International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

SECRET SHARING BASED METHOD FOR AUTHENTICATION OF GRAY SCALE DOCUMENT IMAGES USING PNG IMAGE WITH DATA REPAIR CAPABILITY

Mrs. Sonal Kokate

Final year student, M.E. in "Electronics and Telecommunication Engineering"

Saraswati College of Engineering, Kharghar, Navi Mumbai (India)

ABSTRACT

A new blind authentication method based on the secret sharing technique with a data repair capability for grayscale document images via the use of the Portable Network Graphics (PNG) image is proposed. An authentication signal is generated for each block of a grayscale document image, which, together with the binarized block content, is transformed into several shares using the Shamir secret sharing scheme. The involved parameters are carefully chosen so that as many shares as possible are generated and embedded into an alpha channel plane. The alpha channel plane is then combined with the original grayscale image to form a PNG image. During the embedding process, the computed share values are mapped into a range of alpha channel values near their maximum value of 255 to yield a transparent stego-image with a disguise effect. In the process of image authentication, an image block is marked as tampered if the authentication signal computed from the current block content does not match that extracted from the shares embedded in the alpha channel plane. Data repairing is then applied to each tampered block by a reverse Shamir scheme after collecting two shares from unmarked blocks. Measures for protecting the security of the data hidden in the alpha channel are also proposed. Good experimental results prove the effectiveness of the proposed method for real applications.

I. INTRODUCTION

Digital images are used to preserve important information. But providing integrity and authentication to these images is a challenging task. In this era with the use of fast advanced technologies it is easy to modify the contents of these digital images. It is important to make an effective method to solve image authentication problem [1] [2], particularly for document images such as important certificates. Scanned checks, art drawings, signed documents, circuit diagrams, design drafts, testaments etc. In the case of binary document images, it is difficult to authenticate because of its simple binary nature that lead to perceptible changes after authentication signal are embedded in the image pixel. So in this paper we are performing authentication of grayscale document images. Gray scale images are looking like a binary image, because of this reason it is called as binary like gray scale image. Grayscale images overcome the visual quality problem of binary images. In this paper we are proposing a new method for authentication of document images with a supplementary self repairing capability for fixing tampered image data. The input cover image is taken as binary like grayscale image. After applying the proposed method, the input cover image is transformed into PNG format, with scrambled form in a

supplementary alpha channel for transmission on networks or archiving in database. By using proposed method the stego-image retrieved or received may be verified for its authenticity. Integrity modification 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 destroyed from the stego-image, the entire resulting image is regarded as inauthentic, that means fidelity check of the image failed. The proposed method is based on (t, n) threshold secret sharing Scheme proposed by Shamir [3] and also with encryption based on chaotic logistic map [4]. By secret sharing scheme the secret message is transformed into n shares, and when t of the n shares is collected the secret message can be recovered without data loss. Using logistic map we generate a random key, by this key the PNG image formed with alpha channel plane is scrambled, and made ready for transmission. For highly confidential document image transmission this authentication method can be used, this method provides two layers of security to the document by keeping shares in the alpha channel and encrypting the PNG image. Secret sharing helps the reconstruction of tampered image content and encryption scrambles the image thereby hiding the data's of the document image

Gray scale document image + Alpha channel plane= PNG image.

Several methods for image authentication have been proposed in the past. Chih - Hsuan Tzeng and Wen - Hsiang Tsai [5] proposed, Authentication with embedding special codes. Embedding randomly-generated codes, into the blocks of a given cover image, producing a stego-image. Authentication is achieved by verifying the codes in the blocks of a given stego image. H. Yang & A.C. Kot [6] proposed a method for Authentication with cryptographic signature and block identifier provides a two layer image authentication in which the first layer provides the overall authentication by hiding the cryptographic signature (CS) of the image and the localization of the tampering is obtained from the second layer by embedding the block identifier (BI) in the "qualified" or "self-detecting" macro-blocks (MBs). M Wu and B. Liu [7] proposed Authentication by manipulating flippable pixels. In this method images are partitioned into blocks and significant amount of data will be embedded into each block maintaining a block based relationship and without introducing noticeable artifacts. Che-Wei Lee and Wen-Hsiang Tsai [8] proposed a secret sharing based method for authentication of grayscale document images. This method provides data repairing capability via the use of PNG image. Niladri B. Puhan, Anthony T. S. Ho [9] proposed authentication using Perceptual Modeling, estimates the distortion resulting from flipping of a pixel by finding the curvature-weighted distance difference (CWDD) measure between original and watermarked contour segments.

II. GRAY SCALE IMAGE AUTHENTICATION

Digital information is a form of preserving data for which authentication is necessary to overcome the tampering attacks. In multimedia applications, it is necessary to authenticate the source image which might be subjected to tampering. So, for such content authentication watermarking technique is used. It is one among the emerging fields that are used for content authentication. As, the content authentication is being the hottest topics now a days, it is necessary to assure that the delivering of image to somewhere is delivered as it is. However, with the fast advance of digital technologies it is easy to make the modifications to the images. Thus integrity of image becomes a serious concern. To solve those image authentication problem particularly digitized documents, digital signatures, tables, texts, etc., whose security must be protected. In this paper we are performing image

authentication of grayscale images. Grayscale image has two gray values i.e. foreground and background. Grayscale images look like binary ones. So, we can call a grayscale image as binary like grayscale image. Binary image consists of two colors black and white. Using binary images can cause some problems. As the binary images are simple in nature many unpleasant strokes can encounters. So using grayscale images can solve the problem of visual quality which the binary one cannot do. Many conventional methods have been proposed for authentication of grayscale images. In our proposed method, the image is first watermarked and divided into shares using Shamir secret sharing scheme. Later, those shares are reconstructed using inverse Shamir secret sharing scheme and watermark has been extracted if the image is authentic. Data loss during transmission is marked as gray blocks.

