IN410
Cryptography and Secure Communications

Ahmad Fadlallah

Credits

- The following slides are the adaptation of
  - Lecture slides of Lawrie Brown based on William Stallings book “Cryptography and Network Security: Principles and Practice”
  - Lecture Slides of Dan Boneh – Stanford University
  - Lecture slides of Pawel Wocjan – University of Central California
  - Lecture Slides of Ahmed Serhrouchni – Telecom ParisTech
  - Others (References in the note section)

Course Syllabus

- Context
- **History/ Classical ciphers ←**
- Symmetric Encryption
- Asymmetric Encryption
- Cryptographic hash functions
- Key Distribution
- Security Protocols

Learning objectives

- Topics to be covered in this lecture:
  - Classical Ciphers
    - Principles and Basic terminology
    - Substitution ciphers
    - Transposition ciphers
    - Rotor Machines

Classical Ciphers

Classical Ciphers

Principles and Basic terminology

Substitution Ciphers

Transposition Ciphers

Rotor Machines

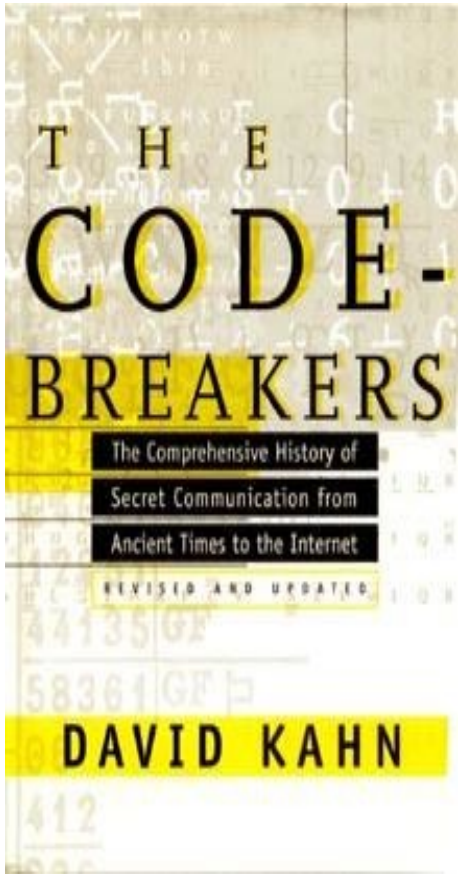
Good Reads

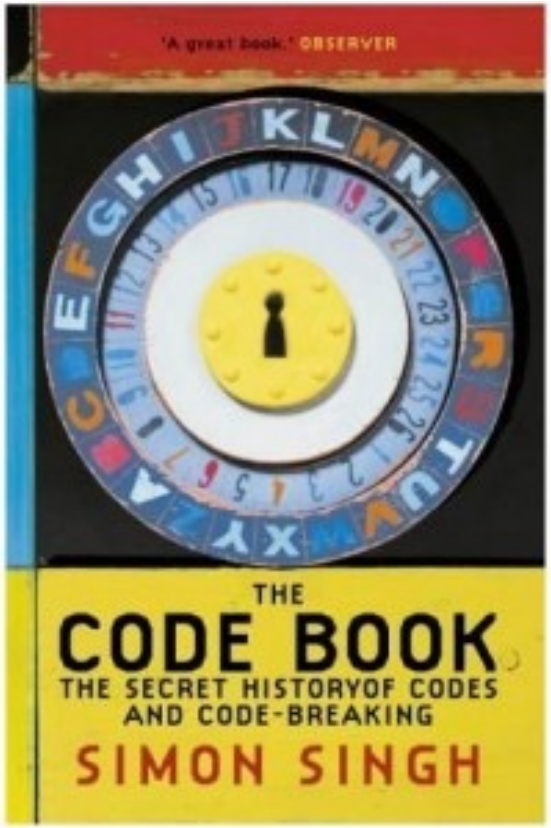
"The code breakers"

David Kahn

"The codebook"

Simon Singh





Basic Terminology

- **Plaintext** : the original message
- **Ciphertext**: the coded message
- **Encryption**: Process of converting from plaintext to ciphertext
- **Decryption**: Restoring the plaintext from the ciphertext
- **Cryptography**: Study of encryption

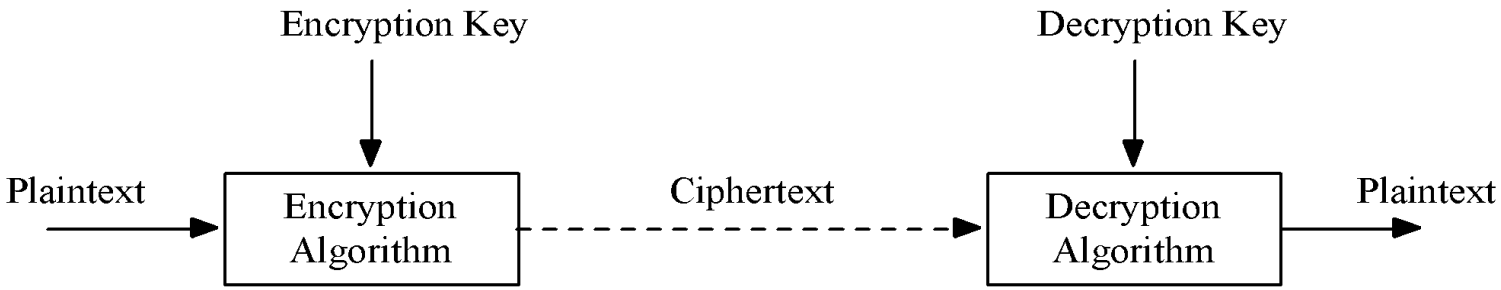
Basic Terminology (cont'd)

- **Cryptanalysis**: Techniques used for decrypting a message without any knowledge of the encryption details

- **Cryptology**: Areas of cryptography and cryptanalysis together

Basic Terminology (cont'd)

- Cryptographic system (Cryptosystem)
  - Schemes used for encryption/decryption
  - Five-tuple (P, C, K, <E, D>)
- Key
  - Secret associated to the cryptosystem
  - Often of fixed size independent from the plaintext size



Types of cryptosystems

- **Restricted-usage cryptosystem**
  - Security is based on the Encryption/Decryption operations
  - All encrypted messages are revealed once the process is revealed
- **General usage cryptosystem**
  - Security is based on the the secrecy of the key
  - All encrypted messages with a certain key are revealed once the key is revealed

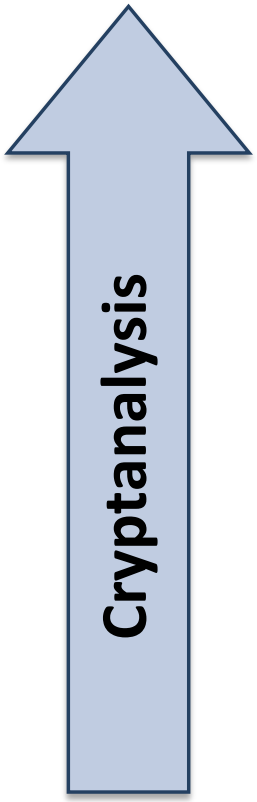
Which one is safer?

Kerckhoffs principles

- Auguste Kerckhoffs (1835-1903)
  - Born in Holland and taught in France
  - Lays the foundations of modern cryptography in a military newspaper in 1883 (“la cryptographie militaire”)

The security of a cryptosystem depends on the secrecy of the key

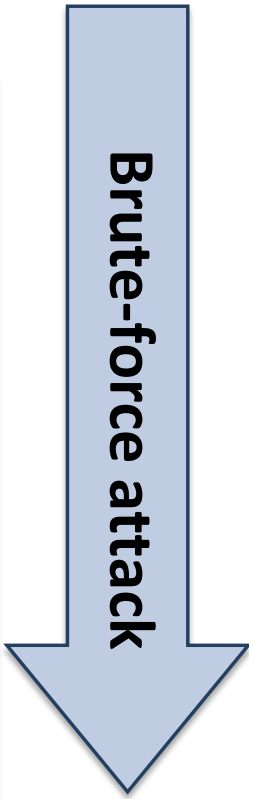
Cryptanalysis vs. Brute-Force Attack



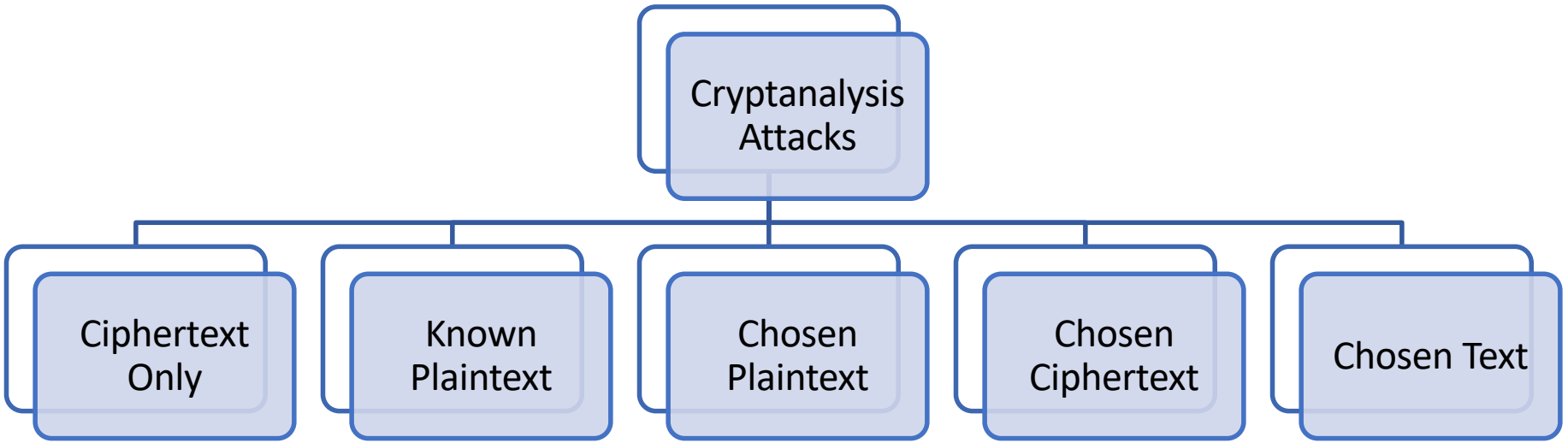
Attack relies on the **nature of the algorithm** plus some **knowledge of the general characteristics of the plaintext and/or ciphertext**

Attack exploits the **characteristics of the algorithm** to attempt to deduce a specific plaintext or to deduce the key being used

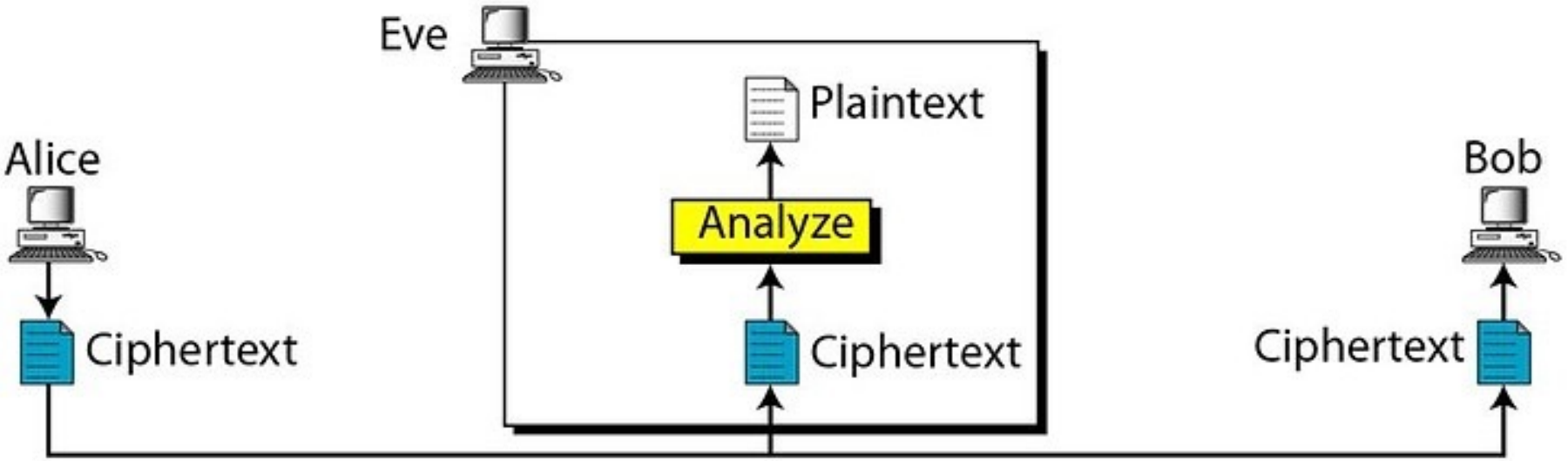
Attacker **tries every possible key** on a piece of ciphertext until an intelligible translation into plaintext is obtained



