Vietnam National University Ho Chi Minh City, University of Science
Department of Information Technology

Topic 2:
Symmetric Cipher

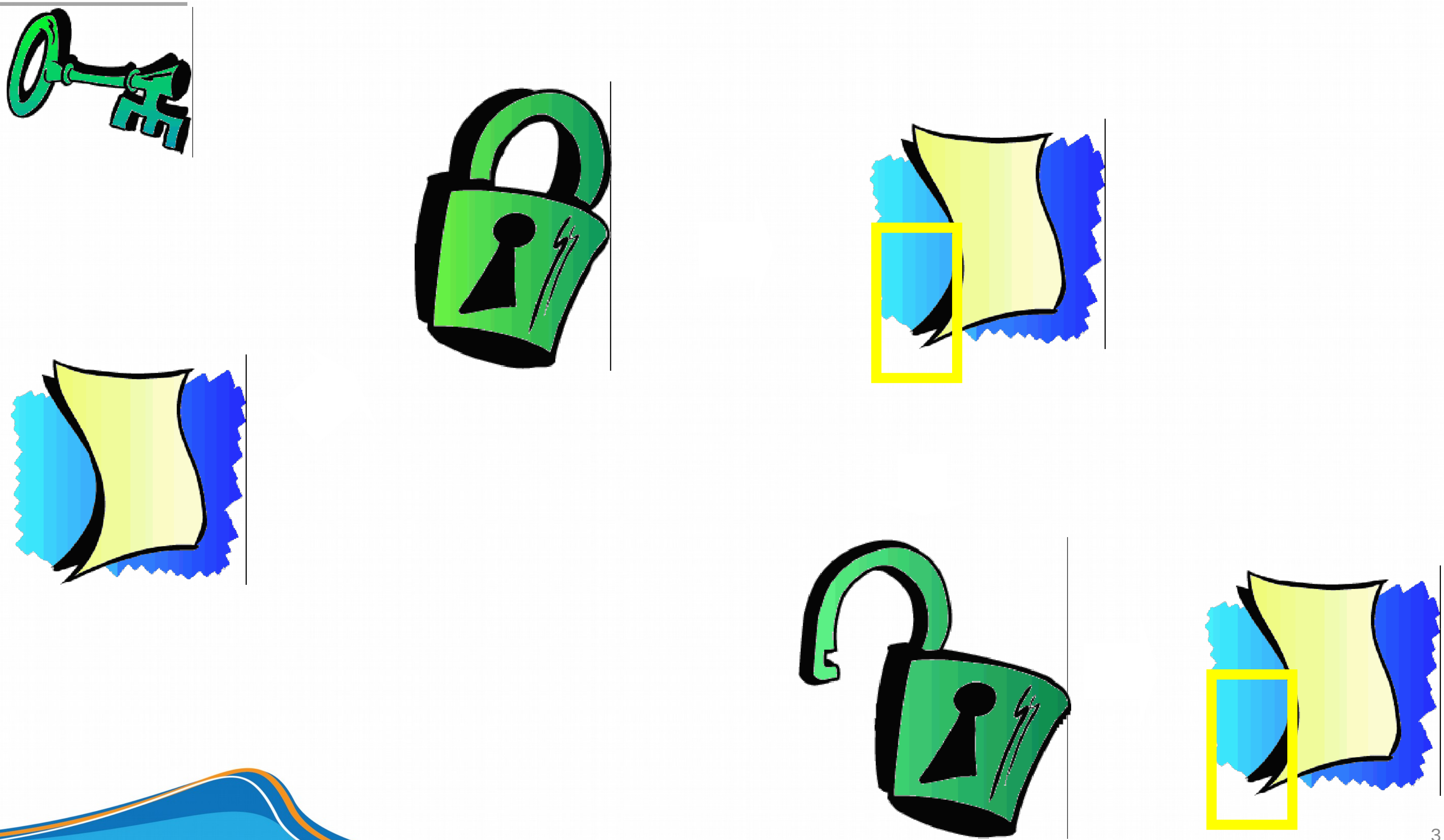
Assoc. Prof. Trần Minh Triết
PhD. Trương Toàn Thịnh

fit@hcmus
KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Introduction

- Symmetric cryptosystem
  - Conventional cryptosystem
  - An encryption system in which the encryption and decryption processes both use the same key – *secret key*.
  - The security of information depends on the security of the key.
- Traditional methods use:
  - Substitution: replace 1 word/character with another word/character
  - Transposition: characters are changed their positions
- Substitution/Transposition can be done with:
  - Mono-alphabetic
  - Poly-alphabetic

