密码学报 ISSN 2095-7025 CN 10-1195/TN
Journal of Cryptologic Research, 2017, 4(5): 498–505

E-mail: jcr@cacrnet.org.cn
http://www.jcr.cacrnet.org.cn
Tel/Fax: +86-10-82789618

后量子密码专栏

基于联立丢番图逼近的子集和问题启发式求解算法*

王保仓[1,2], 卢 珂[1]

1. 西安电子科技大学 综合业务理论与关键技术国家重点实验室, 西安 710071
2. 桂林电子科技大学 认知无线电与信息处理省部共建教育部重点实验室, 桂林 541004

通讯作者: 王保仓, E-mail: bcwang79@aliyun.com

摘 要: 子集和问题是计算机科学中的一个重要问题, 也被应用于公钥密码和伪随机函数的设计. 学界已提出多个求解一般子集和问题的通用求解算法及求解特定子集和问题的特殊求解算法. 本文通过建立子集和问题和联立丢番图逼近问题之间的联系, 提出一种新的子集和问题启发式求解算法. 该算法由给定的子集和问题构造联立丢番图逼近问题, 使用格归约算法寻找该联立丢番图逼近问题的解, 由此构造与原始子集和问题线性无关的新的子集和问题, 从而达到降低原始子集和问题维数的目的; 最后, 通过 $n-1$ 个联立丢番图逼近问题的解来构造 $n-1$ 个线性无关的子集和问题, 并通过求解一个由 n 个变量和 n 个线性方程构成的方程组来求解原始子集和问题. 基于联立丢番图逼近的子集和问题启发式求解算法为子集和问题研究提供了新的思路.

关键词: 子集和问题; 联立丢番图逼近; 启发式算法; 公钥密码; 格归约

中图分类号: TP309.7 **文献标识码:** A DOI: 10.13868/j.cnki.jcr.000201

中文引用格式: 王保仓, 卢珂. 基于联立丢番图逼近的子集和问题启发式求解算法[J]. 密码学报, 2017, 4(5): 498–505.

英文引用格式: WANG B C, LU K. Heuristic algorithm for the subset sum problem based on simultaneous diophantine approximation[J]. Journal of Cryptologic Research, 2017, 4(5): 498–505.

Heuristic Algorithm for the Subset Sum Problem based on Simultaneous Diophantine Approximation

WANG Bao-Cang[1,2], LU Ke[1]

1. State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China
2. Key Laboratory of Cognitive Radio and Information Processing, Ministry of Education (Guilin University of Electronic Technology), Guilin 541004, China

Corresponding author: WANG Bao-Cang, E-mail: bcwang79@aliyun.com

Abstract: The subset sum problem is one of the most significant problems in computer science, and the problem has been used in designing public key cryptographic schemes and pseudorandom functions due to the NP-completeness nature. Therefore, the study of the subset sum problem is

* 基金项目: 国家重点研发计划项目 (2017YFB0802000); 国家自然科学基金项目 (61572390); 宁波市自然科学基金项目 (201601HJ-B01382); 桂林电子科技大学认知无线电与信息处理省部共建教育部重点实验室开放基金 (CRKL160202)
收稿日期: 2017-07-12 定稿日期: 2017-10-10

of important significance in both computer science and cryptography. Under the basic $P \neq NP$ hypothesis, there must exist no polynomial-time algorithms to solve the subset sum problem. Some general-purpose algorithms for solving generic subset sum problems and special-purpose algorithms for solving special subset sum problems have been proposed in the literature. This paper proposes a novel heuristic algorithm for solving the subset sum problem by establishing a connection between the subset sum problem and the simultaneous Diophantine approximation problem. The basic idea of the proposed algorithm is to firstly construct simultaneous Diophantine problems from the given subset sum problem, then find the solutions to the simultaneous Diophantine problems via lattice reduction algorithms, construct new subset sum problems independent of the original subset sum problem and from the solutions to the simultaneous Diophantine problems, ultimately reduce the dimension of the original subset sum problem. Finally, the binary solution to the original subset sum problem is determined by solving an n-variable system of n linear equations, which is derived from $n-1$ linearly independent subset sum problems constructed by solving simultaneous Diophantine approximation problems. The significance of the new approach in this paper provides new insights into the subset sum problem.

