Phát triển thuật toán mật mã khóa công khai dựa trên hệ mật Elgamal

Development of Public Key Cryptographic Algorithm Based on Elgamal Cryptosystem

Lưu Hồng Dũng

Abstract: *This paper proposed a new public key cryptographic algorithm based on the ElGamal cryptosystem. This algorithm has the capacity of information security and anthentication information. The paper also offers analysis on the safety of the proposed schemes, has shown the ability to apply it in practice.*

I. ĐẶT VẤN ĐỀ

Trong thực tế, nhiều ứng dụng đòi hỏi việc bảo mật thông tin cần phải được thực hiện đồng thời với việc xác thực về nguồn gốc và tính toàn vẹn của thông tin. Nội dung bài báo đề xuất một thuật toán mật mã khóa công khai được phát triển dựa trên hệ mật ElGamal [1] cho phép giải quyết tốt các yêu cầu nêu trên.

II. PHÁT TRIỂN THUẬT TOÁN MẬT MÃ KHÓA CÔNG KHAI DỰA TRÊN HỆ MẬT ELGAMAL

1. Hệ mật Elgamal

Hệ mật Elgama được đề xuất vào năm 1984 trên cơ sở bài toán logarith rời rạc. Sau đó, các chuẩn chữ ký số DSS [2] của Mỹ và GOST R34.10-94 [3] của Liên bang Nga đã được hình thành trên cơ sở hệ mật này.

1.1 Thuật toán hình thành tham số và khóa

Các thành viên trong hệ thống muốn trao đổi thông tin mật với nhau bằng thuật toán mật mã Elgamma thì trước tiên thực hiện quá trình hình thành khóa như sau:

1- Chọn số nguyên tố đủ lớn p sao cho bài toán logarit trong Z_p là khó giải.

2- Chọn $g \in Z_p^*$ là phần tử nguyên thủy.

3- Chọn khóa mật x là số ngẫu nhiên sao cho: $1 < x < p$. Tính khóa công khai y theo công thức: $y = g^x \bmod p$.

1.2 Thuật toán mã hóa

Giả sử người gửi là A, người nhận là B. Người gửi A có khóa bí mật là xa và khóa công khai là ya. Người nhận B có khóa bí mật là xb và khóa công khai là yb. Khi đó, để gửi bản tin M cho B, với: $0 \le M < p$, người gửi A sẽ thực hiện các bước như sau:

1- Chọn số ngẫu nhiên k thỏa mãn: $1 < k < p$. Tính giá trị R theo công thức:

$$R = g^k \bmod p.$$

2- Sử dụng khóa công khai của B để tính:

$$C = M \times (y_b)^k \bmod p$$

3- Gửi bản mã gồm (C, R) đến người nhận B.

1.3 Thuật toán giải mã

Để khôi phục bản tin ban đầu (M) từ bản mã (C, R) nhận được, người nhận B thực hiện các bước như sau:

1- Tính giá trị Z theo công thức:

$$Z = R^{x_b} \bmod p = g^{k.x_b} \bmod p$$

2- Tính gía trị nghịch đảo của Z:

$$Z^{-1} = \left(g^{k.x_b}\right)^{-1} \bmod p = g^{-k.x_b} \bmod p$$

3- Khôi phục bản tin ban đầu (*M*):

$$M = C \times Z^{-1} \bmod p$$

1.4 Tính đúng đắn của thuật toán mật mã Elgamal

Giả sử bản tin nhận được sau quá trình giải mã (C, R) là $\overline{M}$, thế thì:

$$\begin{aligned}\overline{M} &= C \times Z^{-1} \bmod p = \\ &= [(M \times (y_b)^k \bmod p) \times (g^{-k.x_b} \bmod p)] \bmod p \\ &= M \times g^{k.x_b} \times g^{-k.x_b} \bmod p \\ &= M\end{aligned}$$

Như vậy, bản tin nhận được sau giải mã ($\overline{M}$) chính là bản tin ban đầu (*M*).

