NGHIÊN CỨU MÔ HÌNH MÃ HÓA THÔNG TIN ỨNG DỤNG CHỮ KÝ ĐIỆN TỬ TRONG QUÁ TRÌNH GỬI VÀ NHẬN VĂN BẢN

RESEARCH INFORMATION ENCRYPTION APPLICATION OF ELECTRONIC SIGNATURES TO THE PROCESS OF SENDING AND RECEIVING TEXTS

Ninh Văn Thọ[1,*]

DOI: https://doi.org/10.57001/huih5804.39

TÓM TẮT

Hiện nay, việc bảo đảm sự an toàn và bí mật của các thông tin tránh mọi nguy cơ bị thay đổi, sao chép hoặc mất mát dữ liệu trong các ứng dụng trên mạng luôn là vấn đề bức xúc, được nhiều người quan tâm. Mạng Internet toàn cầu đã tạo ra những cơ cấu ảo - nơi diễn ra các quá trình trao đổi thông tin trong mọi lĩnh vực chính trị, quân sự, quốc phòng, kinh tế, thương mại… Và chính trong môi trường mở và tiện nghi như thế xuất hiện những vấn nạn, tiêu cực đang rất cần các giải pháp hữu hiệu nhằm đảm bảo an toàn thông tin, chống lại các nạn ăn cắp bản quyền, xuyên tạc thông tin, truy nhập thông tin trái phép... Để tìm giải pháp cho những vấn đề này, bài báo sẽ xây dựng quy trình ứng dụng chữ ký điện tử trong quá trình gửi và nhận văn bản nhằm kiểm tra tính toàn vẹn của dữ liệu trong giao dịch điện tử là một trong những biện pháp bảo mật thông tin.

Từ khoá: *Mã hóa thông tin; chữ ký điện tử; thuật toán RSA; thuật toán MD5.*

ABSTRACT

Nowadays, it is a burning issue to ensure the safety and confidentiality of information to avoid any risk of change, copy or data loss in network applications, which is concerned by many people. Global Internet has created virtual structures where the process of information exchange in every area such as policy, military, defense, economy, trade and so on takes place. Additionally, in such an open and convenient condition, problems and negativeness which arise need effective solutions to ensure information security and fight against copyright theft, information distortion, unauthorized access to information and so on. In order to work out solutions to these problems, the article will build application procedure of electronic signatures in the process of sending and receiving texts to check the data integrity in electronic exchange, which is one of the measures of information security.

Keywords: *Decryption; electronic signatures; RSA algorithm; MD5 algorithm.*

[1]Khoa Điện tử, Trường Đại học Kinh tế - Kỹ thuật Công nghiệp

[*]Email: nvtho@uneti.edu.vn

Ngày nhận bài: 20/5/2022

Ngày nhận bài sửa sau phản biện: 20/8/2022

Ngày chấp nhận đăng: 27/10/2022

1. GIỚI THIỆU

Con người đã sử dụng các hợp đồng dưới dạng điện tử từ hơn 100 năm nay với việc sử dụng mã Morse và điện tín. Tuy nhiên, chỉ với những phát triển của khoa học kỹ thuật gần đây thì chữ ký điện tử mới đi vào cuộc sống một cách rộng rãi. Vào thập niên 1980, các công ty và một số cá nhân bắt đầu sử dụng máy fax để truyền đi các tài liệu quan trọng. Mặc dù chữ ký trên các tài liệu này vẫn thể hiện trên giấy nhưng quá trình truyền và nhận chúng hoàn toàn dựa trên tín hiệu điện tử.

Chữ ký điện tử (electronic signature) là một dạng thông tin được đi kèm theo dữ liệu (bao gồm văn bản, hình ảnh, video,…) nhằm mục đích xác định người chủ của dữ liệu đó.

Trên thực tế, chữ ký điện tử đã được ứng dụng rộng rãi trong các ứng dụng trên mạng. Một trong những ứng dụng quan trọng của chữ ký điện tử là đảm bảo an toàn dữ liệu khi truyền trên mạng. Cụ thể khi ta có một đoạn văn bản muốn gửi cho người nhận, làm thế nào để biết được khi người nhận nhận được đoạn văn bản mà nội dung không bị thay đổi. Để giải quyết vấn đề bảo mật văn bản khi giao dịch trao đổi trên mạng, đến nay đã có nhiều giải pháp liên quan đến vấn đề mã hóa văn bản, bài báo này chọn giải pháp ứng dụng chữ ký điện tử trên cơ sở kết hợp giữa thuật toán băm MD5 và thuật toán mã hóa RSA trong quá trình gửi và nhận văn bản.

