Improved Factoring Attacks on Multi-prime RSA with Small Prime Difference

Mengce Zheng[1,2] Noboru Kunihiro[2] Honggang Hu[1]

[1]University of Science and Technology of China, China

[2]The University of Tokyo, Japan

ACISP2017 June 27, 2017

Outline

RSA and Multi-prime RSA

The RSA cryptosystem:

- $N = pq$ with two distinct prime factors of the same bit-size.
- (e, d) satisfy $ed \equiv 1 \bmod \varphi(N)$, where $\varphi(N) = (p-1)(q-1)$.
- Encryption: $C = M^e \bmod N$.
- Decryption: $M = C^d \bmod N$.

RSA and Multi-prime RSA

The RSA cryptosystem:

- $N = pq$ with two distinct prime factors of the same bit-size.
- (e, d) satisfy $ed \equiv 1 \bmod \varphi(N)$, where $\varphi(N) = (p-1)(q-1)$.
- Encryption: $C = M^e \bmod N$.
- Decryption: $M = C^d \bmod N$.

The multi-prime RSA cryptosystem:

- The modulus is modified as the product of $r (\geq 3)$ primes.
- $N = p_1 p_2 \cdots p_r$ with r distinct prime factors of the same bit-size.
- (e, d) satisfy $ed \equiv 1 \bmod \varphi(N)$, where $\varphi(N) = \prod_{i=1}^{r}(p_i - 1)$.

Prime Difference

Let $\Delta = |p - q|$ be the prime difference of the original RSA scheme.

- Though Δ is close to $N^{\frac{1}{2}}$, there still exist enhanced attacks.
- It is used to enhance small private exponent attack on RSA.[DW02]

Prime Difference

Let $\Delta = |p - q|$ be the prime difference of the original RSA scheme.

- Though Δ is close to $N^{\frac{1}{2}}$, there still exist enhanced attacks.
- It is used to enhance small private exponent attack on RSA.[DW02]

Let $\Delta = \max_{i,j \in \{1,2,\ldots,r\}} |p_i - p_j|$ for the multi-prime RSA scheme.

- It is denoted by N^{γ} for $0 < \gamma < 1/r$.
- The maximal value of difference between every two prime factors.
- It also enhances several attacks on multi-prime RSA.

Main Problem: Factoring with Small Prime Difference

Factoring attack can remove the restriction on the private exponents.

N can be factored under what condition when we are given

- a multi-prime RSA modulus N,
- the number of prime factors r,
- the small prime difference N^{γ}.

Main Problem: Factoring with Small Prime Difference

Factoring attack can remove the restriction on the private exponents.

N can be factored under what condition when we are given

- a multi-prime RSA modulus N,
- the number of prime factors r,
- the small prime difference N^{γ}.

The multi-prime modulus can be factored in polynomial time if $\gamma < \frac{1}{r^2}$.

- Let $p = [N^{\frac{1}{r}}]$ and $x_i = p_i - p$.
- Solve the univariate equation $x_i + p = 0 \bmod p_i$ for $|x_i| < N^{\gamma}$.[ZT13]

Formulation of Main Problem

Find all solutions of the following simultaneous equations.

$$\begin{cases} y_1 + p = 0 \bmod p_1, \\ y_2 + p = 0 \bmod p_2, \\ \qquad \vdots \\ y_r + p = 0 \bmod p_r. \end{cases}$$

- Given N, r and γ.
- Let $p = [N^{\frac{1}{r}}]$ and $y_i = p_i - p$.
- $|y_i| < N^\gamma$ for $1 \le i \le r$.

The factoring problem is similar to multi-prime Φ-hiding problem.[KOS10]

Lattice Based Method

Recover small roots of modular equations using lattice reduction algorithm.

- Construct a set of shift polynomials sharing the common roots,
- Transform polynomials' coefficients into a lattice basis matrix,
- Compute short lattice vectors by the LLL algorithm,
- Transform lattice vectors into equations over the integers,
- Solve the desired roots by Gröbner basis computations.

