Detecting Fake Profile Using Java Static Watermarking and Image Compression Using Huffman Coding

F. Mary Harin Fernandez, B.Manikandan, K. Manojkumar

Assistant Professor, Department of Computer Science and Engineering, Jeppiaar SRR Engineering College.

Abstract-Users all around the world are using online social network, such as Facebook, Twitter, Tumbler and LinkedIn. These are the weaknesses that makes effort less for frauds to misuse user's information and to create identity cloning attack to form fake profile. In this proposed system, the data hiding techniques are being used to hide some information in pictures, so that it can be used to detect botnets and fake profiles. This project presents a classification and analysis of detection mechanisms of clone attacks on online social network, based on attribute similarity, friend network similarity, and profile analysis for a time interval and record of Internet Protocol sequences. In this project, have proposed discrete wavelet transform algorithm for data hiding. Thus this would provide a solution to the clone attacks and providing complete user data privacy preserving. In Huffman coding an input image is split into equal rows & columns and at final stage sum of all individual compressed images which not only provide better result but also the information content of authorized user will be kept secure. It has been shown that the image compression using Huffman coding provides better image quality than image compression using BTC and AMBTC. Moreover, the Huffman coding is quite faster compared to BTC.

Keywords- clone attacks, DWT, Huffman encoding.

I. INTRODUCTION

Social networks play a bigger role in every day's life for sharing personal information's and knowledge sharing. Social network plays a major role in creating new contacts and maintaining the existing contacts. Also social network is used to see the everyday activities, photos, videos and political activities happening. Thus the major concern in social network is fake users misusing the authorized users information's like text, photos and videos with or without the authorized user's permission. Currently many fake profiles are been created for fraudulent activities like money making, malware / virus / Trojan virus distribution to track the user behaviors. Also a research states that many fake Profiles are been created for online social games. In existing social network, the individual information's are kept secret and transmitted in a secure manner with user privacy preserving. Thus even intruders have breached the privacy preserving in many scenarios. Thus many privacy attacks have been identified now days to steal the user personal information for malicious activities. Thus this leads to a huge security problems. CNET report in the year 2013 states Facebook has 8.7% fake users which are approx. 83.09 million users. From [1], the field of image compression continues to grow rapidly. Image compression technique involves reducing the size of image data files, while retaining necessary information. Retaining necessary information depends upon the application. Image segmentation methods, which are primarily a data reduction process, can be used for compression. The reduced file created by the compression process is called the compressed file and is used to reconstruct the image, resulting in the decompressed image. The images before any compression techniques is performed, is called the uncompressed image file. The ratio of the original, uncompressed image file and the compressed file is referred to as the compression ratio.

II. RELATED WORK

- Wavelet Transform

From [2] a Wavelet transform is a transform analysis method, which is known as a microscope of signal analysis. It inherits and develops an idea of localization of short Fourier transform, and it overcomes the shortcoming that the size of window is fixed. It is an ideal tool for time-frequency analysis and processing of signals, and can provide time-frequency windows with frequency variations. Wavelet transform can decompose the image into four parts, namely low resolution (LL), the other three corresponds to horizontal (HL), vertical (LH) and diagonal (HH), this can be further decomposed into high frequency parts and vice versa. Wavelet decomposition can be applied to the image compression and detail enhancement.

Quality constrained compression algorithm (QCC) based on discrete wavelet transform (DWT) is proposed. This spatial-frequency decomposition property of DWT provides possibility not only for the new compression algorithm but also for frequency-domain quality assessment method. To perform this new algorithm, a new metric WNMSE is preferred, which is used assess the quality of an image with comparing the weighted sum of normalized mean square errors of the wavelet coefficients. Thus the metric is consistent with the human judgment of visual quality as well as it is able to estimate the quality during the compression process.

- Image Compression based on Huffman encoding

The image compression techniques are classified into two namely Lossy compression techniques and Lossless compression techniques.

Lossless compression: A technique in which the compressed image is reconstructed without any loss of data is called lossless compression. Lossless compression ratio provides good quality of compressed images, but gives only less compression.