Trong bài báo này, tác giả trình bày những nội dung chính như sau: (1) trình bày về mã hóa thông tin, thuật toán băm MD5, thuật toán RSA, (2) xây dựng quy trình ứng dụng chữ ký điện tử trong quá trình gửi và nhận văn bản nhằm kiểm tra tính toàn vẹn của dữ liệu trong giao dịch điện tử là một trong những biện pháp bảo mật thông tin, (3) xây dựng chương trình thực nghiệm minh họa: hoạt động của thuật toán MD5 và hàm băm, minh họa về mô hình chữ ký điện tử trong quá trình gửi nhận văn bản.

Bài báo chọn thuật toán MD5 và RSA vì các đặc điểm sau:

Mã hóa: giữ bí mật thông tin và chỉ có người có khóa bí mật mới giải mã được.

Tạo chữ ký số: cho phép kiểm tra một văn bản có phải đã được tạo với một khóa bí mật nào đó hay không.

Thỏa thuận khóa: cho phép thiết lập khóa dùng để trao đổi thông tin mật giữa 2 bên.

Thông thường, các kỹ thuật mật mã hóa khóa công khai đòi hỏi khối lượng tính toán nhiều hơn các kỹ thuật mã hóa khóa đối xứng nhưng những lợi điểm mà chúng mang lại khiến cho chúng được áp dụng trong nhiều ứng dụng. Có

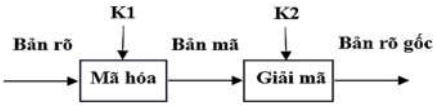
thể hình dung hệ mật này tương tự như sau. A đặt một vật vào một hộp kim loại và rồi khoá nó lại bằng một khoá số do B để lại. Chỉ có B là người duy nhất có thể mở được hộp vì chỉ có người đó mới biết tổ hợp mã của khoá số của mình. Thuật toán mã hóa công khai là thuật toán được thiết kế sao cho. Khóa mã hóa là khác so với khóa giải mã. Mà khóa giải mã hóa không thể tính toán được từ khóa mã hóa. Khóa mã hóa gọi là khóa công khai (public key), khóa giải mã được gọi là khóa riêng (private key).

2. MỘT SỐ VẤN ĐỀ MÃ HÓA DỮ LIỆU

2.1. Khái niệm về mã hóa dữ liệu

Mã hóa dữ liệu là quá trình chuyển đổi các thông tin thông thường (văn bản thường hay văn bản rõ) thành dạng không đọc trực tiếp được, là văn bản mã hóa. Giải mã dữ liệu là quá trình ngược lại, phục hồi lại văn bản thường từ văn bản mã.

Quá trình mã hoá và giải mã được thể hiện trong sơ đồ hình 1.



Hình 1. Quy trình mã hóa dữ liệu

Trong đó:

- Bản rõ (Plaintext or Cleartext): chứa các xâu ký tự gốc, thông tin trong bản rõ là thông tin cần mã hoá để giữ bí mật.

- Mã hóa (Encryption): quá trình chuyển đổi dữ liệu gốc thành dữ liệu được mã hóa sao người khác không thể đọc hiểu được.

- Bản mã (Ciphertext): chứa các ký tự sau khi đã được mã hoá, mà nội dung được giữ bí mật.

- Giải mã (Decryption): quá trình biến đổi trả lại bản mã bản thành bản rõ gọi là giải mã.

- K1 là khóa để mã hóa và K2 khóa để giải mã.

2.2. Thuật toán MD5

Thuật toán MD5 (Message Digest 5) [4], do Ronald Rivest thiết kế năm 1991, là xây dựng một hàm băm để mã hóa một tín hiệu vào có chiều dài bất kỳ và đưa ra một tín hiệu (Digest) ở đầu ra có chiều dài cố định 128 bit (tương ứng với 32 chữ số hệ 16).

