Indonesian Journal of Electrical Engineering and Computer Science
Vol. 2, No. 2, May 2016, pp. 452 ~ 460
DOI: 10.11591/ijeecs.v2.i2.pp452-460

Ensuring Data Integrity Scheme Based on Digital Signature and Iris Features in Cloud

[1]Salah H Refish*, [2]Zaid Ameen Abdul jabbar, [3]Zaid Alaa Hussien
[4]Thair A Kadhim, [2]Ali A Yassin, [2]Mohammed Abdulridha Hussain, [5]Salam Waley
[1]Huazhong University of Science and Technology, Wuhan, China
[2]University of Basrah, Basrah, Iraq
[3]Southern Technical University, Basrah, Iraq
[4]Directorate of Education-Babylon, Iraq
[5]University of Technology, Baghdad, Iraq
*Corresponding author, e-mail: manatheraa@yahoo.com

Abstract

Cloud computing is a novel paradigm that allows users to remotely access their data through web- based tools and applications. Later, the users do not have the ability to monitor or arrange their data. In this case, many security challenges have been raised. One of these challenges is data integrity. Contentiously, the user cannot access his data directly and he could not know whether his data is modified or not. Therefore, the cloud service provider should provide efficient ways for the user to ascertain whether the integrity of his data is protected or compromised. In this paper, we focus on the problem of ensuring the integrity of data stored in the cloud. Additionally, we propose a method which combines biometric and cryptography techniques in a cost-effective manner for data owners to gain trust in the cloud. We present efficient and secure integrity based on the iris feature extraction and digital signature. Iris recognition has become a new, emergent approach to individual identification in the last decade. It is one of the most accurate identity verification systems. This technique gives the cloud user more confidence in detecting any block that has been changed. Additionally, our proposed scheme employs user's iris features to secure and integrate data in a manner difficult for any internal or external unauthorized entity to take or compromise it. Iris recognition is an internal organ that is well protected against damage and wear by a highly transparent and sensitive membrane. Extensive security and performance analysis show that our proposed scheme is highly efficient and provably secure.

Keywords*: Cloud computing; data integrity; iris features; digital signature*

1. Introduction

There are many beneficial characteristics of cloud computing, such as being on-demand self- service, broad network access, resource pooling, rapid elasticity, and measured service [1]. Conversely, there exist many security challenges [2, 3]. Cloud storage which is supplied by the cloud server and provided to the cloud users as a service is considered one of these challenges. On the one hand, cloud infrastructures are more powerful and reliable than personal computing devices, although internal and external threats to data integrity still exist. On the other hand, there exist various incentives for the cloud service provider (CSP) to behave dishonorably towards cloud users, such as financial reasons or reputation. All these issues arise because once the cloud users outsource their data to the CSP they no longer have possession of a local copy of their data. At the same time, cloud users lose the ability to monitor and control their data in the cloud, so it can be easily corrupted, modified, or deleted due to hardware failure or human errors.

Thus, protecting the integrity of data is highly essential and security challenge in the cloud. Additionally, the data stored in the cloud is not only accessed but also frequently updated by cloud user, including insertion, deletion, modification etc. Thus, it is imperative to support the dynamic features of cloud storage. The process of saving data in the remotely located cloud servers is called cloud storage [4].Cloud users can upload their data to the cloud and can access these data anytime and anywhere. There are key characteristics that make cloud storage better than traditional storage. These characteristics are (1) performance: with this

Received January 5, 2016; Revised April 18, 2016; Accepted May 1, 2016

feature the cloud can move huge amounts of data over the global internet; (2) manageability: cloud storage reduces the burden of maintenance at the client side when data is stored remotely in the cloud; (3) availability: in cloud storage, data are retrieved frequently, rapidly, and securely [5].

