Authentication of Grayscale Document Images by using PNG Image with a Data Repair Capability

T.M.Prasanth[1], K.Jansi Lakshmi[2], I.Suneetha[3]
[1]M.Tech (DSCE) Student, [2]Assistant Professor, [3]Associate Professor & Head
Department of ECE, AITS
Annamacharya Institute of Technology and Sciences, Tirupati, India-517520
[1]tmprasanth@gmail.com, [2]jansikaramala@gmail.com, [3]iralasuneetha.aits@gmail.com

Abstract: **-** Digital image is a form for preserving vital information. However, with available digital technologies, one can easily make the visually undetectable modifications to the contents of digital images. Hence ensuring the robustness of a digital image is thus a challenge. This paper proposes an approach for secret sharing of information Shamir scheme is used while sharing a document image and at the receiver reverse Shamir scheme is used. For every block of a grayscale image an authentication signal is generated along with the binarized block content. These binarized contents are transformed into several shares using the Shamir secret sharing scheme. During the course of image authentication the tampered image block is marked and it is matched with the original saved current block. If it does not match with original block, then self-repair mechanism is applied to the tampered block and it is repaired automatically by a reverse Shamir scheme. These type of image content authentication and self-repair capabilities are useful for the security protection of the digital documents in many fields**.**

Index Terms: **-** Portable Network Graphics, Image authentication, secret sharing**.**

I. INTRODUCTION

Nowadays important information is preserved in the form of digital images. It is easy to make modifications in the content of digital images, with the fast advance of digital technologies. To ensure the integrity and authenticity of a digital image while sharing is not secure and so image authentication problem arises[1-2].

The image authentication problem is difficult for a binary document image because of its simple binary nature modifications can be done easily in the binary document image so that the intruders can attack the image easily. Several methods are proposed for the binary image authentication problem [3]. A two layer binary image authentication method in which one layer is used for checking the image fidelity and the other for checking image integrity is proposed. In this method a connectivity-preserving transition criterion for embedding the cryptographic signature and the block identifier [4]. Its time consuming, so another method with three transition criteria are used to determine the flippabilities of pixels in each block to deal with the uneven embedability condition in the host image. It is more complex. A set of pseudorandom pixel in a half tone image are chosen and cleared and authentication codes are computed and inserted in to selected random pixels [5]. A randomly generated authentication codes are embedded into image blocks and code holder is used to reduce the image distortion. A hamming code based data embedding method that flips one pixel in each binary image block for embedding a watermark yields small distortion and low false negative rates.

An image content authentication with a data repair capability for grayscale document image via the use of the Portable Network Graphics (PNG) image is proposed [6]. To create a desired degree of transparency for the image, the alpha channel of the PNG image is used. If the alpha channel effect is less than 255, imperceptible transparency effect is obtained [7].

PNG supports palette-based images, grayscale images and full-color non-palette-based RGB images. PNG images can either use palette-indexed color or be made up of one or more channels. When there is more than one channel in an image all channels have the same number of bits allocated per pixel known as the bit depth of the channel. Although the PNG specification always described the bit depth of channels, most software and users generally described the total number of bits per pixel. It is known as bit depth or color depth. Since multiple channels can affect a single pixel, the number of bits per pixel is often higher than the number of bits per channel. The number of channels will depend on whether the image is grayscale or color and whether it has an alpha channel. PNG allows the following combinations of channels, called the color type [8].

The authentication method deals with binary like grayscale document images instead of pure binary ones and simultaneously solves the problem of image tampering detection, data loss and visual quality keeping. Here a method for the authentication of document images with an additional self-repair capability for fixing tampered image data is used. The input cover image is assumed to be a binary-like grayscale image with two major gray values .The cover image is transformed into a stego-image in the Portable Network Graphics (PNG) format with an additional alpha channel for transmission on networks or archiving in databases. The stego-image, when received or retrieved, may be verified by the proposed method for its authenticity. Integrity modifications of the stego-image can be detected by the method at the block level and repaired at the pixel level. In case the alpha channel is totally removed from the stego-image, the entire resulting image is regarded as inauthentic, meaning that the fidelity check of the image fails.

Grayscale is a range of monochromatic shades from black to white. Therefore, a grayscale image contains only shades of gray and no color. While digital images can be saved as grayscale or black and white images, even color images contain grayscale information. This is because each pixel has a luminance value, regardless of its color. Luminance can also be described as brightness or intensity, which can be measured on a scale from black (zero intensity) to white (full intensity). Authentication is

an added security measure used to prove that someone or something is who or what they say they are before access is granted to personal or confidential information. For secret sharing of information Shamir secret scheme is used to identifying the tampered block. Data repairing is applied to each tampered block by a Reverse Shamir scheme is used.