Fig 1. Gray scale cheque image

III. IMAGE AUTHENTICATION TECHNIQUES

Before presenting and discussing various methods, we start by defining the general requirements that are essential for any authentication system. These requirements are:

Sensitivity: The authentication system must be able to detect any content modification or manipulation. For strict authentication algorithms, detection of any manipulation is required and not only content modification.

Robustness: Also called tolerance. The authentication system must tolerate content preserving manipulations. This property is valid just for algorithms that provide a selective authentication service.

Localization: The authentication system must be able to locate the image regions that have been altered.

Recovery: The authentication system must be able to partially or completely restore the image regions that were tampered.

Security: The authentication system must have the capacity to protect the authentication data against any falsification attempts.

Portability: The authentication system must be able to carry the signature with the protected image during any transmission, storage or processing operation.

Complexity: The authentication system must use real-time implemented algorithms that are neither complex nor slow.

IV. STRICT IMAGE AUTHENTICATION

Strict image authentication methods do not tolerate any changes in the image data. These methods can be further separated in two groups according to the techniques that are used: methods based on conventional cryptography and methods that use fragile watermarking.

V. METHODS BASED ON CONVENTIONAL CRYPTOGRAPHY

Image authentication methods based on cryptography compute a message authentication code (MAC) from images using a hash function [46, 85, 109, 112, 122, 123, 153]. The resulting hash (h) is further encrypted with a secret private key S of the sender and then appended to the image. For a more secure exchange of data between subjects, the hash can be encrypted using public key K1 of the recipient [141]. The verification process is depicted in Fig. 2b. The receiver computes the hash from the received image. The hash that was appended to the received image is extracted and decrypted using private key K1. The extracted hash and the calculated one are then compared. Techniques that are based on the hash computing of image lines and columns are known as line–column hash functions [22]. Separate hashes are obtained for each line and each column of an image. These hashes are stored, and compared afterwards with those obtained for each line and each column of the image to be tested. If any change in the hashes is found, the image is declared manipulated otherwise it is declared authentic.

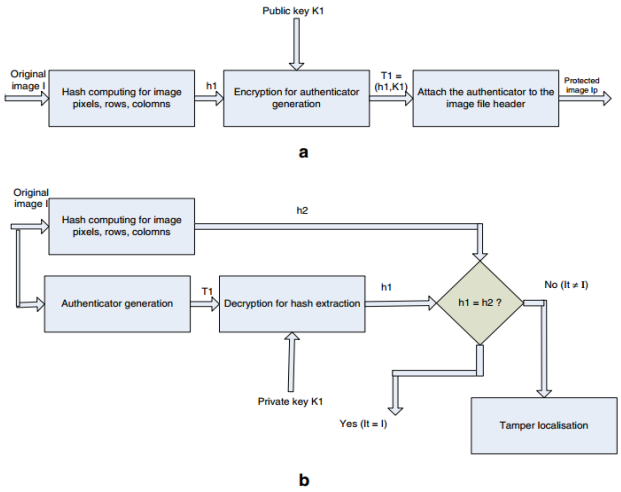

Fig 2. Strict authentication system by conventional cryptography:
(a) Generation of authenticator; (b) verification of authenticity

Distortions localization can be achieved by identifying lines and columns for which the hashes are different. Unfortunately, the localization of changes can be easily lost if more than one region of the image was corrupted. This is called the ambiguity problem of the line–column hash function. To solve this problem, another approach has been proposed by Wolfgang and Delp [145]. This technique consists in obtaining the hash of image blocks, separately. If an image is to be tested, the user calculates the hashes for each block using the same block size, and compares the results with the hashes from the original image to decide whether the image is authentic. Blocks for which hashes are different enable tamper localization. The computation of hashes for each block separately had increased the localization capabilities. However, these techniques are not able to restore image regions that were tampered. Conventional cryptography was developed to solve the problem of message authentication, and had a great success since its appearance. Algorithms based on conventional cryptography show satisfying results for strict image authentication with high tamper detection. Localization performances are not very good but may be acceptable for some applications. Hash functions are very sensitive to any small change in the image pixels or even in the binary image data. In consequence the image is classified as manipulated, when just only one bit of this image is changed; this is very severe for most of applications.

Recently, many teams of researchers have published works where they try to use hash functions while introducing some errors into the images in order to achieve methods that tolerate some desired manipulations. The type of introduced errors automatically determines the kind of manipulations tolerated such as compression [67], and histogram equalization [154]. These methods however, cannot tolerate a combination of several allowed manipulations in the same image. Moreover, they are vulnerable to attacks against the hash functions [54, 111].

VI. FRAGILE WATERMARKING

Watermarking consists in calculating a watermark, hiding it in the image, and then extracting it when it is necessary. In this paper, we choose fragility as the basic criterion for algorithms classification. Fragile watermarking belongs to the strict authentication class, while semi fragile watermarking to the selective authentication class. Some authors define reversible watermarking, also called erasable or invertible [34], as a subgroup of fragile watermarking. The idea behind reversible watermarks is to reconstruct the exact original image when the image is declared as authentic. Thus, it reconstructs the information that was lost during watermarking. Usually, it is a lossless compressed version of the space where the watermark was embedded. This lossless compressed version is thereafter concatenated with the watermark, inserted within the image and extracted for reconstruction purposes only when the image is declared authentic. However, in most image watermarking algorithms, modifications caused by embedding functions are really insignificant. Therefore, reversible watermarks are desired only for specific applications such as for high sensitive images. Moreover, their main goal is to eliminate the distortion artifacts caused by the embedding functions. Interested readers could consult references [34, 40, 64, 133] on this subject. Throughout this paper we compare the restoration capabilities of each algorithm, which is somehow different from reversibility. Restoration is the ability of an algorithm to restore the damaged data. When an algorithm detects and localizes a region with some undesired manipulations, we wish that this algorithm could restore the original data. This requirement is desirable for wide range of applications.