Most researchers have been working to introduce the best options to the cloud users on possession and integrity of data [6-11].Table I shows the main differences between our scheme and other schemes. In this paper, we present a method for ensuring the integrity of data. This method is based on biometric technology, which is considered one of the modern approaches in the security field. Generally, biometric employs physiological or behavioral characteristics to precisely identify each subject. Commonly used biometric features include the face, fingerprints, voice, iris, retina, gait, palm print, hand geometry, dental radiograph, etc. Our work involves the iris. Iris recognition has become a new, emergent approach to individual identification during the last decade [12]. In our proposed method, the iris features are directly obtain from the cloud user. We provide an approach that supplies proof of data integrity which the cloud user can employ it to check the correctness of his or her data in the cloud. Additionally, we propose an efficient and secure possession and data integrity scheme based on feature extraction from cloud user's iris and digital signature to increase the level of security. So, we can summarize our contributions as follows:

- Our work involves the iris. Iris recognition has been considered an effective approach in carrying out individual identification during the last decade.
- The key factor employed in our scheme is based on user' iris features, which has been shown to be more security against known attacks.
- We present a method for ensuring the possession and integrity of data as well as data dynamics.
- The digital signature is introduced to support our scheme in verification phase.

The rest of this paper is organized as follows: Section 2 illustrates design issues and cryptographic primitives. The details of our proposed scheme are presented in Section 3. Section 4 addresses support data dynamics. Security analysis and performance of our work are shown in Section 5. The conclusion in Section 6.

2. Design Issues

2.1. Problem Definitions

We consider a cloud storage system as consisting from three parts as follows: (1) client, who has the data files and he or she wants to be stored in the cloud. The client stores data on the server without keeping a local copy. Hence, it is of critical importance that the client should be able to verify the integrity of the data stored in the remote non-trusted server. (2) Cloud server, which is a managed data storage service. If the server modifies any part of the client's data, the client should be able to detect it. (3) The TPA, who has expertise and capabilities that users don't have and is trusted to estimate the cloud storage security on behalf of the user's requests. In case a third-party auditor verifies the integrity of the client's data, the data should be kept private against the third-party auditor.

Figure (1) illustrates these different entities. Clients rely on the cloud server for data storage and maintenance. They may also frequently access and update their data for various application purposes. To ensure their data is secure in cloud storage the users may resort to contacting the TPA, while the data should be kept private. We assume the cloud server provider may be dishonest in two ways and affect directly the user's data: (1) delete rarely accessed data to decrease the storage cost, (2) decide to hide the corrupted data caused by server hacks or Byzantine failures to keep reputation. The TPA should be able to verify the integrity of data without a local copy of the data. However, any information leaked to the TPA through the verification process should be prohibited. Figure (2) shows our proposed scheme.

