10th International Conference on Information Science and Technology
Bath, United Kingdom, September 9-15, 2020

Cryptanalysis of “Certificateless remote data integrity checking using lattices in cloud storage”

Caihui Lan
School of Electronic and Information Engineering
Lanzhou City University
Lanzhou, China
Email: lan_ch@lzcu.edu.cn

Haifeng Li
School of Software
Dalian University of Technology
Dalian, China
Email: lihaifeng8848@mail.dlut.edu.cn

Caifen Wang
College of Big Data and Internet
Shenzhen Technology University
Shenzhen, China
Email: wangcaifen@sztu.edu.cn

***Abstract*—C. Sasikala et al.’s proposed a certificateless RDIC protocol to audit the outsourced data without the requirement of PKI infrastructure and against quantum computer attacks. They argue the RDIC is correct and possesses desired security properties. However, through cryptoanalysis technique we find that their scheme is neither correct nor secure against two common attacks, that is, signature forgery attack from the malicious cloud service provider (CSP) and public key replacement attack from the curious private key generator (PKG) and other malicious users. Furthermore, we remark that their RDIC is not a rigorous certificateless protocol.**

***Keywords*—Cloud storage, Data integrity, Certificateless, Lattice-based**

I. Introduction

Cloud storage is the fundamental services of cloud computing. By leveraging cloud computing, the data owners can relieve of the burden for local storage, enjoy the conveniently ubiquitous access with independent time and geographical regions, a rising number of people and enterprises intend to outsource data to the cloud server, such as Google Cloud, Alibaba Cloud, etc. By migrating local data to the remote cloud server, people may save vast capital expenditure on large-scale hardware infrastructure, various software, and be freedom from the complicated and burdensome daily maintenance.

Although the cloud storage server can provide enormous appealing benefits to data owners, it also causes unprecedented challenges and serious security problems. For instance, to get more profits, an untrusted Cloud Service Provider (CSP) may deliberately delete the infrequently visited data, or intentionally hide the truth of the serious data corruption accidents only to maintain its reputation, thus, make the data owner missed valuable opportunities to recover data [1]. In addition, despite the cloud storage server is generally much more powerful and reliable than that of ordinary user’s, they still face a serials of security vulnerabilities occasionally, such as failures of system hardware, bugs of system software or application software, the attacks from outside malicious hackers [2]. These security threats may compromise the data with respect of confidentiality, integrity. Thus, the remote data integrity checking is an essential function for the cloud computing paradigm, i.e., the data owner should periodically audit the outsourced data to confirm whether the data are kept intact. In practice, a third-party is often delegated to facilitate public auditing tasks for the owners.

To address the issue of auditing data in remote server, many academic researchers and industrial engineers have conducted the widespread and profound investigation. The seminal practical work of the remote data integrity verifying was developed in 2007 [3]. Following their pioneer work, an increasing number of auditing schemes have been designed under various security paradigms and application scenarios so far[4–8].

It should be pointed that the aforementioned works are built on PKI paradigm that is widely used in public key cryptosystem. Despite PKI is widespread applied and play an important role in public key cryptosystem, it still faces with some serious security issues. For instance, PKI often becomes the main target of hackers and creates a bottleneck for system security. In addition, in the conventional PKI model, the PKI is responsible for managing the public key certificate (like an X.509 certificate) generation, storage, distribution, verification, revocation, renewals, and so on. Hence, the PKI system bares the tough burden of certificates maintenance and it is difficult to deployment of PKI in reality especially in the case of the user scale is relatively large.

To avoid the tough burden of certificates maintenance, an Identity Based Public Key Cryptosystems(ID-PKC) was put forward[9]. Inspired by Shamir’s pioneer work, multiple ID-based public auditing schemes was proposed [10–13].

Unfortunately, these ID-based public auditing schemes suffer from key escrow problem because in IBC based system, it is required a private key generator (PKG) to produce secret key for registrants. Thus, if the PKG is not completely convinced, it could impersonate any legitimate registrant to do any evil things with user’s secrete key.