1.5 Mức độ an toàn của hệ mật ElGamal

Hệ mật ElGamal sẽ bị phá vỡ nếu khóa mật *x* hoặc *k* có thể tính được. Để tính được *x* hoặc *k*, cần phải giải một trong hai bài toán logarit rời rạc, chẳng hạn: $y = g^x \bmod p$ hay: $R = g^k \bmod p$.

Tuy nhiên, trong thực tế việc giải bài toán logarit rời rạc này là việc khó [4].

Một điểm yếu có thể bị tấn công trong hệ mã ElGamal là khi giá trị *k* bị sử dụng lại. Thực vậy, giả sử cùng một giá trị *k* được sử dụng để mã hóa hai bản tin *M* và M^* được các bản mã tương ứng là (C, R) và (C^*, R^*). Khi ấy ta sẽ có:

$$\begin{aligned}&C \times (C^*)^{-1} \equiv \\ &\equiv M \times (g^x)^k \times (M^* \times (g^x)^k)^{-1} \bmod p\end{aligned}$$

Suy ra:

$$M^* \times C \times (C^*)^{-1} \equiv M \bmod p$$

Nghĩa là, chỉ cần biết nội dung của một trong hai bản tin *M* hoặc M^* thì sẽ dễ dàng biết được nội dung của bản tin kia.

Một vấn đề được đặt ra không chỉ với hệ ElGamal nói riêng mà với các hệ mật khóa công khai nói chung là: Giả sử một người gửi A mã hóa bản tin *M* được bản mã C và gửi C cho người nhận B. Sẽ có các tình huống xảy ra như sau:

- Người nhận B không thể biết chắc chắn rằng bản tin nhận được (*M*) có nguồn gốc từ người gửi A.
- Giả sử bản mã *C* đã bị thay đổi thành C^* và người nhận giải mã được bản tin M^*. Trường hợp này, người nhận không thể khẳng định được nội dung bản tin do A gửi đi đã bị thay đổi.

Thuật toán mật mã được đề xuất, phát triển trên cơ sở hệ mật ElGamal sau đây sẽ là giải pháp cho các vấn đề nêu trên.

2. Thuật toán mật mã khóa công khai dựa trên hệ mật Elgamal

2.1. Thuật toán hình thành tham số và khóa

1- Sinh cặp số nguyên tố *p* và *q* đủ lớn và: $q \mid (p-1)$ sao cho bài toán logarit trong Z_p là khó giải.

2- Chọn $g = \alpha^{(p-1)/q} \bmod p$, là phần tử sinh có bậc *q* của nhóm Z_p^*, nghĩa là: $1 < g < p$ và: $g^q \equiv 1 \bmod p$. Ở đây: α là một số nguyên thỏa mãn: $1 < \alpha < (p-1)$.

3- Khóa bí mật *x* là một giá trị được chọn ngẫu nhiên trong khoảng: $1 < x < q$. Khóa công khai *y* được tính theo công thức:

$$y = g^x \bmod p$$

4- Lựa chọn hàm băm (hash) H: $\{0,1\}^* \mapsto Z_q$ (Ví dụ: SHA-1, MD5).

2.2 Thuật toán mã hóa

Giả sử người gửi A có khóa bí mật là x_a, khóa công khai tương ứng là y_a; người nhận B có khóa bí mật là x_b và khóa công khai là y_b. Để gửi bản tin *M* cho B, A thực hiện các bước như sau:

1- Chọn số ngẫu nhiên *k* thỏa mãn: $1 < k < q$.

2- Tính giá trị R theo công thức:

$$R = g^k \bmod p \qquad (1)$$

3- Tính thành phần E theo công thức:

$$E = H(R \| M) \bmod q \quad (2)$$

4- Tính thành phần S theo công thức:

$$S = [k - x_a \times E] \bmod q \quad (3)$$

5- Sử dụng khóa công khai của B để tính thành phần C theo công thức:

$$C = M \times (y_b)^k \bmod p \quad (4)$$

6- Gửi bản mã gồm (C, E, S) đến người nhận B.

