On the Security of Privacy-preserving Keyword Searching for Cloud Storage Services

Fuh-Gwo Jeng[1], Shu-Yuan Lin[2], Bing-Jian Wang[2], Chih-Hung Wang[2], Tzung-Her Chen[2]
(Corresponding author: Tzung-Her Chen)

Department of Applied Mathematics, National Chiayi University[1]
Department of Computer Science and Information Engineering, National Chiayi University[2]
Chiayi City, Taiwan 60004, R.O.C.
(Email: thchen@mail.ncyu.edu.tw)
(Received Apr. 21, 2014; revised and accepted Jan. 5 & Jan. 16, 2015)

Abstract

While the traditional public-key encryption schemes with keyword search (PEKS) were pointed out suffering the performance problems, some high-performance PEKS recently have drawn more attentions. Unfortunately, these performance-enhanced schemes could encounter the same security attacks as the traditional PEKS met before. In this paper, Liu et al.'s privacy-preserving keyword searching scheme for cloud storage services (SPKS) is pointed out suffering the security attack of confidentiality. Precisely, an outside attacker could perform the test process by collecting the transmitted ciphertexts and trapdoors from senders and receivers, respectively. Thus, the relationship between encrypted data and the trapdoors is disclosed. An improved version is presented to avoid the security attacks and, furthermore, to benefit the advantages of SPKS as well.

Keywords: Designated tester, privacy-preserving keyword searching, public-key encryption schemes with keyword search, searchable encryption

1 Introduction

In order to protect the confidentiality of sensitive data in cloud-computing environments, a reliable searchable encryption mechanism is required to encrypt the sensitive data. When a user issued a keyword search onto those encrypted data, only the server (a designated tester) chosen by a sender is able to perform the test by checking the relationship between a ciphertext and a trapdoor. However, an adversary is not allowed to do so [3, 5, 9, 10].

The public-key encryption scheme with keyword search (PEKS) was first proposed by Boneh et al. [1]. Based on Boneh et al.'s scheme, Hwang and Lee [4] proposed another PEKS for a multi-receiver environment. In 2010, the concept of proxy re-encryption was applied in keyword searching by Shao et al. [8] and by Yau and Phan [11] as well. Recently, Rhee et al. [7] enhanced the trapdoor security to prevent from off-line keyword-guessing attacks.

In order to enhance the performance, Liu et al. [6] proposed a secure privacy-preserving keyword searching scheme for cloud storage services (SPKS), which enabled the cloud service provider (CSP) to participate in the partial decipherment to obtain an intermediate result of the decipherment before returning the search results. In such a way, the communication and computational overhead was reduced greatly in decryption process for the user.

However, Liu et al.'s scheme met a potential security problem. That is, anyone can perform the test process by collecting the transmitted ciphertexts and the corresponding trapdoors. Therefore, any outside attacker can further construct the relationship between the encrypted data and the given trapdoors of known keywords. The confidentiality is not guaranteed any more. Thus, consequently, an improved SPKS scheme is presented to prevent from the attacks mentioned above and at the same time to inherit the advantages of SPKS as well.

2 The Security of Liu et al.'s Keyword Searching for Cloud Storage Services

In this section, Liu et al.'s SPKS scheme is briefly reviewed and the readers may refer to [7] for details. Finally, its security problem will be pointed out in the later.

2.1 Review of Liu et al.'s SPKS Scheme

In the scenario, there are three participants including the user (sender), the server (CSP), and the receiver.

1) *Global setup:* Determine two cyclic groups G_1 and G_2 with prime order p , and their admissible bilinear paring function $\hat{e} : G_1 \times G_2 \rightarrow G_2$. Given a random element $g \epsilon G_1$ and let H_1, H_2, H_3, H_4 and H_5 are random

oracles, where $H_1, H_3 : \{0,1\}^* \rightarrow G_1^*, H_2, H_5 : G_2 \rightarrow \{0,1\}^{log^q}$, and $H_4 : G_2 \rightarrow \{0,1\}^n$. The plaintext space includes $m\epsilon\{0,1\}^n$ for some n and $W\epsilon\{0,1\}^*$. The ciphertext space includes $C_m = G_1^* \times \{0,1\}^n$ and $C_W \epsilon G_2$.