For eliminating key escrow issue, Al-Riyami et al. brought forth and provided the concrete certificateless public key cryptosystems (CL-PKC) scheme [14]. Under this model, the secret key of user only partially generated by the PKG. CL-PKC has powerful functions and many charming features that can eliminate the heavy certificate management overheads under PKI model and handle key escrow issue under IBC system simultaneously. Thus, to construct certificateless based public auditing scheme is an excellent approach for solving

978-1-7281-5558-6/20/$31.00 ©2020 IEEE

TABLE I
PARAMETERS AND THEIR MEANING

Parameter	Meaning
q	a prime with $q \geq 3$
n, m	two integers satisfying $m \geq \lceil 5n \log q \rceil$
A	$A \in Z_q^{n \times m}$
D_{Z^m, σ_R}	Gaussian Distribution over Z^m, variance is σ_R
R	$R \in Z_q^{m \times m}$
A'	$A' = AR^{-1}$
δ, σ	two Gaussian parameters
T_A	the optimal bases of $\Lambda_q^{\perp}(A)$
$T_{A'}$	the optimal bases of $\Lambda_q^{\perp}(A')$

data integrity issue in cloud computing.

Following CL-PKC, many certificateless based public auditing protocols[15–20] have been designed in certificateless settings.

Nevertheless, most current public auditing schemes are not secure any longer because the underlying hard problem of them would be broken quickly by the powerful quantum computers attacks in polynomial time according to the security proof by Shor [21]. Fortunately, cryptologist have found several cryptographic primitives that can resist quantum computers attacks, which is named post quantum cryptography [22]. Among which, lattice-based cryptography is an attractive candidate primitive to provide efficient quantum safe.

Inspired by the groundbreaking work of Ajtai[23] in 1996, many lattice-based cryptography schemes with various security features [24–29] have been developed up to now.

II. PRELIMINARIES

A. Definition of lattice

Definition 1 ([27]). Let $B \in R^{n \times n}$ is an invertible matrix, and $b_i \in R^n$ is the i-th column vector. The following set is defined as full-rank lattice, B is a basis of the lattice.

$$\Lambda = \{Bc, c \in Z^n\}$$

Let $\|B\| = \max\limits_{1 \leq i \leq n} \|b_i\|$. $\tilde{B} = \{\tilde{b}_1, \tilde{b}_2, \cdots, \tilde{b}_n\}$ is computed by Gram-Schmidt orthogonal transformation on B.

Definition 2 ([30]). $q \geq 2$ is a natural number, $A \in Z_q^{n \times m}$is a matrix, the three different types of ajtai lattices are described as bellow.

1)$\Lambda_q(A) = \{l \in Z^m : \exists t \in Z_q^n, l = A^T t mod q\}$.

2) $\Lambda_q^{\perp}(A) = \{l \in Z^m : Al = 0 mod q\}$.

3) $\Lambda_q^{x}(A) = \{l \in Z^m : Al = x mod q\}$.

The above defined ajtai Lattices are often employed in the construction of concrete schemes.

Table I is summarized the relevant parameters used for several important lattice algorithm with polynomial time in [31].

Algorithm 1 $(A, T_A) \leftarrow TrapGen(q, n, m)$**([32])**. This algorithm outputs A and T_A satisfied $\|T_A\| \leq O(\log q^n)$ and $\left\|\tilde{T}_A\right\| \leq O(\sqrt{\log q^n})$. Here, A and $Z_q^{n \times m}$ conform uniform distribution with indistinguishability.

Algorithm 2 **(**$(R, T_{A'}) \leftarrow (SampleRwithBasis(q, m, n, A)$ **[24])**. This algorithm returns $T_{A'}$ satisfied $\left\|\tilde{T}_{A'}\right\| \leq \sigma_R / \omega(\sqrt{\log m})$ and R.

