Secure Auditing Scheme with Content-Based Image Retrieval using Human Interaction (RF) in Cloud Computing Environment

Ms. Sonali S. Panchal[1], Ms. Shital Y. Gaikwad[2], Mr. Kamthane A.N[3]

[1]*M.E.(CSE) Final year student, Department of Computer Science and Engineering, Matoshri Prathisthan Group of Institution, Institution of Engineering and Technology, Khupsarwadi, Post Vishnupuri, Nanded-431606, (M.S.)-INDIA*

[2]*Assistant Professor(CSE), Department of Computer Science and Engineering, Matoshri Prathisthan Group of Institution, Institution of Engineering and Technology, Khupsarwadi, Post Vishnupuri, Nanded-431606, (M.S.)-INDIA*

[3]*Head of Department(CSE), Department of Computer Science and Engineering, Matoshri Prathisthan Group of Institution, Institution of Engineering and Technology, Khupsarwadi, Post Vishnupuri, Nanded-431606, (M.S.)-INDIA*

ABSTRACT

"A picture is worth of thousands of words"! Image processing is used for image retrieval system for browsing, searching and retrieving images from a large database of digital images. There are two types of image search, Image Meta Search and Content-Based Image Retrieval (CBIR). CBIR is the mainframe of image retrieval system. To understand query semantics and users expectations so as to communicate faithful results in terms of accuracy, Relevance Feedback is incorporated into CBIR System. This system allows users to progressively refine the search results, to access iteratively the results as "relevant/irrelevant" or even giving him/her the opportunity to specify a degree of relevance, the system creates a new query that better captures the user's needs, hence raising the opportunity to get more relevant image results. In the watermark-based protocol, a unique watermark is directly embedded into the encrypted images by the cloud server before images are sent to the query user. Hence, when an illegal image copy is found, the unlawful query user who distributes the image can be traced by the watermark extraction. Feature vectors get protected by the secure hashing algorithm, auditing and alert generation are used at image users side for verification purpose. TPA (Third Party Auditor) is used to detect fraud or malicious activities performed in a cloud environment.

Keywords *Content-Based Image Retrieval, Relevance Feedback, TPA, encryption, and decryption of images, hashing algorithm, Auditing, cloud computing, watermark-based protocol.*

1. INTRODUCTION

The needs for efficient image storage and retrieval services are reinforced by the increase of large-scale image databases among all kind of areas [1]. For privacy-preserving purposes, sensitive images such as medical and personal images, need to be encrypted before outsourcing, which makes the CBIR technologies in plain text domain to be unusable [13]. Nowadays, it is possible to do retrieval with copy deterrence also so that unauthorized image users will not get access to it. CBIR services typically incur high storage and computation complexities. For increasing the accuracy rate we will be using relevance feedback system. In our proposed system we are including the technique of fraud detection using a hashing algorithm. Considering the case that the authorized query users may illegally copy and distribute the retrieved image to someone unauthorized, a watermark based protocol is used to deter such illegal distributions. In the watermark based protocol, a unique watermark is directly embedded into the encrypted images by cloud server before images are sent to query user. Hence when an illegal image copy is found, the unlawful query user who distributes the image can be traced by the watermark extraction.

The following are the main purpose of this proposed system:

1. To provide high security to multimedia big data

2. To make Content-based image retrieval and retrieve better accuracy over encrypted image retrieval
3. Optimize the resources distribution in multi-streams at the system level
4. To train system by using relevance feedback system for better image retrieval
5. To detect fraud detection using TPA
6. To detect editing of cloud data
7. To compress data to achieve resource constraint
8. To generate watermark-based copy deterrence protocol.

Cloud computing offers a great opportunity for the on-demand access to ample computation and storage resources, which makes it an attractive choice for the image storage and CBIR outsourcing.

2. REVIEW OF LITERATURE

Kui Ren, Zhan Qin, Xinhui wang, Zhihua Xia, and Liangao Zhang all in their paper titled "A Privacy-Preserving and Copy-Deterrence Content-based Image Retrieval Scheme in Cloud Computing" focused on preserving privacy and doing copy deterrence while retrieving content based images in a cloud computing environment. In this system, feature vectors are extracted to represent the corresponding images and to access to unauthorized users get prevented. Watermark certification authority is provided in cloud computing environment for making images more secure [1].