II LITERATURE SURVEY

There are several image authentication methods. In Multipurpose Watermarking for Image Authentication and Protection the application of hiding multimedia information includes ownership protection and content authentication. It is based on two methods, they are robust watermarking and Fragile watermarking Digital signature based on image authentication method. This preserves high robustness for copyright protection [1]. But the detection results are very unstable. The next method is Data Hiding in Binary Image for Authentication and Annotation, here large number of digital image used for business purpose. Cryptographic authentication approach is used for authentication. There are two methods to embed data, they are flappable and shuffling. The disadvantage is it affects visual quality.

In Secret Image Sharing with Stegnography and Authentication, Stegnography is data hiding technique that provides security protection for digital image data [2]. It is used to handle full color images and the quality of the recovery result is nearly lossless. Here the computational cost is high.

In Binary Image Authentication with Tampering Localization by Embedding Cryptographic Signature and Block Identifier, two-layer blind binary image authentication scheme is introduced. The image is partitioned into multiple macro blocks. Block identifier is defined in each block. The macro blocks self detect itself to identify the tampered locations. This occupies large memory. In Pattern-based Data Hiding for Binary Image Authentication by Connectivity –preserving, the connectivity preserving criterion is to assess the flippability of a pixel. It is determined by imposing three transition criteria in 3×3 moving window centered at the pixel. Uneven embeddability of the host image is handled by embedding the watermark only in those embeddable blocks. Here the complexity is high [3]. In Compression Tolerant Image Authentication method, Image authentication scheme based on the extraction of feature points from the image. The set of feature points from a given image is encrypted using public key encryption to generate the digital signature of the image.

The process of image authentication consists of two parts; they are generation of the digital signature and verification of the digital signature. But this method is not secure.

III SYSTEM MODEL

In the image authentication and data repairing method, a PNG image is created from a binary grayscale document image with an alpha channel plane. A PNG image is created from the grayscale document image and alpha channel plane. Here the source document image is converted into stego image. During this conversion, by adding source document image with alpha channel plane, a document image with alpha channel plane is obtained [4]. In addition the source document image is binarized. During binarization, data for authentication and repairing are computed, which is taken as an input for secret sharing scheme to generate secret shares. These share values are mapped with alpha channel value to produce an imperceptibility effect. For security and data protection the mapped secret shares are embedded to alpha channel to produce a stego image.

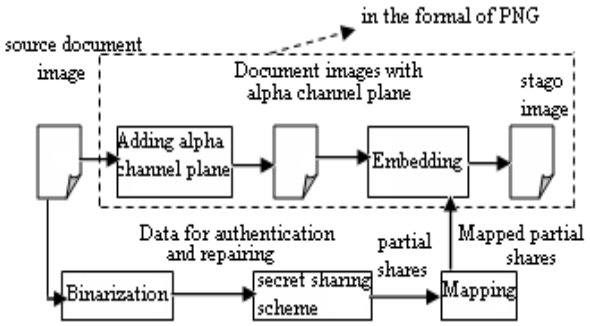

Fig 1: Creating a PNG image from a grayscale document image and alpha channel

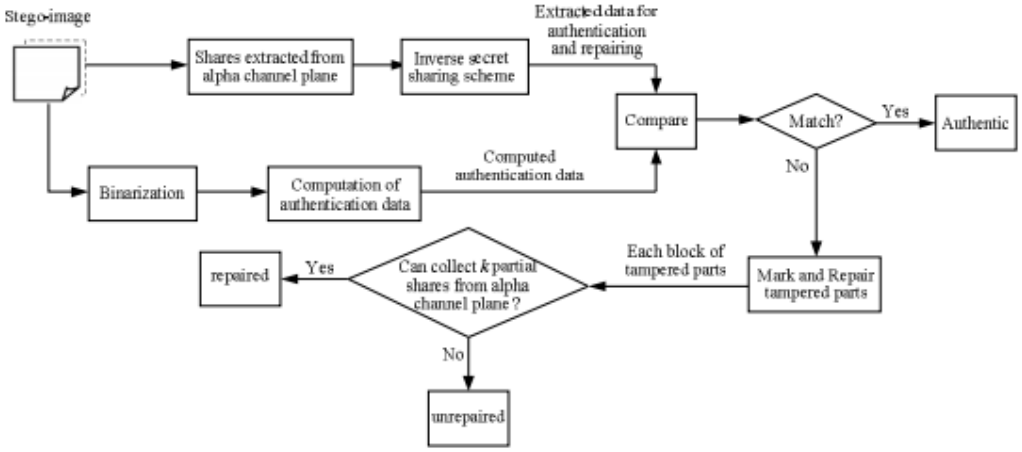

Fig 2: Authentication process including verification and self repairing of a stego image in PNG format.

