UNIVERSITY OF WOLLONGONG AUSTRALIA

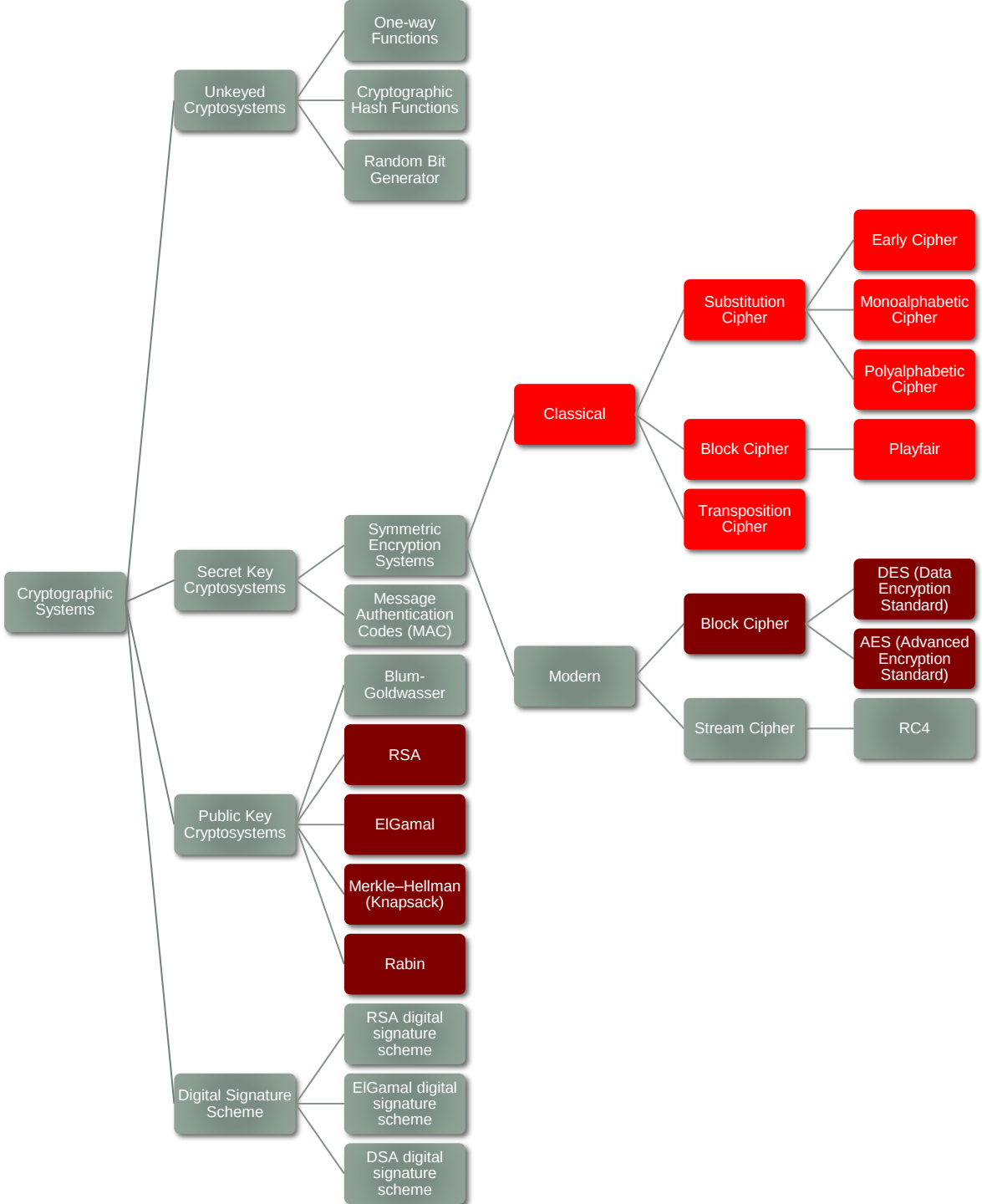
SIM GLOBAL EDUCATION

your attendance ☺ ☺ ☺

TUTORIAL 1

CSCI361 – Computer Security

11 January 2021


Cryptographic Systems
Unkeyed Cryptosystems
One-way Functions
Cryptographic Hash Functions
Random Bit Generator
Secret Key Cryptosystems
Symmetric Encryption Systems
Classical
Substitution Cipher
Early Cipher
Monoalphabetic Cipher
Polyalphabetic Cipher
Block Cipher
Playfair
Transposition Cipher
Modern
Block Cipher
DES (Data Encryption Standard)
AES (Advanced Encryption Standard)
Stream Cipher
RC4
Message Authentication Codes (MAC)
Public Key Cryptosystems
Blum-Goldwasser
RSA
ElGamal
Merkle–Hellman (Knapsack)
Rabin
Digital Signature Scheme
RSA digital signature scheme
ElGamal digital signature scheme
DSA digital signature scheme

Cryptographic system

- A cryptographic system comprises of five components:
  - **Encryption rule** – an algorithm that scramble data to make it unrecognizable.
  - **Decryption rule** – an algorithm that unscrambling data to its original format.
  - **Key** – a finite set of possible keys. A key is a sequence of alpha-numeric characters, produced by an algorithm, that allows us to scramble and unscramble data.
  - **Plaintext** – a finite set of possible unencrypted data; it can be English text, numerical data or anything at all.
  - **Ciphertext** – a finite set of possible data that has been encrypted.

EARLY SUBSTITUTION CIPHER

Early substitution cipher

- It makes use of a simple substitution based on **modular arithmetic**. An example of such cipher is **Caesar Cipher** that used key K = 3. Of course a key K of any other values may be used.
- For example, we would like to encrypt ordinary English text that is made up of 26 alphabets, we would set up a correspondence between alphabetic characters and residues modulo 26 as follows:

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

Early substitution cipher

- Suppose the key for the cipher is K = 24, and the plaintext “a bad day”, to encrypt, we first convert the plaintext to a sequence of integers using the specified correspondence, obtaining the following:

a		**b**	**a**	**d**		**d**	**a**	**y**
0		1	0	3		3	0	24

- Next, we add 24 to each value, reducing each sum modulo 26, and we have

0		**1**	**0**	**3**		**3**	**0**	**24**
24		25	24	1		1	24	22

Early substitution cipher

- Finally, we convert the sequence of integers to alphabetic characters, obtaining the ciphertext “yzybbyw”.
- To decrypt, the recipient will first convert the ciphertext to a sequence of integers, then subtract 24 from each value (reducing modulo 26), and finally convert the sequence of integers to alphabetic characters.

early substitution cipher

Exercise:

Encrypt the following plaintext with key K = 11 using early substitution cipher:

wewillmeetatmidnight

Next decrypt the ciphertext to obtain its plaintext (original text).

Ciphertext →	h	p	h	t	w	w	x	p	p	e	l	e	x	t	o	y	t	r	s	e
Plaintext →	w	e	w	i	l	l	m	e	e	t	a	t	m	i	d	n	i	g	h	t

early substitution cipher

- The problem with early substitution cipher is that it is not secure.
- Early substitution cipher can be **cryptanalyzed** by exhaustive key search. Since there are only 26 possible keys, it is easy to try every possible decryption rule until a “meaningful” plaintext string is obtained.

early substitution cipher

h	p	h	t	w	w	x	p	p	e	l	e	x	t	o	y	t	r	s	e
w	e	w	i	l	l	m	e	e	t	a	t	m	i	d	n	i	g	h	t

h	p	h	t	w	w	x	p	p	e	l	e	x	t	o	y	t	r	s	e
i	q	i	u	x	x	y	q	q	f	m	f	y	u	p	z	u	s	t	f
j	r	j	v	y	y	z	r	r	g	n	g	z	v	q	a	v	t	u	g
k	s	k	w	z	z	a	s	s	h	o	h	a	w	r	b	w	u	v	h
l	t	l	x	a	a	b	t	t	i	p	i	b	x	s	c	x	v	w	i
m	u	m	y	b	b	c	u	u	j	q	j	c	y	t	d	y	w	x	j
n	v	n	z	c	c	d	v	v	k	r	k	d	z	u	e	z	x	y	k
o	w	o	a	d	d	e	w	w	l	s	l	e	a	v	f	a	y	z	l
p	x	p	b	e	e	f	x	x	m	t	m	f	b	w	g	b	z	a	m
q	y	q	c	f	f	g	y	y	n	u	n	g	c	x	h	c	a	b	n
r	z	r	d	g	g	h	z	z	o	v	o	h	d	y	i	d	b	c	o
s	a	s	e	h	h	i	a	a	p	w	p	i	e	z	j	e	c	d	p
t	b	t	f	i	i	j	b	b	q	x	q	j	f	a	k	f	d	e	q
u	c	u	g	j	j	k	c	c	r	y	r	k	g	b	l	g	e	f	r
v	d	v	h	k	k	l	d	d	s	z	s	l	h	c	m	h	f	g	s
w	**e**	**w**	**i**	**l**	**l**	**m**	**e**	**e**	**t**	**a**	**t**	**m**	**i**	**d**	**n**	**i**	**g**	**h**	**t**
x	f	x	j	m	m	n	f	f	u	b	u	n	j	e	o	j	h	i	u
y	g	y	k	n	n	o	g	g	v	c	v	o	k	f	p	k	i	j	v
z	h	z	l	o	o	p	h	h	w	d	w	p	l	g	q	l	j	k	w
a	i	a	m	p	p	q	i	i	x	e	x	q	m	h	r	m	k	l	x
b	j	b	n	q	q	r	j	j	y	f	y	r	n	i	s	n	l	m	y
c	k	c	o	r	r	s	k	k	z	g	z	s	o	j	t	o	m	n	z
d	l	d	p	s	s	t	l	l	a	h	a	t	p	k	u	p	n	o	a
e	m	e	q	t	t	u	m	m	b	i	b	u	q	l	v	q	o	p	b
f	n	f	r	u	u	v	n	n	c	j	c	v	r	m	w	r	p	q	c
g	o	g	s	v	v	w	o	o	d	k	d	w	s	n	x	s	q	r	d

MONOALPHABETIC CIPHER

Monoalphabetic cipher

- This cipher substitutes one (mono) letter or character for another.
- The key of this cipher consists of a permutation of the 26 alphabetic characters (highlighted in green). The number of these permutation is 26!, which is more than 4.0×10^{26}. Thus, an exhaustive key search is infeasible.

For example:

	a	b	c	d	e	…	x	y	z
Key	F	G	N	T	A	…	K	P	L

	a		b	a	d		d	a	y
Plaintext									
Ciphertext	F		G	F	T		T	F	P

Monoalphabetic CIPHER

- Key generation:

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	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
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			
E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z				
F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z					
.	.	.																							
X	Y	Z																							
Y	Z																								
Z																									
26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1

$26! = 4.0 \times 10^{26}$.