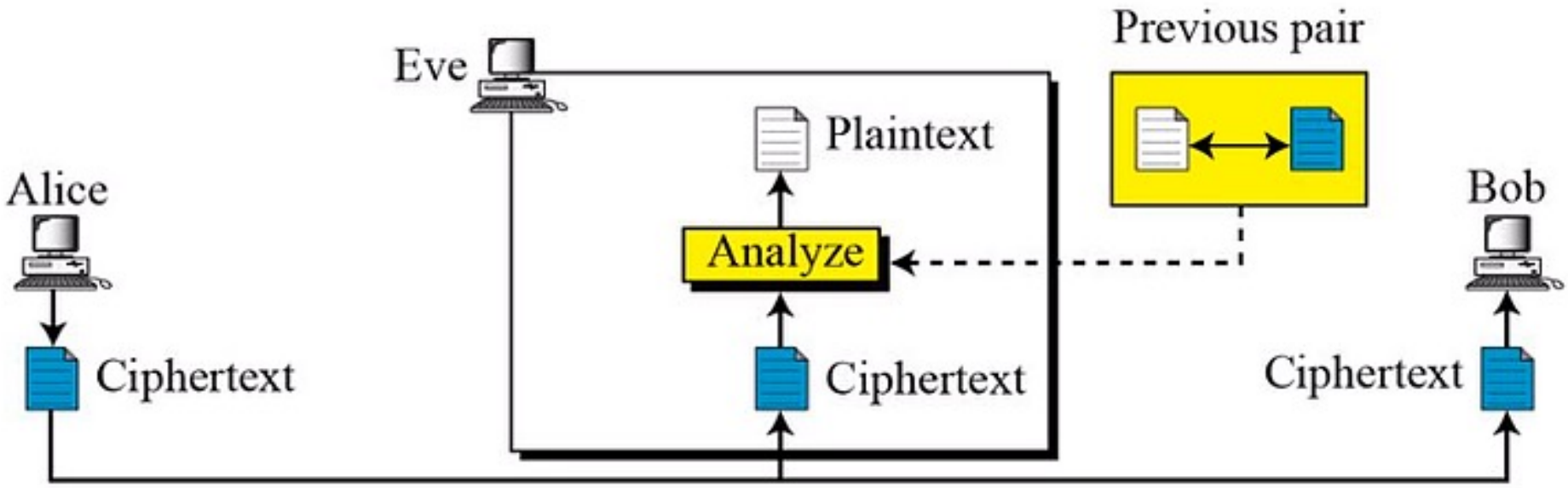
Cryptanalysis Attacks



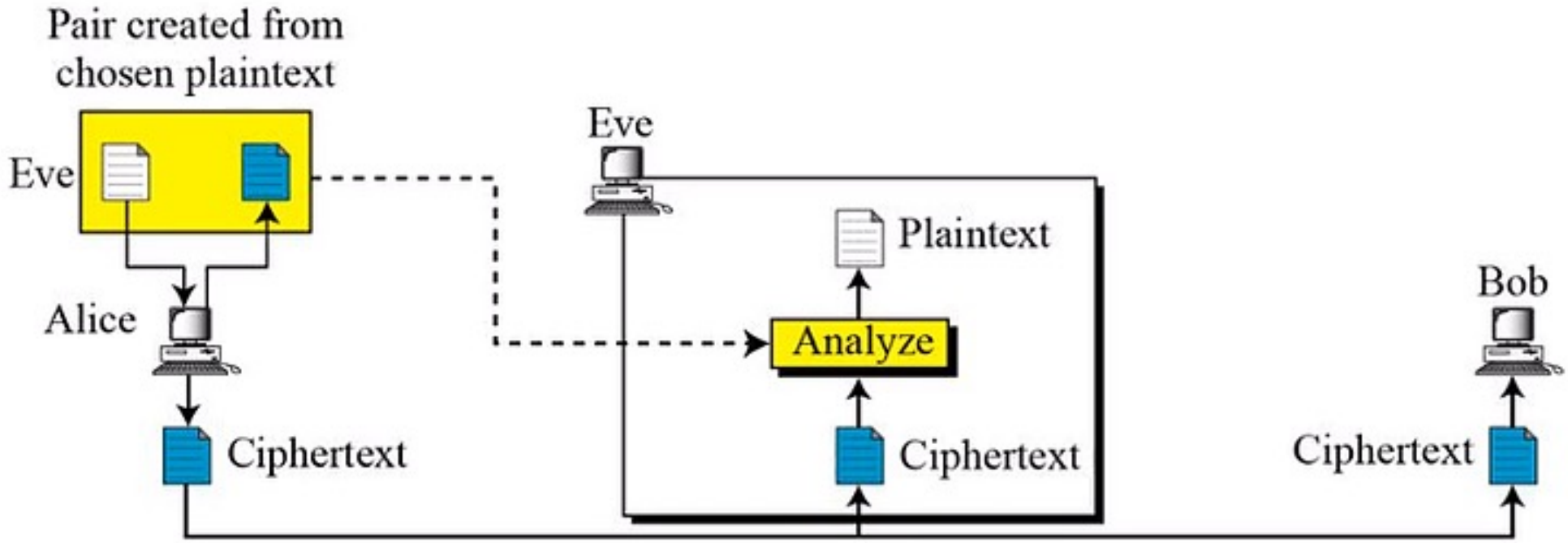
Cryptanalysis attacks
Ciphertext Only



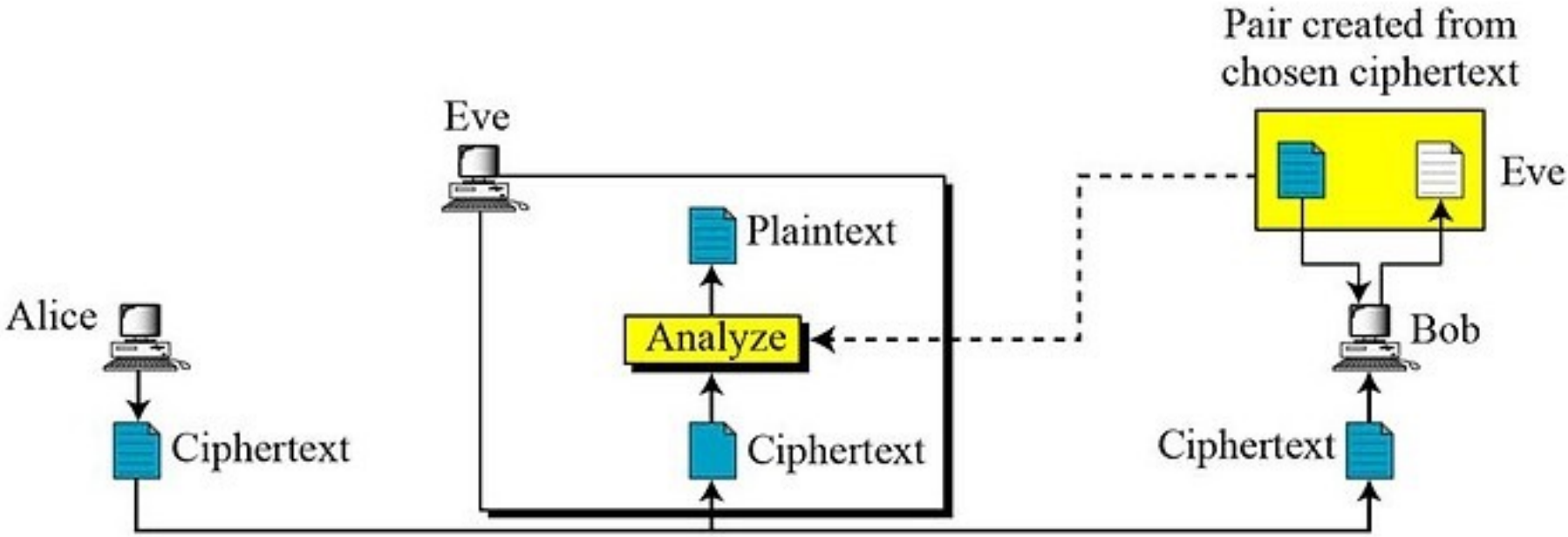
Cryptanalysis attacks
Known Plaintext Attacks



Cryptanalysis attacks
Chosen Plaintext Attacks



Cryptanalysis attacks
Chosen Ciphertext Attacks



Cryptanalysis attacks

Type of Attack	Known to Cryptanalyst
Ciphertext Only	• Encryption algorithm • Ciphertext
Known Plaintext	• Encryption algorithm • Ciphertext • One or more plaintext-ciphertext pairs formed with the secret key
Chosen Plaintext	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen Ciphertext	• Encryption algorithm • Ciphertext • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen Text	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

Brute-Force Attack

Involves trying every possible key until an intelligible translation of the ciphertext into plaintext is obtained

On average, half of all possible keys must be tried to achieve success

To supplement the brute-force approach, some degree of knowledge about the expected plaintext is needed, and some means of automatically distinguishing plaintext from garble is also needed

Exercise

The computing power of a certain machine is the number of instructions it can execute per unit of time. The computing power of a given personal computer (Intel Core i9-9900K at 4.7 GHz) is around 412,000 MIPS where MIPS stand for Millions Instructions per Second.

An optimized algorithm to verify **one**128-bit AES key **needs around 1200 elementary instructions**. Suppose that we have a pair of clear text and encrypted text using AES and that we want to find the encryption key using brute force attack; which means by testing all the keys once after another. We suppose that all **keys are equally probable**.

What time it takes to find the key using a brute force attack?

