Classic cryptography

CRYPTOGRAPHY AND COMPUTER SECURITY

Ana I. González-Tablas Ferreres
José María de Fuentes García-Romero de Tejada
Lorena González Manzano
Sergio Pastrana Portillo

uc3m | Universidad Carlos III de Madrid

COSEC

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- INTRODUCTION

CLASSIC CRYPTOGRAPHY (5th century B.C.)

greek: kryptos = hidden

- Needed a key and a ciphering algorithm.
- Symmetric cipher: both parties must use the same key for encryption and decryption.
- The intention was to guarantee confidentiality concealing the contents of the messages.

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- INTRODUCTION

Two basic techniques are used. Both of them work with characters:

- **Substitution**: each character or letter in the plaintext is modified or substituted by another element in the ciphertext.

- **Transposition or permutation:** all characters or letters in the plaintext are reallocated in the ciphertext, according to some rule, without any modification.

(Shannon formalized these statements some centuries later)

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

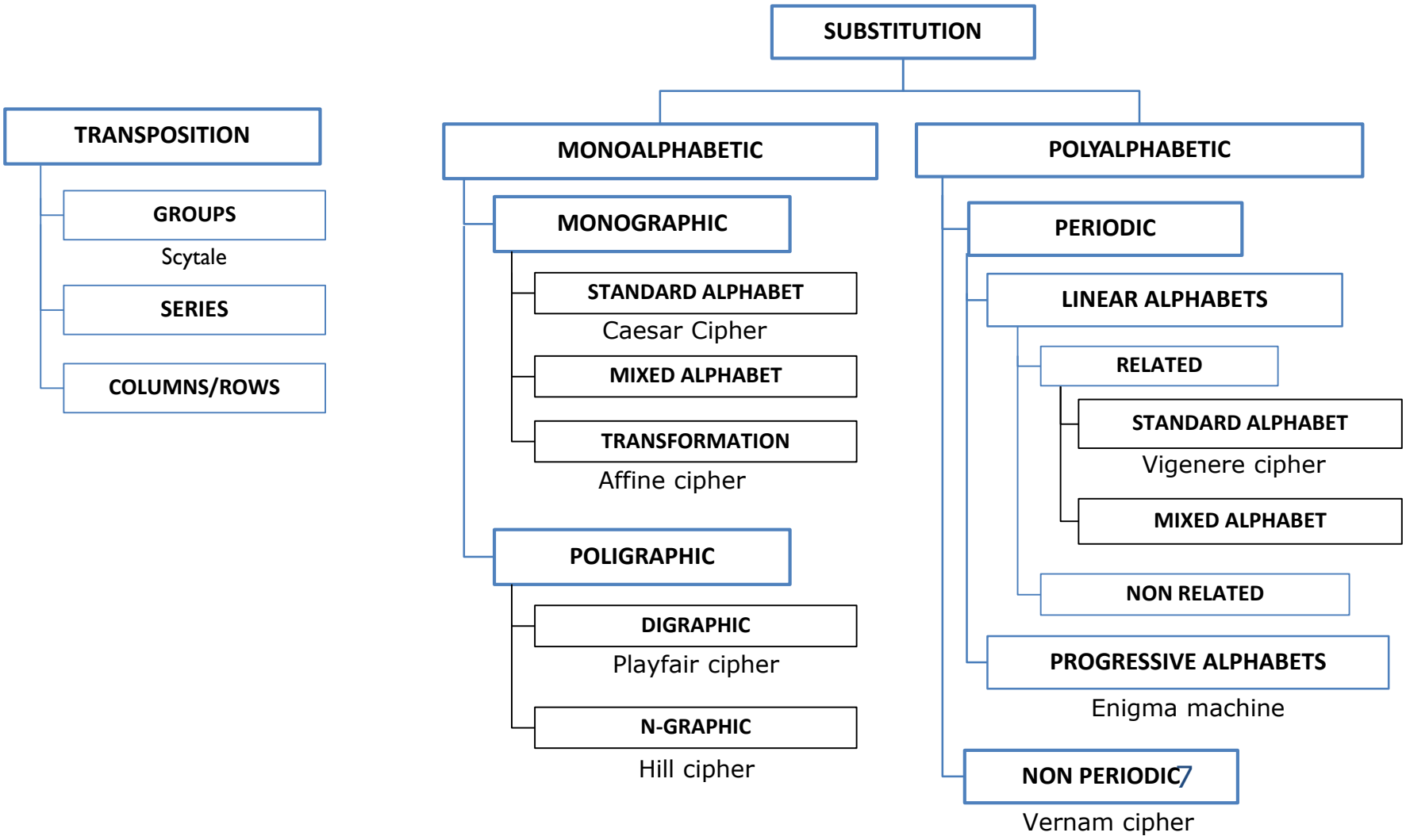
OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- CLASSIFICATION



Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

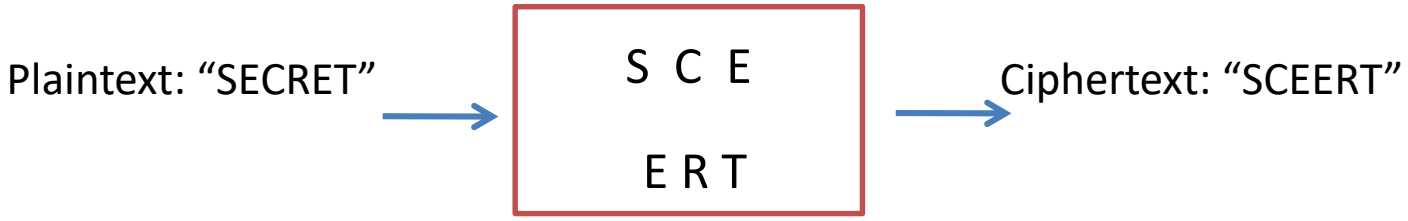
CLASSIC CRYPTOGRAPHY METHODS

- TRANSPOSITION

  The original characters are not altered. The only change happens on their position.

