

TOWARD TO BUILD STRONG IMAGE INTEGRITY SCHEME IN CLOUD COMPUTING ENVIRONMENT

¹HAITHAM ALI HUSSAIN, ²AQEEL A. YASEEN²

¹Management Technical College of Basrah, Southern Technical University, basrah, Iraq, ² Technical Computer Engineering, Al Kunooz University College, Basra, Iraq
E-mail: ¹haitham.ali@stu.edu.iq, ²aay.ali@kunoozu.edu.iq

ABSTRACT

Cloud computing aims to assist users in remotely preserving their data in a cloud server and protect them from malicious attacks, such as an impersonation attack. In addition, cloud computing gives users the ability to store and process their data in an outsource cloud's storage. With the rapidly increasing use of medical images with a growing level of detail due to advances in diagnostic medical images, cloud computing is used to store large amounts of medical images. Although the cloud has considerable potential, it has many security challenges, such as data integrity and unauthorised access of entities to cloud resources. This paper aims to solve the issue of ensuring the integrity of medical image data stored in cloud servers. Thus, we propose a robust scheme that uses cryptohash function and k-nearest neighbour (KNN) to obtain the metadata of images stored in the cloud, thereby protecting the medical images owned by users against any editing, deleting or inserting operations performed by an attacker. Our proposed scheme uses KNN to retrieve medical images to ensure their integrity and has robust security and high performance

Keywords: *Medical Image, Cloud Computing, Meta-Data, Crypto-Hash Function, Knn*

1. INTRODUCTION

The template details Cloud computing is a modern technology that supports several services and applications via the Internet. The advantages of cloud are not restricted to users or individuals but extend to many companies, governments and private institutions. Despite security and privacy issues due to the exploitation of the benefits of cloud, which may slowdown performance and hinder success for many users, cloud components are popular and essential. Security issues should be studied on a large scale to find appropriate solutions to increase the acceptance and success of the cloud. Cloud service providers allow users to store their data in the cloud without worrying about the integrity of user data. Users can upload important data to cloud servers and access their data anytime and anywhere. The main problem of users is that they lose control of their outsourced data in the cloud. Therefore, users require proof that their data are saved in the cloud. Cloud storage does not employ techniques that might help preserve the integrity of user data. Image integrity is considered one of the most critical components in any system; when image integrity deals with a single database, it is responsible for the safety or

preservation of the database through a series of restrictions [1-6].

This case is relatively different in distributed systems because these systems contain and handle numerous applications and databases, thereby causing security or privacy problems. Hence, we will attempt to identify particular ways to avoid such problems. Image integrity in cloud computing means ensuring the integrity of a remote image saved in unreliable cloud servers. A protocol is utilised in this case to obtain ownership of the images in the cloud, and this protocol will prove that the images stored by the real user in the cloud are not modified or updated by any archive; hence, the integrity of the images is considered authoritative. This verification system prevents any cloud storage archive from changing or manipulating images without the permission of the image owner. Medical images outsourced in the cloud can support the necessary information for a health centre, doctors and patients that require management in many branch hospitals, thereby decreasing the information and computational resources used in the hospital. Moreover, obtainable medical equipment can be regenerated to be more resourceful and low-cost as medical terminal units [7-13]. Many schemes have been introduced for the exchange, storage and

distribution of medical images in away that validates data integrity, confidentiality and availability. In this paper, we focus on the following important and necessary cases:

1. Medical image integrity: This characteristic allows users to confirm that the data stored in the cloud are secure and the validity of image integrity is ensured.
2. No data leakage: The content of data is stored in the cloud and can be accessed by authorised users only.
3. We propose an efficient and secure image integrity scheme that is dependent on the cryptohash function and k-nearest neighbours (KNN).

The remainder of this paper is organised as follows: Section 2 describes the design issues. Section 3 provides our proposed scheme. Section 4 resents the experimental result. Section 5 Concludes this paper.

2. DESIGN ISSUES

2.1. Problem Definition

When a user wants to store images with a cloud service provider, a pay-as-you-go service can be used to obtain sufficient storage space. The cloud server must accept those images; hence, stored images in the cloud pass through three main components, as shown in Figure1. The first component is called cloud user (CU) or data owner (DO), which prepares the medical image files to be saved in the cloud. The second component is the cloud server (CS), which contains user's photos. The third component is the cloud service provider (CSP), which is an important component controlled by all cloud servers. Figure 2 shows the basic architecture of our proposed image integrity scheme, which requires the three previously mentioned main components.

In our proposed scheme, we must apply the following two phases:

2.1.1. first phase(configuration phase)

The first phase is divided into three steps. The DO computes the metadata (M) for each medical image to be outsourced in the CS. The DO also maintains the metadata in the server. Figure 3 shows the mechanism of this phase.

