To Apply Watermarking Technique in Cloud Computing To Enhance Cloud Data Security

[1]Neha Khajanchi, [2]Prof. Vishakha Nagrale

Department of CSE, RCERT
Chandrapur (M.S.)

Abstract: The Cloud Computing is a dynamic term which provides dispute free data outsourcing facility which prevent the user from burden of local storage issues. However, the biggest issues to be focused are on providing secure & reliable data archive over unreliable service providers. In this paper, we focus on using watermarking technology of copyright protection for Cloud Computing, Using GLCM & PCA algorithm to extract features of original image and to generate semi blind watermarking image. The collaboration of Digital Watermarking when used for Cloud Computing can significantly result to make the system robust as well as secure user's data.

Keywords: Cloud computing, watermarking, GLCM algorithm

I.INTRODUCTION

Cloud computing is regarded as massively scalable, an on demand configurable resource computing module. It offers the cloud infrastructure in a distributed rather than dedicated infrastructure where clients can have full access to the scalable reliable resources with high performance, everything is provided to the clients as a utility service that is publically accessible via internet. Cloud computing is a collection of computer & servers that are publically accessible via internet. Cloud computing allows consumers & business to use application without installation & access their personal files at any computer with internet access.

Digitization is occurring worldwide, which can be attributed to the rapid progress & advancement in IT. This phenomenon exhibits both the advantages & disadvantages. The problem regarding the ownership of digital media often draws interest from researcher digital information may be copied, attacked or altered during storage or transmission .Thus, effective working methods that protect digital data need to be developed. Thus, the security & confidentiality of such information should be seriously addressed.

II.CLOUD COMMUTING

Cloud Computing is internet based computing whereby shared resources, software & information are provided to consumers on demand, similar to the functioning of the electricity gird. Customers do not own or maintain the physical infrastructure & avoid capital expenditure by renting resources from a third party provider. They consume resources as-a-service and pay only for resources that they consume .Potentially reducing ongoing costs due to the use of infrastructure & technical specialists that are typically shared among many customer to achieve economies of scale, however the cost of applying controls to help address security risks especially associated with shared infrastructure may reduce the potential cost savings of some types of cloud computing.

The cloud of services & applications in internet modem is available from the computer. It is a web-based application which can be used easily by any person who is travelling without having facilities & servers to connect back to office through any kind of virtual private network (VPN). Cloud computing assess you faster & quicker to technology solution & thus reducing the business risks.

III.BENEFITS OF CLOUD COMPUTING

The potential for cost saving is the major reason of cloud services adoption by many organizations. Cloud Computing gives the freedom to use service as per the requirement & pay only for what you use. Due to cloud computing it has become possible to run IT operations as a outsourced unit without much in house resources.

Following are the benefits of Cloud Computing:

1. Lower IT infrastructure & computer costs for users.
2. Improved performance.
3. Fewer maintenance issues.
4. Instant software updates.
5. Improved compatibility between OS.
6. Backup and recovery.
7. Performance and scalability.
8. Increased storage capacity.
9. Increased data safety.

IV. TYPES OF CLOUDS

There are different cloud models that you can subscribe according to business needs:

1. Private cloud: Here computing resources are deployed for one particular organization. This method is more used for intra-business interactions where the computing resources can be governed, owned & operated by the same organization.
2. Community Cloud: Here computing resources are provided for a community & organization.

3. Public Cloud: This type of cloud is usually for B2C (Business to Consumer) type interactions. Here thecomputing resource is owned, governed & operated by government, an academic or businessorganization.

4. Hybrid Cloud: This type of cloud can used for both type of interactions-B2B (Business to Business) and B2C (Business to Consumer). This deployment method is called as different hybrid cloud and the computing resources are bound together bydifferent clouds.

Cloud Service Models*:*

Generally, Cloud Computing services are categorized into three types:

- **Software-as-a-Service (SaaS)**: It is a software distribution model where a third party provider hosts application and makes them available to customers over the high-speed internet connection.
- **Platform-as-a-Service (PaaS):** It is a middle layer which gives the organizations, institutions or companies a freedom and framework for developers to develop their own applications and deploy them and make customers within their company to across the resources.
- **Infrastructure -as-a-service (Iaas)**: Infrastructure is meet vital among the three service models because it is the basic need to launch the organization services over internet in a cloud platform, to make their services available to clients & applications to run them smoothly.

