Đại học Quốc gia Việt Nam – Thành phố Hồ Chí Minh
Đại học Công nghệ Thông tin - ĐHQG-HCM

3

Phòng thí nghiệm

CHỈ DÀNH CHO MỤC ĐÍ CH GIÁO DỤC

Mật mã cổ điển

Giới thiệu về Đảm bảo và Bảo mật Thông tin

tháng 3 năm 2023

Mọi ý kiến phản hồi về tài liệu xin vui lòng gửi về email inseclab@uit.edu.vn

A.TỔNG QUAN

1. Giới thiệu và mục tiêu học tập

Mật mã đóng một vai trò quan trọng trong các hệ thống truyền thông kỹ thuật số hiện đại. Từ điển Oxford1 định nghĩa mật mã là "nghệ thuật viết hoặc giải mã" với "mật mã" ở nơi khác đư ợc định nghĩa là "một hệ thống các tín hiệu đư ợc sắp xếp trư ớc, đặc biệt đư ợc sử dụng để đảm bảo bí mật trong việc truyền tải Về mặt lịch sử , mật mã chỉ tập trung vào việc đảm bảo liên lạc riêng tư giữa hai bên chia sẻ thông tin bí mật trư ớc bằng cách sử dụng "mã". Nó đư ợc sử dụng chủ yếu cho các ứng dụng quân sự, chính phủ và một số ứng dụng thích hợp trong công nghiệp trong nhiều thế kỷ.

Như ng mật mã ngày nay thiên về khoa học nhiều hơn. Chúng ta có thể nói rằng mật mã hiện đại liên quan đến “nghiên cứu các kỹ thuật toán học để bảo mật thông tin, hệ thống và tính toán phân tán chống lại sự tấn công của kẻ thù”.

Mật mã học đã đi từ "một loại hình nghệ thuật xử lý thông tin liên lạc bí mật cho quân đội" thành "một ngành khoa học bảo mật hệ thống cho ngư ời dân bình thư ờng trên toàn cầu". Nó đề cập đến các cơ chế đảm bảo tính toàn vẹn, kỹ thuật trao đổi khóa bí mật, giao thức xác thực ngư ời dùng, đấu giá và bầu cử điện tử, đảm bảo tiền tệ kỹ thuật số, v.v.

Mục tiêu học tập của phòng thí nghiệm này là giúp học sinh làm quen với các khái niệm trong mã hóa khóa bí mật, đặc biệt là mật mã cổ điển, sau khi học xong lab, sinh viên cần có đư ợc trải nghiệm trực tiếp về các thuật toán mã hóa.

Hơn nữa, sinh viên có thể sử dụng các công cụ mật mã và viết các chư ơng trình đơn giản để mã hóa/giải mã tin nhắn. Phòng thí nghiệm này sẽ đề cập đến các chủ đề sau về mật mã cổ điển:

- o Mật mã thay thế dùng một bảng chữ cái - Phân tích tần số.
- o Mật mã đa bảng chữ cái.
- o Mật mã hoán vị.

2. Bối cảnh và điều kiện tiên quyết

Để đảm bảo mọi thứ diễn ra suôn sẻ và bạn đạt đư ợc kết quả tốt hơn trong phòng thí nghiệm này, bạn phải làm quen với các khái niệm mật mã và có đủ kiến thức nền tảng về mật mã đối xứng, đặc biệt là trong mật mã cổ điển.

Ngoài ra, kỹ năng lập trình (với ngôn ngữ ư a thích của bạn, ví dụ: Python, C/C++, Golang) cũng cần thiết.

[1] https://www.oxfordlearnersdictionaries.com/

Phần dưới đây tóm tắt một số khía cạnh chính giúp bạn ôn lại những kiến thức có thể bạn đã có trước khi bắt đầu. Bạn cũng có thể đọc sách giáo khoa và các tài liệu tham khảo liên quan mà tôi liệt kê ở phần 1.4 để biết thêm chi tiết.