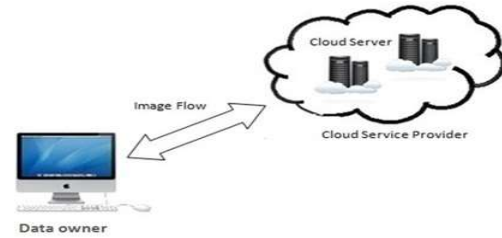


Figure 1: The architecture of cloud image storage service

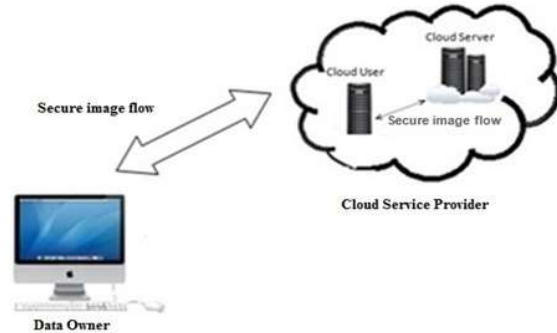


Figure 2: The structure of proposed scheme

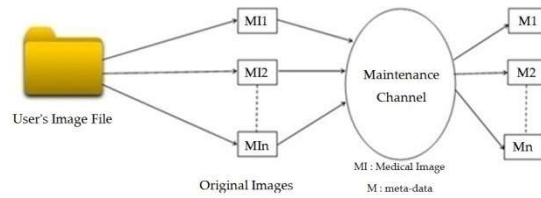


Figure 3: demonstrates the mechanism of computing meta-data

2-1.2 Second Phase(Verification Phase)

The DO aims to ensure the integrity of the medical I mages. Hence, the DO must send a query to the CSP on random images to check whether their images have been compromised. Thereafter, the CSP responds to the DO by sending the queried images. The DO then recomposes new metadata (M') for the queried images and compares M' with M for the same index in the DO side. If the result matches, then the DO's images are secured. Otherwise, the DO detects the non-integrity of images. Notably, the proposed scheme exhibits ensures privacy, good performance and efficiency.

This section may be divided by subheadings. It should provide a concise and precise description of what data is contained, which format, how to read and interpret the data. E.g., for tabular data a note

about what’s contained in each column of the data table.

2.2. Cryptographic Hash Function

Considerable research interest in cryptographic hash functions has been observed in recent years, especially after the popular attacks against Message Digest 5 (MD5) and Secure Hash Algorithm 1 (SHA-1) in 2005. Cryptographic hash functions are frequently called message digest functions and have many information security applications, notably in digital signatures, message authentication codes and other forms of authentication. Cryptographic hash functions are used to extract a fixed-length bit string from a message (image, documents).

A cryptographic hash function is an algorithm and mathematical function that converts a numerical input value into another compressed numerical value. The input to the hash function, Message M (arbitrary values), is returned (produced) by a hash function and are hash values (fixed length) that are mostly small in size. The following picture illustrates the hash function:

Most cryptographic hash functions are designed to take a string of any length as input and produce a fixed-length hash value. Cryptographic hash function has many uses. One of its most important application is integrity testing. Any changes in the input data can change every bit in the resulting hash value. Thus, changing even one bit of the input will change the output dramatically. Hence, hashes are useful in detecting any modification in a data object, such as a message, or various manipulations in multimedia data and digital images, such as compression, enhancement, cropping and scaling [14].

The use of encryption technology is not limited to the introduction of data (texts). Recently, major developments in the use of hash functions with image encryption or other multimedia applications for security and indexing have attracted considerable attention.

A key feature of conventional cryptographic hashing algorithms, such as MD5 and SHA-1, is that they are sensitive to the message, that is, the image hash function should instead consider the changes in the visual domain and produce hash values based on the image’s visual features.

2.3. k-NN(k-nearest neighbors algorithm)

k-NN is one of the algorithms that are simply characterised by their understanding and are considered versatile and have diverse applications. K-NN can be used for classification and predictive

regression problems. However, it is more widely used in classification problems in the industry. A simple working algorithm is stored where all available cases are included and new cases are classified based on a similarity scale (for example, distance functions). K-NN was employed in the early 1970s as a non-parametric technique by using statistical estimation and identifying patterns.

Its work is mainly based on the classification of an issue according to a majority vote of its neighbours, with the situation allocated to the most common and closest group amongst its neighbours by measuring the distance function, where $K = 1$; in the end, the situation is simply defined for the nearest neighbour [15].

3. PROPOSED SCHEME

The notations presented in Table 1 are used in our proposed scheme based on three components: CSP, DO and CU.

The DO should perform some processes on their original images, which represent the medical image files (MIF), before outsourcing the images to the CSP. The DO provides each image with the appropriate metadata, which is employed in the next phase of verification to ensure the integrity of their images stored in the cloud.