Encryption Scheme Security

- **Unconditionally secure**
  - No matter how much time an opponent has, it is impossible for him to decrypt the ciphertext simply because the required information is not there
  - **No encryption algorithm in unconditionally secure**
- **Computationally secure**
  - The cost of breaking the cipher **exceeds** the value of the encrypted information
  - The time required to break the cipher **exceeds** the useful lifetime of the information

Cryptographic Systems

Characterized along three independent dimensions

Type of operations used for transforming plaintext to Ciphertext

- Substitution
- Transposition/ Permutation

Number of keys used

- Symmetric, single-key, secret-key, conventional encryption
- Asymmetric, two-key, or public-key encryption

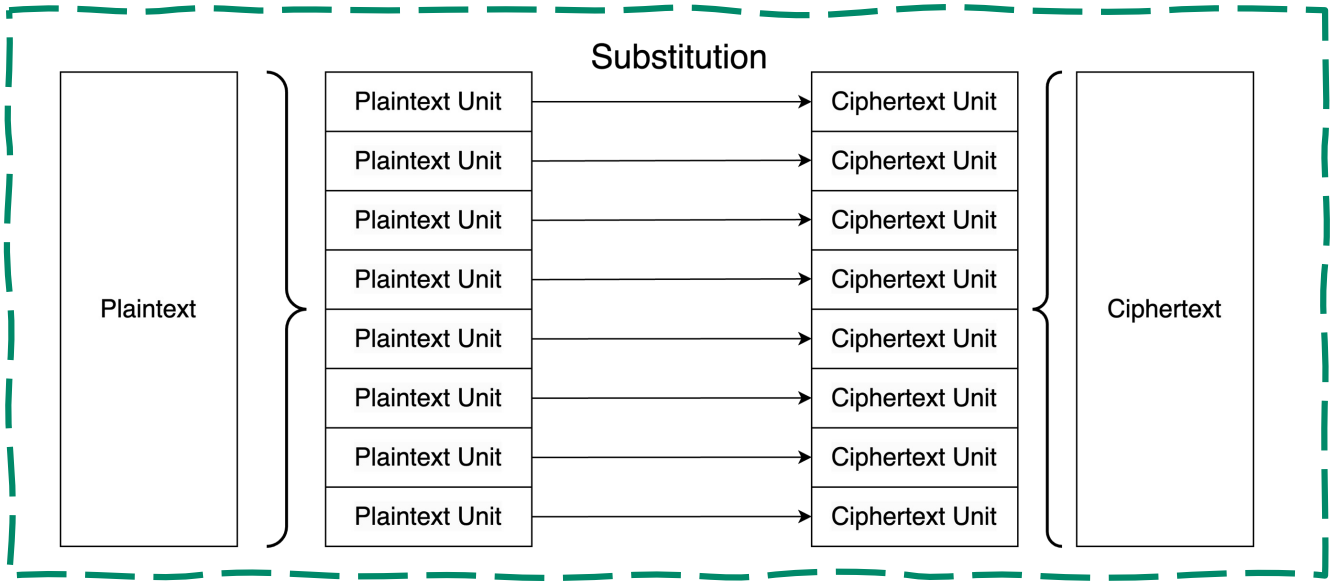
The way in which the plaintext is processed

- Block cipher
- Stream cipher

Classical Ciphers

- Principles and Basic terminology
- **Substitution Ciphers**
- Transposition Ciphers
- Rotor Machines

Substitution Technique



- A **substitution cipher** is a method of encryption by which **units of plaintext** are replaced with **ciphertext units**, according to a regular system
  - "Units" may be single letters, pairs of letters, triplets of letters...
  - Ciphertext unit may not be of the same plaintext alphabet
- The receiver decrypts the text by performing an **inverse substitution**.

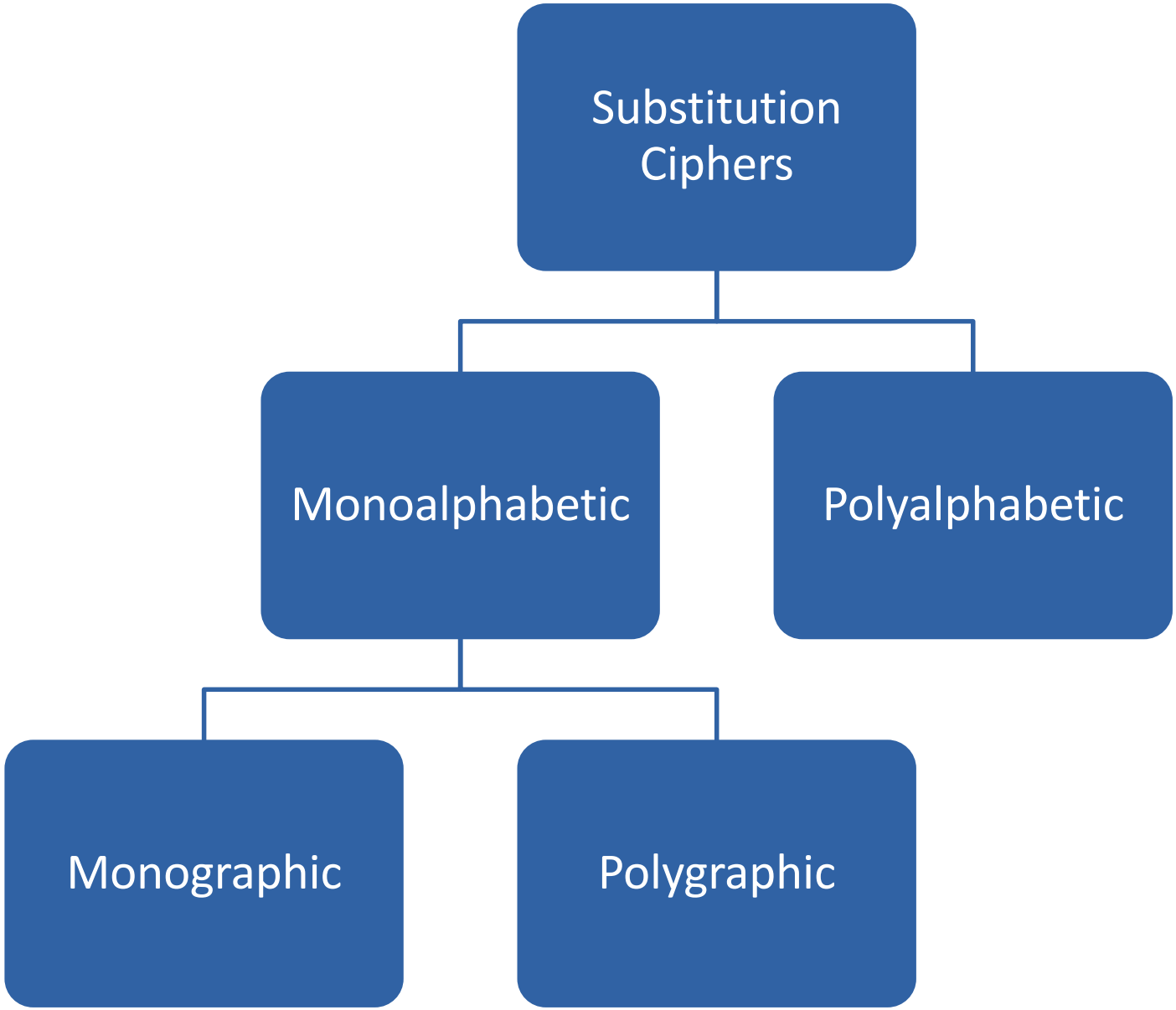
Substitution Ciphers

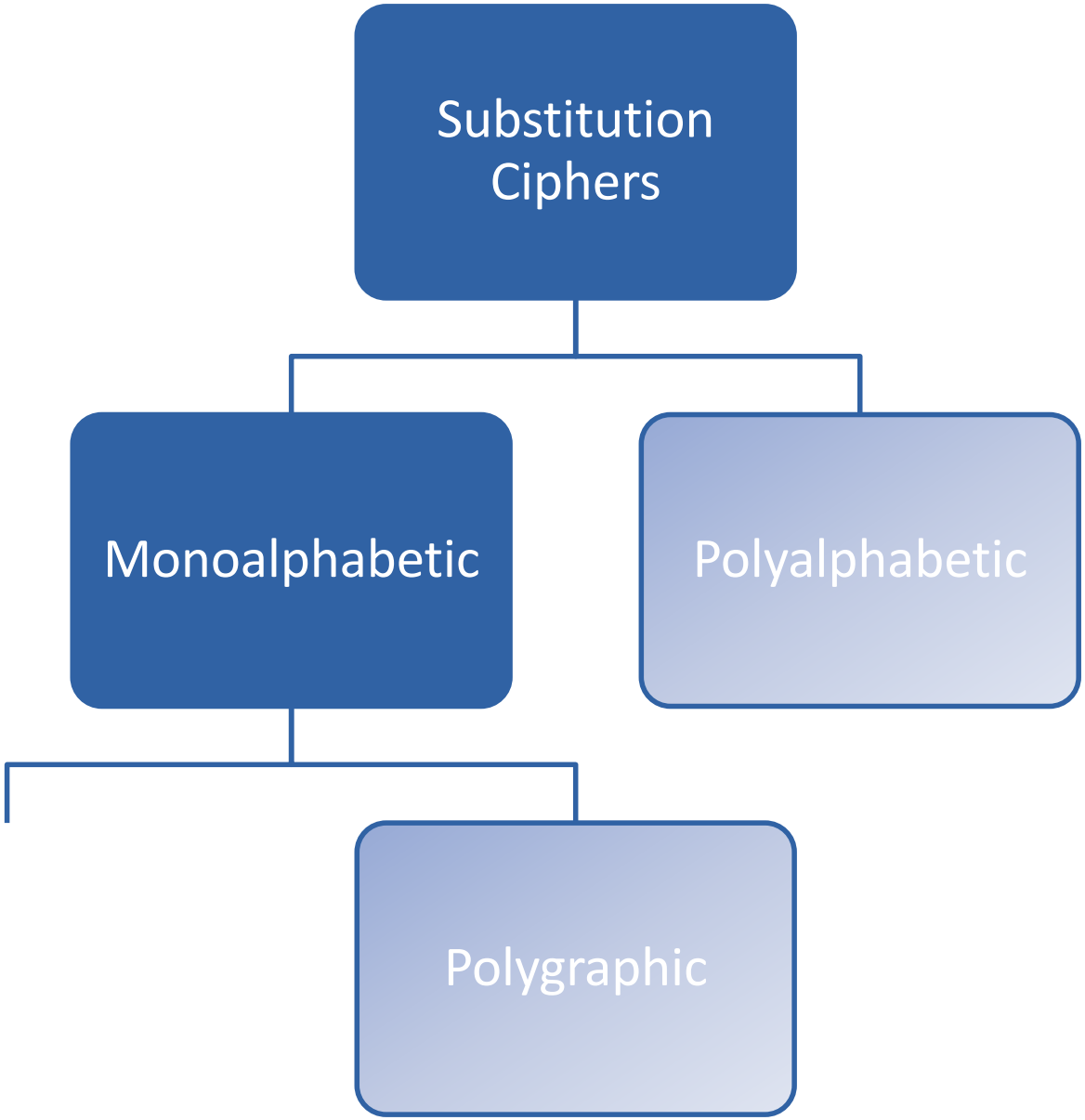
Mono-alphabetic cipher

- Uses fixed substitution over the entire message

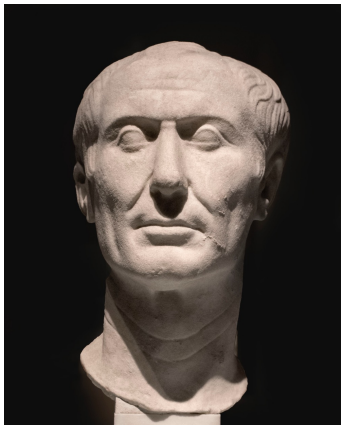
Poly-alphabetic cipher

- Uses a number of substitutions at different positions in the message.
- A unit from the plaintext is mapped to **one of several possibilities** in the ciphertext and vice versa.


Substitution Ciphers
Monoalphabetic
Polyalphabetic
Monographic
Polygraphic


Substitution Ciphers
Monoalphabetic
Polyalphabetic
Polygraphic

Caesar Cipher



- Simplest and earliest known substitution cipher
- Used by **Julius Caesar**
- Involves replacing each letter of the alphabet with the letter standing three places further down the alphabet
- Alphabet is wrapped around so that the letter following Z is A

Plaintext	c	r	y	p	t	o	g	r	a	p	h	y
Ciphertext	F	U	B	S	W	R	J	U	D	S	K	B

Caesar Cipher Algorithm

- **Tabular representation**

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
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

- **Mathematical model**
  - Give each letter a number (0 - 25)
  - Algorithm can be expressed as: $C = E(p) = (p + 3) \; mod \; 26$
- A shift may be of any amount, so that the *general Caesar algorithm* is: $C = E(k, p) = (p + k) \; mod \; 26$
  - k takes on a value in the range 1 to 25
- Decryption algorithm : $p = D(k, C) = (C - k) \; mod \; 26$

Exercise

- Encrypt the plaintext “crypto” using generalized Caesar cipher with a key 10.