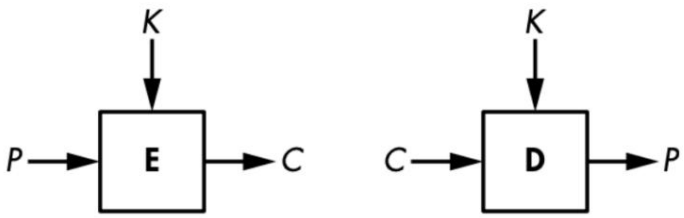
Khái niệm Trước khi bắt đầu, chúng ta định nghĩa một số thuật ngữ. Khi chúng ta mã hóa một tin nhắn, Bản rõ (p) đề cập đến tin nhắn gốc hoặc không được mã hóa. Bản mã (C) đề cập đến tin nhắn được mã hóa hoặc mã hóa. Do đó, mật mã bao gồm hai chức năng: Mã hóa hoặc Mã hóa (E) biến bản rõ thành bản mã Giải mã hoặc Giải mã (D) biến bản mã trở lại bản rõ.

Viết dưới dạng hàm:

$$C = E(Ke,p)$$

$$p = D(Kd, C)$$

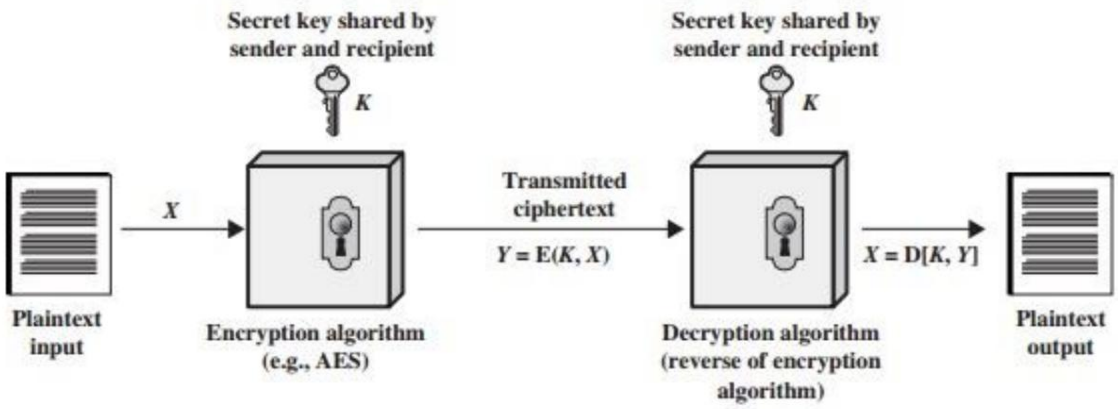
trong đó Ke và Kd lần lượt là khóa mã hóa và khóa giải mã



Hình 1. Mã hóa và giải mã cơ bản.

Các kỹ thuật được sử dụng để giải mã một tin nhắn mà không có bất kỳ kiến thức nào về các chi tiết mã hóa thuộc lĩnh vực phân tích mật mã (phá mã). Các lĩnh vực mật mã và phân tích mật mã cùng nhau được gọi là mật mã học.