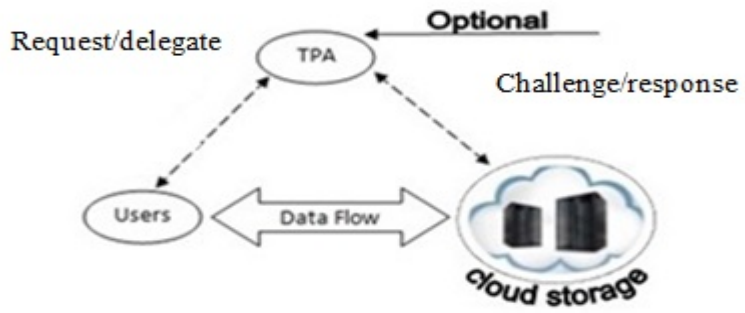

Figure 1. The traditional of auditing schemes

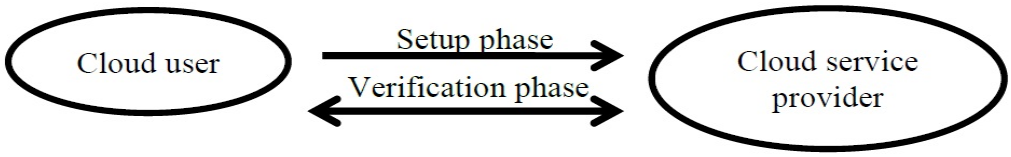

Figure 2. Our scheme architecture

2.2. Iris Recognition

Generally, biometrics employs physiological or behavioral characteristics to precisely identify each subject. Commonly used biometric features include the face, fingerprints, voice, iris, retina, gait, palm print, hand geometry, dental radiograph, etc. Iris recognition has become a new, emergent approach to individual identification in the last decade. The iris of the eye is made up of a series of holes and cracks which are concentrated around each iris which vary from one person to another in terms of the number, shape and even the distance between them. As well, the pigments of the iris vary from one person to another, even if involving the degree of colour, because there are large differences in colour within the same footprint, i.e. what constitutes something distinct and unique to the eye and is the imprint of the iris. This is one of the best methods of security that enables one to confirm the identity of a person. Iris recognition is one of the most accurate identity verification systems. Accurate automatic personal identification is becoming more and more significant to the operation of security systems. A typical iris recognition system is graphically shown in Figure (3). The whole iris recognition process is basically divided into four steps [12, 13]:

- Image acquisition;
- Iris image preprocessing;
- Iris feature extraction; and
- Matching.

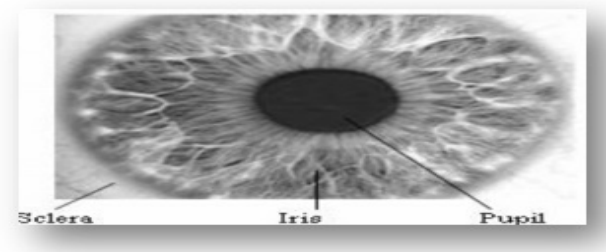

Figure 3. The Human Iris

The iris biometric deals with identifying a human being by his/her iris pattern extracted from theimages of his/her eye. As shown in Figure (3), the human eye consists of three major parts: pupil (the innermost black part), iris (the coloured part) and sclera (the white part). The iris and pupil are said to be non-concentric. The radius of inner border of the iris, i.e. its border with the pupil, is also not constant since the size of the pupil increases and decreases depending on the amount of light incident to the pupil. Each individual has a unique iris pattern. This pattern can be extracted from the image of the eye and encoded. The code can be compared to the codes obtained from the images of other eyes or the same eye. The result of comparison can represent the amount of difference between the compared codes. In that way it can be concluded if the compared eye patterns belong to the same or different eye. Compared to other biometrics, such as voice and facial features, which tend to change over time, the iris biometric is stable and remains the same for a person's lifetime [14]. The use of contact lenses, glasses and even eye surgery cannot affect the iris characteristics.

The iris imprint reader works through reading and storing it as an array: the length of the image is in columns and the width is in rows. This is sent to a computer to get a 512 byte template, according to iris characteristics which will be matched with the stored data. Although, the iris recognition is the strong way in the identification systems as mentioned above, however, it requires software and hardware costs. In our work, we do not need the software and hardware overhead. Just first time the user takes his iris features and then stores it in the USB device. The figure (4) distinguishes between these ways according to the costs.

If we are using the traditional way (A) and suppose we have 1000 users wish to login the system at the same time. It is very difficult to just imagine that, how long time we need to get the iris features as well as the costs of the hardware and software. So, because we are using the cloud environment, we should benefit from its facilities. The term "pay as you go" is a great solution in this field. We can rely on the cloud services providers such as [Google, Amazon, and Microsoft] which areproviding (Iaas, Paas, and Saas). We should say here, using the cloud services do not force us to involve it every time for financial reasons. So, in our work, we exploit cloud provider just to obtain iris features to use it in the cloud server later. (B) Shows the mechanism for generating iris features based on cloud service provider in progress for each user and then he stores it in his USB which is used in login system and verification process later that does not required software and hardware for obtaining iris at first, and do software operations such as preprocessing, feature extraction and classification.

