An Overview of RSA and OAEP Padding

Yutong Zhong

The Affiliated International School of Shenzhen University,Shenzhen, Guangdong, China

Abstract. Encryption is crucial in information communication. The secret data is transformed into secure form and transferred through various channels. It is important for encryption to prevent unauthorized access to data and the encrypted data can't be cracked easily. The RSA algorithm was released by Rivest, Shamir and Adleman in 1978. It was the first practical asymmetric cryptosystem and now it is the most widely used asymmetric cryptosystem in the world, covering security of almost everything such as cellphone communication to online banking. In this article, we review the RSA algorithm and the padding schemes used with RSA encryption to avoid semantical insecurity such as optimal asymmetric encryption padding (OSAP).

Keywords: RSA, OAEP, Encryption, Cryptosystem.

1. Introduction

The RSA algorithm was released by Rivest, Shamir and Adleman in 1978. It was the first practical asymmetric cryptosystem and now it is the most widely used asymmetric cryptosystem in the world, covering security of almost everything such as cellphone communication to online banking. RSA uses two sets of keys for encryption and decryption separately, which are named public key and private key. Data is encrypted by public key and decrypted by private key through various channels like internet. The security of RSA depends on the difficulty to factorize a big integer into two large prime numbers. So far, the size of RSA keys typically ranges from 1024 to 4096. The record of the greatest size of key which has been broken is 829 bits. However, the textbook RSA is a deterministic encryption algorithm and not semantically secure, which means the attacker can distinguish two ciphertexts while the attacker knows the corresponding plaintexts.

Nowadays there are plenty of cryptosystems, which can be mainly divided into two categories: "symmetric and asymmetric". For the symmetric encryption, the simplest form is shown in Figure 1 below. Only one secret key is used for encryption and decryption of data. The secret key is only known to the sender and the recipient.

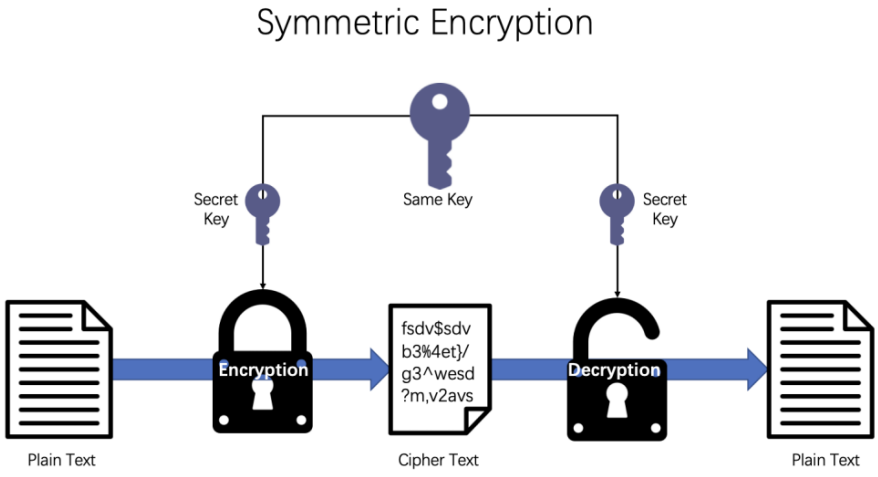

Figure 1. Symmetric Encryption

For Asymmetric Encryption, two different keys are involved. A public key is used to encrypt the plaintext while a secret one is used to decrypt the plaintext. The public key can be made public to anyone who wants to send a secure message to the recipient while the secret key must be kept secret by the recipient. The process for asymmetric encryption is shown in Figure 2. The concept of asymmetric encryption was raised by Diffie, W., and Hellman, M. in 1976.

