Chủ đề 1:

Tổng quan về Mã hóa thông tin và Ứng dụng

Mở đầu

Mở đầu

- Khoa học mật mã đã ra đời từ hàng nghìn năm.
- Trong suốt nhiều thế kỷ, các kết quả của lĩnh vực này hầu như không được ứng dụng trong các lĩnh vực dân sự thông thường của đời sống – xã hội mà chủ yếu được sử dụng trong lĩnh vực quân sự, chính trị, ngoại giao...
- Ngày nay, các ứng dụng mã hóa và bảo mật thông tin đang được sử dụng ngày càng phổ biến trong các lĩnh vực khác nhau trên thế giới, từ các lĩnh vực an ninh, quân sự, quốc phòng…, cho đến các lĩnh vực dân sự như thương mại điện tử, ngân hàng…

Mật mã học

- Mật mã (Cryptography) là ngành khoa học nghiên cứu các **kỹ thuật toán học** nhằm cung cấp các **dịch vụ bảo vệ thông tin.**

  W. Stallings (2003), *Cryptography and Network Security: Principles and Practice, Third Edition*, Prentice Hall

Một số thuật ngữ

- Cryptography
- Cryptanalysis
- Cryptology = Cryptography + Cryptanalysis
- Security
- Steganography

Các vấn đề chính trong Mật mã học

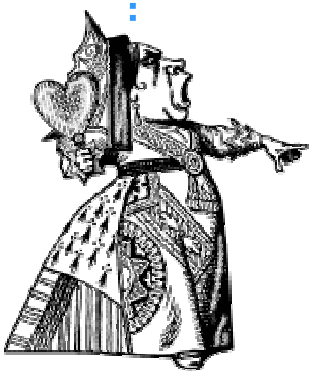
Mật mã học???

Cách hiểu truyền thống: giữ bí mật nội dung trao đổi

Alice và **Bob** trao đổi **với nhau trong khi Eve** tìm cách "nghe lén"

Alice

Bob



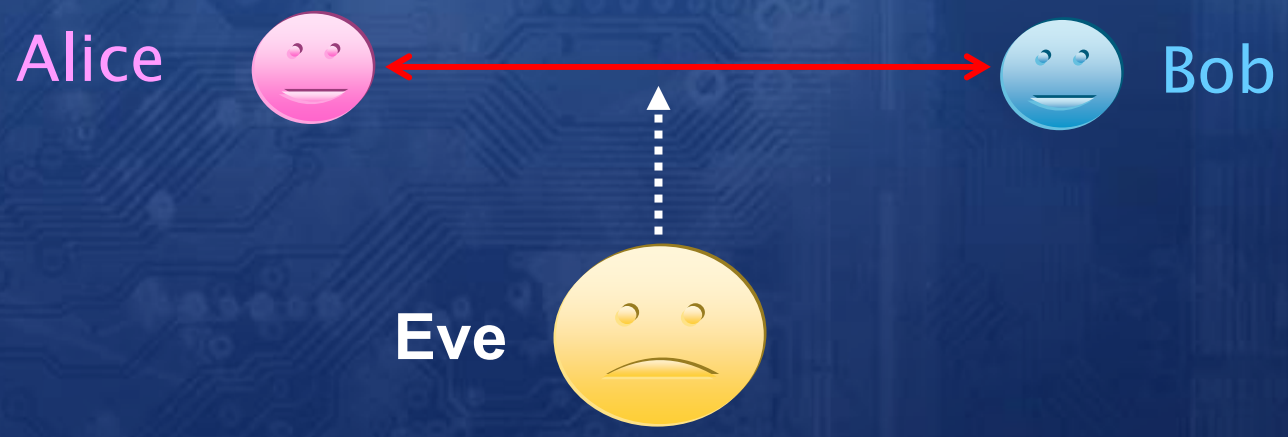
Eve

Một số vấn đề chính trong bảo vệ thông tin

- Bảo mật thông tin (Secrecy): đảm bảo thông tin được giữ bí mật.
- Toàn vẹn thông tin (Integrity): bảo đảm tính toàn vẹn thông tin trong liên lạc hoặc giúp phát hiện rằng thông tin đã bị sửa đổi.
- Xác thực (Authentication): xác thực các đối tác trong liên lạc và xác thực nội dung thông tin trong liên lạc.
- Chống lại sự thoái thác trách nhiệm (Non-repudiation): đảm bảo một đối tác bất kỳ trong hệ thống không thể từ chối trách nhiệm về hành động mà mình đã thực hiện