Algorithm 3 **(**$\theta \leftarrow Sample\Pr e(A, T_A, x, \delta)$ **[27])**. This algorithm returns a preimage $\theta \in Z^m$ is statistically undistinguishable with a sample satisfying normal distribution on $\Lambda_q^{\perp}(A)$ having the variance δ, and such that $A\theta = x \bmod q$. Here $\delta \geq \left\|\tilde{T}_A\right\| .\omega(\sqrt{\log m})$.

Algorithm 4 **(**$T_{A'} \leftarrow NewBasisDel(A, T_A, R, \sigma)$ **[24])**. This algorithm returns $T_{A'}$. Here, $\sigma \geq \left\|\tilde{T}_A\right\| .\sigma_R \sqrt{m} \omega(\sqrt{\log^3 m})$.

B. Hardness assumption

Definition 5 ([33]). Given A,q, and a real constant parameter $\varsigma > 0$, the Small Integer Solution($SIS_{q,n,m,\varsigma}$) Problem aim to obtain $\theta \in Z^m \wedge \theta \neq 0$ over $\Lambda_q^{\perp}(A)$ satisfy $A\theta = 0 \wedge \|\theta\| \leq \varsigma$.

C. System framework

The system framework in [31] involved four entities named the date owner (DO), CS, the TPA and PKG. The detailed meaning of them and the relationship and interaction among them can be referred Fig 1 in literature [31].

III. CRYPTOANALYSIS

Through the cryptoanalysis of RDIC protocol, we find the protocol is either incorrect or insecurity against two common attacks, that is, signature forgery attack and public key replacement attack. The detailed analysis and explanation are described below.

A. Correctness Analysis

We argued RDIC protocol is incorrect according to two main reasons.

1)In the first place, we illustrate that their proven procedure of verification algorithm is incorrect.

In the C. Sasikala et al.'s Certificateless RDIC scheme[31], they prove the correctness of the protocol through proven the verification equation $\gamma A\sigma + S = B \cdot \mu \bmod q$ holds in subsection "5.1 Correctness" in literature[31]. Unfortunately, their proven is incorrect because the proven suffers from a serious mistake in their formula derivation procedure.

In the next, we formally illustrate our argument described in the above as follows.

To facilitate expression, we abstract the correctness proven procedure of C. Sasikala et al.'s as below.

$$\begin{aligned}\gamma \cdot A\sigma + S &= \gamma \cdot A \sum\nolimits_{i=c_1}^{c_r} v_i \sigma_i + S \\ &= \gamma \cdot \left(\sum\nolimits_{i=c_1}^{c_r} v_i A \sigma_i \right) + S \qquad (1.1) \\ &= \gamma \cdot \left(\sum\nolimits_{i=c_1}^{c_r} v_i B f_i \right) + Bw \bmod q \qquad (1.2) \\ &= \gamma \left(B\mu^1 \right) + Bw \bmod q \\ &= B \left(\gamma\mu^1 + w \right) \bmod q \\ &= B\mu \bmod q\end{aligned}$$

Firstly, we elaborate equation (1.1) as follows:

$$\begin{aligned}\gamma \cdot A\sigma + S &= \gamma \cdot A \sum\nolimits_{i=c_1}^{c_r} v_i \sigma_i + S \\ &= \gamma \sum\nolimits_{i=c_1}^{c_r} v_i A \sigma_i + S \\ &= \gamma \sum\nolimits_{i=c_1}^{c_r} v_i h_i + S \\ &= \gamma \sum\nolimits_{i=c_1}^{c_r} v_i \beta_i C + S \\ &= \gamma \sum\nolimits_{i=c_1}^{c_r} v_i \beta_i ({\alpha_1}^T, {\alpha_2}^T, \cdots, {\alpha_n}^T) + S \\ &= \gamma \sum\nolimits_{i=c_1}^{c_r} v_i H_1(id||i)[(H_1(ID||id||1)^T, \\ &\quad (H_1(ID||id||2)^T, \cdots, (H_1(ID||id||n)^T] + S \\ &= \gamma \sum\nolimits_{i=c_1}^{c_r} v_i H_1(id||i)[(H_1(ID||id||1)^T, \\ &\quad (H_1(ID||id||2)^T, \cdots, (H_1(ID||id||n)^T] + Bw \qquad (1.3)\end{aligned}$$