How big is 4×10^{26}?

- A typical computer can operate 10^{15} operations in a second.
- To exhaustively (brute-force) find all possible keys, we need

$$\frac{4.0 \times 10^{26}}{10^{15}} \approx 4.0 \times 10^{11} \text{seconds.}$$

$$\approx 6,721,524,352 \text{ minutes}$$

$$\approx 112,025,405.9 \text{ hours}$$

$$\approx 4,667,725.25 \text{ days}$$

$$\approx 12,788.29 \text{ years}$$

Monoalphabetic CIPHER

- Here is an example of a “random” permutation which could comprise an encryption function. (Note: To avoid confusion, plaintext characters are written in lower case and ciphertext characters are written in upper case.)

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

- The decryption function is the inverse permutation. This is formed by writing the second lines first, and then sorting in alphabetical order. The following is obtained:

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	l	r	y	v	o	h	e	z	x	w	p	t	b	g	f	j	q	n	m	u	s	k	a	c	i

Monoalphabetic CIPHER

Exercise:

Use the following keys, decrypt the ciphertext

MGZVYZLGHCMHJMYXSSFMNHAHYCDLMHA

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

Decryption 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
	d	l	r	y	v	o	h	e	z	x	w	p	t	b	g	f	j	q	n	m	u	s	k	a	c	i

M	G	Z	V	Y	Z	L	G	H	C	M	H	J	M	Y	X	S	S	F	M	N	H	A	H	Y	C	D	L	M	H	A
t	h	i	s	c	i	p	h	e	r	t	e	x	t	c	a	n	n	o	t	b	e	d	e	c	r	y	p	t	e	d

Monoalphabetic CIPHER

- One disadvantage of the permutation of key is that it is **difficult** to remember.
- To overcome this problem, techniques such as **indexed subset** of all possible substitutions is used.
- Additive, multiplicative, and keyword or key-phrase ciphers are examples of such indexed construction.

Monoalphabetic CIPHER

Keyword or key-phrase subset index

Exercise:

- The following is one possible permutation of a key that I can remember easily, but may be not you.

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

Can you create another one that you can remember easily?

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

My key is generated using a keyword (key-phrase) SIONGGO starting from the first letter of my last name Japit.

Monoalphabetic CIPHER

- This is another possible permutation of a key created using a keyword JANE (The person who created this key was Jane Ang).

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

- Is this key good? What is the problem with this 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
J	A	N	E	B	C	D	F	G	H	I	K	L	M	O	P	Q	R	S	T	U	V	W	X	Y	Z

The key is weak. Some of the data (about 46% of characters) may not be scramble (hide). For example, the word toys, port, post, zoo, etc. are not scramble.

Monoalphabetic CIPHER

Multiplicative subset index

Consider an alphabet of $n = 26$ letters/symbols (Latin alphabet) as shown here. Does $f(x) = x \times x \; mod \; n$ or $f(x) = x^2 \; mod \; n$ define a cipher?

a	b	c	d	e	f	g	h	i	j
0	1	2	3	4	5	6	7	8	9
k	l	m	n	o	p	q	r	s	t
10	11	12	13	14	15	16	17	18	19
u	v	w	x	y	z				
20	21	22	23	24	25				

Monoalphabetic CIPHER

Does $f(x) = x^2 \ mod \ 26$ define a cipher? **NO!**

X	a	b	c	d	e	f	g	h	i
X	0	1	2	3	4	5	6	7	8
f(x)	0	1	4	9	16	25	10	23	12
Y	A	B	E	J	Q	Z	K	X	M
X	j	k	l	m	n	o	p	q	r
X	9	10	11	12	13	14	15	16	17
f(x)	3	22	17	14	13	14	17	22	3
Y	D	W	R	O	N	O	R	W	D
X	s	t	u	v	w	x	y	z	
X	18	19	20	21	22	23	24	25	
f(x)	12	23	10	25	16	9	4	1	
Y	M	X	K	Z	Q	J	E	B	

Monoalphabetic CIPHER

- This mapping is **not** one-to-one.
  - Consider plaintext "*moon*". The associated ciphertext is "*ooon*", which we cannot uniquely decrypt.

- The mapping is actually symmetric: $(-x)^2=x^2$.
- What about general x^i?
  - i=5 seems okay.
  - i=7 does too.
  - Other powers less than or equal to 10 does not work.

Monoalphabetic CIPHER

Consider an alphabet of n = 29 letters/symbols (Latin plus a coma, full stop, and a space) as shown here.

Does $f(x) = x \cdot x \bmod n$ or $f(x) = x^2 \bmod n$ define a cipher?

a	b	c	d	e	f	g	h	i	j
0	1	2	3	4	5	6	7	8	9
k	l	m	n	o	p	q	r	s	t
10	11	12	13	14	15	16	17	18	19
u	v	w	x	y	z	,	.		
20	21	22	23	24	25	**26**	**27**	**28**	

Monoalphabetic CIPHER

What about n=29, with x^2? No!

X	a	b	c	d	e	f	g	h	i	j
X	0	1	2	3	4	5	6	7	8	9
f(x)	0	1	4	9	16	25	10	20	6	23
Y	A	B	E	J	Q	Z	K	U	G	X
X	k	l	m	n	o	p	q	r	s	t
X	10	11	12	13	14	15	16	17	18	19
f(x)	13	5	28	24	22	22	24	28	5	13
Y	N	F		Y	W	W	Y		F	N
X	u	v	w	x	y	z	,	.		
X	20	21	22	23	24	25	26	27	28	
f(x)	23	6	20	10	25	16	9	4	1	
Y	X	G	U	K	Z	Q	J	E	B	

Monoalphabetic CIPHER

What about n=29, x^3 then? Yes!

X	a	b	c	d	e	f	g	h	i	j
X	0	1	2	3	4	5	6	7	8	9
f(x)	0	1	8	27	6	9	13	24	19	4
Y	**A**	**B**	**I**	**.**	**G**	**J**	**N**	**Y**	**T**	**E**
X	k	l	m	n	o	p	q	r	s	t
X	10	11	12	13	14	15	16	17	18	19
f(x)	14	26	17	22	18	11	7	12	3	15
Y	**O**	**,**	**R**	**W**	**S**	**L**	**H**	**M**	**D**	**P**
X	u	v	w	x	y	z	,	.		
X	20	21	22	23	24	25	26	27	28	
f(x)	25	10	5	16	20	23	2	21	28	
Y	**Z**	**K**	**F**	**Q**	**U**	**X**	**C**	**V**		

Monoalphabetic CIPHER

- Plaintext "*the big mouse*"
  → ciphertext "**PYG BTN RSZDG**"

- Plaintext "*small elephant swimming*"
  → ciphertext "**DRA,, G,GLYAWP DFTRRTWN**"

Affine CIPHERS

- To increase the number of keys we can combine **additive** and **multiplicative** ciphers. This cipher is also known as ***affine*** ciphers, and it is described as:

$$Y = \alpha \otimes \mathrm{X} \oplus Z \ mod\ 26$$

$where\ X, Y\ and\ Z\ \in \{0,1,2, \ldots, 24,25\}$
$and\ \alpha \in \{1,3,5,7,9,11,15,17,19,21,23,25\}$
$X\ is\ a\ plaintext\ character\ and$
$Y\ is\ the\ corresponding\ ciphertext\ character.$

Affine CIPHERS: $Y = \alpha \otimes X \oplus Z \ mod \ 26$

- How good is this cipher? How many unique pair of (α, Z) that can be used such that decryption always exist?

In order to have a unique pair of (α, Z), α must be **relatively prime** to 26. This is to prevent different plaintext from obtaining the same ciphertext through encryption.

How many values of α that is relatively prime to 26?

Using Euler's totient function, we can determine the number of α that is relatively prime to 26; that is, $26 = 2 \times 13 = (2-1) \times (13-1) = 12$.

Affine CIPHERS: $Y = \alpha \otimes X \oplus Z \ mod \ 26$

- So, what are the 12 numbers that are relatively prime to 26?

Numbers that do not have common factor to 26 in Z_{26}^* are relatively prime to 26; they are {1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23 and 25}.

Since the values of X, Y and Z are $\in \{0,1,2, \ldots, 25\}$, there are a total of $12 \times 26 = 312$ unique pair of keys can be used.

312 unique key-pair are still too small!

Affine CIPHERS: $Y = \alpha \otimes X \oplus Z \ mod \ 26$

- The 312 unique key-pair are as follows:
  - $(1,0), (1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (1,7), (1,8), (1,9), (1,10), (1,11), (1,12), (1,13), (1,14), (1,15), (1,16), (1,17), (1,18), (1,19), (1,20), (1,21), (1,22), (1,23), (1,24), (1,25)$
  - $(3,0), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (3,7), (3,8), (3,9), (3,10), (3,11), (3,12), (3,13), (3,14), (3,15), (3,16), (3,17), (3,18), (3,19), (3,20), (3,21), (3,22), (3,23), (3,24), (3,25)$
  - …

Affine CIPHERS: $Y = \alpha \otimes X \oplus Z \ mod \ 26$

- The 312 unique key-pair are as follows:
  - $(25,0), (25,1), (25,2), (25,3), (25,4), (25,5), (25,6), (25,7), (25,8), (25,9), (25,10), (25,11), (25,12), (25,13), (25,14), (25,15), (25,16), (25,17), (25,18), (25,19), (25,20), (25,21), (25,22), (25,23), (25,24), (25,25)$

There are a total of $12 \times 26 = 312$ key-pairs.

What is special about key-pair (1,0), the first key-pair?

Affine CIPHERS: $Y = \alpha \otimes X \oplus Z \ mod \ 26$

- This key cannot be used to do encryption, because 1 (α) multiply any numbers, we get back the same numbers and since Z is 0, the number will not be transformed to any other numbers. In other words the ciphertext will be exactly the same as the plaintext; nothing is being transformed.
- The effective key-pairs will then be only 311.

Affine Ciphers

Alice wishes to send the following message to Bob using the Affine cipher:

I LOVE UOW

Alice and Bob agreed to use a key a = 3 and b = 17. Encrypt the message.

You can assume the Affine cipher used in this encryption uses only 26 alphabetic characters, and you can ignore the spaces.