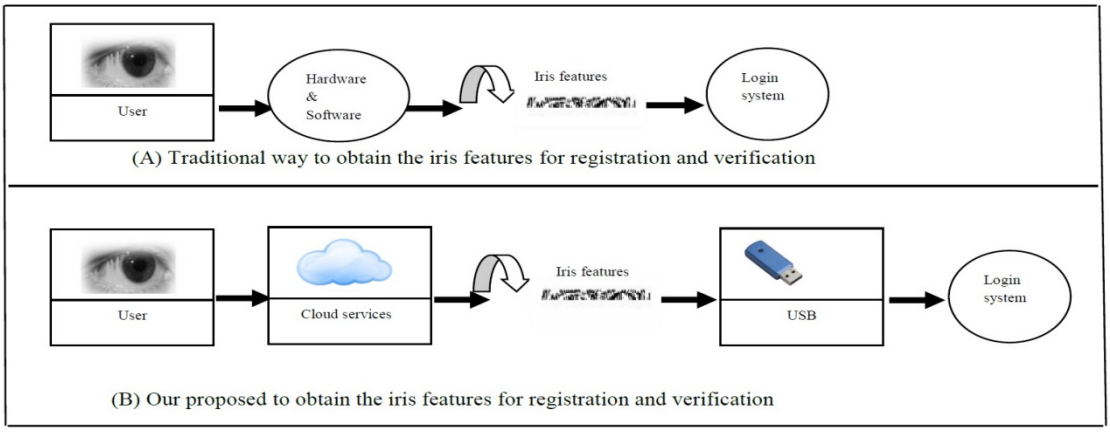

Figure 4. The main difference between traditional way (A) and our proposed (B) for obtaining iris features

2.3. Schnorr Digital Signature

Generally, Schnorr Digital Signature presented a scheme relying on ElGamal digital signature, but with minimizes signature size. It is very attractive security, qualified and generates short signatures. We review Schnoor signature scheme as follows [15].

Key Gen: Let *p, q* are large primes. User selects an element e_1 of prime order *q*, and computes $e_2=e_1^{d \bmod p}$. User's public key is (e_1, e_2, *p, q*) and his private key represents by *d*, where $1 \leq d \leq q$ is uniformly selected.

Sign (e_1, d, M): He picks a random number *r*, when he wants to sign message, he needs to change *r*. The value of *r* is between 1 and *q*. Cloud user computes the first signature $S_1 = H(M||e_1^r \bmod p)$. The signing message (*M*) is based on the value of $e_1^r \bmod p$, where hash function is applied to the concatenation function (||) of *M* and $e_1^r \bmod p$. Then, cloud user computes the second signature $S_2 = r + d\, S_1 \bmod q$ and sends (M, S_1, S_2) to the Verifier.

Verify (M, e_1, e_2, S_1, S_2): Verifier receives *M,* S_1 and S_2, then computes $V = H(M||e_1^{S_2} e_2^{-S_1} \bmod p)$. Finally, Verifier compares *V mod p* with S_1, if the result is true, the massage is accepted; otherwise, it is rejected.

3. Our Proposed Scheme

The generate meta-data which is based on the features of the iris combined with the original data to produce scramble data. In the verification process, the client wants to ensure that his data in safe or not. The verifier (client) prepares a challenge for the target server and asks the server to respond. The challenge detects the number of the original block as well as the related signature that is possessed to be verified. The specified server replies with two values: the original data block and the signature. Here, this time, the verifier uses his or her iris features to decrypt the meta-data and ensures that the decrypted value matches with the original data. If the result is true the integrity of data is confirmed.

3.1. Installation Process

In this process we have two stages as follows: (Figure 5 Shows the Installation and verification processes).

- The cloud user wants to store his data in the cloud server, so he should make some operations as follows:
- Assume that the input file *F* is divided into *m* data blocks by using the data fragment technique where each of the blocks involves \emph{n} sectors.
- The cloud user generates random key *k*, and then uses any symmetric algorithm to encrypt the data blocks. $F'=(k, F)$.
- The metadata are generated from the equation (1) as : $(T_i)=(H(m_i||F_{id})||fi) \rightarrow (1)$, where F_{id} is the identifier of the file..
- The cloud user then sends (F', T_i) to the CSP.