Additionally, we classify an algorithm as symmetric or asymmetric. This mainly depends on the security key model. The asymmetric algorithms are generally more secure as they provide different private and public keys for encoding and decoding. However, they are much slower than symmetric algorithms [112, 123, 153]. The basic idea behind fragile watermarking techniques is to generate a watermark and to insert it in the image to be protected in such a way that any modification made to the image is also reflected in the inserted watermark. Simply verifying the presence of the inserted watermark allows the image authenticity verification and eventually localization of tampered regions. This type of watermarking does not tolerate any image distortion. The image is considered authentic if and only if all its pixels remain unchanged. The first algorithms of fragile watermarking were based on watermark generation from image information only [142] as shown in Fig. 3a. The watermark is computed from a set of image pixels. The computation of the watermark differs between the various authentication methods. The set of pixels may be chosen with the help of a secret key K1. The computed watermark may be encrypted with a key K3. It is then inserted in the least significant bits of another set of pixels. In order to increase the algorithm security, the set of pixels where the watermark is embedded may be determined with another secret key K2. Similarly, the verification schema is shown in Fig. 3b. The secret keys must be known to the receiver, as well. The receiver uses the same key K2 to determine the set of pixels where

the watermark is dissimulated in order to extract it. Also, the receiver uses the same algorithms to calculate the watermark from the received image and then compares the calculated watermark with the dissimulated one to decide whether the image is authentic or not. One of the first techniques that used image authentication by fragile watermarking was proposed by Walton [142]; it used only image information to generate the watermark. This technique is based on the insertion, in the least significant bits (LSB), the checksum calculated with the grey level of the seven most significant bits of pseudo-randomly selected pixels. This method was able to detect and localize manipulations but with no restoration capabilities. Various algorithms were proposed for the realization of this technique [5, 16, 31]. The algorithm that attracted most attention was proposed by Fridrich [31]. A sufficient large number N is chosen to be used for the calculation of the checksums. The size of N directly impacts the probability of detecting manipulations. The original image is first subdivided into blocks of size 8×8; in each block, a pseudo-random walk through its 64 pixels is generated. The checksum S is calculated by the following equation:

$$S = \sum_{j=1}^{64} (a_i . g(p_j)) \bmod N$$

Where g (pj) is the grey level of the pixel pj obtained with the seven most significant bits only and aj is a pseudo random sequence of 64 integers. Then the binary format of S is encoded using a coding algorithm, and inserted into the LSB of the block pixels. To increase the security of the system, the coefficients aj and the pseudo random walk can be dependent on the blocks by using secret keys. The procedure of checking an image authenticity consists in extracting the inserted checksums, recalculating the checksums in a similar way, and finally comparing the two checksums to decide about the image authenticity. This method has the advantage of being very simple and fast. Moreover, it detects and localizes tampering. However, the algorithm cannot detect the manipulation if blocks from the same position of two different images, which are protected with the same key were exchanged. To avoid this type of attack, several improvements were made to this method by extracting more robust bits [16]. The method is not able to restore the damaged data.

In a more general schema, the watermark that is inserted in the image to be authenticated is obtained by combining information from the image with a predefined logo as depicted in Fig and b. A secret key K1 can be used to extract specific image information from the image. In order to generate the watermark, the extracted image information is combined with a binary logo by using another secret key K2. The computed watermark may be encrypted with a key K4. It is then inserted in the least significant bits of a set of pixels that may be determined with a secret key K3. The secret keys must be known to the receiver, as well. The receiver uses the appropriate key to determine the set of pixels where the watermark was dissimulated in order to extract it. Also, the receiver uses the same algorithms to calculate the watermark from the received image and then compares the computed watermark with the dissimulated one to decide whether the image is authentic or not.

International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

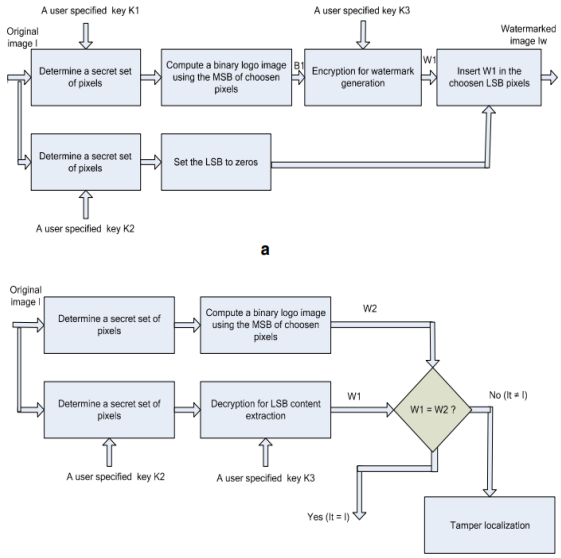

Fig 3. Strict authentication system by fragile watermarking using image information; (a) Generation of authenticator; (b) verification of authenticity

Strict image authentication is appropriate for many applications. For example, a modification of just one or two pixels in some medical or military images can dramatically change the decisions of doctors or war strategists, respectively, and can result in costly. Most existing image applications use image processing operations that preserve the content in order to save memory space and bandwidth or to enhance image quality: compression, filtering, geometrical transformations and image enhancement techniques. Therefore, some tolerant image authentication algorithms are needed.