Cloud Computing is an attractive & exciting paradigm that comes with innumerous benefits, its flexibility agility & advantageous features make it the first priority to adept it. Some of the advantages are:

-Desirable costs: It's beneficial for cloud clients that when using cloud it avoid in investing in infrastructure like hardware & their upgradation. Thus cost efficiency is improved.

-Feasible with Demand: The demands are unpredictable, thus makes of easy to avail services like infrastructure, software & platform as the demand arises is order to match up the required demand of users. Resources when no more required can be withdraw at any point. Smooth Running of the Business, Cloud provides the infrastructure 24X7& monitors it at the back end, thus everything is maintain by cloud so that client doesn't suffer. And keeps data safe & secure so that customers business runs smoothly.

-Energy efficient Paradigm: Cloud Computing is energy efficient as it offers the solution which can protect our environment & save the deforestation & other unfriendly environmental activities. The cloud provides the online secure transactions which minimize the use of papers.

-Scalable Storage : The storage is no more a limitation when client are using cloud platform and they dont'have to buy now the blocky and costly hardware components like servers & storage devices etc. Scalability is the unique feature of Cloud Computing where dynamic provisioning of the resources is being done by the clients themselves within real time slice.

-Mobility: It is the best "on the go" feature provided by cloud. It makes cloud easy to operate from anywhere on the globe & clients can access their application & other resources from various devices.

-Software compatibility: Cloud provides typically support a specific set of software vendors and various software as service providers offers the compatible software to their customers in order to maintain the well defined software standard.

V.WATERMARKING

Digitization is occurring worldwide which can be attributed to the rapid progress & advancement in information technology .This phenomenon exhibits both advantages of disadvantages. The security of these digital data becomes an important problem. Digital information may be copied attacked or altered during storage or transmission. As efficient solution is digital watermarking which has been widely studied during recent years for the purpose of copyright protection, authentication, fingerprinting, copy protection, etc. Generally, Water marking is the technology of embedding a useful data (watermarking information) within a host signal. This embedding should not substantially degrade the perceptual quality of host signal.

A digital medium can refer to any kind of digital data, such as text, image, video or audio. Digital watermarking protects digital media & verifies its legitimate owner. Watermarking was developed from steganography. Bothtechniques use the concept of embedding information into cover data media.

Basic of Watermarking:

The embedding of information directly within the digital data which is also knows as raw or host data in order to generate a watermarked data from it is known as digital watermarking technique. The group of bits when inserted within the digital data file is known as watermarking. This can help in copywriting the information present on the file. The original data can be hidden from the external users with the presence of watermarked data on that source.

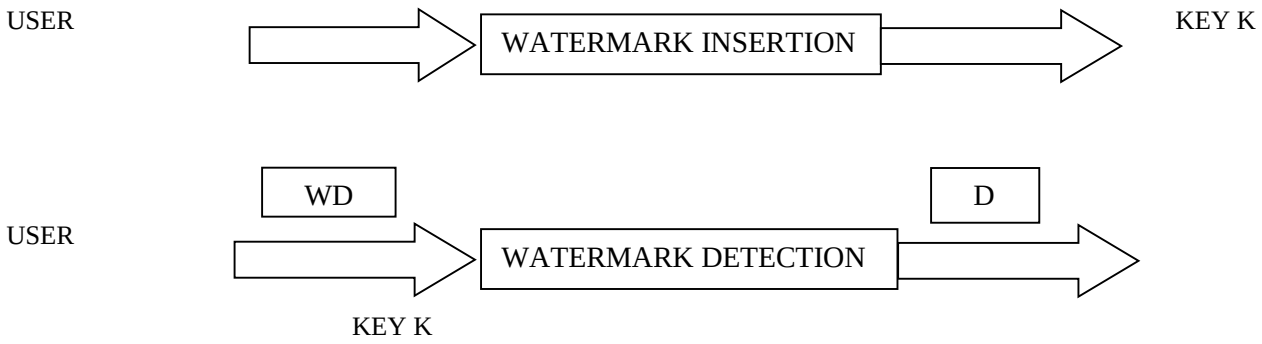

Figure.1: Watermarking Technique

In order to provide copyright protection to the intellectual property the digital watermarks are used. There are two phases present within the watermarking technique as shown in the figure1. The dataset (D) is added by the user with the presence of private key (K) in the first phase. The data achieved here is known as watermark data (WD). Further, the next stage which is phase2, the watermark dataset and the private key of the receiver is utilized in order to extract the embedded watermark. Thus, with the help of proof of the ownership, the original data (D) will be extracted.