When the DO wishes to verify the integrity of MIF, the DO sends a challenge to the authenticated CSP and waits for its response. The CSP replies to the DO by sending the queried image. Hence, the CSP computes a new metadata for the queried image and checks if the new metadata value matches that of the existing original image in the DO’s server. If the result is correct, then the image is regarded as secured. Otherwise, the DO concludes that the CSP has been hacked. Therefore, our proposed scheme has two phases: setup and verification phase.

Table 1: Notification of symbols

Symbol	Definition
CU	Cloud user
DO	Data owner
CS	Cloud server
CSP	Cloud service provider
MIF	Medical image file
MD	Metadata
IF	Index file
MI	Medical images
Meta(I,j)	It refers to the j'th byte in the I'th block of a meta-image file

3.1 Setup Phase

The DO has a set of medical images that are outsourced to the CSP. The DO wants to offer a service to cloud users (CUs), such as doctors, researchers and programmers. The DO also checks the performance of his images from time to time to ensure image integrity and protects images against malicious attacks. The setup phase consists of the following three steps:

3.1.1. The DO collects his medical images into a single file called (Mif).

$$Mif = \{MI1, MI2, MI3, \dots, \dots, MIn\}$$

3.1.2. The DO computes the metadata for each image based on the following equation:

$$MDi = h\left(\sum_{i=1}^n \sum_{j=1}^m MIk(i, j) * (i + j)\right) \quad (1)$$

Where:

n = number of rows in MIi

m = number of columns in MIi

k = selected image

h = cryptography hash function (MD5)

Then, the DO computes index file (IF), which contains the sequence, secure name of each image, metadata and image. Next, the DO outsources IF without metadata to CSP. The DO stores the metadata in his server. The secure name of each image is computed based on the following equation:

$$Secureimg_i = h(name_i, MD_i, i) \quad (2)$$

The information of current image (i, Secureimg_i, MD_i, MI_i) is added to the IF as a new record.

3.1.3. The DO sends the image file (Mif) with IF to the CSP, but IF has no metadata, whereas DO stores metadata in CS. The DO's image will then become available for use from the CU (such as doctors and researchers). Figure 4 explains the main steps of the setup phase.

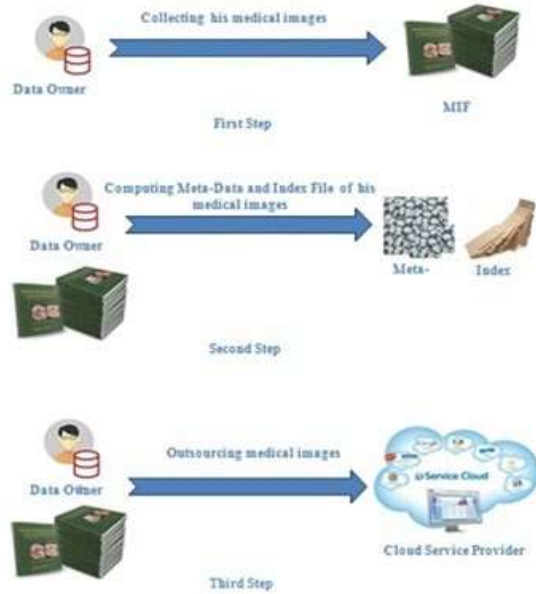


Figure 4: Setup Phase

3.2. Verification Phase

This phase is the most important phase in our work because it checks the integrity of DO images in the CSP. This phase consists of three steps:

3.2.1. first step

The DO randomly selects an image and sends a challenge to the CS (contain the index of image[K]). Upon receiving the request of the DO, the CSP will use the KNN algorithm to perform matching between the received image and the images in the MIF. If a match is found, then the requested image (M'_K) is sent to the DO; if no match is found, then the DO receives a notification that this image is unavailable. Figure 5 explains this step.

$$1- DO \xrightarrow{K} CSP$$

$$2- CSP \xrightarrow{M'_K} DO$$

3.2.2. second step

The DO then computes new metadata (MD'_i) for the selected image (M'_K) based on the following equation:

$$MD'i = h\left(\sum_{i=1}^n \sum_{j=1}^m M'_K(i, j) * (i + j)\right) \quad (3)$$

Thereafter, metadata (MDi) is retrieved from IF saved in the DO. Then, the DO compares two values (MDi,MD'i).If these values are equal, then the integrity of the image indexed K is secure in the CSP. Otherwise, the DO concludes that the images are compromised in the CSP.

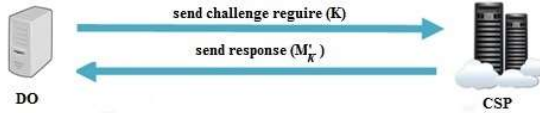


Figure 5. Explains exchange information between DO and CSP

3. EXPERIMENTAL RESULTS AND ANALYSIS

Our proposed scheme was evaluated by using a standard medical dataset from Cyprus University, Cyprus. The dataset contains 1,000 images, which were used to perform our experiments and represent data on the common carotid artery of 80 patients [16-18]. Figure 6 shows some sample training images. The experiments are implemented by using MATLAB R2013a (8.1.0.604) running on Windows 8.1 with an Intel Core i7 processor, 8 GB RAM and 2.4 GHz CPU. The proposed scheme was divided into three phases: setup, verification, and insertion and deletion phases.