2) *KeyGen:* The server (resp. receiver) generates his private key by randomly choosing $sk_{CSP} = x\epsilon\mathbb{Z}_q$ (resp. $sk_R = y\epsilon\mathbb{Z}_q$) and the corresponding public key by computing $pk_{CSP} = g^x$ (resp. $pk_R = g^y$).

3) *EMBEnc:* To encrypt an email m under a receiver's public key g^y and CSP's public key g^x, the user selects a random element $r, t\epsilon\mathbb{Z}_q$, computes

$$u_1 = g^r, u_2 = t \oplus H_5(\hat{e}(g^y, g^x)^r),$$
$$u_3 = m \oplus H_4(\hat{e}(H_3(t), (g^y)^r)),$$

and sets the ciphertext $C_m = \langle u_1, u_2, u_3 \rangle$.

4) *KWEnc:* To encrypt m's keywords $W_1, ..., W_k (k\epsilon\mathbb{Z}^+)$ under a receiver's public key g^y, the user computes $C_{W_i} = H_2(\hat{e}(g^y, H_1(W_i)^r))$, where $W_i\epsilon\{W_1, ..., W_k\}$, and sends $\langle C_m, C_{W_1}, ..., C_{W_k} \rangle$ to CSP.

5) *TCompute:* To retrieve only the emails containing keyword $W_j (j\epsilon\mathbb{Z}^+)$, the receiver computes the trapdoor $T_{W_j} = H_1(W_j)^y \epsilon G_1$ under a receiver's private key , and sends it to CSP.

6) *KWTest:* To determine whether a given email contains keyword W_j, CSP tests whether the equation $C_{W_i} = H_2(\hat{e}(u_1, T_{W_j}))$ holds.

7) *PDecrypt:* To obtain an intermediate result of the decipherment, CSP calculates $t = u_2 \oplus H_5(\hat{e}(g^y, u_1)^x)$, computes $C_t = \hat{e}(H_3(t), u_1)$, and sends $\langle C_m, C_{W_1}, \cdots, C_{W_k}, C_t \rangle$ to the receiver.

8) *Recovery:* Given the ciphertext $C_m = \langle u_1, u_2, u_3 \rangle$ and C_t, the receiver computes $m = u_3 \oplus H_4((C_t)^y)$ to recover the message m.

2.2 Security Problem

Since an outside attacker may intercept the transmitted ciphertexts from senders and the trapdoors from receivers. Without the key, the outside attacker also can easily check if the equation $C_{W_i} = H_2(\hat{e}(u_1, T_{W_j}))$ holds.

The scheme called secure means that attackers have no feasible way to deduce any information about secret. However, this attack does not need to face the hard problem on which a cryptosystem relies. It's worthwhile to note that the cost of this attack by intercepting and checking the above equation is low. That is, the attack is feasible. Therefore, the relationship between encrypted data and the given trapdoors of known keywords is able to be constructed. That is, the security information of linkage between them is revealed. This is why the designated-tester scheme is essential for searchable encryption schemes.

3 Improvement of Liu et al.'s SPKS Scheme

Inspired by Hu and Liu's scheme [2], the enhanced SPKS scheme consisting of the following processes is proposed.

1) *Global setup:*The first process is the same as that setup in Liu et al.'s.

2) *KeyGen:*CSP generates his private key by randomly choosing $sk_{CSP} = x\epsilon\mathbb{Z}_q$ and the corresponding public key by computing $pk_{CSP} = g^x$. The receiver generates his private key by randomly choosing $sk_R = \langle y, z \rangle \epsilon\mathbb{Z}_q$ and the corresponding public key by computing $pk_R = \langle k_{R1}, k_{R2}, k_{R3}, k_{R4} \rangle = \langle g^y, g^{zy^2}, g^{yz}, (pk_{CPS})^z \rangle$.

3) *EMBEnc:*To encrypt an email m under a receiver's public key $k_{R1} = g^y$ and CSPs public key g^x, the sender selects a random element $r, t\epsilon\mathbb{Z}_q$, and computes
$u_1 = g^r, u_2 = t \oplus H_5(\hat{e}(g^y, g^x)^r)$,
$u_3 = m \oplus H_4(\hat{e}(H_3(t), (g^y)^r))$,
and sends the ciphertext $C_m = \langle u_1, u_2, u_3 \rangle$.