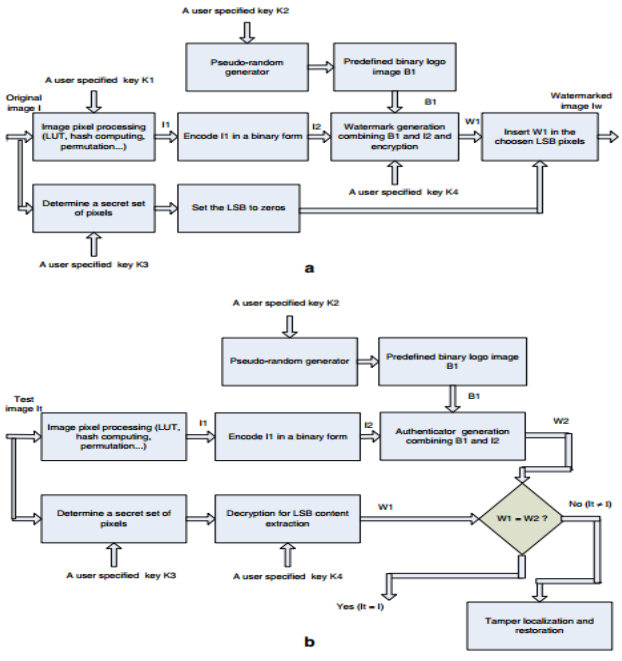

Fig 4. Strict authentication system by fragile watermarking where the watermark is obtained from the image and a logo: (a) Generation of authenticator; (b) verification of authenticity

6.1 Content-Based Image Authentication

Or Selective Authentication

We defined a content modification as an object appearance or disappearance, a modification to an object position, or changes to texture, color or edges. We have also listed the image processing operations that preserve the image content. Thus, lot of applications that base their decisions on images need authentication methods that can tolerate content preserving manipulations while at the same time detect any manipulation that change the image content. This leads to new watermarking methods known as semi fragile watermarking, and to new approaches known as content-based signatures. In this section we will present and compare semi-fragile techniques and content-based signatures approaches that provide selective image authentication service.

VII. SEMI-FRAGILE WATERMARKING

Robust watermarking is designed to resist all attempts to destroy the watermark. Its main application includes the intellectual property protection and owner identification. The robustness of the embedded watermark is crucial to resist any intentional and even unintentional manipulation. The goal of these techniques is not the verification of the image authenticity, but rather the verification of their origins. Conversely, fragile watermarking is designed to easily destroy the embedded watermark following any kind of manipulations of the protected image. It is useful for applications where strict authentication is needed, that is where the main objective is to determine whether the image has been modified or not, with the possibility of locating and reconstructing image regions that have been tampered. On the other hand, semi-fragile watermarking [29, 30, 37] combines characteristics of fragile and robust watermarking techniques. Basically, the idea of semi-fragile watermarking is to insert a watermark in the original image in such a way that the protected image can undergo some specific image processing operations while it is still possible to detect malevolent alterations and to locate and restore image regions that have been altered. For image authentication purposes watermarking algorithms should be invisible. Visible watermarking algorithms are applied for on-line content distribution, transaction tracking or owner identification. The procedures of generating a watermark and embedding it into the image can be dependent on a private or public, symmetric or asymmetric, key system in order to increase the overall system security. This is a trade-off between security and computational time [46, 109, 112, 123, 153]. Generally, symmetric key systems are less secure than asymmetric ones, and asymmetric key systems consume more resources and consequently need more computing time.

The general schema for semi-fragile watermarking methods is shown in Fig. The watermark is computed from the result of an image-processing algorithm applied on the image pixels. The computation of the watermark varies as different image processing algorithms can be used. A secret key K1 can be used to extract specific information from the image. In order to generate the watermark, the extracted image information is often combined with a binary logo using another secret key K2. Usually, the generated watermark is then inserted in a set of frequency coefficients that are in the middle range. The set of coefficients where the watermark is inserted may be determined with the help of a secret key K3. The computed watermark may be encrypted with a key K4. Similarly, the general verification schema is shown in Fig. The secret keys must be known to the receiver, as well. The receiver uses the same key to determine the set of pixels where the watermark is dissimulated in order to extract it. Also, the receiver uses the same algorithms to compute the watermark from the received image and

then compares the computed watermark with the dissimulated one to decide whether the image is authentic or not.