Symmetric cipher



Shift cipher

- **Shift Cipher:**
  - One of the oldest methods used for encryption
  - The message is encrypted by rotating each character by *k* places in the alphabet
  - The case with $k = 3$ is called the Caesar encryption method.
- Let $\mathsf{P} = \mathsf{C} = \mathsf{K} = Z_n$. For each $k \in \mathsf{K}$ we have:
  - $e_k(x) = x + k \bmod n$ and $d_k(y) = y - k \bmod n$, for $x, y \in Z_n$
  - $\mathsf{E} = \{e_k, k \in \mathsf{K}\}$ and $\mathsf{D} = \{d_k, k \in \mathsf{K}\}$
- Properties:
  - Simple
  - Encryption and decryption processing is done quickly
  - Key-space $K = \{0, 1, 2, \ldots, n - 1\} = Z_n$
  - Easily broken by trying every possible key

Shift cipher

- Example: to encrypte a message represented by the letters A to Z (26 letters), we use Z_{26}.
  - Encrypted messages are not secure and can be easily decrypted by trying one after another, 26 *keys*.
  - On average, an encrypted message can be decrypted in about 26/2 = 13 tries.
- Ciphertexts: JBCRCLQRWCRVNBJENBWRWN
- Try $k = 0, 1, 2, \ldots 25$

k = 0	**jbcrclqrwcrvnbjenbwrwn**	k = 5	**ewxmxglmrxmqiweziwrmri**
k = 1	**iabqbkpqvbqumaidmavqvm**	k = 6	**dvwlwfklqwlphvdyhvqlqh**
k = 2	**hzapajopuaptlzhclzupul**	k = 7	**cuvkvejkpvkogucxgupkpg**
k = 3	**gyzozinotzoskygbkytotk**	k = 8	**btujudijoujnftbwftojof**
k = 4	**fxynyhmnsynrjxfajxsnsj**	k = 9	**astitchintimesavesnine**

Substitution cipher

- **Substitution Cipher**:
  - Well-known and widely used encryption method for hundreds of years
  - Encrypt the message by permuting the elements of the alphabet or, more generally, permuting the elements in the source set P.
- Let P = C = Z_n, K are the sets of permutations of n elements 0, 1, …, n – 1. So, for each $\pi \in$ K, a permutation of n elements 0, 1, …, n – 1. For each key $\pi \in$ K, define:
  - $e_\pi(x) = \pi(x)$ and $d_\pi(y) = \pi^{-1}(y)$, for $x, y \in Z_n$
  - E = $\{e_\pi, \pi \in$ K$\}$ and D = $\{d_\pi, \pi \in$ K$\}$

Really secure???

- Properties:
  - Simple, encryption and decryption are done quickly
  - Key-space K has *n! keys*
  - Overcoming the limitation of the Shift-Cipher method: It is impossible to attack by exhausting the key values $k \in K$

Substitution cipher

AO VCO JO IBU RIBU

AO VCO JO I...

?A H?A ?A ...

MA HOA VA UNG DUNG

Attacks based on the occurrence of characters in the language

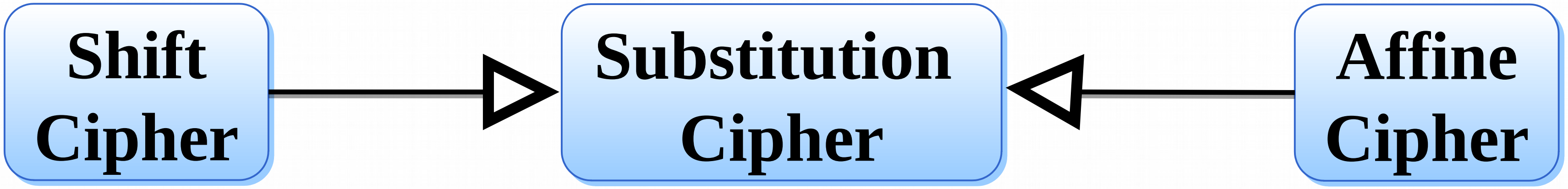
Substitution cipher

```
L FDPH L VDZ L FRQTXHUHG
L FDPH L VDZ L FRQTXHUHG
i ?a?e i ?a? i ?????e?e?
i came i saw i conquered
```

