LECTURE 3: Cryptography

Dr. Esam A. Alwagait
alwagait@ksu.edu.sa

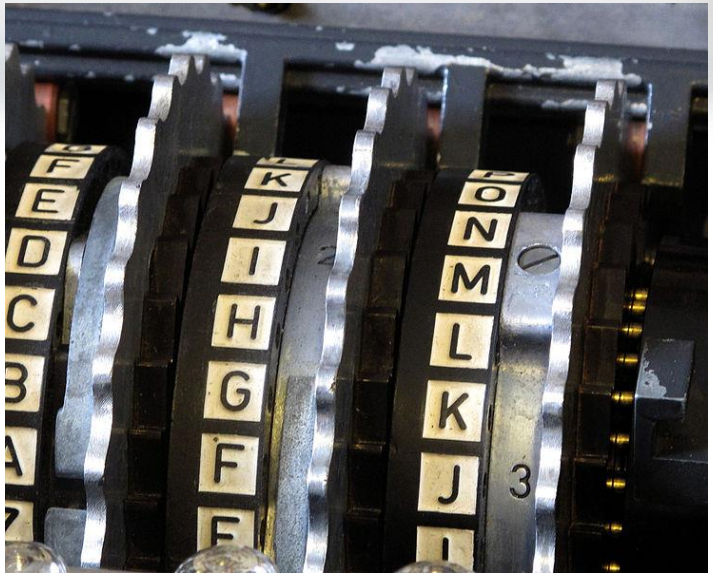
Other substitutions?

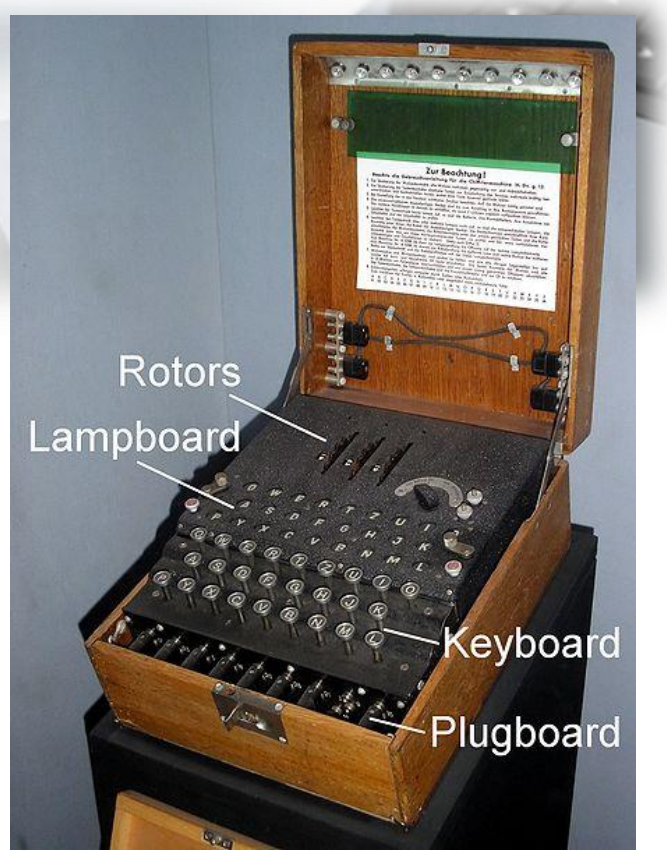
- Write substitutions in a function of permutations
  - $n_1(3) = 5$ means that C is transformed into e
- Use a key (say *word*)
  - P: ABCDEFGHIJKLMNOPQRSTUVWXYZ
  - K: word
  - C: wordabcefghijklmnpqstuvxyz