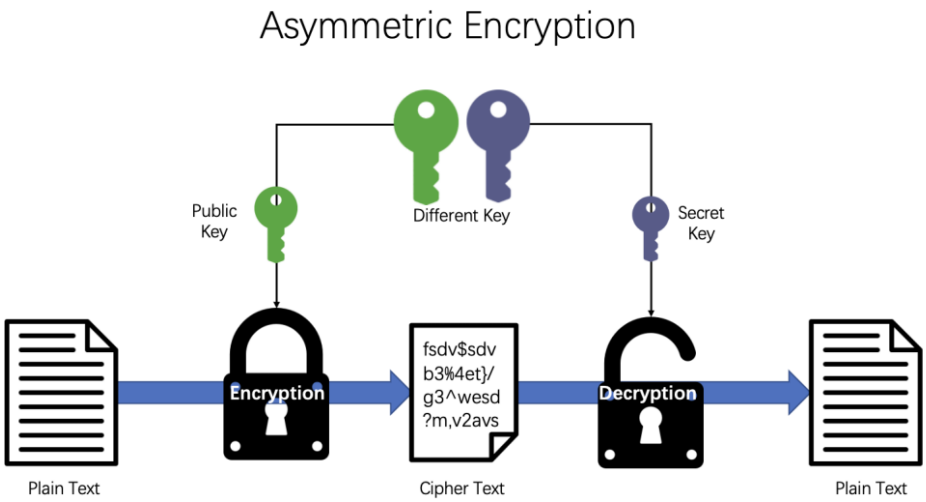

Figure 2. Asymmetric Encryption

Symmetric and asymmetric encryption are both widely used in information communication, such as internet. To date, RSA is one of the most widely used asymmetric cryptosystem in the world. It was released by Rivest, Shamir and Adleman in 1978, based on the concept of "asymmetric encryption".

2. RSA Cryptosystem

The original RSA algorithm or the textbook RSA contains one public key and one secret key. The encryption process is shown as follows:

A. Key Generation
   a) Choose two random big prime numbers p and q, and $N = p \times q$. Normally the size of N is $n = 1024$ bits ;
   b) Calculate $\varphi(N) = (p-1) \times (q-1)$ where $\varphi(N)$ is the Euler's totient function.
   c) Choose an integer e where $1 < e < \varphi(N)$ and $\gcd(e, \varphi(N)) = 1$;
   d) Based on the e chosen, determine d, where $e \times d \equiv 1 (\mathrm{mod}\ \varphi(N))$;

B. Key distribution
   a) The public key (N, e) is sent to the receiver.
   b) The secret key $(\varphi(N), d)$ is kept secret.

C. Encryption
   a) The plaintext m is encrypted into c, corresponding to $m^e \equiv c \ (\mathrm{mod}\ N)$
   b) c is transmitted by the receiver to the sender.

D. Decryption
   a) The sender can recover m from c by using the private key d:

$$c^d \equiv (m^e)^d \equiv m^{ed} \equiv m \ (\mathrm{mod}\ N)$$

The RSA function $m \rightarrow m^e$ is called a trapdoor one-way function, which can be computed very easily but can't be inverted within polynomial time without the trapdoor namely the secret key d. To find out d, the attacker has to factorize n into two big prime numbers, which is the "factoring problem" . For a sufficiently large integer n, there is no algorithm to factorize n within polynomial time, where the security of RSA relies on. So far, the record of the greatest size of key which has been broken is 829 bits by Paul, Z. in 2020. Compared to 829 bits, the normal size of n is 1024 bits.

3. The vulnerability of textbook RSA

RSA is a deterministic encryption which produces the same ciphertext given the same plaintext and key. An adversary can launch a chosen plaintext attack against the RSA encryption. By

encrypting selected plaintext with known public keys, the adversary can compare the intercepted cyphertext with the output of selected plaintext. This attack would lead the leakage of important information of the plaintext. So the original RSA algorithm or textbook RSA algorithm is not semantically secure, which means an adversary can't tell two encrypted ciphertext from each other even the adversary knows the original plaintexts. In addition to the chosen plaintext attack, an adversary can launch blinding attack using the multiplicative property of RSA. For example, an adversary wants to know the signature S of the plaintext M. He asks the private key holder to sign $M' = r^e M (\mathrm{mod}\ N)$ which seems harmless. The private key holder might be willing sign M' and send the adversary the signature S'. The adversary will simply have $S = S'/r$. Indeed,

$$S^e = \frac{S'^e}{r^e} = \frac{M'^{ed}}{r^e} = M^{ed} = M (\mathrm{mod}\ N)$$

To avoid these problems of RSA, the practical RSA encryption needs a randomized padding scheme to the plaintext before encryption.