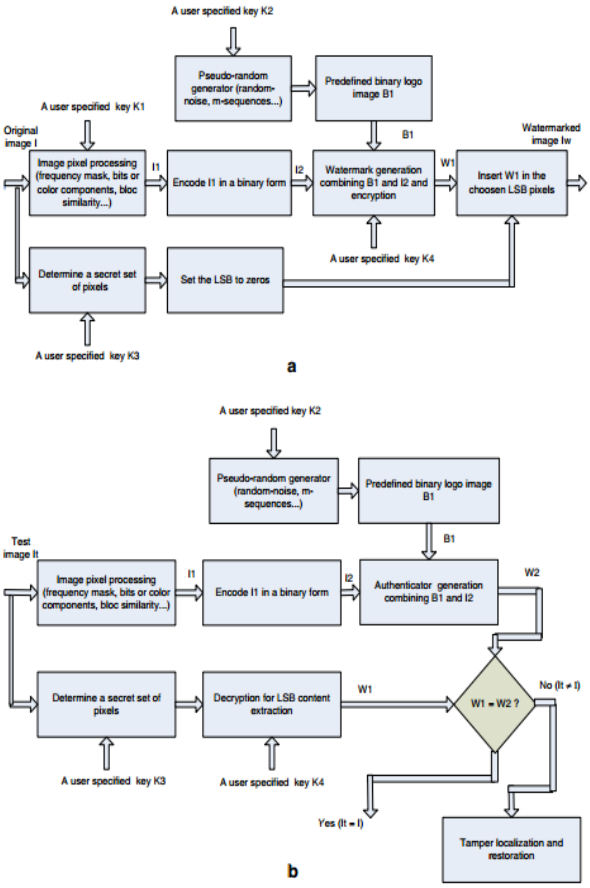

Fig 5. The general schema for semi-fragile watermarking methods

7.1 Image Authentication by Digital Signatures Based on the Image Content

Most recent investigations in the domain of image authentication were concentrated on digital signatures applied to the image content; these approaches offer high performance and promise additional breakthroughs in the near future. Generic diagram of an authentication system based on image content Image authentication systems that use a digital signature based on the semantic content of images could be described in a generic diagram (Fig.5). Such systems consist in (1) extracting specific high level characteristics from the original image; (2) applying a hash function to these characteristics in order to reduce their size; (3) digitally signing the hash value using an existing digital signature algorithm such as a private or public key system to increase the overall security; (4) attaching the signature to the original image or inserting it in the image using techniques for data dissimulation. Likewise, the verifying procedure of an image authenticity consists in (1) generating the image signature using the same algorithm; (2) extracting the attached or dissimulated signature; (3) comparing these two signatures using a comparison algorithm to decide whether the image was altered or not; (4) determining the image regions that were manipulated. When the image is declared as not authentic, information from the original signature could be used to partially or even completely restore the regions that were corrupted. Several parameters directly affect the performance of an image authentication system based on image content signature. These parameters include the choice of the appropriate characteristics, the choice of the hash function and the digital signature

algorithm, the choice of the data dissimulation method in images as well as the choice of the algorithm that compares the signatures to decide about the authenticity of an image. Among these parameters, the image features that represent the image content and the data dissimulation method mostly affect the performance of image authentication methods. In fact, sensitivity, robustness, recovery, portability, safety and complexity are directly affected by the choice of the characteristics that are used to generate a content-based signature; they are affected as well by the choice of the data dissimulation method. The hash function and the digital signature algorithms are almost the same for all techniques. The algorithm used to compare the signatures directly depends on the selected characteristics and the dissimulation method. Therefore, we will use these two parameters, the choice of the appropriate characteristics and the data dissimulation algorithm, to classify and compare existing image authentication systems based on image content signatures.

VIII. ADVANTAGES AND LIMITATIONS OF VARIOUS METHODS

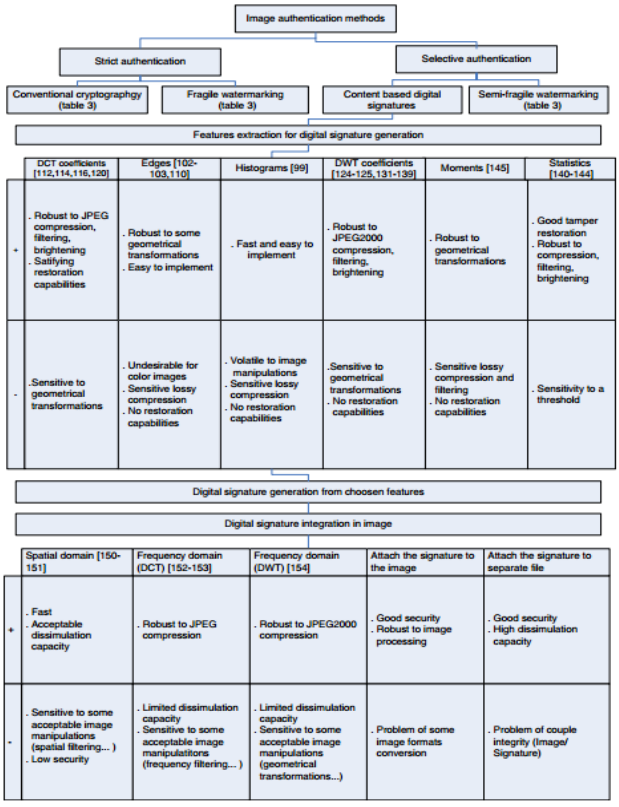

Fig 6. Classification of image authentication methods; plus sign indicates advantages; minus sign indicates disadvantages

Table presents a summarized comparison of image authentication methods discussed in this paper: methods based on conventional cryptography, fragile watermarking, and semi-fragile watermarking and on image content signatures. For each group of methods we have shown the type of the authentication tag, the dependency of this authentication tag on the image, the type of the authentication service provided, that is: strict or content-based (selective) image authentication service, the localization capacity of the altered regions, as well as the possibility of restoration of image corrupted regions. Algorithms are also grouped according to the authentication tag that is used, and references are included. It can be noticed that one principal property of an image authentication