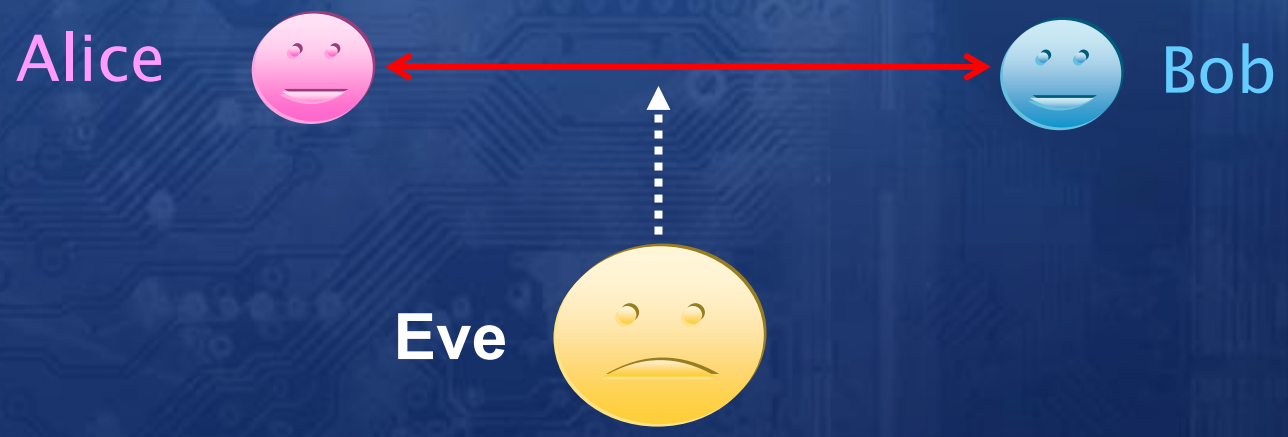
Xác thực (Authentication)

- Ví dụ:
  - Bob chờ Alice "xác nhận" khi đến thời điểm thực hiện công việc
  - Cần đảm bảo rằng Eve không can thiệp để tạo "xác nhận" giả
- Xác thực (Authentication), Định danh (identification)



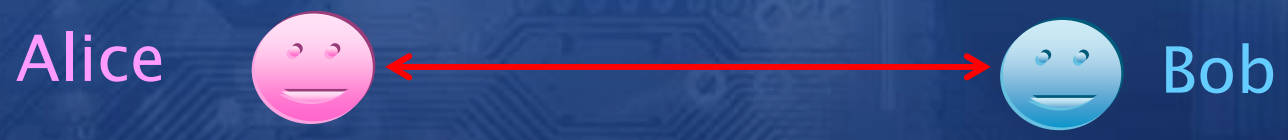
Tính toàn vẹn thông tin (Integrity)

- Ví dụ:
  - Bob cần đảm bảo là nhận chính xác nội dung mà Alice đã gửi
  - Cần đảm bảo rằng Eve không can thiệp để sửa nội dung thông điệp mà Alice gửi cho Bob
- Tính toàn vẹn thông tin (Integrity)



Chống lại sự thoái thác trách nhiệm

- Ví dụ:
  - Bob nhận được 1 thông điệp mà Alice đã gửi
  - Alice không thể "chối" rằng không gửi thông điệp này cho Bob
- Chống lại sự thoái thác trách nhiệm (Non-repudiation)



Các giải pháp

- Authentication (Identity verification)
- Access control (Authorization)
- Data confidentiality (Privacy)
- Data integrity (Tamper-proofing)
- Non-repudiation (Proof of transaction)

Authentication

- Kiểm tra danh tính
  - Đảm bảo rằng một người dùng (có thể là một tổ chức, một phần mềm, ...) phải chứng minh được mình là ai.
  - Ngăn ngừa hiện tượng giả mạo danh tính (Identity interception) và hiện tượng sử dụng username và password dạng clear-text của người khác do dòm ngó được trên mạng vào mục đích xấu xa (Masquerading)
- Giải pháp
  - Chữ ký điện tử

Access Control (Authorization)

- Cho phép một người dùng hợp lệ đăng nhập vào hệ thống và thực hiện các chức năng tương ứng với vai trò của mình
- Access Control có thể hiện thực bằng các dịch vụ File và database
- Giải pháp
  - ACLs (Access Control Lists): danh sách các người dùng, thuộc về các nhóm nào đó, và quy định về các quyền tương ứng cho từng nhóm trên LDAP server.