Tương tự thuật toán Elgamal, thuật toán mật mã mới đề xuất có thể bị tấn công khi sử dụng lại giá trị của *k*. Có thể khắc phục yếu điểm trên nếu sử dụng giá trị $H(x \| M)$ thay cho *k*, với *H(.)* là hàm băm kháng va chạm (chẳng hạn SHA-1, MD5,...) và "||" là toán tử nối/xáo trộn xâu. Khi đó thuật toán mã hóa sẽ được mô tả lại như sau:

1- Tính giá trị R theo công thức:

$$R = g^{H(x_a \| M)} \bmod p$$

2- Tính thành phần E theo công thức:

$$E = H(R \| M) \bmod q$$

3- Tính thành phần S theo công thức:

$$S = [H(x_a \| M) - x_a \times E] \bmod q$$

4- Sử dụng khóa công khai y_b của người nhận để tính thành phần C theo công thức:

$$C = M \times y_b^{H(x_a \| M)} \bmod p$$

5- Gửi bản mã gồm (C, E, S) đến người nhận B.

2.3 Thuật toán giải mã

Từ bản mã (C, E, S) nhận được, B khôi phục và kiểm tra nguồn gốc cũng như tính toàn vẹn của bản tin ban đầu (M) như sau:

1- Tính giá trị $\overline{R}$ theo công thức:

$$\overline{R} = g^S \times (y_a)^E \bmod p \quad (5)$$

2- Khôi phục bản tin ban đầu theo công thức:

$$\overline{M} = C \times (\overline{R})^{-x_b} \bmod p \quad (6)$$

3- Tính thành phần $\overline{E}$ theo công thức:

$$\overline{E} = H(\overline{R} \| \overline{M}) \bmod q \quad (7)$$

4- Kiểm tra nếu $\overline{E} = E$ thì $\overline{M} = M$ và *M* có nguồn gốc từ người gửi A.

Chú ý:

Trường hợp sử dụng giá trị $H(x \| M)$ thay cho *k*, thuật toán giải mã vẫn được giữ nguyên.

2.4 Tính đúng đắn của thuật toán mới đề xuất

Tính đúng đắn của hệ mật mới đề xuất là sự phù hợp giữa thuật toán giải mã với thuật toán mã hóa. Điều cần chứng minh ở đây là:

Cho: p, q là 2 số nguyên tố độc lập, $g = \alpha^{(p-1)/q} \bmod p$, $g = \alpha^{(p-1)/q} \bmod p$, $\alpha \in Z_p^*$, $x_a, x_b \in [1, q-1]$, $y_a = g^{x_a} \bmod p$, $y_b = g^{x_b} \bmod p$, $H : \{0,1\}^* \mapsto Z_p$, $R = g^{H(x_a \| M)} \bmod p$, $E = H(R \| M) \bmod q$, $S = [H(x_a \| M) - x_a \times E] \bmod q$, $C = M \times (y_b)^{H(x_a \| M)} \bmod p$.

Nếu: $\overline{R} = (g)^S \times (y_a)^E \bmod p$, $\overline{M} = C \times (\overline{R})^{-x_b} \bmod p$, $\overline{E} = H(\overline{R} \| \overline{M}) \bmod q$ thì: $\overline{M} = M$ và $\overline{E} = E$.

Chứng minh:

Thật vậy, thay (3) vào (5) ta có:

$$\begin{aligned} \overline{R} &= g^S \times {y_a}^E \bmod p \\ &= g^{k - x_a \times E} \times {y_a}^E \bmod p \\ &= g^k \times g^{-x_a \times E} \times {y_a}^E \bmod p \\ &= R \times {y_a}^{-E} \times {y_a}^E \bmod p \\ &= R \end{aligned} \quad (8)$$

Từ (8) suy ra:

$$\overline{R} = g^k \bmod p \quad (9)$$

Thay (4) và (9) vào (6) ta có:

$$\begin{aligned} \overline{M} &= C \times (\overline{R})^{-x_b} \bmod p \\ &= \{M \times {y_b}^k \bmod p \times \\ &\times (g^k)^{-x_b} \bmod p\} \bmod p = \\ &= M \times {y_b}^k \times {y_b}^{-k} \bmod p = M \end{aligned} \quad (10)$$

Thay (8) và (10) vào (7) ta được:

$$\begin{aligned} \overline{E} &= H(\overline{R} \| \overline{M}) \bmod q \\ &= H(R \| M) \bmod q \\ &= E \end{aligned}$$

Đây là điều cần chứng minh.