Lossy compression: A technique in which the compressed image is reconstructed with loss of data is called lossy compression. The lossy compression techniques leads to loss of information with higher compression ratio. Huffman coding is loss less technique with more attractive features in various application such as medical survey and analysis, technical drawing etc. Huffman coding has better characteristics of image compression. Huffman's coding algorithm is a step by step process and involves the variable length codes to input characters & it is helpful in finding the entropy and probability of the state [3]. It is very easy to calculate quality parameter in Huffman algorithm. Original image can be reconstructed with the help of digital image restoration.

- Image Watermarking based on DWT

Recent researchers on secure digital watermarking techniques from [5], have revealed the fact that the content of the images could be used to improve the invisibility and robustness of a watermarking scheme. In this approach, watermark of an image is created from the content of the host image and discrete wavelet transform (DWT) is used for embedding watermarks on user images, since it is an excellent time frequency analysis method which can be adapted well for extracting the information content of the image.

It adopt a key dependent wavelet transform. In this approach, the host image is transformed into wavelet coefficients using a discrete-time wavelet transform (DWT). Each watermark bit is embedded using two super trees. One of the super tree is quantized based on quantization index in such a way that the two super trees exhibit a large enough statistical difference, which can be extracted for obtaining decision. The watermarking technique is robust to attacks in both frequency and time domains, since watermark bit is embedded in various frequency bands and the information of the watermark bit is spread throughout large spatial regions. This technique is useful for removal of high-pass details in JPEG compression and robust to time domain attacks such as pixel shifting and rotation as shown in fig 1. In addition to copyright protection, the proposed watermarking scheme can also be applied to data hiding or image authentication.

And further proposes a discrete-wavelet transform based multiple watermarking algorithms. In this approach, two important tools encryption and java static watermarking can be used to prevent unauthorized data or information. To improve the robustness of an image, watermarks are being embedded into LL and HH sub bands. This approach is used in such a way that embedding the watermark in lower frequencies is used to robust a group of attacks such as JPEG compression, blurring and adding Gaussian noise and also embedding the watermark in higher frequency domain is used to robust another set of attacks such as histogram equalization, intensity adjustment, gamma correction. Introduced an integer wavelet based watermarking techniques to protect the copyright of an image compression technique to enhance the security. For DEM (digital elevation mode) data, this technique is very useful for digital watermarking in which effectively protects the copyright of DEM data and avoids the unauthorized user. As lifting based scheme is added to construct the compactly supported wavelets whose coefficients are composed of a free variable therefore, it uses only integral addition and shift which is fast and easily realized via hardware.

Fig. 2.1: Steganography image using DWT algorithm.

- Wavelet-Base Watermarking Algorithm for Ownership Verification of Digital Image

Embedding the watermarks into the images, filter bands can be saved and the middle-frequency band to insert the watermark is chosen. The coefficient in that band of the image is used to replace the watermark image as shown in fig 2.

Fig 2.2. Watermarked image.

- A Hybrid Watermark for Tamper Detection in Digital Image

From [6] the fragile watermark image has the advantages such as good localization and security properties. The hybrid watermarks are used to identify changes as well as distinguish malicious data.

The authentication can be done without accessing any information about the original image. For Effective Hybrid Digital Watermarking, Direct Sequence-Spread Spectrum Method is used and watermark image is produced using the personal ID of copyright which is inserted into the original images and the watermark image is detected. This technique is an extension of the spread-spectrum watermarking scheme which combine key with logo method. Values of PSNR of the watermark images are used to check the degradation of quality of the original and watermark image to confirm required invisibility and watermark robustness is applied to protect an attack from the outside.

III. PROPOSED WORK

From [7], In the proposed approach, steganography techniques and methods will be used to detect and identify such fake profiles. In this method, whenever a user uploads user's pictures, some exclusive and information's like email or username and also date of upload would be attached to pictures by using the watermarking methods. In our proposed system the use discrete wavelet transform algorithm for data hiding. Thus this would provide a solution to the clone attacks and providing complete user data privacy preserving. Also when users upload the profile picture or photos it would be watermarked and updated. For watermarking, a technique called Java static watermarking systems and algorithms is been used.

