Tên: Nguyễn Tùng Lâm

MSSV: 20204998

***Hướng đi và phân tích bài many – time pad attack:***

* 1 ký tự thường được encode dưới ASCII ⬄ nó sử dụng 1 byte và là 1 số thuộc [0;255], và trong hệ 16 nó sẽ chạy từ 0x00 -> 0xFF. Hướng đi của bài em sẽ đi theo hướng sau:
* Giả sử ta có 2 msg(1) msg(2) và key (với msg là thông tin gốc, key là khóa và cip là t.tin đã mã hóa):

msg(1) ^ key = cip(1)

msg(2) ^ key = cip(2)

cip(1) ^ cip(2) = msg(1) ^ key \* msg(2) ^ key = msg(1) ^ msg(2)

* Như vậy ta thấy khi XOR 2 bản mã ta sẽ được kết quả tương đương với 2 bản gốc XOR với nhau. Như vậy trong các bước tiếp được nêu sẽ liên hệ mật thiết với bản gốc (dù ta làm việc chính trên bản mã). Mặt khác:

Key = msg ^ cip

* Nhận thấy thêm: 1 văn bản của chúng ta sẽ thường gồm các ký tự a -> z và A -> Z và dấu “ “,

Từ gợi ý của thầy e nhận thấy khi XOR 1 ký tự (hoặc thuộc a -> z hoặc A -> Z ) thì nếu a -> Z sẽ thành ký tự thuộc A -> Z và ngược lại. Tại sao? Do ký tự space trong bảng mã ASCII là 32 (dạng 10) như vậy khi XOR ta có – VD: ‘A’ – 65 – 01000001 – có thể thấy vị trí 2 mũ 5 trong mã 2 của ‘A’ là = 0 => khi XOR vs space – 32 – 00100000 ta được 01100001 = 65 + 32 (ở hệ 10) = 97 = ‘a’…)

* Ta nhận thấy 1 điều là ký tự space và các chữ cái đều có giá trị mã 10 gồm 2 chữ số trở lên, như vậy ta sẽ xét 2 ký tự trong bản mã (2 ký tự bản mã ta được 1 ký tự) VD: 315c4e…
* Quay lại ở trên, mục đích của việc XOR là gì? Mục đích là để tìm vị trí các dấu cách – ký tự space
* Tìm ký tự space thì được gì? Do sự đặc biệt của ký tự space (đã nêu trên), ta sẽ sử dụng ký tự space để tìm khóa. VD: ta có 2 ciphertexts sau

48A6483C

76920AA2

XOR ký tự in đỏ: 0x48 ^ 0x0A = 0x42

Trong những ký tự có thể xuất hiện khi XOR với ký tự space, ta thấy 0x42 có trong đó. Điều này có nghĩa là: khi đó tại vị trí ta đang xét mà có thể xuất hiện ký tự space chỉ có 2 trường hợp xra: hoặc ký tự đó là chữ cái hoặc ký tự đó là space (cụ thể hoặc đó là ký tự “ “ – 0x20 hoặc là chữ cái ‘b’ – 0x62).

* Bởi vì key có thể ngẫu nhiên và đa dạng nhiều trường hợp nên ta sẽ tập trung vào tìm vị trí dấu cách sẽ dễ hơn. Tuy nhiên ta sẽ ko chỉ XOR 2 bản mã mà ta sẽ lặp lại quá trình này với tất cả bản mã. Mục đích là để tìm tần suất xuất hiện của dấu cách (VD: ta phát hiện ra sau quá trình XOR vị trí số 7 có tần số xhien phím space ~ 7 – 8.. lần) thì khi đó ta có thể khẳng định ký tự space đã được encrypted.
* Khẳng định ký tự space đã được encrypted tức là ta đã biết được ký tự trong msg. Như vậy ta XOR space với bản mã ở vị trí tương ứng ta sẽ được key (Key = msg ^ cip với msg là “ “)
* Vậy các bước thực hiện là gì?
* B1: Chuyển từ hex về mã 10.
* B2: ..

Các bước em đã chú thích trong code. Sau tất cả kết quả cuối cùng ta được (kết hợp code chạy bằng máy và giải mã bằng cơm):

Máy:

Text

Description automatically generated

Kết quả cuối:

The secret message is: When using a stream cipher, never use the key more than once :D