Secure Dissemination and Protection of Raster Data using Merkle Hash Tree based Cryptowatermarking

Sangita Santosh Chaudhari

Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India

Abstract. Due to rapid development in the Internet and other communication technologies, it becomes quite easy to copy and distribute multiple illegal copies of high value and sensitive data. Raster data is one of the high voluminous data and it requires huge efforts to sense and generate this data. Therefore, ownership protection as well as its integrity become one of the key problems in spatial information service. There are lot many schemes are available for watermarking and encryption individually, but if both are combined gives manifold advantages. This paper presents a cryptowatermarking scheme by combining watermarking and encryption to protect the copyright of raster data as well as to provide security dissemination level. We have proposed a scheme by employing double transposition, LSB substitution watermarking and Merkle Hash Tree for encryption and watermarking. It has been observed that the proposed scheme is not only robust against encryption attacks, but also has transparency, strongness, large data hiding capacity and correct extraction of watermark.

Keywords: Cryptowatermarking, Merkle Hash tree, Robust watermarking, Raster Data

1. Introduction

Due to rapid growth in communication technology, it becomes extremely easy to digitise data and generate other multimedia contents. It becomes easy to distribute data to multiple recipients over the Internet. However, this cause misuse of all kinds of data which is in digital format. The sensing, acquisition and pre-processing of raster data are cost and manpower intensive tasks. Also, this data is of high volume and it is sensitive sometimes. It is not advisable to store the data at any untrusted server or disseminate it directly without enforcing any safeguarding mechanism. As the data is especially important in lots of analysis applications and not freely available, it may be duplicated and distributed illegally. Therefore, copyright protection is crucial and necessary in today's era. In most of the applications, cryptographic mechanisms such as encryption and digital signatures are used to achieve confidentiality, integrity, and non-repudiation services.

All these listed security mechanisms generally help users to safeguard their data at storage and communication. Although, Watermarking is being used since years as a complementary security mechanism, it fails to secure data completely. It is mainly used for copyright protection to prove the ownership of data.

Security of raster data is required at storage and dissemination level. Watermarking and encryption techniques can be useful to do secure them at these levels. There are many different approaches/schemes are available for watermarking of raster data in spatial as well as frequency domain [1-6]. However, no one have paid attention towards security watermarked raster data at dissemination level. Attackers can use many simple manipulative attacks to tamper watermarked data. Standard algorithms AES, DES, and RSA [7] are very popular for many different types of data, but they may not be suitable for raster data due to its voluminous nature. Also, these well-known algorithms take large amount of processing time for such voluminous data. Cryptography and watermarking can be combined and can provide overall security for raster data.

In the scheme proposed by Xu et al. [8], watermarking and encryption are combined, and the watermark is embedded in the image at the time of decryption of the image. One of the major problems of this scheme is that the encrypted image does not carry watermark. Jiang and Xu [9] depicted simultaneous encryption and watermarking operations in their scheme which they proposed for remote sensing data. Enhanced Arnold scrambling is used for encryption which utilises only one key for complete image. The issue here is that If a key is known to attacker, the whole encryption system fails. Encryption and watermarking are integrated in the scheme proposed by Jiang et al. [10]. They have utilised orthogonal decomposition of remote sensing data. This scheme does not give proper output and the edges get blurred in the resultant image. Ding et al. [11] have proposed integrity authentication scheme for remote sensing images wherein they have utilised perceptual hash scheme based on deep learning concepts. PCA is used to extract the features and perceptual hash is

[*] Corresponding author: sangita.chaudhari@rait.ac.in

calculated on it. This will help to detect local tampering in high resolution remote sensing images. In all existing systems, complete security analysis as well as watermark robustness analysis are not done.

This paper presents the idea of combining cryptography and watermarking to secure raster data. Simple yet effective transposition cipher, LSB based watermarking and robust Merkle hash tree is combined to provide secure solution for safeguarding raster data.

The paper is organized as follows. Proposed crypto-watermarking scheme is presented in section 2. Section 3 describes error analysis measures. Experimental results are discussed in section 4. Conclusions are drawn in section 5.

2. Proposed Cryptowatermarking Scheme

Remote sensing images includes Multispectral, hyperspectral, active/passive microwave, LIDAR images and aerial photographs etc. Due to the enhanced computing power and large number of applications capable to use these data, its demand is increased dramatically. Digital watermarking and cryptography are complementary to each other and used separately from many years to achieve copyright protection and security, but when combined as Cryptowatermarking results in vary robust and secure solution to safeguard the precious data. Simple yet strong and secure solution is proposed to provide to achieve protection and secure dissemination of remote sensing data.