- Frequency analysis
  - Character: E > T > R > N > I > O > A > S
  - Digraph: TH > HE > IN > ER > RE > ON > AN > EN
  - Trigraph: THE > AND > TIO > ATI > FOR > THA > TER > RES

Affine cipher

- Let $P = C = Z_n$, $K = \{(a, b) \in Z_n \times Z_n: gcd(a, n) = 1\}$. For each key $k = (a, b) \in K$, define:
  - $e_k(x) = (ax + b) \bmod n$ and $d_k(y) = a^{-1}(y - b) \bmod n$, for $x, y \in Z_n$
  - $E = \{e_k, k \in K\}$ and $D = \{d_k, k \in K\}$
- For correct decrypt then e_k must be a bijection $\Rightarrow gcd(a, n) = 1$



Affine cipher

- Let $\varphi(n)$ be a number of elements in Z_n and coprime with n
- If $n = \prod_{i=1}^{m} p_i^{e_i}$ where p_i are distinct prime numbers and $e_i \in Z^+$, $1 \le i \le m$ then $\varphi(n) = \prod_{i=1}^{m} p_i^{e_i} - p_i^{e_i-1}$
- We have
  - m ways of choosing b
  - $\varphi(m)$ ways of choosing a
  - $m \times \varphi(n)$ ways of choosing key $k = (a, b)$

Euclide algorithm

- Consider 2 prime numbers *a* and *b* ($a > b$) we have:
  - $a = q_0 b + r_0$ $(0 < r_0 < b)$
  - $b = q_1 r_0 + r_1$ $(0 < r_1 < r_0)$
  - $r_0 = q_2 r_1 + r_2$ $(0 < r_2 < r_1)$
  - $r_1 = q_3 r_2 + r_3$ $(0 < r_3 < r_2)$
  - …
  - $r_{m-2} = q_{m-1} r_{m-1} + r_m$ $(0 < r_m < r_{m-1})$
  - $r_{m-1} = q_m r_m$ $(0 = r_{m+1} < r_m)$
- Easily see:
  - $gcd(a, b) = gcd(b, r_0) = gcd(r_0, r_1) = \ldots = gcd(r_{m-1}, r_m) = r_m$.
  - Example: $gcd(1071, 462) = gcd(462, 147) = gcd(147, 21) = 21$

https://en.wikipedia.org/wiki/Euclidean_algorithm

Extended Euclide algorithm

- Consider 2 prime numbers *a* and *b*, let's build
  - $r_0 = a$ $\quad r_1 = b$
  - $s_0 = 1$ $\quad s_1 = 0$
  - $t_0 = 0$ $\quad t_1 = 1$
- So, we have:
  - $r_2 = r_0 - q_0 r_1$
  - $s_2 = s_0 - q_0 s_1$
  - $t_2 = t_0 - q_0 t_1$
  - …
  - $r_{i+1} = r_{i-1} - q_i r_i \ (i \geq 1)$
  - $s_{i+1} = s_{i-1} - q_i s_i$
  - $t_{i+1} = t_{i-1} - q_i t_i$

Algorithm stops when $r_{k+1} = 0$ & $r_k = gcd(a, b) = as_k + bt_k$

Extended Euclide algorithm

- Example $a = r_0 = 240$ and $b = r_1 = 46$

i	q_{i-1}	r_i	s_i	t_i
0		240	1	0
1		46	0	1
2	240 / 46 = 5	240 − 5 × 46 = 10	1 − 5 × 0 = 1	0 − 5 × 1 = −5
3	46 / 10 = 4	46 − 4 × 10 = 6	0 − 4 × 1 = −4	1 − 4 × −5 = 21
4	10 / 6 = 1	10 − 1 × 6 = 4	1 − 1 × −4 = 5	−5 − 1 × 21 = −26
5	6 / 4 = 1	6 − 1 × 4 = 2	−4 − 1 × 5 = −9	21 − 1 × −26 = 47
6	4 / 2 = 2	4 − 2 × 2 = 0	5 − 2 × −9 = 23	−26 − 2 × 47 = −120

