Implicit Related-Key Factorization Problem on the RSA Cryptosystem

Mengce Zheng Honggang Hu

University of Science and Technology of China

December 5, 2019

Outline

Outline

The RSA Cryptosystem

One RSA instance consists of four parameters: N, e, d and $\varphi(N)$

- $N = pq$ with two large prime factors of the same bit-size
- Public and private keys (e, d) satisfy $ed \equiv 1 \bmod \varphi(N)$
- Euler's totient function $\varphi(N) = (p-1)(q-1)$
- Encryption is $c = m^e \bmod N$ and decryption is $c^d \bmod N$

Key equation of the RSA cryptosystem: $ed \equiv 1 \bmod \varphi(N)$

- $ed = k(N + 1 - p - q) + 1$ for an unknown positive integer k
- Many attacks have been proposed to solve RSA key equation
- Using small private key is always considered in attacks

Two Attacks on RSA

Partial key exposure attack

- Given a small fraction of the private key bits
- Eg: $d = \bar{d} + d'$ with known MSBs $\bar{d}$ and unknown LSBs d'
- The goal is to reconstruct the entire private key d

Implicit factorization problem

- Given an oracle providing implicit information about prime factors
- Eg: $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ with p_1, p_2 sharing some LSBs
- The goal is to find q_1, q_2 and then factor N_1, N_2

Our Research Problem

What if an attacker knows implicit information about private keys

- Given amounts of shared MSBs and LSBs of implicitly related keys
- Eg: $d_2 - d_1 = d_{21}D$ with known D and unknown middle bits d_{21}
- The goal is to factor RSA moduli using given implicit relation

We study this new attack scenario mainly from theoretical interests

- Disclose the vulnerability of RSA under weaker hypothesis
- Extend attacks and enrich cryptanalyses in the literature
- Private keys may be generated with imperfect randomness

Our Research Problem

Implicit Related-Key Factorization Problem

Let $(N_1, e_1, d_1), \ldots, (N_n, e_n, d_n)$ be n distinct RSA key pairs, where $N_1, \ldots, N_n$ are of the same bit-size and the prime factors are also all of the same bit-size. Given the implicit information that certain portions of the bit pattern in private keys $d_1, \ldots, d_n$ are common, under what condition is it possible to efficiently factor RSA moduli.

We consider the full size case for $e_i \approx N$. (N denotes an integer of the same bit-size as N_i for simplicity.) Assume $d_i \approx N^{\delta}$ satisfy the implicit relation $d_j = d_i + d_{ji}D$ for $1 \leq i < j \leq n$, where D denotes the bit-length of shared LSBs and d_{ji} denotes the difference between every two unknown middle bits with $|D| \approx N^{\gamma}$ and $|d_{ji}| \approx N^{\beta}$.

Lattice and Lattice Reduction

Lattice $\mathcal{L}$ is spanned by linearly independent vectors $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{R}^n$

- $\mathcal{L}(\vec{b}_1, \ldots, \vec{b}_n) = \{z_1\vec{b}_1 + \cdots + z_n\vec{b}_n : z_i \in \mathbb{Z}\}$
- The basis vectors generate a lattice basis matrix B
- The lattice determinant is $\det(\mathcal{L}) = |\det(B)|$ for full-rank B

Lattice reduction algorithm is used for finding approximately short vectors

- Eg: the famous LLL algorithm and its improved variants
- The running time is polynomial in n and $\log_2 \max_{1 \le i,j \le n} |B_{ij}|$
- Use Gaussian Heuristic to estimate the size of reduced vectors
- Application in cryptanalysis of RSA and other cryptosystems

Lattice-Based Technique

Find small roots of modular polynomial equations

1. Modular equation derived from known and unknown parameters
2. Construct shift polynomials sharing the common root modulo R
3. Transform coefficient vectors into a lattice basis matrix B
4. Reduce m-dimensional lattice $\mathcal{L}(B)$ using the LLL algorithm
5. Transform reduced short vectors into integer equations
6. Extract the common root of equations over the integers