3.2. Verification Process

- The cloud user (or third party auditor) is required to generate a challenge message. This message consists of c data blocks randomly as a challenge message ($chal = CSi_{i=1}^{c}$) by using pseudo-random permutation [16] keyed with a fresh randomly-chosen key to prevent the server from anticipating the block indices.
- Cloud server sets up public key of Schnorr's digital signature $PK_{Schnorr} = (e_1, e_2, p, q)$ and private key: $SK_{Schnorr} = d \in Z_q^*$.
- When the challenge message is received by the cloud server, the proof message, including aggregation authenticator tags T_i and a linear combination of the blocks $\sigma=\sum_{i=CS1}^{CSc} F_i'$, where *i* is the index of the block. This proof is generated based on the challenge message$T'=.\sum_{i=CS1}^{CSc} T_i || \sigma$
- The cloud server will Apply digital signature schnorr to produce two values S_1, S_2: $S_1=h(T'||e_1^{r_i} \bmod p)$, $S_2=r_i+ds_1 \bmod q$.Then send (S_1, S_2) to the cloud user.
- The cloud user will do some operations as follows: Compute $V=h(T||e_1^{S_2} e_2^{-S_1} \bmod p$.Verify whether $V=S_1$. If true, the cloud user knows that the data is stored correctly.

