd)

- Are there disadvantages to this?
  - ciphertext longer than the plaintext
- How many homophones per plaintext char?
  - fixed number
  - variable: more for frequent plaintext characters

Polygram Ciphers

- simple substitution ciphers, *e.g.* shift ciphers, keyword mixed alphabet, (even Vigenere tableau) ... substitute one character for another character
- **polygram ciphers** substitute a group of characters for another group of characters
- goal: make frequency analysis more difficult
- examples:
  - playfair cipher
  - hill cipher

One time pad

- Idea:
  - Take a stream of random data (keystream)
    - used to be physically on a pad.
    - rip out as many random pages as you need.
  - Combine it with plaintext to form ciphertext
- Message receiver uses same keystream to recover plaintext
- If the stream is truly random → perfect security
- Why don't we use this all the time?

One time pad

- Idea:
  - Take a stream of random data (keystream)
  - Combine it with plaintext to form ciphertext
- Message receiver uses same keystream to recover plaintext
- If the stream is truly random → perfect security
- Why don't we use this all the time?
  - How do we get unlimited truly random stream?
  - If we could get it, how do we distribute it?
  - What if sender and receiver aren't synchronized?

Playfair encryption (cont'd)

- same row
  - substitute with chars to right
  - examples: *MC→AR, RM→OA,SU→TP*
- same col
  - substitute with chars below
  - examples: *EU→LZ, GW→QA*
- different row and col → *tricky*

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

Playfair – different row and col

- substitute plaintext letter with letter that
  - is in its own row
  - and is in the column of the other plaintext letter
- example, AT becomes RQ

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

M	A	C	R	O
			D	
			K	
			T	
			Y	

	A			
	I/J			
	G			
P	Q	S	T	U
	W			

Reusing the stream

- So why can't we just reuse the DVD?
  - It's very insecure: more on this later

Combining Plaintext with Keystream

- Can do it a different ways:
  - XOR
  - If text, can add to key (mod 26)

Quick Quiz

- I have:
  - DVD (≈ 5 GBytes) of random data
  - a 1.5 Mbps DSL
- If I copy the DVD and give it to a friend, how long can we use it as a one-time pad?

Quick Quiz

- I have:
  - DVD (≈ 5 GBytes) of random data
  - a 1.5 Mbps DSL
- If I copy the DVD and give it to a friend, how long can we use it as a one-time pad?

$$\frac{(5\, GBytes * 1024\, MBytes / GByte * 8\, bits / byte)}{(1.5\, Mbps)} = 27{,}306\, secs. = about\ 7.5\, hours$$

XOR (cont'd)

$$a \oplus a = ?$$

$$a \oplus b \oplus b = ?$$

XOR (cont'd)

$$a \oplus a = 0$$

$$a \oplus b \oplus b = a$$

XOR is an involution

XOR

- *Review:* XOR logic operator

A	B	A XOR B
FALSE	FALSE	FALSE
FALSE	TRUE	TRUE
TRUE	FALSE	TRUE
TRUE	TRUE	FALSE

XOR

- XOR often denoted $\oplus$
- Don't have to write the words TRUE, FALSE
- The following are equivalent:

A	B	$A \oplus B$
FALSE	FALSE	FALSE
FALSE	TRUE	TRUE
TRUE	FALSE	TRUE
TRUE	TRUE	FALSE

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

XOR Example

- can encrypt by XORing plaintext with keystream
- Example: plaintext = “Chappelle”

c	C	h	a	p	p	e	l	l	e
c(bin)	01000011	01101000	01100001	01110000	01110000	01100101	01101100	01101100	01100101
key	33	72	31	79	82	74	126	89	2
key(bin)	00100001	01001000	00011111	01001111	01010010	01001010	01111110	01011001	00000010
c XOR k	01100010	00100000	01111110	00111111	00100010	00101111	00010010	00110101	01100111

- Question: if I have the keystream, how do I decrypt?
  - XOR it the keystream with the ciphertext.

XOR in Java

- a XOR b in Java is *a ^ b*
  - same in C, C++, Perl
- Code that produced the example:

```
System.out.println("c\tc(bin)\tkey\tkey(bin)\tc XOR k");
for (int i=0; i<plaintext.length(); i++)
    System.out.println(plaintext.charAt(i) + "\t" +
        Integer.toBinaryString((int)plaintext.charAt(i))
        + "\t" + key[i] + "\t" + Integer.toBinaryString(key[i])+ "\t"
        + Integer.toBinaryString((int)plaintext.charAt(i) ^ key[i]));
```


Aside: Involutions

Let:

- S be a finite set
- f a bijection (1→1, onto) from S to S (*i.e.* $f: S \rightarrow S$)

- f is an involution if $f = f^{-1}$
  - *i.e.* $f(f(x)) = x$
- So XOR is an involution

XOR Example

- can encrypt by XORing plaintext with keystream
- Example: plaintext = “Chappelle”

c	C	h	a	p	p	e	l	l	e
c(bin)	01000011	01101000	01100001	01110000	01110000	01100101	01101100	01101100	01100101
key	33	72	31	79	82	74	126	89	2
key(bin)	00100001	01001000	00011111	01001111	01010010	01001010	01111110	01011001	00000010
c XOR k	01100010	00100000	01111110	00111111	00100010	00101111	00010010	00110101	01100111

- Question: if I have the keystream, how do I decrypt?

