An algorithm to obtain an RSA modulus with a large private key

L. Hernández Encinas, J. Muñoz Masqué
and A. Queiruga Dios

Instituto de Física Aplicada, CSIC
C/ Serrano 144, 28006-Madrid, Spain
E-mails: {luis, jaime, araceli}@iec.csic.es

Abstract

Sufficient conditions are obtained on the prime factors of an RSA modulus in order to avoid Wiener and Boneh-Durfee attacks. The public exponent can be chosen arbitrarily.

Key words and phrases. Boneh-Durfee attack, Decryption exponent, Public key cryptography, RSA cryptosystem, Wiener attack.

Mathematics Subject Classification 2000. Primary 94A60; Secondary 14G50, 68P25.

1 Introduction

Let $n = pq$ be the modulus of a RSA cryptosystem with private key d and public exponent e. A classical attack to RSA ([11], also see [9]) shows that the cryptosystem becomes insecure if $d < \sqrt[4]{n}$. In the last years, several improvements to the RSA short decryption exponent attack have been obtained (*e.g.*, see [1]). In the Boneh-Durfee attack ([2, 3]) it is shown that RSA is insecure if the decryption exponent $d < n^{0.292}$. Furthermore, the authors conjectured that, indeed, this cryptosystem is insecure for $d < \sqrt{n}$. Moreover, Sun *et al.* ([8]) have proposed three variants of the RSA with small private keys for resisting the Boneh-Durfee attack. These variants suggest to use unbalanced factor primes p, q of the RSA modulus. Nevertheless, Durfee and Nguyen ([4]) have broken two of these three new proposals. More recently, Weger ([10]) has proved that if the prime numbers p, q are chosen in such a way that its difference $|p - q|$ is small enough, then one obtains improvements on the Wiener and Boneh-Durfee attacks.

These cryptanalyses have increased the interest of using decryption exponents as large as possible, in order to avoid such attacks, although then the decryption process is slower. In fact, it is usually recommended that the size of the decryption exponent d must be almost equal to the RSA modulus n (*e.g.*, see [5, §8.2.2], [7, §12.4]).

In this communication we show that, once the public exponent e has been chosen, we can select the prime factors p, q in such a way that the decryption exponent d has a bitlength almost equal to the bitlength of the RSA modulus. Hence we obtain sufficient conditions on p and q which prevents the cryptosystem against the aforementioned attacks.

2 Basic results

Proposition 1 *Assume d, e are the decryption and encryption exponents respectively of the RSA cryptosystem with modulus $n = pq$. If $de = 1+k\phi(n)$, then $k < e$. If $k = e - 1$, then*

$$d \geq \frac{2}{3}\phi(n).$$

Proof. If e denotes the encryption exponent, then there exists a positive integer k such that $de = 1 + k\phi(n)$. As

$$\frac{1 + k\phi(n)}{e} = d < \phi(n),$$

we have

$$k < \frac{\phi(n)e - 1}{\phi(n)} = e - \frac{1}{\phi(n)} < e.$$

Furthermore, as $e \geq 3$, for $k = e - 1$ we obtain

$$\begin{aligned} d &= \frac{1 + (e-1)\phi(n)}{e} \\ &= \frac{1}{e} + \left(1 - \frac{1}{e}\right)\phi(n) \\ &> \left(1 - \frac{1}{e}\right)\phi(n) \\ &\geq \frac{2}{3}\phi(n). \end{aligned}$$

■

Hence, when k is as large as possible, the bitlength of d can be bounded as follows:

Corollary 2 *With the same hypotheses as in* Proposition 1*, we have*

$$\text{bitlength}(d) \geq \text{bitlength}(n) - 1.$$

Proof. As $d \geq \frac{2}{3}\phi(n)$, from the very definition of the Euler function we deduce,

$$\begin{aligned} \lfloor \lg d \rfloor &> \lfloor \lg(p-1) + \lg(q-1) + \lg(2/3) \rfloor \\ &\geq \lfloor \lg(p-1) \rfloor + \lfloor \lg(q-1) \rfloor - 1 \\ &\geq \lfloor \lg n \rfloor - 2. \end{aligned}$$

Accordingly, $\text{bitlength}(d) = \lfloor \lg d \rfloor + 1 > \lfloor \lg n \rfloor - 1 = \text{bitlength}(n) - 2$. ■

Below, we analyze the conditions that p, q must satisfy for the equation $k = e - 1$ to hold. Then, Proposition 1 will ensure that the bitlength of the decryption exponent is almost the same than that of n.