Affine Ciphers

Convert the plaintext to its numerical equivalent using the following mapping:

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

Affine Cipher

Hence,

I	L	O	V	E	U	O	W
8	11	14	21	4	20	14	22

For each of the integers (8, 11, 14, 21, 4, 20, 14, 22), we compute their ciphertext in numerical form as follows:

$8: \quad c = aX + b\ (mod\ 26) = (3) \times (8) + 17\ (mod\ 26) = \ 41\ mod\ 26 = 15$
$11: \ c = aX + b\ (mod\ 26) = (3) \times (11) + 17\ (mod\ 26) = 50\ mod\ 26 = 24$
$14: \ c = aX + b\ (mod\ 26) = (3) \times (14) + 17\ (mod\ 26) = \ 59\ mod\ 26 = 7$
$21: \ c = aX + b\ (mod\ 26) = (3) \times (21) + 17\ (mod\ 26) = 80\ mod\ 26 = 2$
$4: \quad c = aX + b\ (mod\ 26) = (3) \times (4) + 17\ (mod\ 26) = 29\ mod\ 26 = 3$
$20: \ c = aX + b\ (mod\ 26) = (3) \times (20) + 17\ (mod\ 26) = 77\ mod\ 26 = 25$
$14: \ c = aX + b\ (mod\ 26) = (3) \times (14) + \ 17\ (mod\ 26) = 59\ mod\ 26 = 7$
$22: \ c = aX + b\ (mod\ 26) = (3) \times (22) + 17\ (mod\ 26) = 83\ mod\ 26 = 5$

Next, convert the integers (15, 24, 7, 2, 3, 25, 7, 5) back to letters, and we have:
Hence, the ciphertext is PYHCDZHF.

15	**24**	**7**	**2**	**3**	**25**	**7**	**5**
P	Y	H	C	D	Z	H	F

Statistical cryptanalysis

- In the early substitution, monoalphabetic substitution, and affine ciphers once a key is chosen, each alphabetic character is mapped to a unique alphabetic character. These ciphers are insecure against statistical cryptanalysis.
- Due to the properties of language, relationship found in plaintext is not broken in ciphertext; in other words, the properties of language found in the plaintext can also be found (determined) in the ciphertext.

Statistical cryptanalysis

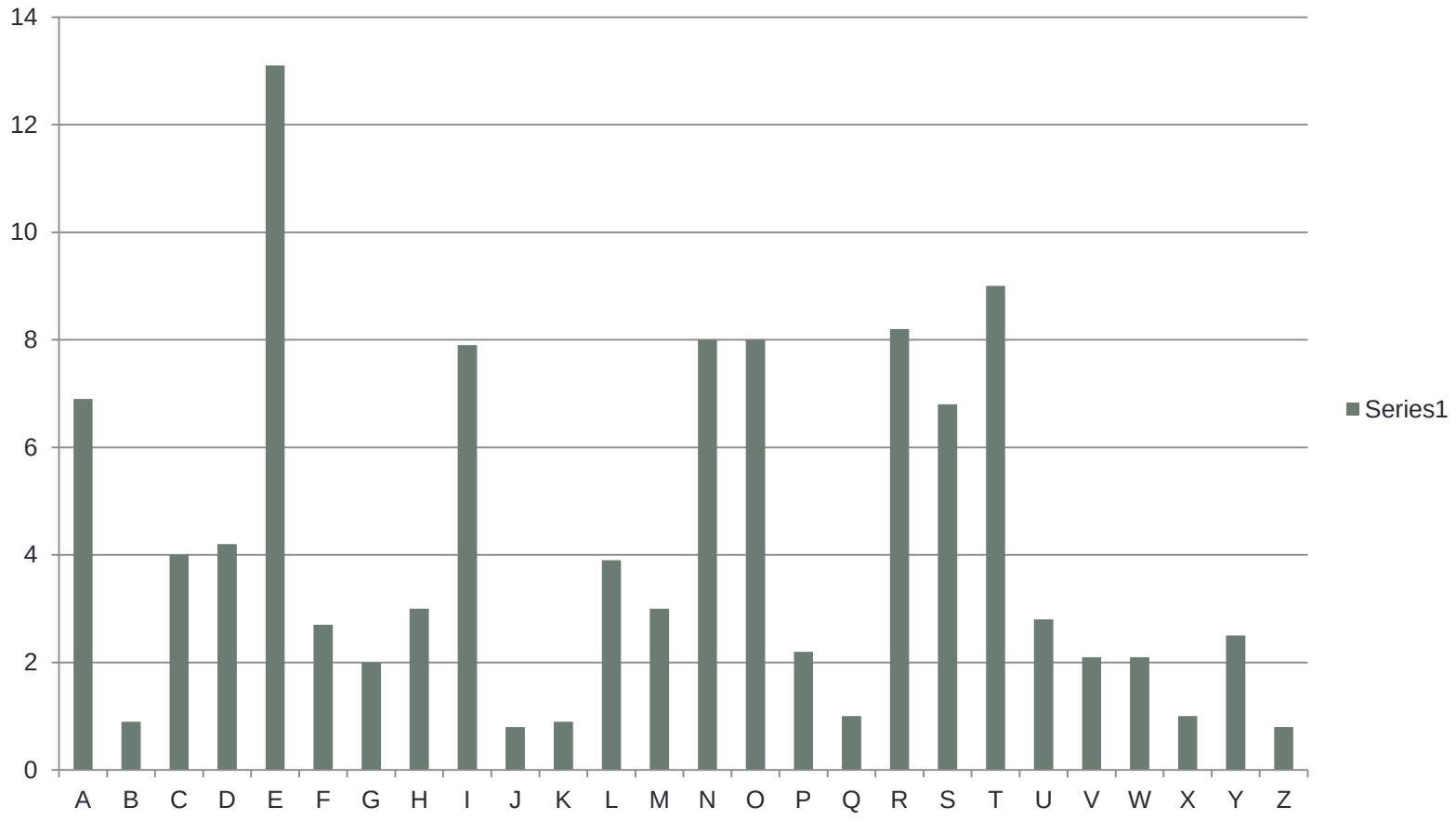
Is the frequency distribution for a language fixed?

No it varies from text to text. Even the frequency order according to large samples of English text differs between studies.

etaoinshrdlucmfwypvbgkqjxz	(Mergenthaler, 1884)
etoanurshdlcfumpywgbvkxjqz	(Valerio, 1893)
etaonisrhldcupfmwybgvkqxjz	(Gaines & Meaker,1939)
etoanirshdlcwumfygpbvkxqjz	(Smith, 1943)
etoanirshdlufcmpywgbvkxzjq	(Sacco, 1951)
etaonirshdlucmpfywgbvjkqxz	(Kahn, 1967)
etaonrishdlfcmugpywbvkxjqz	(Konheim, 1981)
etaoinsrhldcumfpgwybvkxjqz	(Meyer & Matyas, 1982)

(From *Decrypted Secrets: Methods and Maxims of Cryptology,* F.L Bauer)

STATISTICAL CRYPTANALYSIS



English letters frequencies calculated from usual text

Statistical cryptanalysis

Apart from the frequency with which letters appear, what other patterns are useful?

Other than the already mentioned bigrams and trigrams in the lecture notes, the following are some other useful patterns.

- **One-letter English words:** a, i
- **Two-letter English words:** an, at, as, he, be, in, is, it, on, or, to, of, do, go, no, so, my
- **Most frequent English words:** the, of, and, to, a, in, that, it, is, I, for, as, with, was, his, he, be, not, by, but, have, you, which, are, on, or, her
- **Average word length in English =** 4.5 letters
- **Most frequent English bigrams:** th, he, an, in, er, re, on, es, ti, at, st, en, or, nd, to, nt, ed, is, ar
- **Common reversals:** er-re, es-se, an-na, ti-it, on-no, in-ni, en-ne, at-ta, te-et, or-ro, to-ot, ar-ra, st-ts, is-si, ed-de, of-fo

Statistical cryptanalysis

- **Most frequent English trigrams:**

  the, ing, and, ion, tio, ent, ere, her, ate, ver, ter, tha, ati, for, hat, ers, his, res, ill, are, con, nce, all, eve, ith, ted, ain, est, man, red, thi, ive, rea, wit, ons, ess, ave, per, ect, one, und, int, ant, hou, men, was, oun, pro, sta, ine, whi, ove, tin, ast, der, ous, rom, ven, ard, ear, din, sti, not, ort, tho, day, ore, but, out, ure, str, tic, ame, com, our, wer, ome, een, lar, les, san, ste, any, art, nte, rat, tur, ica, ich, nde, pre, enc, has, whe, wil, era, lin, tra.
- **Frequencies of initial letters:**

  t, a, s, o, i, c, w, p, b, f, h, m, r, d, e, n, l, g, u, y, v, j, k, q, x, z
- **Frequencies of final letters**

  e, s, d, n, t, r, y, o, f, l, a, g, h, m, w, k, c, p, i, x, u, b, v, j, z, q

Statistical cryptanalysis

How does not using English effect the index of coincidence?

Arabic	0.075889	Italian	0.073294
Danish	0.070731	Japanese	0.077236
Dutch	0.079805	Malay	0.085286
English	0.066895	Norwegian	0.069428
Finnish	0.073796	Portuguese	0.074528
French	0.074604	Russian	0.056074
German	0.076667	Spanish	0.076613
Greek	0.069165	Swedish	0.064489
Hebrew	0.076844	Random	0.038461

Statistical cryptanalysis

Let’s do some cryptalanalysis using the “Krypto”.

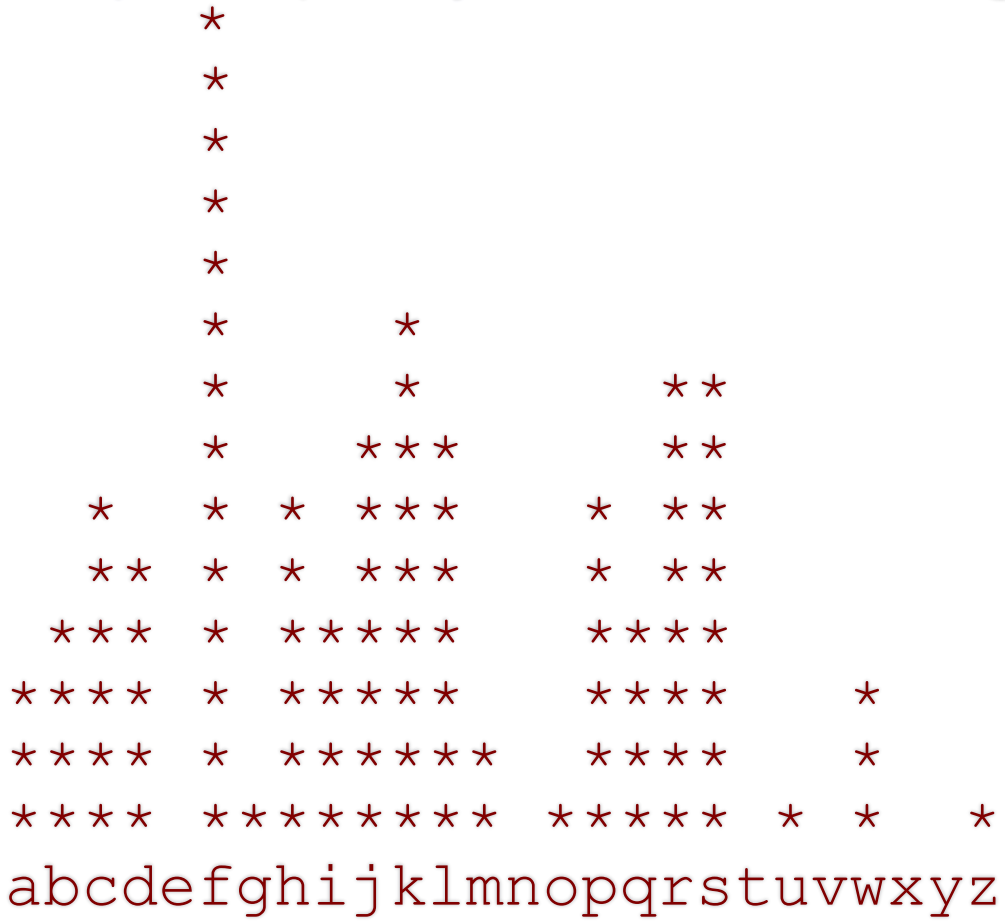
Statistical cryptanalysis

- Load data file **Data1**: **-> r Data1**
- Display code: **-> p**

  rcf odb ikuqf ljs rcf qilhh fhfmcljr wfjr qwdiidjb dj rcf mkkh
  rcf odb ikuqf ljs rcf qilhh fhfmcljr wfjr qwdiidjb dj rcf mkkh

