CHƯƠNG 2: MÃ HÓA ĐỐI XỨNG CỔ ĐIỂN

Mục đích: *Giới thiệu chung về các hệ mã hóa đối xứng cổ điển: nguyên tắc mã hóa, giải mã. Đưa các ví dụ minh họa để sinh viên nắm được các hệ mã hóa.*

Yêu cầu: *Sinh viên nắm được nguyên tắc mã hóa, giải mã của các hệ mã hóa cổ điển. Vận dụng được vào thực hiện các bài tập mã hóa.*

Mã hoá cổ điển là phương pháp mã hoá đơn giản nhất xuất hiện đầu tiên trong lịch sử ngành mã hoá. Thuật toán đơn giản và dễ hiểu. Những phương pháp mã hoá này là cở sở cho việc nghiên cứu và phát triển thuật toán mã hoá đối xứng được sử dụng ngày nay. Trong mã hoá cổ điển có hai phương pháp nổi bật đó là:

Mã hoá thay thế

Mã hoá hoán vị

2.1 Mã đối xứng

2.1.1 Các khái niệm cơ bản

Mật mã đối xứng sử dụng cùng một khóa cho việc mã hóa và giải mã. Có thể nói mã đối xứng là mã một khoá hay mã khóa riêng hay mã khoá thỏa thuận. Ở đây người gửi và người nhận chia sẻ khoá chung K, mà họ có thể trao đổi bí mật với nhau. Ta xét hai hàm ngược nhau: E là hàm biến đổi bản rõ thành bản mã và D là hàm biến đổi bản mã trở về bản rõ. Giả sử X là văn bản cần mã hóa và Y là dạng văn bản đã được thay đổi qua việc mã hóa. Khi đó ta ký hiệu:

$$Y = E_K(X)$$

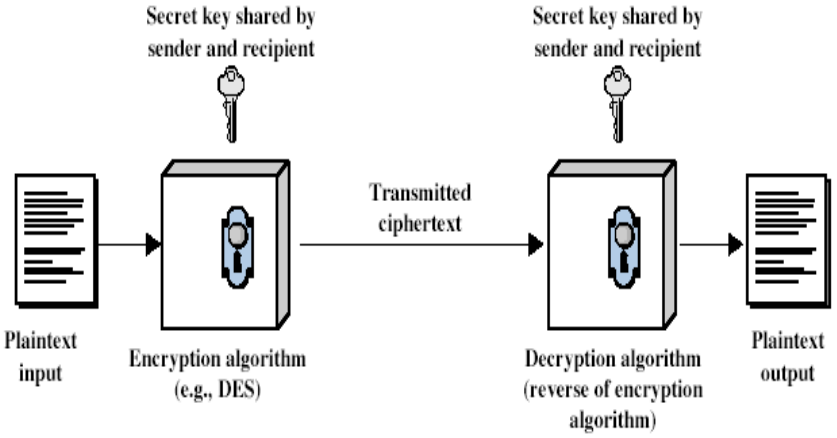
$$X = D_K(Y)$$

Mọi thuật toán mã cổ điển đều là mã khoá đối xứng, vì ở đó thông tin về khóa được chia sẻ giữa người gửi và người nhận. Mã đối xứng là kiểu duy nhất trước khi phát minh ra khoá mã công khai (còn được gọi là mã không đối xứng) vào những năm 1970. Hiện nay các mã đối xứng và công khai tiếp tục phát triển và hoàn thiện. Mã công khai ra đời hỗ trợ mã đối xứng chứ không thay thế nó, do đó mã đối xứng đến nay vẫn được sử dụng rộng rãi.

Sau đây ta đưa ra định nghĩa một số khái niệm cơ bản về mã hóa.