- Analysis of Huffman Encoding Algorithm

This matrix represent digital image N x N. These matrix arrays given in matrix are the elements of image.

$$f(x,y)= \begin{matrix} b_{0,0} & b_{0,1} \cdots \cdots & b_{0,M-1} \\ b_{1,0} & b_{1,1} \cdots \cdots & b_{1,M-1} \\ b_{M-1,0} & b_{M-1,1} \cdots \cdots & b_{M-1,M-1} \end{matrix} \qquad (1)$$

In equation (1), this digital image f(x, y) is break into a set of non-overlapping four sub images i.e. two row and two column this can be represented as

f(x, y) which is a digital image is divided into four small images. These small images is also called as non-overlapping sub images.

$$f(x,y)=\begin{bmatrix} f_{1(x,y)} & f_{2(x,y)} \\ f_{3(x,y)} & f_{4(x,y)} \end{bmatrix} \qquad (2)$$

In equation (2), $f_{1(x,y)}$, $f_{2(x,y)}$,$f3(x,y)$, $f4(x,y)$ are the sub-matrix of original image after applying Huffman coding on these sub-matrix they gives f1` (x, y),f2 `(x, y),f3 `(x, y),f4 `(x, y) respectively the compressed image can be obtain by adding these matrixes.

- Watermark-embedding algorithm of an image

The watermark-embedding method is shown in fig 3. This algorithm mandatorily includes the following steps: wavelet decomposition, block splitting, watermark embedding, wavelet reconstruction, and watermarked image testing.

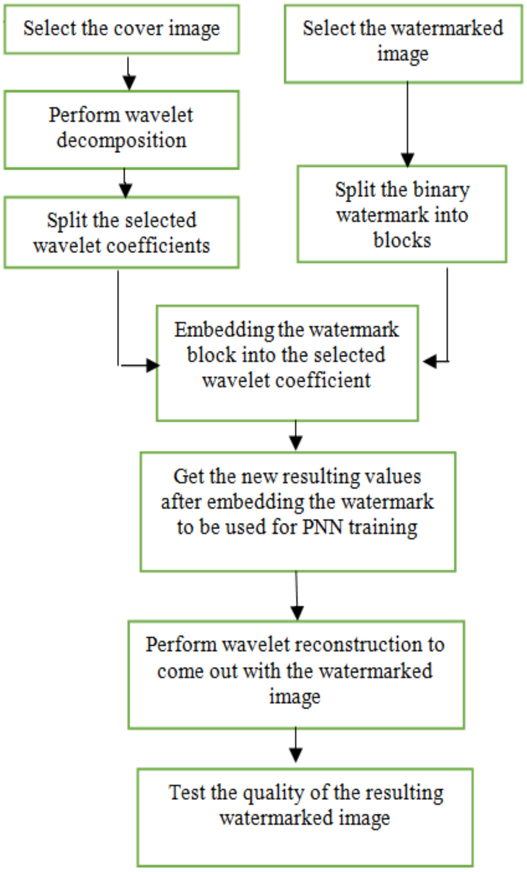

Select the cover image
Select the watermarked image
Perform wavelet decomposition
Split the selected wavelet coefficients
Split the binary watermark into blocks
Embedding the watermark block into the selected wavelet coefficient
Get the new resulting values after embedding the watermark to be used for PNN training
Perform wavelet reconstruction to come out with the watermarked image
Test the quality of the resulting watermarked image

Fig. 3.1 Steps for applying steganography and watermark.

- Architecture Diagram

The original user logins to account and user's IP and MAC address is stored in a database. When a user uploads an image the image is first watermarked and then uploaded to the server. When the second user downloads the image of the original user the respected systems IP and MAC address is also stored in the database. If the second user modifies and uploads the image to the server, the server using the watermark and IP and MAC address a notification of allow or block is sent to the original user to allow or block the image to upload as shown in fig 4.

