Comments on Matrix-Based Secret Sharing Scheme for Images

Esam Elsheh and A. Ben Hamza

Concordia Institute for Information Systems Engineering
Concordia University, Montréal, QC, H3G 2W1, Canada
{e_elshe,hamza}@ciise.concordia.ca

Abstract. Several attempts have been made to propose efficient secret sharing schemes for 2D images. Rey M.D (Iberoamerican Congress on Pattern Recognition, 2008) proposed a relatively fast image secret sharing scheme based on simple binary matrix operations. In this work, we show that care should be taken when choosing the matrices that corresponding to the shares, in particular if the rank of these singular matrices is not low enough then one can recover the secret image from only one share. Experimental results are provided to demonstrate the practicality of the recovery procedure on various 2D images.

Keywords: Secret sharing; Image processing; Cryptanalysis.

1 Introduction

The continuing advancements in computer technologies and the rapid increase in internet users have led to the increasing usage of network-based data transmission. In numerous applications, such as military documents and sensitive business data, this information must be kept secret and safe. Recently, 2D images are considered as important as any other text sensitive information. As a result, several 2D image-protection techniques, such as data encryption in [1,2] and steganography in [3,4], have been proposed to insure the security of secret images. One major disadvantage of the traditional protection techniques, such as encryption, is their policy of centralized storage, in that an entire protected model is usually maintained in a single information storage. If an intruder detects security vulnerability in the information storage in which the protected images resides, then s/he may attempt to decipher the secret model inside, or simply damage the entire information storage. Hence, the secret sharing is a defense mechanism to protect the secret that does not suffer from these problems. It works by splitting the secret into n shares that are transmitted and stored separately. One can then reconstruct the original secret if at least a preset number t ($1 \leq t \leq n$) of these n shares are obtained. However, knowledge of less than t shares is insufficient for revealing the secret.

The first secret sharing scheme was introduced independently in [5] and [6]. Both schemes are based on the use of Lagrange interpolation polynomial and the intersection of affine hyperplanes, respectively. Since after, several studies

have investigated different implementations of the (t, n) -threshold schemes and their usage in the keys communication of the cryptosystems. The majority of these schemes are based on different mathematical primitives, such as matrix theory and prime numbers [7]. These protocols are specifically designed for text and numeric data. Due to the main distinctive nature of multimedia, in the sense that they have a large amount of data and the difference between two neighboring values is typically very small, it is considerably difficult to apply directly traditional secret sharing schemes to digital images. Thus, various secret sharing protocols have been designed exclusively for digital images, some based on vector quantization [9], Shamir-based schemes [10, 11, 12], sharing circle [7, 8], binary matrices [13], or cellular automata [1, 2]. In this paper, we show that the scheme proposed in [13] is not an ideal image secret sharing scheme due to the fact that any participant is able to recover the secret image using only his share.

The rest of the paper is organized as follows. In Section 2, we briefly review the description of the image secret sharing scheme proposed in [13]. In Section 3, we present the theoretical steps to recover the secret image from a single share only. In Section 4, we provide experimental results on different images to validate the effectiveness of the recovering process. Finally, we conclude in Section 5.

2 Description of $(2, n)$ -Threshold Image Secret Sharing Scheme Proposed by Rey.M.D

The proposed secret sharing scheme in [13] is based on binary operations of two matrices $\mathbf{A}$ and $\mathbf{B}$ that satisfy the following algebraic property, let $\mathbf{A}$ and $\mathbf{B}$ two binary matrices such that $\mathbf{A} \oplus \mathbf{B} = \mathbf{Id}$, then the following theorem holds,

Theorem 1. $\mathbf{A}^m \oplus \mathbf{B}^m = \mathbf{Id}$ if and only if $m = 2^e$, with $e \in \mathbb{Z}^+$

The reader can refer to [13] for the proof of theorem 1.