system, the detection of malevolent manipulations, is not included in this table for the following reason: All described methods can detect malevolent manipulations. Moreover, the robustness against content preserving manipulations is not offered by the first two categories since they provide a strict authentication services and do not tolerate any modification to the original image. According to this summary table, algorithms performances are very similar. In fact, most of algorithms offer acceptable detection and localization of image manipulations while restoration performances still need to be improved. For strict authentication applications, where no modification to the original image is allowed, fragile watermarking algorithms perform better than algorithms based on conventional cryptography. Fragile watermarking algorithms offer high detection and localization capabilities. Moreover, some of them could provide an acceptable restoration level of damaged regions. On the other hand, selective authentication methods tolerate some desired manipulations while detecting any malevolent operations. Semi-fragile algorithms show good results for detecting and locating any malevolent manipulations while providing acceptable reconstruction performances. Unfortunately, their tolerance against desired manipulations includes mainly compression, noise addition and rotation by small angles, whereas, many of the desired manipulations need to be tolerated in practice. Since algorithms based on digital signature show more interesting results, we present them and compare their performances along with references in Fig. Figure presents a classification of image authentication methods with a detailed comparison of signature content-based methods. The comparison is made according to two important properties: the domain from which features are extracted to provide a content-based signature and the domain used to dissimulate or attach this signature. Moreover, for the sake of simplicity, only the most important weakness and strength for each group are highlighted. Every image-extracted feature used to generate the image signature has its weakness and force. The comparison of these features, their weaknesses and forces, help choosing the right method for a specific application. For example, if an application needs to tolerate compression with JPEG or JPEG2000 standard, the DCT domain or DWT domain, respectively, are best suited to generate the signature. If geometrical transformations need to be tolerated, the use of moments would be the best choice. If restoring the damaged data is important, statistical features could help well. Moreover, they are able to survive lossy image compression and a predefined set of content preserving manipulations (filtering, brightening...). On the other hand, using edges for content-based signature is undesirable for color images since one may change colors without affecting edges. This could result in an error where an image is declared authentic while some undesirable changes were introduced to it. Dissimulating signatures or attaching them to the image depends on the application and user requirements. A big dissimulation capacity and a high security can be achieved by attaching the signature to the image or to a separate file. However, the latter solution suffers from the problem of ensuring the couple image-signature integrity.

IX. PROBLEM DEFINITIONS

The image authentication problem is difficult for a binary document image because of its simple binary nature that leads to perceptible changes after authentication signals are embedded in the image pixels. Such changes will arouse possible suspicions from attackers. A good solution to such binary image authentication should thus take into account not only the security issue of preventing image tampering but also the necessity of keeping the visual quality of the resulting image. In this paper, we propose an authentication method that deals with binary-

like grayscale document images instead of pure binary ones and simultaneously solves the problems of image tampering detection and visual quality keeping.

X. ADVANTAGES OF THE PROPOSED METHOD

The proposed method has several other merits, which are described in the following:

1) Providing pixel-level repairs of tampered image parts-As long as two untampered partial shares can be collected, a tampered block can be repaired at the pixel level by the proposed method. This yields a better repair effect for texts in images because text characters or letters are smaller in size with many curved strokes and need finer pixel-level repairs when tampered with.

2) Having higher possibility to survive image content attacks- By skillfully combining the Shamir scheme, the authentication signal generation, and the random embedding of multiple shares, the proposed method can survive malicious attacks of common content modifications, such as superimposition, painting, etc., as will be demonstrated by experimental results subsequently described.

3) Making use of a new type of image channel for data hiding- Different from common types of images, a PNG image has the extra alpha channel plane that is normally used to produce transparency to the image. It is differently utilized by the proposed method for the first time as a carrier with a large space for hiding share data. As a comparison, many other methods use LSBs as the carriers of hidden data.

4) Causing no distortion to the input image- Conventional image authentication methods that usually embed authentication signals into the cover image itself will unavoidably cause destruction to the image content to a certain extent. Different from such methods, the proposed method utilizes the pixels' values of the alpha channel for the purpose of image authentication and data repairing, leaving the original image (i.e., the grayscale channel) untouched and thus causing no distortion to it. The alpha channel plane may be removed after the authentication process to get the original image.

5) Enhancing data security by secret sharing- Instead of hiding data directly into document image pixels, the proposed method embeds data in the form of shares into the alpha channel of the PNG image. The effect of this may be regarded as double-fold security protection, one fold contributed by the shares as a form of disguise of the original image data and the authentication signals and the other fold contributed by the use of the alpha channel plane.

XI. PROPOSED METHOD

A method for the authentication of document images with an additional self-repair capability for fixing tampered image data is proposed. The input cover image is assumed to be a binary-like grayscale image with two major gray values like the one shown in Fig. After the proposed method is applied, 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. The proposed method is based on the so-called (k,n) -threshold secret sharing scheme

proposed by Shamir in which a secret message is transformed into shares for keeping by participants, and when of the shares, not necessarily all of them, are collected, the secret message can be lossless recovered. Such a secret sharing scheme is useful for reducing the risk of incidental partial data loss.