Crucial Condition

$$\det(\mathcal{L}) < R^m$$

Outline

Concrete Attack Scenario

Given (N_1, e_1, d_1) and (N_2, e_2, d_2) with implicitly related keys d_1, d_2

- N_1, N_2 are of the same bit-size denoted by $\log_2 N$
- $e_1 \approx e_2 \approx N$ are approximately of the same bit-size of moduli
- $d_1 \approx d_2 \approx N^{\delta}$ satisfy $d_2 = d_1 + d_{21}D$ for $|d_{21}| \approx N^{\beta}$ and $|D| \approx N^{\gamma}$

Perform the splitting technique based on Gaussian Heuristic

- Construct a two-dimensional lattice generated by $\begin{bmatrix} a_0 & e_1 \\ 0 & N_1 \end{bmatrix}$
- $d_1 = a_1 c_1 + a_2 c_2$ for known a_1, a_2 and unknown c_1, c_2
- $|a_1| \approx |a_2| \approx N^{\frac{1}{4}}$ and $|c_1| \approx |c_2| \approx N^{\delta - \frac{1}{4}}$

Attack – Modular Equation

Combine $d_1 = a_1c_1 + a_2c_2$ with given implicit relation $d_2 = d_1 + d_{21}D$

- $d_2 = a_1c_1 + a_2c_2 + d_{21}D$ for unknown variables c_1, c_2, d_{21}
- Substitute d_2 in key equation $e_2d_2 = k_2(N_2 + 1 - p_2 - q_2) + 1$
- $e_2(a_1c_1 + a_2c_2 + d_{21}D) - k_2(N_2 + 1 - p_2 - q_2) - 1 = 0$

Find small roots of modular equation f in four variables

- $f(x, y, z, w) := x(y - N_2 - 1) + e_2a_1z + e_2Dw - 1 \bmod e_2a_2$
- Unknown variables: $x = k_2$, $y = p_2 + q_2$, $z = c_1$ and $w = d_{21}$
- Apply the linearization technique to let $u := xy - 1$
- $\bar{f}(x, z, w, u) := u - (N_2 + 1)x + e_2a_1z + e_2Dw \bmod e_2a_2$

Attack – Shift Polynomials

Define shift polynomials $g_{[i,j,k,l_1,l_2]}$ for $s \in \mathbb{Z}_+$ and $i, j, k, l_1, l_2 \in \mathbb{N}$

$g_{[i,j,k,l_1,l_2]}(x, y, z, w, u) := x^i y^j z^{l_1} w^{l_2} \bar{f}^k(x, z, w, u) E^{s-k}$ for $E = e_2 a_2$

Construct the set of shift polynomials using $\mathcal{G} := \mathcal{G}_1 \cup \mathcal{G}_2$ for $0 \leq \tau \leq 1$

$$\mathcal{G}_1 := \{g_{[i,0,k,l_1,l_2]}(x, y, z, w, u) : k = 0, \ldots, s; i = 0, \ldots, s-k; \\ l_1 = 0, \ldots, s-k-i; l_2 = 0, \ldots, s-k-i-l_1.\}$$

$$\mathcal{G}_2 := \{g_{[0,j,k,l_1-l_2,l_2-k]}(x, y, z, w, u) : l_1 = 0, \ldots, s; j = 1, \ldots, \tau l_1; \\ l_2 = 0, \ldots, l_1; k = 0, \ldots, l_2.\}$$

The common root is $(k_2,\ p_2 + q_2,\ c_1,\ d_{21},\ k_2(p_2 + q_2) - 1)$ modulo E^s

Attack – Lattice Construction

Coefficient vectors of $g_{[i,j,k,l_1,l_2]}(xX, yY, zZ, wW, uU)$ generate B

- X, Y, Z, W and U denote upper bounds on unknown variables
- The lattice basis matrix B is square and lower triangular