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

Exercise - Solution

- Plaintext: “crypto” , K = 10
- C = P + 10 mod 26

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

PT Char.	C	R	Y	P	T	O
PT #	2	17	24	15	19	14
CT #	2 + 10 =12	27%26=1	34%26=8	25	29%26=3	24
CT Char.	M	B	I	Z	D	Y

Polybius square

- Also known as the Polybius checkerboard
- Invented by the Ancient Greek historian and scholar Polybius (in 150 BC)
- Substitute each line by its position in the table
- Example:
  - L <-> 31
  - CRYPTO <-> 13 42 54 44 34
- Possibility to add symbols
- Possibility to add a key

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

Exercise

- Encrypt the message “university” using Polybius Square

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

Exercise - Solution

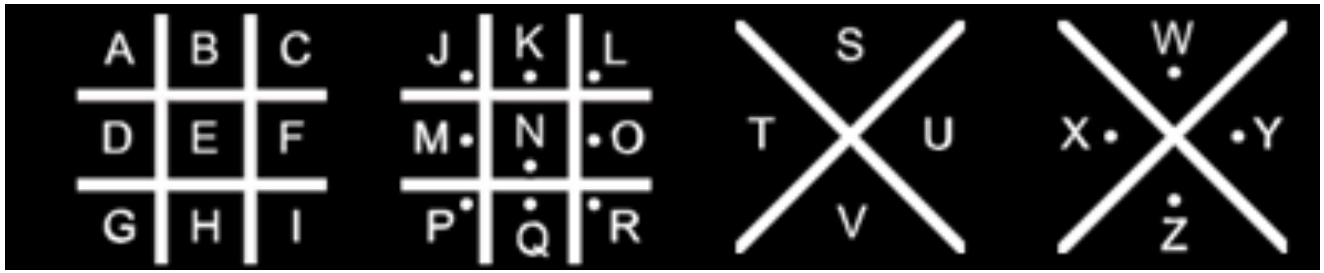
- Encrypt the message “university” using Polybius Square

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

U	15
N	33
I	24
V	51
E	15
R	42
S	43
I	24
T	44
Y	54

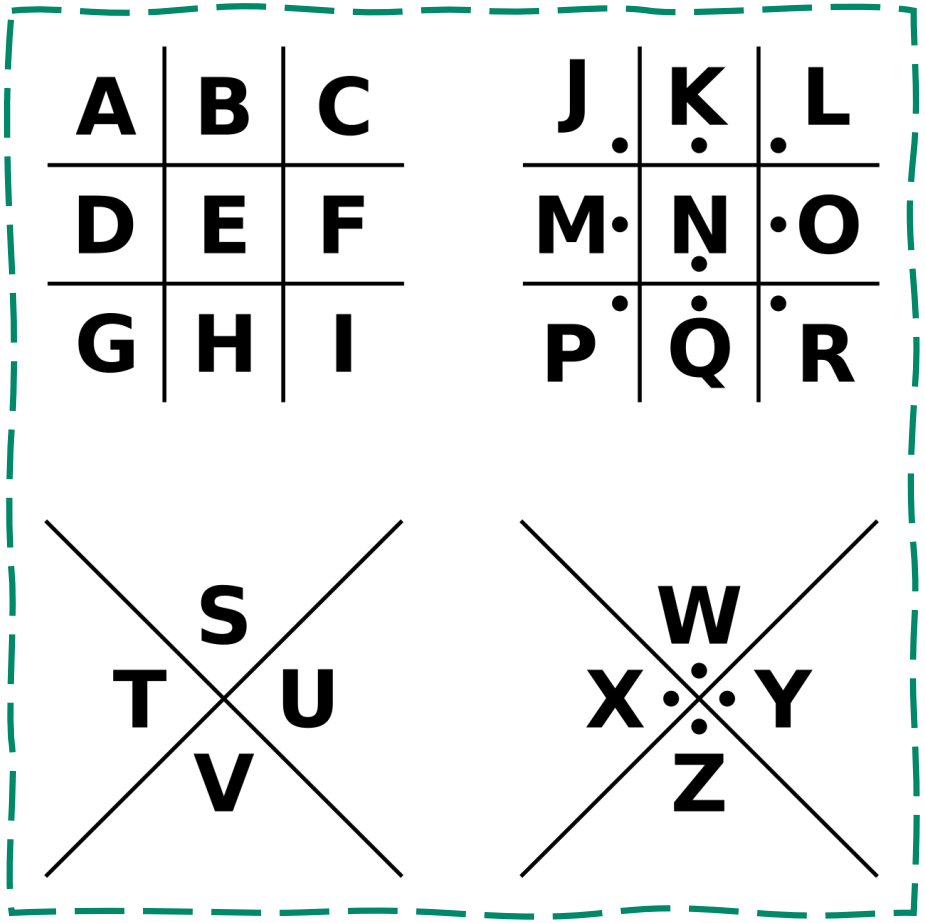
Pigpen Cipher

- Used by Freemasons in the 18th Century to keep their records private
- The cipher does not substitute one letter for another; rather it **substitutes each letter for a symbol**.
- Each letter is enciphered by replacing it with a symbol that corresponds to the portion of the Pigpen grid that contains the letter.



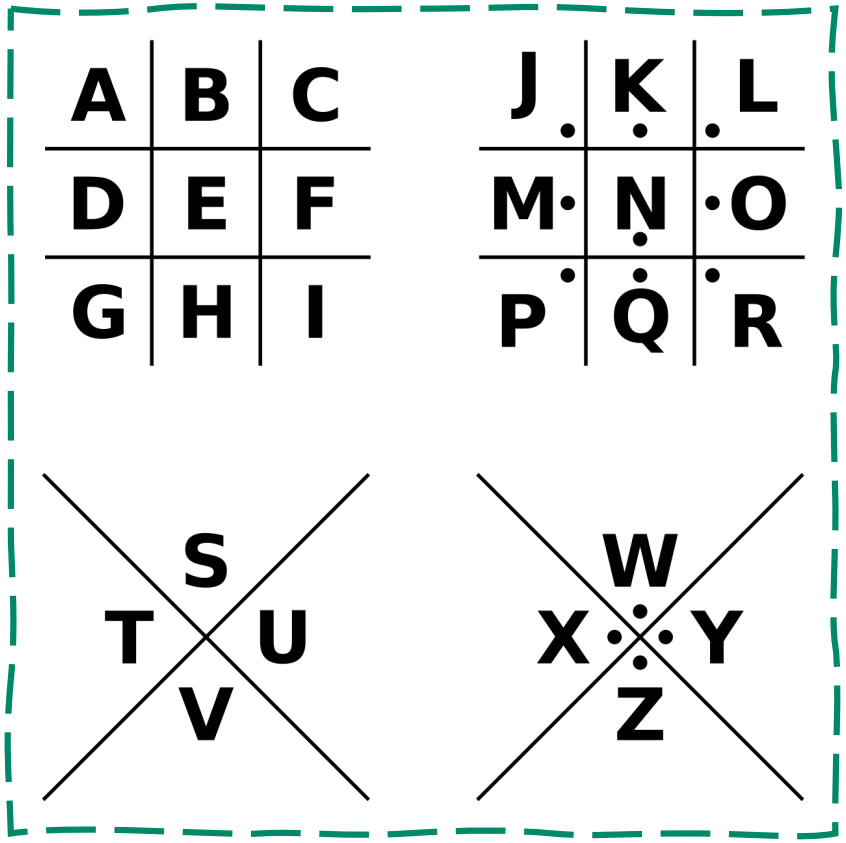
Exercise

Encrypt ”I LOVE CRYPTOGRAPHY” using Pigpen cipher



Exercise - Solution

- PT: ”I LOVE CRYPTOGRAPHY”



- Solution:

Affine Cipher

- $E(x) = (ax + b) mod\ n$
  - n: size of the alphabet
  - (a, b): key of the cipher.
  - **Condition**: a and n must be co-prime.
- $D(x) = a^{-1} \times (x - b) mod\ n$
  - a^{-1} : *multiplicative inverse of* $a (mod\ n)$
  - a^{-1} satisfies the equation: $a \times a^{-1} = 1\ mod\ n$

Exercise

- Let us consider:
  - $E(x) = ax + b \ mod \ 26$, with $a = 9, b = 2$
    - Is 9 a valid choice for parameter “a”? Why?
    - Is b=2 a valid choice? Why?
    - Encrypt the plaintext “affine”
    - Find the decryption function D(y)

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

Monoalphabetic Ciphers

- Plenty of other monoalphabetic substitution codes
- **Word substitution codes**
  - Each word is substituted with a arbitrary chosen word/symbol
  - Need for a dictionary (or code book)
  - Example: “the wind talkers”
  - Problems
    - Intercepting the code book
    - Changing the code book

Breaking Monoalphabetic Ciphers

- Permutation of a finite set of elements S is an ordered sequence of all the elements of S, with each element appearing exactly once
- If the “cipher” line can be any permutation of the 26 alphabetic characters, then there are 26! (27! With the space character) or greater than 4×10^{26} possible keys
- How to break such cipher?

Breaking Monoalphabetic Ciphers

- Property of Monoalphabetic cipher
  - Ciphertext length = Plaintext length
  - Bijective substitution
  - The values of character frequencies are the same between plaintext and ciphertext
- How to cryptanalyze?

Breaking Monoalphabetic Ciphers

- Letter frequency analysis
- Proposed by Al-Kindi in 9th century

Breaking monoalphabetic ciphers

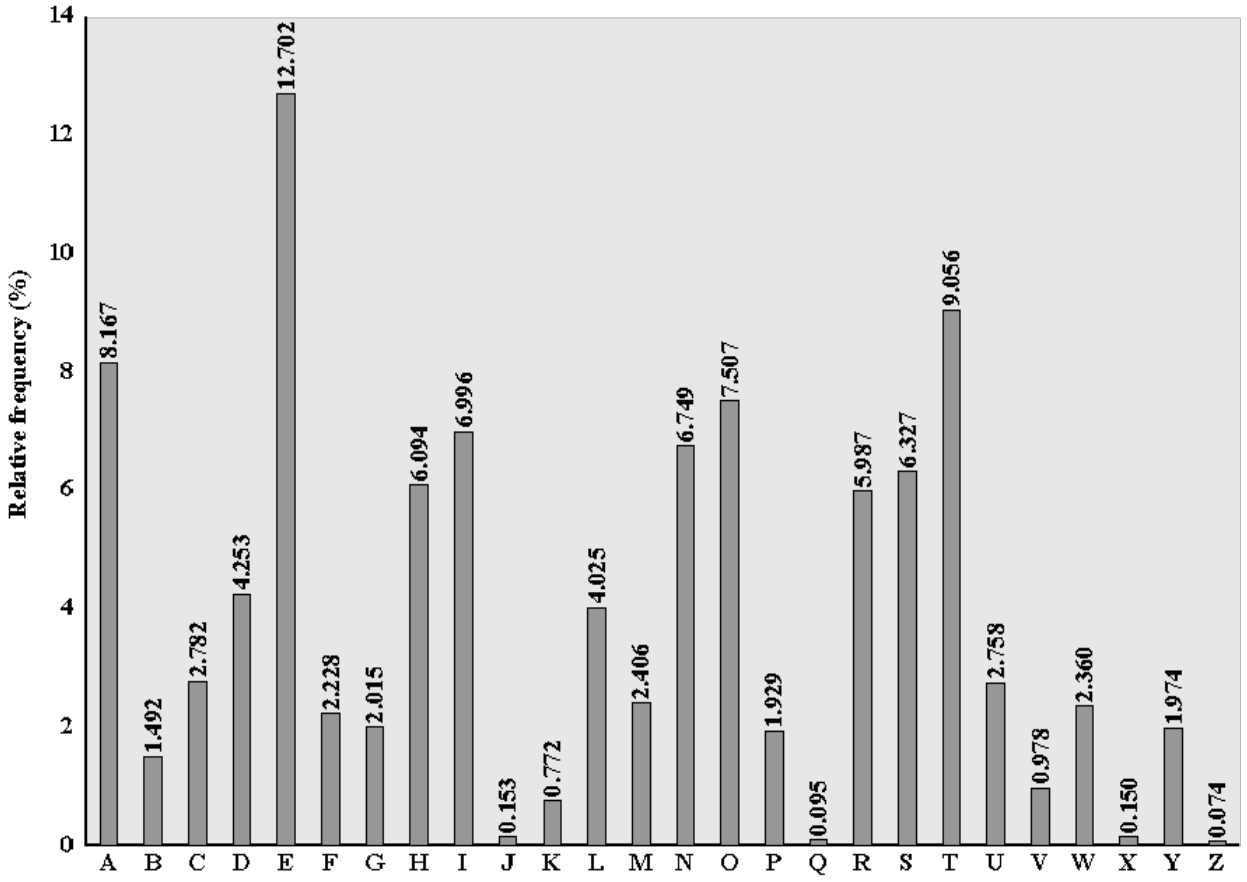
"The Arab scholars invented cryptanalysis, the science of unscrambling a message without knowledge of the key. They cracked the monoalphabetic substitution cipher after several centuries of its successful use. This would not have been possible in a society until it had reached a sufficiently sophisticated level of scholarship in mathematics, statistics, and linguistics."