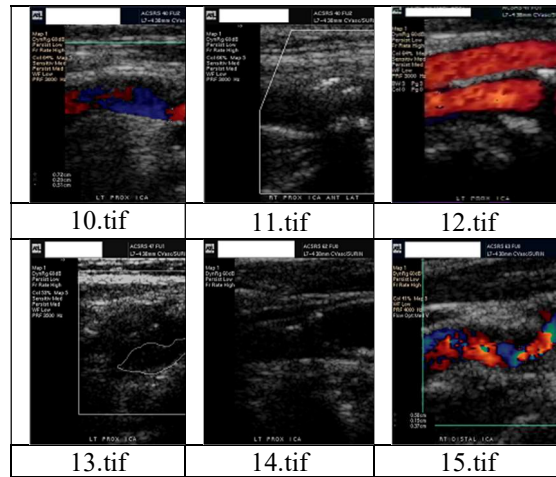
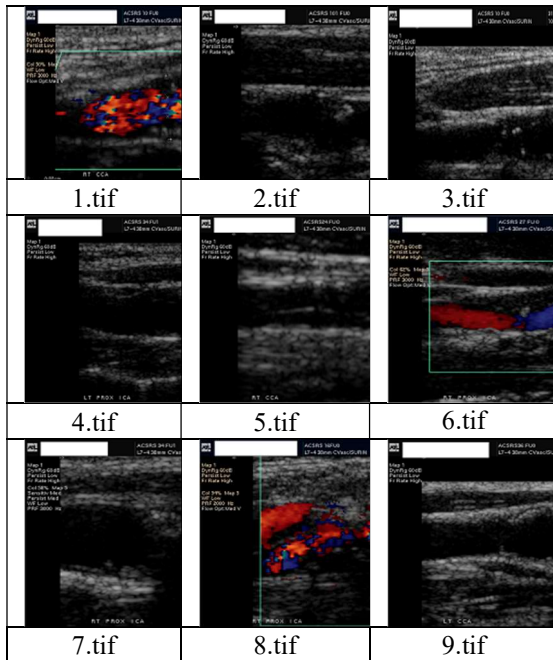


Figure 6: Training image samples

4.1. Setup Phase

This phase includes the creation of the index file. The index file consists of the index image, secure name, metadata and image path. This phase is evaluated in terms of processing times of index file creation and hash function for metadata.

4.1.1. Metadata calculation

We utilised hash function SHA-256 from the processing time to calculate the metadata of the images. Four images were randomly chosen in our experiments to apply the method of hash function, as shown in Table 2, Figure 7 indicates the processing time of the applied hash function

Table 2: Processing time of applied hash function method

Images	MD5	SHA-1	SHA-256	SHA-384	SHA-512
1	0.00718	0.008149	0.000848	0.013563	0.007258
2	0.00054	0.000419	0.018035	0.054113	0.043523
3	0.036017	0.036017	0.00968	0.011425	0.009776
4	0.00044	0.00835	0.007292	0.021496	0.040453