Eg: Lattice basis matrix for $s = 1$ and $\tau = 1$ with $C := -(N_2 + 1)$

	1	x	z	yz	w	yw	u	yu
$g_{[0,0,0,0,0]}$	E							
$g_{[1,0,0,0,0]}$		EX						
$g_{[0,0,0,1,0]}$			EZ					
$g_{[0,1,0,1,0]}$				EYZ				
$g_{[0,0,0,0,1]}$					EW			
$g_{[0,1,0,0,1]}$						EYW		
$g_{[0,0,1,0,0]}$		CX	e_2a_1Z		e_2DW		U	
$g_{[0,1,1,0,0]}$	C			e_2a_1YZ		e_2DYW	CU	YU

Attack – Lattice Reduction

$\det(\mathcal{L})$ is the product of diagonal entries in the lattice basis matrix B

- Diagonal entries for polynomials in $\mathcal{G}_1$ is $X^i Z^{l_1} W^{l_2} U^k E^{s-k}$
- Diagonal entries for polynomials in $\mathcal{G}_2$ is $Y^j Z^{l_1-l_2} W^{l_2-k} U^k E^{s-k}$

$$\begin{aligned}\det(\mathcal{L}) &= \left(\prod_{k=0}^{s} \prod_{i=0}^{s-k} \prod_{l_1=0}^{s-k-i} \prod_{l_2=0}^{s-k-i-l_1} X^i Z^{l_1} W^{l_2} U^k E^{s-k} \right) \\ &\times \left(\prod_{l_1=0}^{s} \prod_{j=1}^{\tau l_1} \prod_{l_2=0}^{l_1} \prod_{k=0}^{l_2} Y^j Z^{l_1-l_2} W^{l_2-k} U^k E^{s-k} \right) \\ &= X^{s_x} Y^{s_y} Z^{s_z} W^{s_w} U^{s_u} E^{s_E}\end{aligned}$$

Attack – Lattice Reduction

Figure out the lattice dimension m and omit the negligible terms

$$m = \sum_{k=0}^{s} \sum_{i=0}^{s-k} \sum_{l_1=0}^{s-k-i} \sum_{l_2=0}^{s-k-i-l_1} 1 + \sum_{l_1=0}^{s} \sum_{j=1}^{\tau l_1} \sum_{l_2=0}^{l_1} \sum_{k=0}^{l_2} 1 = \frac{1+3\tau}{24} s^4 + o(s^4)$$

Apply the crucial condition $\det(\mathcal{L}) < R^m$ with $R = E^s$

$$X^{s_x} Y^{s_y} Z^{s_z} W^{s_w} U^{s_u} E^{s_E} < E^{\frac{1+3\tau}{24} s^5}$$

Figure out contributions of diagonal entries: s_x, s_y, s_z, s_w, s_u and s_E

$$s_x = \frac{1}{120} s^5,\ s_y = \frac{\tau^2}{20} s^5,\ s_z = s_w = s_u = \frac{1+4\tau}{120} s^5,\ s_E = \frac{4+11\tau}{120} s^5$$

Attack – Condition Derivation

Figure out upper bounds on unknown variables: X, Y, Z, W, U and E

$$X = N^{\delta},\ Y = N^{\frac{1}{2}},\ Z = N^{\delta-\frac{1}{4}},\ W = N^{\beta},\ U = N^{\delta+\frac{1}{2}},\ E = N^{\frac{5}{4}}$$

Substitute in simplified condition
$X^{s_x} Y^{s_y} Z^{s_z} W^{s_w} U^{s_u} E^{s_E - \frac{1+3\tau}{24} s^5} < 1$

$$\delta + 3\tau^2 + (1+4\tau)\left(\delta - \frac{1}{4} + \beta + \delta + \frac{1}{2}\right) + \frac{5}{4}(4 + 11\tau - 5 - 15\tau) < 0$$

