CLOSEST VECTOR PROBLEM FOR P-ADIC CRYPTOGRAPHY

ARJUN AGARWAL

1. Introduction to Public-Key Cryptography

In the early 1970's messages were encoded using a symmetric encryption scheme. In such schemes the same secret key was used to encode and decode messages. The need to transmit this key over potentially unsecure network makes symmetric encryption schemes vulnerable from a security point of view. The RSA (Rivest-Shamir-Adleman) algorithm, first described in 1977 by Ron Rivest, Adi Shamir and Leonard Adleman of the Massachusetts Institute of Technology. This algorithm was an assymetric encryption scheme which relies on two mathematically linked keys, a public key and a private key. For example, a receiver, Alice, chooses two prime numbers p, q and $n = pq$. The Euler-totient function $\phi(n) = (p-1)(q-1)$. Alice chooses a public key $e < \phi(n)$ such that e is relatively prime to $\phi(n)$. The private key $d = e^{-1} \pmod{\phi(n)}$. Since e is relatively prime to $\phi(n)$, its inverse d exists. Alice now shares the pair (n, e) with Bob. Bob encrypts a message M to Alice as $C = M^e \pmod{n}$. It can be shown that when Alice decrypts the message as $C^d \pmod{n}$ she recovers the original message. The security of the system relies on the difficulty in factoring n for large primes p and q. Without factoring n the private key d cannot be obtained to decrypt the message. Such functions in computer science are known as one way functions. A one way function in computer science is a function that is easy to compute on any given input, but hard to get an inverse of. In this case, given two primes p and q we can easily find $n = (p-1)(q-1)$ but its very difficult to find p, q given n.

2. Lattice Based Cryptosystems

The security of traditional assymetric cryptography schemes is based on the fact that given two primes p and q we can easily find $n = (p-1)(q-1)$ but its very difficult to find p, q given n. In other words factorization of large integers is a difficult problem. Advances in quantum computing and quantum algorithms have made the problem of finding inverses of one way functions relatively easy. For example an integer n can be factored on a quantum computer using Shor's algorithm in $log(n)$ time. Lattice-based cryptography is expected to be resistant to attacks by both classical and quantum computers. This is the reason to study cryptography schemes based on lattice constructions. The longest vector problem and closest vector problems are p-adic analogues of shortest vector problems and closest vector problems in lattices of Euclidean space. These problems can be used to create a public key cryptosystem. I will review one such scheme proposed by Deng et-al.

3. Key Definitions

This section goes over some key definitions and terms used in this report.

- **Definition 1.1** p - prime

Date: August 15, 2022.

- **Definition 1.2** $f(x)$ - an irreducible polynomial in $\mathbb{Q}_p$ of degree n
- **Definition 1.3**K is an extension field of $\mathbb{Q}_p$ of degree n.
- **Definition 1.4**$\alpha_1, \alpha_2, \ldots, \alpha_m \in K$ are $\mathbb{Q}_p$-linearly independent vectors.
- **Definition 1.5**$L(\alpha_1, \alpha_2, \ldots, \alpha_m) = \{\sum_{i=1}^{m} a_i\alpha_i | a_i \in \mathbb{Z}_p, 1 \leq i \leq m\}$ is a lattice with basis vectors $(\alpha_1, \alpha_2, \ldots, \alpha_m)$. m and n are called the rank and dimension of the lattice respectively.
- **Definition 1.6**Given a lattice L in K we recursively define $\lambda_1, \lambda_2, \ldots$ as $\lambda_1 = \max_{1\leq i\leq m} |\alpha_i|_p$ and $\lambda_{j+1} = \max\{|x|_p \; | x \in L, |x|_p < \lambda_j\}$ for $j \geq 1$. By definition $\lambda_1 > \lambda_2 > \lambda_3 > \ldots$ and $\lim_{j\to\infty} \lambda_j = 0$.
- **Definition 1.7** Given a lattice L in K. Let $t \in K - L$ be a target vector, Define s positive real numbers $\mu_1 > \mu_2 > \mu 3, \ldots \mu_s$ as $\{\mu_1, \mu_2, \ldots, \mu_s\} = \{|t - v|_p, \, v \in L\}$. Thus $\mu_{max} = \mu_1$ and $\mu_{min} = \mu_s$