  larfpwlpsq rcf pfs skb ljs rcf bpffj hdzlps hkkgfs akp akks
  larfpwlpsq rcf pfs skb ljs rcf bpffj hdzlps hkkgfs akp akks

- Display modified only: **-> l**
- Reset code using: **-> z**
- Frequency distributions: **-> f [num]**
- Graph frequency distribution: **-> g**




```
-> f 2
cf      5
rc      5
jr      3
kk      3
lj      3
ak      2
dj      2
fj      2
fm      2
fp      2
fs      2
hf      2
js      2
lp      2
pf      2
ps      2
sr      2
ar      1
bd      1
bi      1
bl      1
bp      1
```

- Use frequency analysis of the English language. Consider the most common letters and combinations of letters.
- The most common letters are **e, t, a, o, n, th** is a common pair and **the** is a common triplet.
- Replace symbols in the code with other symbols.
  - For example, ciphertext “f” with plaintext “e” using

->s f e

-> l

rce odb ikuqe ljs rce qilhh ehemcljr wejr qwdiidjb dj rce mkkh

larepwlpsq rce pes skb ljs rce bpeej hdzlps hkkges akp akks

- Continue in this way…

- **The plaintext…**

the big mouse and the small elephant went swimming in the pool

afterwards the red dog and the green lizard looked for food

- **The key…**

We can produce a table of the mapping.

a	b	c	d	e	f	g	h	i	j
L	**O**		**S**	**F**	A	B	C	D	
k	l	m	n	o	p	q	r	s	t
G	H	I	J	K	M		P	Q	R
u	v	w	x	y	z				
U		W							

- There are gaps because not every letter appears in the plaintext.
- The “keyword” is LOTSOF. Duplicate letters are removed, and the remaining alphabet fills the space.
- You may not be able to identify the keyword, but the substitution alphabet is the more important thing.

Another example

-> r Ctext-1

-> l

tgg ts mkcn iqmj sln htqf gbkn mi sln lmqbzmk wlbslnq bs qnsbqnh
sm atbk bsr jmjnksuj sln jmkrsnq qurlnh ruhhnkgy smwtqhr sln teqtltj
gbkcmgk wbsl tgtqjbka qtobhbsy rsmoonh ruhhnkgy temus swnksy inns
iqmj sln lugg tkh hbnh mus kms hbvbka ukhnq sln wtsnq imq bsr
eqbggbtkcy hbh kms tetsn eus ruhhnkgy tkh tr bi sln rmuqcn mi slbr
eqbggbtks njtktsbmk wtr nxltursnh slnk bs qntoontqnh mk sln mslnq
rbhn mi sln vnrrng tr bi bs lth suqknh tkh rgbh ukhnq sln lugg
tky jmjnks t cmggbrbmk jbals ltvn mccuqqnh wlbcl wmugh ltvn ennk

Another example

Determine the index of coincidence

-> I

IC = 0.065

Average = 0.065

- The index for the English language is about 0.065, while random is about 0.038. This suggests we have a mono-alphabetic substitution (which we know we do).

```
-> f 1
n    49
s    44
b    33
t    33
k    32
h    30
l    28
m    28
q    26
g    22
r    20
u    17
j    12
i     9
c     8
w     8
e     7
y     7
o     5
a     4
v     4
f     1
```

sl	17	sln	13	hhnk	3
ln	14	isl	4	hnkg	3
nh	9	bsl	3	isln	3
bs	8	ggb	3	jsln	3
nk	8	ggt	3	misl	3
tk	7	hhn	3	nkgy	3
gb	6	hnk	3	nmis	3
gg	6	jsl	3	ruhh	3
hn	6	kgy	3		
lt	6	lnl	3		

- Examine the graph and frequency distributions.
- Identify the most frequent letter with “e”: Here **n** or **s**.
- Identify the most frequent trigram with “the”. Here “sln” is by far the most frequent trigram.
  - Note that the trigram supports the identification of “e” with “n”.
- This suggests three letters straight away. We substitute them using **-> s [ch1] [ch2]**.

```
-> s n e
-> s s t
-> s l h
```

What now?

-> p
tgg **tt** mkce iqmj the htqf gbke mi the hmqbzmk whbtheq **bt** qetbqeh
tgg ts mkcn iqmj sln htqf gbkn mi sln lmqbzmk wlbslnq bs qnsbqnh

tm atbk btr jmjektuj the jmkrteq qurheh ruhhekgy tmwtqhr the teqthtj
sm atbk bsr jmjnksuj sln jmkrsnq qurlnh ruhhnkgy smwtqhr sln teqtltj

gbkcmgk wbth tgtqjbka qtobhbty rtmooeh ruhhekgy temut twekty ieet
gbkcmgk wbsl tgtqjbka qtobhbsy rsmoonh ruhhnkgy temus swnksy inns

iqmj the hugg tkh hbeh mut kmt hbvbka ukheq the wtteq imq btr
iqmj sln lugg tkh hbnh mus kms hbvbka ukhnq sln wtsnq imq bsr

eqbggbtkcy hbh kmt tette eut ruhhekgy tkh tr **bi** the rmuqce mi thbr
eqbggbtkcy hbh kms tetsn eus ruhhnkgy tkh tr bi sln rmuqcn mi slbr

eqbggbtkt ejtkttbmk wtr exhturteh thek **bt** qetooetqeh mk the mtheq
eqbggbtks njtktsbmk wtr nxltursnh slnk bs qntoontqnh mk sln mslnq

rbhe mi the verreg tr **bi bt** hth tuqkeh tkh rgbh ukheq the hugg
rbhn mi sln vnrrng tr bi bs lth suqknh tkh rgbh ukhnq sln lugg

tky jmjekt t cmggbrbmk jbaht htve mccuqqeh whbch wmugh htve eeek
tky jmjnks t cmggbrbmk jbals ltvn mccuqqnh wlbcl wmugh ltvn ennk

- Returning to the distribution we see “b” and “t” were the most frequently occurring letters after “s” and “n”, which we have just assigned.
- After “e” and “t” the most frequently occurring English letters are “a” and “o”.
- We note that with the substitutions we have made there are words “b**t**” and “t**t**”, where the bold letters are in plaintext, and the ordinary letters are not.
- The most common “?t” words are “at” and “it”; so “b” and “t” probably correspond to “**a**” and “**i**”, in some order. We note also that the letter “t” occurs by itself in the ciphertext (on the last line) supporting the hypothesis that it is probably “**a**” or “**i**”. Remember that it *might* be an initial though.
- We could guess which matches, or be more careful and examine where else “b” and “t” occur.
- If you have words with a few known (plaintext) letters and you aren’t sure what words fit the pattern you can use a crossword solving websites: such as http://www.oneacross.com/
- If you know how, you can also grep for words matching the pattern in **/usr/dict/words**.

- We determine “t” probably corresponds to “**a**”.

-> s t a

-> s b i

Observe the word “**it**r”. This suggests “r” is “**s**” so …

-> s r s

- Analysis continues and we build up a table of results.
- Keeping a table as you go helps.

a	b	c	d	e	f	g	h	i	j
T	**E**	**C**	**H**	**N**	**I**	**A**	**L**	B	D
k	l	m	n	o	p	q	r	s	t
F	G	J	K	M	O	P	Q	R	S
u	v	w	x	y	z				
U	V	W	X	Y	Z				

Keyword: **technical**

- After all the substitutions the plaintext is found to be:

all at once from the dark line of the horizon whither it retired

to gain its momentum the monster rushed suddenly towards the abraham

lincoln with alarming rapidity stopped suddenly about twenty feet

from the hull and died out not diving under the water for its

brilliancy did not abate but suddenly and as if the source of this

brilliant emanation was exhausted then it reappeared on the other

side of the vessel as if it had turned and slid under the hull

any moment a collision might have occurred which would have been

- Note:
  - No punctuation!
  - Names: Abraham Lincoln.
  - From: Twenty thousand leagues under the sea: Jules Verne (1870).

Monoalphabetic cipher

Summary: How to perform cryptanalysis on Monoalphabetic Cipher?