Kui Ren, Zhihua Xia, Zhan Qin, and Yi Zhu all in their paper titled "Towards Privacy-Preserving Content-Based Image Retrieval in Cloud Computing focused on CBIR applications" which are developed very fast along with the improvement in the availability, quantity, and importance of images which are present in daily life. In this system, privacy is preserved of retrieval process to control the access to images by authorized users only. There is data owner who sends the CBIR services and image database to the cloud, without giving any idea about the original contents of the image database to the server [2].

Xinhui Wang, Zhihua Xia, Zhan Qin, Liangao Zhang, Kui Ren all in their paper titled "A Privacy-Preserving and Copy deterrence Content-based Image Retrieval Scheme in Cloud Computing" describes retrieval process of content based images with copy deterrence and preserving privacy in cloud computing. For preserving privacy, sensitive images, like personal and medical images, required to convert in encrypted form before outsourcing because of this CBIR technology present in plain text domain becomes unusable. Moreover, secure kNN algorithm is used to protect the feature vectors, and standard stream cipher encrypts the image pixels [4].

Nasir Memon, K. Gopalakrishnan, Poorvi L. Vora all in their paper titled "Protocols for Watermark Verification" focused on adding a watermark signal into the digital image which is later deleted or extracted for making an assertion about the particular image. There are two categories of watermarks present: invisible and visible. Evidently, company logos or visible messages are presently invisible watermarks which indicate the image ownership. On the other hand, invisible watermarks contain discrete modifications to the image and the invisibly watermarked image which visually appears very similar to the original image [5].

P.T. Boufounos and S. Rane all in their paper titled "Privacy-Preserving nearest neighbor methods" comparing signals without revealing them, focused on the privacy-preserving NN (PPNN) method in which the reader will come to know that it convenient to make partition of this in two different problems: privacy-preserving minimum finding method follow another method that is privacy preserving distance computation. Under certain considerations, privacy model dictates that which mathematical tools should be useful for PPNN. It also defines the complexity and structure of resulting protocols. These models make assumptions upon requirement, behavior, sharing. These assumptions have a main focus on participating entities behavior, the amount of possible information that gets shared among participants and privacy requirements [12].

3. LIMITATIONS OF EXISTING SYSTEM

Despite the tremendous benefits, image privacy becomes the main concern with CBIR outsourcing [1]. For example, patients may not want to disclose their medical images to any others except to a specific doctor in medical CBIR applications. To formulate the problem, this paper considers two types of privacy threats. Firstly, a curious cloud

server may look into the owner's database for additional information. Secondly, after receiving the retrieved images, the query user may illegally distribute these images to someone unauthorized for benefits [1]. Following are some of more limitations:

- The system requires proper image inputs.
- Internet connection is compulsory.
- Less robustness and embedding capacity
- PCBIR scheme which protects the privacy of the query image, but exposes the unencrypted image database to the server directly.
- High computation and storage burden Image retrieval accuracy is low.
- None of the existing schemes consider the dishonest query users who may illegally distribute the retrieved images.

4. PROPOSED SYSTEM ARCHITECTURE

The needs for efficient image storage and retrieval services are reinforced by the increase of large-scale image databases among all kind of areas [1]. For privacy-preserving purposes, sensitive images such as medical and personal images, need to be encrypted before outsourcing, which makes the CBIR technologies in plain text domain to be unusable [13]. Nowadays, it is possible to do retrieval with copy deterrence also so that unauthorized image users will not get access to it. CBIR services typically incur high storage and computation complexities [13]. For increasing the accuracy rate we will be using relevance feedback system.