- **Bản rõ** X được gọi là là bản tin gốc. Bản rõ có thể được chia nhỏ có kích thước phù hợp.
- **Bản mã** Y là bản tin gốc đã được mã hoá. Ở đây ta thường xét phương pháp mã hóa mà không làm thay đổi kích thước của bản rõ, tức là chúng có cùng độ dài.
- **Mã** là thuật toán E chuyển bản rõ thành bản mã. Thông thường chúng ta cần thuật toán mã hóa mạnh, cho dù kẻ thù biết được thuật toán, nhưng không biết thông tin về khóa cũng không tìm được bản rõ.
- **Khoá K** là thông tin tham số dùng để mã hoá, chỉ có người gửi và người nhận biết. Khóa là độc lập với bản rõ và có độ dài phù hợp với yêu cầu bảo mật.
- **Mã hoá** là quá trình chuyển bản rõ thành bản mã, thông thường bao gồm việc áp dụng thuật toán mã hóa và một số quá trình xử lý thông tin kèm theo.
- **Giải mã** chuyển bản mã thành bản rõ, đây là quá trình ngược lại của mã hóa.
- **Mật mã** là chuyên ngành khoa học của Khoa học máy tính nghiên cứu về các nguyên lý và phương pháp mã hoá. Hiện nay người ta đưa ra nhiều chuẩn an toàn cho các lĩnh vực khác nhau của công nghệ thông tin.
- **Thám mã** nghiên cứu các nguyên lý và phương pháp giải mã mà không biết khoá. Thông thường khi đưa các mã mạnh ra làm chuẩn dùng chung giữa các người sử dụng, các mã đó được các kẻ thám mã cũng như những người phát triển mã tìm hiểu nghiên cứu các phương pháp giải một phần bản mã với các thông tin không đầy đủ.
- **Lý thuyết mã** bao gồm cả mật mã và thám mã. Nó là một thể thống nhất, để đánh giá một mã mạnh hay không, đều phải xét từ cả hai khía cạnh đó.

Các nhà khoa học mong muốn tìm ra các mô hình mã hóa khái quát cao đáp ứng nhiều chính sách an toàn khác nhau.



Hình 2.1: Mô hình mã đối xứng

2.1.2 Các yêu cầu

Một mã đối xứng có các đặc trưng là cách xử lý thông tin của thuật toán mã, giải mã, tác động của khóa vào bản mã, độ dài của khóa. Mối liên hệ giữa bản rõ, khóa và bản mã càng phức tạp càng tốt, nếu tốc độ tính toán là chấp nhận được. Cụ thể hai yêu cầu để sử dụng an toàn mã khoá đối xứng là:

- Thuật toán mã hoá mạnh. Có cơ sở toán học vững chắc đảm bảo rằng mặc dù công khai thuật toán, mọi người đều biết, nhưng việc thám mã là rất khó khăn và phức tạp nếu không biết khóa.
- Khoá mật chỉ có người gửi và người nhận biết. Có kênh an toàn để phân phối khoá giữa các người sử dụng chia sẻ khóa. Mối liên hệ giữa khóa và bản mã là không nhận biết được.

2.1.3 Mật mã

Hệ mật mã được đặc trưng bởi các yếu tố sau:

- Kiểu của thao tác mã hoá được sử dụng trên bản rõ:
  - Phép thế - thay thế các ký tự trên bản rõ bằng các ký tự khác
  - Hoán vị - thay đổi vị trí các ký tự trong bản rõ, tức là thực hiện hoán vị các ký tự của bản rõ.

  - Tích của chúng, tức là kết hợp cả hai kiểu thay thế và hoán vị các ký tự của bản rõ.
- Số khoá được sử dụng khi mã hóa: một khoá duy nhất - khoá riêng hoặc hai khoá - khoá công khai. Ngoài ra còn xem xét số khóa được dùng có nhiều không.
- Một đặc trưng của mã nữa là cách mà bản rõ được xử lý, theo:
  - Khối - dữ liệu được chia thành từng khối có kích thước xác định và áp dụng thuật toán mã hóa với tham số khóa cho từng khối.
  - Dòng - từng phần tử đầu vào được xử lý liên tục tạo phần tử đầu ra tương ứng.

2.1.4 Thám mã

Có hai cách tiếp cận tấn công mã đối xứng.

- Tấn công thám mã dựa trên thuật toán và một số thông tin về các đặc trưng chung về bản rõ hoặc một số mẫu bản rõ/bản mã. Kiểu tấn công này nhằm khai phá các đặc trưng của thuật toán để tìm bản rõ cụ thể hoặc tìm khóa. Nếu tìm được khóa thì là tai họa lớn.
- Tấn công duyệt toàn bộ: kẻ tấn công tìm cách thử mọi khóa có thể trên bản mã cho đến khi nhận được bản rõ. Trung bình cần phải thử một nửa số khóa mới tìm được.

Các kiểu tấn công thám mã.

- Chỉ dùng bản mã: biết thuật toán và bản mã, dùng phương pháp thống kê, xác định bản rõ.