"Book" Ciphers

- construct a poor man's one-time pad
- get "randomness" from:
  - novels
  - newspapers
  - telephone books
  - pieces of music
  - decks of cards

Key Reuse

- What happens if you use the same key twice?

$$C_1 = P_1 \oplus K \qquad C_2 = P_2 \oplus K$$

Combining Plaintext with Keystream

- Besides XOR, for text, you can add key to data mod 26
- Example:

Plaintext	D	A	V	E	A	T	T	E	L
	3	0	21	4	0	19	19	4	11
Key	J	O	M	P	K	R	L	Q	E
	9	14	12	15	10	17	11	16	4
P + K (mod 26)	12	14	7	19	10	10	4	20	15
ciphertext	M	O	H	T	K	K	E	U	P

Vernam Cipher

- type of one-time pad
- combine an arbitrarily long nonrepeating series of numbers with the plaintext to form ciphertext
- originally implemented as a paper tape attached to a teletype machine

5 rotor machine

- For a 5 rotor machine, 26^5 substitution alphabets before machine repeats
- For a practical break based on letter frequency:
  - "The ciphertext would have to be as long as all the speeches made on the floor of the Senate and House of Representatives in three successive sessions of Congress" -- Kahn, *The Codebreakers*

Enigma Exhibit at NSA

- enigma exhibit at the NSA
  - http://www.nsa.gov/museum/museu00007.cfm
- java enigma simulator
  - http://russells.freeshell.org/enigma/

Key Reuse

- What happens if you use the same key twice?

$$C_1 = P_1 \oplus K \qquad C_2 = P_2 \oplus K$$

$$C_1 \oplus C_2 = P_1 \oplus K \oplus P_2 \oplus K$$

$$\ldots = P_1 \oplus P_2$$

much easier to solve

Rotor Machines

- Implements a kind of Vigenere tableau
- Physically:
  - keypad
  - several rotors
  - keypad wired to a rotor, and rotors wired to each other
- After each key is pressed, at least one rotor spins
- rotors positions don't repeat until $26^{\#rotors}$ keys have been pressed
- effect: $26^{\#rotors}$ substitution alphabets
- WWII examples:
  - Enigma
  - Purple

2.3 Transposition Ciphers

- Rearrange P to get C
- Example:
  - P = BOREDOM
  - C = MOODERB

Columnar Transposition

- Use a two-dimensional array (*matrix*)
- P = “NARCOLEPTIC”

1	**2**	**3**	**4**
N	A	R	C
O	L	E	P
T	I	C	

- C formed by reading down columns
  - “NOTALIRECCP”
- Can also reorder columns
  - 2, 1, 3, 4 becomes “ALINOTRECCP”

Rotor Machines. Why?

- Why mention rotor machines?
  - They're not used but lead to DES

Chapter Outline

Breaking a Transposition Cipher

1) Figure out that it's a transposition cipher

- ciphertext chars will have same frequency as plaintext

2) Break the transposition

- use common letter pairs (digrams), triples (trigrams) to figure out *d*

Common English Digrams and Trigrams

Digrams	Trigrams
EN	ENT
RE	ION
ER	AND
NT	ING
TH	IVE
ON	TIO
IN	FOR
TF	OUR
AN	THI
OR	ONE

More General Transposition

- many transpositions use fixed period *d*
- Let Z_d be the integers from 1 to *d*
- Let $f: Z_d \rightarrow Z_d$ be a permutation over Z_d
- Key for the cipher is K=(d,f)
- message $M=m_1, m_2, \ldots, m_d, m_{d+1}, \ldots, m_{2d}, \ldots$
- ciphertext $C=m_{f(1)}, m_{f(2)}, \ldots, m_{f(d)}, m_{d+f(1)}, \ldots, m_{d+f(d)}, \ldots$
- this is easier to see with an example

General Transposition Example

- suppose that the period *d* = 4
- suppose that f is:

i	**1**	**2**	**3**	**4**
f(i)	2	4	1	3

- M = "AGGRAVATION"

M	A G G R	A V A T	I O N
E(M)	G R A G	V T A A	O I N

- short block at the end:
  - chars in C in relative position in permutation
  - *e.g.* 2 is before 3 in *f(i)* so I is before N

Diffusion

- Other than simple permutations, is there anything else that we can do to provide diffusion?
  - Anything else that we can do to spread the information around, *e.g.*
    - add redundant information
    - steganography

Combinations of Approaches

- If it's not too difficult to break:
  - basic substitutions
  - basic permutations
- Use a combination of the two → **product cipher**
  - composition of functions
  - stronger than the separate parts
- substitution adds **confusion**
- transposition adds **diffusion**

Diffusion

- Other than simple permutations, is there anything else that we can do to provide diffusion?