Proposition 3 *Let d, e be as in* Proposition 1*, where we further assume $k = e - 1$. Let $r_p = p \pmod e$ and $r_q = q \pmod e$ be the residues of p and q, respectively modulo e, and let S_e be the set of elements $r \in \mathbb{Z}_e$ such that $r(r-1)$ is invertible modulo e; that is, $r(r-1) \in \mathbb{Z}_e^*$.*

(i) *We have $r_p, r_q \in S_e$. Hence $r_p, r_q \in \mathbb{Z}_e^* - \{0, 1\}$.*

(ii) *In $\mathbb{Z}_e$, we have*

$$r_q = \frac{r_p}{r_p - 1}. \tag{1}$$

(iii) *If $e = p_1^{m_1} \cdots p_t^{m_t}$ is the prime factorization of e, then*

$$\#S_e = \prod_{i=1}^{t} (p_i - 2) p_i^{m_i - 1}.$$

Conversely, if p and q are arbitrary primes satisfying (ii), then $k = e - 1$.

Proof.

(i) Assume $p = ec_p + r_p$, where $c_p \in \mathbb{N}$. Then $r_p \in \mathbb{Z}_e^*$, because p is a prime number. We also have $p-1 = ec_p + r_p - 1$ and $\gcd(e, p-1) = 1$. Hence we conclude $\gcd(e, r_p - 1) = 1$.

(ii) Taking account of the identities

$$\begin{aligned} d &= \frac{1 + (e-1)\phi(n)}{e} \\ &= \phi(n) - \frac{\phi(n) - 1}{e}, \end{aligned}$$

it follows that e divides $\phi(n) - 1 = pq - p - q$; hence $(r_p - 1)r_q = r_p$ in $\mathbb{Z}_e$.

(iii) The formula follows taking into account that an integer s is invertible in $\mathbb{Z}/p^m$ if and only if it is invertible in $\mathbb{Z}/p$.

Conversely, if p and q satisfy (ii), then

$$q(p-1) \equiv p (\mathrm{mod}\, e),$$

or equivalently,

$$n \equiv p + q (\mathrm{mod}\, e);$$

say,

$$n = p + q + he,$$

for an integer $h > 0$.

Then, we have

$$\begin{aligned} f &= \frac{1 + (e-1)\phi(n)}{e} \\ &= \phi(n) - \frac{\phi(n) - 1}{e} \\ &= \phi(n) - h. \end{aligned}$$

Hence, f is an integer, and $ef \equiv 1 (\mathrm{mod}\, \phi(n))$. Therefore, $d = f$ and $k = e - 1$. ■

The next corollary measures the percent of elements in $\mathbb{Z}_e$ that belong to S_e.

Corollary 4 *Letting $N_e = \#S_e$, we have*

$$\frac{N_e}{\phi(e)} = \prod_{i=1}^{t} \frac{p_i - 2}{p_i - 1}.$$

Hence, if every prime factor of e goes to ∞, then the probability of finding a residue in S_e goes to 1.

Some simple consequences of the previous results are the following:

1. Note that by virtue of the Theorem of Dirichlet ([6, Chapter 2]) the primes p of the form $p = ec_p + r_p$ are approximately equi-distributed among the series $ec_p + r_p$ for a fixed c_p, and that every arithmetic series $ec_p + r_p$ with $\gcd(c_p, r_p) = 1$ contains infinitely many primes.

2. If e is a prime, then $N_e = e - 2$.

3. If $e = 3^m$, then
$$\frac{N_e}{\phi(e)} = \frac{1}{2}, \tag{2}$$
and if $e = l^m$, where $l > 3$ is a prime, then
$$\frac{N_e}{\phi(e)} > \frac{1}{2}. \tag{3}$$

4. Assume $t \geq 2$ and that the least prime factor of e is greater than 3; that is, $3 < p_1 < \ldots < p_t$. Then, we have
$$\frac{N_e}{\phi(e)} > \frac{1}{2}. \tag{4}$$

In fact, if we set

$$\begin{aligned} \varepsilon_i &= \frac{1}{p_i - 1}, \quad 1 \leq i \leq t, \\ f(x) &= (x - \varepsilon_1) \cdots (x - \varepsilon_t), \end{aligned}$$

then it is easily seen that (4) is equivalent to saying $2f(1) - 1 > 0$, and this inequality can be proved in two steps,

(a) If $p_1 = 5$, $p_2 = 7$, and $p_i \geq 11$, $3 \leq i \leq t$, the result follows from a direct computation, and

(b) If $p_i \geq 13$ for $3 \leq i \leq t$, the result follows by simply bounding $\varepsilon_1 + \ldots + \varepsilon_t$.