4. The Padding Scheme for RSA

In cryptography, padding is a number of operations including appending data to anywhere of the plaintext before encryption. The purpose of a padding scheme is to avoid adversary to retrieve information of the primitive, for example, a chosen plaintext attack or an adaptive chosen ciphertext attack in RSA.

Optimal Asymmetric Encryption Padding

Optimal Asymmetric Encryption Padding(OAEP) was invented by Mihir Bellare and Phillip Rogaway in 1994 and enchanted by Don Johnson and Stephen Matyas in 1996. It was standardized as RSAES-OAEP in PKCS#1 Version 2 and lately republished as RFC 2437. OAEP combined with RSA is good at performance and provides good security especially against adaptive chosen ciphertext attack.

There are two aims of OAEP:

A. Adding random padding to plaintext can convert RSA from a deterministic scheme into a probabilistic one.
B. Prevent leaking any encryption structure information caused by chosen plaintext attack.

The padding process of OAEP is shown as below:

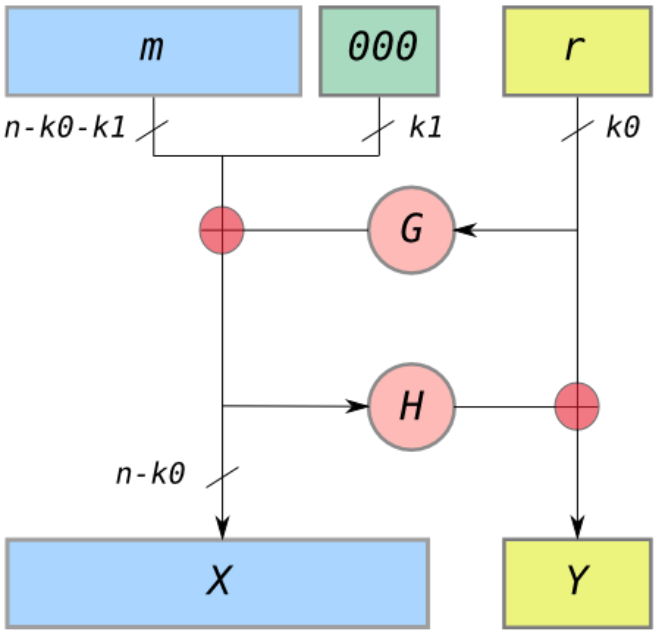

Figure 3. Scheme of OAEP

Where

- n: the length of bits of RSA modulus
- k_0 and k_1: numbers defined by OAEP protocol
- m: the plaintext with a length of $n - k_0 - k_1$ bits
- G and H are two cryptographic hash functions
- $\oplus$: xor operation
- r: a random generated string of k_0 bits

Encoding of OAEP:

a) The plaintext m is padded with k_1 zeros appending m to m' with $n - k_0$ bits length.
b) r is converted into a string of $n - k_0$ bits by a cryptographic hash function G.
c) $X = m' \oplus G(r)$.
d) X is reduced to k_0 bits by H.
e) $Y = r \oplus H(X)$.
f) The result of padding is X and Y.

Decoding of OAEP:

a) r is recovered by $r = Y \oplus H(X)$.
b) m' is recovered by $m' = X \oplus G(r)$.

Security of OAEP

The OAEP provides semantic security against chosen ciphertext attack, though Victor Shoup raised doubt about whether OAEP could provide such security. In 2001, Eiichiro Fujisaki's team proved that RSA-OAEP is semantical secure in the random oracle model.