Input: thông điệp với độ dài bất kỳ

Output: giá trị băm (message digest) 128 bits

Giải thuật gồm 5 bước:

Bước 1: nhồi dữ liệu

Nhồi thêm các bits sao cho dữ liệu có độ dài l ≡ 448 mod 512 hay l = n * 512 + 448 (n,l nguyên). Số lượng bit nhồi thêm nằm trong khoảng 1 đến 512. Các bit được nhồi gồm 1 bit "1" và các bit 0 theo sau.

Bước 2: thêm vào độ dài

Độ dài của khối dữ liệu ban đầu được biểu diễn dưới dạng nhị phân 64bit và được thêm vào cuối chuỗi nhị phân kết quả của bước 1.

Nếu độ dài của khối dữ liệu ban đầu > 2^{64}, chỉ 64 bits thấp được sử dụng, nghĩa là giá trị được thêm vào bằng K mod 2^{64}. Kết quả có được từ 2 bước đầu là một khối dữ liệu có độ dài là bội số của 512.

Bước 3: khởi tạo bộ đệm MD (MD buffer)

Một bộ đệm 128bit được dùng lưu trữ các giá trị băm trung gian và kết quả. Bộ đệm được biểu diễn bằng 4 thanh ghi 32 bit với các giá trị khởi tạo ở dạng littleiendian (byte có trọng số nhỏ nhất trong từ nằm ở địa chỉ thấp nhất) như sau:

A = 67 45 23 01

B = EF CD AB 89

C = 98 BA DC FE

D = 10 32 54 76

Các giá trị này tương đương với các từ 32 bit sau:

A = 01 23 45 67

B = 89 AB CD EF

C = FE DC BA 98

D = 76 54 32 10

Bước 4: Xử lý các khối dữ liệu 512 bit

Trọng tâm của giải thuật là hàm nén (compression function) gồm 4 "vòng" xử lý. Các vòng này có cấu trúc giống nhau nhưng sử dụng các hàm luận lý khác nhau gồm F, G, H và I như sau:

$F(X,Y,Z) = X \wedge Y \vee \neg X \wedge Z$

$G(X,Y,Z) = X \wedge Z \vee Y \wedge \neg Z$

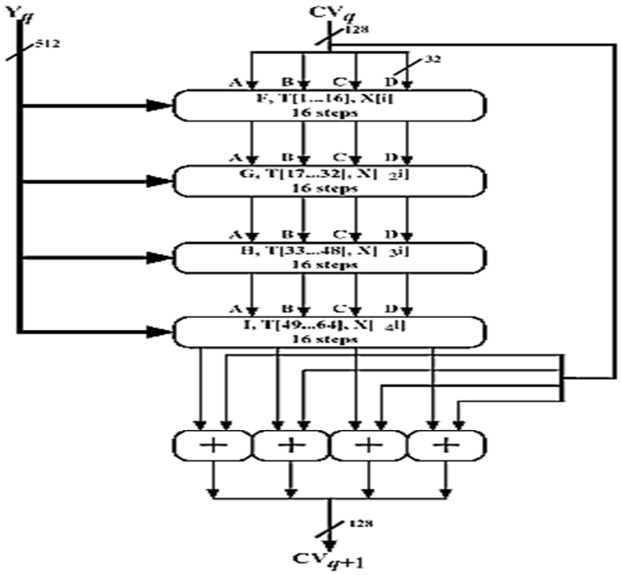
$H(X,Y,Z) = X \text{ xor } Y \text{ xor } Z$

$I(X,Y,Z) = Y \text{ xor } (X \vee \neg Z)$

Mảng 64 phần tử được tính theo công thức:

$T[i] = 2^{32} \times abs(\sin(i))$, i được tính theo radian.

Kết quả của 4 vòng được cộng (theo modulo 2^{32} với đầu vào CV_q để tạo CV_{q+1}



Hình 2 . Thuật toán MD5 xử lý các khối dữ liệu 512 bit [4]

Các giá trị trong bảng T (bảng 1).