4) *KWEnc:* To encrypt m's keywords $W_1, ..., W_k (k\epsilon\mathbb{Z}^+)$ under the receiver's public key $k_{R2} = g^{zy^2}$ and $k_{R4} = g^{zx}$, the user computes $C_{W_i} = \langle A, B \rangle = \langle (k_{R2}^r), H_2(\hat{e}(k_{R4}, H_1(W_i)^r)) \rangle$, where $W_i\epsilon\{W_1, ..., W_k\}$, and sends $\langle C_m, C_{W_1}, ..., C_{W_k} \rangle$ to CSP.

5) *TCompute:* To retrieve the emails containing keyword $W_j (j\epsilon\mathbb{Z}^+)$, the receiver computes the trapdoor $T_{W_j} = \langle T_1, T_2 \rangle = \langle (pk_{CSP})^{r'}, H_1(W_j)^{1/y^2} \cdot g^{r'} \rangle$ where $r'\epsilon\mathbb{Z}_q$ is randomly chosen by the receiver, and sends it to CSP.

6) *KWTest:* To determine whether a given email contains keyword W_j, CSP should compute $T_3 = (T_2)^x / T_1 = H_1(W_j)^{x/y^2}$ with the private key x, and then check if $H_2(\hat{e}(A, T_3))$ is equal to B.

7) *PDecrypt:* To obtain an intermediate result of the decipherment, CSP calculates $t = u_2 \oplus H_5(\hat{e}(g^y, u_1)^x)$ and $C_t = \hat{e}(H_3(t), u_1)$, and sends $\langle C_m, C_{W_1}, \cdots, C_{W_k}, C_t \rangle$ to the user.

8) *Recovery:* Given the ciphertext $C_m = \langle u_1, u_2, u_3 \rangle$ and C_t, the receiver computes $m = u_3 \oplus H_4((C_t)^y)$ to recover the message m.

4 Discussions

4.1 Correctness

The correctness of searchable encryption, i.e. *KWTest*, is described as follows.

$$\begin{aligned} B &= H_2(\hat{e}(k_{R4}, H_1(W_i)^r)) \\ &= H_2(\hat{e}(g^{xz}, H_1(W_i)^r)) \\ &= H_2(\hat{e}(g, H_1(W_i))^{xzr}) \\ &= H_2(\hat{e}(g^{y^2zr}, H_1(W_i)^{x/y^2})) \\ &= H_2(\hat{e}(A, T_3)). \end{aligned}$$

If $W_i = W_j, H_1(W_i)^{x/y^2} = H_1(W_j)^{x/y^2} = T_3$. Hence, $B = H_2(\hat{e}(A, H_1(W_i)^{x/y^2})) = H_2(\hat{e}(A, T_3))$.

4.2 Security

Since the outside attacker doesn't have CSP's private key sk_{CSP} to compute $T_3 = (T_2)^x/T_1 = H_1(W_j)^{x/y^2}$,even if an outside attacker obtains the ciphertext C_m and the trapdoor T_{W_j}, (s)he still cannot perform the test process. Suppose *Alice* is an attacker. Assume that $T_w = \langle T_1, T_2 \rangle$ is a trapdoor. To retrieve a correct keyword w from the given T_w, it should be possible if *Alice* obtains $H_1(w)^{1/y^2}$ or $H_1(w)$ from T_w.

Because a discrete logarithm problem is hard, *Alice* has no feasible way to obtain the unknown r' or $x \epsilon \mathbb{Z}_q$ from $T_1 = (pk_{CSP})^{r'}$ where $pk_{CSP} = g^x$.

Furthermore, even though *Alice* can compute

$$\begin{aligned} & e(pk_{CSP}, T_2)/e(g, T_1) \\ = \; & e(g^x, H_1(w)^{1/y^2} \cdot g^{r'})/e(g, (g^x)^{r'}) \\ = \; & e(g^x, H_1(w)^{1/y^2}). \end{aligned}$$

Alice has no feasible way to guess keyword w by computing $e(pk_{CSP}, H_1(w)^{1/y^2})$ without knowing receiver's secret key y or CSP's secret key x.

Even CSP can obtain $g^{r'}$ from T_1 using its secret key $pk_{CSP} = x$, CSP cannot successfully guess the keyword by checking if $e(k_{R4}, T_2) = e(k_{R4}, g^{r'})e(g, H_1(w'))$.

5 Conclusions

In this paper, the security weakness in Liu et al.s scheme is pointed out. To benefit from Liu et al.'s scheme, i.e., efficiency, an improved version is proposed to solve their weakness and to keep the advantages of Liu et al.'s scheme as well.