VI.DIGITAL WATERMARKING TECHNIQUES:

On the basis of different types of documents present, the digital watermarking techniques are classified into various categories. They are:

1. Text Watermarking: In order to provide copyright protection to the text document, this approach is used. These are three types of digital watermarking provided here:
   - Line shift coding: The vertical shifting of location of text lines in order to encode the document is provided by this method.
   - Word shift coding: The document can be encoded through the horizontal shifting of the location ofwords.
   - Feature coding: Specific features are selected here and altered within his method.
2. Image Watermarking: The watermark in images is added by this method. It is not easy to remove the watermark from the image as it is already a part of that image.
3. Video Watermarking: The cryptographic information that is generated from the frames of digital video is used to provide the cryptographic information. This generated cryptographic information is embedded within the some video is this process amongst the original, unmarked and marked video, the user cannot easily differentiate
4. Audio Watermarking: An electronic identifier is embedded within the audio signal using this method. Within the audio file, the text or image are used to be embedded in such a manner within various techniques which can help is recovery of the text.

Characteristics of digital watermarking:

Digital watermarking can be classified as visible and invisible. The visible watermarks are viewable to the normal eye such as bills, company loges & television channel loges etc. This type of watermark is easily viewable without any mathematical calculation but these embedded watermarks can be destroyed easily. In the case of invisible watermark, the locations in which the watermark is embedded are secret, only the authorized person extracts the watermark. Some mathematical calculations are required to retrieve the watermark. This kind of watermark is not viewable by anordinary eye. Invisible watermarks are more secure and robust thanvisible watermarks.

The main characteristics of digital watermark are:

- Robustness: The watermark should be able to withstand after normal signal processing operations such as image cropping transformation, compression, etc.
- Imperceptibility: The watermarked image should look like same as the original image to the normal eye. The views cannot detectthe watermark embedded in it. Imperceptibility also known as invisibility or fidelity is the most significant requirement in watermarking system.
- Security: An unauthorized person cannot detect, retrieve or modify the embedded watermark. Security is the ability to resist against intentional attacks.
- Capacity: Capacity (also known as payload) refers to the number of bits embedded into the image. The capacity of an image could be different according to the application that watermark is designed for

Applications of digital Watermarking:

1. Copyright protection: The copyright information can be embedded as watermark into the production. Once hire is a dispute on the ownership the watermark can be extracted to provide the evidence of who is the owner of this product.

2. Image and Content Authentication: In an image authentication application the intent is to detect modification to the data. The characteristics of the image such as its edges are embedded & compared with the current image for differences. A solution to this problem could be borrowed from cryptography. Where digital signature has been studied as a message authentication method.

3. Tamper Proofing: Digital watermarks which are fragile in nature can be used for tamper proofing. Digital content can be embedded with fragile watermarks that get destroyed whenever any sort of modification is made to the content. Such watermarks can be used to authenticate the content.

4. Fingerprinting: A fingerprinting is a technique by which a work can be assigned a unique identification by storing some digital information is it in the form of watermark. Detecting the watermark from any illegal copy can lead to the identification of the person who has leaked the original content.

5. Broadcast monitoring: Broadcasting of TV channels and radio news is also monitoring by watermarking. It is generally done with the paid media like sports broadcast or news broad cost.

6. Media Forensics: Forensic Watermark application enhances a content owner's ability to detect and respond to misuse of its assets. Forensics watermarking is used not only to gather evidence for criminal proceedings, but also to enforce contractual usage agreements between a content owner and the people or companies with which it shares its content.

7. Medical Applications: Medical media and documents also digitally verified, having the in- formation of patient and the visiting doctors. These watermarks can be both visible and invisible. This watermarking helps doctors and medical application to verify that reports are not edited by illegal means.

VII. WATERMARKING TECHNIQUES:

Digital image watermarking schemes mainly fall into two broad categories:
Spatial-Domain and frequency domain techniques.

A. **Spatial Domain Techniques**:- Some of the spatial domain techniques of watermarking are follows.

i. Least Significant Bit (LSB): It is the most straight-forward technique of watermarking embedding. This scheme embeds watermark in the LSB of the pixel. Given an image with pixel, and each pixel being represented by an 8-bit sequence, the watermarking are embedded in the last (that is least significant), bit of selected pixels of the image. This method is easy to implement and does not generate serious distortion to the image; however, it is not very robust against attacks. For instance, attacks could simply randomize all LSBS, which effectively destroys the hidden information.

ii. SSM-Modulation-Based Technique: In Spread-spectrum modulation techniques the energy generated at different discrete frequencies is purposively dispersed or appropriated in time, for secure communication establishment, increasing resistance to natural interference and jamming and for preventing detection. SSM watermarking algorithm embeds information in content of image watermarking, it embed message by combining the cover image with a small pseudo noise signal modulated by the added watermark.