In the proposed Cryptowatermarking scheme, two stage transposition cipher, least significant bit (LSB) based watermarking on zigzag pattern matrix, Merkle hash tree-based indexing of the encrypted image and RSA digital signature are effectively combined to make it robust and strong. These cipher when used as individual are prone to attacks. However, proper cascading of these ciphers makes the overall system robust and strong and it is impossible to break the security provided by this scheme.

It is a common practice in communication systems to use secure hash algorithms to verify the integrity of content transferred over insecure communication channel. The employment of secure hash algorithms enables user to securely disseminate the content over the Internet. Typically, algorithms such as Secure Hash Algorithm (SHA) and Message Digest (MD) have been used to check the content integrity after receiving the entire data from sender. The Merkle hash tree (MHT) can be used to verify content delivered from sender. MHT is a binary tree that organizes hashes hierarchically. Figure 6.1 illustrates MHT having 8 objects O1 to O8 representing leaf nodes of a tree. A node N contains hash value hN which is computed as: if N is a leaf node then hN = H(O1) otherwise hN = H(hN:lc || hN:rc). Here, N:lc and N:rc are left and right child of N respectively, and "||" represents concatenation of two binary strings. The complete tree is formed using iterative hashing. Finally, root of the MHT is signed by the data owner using public key digital signature scheme. Here, each O_i corresponds to 8x8 non overlapping block of encrypted watermarked raster data.

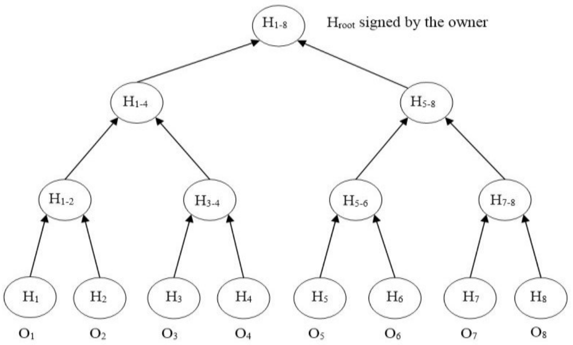

Fig. 1. Merkle Hash Tree (MHT)

Algorithm 1 and Algorithm 2 shows the Cryptowatermarking and decrpyowatermarking process respectively.

Algorithm 1: Cryptowatermarking
Input: *MxN* Raster Data (*R*), Binary Watermark(*W*), private key of sender(*S.Pr*)
Output: Encrypted Watermarked Raster Data (*EWR*), signature (*SG.EWR*)

```
1. For each of the bands (R, G and B) of input Raster Data(R) repeat the following process:
   1.1 Shift even rows by $EK_{SR}$ mod M.
       // $EK_{SR}$ is selected in such a way that $EK_{SR}$ is relatively prime to M.
   1.2 Shift even columns by $EK_{SC}$ mod N.
       // $EK_{SC}$ is selected in such a way that $EK_{SC}$ relatively prime to N.
   1.3 Perform watermarking on the Raster Data (SR) generated in step 1.2
       (i) Divide the Raster Data (SR) into 8x8 non overlapping blocks
       (ii) Read the block data using zigzag reading pattern and substitute LSB bit of each element on the pattern by each pixel of binary watermark in sequence.
       (iii) Repeat the procedure by embedding binary multiple time till all non-overlapping blocks are covered.
   1.4 Shift odd rows by $OK_{SR}$ mod M
       //$OK_{SR}$ is selected in such a way that $OK_{SR}$ is relatively prime to M.
   1.5 Shift odd columns by $OK_{SC}$ mod N.
       // $OK_{SC}$ is selected in such a way that $OK_{SC}$ is relatively prime to N.
   1.6 Concatenate all the resultant bands to get Encrypted watermarked Raster Data (EWR)
2. Perform signature generation on EWR generated in step 1.6
   2.1 Divide EWR into 8x8 non overlapping bocks
   2.2 Construct Merkle Hash tree and calculate hash of root
   2.3 Sign the hash of root by private key of sender to get signature (SG.EWR)
```