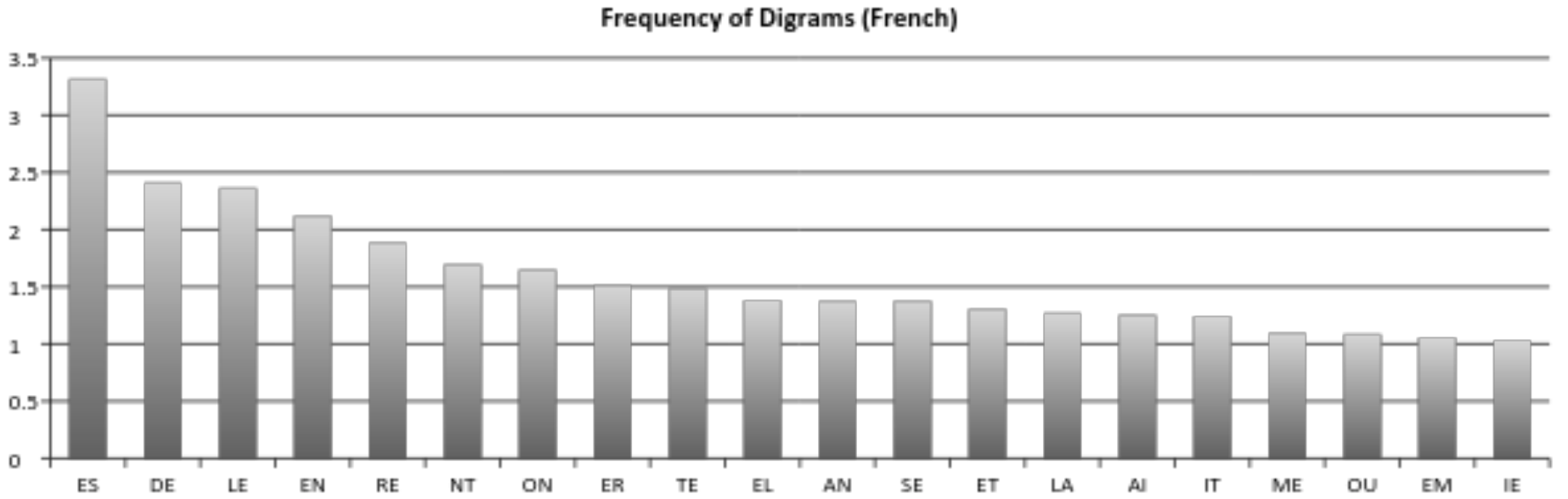
"The innocuous observation that some letters are more common than others in written documents would lead to the first great breakthrough in cryptanalysis. The method, called frequency analysis is described in a treatise by Abu Yusuf Ya'qub ibn Is-haq ibn as-Sabbah ibn 'omran ibn Ismail al-Kindi (let's call him ***al-Kindi*** *for short) in the ninth century."*

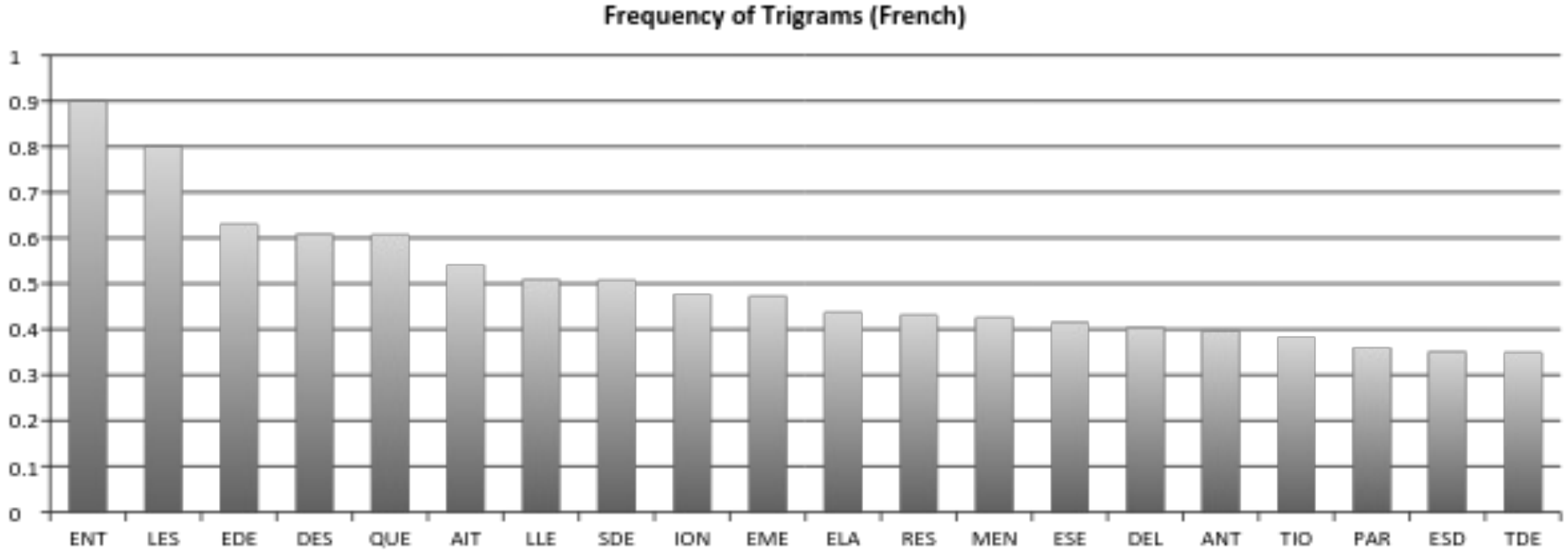
-The codebook – Simon Singh

Breaking monoalphabetic ciphers

- Letter Frequency Analysis (English language)




Frequency of Digrams (French)
3.5
3
2.5
2
1.5
1
0.5
0
ES
DE
LE
EN
RE
NT
ON
ER
TE
EL
AN
SE
ET
LA
AI
IT
ME
OU
EM
IE


Frequency of Trigrams (French)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
ENT
LES
EDE
DES
QUE
AIT
LLE
SDE
ION
EME
ELA
RES
MEN
ESE
DEL
ANT
TIO
PAR
ESD
TDE

Breaking Monoalphabetic Ciphers

- How to break?
  - **Calculate the letter frequency** in ciphertext and **compare them to the corresponding language letter frequency**
  - Use single letter, digram and trigram frequency analysis
  - Apply the results on the ciphertext

Breaking Monoalphabetic Ciphers

- How to break? (Additional Clues)
  - Identify common pairs of letters
  - Identify the smallest words first
  - Tailor made frequency tables
  - Play the guessing game
    - “Al-Khalil, an early Arabian cryptanalyst, demonstrated this talent when he cracked a Greek ciphertext. He guessed that the ciphertext began with the greeting 'In the name of God'. Having established that these letters corresponded to a specific section of ciphertext, he could use them as a crowbar to prise open the rest of the ciphertext. This is known as a crib”

Breaking Monoalphabetic Ciphers

- Test your Skills

https://www.101computing.net/frequency-analysis/

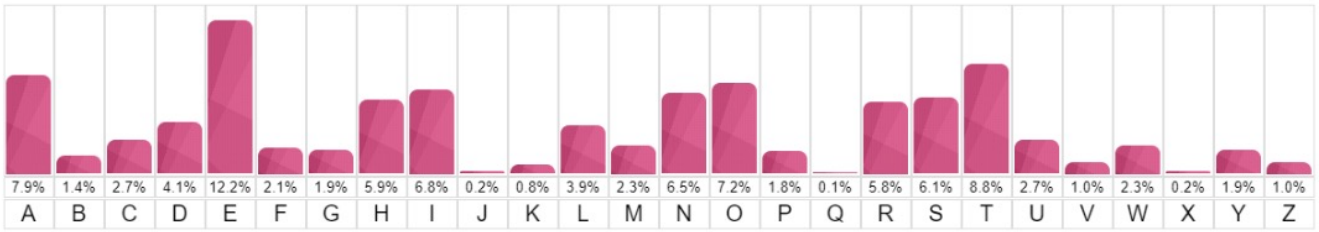
Frequency Analysis

Posted on November 9, 2019 | Posted in Computer Science, Computing Concepts, Cryptography

In cryptography, frequency analysis is the study of the **frequency of letters** or groups of letters in a ciphertext. The method is used as an aid to breaking **substitution ciphers** (e.g. mono-alphabetic substitution cipher, Caesar shift cipher, Vatsyayana cipher).

Frequency analysis consists of **counting the occurrence of each letter** in a text. Frequency analysis is based on the fact that, in any given piece of text, certain letters and combinations of letters occur with varying frequencies. For instance, given a section of English language, letters **E, T, A and O** are the most common, while letters Z, Q and X are not as frequently used.