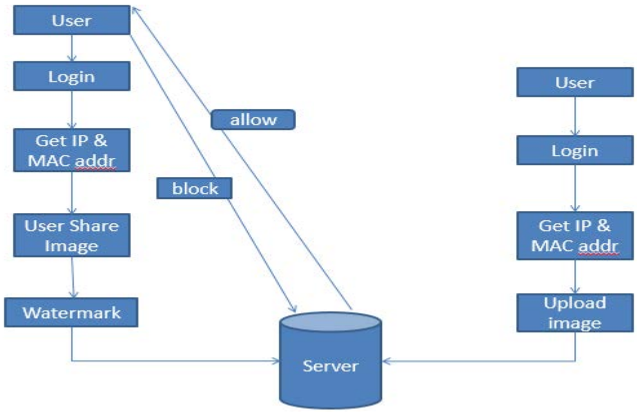


Fig 3.2: Providing security for user profile.

The image which is uploaded securely is then uploaded to the third party storage which is the cloud. In order to provide additional security to the user files a method called steganography and compression is used. Where steganography is provided using discrete wavelet transform algorithm (DWT) and compression using Huffman encoding. The image is provided with steganography, which is adding a cover image to the original image using DWT algorithm. Now the image which is steganographed is then compressed using Huffman encoding algorithm and then uploaded to the cloud.

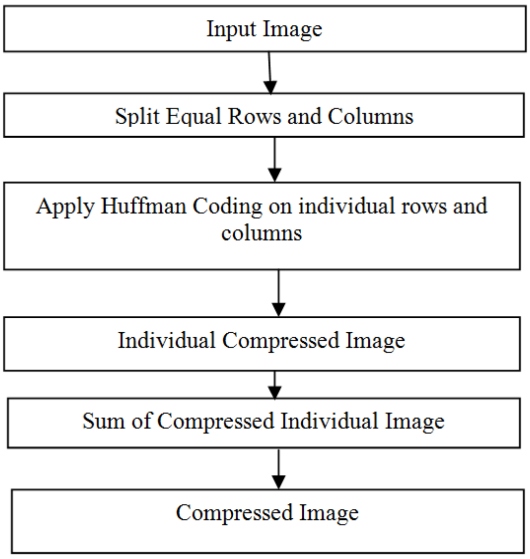

Input Image
Split Equal Rows and Columns
Apply Huffman Coding on individual rows and columns
Individual Compressed Image
Sum of Compressed Individual Image
Compressed Image

Fig 3.3: Block Diagram of Huffman algorithm

When the image is retrieved by the user first the cover image is downloaded then using Huffman encoding algorithm the original image is obtained as shown in fig 5.

IV. EXPERIMENTAL RESULTS

In the experiment, the used 512×512 pixel image as cover image. Two different watermark images of 64×64 pixel size image is used as watermark.

- Image Quality Parameter

There are four major important parameters measure between uncompressed image and compressed image [3], these are following

A. Compression ratio (CR)

Compression ratio is used to measure the compression efficiency. Compression ratio is the ratio of original image and compressed image. As compression ratio increases the image quality increases. CR=Size of original image/Size of compressed image.

B. Bit Rate:

It is information (bits) stored per pixel of an image. This is ratio of number of bits in the compressed image to total number of pixel in original image. Number of bits per pixel required by the compressed image. BR= (b/CR) b= No. of bits per pixel. When bit rate is large it means large memory required to store an image. High bit rate indicate that image acquire more colors so bit rate should be less.

C. The Mean Squared Error (MSE):

The difference between original image data and compressed image data is called mean square error (MSE). MSE is inversely proportional to PSNR, as MSE decreases the PSNR increases. PSNR indicate quality of image. Image compression is lossless when MSE is zero. Its better to have less MSE

D. Peak Signal to noise Ratio (PSNR):

PSNR is the ratio between maximum signal powers to noise appear in signal. PSNR is related to quality of image. For good quality of image the PSNR of image should be high. PSNR is depends upon the mean square error (MSE) of image. When the difference between the original image and compressed is less the PSNR is high so eventually the quality of image is also high.

