

针对离散私钥比特泄漏的 RSA 格攻击方法

刘向辉^{1,2}, 韩文报^{1,2}, 王 政^{1,2}, 权建校³

(1. 解放军信息工程大学四院, 郑州 450002; 2. 数学工程与先进计算国家重点实验室, 郑州 450002;

3. 江南计算技术研究所, 江苏 无锡 214083)

摘 要: RSA 算法是目前应用最广泛的公钥密码体制之一, 而格攻击是针对 RSA 体制的一类重要攻击方法。为此, 将 RSA 算法的部分私钥泄漏问题转化为多变元线性同余方程的求解问题, 基于同余方程构造出特定的格, 利用 LLL 格基约化算法进行约化, 从而以一定的概率求得同余方程的小根。以上述多变元线性同余方程的小根求解技术为基础, 提出一种针对离散私钥比特泄漏的 RSA 格攻击方法。在该方法下, 如果 RSA 算法的公钥参数 $e=N^\beta \leq N^{1/2}$, 并且私钥 d 的未知部分 $N^\alpha \leq N^{1/2-\beta}$, 则能以高概率恢复出 RSA 算法的私钥 d 。通过 NTL 包对长度为 1 024 bit 的大整数进行实验, 结果验证了该攻击方法的有效性。

关键词: RSA 算法; 格攻击; 离散私钥比特泄漏; 线性同余方程; 小根; 格基约化算法

RSA Lattice Attack Method for Discrete Private Key Bit Leakage

LIU Xiang-hui^{1,2}, HAN Wen-bao^{1,2}, WANG Zheng^{1,2}, QUAN Jian-xiao³

(1. The Fourth Institute, PLA Information Engineering University, Zhengzhou 450002, China;

2. State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450002, China;

3. Jiangnan Institute of Computing Technology, Wuxi 214083, China)

【Abstract】 RSA algorithm is one of the most widely used public key cryptosystems at present and lattice attacks play an important role for the analysis of RSA system. The problem of partial discrete private key bit leakage is transformed into the solution of multivariate linear congruence equations and a special lattice is constructed. And then by the lattice reduction algorithms such as LLL algorithm, the small roots of multivariate linear congruence equations can be obtained with a high probability. Based on the above technology, this paper proposes a lattice attack method on RSA for discrete private key bit leakage. With this method, if the public parameter satisfies $e=N^\beta \leq N^{1/2}$ and the unknown part of private key d satisfies $N^\alpha \leq N^{1/2-\beta}$, it can recover the private key d with a high probability. The experiment on 1 024 bit number is given with NTL package and the results verify the availability of the attack method.

【Key words】 RSA algorithm; lattice attack; discrete private key bit leakage; linear congruence equation; small root; lattice base reduction algorithm

DOI: 10.3969/j.issn.1000-3428.2014.03.033

1 概述

自 1976 年公钥密码的思想提出以来, 各种公钥密码体制不断涌现, 但公认安全的且应用广泛的却并不多, 其中较著名的是 RSA 公钥密码体制、ElGamal 公钥密码体制和椭圆曲线公钥密码体制。而 RSA 公钥密码体制由于具有简单易用、明文长度相同等优点, 在各种秘密通信中得到广泛应用, 一直是公钥密码学研究的热点^[1]。一般来讲, 公钥密码体制往往利用数学中已经得到证明的难题或公认的难题来设计方案, RSA 算法就是建立在大数分解问题上的。目前, 针对大数分解问题最好的通用攻击算法是一般数域

筛法, 它是一个亚指数时间算法, 在当前的计算能力下无法对实际使用的 RSA 模数进行分解。也就是说, 在不借助其他条件下直接通过大数分解对 RSA 体制进行攻击是困难的。