  The frequency of appearance for each letter is not modified (comparing plaintext and ciphertext) → Frequence analysis

  - EXAMPLES:
    - RAIL FENCE TRANSPOSITION
      - The Rail Fence Cipher involves writing messages so that alternate letters are written on separate upper and lower lines.
      - The sequence of letters on the upper line is then followed by the sequence on the lower line, to create the final encrypted message.



Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- GROUP TRANSPOSITION

  Permutation Π_M describes the order for a group of p letters

Example:

Π_M=24531

M = MANOS ARRIB AESTO ESUNA TRACO

C = AOSNM RIBRA ETOSA SNAUE RCOAT

- By increasing the period length p, the cipher will be less vulnerable
- p = message length => transposition by series

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

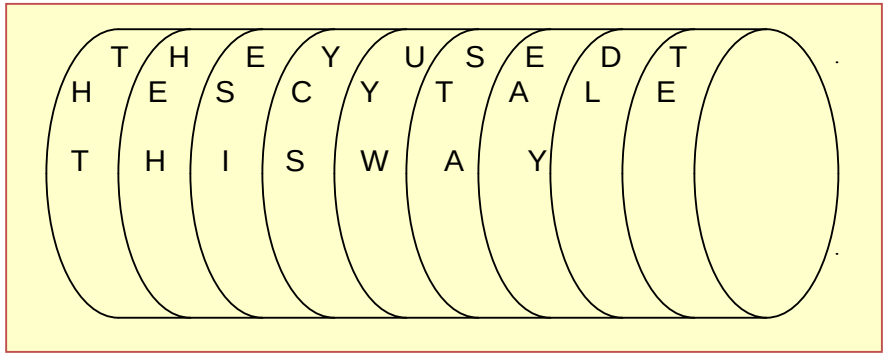
CLASSIC CRYPTOGRAPHY METHODS

- GROUP TRANSPOSITION. EXAMPLES:

- SCYTALE
  - Tapered baton + strip of parchment or leather
  - With the strip of leather wound, the message is written across the baton, longitudinally

- The plaintext is recovered winding the strip on another baton with the same diameter
  - System key is the diameter of the baton

M = THEY USED THE SCYTALE THIS WAY



C = THTHEHESIYCSUYWSTAEAYDLTE

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- TRANSPOSITION BY SERIES

The message is ordered as a chain of submessages:

$$M' = M_{S1}M_{S2}M_{S3}..., \text{ with } M_{Sx} \text{ functions or series.}$$

Example:

M_{S1} = 1,2,3,5,7,11,13,17,19,23 (primes)
M_{S2} = 4,6,8,10,12,14,16,18,20,22,24,26 (even numbers)
M_{S3} = 9,15,21,25,27 (odd numbers)

M = ERRAR ES HUMANO, PERDONAR DIVINO
C = ERRRSAODNI AEHMNPROADV NUERIO

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- TRANSPOSITION BY COLUMNS/ROWS
- How it works:
  1. Symbols are placed following a certain geometric pattern,
  2. And then extracted according to a certain path.

  Bidimensional pattern (matrix).
  - Symbols are placed in consecutive rows (columns) and then extracted column by column (row by row) from the first to the last.

M = ESTE ES UN EJEMPLO DE TRANSPOSICIÓN COLUMNAR

C = ESEDNICN SUMESCOA TNPTPILR EELROOUX EJOASNMX

E	S	T	E	E
S	U	N	E	J
E	M	P	L	O
D	E	T	R	A
N	S	P	O	S
I	C	I	O	N
C	O	L	U	M
N	A	R	X	X

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- AN EXAMPLE OF COLUMNAR TRANSPOSITION, WITH A KEY

Key = ESPÍA (alphabetical order: A,E,I,P,S)

M = EJEMPLO DE TRANSPOSICIÓN COLUMNAR CON CLAVE

E	S	P	I	A
E	J	E	M	P
L	O	D	E	T
R	A	N	S	P
O	S	I	C	I
O	N	C	O	L
U	M	N	A	R
C	O	N	C	L
A	V	E	X	X

A	E	I	P	S
P	E	M	E	J
T	L	E	D	O
P	R	S	N	A
I	O	C	I	S
L	O	O	C	N
R	U	A	N	M
L	C	C	N	O
X	A	X	E	V

C = PTPILRLX ELROOUCA MESCOACX EDNICNNE JOASNMOV

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- SUBSTITUTION

NUMERICAL REPRESENTATION OF THE ALPHABETS

- 27 letters alphabet: (A, B,..., Z) → (0, 1,...,26)
- 37 letters alphabet: (A, B,..., Z, 0, 1, ...9) → (0, 1,...,36)

0	A
1	B
2	C
3	D
4	E
5	F
6	G

7	H
8	I
9	J
10	K
11	L
12	M
13	N

14	Ñ
15	O
16	P
17	Q
18	R
19	S
20	T

21	U
22	V
23	W
24	X
25	Y
26	Z

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- SIMPLE MONOALPHABETIC SUBSTITUTION (MONOGRAPHIC)

Substitution of 1 character of plaintext by 1 character of ciphertext

$$E(m_i)=(am_i+b) \bmod n$$

a : decimation constant

b : shift constant

n : number of letters of the alphabet (27 for Spanish)

Key = (a,b)

gcd(a,n)=1 (condition for the existence of solution of the congruential equation)

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

MONOGRAPHIC SUBSTITUTION. PARTICULAR CASES

Shift cipher (Caesar, ROT 13, …)

$$E(m_i)=(m_i+b) \bmod n$$

Caesar cipher

$$E(m_i)=(m_i+3) \bmod n$$

Decimation cipher

$$E(m_i)=(am_i) \bmod n$$

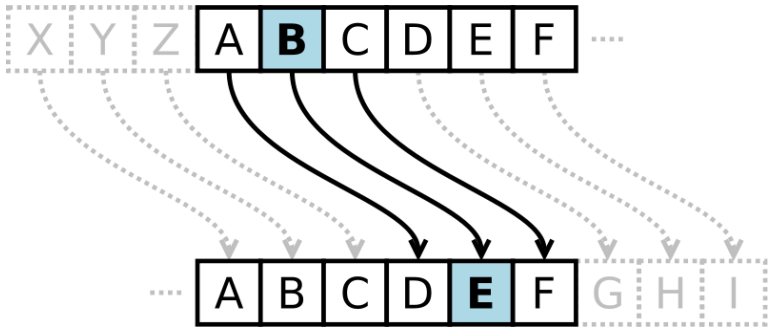
Affine cipher

$$E(m_i)=(am_i+b) \bmod n$$

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

– EXAMPLE. CAESAR CIPHER



$E_3(x) = (x + 3) \bmod 27$

$D_3(x) = (x - 3) \bmod 27$

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

M = NUNCA VI NEVAR TANTO

C = PXPFD YL PHYDU WDPWR

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- POLYGRAPHIC MONOALPHABETIC SUBSTITUTION

  Substitution of n (n >=2) plaintext characters by n ciphertext characters

  $\mathbf{M = m_1m_2 \cdot m_3m_4 \cdot \ldots \cdot m_{N-1}m_N}$

  $\mathbf{E_k(M) = E_k(m_1 \cdot m_2) \cdot E_k(m_3 \cdot m_4) \cdot \ldots \cdot E_k(m_{N-1} \cdot m_N)}$

  $\mathbf{E_k(M) = c_1c_2 \cdot c_3c_4 \cdot \ldots \cdot c_{N-1}c_N}$

  Methods:

  - Playfair (Wheatstone)
  - Hill

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- PLAYFAIR
  - Digraphic substitution. Digraphs
  - Matrix of 5x5 characters (taking J and Ñ out) with key starting in first row, without repeating characters
  - m_1m_2 same row, $c_1c_2 \rightarrow$ right
  - m_1m_2 same column, $c_1c_2 \rightarrow$ below
  - $m_1m_2 \neq$ row $\neq$ column, $c_1c_2 \rightarrow$ diagonal
  - Repeated digraphs $\rightarrow$ a pad character must be inserted to split the digraph

Playfair Matrix, adapted to Spanish, no key

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

Playfair Matrix, adapted with key: PRIMAVERA

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

RI → IM
BI → HB
ES → CO
BE → CB
FU → NP
OC → SE
OT → QN
AL → DT

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- HILL
  - Ciphers N characters at a time (example: “pan” → “dyj”)
  - Makes use of simple linear equations
    - NxN linear matrix transformations
    - K_E (NxN) must have an inverse in the cipher field
    - Pad characters if text length is not multiple of N
  - It is an example of a ‘block cipher’

$$C = K_E * M \pmod{N}$$

$$M = K_D * C \pmod{N}$$

$$K_D = K_E^{-1} \pmod{N}$$

$$\begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{pmatrix} = \begin{pmatrix} k_{1,1} & k_{1,2} & \dots & k_{1,n} \\ k_{2,1} & k_{2,2} & \dots & k_{2,n} \\ \dots & \dots & \dots & \dots \\ k_{n,1} & k_{n,2} & \dots & k_{n,n} \end{pmatrix} \mathbf{x} \begin{pmatrix} m_1 \\ m_2 \\ \dots \\ m_n \end{pmatrix}$$

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CLASSIC CRYPTOGRAPHY METHODS