- Extraction Tests

The performance of the extraction method is evaluated by means of measuring imperceptibility and robustness. The NCC value is used to measure the image quality of the watermark after extraction.

$$NCC=\sum i\sum jW(i,j)\cdot W'(i,j)\sum i\sum jW2(i,j) \qquad (3)$$

In equation (3), $W(i, j)$ and $W'(i, j)$ are pixel values at the i, j locations of the original watermark and the extracted watermark image, respectively.

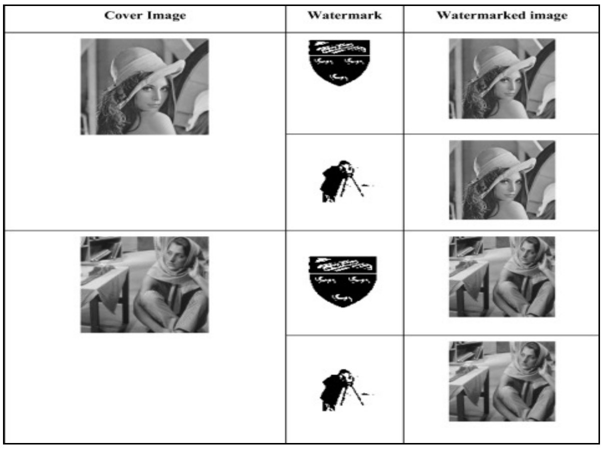
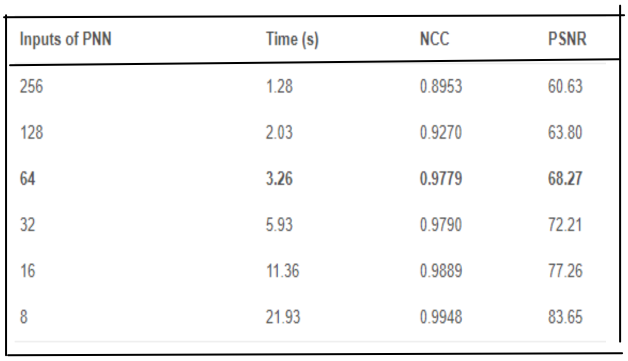


Fig 4.1: Images after applying cover image and watermarking.

To test effectiveness of the extraction algorithm, the inputs of PNN response to the watermarked image should be considered. The below table summarizes the results obtained using different input vector sizes of PNN in the attack-free case. In this test, a 512 × 512 pixel gray Lena image is used as a cover image, whereas a 64 × 64 pixel binary image of the UM Logo is taken as a watermark as shown in fig 6.

Table 1: Quality of Watermark Extraction



Inputs of PNN	Time (s)	NCC	PSNR
256	1.28	0.8953	60.63
128	2.03	0.9270	63.80
64	**3.26**	**0.9779**	**68.27**
32	5.93	0.9790	72.21
16	11.36	0.9889	77.26
8	21.93	0.9948	83.65

According to Table 1, the extraction quality of the watermark images is degraded as the number of PNN inputs increases. This effect is observed because of increasing number of PNN value inputs indicate fewer neural networks are available for the extraction of watermark from the image, such that, the accuracy and execution time of PNN will get decrease. However, when number of PNN inputs decreases, more neural networks can be used for extraction of the watermark, so that the accuracy and extraction time of PNN will get increased.

- Robustness Tests

To prove the robustness of proposed algorithm, the investigated the effect of the following attacks on watermarked images:

(a)JPEG compression (quality factor = 70, 50, and 10).

(b)Image rotation (with rotation angle = 5° and 45°).

(c)Gaussian noise (with standard deviation $\sigma = 20$ and 50).

(d)Image cropping (up to 25%).

(e)Median filter.

The extracted mages are partially degraded after JPEG, rotation, and cropping attacks. However, the extracted watermark images remain recognizable because these attacks change the indexed reference values, which might contain the watermark location values of the embedded values. Thus, during the retrieval of data, the extracted values may not be the watermarked values of indexed location values, which have been changed.