Algorithm 2: Decrpyowatermarking

```
Input: Encrypted Watermarked Raster Data (EWR), signature (SG.EWR), Public key Sender (S.Pu)
Output: MxN Decrypted Raster Data (DR), Extracted Binary Watermark(W′)
1.  Perform signature verification for the received data EWR
    1.1  Divide R into 8x8 non overlapping bocks
    1.2  Construct Merkle Hash tree and calculate hash of root
    1.3  Verify the signature using public key of sender(S.Pu). Successful verification denotes intact data whereas unsuccessful verification indicates tampering of data in communication.
2.  If signature verification is successful go to step 3 else notify sender to resend the data.
3.  For each of the bands (R, G and B) of EWR repeat the following process:
    3.1  Shift odd columns by $OK_{Sc}$ mod N.
    3.2  Shift odd rows by $OK_{SR}$ mod M.
    3.3  Extract the watermark from the Raster Data generated in step 3.2.
         (i)   Divide the Raster Data (SR) into 8x8 non overlapping blocks.
         (ii)  Read the block data using zigzag reading pattern and extract LSB bit of each element from each element on zigzag pattern.
         (iii) Repeat the procedure of LSB bit extraction for all non-overlapping blocks to get the binary watermark (W′)
    3.4  Shift even columns by $EK_{SC}$ mod N.
    3.5  Shift even rows by $EK_{SR}$ mod M.
4.  Concatenate all the resultant bands to obtain Decrypted Raster Data (DR)
```

Two stage transposition ciphers employed in the schemes provide good diffusion whereas the watermarking process itself provides good confusion and hence making the scheme robust. The encrypted and watermarked raster data is divided in to 8x8 blocks. All the pixels in 8x8 blocks are concatenated and represents one object. All the blocks are represented in object as concatenation of all pixels in corresponding blocks. They can be treated as leaf node of MHT (figure 1). Hash of the root is calculated using bottom up hashing approach and finally root is signed by sender with his private key. Sender sends the key set as {*S.Pu*, EK_{SC}, EK_{SR}, OK_{SC}, OK_{SR}, *M, N, size of watermark*} to the receiver. Encrypted watermarked data (EWR) along with the calculated signature (*SG.EWR*) is also sent to receiver. RSA digital signature is used to generate signature on EWR.

At receiver side, receiver first verifies the signature *SG.EWR*. Successful verification indicates tamperproof communication of encrypted watermarked data, else the sender is requested to send the data again. Decryption and watermark extraction are carried out as per the steps illustrated in Algorithm 2 and the required keys received from sender.

3. Evaluation Measures

Normalised correlation is used to indicated similarity between original watermark (W_k) and extracted watermark (W_k'). It is evaluated using equation 1. A Complete similarity is indicated by a coefficient with value one. Normalised correlation ranges between 0 to 1.

$$NC = \frac{\sum_{k} (w_k \cdot w_k')}{\sqrt{\sum_{k} (w_k')^2}} \qquad (1)$$

Equation 2 shows the Entropy (E) of an image I where $P_r(\alpha_i)$ is the probability of occurrence of the grey level α_i and G is the total number of grey levels. Entropy is a measure indicating randomness of the pixels in the encrypted image. The value of entropy is 8 if every pixel has equal probability in encrypted image. It shows complete randomness in the output image. Low value makes the encrypted image susceptible to various statistical attacks.

$$E(I) = \sum_{i=1}^{G} p_r(\alpha_i)\log_2(p_r(\alpha_i)) \qquad (2)$$

4. Results and Discussion

The proposed Cryptowatermarking system is evaluated using various remote sensing images of varying sizes. Figure 2 shows 6 such images 1 to 6 of sizes 1280x1280, 1600x1080, 800x800, 960x840, 760x640 and 640x640 respectively. The watermark of size 100x50 pixels is used for watermarking. In the experiment, different keyset for double transposition cipher is used as shown in Table 1. Encrypted watermarked raster data and decrypted raster data are shown in figure 3 and 4. The original watermark and extracted watermark are shown in figure 5. The corresponding normalized coefficient between original and extracted watermark is obtained as one. SHA 256 is used for hashing in Merkle Hash Tree and RSA digital signature is employed for signing the root of Merkle Hash Tree. Public modulus(n) is considered as "*d94d889e88853dd89769a18015a0a2e6bf82bf356fe14f251fb4f5e2df0d9f9a94a68a30c428b39e3362fb3779a497eceaea37100f264d7fb9fb1a97fbf621133de55fdcb9b1ad0d7a31b379216d79252f5c527b9bc63d83d4ecf4d1d45cbf843e8474babc655e9bb6799cba77a47eafa838296474afc24beb9c825b73ebf549*" and public exponent (*e*) is considered as *10001*. Private exponent (*d*) was calculated using values of *n* and *e*.