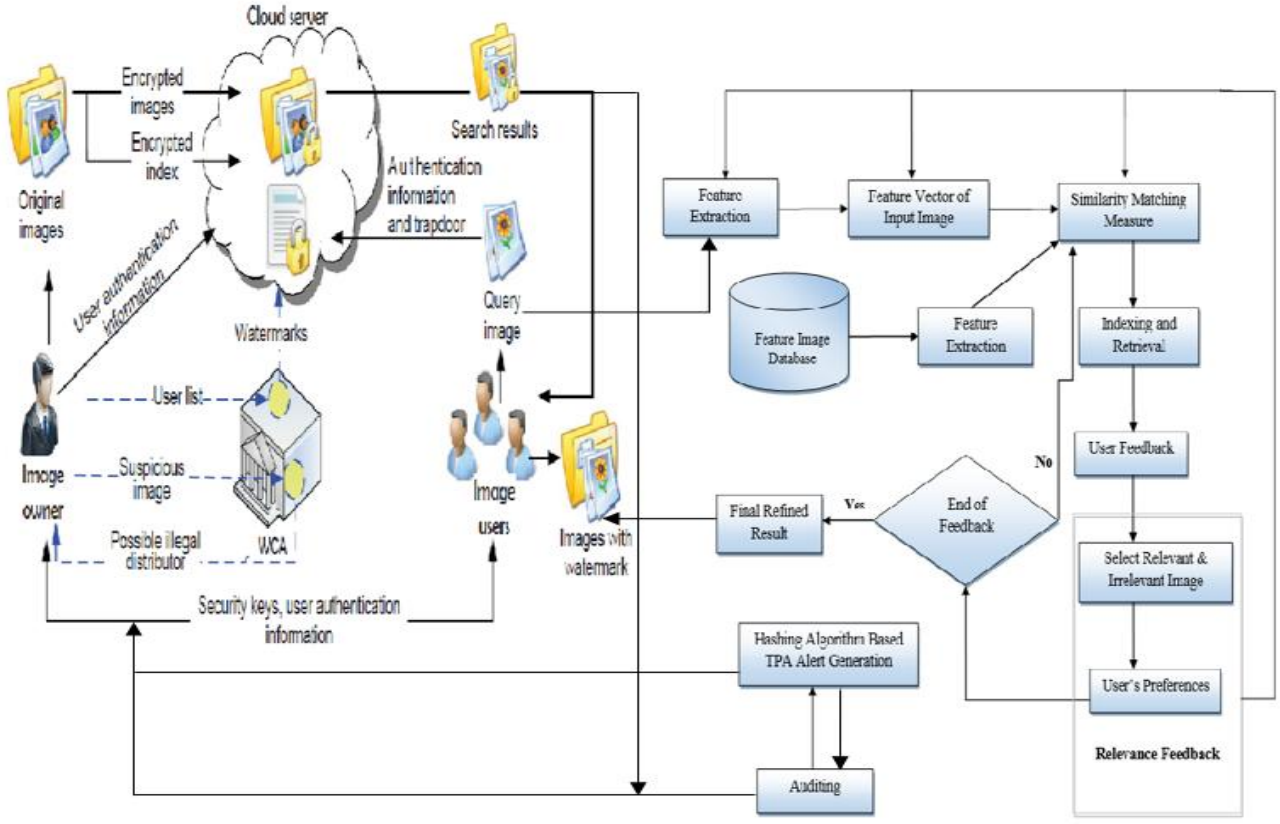

Fig -1: Architecture of Proposed System

The main purpose of this system is to propose a scheme that supports CBIR over the encrypted image without leaking the sensitive information to the cloud server. While uploading images for security purpose so that no intruders can hack images for unauthentic use of images and securing images from hackers. Among users and owner

security keys and user authentication information is shared. Using those keys and information, encryption and decryption are done at image owner and user level [13]. Following are the phases which describe how the proposed system works gradually:

1. Original images are encrypted by image owner on cloud server, meanwhile encrypted index of those images are saved on cloud server along with user authentication information.
2. The watermark-based protocol is applied to those images.
3. Image owner sends the suspicious image to Watermark Certification Authority (WCA) and WCA checks for the possible illegal distributor. User lists are stored with WCA by image owner.
4. WCA generates watermark on those secured images and transfers them towards cloud server.
5. After that encrypted images are embedded with a watermark on a cloud server.
6. Once the watermark is embedded then the search image result is forwarded to image users. Image user now will fire image query and features will be extracted.
7. CBIR will check similarity measures between fired image query and stored database of images.
8. The image is processed by indexing and image retrieval scheme with the help of CBIR.
9. Then human interaction system i.e. Relevance Feedback System will give the result as relevant or irrelevant images iteratively with the help of users preferences.
10. After that user will give proper feedback for the retrieved result, if yes then the final refined result will be displayed and if not, then again the system will check similarities matching measure and will again ask for relevance feedback iteratively.
11. Final refined results will be images with a watermark.
12. Watermark will be extracted and images will is decrypted as per user's requirement.

