PHÂN TÍCH ĐỘ AN TOÀN CỦA THUẬT TOÁN MẬT MÃ NTRU

Phạm Quốc Hoàng, Phạm Thị Hiền
Học viện Kỹ thuật mật mã

Thuật toán NTRU (Nth degree Truncated polynomial Ring Units) là thuật toán mật mã khóa công khai dựa trên lý thuyết lưới với độ an toàn phụ thuộc vào độ khó của tìm phân tích "ngắn" cho các đa thức đại số trên vành [1][6]. Bài toán này tương đương với tìm véc-tơ ngắn nhất SVP (Shortest Vector Problem) trong một lưới thực sự 2N chiều [2][3]. Bài viết trình bày về thuật toán NTRU và một số phương pháp phá vỡ thuật toán này [5].

Hệ mật khóa công khai dựa trên lý thuyết lưới đã và đang trở thành một xu hướng chính trong nghiên cứu mật mã hiện đại, do ưu điểm về tốc độ mã hóa và giải mã, cũng như việc lưu trữ khóa cần ít tài nguyên mà độ an toàn vẫn được đảm bảo. Đặc biệt, các hệ mật này được phỏng đoán có khả năng kháng lại tấn công sử dụng máy tính lượng tử, được chứng minh an toàn dựa trên các giả thiết trong trường hợp xấu nhất. Năm 1998, NTRU là hệ mật khóa công khai được đề xuất bởi J.Hoffstein, J.Pipher và J. Silverman.

Từ đó tới nay, hệ mật này được nghiên cứu phát triển cải tiến và đã thể hiện được những ưu điểm mà người sử dụng kì vọng về hệ mật sử dụng lưới. Độ an toàn của NTRU vẫn được bảo đảm, trong đó đặc biệt theo đánh giá của NIST, hệ mật NTRU là một trong những hệ mật có khả năng kháng lại tấn công dựa trên tính toán lượng tử tốt nhất. Khi có máy tính lượng tử đủ mạnh, thuật toán Shor sẽ được dùng để thám hệ mật RSA. Hệ mật NTRU có độ an toàn dựa trên lưới, nên hiện chưa có thuật toán nào dùng cho máy tính lượng tử thám được NTRU. Năm 2009, NTRU đã được đưa vào chuẩn mật mã khóa công khai IEEE P1363.1. Năm 2020, NTRU tiếp tục được cải tiến và lọt vào vòng 3 cuộc thi mật mã hậu lượng tử của NIST [5][7].

Thuật toán NTRU được ứng dụng trong hệ thống kiểm soát và thu thập dữ liệu (Supervisory Control and Data Acquisition System) [8]. Lý do nó được ứng dụng vì có tốc độ làm việc nhanh hơn so với thuật toán RSA và ECC.

THUẬT TOÁN MẬT MÃ NTRU

Các khái niệm và định nghĩa liên quan đến hệ mật

Các kí hiệu và thông số dưới đây sử dụng số là số nguyên tố lẻ, với các định nghĩa như sau: modulo lớn mà mỗi hệ số được rút gọn (không bí mật);

1. q: modulo lớn mà mỗi hệ số được rút gọn (không bí mật);

2. p: modulo nhỏ mà mỗi hệ số được rút gọn (không bí mật);

3. f: Khóa riêng, hay còn gọi là khóa bí mật (là một đa thức);

4. g: Một đa thức được sử dụng để tạo khóa công khai từ (bí mật nhưng bị loại bỏ sau lần sử dụng ban đầu);

5. h: khóa công khai (là một đa thức);

6. r: đa thức "làm mù" ngẫu nhiên (bí mật nhưng bị loại bỏ sau lần sử dụng đầu tiên);

7. d_f, d_g, d_r: tương ứng là số các hệ số bằng 1 trong các đa thức f,g,r.

Tập hợp tham số NTRU là $(N, p, q, L_f, L_g, L_r, L_m)$, với N là số nguyên tố đủ lớn, p,q là các số nguyên dương nguyên tố cùng nhau, q lớn hơn đáng kể so với p; L_f, L_g, L_r, L_m là tập hợp các đa thức bậc $N-1$ với hệ số nguyên.