Key words: subset sum problem; simultaneous Diophantine approximation; heuristic algorithm; public key cryptography; lattice reduction

1 引言

子集和问题是计算机科学中的一类重要的问题, 在 $P \neq NP$ 假设下不存在求解该问题的多项式时间算法. 在密码学中, 子集和问题也被称作背包问题. 子集和问题因其 NP 完全性被认为是设计后量子公钥密码的重要数学资源 [1,2]. 背包公钥密码描述简洁、加解密速度快, 一度被认为是非常有发展前途的密码算法. 子集和问题的困难假设是密码学中的一类重要的困难假设, 比如, 2009 年 Gentry 首次提出的全同态密码 [3] 以及后来的基于近似最大公因数问题的全同态密码 [4] 的安全性就部分依赖于稀疏子集和问题的困难性假设. 因此, 对子集和问题求解算法的研究对深入理解子集和问题相关的密码算法的安全性至关重要.

研究人员一直在试图寻找子集和问题的通用求解算法, 其最直接的求解算法是穷举搜索算法. 对于 n 维子集和问题, 其计算复杂度为 $O(2^n)$. 1974 年, Horowitz 和 Sahmir [5] 提出了一种新的求解算法, 所需时间复杂度为 $O\left(n2^{n/2}\right)$, 空间复杂度为 $O\left(2^{n/2}\right)$. 1979 年, Schroeppel 和 Shamir [6,7] 提出一种基于生日悖论技术的子集和求解算法, 所需时间复杂度为 $O\left(n2^{n/2}\right)$, 空间复杂度为 $O\left(n2^{n/4}\right)$. 相对 Horowitz 和 Sahmir 提出的算法, 该算法极大降低了空间复杂度, 并在此后的 31 年里都是最好的子集和问题通用求解算法. 在 2010 年的 Eurocrypt 会议上, Howgrave-Graham 和 Joux [8] 提出了一种新的子集和问题的求解算法, 该算法在 Schroeppel 和 Shamir 算法的基础上进行改进, 将时间复杂度降低至 $\widetilde{O}\left(2^{0.385n}\right)$ 或 $\widetilde{O}\left(2^{0.3113n}\right)$, 显著提升了求解效率. 2011 年, Becker 等人 [9] 对 Howgrave-Graham 和 Antoine Joux 的算法进一步改进, 将时间复杂度降低至 $\widetilde{O}\left(2^{0.291n}\right)$, 同时为了降低算法的空间复杂度, 他们提出了一种基于循环查找的算法, 该算法时间复杂度降低至 $\widetilde{O}\left(2^{0.75n}\right)$, 空间复杂度为常数. 随后, 在确保空间复杂度依然为常数的情况下, Becker 等人利用 Howgrave-Graham 和 Antoine Joux 的算法, 将时间复杂度进一步降低为 $\widetilde{O}\left(2^{0.72n}\right)$.

针对密码学中所使用的子集和问题, 研究人员设计了一些特殊的求解算法. 在基于子集和问题的公钥密码中, 由于解密和陷门嵌入的需要, 底层的子集和问题往往具有一些特殊的结构. 这些特殊结构使得公钥密码中的子集和问题不是随机生成的, 这些问题就可能存在有效的求解算法. 比如, 公钥密码中的子集和问题的子集和密度往往不能太高, 普遍要低于 1. 因此, 就存在针对低密度子集和问题的求解算法 [10–16]. Coster 等人就证明了, 当子集和密度小于 0.9408 时, 子集和问题可以归约到格上的最短向量问题 [10–12]. 因此, 人们就希望通过降低子集和问题解的汉明重量的方式来提高公钥密码中子集和问题的密度 [17–22], 这一类公钥密码被称为低重量背包密码 [23–25]. 虽然低重量背包密码体制能够在一定程度上抵

抗低密度子集和攻击, 但低重量背包公钥密码仍然不能够阻挡把底层的背包 (子集和) 问题归约到相应的格上最近向量或最短向量问题 [23–26].

本文提出一种基于联立丢番图逼近的子集和问题启发式算法, 为子集和问题求解提供一种新的研究思路. 本文算法的基本思想是建立子集和问题和联立丢番图逼近问题之间的联系, 通过将原始子集和问题转换为联立丢番图逼近问题来构造一个与原始子集和问题线性无关的新的子集和问题, 从而达到降低原始子集和问题维数的目的. 具体来讲, 首先我们对原始子集和问题做一系列转换, 构造 $n-1$ 个线性无关的子集和问题. 这样, 我们就得到了一个由 n 个变量构成的 n 维线性方程组. 最后, 通过使用高斯消元法求解该线性方程组, 就可以获得了原始子集和问题的解.

