Attacks on Shamir's 'RSA for paranoids'

Henri Gilbert
France Télécom
CNET DTL SSR
38-40 Rue du Général Leclerc
92131 Issy les Moulineaux
France
henri.gilbert@cnet.francetelecom.fr

Dipankar Gupta
Hewlett-Packard Laboratories
Filton Road, Stoke Gifford
Bristol BS12 6QZ
United Kingdom
Dipankar_Gupta@hp.com

Andrew Odlyzko
AT&T Labs - Research
Florham Park, NJ 07932
USA
amo@research.att.com

Jean-Jacques Quisquater
UCL Crypto Group
Université catholique de Louvain
Place du Levant, 3
B-1348 Louvain-la-Neuve
Belgium
quisquater@dice.ucl.ac.be

August 27, 1998

ABSTRACT

In order to allow for efficient use of extremely large moduli, Adi Shamir has proposed a variant of RSA in which one of the two prime factors is much smaller than the other. This note points out that unless special precautions are taken, simple implementations of Shamir's idea are subject to protocol attacks that recover the secret keys.

Attacks on Shamir's 'RSA for paranoids'

H. Gilbert
France Télécom
CNET DTL SSR
38-40 Rue du Général Leclerc
92131 Issy les Moulineaux
France
henri.gilbert@cnet.francetelecom.fr

D. Gupta
Hewlett-Packard Laboratories
Filton Road, Stoke Gifford
Bristol BS12 6QZ
United Kingdom
Dipankar_Gupta@hp.com

A. M. Odlyzko
AT&T Labs - Research
Florham Park, NJ 07932
USA
amo@research.att.com

J.-J. Quisquater
UCL Crypto Group
Université catholique de Louvain
Place du Levant, 3
B-1348 Louvain-la-Neuve
Belgium
quisquater@dice.ucl.ac.be

1. Introduction

The most popular public key cryptosystems rely for their presumed security on the difficulty of factoring integers or computing discrete logarithms in finite fields. Unfortunately for cryptosystem designers, advances in computational number theory and in the availability of computing power are rapidly increasing the sizes of moduli that are required for safety (see [Odlyzko], for example). In many applications, this is not a major problem, since computational resources available for legitimate encryption and decryption are also increasing rapidly. In some cases, though, such as those of smart cards and mobile wireless devices, power and size constraints do limit the amount of computing that can be done, and this often forces designers to use uncomfortably large moduli.

Shamir [Shamir] has proposed a variant of the RSA cryptosystem that allows for the use of large moduli while keeping computing requirements substantially lower than those in conventional RSA. Usually one chooses the public modulus n to be the product of two roughly equal primes p and q, $n = p \cdot q$. The public encryption exponent e is then chosen so that $\gcd(e, (p-1)(q-1)) = 1$ (often e is chosen relatively small, say $e = 3$ or $e = 2^{16} + 1$, in order to speed up the encryption operation), and the secret decryption exponent d is then computed from

$$de \equiv 1 \mod (\mathrm{lcm}(p-1, q-1)) \ . \tag{1.1}$$

A plaintext m, represented as an integer $\in [0, n-1]$ is transformed into the ciphertext c by

$$c \equiv m^e (\mathrm{mod}\ n) \ , \tag{1.2}$$

and the decryption operation is

$$m \equiv c^d (\mathrm{mod}\ n) \ . \tag{1.3}$$

The decryption operation (1.3) is frequently speeded up by computing

$$m_1 \equiv c^{d_1} (\mathrm{mod}\ p) \ , \tag{1.4}$$

$$m_2 \equiv c^{d_2} (\mathrm{mod}\ q) \ , \tag{1.5}$$

where

$$d_1 \equiv d\ (\mathrm{mod}\ p-1), \quad 0 < d_1 < p-1 \ , \tag{1.6}$$

$$d_2 \equiv d\ (\mathrm{mod}\ q-1), \quad 0 < d_2 < q-1 \ , \tag{1.7}$$

and m is recovered from m_1 and m_2 via an easy Chinese Remainder Theorem (CRT) operation.

Shamir's observation was that if $0 \leq m < p$, it is unnecessary to compute m_2, since $m = m_1$. RSA is usually used to convey only short messages (such as keys for symmetric cryptosystems or authentication information), so limiting m to the range $[0, p-1]$ is not a serious constraint. Since the legitimate recipient only needs to carry out operations modulo p, the other prime q can be chosen large enough to prevent attacks by general integer factorization algorithms. For example, for extremely cautious users, Shamir suggests [Shamir] choosing p of 500 bits and q of 4500 bits. (Given the current state of the art in integer factorization, and especially the lack of progress in algorithms that take advantage of existence of small prime divisors, such moduli appear to offer about as much security in the long run as 5000-bit integers composed of two 2500-bit primes, cf.

[Odlyzko, Shamir]. The current record for the largest prime found by the elliptic curve factoring method is 48 decimal digits.) The ciphertext c is then 5000 bits long, but it can easily be reduced modulo p, and the decryption operation (1.4) involves exponentiating a 500-bit integer to a 500-bit exponent modulo a 500-bit integer. For less sensitive communication, one might replace 500 by 300, say. (Shamir's paper also discusses some additional optimization. For example, it presents methods for reduction in sizes of public key directories through special choices of p and q. Some additional modifications of Shamir's idea have been suggested by Jim Reeds and the third author of this note. For example, for situations where the decryption effort has to be minimized as much as possible, it is possible to choose the decryption exponent so that it is small modulo $p-1$. Cf. [Wiener].)

