Authenticity Protection in Outsourced Database

Jun Ye
School of Mathematics and Statistics,
Artificial Intelligence Key Laboratory of Sichuan Province,
Sichuan University of Science & Engineering,
Guangxi Key Laboratory of Cryptography
and Information Security,
Zigong, Sichuan, P. R. China
yejun@suse.edu.cn

Zheng Xu *
The Third Research Institute of
the Ministry of Public Security,
Shanghai, China
zhengxu@shu.edu.cn

Yong Ding
Guangxi Key Laboratory of Cryptography
and Information Security,
School of Computer Science
and Information Security,
Guilin University of Electronic Technology,
Guilin, Guangxi, P. R. China
stone_dingy@126.com

Qin Wang
School of Automation
& Information Engineering,
Sichuan University of Science & Engineering,
Zigong, Sichuan, P. R. China
438096157@qq.com

***Abstract*— In this paper, we focus on the security of outsourced database. A verification scheme for outsourced database is proposed based on the verifiable polynomial technique. In this scheme, we consider the encrypted database. The outsourced high degree polynomial will enhance the authenticity of the data. If the cloud server returns fake data, it will be detected by the clients easily.**

Keywords– Verifiable, Polynomial, Outsourced Database

1 Introduction

With the rapid development of Internet and network technology, more and more data will be used. The fast growth of data brings much trouble to people. Cloud computing is powerful, it helps us to solve the big data problem. Cloud computing is the development and application of distributed computing, parallel computing and grid computing. With the help of the powerful cloud server, the client can easily accomplish the complex work easily.

Outsource service allows resource constrained clients to outsource the complex tasks to the powerful cloud server with the manner of pay-per-use. In order to save the own storage space, people will outsource the huge database to a remote cloud server. With the development of cloud computing, outsourcing is becoming an important part in modern business. By utilizing outsourcing, clients can concentrate on their own work and operate their business applications via the Internet, rather than maintaining the substantial hardware, software, and applications.

However, the powerful cloud server is untrusted. When data is outsourced, clients loose the controllability of the outsourced data. It is a hard work to verify the authenticity of the data. The untrusted cloud server may tamper the data. When clients query the data, they will get the fake data. Thus, it is important to study the verifiable outsourced database.

Our Contributions. This paper focus on the correctness of the outsourced data. A new and simple scheme to verify the encrypted data is proposed. In the scheme, the secure oursourcing computation of high degree polynomials is used to help the clients accomplish the verification. Clients can verify the required parts of the outsourced database by checking the *proof* information. The outsourced data is encrypted, and it will not be revealed.

1.1 Related Work

In 1980s, Ben-Or et al. proposed a outsource computation scheme with an honest-but-curious oracle [3, 4]. Then some outsourcing computation

DOI reference number: 10.18293/SEKE2017-027

schemes come out [7, 11]. In 2002, Atallah, Pantazopoulos and Rice [1] proposed secure outsourcing scheme for scientific computing and numerical calculations. However, the verification phase was not considered. In 2008 Benjamin er al. [6] proposed a verifiable outsourcing computation scheme for linear algebraic calculation. And In 2010, Gentry et al. [10] expanded verifiable outsourcing computation to arbitrary function F. However, the efficiency is low. In 2016, Ye et al.[15] proposed a verifiable delegation scheme for polynomials, which improved the efficiency.

There are a lot of work on the verification of outsourced database, such as, [8, 2]. The homomorphic encryption was used in some schemes, however, this reduced the computation efficiency. Then some schemes without homomorphic encryption comes out. Some schemes are based on based on Message Authentication Code [5, 9], and some are based on Merkle hash trees [13, 12]. An index tree for the database is generated by using hash functions, by which the authenticity auditing can be achieved. However, lots of information for verification has to be stored. In 2009, Pang et al. [14] proposed an outsourced database scheme based on the signature chaining technique, in which the computation can be used for verification. In 2013 Catalano and Fiore [9] used the vector commitment to generate a verifiable database with efficient update.

1.2 Organization

The organization of this paper is as follows. Some preliminaries are given in Section 2. The proposed scheme are given in Section 3. Finally, the conclusion is made in Section 4.