Các phép toán của thuật toán mật mã NTRU được thực hiện trên vành đa thức $Z[x]/(x^N-1)$ gồm các đa thức có hệ số nguyên và bậc nhỏ hơn hoặc bằng $N-1$. Ví dụ $a(x)$, $b(x)$ là hai đa thức thuộc R thì có 2 phép toán đó là:

$a(x)+b(x)=c(x) mod (x^N-1)$ và

$a(x).b(x)=c'(x) mod (x^N-1)$.

Trong đó, một phần tử $f \in Z[x]/(x^N-1)$ có thể được viết dưới dạng một đa thức hoặc một véc-tơ như sau:

$$f = \sum_{i=0}^{N-1} f_i x^i = (f_0, f_1, \dots, f_{N-1})$$

Phép nhân được xác định trên vành này sẽ được kí hiệu là $\otimes$ và được xác định theo biểu thức sau:

$$f \otimes g = h \text{ với } h_k = \sum_{i=0}^{k} f_i g_{k-i} + \sum_{i=k+i}^{N-1} f_i g_{N+k-i} = \sum_{i+j \equiv k \bmod N} f_i g_j$$

Một vành đa thức $R=Z_q[x]/(x^N-1)$ có các phép toán tương tự như trên vành đa thức $Z[x]/(x^N-1)$ nhưng tất các các hệ số được rút gọn theo modulo q.

Quá trình tạo khóa

Để tạo ra một khóa cho NTRU, Bob chọn ngẫu nhiên hai đa thức $F \in R=Z_q[x]/(x^N-1)$ và $g \in L_g$. Đặt $f=1+2F$. Các đa thức f, g với ngẫu nhiên d_f, d_g hệ số bằng 1 và phần còn lại bằng 0; f, g phải có nghịch đảo theo modulo q và modulo p.

Đối với các lựa chọn tham số thích hợp, điều này sẽ đúng với hầu hết các lựa chọn của f và việc tính toán các nghịch đảo này sẽ được thực hiện dễ dàng dựa trên thuật toán Euclid mở rộng. Chúng ta sẽ kí hiệu các nghịch đảo này là f_q^{-1} và f_p^{-1}, tức là:

$$f \otimes f_q^{-1} \equiv 1 \bmod q$$

$$f \otimes f_p^{-1} \equiv 1 \bmod p$$

Sau đó, Bob tính đại lượng là một đa thức trên vành sao cho:

$$f \otimes h \equiv pg \ (mod\ q) \leftrightarrow$$

$$h \equiv f_q^{-1} \otimes pg \ (mod\ q).$$

Khóa công khai của Bob chính là đa thức h. Còn khóa bí mật của Bob là đa thức f. Đa thức g chỉ cần dùng trong quá trình tạo khóa của hệ mật nhưng cũng cần giữ bí mật bởi vì kết hợp nó với khóa công khai có thể suy ra khóa bí mật.

Quá trình mã hóa

Giả sử Alice muốn gửi một thông báo cho Bob. Ban đầu, Alice sẽ lựa chọn một thông báo từ tập bản rõ $\mathcal{L}_m$. Do NTRU là hệ mật khóa công khai xác suất, nên để mã hóa bản rõ m ta cần chọn một bản đại diện i (message representative) và chọn ngẫu nhiên một đa thức che (blinding polynomial) $r \in \mathcal{L}_r$. Sau đó, Alice sử dụng khóa công khai của Bob để tính:

$$e \equiv r \otimes h + i (mod\ q)$$

Đây chính là bản mã mà Alice sẽ gửi cho Bob (được mã hóa bằng khóa công khai h của Bob, trong đó:

$$h \equiv f_q^{-1} \otimes pg \ (mod\ q)).$$

Quá trình giải mã

Bob nhận được một thông điệp e từ Alice và muốn giải mã nó bằng khóa bí mật f. Để làm điều này một cách hiệu quả, Bob nên tính trước (và lưu trữ) f_p^{-1} như đã nói ở trên. Đầu tiên Bob tính:

$$a \equiv f \otimes e (mod\ q)$$

Trong đó, Bob lựa chọn sao cho các hệ số của $a \in [-q/2; q/2]$. Coi a là một đa thức với hệ số nguyên, Bob khôi phục bản rõ bằng cách tính:

$$f_p^{-1} \otimes a (mod\ p).$$

PHÂN TÍCH ĐỘ AN TOÀN CỦA THUẬT TOÁN MẬT MÃ NTRU

Xét bài toán 1:

Với N, q, p và đa thức h có bậc lớn nhất là $N-1$, tìm đa thức f và g trong vành $R=Z_q[x]/(x^N-1)$ với các hệ số nhỏ sao cho $f \otimes h \equiv pg \ (mod\ q)$

Nếu ta có thể giải được bài toán này thì coi như ta đã phá được hệ mật NTRU. Một nghiệm trong các nghiệm của bài toán trên sẽ là khóa bí mật tương ứng với khóa công khai h.

Xét bộ tham số cụ thể cho NTRUEncrypt $N = 251$, $q = 239$, $p = 2$, $d_f = d_g = d_r = 72$. Trong [5], các

nhà nghiên cứu NTRU đã giới thiệu phương pháp để chọn các hệ số cho các mức an toàn khác nhau. Khi đó $f \otimes h \equiv 2g \pmod q$ suy ra $f \otimes x^i \otimes h \equiv 2g \otimes x^i \pmod q$. Do đó nếu tồn tại một nghiệm của bài toán thì ta sẽ có một số nghiệm, cụ thể là các dạng xoay của f và g. Rõ ràng chúng đều có cùng độ dài. Mục đích là tìm một trong những nghiệm đó.

Xét bài toán 2: Tìm bản rõ từ bản mã bằng cách sử dụng lý thuyết lưới.

Giả sử r là đa thức che mà ta sử dụng trong việc tạo ra bản mã e và h là khóa công khai. Trong lưới $L(B)$ có véc-tơ $\begin{pmatrix} r & r \otimes h \end{pmatrix}$. Nếu gắn vào cơ sở của lưới đó véc-tơ $\begin{pmatrix} 0 & e \end{pmatrix}$ thì nhận được ma trận cơ bản sau:

$$B_e = \begin{pmatrix} I & H & \overline{0}^T \\ 0 & qI & \overline{0}^T \\ \overline{0} & e & 1 \end{pmatrix}$$

Ở đây, chúng ta làm tăng số chiều lên 1 và giữ các véc-tơ độc lập tuyến tính với nhau. Thấy rằng trong lưới mới ở trên có véc-tơ ngắn là $\begin{pmatrix} r & -i & -1 \end{pmatrix}$, trong đó i là bản đại diện được dùng, với đa thức $s \in R$:

$$\begin{pmatrix} r & s & -1 \end{pmatrix} \begin{pmatrix} I & H & \overline{0}^T \\ 0 & qI & \overline{0}^T \\ \overline{0} & e & 1 \end{pmatrix}$$

$$= \begin{pmatrix} r & r \otimes h - e + qs & -1 \end{pmatrix} = \begin{pmatrix} r & -i & -1 \end{pmatrix}$$

Tìm bản đại diện i sẽ phát hiện bản rõ. Chú ý rằng cả r và i là các dãy nhị phân và kẻ tấn công không biết.

Hằng số lưới

Với một lưới L "ngẫu nhiên", phỏng đoán của Gauss (Gaussian heuristic) chỉ ra độ dài của véc-tơ ngắn nhất của các véc-tơ khác không xấp xỉ bằng:

$$\sigma(L) \approx \sqrt{\frac{\dim(L)}{2\pi e}} \det(L)^{\frac{1}{\dim(L)}}$$

Kí hiệu véc-tơ ngắn nhất thực sự của lưới L là $\lambda(L)$. Hằng số lưới có dạng:

$$c = \frac{\lambda(L)}{\sigma(L)} \sqrt{\dim L}$$

Các kinh nghiệm thực hành chỉ ra rằng các kỹ thuật tìm véc-tơ ngắn nhất thực sự sẽ nhanh hơn khi hằng số lưới c nhỏ đi.