Data Confidentiality (Privacy)

- Bảo vệ thông tin trên đường truyền khỏi những con mắt liếc ngang liếc dọc
- Giải pháp
  - Mã hóa dữ liệu (Mã hóa đối xứng và mã hóa bất đối xứng)

Data Integrity (Tamper- Proofing)

- Ngăn ngừa dữ liệu bị thay đổi trên đường truyền
  - Đảm bảo dữ liệu từ người gửi đến người nhận phải giống nhau
- Giải pháp
  - Hàm băm mật mã (Message Digest)

Non-repudiation (Proof of Transaction)

- Chứng minh với tổ chức thứ 3 một giao tác nào đó đã thật sự xảy ra
  - Bảo vệ cả người gửi lẫn người nhận
- Giải pháp
  - Chữ ký điện tử

Lịch sử phát triển của Mật mã học

Sơ lược lịch sử phát triển của mật mã học

Nguồn: http://www.cqrsoft.com/history/scytale.htm

Dẫn nhập

- Ấn/con dấu được sử dụng để đóng lên các tài liệu quan trọng
- Mật khẩu (Password) được sử dụng để định danh người trong tổ chức
- …

Nguồn:

http://images.encarta.msn.com/xrefmedia/sharemed/targets/images/pho/t025/T025102A.jpg

Mã hóa thời kỳ cổ đại

ת	ש	ר	ק	צ	פ	ע	ס	נ	מ	ל	כ	י	ט	ח	ז	ו	ה	ד	ג	ב	א
א	ב	ג	ד	ה	ו	ז	ח	ט	י	כ	ל	מ	נ	ס	ע	פ	צ	ק	ר	ש	ת

- Phương pháp mã hóa **Atbash**:
  - Được sử dụng trong tiếng Hebrew cổ "בבל = ששך"
- Phương pháp **Caesar**

A	B	C	…	X	Y	Z
D	E	F	…	A	B	C

- Bất kỳ ai biết được quy tắc mã hóa này để dễ dàng giải mã thông điệp

Mã hóa thời kỳ cổ đại

- Phương pháp Caesar là một trường hợp đặc biệt của phương pháp mã hóa bằng cách dịch chuyển (Shift Ciphers).
- Phương pháp Shift Cipher: các ký tự được xoay vòng đi K vị trí trong bảng chữ cái. K được xem là khóa để giải mã

A	B	C	…	X	Y	Z
D	E	F	…	A	B	C

- Cả phương pháp Atbash và Shift Cipher đều là trường hợp đặc biệt của phương pháp tổng quát được sử dụng trong thời cổ đại: Phương pháp Thay thế đơn ký tự (MonoAlphabetic Substitution Cipher)

Ceasar Shift Ciphers

- VD: dịch chuyển 3 ký tự trong bảng chữ cái, ta có:

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

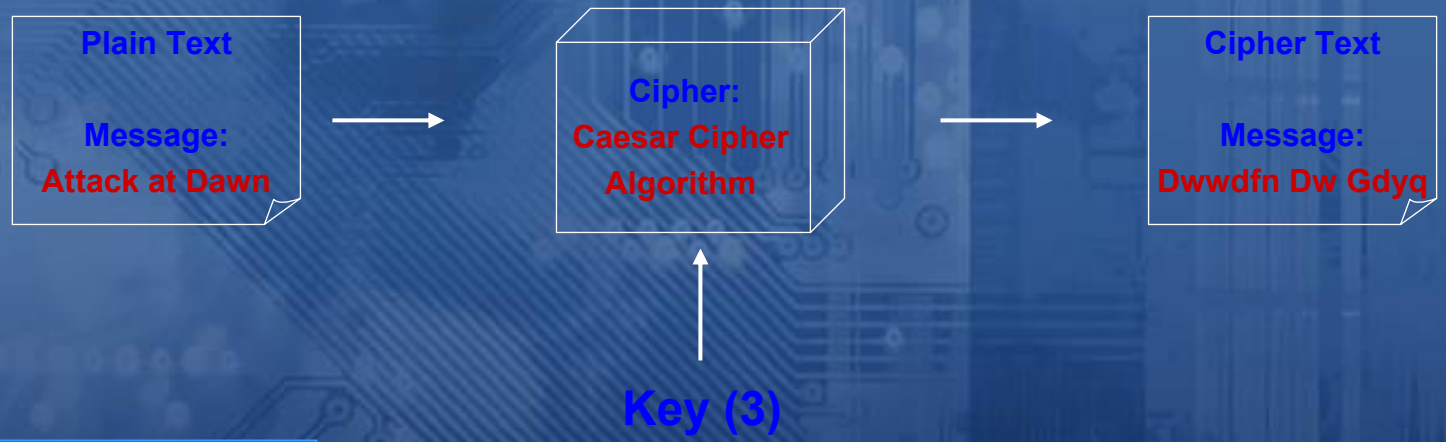
↓

D 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

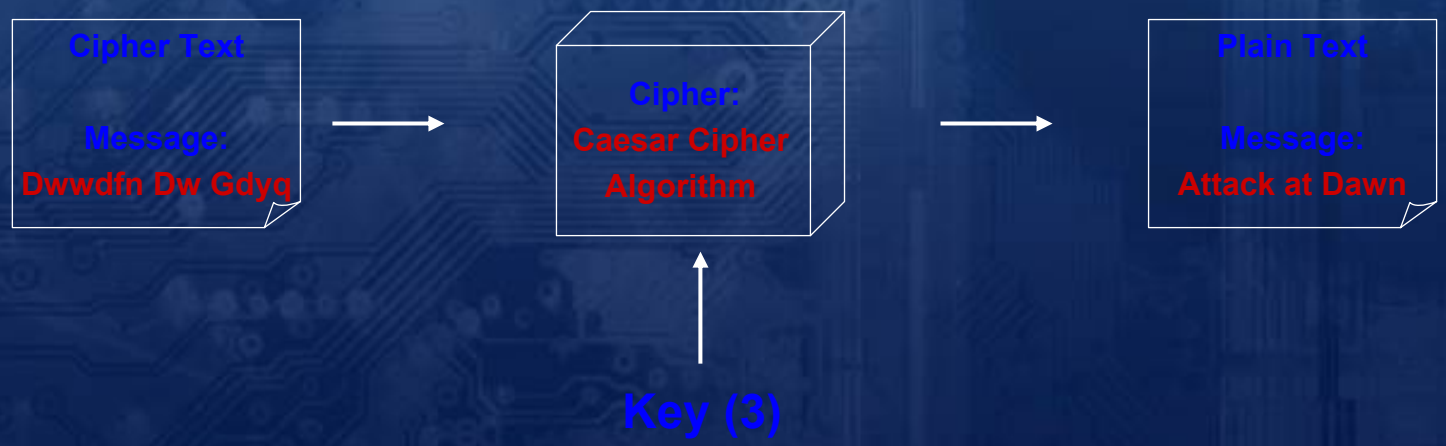
- Mã hóa cụm từ: Attack at Down

Quá trình mã hóa và giải mã

Mã hóa



Giải mã



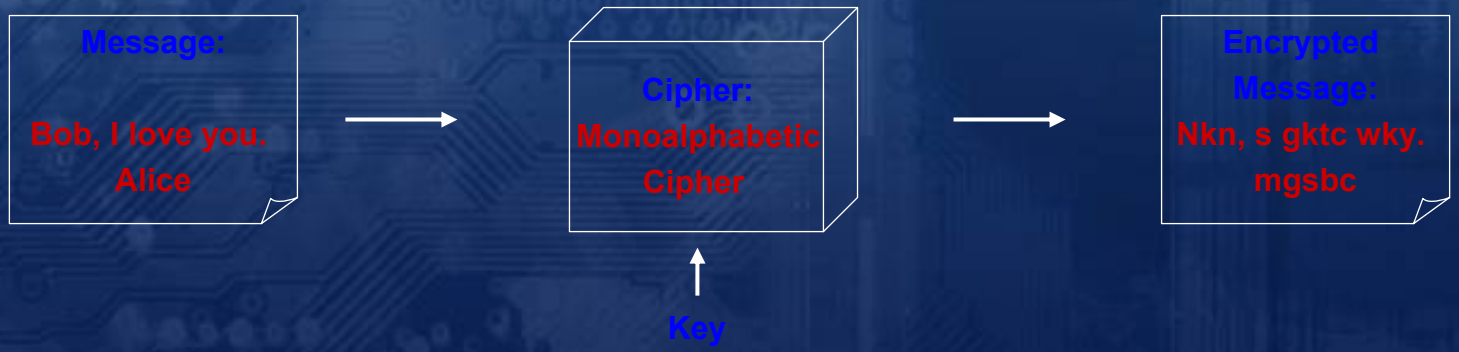
Monoalphabetic Substitution Cipher

- Mỗi ký tự thay thế bằng 1 ký tự khác duy nhất trong bảng chữ cái

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

↓

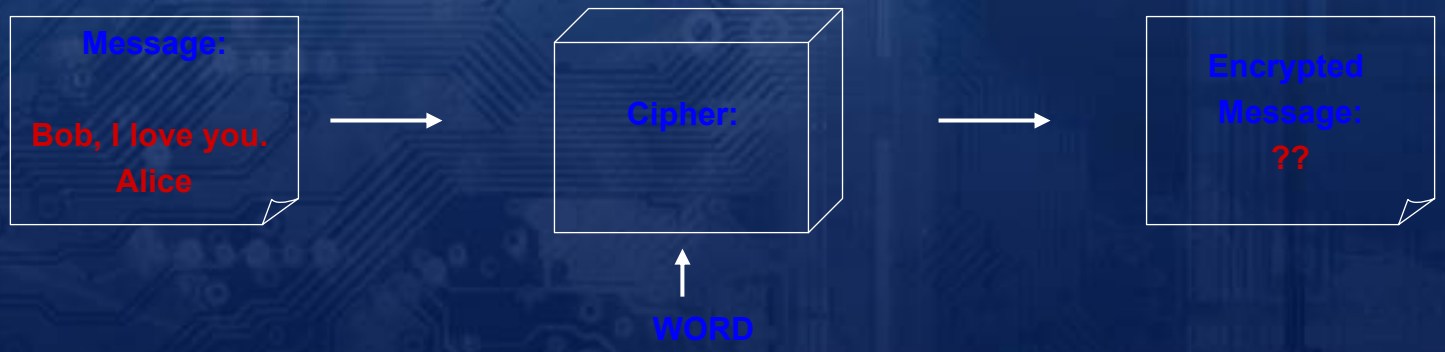
M N B V C X Z A S D F G H J K L P O I U Y T R E W Q



Using a key to shift alphabet

- Sử dụng khóa là một từ nào đó, ví dụ WORD
- Đảm bảo ánh xạ một-một

