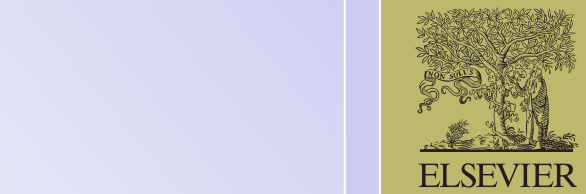
HNP及其格求解方法

2022数学密码赛赛题之二

2022年04月

HNP及其格求解方法

- §1 HNP的研究背景及数学描述
- §2 HNP的格求解模型
- §3 格算法简介



§1 HNP的研究背景及数学描述

HNP是隐藏数问题(Hidden Number Problem)的简称, 最早由D.Boneh和R.Venkatesan等人引入来研究Diffie-Hellman密钥交换体制的比特安全性(bit-security).

后来, P.Q.Nguyen、E.Shparlinski等人把数字签名算法(Digital Signature Algorithm, 简记为DSA)在部分私钥比特已知时的体制安全性归于HNP问题的求解.

可以使用HNP求解方法攻击的私钥存在熵漏的公钥密码和签名体制中包括

DSA

ECDSA

ElGamal

.....

比特泄露的原因包括但不限于:

随机数生成算法缺陷、

侧信道攻击暴露、

.....

• HNP的数学描述

设q是一固定的m比特正整数(已知), $m = [\log_2(q)]$. 固定用$\{0, 1, \cdots, q-1\}$来表示整数模q剩余类环$\mathbb{Z}/q\mathbb{Z}$中的元素.

设$x_0 \in \mathbb{Z}/q\mathbb{Z}$为未知变元, 即为**隐藏数**.

在$\mathbb{Z}/q\mathbb{Z}$中随机一致的选取n个元素(已知), 记为列向量$(\alpha_1, \cdots, \alpha_n)^t \in (\mathbb{Z}/q\mathbb{Z})^n$. 对$1 \le i \le n$, 令$\beta_i = \alpha_i x_0 \in \mathbb{Z}/q\mathbb{Z}$, $(\beta_1, \cdots, \beta_n)^t$是线性方程组

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} x_0 \equiv \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \pmod{q}$$

的常数项.

- HNP的数学描述(续)

假设每个分量β_i的m比特分位信息只有部分固定分位的比特已知, 即已知方程组常数项分量的部分比特信息, 试求解隐藏数x_0.

- HNP的主要参数

模数q, q的比特数m, 方程量n, β_i的已知比特数s

注记: 数据文件中的Coeff, KnownNonce向量分别给出了方程组系数向量与常数项已知部分比特的向量.

- **HNP的可解性**

如果β的全部信息已知, 就可以用很少的方程还原出x_0(求解模q上单变元线性方程组).

而β只有部分信息(s比特)已知时, 比值s/m决定了一个方程给出的信息熵. 当方程量n充分大时, 隐藏数x_0的唯一确定的.

在方程量足够的条件下, 需要通过构建数学模型, 寻找合适的算法, 实现求出隐藏数x_0的目的.

§2 HNP的格求解模型

当HNP中常数项β分量的高s比特已知时, 通过构建格模型, 使用格算法可以求解隐藏数x_0.

首先, 把模q的方程$\alpha_i x_0 \equiv \beta_i (\mod q)$转化为整方程$\alpha_i x_0 + k_i q = \beta_i$, 其中$k_i \in \mathbb{Z}$未知.

假设给出n个隐藏数方程, 考虑由下面矩阵A的列向量张成的格Λ,

$$A = \underbrace{\begin{pmatrix} \alpha_1 & q & 0 & \cdots & 0 \\ \alpha_2 & 0 & q & \cdots & 0 \\ \vdots & & \cdots & \ddots & \\ \alpha_n & 0 & 0 & \cdots & q \end{pmatrix}}_{n+1}$$

格Λ中包含向量

$$b = A\begin{pmatrix} x_0 \\ k_1 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} \alpha_1 x_0 + k_1 q \\ \alpha_2 x_0 + k_2 q \\ \vdots \\ \alpha_n x_0 + k_n q \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \in \Lambda.$$

而每个β_i的高s比特已知, 记为y_i(数据中的KnownNonce), 可设$\beta_i = 2^{m-s}y_i + \varepsilon_i, 0 \le \varepsilon_i \le 2^{m-s} - 1$. 那么向量$t = \begin{pmatrix} 2^{m-s}y_1 \\ 2^{m-s}y_2 \\ \vdots \\ 2^{m-s}y_n \end{pmatrix}$与格点$b$的差为$b - t = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$.