Bảng 1. Giá trị của T

T[1] = d76aa478	T[17] = f61e2562	T[33] = fffa3942	T[49] = f4292244
T[2] = e8c7b756	T[18] = c040b340	T[34] = 8771f681	T[50] = 432aff97
T[3] = 242070db	T[19] = 265e5a51	T[35] = 6d9d6122	T[51] = ab9423a7
T[4] = c1bdceee	T[20] = e9b6c7aa	T[36] = fde5380c	T[52] = fc93a039
T[5] = f57c0faf	T[21] = d62f105d	T[37] = a4beea44	T[53] = 655b59c3
T[6] = 4787c62a	T[22] = 2441453	T[38] = 4bdecfa9	T[54] = 8f0ccc92
T[7] = a8304613	T[23] = d8a1e681	T[39] = f6bb4b60	T[55] = ffeff47d
T[8] = fd469501	T[24] = e7d3fbc8	T[40] = bebfbc70	T[56] = 85845dd1
T[9] = 698098d8	T[25] = 21e1cde6	T[41] = 289b7ec6	T[57] = 6fa87e4f
T[10] = 8b44f7af	T[26] = c33707d6	T[42] = eaa127fa	T[58] = fe2ce6e0
T[11] = ffff5bb1	T[27] = f4d50d87	T[43] = d4ef3085	T[59] = a3014314
T[12] = 895cd7be	T[28] = 455a14ed	T[44] = 4881d05	T[60] = 4e0811a1
T[13] = 6b901122	T[29] = a9e3e905	T[45] = d9d4d039	T[61] = f7537e82
T[14] = fd987193	T[30] = fcefa3f8	T[46] = e6db99e5	T[62] = bd3af235
T[15] = a679438e	T[31] = 676f02d9	T[47] = 1fa27cf8	T[63] = 2ad7d2bb
T[16] = 49b40821	T[32] = 8d2a4c8a	T[48] = c4ac5665	T[64] = eb86d391

Bước 5: Xuất kết quả

Sau khi xử lý hết L khối 512 bit, đầu ra của lần xử lý thứ L là giá trị băm 128 bits.

Dưới đây là các ví dụ mô tả các kết quả qua chương trình thực nghiệm thuật toán MD5

MD5("chữ ký điện tử") =
bee46aae61a033c5e3d0e15258ca3b97

Chỉ một thay đổi (chẳng hạn viết hoa chữ c thành C) cũng làm thay đổi hoàn toàn kết quả:

MD5("Chữ ký điện tử") =
24535bc1b2e19d148d8beee62797008a

2.3. Thuật toán RSA (RSA - Ron **R**ivest, Adi **S**hamir và Leonard **A**dleman).

Phương pháp sử dụng thuật toán RSA được đặt tên dựa theo chữ cái đầu của 3 tác giả của hệ mã Rivest, Shamir và Adleman. Đây là thuật toán mã hóa nổi tiếng nhất và cũng là thuật toán được ứng dụng thực tế nhất. Dựa trên nền tảng lý thuyết về phân tích thừa số nguyên tố của số nguyên lớn, phương pháp RSA được ứng dụng vào mô hình mã hóa, mô hình truyền nhận khóa và mô hình chữ ký điện tử.

Thuật toán RSA có hai khóa: khóa công khai (hay khóa công cộng) và khóa bí mật (hay khóa cá nhân). Mỗi khóa là những số cố định sử dụng trong quá trình mã hóa và giải mã. Khóa công khai được công bố rộng rãi cho mọi người và được dùng để mã hóa. Những thông tin được mã hóa bằng khóa công khai chỉ có thể được giải mã bằng khóa bí mật tương ứng.

Để cài đặt RSA ban đầu mỗi người dùng sinh khóa công khai và khóa bí mật của mình bằng cách:

Tạo khóa công khai và khóa bí mật:

• Chọn hai số nguyên tố lớn ngẫu nhiên (cỡ gần 100 chữ số) khác nhau p và q

• Tính N = p*q

• Chọn một số e nhỏ hơn N và (e, ∅(N)) = 1, e được gọi là số mũ lập mã

• Tìm phần tử ngược của e trên vành module ∅ (N), d là số mũ giải mã

khóa công khai là K_P = (e, N)

khóa bí mật là $K_S = K_P^{-1}$ =(d, N)

Quá trình mã hóa và giải mã:

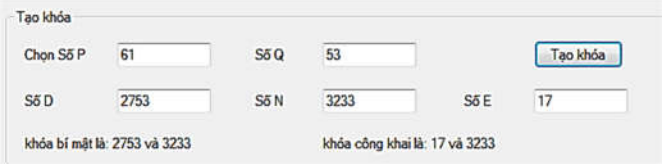
Giả sử khối bản gốc của người gửi là M, khối bản mã của người nhận là C.