- HILL. EXAMPLE

M= I CAN'T DO IT

8 2 0 13 19 3 14 8 19

C= EOM TMY SVJ

4 14 12 19 12 14 18 21 9

$$\begin{pmatrix} 4 \\ 14 \\ 12 \end{pmatrix} = \begin{pmatrix} 9 & 18 & 10 \\ 16 & 21 & 1 \\ 5 & 12 & 23 \end{pmatrix} \times \begin{pmatrix} 8 \\ 2 \\ 0 \end{pmatrix} \quad \text{(mod. 26)}$$

$$\begin{pmatrix} 19 \\ 12 \\ 14 \end{pmatrix} = \begin{pmatrix} 9 & 18 & 10 \\ 16 & 21 & 1 \\ 5 & 12 & 23 \end{pmatrix} \times \begin{pmatrix} 13 \\ 19 \\ 3 \end{pmatrix} \quad \text{(mod. 26)}$$

$$\begin{pmatrix} 18 \\ 21 \\ 9 \end{pmatrix} = \begin{pmatrix} 9 & 18 & 10 \\ 16 & 21 & 1 \\ 5 & 12 & 23 \end{pmatrix} \times \begin{pmatrix} 14 \\ 8 \\ 19 \end{pmatrix} \quad \text{(mod. 26)}$$

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

POLYALPHABETIC SUBSTITUTION

- Blaise de Vigenére (french cryptographer, 1523-1596)
  - 26 alphabets
  - 26 Caesar monoalphabetic substitutions
  - Key length m

$$E(m_j) = (m_j + k_{(j \bmod m)}) \bmod 26$$

where:

k_i = shift for alphabet i

m_j = letter in the j-th position in the text

$E(m_j)$ = Encrypted character

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

POLYALPHABETIC SUBSTITUTION

Vigenére's table

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

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

POLYALPHABETIC SUBSTITUTION

Vigenére's key:

- The key defines the shift used for each letter in the text,
  i.e.: **SOL**

- Encryption:

Message:	H	E	L		L	O	M		A	T	E
Periodic key:	S	O	L		S	O	L		S	O	L
Chipertext:	Z	S	W		D	C	X		S	H	P

- Using the table:

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

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

Polyalphabetic Substitution

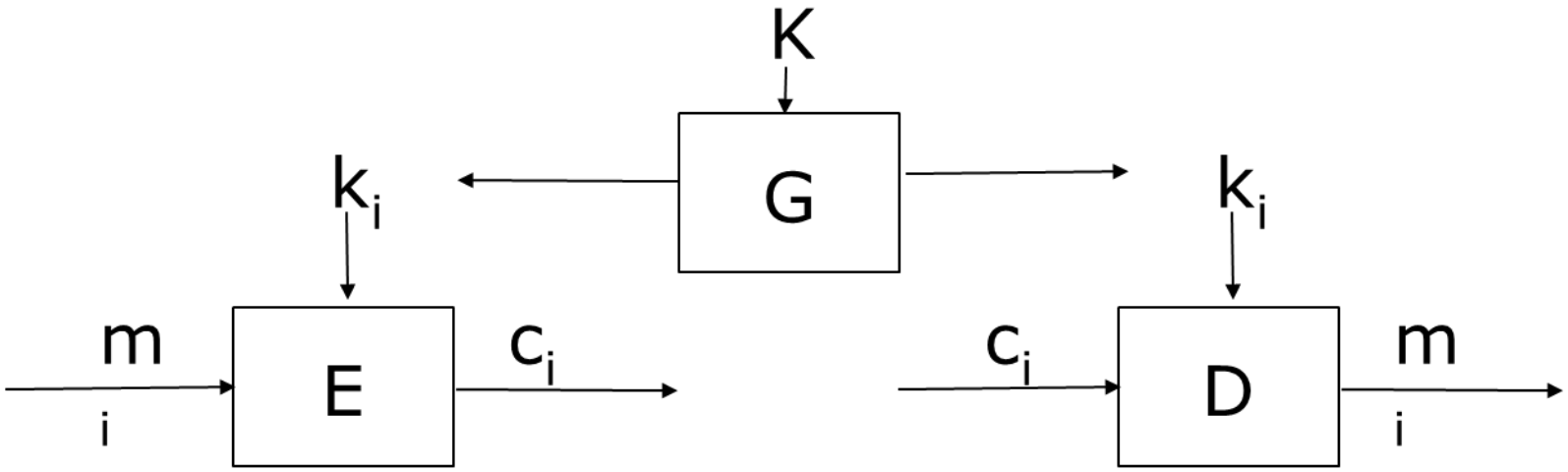
- Vigenère with autokey
- Instead of repeating the key over and over in order to encrypt the text, the key is used once and the cleartext is used to decrypt or encrypt the text:
- Key: SOL

A	T	O	M	I	C	P	L	A	N
S	O	L	A	T	O	M	I	C	P
S	H	Z	M	B	Q	B	T	C	C

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

Polyalphabetic Substitution

- Non-periodic polyalphabetic substitution: Vernam



$$c_i = E(m_i) = m_i \oplus k_j \qquad m_i = D(c_i) = c_i \oplus k_j$$

- Perfect secrecy if the key:
  - is truly random,
  - as large as or greater than the plaintext,
  - never reused in whole or part, and
  - kept secret.