Obtain the inequality of δ with τ to be optimized later

$$\delta < \frac{(1-\beta)(1+4\tau) - 3\tau^2}{3 + 8\tau}$$

Attack – Result Illustration

Set $\tau = (\sqrt{177 - 96\beta} - 9)/24$ and obtain the final condition

$$\delta < \frac{25 - 16\beta - \sqrt{177 - 96\beta}}{32}$$

Main Result

Implicit Related-Key Factorization Attack

Let $N_1 = p_1 q_1$ and $N_2 = p_2 q_2$ be given two RSA moduli of the same bit-size, where p_1, q_1, p_2, q_2 are large primes of the same bit-size. Let e_1, d_1, e_2, d_2 be some integers satisfying $e_1 d_1 \equiv 1 \bmod (p_1 - 1)(q_1 - 1)$ and $e_2 d_2 \equiv 1 \bmod (p_2 - 1)(q_2 - 1)$ such that $e_1 \approx e_2 \approx N$ and $d_1 \approx d_2 \approx N^{\delta}$. Given the implicit information that $d_2 = d_1 + d_{21} D$ for $|d_{21}| \approx N^{\beta}$. Then N_1 and N_2 can be factored in polynomial time if

$$\delta < \frac{25 - 16\beta - \sqrt{177 - 96\beta}}{32}.$$

Concrete Attack Scenario

Perform the splitting technique using the following basis matrix

$$\begin{bmatrix} a_0 & 0 & \cdots & 0 & e_2 & \cdots & e_n \\ 0 & b_0 & \cdots & 0 & e_2 D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_0 & 0 & \cdots & e_n D \\ 0 & 0 & \cdots & 0 & N_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & N_n \end{bmatrix}$$

d_1 is an integer linear combination of $(2n-1)$ unknown variables

$$d_1 = a_1 c_1 + a_2 c_2 + \cdots + a_{2n-1} c_{2n-1}$$

Main Result

Implicit Related-Key Factorization Attack

Let $N_i = p_i q_i$ for $1 \leq i \leq n$ be given RSA moduli of the same bit-size, where p_i and q_i are large primes of the same bit-size. Let e_i and d_i be some integers satisfying $e_i d_i \equiv 1 \bmod (p_i - 1)(q_i - 1)$ such that $e_i \approx N$ and $d_i \approx N^{\delta}$. Given the implicit information that $d_j = d_i + d_{ji} D$ for $1 \leq i < j \leq n$ with $|d_{ji}| \approx N^{\beta}$. Then given RSA moduli can be factored in polynomial time (but exponential in n) if

$$\delta < \frac{1}{2} - \beta + \frac{2n^2 + n - 1 + 4n^2\beta}{4n^3} - \frac{\sqrt{(2n-1)(6n^3 + 3n^2 - 1 - 8n^2(n-1)\beta)}}{4n^3}.$$

Experimental Results

Randomly generate two RSA instances with implicitly related keys

1. Generate two 1024-bit RSA moduli N_1 and N_2
2. Generate implicit related-keys d_1 and d_2 according to β and γ
3. Compute respective public keys e_1 and e_2 from N_1, d_1 and N_2, d_2

Table: Asymptotic bounds and experimental results for given two RSA instances

γ	β	δ_∞	δ_e	s	τ	m	TL	TG
0.117	0.048	0.346	0.292	5	0.200	136	112.834	0.121
0.117	0.039	0.350	0.315	6	0.166	225	1933.569	0.151
0.043	0.058	0.342	0.295	5	0.200	136	140.853	0.122
0.034	0.063	0.340	0.296	6	0.166	225	1647.016	0.176
0.078	0.092	0.329	0.290	5	0.200	136	174.732	0.146
0.078	0.097	0.327	0.293	6	0.166	225	2336.019	0.162

Outline

Conclusion

Focus on a new problem concerning implicitly related keys

- Factor RSA moduli with the help of given implicit relation
- Apply lattice-based method for solving modular equations
- Propose lattice-based implicit related-key factorization attacks
- Verify the validity of our proposed attack by numerical experiments

Further improvements remain as future works

- More efficient lattice constructions
- Generic attacks for arbitrary public keys

THANK YOU FOR YOUR ATTENTION!

Q & A