Plain Text	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
	↓
C1	W O R D A B C E F G H I J K L M N P Q S T U V X Y Z



Mã hóa thời kỳ cổ đại

- Không phải tất cả các phương pháp mã thời cổ đại đều sử dụng phương pháp thay thế.
- Thiết bị mã hóa đầu tiên: Spartan scytale

Nguồn:
http://plus.maths.org/issue34/features/ekert/

- Sử dụng thiết bị này, các chữ cái trong thông điệp không bị thay đổi, mà chỉ thay đổi vị trí xuất hiện của các thông điệp (Transposition)

Columnar Transposition

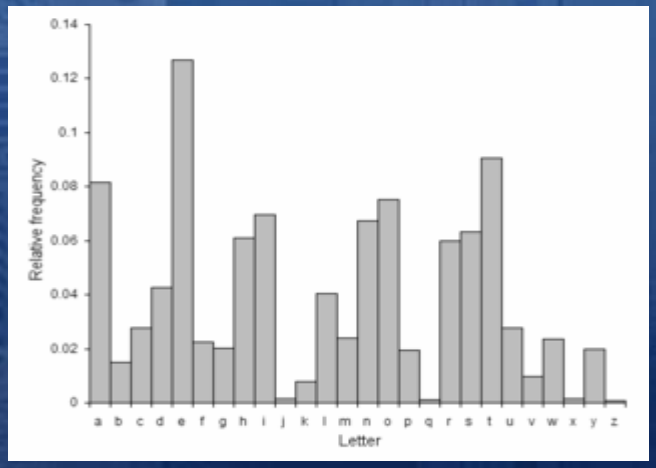
- Sắp xếp các ký tự của dữ liệu plaintext vào các cột
- Nếu các ký tự của plaintext không là bội số của cột, thay bằng các ký tự ít xuất hiện như x, z.