Acknowledgments

This research was partially supported by National Science Council, Taiwan, R.O.C., under contract no. NSC 102-2221-E-415-014- and NSC 102-2221-E-415-007-.

References

[1] D. Boneh, G. D. Crescenzo, R. Ostrovsky, G. Persiano, "Public-key encryption with keyword search," in *Proceedings of EUROCRYPT'04*, LNCS 3027, pp. 506-522, 2004.

[2] C. Hu, P. Liu, "A secure searchable public key encryption scheme with a designated tester against keyword guessing attacks and its extension," in *Communications in Computer and Information Science*, vol. 215, pp. 131-136, 2011.

[3] S. T. Hsu, C. C. Yang, and M. S. Hwang, "A study of public key encryption with keyword search," *International Journal of Network Security*, vol. 15, no. 2, pp. 71–79, 2013.

[4] Y. H. Hwang, P. J. Lee, "Public-key encryption with conjunctive keyword search and its extension to a multi-user system," in *Proceedings of Pairing'07*, LNCS 4575, pp. 2-22, 2007.

[5] C. C. Lee, S. T. Hsu, and M. S. Hwang, "A study of conjunctive keyword searchable schemes," *International Journal of Network Security*, vol. 15, no. 5, pp. 321–330, 2013.

[6] Q. Liu, G. Wang, J. Wu, "Secure and privacy preserving keyword searching for cloud storage services," *Journal of Network and Computer Applications*, vol. 35, no. 3, pp. 927–933, 2012.

[7] H. S. Rhee, J. H. Park, W. Susilo, D. H. Lee, "Trapdoor security in a searchable public-key encryption scheme with a designated tester," *Journal of Systems and Software*, vol. 83, no. 5, pp. 763-771, 2010.

[8] J. Shao, Z. F. Cao, X. H. Liang, H. Lin, "Proxy reencryption with keyword search," *Information Sciences*, vol. 180, no. 13, pp. 2576-2587, 2010.

[9] J. Wang, X. Yu, and M. Zhao, "Fault-tolerant verifiable keyword symmetric searchable encryption in hybrid cloud ," *International Journal of Network Security*, vol. 17, no. 4, pp. 471–483, 2015.

[10] Y. Wang, W. Bao, Y. Zhao, H. Xiong, and Z. Qin, "An ElGamal encryption with fuzzy keyword search on cloud environment," *International Journal of Network Security*, vol. 18, no. 3, pp. 481–486, 2016.

[11] W. C. Yau, R. C. W. Phan, S. H. Heng, B. M. Goi, "Proxy re-encryption with keyword search: new definitions and algorithms," *Communications in Computer and Information Science*, vol. 122, pp. 149-160, 2010.

Fuh-Gwo Jeng received his M.S. in computer and information science from National Chiao Tung University and Ph.D. degree at the Institute of Computer Science, National Chung Hsing University, Taiwan. He is presently an associated professor of Department of Applied Mathematics, National Chiayi University. His research interests include information security and computer graphics.

Shu-Yuan Lin received her M.S. in Department of Computer Science and Information Engineering from

National Chiayi University in 2013. Her research interest is multimedia security.

Bing-Jian Wang received his M.S. in Department of Computer Science and Information Engineering from National Chiayi University in 2012. His research interest is visual crypotography and information security.

Chih-Hung Wang was born in Kaohsiung Taiwan, in 1968. He received his BS degree in information science from Tunghi University and MS degree in information engineering from National Chung-Cheng University, Taiwan, R.O.C., in 1991 and 1993, respectively. He received the PhD degree in information engineering from National Cheng Kung University, Taiwan, R.O.C. in 1998. He is presently an associated professor of Department of Computer and Information Engineering, Nation Chiayi University, Taiwan, R.O.C. His research interests include cryptography, and information security.

Tzung-Her Chen was born in Tainan, Taiwan, Republic of China, in 1967. He received the B.S. degree in Department of Information & Computer Education from National Taiwan Normal University in 1991 and the M.S. degree in Department of Information Engineering from Feng Chia University in 2001. In 2005, he received his Ph.D. degree in Department of Computer Science from National Chung Hsing University. He has been with Department of Computer Science and Information Engineering at National Chiayi University as Professor since August 2011. His research interests include information hiding, multimedia security, digital rights management, network security. He is an honorary member of the Phi Tau Phi Scholastic Honor Society.