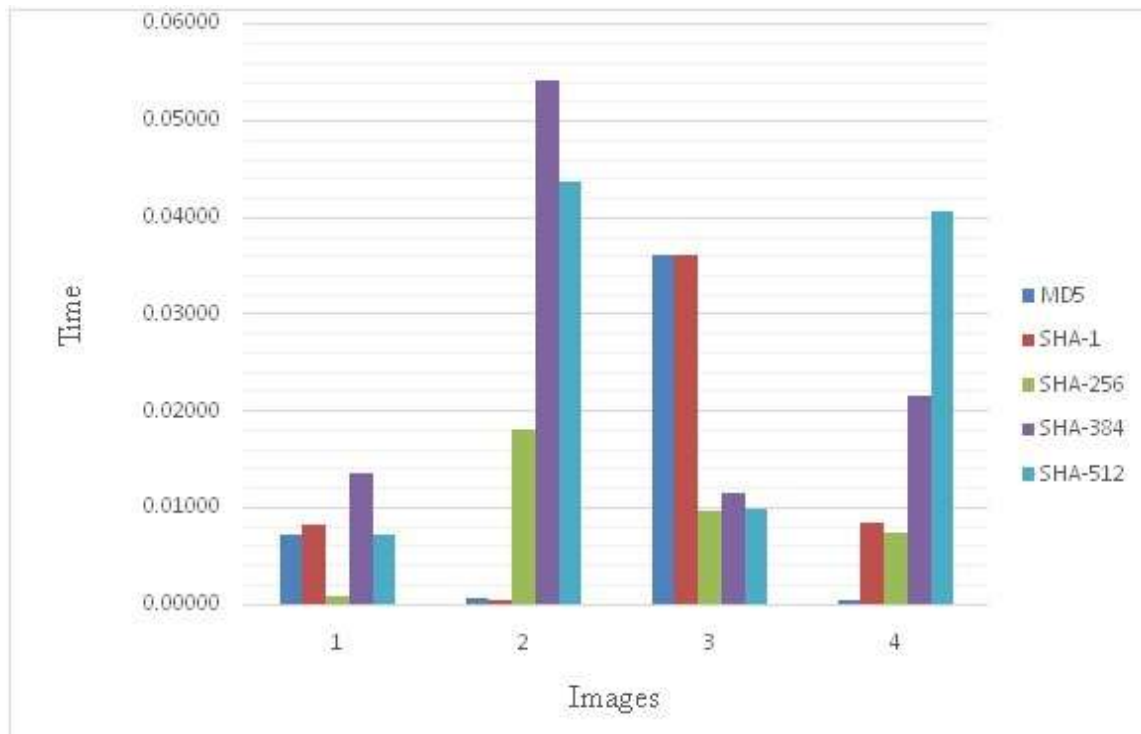
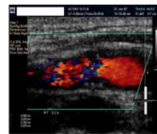
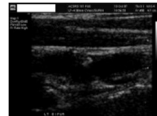
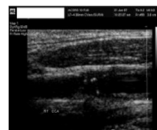
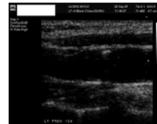
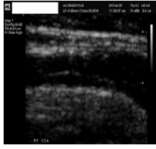
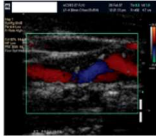
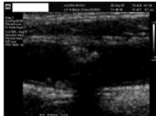
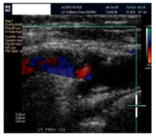


Figure 7: Processing time of applied hash function methods

Table 3 shows a metadata sample for 10 images. This table contains the metadata for the first 10 images created and stored in the index file.

Table 3: Sample of metadata for 10 images

Image name	Size (kb)	Meta Data	Images
1	4	085abe282681c3573e3263d2ff8f059efcb0636c857448f16d701bfabbc57d40	
2	7.27	56e367066bc3458b6258434f27c1ef7979984edf3f359fdde5c88c774ed07446	
3	11.59	573c7b6c28c41dde6bab6934e09a25675d2e94d3c691f39615aff465e33bf b6e	
4	14.27	968c80fa37b15751d6388cb933eba2aa720b14b9c89a6e8dd7ef5aa82c1d6a08	

5	17.34	e0f0ccee0456ecc0fdf7cec65e5edbaa10a6a72c60c3a145dcf9f45df6d21d6	
6	20.37	8ff62bdd27bdf1c8eb8231a3b6320ac95505ae815c6ac131dd7be83bff4ba0ac	
7	23.53	7054b2c054362eae49bcf1c8c425a1685891303a8d50db8c54debe3c1f805b95	
8	28.47	37adefcd8503a0d331752b8adecc9bba78c06cc3a61ac6c2eb4f702053d9f954	
9	32.31	8f2e4c988cfef72450925aa93d7ccc6ce845c40ba261ce2bf2be06fddea46a76	
10	36.19	2af96d48131302084cccbb8bf4a697d6f0a2b7bb1d8138304f0ce169125f1a09	

4.1.2. Processing time calculation of index file creation

We utilised 1,000 images to create the preceding index file. The images were grouped into 100 images to calculate the processing time, and the image size is in kilobytes. Table 4 and Figure 8 show the result of processing time calculation of index file creation.

Table 4: Processing time of index file creation

Image No.	Size (Kb)	Time
1-100	375.2910	9.377999
101-200	375.5439	7.832183
201-300	382.9033	6.5083
301-400	409.6230	6.491217
401-500	379.0684	6.33933
501-600	336.9824	7.569229
601-700	356.1689	7.304986
701-800	275.7891	6.404858
801-900	331.1777	6.480049
901-1000	413.4648	6.623239

4.2. Verification phase

This phase is one of the most important phases, in which the integrity of stored images in the cloud server is checked. It is considered the basic idea of our proposed scheme. This phase is divided into three steps.

4.2.1. Processing time calculation of (sending/receiving) image

The DO randomly selects an image from MIF, sends a challenge to the CS (containing the index of the image) and requests a response from the CS. The CS matches the received index with the indices of the images in the index file. It retrieves the required image for the DO for matching.

If no match is found, then the CS responds that the required images are not available. Table 5 and Figure 9 show the result of processing time calculation of (sending/receiving) images, where:

- Time1 (sdo): processing time of sending from DO to CSP;
- Time2 (csp): processing time inside CSP;
- Time3 (rdo): processing time from CSP to DO;
- V.T.: total verification of processing time (sending, matching and receiving).