POLYALPHABETIC SUBSTITUTION

- Non-periodic polyalphabetic substitution: Vernam
  - Disadvantages

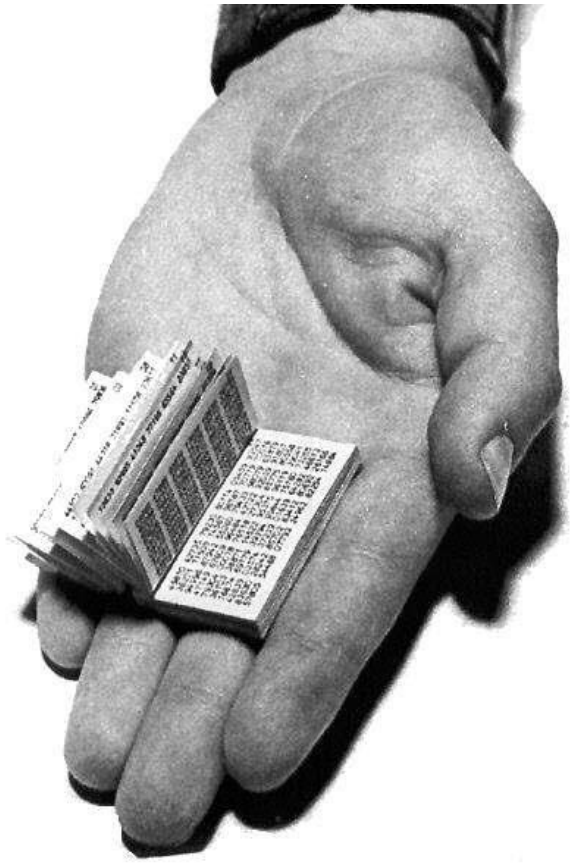
    Key size,

    Randomness (key generation)

    Can only be used once (key distrib.)

  - Advantages

    Perfect secrecy, impossible to cryptoanalyze



Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

POLYALPHABETIC SUBSTITUTION

- Enigma machine
- Used in World War II by Germany for encryption and decryption of top secret documents.

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

POLYALPHABETIC SUBSTITUTION

Functioning:

Enigma Machine looks like an ordinary typewriter.

With keys for all letters. When a letter is pressed, a bulb under the letter corresponding to the output is lit.

In between the key and the bulb the wires went through some wheels. The connections between these wheels were random but the same in all the machines.

So when a key is hit, the current goes through these wheel and cause an entirely different letter to be lit up.

At every keystroke, the first wheel turns one time, so that even if the same letter is input again, the result will be a different letter. When the first wheel completes a full turn, the second wheel will turn once. When it completes its turn, the third wheel will turn once and so on. (up to 4-16 wheels).

A wheel should not have to start at the 'A' letter. It could start at any letter. This position was called the key and it was extremely necessary for the correct encryption and decryption of the message. This key was changed every day and generals who where to use this machine where given books to find out which key should be used in a particular day.

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

OUTLINE

- 2. Classic cryptography
  - Introduction
  - Classification
  - Transposition methods
    - By groups
    - By series
    - By columns/ rows
  - Substitution methods
    - Monoalphabetic methods
      - Monographic
      - Polygraphic
    - Polyalfabethical methods
  - Cryptanalysis – classic cryptography

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- Breaking shifting ciphers

Going through all possible of keys:

Brute force

$$E_n(x) = (x + n) \bmod 26$$

$$D_n(x) = (x - n) \bmod 26$$

EXAMPLE: CIPHER TEXT: HAAH JRHA KHDU

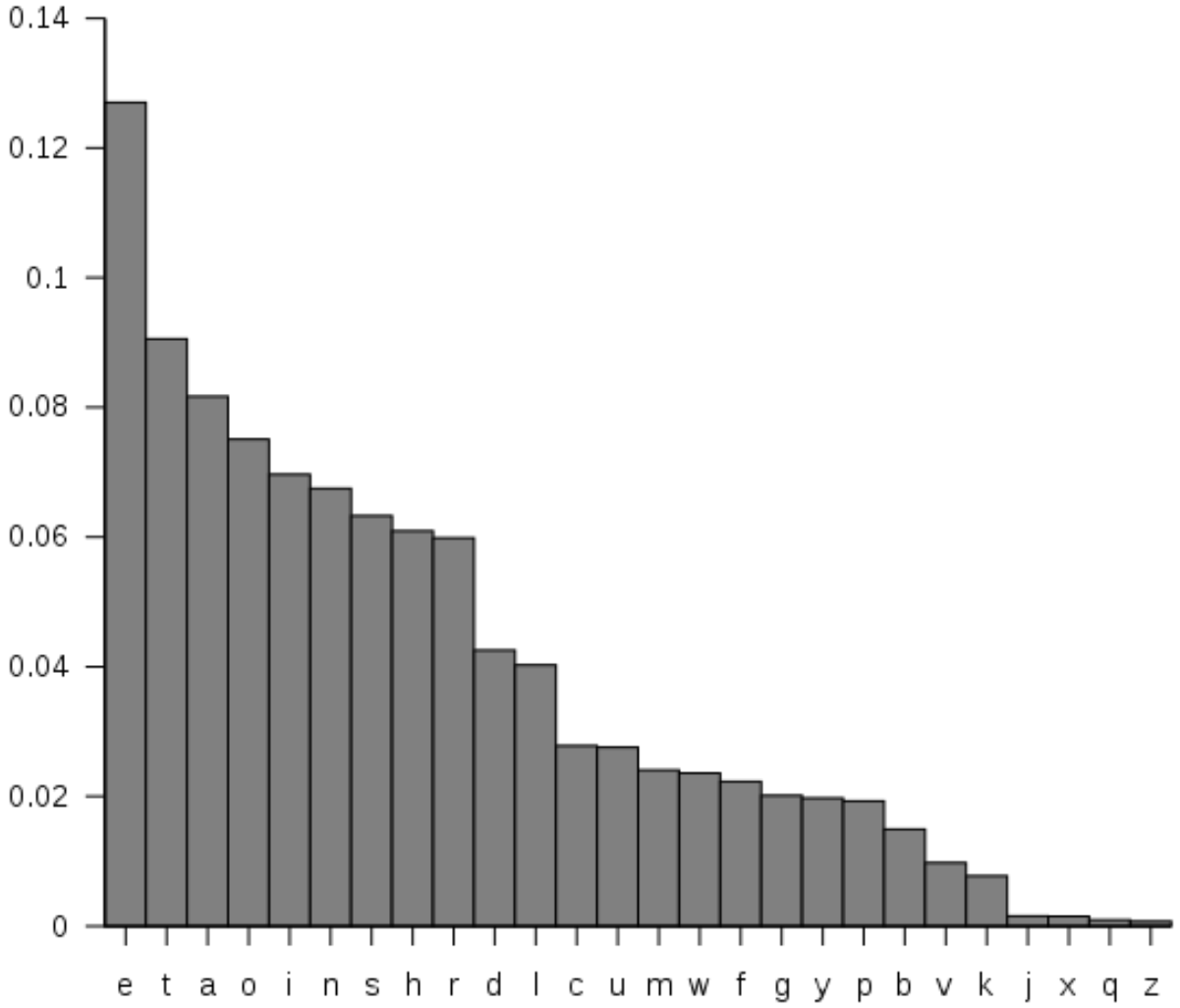
Shift: 0: haah jrha khdu
Shift: 1: ibbi ksib liev
Shift: 2: jccj ltjc mjfw
Shift: 3: kddk mukd nkgx
Shift: 4: leel nvle olhy
Shift: 5: mff mowmf pmiz
Shift: 6: nggn pxng qnja
Shift: 7: ohho qyoh rokb
Shift: 8: piip rzpi splc
Shift: 9: qjjq saqj tqmd
Shift: 10: rkkr tbrk urne
Shift: 11: slls ucsl vsof
Shift: 12: tmmt vdtm wtpg
Shift: 13: unnu weun xuqh
Shift: 14: voov xfvo yvri
Shift: 15: wppw ygwp zwsj
Shift: 16: xqqx zhxq axtk
Shift: 17: yrry aiyr byul
Shift: 18: zssz bjzs czvm
Shift: 19: atta ckat dawn
Shift: 20: buub dlbu ebxo
Shift: 21: cvvc emcv fcyp
Shift: 22: dwwd fndw gdzq
Shift: 23: exxe goex hear
Shift: 24: fyyf hpfy ifbs
Shift: 25: gzzg iqgz jgct

So the decryption key is 19, the encryption key is therefore 7.

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- Frequency Analysis



CRYPTANALYSIS

- Frequency Analysis

English also has a number of common letter patterns
that we can also use to help decrypt monoalphabetic ciphers:

Common pairs	TH, EA, OF, TO, IN, IT, IS, BE, AS, AT, SO, WE, HE, BY, OR, ON, DO, IF, ME, MY, UP
Common repeated letters	SS, EE, TT, FF, LL, MM and OO
Common triplets	THE, EST, FOR, AND, HIS, ENT or THA

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- Breaking monoalphabetic substitution ciphers

  Affine or linear encryption:

$$E(m_i) = (a\ m_i\ + b\ ) \bmod . n$$

- Frequency analysis attack

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- Example:

QWMMPQDVKUVFDTXJQVDBOPIDUHDQQUGDLAMWJGXBGURRB
PBURMKULDVXOOKUJUOVDJQDGBWHLDJQQMUODQUBIMWBOV
WUVXPBUBIOKUBGXBGURROKUJUOVDJQVPWMMDJOUQDVKDBV
KDCDAQXEDFKXOKLPWBIQVKDQDOWJXVAPTVKDQAQVDHXQU
RMKULDVXOOKUJUOVDJQVKDJDTPJDVKDVPVURBWHLDJPTCDAQ
XQPTDBPJHPWQQXEDBDNDJVKDRDQQFDFXRRQDDVKUVQXHM
RDQWLQVXVWVXPBXQNDJAQWQODMVXLRDVPOJAMVUBURAVX
OUVVUOCQ

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

We find the following character counts:

Nº	Letter	Frequency (%)	Frequency
1	**D**	**13.1498**	**43**
2	**V**	**10.0917**	**33**
3	Q	9.7859	32
4	U	8.2569	27
5	O	5.8104	19
6	X	5.8104	19
7	B	5.5046	18
8	J	5.5046	18
9	K	5.1988	17
10	P	4.5872	15
11	R	3.9755	13
12	W	3.9755	13
13	M	3.6697	12
14	A	2.4465	8
15	L	2.4465	8
16	G	1.8349	6

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

We guess that the encryption takes
E→D and T→V

Supposition: E -> D, T -> V

$$3 = a \cdot 4 + b \bmod 26$$

$$21 = a \cdot 19 + b \bmod 26$$

Then:

b= 3-4a and,

21=19a-4a+3=15a+3 mod 26

Then:

15a=18mod26; gcd(15,26)=1, applying Euler, a=22

So:

$$b = 19$$

$$a = 22$$

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

Cleartext

THE CALL OF DEATH IS A CALL OF LOVE. DEATH CAN BE SWEET IF WE ANSWER IT IN THE AFFIRMATIVE, IF WE ACCEPT IT AS ONE OF THE GREAT ETERNAL FORMS OF LIFE AND TRANSFORMATION.