Phân loại



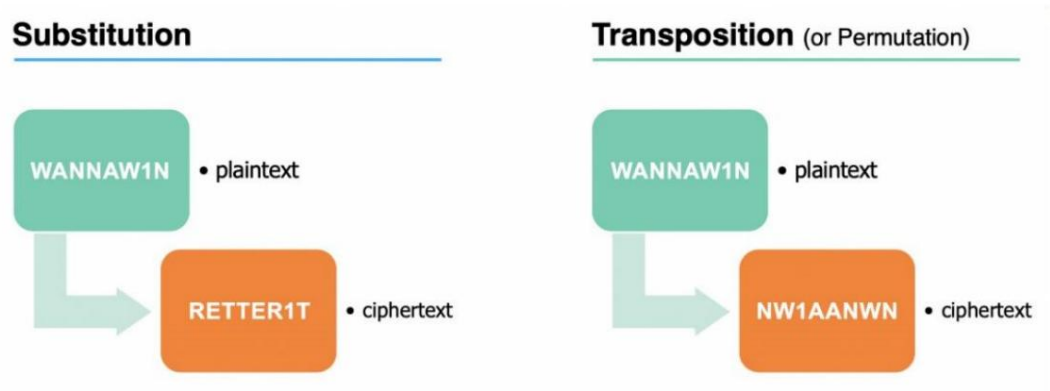
Hình 2. Mô hình mã hóa đối xứng.

Về phương pháp mã hóa, có hai loại mật mã:

1. Mật mã đối xứng (hoặc mật mã khóa bí mật): Sử dụng một khóa để mã hóa và giải mã, là loại mã hóa duy nhất được sử dụng trước khi mã hóa khóa công khai phát triển vào những năm 1970 (Hình 2).
2. Mật mã bất đối xứng (hoặc mật mã khóa công khai): Sử dụng 2 khóa liên quan là khóa chung và khóa riêng, dùng để thực hiện các hoạt động bổ sung cho nhau như mã hóa và giải mã hoặc tạo chữ ký và xác minh chữ ký.

Mật mã cổ điển là mật mã có trước máy tính, do đó hoạt động trên các chữ cái thay vì trên bit và gần như là mật mã đối xứng.

Hai khối xây dựng cơ bản của tất cả các kỹ thuật mã hóa là thay thế và chuyển vị (Hình 3).



Hình 3. Ví dụ về thay thế và chuyển vị.

1. Mật mã thay thế: thay các chữ cái trong bản rõ bằng các chữ cái khác hoặc bằng số hoặc ký hiệu.
2. Mật mã chuyển vị (hoặc hoán vị) : thực hiện một số loại hoán vị (hoặc thay đổi thứ tự) trên các chữ cái trong văn bản gốc.

B. NHIỆM VỤ THÍ NGHIỆM

1. Khởi động: Giải mã

Nhiệm vụ 1: Hãy bắt đầu với một nhiệm vụ đơn giản không sử dụng bất kỳ thuật toán mã hóa nào. Hãy thử giải các mã sau:

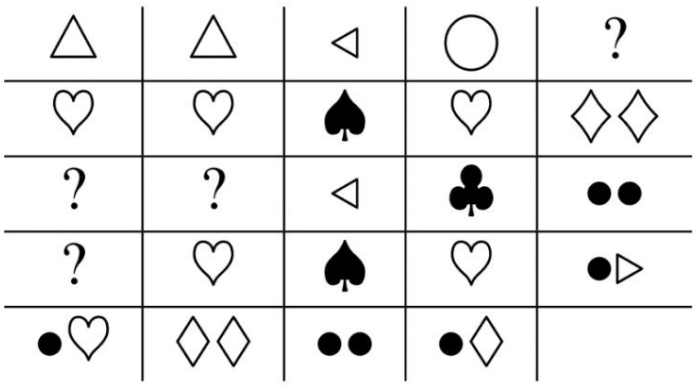
1. Chúng ta cần tìm mật mã để mở ổ khóa như hình 4. Ổ khóa có mã PIN gồm 3 chữ số thỏa mãn 5 điều kiện (gợi ý) được đưa ra. Bạn có thể bẻ khóa mã này được không? Nếu có thể hãy giải thích cách thực hiện.
2. Tìm bảng mã tương ứng cho các số từ 1 đến 9 theo gợi ý ở bảng 1.2.1

- Mỗi ký hiệu trong bộ (•) mã hóa duy nhất cho một trong các số từ 1 đến 9.
- Cột ngoài cùng bên phải là tổng các số ở mỗi hàng.
- Hàng dưới cùng là tổng các số ở mỗi cột.