- Generate the frequency table for the ciphertext. Then guess the letter the “e” is being substituted as.
- Look for any single letter (might be “a” or “i”)
- Search for any possible bigrams or trigrams in the ciphertext. (“an”, “in”, “of”, “to”, “is”, “the”, “and”, “was”, “for”…
- For the rest, some guessing works are to be done. Remember to fill up the substitution table as you go along guessing the substituted letter.
  - Plaintext above ciphertext (not the other way round)

POLY ALPHABETIC CIPHER

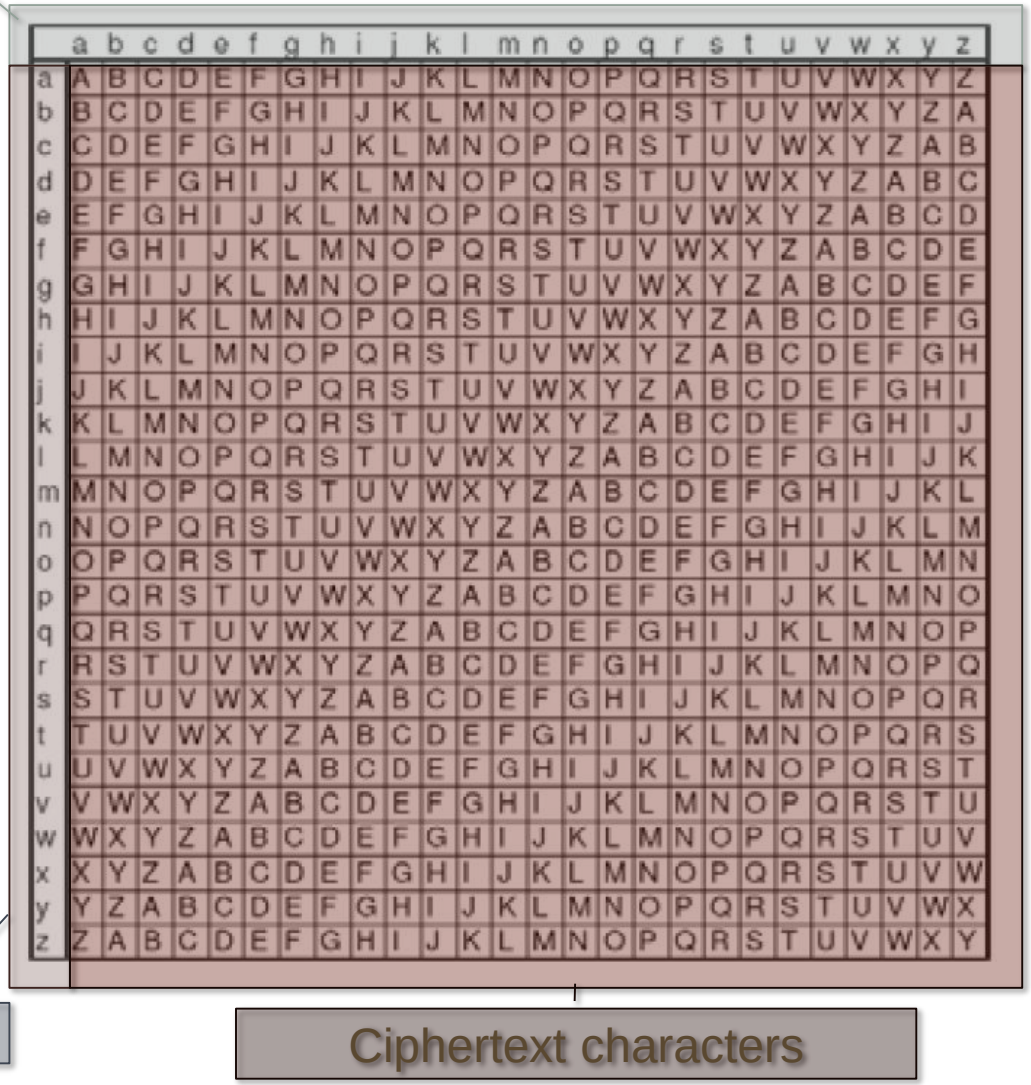
Polyalphabetic ciphers

- Basic idea: One ciphertext character can represent more than one (poly) plaintext characters, and yet decryption is possible; that is, it must be clear that when the ciphertext character can be decrypted to which one.

- An example of polyalphabetic cipher is Vigenere Cipher.

Polyalphabetic cipher

- Vigenere Cipher uses Trithemius table (Vigenere Tableau) to encrypt.
- If we know "modular addition", we do not need to use it (Trithemius Table.)



Polyalphabetic cipher

A quick revision on modular operations:

- Notation: **r mod n**
- Rough meaning: **r** is a remainder of some number when it is divided by **n**. So **r** should be one of 0, 1, 2, …, n-1.

Example:

- 5 mod 7 = 5
- 2 mod 7 = 2
- 10 mod 7 = 3
- 18 mod 7 = 4 mod 7 = 4

Note: if r is bigger than n, we should “reduce” it by dividing and getting a remainder.

Polyalphabetic cipher

- We can do modular addition, subtraction, and multiplication.

Example:

- 2 mod 7 + 3 mod 7 = (2 + 3) mod 7 = 5 mod 7 = 5
- 6 mod 7 + 7 mod 7 = (6 + 7) mod 7 = 13 mod 7 = 6
- 6 mod 7 – 8 mod 7 = (6 – 8) mod 7 = -2 mod 7

  ☺What is a remainder when -2 is divided by 7?
- 5 mod 7 x 3 mod 7 = (5 x 3) mod 7 = 15 mod 7 = 1

Polyalphabetic cipher

Modular addition:

- (a + b) mod n = r mod n
  - If $0 \leq r < n$ ➔ r
  - While ($r \geq n$)

    r = r – n
  - While ($r < 0$)

    r = r + n

How about modular subtraction and multiplication?

Polyalphabetic cipher

Now back to Vigenere cipher!

- Encrypt the plaintext ‘tellmeabout’ using Vigenere cipher with a key ‘name’

P	t	e	l	l	m	e	a	b	o	u	t
p	19	4	11	11	12	4	0	1	14	20	19
K	n	a	m	e	n	a	m	e	n	a	m
k	13	0	12	4	13	0	12	4	13	0	12
(*p*+*k*) mod 26	**6**	**4**	**23**	**15**	**25**	**4**	**12**	**5**	**1**	**20**	**5**
C	G	E	X	P	Z	E	M	F	B	U	F

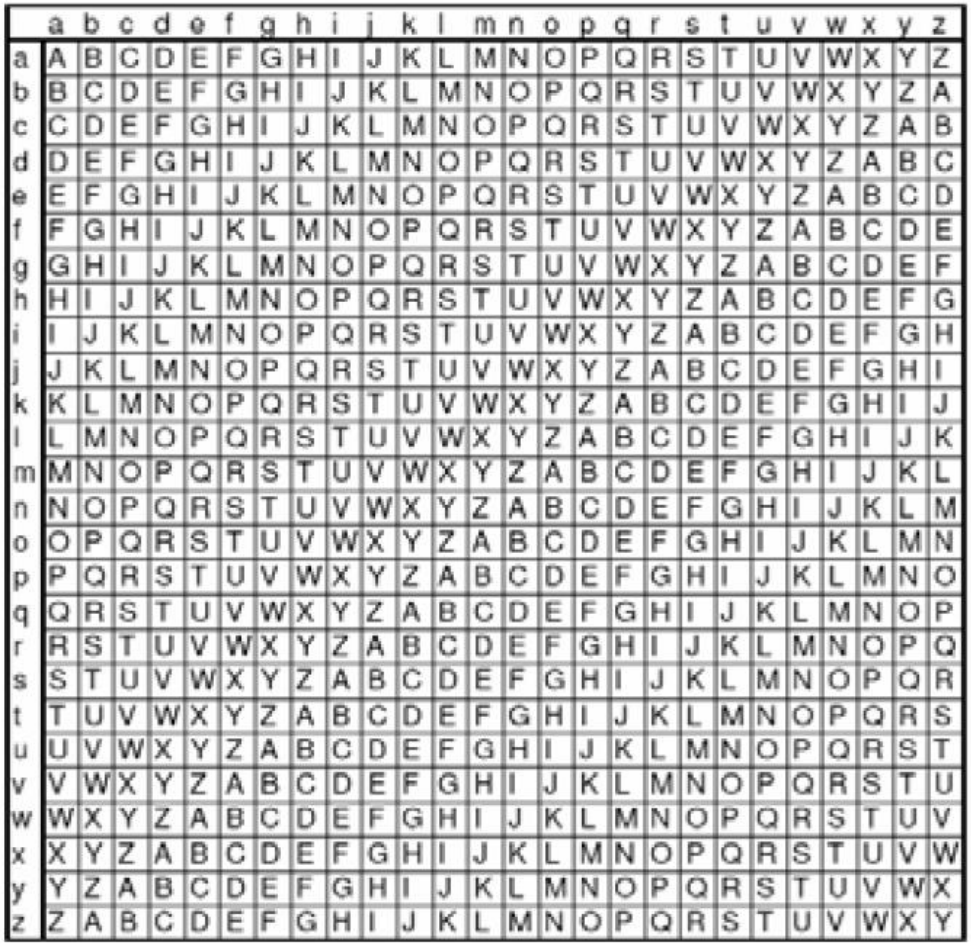
Polyalphabetic cipher

- How to decrypt?

Trithemius table (Vigenere Tableau) to the rescue ☺

…but note that we can decrypt using Trithemius table provided we have the key ☹

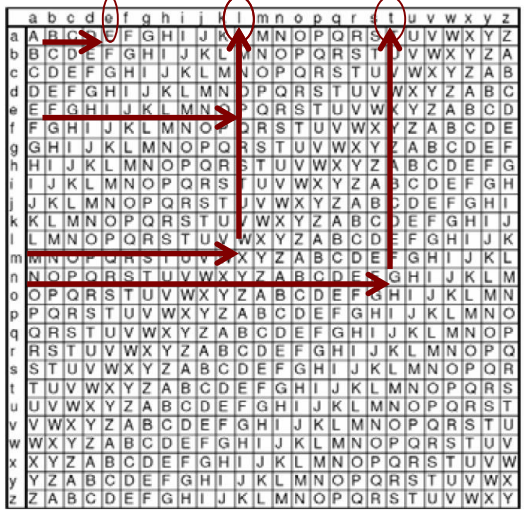
So how to get the key?



	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
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

Polyalphabetic cipher

C	G	E	X	P	Z	E	M	F	B	U	F
c	6	4	23	15	25	4	12	5	1	20	5
K	n	a	m	e	n	a	m	e	n	a	m
k	13	0	12	4	13	0	12	4	13	0	12
(*c*-*k*) mod 26	**19**	**4**	**11**	**11**	**12**	**4**	**0**	**1**	**14**	**20**	**19**
P	T	E	L	L	Z	E	A	B	O	U	T



Decryption process requires a key. This is not an issue because a legitimate user will know the key.

However, polyalphabetic cipher is not secure because the key can be easily obtained as well.

How to obtain the secrete key? Or How to break a polyalphabetic cipher?

Polyalphabetic cipher

- One possible way, we can use Index of Coincidence (IC) method or Kasiski method to deduce key length (period), then we use a substitution method to determine the key. Once we determine the key, we use simple additive cipher or Vigenere Tableau to decrypt.
- Index of coincidence (IC)
  - Probability that two randomly chosen elements of X are identical.
  - It is a measure of roughness.
  - It can be used to estimate the period (length of a key).

Polyalphabetic cipher

- After finding the period, break the ciphertext into pieces (blocks), each of which has the same length as the period.