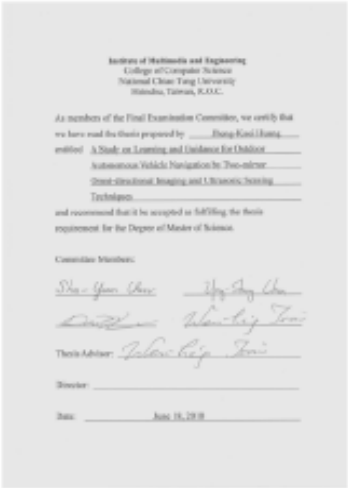

Fig 7. Binary-like grayscale document image with two major gray values

11.1 Design Technologies

Block Diagram

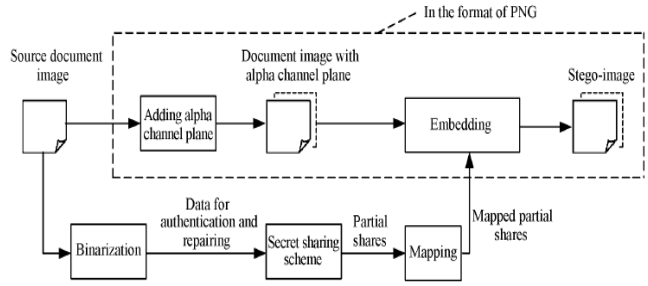

Fig 8. Creating a PNG image from a grayscale document image and an alpha channel

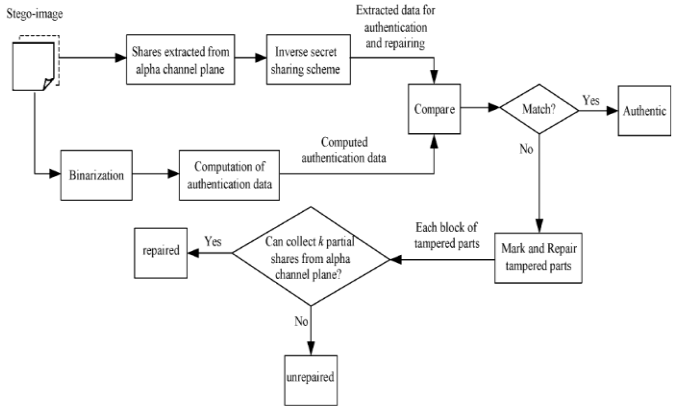

Fig 9. Authentication process including verification and self-repairing of a stego-image in PNG format

International Journal of Electrical and Electronics Engineers
Vol. No.7 Issue 02, July-December 2015
www.arresearchpublication.com
IJEEE
ISSN 2321 - 2055

XII. MATHEMATICAL EQUATIONS

Algorithm 1: (*k,n*)-threshold secret sharing.

Input : secret *d* in the form of an integer, number *n* of participants, and threshold $k \leq n$.

Output: *n* shares in the form of integers for the *n* participants to keep.

Step 1. Choose randomly a prime number *p* that is larger than *d*.

Step 2. Select $k-1$ integer values $c_1, c_2 \ldots c_{k-1}$ within the range of 0 through $p-1$.

Step 3. Select *n* distinct real values $x_1, x_2 \ldots x_n$.

Step 4. Use the following $(k-1)$- degree polynomial to compute n function values $F(x_i)$, called partial shares for $i = 1,2,\ldots n$, i.e.

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

Step 5. Deliver the two-tuple $(x_i, F(x_i))$ as a share to the i^{th} participant where $i=1,2,\ldots n$.

Since there are *k* coefficients, namely, *d* and c_1 *through* c_{k-1} *in (1)* above, it is necessary to collect at least *k* shares from *n* participants to form *k* equations of the form of *(1)* to solve these *k* coefficients in order to recover secret *d*. This explains the term threshold for *k* and the name *(k, n)* - threshold for the Shamir method. Below is a description of the just-mentioned equation-solving process for secret recovery.

Algorithm 2: Secret recovery

Input: *k* shares collected from the *n* participants and the prime number *p* with both *k* and *p* being those used in algorithm 1.

Output: secret *d* hidden in the shares and coefficients c_i used in (1) in Algorithm 1,

where $i = 1, 2 \ldots k-1$.

Step 1, Use the *k* shares

$(x_1, F(x_1)), (x_2, F(x_2)) \ldots, (x_k, F(x_k))$

To set up

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

Where $j = 1, 2 \ldots k$.

Step2. Solve the *k* equations above by Lagrange's interpolation to obtain *d* as follows:

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

Step 3. Compute c1 *through* c_{k-1} by expanding the following equality and comparing the result with (2) in Step 1 while regarding variable *x* in the equality below to be x_j *in* (2):

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

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

Algorithm for Stego-Image Generation

The following algorithm describes the generation of stego-image of the proposed method:

Algorithm 3: Generating stego image in PNG format from a given grayscale image.

Input: A grayscale image document I with two major gray values and secret key K.

Output: Stego image I' in PNG with encrypted format, relevant data embedded, including the authentication signal and the data used for repairing.

Stage 1: Authentication Signal Generation.

Step1 (Binarization of input image) Moment preserving Threshold applied I to obtain two representative gray values g_1 and g_2. Computing the average of g_1 and g_2 to obtain the threshold value. Use this threshold to binaries I, yielding a binary version of I_b.

Step2 (Conversion of cover image into PNG format) Convert I into PNG image with an alpha channel plane I_α by creating new image layer with 100% opacity and no color as I_α and combining it with I using an image processing software package.

Step3 (Starting of loop) Take in an unrefined raster scan order of 2*3 block B_b in I_b with pixels p_1 , p_2…..p_6

Step4 (Authentication signal generation) Generate 2-bit authentication signal $s=a_1a_2$with $a_1=p_1 \oplus p_2 \oplus p_3$ and $a_2=p_4 \oplus p_5 \oplus p_6$.

Stage 2: Design and Embedding of Shares.

Step5 (Creation of data for secret sharing) concatenate the 8 bits of a_1, a_2 and p_1 through p_6 form an 8-bit string, divide this string into two 4-bit segments, and convert the segment into 2 decimal numbers $m_1$and m_2 respectively.