4. Longest Vector Problem (LVP)

Given a lattice $L(\alpha_1, \alpha_2, \ldots, \alpha_m)$ in K. Let $\alpha = \{\sum_{i=1}^{m} a_i\alpha_i | a_i \in \mathbb{Z}_p, 1 \leq i \leq m\}$ be an element of L. Now

$$\begin{aligned} |\alpha|_p &= |\sum_{i=1}^{m} a_i\alpha_i|_p \\ &\leq \max_{1\leq i\leq m}(|a_i\alpha_i|_p) \\ &\leq \max(|\alpha_i|_p) \end{aligned}$$

This means that the length of any element in L is bounded above and as the valuation group of K is discrete the elements of L can take discrete values and have an upper bound. We recursively define $\lambda_1, \lambda_2, \ldots$ as

$$\begin{aligned} \lambda_1 &= \max_{1\leq i\leq m} |\alpha_i|_p \\ \lambda_{j+1} &= \max\{|x|_p \; | x \in L, |x|_p < \lambda_j\} \text{ for } j \geq 1 \end{aligned}$$

By definition $\lambda_1 > \lambda_2 > \lambda_3 > \ldots$ and $\lim_{j\to\infty} \lambda_j = 0$. The Longest Vector problem can now be defined as the problem of finding $v \in L$ such that $|v|_p = \lambda_2$. For a given j the algorithm takes $O(p^{m(j-1)})$ p-adic absolute value computations.

5. Closest Vector Problem (CVP)

Given a lattice $L(\alpha_1, \alpha_2, \ldots, \alpha_m)$ in K, and a target vector $t \in K - L$ such that $|t|_p \leq \lambda_1$. Define s positive real numbers

$$\mu_1 > \mu_2 > \mu 3 > \ldots \mu_s \text{ as } \{\mu_1, \mu_2, \ldots, \mu_s\} = \{|t - v|_p, \, v \in L\}$$

Basically we take the first s largest distances of elements of L from t. Thus $\mu_{max} = \mu_1$ and $\mu_{min} = \mu_s$ The closest vector problem (CVP) is to find a lattice vector $v \in L$ satisfying $|t - v|_p \leq \mu_{min}$

6. Solving LVP and CVP with Orthogonal Bases

6.1. **Orthogonal bases.** If V is a left vector space over $\mathbb{Q}_p$ of finite dimension $n > 0$ and $|.|$ be a norm on V. $\alpha_1, \alpha_2, \dots, \alpha_n$ is called an orthogonal basis of V over $\mathbb{Q}_p$ if V can be expressed as a direct sum of n 1-dimensional subspaces V_i $(1 \leq i \leq n)$ spanned by α_i such that

$$\left| \sum_{i=1}^{n} v_i \right| = \max_{1 \leq i \leq n} |v_i|, \forall v_i \in V_i, i = 1, 2, \dots, n$$

6.2. **Solving LVP with orthogonal bases.** Suppose we want to find $v_j \in L$ satisfying $|v_j|_p = \lambda_j$. Let $L(\alpha_1, \alpha_2, \dots, \alpha_m)$ be a lattice where $\alpha_1, \alpha_2, \dots, \alpha_m$ are orthogonal bases. Also without loss of generality assume $|\alpha_1|_p > |\alpha_2|_p \cdots > |\alpha_m|_p$. For any vector v in L, since the bases are orthogonal and the lengths in L are discrete the set of discrete lengths in L can be represented as $\{\log_p(|\alpha_i|_p) - k \mid i = 1, 2, \dots, m,\ k = 0, 1, 2, \dots\}$. If these valuations are considered in descending order then $\lambda_1 = |\alpha_1|_p$ and $\lambda_j = \log_p(|\alpha_i|_p) - k$ for some i. The main idea is that we need to compute m p-adic computations when the bases are orthogonal. Also the vector v_j whose length is λ_j can be written as $v_j = p^k \alpha_i$