5. Conclusion

With the rapid development of the new Internet economy, such as network payment and cloud storage, network information leakage has occurred repeatedly, posing a very high challenge to the security of network information. The RSA encryption algorithm is the most secure and widely used public key cryptographic algorithm. In this paper, we review RSA algorithm and one most used padding scheme OAEP with RSA. RSAES-OAEP protects RSA against semantical insecurity.

References:

[1] R. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," Commun. ACM, vol. 26, pp. 96–99, Jan. 1983, doi: 10.1145/359340.359342.

[2] Z. Paul, "Factorization of RSA-250," 2020. [Online]. Available: https://lists.gforge.inria.fr/pipermail/cado-nfs-discuss/2020-February/001166.html.

[3] S. Goldwasser and S. Micali, "Probabilistic Encryption & How to Play Mental Poker Keeping Secret All Partial Information," in Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing, 1982, pp. 365–377, doi: 10.1145/800070.802212.

[4] W. Diffie and M. Hellman, "New directions in croptography," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 22, p. 159, 1976, doi: 10.1007/3-540-44709-1_14.

[5] D. Boneh, "Twenty Years of Attacks on the RSA Cryptosystem," Not. AMS, vol. 46, Feb. 2002.

[6] W. Contributors, "RSA (cryptosystem)," Wikipedia, The Free Encyclopedia., 2021. https://en.wikipedia.org/w/index.php?title=RSA_(cryptosystem)&oldid=1033804079 (accessed Sep. 01, 2021).

[7] W. Contributors, "Padding (cryptography)," Wikipedia, The Free Encyclopedia., 2021. https://en.wikipedia.org/w/index.php?title=Padding_(cryptography)&oldid=1037259171.

[8] D. Boneh, “Simplified OAEP for the RSA and rabin functions,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 2139 LNCS, pp. 275–291, 2001, doi: 10.1007/3-540-44647-8_17.

[9] M. Bellare and P. Rogaway, “Optimal asymmetric encryption,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 950, pp. 92–111, 1995, doi: 10.1007/bfb0053428.

[10] D. B. Johnson and S. M. Matyas, “Asymmetric encryption: Evolution and enhancements,” CryptoBytes, vol. 2, no. 1, p. 3, 1996, [Online]. Available: https://iu.edu.jo/files/FacultyIT/Computer-Science/Courses/IT Security/stalling Computer security PP slides/Papers/AsymmetricEncryption.pdf.

[11] Wikipedia contributors, “PKCS 1,” Wikipedia, The Free Encyclopedia., 2021. https://en.wikipedia.org/w/index.php?title=PKCS_1&oldid=1026886353 (accessed Sep. 01, 2021).

[12] R. S. A. Laboratories, R. S. A. Security, and C. D. Bedford, “RSAES-OAEP Encryption Scheme.” RSA Lab, 2013.

[13] Wikipedia contributors, “Optimal asymmetric encryption padding,” Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Optimal_asymmetric_encryption_padding&oldid=1041789929 (accessed Sep. 03, 2021).

[14] S. Goldwasser and S. Micali, “Probabilistic encryption,” J. Comput. Syst. Sci., vol. 28, no. 2, pp. 270–299, 1984, doi: https://doi.org/10.1016/0022-0000(84)90070-9.

[15] C. Rackoff and D. R. Simon, “Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 576 LNCS, pp. 433–444, 1992, doi: 10.1007/3-540-46766-1_35.

[16] V. Shoup, “OAEP Reconsidered,” J. Cryptol., vol. 15, no. 4, pp. 223–249, 2002, doi: 10.1007/s00145-002-0133-9.

[17] E. Fujisaki, T. Okamoto, D. Pointcheval, and J. Stern, “RSA-OAEP is secure under the RSA Assumption,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 2139 LNCS, pp. 260–274, 2001, doi: 10.1007/3-540-44647-8_16.