Để mã hóa một thông điệp M ta sử dụng công thức sau:

$C = M^e \pmod N$ $(0 <= M < N)$

Để giải mã: $M = C^d \pmod N$

Hình 3 là các kết quả minh họa tìm khóa công khai và khóa bí mật qua chương trình thực nghiệm.Giả sử ta có P = 61 và q = 53, ta tìm được khóa công khai là (2753, 3233), khóa bí mật là (17, 3233).



Hình 3. Minh họa quá trình tạo khóa bí mật và khóa công khai

3. CHỮ KÝ ĐIỆN TỬ

Chữ ký điện tử (electronic signature) hoạt động dựa trên hai khoá là public key và private key và thực hiện qua hai giai đoạn là việc hình thành chữ ký trên tài liệu ở phía người gửi và việc xác nhận tài liệu nhận được chính xác và nguyên vẹn hay không ở phía người nhận. Do đó, thuật toán băm MD5 và thuật toán mã hóa RSA có thể được áp dụng để xây dựng ứng dụng chữ ký điện tử.

Vấn đề bảo mật ở electronic signature không giống với các phương pháp mã hoá cổ điển là chỉ dùng một khoá cho cả việc mã hoá ở người gửi và giải mã ở người nhận mà chữ ký điện tử sử dụng hai khoá: private key để mã hoá và public key để giải mã kiểm tra.

3.1. Quá trình tạo chữ ký điện tử và gửi tệp văn bản

- Tạo khóa bí mật, khóa công khai như đã trình bày trong mục 3 và gửi khóa công khai cho người nhận.

- Từ file dữ liệu ban đầu, chương trình sử dụng hàm băm MD5 để mã hóa thành chuỗi kí tự dài 128 bit. Thông điệp sẽ được băm thành thông điệp rút gọn.

- Tạo chữ ký điện tử bằng cách mã hóa thông điệp bằng khóa bí mật vừa tạo

- Kết hợp file ban đầu với chữ điện tử thành một thông điệp và gửi cho người nhận.

Minh họa kết quả của quá trình tạo chữ ký điện tử qua phần thực nghiệm của chương trình.

Giả sử nội dung của file văn bản như sau:

"Trời đã bắt đầu ấm dần. Từng đàn chim nối đuôi nhau bay lượn, hót ríu rít chào đón xuân sang. Cỏ cây như bừng tỉnh dậy sau những tháng ngày lạnh lẽo của mùa đông, xôn xao khoe chồi non, lộc biếc. Trong vườn, trăm loài hoa đua nở. Ong bướm dập dìu bay lượn quấn quýt trong màu hoa hương hoa. Nắng xuân vàng tươi. Cảnh núi sông đẹp như gấm hoa. Ai cũng thấy lòng mình phơi phới."

Hàm băm MD5 đã mã hóa đoạn văn bản trên thành một thông điệp rút gọn sau:

36 148 150 84 134 246 187 51 73 122 43 153 79 165 237 (*)

Bước tiếp theo tạo chữ ký điện tử cùng với khóa bí mật ta thu được kết quả như sau:

2761 2220 2641 1533 2046 16 966 1076 2630 565 1159 2105 2080 2165 1312 1786.

3.2. Quá trình kiểm tra xác nhận chữ ký trên tài liệu

Kỹ thuật chữ ký điện tử cho phép người nhận message có kèm chữ ký kiểm tra tính xác thực và tính toàn vẹn của nó. Quá trình kiểm tra chữ ký điện tử nhằm mục đích xác định một message gửi đi đã được ký bằng khoá private key đúng với khóa public key gửi đi hay không. Như vậy việc kiểm tra một chữ ký điện tử được thực hiện trong 3 bước:

Bước 1: Tính Current Hash-Value.

Trong bước một, một hash-value của message đã ký được tính. Sử dụng thuật toán băm như đã dùng trong suốt quá trình ký. Hash-value nhận được được gọi là current hash-value bởi vì nó được tính từ trạng thái hiện thời của message.