本文的章节安排如下. 第 2 节给出子集和问题、联立丢番图逼近问题以及格的基本概念; 第 3 节给出求解子集和问题的一个新型算法; 第 4 节进行总结和讨论.

2 预备知识

本节中我们首先给出子集和问题的详细定义, 然后介绍关于格的基本概念及相关知识, 最后给出联立丢番图逼近问题的定义以及相关结论.

2.1 子集和问题

定义 1 (子集和问题) 给定 n 个正整数 $a_1,\cdots,a_n$ 和正整数 s, 要求确定 n 元列向量 $\mathbf{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 的值, $x_i\in\{0,1\}$, $1\leqslant i\leqslant n$, 使得下式成立:

$$\sum_{i=1}^{n}a_ix_i=s$$

2.2 格及相关知识

格是一类定义在实数空间上的离散加法群, 被广泛用于密码分析和密码设计.

定义 2 (格) 设 $\mathbf{V}=(\mathbf{v}_1,\cdots,\mathbf{v}_m)^{\mathrm{T}}$ 是 m 维空间中的一组线性无关的行向量, 则该组向量的整系数线性组合的全体就定义为格, 即,

$$\mathbf{\Lambda}=L(\mathbf{V})=\left\{\mathbf{zV}=\sum_{i=1}^{m}z_i\mathbf{v}_i\middle|\,\mathbf{z}=(z_1,\cdots,z_m)\in\mathbb{Z}^m\right\}$$

这里, $\mathbf{V}=(\mathbf{v}_1,\cdots,\mathbf{v}_m)^{\mathrm{T}}$ 被称为格的一组基. 格中的一个不变量是格的行列式, 记为

$$\det(L(\mathbf{V}))=\sqrt{\det(\mathbf{VV}^{\mathrm{T}})}$$

对于任意 n 维实向量 $\mathbf{a}=(a_1,\cdots,a_n)$, 其 1-和 2-范数定义为,

$$|\mathbf{a}|_1=\sum_{i=1}^{n}|a_i|,\quad |\mathbf{a}|_2=\sqrt{\sum_{i=1}^{n}a_i^2}$$

关于 1-和 2-范数, 容易证明下面的结论.

引理 1 对于任意 n 维实向量 $\mathbf{a}=(a_1,\cdots,a_n)$, 都有 $|\mathbf{a}|_1\leqslant\sqrt{n}\,|\mathbf{a}|_2$.

证明: 只需注意到下式成立即可,

$$|\mathbf{a}|_1^2=\left(\sum_{i=1}^{n}|a_i|\right)^2=\sum_{i=1}^{n}a_i^2+2\sum_{i\neq j}|a_ia_j|\leqslant n\sum_{i=1}^{n}a_i^2=\left(\sqrt{n}\,|\mathbf{a}|_2\right)^2$$

□

定义 3 (最短向量问题) 给定格 $\mathbf{\Lambda}$ 的一组基 $\mathbf{V}$, 求该格中的一个非零向量 $\mathbf{v}$, 使得对于格中的任意非零向量 $\mathbf{u}$ 都有 $|\mathbf{v}|_2 \leqslant |\mathbf{u}|_2$.

对于格中的最短向量的长度, 我们不加证明地引入下面的闵可夫斯基定理.

定理 1 [27] 格 $L(\mathbf{V})$ 中最短向量的长度 $\mathbf{\Lambda}$ 满足 $\mathbf{\Lambda} \leqslant \sqrt{m}\det(L(\mathbf{V}))^{1/m}$.

2.3 联立丢番图逼近

联立丢番图逼近问题是丢番图逼近理论中的一个基本问题, 在密码设计 [28] 和分析 [29] 等领域中有广泛的应用.

定义 4 (联立丢番图逼近) 给定 $n+1$ 个实数 $\alpha_1, \alpha_2, \cdots, \alpha_n, \varepsilon > 0$ 以及一个整数 $\mathcal{Q} > 0$, 寻找整数 $p_1, p_2, \cdots, p_n, q$ 满足 $0 < q < \mathcal{Q}$, 使得 $|\alpha_i - p_i/q| \leqslant \varepsilon/\mathcal{Q}$.

该问题等价为寻找一组具有相同分母的较小分数 $p_1/q, \cdots, p_n/q$, 并且 $p_1/q, \cdots, p_n/q$ 分别逼近 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 当 $\mathcal{Q} \geqslant \varepsilon^{-n}$ 时, 该问题有解 [2]. 虽然尚未发现该问题的有效求解算法, 但通过格归约算法可以找到一个近似解. 稍后我们将阐述如何利用格归约算法来求解联立丢番图逼近问题的近似解, 读者也可参阅文献 [2].