2.5 Mức độ an toàn của thuật toán mới đề xuất

Mức độ an toàn của thuật toán mới đề xuất được đánh giá bằng:

1) Khả năng chống thám mã.
2) Khả năng chống giả mạo về nguồn gốc và nội dung của bản tin.

Về khả năng chống thám mã, có thể thấy rằng mức độ an toàn của thuật toán mật mã mới đề xuất là hoàn toàn tương đương với thuật toán Elgamal. Hệ mật sẽ bị phá vỡ nếu khóa mật *x* có thể tính được từ việc giải bài toán logarit rời rac như: $y = g^x \bmod p$ hay: $S = [H(x \| M) - x \times E] \bmod q$

Khả năng chống giả mạo về nguồn gốc và nội dung bản tin phụ thuộc vào mức độ khó của việc tạo ra các cặp (C, E, S) giả mạo thỏa mãn điều kiện kiểm tra của thuật toán giải mã: $\overline{E} = E$. Giả sử cặp (E^*, S^*) là giả mạo, thế thì nó cần phải thỏa mãn:

$$E^* \equiv H((g^{S^*} \times y^{E^*} \bmod p) \| M) \bmod q$$

Với *H(.)* là hàm băm kháng va chạm, hơn nữa nội dung bản tin M cũng chưa biết thì việc chọn được cặp (E^*, S^*) giả mạo là hầu không thể thực hiện được.

III. KẾT LUẬN

Bài báo đề xuất một thuật toán mật mã khóa công khai được phát triển dựa trên hệ mật ElGamal có khả năng đồng thời:

1) Bảo mật thông tin.
2) Xác thực về nguồn gốc thông tin.
3) Xác thực về tính toàn vẹn của thông tin.

Hơn nữa, mặc dù bản mã được tạo ra bởi thuật toán mới đề xuất bao gồm 3 thành phần (C,E,S) nhưng độ dài của nó không lớn hơn độ dài bản mã 2 thành phần (C,R) mà thuật toán ElGamal tạo ra.

Giả sử |p| = 512 bit, |q| = 160 bit khi đó độ dài bản mã do thuật toán ElGamal tạo ra là: |C| + |R| = 512 bit + 512 bit = 1024 bit, còn độ dài bản mã do thuật toán mới đề xuất tạo ra là: |C| + |E| + |S| = 512 bit + 160 bit + 160 bit = 832 bit.

Những phân tích về mức độ an toàn cho thấy khả năng ứng dụng thuật toán mới đề xuất trong thực tế là hoàn toàn có thể.

TÀI LIỆU THAM KHẢO.

[1]. T. ELGAMAL. *A public key cryptosystem and a signature scheme based on discrete logarithms*. IEEE Transactions on Information Theory. 1985, Vol. IT-31, No. 4. pp.469–472.

[2]. *National Institute of Standards and Technology, NIST FIPS PUB 186-3*. Digital Signature Standard, U.S. Department of Commerce,1994.

[3]. GOST R 34.10-94. Russian Federation Standard. Information Technology. *Cryptographic data Security. Produce and check procedures of Electronic Digital Signature based on Asymmetric Cryptographic Algorithm*. Government Committee of the Russia for Standards, 1994 (in Russian).

[4]. D.R STINSON, *CRYPTOGRAPHY, Theory and Practice*, CRC Press 1995.

Nhận bài ngày: 20/03/2012

SƠ LƯỢC VỀ TÁC GIẢ

LƯU HỒNG DŨNG

Sinh năm 1966 tại Hưng Yên.

Tốt nghiệp Đại học chuyên ngành Vô tuyến điện tử năm 1989 và Cao học ngành Kỹ thuật Điện tử và Thông tin liên lạc năm 2000 tại Học viện Kỹ thuật Quân sự.

Hiện là giảng viên Khoa Công nghệ Thông tin, Học viện Kỹ thuật Quân sự.

Lĩnh vực nghiên cứu: An toàn và bảo mật thông tin, An ninh mạng máy tính.

Email: luuhongdung@gmail.com