Step6 (Generation of partial shares) Set p, c_i, and x_i the following value apply eqn. (1) of Algorithm 1, 1) p=17 (the smallest Prime number larger than 15); 2) d=$m_1$and c_1= m_2; and 3) x_1=1, x_2=2 …x6=6. Perform algorithm 1 as a (2, 6) threshold secret sharing scheme and 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

Step7 (Mapping of partial shares) Add 238 to each of $q_1$through q_6, resulting in the new value of $q_{1'}$ through q_6' respectively, which fall in the nearly total transparency range of 238 through 254 in the alpha channel plane I_α'.

Step 8 (Embedding two fractional shares in the current block) Take block B_α in I_α corresponding to B_b in I_b , select the first two pixels in B_α in the raster scan order and replace their values by q_1' and q_2' respectively.

Step 9 (Embedding remaining partial shares at random pixels) Use key K to select randomly four pixels in I_α but outside B_α , not the first two pixels of any block; in the raster scan order, and replace four pixels values by the remaining four partial shares q_3' through q_6' generated above, respectively.

Step10 (End of loop) If their exist any unprocessed block in I_b , then go to step 3 otherwise take the I in the PNG format.

Stage 3: PNG Image Encryption.

Step11 (Encryption of the PNG image) Encrypt the PNG image using chaotic logistic map, take the final I in PNG with encrypted format as the desired stego-image I′.

The prime number p used here is 17, so the values of q_1 through q_6 yield by equation (3) are between 0 and 16. After executing step 7 of above algorithm, they become q_1' through q_6' respectively. Which all fall into the small interval of integers ranging from 238 to 254 with a width of 17 (the value of the prime number). Consequent embedding of q_1' through q_6' in a narrow interval into the alpha channel plane means that very alike values will appear everywhere in the plane, resulting in a nearly uniform transparency effect, which will not stimulate notice from an attacker. We choose prime number to be 17 in the above algorithm because, if it was chosen instead to be larger than 17, then the above mentioned interval will be enlarged and the values of q_1' through q_6' will become possibly smaller than 238, creating visually whiter stego image. In contrast, the 8 bits mentioned in steps 5 and 6 above are transformed into two decimal numbers m_1 and m_2 with their maximum values being 15(step 5 above) , which are forced to lie in the range of 0 through p-1 (step 2 in algorithm 1). Therefore p should not be chosen to be smaller than 16, i.e.; p=17 is the best possible answer.

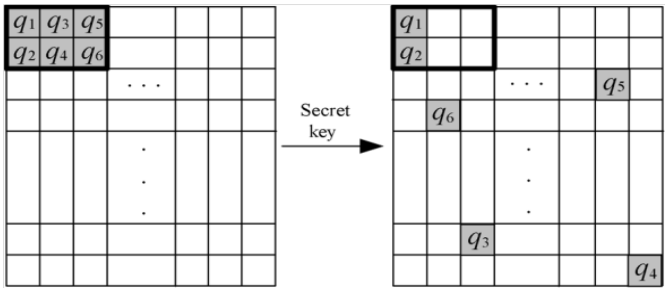

Fig 10. Illustration of embedding six shares created for a block:

Two shares embedded at the current block, and the other four in four randomly selected pixels outside the block, with each selected pixel not being the first two ones in any block.

XIII. CONCLUSION

We have proposed a secure authentication scheme for grayscale document images by the use of secret sharing method and chaotic logistic map. In this scheme security is provided by, secret sharing and encryption. Using Shamir secret sharing method both the generated authentication signal and the content of a block are

transformed into partial shares. Which are then distributed in an elegant manner into an alpha channel plane to create a PNG image. This image is encrypted by using chaotic logistic map and forms a stego image. In the authentication process, if it seen that the data is tampered then self-repairing is done in the content of the tampered block by reverse Shamir scheme. This method enhances the security by embedding the data in the alpha channel plane and encrypting the PNG image.

REFERENCES

[1] M. U. Celik, G. Sharma, E. Saber, and A.M. Tekalp, "Hierarchical watermarking for secure image authentication with localization," IEEE Trans. Image Processing, vol.11, no.6, pp.585-595, june.2002.

[2] C Yu, X Zhang "Watermark embedding in binary images for authentication", IEEE Trans. Signal Processing, vol.01, no.07, pp.865-868, September. 2004.

[3] A. Shamir, "How to share a secret," Commun.ACM, vol.22, no.11, pp.612-613, Nov, 1979.

[4] P.Jhansi Rani, S. DurgaBhavani1Int'1Conf on Recent Advances in Information Technology RAIT-2012.

[5] Chih-HsuanTzeng and Wen-Hsiang Tsai, "A new approach to authentication of Binary image for multimedia communication with distortion reduction and security enhancement. IEEE communication letters VOL.7.NO.9 2003

[6] H. Yang and A. C. Kot, "Binary image authentication with tampering localization by embedding cryptographic signature and block identifier," IEEE Signal Processing Letters, vol. 13.

[7] M. Wu and B. Liu, "Data hiding in binary images for authentication and annotation," IEEE Trans.on Multimedia, vol. 6, no. 4, pp. 528–538, Aug. 2004.

[8] Che- Wei Lee and Wen-Hsiang Tsai "A secret-sharing-based method for authentication of grayscale document images via the use of the png image with data repair capability" IEEE Trans. Image Processing., vol.21, no.1, january.2012.

[9] Niladri B. Puhan, Anthony T. S. Ho "Binary Document Image Watermarking for Secure Authentication Using Perceptual Modeling" IEEE International Symposium on Signal Processing and Information Technology2005.

[10] W.H. Tsai, "Moment-Preserving thresholding: a new approach." Computer Vision, Graphics, and Image Processing, vol. 29, no.3, pp.377-393, 1985.