B. Frequency Domain Technique: The aim of this technique is to embed the watermarks in the spectral coefficients of the image. The most commonly used transforms are the discrete cosine transform (DCT), discrete Fourier transform in the frequency domain is that the characteristics of the human virtual system (HVS) are better capture by the spectral coefficients. This techniques provide, more information hiding capacity and high robustness against various geometrical attacks. In this paper we applied DWT techniques.

Discrete Wavelet transforms (DWT):

Wavelet transform is a modern technique frequently used in digital image processing, compression, and watermarking, etc. DWT is a mathematical tool for hierarchically decomposing an image. It is useful for processing of non-stationary signals. The transform is based on small, called wavelets, of varying frequency and limited duration.

DWT is the multiresolution description of image the decoding can be processed sequentially from a low resolution to the higher resolution. The DWT splits the signal into high and low frequency parts. The high frequency components are usually used for watermarking since the human eye is less sensitive to changes in edges. In two dimensional applications, for each level of decomposition, we first person the DWT in vertical direction, followed by the DWT in the horizontal direction. After the first level of decomposition, there are 4 sub-bands: LL1, LH1, HL1, and HH1, for each successive level of decomposition, the LL sub-band of the previous level is used as the input. To perform second level decomposition, the DWT is applied to LL1 band which decomposes the LL1 band into the four sub band. To perform third level decomposition, the DWT is applied to LL2 band which decompose this band into the four sub band, LL3, LH3, HL3, and HH3. This result in 10 sub band per component LH1, HL1, and HH1 contain the highest frequency band. Again we decompose to four level of decomposition, so DWT is currently used in wide range of signal processing application, such as audio and video compression, removal of noise suits many applications very well.

VIII.PROPOSED METHODOLOGY

The watermarking is the efficient technique to provide security to the image data. The watermarking techniques are broadly classified into blind and semi-blind watermarking technique. The main objectives of proposed technique are:

1. To proposed improvement in DWT technique for the generation of semi-blind watermarking.
2. The proposed technique will be based on the GLCM algorithm to analyze feature of the original image.
3. To implement proposed technique and compare with existing in terms of PSNR, BER, and MSSIM.

The proposed algorithm includes there steps: Decomposing the cover image, embedding and extraction. A binary watermark image will be used as the watermark for embedding. The textual feature of the image is analyzed with the help of DWT algorithm, with the help of GLCM algorithm. This algorithm is utilized as it is less complex in nature and optimal for generation of sets for blind watermarks.

IX.GRAY-LEVEL CO-OCCURRENCE MATRIX (GLCM):

In this paper, gray level co-occurrence matrix is formulated to obtain statistical texture features. A number of texture features may be extracted from the GLCM. In statistical texture analysis, texture features are computed from the statistical distribution of observed combination of intensities at specified positions relative to each other in the image. According to the number of intensity points (pixel) in each combination statistics are classified into first order, second order and higher order statistics. The GLCM is a way of extracting second order statistical texture features. Third and higher order texture consider the relationships among three or more pixel. These are theoretically possible but not commonly implemented due to calculation time and interpretation difficulty.

GLCM is a tabulation of how often different combination of pixel brightness values occur in an image. GLCM contains the information about the position of pixel having similar gray level values. GLCM calculation units receive pairs of gray level values as input. The GLCM calculation unit consists of the different combination of gray value like a0b1, a2b3, a10b21, etc. This gives the deviation present in the image when compared with original image by predictive image.

This gray-level co-occurrence matrix (GLCM) considers the relationship between two neighboring pixel, the first pixel is known as a reference and the second is known as a neighbor pixel. The GLCM is a square matrix with Ng dimension, where Ng equals the number of gray level in the image. Each element of the matrix is the numbers of occurrence of the pair of pixel with value i and a pixel with value j. A co-occurrence matrix is a two dimensional array in which both rows and columns represent a set of possible image values. For example consider the 4 by 5 matrix of image I.