Tool Used in Our Methods

All solutions of the linear equation can be found.[HM08]

$$\sum_{i=1}^{n} \eta_i < 1 - (n+1)(1-\beta) + n(1-\beta)^{\frac{n+1}{n}}$$

- $a_1 x_1 + \cdots + a_n x_n + a_{n+1} = 0 \bmod p$.
- $a_1, \ldots, a_n$ and a_{n+1} are some integers.
- p $(\geq N^{\beta})$ is a divisor of N.
- N is a known large composite integer (of unknown factorization).
- Solutions $(x_1^{(0)}, \ldots, x_n^{(0)})$ satisfy $|x_i^{(0)}| \leq N^{\eta_i}$.

The time complexity is polynomial in $\log N$ and exponential in n.

Our Ideas

Solve an r-variate equation instead of the univariate equation.

- Using r equations is better than using only one equation.
- Combining all equations together provides an r-variate equation.
- However, the time complexity is exponential in r.

Our Ideas

Solve an r-variate equation instead of the univariate equation.

- Using r equations is better than using only one equation.
- Combining all equations together provides an r-variate equation.
- However, the time complexity is exponential in r.

Solve an l-variate equation by the optimal linearization technique.

- Using k $(2 \leq k \leq r-1)$ equations will provide better bound.
- Recover the modulus rather than the unknown variables.
- Apply the optimal linearization technique [TK12] for l variables.
- The consumption is lower and time complexity is polynomial in r.

Our Results

The multi-prime modulus can be factored in polynomial time if

- for $r \leq 6$,

$$\gamma < \frac{2}{r(r+1)}$$

Our Results

The multi-prime modulus can be factored in polynomial time if

- for $r \leq 6$,

$$\gamma < \frac{2}{r(r+1)}$$

- for $r \geq 7$ with an optimal l,

$$\gamma < \frac{2}{l+1}\left(\frac{1}{r}\right)^{\frac{l+1}{l}}$$

Our Results

The multi-prime modulus can be factored in polynomial time if

- for $r \le 6$,

$$\gamma < \frac{2}{r(r+1)}$$

- for $r \ge 7$ with an optimal l,

$$\gamma < \frac{2}{l+1}\left(\frac{1}{r}\right)^{\frac{l+1}{l}}$$

- for much larger r,

$$\gamma < \frac{2}{\mathrm{e}r(\log r + 1)}$$

Notations Used in Our Methods

$p = [N^{\frac{1}{r}}]$ denotes the value of rounding $N^{\frac{1}{r}}$ to the nearest integer.

Elementary symmetric polynomial in k variables $y_1, \ldots, y_k$ of degree i.

$$\sigma_i^k = \sum_{\substack{|\lambda|=i \\ \lambda \subset \{1,\ldots,k\}}} \left(\prod_{j \in \lambda} y_j \right)$$

Notations Used in Our Methods

$p = [N^{\frac{1}{r}}]$ denotes the value of rounding $N^{\frac{1}{r}}$ to the nearest integer.

Elementary symmetric polynomial in k variables $y_1, \ldots, y_k$ of degree i.

$$\sigma_i^k = \sum_{\substack{|\lambda|=i \\ \lambda \subset \{1,\ldots,k\}}} \left(\prod_{j \in \lambda} y_j \right)$$

Q_k denotes the product of k distinct prime factors chosen from $p_1, \ldots, p_r$.

Q'_k denotes the numerical value of the left side of the modular equation.

The Direct Method (1)

Let e be the inverse of p modulo N, that is $ep = 1 \bmod N$.

$$y_i + p = 0 \bmod p_i \quad \rightarrow \quad ey_i + 1 = 0 \bmod p_i$$

Collect the modular equations as many as possible.

$$\begin{cases} ey_1 + 1 = 0 \bmod p_1, \\ \qquad \vdots \\ ey_r + 1 = 0 \bmod p_r. \end{cases}$$

The Direct Method (2)

Combine all equations together by multiplication.

$$
\begin{aligned}
\prod_{i=1}^{r}(ey_i+1) = 0 \bmod N \quad & \to \quad \sum_{i=1}^{r} e^i \sigma_i^r + 1 = 0 \bmod N \\
& \to \quad \sum_{i=1}^{r} e^i \sigma_i^r + ep = 0 \bmod N \\
& \to \quad \sum_{i=1}^{r} e^{i-1} \sigma_i^r + p = 0 \bmod N \\
& \to \quad e^{r-1}\sigma_r^r + \cdots + e\sigma_2^r + \sigma_1^r + p = 0 \bmod N
\end{aligned}
$$