6.3. **Solving CVP with orthogonal bases.** Let the target vector $t \in K - L$ be such that $|t|_p \leq \lambda_1$. Let V $(\supset L)$ be a k-dimensional (with $k \geq m$) $\mathbb{Q}_p$-vector subspace of the field K with $\alpha_1, \alpha_2, \dots, \alpha_m, \alpha_{m+1}, \dots, \alpha_k$ being an orthogonal basis of V. Let $t \in V$ be represented as $t = \sum_{i=1}^{k} b_i \alpha_i,\ b_i \in \mathbb{Q}_p, i = 1, 2, \dots, k$ and a lattice vector $v \in L$ be $v = \sum_{i=1}^{m} a_i \alpha_i,\ a_i \in \mathbb{Z}_p, i = 1, 2, \dots, m$ Now since the bases are orthogonal

$$\begin{aligned} |t - v|_p &= \max\{ & |b_i - a_i|_p \cdot |\alpha_i|_p \ (1 \leq i \leq m), \\ & & |b_j \alpha_j|_p \ (m + 1 \leq j \leq k)\} \end{aligned}$$

Consider two cases

- **Case 1** when $b_i \notin \mathbb{Z}_p$. In this case $|b_i - a_i|_p = |b_i|_p > 1$
- **Case 2** when $b_i \in \mathbb{Z}_p$. In this case $\{|b_i - a_i|_p\} = \{0, p^{-c} \ |c = 0, 1, 2, \dots\}$

Let $b_i \notin \mathbb{Z}_p$ for $1 \leq i \leq l$ and $b_i \in \mathbb{Z}_p$ for $l + 1 \leq i \leq m$ Then for all $v \in L$ the distance from the target vector is the maximum of three computations as given below

$$\begin{aligned} \{|t - v|_p\} &= \{\max\{ & |b_i|_p |\alpha_i|_p \ (1 \leq i \leq l); \\ & & |b_j \alpha_j|_p \ (m + 1 \leq j \leq k); \\ & & p^{-c_u} \cdot |\alpha_u|_p, c_u = 0, 1, 2 \dots (l + 1 \leq u \leq m)\}\} \end{aligned}$$

We can show that we just need $O(n)$ p-adic computations to find closets vector to t in the lattice L. The computation time is exponential if we do not start with an orthogonal basis. Key Theorem on Solving CVP with Orthogonal Bases. Let K be an extension of $\mathbb{Q}_p$ of degree n. Given a lattice $L(\alpha_1, \alpha_2, \dots, \alpha_m)$ in K, and a target vector $t \in K - L$ such that $|t|_p \leq \lambda_1$. V $(V \supset L)$ be a $k-$dimensional $\mathbb{Q}_p$-vector subspace of the field K. Let $\alpha_1, \alpha_2, \dots \alpha_m, \alpha_{m+1}, \dots, \alpha_k$ $(k \geq m)$ be an orthogonal basis for V. Let $t \in V$. There is an algorithm to find $v_i \in L$ satisfying $|t - v_i|_p = \mu_i$ for each $1 \leq i \leq s$. This algorithm takes $O(n)$ p-adic absolute value computations of elements of K.