3 基于联立丢番图逼近的求解算法

我们提出一种基于联立丢番图逼近的子集和问题启发式求解算法. 在本节中我们首先介绍算法的基本思想, 随后给出利用联立丢番图逼近生成一个新的子集和问题的过程以及详细算法, 最后给出子集和问题具体的求解算法.

3.1 基本思想

给定一个子集和问题 $\sum_{i=1}^{n} a_i x_i = s$, 已知 n 个正整数 $a_1, \cdots, a_n$ 和正整数 s, 求解使等式成立的 n 元列向量 $\mathbf{x} = (x_1, \cdots, x_n)^{\mathrm{T}}$ 的值, $x_i \in \{0, 1\}$, $1 \leqslant i \leqslant n$. 我们可以利用联立丢番图逼近算法生成 $n-1$ 个新的子集和问题, $\sum_{i=1}^{n} a_i^{(j)} x_i = s^{(j)}$, $j = 1, \cdots, n-1$, 要求 n 个行向量 $(a_1, \cdots, a_n), \left(a_1^{(1)}, \cdots, a_n^{(1)}\right), \cdots, \left(a_1^{(n-1)}, \cdots, a_n^{(n-1)}\right)$ 构成的向量组线性无关. 令矩阵

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_1^{(1)} & a_2^{(1)} & \cdots & a_n^{(1)} \\ \vdots & \vdots & & \vdots \\ a_1^{(n-1)} & a_2^{(n-1)} & \cdots & a_n^{(n-1)} \end{pmatrix}$$

并令列向量

$$\mathbf{s} = \left(s, s^{(1)} \cdots, s^{(n-1)}\right)^{\mathrm{T}}$$

由于 $\mathbf{Ax} = \mathbf{s}$ 有唯一解, 因此用高斯消元法求解该线性方程组, 即可求出子集和问题 $\sum_{i=1}^{n} a_i x_i = s$ 的解 $\mathbf{x} = (x_1, \cdots, x_n)^{\mathrm{T}}$.

3.2 生成新的子集和问题

给定一个子集和问题 $\sum_{i=1}^{n} a_i x_i = s$ 已知 n 个正整数 $a_1, \cdots, a_n$ 和正整数 s, 我们在整数环 $\mathbb{Z}$ 上选取一个整数 M(暂时要求 $(n+1)^{n+1} < M < 2(n+1)^{n+1}$, 稍后将给出证明), 希望能够在模 M 的剩余类环 $\mathbb{Z}_M$ 上找到一个整数 w 使得所有的 $b_i \equiv wa_i \mod M$ 都比较小 (所有的 b_i 均小于 M/n). 所以 $\sum_{i=1}^{n} b_i x_i \equiv \sum_{i=1}^{n} wa_i x_i \equiv ws \mod M$, 令 $c \equiv ws \mod M$, $\sum_{i=1}^{n} b_i x_i \equiv c \mod M$. 又由于 $\sum_{i=1}^{n} b_i x_i \leq \sum_{i=1}^{n} b_i < \sum_{i=1}^{n} M/n = M$, 即得一个新的子集和问题 $\sum_{i=1}^{n} b_i x_i = c$.

注意到 $b_i \equiv wa_i \mod M$, 一定存在整数 k_i 使得 $wa_i - k_i M = b_i$. 在前面的讨论中, 我们要求所有的 $b_i \equiv wa_i \mod M$ 都比较小, 即所有的 $wa_i - k_i M = b_i$ 都比较小; 等式 $wa_i - k_i M = b_i$ 两边同时除以 wM, 得到等式 $|a_i/M - k_i/w| = b_i/wM$. 又由于 $b_i < M/n$, $w < M$, 可得 $|a_i/M - k_i/w| = b_i/wM <$

$1/nw$, 因此 $k_1/w,\cdots,k_n/w$ 就是一组具有相同分母 w 的较小分数, 且能够分别逼近 $a_1/M,\cdots,a_n/M$. 这样一来, 寻找整数 w 的问题就转换成了联立丢番图逼近求解问题.