然而, 在实际使用过程中可能会泄漏 RSA 体制的部分信息, 例如泄漏私钥 p 或者 d 的若干比特信息。同时, 由于旁道攻击等手段的发展, 攻击者往往能够获得部分密钥信息。如何利用这些已知信息对 RSA 体制进行攻击成为密码学的一个重要研究课题。文献[2-3]提出利用 LLL 算法求解整数系同余方程及多项式方程小根的方法, 此后该方法被广泛应用于 RSA 算法的私钥泄漏攻击中, 例如, 在泄漏

基金项目: 国家自然科学基金资助项目(61003291); 数学工程与先进计算国家重点实验室开放基金资助项目(2013A03)。

作者简介: 刘向辉(1984—), 男, 博士研究生, 主研方向: 密码学, 信息安全; 韩文报, 教授、博士、博士生导师; 王 政, 副教授、博士; 权建校, 助理研究员、硕士。

收稿日期: 2013-03-07 **修回日期:** 2013-04-07 **E-mail:** lxhkz2002@163.com

私钥 d 的低 $n/4$ 比特, 同时加密指数较小的条件下 RSA 的私钥恢复^[4]; 在加密指数较大情况下的 RSA 部分私钥泄漏攻击等^[5]; 私钥指数 p 的连续比特泄漏攻击等^[6]。

早期的 RSA 私钥泄漏攻击都建立在文献[2-3]求解多变量模方程或者求解多变量多项式方程小根的基础上, 主要集中在私钥 d 的高位连续比特或者低位连续比特泄漏的情形。在实际环境中, 攻击者获得的部分私钥信息通常是不连续的, 特别是旁道攻击^[7]的存在使得 RSA 体制离散比特私钥泄露攻击也显得更有意义。文献[8]通过构造多变量线性模方程并利用格基约化算法进行求解, 提出针对泄漏私钥 p 部分离散比特的攻击方法, 在泄漏私钥 p 的 70% 比特信息的条件下该方法能够有效分解 RSA 的模数。文献[9]利用变换多项式的格构造方法给出针对私钥指数 d 的离散比特泄漏攻击, 但该方法要求格的维数较高。

本文通过构造多变量线性模方程, 并利用典型的格构造方法, 针对 RSA 算法私钥 d 部分离散比特泄漏的情况进行分析。在公钥指数较小的条件下, 如果泄漏私钥 d 的部分离散比特, 则可以有效恢复出 RSA 算法的私钥参数 d 。

2 准备知识

本节给出 RSA 格攻击所用到的基础知识, 包括格的基本概念、Minkowski 定理等格的基本理论以及 LLL 算法等格基约化算法, 具体细节可参考文献[10]。

定义 1 设 $b_1, b_2, \dots, b_m \in R^n$ 为一组线性无关的向量, 其中, R^n 为实数域上的 n 维向量空间; $L(b_1, b_2, \dots, b_m) = \{z = \sum_{i=1}^m \lambda_i b_i \mid \lambda_i \in Z\}$ 称为以 (b_1, b_2, \dots, b_m) 为基的格, m 称为格的维数。若 $m=n$, 则格 L 称为满秩。如果 $R=Z$, 也即格中元素取自整数环, 则称其为整格。

定义 2 令 $b_1^*, b_2^*, \dots, b_m^*$ 为 b_1, b_2, \dots, b_m 进行 Gram-Schmidt 正交化后所得的向量, 则对以 (b_1, b_2, \dots, b_m) 为基的 m 维格 L , 称其行列式 $\det(L) = \prod_{i=1}^m \|b_i^*\|$ 为格的范数。其中, $\|b\|$ 表示向量 b 的欧氏范数, 又称为向量的长度, 也即若 $b = (b_1, b_2, \dots, b_n)$, 则 $\|b\| = (b_1^2 + b_2^2 + \dots + b_n^2)^{1/2}$ 。

显然, 一个格有多组基, 在解决格上相关问题时, 希望能找到一组特定的基有利于问题的解决, 寻找这组基的过程就称为格基约化, 这组基就称为约化基。拥有长度较短向量的基往往具有一些良好的性质, 如何寻找具有短向量的基一直受到人们的关注。定理 1 给出了格中最短向量长度的上界。

