[Bài đọc] Mã xác thực thông báo sử dụng hàm băm (HMAC)

Giới thiệu chung về mã xác thực thông báo MAC

Mã xác thực thông báo (Message Authentication Code) là một đoạn mã được tạo ra nhằm mục đích xác thực tính toàn vẹn và nguồn gốc của thông báo. Đây là một cơ chế quan trọng để duy trì tính toàn vẹn và không thể chối bỏ dữ liệu. Với các giao thức trực tuyến, mã xác thực thông báo được sử dụng phổ biến và với việc kết hợp với các cơ chế khác (như chữ ký số) để đảm bảo tính xác thực giữa các bên tham gia giao dịch.

Mục đích của hàm MAC là đảm bảo để hai (hay nhiều) bên tham gia giao dịch khi có chung khóa bí mật có thể giao dịch với nhau, kèm theo khả năng có thể phát hiện được thay đổi của thông báo trong quá trình vận chuyển, nhằm tránh các tấn công làm thay đổi thông báo.

Thuật toán MAC dựa trên thông báo đầu vào và khóa mật để tạo ra thẻ MAC (MAC tag). Thông báo và thẻ MAC được gửi tới người nhận, người nhận tính lại giá trị MAC và so sánh nó với giá trị thẻ MAC nhận được. Nếu hai giá trị thẻ MAC trùng nhau thì coi như thông báo chính xác, ngược lại thông báo coi là đã bị thay đổi.

Đối với kẻ tấn công giả mạo thông báo, mục tiêu đặt ra là phải phá được hàm MAC, việc này khó tương đương với việc phá khóa bảo vệ thông báo. Trên thực tế, các giao thức thường chia thông báo dài làm nhiều đoạn nhỏ và chúng được xác thực độc lập với nhau, dẫn đến phát sinh tấn công lặp lại (Replay Attack). Do vậy khi thiết kế giao thức có sử dụng MAC cần rất thận trọng.

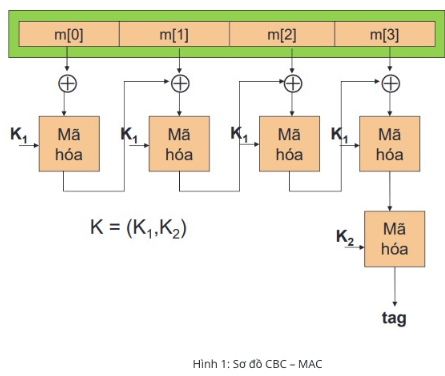
Để giúp các nhà phát triển ứng dụng tích hợp MAC vào sản phẩm khác nhau, viện tiêu chuẩn và công nghệ quốc gia của Mỹ (NIST) đưa ra hai chuẩn về hàm MAC. Tiêu chuẩn thứ nhất là mã xác thực thông báo sử dụng hàm một chiều có khóa HMAC (Keyd-Hash Messasge Authentication Code). Chuẩn này mô tả phương pháp an toàn chuyển hàm một chiều kháng va chạm Hash thành hàm MAC. Chuẩn thứ hai NIST đưa ra là mã xác thực thông báo mã hóa (Cipher Message Authentication Code - CMAC).

Không giống HMAC, CMAC sử dụng mã khối để thực hiện chức năng MAC, nó rất phù hợp với các ứng dụng bộ nhớ hạn chế chỉ đủ để dùng cho mã hóa dữ liệu.

Hàm MAC khác với hàm một chiều hash, nó bảo đảm xác thực thông báo chứ không phải toàn vẹn thông báo, tuy nhiên cả hai dựa trên nguyên lý rất chung. Trong cả hai trường hợp chúng ta đều tìm cách xác định tính chính xác, cụ thể hơn là sự toàn vẹn của thông báo. Tất nhiên mục tiêu của xác thực là tìm cách trả lại tính ban đầu của thông báo.

Mã xác thực thông báo mã hóa

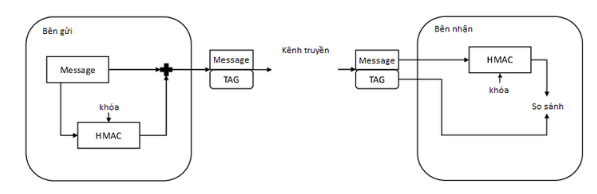
CMAC được lấy từ dự thảo mã xác thực thông báo một khóa (One - Key Message Authentication Code - OMAC). Nó dựa trên mã xác thực thông báo chế độ mã CBC ba khóa (Three-key Cipher Block Chaining MAC). MAC mã hóa nguyên thủy của NIST là CBC - MAC. Trong đó người gửi chỉ cần chọn khóa độc lập với khóa mã và sử dụng mã theo chế độ CBC. Người gửi bỏ qua mọi bản mã giữa, chỉ quan tâm đến bản mã cuối cùng là thẻ MAC. Khóa dùng cho CBC-MAC không trùng với khóa mã của bản rõ, khi đó MAC là an toàn.