2 Preliminaries

2.1 Hash Function

A hash function can take an arbitrary input and output a fixed-size string, which satisfies the following properties.

- It is easy to compute the hash value for any given input.
- It is infeasible to compute an input such that the hash value equals a given hash value.
- It is infeasible to find two different inputs that can get the same hash value.

2.2 Verifiable Outsourced Database

The outsourced database is an example of client-server model. In the outsoutced database model, the service provider is powerful, who has the infrastructure for outsourced databases, and can provide efficient data processing, such as, store, update and query the database.

Verifiable outsourced database allows clients to authenticate database operations. The clients can query the outsourced database, when get the returned results, the clients can verify the correctness of the outsourced data with the help of the proof.

3 The Proposed Scheme

The are three parts in the system, client, cloud server 1 and cloud server 2.

Initialization. The client generates a high degree polynomial, $f(x)$.

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d$$

where, $a_i \in \mathbb{Z}_p$, $i = 0, 1, \cdots, d$, P is a large prime, and g is a generator.

Then, client selects a prime q, such that $|q| = |p|$, and computes $n = pq$. Client chooses $q_i \in_R \mathbb{Z}_n^*$, and computes

$$r_i = (1+n)^{a_i} q_i^n \mod n^2$$

where, $i = 0, 1, \cdots, d$.

The polynomial is transformed as

$$F(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_d x^d.$$

Client randomly selects $a, k_0, k_1 \in \mathbb{Z}_p^*$, and computes

$$t_i = g^{k_0 k_1^i} g^{a r_i}$$

where, $i = 0, 1, \cdots, d$.

At last client sends

$$(r_0, r_1, \cdots, r_d)$$

and

$$(t_0, t_1, \cdots, t_d)$$

to cloud server 1.

Ciphertext Generation.The client selects an encryption algorithm $En(\cdot)$ and encrypts the original data x as $\sigma_x = En(x)$, and sends it to the cloud server 1.

When receiving the data σ_x, the cloud server 1 computes

$$\sigma_y = F(\sigma_x) = \prod_{i=0}^{d} r_i^{\sigma_x^i}$$

and

$$T = \prod_{i=0}^{d} t_i^{\sigma_x^i}.$$

Then, cloud server 1 sends σ_y and T to the client.

Client computes

$$Z = \prod_{i=0}^{d} (g^{k_0 k_1^i})^{\sigma_x^i},$$

and

$$y = \frac{(\sigma_y)^{\Phi(n)} - 1}{n} \Phi^{-1}(n).$$

Then, client verifies whether the following equation holds

$$T = Z \cdot g^{ay}.$$

If not, client outputs $\perp$. Otherwise, client computes

$$proof = H(i||\sigma_x||y)$$

where, $H(\cdot)$ is a non-collision hash function.

At last, client sends

$$(i, \sigma_x, proof)$$

to cloud server 2.

Query. The client queries for the data in the position i. Then the cloud server returns the ciphertext

$$C_i = (i, \sigma_{x_i}, H(i||\sigma_{x_i}||y_i))$$

to the client.

Verify. Client sends σ_x to the cloud server 1. Cloud server 1 computes

$$\sigma_{y_i} = F(\sigma_{x_i}) = \prod_{i=0}^{d} r_i^{\sigma_{x_i}^i}$$

and

$$T_i = \prod_{i=0}^{d} t_i^{\sigma_x^i}.$$

Then, cloud server 1 sends σ_{y_i} and T_i to the client.

Client computes Z_i with σ_{x_i},

$$\begin{aligned} Z_i &= \prod_{i=0}^{d} (g^{k_0 k_1^i})^{\sigma_x^i} \\ &= g^{k_0 \frac{1-(k_1\sigma_{x_i})^{n+1}}{1-k_1\sigma_{x_i}}} \end{aligned}$$

and

$$y_i^* = \frac{(\sigma_{y_i})^{\Phi(n)} - 1}{n} \Phi^{-1}(n).$$

And then verifies whether the following equation holds

$$T_i = Z_i g^{a y_i^*}.$$

If not, client outputs $\perp$. Otherwise, client computes

$$proof_i^* = H(i||\sigma_{x_i}||y_i).$$

and verifies

$$proof_i^* \stackrel{?}{=} proof.$$

If the equation holds, the data σ_x is correct, and client can get the original data $x = Dec(\sigma_x)$. Otherwise, client outputs $\perp$.