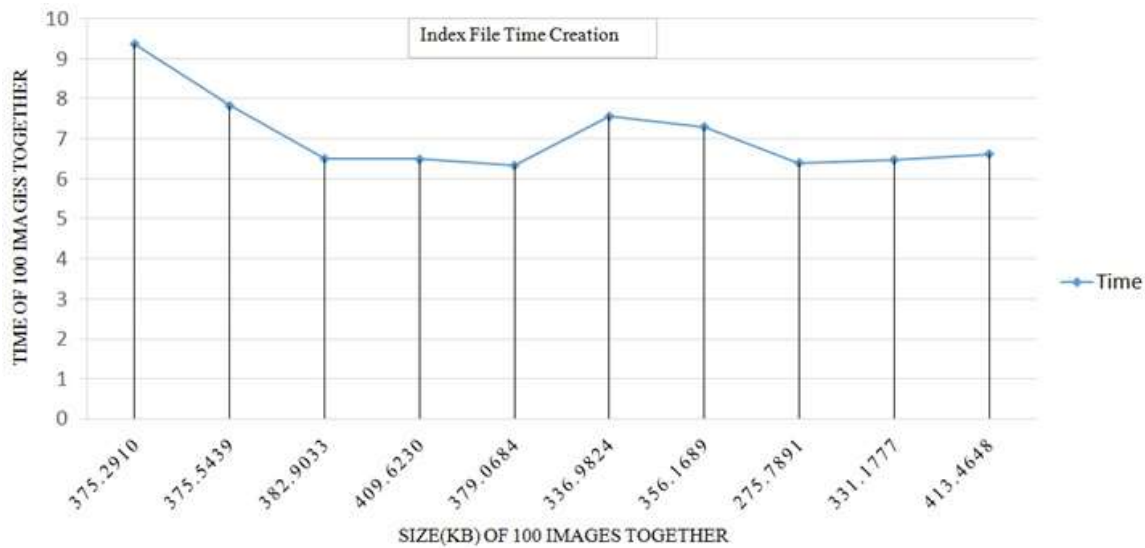


Figure 8: Processing time of index file creation

Table 5: Processing time of (sending/receiving)

Random	Time1 (sdo)	Time2 (csp)	Time3 (rdo)	V.T.
913	8.37E-04	4.32E-04	1.83E-04	0.001452
128	4.48E-05	3.07E-05	7.38E-05	0.000149
920	4.22E-05	2.99E-05	1.15E-05	0.000084
637	4.18E-05	3.03E-05	1.15E-05	0.000084
99	4.18E-05	2.94E-05	1.15E-05	0.000083
281	4.22E-05	3.03E-05	1.15E-05	0.000084
551	4.14E-05	2.90E-05	1.11E-05	0.000081
965	3.88E-05	4.05E-05	1.15E-05	0.000091
972	5.16E-05	3.11E-05	1.15E-05	0.000094
159	4.18E-05	3.16E-05	1.24E-05	0.000086

4.2.2. Calculation of new metadata

Once the DO receives the required image from the CS, it calculates new metadata (MD') for the received image, as mentioned in our proposed scheme (verification phase).

4.2.3. Comparison of MD and MD'

The DO compares MD and MD'. If a match is found between them, then integrity has been

achieved. If no match is found, then the requested image has no integrity.

4.3. Insertion/Deletion Phase

4.3.1. In the insertion process, the DO inserts an image to the CS. The insertion of an image includes the calculation of the index, secure name and metadata. The index file size is updated after the insertion process.

4.3.2. In the deletion process, the DO deletes an image from the CS. The DO selects the secure name of the image to be deleted from IF and then sends the selected secure name to the CS. The CS searches for the received secure name inside IF and then deletes the image according to the search result. If the secure name exists in IF, then the image will be deleted; otherwise, CS notifies the DO regarding the occurrence of an error in the secure name. The size of IF in CS is automatically updated. The processing time of insertion and deletion phases is calculated as shown in Table 6. Figure 10 shows the processing time of the insertion/deletion phase.