Plain Text	Cipher Text
T H I S I	T S S O H
S A M E S	O A N I W
S A G E T	H A A S O
O S H O W	L R S T O
H O W A C	I M G H W
O L U M N	U T P I R
A R T R A	S E E O A
N S P O S	M R O O K
I T I O N	I S T W C
W O R K S	N A S N S

Mã hóa thời kỳ cổ đại

- Theo các tài liệu ghi nhận lại, phương pháp phân tích tần số sử dụng được sử dụng từ thế kỷ thứ 9

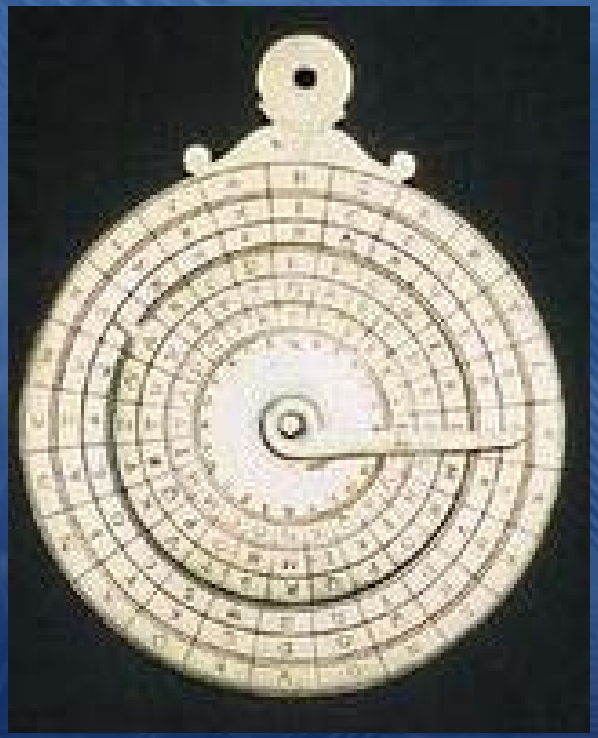
http://plus.maths.org/issue34/features/ekert/



http://en.wikipedia.org/wiki/Caesar_cipher

- Mã hóa ở Châu Âu gần như ít có sự phát triển từ thời cổ đại đến thế kỷ 14!!!

Mã hóa thời kỳ phục hưng



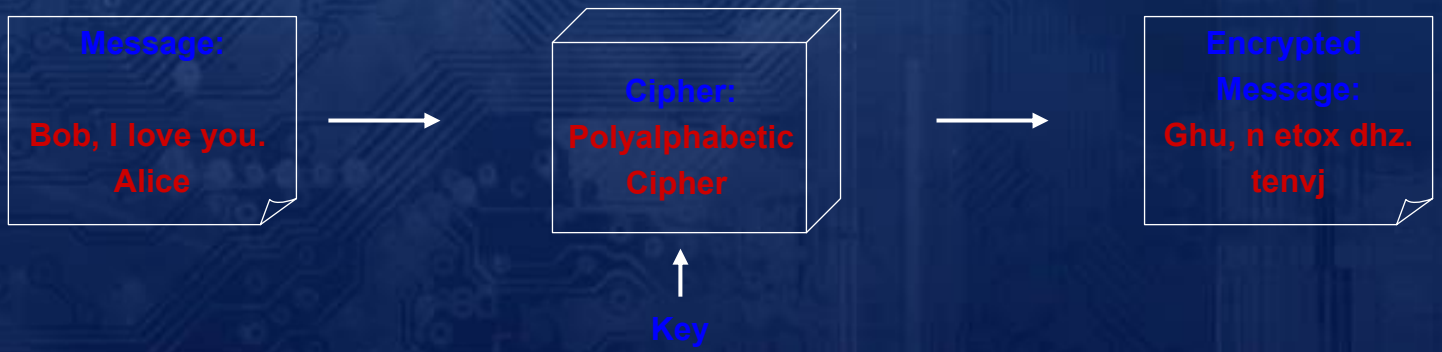
- Ở Ý, cũng như các nước Châu Âu khác, mật mã học bắt đầu được phát triển trở lại
- Các quốc gia, các thành phố bắt đầu tìm kiếm các chuyên gia về mật mã và phá mã để mã hóa và giải mã các bức thư.
- Phương pháp mã hóa giai đoạn này thường là Thay thế đa ký tự (PolyAlphabetic Substitution Cipher).
- Nhiều dụng cụ mã hóa được chế tạo và sử dụng

Polyalphabetic Substitution Cipher

- Sử dụng một chuỗi Monoalphabetic Cipher
  - C1, C2, C2, C1, C2

Plain Text	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
	↓
C1(k=6)	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
C2(k=20)	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