The steps of the image secret sharing scheme proposed in [13] are described as follows,

The setup phase. Let the matrix $\mathbf{J} = (p_{i,j})$ be a gray-level image defined by $n \times n$ pixels such that $p_{i,j}$ is the numeric value of the gray color of the (i, j) -th pixel, where $p_{i,j} = (q_{i,j}^1, q_{i,j}^2, \dots, q_{i,j}^8) \in \mathbb{Z}_2^8$, with $1 \leq i, j \leq n$. Consequently, eight binary matrices with coefficients in $\mathbb{Z}_2$ are extracted from $\mathbf{J} = (q_{i,j}^k)$, where $1 \leq k \leq 8$. For simplicity, we write $\mathbf{J} = \mathbf{J}^1 \parallel \mathbf{J}^2 \parallel \dots \parallel \mathbf{J}^8$ as a concatenation of eight binary matrices, where each of them represents a black and white subimage. Fig. 1 and Fig. 2 correspond to Lena gray image defined by 128×128 pixels and its eight subimages, respectively.

The sharing phase. We choose two singular binary matrices $\mathbf{A}$ and $\mathbf{B}$ satisfying Theorem 1. Then, we randomly choose $n/2$ integer numbers $e_1 \leq e_2 \leq \dots \leq e_{n/2}$ and computes $m_i = 2^{e_i}$. The shares $\mathbf{S}_i^1, \mathbf{S}_i^2$, where $1 \leq i \leq n/2$, are computed as follows,

$$\begin{aligned}\mathbb{S}_i^1 &= \mathbf{A}^{m_i} \cdot \mathbf{J}^1 \parallel \mathbf{A}^{m_i} \cdot \mathbf{J}^2 \parallel \dots \parallel \mathbf{A}^{m_i} \cdot \mathbf{J}^8, \\ \mathbb{S}_i^2 &= \mathbf{B}^{m_i} \cdot \mathbf{J}^1 \parallel \mathbf{B}^{m_i} \cdot \mathbf{J}^2 \parallel \dots \parallel \mathbf{B}^{m_i} \cdot \mathbf{J}^8.\end{aligned}\quad (1)$$

The recovery phase. The users P_i^1 and P_i^2 combine their shares, $\mathbb{S}_i^1$, $\mathbb{S}_i^2$, and computes the original image as follows,

$$\mathbf{J} = (\mathbf{A}^{m_i} \cdot \mathbf{J}^1) \oplus (\mathbf{B}^{m_i} \cdot \mathbf{J}^1) \parallel (\mathbf{A}^{m_i} \cdot \mathbf{J}^2) \oplus (\mathbf{B}^{m_i} \cdot \mathbf{J}^2) \parallel \dots \parallel (\mathbf{A}^{m_i} \cdot \mathbf{J}^8) \oplus (\mathbf{B}^{m_i} \cdot \mathbf{J}^8).$$

The drawback of this scheme is that every participant P_i^1 has only one qualified participant from the pool of participants that s/he can collude to recover the original image. If the share of the qualified participant P_i^2 is altered or modified, the participant P_i^1 will never be able to recover the original image. Obviously, this scheme does not provide the basic property of the secret sharing idea, which is the safe, that is, the original secret can be recovered even if more than one share were destroyed.

Thus, the author in [13] proposed a generalization of the protocol. Such that any participant P_i^1 , $1 \leq i \leq n/2$, can combine his share with any other participant P_j^2 , $1 \leq j \leq n/2$ to recover the secret image. The generalization of the protocol is exactly the same as in the basic proposal except for the sharing phase where the data $\{\mathbb{S}_i^1, m_i = 2^{e_i}\}$, $\{\mathbb{S}_i^2, m_i = 2^{e_i}\}$ are distributed to the participants P_i^1 , P_i^2 respectively. Finally, in the recovery phase the participants P_i^1 , P_j^2 recover the secret $\mathbf{J}$ image as follows:

- i) They compare the integer numbers m_i and m_j .
If $m_i < m_j$, P_i^1 computes

$$\mathbf{A}^{m_i+m_j-m_i} \cdot \mathbf{J}^1 \parallel \dots \parallel \mathbf{A}^{m_i+m_j-m_i} \cdot \mathbf{J}^8 = \mathbf{A}^{m_j} \cdot \mathbf{J}^1 \parallel \dots \parallel \mathbf{A}^{m_j} \cdot \mathbf{J}^8.$$

otherwise, if $m_i > m_j$, P_j^2 computes

$$\mathbf{B}^{m_j+m_i-m_j} \cdot \mathbf{J}^1 \parallel \dots \parallel \mathbf{B}^{m_j+m_i-m_j} \cdot \mathbf{J}^8 = \mathbf{B}^{m_i} \cdot \mathbf{J}^1 \parallel \dots \parallel \mathbf{B}^{m_i} \cdot \mathbf{J}^8.$$

- ii) The recovery of the original image is carried as follows

$$\mathbf{J} = (\mathbf{A}^{m_j} \cdot \mathbf{J}^1) \oplus (\mathbf{B}^{m_j} \cdot \mathbf{J}^1) \parallel \dots \parallel (\mathbf{A}^{m_j} \cdot \mathbf{J}^8) \oplus (\mathbf{B}^{m_j} \cdot \mathbf{J}^8)$$

or

$$\mathbf{J} = (\mathbf{A}^{m_i} \cdot \mathbf{J}^1) \oplus (\mathbf{B}^{m_i} \cdot \mathbf{J}^1) \parallel \dots \parallel (\mathbf{A}^{m_i} \cdot \mathbf{J}^8) \oplus (\mathbf{B}^{m_i} \cdot \mathbf{J}^8)$$

We should highlight here that in order for the participants P_i^1 and P_j^2 to carry out the recovery phase of the generalized scheme, they must know the matrices $\mathbf{A}$ and $\mathbf{B}$.

We should also note that the most significant bits of the 8-bit values of grey images are the most important bits. One can distinguish the original image from the first subimages. For instance, if we can recover $\mathbf{J}^1$ or $\mathbf{J}^2$ even partly we can recognize the original image. Recall that in cryptography, the distinguishing attack is an attack in which the attacker is given an encrypted secret, which is here the secret image, and try to determine if this secret is random or it comes from a specific cryptosystem.

Fig. 1. 128×128 Lena gray image

Fig. 2. (a) Subimage J^1 (b) Subimage J^2 (c) Subimage J^3 (d) Subimage J^4 (e) Subimage J^5 (f) Subimage J^6 (g) Subimage J^7 (h) Subimage J^8

3 Recovering the Original Image from a Single Share

It is proven in ([14], Ch. A.3.3) that a real-valued matrix filled with independent and identically distributed random variables, with continuous probability distribution function, will be singular with probability zero. On the contrary, in $GF(2)$ the probability a square random binary matrix is singular as its dimension tends to infinity is 71.1 [15].

The scheme proposed in [13] stated that the random matrices $\mathbf{A}$ and $\mathbf{B}$ must be singular without specifying the value of the ranks. Furthermore, generating a singular binary matrix randomly will produce a matrix that has high rank with high probability. Consequently, solving a set of linear equations when the coefficient matrix has a high rank will result in decreasing the number of independent/depended variables and increasing the number of determined variables.

The main weakness in the generalization of the image secret sharing scheme proposed in [13] is that for the participants to reconstruct the secret image they must know $\mathbf{A}$ and $\mathbf{B}$ along with the integer numbers m_i and m_j . For instance, the participants P_i^1 and P_j^2 must know $(\mathbf{A}, \mathbf{B}, m_i, m_j)$ to perform the comparison between m_i and m_j and the computation $\mathbf{A}^{m_i+m_j-m_i} \cdot \mathbf{J}^1 \parallel \dots \parallel \mathbf{A}^{m_i+m_j-m_i} \cdot \mathbf{J}^8$ or $\mathbf{B}^{m_j+m_i-m_j} \cdot \mathbf{J}^1 \parallel \dots \parallel \mathbf{B}^{m_j+m_i-m_j} \cdot \mathbf{J}^8$, as stated in the generalization of the scheme.