4 Conclusion

The rapid increasing of data brings much trouble to people. More storage space is needed. Cloud computing gathers a lot of resource together, and provides huge storage space for users. In order to save the local resource, clients often outsource their database to the remote cloud server. And for the security, the outsourced data should be encrypted. As Large amounts of data is outsourced to the cloud server, the users have to keep a little information for verification. In this paper, a new algorithm for outsourced database verification is proposed. The actual data is encrypted, and it will not be revealed. The computation results can be easily verified by the client. The clients can easily check the authenticity of outsourced data.

ACKNOWLEDGMENT

This work was supported by Guangxi Key Laboratory of Cryptography and Information Security (No. GCIS201607); the Fund of Lab of Security Insurance of Cyberspace, Sichuan Province (szjj2016-091); the Talent Project of Sichuan University of Science & Engineering (2017RCL23).

References

[1] M.J. Atallah, K. N. Pantazopoulos, J.R. Rice, and E.H. Spafford. Secure outsourcing of scientific computations. *Advances in Computers*, 54:215–272, 2002.

[2] M. Backes, D. Fiore, and R. M. Reischuk. Verifiable delegation of computation on outsourced data. In *2013 ACM SIGSAC Conference on Computer and Communications Security, CCS'13, Berlin, Germany, November 4-8, 2013*, pages 863–874, 2013.

[3] D. Beaver and J. Feigenbaum. Hiding instances in multioracle queries. In *Proc. of STAC' 90*, volume 415, pages 37–48. Springer-Verlag, 1990.

[4] D. Beaver, J. Feigenbaum, J. Kilian, and P. Rogaway. Locally random reductions: Improvements and applications. *Journal of Cryptology*, 10(1):17–36, 1997.

[5] S. Benabbas, R. Gennaro, and Y. Vahlis. Verifiable delegation of computation over large datasets. In *Advances in Cryptology - CRYPTO 2011 - Proceedings of the 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011*, pages 111–131, 2011.

[6] D. Benjamin and M.J. Atallah. Private and cheating-free outsourcing of algebraic computations. In *Sixth Annual Conference on Privacy, Security and Trust, PST 2008, Fredericton, New Brunswick, Canada*, pages 240–245. Springer-Verlag, October 2008.

[7] M. Blum and S. Kannan. Designing programs that check their work. *Journal of the ACM (JACM)*, 42(1):269–291, 1995.

[8] D. Boneh and B. Waters. Conjunctive, subset, and range queries on encrypted data. In *Theory of Cryptography, Proceedings of the 4th Theory of Cryptography Conference, TCC 2007, Amsterdam, The Netherlands, February 21-24, 2007*, pages 535–554, 2007.

[9] D. Catalano and D. Fiore. Vector commitments and their applications. In *Public-Key Cryptography - PKC 2013 - Proceedings of the 16th International Conference on Practice and Theory in Public-Key Cryptography, Nara, Japan, February 26 - March 1, 2013*, pages 55–72, 2013.

[10] Rosario Gennaro, Craig Gentry, and Bryan Parno. Non-interactive verifiable computing: Outsourcing computation to untrusted workers. In *Advances in Cryptology - CRYPTO 2010, 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings*, pages 465–482, 2010.

[11] O.J. Akomode B. Lees and C. Irgens. Constructing customised models and providing information to support it outsourcing decisions. *Logistics Information Management*, 11(2):114–127, 1998.

[12] E. Mykletun, M. Narasimha, and G. Tsudik. Authentication and integrity in outsourced databases. *TOS*, 2(2):107–138, 2006.

[13] H. Pang, A. Jain, K. Ramamritham, and K. Tan. Verifying completeness of relational query results in data publishing. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, Baltimore, Maryland, USA, June 14-16, 2005*, pages 407–418, 2005.

[14] H. Pang, J. Zhang, and K. Mouratidis. Scalable verification for outsourced dynamic databases. *PVLDB*, 2(1):802–813, 2009.

[15] J. Ye, H. Zhang, and C. Fu. Verifiable delegation of polynomials. *International Journal of Network Security*, 18(2):283–290, 2016.