1	1	5	6	8
2	3	5	7	1
4	5	7	1	2
8	5	1	2	5

fig: Image Matrix

Equivalent GLCM matrix for the above image I is

1	2	0	0	1	0	0	0
0	0	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
1	0	0	0	0	1	2	0
0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0

fig : GLCM Image Matrix

After you create the GLCMs, you can derive several statistics from them using matlab function. These statistics provide information about the texture of an image. According to co-occurrence matrix, there are fourteen textural features measured from the probability matrix to extract the characteristics of texture statistics of remote sensing image. Some of the features described below.

1. Angular second moment: - It is also known as uniformity or energy. It is the sum of squares of entries in the GLCM angular second moment measures the image homogeneity. It is high when image has very good homogeneity or when pixels are very similar.
2. Inverse difference moment (IDM): It is the local homogeneity. It is high when local gray level is uniform and inverse GLCM is high.
3. Entropy:- It shows the amount of information of the image that is needed for the image compression. Entropy measures the loss of information or message in a transmitted signal and also measures the image information.
4. Correlation: It measures the linear dependency of gray level of neighboring pixels.

The Proposed Algorithm Can Applied In Following Steps:

1. Input the gray scale cover image of size 512x512.
2. Apply DWT of level-4 to obtain matrix of size 32x32 with LL4 coefficients.
3. Referenced the obtained image into vector Ui of size 1024x1.
4. Now argument the Vi vector with its quantized vector V of same size in order to construct set of size 1024x2.

5. Generate an output of size 1024x1.
6. Quantize the data of image to obtain the pixels in order to create GLCM square matrix of size NxN.
7. Normalize the GLCM is each element (i,j) of matrix is divided by its total number of elements then obtained elements are considered as probabilities of finding relation and extract the texture feature.
8. Now use PCA algorithm to select feature from extracted features of step 7 and it dynamically select the scaling factor R .
9. Use following equation to embed the watermarking image (yi) into cover image according to the predicted output of step 5, by using R obtained from step 8.

$$Ui = Zi + \alpha * Yi$$

10. Final semi blind watermarked image is generated.

X.RESULT ANALYSIS

In this section, we demonstrate the effectiveness of our proposed methodology. The simulation is done on MATLAB and analysis of PSNR and robustness of image. This method is applied to several images having different types of pixels. We measure the quality of watermarked images in terms of PSNR, MSE, BER, and MSSIM.

PSNR: Peak to signal ratio: It is used to measure the imperceptibility of the watermarked and the extracted watermark images. It is defined by the mean squared error between the corresponding pixel values of the cover image and the watermarked image. This function is widely used because of its simplicity and clarity. If the value of PSNR is higher it means the reconstructed image had better quality.

MSE: Mean Square Error: It represents the cumulative squared error between the compressed and the original image. The lower the value of MSE, the lower will be the error.

BER: Bit Error Rate : The Bit Error Rate is the number of bit errors per unit time. The BER is the number of bit errors divided by the total number of transferred bits during a studied time interval.

MSSIM: Multi Scale Structural Similarity Index Measurement: SSIM is a perceptual metric that quantifies image quality degradation caused by processing such as data compression. It is a method for measuring the similarity between two images. Using SSIM as a basis, MS-SSIM extends the technique by making multiple SSIM image evaluations at different image scales. This is accomplished by repeatedly performing the image analysis in multiple iterations, with each successive image pair that is downsampled by a factor of 2 from the previous iteration.

Performance Analysis

	Parameter Values	Leena	Taj	Cat	Rcert
WATERMARKED IMAGE	PSNR	24.40	27.36	24.70	25.17
	MSE	238.16	120.39	221.88	199.27
CONTRAST ATTACK	PSNR	24.38	27.80	25.16	25.06
	MSE	239.19	108.78	199.73	204.47
SHARPENED ATTACK	PSNR	19.82	25.78	24.52	14.69
	MSE	683.63	173.30	231.53	226.58
SALT & PEPPER ATTACK	PSNR	30.06	31.34	30.28	30.31
	MSE	64.65	48.16	61.43	61.01

Following is the analysis of values when we applied our proposed methodology :

	Parameter Values	Leena	Taj	Cat	Rcert
WATERMARKED IMAGE	BER	0.07	0.08	0.09	0.05
	MSSIM	160.08	64.09	90.69	224.09
CONTRAST ATTACK	BER	0.00	0.18	0.06	0.06
	MSSIM	255.00	133.70	239.52	233.65
SHARPENED ATTACK	BER	0.02	0.08	0.07	0.21
	MSSIM	253.81	125.80	231.96	219.69
SALT & PEPPER ATTACK	BER	0.01	0.9	0.03	0.04
	MSSIM	254.57	157.54	234.56	232.72

The Proposed algorithm is implemented in MATLAB by considering the authentic dataset.