Encryption and decryption results are highly dependent on the order of ciphers and proper usage of utilized key space. Cryptosystem is said to be secure if the image cannot be decrypted correctly even if there is a change in sequence of ciphers and keys. Keys used here are more sensitive they should be relatively prime to number of pixels in rows and columns. Decryption many are not successful if there is a small change in the keys used in encryption and decryption.

In cryptography, it is very crucial to design the algorithms which provides robustness against all the attacks like brute force and statistical attacks. One of the best ways to check the security of the image security algorithm is histogram analysis. When histogram of encrypted image is uniform, it becomes difficult to get information from image and thus launching statistical attacks. As Merkle hash tree, hashing and digital signature are used, it is difficult to launch an attack and compromise key or raster data.

Table 1. Keyset for double transposition cipher

Image	EK_{SR}	OK_{SR}	EK_{SC}	OK_{SC}
Sat_Image1	347	991	103	1173
Sat_Image2	563	1093	1051	163
Sat_Image3	401	191	571	73
Sat_Image4	199	311	31	839
Sat_Image5	47	419	541	313
Sat_Image6	463	557	13	97

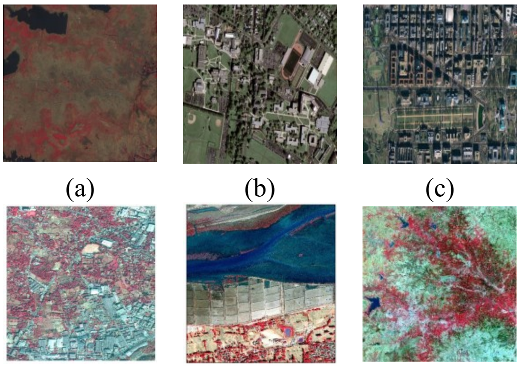
(a) (b) (c)
(d) (e) (f)

Fig. 2. Original Raster Data: (a) Sat_Image1; (b)Sat_Image2; (c)Sat_Image3; (d) Sat_Image4; (e)Sat_Image5; (f) Sat_Image6

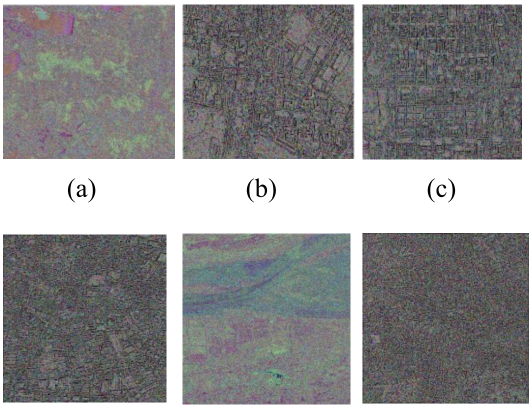
(a) (b) (c)
(d) (e) (f)

Fig. 3. Encrypted Watermarked Raster Data: (a) EWSat_Image1; (b)EWSat_Image2; (c)EWSat_Image3; (d) EWSat_Image4; (e)EWSat_Image5; (f) EWSat_Image6

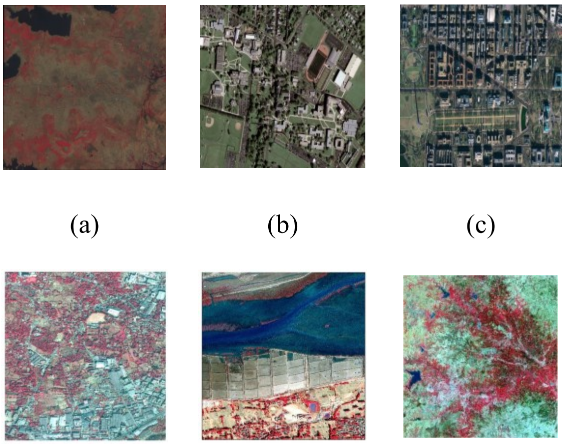
(a) (b) (c)
(d) (e) (f)

Fig. 4. Decrypted Raster Data: (a) DSat_Image1; (b)DSat_Image2; (c)DSat_Image3; (d) DSat_Image4; (e)DSat_Image5; (f) DSat_Image6

(a) (b)

Fig. 5. (a) Original Watermark, and (b) Extracted watermark with NC=1

Double transposition cipher and watermarking algorithm provides exceptionally good diffusion and confusion in the encrypted watermarked image. Histogram analysis can be used to check confusion and diffusion of the proposed scheme. Figure 6 shows the RGB histogram of both the original and encrypted watermarked raster data. From the figure, it is observed that the histogram of the encrypted watermarked data is nearly uniformly distributed, and very different from the histogram of the original data and hence there will be no clue from the encrypted watermarked image to launch statistical attack on resultant data in the proposed scheme. Entropy for encrypted watermarked data are shown in Table 2 and it is observed that it is nearly equal to 8 which shows that the scheme is robust against statistical attacks.