  - Line 6 finds stop-condition $r_6 = 0$
  - Result: $gcd(240, 46) = 2 = -9 \times 240 + 47 \times 46$
- Note: $gcd(a, b) = 1 = as + bt$ $(a \perp b)$
  - If $bt \bmod a \cong 1 \Rightarrow t = b^{-1} \bmod a$
  - If $as \bmod b \cong 1 \Rightarrow s = a^{-1} \bmod b$

https://en.wikipedia.org/wiki/Extended_Euclidean_algorithm

Vigenere cipher

- Choose a positive integer m. Let $\mathcal{P} = \mathcal{C} = \mathcal{K} = (Z_n)^m$.
  - $\mathcal{K} = \{(k_1, k_2, \ldots, k_m) \in (Z_n)^m\}$
  - For each key $k = (k_1, k_2, \ldots, k_m) \in \mathcal{K}$ and $\forall x, y \in (Z_n)^m$, define:
    - $e_k(x_1, x_2, \ldots, x_m) = ((x_1 + k_1) \bmod n, (x_2 + k_2) \bmod n, \ldots, (x_m + k_m) \bmod n)$
    - $d_k(y_1, y_2, \ldots, y_m) = ((y_1 - k_1) \bmod n, (y_2 - k_2) \bmod n, \ldots, (y_m - k_m) \bmod n)$
- Subtitution cipher: for each key k, plain-text $x \in P$ is mapped to only one $y \in C$.
- Vigenere cipher uses key with length m.
  - Named after Blaise de Vigenere (Century 16)
  - The Vigenere cipher can be viewed as consisting of m displacement ciphers that are applied alternately on a periodic basis.
  - Key-space K of Vigenere cipher is n^m
  - For example: $n = 26$, $m = 5$ then key-space has $\sim 1.1 \times 10^7$ keys

Vigenere cipher

- Example: $m = 6$ and keyword CIPHER
- Then, key $k = (2, 8, 15, 7, 4, 17)$
- Let plaintexts: **thiscryptosystemisnotsecure**

t	h	i	s	c	r
19	7	8	18	2	17
2	8	15	7	4	17
21	15	23	25	6	8

y	p	t	o	s	y
24	15	19	14	18	24
2	8	15	7	4	17
0	23	8	21	22	15

s	t	e	m	i	s
18	19	4	12	8	18
2	8	15	7	4	17
20	1	19	19	12	9

n	o	t	s	e	c
13	14	19	18	4	2
2	8	15	7	4	17
15	22	8	25	8	19

u	r	e
20	17	4
2	8	15
22	25	19

Hill cipher

- Hill cipher (1929), author: Lester S. Hill
- Main idea: use m linear-combinations of m characters in plaintext to produce m characters in ciphertext
- Example:

$$(y_1, y_2) = (x_1, x_2) \times \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} \Rightarrow \begin{cases} y_1 = 11x_1 + 3x_2 \\ y_2 = 8x_1 + 7x_2 \end{cases}$$

$$(x_1, x_2) = (y_1, y_2) \times \begin{pmatrix} \frac{7}{53} & \frac{-8}{53} \\ \frac{-3}{53} & \frac{11}{53} \end{pmatrix} \Rightarrow \begin{cases} x_1 = \frac{7}{53} y_1 + \frac{-3}{53} y_2 \\ x_2 = \frac{-8}{53} y_1 + \frac{11}{53} y_2 \end{cases}$$

Hill cipher

- Choose a positive integer m. Define
  - $\mathcal{P} = \mathcal{C} = (Z_n)^m$
  - $\mathcal{K}$ is a set of inverse matrixes $m \times m$, for each key $k \in \mathcal{K}$, Let:
    - $e_k(x) = xk = (x_1, x_2, \ldots, x_m) \times$ where $x = (x_1, x_2, \ldots, x_m) \in \mathcal{P}$
    - $d_k(y) = yk^{-1}$ where $y \in \mathcal{C}$.
  - All arithmetic operations are performed on Z_n

- $e_k(x) = xk = (x_1, x_2, \ldots, x_m) \times \begin{pmatrix} k_{1,1} & \cdots & k_{1,m} \\ \vdots & \ddots & \vdots \\ k_{m,1} & \cdots & k_{m,m} \end{pmatrix}$ where $x = (x_1, x_2, \ldots, x_m) \in \mathcal{P}$
- $d_k(y) = yk^{-1}$ where $y \in \mathcal{C}$.
- All arithmetic operations are performed on Z_n