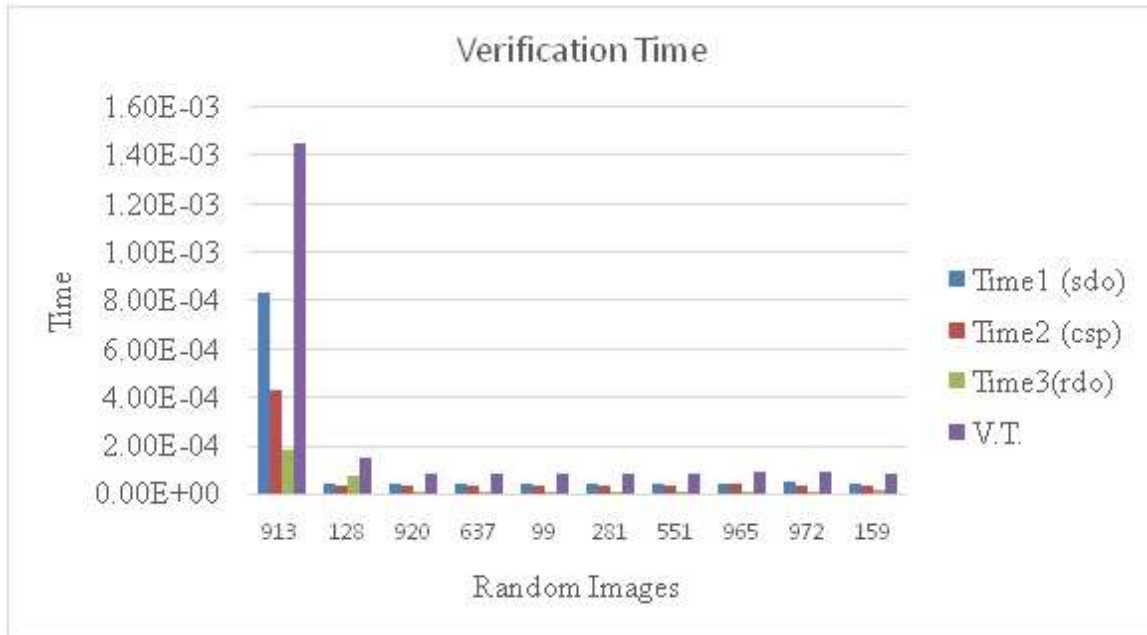


Figure 9: Processing time of (sending/receiving)

Table 6: Processing time of insertion/deletion phase

Images	Time(insert)	Time(delete)
1	0.021137	0.002803
2	0.042053	0.000592
3	0.028812	0.00019
4	0.038208	0.000189
5	0.01971	0.000168
6	0.033564	0.000173
7	0.023247	0.000169
8	0.030421	0.000174
9	0.032443	0.000164
10	0.03189	0.000181

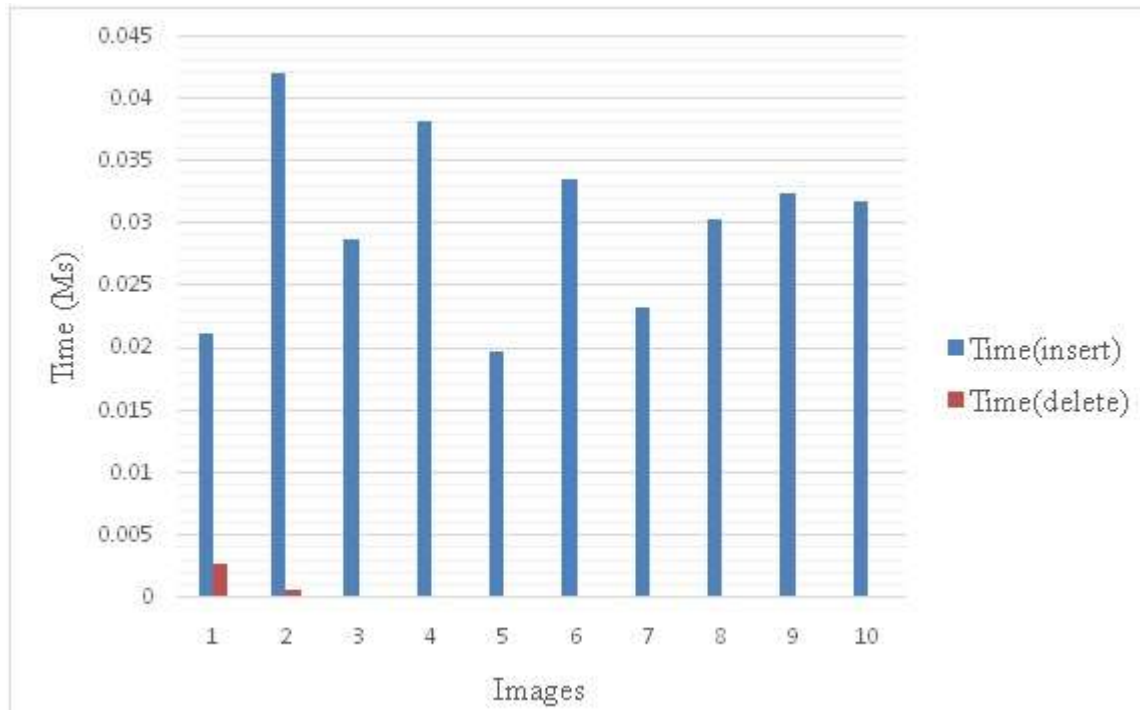


Figure 10: Processingtime of insertion/deletion phase

5. CONCLUSION

The considerable development in information technology and cloud computing nowadays enables users to access data anytime and anywhere. Many applications, images and data in the medical field are outsourced to cloud computing. The transfer of information between cloud entities must be secured. However, security remains a major challenge faced by cloud users. In this paper, we proposed a secure scheme to protect sensitive medical images stored on a CSP. Our proposed scheme comprises two phases: setup and verification. In the setup phase, the DO can store images in the CSP based on IF containing sensitive information used in the verification phase. In the second phase, the DO checks the integrity of images by comparing the metadata of the query image and that of the original one. The proposed scheme can examine the integrity of medical images in CSP and then detect the authority of the CSP. Experimental results indicate that our proposed scheme is flexible, efficient, secure and accurate