Now, comparing equation (1.3) in our formula derivation procedure with equation (1.2) in C. Sasikala et al.'s proven procedure, we observed that equation (1.3) and equation (1.2) both are the sum of two terms and the second term of them is the same Bw. Therefore, if equation $(1.1) = (1.2)$ holds, it can be inferred that equation $(1.3) = (1.2)$ should hold. In addition, we could get the first term of equation (1.3) should equal the first term of (1.2). However, the first term of equation (1.3) apparently has nothing to do with the data file block f_i and the cloud's public key B, so we can safely draw the conclusion that (1.3)= (1.2) will not hold in any case, i.e., in C. Sasikala et al.'s proven procedure (1.1)= (1.2) is not hold. Thus, their proven is incorrect.

2)In the second place, we further illustrate that signature algorithm is also incorrect because the precondition for applying the SamplePre algorithm not holds.

It is a common knowledge that when we want to utilize one formula, theorem or algorithm, we should first check whether the required applicable preconditions of them is satisfied. Take the preimage sampling algorithm $\sigma_i = Samplepre$ (A, sk, h_i, s) for example, there are four parameters in it and their relation should be meet the requirements that $A\sigma_i = h_i$ and $Ask = 0$ are hold, i.e., it should be satisfied that sk is the trapdoor of matrix A according to the definition of SamplePre algorithm. However, from the private key generation stage $Set - Privatekey(PParams, P_{ID}, S_{ID})$ it can be obtained that $sk = (P_{ID}, S_{ID})$,$AP_{ID} \neq 0$, $AS_{ID} \neq 0Ask \neq 0$, i.e., none of the three equations is equal to zero. Apparently, C. Sasikala et al.'s cannot satisfy the precondition of SamplePre algorithm, and thus, they are wrong in using the SamplePre algorithm to design the Certificateless RDIC scheme.

B. Security Analysis

Now we begin to demonstrate the following two concrete attacks of C. Sasikala et al.'s certificateless RDIC scheme, that is, signature forgery attack from the malicious CSP and public key replacement attack from the curious PKG.

1) signature forgery attack

To resist signature forgery attack or say unforgeability is the basic requirement for a RDIC scheme. If the CS can forge DO's signature, it is bound to defraud the TPA to convince that data is integrated, then the auditing for the CS become useless.

Unfortunately, in Certificateless RDIC protocol cannot satisfy the unforgeability because at the stage of SignGen, their SignGen algorithm is only using the index of file block and has nothing to do with the file block itself. Thus, anyone including the CSP can forge signature of the data only by data block's index and without the possess of DO's data, i.e., any malicious CSP can generate and response the valid proof without user's original data to deceive TPA by passing the data checking even if the data blocks are partial corruption. What is even worse, the CSP may deliberately erase the seldom visited data to gain more economic benefits. Therefore, their scheme fails to resist signature forgery attack. Therefore, their scheme cannot detect the data whether is tempered by the evil CSP.

2) public key replacement attack

In CL-PKC system, each secret key constitutes two components. The first part is produced by a partially trusted authority named Key Generation Center (KGC) or Private Key Generator (PKG), which can act as an implicit certificate for authenticate user and be verified while verifying the signature simultaneously, and the second part is self-selected random secret number by registrant. However, in CL-PKC system, the public key is not a unique identity information (like that in the ID-based cryptosystem) but an uncertainty secret number generated by himself randomly. Moreover, no certificate (like X.509) could certify the public key legitimate or valid. Therefore, the hostile could substitute it to conduct the public key replacement attack for capturing a typical CL-PKC system.