The following chart shows the frequency of each letter of the alphabet for the English language:



Homophonic Substitution Ciphers

- Objective: **Obscure the letter frequencies**
- Provide multiple substitutes (homophones) for a single letter
- The number of substitutes (for a given letter) can be proportional to the letter frequency
- Homphones assignment can be in rotation or randomly

Homophonic Substitution Ciphers

- Example

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	X	S	F	Z	E	H	C	V	I	T	P	G	A	Q	L	K	J	R	U	O	W	M	Y	B	N
9				7				3					5	0				4	6						
				2																					
				1																					

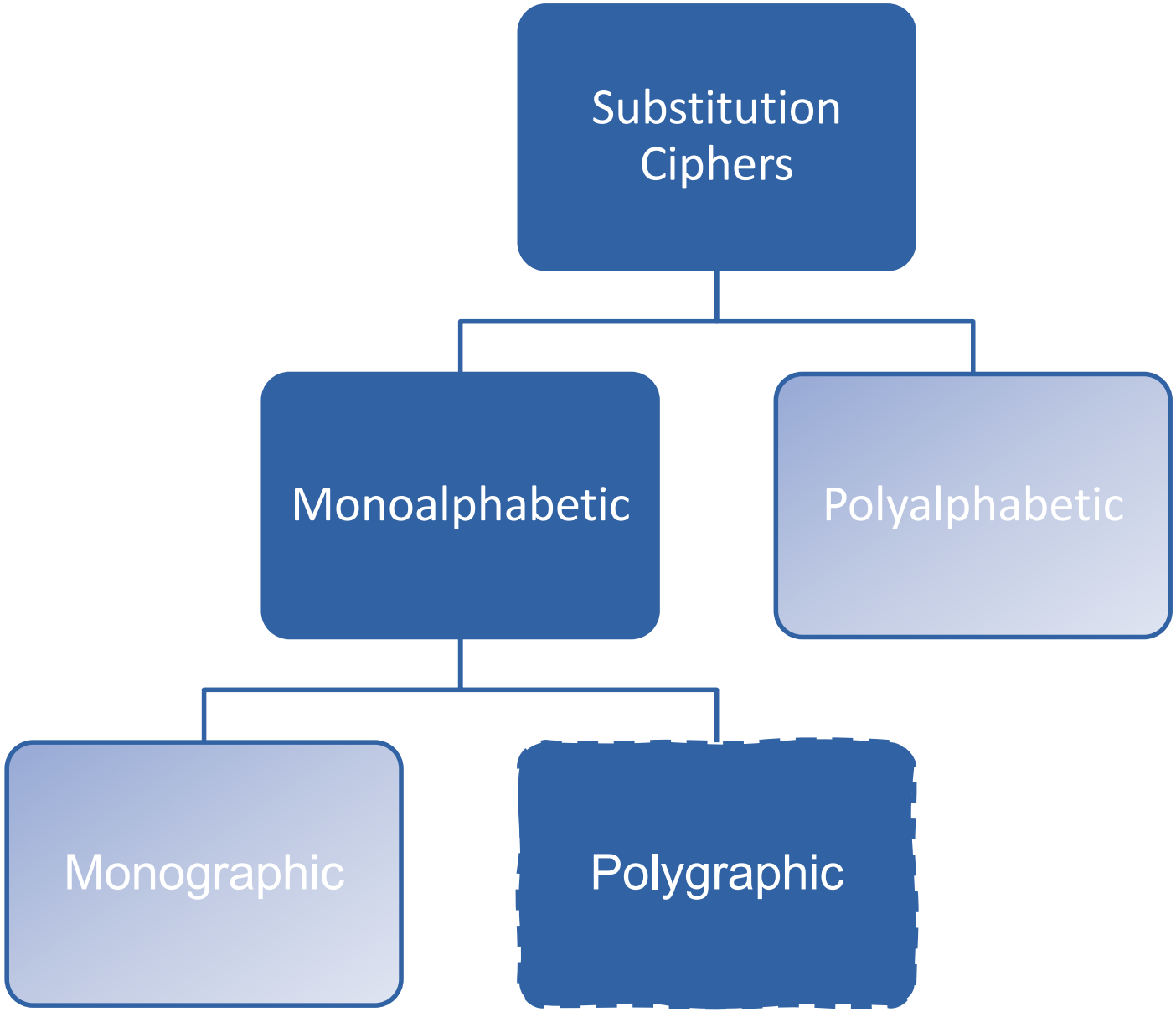
```
plaintext:  DEFEND THE EAST WALL OF THE CASTLE
ciphertext: F7EZ5F UC2 1DR6 M9PP 0E 6CZ SD4UP1
```

Homophonic Substitution Ciphers

- It it breakable?
  - While this makes analysis a bit harder, it doesn't hide all statistical properties
  - Cryptanalysis reference:
    - "Efficient Cryptanalysis of Homophonic Substitution Ciphers", Amrapali Dhavare, Richard M. Low and Mark Stamp

Substitution ciphers - Enhancements

- Objective: lessen the extent to which the structure of the plaintext survives in the ciphertext
- Two approaches
  - Encrypt multiple letters (Poly-Graphic substitution)
  - Use multiple cipher alphabet (Poly-Alphabetic ciphers)


Substitution Ciphers
Monoalphabetic
Polyalphabetic
Monographic
Polygraphic

Polygraphic Substitution Cipher

- Instead of encoding single characters, a polygraphic substitution cipher encrypts groups of letters (2, 3 and more)
- Examples:
  - Playfair cipher
  - Hill cipher

Playfair Cipher

- Invented by British scientist Sir Charles Wheatstone in 1854
- Named after Lord Playfair who promoted the use of the cipher
- Best-known multiple-letter encryption cipher
- Treats digrams in the plaintext as single units and translates these units into ciphertext digrams
- Based on the use of a 5 x 5 matrix of letters constructed using a keyword
- Used as the standard field system by the British Army in World War I and the U.S. Army and other Allied forces during WWII

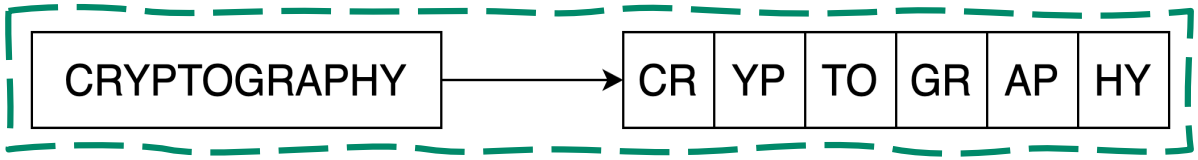
Playfair Cipher – Key Setup

1. Fill in letters of keyword (remove duplicates) from left to right and from top to bottom, then fill in the remainder of the matrix with the remaining letters in alphabetic order
   - Example: Using the keyword MONARCHY:

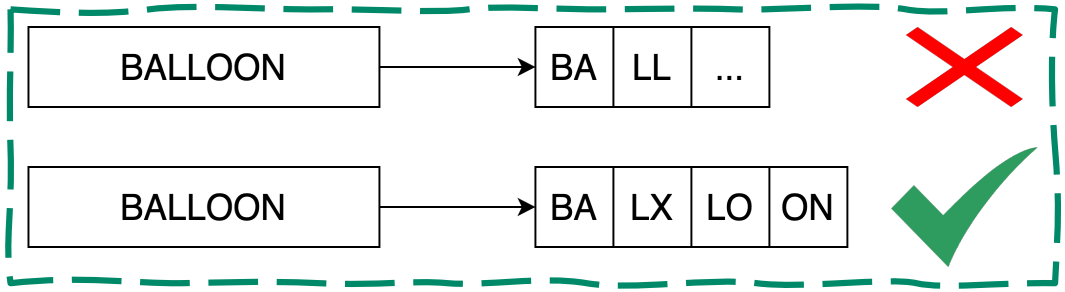
M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Playfair Cipher – Preparing the diagrams

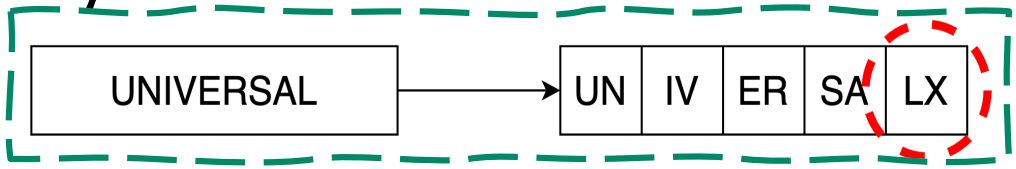
1. Split the text into digrams



2. Repeating plaintext letters in the same pair are separated with a filler letter (e.g., X)



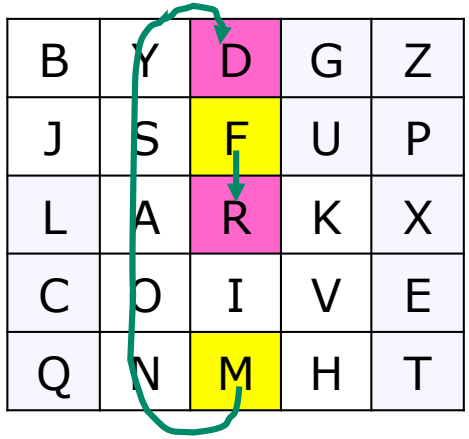
3. Possibility to add filler letter at the end

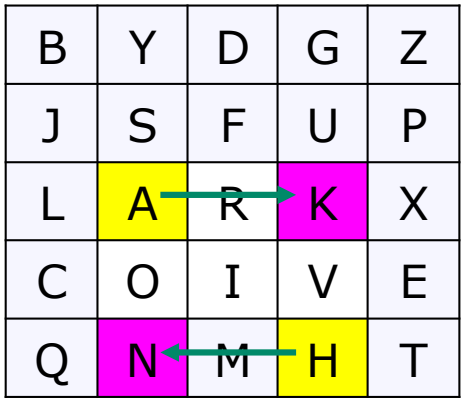


Playfair Cipher – Encryption

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

1. Two plaintext letters that fall in the **same row** of the matrix are each replaced by the letter to the right, with the first element of the row circularly following the last.
2. Two plaintext letters that fall in the **same column** are each replaced by the letter beneath, with the top element of the column circularly following the last.
3. **Otherwise**, each plaintext letter in a pair is replaced by the letter that lies in its own row and the column occupied by the other plaintext letter.





Playfair Cipher - Example

- Key: Monarchy, Plaintext: Instruments
- Z is the filler letter
- Plaintext split: 'in' 'st' 'ru' 'me' 'nt' ‘sz'
  - Key table

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

Playfair Cipher - Example (cont'd)

- Key: Monarchy, Plaintext: Instruments

in:

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

st:

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

ru:

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

me:

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

nt:

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

sz:

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

Exercise

- Encrypt the plaintext "message” using Playfair cipher using the password “crypto”

Hill Cipher

- Developed by the mathematician Lester Hill in 1929.
- The encryption algorithm takes m successive plain text and substitute for them m cipher text letters.
- Each character is assigned a numerical value ($a = 0, \dots z = 25$).
  - Alphabet size $n = 26$, but can be extended with more characters
- Passwords can be used to construct the Key Matrix K ($m \times m$ square matrix)
- $K = \begin{pmatrix} K_{11} & \cdots & K_{1m} \\ \vdots & \ddots & \vdots \\ K_{m1} & \cdots & K_{mm} \end{pmatrix}$
- K **must be invertible** $mod\ n \Rightarrow K^{-1}\ must\ exist$

$$C = KP \bmod n$$

$$P = K^{-1}.C \bmod n = K^{-1}.K.P \bmod n = P$$

Hill Cipher – Key Setup

- Given a key phrase, fill the matrix with the corresponding letters (converted to numbers)
- If the key length $\mathrm{l} > m \times m$
  - Consider the first $m \times m$ letters
- If the key length $l < m \times m$
  - Fill up the remaining elements (from the alphabet)

Hill Cipher – Preparing the m-grahs

- Split the plaintext into m-graphs
  - Every m characters constitute one column of the plaintext matrix
  - Add filler letter to complete the last column if needed

Exercise

- Encrypt the plaintext message "retreat now" using the key-phrase "backup" and a 3 x 3 matrix. Fill the key with consecutive alphabet letters starting from a whenever needed

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

Exercise - Solution

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

- Hill cipher parameters: $m = 3, n = 26$
- PT: “retreatnow” key-phrase =”backup”
- Key K is a 3×3 matrix
  - $K = \begin{pmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{pmatrix}$
- $K = \begin{pmatrix} "b" & "a" & "c" \\ "k" & "u" & "p" \\ "a" & "b" & "c" \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 10 & 20 & 15 \\ 0 & 1 & 2 \end{pmatrix}$