构造如下矩阵

$$\mathbf{V}=\begin{pmatrix}-M & 0 & \cdots & 0 & 0\\ 0 & -M & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & -M & 0\\ a_1 & a_2 & \cdots & a_n & 1\end{pmatrix}=\begin{pmatrix}\mathbf{v}_1\\ \mathbf{v}_2\\ \vdots\\ \mathbf{v}_n\\ \mathbf{v}_{n+1}\end{pmatrix} \tag{1}$$

矩阵 $\mathbf{V}=(\mathbf{v}_1,\cdots,\mathbf{v}_{n+1})^{\mathrm{T}}$ 中行向量的整系数线性组合构成了一个格 $\mathbf{\Lambda}$, 可以利用格归约算法在多项式时间内寻找到格 $\mathbf{\Lambda}$ 的一组归约基. 令该组归约基中的最短向量为 $\mathbf{v}$, 并令 $\mathbf{v}$ 由基 $(\mathbf{v}_1,\cdots,\mathbf{v}_{n+1})^{\mathrm{T}}$ 的整系数线性组合表示的系数为 $k_1,\cdots,k_n,w$, $\mathbf{v}=\sum_{i=1}^{n}k_i\mathbf{v}_i+w\mathbf{v}_{n+1}=(wa_1-k_1M,\cdots,wa_n-k_nM,w)$. 由于向量 $\mathbf{v}$ 是格 $\mathbf{\Lambda}$ 中的一个短向量, 可以断定所有的分量 $w_1a_1-k_1M,\cdots,w_na_n-k_nM,w$ 都比较小. 令 $wa_i-k_iM=b_i$, 这样一来, 所有的 b_i 都比较小. 最后, 我们就把联立丢番图逼近问题的求解转化到了格上最短向量问题的求解.

对于一个给定的子集和问题 $\sum_{i=1}^{n}a_ix_i=s$, 新的子集和问题生成算法 NewSSPGen 见算法 1.

算法 1 New subset sum problem generation algorithm

Input: $a_1,\cdots,a_n,s$

Output: $b_1,\cdots,b_n,c\equiv ws \mod M$

1 在区间 $((n+1)^{n+1},2(n+1)^{n+1})$ 均匀且随机选取整数 M

2 对由矩阵 (1) 的行向量生成的格 $\mathbf{\Lambda}$ 实施格归约算法

3 求出一组归约基, 令该组归约基中的最短向量为 $\mathbf{v}$, 并令

$$\mathbf{v}=\sum_{i=1}^{n}k_i\mathbf{v}_i+w\mathbf{v}_{n+1}=(wa_1-k_1M,\cdots,wa_n-k_nM,w)$$

```
4  for i = 1, ..., n do
5  |   令 b_i = w a_i - k_i M;
6  |   if b_i < M/n then
7  |   |   i++
8  |   end
9  |   else
10 |   |   goto 2
11 |   end
12 end
```

在新的子集和问题生成算法 NewSSPGen 中, 我们有如下结论.

定理 2 设格 $L(\mathbf{V})$ 中的最短向量为

$$\mathbf{v}=(v_1,\cdots,v_n,v_{n+1})=\sum_{i=1}^{n}k_i\mathbf{v}_i+w\mathbf{v}_{n+1}=(wa_1-k_1M,\cdots,wa_n-k_nM,w)$$

若 $M>(n+1)^{n+1}$, 则必有, $0\leqslant\sum_{i=1}^{n}v_ix_i<M$.

证明: 设 $(x_1,\cdots,x_n)$ 是子集和 $\sum_{i=1}^{n}a_ix_i=s$ 的解. 我们不妨设 $\sum_{i=1}^{n}v_ix_i\geq 0$ (因为若 $\sum_{i=1}^{n}v_ix_i\leqslant 0$, 则 $\sum_{i=1}^{n}(-v_i)x_i\geq 0$). 又 $0\leqslant\sum_{i=1}^{n}v_ix_i\leqslant\sum_{i=1}^{n}|v_i|x_i$, 由于 $x_i\in\{0,1\}$, 所以 $0\leqslant\sum_{i=1}^{n}v_ix_i\leqslant\sum_{i=1}^{n}|v_i|x_i\leqslant\sum_{i=1}^{n}|v_i|$. 又因为 $\sum_{i=1}^{n}|v_i|\leqslant\sum_{i=1}^{n+1}|v_i|=|\mathbf{v}|_1$, 所以由引理 1 我们可得 $|\mathbf{v}|_1\leqslant\sqrt{n+1}|\mathbf{v}|_2$. 又因为 $\mathbf{v}$ 是格上的最短向量, 由闵可夫斯基定理我们可知 $|\mathbf{v}|_2\leqslant\sqrt{n+1}\det(L(\mathbf{V}))^{1/n+1}$, 而 $\det(L(\mathbf{V}))$ 等于矩阵 $\mathbf{V}$ 的行列式的绝对值, 即 $\det(L(\mathbf{V}))=M^n$, 故 $|\mathbf{v}|_2\leqslant\sqrt{n+1}(M^n)^{1/n+1}$, 所以 $\sum_{i=1}^{n}v_ix_i\leqslant(n+1)M^{n/n+1}$. 因此当 $M>(n+1)^{n+1}$ 时,

