Knapsack Public-Key Encryption

OAlienO

January 9, 2019

Table of Contents

1 Knapsack Public-Key Encryption
- Introduction
- Subset Sum Problem
- Merkle-Hellman Knapsack Encryption

Knapsack Public-key Encryption 簡介

- 與 RSA 同年於 1978 年被提出
- 為最早被提出的其中一個 public-key cryptosystem
- 其安全性建立在 subset sum problem

Subset Sum Problem

- Given a set $\{a_1, a_2, \cdots, a_n\}$ called knapsack set
- And a positive integer s
- Determine whether $\exists\ x_i \in \{0, 1\}, 1 \leq i \leq n : \sum_{i=1}^{n} a_i x_i = s$
- 翻譯: 是否存在一個子集合的合等於 s

Naive Approach

- 暴力嘗試所有可能子集合
- 時間複雜度: $O(2^n)$

Meet In The Middle Approach

- 將集合切一半
- 前半段的所有子集合建表
- 後半段所有可能子集合去搜尋 $s - \sum_{i=\frac{n}{2}}^{n} a_i x_i$ 是否在表中
- 時間複雜度: $O(2^{\frac{n}{2}})$

Dynamic Programming Approach

- dp[i][j] : i 是集合前 i 個數, j 是目標的總和
- dp[i][j] = dp[i-1][j] | dp[i-1][j - a[i]]
- 時間複雜度: $O(ns)$

Solving Subset Sum Poblems of Low Density

- The density of a knapsack set $S = \{a_0, a_1, \cdots, a_n\}$ is

$$d = \frac{n}{max\{log(a_i)|1 \leq i \leq n\}}$$

- If the knapsack set has **low density**, we can use L^3-lattice basis reduction algorithm to solve the corresponding subset sum problem

L^3 Algorithm Approach

- Let $m = \left\lceil \frac{1}{2}\sqrt{n} \right\rceil$
- Form an $(n+1)$-dimensional lattice L with basis consisting of the rows of the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & ma_1 \\ 0 & 1 & 0 & \cdots & 0 & ma_2 \\ 0 & 0 & 1 & \cdots & 0 & ma_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & ma_n \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} & ms \end{pmatrix}$$

L^3 Algorithm Approach (cont.)

- Find a reduced basis B of L (using L^3 Algorithm)
- For each row $y = (y_1, \cdots, y_{n+1})$ in B, do the following to $\pm y$
  - If $y_{n+1} = 0$ and $y_i \in \{-\frac{1}{2}, \frac{1}{2}\}$ for all $i = 1, 2, \cdots, n$
  - Set $x_i = y_i + \frac{1}{2}$
  - Test whether $\sum_{i=1}^{n} a_i x_i = s$

L^3 Algorithm Approach Explanation

- If $(x_1, x_2, \cdots, x_n)$ is a solution to the subset sum problem
- Consider $y = \sum_{i=1}^{n} x_i b_i - b_{n+1}$ in L
- $y_i \in \{-\frac{1}{2}, \frac{1}{2}\}$ for $1 \leq i \leq n$
- $y_{n+1} = 0$
- y is a vector of short length in L
- With high probability y will be in the reduced basis B
- Coster [1] show that id density < 0.9408, it can almost always be solved

Lattice

inner product

- Let $x = (x_1, x_2, \cdots, x_n)$ and $y = (y_1, y_2, \cdots, y_n)$ be two vectors in $\mathbb{R}^n$
- $< x, y >= x_1y_1 + x_2y_2 + \cdots + x_ny_n$

length

- $\|y\| = \sqrt{< y, y >} = \sqrt{y_1^2 + y_2^2 + \cdots + y_n^2}$

Lattice

Lattice

- Given n linearly independent vectors $\mathbf{b_1}, \mathbf{b_2}, \cdots, \mathbf{b_n} \in \mathbb{R}^m$, the lattice generated by them is defined as

$$L(\mathbf{b_1}, \cdots, \mathbf{b_n}) \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{n} x_i \mathbf{b_i} \mid x_i \in \mathbb{Z} \right\}$$