CRYPTANALYSIS

- Monoalphabetic substitution:
  - Any particular letter in the plaintext will always be transformed into the same letter in the ciphertext
  - The length of the chipertext must be sufficiently large to study frequencies
  - Pure shiffting: (a=1)

    Caesar type
  - Affine transformation: (*a >1* and *b>0)*

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- Polyalphabetic substitution:

KASISKI METHOD

Example:

Message	THES UNAN DTHE MANI NTHE MOON
Key	KING KING KING KING KING KING
Ciphertext	DPRY EVNT NBUK WIAO XBUK WWBT

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- KASISKI method

Steps:

- Look for repeated strings of characters (eg.: BUK)
- Compute the distance between them which is a multiple of the length of the key

  (eg-: 8 = 2*4)

  →L = Length of the key= g.c.d (of all distances)
- Divide the ciphertext into subblocks of length L:

  DPRY
  EVNT
  NBUK
  WIAO
  XBUK
  WWBT
- Frequency analysis of each group of letters which has been encrypted using the same alphabet

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- KASISKI method. Example

```
OOEXQGHXINMFRTRIFSSMZRWLYOWTTWTJIWMOBEDAXHVHSFTRIQKMENXZ
PNQWMCVEJTWJTOHTJXWYIFPSVIWEMNUVWHMCXZTCLFSCVNDLWTENUHSY
KVCTGMGYXSYELVAVLTZRVHRUHAGICKIVAHORLWSUNLGZECLSSSWJLSKO
GWVDXHDECLBBMYWXHFAOVUVHLWCSYEVVWCJGGQFFVEOAZTQHLONXGAHO
GDTERUEQDIDLLWCMLGZJLOEJTVLZKZAWRIFISUEWWLIXKWNISKLQZHKH
WHLIEIKZORSOLSUCHAZAIQACIEPIKIELPWHWEUQSKELCDDSKZRYVNDLW
TMNKLWSIFMFVHAPAZLNSRVTEDEMYOTDLQUEIIMEWEBJWRXSYEVLTRVGJ
KHYISCYCPWTTOEWANHDPWHWEPIKKODLKIEYRPDKAIWSGINZKZASDSKTI
TZPDPSOILWIERRVUIQLLHFRZKZADKCKLLEEHJLAWWVDWHFALOEOQW
```

- Step 1, Look for repetitions:

SYE: 122, 196, 383
ZKZA: 252, 439, 472

196−122 = 2·37
383−196 = 11·17
383−122 = 9·29
439−252 = 11·17
472−439 = 3·11
472−252 = 22·5·11

- Step 2: Look for the most repeated divisor 11

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- KASISKI method
  - Step 3: subcryptograms. Estimated key length=11

```
1   OOEXQGHXINM FRTRIFSSMZR WLYOWTTWTJI WMOBEDAXHVH
2   SFTRIQKMENX ZPNQWMCVEJT WJTOHTJXWYI FPSVIWEMNUV
3   WHMCXZTCLFS CVNDLWTENUH SYKVCTGMGYX SYELVAVLTZR
4   VHRUHAGICKI VAHORLWSUNL GZECLSSSWJL SKOGWVDXHDE
5   CLBBMYWXHFA OVUVHLWCSYE VVWCJGGQFFV EOAZTQHLONX
6   GAHOGDTERUE QDIDLLWCMLG ZJLOEJTVLZK ZAWRIFISUEW
7   WLIXKWNISKL QZHKHWHLIEI KZORSOLSUCH AZAIQACIEPI
8   KIELPWHWEUQ SKELCDDSKZR YVNDLWTMNKL WSIFMFVHAPA
9   ZLNSRVTEDEM YOTDLQUEIIM EWEBJWRXSYE VLTRVGJKHYI
10  SCYCPWTTOEW ANHDPWHWEPI KKODLKIEYRP DKAIWSGINZK
11  ZASDSKTITZP DPSOILWIERR VUIQLLHFRZK ZADKCKLLEEH
12  JLAWWVDWHFA LOEOQW
```


Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- KASISKI method
  - Step 4: Frequency analysis in each cryptogram.
  - We assume letter E to be substituted into the most frequent character of each group

subcryptogram	Most frequent characters
1	S,W,Z,V
2	L,A
3	F,T
4	D,O, R
5	L,I,W
6	W,L
7	T,H,W
8	I,S,X,E
9	E,H
10	Z,E,Y
11	I

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS

- KASISKI.

The partial key
I****IL*A*W gives the following text which is suggestive of English:

```
1   W****OS*I*I N****ND*M*N E****BE*T*E E****LL*H*D
2   A****YV*E*T H****UN*E*P E****BU*W*E N****EP*N*R
3   E****HE*L*O K****EE*N*D A****BR*G*T A****IG*T*N
4   D****IR*C*E D****TH*U*H O****AD*W*H A****DO*H*A
5   K****GH*H*W W****TH*S*A D****OR*F*R M****YS*O*T
6   O****LE*R*A Y****TH*M*C H****RE*L*G H****NT*U*S
7   E****EY*S*H Y****ES*I*E S****WW*U*D I****IN*E*E
8   S****ES*E*M A****LO*K*N G****EE*N*H E****NG*A*W
9   H****DE*D*I G****YF*I*I M****EC*S*A D****OU*H*E
10 A****EE*O*S I****ES*E*E S****ST*Y*L L****AR*N*G
11 H****SE*T*L L****TH*E*N D****TS*R*G H****SW*E*D
12 R****DO*H*W T****E
```

Do you want to finish the puzzle?