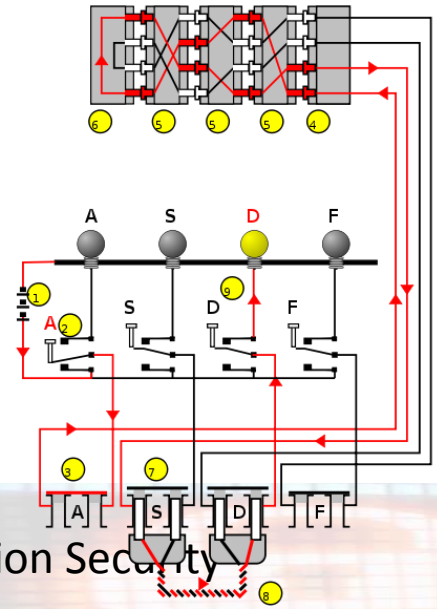
  - How about (*professional*)
  - P: ABCDEFGHIJKLMNOPQRSTUVWXYZ
  - C: profesinalbcdghjkmqtuvwxyz
- Counting and rearranging (say by threes?)
  - P: ABCDEFGHIJKLMNOPQRSTUVWXYZ
  - C: adgjmpsvybehknqtwzcfilorux
- A polyalphabetic cipher
  - Any cipher based on substitution, using multiple substitution alphabets

Enigma machine!

- Invented by the German engineer Arthur Scherbius at the end of World War I
- Electro-mechanical rotor cipher machines
- First broken by the Polish Cipher Bureau
- Broken due to German procedural flaws, operator mistakes, laziness, failure to systematically introduce changes in encipherment procedures

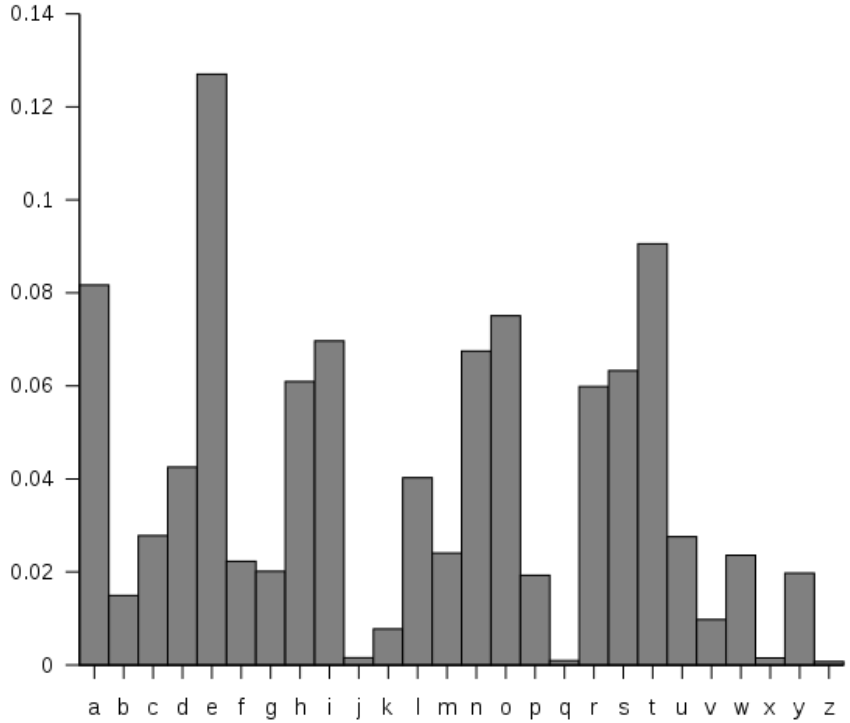






Cryptanalysis of Substitution Ciphers

- Tools:
  - Guess, strategy, and mathematical Skill
- Seems secure on the surface, why?
- There're 26! Possible different decipherments, how?
  = 26*25*24…..*2*1
- Using brute force attack? (on the surface)
  - One permutation per microsecond leads to a thousand years to test all 26! possibilities!
- How about frequency distribution analysis of English language?
  - The letters E, T, O, A occur more often than J, Q, X
  - With a good computer and enough ciphertext, may take an hour!
  - So, short ciphertext messages give attacker little to work with! (more secure)