- We call $\{\mathbf{b_1}, \cdots, \mathbf{b_n}\}$ a basis of lattice L

Lattice

Some Definition

- $\mu_{i,j} = \frac{<b_i, b_j^*>}{<b_j^*, b_j^*>}, 1 \leq j < i \leq n$
- $b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{i,j} b_j^*, 1 \leq i \leq n$

Lattice

Reduced Basis

- $|\mu_{i,j}| \leq \frac{1}{2}$, for $1 \leq j < i \leq n$
- $\|b_i^*\|^2 \geq (\frac{3}{4} - \mu_{i,i-1}^2)\|b_{i-1}^*\|^2$, for $1 < i \leq n$

L^3 Algorithm

- L^3 algorithm is a polynomial-time algorithm for finding a reduced basis

Merkle-Hellman Knapsack Encryption

superincreasing sequence

- A sequence $(b_1, b_2, \cdots, b_n)$
- Satisfy $b_i > \sum_{j=1}^{i-1} b_j$ for every $2 \leq i \leq n$

Merkle-Hellman Knapsack Encryption

- Solving superincreasing subset sum problem is easy
- 倒著從 n 掃回 1，只要 $s > b_i$ 就代表 $x_i = 1$，並把 s -= b_i
- Merkle-Hellman Knapsack Encryption 就是將 superincreasing sequence 藏起來，讓有私鑰的人才可以解 superincreasing subset sum problem

Merkle-Hellman Knapsack Encryption

8.36 Algorithm Key generation for basic Merkle-Hellman knapsack encryption

SUMMARY: each entity creates a public key and a corresponding private key.

1. An integer n is fixed as a common system parameter.
2. Each entity A should perform steps 3 – 7.
3. Choose a superincreasing sequence $(b_1, b_2, \ldots, b_n)$ and modulus M such that $M > b_1 + b_2 + \cdots + b_n$.
4. Select a random integer W, $1 \leq W \leq M - 1$, such that $\gcd(W, M) = 1$.
5. Select a random permutation π of the integers $\{1, 2, \ldots, n\}$.
6. Compute $a_i = W b_{\pi(i)} \bmod M$ for $i = 1, 2, \ldots, n$.
7. A's public key is $(a_1, a_2, \ldots, a_n)$; A's private key is $(\pi, M, W, (b_1, b_2, \ldots, b_n))$.

Merkle-Hellman Knapsack Encryption

8.37 Algorithm Basic Merkle-Hellman knapsack public-key encryption

SUMMARY: B encrypts a message m for A, which A decrypts.

1. *Encryption*. B should do the following:
   (a) Obtain A's authentic public key $(a_1, a_2, \ldots, a_n)$.
   (b) Represent the message m as a binary string of length n, $m = m_1 m_2 \cdots m_n$.
   (c) Compute the integer $c = m_1 a_1 + m_2 a_2 + \cdots + m_n a_n$.
   (d) Send the ciphertext c to A.
2. *Decryption*. To recover plaintext m from c, A should do the following:
   (a) Compute $d = W^{-1} c \bmod M$.
   (b) By solving a superincreasing subset sum problem (Algorithm 8.35), find integers $r_1, r_2, \ldots, r_n$, $r_i \in \{0, 1\}$, such that $d = r_1 b_1 + r_2 b_2 + \cdots + r_n b_n$.
   (c) The message bits are $m_i = r_{\pi(i)}$, $i = 1, 2, \ldots, n$.

Merkle-Hellman Knapsack Encryption

- Merkle-Hellman Knapsack Encryption is insecure
- It can be solve by L^3 algorithm approach with high probability
- Chor-Rivest knapsack encryption is designed to resist the attack
- But Chor-Rivest knapsack encryption also has been broken in 1995 [2]

M. J. Coster, A. Joux, B. A. LaMacchia, A. M. Odlyzko, C.-P. Schnorr, and J. Stern, "Improved low-density subset sum algorithms," *computational complexity*, vol. 2, no. 2, pp. 111–128, 1992.

C.-P. Schnorr and H. H. Hörner, "Attacking the chor-rivest cryptosystem by improved lattice reduction," in *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 1–12, Springer, 1995.