Shamir proposed using e in the range $20 < e < 100$ for p of 500 bits and q of 4500 bits in order to lower the computational burden of encryption. There are dangers in case of small exponents in RSA, but they are understood [Hastad], and with proper precautions one can even use $e = 3$ with conventional RSA, as proposed originally by Knuth. However, with all variants of RSA, it is important that the encryption operation should involve modular reduction, to avoid the occurrence of $c = m^e$ (as would happen for $e < 10$ with p of 500 bits and q of 4500 bits in Shamir's scheme).

Shamir's scheme offers substantial computational advantages. Even for p of 256 bits and q of 768 bits, there is a 16-fold speedup over conventional RSA with p and q of 512 bits each.

In this note we point out that Shamir's scheme is insecure if used in some common modes. We also show how to protect against such protocol failures.

2. Protocol attacks

Suppose that n and e are Bob's public key. If ordinary RSA protocols are used, and Alice can get Bob to send her a decryption of a ciphertext, she can recover the secret p and q. What Alice does is to encrypt a message m which is bigger than p and send the ciphertext $m^e \bmod n$ to Bob. Bob decrypts the ciphertext to get $\widetilde{m}$. As $m > p$, $\widetilde{m} \neq m$. Now, if Alice can access $\widetilde{m}$, she knows that $p \mid (m - \widetilde{m})$. We also know (using Chinese Remainder theorem) that $q \nmid (m - \widetilde{m})$ under the constraints $0 < m, \widetilde{m} < n$ and $m \neq \widetilde{m}$. Alice can therefore recover p by computing $\gcd(m - \widetilde{m}, n)$. Note that under ordinary circumstances Bob might easily be tempted to send $\widetilde{m}$ to Alice (as in "What is this junk $\widetilde{m}$ you have just sent me?"). (See [JoyeQ] for another instance where such as attack is used.) The possibility of revealing $\widetilde{m}$ are increased if Bob is a smart card, say, which automatically outputs $\widetilde{m}$ as a cryptographic key to be used for further communications.

Adi Shamir has pointed out (personal communication) that the above attack was also observed by him and several other people. He also has noted that it is similar to the known ciphertext attack on Rabin's variant of RSA. The standard defence against such attacks is to require redundancy in the message, such as appearance of a specified bit pattern in the message m. If $\widetilde{m}$ does not contain such a pattern, it has to be rejected.

We next show that it does not suffice to add just any redundancy to the message m. At the very least, it is necessary to specify a substantial number of the high order bits of m. Otherwise, even if Bob does not respond to decrypted messages $\widetilde{m}$ that do not make sense, he can still betray his secret keys through his action. For example, suppose that Alice creates m by taking a message of the sort "Pay $101.74 to the order of Bob from my account no. 123 at Bank of Podunk" (including her digital signature) and pad it to create a plaintext m that is of about the suspected size of p. Then, if $m > p$, $\widetilde{m}$ will be discarded by a cautious Bob who has been indoctrinated in proper use of the cryptosystem. However, if $m < p$, then $\widetilde{m} = m$, and Bob will receive a legitimate check which he will surely be tempted to cash. Once he does that, though, Alice will know that $m < p$. Further probes can then reveal additional bits of p.

In the attack presented above, the cost per bit of p can be reduced, even if Alice thinks that $100.00 is the least amount that is guaranteed to make Bob cash a check. For example, suppose that Alice has determined through prior probes that $6 * 2^k < p < 7 * 2^k$. In the standard binary search, Alice would select $\widetilde{m}$ close to $13 * 2^{k-1}$. Instead, she can first try $\widetilde{m}$ close to $111 * 2^{k-4}$. If this probe reveals that $p < 111 * 2^{k-4}$, she can then try $\widetilde{m}$ close to $110 * 2^{k-4}$, and so on. Thus Alice will learn four bits of p instead of one at the expense of at most one single cashed check. This approach does require more probes than the standard one, though, and so risks arousing Bob's suspicions.

The above attack might seem fanciful, but similar attacks could arise in any situation where Bob will behave differently depending on whether $m > p$ or not. For instance, if Shamir's unbalanced RSA is used for session key exchange, Alice's attack can consist of providing Bob with $m^e \bmod n$ and testing whether Bob then encrypts session messages with m.

Our attacks do not break Shamir's scheme completely, but they do point out the need for carefully introduced redundancy in plaintexts, and in ensuring that recipients never reveal any decrypted messages.

References

[Hastad] J. Hastad, On using RSA with low exponent in a public key network, *Advances in Cryptology — CRYPTO '85*, H. C. Williams, ed., Lecture Notes in Comp. Sci #218, Springer, 1986, pp. 403–408.

[JoyeQ] M. Joye and J.-J. Quisquater, On the importance of securing your bins: The garbage-man-in-the-middle attack, *4th ACM Conf. Computer Comm. Security,* T. Matsumoto, ed., ACM Press, 1997, pp. 135–141.

[Odlyzko] A. M. Odlyzko, The future of integer factorization, *CryptoBytes* (*The technical newsletter of RSA Laboratories*), vol. 1, no. 2 (1995), pp. 5–12. Available at ⟨http://www.rsa.com/rsalabs/cryptobytes/⟩ and at ⟨http://www.research.att.com/∼amo⟩.

[Shamir] A. Shamir, RSA for paranoids, *CryptoBytes* (*The technical newsletter of RSA Laboratories*), vol. 1, no. 3 (1995), pp. 1,3, and 4. Available at ⟨http://www.rsa.com/rsalabs/cryptobytes/⟩.

[Stinson] D. R. Stinson, *Cryptography: Theory and Practice*, CRC Press, 1995.

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