XI.CONCLUSION

With the development of Cloud computing, cloud security remains as a hot spot issue. This paper Tries to provide a new insight into the essence of cloud security and proposes a new method based on watermarking which can be applied to solving the trust management between data owners and service providers. In this paper, efficiency of watermarking approach is concluded as it hides all the sensitive information which is stored in the form of images. Here GLCM and PCA algorithm has been utilized in order to improve the working capability based watermarking technique .The extracted feature of an image are selected by PCA algorithm and features of the original image are extracted by the GLCM algorithm. On the basis of simulation results, it is concluded that proposed algorithm performs well in terms of PSNR and MSE. In the future, we will conduct more experiments in cloud computing and test the performances of our approach which may provide a reference solution for cloud security.

References:

[1] U. Yadav, J.P. Sharma, D. Sharma and P.K. Sharma, "Different WatermarkingTechniques and its Applications:A Review", International Journal ofScintilla and Engineering Research, Vol. 5, no.4, (2014) April.

[2] Chih-Chin Laiand Cheng-ChihTsai, "Digital Image Watermarking Using DiscreteWavelet Transform and Singular Value Decomposition", IEEE TRANSACTIONS ON INSTRU- MENTATION AND MEASUREMENT, VOL.59, NO.11, NOVEMBER 2010.

[3] C.-c. Lai and c.-c. Tsai, "Digital Image Watermarking Using Discrete Wavelet Trans- formand Singular Value Decomposition", IEEE Transactions on Instrumentation and Measurement, vol. 59, no.11, (2010) November.

[4] M. Narang and S. Vashisth, "Digital Watermarking using Discrete Wavelet Transform" International Journal of Computer Applications (1975-88887) vol.74, no.20, (2013) July.

[5] Salama, A., Atta, R. , Rizk, R. Wanes, F., "A robust digital image watermarking technique based on wavelet transform". In: IEEE Int. Conf. on Sys. Eng. and Tech., pp. 100-104 (2011).

[6] Ahmed S. Salama, Mohammed A. AI- Qodah, Abdullah M. Tliyasu, Awad Kh. AI-Asmari and Fei Yan: A Hybrid Fusion Technique for watermarking Digital Images: Advances in Intelligent Systems and Computing Volume 240, pp 207-217, (2014).

[7] lliyasu, A., Le, P., Dong, F., Hirota, K.: Watermarking and authentication of quantum images based on restricted geometric transformations. Information Sciences 186(1), 126-149 (2012). 562.

[8] AI-Asmari, A., Salama, A., T1liyasu, A., AI-Qodah, M.: A DWT ordering scheme for hiding data in images using pixel value difference. In: IEEE Eighth Int. Conf. on Computational Intelligence and Security(CIS), pp. 553-557 (2012).

[9] Abid Khan, Ayyaz Yaqoob, Kinza Sarwar, Mouzna Tahir, Mansoor Ahmed, "Secure Logging as a Service Using Reversible Watermarking", The 12th International Conference on Future Networks and Communications, (FNC-2017)

[10] Rita Choudhary, Girish Parmar, "A Robust image Watermarking Technique using 2-level Discrete Wavelet Transform (DWT) ", IEEE 2nd International Conference on Communication, Control and Intelligent Systems (CCIS)

[11] Mr. Y. Gangadhar, Dr. V.S. Giridhar Akula, Dr. P. Chenna Reddy, " A Survey on Geo metric Invariant Watermarking Techniques" ,2016 IEEE

[12] Ahmed S. Salama, Mohamed Amr Mokhtar, "Combined Technique for Improving Digital Image Watermarking", 2016 2nd IEEE International Conference on Computer and Communications

[13] Mr. R. D. Shelke, Dr. Milind U. Nemade, "Audio Watermarking Technique Protection : A Review", 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication

[14] Chengxiang Yin, Jin Hu, Xuejun Zhang, Xiang Xie, "Advertising system based on cloud computing and audio watermarking", 2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing

[15] Muhammad Imran and Bruce A. Harvey, Adnan Ali Memon, "A Novel Blind Color Image Watermarking Technique Based on Singular Value Decomposition and Principal Component Analysis", 2016, The Sixth International Conference on Innovative Computing Technology.