定理 1 设格 $L \subseteq R^n$ 是一个满秩格, 那么必然存在非零向量 v 满足 $\|v\| \leq \sqrt{n} \det(L)^{1/n}$ 。

Minkowski 定理说明格中最短向量的长度 $\|v\| \leq \sqrt{n} \det(L)^{1/n}$, 但如何寻找格中范数最小非零向量的最短向量问题(the Shortest Vector Problem, SVP)是格上的著名难题,

格理论研究表明它是 NP 问题。1982 年, 著名的 LLL 算法由 Lenstra 等人提出^[11], 可以在多项式时间求取格中长度为 $\|v\| \leq 2^{(n-1)/4} \det(L)^{1/n}$ 的近似最短向量。实际上, LLL 算法所求得的向量长度比理论结果要好, 往往能够求得需要的短向量。

3 一种离散私钥比特泄漏的 RSA 格攻击方法

本节根据 RSA 算法的已知信息建立多变量线性同余方程, 并利用格基约化算法对方程进行求解, 从而给出离散私钥比特泄漏的 RSA 格攻击方法。

设 RSA 算法的模数为 $N = pq$, 公钥为 e , 私钥为 d , 则满足 $ed = 1 \bmod \varphi(N)$, 其中, $\varphi(N) = (p-1)(q-1)$ 为欧拉函数。假设私钥 d 有 n 个比特块未知, 设其值分别为 x_1, x_2, \dots, x_n , 于是得:

$$d = a_1 x_1 + a_2 x_2 + \dots + a_n x_n + a_{n+1}$$

其中, $a_k = 2^l$ 表示第 k 个未知块 x_k 在第 l 比特的位置。

对于 $ed = 1 \bmod \varphi(N)$, 显然存在正整数 k 使得 $ed - 1 = k\varphi(N)$ 。将 d 代入可得:

$$e(a_1 x_1 + a_2 x_2 + \dots + a_n x_n) + ea_{n+1} - 1 = k(N - (p+q) + 1)$$

将上述方程模 N 可得 $n+1$ 变元的线性同余方程如下:

$$ea_1 x_1 + ea_2 x_2 + \dots + ea_n x_n + k(p+q-1) + ea_{n+1} - 1 = 0 \bmod N$$

将同余方程乘以 $e^{-1} \bmod N$ 可得:

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n + e^{-1}k(p+q-1) + (a_{n+1} - e^{-1}) = 0 \bmod N$$

方程的解为:

$$(y_1, y_2, \dots, y_n, y_{n+1}) = (x_1, x_2, \dots, x_n, k(p+q-1))$$

下面对上述方程进行求解。为描述方便, 令 $b_1 = a_1, \dots, b_n = a_n, b_{n+1} = e^{-1}, b_{n+2} = a_{n+1} - e^{-1}$, 同时, $x_{n+1} = k(p+q-1)$, 那么方程具有以下形式:

$$f_N(x_1, x_2, \dots, x_{n+1}) = b_1 x_1 + b_2 x_2 + \dots + b_{n+1} x_{n+1} + b_{n+2} = 0 \bmod N$$

对于一般的多变量线性同余方程, 解的个数以及解的结构是一个较困难的问题, 但是在某些限定条件下, 能够得到上述同余方程的唯一解。定理 2 给出了一种求解情况, 能够在一些限定条件下完成对 RSA 算法私钥 d 的恢复。

定理 2 令 N 是 RSA 模数, 私钥 d 满足 $f_N(x_1, x_2, \dots, x_{n+1}) = b_1 x_1 + b_2 x_2 + \dots + b_{n+1} x_{n+1} + b_{n+2} = 0 \bmod N$ 是线性同余方程, 其中, 方程的解 $y_1 < N^{\gamma_1} = X_1, \dots, y_n < N^{\gamma_n} = X_n$, $\gamma_1 + \gamma_2 + \dots + \gamma_n = \alpha$ 为私钥 d 未知的 n 个比特块, 同时假设 $e = N^\beta$ 。如果 $\alpha + \beta \leq 1/2$, 那么可以恢复 RSA 算法的私钥 d 的未知部分。