Với lưới L có ma trận cơ sở là: $\begin{pmatrix} \lambda I & H \\ 0 & qI \end{pmatrix}$

Và hằng số cân bằng (balancing constant) $\lambda \neq 0$, ta có:

$$\sigma(L) = \sqrt{\frac{2N}{2\pi e}} (\lambda q)^{N\frac{1}{2N}} = \sqrt{\frac{Nq\lambda}{\pi e}} \text{ và}$$

$$\lambda(L) \approx \sqrt{\|\lambda f\|^2 + \|2g\|^2}$$

Nếu ta giả thiết $\|f\| \approx \|2g\|$, khi đó:

$$c = \sqrt{\frac{2\pi e}{q\lambda} \left(\lambda^2 \|f\|^2 + \|2g\|^2 \right)}$$

là nhỏ nhất khi $\lambda = 1$. Giá trị nhỏ nhất với các tham số của NTRU là $c \approx 6{,}4$.

Đối với lưới $L(B_e)$, độ dài của véc-tơ ngắn nhất mong muốn là khoảng $\left(\sqrt{\frac{Nq}{\pi e}} \right)$ nhưng độ dài của véc tơ ngắn nhất thực tế là $\lambda\left(L\left(B_e\right)\right) \approx \sqrt{N}$. Khoảng một nửa các thành phần là 1 hoặc -1 và nửa còn lại là 0. Trong trường hợp này, hằng số c xấp xỉ là 4,2.

Rút gọn hằng số lưới

Chúng ta vận dụng lưới theo một số thông tin thêm đã có về f và g. Giả sử h là khóa công khai sinh theo các tham số của hệ mã NTRU đã đề cập ở trên. Tồn tại đa thức $F, g \in R$ với d hệ số bằng 1 và $N - d$ hệ số bằng 0 sao cho $(1 + 2F) \otimes h \equiv 2g \pmod q$. Sau đây là một số kĩ thuật để nhận được các lưới có hằng số lưới nhỏ hơn ban đầu được rút ra từ các tấn công vào hệ mã [6].

Kỹ thuật 1: Giảm hằng số lưới dựa trên độ dài của đa thức f và g.

Véc-tơ ngắn nhất của lưới càng ngắn thì càng dễ tìm được. Độ dài của véc-tơ $(f \quad 2g)$ là:

$$\sqrt{\|f\|^2 + \|2g\|^2} \approx 2\sqrt{2d}$$

Chú ý rằng ta có các véc-tơ f và $2g$ đều có d tọa độ là 2. Từ đây sẽ đi tìm một véc-tơ luôn gần f và $2g$ do chúng tương tự nhau. Xét véc-tơ có dạng $a\overline{1} = \begin{pmatrix} a & a & \cdots & a \end{pmatrix}$.Tại đây cần cực tiểu hóa khoảng cách $\begin{pmatrix} f & 2g \end{pmatrix}$ và $\begin{pmatrix} a\overline{1} & a\overline{1} \end{pmatrix}$:

$$\sqrt{\|f - a\|^2 + \|2g - a\|^2} \approx$$

$$\sqrt{2\left(d(2-a)^2 + (N-d)a^2\right)} \approx$$

$$\sqrt{2Na^2 - 8da + 8d}$$

Khoảng cách nhỏ nhất khi $a \approx \frac{2d}{N}$. Do đó, khoảng cách nhỏ nhất xấp xỉ là:

$$\sqrt{2N\left(\frac{2d}{N}\right)^2 - 8d\frac{2d}{N} + 8d} = 2\sqrt{2d - \frac{2d^2}{N}}.$$

Như vậy, sẽ có lợi hơn khi thêm véc-tơ $a\overline{1}$ vào lưới và giải bài toán tìm véc-tơ ngắn nhất đối với lưới mới đó. Điều này được gọi là tấn công nhúng (embedding attack). Trong trường hợp đang xét, $\frac{2d}{N} \approx \frac{1}{2}$, chúng ta nhân tất cả các phân tử của lưới thêm 2 để nhận được kết quả là số nguyên. Khi đó ma trận có dạng:

$$\begin{pmatrix} f & r & -1 \end{pmatrix}\begin{pmatrix} 2I & 2H \\ 0 & 2qI \\ \overline{1} & \overline{1} \end{pmatrix} = \begin{pmatrix} 2f - \overline{1} & 4g - \overline{1} \end{pmatrix}$$

Với một số $r \in Z^N$. Khi đó,

$$\sigma(L) \approx 2\sqrt{\frac{Nq}{\pi e}},\ \ \lambda(L) \approx \sqrt{2\left((N-d).1^2 + d.3^2\right)}$$

và $c \approx 5{,}4$.