Polyalphabetic cipher

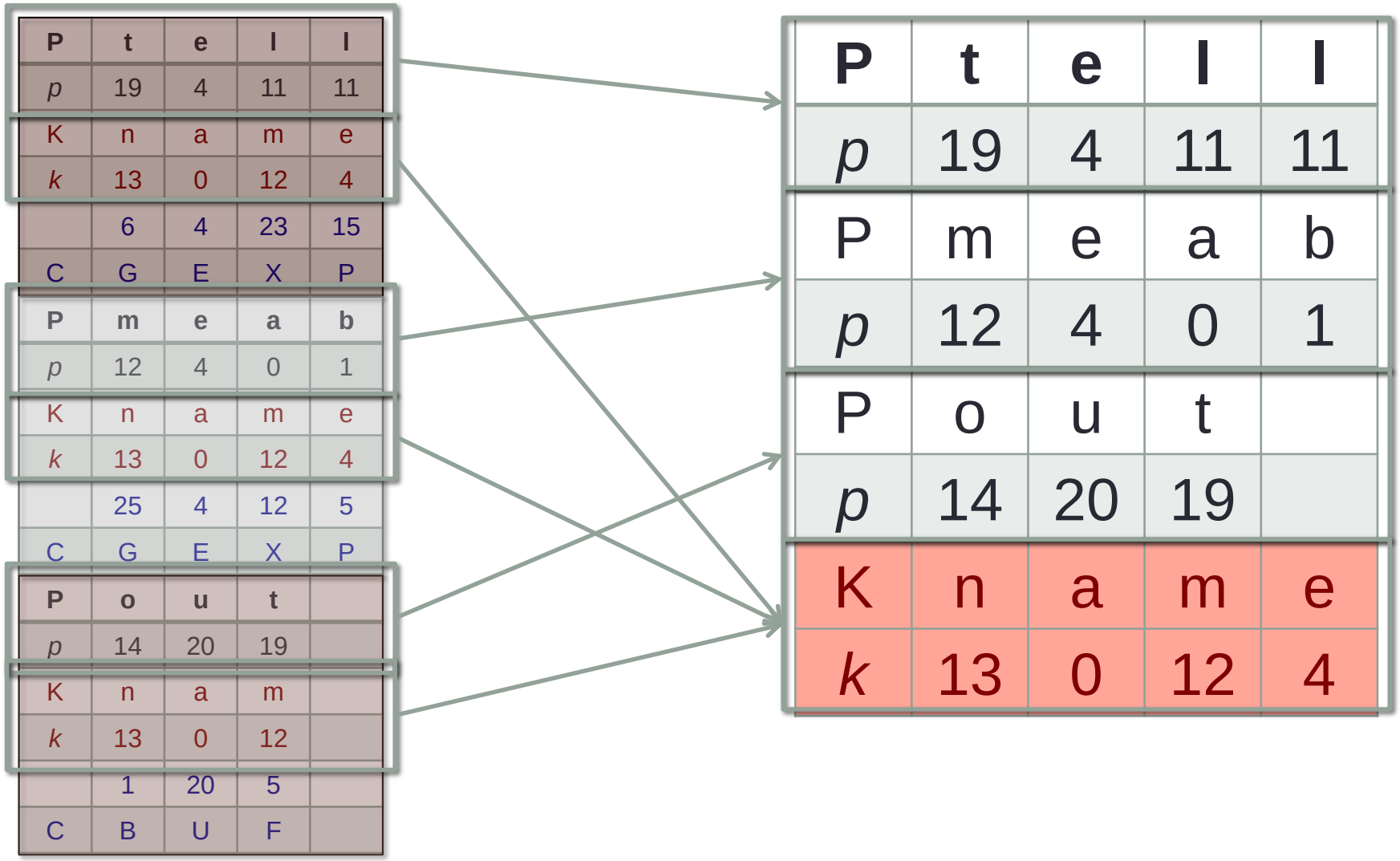
- Using the same example, we know that the key is ‘name’, thus the period is 4. (We will see how to determine the period if we do not know the key in a while.)
- Now we split the ciphertext into groups of 4 (the period or the size of the key.)

Polyalphabetic cipher

P	t	e	l	l	m	e	a	b	o	u	t
p	19	4	11	11	12	4	0	1	14	20	19
K	n	a	m	e	n	a	m	e	n	a	m
k	13	0	12	4	13	0	12	4	13	0	12
(*p*+*k*) mod 26	6	4	23	15	25	4	12	5	1	20	5
C	G	E	X	P	Z	E	M	F	B	U	F

…and arrange the blocks vertically on top of another as shown next.

Polyalphabetic cipher



Polyalphabetic cipher

P	t	e	l	l
p	19	4	11	11
P	m	e	a	b
p	12	4	0	1
P	o	u	t	
p	14	20	19	
K	n	a	m	e
k	13	0	12	4

P	t	e	l	l
C	6/G	4/E	23/X	15/P
P	m	e	a	b
C	25/Z	4/E	12/M	5/F
P	o	u	t	
C	1/B	20/U	5/F	

What is the implication? Or what is your observation?

Each group is an additive cipher; i.e., offset with a key K (Similar to Caesar cipher).

Krypto for polyalphabetic ciphers

-> r Ctext-2

-> l

bamec po nw ykwxzvn i pej wsmseesaiu wt mlv iagsvqvkij kf mlv brbkrpe

llv blxh rjd wmu jom ekpavo

ykulmec wbxy iuvl dkrx iolegwzre uyzhdbrxo wamtd tai fnizmewl miewnmw

twnm ewbokh kk rxstyuic fqr lsteemc rzoixj iagc jqca wllekjzyitp

twptfzhimc rjd hvzcigec ocbiepiymt yoggvltbseenz qrphxqrpivec okbpc

en gemegtxzkn tru axipfnammfj txgyjijyvo fhv tkpbrx en ysx jizlk wnw wkkrf

azph mlv enomjebei ywztvuo oy vfykl wykaew rjd vyinegxj phx kiaam wvw

vxrkqrxvj daw xf xe tfca th gfimtru wle xya pxsghe br kdebv uny eeez

-> i

IC = 0.042

Average = 0.042

- This Implies we have a polyalphabetic substitution cipher.

Krypto for polyalphabetic ciphers

- If we graph the frequencies this time we see it is a little flatter.
- The first thing we want to do is find the period, that is the number of alphabets used in the substitution. We can start by using the Friedman or Kappa test, for n=432.

$$\kappa = \frac{0.027n}{(IC)(n-1) - 0.038n + 0.065}$$

- This gives a value of about **6.6**, which is very close to 6.5.

Krypto for polyalphabetic ciphers

- Use the command **-> i <p>** where **<p>** is the period being tried.

-> i 8
IC = 0.035
IC = 0.035
IC = 0.046
IC = 0.038
IC = 0.034
IC = 0.041
IC = 0.034
IC = 0.041
Average = 0.038

- These indexes correspond to random so the period isn't 8. Let us build up a table of the IC values to try and find the length.

Krypto for polyalphabetic ciphers

Length	Kappa (K)
1	0.042
2	0.041, 0.041
3	0.041, 0.040, 0.041
4	0.040, 0.039, 0.042, 0.040
5	**0.059, 0.053, 0.066, 0.054, 0.062**
6	0.041, 0.042, 0.038, 0.037, 0.038, 0.038
7	0.044, 0.050, 0.038, 0.067, 0.044, 0.038, 0.041
8	0.035, 0.035, 0.046, 0.038, 0.034, 0.041, 0.034, 0.041

Krypto for polyalphabetic ciphers

- This table suggests we should try a length of 5.
- The keywords for the polyalphabetic assignment will have length between 5 and 10 inclusive. Note that duplicated letters are **not** removed.
- The keyword for the monoalphabetic may be longer (up to about 12) although there will almost certainly be duplicate letters so the effective keyword will be shorter.

Krypto for polyalphabetic ciphers

- The i^{th} letter in the keyword corresponding to the i^{th} substitution alphabet identifies a shift cipher. The plaintext letter “a” is mapped to the i^{th} keyword letter in the i^{th} alphabet, and this determines the complete substitution alphabet.
- Recall the most common letters in the English alphabet are “etao”. We can use their relative positions in the frequency graph to determine a possible correspondence.
  - That is, they should all occur a reasonable number of times.

Letter	a	e	o	t
Position	1	5	15	20

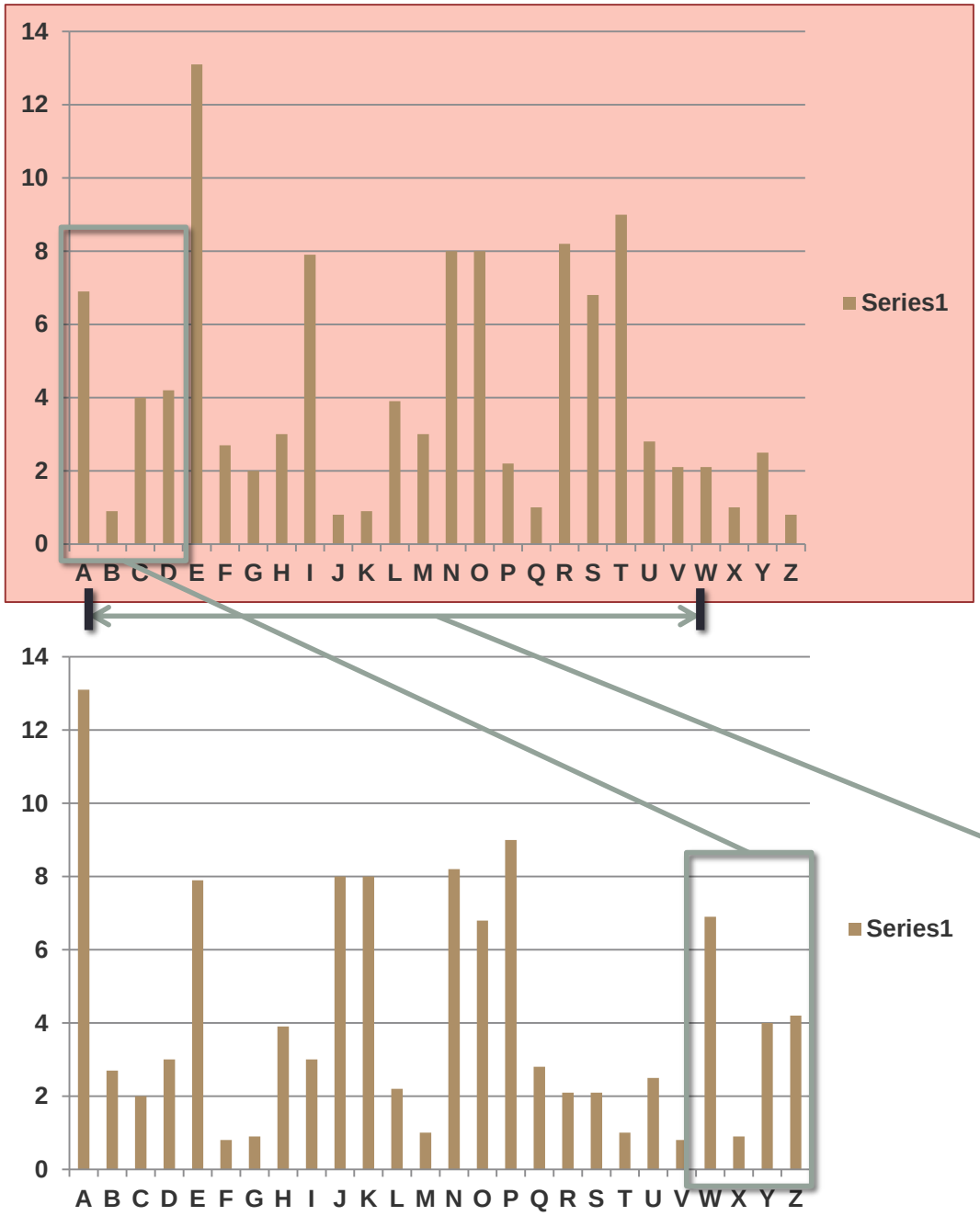
Krypto for polyalphabetic ciphers

- We can use

-> g i 5

i=0, 1, 2, 3, 4 to give the five sub alphabets.