Hill Cipher

- Developed by mathematician Lester Hill in 1929
- Encrypts ***m*** plaintext letters to ***m*** ciphertext letters
- For ***m=3***, let the plaintext be $\boldsymbol{P = (p_1, p_2, p_3)}$ and the ciphertext be $\boldsymbol{C = (c_1, c_2, c_3)}$
- For encryption
  - ***C = KP mod 26***, where ***C*** is the ciphertext matrix, ***K*** is the key matrix and ***P*** is the plaintext matrix
- Decryption uses the inverse of ***K***: Thus
  - $\boldsymbol{P = K^{-1} C \bmod 26}$ (where $\boldsymbol{KK^{-1} = K^{-1}K}$ = the ***identity matrix I***)

- The system can be described as a set of linear equations
- Cryptanalysis using frequency analysis is difficult, especially as ***m*** gets larger. This cipher hides single-letter frequencies, as well as **diagram**, **trigram**, and so on, up to **(m-1)-grams**, for any chosen block length ***m***
- Thus, the Hill cipher is strong against a ciphertext-only attack
  - However, it falls easily (almost trivially) to a known-plaintext attack

Review: inversion of 2X2 matrices

- In linear algebra an n-by-n (square) matrix A is called invertible if there exists an n-by-n matrix B such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

- where $\mathbf{I}_n$ denotes the *n*-by-*n* identity matrix
- matrix **B** is uniquely determined by **A** and is called the ***inverse*** of **A**, denoted by $\mathbf{A}^{-1}$
- A matrix is invertible if it is a square matrix and its determinant is not zero
- The key matrix K must satisfy two conditions to be correct:
  - The determinant is not zero
  - The determinant has a multiplicative inverse in Z_{26} (exist in the reciprocals module 26 table below)

Determinants' Reciprocals Modulo 26												
Determinant	1	3	5	7	9	11	15	17	19	21	23	25
Reciprocal Modulo 26	1	9	21	15	3	19	7	23	11	5	17	25

Hill Cipher example

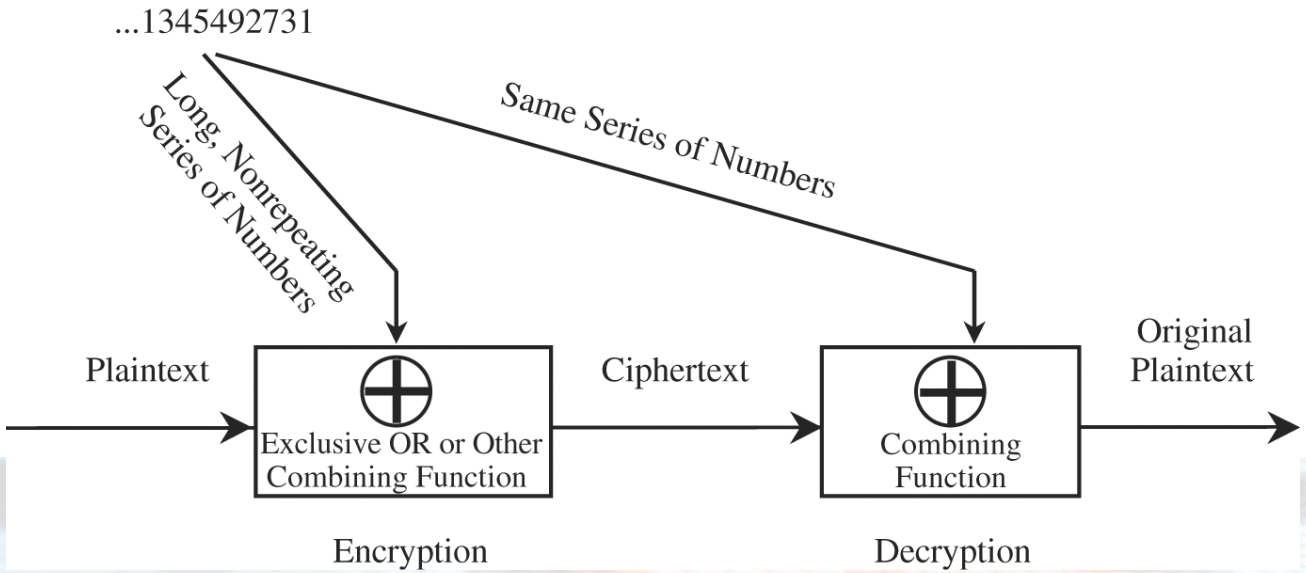
- $K = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$
- $\det(k) = -5 = 21 \pmod{26}$
  - note it is non-zero and has multiplicative inverse in Z_{26} reciprocals table
- $P = AT$, this is represented by $\begin{bmatrix} A \\ T \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ 19 \end{bmatrix}$
- To encrypt, compute $C = KP = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 19 \end{bmatrix} = \begin{bmatrix} 57 \\ 19 \end{bmatrix} = \begin{bmatrix} 5 \\ 19 \end{bmatrix} \pmod{26} = \begin{bmatrix} f \\ t \end{bmatrix}$
- To decrypt, find
- $\det(k)^{-1} = 21^{-1} = 5 \pmod{26}$, see the reciprocals table of modulo 26
- $K^{-1} = \det(k)^{-1} \begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix} = 5 \begin{bmatrix} 1 & 23 \\ 24 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 115 \\ 120 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 11 \\ 16 & 5 \end{bmatrix} \pmod{26}$
- Finally, compute $P = K^{-1}C = \begin{bmatrix} 5 & 11 \\ 16 & 5 \end{bmatrix} \begin{bmatrix} 5 \\ 19 \end{bmatrix} = \begin{bmatrix} 234 \\ 175 \end{bmatrix} = \begin{bmatrix} 0 \\ 19 \end{bmatrix} (mod\ 26) = \begin{bmatrix} A \\ T \end{bmatrix}$