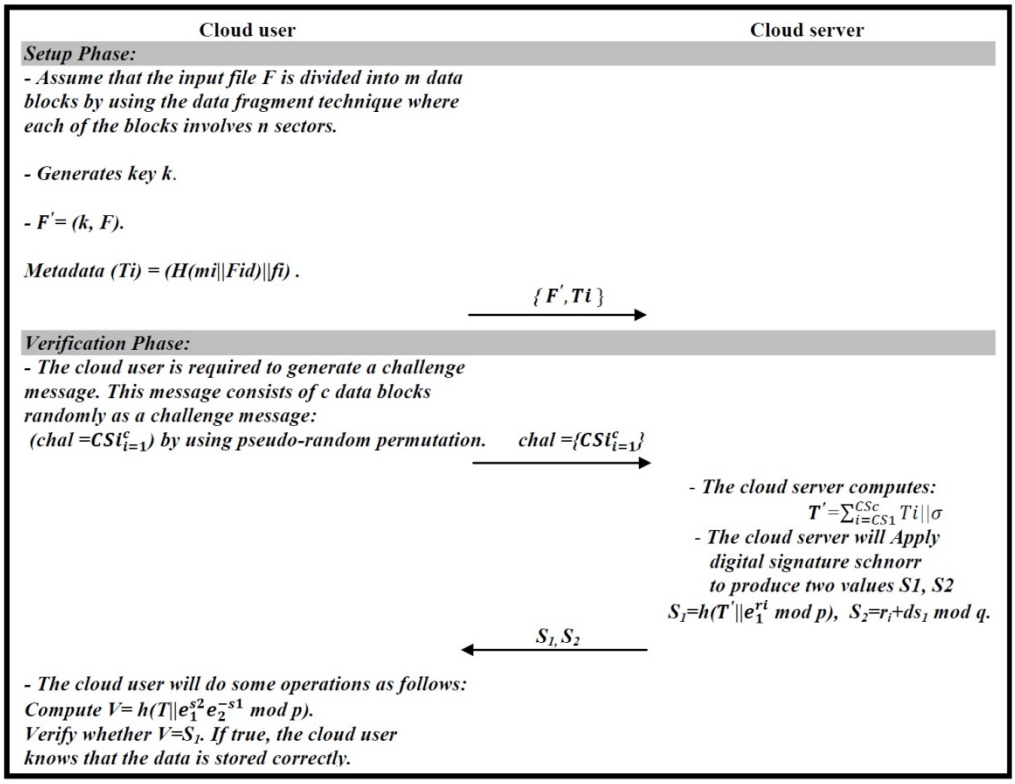

Figure 5. The Verification process in our proposed scheme

4. Providing Data Dynamics

We supposed that *F* represents static or archived data. This approach can be adapted to some application scenarios, such as libraries and scientific databases. However, in cloud data storage, there are many possible scenarios where the data stored in the cloud is dynamic, as for instance e-documents, photos, log files, etc. Therefore, it is crucial to consider the dynamic case, where a user may wish to perform multiple functions (block level update, deletions and additions that modify the data), while keeping corrections, ensuring security of storage.

Data modification: We set out from data modification, which is one of the most repeatedly used operations in cloud data storage. A fundamental data modification operation indicates the replacement of particular blocks with new ones. Imagine that the cloud user demands the modification of the data block m_i. Firstly, this depend on the new data blockNm_i, the cloud user produces corresponding metadata (T_i)=($H(Nm_i||F_{id})||fi$). Then, he or she generates an *update challenge=(M,i,* T_i'*, Nm$_i$)* and sends it to the cloud server, where *M* denotes the modification operation. Upon receiving the request, the cloud server (1) replaces $m_i$withNm_i; (2) replaces$T_i$withT_i'.

Data insertion: Compared to the modification process, which does not manipulate the logic structure of the cloud user's data file, there is one form of data operations, data insertion, which indicates by inserting other blocks following some specified positions in the original data file. For instance, the cloud user requests to add block Nm_i after *i*th block m_i. This procedure is similar to the modification process. Firstly, depending on $Nm_i$the cloud user produces an *update challenge= (I, i,* T_i'*, Nm$_i$)* and sends it to the cloud server, where *I* refers to the insertion operation. Upon receiving the request, the cloud server (1) adds Nm_iafterm_i; (2) adds T_i' after T_i.

Data deletion: This is the opposite operation to data insertion. It indicates removing the particular block and moving all the following blocks one block forward. For instance, the cloud server receives the *update challenge=(D,i)* for deleting of m_i, where *D* denotes the deletion operation. The cloud server deletes each of m_i, T_i from its storage space.

Table 1. Comparison of different integrity verification schemes with our scheme

Proposed scheme	Privacy-preserving	Dynamic operations	Unlimited number of queries	Public verifiability	Recoverability	Non-trusted server
Our scheme	Yes	Yes	Yes	Yes	Yes	Yes
[6]	No	No	No	Yes	No	Yes
[7]	No	No	No	No	No	Yes
[8]	No	No	No	No	No	Yes
[11]	No	Yes	No	No	No	Yes
[17]	No	Yes(partially)	No	No	No	Yes
[18]	No	Yes	Yes	Yes	No	Yes
[19]	Yes	Yes	Yes	Yes	No	Yes

5. Security Analysis of Our Proposal

In this section, we review a formal analysis of the security features of our scheme.

Theorem 1. Our proposed scheme provides privacy protection.

Proof. The input file *F* is divided into *m* data blocks by using the data fragment technique where each of the blocks involves *n* sectors. After that, the cloud user generates random key *k*, and then uses any symmetric algorithm to encrypt the data blocks. $F'=(k, F)$. The metadata are generated, the cloud user then sends (F', T_i) to the CSP. So, there is no way to learn the content of the data file. Therefore, our scheme provides privacy protection.

Theorem 2. If the CSP can generate a valid proof that passes the *VerifyProof* phase of the verifier, then it must indeed possess the specified intact data. So, our proposed scheme provides data storage correctness.

Proof. The metadata are generated from $T_i=(H(m_i||F_{id})||fi)$, where F_{id} is the identifier of the file. Any adversary or attacker will have difficulty forging (T_i). Thus, a malicious CSP cannot tamper with a valid response $T'=\sum_{i=CS1}^{CSc} T_i||\sigma$ to pass the verification phase by the verifier because the cloud server will apply digital signature schnorr to produce two values S_1, S_2, $S_1=h(T'||e_1^{r_i} mod\ p)$, $S_2=r_i+ds_1 mod\ q$. Then send (S_1, S_2) to the cloud user (verifier). After that, the cloud user computes $V=h(T||e_1^{S_2}e_2^{-S_1} mod\ p)$, and verify whether $V=S_1$. If true, the cloud user knows that the data is stored correctly. So, our scheme provides data storage correctness.