III. SECRET SHARING USING SHAMIR METHOD

Shamir method for secret sharing performed by following algorithms.

1. Algorithm for Threshold secret sharing.

Input: secret d in the form of an integer [5], number of n participants and threshold $k \leq n$.
Output: n shares in the form of integers for then participants.

i. Select a prime number p that is larger than d randomly.
ii. Select k -1 integer values c1, c2,…, c_{k-1} within the Range of 0 through p -1.
iii. Select n distinct real values $x_1, x_2, ..., x_n$.
iv. Use the following (k-1)-degree polynomial to compute n function values F(xi), called partial shares for i= 1, 2, …, n:

$$F(x_i) = (d = c_1 x_i + c_2 x_i^2 + ... + c_{k-1} x_i^{k-1}) \bmod p, \quad (1)$$

v. Send the two-tuple ($(x_i, F(x_i))$ as a share to the i th participant where i= 1, 2, .., n.

There are k coefficients denoted d and c through, k shares are collected from n participants to form k equation to recover secret d.

2. Algorithm for Secret recovery.

Input: k shares collected from n participants and the prime number p with both k and p used in threshold secret sharing algorithm.
Output: Secret d hidden in the shares and coefficients c_i used in threshold secret sharing algorithm.
where i = 1, 2, …, k-1.

i. use the k shares $(x_1, F(x_1)), (x_2, F(x_2)), .., (x_k, F(x_k))$, to set up the following equations:

$$F(x_j) = (d + c_1 x_j + c_2 x_j^2 + ... + c_{k-1} x_j^{k-1}) \bmod p, \quad (2)$$

Where j=1,2,…,k.

ii. Solve the above equations by Lagrange's interpolation to obtain d as follows [6]

$$d = (-1)^{k-1}[F(x_1)\frac{x_2 x_3 ... x_k}{(x_1 - x_2)(x_1 - x_3)...(x_1 - x_k)} +$$
$$F(x_2)\frac{x_1 x_3 ... x_k}{(x_2 - x_1)(x_2 - x_3)...(x_2 - x_k)} + ...$$
$$... + F(x_k)\frac{x_1 x_2 ... x_{k-1}}{(x_k - x_1)(x_k - x_2)...(x_k - x_{k-1})}] \bmod p$$

ii. Compute $c_1$through $c_{k-1}$by expanding the following equality and comparing the result with (2) in step i while regarding the variable x in the equality below to be x_j in (2):

$$F(x) = [F(x_1)\frac{(x - x_2)(x - x_3)...(x - x_k)}{(x_1 - x_2)(x_1 - x_3)...(x_1 - x_k)} +$$
$$F(x_2)\frac{(x - x_1)(x - x_3)...(x - x_k)}{(x_2 - x_1)(x_2 - x_3)...(x_2 - x_k)} + ...$$
$$... + F(x_k)\frac{(x - x_1)(x - x_2)...(x - x_{k-1})}{(x_k - x_1)(x_k - x_2)...(x_k - x_{k-1})}] \bmod p$$

iii. In the above algorithm is included additionally for the purpose of computing the values of the parameters c_i in the proposed method. In other applications, if only the secret value d need be recovered, this step may be eliminated.

IV. PROCESS OF IMAGE AUTHENTICATION AND DATA REPAIRING

Image authentication and data repairing based on the following algorithms.

1. Algorithm for generation of stego-image.

It performed by two stages. They are

- Generation of authentication signals.
- Creation and embedding of shares.

Input: a grayscale document image with two gray values and secret key k.
Output: stego image in the PNG format.
Stage I: Generate authentication signals.

To generate the authentication signals the following
Steps are performed.

1. *Binarization:* In binarization movement threshold technique is used. There are two gray values in this technique and the average is calculated for the two gray values [7]. If the gray value greater than the threshold value then it is called as foreground and if the gray value less than the threshold value then it is called as background.
2. *Transform the image into the PNG format:* Transform image into a PNG image with an alpha channel plane by creating a new image layer with 100% opacity and no color.
3. *Beginning of loop:* Block of 2×3 to get 6 pixels p_1, p_2,…,p_6.
4. *Creation of authentication signals:* Take two bit authentication signals. One is a_1 and another is a_2.

$$a_1 = p_1 \oplus p_2 \oplus p_3 \text{ and}$$
$$a_2 = p_4 \oplus p_5 \oplus p_6$$

Where $\oplus$ denotes the exclusive- OR operation.

$$s = a_1 \times a_2$$

Stage II: create and embedded of partial shares.