One-Time Pads

- Each bit or character from the plaintext is encrypted by a modular addition with a bit or character from a secret random key (or pad) of the same length as the plaintext, resulting in a ciphertext
  - a large, nonrepeating set of keys is written on sheets of paper, glued together into a pad
  - if the keys are 20 characters long and a sender must transmit a message 300 characters in length, the sender would tear off the next 15 pages of keys
- Sometimes considered the perfect cipher
- Has two major problems
  - The need for absolute synchronization between sender and receiver
  - The need for an unlimited number of keys!

One-Time Pads: Vernam Cipher

- Vernam's invention used an arbitrarily long punched paper tape that fed into a teletype machine
- The tape contained random numbers that were combined with characters typed into the teletype.
- The sequence of random numbers had no repeats, and each tape was used only once



One-Time Pads: Vernam Cipher

- The letters would first be converted to their numeric equivalents

V	E	R	N	A	M	C	I	P	H	E	R
21	4	17	13	0	12	2	8	15	7	4	17

- Next, we generate random numbers
  - 76 48 16 82 44 03 58 11 60 05 48 88
- The encoded message is the sum mod 26 of each coded letter with the random number

Plaintext	V	E	R	N	A	M	C	I	P	H	E	R
Numeric Equivalent	21	4	17	13	0	12	2	8	15	7	4	17
+ Random Number	76	48	16	82	44	3	58	11	60	5	48	88
= Sum	97	52	33	95	44	15	60	19	75	12	52	105
= mod 26	19	0	7	17	18	15	8	19	23	12	0	1
Ciphertext	t	a	h	r	s	p	i	t	x	m	a	b

The message VERNAM CIPHER is encoded as tahrsp itxmab

One-Time Pads: Book Ciphers

- Another source of supposedly "random" numbers is any book, piece of music, or other object of which the structure can be analyzed
- Both the sender and receiver need access to identical objects
- Example: a telephone book
  - The sender and receiver might agree to start at page 35 and use two middle digits (ddd-DDdd) of each seven-digit phone number, mod 26, as a key letter for a substitution cipher
  - Ken Follett's novel: *The Key to Rebecca*, used as the source of keys for spies in World War II