$\sum_{i=1}^{n} v_i x_i \leqslant (n+1)\, M^{n/n+1} < M$ 成立. □

因为 $v_i = wa_i - k_i M$, 所以 $\sum_{i=1}^{n} v_i x_i \equiv ws \mod M$. 令 $s' \equiv ws \mod M$, 由定理 2 我们可知当 $M > (n+1)^{n+1}$ 时, $0 \leqslant \sum_{i=1}^{n} v_i x_i < M$, 所以 $s' = \sum_{i=1}^{n} v_i x_i$, 因此我们构造出一个新的子集和问题. 因此, 在算法的参数设置中, 我们要求 $(n+1)^{n+1} < M$, 为了限定随机选取的 M 的上界, 我们限定 $M < 2(n+1)^{n+1}$.

3.3 子集和问题具体求解算法

本小节我们将给出子集和问题求解算法的详细描述. 对于一个给定的待求解子集和问题 $\sum_{i=1}^{n} a_i x_i = s$, 该算法的输入是 $a_1, \cdots, a_n, s$, 输出是该子集和问题的解 $\mathbf{x} = (x_1, \cdots, x_n)^{\mathrm{T}}$. 具体算法流程见算法 2.

算法 2 Subset sum algorithm

Input: $a_1, \cdots, a_n, s$

Output: $\mathbf{x} = (x_1, \cdots, x_n)^{\mathrm{T}}$

初始化 $\mathbf{a}^{(0)} = \left(a_1^{(0)}, \cdots, a_n^{(0)}\right) = (a_1, \cdots, a_n)$, $s^{(0)} = s$

for $i = 1, \cdots, n-1$ **do**

  运行

$$\text{NewSSPGen}\left(\mathbf{a}^{(0)}, s^{(0)}\right) = \left(\mathbf{a}^{(i)}, s^{(i)}\right) = \left(\left(a_1^{(i)}, \cdots, a_n^{(i)}\right), s^{(i)}\right)$$

  if $i+1$ 个行向量构成的向量组 $\mathbf{a}^{(0)}, \cdots, \mathbf{a}^{(i)}$ 线性无关 **then**

    $i{+}{+}$

  end

  else

    goto 4

  end

end

求解线性方程组 $\mathbf{A}\mathbf{x} = \mathbf{s}$ 得解 $\mathbf{x} = (x_1, \cdots, x_n)^{\mathrm{T}}$

3.4 讨论

本小节对上述给出的子集和问题的启发式求解算法做如下说明.

评述 1: 关于算法的计算复杂度. 我们并未给出该启发式求解算法的计算复杂度分析结果. 事实上, 也很难给出该算法具体的计算复杂度. 原因大致可以分为两个方面. 其一, 该算法的具体计算复杂度依赖于所使用的格归约算法 (比如 LLL 算法 [30], BKZ 算法 [31], 筛法算法 [32]) 的计算复杂度. 其二, 格归约算法的解和联立丢番图逼近问题的解之间可能存在缝隙. 因此, 使用格归约算法的计算复杂度来评估联立丢番图逼近问题的求解计算复杂度会不准确.

评述 2: 关于求解的具体效果. 本文只是在一定意义下把子集和问题的求解归约到了联立丢番图逼近问题的求解, 而在求解过程中最核心的算法是格归约算法. 在子集和公钥密码的密码分析历史中, 很大一部分工作就是把子集和公钥密码的安全性归约到格上的最短向量问题 [10–16, 23–26]. 须注意, 格上的最短向量问题本身也是一个困难问题, 只有当格的维数较小时 (比如说 < 300), 格归约算法才能被认为可以寻找到格上的一个短向量, 当格的维数较大 (比如说 > 500) 时, 格归约算法不能够求出格上的一个短向量. 因此, 该启发式求解算法的具体求解效果依赖于待求解的子集和问题的维数.