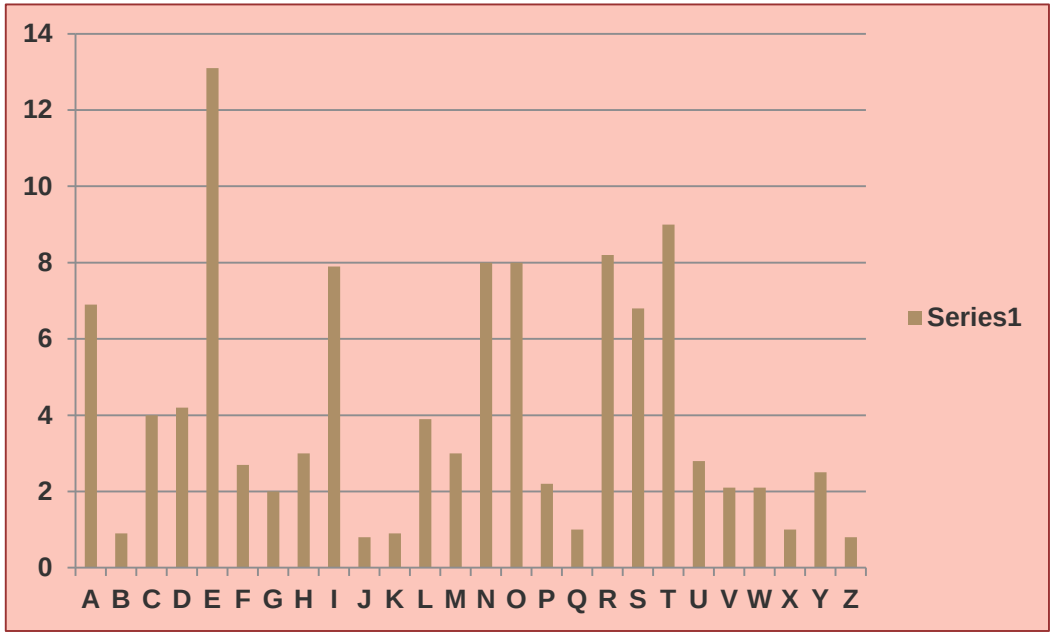
- Consider **-> g 0 5**.
- In this group, the most frequently used letter is “A”. Making assumption that this letter “A” (in ciphertext) is actually letter ‘e’ (in plaintext), we then check the rest to verify that the corresponding letters are sensible. For example,



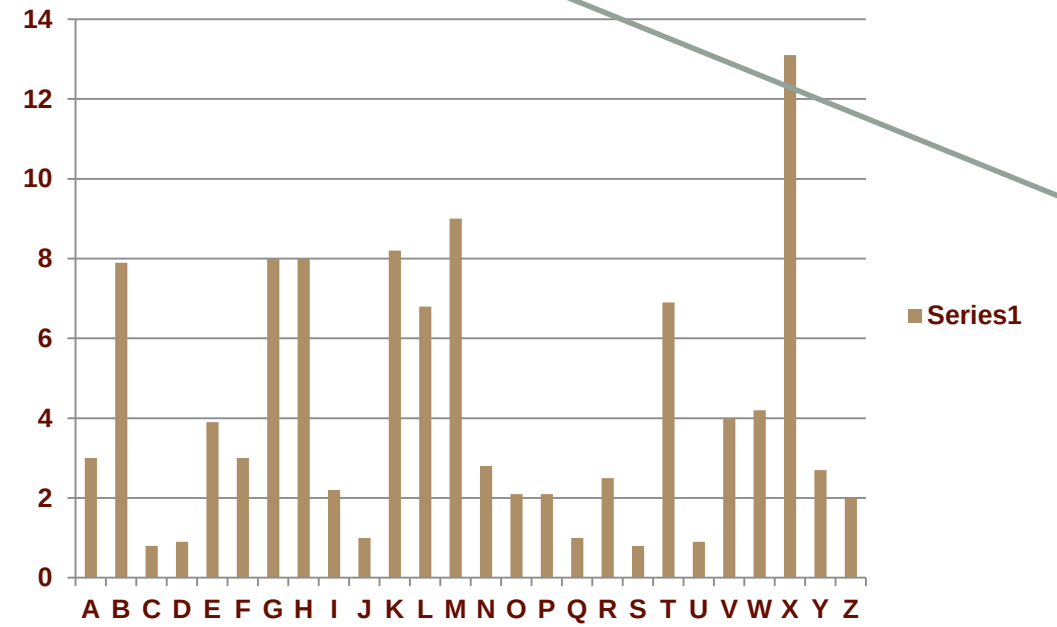
The letters frequencies of the usual English text.

The letter ‘A’ is shifted 23 places to the right and replaced with the letter ‘W’. Thus the key is 23 or ‘W’

The letters frequencies of the ciphertext.

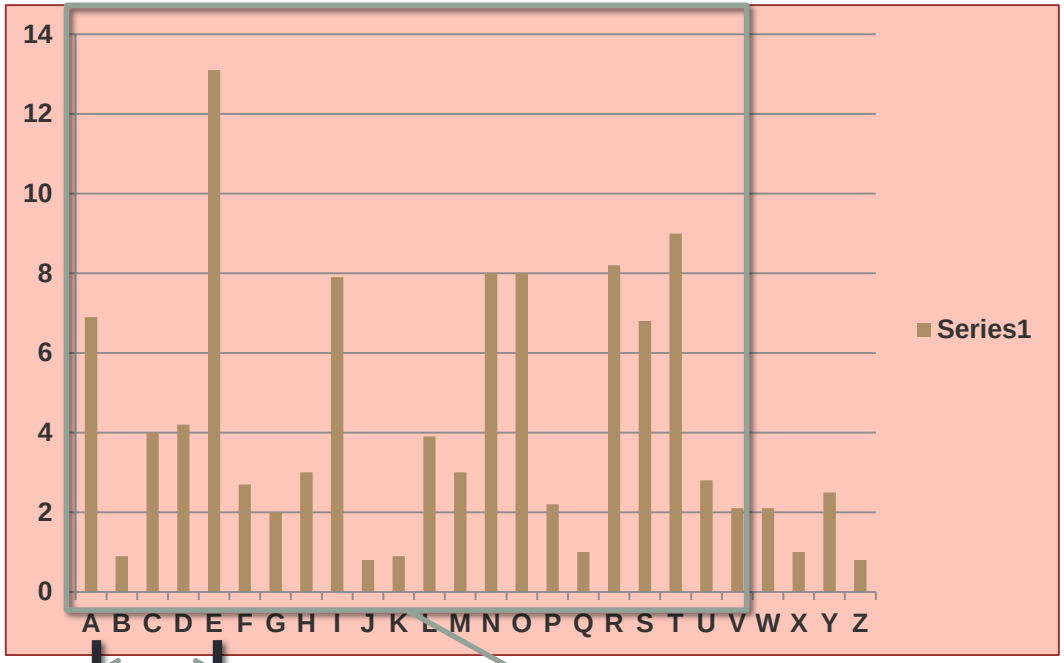


The letters frequencies of the usual English text.

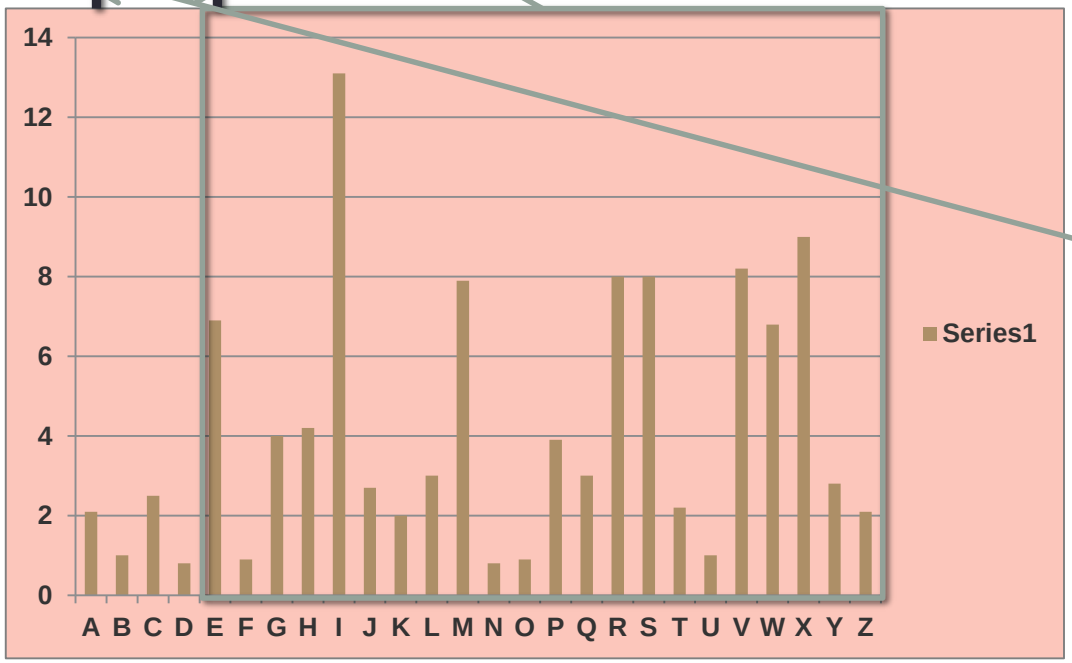


The letter ‘A’ is shifted 20 places to the right and replaced with the letter ‘T’. Thus the key is 20 or ‘T’

The letters frequencies of the ciphertext.