One-Time Pads: Book Ciphers Vigenère Tableau

- Select a passage from Descarte's meditation: *What of thinking? I am, I exist, that is certain*
- Write the message (MACHINES CANNOT THINK) under enough of the key and encode the message by selecting the substitution in row p_i, column k_i
- ***K**: iamie xistt hatis cert*
- ***P**: MACHI NESCA NNOTT HINK*
- ***C**: uaopm kmkvt unhbl jmed*

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

	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

Summary of substitutions

- The goal of substitution is *confusion*
  - Making it difficult for a cryptanalysis to determine how a message and a key were transformed into ciphertext
- Substitutions are effective cryptographic devices
- Formed the basis of many cryptographic algorithms used for diplomatic communication through the first half of the twentieth century
- But substitution is not the only kind of encryption technique!
  - Transposition (or sometimes permutation)

Transpositions (Permutations)

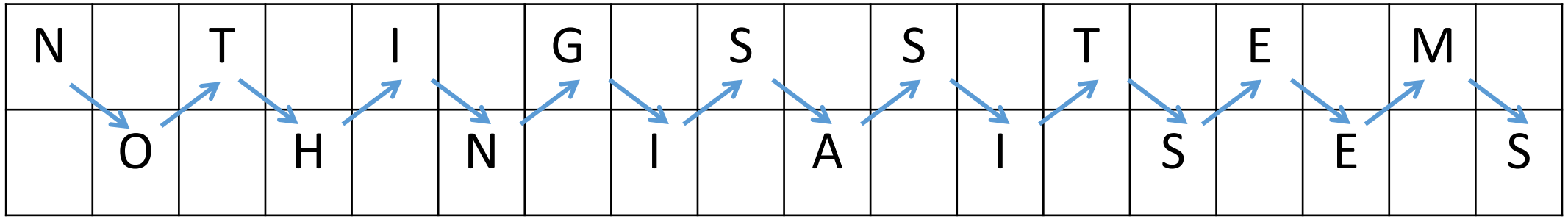
- Rearranging the letters of the plaintext message
- The goal is *diffusion*
  - i.e., scrambling the text so that adjacent-character analysis fails
- Widely spreading the information from the message or the key across the ciphertext
- The algorithm requires a constant amount of work per character, and the time needed to apply the algorithm is proportional to the length of the message
- algorithm requires storage for all characters of the message, so the space required is variable depending on the length of the message as well

Transposition Cipher

- Letters stay the same
- Order is different

Rail Fence Cipher

- Rail Fence Cipher
  - Encryption: plaintext is written down as a sequence of diagonals and then read off as a sequence of rows. The key of the rail fence is the number of rails , i.e., the depth of the diagonal.
- Example:
  - Key: rail fence of depth 2
  - Plaintext: NOTHING IS AS IT SEEMS
  - Encrypt: First writing NOTHING IS AS IT SEEMS on two lines in a zig-zag pattern (or rail fence). The ciphertext is produced by transcribing the first row followed by the second row.



- Ciphertext: NTIGS STEMO HNIAI SES

Rail Fence Cipher

- Decryption: In case of rail fence depth 2, write half the letters on one line, half on the second.(Note that if there are an odd number of letters, include the “middle” letter on the top line.)

  - **Example:** Decipher **MKHSE LWYAE ATSOL.**
  - **Solution:** Since there are 15 letters, we write 8 on the top line and 7 on the bottom line so that