- Mỗi ? có thể đại diện cho bất kỳ số có một hoặc hai chữ số nào và có thể giống hoặc khác nhau từ nhau.

Hình 4. Crack mã để mở khóa.

Bảng 1. Tìm bảng mã tương ứng cho mỗi số.



2. Mật mã Caesar

Nhiệm vụ 2: Trong nhiệm vụ này, bạn phải tạo một ứng dụng bằng ngôn ngữ lập trình đã chọn để thực hiện mã hóa và giải mã bằng mật mã Caesar . Ứng dụng phải đáp ứng các tiêu chí sau:

- Cho phép người dùng nhập vào khóa và văn bản gốc để mã hóa hoặc văn bản mã hóa để giải mã.

Cung cấp khả năng thực hiện một cuộc tấn công vũ phu bằng cách thử tất cả các khóa có thể để giải mã một bản mã nhất định mà không cần biết khóa.

- Để xác thực chương trình của bạn, hãy kiểm tra chương trình bằng một thông báo ít nhất 100 từ và so sánh kết quả với các công cụ mã hóa trực tuyến khác, chẳng hạn như dcode (https://www.dcode.fr), CrypTool 2 trực tuyến. Ngoài ra, hãy sử dụng chương trình của bạn để bẻ khóa văn bản mật mã sau:

Mfwzpn Rzwfpfrn bfx gtws ns Pdtyt ns 1949 fsi stb qnajx sjfw Ytpdt. Mj nx ymj fzymtw tk rfsd stajqx fx bjqq fx x mtwy xytwnjx fsi sts-knhynts. Mnx btwpx nshqzij Stwbjlnf s Btti , Ymj Bnsi-Zu Gnwi Hmwtsnhqj, Pfkpf ts ymj Xmtwj , Fkyjw Mnx btwp mfx gjjs ywfsxqfyji nsyt rtwj ymfs ktwyd qfslzfljx, fsi ymj rtxy wjhjsy tk mnx rfsd nsyjwsfyntsfq mt stzwx nx ymj Ojwzxfqjr Uwnej, bmtxj uwjantzx wjhnunny x nshqzij HOẶC Htjyejj, Rnqfs Pzsijwf, fsi AX Sfnufzq .

Bạn có tìm thấy điều gì đặc biệt liên quan đến khóa dùng để mã hóa bản mã này không?

Lời khuyên: Sử dụng thuật toán Caesar

Mã hóa: C = E(k, p) = (p + k) mod 26

Giải mã: p = D(k,C) = (C- k) mod 26

trong đó C = Bản mã, p = bản rõ, k là khóa

3. Mật mã thay thế đơn chữ cái và phân tích tần số

(Nhiệm vụ này dựa trên tài liệu phòng thí nghiệm SEED Labs của Wenliang Du, Đại học Syracuse.)

Ngư ời ta biết rằng mật mã thay thế dùng một bảng chữ cái (còn đư ợc gọi là mật mã dùng một bảng chữ cái) không an toàn vì nó có thể phải chịu sự phân tích tần số.

Bạn đư ợc cung cấp một văn bản mật mã đư ợc mã hóa bằng mật mã dùng một bảng chữ cái trong phòng thí nghiệm này. Mỗi chữ cái trong văn bản gốc đư ợc thay thế bằng một chữ cái khác, trong đó sự thay thế không thay đổi (tức là một chữ cái luôn đư ợc thay thế bằng cùng một chữ cái trong quá trình mã hóa.

Nhiệm vụ 3: Bấm vào đây để tải xuống tệp văn bản mã hóa 2.

Công việc của bạn là tìm ra văn bản gốc bằng phư ơng pháp phân tích tần suất, đư ợc biết văn bản gốc là một bài báo tiếng Anh.

Mô tả cách tìm văn bản thuần túy một cách chi tiết (từng bư ớc).