5. If $t \geq 2$ and $p_1 = 3$, then $N_e/\phi(e)$ is slightly less than $\frac{1}{2}$.

3 Algorithm

From the previous results, an algorithm to generate the keys for the RSA and modifying the algorithm proposed in [5, § 8.2.1] is as follows:

1. Choose the encryption exponent $e > 2$.

2. Generate a large random prime p such that $r_p(r_p - 1) \in \mathbb{Z}_e^*$, where $r_p = p (\mathrm{mod}\, e)$.

3. Compute a large prime number $q = r_p \cdot (r_p - 1)^{-1} (\mathrm{mod}\, e) + k \cdot e$, for some k.

4. Compute $n = p \cdot q$, $\phi = (p - 1)(q - 1)$, and verify that $1 < e < \phi$ and $\gcd(e, \phi) = 1$.

5. Use the extended Euclidean algorithm to compute the unique integer d, $1 < d < \phi$, such that $e \cdot d \equiv 1 \,(\mathrm{mod}\, \phi)$, or compute directly $d = (1 + (e-1)\,\phi)\,/e$.

4 Conclusions

The probability that a random chosen d would be sufficiently large is high, but, then the probability of the corresponding encryption exponent e would be large is also high. In practice, however, the exponent e is wanted to belong to a small predetermined set of values; for example, 3 or $2^{16}+1$, in order to improve the efficiency of the encryption process. The interest of the algorithm above is that encryption exponents whose associated decryption exponent is of the same size as the RSA modulus, can be chosen arbitrarily. Once the encryption exponent e is selected, the only constraints for p and q are the items (i) and (ii) in the Proposition 3. Such conditions only affect the residues of p and q modulo e.

As the map

$$x \mapsto \frac{x}{x-1}$$

is an involution, the residues r_p and r_q play a completely symmetric role. Hence every $r_p \in S_e$ corresponds to a unique $r_q \in S_e$ such that (r_p, r_q) satisfies the items in Proposition 3. For example, $r_p = 2$ if and only if $r_q = 2$. This is the only case in which $r_p = r_q$.

If e is a prime number, then p can be chosen arbitrarily and q is only conditioned to satisfy the equation (1). We remark that N_e is as large as possible when e is a prime. On the other hand, if e is a composite number, then N_e may be much smaller than $e-2$, thus imposing more constrains to select p and q than in the prime case.

If $e = 3$, there are no constrain on p and q, as in this case, we have $r_p = r_q = 2$ for every pair p, q satisfying the RSA conditions.

By setting $k = e - 1$ does not provide a factoring advantage to an attacker, as the prime factors, p, q, can still be taken at random as follows from the formulas (2)-(4).

Finally, as the algorithm above shows, choosing r_p and r_q in S_e does not increase the running time for key generation, as the modification runs in polynomial time.

Acknowledgement. Supported by Ministerio de Ciencia y Tecnología of Spain, under grant TIC2001–0586.

References

[1] D. Boneh, Twenty years of attacks on the RSA cryptosystem, *Notices Amer. Math. Soc.* **46**, 2 (1999), 203–213.

[2] D. Boneh and D. Durfee, Cryptanalysis of RSA with private key d less than $N^{0.292}$, Advanced in Cryptology-EUROCRYPT'99, LNCS **1592** (1999), 1–16.

[3] D. Boneh and D. Durfee, Cryptanalysis of RSA with private key d less than $N^{0.292}$, *IEEE Trans. Inform. Theory* **46**, 4 (2000), 1339–1349.

[4] G. Durfee and P.Q. Nguyen, Cryptanalysis of the RSA schemes with short secret exponent from Asiacrypt'99, Proceedings of Asiacrypt'00, LNCS **1976** (2000), 14–29.

[5] A. Menezes, P. van Oorschot, and S. Vanstone, Handbook of applied cryptography, CRC Press, Boca Raton, FL., 1997.

[6] H. Riesel, Prime numbers and computer methods for factorization, Birhäuser, Boston, 1994.

[7] B. Schneier, Applied cryptography, John Wiley & Sons, New York, 1994.

[8] H.M. Sun, W.C. Yang, and C.S. Laigh, On the desgin of RSA with short secret exponent, Proceedings of Asiacrypt'99, LNCS **1716** (1999), 150-164.

[9] E.R. Verheul and H.C.A. van Tilborg, Cryptanalysis of 'less short' RSA secret exponents, *Appl. Algebra Engrg. Comm. Comput.* **8** (1997), 425–435.

[10] B. de Weger, Cryptanalysis of RSA with small prime difference, *Appl. Algebra Engrg. Comm. Comput.* **3** (2002), 17–28.

[11] M.J. Wiener, Cryptanalysis of short RSA secret exponents, *IEEE Trans. Inform. Theory* **36**, 3 (1990), 553–558.