由于差向量$\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$的长度较短($s$越大, ε越短), 格点b可以认为是与已知的目标向量t比较接近的格点.

当m, s, n的取定时, 我们希望能够通过求解格中最近向量问题(CVP)的算法, 求出距离目标向量t较近的格点b. b对A中第一个向量的系数x_0就给出HNP的解.

使用该方法能否求解HNP, 取决与格Λ的性质及CVP算法的求解能力. 对所需方程量n的定量分析是非常有必要的.

§3 格算法简介

3.1. 格的基本概念

直观的讲, 格是在空间中**“规则排列”**的**离散点**的集合.

定义 格Λ是n维欧氏空间$\mathbb{E}^n$中离散的加法子群.

欧氏空间: 带有内积结构的有限维实线性空间;

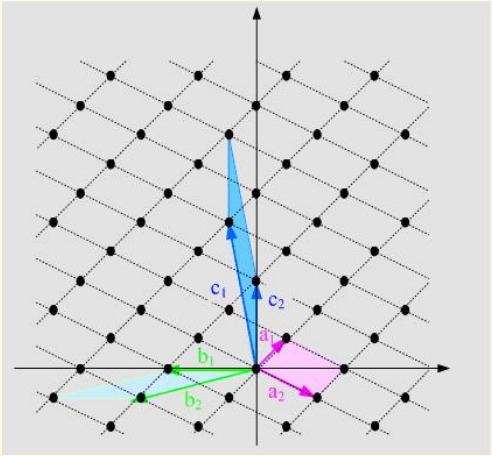
加法运算: 线性空间中向量的加法.

欧氏空间$\mathbb{E}^n$中(满秩)格Λ是由$\mathbb{R}$-线性无关的向量组$B = \{b_1, b_2, \cdots, b_n\}$生成的(加法)Abelian子群, 即

$$\Lambda = \mathbb{Z}b_1 \oplus \cdots \oplus \mathbb{Z}b_n = \{\Sigma_{i=1}^n x_i b_i \mid x_i \in \mathbb{Z}, 1 \le i \le n\}.$$

B称为Λ的一组**基**. 对Λ中任意的格点b, 存在唯一的整系数向量$x \in \mathbb{Z}^n$, 使得$b = Bx$.

当$n \ge 2$时, 一个n维格Λ具有无限多的格基.



• 格的基本参数

格的维数: 格Λ一组基的元素个数, $n = \dim(\Lambda)$.

格的体积: 设$B = \{b_1, b_2, \cdots, b_n\}$是格$\Lambda$的一组基, $G = ((b_i, b_j))_{1 \le i,j \le n}$是$B$的Gram矩阵. 格的体积定义为:

$$\mathrm{vol}(\Lambda) = \sqrt{\det G},$$

格体积的定义与格基的选择无关.

体积的几何意义: 格的体积度量了一个格点周围空间的大小, 反应了空间中格点分布的疏密, 体积越大, 格点分布越稀疏.

格中(非零)最短向量长度: 格中最短的非零格向量长度记为:

$$\lambda_1(\Lambda) = \min\{\|b\| > 0 \mid b \in \Lambda - \{0\}\},$$

长度为$\lambda_1(\Lambda)$的格向量称为格中(非零)最短向量.

- $\lambda_1(\Lambda)$与$\mathrm{vol}(\Lambda)$的关系

Gauss直观(Gaussian Heuristic):

设D是$\mathrm{span}_{\mathbb{R}}(\Lambda)$中一个$n$维凸体, Gauss直观“预测”$D$中格点**粗略**的个数约为$\mathrm{vol}(D)/\mathrm{vol}(\Lambda)$.