- Biết bản rõ: biết thuật toán, biết được bản mã/bản rõ tấn công tìm khóa.

- Chọn bản rõ: chọn bản rõ và nhận được bản mã, biết thuật toán tấn công tìm khóa.

- Chọn bản mã: chọn bản mã và có được bản rõ tương ứng, biết thuật toán tấn công tìm khóa.

- Chọn bản tin: chọn được bản rõ hoặc mã và mã hoặc giải mã tuong ứng, tấn công tìm khóa.

2.1.5 Tìm duyệt tổng thể (Brute-Force)

Về mặt lý thuyết phương pháp duyệt tổng thể là luôn thực hiện được, do có thể tiến hành thử từng khoá, mà số khoá là hữu hạn. Phần lớn công sức của các tấn công đều tỷ lệ thuận với kích thước khoá. Khóa càng dài thời gian tìm kiếm càng lâu và thường tăng theo hàm mũ. Ta có thể giả thiết là kẻ thám mã có thể dựa vào bối cảnh để biết hoặc nhận biết được bản rõ.

Sau đây là một số thống kê về mối liên hệ giữa độ dài khóa, kích thước không gian khóa, tốc độ xử lý và thời gian tìm duyệt tổng thể. Chúng ta nhận thấy với độ dài khóa từ 128 bit trở lên, thời gian yêu cầu là rất lớn, lên đến hàng tỷ năm, như vậy có thể coi phương pháp duyệt tổng thể là không hiện thực.

2.1.6 Độ an toàn

Có thể phân loại an toàn thành hai kiểu như sau:

- An toàn không điều kiện: ở đây không quan trọng máy tính mạnh như thế nào, có thể thực hiện được bao nhiêu phép toán trong một giây, mã hoá không thể bị bẻ, vì bản mã không cung cấp đủ thông tin để xác định duy nhất bản rõ. Việc dùng bộ đệm ngẫu nhiên một lần để mã dòng cho dữ liệu mà ta sẽ xét cuối bài này được coi là an toàn không điều kiện. Ngoài ra chưa có thuật toán mã hóa nào được coi là an toàn không điều kiện.
- An toàn tính toán: với nguồn lực máy tính giới hạn và thời gian có hạn (chẳng hạn thời gian tính toán không quá tuổi của vũ trụ) mã hoá coi như không thể bị bẻ. Trong trường hợp này coi như mã hóa an toàn về mặt tính toán. Nói chung từ nay về sau, một thuật toán mã hóa an toàn tính toán được coi là an toàn.

2.2 Các mã thế cổ điển thay thế

Có hai loại mã cổ điển là mã thay thế và mã hoán vị (hay còn gọi là dịch chuyển).

- **Mã thay thế:** từng kí tự (nhóm kí tự) trong bản rõ được thay thế bằng một kí tự (một nhóm kí tự) khác để tạo ra bản mã. Bên nhận chỉ cần thay thế ngược lại trên bản mã để có được bản rõ ban đầu.

- **Mã hoán vị:** các kí tự trong bản rõ vẫn được giữ nguyên, chúng chỉ được sắp xếp lại vị trí để tạo ra bản mã. Tức là các kí tự trong bản rõ hoàn toàn không bị thay đổi bằng kí tự khác mà chỉ đảo chỗ của chúng để tạo thành bản mã.

Trước hết ta xét các mã cổ điển sử dụng phép thay thế các chữ của bản rõ bằng các chữ khác của bảng chữ để tạo thành bản mã.

- Ở đây các chữ của bản rõ được thay bằng các chữ hoặc các số hoặc các ký tự khác.
- Hoặc nếu xem bản rõ như môt dãy bít, thì phép thế thay các mẫu bít bản rõ bằng các mẫu bít bản mã.

2.2.1 Mã Ceasar

Đây là mã thế được biết sớm nhất, được sáng tạo bởi Julius Ceasar. Lần đầu tiên được sử dụng trong quân sự. Việc mã hoá được thực hiện đơn giản là thay mỗi chữ trong bản rõ bằng chữ thứ ba tiếp theo trong bảng chữ cái.

Ví dụ:

Meet me after the toga party

PHHW PH DIWHU WKH WRJD SDUWB

Ở đây thay chữ m bằng chữ đứng thứ 3 sau m là p (m, n, o, p); thay chữ e bằng chữ đứng thứ 3 sau e là h (e, f, g, h).