7. Implementation of Public Key Cryptosystem

We first start with some notations

- $f(x) = x^n + f_1x^{n-1} + \cdots + f_{n-1}x + f_n \in \mathbb{Z}_p[x]$ be a non reducible polynomial such that $|f_n|_p = p^{-1}$ and $|f_i|_p < 1$ for $1 \leq i \leq n-1$
- K is a totally ramified extension field of degree $n > 1$
- θ is a root of $f(x) = 0$.
- $O_K = x \in K, |x|_p \leq 1$ is a discrete valuation ring
- $\zeta \in O_K$ such that $\mathbb{Z}_p[\zeta] = \mathbb{Z}_p[\theta] \to K = \mathbb{Q}_p(\zeta)$
- $F(x) \in \mathbb{Z}_p[x]$ be a minimum pol;ynomial of ζ over $\mathbb{Q}_p$ that is also monic and of degree n
- $j_i \in \mathbb{Z}$ are n non-negative integers such that the $j_i \pmod{n}$ for $1 \leq i \leq n$ are distinct
- $\alpha_i = \theta^{j_i}$ $(1 \leq i \leq n)$ are linearly independent over $\mathbb{Q}_p$ and form an orthogonal basis.

Chose $0 < m \leq n$ and $\delta \in \mathbb{R}^+$ Choose a matrix $A \in GL_m(\mathbb{Z}_p)$ and

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = A \times \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}$$

The matrix A is chosen such that the vectors $\beta_1, \ldots, \beta_m$ have the same length or almost the same length.

Public Key Generation Bob publishes a public key is set to $(F(x), \delta, (\beta_1, \beta_2, \ldots, \beta_m))$

Private Key Generation Bob keeps a private key set to $(A, (\alpha_1, \alpha_2, \ldots, \alpha_m))$

Encryption Let $(a_1, a_2, \ldots, a_m)$ where each $a_i \in \{0, 1, \ldots, p-1\}$ be the plaintext that is to be encoded. Alice chooses randomly $r \in K$ with $|r|_p < p^{-\delta}$ and computes a ciphertext $C = a_1\beta_1 + a_2\beta_2 + \cdots + a_m\beta_m + r \in K$ This ciphertext is sent to Bob

Decryption Bob uses the orthogonal basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. He then finds the lattice vector $v \in L$ that is closest to C. This lattice vector can be found computationally quickly as long as the bases are orthogonal. Thus

$$v = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m, \; b_i \in \mathbb{Z}_p$$

The original plaintext can be recovered as $(b_1, \ldots, b_m) \cdot A^{-1} \pmod{p}$. Since $A \in GL_m(\mathbb{Z}_p)$ its inverse exists.

Correctness The key idea is that we generate a ciphertext C from a plaintext P that is a distance $p^{-\delta}$ from the vector $A \times P \pmod{p}$. With the right constraints on δ the CVP gets us back the solution $A \times P$ that is the closest vector to the target C. P can now be retrieved from $A \times P$ using $A^{-1} \times A \times P$. As long as $j_i \leq \delta n$ for all $1 \leq i \leq m$ we can guarantee the uniqueness of the closest vector solution and therefore recover the original plaintext.

8. Conclusion

Traditional public key cryptography schemes are based on generating keys from factorization of large integers. With advances in quantum computing and algorithms, such cryptography systems might not be very secure as integer factorization is a linear time problem with quantum algorithms. Encoding using p-adic space offers an alternate to such traditional schemes. The Longest Vector Problem (LVP) and Shortest Vector Problem (CVP) in p-adic lattices are p-adic analogues of the shortest vector problem and Closest Vector Problem in lattices in Euclidean space. A public key cryptography scheme based on Closest Vector Problem (CVP) in p-adic lattices is studied in this report. Several open topics remain to be solved in this study. For example, it is not yet known if LVP and CVP are NP-hard problems. However, this is a promising approach, that utilizes the properties of p-adic spaces to implement a public key cryptography system.

9. References

[1] R.L. Rivest, A. Shamir, and L. Adleman, A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, *https://people.csail.mit.edu/rivest/Rsapaper.pdf*

[2] Yingpu Deng1, Lixia Luo1, Yanbin Pan1, Zhaonan Wang1 and Guanju Xiao2. Cryptosystems and Signature Schemes from p-adic Lattices. Cryptology ePrint Archive, Paper 2021/522, 2021.

Email address: arjunagarwal010@gmail.com