这启发我们把体积为$\mathrm{vol}(\Lambda)$的n维球的半径作为$\lambda_1(\Lambda)$的估计, 此时有

$$\lambda_1(\Lambda) \approx \left(\frac{\mathrm{vol}(\Lambda)}{\sigma_n}\right)^{\frac{1}{n}} \approx \sqrt{\frac{n}{2\pi e}}\mathrm{vol}(\Lambda)^{\frac{1}{n}}.$$

- Gauss直观(Gaussian Heuristic)

格Λ的Gauss直观

$$\sqrt{\frac{n}{2\pi e}}\mathrm{vol}(\Lambda)^{\frac{1}{n}}$$

的值对Λ的性质特别是其中计算问题具有非常重要的含义

对于给定的初始格基, 格的体积进而Gauss直观是容易计算的.

一般对于“随机”产生的维数较高的格Λ, Gauss直观给出了$\lambda_1(\Lambda)$较好的估计.

另一方面, 对于随机给定由格张成的空间中的目标向量t, 距t最近的格点距离大致也在Gauss直观的附近.

- 格中两个计算困难问题

最短向量问题(SVP): 求取Λ中非零的最短向量b, 即$b \in \Lambda$满足$\|b\| = \min\{\|c\| \in \mathbb{R}_{\geq 0} \mid c \in \Lambda \setminus \{0\}\} = \lambda_1(\Lambda)$.

最近向量问题(CVP): 给定欧氏空间中的任一目标向量t, 求取格向量$b \in \Lambda$, 使得$\|t - b\|$最短, 即

$$\|t - b\| = \mathrm{dist}(t, \Lambda) = \min\{\|t - c\| \in \mathbb{R}_{\geq 0} \mid c \in \Lambda\}.$$

现有的研究结果表明:

– 确切CVP是NP-完全的;

– 确切SVP在随机归约下是NP-完全的.

虽然SVP和CVP都基本上是NP困难的, 但有很多有效算法可以求解近似SVP和CVP.

SVP算法举例:

(1)格基约化算法: 从格的原始基出发, 通过做格基的线性变换, 求出一组约化基, 从约化基寻找出短的格向量, 作为近似SVP的解. 知名的格算法有: LLL、BKZ、BKZ2.0 · · ·

(2)直接从格基出发, 搜索出短向量. 例如: ENUM、ENUM with pruning、Sieve algorithm · · ·

注记: 如果格Λ由生成元给出(未必是基), 计算得到格的一组基是容易的.

CVP算法举例:

(1)最近平面法(Nearest Plane Algorithm): 给定目标向量t, 通过格基的正交化, 寻找距离t较近的格点; 使用ENUM算法的CVP版本, 搜索最近向量(格的维数较大时, 需要剪枝策略); 或者筛法的CVP版本···

(2)嵌入SVP求解: 设B是格基矩阵, t是目标向量. 考虑矩阵$\begin{pmatrix} B & t \\ 0 & K \end{pmatrix}$列张成的格$\Lambda'$. 其中$K$是合适的因子. 当$K$较大时, Λ'的约化基中含有短向量$\pm \begin{pmatrix} t \\ K \end{pmatrix} - \begin{pmatrix} b \\ 0 \end{pmatrix}$. b可以作为距离t较近的格点.

格算法的时间复杂度

影响格算法的时间复杂度的参数主要是格的维数n和原始的格向量长度的规模$B = \|b_i\|, 1 \leq i \leq n$.

根据实际问题的需求, 构建更高效的格模型, 制定针对性策略, 采用合适的格算法, 选取恰当的参数, 提高问题的解决能力是非常有挑战性的工作

最后, 欢迎大家改进现有的HNP格求解方法提高求解能力, 也鼓励尝试创设求解HNP的非格模型方法.

References

[1] Boneh, D., and Venkatesan, R. , *Hardness of Computing the Most Significant Bits of Secret Keys in Diffie-Hellman and Related Schemes*, in Koblitz, N. (ed.) Advances in Cryptology - CRYPTO '96. LNCS, vol. 1109, pp. 129-142. Springer, Heidelberg, 1996.

[2] Nguyen P.Q., and Shparlinski I.E., *The Insecurity of the Digital Signature Algorithm with Partially Known Nonces*, in Journal of Cryptology, 15(3), 151-176, 2002.

[3] Akavia, A., *Solving Hidden Number Problem with One Bit Oracle and Advice*, in Halevi, S. (ed.) Advances in Cryptology - CRYPTO 2009. LNCS, vol. 5677, pp. 337-354. Springer, Heidelberg, 2009.

欢迎大家踊跃讨论!