评述 3: 关于模数 M 的选取. 在本文中我们希望在模数 M 的的非负最小完全剩余系中找到满足条件的 w. 而在实际操作中, 我们可以选择模数 M 的绝对最小完全剩余系, 这样可能会提高我们得到满足条件的 w 的概率, 从而提高生成一个新的子集和问题的概率, 同时也会提高求解子集和问题的计算效率.

评述 4: 关于联立丢番图逼近问题. 本文提出一种基于联立丢番图逼近的子集和求解算法, 而联立丢番图逼近问题的求解在本文中起到一个连接过渡的作用. 通过联立丢番图比较难问题求解的过渡, 我们将求解子集和问题转换到求格上最短向量的问题. 求解出格上的最短向量来构造一个新的子集和问题. 因此, 本文算法的核心是格归约算法.

评述 5: 关于新子集和向量的线性相关性. 我们要求向量组 $\mathbf{a}^{(0)}, \cdots, \mathbf{a}^{(n)}$ 线性无关, 为了说明这一点, 我们看一个小例子. 设 $\mathbf{a} = (511, 383, 72, 452, 190)$; 随机选取 $M = 124785$ 时, 求解 $W = 88673$, 并且计算得 $\mathbf{a}^{(1)} = w\mathbf{a} \mod M = (14948, 20239, 20421, 24211, 1895)$; 随机选取 $M = 196784$ 时, 求解

$W = 136755$, 并且计算得 $\mathbf{a}^{(2)} = w\mathbf{a} \mod M = (23485, 32621, 7160, 23084, 7962)$. 而且, 不难验证向量 $\mathbf{a}$, $\mathbf{a}^{(1)}$, $\mathbf{a}^{(2)}$ 就是线性无关的.

4 结论

本文提出了一种基于联立丢番图逼近的子集和问题启发式求解算法, 利用联立丢番图逼近生成 $n-1$ 子集和问题, 这新生成的 $n-1$ 子集和问题与原始的子集和问题是线性无关的, 这样就可以构造一个 n 维的线性方程组. 通过求解这个线性方程组, 进而得到原子集和问题的解. 该算法为求解子集和问题提供新的研究思路. 该算法的计算复杂度分析是今后可以继续开展的研究工作.

5 致谢

作者感谢审稿人的宝贵意见, 这些意见极大改进了论文的内容. 特别是论文的定理 2 关于子集和问题构造的存在性证明就是来自审稿人的建议.

References

[1] OKAMOTO T, TANAKA K, UCHIYAMA S. Quantum public-key cryptosystems[C]. In: Advances in Cryptology—CRYPTO 2000. Springer Berlin Heidelberg, 2000: 147–165.

[2] WANG B, WU Q, HU Y. A knapsack-based probabilistic encryption scheme[J]. Information Sciences, 2007, 177(19): 3981–3994.

[3] GENTRY C. Fully homomorphic encryption using ideal lattices[C]. In: ACM Symposium on Theory of Computing, STOC. 2009: 169–178.

[4] VAN DIJK M, GENTRY C, HALEVI S, et al. Fully homomorphic encryption over the integers[C]. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer Berlin Heidelberg, 2010: 24–43.

[5] HOROWITZ E, SAHNI S. Computing partitions with applications to the knapsack problem[J]. Journal of the ACM, 1974, 21(2): 277–292.

[6] SCHROEPPEL R, SHAMIR A. A $TS^2 = O(2^n)$ time/space tradeoff for certain NP-complete problems[C]. In: 20th Annual Symposium on Foundations of Computer Science 1979. IEEE, 1979: 328–336.

[7] SCHROEPPEL R, SHAMIR A. A $T = O(2^{n/2})$, $s = O(2^{n/4})$ Algorithm for Certain NP-Complete Problems[J]. SIAM Journal on Computing, 1981, 10(3): 456–464.

[8] HOWGRAVE-GRAHAM N, JOUX A. New Generic Algorithms for Hard Knapsacks[C]. In: EUROCRYPT 2010: 235–256.

[9] Becker A, Coron J S, Joux A. Improved Generic Algorithms for Hard Knapsacks[C]. In: Advances in Cryptology—EUROCRYPT 2011. Springer Berlin Heidelberg, 2011: 364–385.

[10] JOUX A, STERN J. Improving the critical density of the Lagarias-Odlyzko attack against subset sum problems[C]. In: Fundamentals of Computation Theory. Springer Berlin Heidelberg, 1991: 258–264.

[11] COSTER M J, LAMACCHIA B A, ODLYZKO A M, et al. An improved low-density subset sum algorithm[C]. In: Workshop on the Theory and Application of of Cryptographic Techniques. Springer Berlin Heidelberg, 1991: 54–67.