Inverse matrix

- Let inverse matrix K, define K^{-1}
- Steps:
  - Convert from matrix $(K \mid I_n)$ to $(I_n \mid K^{-1})$
  - Elementary transformations:
    - Multiply 1 line by 1 a number $\neq 0$
    - Replace 1 line by using that line adding/subtracting α times to/from other lines

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[(1)\to(1)-(2)]{(3)\to(3)-1} \left(\begin{array}{ccc|ccc} 2 & 0 & -4 & 1 & -1 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 & 0 & 1 \end{array}\right) \xrightarrow[\substack{(1)\to(1)+2(3) \\ (3)\to0.5(1) \\ (2)\to(2)-2(3)}]{}$$

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & -1 & -1 & 2 \\ 0 & 1 & 0 & \frac{3}{2} & 1 & \frac{-3}{2} \\ 0 & 0 & 1 & \frac{-1}{2} & 0 & \frac{1}{2} \end{array}\right) \xrightarrow[(1)\to0.5(1)]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-1}{2} & \frac{-1}{2} & 1 \\ 0 & 1 & 0 & \frac{3}{2} & 1 & \frac{-3}{2} \\ 0 & 0 & 1 & \frac{-1}{2} & 0 & \frac{1}{2} \end{array}\right)$$

Permutation cipher

- The idea of the presented methods: replace each character in the source message with another character to form the encrypted message.
- The main idea of the Permutation Cipher method is to keep the characters in the source message the same, but only change the position of the characters.
- Choose a positive integer m. Let
  - $\mathcal{P} = \mathcal{C} = (\mathbb{Z}_n)^m$
  - $\mathcal{K}$ is a set of permutations of m elements $\{1, 2, \ldots, m\}$. For each key $\pi \in \mathcal{K}$, define:
    - $e_\pi(x_1, x_2, \ldots, x_m) = (x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(m)})$
    - $d_\pi(y_1, y_2, \ldots, y_m) = (y_{\pi^{-1}(1)}, y_{\pi^{-1}(2)}, \ldots, y_{\pi^{-1}(m)})$, where π^{-1} is an inverse permutation of π

Permutation cipher

- The permutation encryption method is a special case of Hill cipher.
- Example: choose $m \equiv 3$, so $\pi \equiv$

1	2	3
3	1	2

& $\pi^{-1} \equiv$

1	2	3
2	3	1

  - Let plain-text = EAT = (4, 0, 19)
    - Compute $(y_1, y_2, y_3) = (x_1, x_2, x_3) \times \pi = (4, 0, 19) \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = (19, 4, 0) =$ TEA
    - So, cipher-text = TEA
  - To decrypt to plain-text, we need an inverse matrix $\pi^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$
    - Compute $(x_1, x_2, x_3) = (y_1, y_2, y_3) \times \pi^{-1} = (19, 4, 0) \times \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = (4, 0, 19) =$ EAT

The permutation encryption method is a special case of Hill cipher.
Example: choose $m \equiv 3$, so $\pi \equiv$ & $\pi^{-1} \equiv$
Let plain-text = EAT = (4, 0, 19)
Compute $(y_1, y_2, y_3) = (x_1, x_2, x_3) \times \pi = (4, 0, 19) \times = (19, 4, 0) =$ TEA
So, cipher-text = TEA
To decrypt to plain-text, we need an inverse matrix $\pi^{-1} =$
Compute $(x_1, x_2, x_3) = (y_1, y_2, y_3) \times \pi^{-1} = (19, 4, 0) \times = (4, 0, 19) =$ EAT

Permutation cipher

- Example: choose $m = 6$
- So $\pi =$

1	2	3	4	5	6
3	5	1	6	4	2

& $\pi^{-1} =$

1	2	3	4	5	6
3	6	1	5	2	4

- Assume plain-texts = shesellsseashellsbytheseashore

s	h	e	s	e	l
e	e	s	l	s	h

l	s	s	e	a	s
s	a	l	s	e	s

h	e	l	l	s	b
l	s	h	b	l	e

y	t	h	e	s	e
h	s	y	e	e	t

a	s	h	o	r	e
h	r	a	e	o	s

- So, cipher-texts = eeslshsalseslshblehsyeethraeos