I. Nguyên tắc giải mã :

Với a và b là các bản rõ, xa và xb là các bản mã, ta có a xor b = xa xor xb.

Thật vậy để chứng minh ta có: giả sử ta có khóa k khi đó

a xor k = xa

b xor k = xb

=> xa xor xb = a xor k xor b xor k = a xor b xor 0 = a xor b

- Khi xor khoảng trắng với một chữ cái, ta được kết quả là chữ cái đó sẽ được viết hoa (up case) hoặc viết thường (low case). Còn nếu xor 2 khoảng trắng với nhau ta sẽ ra 0 (chỉ quan tâm đến chữ cái và khoảng trắng vì tần suất xuất hiện lớn và mọi ký tự khác xor với chữ cái đều không tạo ra được chữ cái)

Ta có cách giải mã khóa như sau : xor lần lượt từng cipher text với nhau, đối với từng vị trí (8 bit tương đương với 2 ký tự hex) mà kết quả là một chữ cái hoặc số 0 khi đó ta biết một trong hai cipher text đó có một cái là mã hóa của space tính khả năng xuất hiện space đối với từng cipher text, lấy cái cao nhất rồi suy ngược ra khóa. Sau khi có key ta đem xor với cipher text (cipher text11) sẽ ra được message.

II. Thuật toán :

1.

- Đọc các cipher text từ file data.txt lưu vào mảng

- Đọc target cipher từ file target.txt

2. Xor lần lượt từng cipher text với nhau (không lặp lại). Giả sử cipher text là i,j. Nếu kết quả của phép xor là một chữ cái (‘a’ -> ‘z’ hoặc ‘A’ -> ‘Z’) hoặc 0 thì tăng is_space[i][vị trí] lên một đơn vị.

3. Sinh key : Giả sử cho key có độ dài bằng cipher text dài nhất. Với mỗi vị trí, tìm cipher text sao cho is_space của cipher text lớn nhất tại vị trí đó (Giá trị is_space >= 7 để đảm bảo tính chính xác). Ta tìm được cipher text(max_index) mà tại vị trí đó có khả năng là mã hóa dấu space nhất khi này để tìm key tại vị trí này chỉ cần lấy cipher_text[max_index][vị trí] xor 32 (khoảng trắng)

4. Sau khi tìm được khóa (khóa dài nhất có thể) thì lấy khóa xor với cipher text 11 thì ra được message cần tìm

5. Khóa tìm được sẽ không đúng 100% (do các ký tự đặc biệt, hoặc là do không tìm được toàn bộ key) nên cần sửa lại khóa sau khi đã đoán được message cho đúng. Message chương trình cho ra sẽ đánh dấu các vị trí không biết(vì key bị khuyết) bằng ký tự “*”.

< Các bản mã, key khi thực hiện XOR đều thực hiện ở dạng nhị phân>

III. Áp dụng vào bài toán

- Sau khi chạy chương trình (ManyTimePad.cpp) cho ra kết quả bản rõ là:

The secuet message*is: Wh*n using a ~tream cipher, never*use the key more than once

- Chỉnh sửa thủ công kết quả chương trình cho ra ta được bản rõ hoàn chỉnh như sau:

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