

Mật mã học: Phần 1 - Mã hóa Caesar

Khái niệm mã hóa dữ liệu và giải mã

Mã hóa dữ liệu là tiến trình che dấu dữ liệu thật (plaintext), nghĩa là chuyển dữ liệu thật thành dữ liệu không có ý nghĩa hoặc có ý nghĩa khác xa với dữ liệu thật. Tiến trình đó gọi là mã hóa (encryption). Kết quả của tiến trình gọi là bản mã (ciphertext). Từ "encryption" được tạo ra từ "cryptography" (mật mã) xuất phát từ tiếng Hi Lạp cổ xưa "Kryptos" (Che dấu) và từ "graphia" (viết). Tiến trình mã hóa dữ liệu có thể được thực hiện bằng cách hoán vị dữ liệu thật hoặc thay thế chúng bằng dữ liệu khác.

Tiến trình ngược với tiến trình mã hóa tức là chuyển từ bản mã thành dữ liệu ban đầu gọi là giải mã.

Hệ mã Caesar

Trong mật mã học, mật mã Caesar (Xê da), còn gọi là mật mã dịch chuyển, là một trong những mật mã đơn giản và được biết đến nhiều nhất. Hệ mã Caesar là một hệ mã hóa thay thế đơn âm, làm việc trên bảng chữ cái tiếng Anh 26 ký tự. Đó là một dạng của mật mã thay thế, trong đó mỗi ký tự trong văn bản được thay thế bằng một ký tự cách nó một đoạn trong bảng chữ cái để tạo thành bản mã.

Ví dụ:

Đối với bảng mã tiếng anh (ABCDEFGHIJKLMNOPQRSTUVWXYZ...), nếu độ dịch là 3, A sẽ được thay bằng D, B sẽ được thay bằng E, ..., W sẽ thay bằng Z, X sẽ thay bằng A, Y sẽ thay bằng B và Z thay bằng C. Phương pháp được đặt tên theo Caesar, vị hoàng đế đã sử dụng nó thường xuyên trong công việc. Không gian bản rõ P là các từ cần được mã hóa được tạo từ bảng chữ cái A. Không gian bản rõ C là các từ đã được mã hóa.

Để mã hóa, người ta đánh số các chữ cái từ $0 \text{ đến } N-1$ (N là tổng số phần tử của bản chữ cái). Không gian khóa $K = \mathbb{Z}_N$. Với mỗi khóa $K \in K$ hàm mã hóa và giải mã một ký tự có số thứ tự là i sẽ được biểu diễn như sau:

a. Mã hóa : $E_K(i) = (i+k) \bmod N$ b. Giải mã : $D_K(i) = (i-k) \bmod N$

Trong đó: $N = 26$ nếu hệ mã Caesar sử dụng trên bảng chữ cái tiếng Anh (nếu sử dụng trên bảng chữ cái khác thì N sẽ thay đổi). k : tương ứng với số thứ tự chữ cái trong bảng mã (ví dụ : $a = 0, b = 1 \dots$)

Xét ví dụ sau: Cho bản rõ : TOIYEUVIETNAM Khóa $k = 4$

Tìm bản mã ? Theo công thức thay vào ta sẽ có kết quả sau.

Bản rõ	T	O	I	Y	E	U	V	I	E	T	N	A	M
i	19	14	8	23	4	20	21	8	4	19	13	0	12
$(i+k) \bmod N$	23	18	12	1	8	24	25	12	8	23	17	4	16
Bản mã	Y	S	M	B	Y	Z	W	M	I	Y	R	E	Q

Vậy bản mã là : YSMBYZWMIYREQ

Quá trình giải mã thì ngược lại với quá trình mã hóa.

Cài đặt

Mỗi ngôn ngữ khác nhau sẽ có cách cài đặt khác nhau. Mình sẽ sử dụng Python để cài đặt thuật toán:

Mặc định bạn có thể dùng bảng chữ cái Tiếng Anh nhưng trong phần cài đặt này. Để thực tế hơn. Mình sẽ sử dụng bảng mã Tiếng Việt.

1. Bảng mã

[illegible]

1. Hàm mã hóa

```
def encrypt(self, n, plaintext):
    """Encrypt the string and return the ciphertext"""
    result = ''
```

```

for l in plaintext:
    try:
        i = (self.key.index(l) + n) % len(self.key)
        result += self.key[i]
    except ValueError:
        result += l
return result

```

Nếu đã đọc xong đồng lý thuyết bên trên thì chắc hẳn bạn sẽ không gặp khó khăn gì trong đoạn code cài đặt trên:

Mã hóa : $EK(i) = (i+k) \bmod N$

Trong đó:

- `self.key.index(l)` : tương ứng với `i`
- `n`: tương ứng với khóa `k`
- `len(self.key)`: tương ứng với độ dài của bản mã.

2. Hàm giải mã

```

def decrypt(self, n, ciphertext):
    """Decrypt the string and return the plaintext"""
    result = ''

    for l in ciphertext:
        try:
            i = (self.key.index(l) - n) % len(self.key)
            result += self.key[i]
        except ValueError:
            result += l

    return result

```

Áp dụng với công thức giải mã. Ta cũng sẽ viết được hàm tương tự.

3. Gọi hàm

```

def show_result(self, plaintext, n):
    """Generate a resulting cipher with elements shown"""
    encrypted = self.encrypt(n, plaintext)
    decrypted = self.decrypt(n, encrypted)

```

```
print('Rotation: ', n)
print('Plaintext: ', plaintext)
print('Encrytped: ', encrypted)
print('Decrytped: ', decrypted)
```

Tài liệu

Đến đây bài viết phần đầu về mã hóa cũng kết thúc. Mong rằng anh chị em đừng gạch đá nhiều quá. Và đưa ra những góp ý nhiều hơn để em có thể vững tâm trong quá trình tìm hiểu các hệ mã hóa .

Mọi người nếu lười cài đặt cũng có thể tham khảo link github bài em đã làm sẵn. Chỉ việc kéo về và chạy <https://github.com/tuanquanghpvn/cipher>