When image users get search results, hashing and auditing get performed for verifying that search results are satisfactory or not. If it is not then alert message is sent to appropriate image owner. Once, the image user obtains image search outcomes, hashing algorithm, and auditing with the help of TPA alert generation is performed for verification of refines result which is satisfactory or not. If not then TPA will generate an alert message and it will be directed to suitable image owner. While developing this system robustness is considered and by performing strong literature survey we are able to develop robustness and privacy of content-based image retrieval using various algorithms, also at the same time to deal with the problems.

4.1 Models of Proposed System

Mainly, there are 10 types of entities are involved:

1. Image owner
2. Image users
3. Preprocessing
4. User authentication
5. Watermark generation
6. Hashing algorithm
7. Signature generation
8. Image encryption and decryption
9. Auditing
10. Alert generation

4.2 Procedure to perform Hashing Algorithm, auditing and Alert Generation [13]

1. Start
2. Read data owner id (udoid)
3. If (doid ≠ udoid)
4. Stop
5. Read filename from AWS
6. Retrieve the number of blocks from TPA XML
7. Select the blocks number that user wants to verify
8. Get the auxiliary information for block chal from TPA XML
9. Based on auxiliary information generate new root for MHT
10. If (new root ≠ root) file modified
11. Else File not modified
12. Stop

4.3 Framework for Watermark based protocol

For security and efficiency, a CEW algorithm proposed by Zhang [1] is exploited in this scheme. Since the data owner can easily get the original images, we modify the embedding and extraction procedure to achieve a better watermark extraction accuracy. Image owner encrypts the original image with the help of a key. Encrypted images are retrieved to a query request on a cloud server with the help of watermark embedding. Encrypted and watermarked images are sent to query the user for image decryption. Decrypted and watermarked images are obtained. WCA will extract watermark using the key with the help of watermark bits. Watermark bits are associated with users using secret keys in watermark algorithm [1].

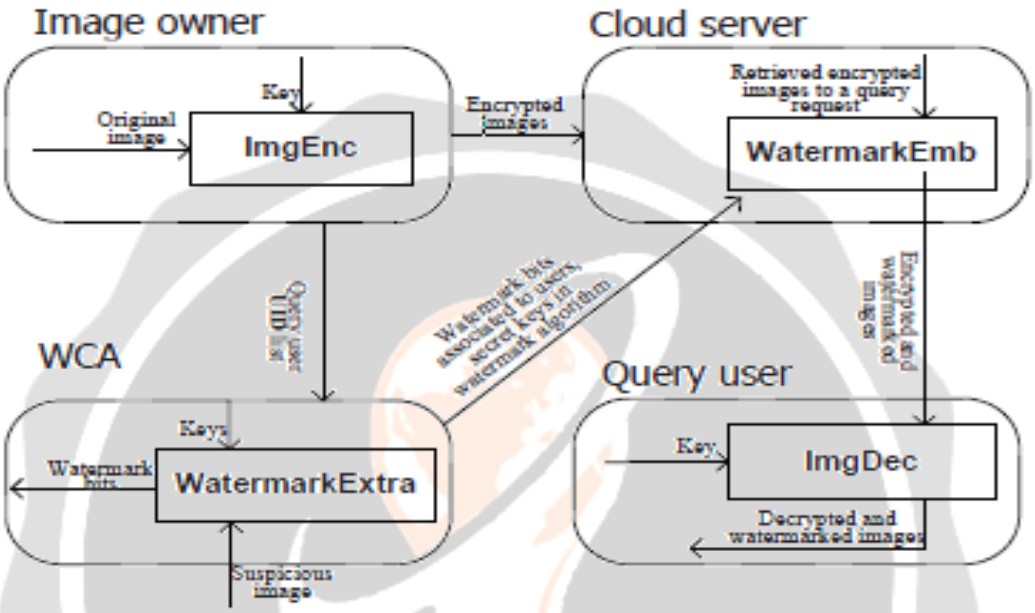

Fig -2: Framework of watermark-based protocol

4.4 Applications of Proposed System

1. Search for a picture to go with a broad story or search to illustrate a document
2. General browsing to make an interactive choice
3. Simple users search for one specific image on the web
4. Education and Training
5. Medical Diagnosis
6. Fashion and Publishing
7. Police Crime Branch for photo recognition in crime prevention
8. Cartography- Mapmaking from photographs, synthesis of weather maps
9. Multimedia like Journalism and Advertising

5. CONCLUSION

The system supports Content-Based Image Retrieval over the encrypted image without leaking the sensitive information to the cloud server. The secure CEW algorithm proposed by Zhang is applied to encrypt the visual features along with indexing and image retrieval using relevance feedback system to increase the accuracy of the image query fired by the user [1]. Hashing algorithm and auditing is used for verification purpose. If the results are satisfactory to image users, the process terminates otherwise feedback module sent back to image owner. We also consider the dishonest users and proposed a watermark-based protocol using WCA to deter the illegal distribution of images and thus provides security to images. In this system, encryption on images is done before outsourcing and with copy deterrence, access from unauthorized users are get prevented. The trusted agency is used to generate a

watermark based protocol for avoiding illegal distribution [13]. The performance of this system based on factors like how efficiently the watermark is get embedded and extracted from the image, how smoothly the encryption and decryption process carried out without losing image pixels, and the rate of verification [13]. Unlike Zhang's algorithm, which has been proved to be effective in encryption and decryption, my method is to retrieve appropriate images that are saved on cloud server by image owners [1] [13]. The main focus of my paper in on hashing algorithm, auditing, and alert generation.

6. Acknowledgment

I would like to express my gratitude to P.G. Dept. of Computer Science and Engineering, MPGI, Nanded. I am also thankful to my guide Assistant Professor Ms. Shital Y. Gaikwad for her guidance and encouragement. Their expert suggestions and scholarly feedback had greatly enhanced the effectiveness of this work. I would like to express the deepest appreciation to authors Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun, and Kui Ren for their beneficial information and knowledge.

7. REFERENCES

[1] Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun, and Kui Ren, "A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing", IEEE TRANSACTIONS ON INFORMATION FORENSIC AND SECURITY, vol.11, 2016, pp. 2594 - 2608.
[2] Z. Xia, Y. Zhu, X. Sun, Z. Qin, and K. Ren, "Towards privacy-preserving content-based image retrieval in cloud computing, IEEE Transactions on Cloud Computing", vol. PP, no. 99, 2015, pp. 1-1.
[3] Z. Qin, J. Yan, K. Ren, C. W. Chen, and C. Wang, "Towards efficient privacy-preserving image feature extraction in cloud computing", in ACM International Conference on Multimedia. ACM, 2014, pp. 497-506.
[4] Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun, and Kui Ren, "A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing", IEEE TRANSACTIONS ON INFORMATION FORENSIC AND SECURITY, VOL. , NO. , SEPTEMBER 2016.
[5] K. Gopalakrishnan, N. Memon, and P. L. Vora, "Protocols for watermark verification", IEEE MultiMedia, no. 4, pp. 66-70, 2001.
[6] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data", in Proc. of 28th International Conference on Data Engineering. IEEE, 2012, pp. 1156-1167.
[7] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, 2001, pp. 703-715.
[8] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search, in Advances in Cryptology-Eurocrypt". Springer, 2004, pp. 506-522.
[9] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data, in Proc. of 28th International Conference on Data Engineering". IEEE, 2012, pp. 1156-1167.
[10] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data", IEEE Transactions on Parallel and Distributed Systems, vol. PP, no. 99, 2015, p. 1.
[11] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data", IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 1, 201, pp. 222-233.
[12] S. Rane and P. T. Boufounos, "Privacy-preserving nearest neighbor methods: comparing signals without revealing them", IEEE Signal Processing Magazine, vol. 30, no. 2, 2013, pp. 18-28.
[13] S. R. Lahane, Sonal H. Kunte, "Using Watermark-Based Protocol, Increase robustness and Privacy of Content-Based Image Retrieval in Cloud Computing Environment", IJARIIE-ISSN (0)-2395-4396, Vol-3 Issue-4 2017.