Exercise - Solution

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

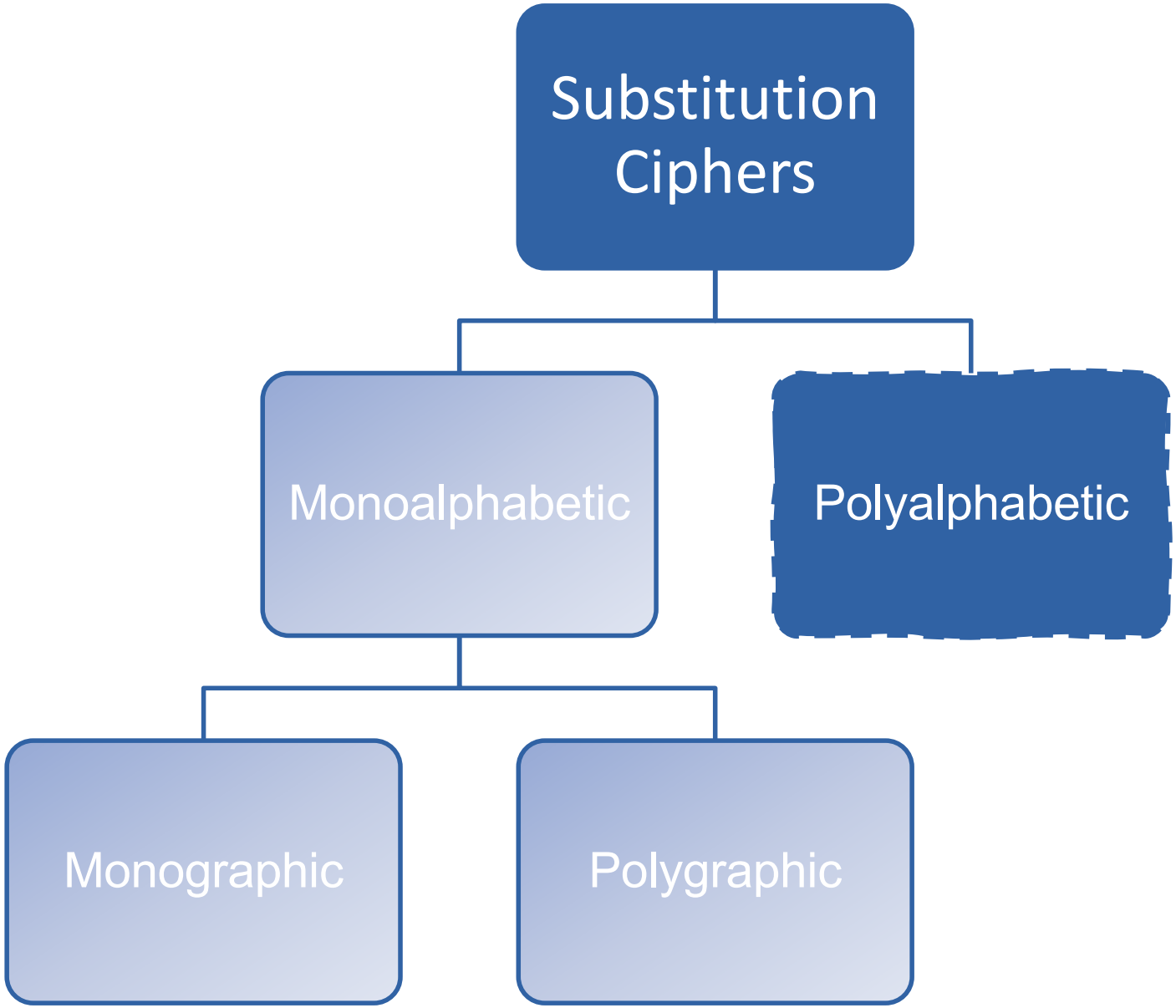
- $PT = \begin{pmatrix} "r" \\ "e" \\ "t" \end{pmatrix}\begin{pmatrix} "r" \\ "e" \\ "a" \end{pmatrix}\begin{pmatrix} "t" \\ "n" \\ "o" \end{pmatrix}\begin{pmatrix} "w" \\ "x" \\ "x" \end{pmatrix} = \begin{pmatrix} 17 \\ 4 \\ 19 \end{pmatrix}\begin{pmatrix} 17 \\ 4 \\ 0 \end{pmatrix}\begin{pmatrix} 19 \\ 13 \\ 14 \end{pmatrix}\begin{pmatrix} 22 \\ 23 \\ 23 \end{pmatrix}$
- $C_1 = KP_1 = \begin{pmatrix} 1 & 0 & 2 \\ 10 & 20 & 15 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} 17 \\ 4 \\ 19 \end{pmatrix} mod\ 26 = \begin{pmatrix} 55 \\ 535 \\ 42 \end{pmatrix} mod\ 26 = \begin{pmatrix} 3 \\ 15 \\ 16 \end{pmatrix} = \begin{pmatrix} "D" \\ "P" \\ "Q" \end{pmatrix}$
- $C_2 = KP_2 = \begin{pmatrix} 1 & 0 & 2 \\ 10 & 20 & 15 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} 17 \\ 4 \\ 0 \end{pmatrix} mod\ 26 = \begin{pmatrix} 17 \\ 16 \\ 4 \end{pmatrix} = \begin{pmatrix} "R" \\ "Q" \\ "E" \end{pmatrix}$
- $C_3 = KP_3 = \begin{pmatrix} 1 & 0 & 2 \\ 10 & 20 & 15 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} 19 \\ 13 \\ 14 \end{pmatrix} mod\ 26 = \begin{pmatrix} 3 \\ 15 \\ 16 \end{pmatrix} = \begin{pmatrix} "V" \\ "K" \\ "P" \end{pmatrix}$
- $C_4 = KP_4 = \begin{pmatrix} 1 & 0 & 2 \\ 10 & 20 & 15 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} 22 \\ 23 \\ 23 \end{pmatrix} mod\ 26 = \begin{pmatrix} 3 \\ 15 \\ 16 \end{pmatrix} = \begin{pmatrix} "Q" \\ "L" \\ "R" \end{pmatrix}$

Hill Cipher - Security

- It completely hides single-letter frequencies
- The use of a larger matrix hides more frequency information
  - A 3 x 3 Hill cipher hides not only single-letter but also two-letter frequency information
- Strong against a ciphertext-only attack
- Easily broken with a known plaintext attack

Polygraphic Substitution Cipher

- Still not a very secure way of encrypting data
  - hides the frequencies of individual letters
  - However, natural languages also show typical frequencies for n-grams (although the curve is flattened)


Substitution Ciphers
Monoalphabetic
Polyalphabetic
Monographic
Polygraphic

Poly-alphabetic Ciphers

- Use different monoalphabetic substitutions as one proceeds through the plaintext message
- General rules
  - A set of related Monoalphabetic substitution rules is used
  - A key determines which particular rule is chosen for a given transformation

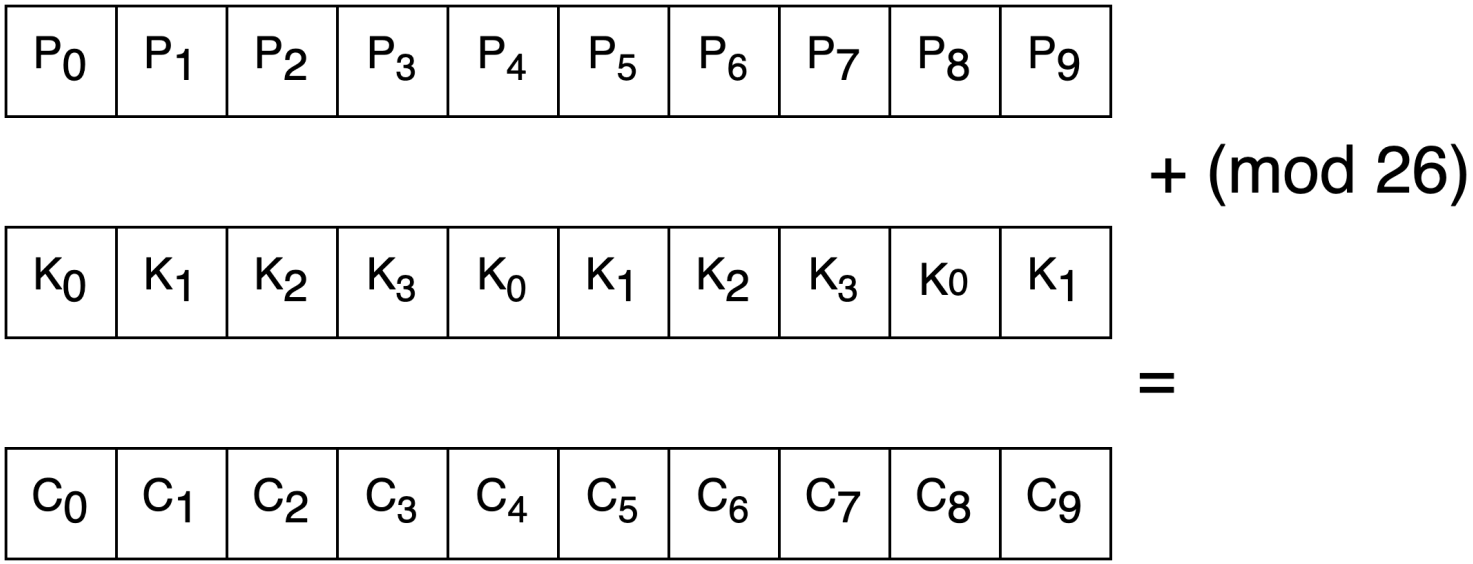
Poly-alphabetic substitution

- How it works?
  - There are multiple one-letter keys
  - The first key encrypts the first letter of the plaintext, the second key encrypts the second letter of the plaintext, and so on
  - After all keys are used, you start over with the first key
  - The number of keys determines the period of the cipher

Vigenère cipher

- Invented by Blaise de Vigenère (1523 – 1596)
- The key is now a string, not just a character
- Substitution based on the position of the character and on the key
- To encrypt, shift each character in the plaintext by the amount dictated by the next character of the key
  - Wrap around in the key as needed
- Decryption just reverses the process

Vigenère cipher



$$l \Rightarrow Key\ length, n \Rightarrow Message\ length$$

$$C_i = P_i + K_{i\ mod\ l}\ mod\ 26, 0 \leq i < n$$

The Vigenère cipher

Tabula recta

	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

Exercise

- Encrypt the message “ I love Cryptography” using Vigenère Cipher with the keyword “secret"

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

Exercise

Consider the Vigenère cipher over the lowercase English alphabet, where the key length can be anything from 8 to 12 characters.
What is the size of the key space for this scheme?

- Answer:

$$Key\ space\ size\ = 26^8 + 26^9 + 26^{10} + 26^{11} + 26^{12}$$

The Vigenère cipher

- Size of key space?
  - If keys are 14-character strings; then key space has size $26^{14} \approx 2^{66}$
  - Brute-force search expensive/impossible
- Believed to be secure for many years…
- But, Is it really secure?

Attacking the Vigenère Cipher - Example

- The following ciphertext was encrypted with Vigenère Cipher

```
LLGTSNJKIVLXHJSLORHBKLLGSILLLYVFVBFLLGWEV
MDXKVWHXLLGLRBNWVUZXR
```

Look for repeated patters (the longest the better)

Attacking the Vigenère Cipher - Example

- The following ciphertext was encrypted with Vigenère Cipher

```
LLGTSNJKIVLXHJSLORHBKLLGSILLLYVFVBFLLGWEV
MDXKVWHXLLGLRBNWVUZXR
```

What is the key size?

Attacking the Vigenère Cipher

- **Phase 1: Find the Key length**
- Search the ciphertext for repeated strings of letters; the longer strings you find the better
- For each occurrence of a repeated string, count how many letters are between the first letters in the string and add one
- Factor the number you got in the above computation
- Repeat this process with each repeated string you find and make a table of common factors.
- The most common factor (n) is probably the length of the keyword that was used to encipher the ciphertext

Attacking the Vigenère Cipher

- **Phase 2: Find the Key**
- Do a frequency count on the ciphertext, on every n^{th} letter.
- This results in **n different frequency counts**.
- Compare these counts to standard frequency tables to figure out how much each letter was shifted by.
- Undo the shifts and read off the message!