M		K		H		S		E		L		W		Y
	A		E		A		T		S		O		L	

  - Plaintext: MAKE HASTE SLOWLY

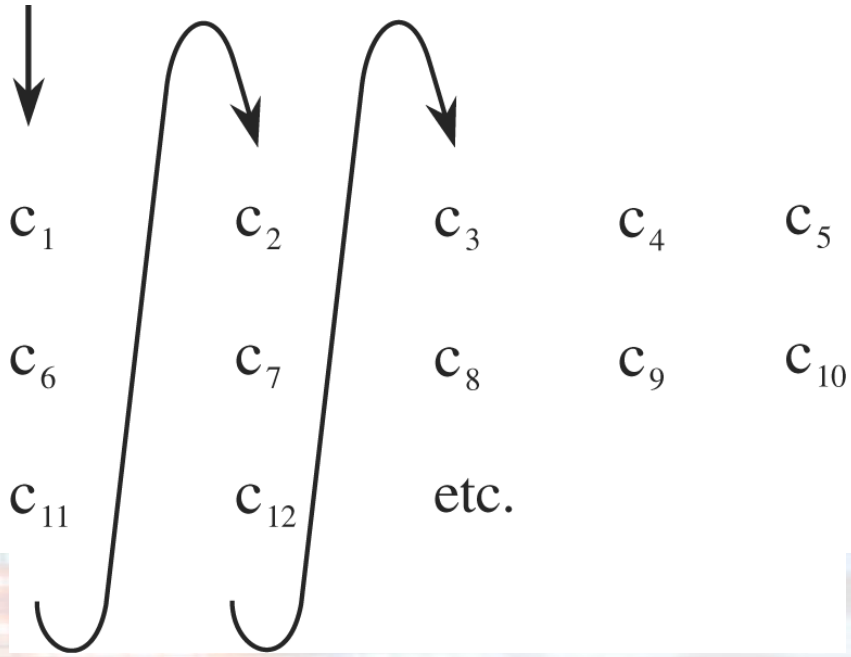
Columnar Transposition

- Key: The number of columns (e.g., k) is the key information.
- Encryption: Plaintext is written horizontally in k columns, and is then transcribed vertically *column-by-column*
- Decryption: Suppose that the length of the ciphertext is ***n*** and the key is ***k***. Then the letters will fill ***n DIV k*** full rows, and there will be one partial row at the end with ***n MOD k*** letters. Transcribing ***row-by-row*** will then yield the plaintext

Columnar Transposition

- Rearrangement of the characters of the plaintext into columns
- How about decipherment/decipherment complexity? (storage space and time)
- Example: five-column transposition:

C_1	C_2	C_3	C_4	C_5
C_6	C_7	C_8	C_9	C_{10}
C_{11}	C_{12}	etc.		



Columnar Transposition (Example:Encryption)

- Question:
  - Encrypt "NOTHING IN THE WORLD IS MORE DANGEROUS THAN SINCERE IGNORANCE AND CONSCIENTIOUS STUPIDITY" with a key of k = 9 columns.
- Solution:
  - We write the plaintext horizontally in 9 columns as follows:

N	O	T	H	I	N	G	I	N
T	H	E	W	O	R	L	D	I
S	M	O	R	E	D	A	N	G
E	R	O	U	S	T	H	A	N
S	I	N	C	E	R	E	I	G
N	O	R	A	N	C	E	A	N
D	C	O	N	S	C	I	E	N
T	I	O	U	S	S	T	U	P
I	D	I	T	Y				

- The cipher text is therefore:

NTSES NDTIO HMRIO CIDTE OONRO

OIHWR UCANU TIOES ENSSY NRDTR

CCSGL AHEEI TIDNA IAEUN IGNGN NP

Columnar Transposition(Example: Decryption)

- Question:
  - Suppose the ciphertext is: GPSDO AILTI VRVAA WETEC NITHM EDLHE TALEA ONME.
  - It is known that the key is k = 7 columns
  - What is the plaintext ?

- Solution:
  - There are 39 letters in the ciphertext which means that there are ***39 DIV 7 = 5*** full rows and one partial row with ***39 MOD 7 = 4*** letters

Columnar Transposition(Example: Decryption) Cont…

G	I	V	E	M	E	A
P	L	A	C	E	T	O
S	T	A	N	D	A	N
D	I	W	I	L	L	M
O	V	E	T	H	E	E
A	R	T	H	X	X	X

Plain Text:

GIVE ME A PLACE TO STAND AND I WILL MOVE THE EARTH