Lư u ý bạn không đư ợc phép sử dụng chế độ tự động của bất kỳ công cụ nào (như CrypTool 2, dCode, quipqiup) để giải mã.

Mẹo: Sử dụng phân tích tần số, bạn có thể tìm ra văn bản thuần túy của một số ký tự khá dễ dàng. Đối với những ký tự đó, bạn có thể muốn thay đổi chúng trở lại văn bản thuần túy, vì bạn có thể nhận đư ợc nhiều manh mối hơn. Tốt hơn nên sử dụng chữ in hoa cho văn bản thuần túy, vì vậy với cùng một chữ cái, chúng ta biết đâu là văn bản thuần túy và đâu là văn bản mật mã

Nếu bạn đang sử dụng Linux hoặc MacOS, bạn có thể sử dụng lệnh tr để thực hiện việc này. Ví dụ: trong phần sau, chúng ta thay thế các chữ cái a, e và t trong in.txt bằng các chữ cái X, G, E tư ơng ứng, kết quả là đư ợc lư u trong out.txt.

```
$ tr 'aet' 'XGE' < in.txt > out.txt
```

[2] https://seedsecuritylabs.org/Labs_16.04/Crypto/Crypto_Encryption/files/ciphertext.txt

Bạn cũng có thể sử dụng Cryptool hoặc dCode để giải mã thủ công bằng cách phân tích và thay thế các chữ cái. Có rất nhiều tài nguyên trực tuyến mà bạn có thể sử dụng. Hai liên kết có giá trị như sau:

Phân tích N-gram trực tuyến của Cryptool: Trang web này có thể tạo ra số liệu thống kê từ dãy số), tần số bát quái (chuỗi 3 chữ cái), v.v.

http://norvig.com/mayzner.html: Trang web này cung cấp tần số cho một văn bản gốc tiếng Anh điển hình, bao gồm tần số unigram, bigram và trigram.

Nhiệm vụ nâng cao 3.1: Giải mã văn bản mật mã từ cuốn The Gold-Bug của Edgar Allan Poe và giải thích cách thực hiện:

53‡‡†305))6*;4826)4‡.)4‡);806*;48†8¶60))85;;]8*;:‡*8†83(88)5*†; 46(;88*96

?;8)‡(;485);5*†2:*‡(;4956*2(5*—4)8¶8*;4069285);)6†8)4 ‡‡;1(‡9;48081;8:8‡

1;48†85;4)485†528806*81(‡9;48;(88;4(‡?34;48)4‡;161;: 188;‡?;

Được biết rằng:

- Nguyên văn là một bài viết bằng tiếng Anh.
- Văn bản mật mã không bao gồm dấu câu và dấu cách.
- Mỗi ký hiệu tương ứng với một chữ cái trong bảng chữ cái tiếng Anh.

4. Mật mã Playfair

Nhiệm vụ 4: Viết một ứng dụng bằng ngôn ngữ lập trình của riêng bạn để mã hóa và giải mã tin nhắn bằng mật mã Playfair, ứng dụng của bạn phải đáp ứng các yêu cầu sau

yêu cầu:

- Cho phép nhập khóa và văn bản thuần để mã hóa hoặc văn bản mật mã để giải mã bằng khóa đã cho.

- Hiển thị ma trận Playfair (5x5) tương ứng với key đã cho.

1. Kiểm tra chương trình của bạn bằng thông báo ít nhất 100 từ và so sánh kết quả với các công cụ mã hóa khác (như Cryptool 2) để xác minh.

2. Sử dụng ma trận Playfair bên dưới (Bảng 1.2) để mã hóa thông báo sau.

Thông điệp: Tôi chỉ tiếc rằng tôi chỉ có một cuộc đời để cống hiến cho đất nước mình.

Tin nhắn này là của Nathan Hale, một người lính trong Chiến tranh Cách mạng Hoa Kỳ.

Bảng 2. Lưới Playfair cho Nhiệm vụ 4 - Phần 2.