REFERENCES:

- [1] A. Singh ; K. Chatterjee. Cloud security issues and challenges: A survey. Journal of Network and Computer Applications, 2017, Volume 79, pp. 88-115.
- [2] P.Ravi Kumarm ;P. Herbert Raj ;P. Jelciana . Exploring Data Security Issues and Solutions in Cloud Computing, The 6th International Conference on Smart Computing and Communications, Procedia Computer Science 125 (2018),Kurukshehra, India, 7-8 December 2017 ,IEEE,2018,pp. 691–697.
- [3] Ali A. Yassin ; Hikmat Z. Neima ; Haider Sh. Hashim. Security and integrity of data in cloud computing based on feature extraction of handwriting signature. International Journal of Cyber-Security and Digital Forensics2014,Volume 3, No. 2, pp. 93-105.
- [4] .N. Garg ;S. Bawa. Comparative analysis of cloud data integrity auditing protocols.Journal of Network and Computer Applications2016, Vol.ume 66, pp. 17-32.
- [5] Chen Lin;Zhidong Shen; Qian Chen;Frederick T.Sheldon. A data integrity verification scheme in mobile cloud computing . Journal

- of Network and Computer Applications2017, Volume 77, pp. 146-151.
- [6] Rajat Saxena ; Somnath Dey.Cloud Audit: A Data Integrity Verification Approach for Cloud Computing. Procedia Computer Science 2016, Volume 89, pp. 142-151.
- [7] J. Cox, J. Kilian ; F. T. Leighton ; T. Shamoon . Secure Spread Spectrum Watermarking for Multimedia . IEEE Transactions on Image Processing1997, Volume 6, no. 12, pp.1673-1687.
- [8] Xu Li; Xingming Sun; Quansheng Liu.Image Integrity Authentication Scheme Based on Fixed Point Theory. IEEE Transactions on Image Processing2014, Volume 24, no. 2, pp. 632 – 645.
- [9] J. C. Dagadu ; L. Jian-Ping ; F. Shah ; N. Mustafa; ; K. Kumar. DWT based encryption technique for medical images, The 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP),Chengdu, China, 16-18 Dec. 2016,IEEE, 23 October 2017.
- [10] R. Sreejith; S. Senthil.A novel tree based method for data hiding and integrity in medical images, IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE), Karur, India, 27-28 April 2017,IEEE,2017, pp.1-4.
- [11] W. Cao ; Y. Zhou ; C. L. Philip Chen ; L. Xia. Medical image encryption using edge maps. Signal Processing2017, Volume 132, pp. 96-109,
- [12] M. Ghadi ; L. Laouamer ; T. Moulahi. Enhancing digital image integrity by exploiting JPEG bitstream attributes.Journal of Innovation in Digital Ecosystems2015, Volume 2, Issues 1–2,, pp. 20-31.
- [13] J. B. Lima ; F. Madeiro ; F. J. R. Sales. Encryption of medical images based on the cosine number transform.Signal Processing: Image Communication2015, Volume 35, pp. 1-8.
- [14] Ali A. Yassin ; Abdullah Mohammed Rashid ; Zaid Ameen Abduljabbar ; Hamid Ali Abed Alasadi ; Abdulla JyAldarwish. Toward For Strong Authentication Code In Cloud Of Internet Of Things Based On Dwt And Steganography.Journal Of Theoretical & Applied Information Technology2018, Volume 96,pp.2922-2935.
- [15] Si-BaoChena ;Yu-LanXua;Chris H.Q.Dingb; BinLuoa.A Nonnegative Locally Linear KNN model for image recognition .Pattern Recognition2018,Volume 83, pp. 78-90.
- [16] Christos P. Loizou, Constantinos S. Pattichis, Senior Member, IEEE, Marios Pantziaris, and Andrew Nicolaides . An integrated system for the segmentation of atherosclerotic carotid plaque.IEEE transactions on information technology in biomedicine2007,Volum 11, no. 6, pp. 661-667.
- [17] C.P. Loizou, C.S. Pattichis, “Despeckle filtering algorithms and Software for Ultrasound Imaging,” Synthesis Lectures on Algorithms and Software for Engineering, Ed. Morgan & Claypool Publishers, USA, 2008.
- [18] Laboratory of eHealth of the University of Cyprus : www.medinfo.cs.ucy.ac.cy.