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

CRYPTANALYSIS TECHNIQUES

Remember:

- **Cryptanalysis of monoalphabetic substitution**
  - Language redundancy
  - Frequency analysis
  - Affine or linear substitution

- **Cryptanalysis of periodic polylphabetic substitution**
  - Determine the number of alphabets, i.e. the length of the encryption key.
  - Use Kasiski method to determine the length of the key:
    - Look for repeated strings of characters
    - The length of the key can be estimated computing the g.c.d of the distances between repetitions.
  - Create subcryptograms with those characters which have been encrypted using the same alphabet. Analyze letter frequencies in those subcryptograms.

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE

- IC is a statistical method for determining whether a cipher is monoalphabetic or polyalphabetic

- IC was developed by William Friedman (1891 – 1969) and published in The Index of Coincidence and its Applications in Cryptography *(1920)*

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE

- IC is the probability of randomly selecting two letters from a given text such that they are the same letter

E.g.:

26 letters in the alphabet, 2 independent experiments:

→ Prob("choose letter A ")=1/26

→ Prob("Choose letter A twice")=1/26*1/26

→ Prob(Choose any two letters and they coincide")=

$$26*(1/26)*(1/26) = (1/26) = 0.038$$

Then the IC of a random set of letters of the alphabet is: 0.038

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE

The probability of choosing the same two letters from a ciphertext of n characteres (i.e., two As or two Bs or… two Zs) would be:

$$f_A/n^* (f_A-1)/(n-1) + f_B/n^* (f_B-1)/(n-1) +\ldots+ f_Z/n^* (f_Z-1)/(n-1)$$

Where:

f_A is the frequency of the letter A in the ciphertext,

f_B is the frequency of the letter B in the ciphertext,

…

f_Z is the frequency of the letter Z in the ciphertext

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE

Each language has a different IC:

Spanish:	0.0775	English:	0.0667
Russian:	0.0529	German:	0.0762

Computing the IC of a text:

$$\frac{\sum f_i * (f_i - 1)}{N*(N-1)}$$

for $0 \leq i \leq 26$,

f_i frequency of letter I in the text

N is the number of characters in the text

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE

- EXAMPLE:
  - The IC of the text "THE COINCIDENCE INDEX IS" is :

$$\sum f_i(f_i-1)= (4*3)_E+(4*3)_I+(3*2)_C+(3*2)_N+(2*1)_D= 38$$

divided by N*(N-1) = 21*20= 420

IC= 38/420 = 0.090

  - The IC of the ciphertext "KYV TFZETZUVETV ZEUVO ZJ" is :

$$\sum f_i(f_i-1)= (4*3)_V+(4*3)_Z+(3*2)_E+(3*2)_T+(2*1)_U= 38$$

divided by N*(N-1) = 21*20= 420

IC= 38/420 = 0.090

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE ISSUES

IT IS USED:

- IC is a statistical method for determining whether a cipher is monoalphabetic or polyalphabetic
  - A monoalphabetic cipher does NOT alter the IC of the original text
- It is used to estimate the length of the key in the cryptoanalaysis of polyalphabetical ciphers.
- It is used to distinguish ciphertext from plaintext

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE ISSUES, EXAMPLE

- Ciphertext from text in English:

Remember English IC: 0.0667

```
xwlfoqr2meziizwñ5oejxwzzisu7eesst3wñ("etkssxdg5mezg3tj2yrx
kdl@nvyk5lniyr1zkmw1ydgrutqustulgtnymirlntnw0eesviwiqs2nzk
37eys0ñtvrrñwiw{nrgñw1ew0yojwzrjetu37fovzjgiz0legñ8pjv'r2m
ifx0nij4jrdzw1ydñ(ifdy)jymy)qheqr5f2f7ñeou7snqmrñtvf8nfñyr
jrhf}jpok=1ezk@1zwf{vnju@uejxwzzisuqjw'
```


Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE ISSUES, EXAMPLE

Remember English IC: 0.0667

Polyalphabetic

Period	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5
1	0,033				
2	0,032	0,032			
3	0,038	0,032	0,031		
4	0,029	0,031	0,030	0,032	
5	0,051	0,069	0,046	0,085	0,061
6	0,033	0,024	0,035	0,035	0,036
7	0,031	0,036	0,024	0,042	0,032

Estimated length key or period

Ana I. González-Tablas Ferreres, José M. de Fuentes García-Romero de Tejada, Lorena González Manzano, Sergio Pastrana Portillo

INDEX OF COINCIDENDE ISSUES, EXAMPLE

- Clear text:

  Recall that, using frequency analysis, peaks and valleys of frequencies suggest a monoalphabetic cipher and relatively uniform frequencies suggest a polyalphabetic cipher. The Friedman test is a statistical way of "looking for peaks and valleys versus uniform frequencies."

- Key: fried

CRYPTOGRAPHY AND COMPUTER SECURITY

COSEC

uc3m | Universidad Carlos III de Madrid