The Direct Method (3)

$\sum_{i=1}^{n} \eta_i < 1$ for $n = r$ and $\eta_i = i\gamma$.

$$\sum_{i=1}^{r} i\gamma < 1 \quad \rightarrow \quad \gamma < \frac{2}{r(r+1)}$$

The Direct Method (3)

$\sum_{i=1}^{n} \eta_i < 1$ for $n = r$ and $\eta_i = i\gamma$.

$$\sum_{i=1}^{r} i\gamma < 1 \quad \rightarrow \quad \gamma < \frac{2}{r(r+1)}$$

After solving $e^{r-1}\sigma_r^r + \cdots + e\sigma_2^r + \sigma_1^r + p = 0 \bmod N$, we obtain

- $\sigma_1^r, \ldots, \sigma_r^r$,
- $x_1, \ldots, x_r$ by solving $x^r - \sigma_1^r x^{r-1} + \cdots + (-1)^r \sigma_r^r = 0$,
- $p_1, \ldots, p_r$ by computing $p_i = x_i + p$.

The direct method works in time polynomial in $\log N$ but exponential in r.

The Optimized Method (1)

Take fewer equations such as k $(2 \leq k \leq r-1)$ equations.

- The equation is $\prod_{i=1}^{k}(y_i + p) = 0 \bmod Q_k$.
- It is not necessary to know the values of $y_1, \ldots, y_k$.
- It is enough to know the numerical value of $\prod_{i=1}^{k}(y_i + p)$, Q'_k.
- Computing $\gcd(Q'_k, N)$ provides all products of k prime factors Q_k.
- Apply the optimal linearization technique for l $(2 \leq l \leq k)$ variables.

The advantage is lower consumption with fewer variables.

The Optimized Method (2)

Expand the product of k equations.

$$\prod_{i=1}^{k}(y_i + p) = 0 \bmod Q_k \quad \rightarrow \quad \sum_{i=0}^{k} p^i \sigma_{k-i}^k = 0 \bmod Q_k$$

$$\rightarrow \quad \sigma_k^k + p\sigma_{k-1}^k + \cdots + p^k = 0 \bmod Q_k$$

Search for the optimal linearization when it can be efficiently solved.

The Optimized Method (3)

Perform a linearization for the case of l $(2 \le l \le k)$ variables.

$$p^{k-t_1}u_1 + p^{k-t_2}u_2 + \cdots + p^{k-t_l}u_l + p^k = 0 \bmod Q_k$$

- Let $t_1, \ldots, t_{l+1}$ be integers satisfying $t_1 = k > t_2 > \cdots > t_{l+1} = 0$.
- $u_i = \sum_{j=t_{i+1}+1}^{t_i} p^{t_i - j}\sigma_j^k$ for $1 \le i \le l$.
- Apply theorem with $\beta = k/r$ and $\eta_i = (t_i - t_{i+1} - 1)/r + (t_{i+1}+1)\gamma$.

Obtain the condition with $\sum_{i=2}^{l} t_i$, k and l that are optimized later.

$$\gamma < \frac{l}{l + \sum_{i=2}^{l} t_i}\left(\frac{k+1}{r} + (1 - \frac{k}{r})^{\frac{l+1}{l}} - 1\right)$$

The Optimized Method (4)

Optimize $\sum_{i=2}^{l} t_i$ for $(t_1, t_2, t_3, \ldots, t_l) = (k, l-1, l-2 \ldots, 1)$.

$$\gamma < \frac{2}{l+1}\left(\frac{k+1}{r} + (1-\frac{k}{r})^{\frac{l+1}{l}} - 1\right)$$

The condition is further optimized by taking $k = r - 1$.

$$\gamma < \frac{2}{l+1}\left(\frac{1}{r}\right)^{\frac{l+1}{l}}$$

The Optimized Method (5)

The optimal value of l for each positive integer r (≤ 10).

- $l = 2$ for $r = 3, 4, 5$,
- $l = 3$ for $r = 6, 7, 8, 9, 10$.