In the test extraction process, each attack has been tested with each previously described watermarked image. After JEPG compression, cropping, Gaussian noise, rotation, and median filter attacks it shows the watermarked image, as well as the extraction results for the two watermark binary image,

CONCLUSION

This project presents a brief knowledge about the attacks and defense mechanisms which are prominent on Online Social networks. It also explains the work which had been performed in the field of detecting and preventing clone profiles and secure transmission of images from sender to receiver.

FUTURE WORK

In future, a new hybrid dynamic approach can be developed by integrating the content free approach with content based. Additional features like evaluation of login action time pattern and click pattern of users or observing users activities (content free) can be included, along with the threshold put on similarity of users value attributes (content based). Another framework can be implemented by analyzing the genuine user activity of sending friend request and putting a threshold on suspicious cloned profile (having same name and other attributes to that of user) for sending invitation to join clone accent.

REFERENCES

[1]. Nehal Markandeya, Prof.Dr.Sonali Patil, “Image Compression Using Huffman coding”,International journal of Engineering and Computer Science, Volume:6, Issue:1, pp.19999-20002, 2017.

[2]. Chunyan Zhang, Shuangshuang Wang,“An encrypted medical image retrieval algorithm based on DWT-DCT frequency domain ”,IEEE 15th International Conference on Software Engineering Research, Management And Application, London, ISBN:978-1-5090-5756-6, pp.135-141, June 2017.

[3]. Rachit Patel, Virendra Kumar, Vaibhav Tyagi, Vishal Asthana, “A Fast and Improved Image Compression Technique Using Huffman Coding”, International Conference on Wireless Commnication, Signal Processing and Networking, India, ISBN:978-1-4673-9338-6, pp. 2283-2286, March 2016.

[4]. M.Barek Marwan, Ali Kartit and Hassan Ouahmane, “Applying Homomorphic Encryption for securing Cloud Database”, International Conference on Information Science and Technology, Morocco, ISSN:2327-1884, pp.658-664, October 2016.

[5]. Yahya Al-Nabhani, Anuddin Wahid, Rafidah MD Noor, “Robust Watermarking Algorithm for Digital Images Using Discrete Wavelet and Probabilistic Neural Network”, Journal of King Saud University – Computer and Information Sciences, Volume: 27, Issue:4, pp.393-401, 2015.

[6]. Rayachoti Eswaraiah, Edara Sreenivasa Reddy, “Robust Medical Image Watermarking Technique for Accurate Detection of Tampers Inside Region of Interest and Recovering Original Region of Interest”, Institute of Engineering and Technology Image Processing, Volume:9, pp.615-625, 2015.

[7]. Sangita Zope-Chaudhari, Parvatham Venkatachalam, Krishna Mohan Buddhiraju, “Secure Dissemination and Protection of Multispectral Images using crypto Watermarking”, IEEE Journal of Selected Topics in

Applied Earth Observations And Remote Sensing, Volume:8, Issue:11, pp.5388-5394, 2015.

[8]. Chin-chen chang, chin-yu sun, ”Polynomial-Based Secret Sharing Scheme Based on the Absolute Moment Block Truncation Coding Technique”,2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Japan, ISBN:978-1-4799-5390-5, pp. 485-488, December 2014.

[9]. A.M.Raid, W.M.Khedhr, M.A.El-Dosuky and Wesam Ahmed, “Jpeg Image Compression Using Discrete Cosine Transform- A Survey,” International Journal of Computer Science and Engineering Survey, Volume: 5, Issue: 2, pp.39-47, 2014.

[10]. Jagdish H, Pujar, Lohit M and K. Kadlaskar "A New Lossless Method of Image Compression and Decompression Using Huffman Coding Technique", Journal of Theoretical and Applied Information Technology, Volume:15, pp.18-23, 2010.

[11]. Zhigang Gao, and Yuan F. Zheng, “Quality Constrained Compression Using DWT-Based Image Quality Metric”, IEEE Transactions on Circuits And Systems for Video Technology, Volume:18, Issue:7, pp.910-922, July 2008.