1. *Creation of secret data sharing:* Due to security 8 bit string are divided into 4 bit and mentioned as m1, m2 respectively
2. *Partial share generation: T*o generate six partial shares q_1 through q_6 using the following equations:

$$q_i = F(x_i) = (d + c_1 x_i) \bmod p, \quad (3)$$

Where i=1,2,…6.

$F(x_i)$ = partialshares
d = secretdata
c = coefficients
x_i = distinct real values.

3. *Mapping of the partial shares:* Due to mapping just add 238 to six partial shares q_1 through q_6 [8].The total transparency range is 238 to 254. Embedding of two partial shares in the current block and embedding remaining pixels at random pixels.
4. *End of looping:* If there are any unprocessed block then it will return back to step 3.

2. Algorithm for authentication of stego image in the PNG format

Input: stego image with two gray value and secret key k.
Output: image with tampered block and their data repaired.
Step 1: Two representative gray values are extracted.
Step 2: Verify the stego image

- Extract the hidden signal.
- Compute the signal from the current block content.
- Match both the hidden signal and compute signal, and then marking the tampered block.

Step 3: Original image content (self repairing)

- Extract the remaining partial shares.
- Repair the tampered block, and then it automatically self repairing.

V. RESULTS AND DISCUSSION

For editing the image two common operations are used. They are superimposing and painting. The superimposing operation destroys the content of the alpha channel values. It replaces all the original alpha channel values at the attacked part with the new values of 255.The Authentication result of the document image of a signed paper attacked by superimposing operation performed by the below fig.

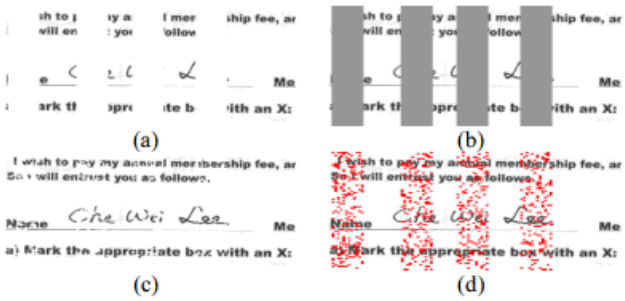

Fig.3. Authentication result of document image by superimposing operation. (a)Tampered image yielded by the superimposing operation. (b) Tampered blocks detected and marked as gray. (c) Data repair result. (d) Data repair result with red dots indicating unrepaired tampered blocks.

If the tampered areas grows the data repair results becomes worse. When the tampered area becomes larger, fewer partial shares still survive for data repairing. The superimposing operation is performed in the tampered image. After performing the superimposing operation, the tampered blocks are detected and marked as gray in (b). The results after data repairing are shown in (c). In data repair results, the red dots indicate the unrepaired tampered blocks (d).

Authentication result of a document image in the form of PNG, attacked by added noises are shown in fig (4). The original cover image is generated in (a). The data embedded into the stego image is shown in (b). The tampered image with added noises is shown in (c). The tampered blocks are detected and marked as gray in (d). Data repair result (e). Red dots indicate unrepaired tampered blocks in (f).

V.CONCLUSION

A secret sharing method for a binary grayscale document images has been proposed. With this approach if the document has been illicitly tampered, it has ability to identify the tampered block and has the self-repair capability by using Shamir method. The undesired opaque effect visible in the stego-image coming from embedding the partial shares has been eliminated by mapping the share values into a small range of alpha channel values near their maximum transparency value of 255.

REFERENCES

[1]. Shamir, Adi (1979), "How to share a secret", Communications of the ACM
[2]. Liu, C. L. (1968), Introduction to Combinatorial Mathematics, New York: McGraw-Hill.
[3]. Dawson, E.; Donovan, D. (1994), "The breadth of Shamir's secret-sharing scheme", Computers & Security
[4]. Knuth, D. E. (1997), The Art of Computer Programming, II: Seminumerical Algorithms (3rd ed.), Addison-Wesley, p.
[5]. Lu.Z M, Xu.D.G, and Sun.S.H, "Multipurpose image watermarking algorithm based on multistage vector quantization," IEEE Trans. Image Process., vol. 14, no. 6, pp. 822–831, Jun. 2005.
[6]. Wu.D.C, Tsai.W.H, 1998. "Data hiding in images via multiple-based number conversion and lossy compression".IEEE Transaction on Consumer Electronics 44(4), 1406-1412.
[7]. Wu.M and Liu.B, "Data hiding in binary images for authentication and annotation," IEEE Trans. Multimedia, vol. 6, no. 4, pp. 528–538, Aug. 2004.
[8]. Yang.H and Kot.A.C., "Binary image authentication with tampering localization by embedding cryptographic signature and block identifier," IEEE Signal Process. Lett., vol. 13, no. 12, pp. 741–744, Dec. 2006.

*****