Bước 2: Tính Original Hash-Value.

Trong bước hai của quá trình kiểm tra chữ ký điện tử, electronic signature được giải mã với cũng với thuật toán mã hoá đã được sử dụng trong suốt quá trình ký. Việc giải mã được thực hiện bằng khoá public key tương ứng với khoá private key được dùng trong suốt quá trình ký của message. Kết quả chúng ta nhận được original hash-value mà đã được tính từ message gốc trong suốt bước một của quá trình ký (original message digest).

Bước 3: So sánh current hash-value với original hash-value.

Trong bước ba, chúng ta đối chiếu current hash-value nhận được trong bước một với original hash-value nhận được trong bước hai. Nếu hai giá trị này giống hệt nhau thì việc kiểm tra sẽ thành công nếu chứng minh được message đã được ký với khoá private key đúng với khoá public key đã được dùng trong quá trình kiểm tra và ngược lại.

Minh họa kết quả của các bước kiểm tra xác nhận chữ ký trên tài liệu.

Ví dụ: Đoạn văn bản dưới đây và chữ ký điện tử vừa được gửi tới người nhận.

"Trời đã bắt đầu ấm dần. Từng đàn chim nối đuôi nhau bay lượn, hót ríu rít chào đón xuân sang. Cỏ cây như bừng tỉnh dậy sau những tháng ngày lạnh lẽo của mùa đông, xôn xao khoe chồi non, lộc biếc. Trong vườn, trăm loài hoa đua nở. Ong bướm dập dìu bay lượn quấn quýt trong màu hoa hương hoa. Nắng xuân vàng tươi. Cảnh núi sông đẹp như gấm hoa. Ai cũng thấy lòng mình phơi phới."

(chữ ký điện tử: 2761 2220 2641 1533 2046 16 966 1076 2630 565 1159 2105 2080 2165 1312 1786)

Người nhận sẽ tiến hành kiểm tra xác nhận chữ ký trên tài liệu.

- Mở nội dung đoạn văn bản

- Dùng hàm băm băm đoạn văn bản và thu được kết quả sau khi băm là:

149 89 124 229 166 107 174 73 172 18 15 253 108 197 101 137

- Nhập khóa giải mã (e, N) ta thu được kết quả chữ ký số sau giải mã là:

12 236 148 150 84 134 246 187 51 73 122 43 153 79 165 237 (**)

Kết luận: Kết quả mã hóa thông điệp (*) và kết quả thu được sau khi giải mã (**) hoàn toàn trùng khớp, vậy nội dung file văn bản của người gửi và nội dung file văn bản của người nhận hoàn toàn được bảo mật qua xác minh chữ ký điện tử.

4. KẾT LUẬN

Bài báo đã nêu được quy trình ứng dụng chữ ký điện tử trên cơ sở kết hợp giữa thuật toán băm MD5 và thuật toán mã hóa RSA. Từ đó, bài báo đã minh họa chương trình thực nghiệm ứng dụng chữ ký điện tử trong quá trình gửi và nhận các tệp văn bản nhằm đảm bảo tính bảo mật của dữ liệu. Mục đích của chữ kí điện tử: đảm bảo thông điệp gửi đi không bị mất mát hay thay đổi, giữ nguyên tình trạng ban đầu khi gửi đồng thời xác thực các đối tượng liên quan và xác thực nội dung thông điệp được gửi đi và nhận được là chính xác và không bị thay đổi.

TÀI LIỆU THAM KHẢO

[1]. Hartini Saripana , Zaiton Hamin, 2011. *The application of the digital signature law in securing internet banking: some preliminary evidence from Malaysi.* Procedia Computer Science 3, Published by Elsevier Ltd, 248–253.

[2]. S. Mason, 2005. *Digital Signatures: Is that really you*?. IEE Engineering Management, 9 - 12.

[3]. Suranjan Choudhury, Kartik Bhanagar, Wasim Haque, NIIT, 2002. *Public key infrastructure Implemetion and Design*. M & T Books.

[4]. R. Rivest, 1992. *The MD5 Message-Digest Algorithm*. MIT Laboratory for Computer Science and RSA Data Security, Inc.

AUTHOR INFORMATION

Ninh Van Tho

University of Economics - Technology for Industries