Theorem 3. Our proposed scheme can withstand the off-line guessing and forgery attacks.

Proof. In our proposed scheme, the active attacker demeanors such as impersonation do not gain him/her any profit byapplying off-line guessing attack, because the cloud user will not replay unless he checks the honest of the CSP. Anattacker is not able to compute *V* since he does not have the ability to get the values of (f_i, e_1, e_2, p, q), because our proposed scheme prevents disclosing any information through the communication protocol between cloud user and CSP. Therefore, our proposed scheme resists the off-line and forgery attacks.

Theorem 4. Our work can supply recoverability.

Proof. Generally, when the verifier can detect the corrupted data, he or she executes the data recovery process for salvaging and handling of the data. In our proposed scheme the cloud user calculates $V=h(T||e_1^{S_2}e_2^{-S_1} mod\ p)$, then Verify whether $V=S_1$. If false, the cloud user knows that the data is not stored correctly; and the data was tampered with illegally. At that time, the cloud user should return the m_i,T_i to the cloud server to recover the original data which is modified. So, this scheme can supply recoverability.

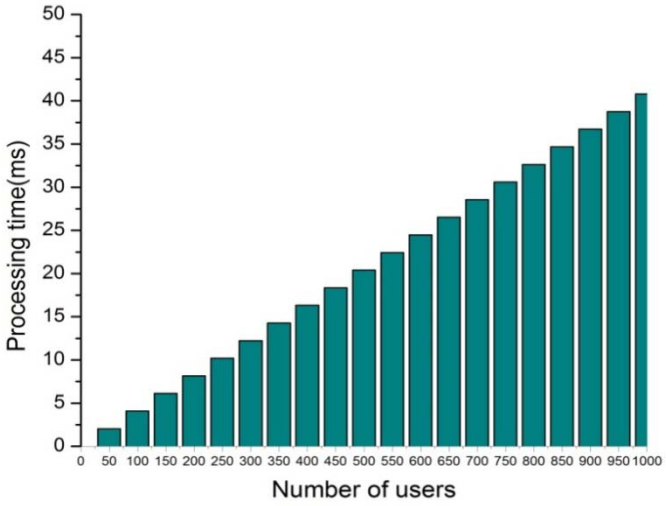

Figure 6. The performance of our auditing scheme

The cloud user generates and encrypts the data files before storing them in the cloud server. The iris feature extraction ensures efficiency, high performance, and security. The efficiency of our work has been tested by measuring the response time of the cloud server. Our work has been executed and tested on a database containing iris features of many users. These iris features were acquired randomly from the cloud user. We show the performance of our scheme in Figure 6 which is more efficient and suitable in cloud environment.

6. Conclusion

In this paper, we produced an approach for data integrity in cloud computing and involved the iris features as well as the digital signature to achieve the correctness of data. In such a way, the cloud user is assisted in confirming that the data is not accessed from unauthorised entities that utilise the cloud server. Additionally, the cloud user can trust uploaded data in any situation. The notion of our proposal involves obtaining integrity in cloud data storage with powerful reliability in order for the users to not have to worry about uploading their data. In this proposed scheme we have important features compared to prior, related work (Table I). Additionally, this work enjoys many security features such as: forward secrecy, data storage correctness, biometric agreement, and privacy protection. The performance and security analysis show that our scheme is efficient and secure against unauthorised servers and users. So, it is extremely convenient for cloud storage systems.