Có thể định nghĩa việc mã hoá trên qua ánh xạ trên bảng chữ cái sau: các chữ ở dòng dưới là mã của các chữ tương ứng ở dòng trên:

a b c d e f g h i j k l m n o p q r s t u v w x y z

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

Về toán học, nếu ta gán số thứ tự cho mỗi chữ trong bảng chữ cái. Các chữ ở dòng trên có số thứ tự tương ứng là số ở dòng dưới:

a b c d e f g h i j k l m

0 1 2 3 4 5 6 7 8 9 10 11 12

n o p q r s t u v w x y z

13 14 15 16 17 18 19 20 21 22 23 24 25

thì mã Ceasar được định nghĩa qua phép tịnh tiến các chữ như sau:

$$c = E(p) = (p + k) \bmod (26)$$

$$p = D(c) = (c – k) \bmod (26)$$

Ở đây, p là số thứ tự của chữ trong bản rõ và c là số thứ tự của chữ tương ứng của bản mã; k là khoá của mã Ceasar. Có 26 giá trị khác nhau của k, nên có 26 khoá khác nhau. Thực tế độ dài khoá ở đây chỉ là 1, vì mọi chữ đều tịnh tiến đi một khoảng như nhau.

Thám mã Ceasar là việc làm đơn giản, do số khoá có thể có là rất ít.

Chỉ có 26 khoá có thể, vì A chỉ có thể ánh xạ vào một trong số 26 chữ cái của bảng chữ cái tiếng Anh: A, B, C, …Các chữ khác sẽ được xác định bằng số bước tịnh tiến tương ứng của A. Kẻ thám mã có thể thử lần lượt từng khoá một, tức là sử dụng phương pháp tìm duyệt tổng thể. Vì số khoá ít nên việc tìm duyệt là khả thi. Cho trước bản mã, thử 26 cách dịch chuyển khác nhau, ta sẽ đoán nhận thông qua nội dung các bản rõ nhận được.

Ví dụ. Bẻ bản mã "GCUA VQ DTGCM" bằng cách thử các phép tịnh tiến khác nhau của bảng chữ, ta chọn được bước tịnh tiến thích hợp là 24 và cho bản rõ là "easy to break".

2.2.2 Các mã bảng chữ đơn

Bây giờ ta khắc phục nhược điểm của mã Ceasar bằng cách mã hoá các chữ không chỉ là dịch chuyển bảng chữ, mà có thể tạo ra các bước nhảy khác nhau cho các chữ. Trong một mã mỗi chữ của bản rõ được ánh xạ đến một chữ khác nhau của bản mã. Do đó mỗi cách mã như vậy sẽ tương ứng với một hoán vị của bảng chữ và hoán vị đó chính là khoá của mã đã cho. Như vậy độ dài khoá ở đây là 26 và số khoá có thể có là 26!. Số khoá như vậy là rất lớn.

Ví dụ. Ta có bản mã tương ứng với bản rõ trong mã bảng chữ đơn như sau:

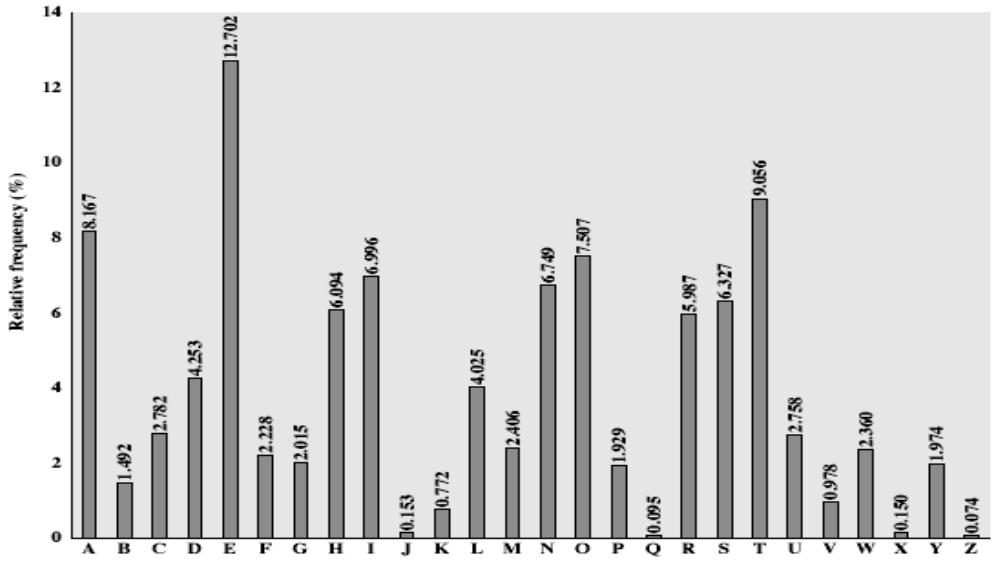
Plain: abcdefghijklmnopqrstuvwxyz

Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext: ifwewishtoreplaceletters

Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA

- Tính an toàn của mã trên bảng chữ đơn. Tổng cộng có 26! xấp xỉ khoảng 4 x 1026 khoá. Với khá nhiều khoá như vậy nhiều người nghĩ là mã trên bảng chữ đơn sẽ an toàn. Nhưng không phải như vậy. Vấn đề ở đây là do các đặc trưng về ngôn ngữ. Tuy có số lượng khoá lớn, nhưng do các đặc trưng về tần suất xuất hiện của các chữ trong bản rõ và các chữ tương ứng trong bản mã là như nhau, nên kẻ thám mã có thể đoán được ánh xạ của một số chữ và từ đó mò tìm ra chữ mã cho các chữ khác. Ta sẽ xét khía cạnh này cụ thể trong mục sau.
- Tính dư thừa của ngôn ngữ và thám mã. Ngôn ngữ của loài người là dư thừa. Có một số chữ hoặc các cặp chữ hoặc bộ ba chữ được dùng thường xuyên hơn các bộ chữ cùng độ dài khác. Chẳng hạn như các bộ chữ sau đây trong tiếng Anh "th lrd s m shphrd shll nt wnt". Tóm lại trong nhiều ngôn ngữ các chữ không được sử dụng thường xuyên như nhau. Trong tiếng Anh chữ E được sử dụng nhiều nhất; sau đó đến các chữ T, R, N, I, O, A, S. Một số chữ rất ít dùng như: Z, J, K, Q, X. Bằng phương pháp thống kê, ta có thể xây dựng các bảng các tần suất các chữ đơn, cặp chữ, bộ ba chữ.



Hình 2.2: Bảng tần suất chữ cái tiếng Anh

Sử dụng bảng tần suất vào việc thám mã

Điều quan trọng là mã thế trên bảng chữ đơn không làm thay đổi tần suất tương đối của các chữ, có nghĩa là ta vẫn có bảng tần suất trên nhưng đối với bảng chữ mã tương ứng. Điều đó được phát hiện bởi các nhà khoa học Ai cập từ thế kỷ thứ 9. Do đó có cách thám mã trên bảng chữ đơn như sau:

- Tính toán tần suất của các chữ trong bản mã
- So sánh với các giá trị đã biết
- Tìm kiếm các chữ đơn hay dùng A-I-E, bộ đôi NO và bộ ba RST; và các bộ ít dùng JK, X-Z..
- Trên bảng chữ đơn cần xác định các chữ dùng các bảng bộ đôi và bộ ba trợ giúp.

Ví dụ. Thám mã bản mã trên bảng chữ đơn, cho bản mã:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSXEPYEP
OPDZSZUFPOUDTMOHMQ

- Tính tần suất các chữ
- Đoán P và Z là e và t.
- Khi đó ZW là th và ZWP là the.
- Suy luận tiếp tục ta có bản rõ:

it was disclosed yesterday that several informal but
direct contacts have been made with political
representatives in moscow

2.2.3 Mã Playfair

Như chúng ta đã thấy không phải số khoá lớn trong mã bảng chữ đơn đảm bảo an toàn mã. Một trong các hướng khắc phục là mã bộ các chữ, tức là mỗi chữ sẽ được mã bằng một số chữ khác nhau tùy thuộc vào các chữ mà nó đứng cạnh. Playfair là một trong các mã như vậy, được sáng tạo bởi Charles Wheastone vào năm 1854 và mang tên người bạn là Baron Playfair. Ở đây mỗi chữ có thể được mã bằng một trong 7 chữ khác nhau tùy vào chữ cặp đôi cùng nó trong bản rõ.