Vigenère Autokey System

- A keyword is concatenated with the plaintext itself to provide a running key
- Example: encrypt the plaintext “we are discovered save yourself” with a keyword *deceptive*

PT	W	E	A	R	E	D	I	S	C	O	V	E	R	E	D	S	A	V	E	Y	O	U	R	S	E	L	F
Key	D	E	C	E	P	T	I	V	E	W	E	A	R	E	D	I	S	C	O	V	E	R	E	D	S	A	V
CT	Z	I	C	V	T	W	Q	N	G	K	Z	E	I	I	G	A	S	X	S	T	S	L	V	V	W	L	A

This scheme is still vulnerable to cryptanalysis

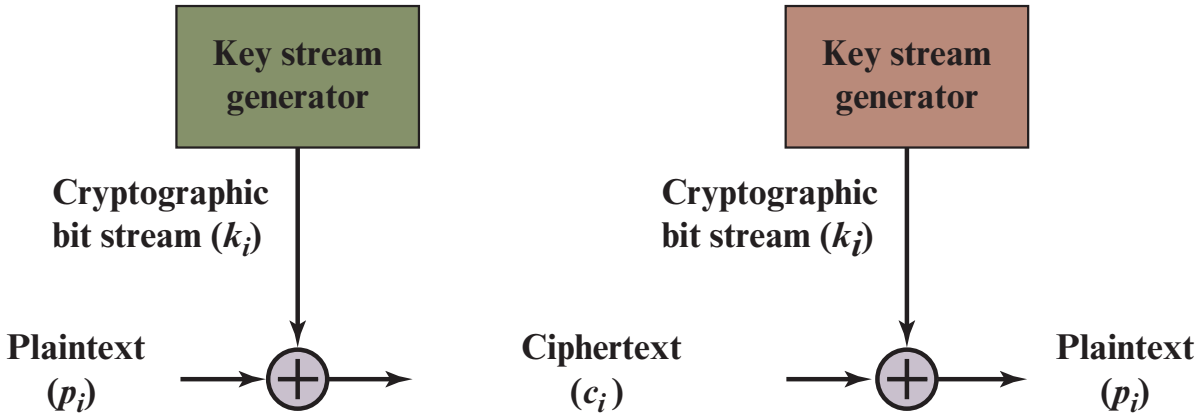
- The key and the plaintext share the same frequency distribution of letters,
- A statistical technique can be applied

How to improve substitution ciphers?

- Use a key that
  - is as long as the plaintext
  - has no statistical relationship to it.

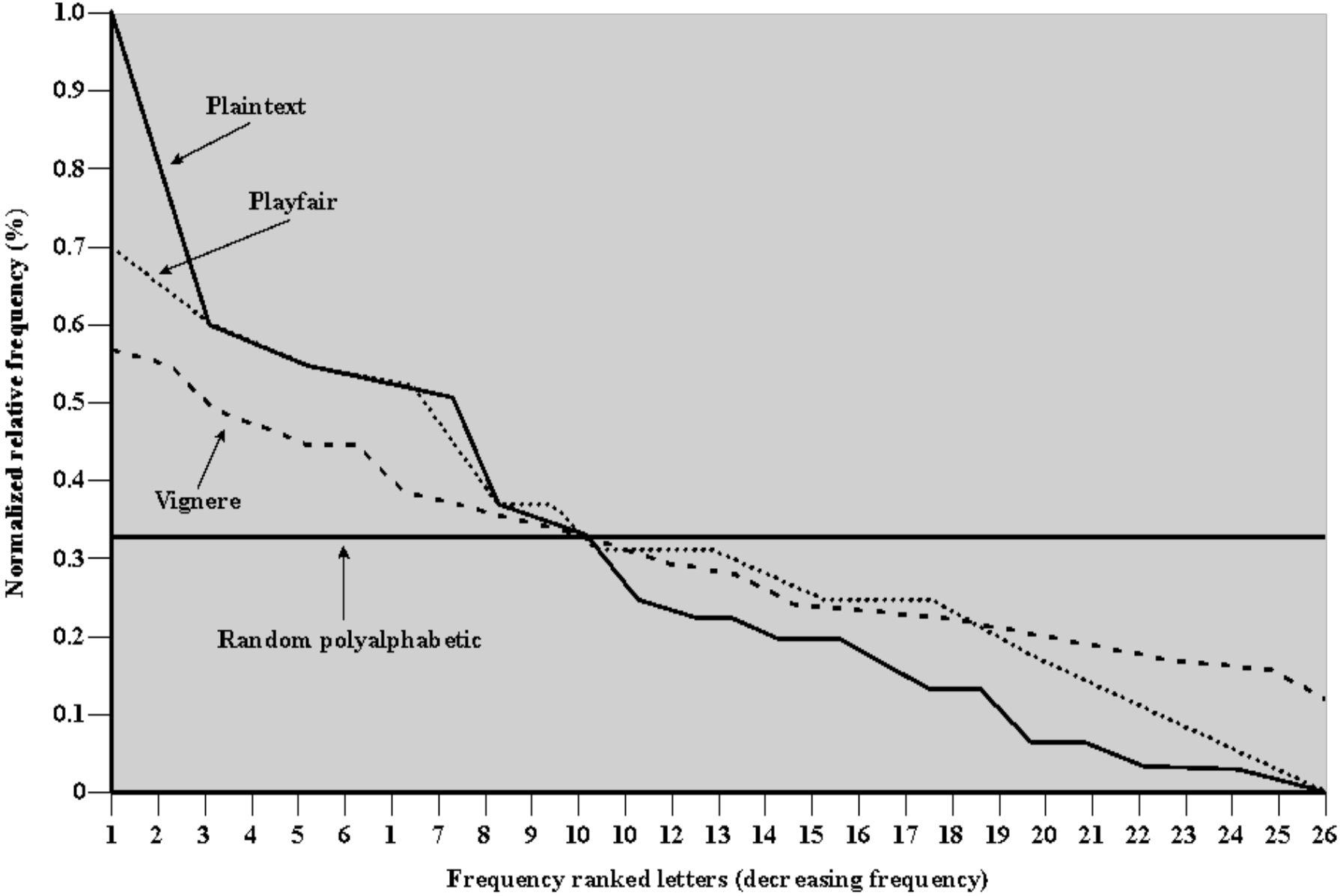
Vernam Cipher

- Proposed by Gilbert Vernam in 1918
- Works on binary data rather than letters
- $C_i = P_i \oplus K_i \Rightarrow P_i = C_i \oplus K_i$
- Essence of Vernam cipher: the construction of the key
  - Vernam proposed to repeat the keyword continuously



One-Time Pad (OTP)

- Improvement to Vernam cipher proposed by Joseph Mauborgne
- Use a random key that is as long as the message so that the key need not be repeated
- Key is used to encrypt and decrypt a single message and **then is discarded**
- Each new message requires a new key of the same length as the new message
- **Scheme is unbreakable**
  - § Produces random output that bears no statistical relationship to the plaintext


Plaintext
Playfair
Vignere
Random polyalphabetic
Normalized relative frequency (%)
Frequency ranked letters (decreasing frequency)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 2 3 4 5 6 1 7 8 9 10 10 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Classical Ciphers

- Principles and Basic terminology
- Substitution Ciphers
- **Transposition Ciphers**
- Rotor Machines

Transposition ciphers

- Apply some sort of permutation on the plaintext letters
- Keep the same characters but re-arrange them
- Example:
  - Scytales (500 B.C.)
  - Rail fence cipher

Rail Fence Cipher

- Simplest transposition cipher
- Plaintext is written down as a sequence of diagonals and then read off as a sequence of rows
- Example:
  - Plaintext: this is a secret message
  - Encryption with Key=2

T		I		I		A		E		R		T		E		S		G	
	H		S		S		S		C		E		M		S		A		E

  - Ciphertext: TIIAERTESGHSSSCEMSAE

Row Transposition Cipher

- Is a more complex transposition
  - Steps:
    - Write the message in a rectangle row by row
    - read the message off, column by column, but permute the order of the columns
  - The order of the columns then becomes the key to the algorithm

Row Transposition Cipher (cont'd)

- Example
  - Key: 3 4 2 1 5 6 7
  - PT: attack postponed until two am

a	t	t	a	c	k	p
o	s	t	p	o	n	e
d	u	n	t	i	l	t
w	o	a	m	x	y	z

  - CT: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Transposition ciphers

- Transposition ciphers can be made more secure by performing more than one stage of transposition
- Makes the algorithm more difficult to cryptanalysis

Classical Ciphers

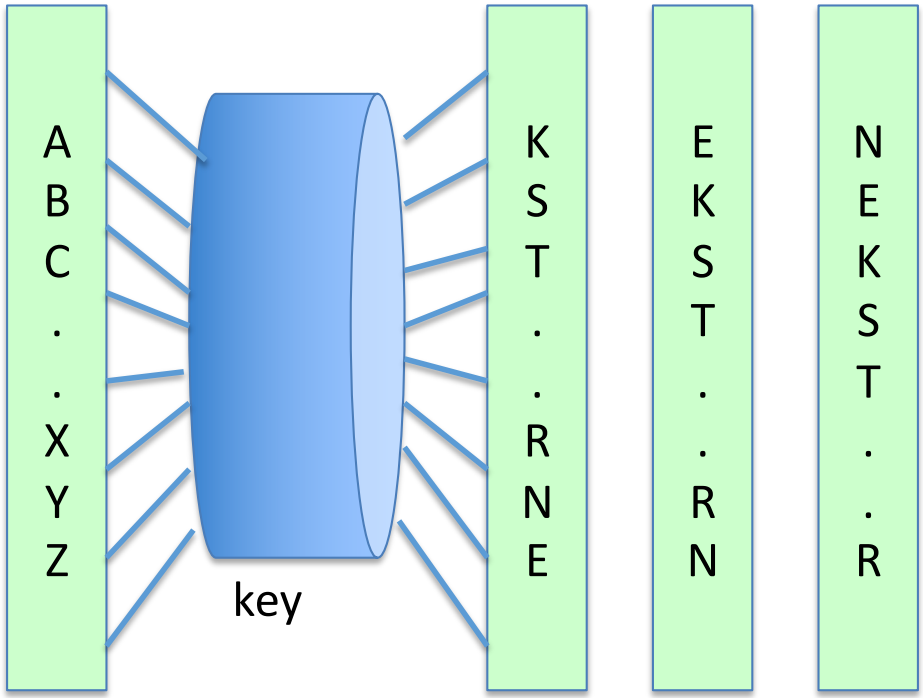
- Principles and Basic terminology
- Substitution Ciphers
- Transposition Ciphers
- **Rotor Machines**

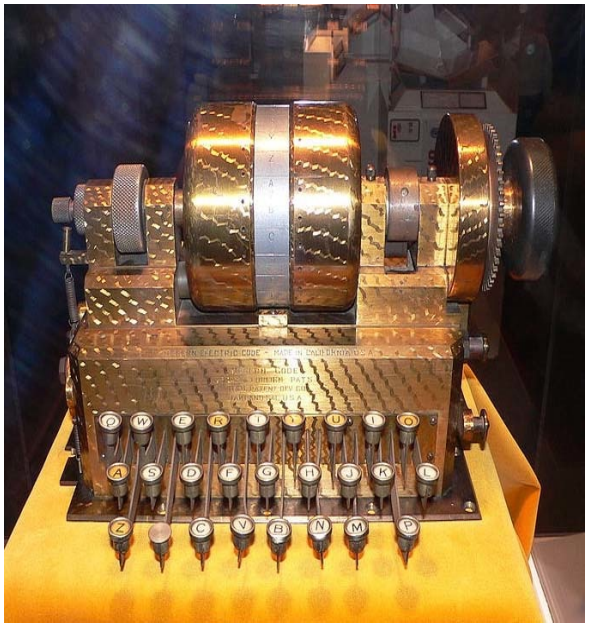
Rotor Machines

- Typical application of multi-stage encryption
- Appeared in the 19th century
- Single rotor and multiple rotor machines

Rotor Machines

- Early example: the Hebern machine (single rotor)
- The disk rotates at each character input and changes the substitution table
- Broken using frequency analysis with enough ciphertexts





Rotor Machines

- Most famous: the Enigma (3-5 rotors)
- Key: initial settings of the rotors
- # keys = $26^4 = 2^{18}$ (for 4- rotors machine)

Rotor Machines