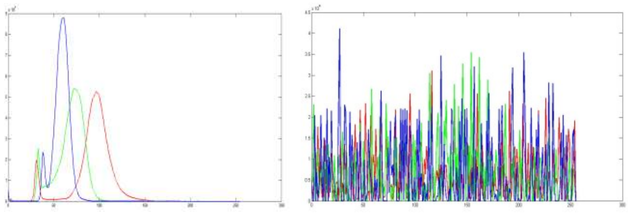
(a) (b)

Fig. 6. Histogram Analysis (a) Original Raster Data (Sat_Image1), and (b) Encrypted watermarked Raster Data

Table 2. Entropy of Encrypted Raster Data

Image	Entropy
EWSat_Image1	7.9990
EWSat_Image 2	7.9904
EWSat_Image 3	7.9900
EWSat_Image 4	7.9921
EWSat_Image 5	7.9981
EWSat_Image 6	7.9823

Perceptual quality of the image can be checked using peak to signal noise ratio (PSNR). If its value is small, image quality is poor and vice versa. From table 3, it is observed that the proposed system results into small PSNR values thus resulting into true randomness in the encrypted image.

Table 3. PSNR of Encrypted Raster Data

Image	PSNR
EWSat_Image1	12. 7021
EWSat_Image 2	12.2600
EWSat_Image 3	12.0306
EWSat_Image 4	11.9327
EWSat_Image 5	12.3943
EWSat_Image 6	12.2846

5. Conclusion

In this paper, we have proposed cryptowatermarking, a combination of watermarking and encryption to provide copyright protection for raster data and for secure dissemination copyrighted raster data. During transmission, Merkle Hash tree-based encryption and RSA digital signature helps in preventing attacks and data loss. Watermark can be efficiently extracted at receiver side which can be used to prove ownership of the data. The proposed scheme satisfies the security of encryption, the invisibility and robustness against various brute force and statistical attacks.

References

[1] M. Barni, F. Bartolini, V. Cappellini, E. Magli, and G. Olmo. Near-lossless digital watermarking for copyright protection of remote sensing images. IEEE International Geoscience and Remote Sensing Symposium (2002).

[2] Y. Chauhan, P. Gupta, and K. Majumder. Digital Watermarking of Satellite Images, India Conference on Computer Vision, Graphics, and Image Processing (2002).

[3] T. Hemalatha, V. Joevivek, K. Sukumar K. Soman. Robust watermarking of remote sensing images without the loss of spatial information, 10th ESRI India User Conference (2009).

[4] B. Kumari and V. Rallabandi. Modified patchwork-based watermarking scheme for satellite imagery, Signal Processing (2008).

[5] P. Zhu, and C. Chen, A copyright protection watermarking algorithm for remote sensing image based on binary image watermark, International Journal Light and Electron Optics (2013).

[6] B. Ziegeler, H. Tamhankar, J. Fowler, and L. Bruce. Wavelet-Based watermarking of remotely sensed imagery tailored to classification performance, IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data (2003).

[7] B. Schneier. Applied cryptography. Wiley (1995).

[8] Y. Xu, Z. Xu, and Y. Zhang. Content security protection for remote sensing images integrating selective content encryption and digital fingerprint, Journal of Applied Remote Sensing (2012).

[9] L. Jiang, and Z. Xu, Commutative encryption and watermarking for remote sensing image, International Journal of Digital Content Technology, and its Applications (2012).

[10] L. Jiang, Z. Xu, and Y. Xu. A new comprehensive security protection for remote sensing image based on the integration of encryption and watermarking, IEEE International Geoscience and Remote Sensing Symposium (2013).

[11] K. Ding, Z. Yang, Y. Wang, and Y. Liu. An improved perceptual hash algorithm based on u-net for the authentication of high-resolution remote sensing image. Applied Sciences (2019).