CMAC nhận khóa mật K là đầu vào của quá trình khởi tạo. Nó sử dụng khóa để sinh hai khóa bổ sung. Để thuận tiện, các khóa được tạo ra phụ thuộc vào nhau, Điều này phải trả giá là nếu kẻ tấn công biết được một khóa, nó có thể tìm ra khóa kia. Tấn công đối với hàm MAC là các tấn công trực tuyến. Vì tất cả dữ liệu cần xác thực, kẻ tấn công không dễ gì truy vấn tới thiết bị, nhưng nó có thể biết được dữ liệu nào đưa vào MAC. Trên thực tế kẻ tấn công có rất ít thời gian để thực hiện giả mạo mà không bị phát hiện, xác xuất thực hiện thành công sự giả mạo rất nhỏ.

Nhìn chung tốc độ của CMAC phụ thuộc vào tốc độ mã hóa, với sự tối ưu hóa hàm tính toán.

## Mã xác thực thông báo sử dụng hàm băm



Khi đã có các hàm hash tốt, chúng ta muốn có các mã xác thực thông báo MAC dựa trên các hàm hash đó. Vì hàm hash thông thường nhanh hơn và mã nguồn của hàm hash được phổ biến rộng rãi hơn, nên việc sử dụng chúng tạo nên MAC sẽ hiệu quả hơn.

HMAC được thiết kế theo chuẩn Internet RFC2104, sử dụng hàm hash trên mẩu tin:

*HMACK = Hash [(K+ XOR opad) || Hash [(K+ XOR ipad) || M)]]*

Trong đó K+ là khóa đệm mở rộng của K và opad, ipad là các hằng bộ đệm đặc biệt, M là thông báo.

Đặc điểm:

* Dùng hàm băm nguyên mẫu.
* Cho phép thay thế dễ dàng hàm băm được nhúng vào trong trường hợp các hàm băm nhanh hơn hoặc nhiều bảo mật được tìm ra hoặc yêu cầu.
* Dùng và quản lý các khóa một cách dễ dàng

MAC được thiết kế để đảm bảo xác thực giữa các bên trong kênh liên lạc, giúp việc gửi và nhận thông báo được xác thực với nhau và khả năng bị kẻ tấn công giả mạo là rất thấp. Tuy vậy để sử dụng chúng an toàn, hiệu quả một số vấn đề sau đây cần đặc biệt lưu ý. Trên thực tế, MAC được triển khai sử dụng trong các giao thức SSL và TLS trên internet.

Giống như đối với mã khối, hàm hash cũng có tấn công vét cạn, cụ thể: Hash chống va chạm mạnh có 2m/2 giá trị cần phải duyệt để tìm ra va chạm, có nghĩa là với m là độ dài mã hash, thì 2m/2 xác định sức mạnh của nó chống đối lại tấn công vét cạn. Ta cần lựa chọn m đủ lớn để việc duyệt tìm 2m/2 phương án là không khả thi. Theo các tài liệu công bố hiện nay, hàm MD5 128 bit là không còn an toàn cho nên khuyến nghị sử dụng các hàm hash an toàn hơn như SHA-256,… Tấn công vét cạn trên MAC khó hơn, do nó phụ thuộc thêm vào khóa. Có thể tấn công vào không gian khóa (như là tìm khóa) hoặc MAC. Độ dài ít nhất 128 bit MAC là cần thiết để đảm bảo an toàn.

Việc sử dụng hàm MAC không nên quá lạm dụng trong các ứng dụng yêu cầu bảo mật, ít nhất không sử dụng nó như hàm hash hay hàm mã hóa. Một câu hỏi đặt ra là nên sử dụng CMAC hay HMAC? Dùng CMAC hay HMAC phải dựa trên vấn đề mà ta cần giải quyết. CMAC là hợp lý nếu như thuật toán mật mã, các giao thức phân phối, thỏa thuận khóa, hạ tầng đảm bảo an toàn. CMAC dựa trên mã khối nhưng với đầu vào nhỏ (so với hash) và đầu ra ngắn gọn, thời gian trễ cho tính toán nhỏ. HMAC thắng thế khi áp dụng cho thông báo kích thước lớn. Hash xử lý dữ liệu với số vòng/byte ít hơn mã hóa. Hash cũng thường tạo ra dữ liệu thẻ MAC dài theo yêu cầu cụ thể.

Ứng dụng mã xác thực thông báo

Các hàm mã xác thực thông báo được sử dụng rông rãi trong các sản phẩm bảo mật và an toàn thông tin để đảm bảo tính xác thực và toàn vẹn của thông tin. Các hàm mã xác thực thông báo thông dụng hay được sử dụng là: HMAC-SHA-1, HMAC-SHA-224, HMAC-SHA-256, HMAC-SHA-384 và HMAC-SHA-512.