Solve the following linear equation with an optimal l.

$$u_1 + p^{r-l}u_2 + \cdots + p^{r-2}u_l + p^{r-1} = 0 \bmod Q_{r-1}$$

For much larger r and $l \approx \log r$, the condition is approximated

$$\gamma < \frac{2}{er(\log r + 1)}$$

The Optimized Method (6)

After solving $u_1 + p^{r-l}u_2 + \cdots + p^{r-2}u_l + p^{r-1} = 0 \bmod Q_{r-1}$, we obtain

- $u_1, \ldots, u_l$,
- $Q'_{r-1} = u_1 + p^{r-l}u_2 + \cdots + p^{r-2}u_l + p^{r-1}$,
- Q_{r-1} by computing $\gcd(Q'_{r-1}, N)$,
- $p_1, \ldots, p_r$ by computing N/Q_{r-1}.

The optimized method works in time polynomial in $\log N$ and r.

Further Improvement

Applying better lattice constructions [TK13] since u_i are unbalanced.

$$u_1 + p^{r-2}u_2 + p^{r-1} = 0 \bmod Q_{r-1}$$

r	DM	OM	FI	ZT
3	**0.1666**	0.1283	—	0.1111
4	**0.1000**	0.0833	0.0835	0.0625
5	**0.0666**	0.0596	0.0608	0.0400
6	**0.0476**	0.0458	0.0474	0.0277
7	0.0357	0.0373	**0.0387**	0.0204
8	0.0277	0.0312	**0.0327**	0.0156
9	0.0222	0.0267	**0.0282**	0.0123
10	0.0181	0.0232	**0.0248**	0.0100

Experimental Results

The experiments for $r = 3$ with a 1536-bit multi-prime RSA modulus.

- The direct method performs better with similar lattice dimension.
- The optimized method runs much faster as predicted.

Experimental Results

The experiments for $r = 3$ with a 1536-bit multi-prime RSA modulus.

- The direct method performs better with similar lattice dimension.
- The optimized method runs much faster as predicted.

The experiments for $4 \leq r \leq 7$ with around 300-dimensional lattices.

- Use the optimized method since it is more efficient.
- Our results are superior to the previous experimental bounds.

r	4	5	6	7
OM	0.0750	0.0533	0.0337	0.0286
ZT	0.0620	0.0396	0.0275	0.0202

Conclusions

We propose improved factoring attacks on multi-prime RSA.

- Factoring attack works better with much smaller prime difference.
- Factoring attack removes the restriction on the private exponents.

Conclusions

We propose improved factoring attacks on multi-prime RSA.

- Factoring attack works better with much smaller prime difference.
- Factoring attack removes the restriction on the private exponents.

We use lattice based method to solve the factoring problem.

- Apply the optimal linearization technique to reduce the consumption.
- Obtain further improvement by better lattice constructions.
- Verify two methods by the experiments.

References I

[DW02] Benne De Weger.
Cryptanalysis of RSA with small prime difference.
Applicable Algebra in Engineering, Communication and Computing, 13(1):17–28, 2002.

[HM08] Mathias Herrmann and Alexander May.
Solving linear equations modulo divisors: On factoring given any bits.
In J. Pieprzyk, editor, *ASIACRYPT 2008*, volume 5350 of *LNCS*, pages 406–424. Springer, Heidelberg, 2008.

[KOS10] Eike Kiltz, Adam O'Neill, and Adam Smith.
Instantiability of RSA-OAEP under chosen-plaintext attack.
In T. Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 295–313. Springer, Heidelberg, 2010.

References II

[TK12] Kaori Tosu and Noboru Kunihiro.
Optimal bounds for multi-prime Φ-hiding assumption.
In Willy Susilo, Yi Mu, and Jennifer Seberry, editors, *ACISP 2012*, volume 7372 of *LNCS*, pages 1–14. Springer, Heidelberg, 2012.

[TK13] Atsushi Takayasu and Noboru Kunihiro.
Better lattice constructions for solving multivariate linear equations modulo unknown divisors.
In C. Boyd and L. Simpson, editors, *ACISP 2013*, volume 7959 of *LNCS*, pages 118–135, Heidelberg, 2013. Springer.

[ZT13] Hui Zhang and Tsuyoshi Takagi.
Attacks on multi-prime RSA with small prime difference.
In C. Boyd and L. Simpson, editors, *ACISP 2013*, volume 7959 of *LNCS*, pages 41–56, Heidelberg, 2013. Springer.