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

5. Mật mã đa bảng chữ cái – Vigenère

> Nhiệm vụ 5: Viết một ứng dụng sử dụng ngôn ngữ lập trình bạn đã chọn để mã hóa và giải mã tin nhắn bằng mật mã Vigenère.
>
> Kiểm tra ứng dụng của bạn bằng một tin nhắn có ít nhất 100 từ và khóa khoảng 10-20 chữ cái, sau đó xác minh kết quả bằng các công cụ mã hóa khác (ví dụ: Cryptool 2, dCode, v.v.)

6. Mật mã khác – Capture The Flag

> Bài 6: Giải mã đoạn văn bản sau để tìm lá cờ, được biết tin nhắn đã bị được mã hóa bằng mã ASCII:
>
> 87 101 108 99 111 109 101 32 116 111 32 67 114 121 112 116 111 32 73 115
>
> 108 97 110 100 33 33 33 32

Mẹo: Bạn có thể tham khảo bảng ASCII tại http://www.ascii-code.com

> Nhiệm vụ 7: Có thể tải xuống tệp crypto01.jpg từ đây, và nó chứa một lá cờ.
> Biết rằng hình ảnh này đã được mã hóa bằng mật mã XOR bằng khóa gồm 6 chữ cái. Hãy cùng tìm cờ và mô tả cách thực hiện

Lời khuyên: Bạn có thể tham khảo https://www.dcode.fr/xor-cipher để tìm hiểu cách hoạt động của mật mã XOR.

> Nhiệm vụ nâng cao 7.1: Mô tả một mật mã cổ điển khác chưa được đề cập trong bài thí nghiệm này Viết một ứng dụng và đưa ra một minh họa để minh họa cách thức hoạt động của nó.

C. YÊU CẦU & ĐÁNH GIÁ

Học viên được yêu cầu học tập và thực hành theo hướng dẫn, thực hiện các nhiệm vụ trong nhóm được phân công.

Gửi báo cáo kết quả bao gồm mã code, cơ sở dữ liệu đã xuất và chi tiết các nhiệm vụ (báo cáo) mà nhóm đã hoàn thành, bao gồm các quan sát và ảnh chụp màn hình kết quả nếu có, giải thích các quan sát (nếu có).

Định dạng báo cáo:

- File PDF tập trung vào nội dung, tránh mô tả mang tính lý thuyết.
- Quy ước đặt tên: [Class Code]-LabX_StudentID1.
- Ví dụ: [NT219.K11.ANTN.1]-Lab1_1852xxxx-.
- Nếu báo cáo bao gồm nhiều tệp, hãy nén tất cả các tệp vào một tệp ZIP với phần mở rộng cùng tên báo cáo.
- Gửi báo cáo trong khung thời gian đã thỏa thuận trên Courses.uit.edu.vn.

Các trường hợp đạo văn, chậm trễ, v.v. sẽ được xử lý tùy theo mức độ nghiêm trọng của lỗi.
sự vi phạm.

D. TÀI LIỆU THAM KHẢO

[1] William Stallings, Mật mã và an ninh mạng: Nguyên tắc và thực tiễn, lần thứ 7
ed, Pearson Education, 2017. Chương 3. Kỹ thuật mã hóa cổ điển

[2] Wenliang Du (Đại học Syracuse), Phòng thí nghiệm mật mã SEED https://
seedsecuritylabs.org/Labs_16.04/Crypto/

[3] Bernhard Esslinger và cộng sự, Sách CrypTool: Học và trải nghiệm mật mã với CrypTool và
SageMath, tái bản lần thứ 12, 2018. Có sẵn: https://www.cryptool.org/en/ctp-documentation

Nền tảng đào tạo và tài liệu liên quan:

- Tiền điện tử - https://cryptopals.com/

KẾT THÚC
Chúc các bạn hoàn thành xuất sắc nhiệm vụ!