Kỹ thuật 2: Giảm hằng số lưới dựa vào quá trình tạo khóa bí mật.

Xuất phát quan điểm cố gắng bẻ khóa hệ NTRU, kẻ tấn công đã biết rằng tồn tại đa thức f sao cho $f = 1 + 2F$. Ta có:

$$2F.2 - \sum_{i=0}^{N-1} x^i = 4F - \sum_{i=0}^{N-1} x^i$$

do

$$(1+2F) \otimes 2h + 2qr - \sum_{i=0}^{N-1} x^i = 4g - \sum_{i=0}^{N-1} x^i,$$

nên

$$2F \otimes 2h + 2qr - \left(\sum_{i=0}^{N-1} x^i - 2h\right) = 4g - \sum_{i=0}^{N-1} x^i.$$

Suy ra

$$F \otimes 4h + 2qr + 2h = 4g$$

và do kẻ tấn công biết h nên cũng biết các cặp của r nếu $2 \nmid q$. Giả sử $r = 2r' + r''$; $r'' \in Z_2[x]/\left(x^N - 1\right)$ đã biết bởi kẻ tấn công. Ta có:

$$\begin{pmatrix} F & r' & -1 \end{pmatrix}\begin{pmatrix} 4I & 4H \\ 0 & 2qI \\ \overline{1} & \overline{1} - 2h - 2qr'' \end{pmatrix}$$
$$= \begin{pmatrix} 4F - \overline{1} & 4g - \overline{1} \end{pmatrix}$$

suy ra:

$$\sigma(L) \approx 4\sqrt{\frac{Nq}{\pi e}}, \lambda(L) = \sqrt{2\left((N-d).1^2 + d.3^2\right)}$$

và $c \approx 2{,}7$.

KẾT LUẬN

Bài báo đã trình bày mô tả thuật toán mật mã NTRU và giới thiệu ứng dụng của thuật toán NTRU, sau đó phân tích một số cách thức, phương pháp phân tích mã để phá vỡ thuật toán NTRU. Các phân tích này được áp dụng đối với bộ tham số của hệ mật NTRU trong chuẩn IEEE P.1363.1, 2008. Việc phân tích độ an toàn của hệ mật với các bộ tham số khác trong các phiên bản của NTRU lọt vào vòng ba cuộc thi mật mã hậu lượng tử của NIST sẽ là định hướng nghiên cứu tiếp theo của nhóm tác giả.❖

TÀI LIỆU THAM KHẢO

1. A. K. Lenstra, H. W. Lenstra, and L. Loavasz: "Factoring polynomials with rational coeffcients". Math. Ann., 261:515-534, 1982.
2. Cynthia Dwork "Lattices and their application to cryptography", Lecture notes of a course given in Stanford University, 1998.
3. M.Ajtai, "The Shortest Vector Problem in L2 is NP-Hard for Randomized Reductions", Proceedings 30th Annual ACM Symposium on Theory of Computing, 1998.
4. A. May, "Cryptanalysis of NTRU" Unpublished preprint, 1999. Available at http://www.informatik.uni-frankfurt.de/~alex/ntru.ps.
5. IEEE P1363.1TM/D10, Draft Standard for Public-Key Cryptographic Techniques Based on Hard Problems over Lattices. July 2008.
6. Tommi Meskanen: "On the NTRU Crytosystem", luận văn tiến sĩ, Khoa Toán, Đại học Turku, Phần Lan, 2005.
7. NTRU Algorithm Specifications and Supporting Documentation, 2020.
8. Amritha Puliadi Premnath "Application of NTRU Cryptographic Algorithm for securing SCADA communication", University Libraries, 2013.