- Ví dụ



Mã hóa thời kỳ phục hưng

- Phương pháp mã hóa bằng cách thay thế đa ký tự có thể được xem như sử dụng nhiều lần thay thế đơn ký tự liên tiếp nhau.
- Thường dùng dụng cụ Cipher Disk, hoặc dùng bảng tra để giúp mã hóa và giải mã

- Kỹ thuật chính (kinh điển) dùng để phá vỡ hệ mã Thay thế đa ký tự gồm 2 bước:
  - Tìm ra độ dài của chu kỳ
  - Áp dụng kỹ thuật phân tích (cho phương pháp mã hóa thay thế đơn ký tự) + thông tin thu được từ các ký tự trước

Mã hóa trong thế kỷ 19 và đầu thế kỷ 20

- Mã hóa được sử dụng phổ biến trong Thế chiến I
- Sự phát triển của sóng vô tuyến và điện đài giúp việc liên lạc trong quân đội được thực hiện dễ dàng và nhiều hơn.
- Đòi hỏi các thiết bị hỗ trợ việc mã hóa và giải mã
- Các máy mã hóa ra đời

Mã hóa trong thế kỷ 19 và đầu thế kỷ 20

- Thế chiến thứ 2: cuộc chiến trên lĩnh vực khoa học, trong đó có cả khoa học mật mã.
- Máy mã hóa Enigma (của Đức) bị quân đội Anh giải mã
- Máy mã hóa “Purple” của Nhật bị quân đội Mỹ giải mã

Hệ thống mã hóa

Hệ thống mã hóa

Hệ thống mã hóa

điều kiện sau:

1. *Tập nguồn*

2. *Tập đích C là tập*

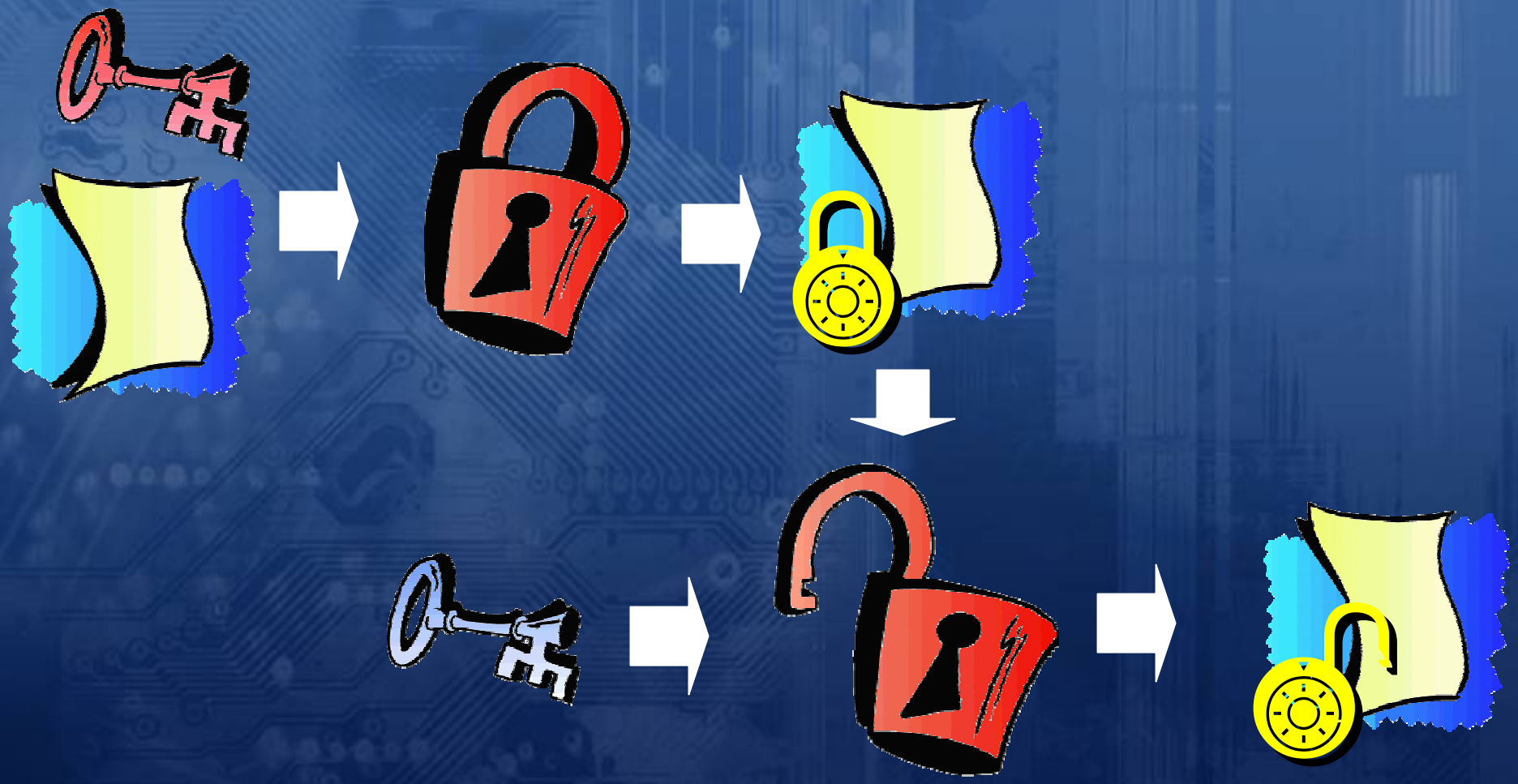
3. *Tập khóa K là tập hữu hạn*

4. *E và D lần lượt là tập luật mã hóa và giải mã. Với mỗi khóa* $k \in K$ *, tồn tại luật mã hóa* $e_k \in E$ *và luật giải mã* $d_k \in D$ *tương ứng. Luật mã hóa* $e_k : P \to C$ *và luật giải mã* $e_k : C \to P$ *là hai ánh xạ thỏa mãn* $d_k(e_k(x)) = x, \forall x \in P$

Bảo đảm một mẩu tin *x* được mã hóa bằng luật mã hóa e_k có thể được giải mã chính xác bằng luật d_k

Hệ thống mã hóa đối xứng

Mã hóa khóa công cộng



Mã đối xứng VS mã bất đối xứng

Tốc độ xử lý nhanh	Tốc độ xử lý chậm
Mã khóa ngắn	Mã khóa dài
Khó trao đổi mã khóa	Trao đổi mã khóa dễ dàng

Trường Z_m

Khái niệm về Z_m

- Z_m được định nghĩa là tập hợp $\{0, 1, \ldots, m\text{-}1\}$, được trang bị phép cộng (ký hiệu +) và phép nhân (ký hiệu là $\times$).
- Phép cộng và phép nhân được thực hiện trong Z_m tương tự như trong Z, ngoại trừ kết quả tính theo modulo m
- **Ví dụ**:
  - Giả sử ta cần tính giá trị trong Z_{16}.
  - Trong Z, ta có kết quả của phép nhân $11 \times 13 = 143$
  - Do $143 \equiv 15 \pmod{16}$ nên $11 \times 13 = 15$ trong Z_{16}.

Tính chất của Z_m

1. Phép cộng đóng trong $\mathbf{Z}_m$, $\forall a,b \in \mathbf{Z}_m$, $a+b \in \mathbf{Z}_m$
2. Tính giao hoán của phép cộng trong $\mathbf{Z}_m$, $\forall a,b \in \mathbf{Z}_m$, $a+b=b+a$
3. Tính kết hợp của phép cộng trong $\mathbf{Z}_m$, $\forall a,b,c \in \mathbf{Z}_m$, $(a+b)+c=a+(b+c)$
4. $\mathbf{Z}_m$ có phần tử trung hòa là 0, $\forall a,b \in \mathbf{Z}_m$, $a+0=0+a=a$
5. Mọi phần tử a trong $\mathbf{Z}_m$ đều có phần tử đối là $m-a$

Tính chất của Z_m (tt)

6. Phép nhân đóng trong $\mathbf{Z}_m$, $\forall a,b \in \mathbf{Z}_m$, $a \times b \in \mathbf{Z}_m$

7. Tính giao hoán của phép nhân trong $\mathbf{Z}_m$, $\forall a,b \in \mathbf{Z}_m$, $a \times b = b \times a$

8. Tính kết hợp của phép nhân trong $\mathbf{Z}_m$, $\forall a,b,c \in \mathbf{Z}_m$, $(a \times b) \times c = a \times (b \times c)$

9. $\mathbf{Z}_m$ có phần tử đơn vị là 1, $\forall a,b \in \mathbf{Z}_m$, $a \times 1 = 1 \times a = a$

10. Tính phân phối của phép nhân đối với phép cộng, $\forall a,b,c \in \mathbf{Z}_m$, $(a+b) \times c = a \times c + b \times c$