References

[1] Mykletun E, Narasimha M, Tsudik G. Authentication and integrity in outsourced databases. *Trans. Storage*. 2006; 2(2): 107-138.
[2] Ashish Kumar. World of Cloud Computing and security. *International Journal of Cloud Computing and Services science (IJ-CLOSER)*. 2012; 1(2): 53-58.
[3] *Sean Carlin, Kevin Curran.* Cloud Computing Technologies. *International Journal of Cloud Computing and Services science (IJ-CLOSER)*. 2012; 1(2): 59-65.
[4] Wang C, Wang Q, Ren K, Cao N, Lou W. Towards Secure and Dependable Storage Services in Cloud Computing. *IEEE Transactions on Services Computing*. 2012; 5(2): 220-232.
[5] EswaranS.Abburu S. Identifying Data Integrity in the Cloud Storag. *International Journal of Computer Science Issues*. 2012; 9(2): 403-408.
[6] Ateniese G, Burns R, Curtmola R, Herring J, Kissner L, Peterson Z, Song D. *Provable Data Possession at Untrusted Stores*. Proceedings of the 14th conference on Computer and Communication Security (CCS '07). Alexandria, USA, ACM. 2007; 598-609.
[7] Juels A, Kaliski BS Jr. *Pors: Proofs of Retrievability for Large Files.* Proceedings of the 14th conference on Computer and Communicatio Security (CCS '07), Alexandria, USA, ACM. 2007: 584-597.

[8] Bowers KD, Juels A, Oprea A. *HAIL: A High-Availability and Integrity Layer for Cloud Storage*. Proceedings of the 16th conference on Computer and Communications Security (CCS '09), Chicago, IL, USA, ACM. 2009: 187-198.
[9] Lin HY, Tzeng WG. A secure erasure code-based c1oud storage system with secure data forwarding. *IEEE Transaction onParallel and Distributed Systems*. 2012; 23(6): 995-1003.
[10] WWang Q, Wang C, Li J, Ren K, Lou W. *Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing.* Proceedings of the 14th Europeanconference. Research in Computer Security (ESORICS'09), Saint Malo, France. 2009: 355-370.
[11] E Erway C, Kupcu A, Papamanthou C, Tamassia R. *Dynamic Provable Data Possession.* Proceedings of the 16th conference on Computer and Communications Security (CCS '09), Chicago, IL, USA, ACM. 2009: 213-222.
[12] P Pravin S. Patil, Kolhe SR, Patil RV. The Comparison of Iris Recognition using Principal Component Analysis, Log Gabor and Gabor Wavelets. *International Journal of Computer Applications*. 2012; 43(1): 29-33.
[13] Kalka ND, Jinyu Z, Natalia AS, Bojan CL. Image quality assessment for iris. *Defense and Security Symposium*. 2006: 1-11.
[14] J Daugman J. Recognizing persons by their iris patterns. *Advances in Biometirc Person Authentication*, Springer. 2005: 5-25.
[15] BForouzan B. A. Cryptography and Network Security. 1st Edition. McGraw-Hill companies. Inc. 2008.402-403.
[16] R Impagliazzo R, Levin LA, Luby M. *Pseudo-random generation from one-way functions*. Proceedings of the Twenty-First Annual ACM Symposium On Theory of Computing, ACM, Seattle,Washington, United States. 1989: 12–24.
[17] A Ateniese G, Pietro RD, Mancini LV, Tsudik G. *Scalable and Efficient Provable Data Possession.* Proceedings of the 4th International Conference on Security and Privacy in Communication Networks (Secure Comm'08), Istanbul, Turkey. 2008: 1-10.
[18] C Wang C, Wang Q, Ren K,Lou W. *Ensuring data storage security in cloud computing.* Proceedings of the17th International Workshop on Quality of Service (IWQoS '09), Charleston, South Carolena, IEEE. 2009: 1-9.
[19] Z Hao Z, Zhong S, Yu N. A Privacy-Preserving Remote Data Integrity Checking Protocol with DataDynamics and Public Verifiability. *IEEE Transactions on Knowledge and Data Engineering.* 2011; 23(9): 1432-1437.