In Certificateless RDIC protocol of C. Sasikala et al. [31], the stage of signature verification has nothing to do with the public keyPk related to the secret numberss_{ID} which is selected by user himself, hence, it is not a rigorous certificateless scheme. Actually, it degenerates into identity-based signature scheme because the PKG have known the , which is generated by PKG, therefore, the malicious PKG can impersonate the legitimate user to forge the signature and the forged signature by PKG must be passed in the signature verification stage. Particularly, because CL-PKC system overcome the key escrow problem, it is generally supposed that PKG is no longer fully trusted but semi-trusted or partial-trusted. Even worse, other malicious users and adversaries can also easily substitute the targeted Pk to Pk' , where Pk' is generated by using ${s_{ID}}'$ which arbitrarily chosen by the hostile without knowing the s_{ID} and launch the public key replacement attack successfully since the stage of signature verification has nothing to do with Pk.

IV. Conclusion

We present cryptoanalysis of the Certificateless RDIC protocol proposed by C. Sasikala et at. Firstly, we formally prove that their scheme is incorrect from two aspect, i.e., neither their proven procedure of verification algorithm nor the precondition for applying the SamplePre algorithm in their signature algorithm is correct. Secondly, with respect to the security, we demonstrate that their protocol has two serious security vulnerabilities through two concrete attacks, i.e., signature forgery attack from the malicious CSP and public key replacement attack from the curious PKG. We further point out that C. Sasikala et al.'s protocol is not a rigorous certificateless protocol. It is well-known that design a secure and efficient lattice-based algorithm is an open and hard research subject full of challenge we expect that our cryptoanalysis is of great referential significance for cryptography researcher. Our future work will be aimed at exploring secure and efficient RDIC schemes for cloud computing from lattice.

References

[1] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling public auditability and data dynamics for storage security in cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 5, no. 22, pp. 847–859, 2011.

[2] M. Ali, S. U. Khan, and A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges," *Information sciences*, vol. 305, pp. 357–383, 2015.

[3] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in *Proceedings of the 14th ACM conference on Computer and communications security*. Acm, 2007, pp. 598–609.

[4] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, "Toward secure and dependable storage services in cloud computing," *IEEE Transactions on Services Computing*, vol. 5, no. 2, pp. 220–232, 2011.

[5] K. Yang and X. Jia, "An efficient and secure dynamic auditing protocol for data storage in cloud computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 9, pp. 1717–1726, 2012.

[6] X. Liu, Y. Zhang, B. Wang, and J. Yan, "Mona: Secure multi-owner data sharing for dynamic groups in the cloud," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1182–1191, 2012.

[7] B. Wang, B. Li, and H. Li, "Oruta: Privacy-preserving public auditing for shared data in the cloud," *IEEE Transactions on Cloud Computing*, vol. 2, no. 1, pp. 43–56, 2014.

[8] Y. Zhu, H. Hu, G.-J. Ahn, and M. Yu, "Cooperative provable data possession for integrity verification in multicloud storage," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 12, pp. 2231–2244, 2012.

[9] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Workshop on the theory and application of cryptographic techniques*. Springer, 1984, pp. 47–53.

[10] H. Wang, Q. Wu, B. Qin, and J. Domingo-Ferrer, "Identity-based remote data possession checking in public clouds," *IET Information Security*, vol. 8, no. 2, pp. 114–121, 2013.

[11] H. Wang, "Identity-based distributed provable data possession in multicloud storage," *IEEE Transactions on Services Computing*, vol. 8, no. 2, pp. 328–340, 2014.

[12] Y. Yu, M. H. Au, G. Ateniese, X. Huang, W. Susilo, Y. Dai, and G. Min, "Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 4, pp. 767–778, 2016.

[13] Y. Yu, L. Xue, M. H. Au, W. Susilo, J. Ni, Y. Zhang, A. V. Vasilakos, and J. Shen, "Cloud data integrity checking with an identity-based auditing mechanism from rsa," *Future Generation Computer Systems*, vol. 62, pp. 85–91, 2016.