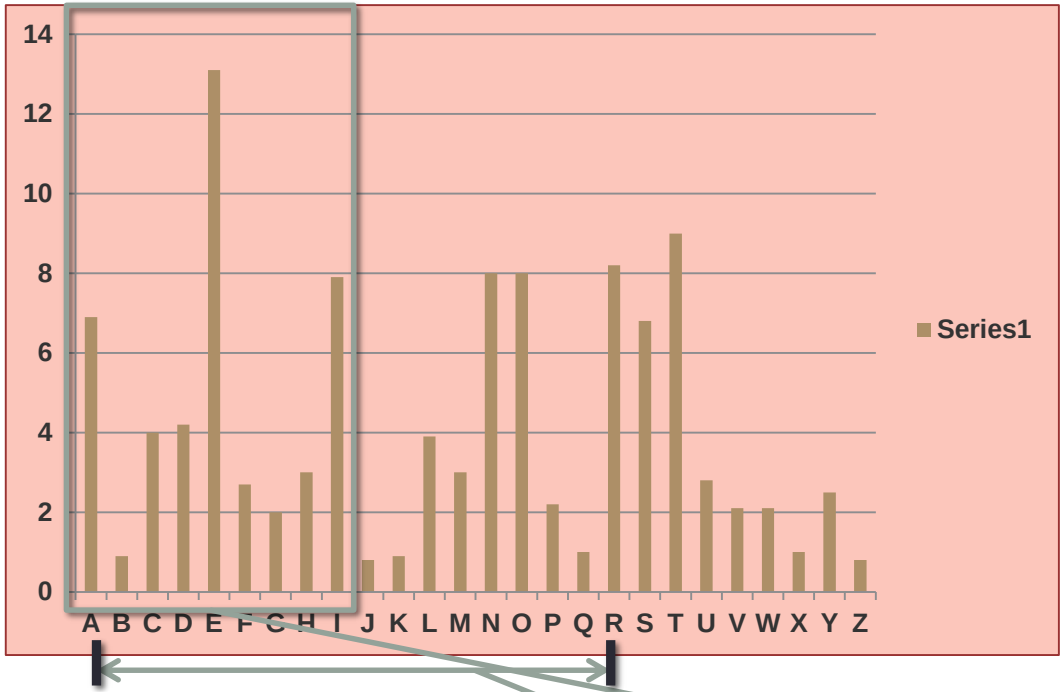


The letters frequencies of the usual English text.

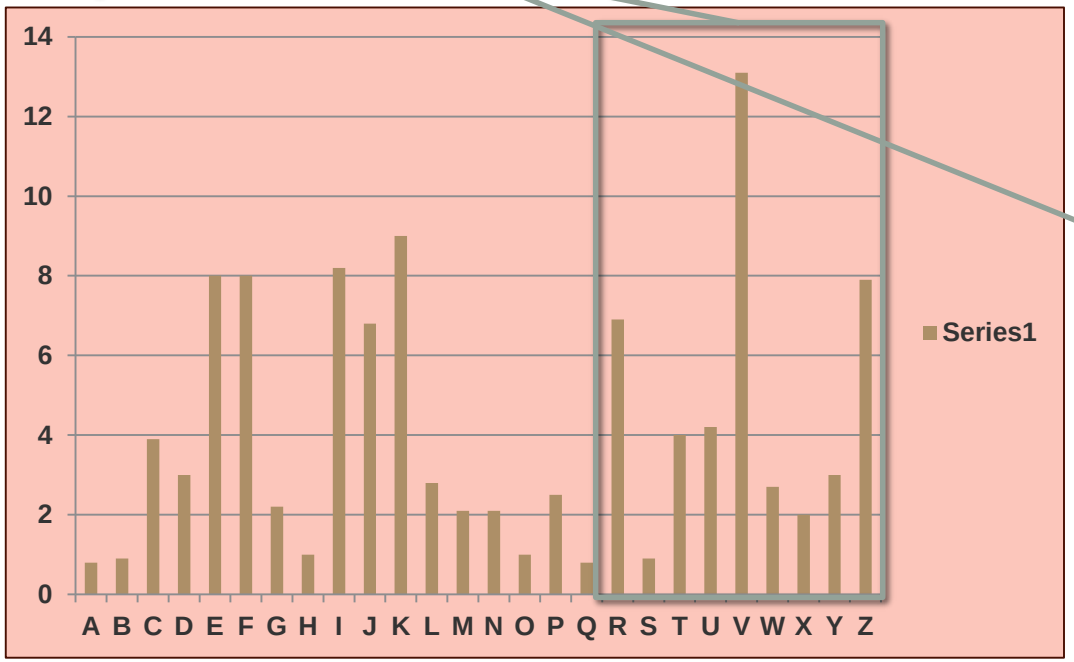


The letter ‘A’ is shifted 5 places to the right and replaced with the letter ‘E’. Thus the key is 5 or ‘E’

The letters frequencies of the ciphertext.



The letters frequencies of the usual English text.



The letter ‘A’ is shifted 18 places to the right and replaced with the letter ‘R’. Thus the key is 18 or ‘R’

The letters frequencies of the ciphertext.

- The keyword is **water**. We can check this using the substitution…

->S -B water
-> I
fatal to us however i was astonished at the manoeuvres of the frigate she fled and did not attack
housing with much more expensive buildings which the original tenants cant afford to reoccupy our society adopts many such superficial capability and original scientific conceptioning mathematical skill
in navigation and exploration techniques for coping in fog night and storm with the invisible hazards of rocks shoals and currents the great sea venturers had to be able to command all the people in their dry land

- Notes:
  - No punctuation again. e.g. can't becomes cant.
  - There could be spelling mistakes.
  - The sentences are from mixed sources.
    - You could have mixed sources in your ciphertext! They may not even form complete sentences.
  - Doesn't make much sense as a paragraph.

POLYALPHABETIC CIPHER

Summary: How to perform cryptanalysis on Vigenere Cipher?

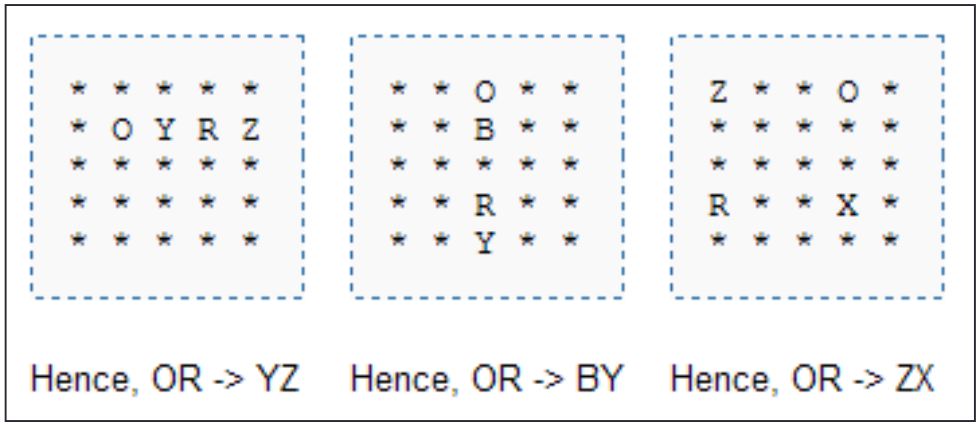
- Group the ciphertext with respect to the period size.
  - For example, if period is 3,
    - Group all the 1^{st}, 4^{th}, 7^{th}, … characters together
    - Group all the 2^{nd}, 5^{th}, 8^{th}, … characters together
    - Group all the 3^{rd}, 6^{th}, 9^{th}, … characters together
- Obtain the period by finding the IC close to 0.065 (for English language)
- Perform mono cipher on each group to get the starting letter.

CLASSICAL BLOCK CIPHER

Playfair

Playfair cipher

- Playfair Cipher
  - 5x5 key matrix (J is omitted)
  - 4 basic rules:
    - Same row: 1 letters right
    - Same col: 1 letters down
    - Same letter: insert a constant char in between the 2
    - Different col and row: first element anticlockwise



Playfair cipher

- Playfair Cipher
  - Same row: 1 char right
  - Same column: 1 char below
  - Rectangular: anti-clockwise
  - Encrypt “ComputerSecurity”

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

CO MP UT ER SE CU RI TY

OD TH MU GH HL NB HO YP

TRANSPOSITION CIPHER

Transposition cipher

In the lecture we considered transposition "ciphers" where we wrote the plaintext in row order and read the ciphertext off column by column.

- ***Decrypt the following ciphertext obtained using that method.***

TIOXSTZHCWJOHYEKNUVEDQBFMELO UROPRAG

Transposition cipher

There is a file **Ctext.txt** in the tutorial directory. The message is 35 characters long. We can load it into Krypto and use

-> l n 1

to look at breaking it into blocks of size n, and reading down the columns.

Transposition cipher

```
-> | 4 1
TIOX
STZH
CWJO
HYEK
NUVE
DQBF
MELO
UROP
RAG
```

```
-> | 5 1
TIOXS
TZHCW
JOHYE
KNUVE
DQBFM
ELOUR
OPRAG
```

```
-> | 6 1
TIOXST
ZHCWJO
HYEKNU
VEDQBF
MELOUR
OPRAG
```

Transposition cipher

-> | 7 1

TIOXSTZ

HCWJOHY

EKNUVED

QBFMELO

UROPRAG

THEQUICKBROWNFOXJUMPSOVERTHELAZY DOG

The quick brown fox jumps over the lazy dog

Transposition cipher

This method can be extended by specifying a key determining the order in which columns are read off.

- **A further extension are the double (or more) transposition ciphers, where the output of the first transposition array is re-encrypted through the same array.**
- **Use the double transposition cipher to encrypt the plaintext: “this is the test message for computer security this autumn”, using the key [3 2 5 1 7 4 6].**

Transposition cipher

thisisthetestmessageforcomputersecuritythisautumn

t	h	i	s	i	s	t
h	e	t	e	s	t	m
e	s	s	a	g	e	f
o	r	c	o	m	p	u
t	e	r	s	e	c	u
r	i	t	y	t	h	i
s	a	u	t	u	m	n

Key: 3 2 5 1 7 4 6

Output: itscrtuhesreiaisgmetutheotrstmfuuinseaosytstepchm

Transposition cipher

itscrtuhesreiaisgmetutheotrstmfuuinseaosytstepchm

i	t	s	c	r	t	u
h	e	s	r	e	i	a
i	s	g	m	e	t	u
t	h	e	o	t	r	s
t	m	f	u	u	i	n
s	e	a	o	s	y	t
s	t	e	p	c	h	m

Key: 3 2 5 1 7 4 6

Output: ssgefaeteshmetreetuscihittssuausntmcrmouoptitrith