证明: 对于方程 $f_N(x_1, x_2, \dots, x_{n+1}) = b_1 x_1 + b_2 x_2 + \dots + b_{n+1} x_{n+1} + b_{n+2} = 0 \bmod N$, 首先假设 b_i 和 N 互素, 也即

$\gcd(b_i, N) = 1, i = 1, 2, \dots, n+2$, 否则可以直接分解 N , 从而求出私钥 d 。

由已知条件 $y_1 < N^{\gamma_1} = X_1, \dots, y_n < N^{\gamma_n} = X_n, k = (ed-1)/\varphi(N) < e = N^\beta$, 且 $p+q-1 \leq 3N^{1/2}$ 可得, $y_{n+1} < 3N^{\beta+1/2} = X_{n+1}$ 。对系数 b_{n+2} , 引入变元 x_{n+2} 满足 $y_{n+2} \leq X_{n+2} = 1$, 也即将 $n+1$ 变元的线性同余方程当作 $n+2$ 变元方程:

$$f_N(x_1, x_2, \dots, x_{n+2}) = b_1x_1 + b_2x_2 + \dots + b_{n+2}x_{n+2} = 0 \pmod{N}$$

方程的解 $(y_1, y_2, \dots, y_{n+2})$ 满足 $y_i < N^{\gamma_i} = X_i, i = 1, 2, \dots, n, y_{n+1} < 3N^{\beta+1/2} = X_{n+1}, y_{n+2} = 1$ 。

将方程两边同乘以 $-b_1^{-1}$ 得新的线性同余方程 $x_1 + c_2x_2 + c_3x_3 + \dots + c_{n+2}x_{n+2} = 0 \pmod{N}$, 对方程的解 $(y_1, y_2, \dots, y_{n+2})$, 必存在正整数 y 满足 $c_2y_2 + c_3y_3 + \dots + c_{n+2}y_{n+2} = y_1 - yN$ 。令 $Y_i = N/X_i$, 考虑如下矩阵生成的格 L :

$$L = \begin{bmatrix} Y_1N & 0 & 0 & \dots & 0 & 0 \\ Y_1c_2 & Y_2 & 0 & \dots & 0 & 0 \\ \vdots & 0 & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \dots & Y_{n+1} & 0 \\ Y_1c_{n+2} & 0 & 0 & \dots & 0 & Y_{n+2} \end{bmatrix}$$

显然, 它是一个 $n+2$ 维的格, $\mathbf{v} = (y, y_2, \dots, y_{n+1}, y_{n+2})$, $\mathbf{B} = (Y_1y_1, Y_2y_2, \dots, Y_{n+2}y_{n+2})$ 是格 L 中的向量。由于 $Y_iy_i = \frac{y_i}{X_i}N \leq N$, 从而可得 $\|\mathbf{v}\| \leq \sqrt{n+2}N$ 。如果向量 \mathbf{v} 是格 L 中的短向量, 那么可以通过格基约化算法将其求出, 从而得到方程的解, 下面利用 Minkowski 定理说明 \mathbf{v} 是格 L 中的短向量。

易求格 L 的范数为:

$$\det(L) = N \prod_{i=1}^{n+2} Y_i = N \prod_{i=1}^{n+2} \frac{N}{X_i} = N^{n+3} \prod_{i=1}^{n+2} \frac{1}{X_i}$$

如果 $\prod_{i=1}^{n+2} X_i \leq N$, 那么向量 $\|\mathbf{v}\| \leq \sqrt{n+2}N \leq \sqrt{n+2} \det(L)^{1/(n+2)}$, 根据 Minkowski 定理可知向量 \mathbf{v} 是格 L 中的一个短向量, 利用格基约化算法可以求出。由已知条件, $\prod_{i=1}^n X_i \leq N^{\gamma_1+\gamma_2+\dots+\gamma_n} = N^\alpha, X_{n+1} = 3N^{\beta+1/2}, X_{n+2} = 1$, 省略小常数 3, 那么当 $\alpha + \beta \leq 1/2$, 可得 $\prod_{i=1}^{n+2} X_i \leq N$ 。结论得证。