[14] S. S. Al-Riyami and K. G. Paterson, "Certificateless public key cryptography," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2003, pp. 452–473.

[15] B. Wang, B. Li, H. Li, and F. Li, "Certificateless public auditing for data integrity in the cloud," in *2013 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2013, pp. 136–144.

[16] Y. Zhang, C. Xu, S. Yu, H. Li, and X. Zhang, "Sclpv: Secure certificateless public verification for cloud-based cyber-physical-social systems against malicious auditors," *IEEE Transactions on Computational Social Systems*, vol. 2, no. 4, pp. 159–170, 2015.

[17] D. He, N. Kumar, H. Wang, L. Wang, and K.-K. R. Choo, "Privacy-preserving certificateless provable data possession scheme for big data storage on cloud," *Applied Mathematics and Computation*, vol. 314, pp. 31–43, 2017.

[18] D. He, N. Kumar, S. Zeadally, and H. Wang, "Certificateless provable data possession scheme for cloud-based smart grid data management systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1232–1241, 2017.

[19] D. He, S. Zeadally, and L. Wu, "Certificateless public auditing scheme for cloud-assisted wireless body area networks," *IEEE Systems Journal*, vol. 12, no. 1, pp. 64–73, 2015.

[20] J. Li, H. Yan, and Y. Zhang, "Certificateless public integrity checking of group shared data on cloud storage," *IEEE Transactions on Services Computing*, 2018.

[21] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, 1997.

[22] Y.-K. L. D. M. R. P. R. P. D. S.-T. Lily Chen, Stephen Jordan. (2016) Report on Post-Quantum Cryptography. [Online]. Available: http://nvlpubs.nist.gov/nistpubs/ir/2016/NIST.IR.8105.pdf

[23] M. Ajtai, "Generating hard instances of lattice problems," in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, 1996, pp. 99–108.

[24] S. Agrawal, D. Boneh, and X. Boyen, "Lattice basis delegation in fixed dimension and shorter-ciphertext hierarchical IBE," in *Annual Cryptology Conference*. Springer, 2010, pp. 98–115.

[25] D. Boneh and D. M. Freeman, "Linearly homomorphic signatures over binary fields and new tools for lattice-based signatures," in *International Workshop on Public Key Cryptography*. Springer, 2011, pp. 1–16.

[26] W. Chen, H. Lei, and K. Qi, "Lattice-based linearly homomorphic signatures in the standard model," *Theoretical Computer Science*, vol. 634, pp. 47–54, 2016.

[27] C. Gentry, C. Peikert, and V. Vaikuntanathan, "Trapdoors for hard lattices and new cryptographic constructions," in *Proceedings of the fortieth annual ACM symposium on Theory of computing*. ACM, 2008, pp. 197–206.

[28] X. Zhang, H. Wang, and C. Xu, "Identity-based key-exposure resilient cloud storage public auditing scheme from lattices," *Information Sciences*, vol. 472, pp. 223–234, 2019.

[29] X. Zhang, C. Xu, H. Wang, Y. Zhang, and S. Wang, "Fs-peks: Lattice-based forward secure public-key encryption with keyword search for cloud-assisted industrial internet of things," *IEEE Transactions on Dependable and Secure Computing*, 2019.

[30] M. Ajtai, "Generating hard instances of the short basis problem," in *International Colloquium on Automata, Languages, and Programming*. Springer, 1999, pp. 1–9.

[31] C. Sasikala and C. S. Bindu, "Certificateless remote data integrity checking using lattices in cloud storage," *Neural Computing and Applications*, vol. 31, no. 5, pp. 1513–1519, 2019.

[32] J. Alwen and C. Peikert, "Generating shorter bases for hard random lattices," *Theory of Computing Systems*, vol. 48, no. 3, pp. 535–553, 2011.

[33] D. Micciancio and O. Regev, "Worst-case to average-case reductions based on gaussian measures," *SIAM Journal on Computing*, vol. 37, no. 1, pp. 267–302, 2007.