Ma trận khoá Playfair. Cho trước một từ làm khoá, với điều kiện trong từ khoá đó không có chữ cái nào bị lặp. Ta lập ma trận Playfair là ma trận cỡ 5 x 5 dựa trên từ khoá đã cho và gồm các chữ trên bảng chữ cái, được sắp xếp theo thứ tự như sau:

- Trước hết viết các chữ của từ khoá vào các hàng của ma trận bắt từ hàng thứ nhất.
- Nếu ma trận còn trống, viết các chữ khác trên bảng chữ cái chưa được sử dụng vào các ô còn lại. Có thể viết theo một trình tự qui ước trước, chẳng hạn từ đầu bảng chữ cái cho đến cuối.
- Vì có 26 chữ cái tiếng Anh, nên thiếu một ô. Thông thường ta dồn hai chữ nào đó vào một ô chung, chẳng hạn I và J.
- Giả sử sử dụng từ khoá MORNACHY. Lập ma trận khoá Playfair tương ứng như sau:

MONAR
CHYBD
EFGIK
LPQST
UVWXZ

Mã hoá và giải mã: bản rõ được mã hoá 2 chữ cùng một lúc theo qui tắc như sau:

- Chia bản rõ thành từng cặp chữ. Nếu một cặp nào đó có hai chữ như nhau, thì ta chèn thêm một chữ lọc chẳng hạn X. Ví dụ, trước khi mã "balloon" biến đổi thành "ba lx lo on".
- Nếu cả hai chữ trong cặp đều rơi vào cùng một hàng, thì mã mỗi chữ bằng chữ ở phía bên phải nó trong cùng hàng của ma trận khóa (cuộn vòng quanh từ cuối về đầu), chẳng hạn "ar" biến đổi thành "RM"
- Nếu cả hai chữ trong cặp đều rơi vào cùng một cột, thì mã mỗi chữ bằng chữ ở phía bên dưới nó trong cùng cột của ma trận khóa (cuộn vòng quanh từ cuối về đầu), chẳng hạn "mu" biến đổi thành "CM"

- Trong các trường hợp khác, mỗi chữ trong cặp được mã bởi chữ cùng hàng với nó và cùng cột với chữ cùng cặp với nó trong ma trận khóa. Chẳng hạn, "hs" mã thành "BP", và "ea" mã thành "IM" hoặc "JM" (tuỳ theo sở thích)

An toàn của mã Playfair:

- An toàn được nâng cao so hơn với bảng đơn, vì ta có tổng cộng 26 x 26 = 676 cặp. Mỗi chữ có thể được mã bằng 7 chữ khác nhau, nên tần suất các chữ trên bản mã khác tần suất của các chữ cái trên văn bản tiếng Anh nói chung.
- Muốn sử dụng thống kê tần suất, cần phải có bảng tần suất của 676 cặp để thám mã (so với 26 của mã bảng đơn). Như vậy phải xem xét nhiều trường hợp hơn và tương ứng sẽ có thể có nhiều bản mã hơn cần lựa chọn. Do đó khó thám mã hơn mã trên bảng chữ đơn.
- Mã Playfair được sử dụng rộng rãi nhiều năm trong giới quân sự Mỹ và Anh trong chiến tranh thế giới thứ 1. Nó có thể bị bẻ khoá nếu cho trước vài trăm chữ, vì bản mã vẫn còn chứa nhiều cấu trúc của bản rõ.

2.2.4 Các mã đa bảng

Một hướng khác làm tăng độ an toàn cho mã trên bảng chữ là sử dụng nhiều bảng chữ để mã. Ta sẽ gọi chúng là các mã thế đa bảng. Ở đây mỗi chữ có thể được mã bằng bất kỳ chữ nào trong bản mã tùy thuộc vào ngữ cảnh khi mã hoá. Làm như vậy để trải bằng tần suất các chữ xuất hiện trong bản mã. Do đó làm mất bớt cấu trúc của bản rõ được thể hiện trên bản mã và làm cho thám mã đa bảng khó hơn. Ta sử dụng từ khoá để chỉ rõ chọn bảng nào được dùng cho từng chữ trong bản tin. Sử dụng lần lượt các bảng theo từ khóa đó và lặp lại từ đầu sau khi kết thúc từ khoá. Độ dài khoá là chu kỳ lặp của các bảng chữ. Độ dài càng lớn và nhiều chữ khác nhau được sử dụng trong từ khoá thì càng khó thám mã.

2.2.5 Mã Vigenere

Mã thế đa bảng đơn giản nhất là mã Vigenere. Thực chất quá trình mã hoá Vigenere là việc tiến hành đồng thời dùng nhiều mã Ceasar cùng một lúc trên bản rõ với nhiều khoá khác nhau. Khoá cho mỗi chữ dùng để mã phụ thuộc vào vị trí của chữ đó trong bản rõ và được lấy trong từ khoá theo thứ tự tương ứng.

Giả sử khoá là một chữ có độ dài d được viết dạng K = K1K2…Kd, trong đó Ki nhận giá trị nguyên từ 0 đến 25. Khi đó ta chia bản rõ thành các khối gồm d chữ. Mỗi chữ thứ i trong khối chỉ định dùng bảng chữ thứ i với tịnh tiến là Ki giống như trong mã Ceasar. Trên thực tế khi mã ta có thể sử dụng lần lượt các bảng chữ và lặp lại từ đầu sau d chữ của bản rõ. Vì có nhiều bảng chữ khac nhau, nên cùng một chữ ở các vị trí khác nhau sẽ có các bước nhảy khác nhau, làm cho tần suất các chữ trong bản mã dãn tương đối đều.

Giải mã đơn giản là quá trình làm ngược lại. Nghĩa là dùng bản mã và từ khoá với các bảng chữ tương ứng, nhưng với mỗi chữ sử dụng bước nhảy lui lại về đầu.

Ví dụ: Để sử dụng mã Vigenere với từ khóa và bản rõ cho trước ta có thể làm như sau:

- Viết bản rõ ra
- Viết từ khoá lặp nhiều lần phía trên tương ứng của nó
- Sử dụng mỗi chữ của từ khoá như khoá của mã Ceasar
- Mã chữ tương ứng của bản rõ với bước nhảy tương ứng.
- Chẳng hạn sử dụng từ khoá deceptive

key: deceptivedeceptivedeceptive

plaintext: wearediscoveredsaveyourself

ciphertext:ZICVTWQNGRZGVTWAVZHCQYGL

Để mã chữ w đầu tiên ta tìm chữ đầu của khóa là d, như vậy w sẽ được mã trên bảng chữ tịnh tiến 3 (tức là a tịnh tiến vào d). Do đó chữ đầu w được mã bởi chữ Z. Chữ thứ hai trong từ khóa là e, có nghĩa là chữ thứ hai trong bản rõ sẽ được tịnh tiến 4 (từ a tịnh tiến đến e). Như vậy thứ hai trong bản rõ e sẽ được mã bởi chữ I. Tương tự như vậy cho đến hết bản rõ.

Trên thực tế để hỗ trợ mã Vigenere, người ta đã tạo ra trang Saint – Cyr để trợ giúp cho việc mã và giải mã thủ công. Đó là một bảng cỡ 26 x 26 có tên tương ứng là các chữ cái trong bảng chữ tiếng Anh. Hàng thứ i là tịnh tiến i chữ của bảng chứ cái. Khi đó chữ ở cột đầu tiên chính là khoá của bảng chữ ở cùng hàng. Do đó

chữ mã của một chữ trong bản rõ nằm trên cùng cột với chữ đó và nằm trên hàng tương ứng với chữ khoá.

ABCDEFGHIJKLMNOPQRSTUVWXYZ

A ABCDEFGHIJKLMNOPQRSTUVWXYZ
B BCDEFGHIJKLMNOPQRSTUVWXYZA
C CDEFGHIJKLMNOPQRSTUVWXYZAB
D DEFGHIJKLMNOPQRSTUVWXYZABC
E EFGHIJKLMNOPQRSTUVWXYZABCD
F FGHIJKLMNOPQRSTUVWXYZABCDE
G GHIJKLMNOPQRSTUVWXYZABCDEF
H HIJKLMNOPQRSTUVWXYZABCDEFG
I IJKLMNOPQRSTUVWXYZABCDEFGH
J JKLMNOPQRSTUVWXYZABCDEFGHI
K KLMNOPQRSTUVWXYZABCDEFGHIJ
L LMNOPQRSTUVWXYZABCDEFGHIJK
M MNOPQRSTUVWXYZABCDEFGHIJKL
N NOPQRSTUVWXYZABCDEFGHIJKLM
O OPQRSTUVWXYZABCDEFGHIJKLMN
P PQRSTUVWXYZABCDEFGHIJKLMNO
Q QRSTUVWXYZABCDEFGHIJKLMNOP
R RSTUVWXYZABCDEFGHIJKLMNOPQ
S STUVWXYZABCDEFGHIJKLMNOPQR
T TUVWXYZABCDEFGHIJKLMNOPQRS
U UVWXYZABCDEFGHIJKLMNOPQRST
V VWXYZABCDEFGHIJKLMNOPQRSTU
W WXYZABCDEFGHIJKLMNOPQRSTUV
X XYZABCDEFGHIJKLMNOPQRSTUVW
Y YZABCDEFGHIJKLMNOPQRSTUVWX

Z ZABCDEFGHIJKLMNOPQRSTUVWXY

Bảng Saint Cyr

An toàn của mã Vigenere. Như vậy có chữ mã khác nhau cho cùng một chữ của bản rõ. Suy ra tần suất của các chữ bị là phẳng, nghĩa là tần suất xuất hiện các chữ trên bản mã tương đối đều nhau. Tuy nhiên chưa mất hoàn toàn, do độ dài của khoá có hạn, nên có thể tạo nên chu kỳ vòng lặp. Kẻ thám mã bắt đầu từ tần suất của chữ để xem có phải đây là mã đơn bảng chữ hay không. Giả sử đây là mã đa bảng chữ, sau đó xác định số bảng chữ trong từ khoá và lần tìm từng chữ. Như vậy cần tăng độ dài từ khoá để tăng số bảng chữ dùng khi mã để "là" tần suất của các chữ.

2.2.6 Phương pháp thám mã Kasiski

Phương pháp phát triển bởi Babbage và Kasiski. Ta thấy các chữ như nhau trên bản rõ và cách nhau một khoảng đúng bằng độ dài từ khoá (chu kỳ), thì sẽ được mã bằng cùng một chữ. Như vậy từ độ lặp của các chữ trong bản mã có thể cho phép xác định chu kỳ. Tất nhiên không phải khi nào cũng tìm được độ dài từ khoá. Sau đó tìm các chữ trong từ khoá bằng cách tấn công từng bảng chữ đơn với cùng kỹ thuật dựa trên các bảng tần suất của các bộ chữ như trước.

2.2.7 Mã khoá tự động

Lý tưởng nhất là ta có khoá dài như bản tin. Do đó Vigenere đề xuất khoá tự động sinh cho bằng độ dài bản tin như sau: từ khoá được nối tiếp bằng chính bản rõ để tạo thành khoá. Sau đó dùng mã Vigenere để mã bản rõ đã cho. Khi đó biết từ khoá có thể khôi phục được một số chữ ban đầu của bản rõ. Sau đó tiếp tục sử dụng chúng để giải mã cho văn bản còn lại. Sự cải tiến này làm mất khái niệm chu kỳ, gây khó khăn cho việc thám mã, nhưng vẫn còn đặc trưng tần suất để tấn công.

Ví dụ. Cho từ khoá deceptive. Ta viết bản rõ nối tiếp vào từ khoá tạo thành từ khoá mới có độ dài bằng độ dài bản rõ.

key: deceptivewearediscoveredsav

plaintext: wearediscoveredsaveyourself

ciphertext:ZICVTWQNGKZEIIGASXSTSLVVWLA

2.2.8 Bộ đệm một lần

Nếu khoá thực sự ngẫu nhiên được dùng và có độ dài bằng bản rõ thì ta nói đó là bộ đệm một lần. Vì nó chỉ được dùng một lần và ngẫu nhiên, nên mã hoá sẽ an toàn. Mã sẽ không bẻ được vì bản mã không có liên quan thống kê gì với bản rõ, do bộ đệm được sinh ngẫu nhiên. Có thể nói mã bộ đệm một lần là an toàn tuyệt đối, vì với bản rõ bất kỳ và bản mã bất kỳ, luôn tồn tại một khoá để ánh xạ bản rõ đó sang bản mã đã cho. Về mặt lý thuyết, xác suất để mọi mẩu tin (có cùng độ dài với bản rõ) trên bảng chữ mã là mã của một bản rõ cho trước là như nhau. Khoá chỉ sử dụng một lần, nên các lần mã là độc lập với nhau.

Vấn đề khó khăn của mã bộ đệm một lần là việc sinh ngẫu nhiên khóa và phân phối khoá an toàn. Do đó bộ đệm một lần ít được sử dụng và chỉ dùng trong trường hợp đòi hỏi bảo mật rất cao.