注释 1 在证明过程中, 省略小常数 3, 这是因为常数 3 相对于 N 非常微小。如果 N 是一个 1 024 bit 的大整数, 小常数影响 N 的指数为 0.00 155, 而且它并不随着攻击算法参数的改变而改变。于是为了计算方便, 通常忽略这些小常数, 并且它基本不影响攻击算法的结果。

注释 2 格基约化算法通常采用 LLL 算法, 它是一个多项式时间算法, 这样能够在多项式时间内完成攻击, 而且 LLL 算法往往能够得到需要的短向量。

注释 3 定理 2 描述的攻击方法是一个概率算法。也就是说, 如果满足已知条件, 该方法并不能保证一定可以恢复出私钥 d 。但实际结果表明, 该方法往往能够得到方程的解并恢复出私钥。

根据上述描述及定理 2 的证明, 给出如下针对离散私钥比特泄漏的 RSA 格攻击算法。

算法 离散私钥比特泄漏的 RSA 格攻击算法

输入 RSA 算法的模数为 N , 公钥 $e = N^\beta$, 私钥 d 未知比特块的界满足 $y_1 < X_1, y_2 < X_2, \dots, y_n < X_n, \prod_{i=1}^n X_i \leq N^\alpha$ 且 $\alpha + \beta \leq 1/2$

输出 私钥 d 未知比特块的值 y_1, y_2, \dots, y_n

(1) 根据式 $ed = 1 \pmod{\varphi(N)}$ 列出 $n+1$ 元线性同余方程 $f_N(x_1, x_2, \dots, x_{n+1})$;

(2) 根据定理 2 的证明对方程进行变换并构造格 L ;

(3) 利用 LLL 算法求取格 L 中的短向量;

(4) 对短向量进行代入计算出原始方程的解;

(5) 输出私钥 d 未知比特块的值。

假设 RSA 算法的模数 N 的长度为 m , 公钥参数为 $e = N^\beta$, 根据上述算法, 如果私钥 d 的长度为 $(1/2 - \beta)m$ 比特块未知, 那么可以将私钥 d 恢复出来。在实际使用中, 通常选取 N 为 1 024 bit 的数, $e=65\ 537$, 也就是说, $\beta=0.015\ 6$ 。此时, 如果私钥 d 长度为 $0.483m$ 的比特块未知, 那么可以将 d 恢复出来。相对于单块连续比特泄漏的情形, 攻击算法需要泄漏更多的信息。例如, Boneh 等人只需要泄漏私钥 d 的低 $m/4$ 比特, 也即 $0.75m$ 的比特块未知, 那么可以恢复出私钥 d 。但是, 本文攻击算法是基于离散比特的私钥泄漏攻击, 未知的比特信息可以在任意位置, 在实际环境中具有更强的适应性。

4 实验结果与分析

本文算法能够有效恢复出离散私钥比特泄漏情形的私钥 d , 本节对攻击方法进行实现, 给出了部分实验结果。

实验采用 Intel Core2 Duo CPU E7500 2.93 GHz、2 GB 内存、Windows XP 操作系统、C++编程语言、Visual Studio 2005 编程环境。实验基本数据类型以及部分函数使用 NTL 包 5.5.2 版本^[12]。

随机产生 1 024 bit 的大整数如下:

$N=89884656743115795386465259539451236680898848947115$
 32863671504057886633790275048156635423866120376801
 05600569399356966788293948844072084453245030147457
 09298151367448461125728029121649765323616136679383
 49007024304932238762308699491286658762896157592200
 92451208280035185453770595398900240518477232773451
 74851613

选择公钥参数 $e=65\ 537$, 并计算其私钥参数:

