International Journal of Computer Trends and Technology (IJCTT) – volume 11 number 4 – May 2014

Achieving Privacy Assured Outsourcing of Data in Cloud Using Optimalvisual Cryptography

T.Nataraj M.E[1], S.K. Mahalingam, M.E, Ph.D, [2]

[1] PG Scholar,Department of Computer Science and Engineering, JKK Nattraja College of Engg and Technology, Namakkal, Tamilnadu

[2] Research Scholar, Department of Computer Science and Engineering, Anna University, Coimbatore.

Abstract— **Security has emerged as the most feared aspect of cloud computing and a major hindrance for the customers. In existing system for establishing secure and privacy-assured service outsourcing in cloud computing which uses Linear programming and compressed sensing techniques to transform images, which aims to take security, complexity, and efficiency into consideration from the very beginning of the service flow. But it makes more complexity because the data is sent in its raw form to one cloud. The cryptography schemes are computationally more complex. In order to enhance the security and reduce the complexity, to use data obfuscation through a novel visual cryptography. A conventional threshold (k out of n) visual secret sharing scheme encodes one secret image into transparencies (called shares) such that any group of transparencies reveals when they are superimposed, while that of less than ones cannot. In the proposed work, novel multiple secret visual cryptographic schemes are used to encode the secret s images into n shares. Convert the data into basic images and send the encrypted form of image by using multiple visual cryptographic schemes. (k, n, s) -MVCS, in which the superimposition of each group of shares reveals the first, second, s^{th} secret, respectively where s=n-k+1. The proposed system also considers visual cryptography without pixel expansion. A new scheme for visual cryptography is developed and configured for the cloud for storing and retrieving textual data. Testing the system with query execution on a cloud database indicates full accuracy in record retrievals with negligible false positives. In addition, the system is resilient to attacks from within and outside the cloud. An experimental result shows that the Complexity analysis, Security analysis, the system is tested against artificial intelligence/machine learning based attacks.**

Keywords— **Security and privacy, Cloud computing, visual secret sharing, multiple secrets**

I. INTRODUCTION

Cloud computing is a phrase used to describe a variety of computing concepts that involve a large number of computers connected through a real-time communication network such as the Internet. In science, cloud computing is a synonym for distributed computing over a network, and means the ability to run a program or application on many connected computers at the same time. The ever increasing demand for large scale computing, combined with advances in low cost yet fast networking technologies, has helped cloud computing to emerge as a promising computing model.

In the cloud computing Security is an important concern. A novel procedure to send and retrieve data to and from a cloud using database style query without using standard cryptography schemes, and thus offers efficient retrievals while maintaining data confidentiality. Propose to use data obfuscation instead of an encryption/decryption scheme to achieve data confidentiality. In this work, a novel procedure for visual cryptography, which will use to conduct obfuscation. To show that using this procedure, information cannot be understood by the cloud and is only decipherable by the user.

Visual cryptography relies on decryption using only the human visual system where the data is in a visual form, such as, printed text or pictures. Thus, it avoids the huge computational complexity associated with standard encryption schemes discussed in preceding sections. Visual cryptography relies on breaking up an image into multiple shares such that the image can be reconstructed only when all the shares are available. A share is printed separately and when all the shares are superimposed, the original image can be revealed. The main idea of the Extended Visual Cryptography (EVC) proposed is the generation of innocent looking shares, instead of noise-like shares generated above. The innocent looking shares have two advantages over the noise-like ones: the first one is the facility to distinguish the share of one participant from the shares of others and the second advantage is the cheating prevention.

II. LITERATURE REVIEW

Abdalla. M et.al proposed cryptographic primitive that must meet two conditions. One is of course a security condition [1]. The other, which we will here call a consistency condition, ensures that the primitive fulfills its function. The consistency condition is that decryption reverses encryption, meaning that if M is encrypted under public key pk to result in cipher text C, and then decrypting C under the secret key corresponding to pk results in M being returned.

Carswell et.al suggested content-based image retrieval using queries on shape and topology [2]. To focus on the particularities of image databases encountered in typical topographic applications and present the development of a spatial data management system that enables such queries. The query requires user-provided sketches of the shape and spatial configuration of the object or objects. which should appear in the images to be retrieved. This approach combines the design of an integrated database with the development of a feature library and the necessary matching tools.

Juels et.al suggested to explore privacy protection in cloud architectures [3]. In particular, we consider the challenge of having a cloud service run applications over client data while:

(1) Not being able to learn any information itself and (2) Releasing output values to clients in accordance with an access-control policy. To argue that by itself, cryptography—and by implication, any logical layer information security tool can't solve this problem in its full generality.

Cong Wang et.al proposed privacy preserving multi-keyword ranked search over encrypted cloud data (MRSE) [4].To establish a set of strict privacy requirements for such a secure cloud data utilization system. Among various multi keyword semantics, we choose the efficient similarity measure of "coordinate matching", i.e., as many matches as possible, to capture the relevance of data documents to the search query. Further use "inner product similarity" to quantitatively evaluate such similarity measure. First propose a basic idea for the MRSE based on secure inner product computation, and then give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models.

De Santis et.al suggested a secret image sharing scheme and have modified the existing schemes to provide a better and efficient technique [5]. The previous scheme proposed by Dong and Ku makes the use of matrix multiplication property for construction of shares and addition of shares to reconstruct the secret image. Improved the share construction technique by reducing the computational complexity by applying matrix addition instead of matrix multiplication.

M. Bellare et.al Propose a novel outsourced image recovery service (OIRS) architecture with privacy assurance [6]. For the simplicity of data acquisition at data owner side, OIRS is specifically designed under the compressed sensing framework. The acquired image samples from data owners are later sent to cloud, which can be considered as a central data hub and is responsible for image sample storage and provides on-demand image reconstruction service for data users. OIRS aims to shift such expensive computing workloads from data users to cloud for faster image reconstruction and less local resource consumption, yet without introducing undesired privacy leakages on the possibly sensitive image samples or the recovered image content.

Nathan Chenette et.al proposed Order-preserving symmetric encryption (OPE) is a deterministic encryption scheme whose encryption function preserves numerical ordering of the plaintexts [7]. OPE has a long history in the form of one-part codes, which are lists of plaintexts and the corresponding cipher texts, both arranged in alphabetical or numerical order so only a single copy is required for efficient encryption and decryption. A more formal treatment of the concept of order-preserving symmetric encryption (OPE) was proposed in the database community.

Atallah et.al investigates such secure outsourcing for widely applicable sequence comparison problems and gives an efficient protocol for a customer to securely outsource sequence comparisons to two remote agents [8]. The " string editing" problem, i.e., computing the edit distance between two strings. The edit distance is one of the most widely used notions of similarity: it is the least-cost set of insertions, deletions, and substitutions required to transform one string into another.

Bajaj.S et.al Suggested Trusted DB achieves this by utilizing common unsecured server resources to the maximum extent possible [9]. The contributions of this work are two-fold: (i) the introduction of new cost models and insights that explain and quantify the advantages of deploying trusted hardware for data processing, and (ii) the design, development, and evaluation of TrustedDB, a trusted hardware based relational database with full data confidentiality.

C.Gentre et.al Proposed a Verifiable Computation Scheme: this is a protocol between two polynomial-time parties, a *client* and a *worker*, to collaborate on the computation of a function F: $\{0, 1\}^n \rightarrow \{0, 1\}^m$. To introduce and formalize the notion of Verifiable Computation, which enables a computationally weak client to "outsource" the computation of a function on various dynamically-chosen inputs to one or more workers [10]. The workers return the result of the function evaluation, well as a proof that the computation of F was carried out correctly on the given value x_i. The primary constraint is that the verification of the proof should require substantially less computational effort than computing $F(x_i)$ from scratch.

III. CRYPTOGRAPHIC SCHEMES FOR DATA CONFIDENTIALITY

A. Secure and privacy-assured service outsourcing

A secure and privacy-assured service outsourcing is used in cloud computing which uses Linear programming and compressed sensing techniques to transform images, which aims to take security, complexity, and efficiency into consideration from the very beginning of the service flow. Because data explosion is the fast-growing trend to outsource the image management systems to cloud and leverage its economic yet abundant computing resources to efficiently and effectively acquire, store, and share images from data owners to a large number of data users. Although outsourcing the image services is quite promising, in order to become truly successful, it still faces a number of fundamental and critical challenges, among which security is the top concern. To initiate the investigation for these challenges and propose a novel outsourced image recovery service (OIRS) architecture with privacy assurance. For the simplicity of data acquisition at data owner side, OIRS is specifically designed under the compressed sensing framework. The acquired image samples from data owners are later sent to cloud, which can be considered as a central data hub and is responsible for image sample storage and provides on-demand image reconstruction service for data users. But it makes more complexity because the data is sent in its raw form to one cloud. The cryptography schemes are computationally more complex. In order to enhance the security and reduce the complexity, to use data obfuscation through a novel visual cryptography.

B. Multiple Visual Cryptographic Scheme

A visual cryptography to send and retrieve data to and from a cloud using database style query without using standard

cryptography schemes, and thus offers efficient retrievals while maintaining data confidentiality. They propose to use data obfuscation instead of an encryption/decryption scheme to achieve data confidentiality. In this work, we have come up with a novel procedure for visual cryptography, which we will use to conduct obfuscation. Proposed system is an idea of sending all datas in an image forma. This idea is the extensions of Privacy-Assured Outsourcing of Image Reconstruction Service in Cloud. Here we use visual cryptography instead of compressed sensing and Linear programming. On the other hand, Mmultiple visual cryptographic scheme is proposed convert the data into basic images and send the encrypted form of image. The visual secret sharing (VSS) scheme is to encrypt a secret image into n meaningless share images. It cannot leak any information of the shared secret by any combination of the n share images except for all of images. The visual secrets sharing scheme for multiple secrets is called multiple-secret visual cryptographic schemes (MVCSs). Multiple-secret visual cryptographic schemes (MVCSs, for short) using direct superimposition only for decoding.

(1)(k,n,s) -MVCS, in which the superimposition of each group of shares reveals the first, second, s^{th} secret, respectively where s=n-k+1.

(2)(k,n,s,R) -MVCS, where R is a *revealing list* specifying that secret would be revealed by each group of shares.

IV. Methodology and System Development

A. Sending data

The user has some data in a simple text file that is to be uploaded to the clouds. If the data is in another format, it must be converted to text.

Converting text to image: By using the BMP file format for converting text to images. The bits representing the pixels are packed in rows, which allow easy image manipulation for the system. Each image is constructed in Grayscale, 1- bit depth and two colors, thus is in black and white. We fix the resolution as 40×40, with which each image is 382 bytes in size. The image resolution can be increased if number of cloud providers to which image is to be cropped and sent, is high. To allow scalability on the cloud, we try to keep the image file size as low as possible.

Image obfuscation: As data is read, its equivalent images are obfuscated by adding noise to each image. Adding noise to the data severely decreases the possibility of discovering the actual text. We propose to focus on Gaussian and speckle noise for our system since they are two most important and common categories of noises for images. We will use these noises for our system and compare their performance. For Gaussian noise, the noise density follows a normal distribution, also known as the Gaussian distribution.

$$f(x) = \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

variance σ^2 identify the normal distribution. The noise if often represented as N (μ, σ^2). The standard normal distribution has zero mean. Varying the mean introduces skewness in the data, thus makes the distribution asymmetric. When the mean is positive, distribution becomes right skewed and when the mean is negative we get a left skewed distribution. While adding noise to the images, the original pixel is replaced by a black or white pixel. So image size and other attributes remain the same after noise addition. The original image, the text is black on a white background.

Data division across the cloud: After adding noise to the bitmap images, each image is split into as many equal parts as the number of cloud providers. The cloud providers are independent and are not aware of each other's presence or data held by other providers. The data held by them is disjoint, that is, no two providers have any part of the data common between them. Segregating the data among different clouds, maintains the disjoint property and better preserves data confidentiality.

B. Receiving Data

To retrieve records which begin with a certain query, for a query of length *l*, need to evaluate the first *l* locations in a record. Then, we need to check if each of the first *l* images in a record match with the respective equivalent images of the search query. If the query string matches with the images in the first *l* locations of the record, then we retrieve the record and send it to the user. To determine a suitable method to match the noisy data in the server with the unnoisy images of the search query. Then a suitable metric needs to be assigned to determine whether a match is found or not.

Image retrieval from the records: Once the query is received, the system retrieves the images stored at selected locations in the beginning of a record from each provider. Recall that the number of records at each cloud is the same and contain only a part of the original image. The first image from the first record in each server is retrieved. Aligning these images together, we will get the original noisy image of a character which was cropped and sent to individual clouds. Thus, for instance, an original 40×40 noisy image of a character was split into four 10×40 images and sent to four cloud providers. The same is repeated for all characters and they are stored in records of size eigth at each cloud.

Matching the images: The images retrieved from the clouds are noisy. Pattern matching on these images generates extremely chaotic results. Images must be denoised first. We are working on bitmap images and each pixel is simply black or white. The text is black and background is white. Thus, performing a bitwise AND operation between noisy and unnoisy images of the same character will produce an unnoisy image, which we call the *mask*. Then, we can compare the *mask* with the actual image of the character in *lib*, which will result in an almost perfect match. Clearly, the mask is noisy whereas we expected an unnoisy image. This is because bitwise AND produces a 1 (=white) only when both inputs are 1 and in all other cases we get output as 0 (=black). In the background, unnoisy image is all 1, while noisy has both 0 and 1. Thus, the output remains noisy in the background. For the text, original image is all 0 in that part, and in noisy image is 1 and 0, thus we get 0 that is black, as the text color in the mask. We cannot perform a pattern matching with so much

noise in the mask. Instead, if we perform a bitwise NOT on the original and noisy images and then perform AND between them to, and then again NOT the output of the last step we will get a perfect mask. Note that in this mask the background is all white, however, the text will remains noisy for Gaussian noise, while for speckle noise, even the text will be perfectly unnoisy. To perform the matching between two images, we employ the normalized cross correlation (*NCC*) metric, $\frac{1}{n-1}\sum_{x,y}\frac{9f(x,y)-f)(t(\widehat{x,y})-\hat{t}}{\sigma_f\sigma_t}$

Where, the similarity between images f(x,y) and t(x,y) each having n pixels is calculated. We use *NCC* to determine if a mask and the search query image match. As noted before, a mask is created every time an image in the record is accessed during data retrieval. The *NCC* value, labelled as *ncc* henceforth, between the mask and the character which is to be searched, is then calculated.

C. Sending and Retrieving data with multiple visual secrete sharing scheme

By using the the BMP file format for converting text to images. The bits representing the pixels are packed in rows, which allow easy image manipulation for the system. In that each image is constructed in Grayscale, 1-bit depth and two colors, thus is in black and white. After that using the multiple secret visual cryptography technique the data obfuscation is accomplished. The visual secrets sharing scheme for multiple secrets is called multiple-secret visual cryptographic schemes (MVCSs). Consider s secret images shared by n participants. The two threshold multiple-secret visual cryptographic schemes (MVCSs, for short) using direct superimposition only for decoding: (1) (k,n,s) –MVCS in which the superimposition of each group of k,k+1,..n shares reveals the first, second, , th secret, respectively, where s=n-k+1; and (2) (K,n,s,R)-MVCS where R is a revealing list specifying that secret would be revealed by each group of shares. A new pseudo random approach is used to add security to the system. However, the overhead involved in such a scheme must be balanced such that performance does not suffer. After that the image is retrieved from the records. The first image from the first record in each server is retrieved. Aligning these images together, we will get the original noisy image of a character which was cropped and sent to individual clouds. The images retrieved from the clouds are noisy. Pattern matching on these images generates extremely chaotic results.

V. EXPERIMENTAL RESULTS

In the experimental results the performance of the existing and the proposed system is evaluated. Comparative analysis has been done for multiple visual cryptographic schemes (MCVS) with Pixel expansion and multiple visual cryptographic schemes without pixel expansion. These analyses can be done by considering two evaluation parameter which is CPU computation Time and Pixel Density calculation. High usage of CPU computation time results slows down the processing time. High Pixel density results less efficient in processing. The result obtained for the existing and proposed research has been discussed and respective result has been shown.

A. CPU computation time

The computation time varies from different system according to the memory usage of resources.

CPU computation time= E_T (a) - E_T (b)

Where E_T (a) after executing Time, E_T (b) before execution time.

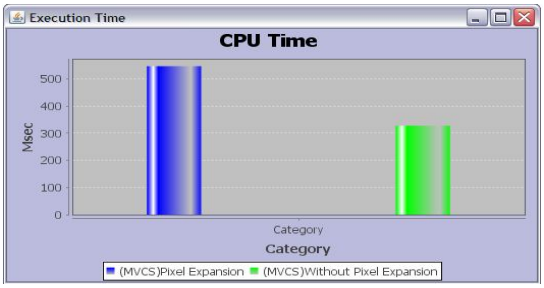


Fig 1 CPU time comparison

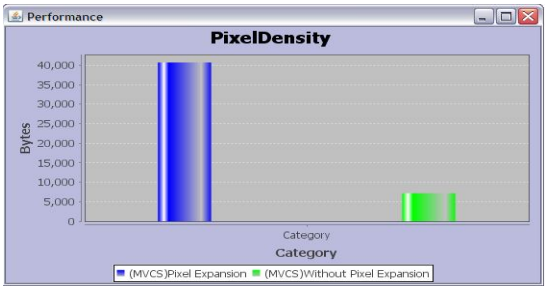
The above graph shows the comparison graph for existing and proposed CPU Execution time .In Figure 1. MVCS with pixel expansion utilizes high computation time of around 500 Msec. The proposed scheme of MVCS without pixel expansion method utilizes around 300 MSec to process .Thus The proposed method achieves less CPU Execution time to run when compare with the Existing research.

B. Pixel Density calculation

Pixel Density can be calculated by below formula.

Pixel Density=Square root ((wp*wp) + (hp*hp))/di

Where wp is width resolution in pixels, hp is height resolution in pixels and di is diagonal size in inche.



.

Fig 2 Pixel size comparison

The above graph shows the pixel size for existing and proposed research. The existing system uses pixel expansion so the result returns largest pixel size. In Figure 2 MVCS with Pixel expansion takes nearly 40,000 bytes for embedding secret image in the shares. The proposed scheme takes around 6000 bytes. Thus the proposed system works on the no pixel expansion and thus returns the smallest pixel size in the result. Less overhead for storage and transmission is achieved to share multiple secrets while using the proposed scheme.

VI. CONCLUSION AND FUTURE WORK

A novel method to achieve data confidentiality in the cloud computing environment. The cloud provider is considered untrustworthy and the data must be concealed not only from an outside attack but the provider itself must not be able to extract meaningful information from the data. Instead of relying on one cloud service provider, we propose to use multiple (untrustworthy) public cloud providers. By using multiple secret visual cryptography to protect the data on the cloud. Standard encryption schemes are avoided, yet we achieve strong privacy of the data. A new visual cryptography scheme for binary images is introduced for our system. The complexity of our approach is shown to be reasonable and much less than standard encryption based schemes. The results are encouraging for our system and indicate that our work has the potential for a large-scale application. To this end, we propose some key improvements and suggestions for future systems.

REFERENCES

[1] Abdalla. M, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange, J. Malone-Lee, G. Neven, P. Paillier, and H. Shi(2005) "Searchable encryption revisited: Consistency properties, relation to anonymous ibe, and extensions,""*CRYPTO 2005. LNCS* vol. 3621, pp. 205–222. Springer, Heidelberg .

[2] Agouris.P, J. Carswell, Stefanidis.A (1999) "An environment for content based image retrieval from large spatial databases," ISPRS J. Photogram. Remote Sens., vol. 54, no. 4, pp. 263_272.

[3] Van Dijk.M, and A. Juels(1010) "On the impossibility of cryptography alone for privacy-preserving cloud computing," in Cryptology ePrint Archive, Report 2010/305, 2010. Available: http://eprint.iacr.org//305.

[4] Ning Cao, Cong Wang, Wenjing Lou(2009) "Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50–55.

[5] De Santis.A (2009) "An Optimal (n,n) Secret Image sharing Scheme," Faculty of Engineering Mechanics and Systems," University of Tsukuba.

[6] Cong Wang, Bingsheng Zhang (2013) "Privacy-Assured Outsourcing of Image Reconstruction Service in Cloud," U.S. National Science Foundation under Grants CNS-1262277.

[7] Alexandra Boldyreva, Nathan Chenette, Younho Lee and Adam O'Neill(2010) "Order-Preserving Symmetric Encryption," Georgia Institute of Technology, Atlanta, GA, USA.

[8] Atallah.M and Ji.L (2005) "Secure outsourcing of sequence comparisons," Int. J. Inf. Security, vol. 4, no. 4, pp. 277_287.

[9] Bajaj.S, and R. Sion(2011) "TrustedDB: a trusted hardware based database with privacy and data confidentiality," in Proceedings of the 2011 international conference on Management of data (SIGMOD '11), ACM, New York, NY, USA, pp. 205-216, [doi=10.1145/1989323.1989346].

[10] Gennaro.R, C.Gentre and B. Parno (2010) "Non-interactive verifiable computing: Outsourcing computation to untrusted workers,"in Proc. CRYPTO, pp. 465_482.