[12] COSTER M J, JOUX A, LAMACCHIA B A, et al. Improved low-density subset sum algorithms[J]. Computational Complexity, 1992, 2(2): 111–128.

[13] BRICKELL E F. Solving low density knapsacks[C]. In: Advances in Cryptology. Springer US, 1984: 25–37.

[14] LAGARIAS J C, ODLYZKO A M. Solving Low-density Subset Sum Problems[M]. IEEE Computer Society press, 1983.

[15] LAGARIAS J C, ODLYZKO A M. Solving low-density subset sum problems[J]. Journal of the ACM(JACM), 1985, 32(1): 229–246.

[16] IZU T, KOGURE J, KOSHIBA T, et al. Low-density attack revisited[J]. Designs Codes & Cryptography, 2007, 43(1): 47–59.

[17] CHOR B, RIVEST R L. A Knapsack type public key cryptosystem based on arithmetic in finite fields (preliminary draft)[C]. In: Advances in Cryptology. Springer Berlin Heidelberg, 1985: 54–65.

[18] CHOR B, RIVEST R L. A knapsack-type public key cryptosystem based on arithmetic in finite fields[J]. IEEE Transactions on Information Theory, 1988, 34(5): 901–909.
[19] ENCINAS L H, MASQUÉ J M, DIOS A Q. Maple implementation of the Chor-Rivest cryptosystem[C]. In: International Conference on Computational Science. Springer-Verlag, 2006: 438–445.
[20] ENCINAS L H, MASQUÉ J M, DIOS A Q. Safer parameters for the Chor–Rivest cryptosystem[J]. Computers & Mathematics with Applications, 2008, 56(11): 2883–2886.
[21] ENCINAS L H, MASQUÉ J M, DIOS A Q. Analysis of the efficiency of the Chor–Rivest cryptosystem implementation in a safe-parameter range[J]. Information Sciences, 2009, 179(24): 4219–4226.
[22] KATE A, GOLDBERG I. Generalizing cryptosystems based on the subset sum problem[J]. International Journal of Information Security, 2011, 10(3): 189–199.
[23] OMURA K, TANAKA K. Density attack to the knapsack cryptosystems with enumerative source encoding[J]. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2004, 87(6): 1564–1569.
[24] NGUYÊN P Q, STERN J. Adapting density attacks to low-weight knapsacks[C]. In: Advances in Cryptology—ASIACRYPT. 2005, 3788: 41–58.
[25] KUNIHIRO N. New definition of density on knapsack cryptosystems[C]. In: International Conference on Cryptology in Africa. Springer Berlin Heidelberg, 2008: 156–173.
[26] HU G, PAN Y, ZHANG F. Solving random subset sum problem by lp-norm SVP oracle[C]. In: Public Key Cryptography. Springer Berlin Heidelberg, 2014: 399–410.
[27] PAN Y B. The Analysis and Design of Lattice-based Public-key Cryptosystems[D]. Beijing: Academy of Mathematics and Systems Science, 2010: 8-13.
潘彦斌. 基于格的公钥密码体体制的分析与设计 [D]. 北京: 中国科学院数学与系统科学研究院, 2010: 8–13.
[28] BAOCANG W, YUPU H. Public key cryptosystem based on two cryptographic assumptions[J]. IEE Proceedings-Communications, 2005, 152(6): 861–865.
[29] WANG B C, HU Y P. Diophantine approximation attack on a fast public key cryptosystem[J]. Information Security Practice and Experience, 2006: 25–32.
[30] LENSTRA A K, LENSTRA H W, LOVÁSZ L. Factoring polynomials with rational coefficients[J]. Mathematische Annalen, 1982, 261(4): 515–534.
[31] SCHNORR C P. A hierarchy of polynomial time lattice basis reduction algorithms[J]. Theoretical Computer Science, 1987, 53(2–3): 201–224.
[32] WANG X, LIU M, TIAN C, et al. Improved Nguyen-Vidick heuristic sieve algorithm for shortest vector problem[C]. In: ACM Symposium on Information, Computer and Communications Security. ACM, 2011: 1–9.

作者信息

王保仓 (1989–), 河南郸城人, 博士, 教授. 主要研究领域为云计算安全, 大数据安全, 数论算法, 全同态加密, 后量子公钥密码.
E-mail:
bcwang79@aliyun.com

卢珂 (1993–), 河南固始人, 硕士. 主要研究领域为云存储安全.
E-mail:
18439925995@163.com