To recover the original $n \times n$ gray image from a single share, the participant P_i^1 has his share $S_i^1 = \mathbf{A}^{m_i} \cdot \mathbf{J}^1 \parallel \dots \parallel \mathbf{A}^{m_i} \cdot \mathbf{J}^8$ and knows $\mathbf{A}$ and m_i . Then, he perform the following to recover $\mathbf{J}^1$,

Let $\mathbf{A} = (a_{i,j})$, $\mathbf{J}^1 = (J_{i,j})$ and $\mathbf{A}^{m_i} \cdot \mathbf{J}^1 = (s_{i,j})$, calculate $\mathbf{A}^{m_i} = (\hat{a}_{i,j})$, $1 \leq i, j \leq n$ and form the set of linear Boolean equations as follows,

For $k = 1$ to n

$$\begin{aligned}
 \bigoplus_{i=1}^n \hat{a}_{k,i} J_{i,1} &= s_{k,1} \\
 \bigoplus_{i=1}^n \hat{a}_{k,i} J_{i,2} &= s_{k,2} \\
 \vdots &\quad \quad \quad \vdots \\
 \bigoplus_{i=1}^n \hat{a}_{k,i} J_{i,n-1} &= s_{k,n-1} \\
 \bigoplus_{i=1}^n \hat{a}_{k,i} J_{i,n} &= s_{k,n}
 \end{aligned} \tag{2}$$

We have $n \times n$ binary linear equations with $n \times n$ variables. Solving this set of linear equations will produce determined variables (pixels) with specific values, depended variables, and independent variables which we can assign any value to them. Due to the nature of the image, there is no need to find the exact values of all the pixels, with only small percentage of the correct values of the pixels we can distinguish the original image, as we can see from the recovered $\mathbf{J}^1$ and $\mathbf{J}^2$ in Fig 3. To recover the original gray image, we perform the previous recovery steps to all eight subimages and then concatenate them together.

4 Experimental Results

We applied the recovery technique on different 128×128 gray images, namely, Lena, F16 and Fishing boat. The singular binary matrix $\mathbf{A}$ was chosen randomly with a rank equals 127 and the integer number m_i equals 2.

It is worthy to mention that we used the BooleanPolynomial class from the SAGE [16] package to do the fast calculations for solving the set of linear Boolean equations depending on 128×128 binary variables.

As an example in the case of Lena, after solving the set of linear binary equations, the number of undetermined variables was 12928, which were set to zeros. On the other hand, the number of determined variables for each of $\mathbf{J}^1, \mathbf{J}^2 \dots, \mathbf{J}^8$ was 3456, which is the number of bits (pixels) that were recovered exactly. Although this number seems too small comparing to 128×128 bits, it is enough to recognize the original image easily as shown in Fig.3.

One can see from Fig. 4 and Fig. 5 the similarity between the recovered images and the corresponding original images.

To see the Matlab code used in the implementation and the experimental results, look at: http://users.encs.concordia.ca/~e_elsh/ImageSecretSharing/.

Fig. 3. The recovered subimages of Lena (a) J^1 (b) J^2

Fig. 4. Original images of (a) Lena (b) F16 (c) Fishing boat (d)-(f) their histograms

Fig. 5. Recovered images of (a) Lena (b) F16 (c) Fishing boat (d)-(f) their histograms

5 Conclusions

In this paper, we showed the proposed scheme for image secret sharing by Rey M.D is insecure. This is because any participant in the scheme can recover the original image without the need to combine his share with any other participant. Specifically, the main flaw is the generalization of the protocol, where both matrices $\mathbf{A}$ and $\mathbf{B}$ are known to all participants.

References

1. Cheng, H., Xiaobo, L.: Partial encryption of compressed images and videos. *IEEE Trans. Signal Process.* 48(8), 2439–2451 (2000)
2. Bourbakis, N., Dollas, A.: Scan-based compression-encryption hiding for video on demand. *IEEE Multimedia Mag.* 10, 79–87 (2003)
3. Marvel, L.M., Boncelet, C.G., Retter, C.T.: Spread spectrum image steganography. *IEEE Trans. Image Process.* 8(8), 1075–1083 (1999)
4. Petitcolas, F.A.P., Anderson, R.J., Kuhn, M.G.: Information hiding-a survey. *Proc. IEEE, Special Issue on Protection of Multimedia Content* 87(7), 1062–1078 (1999)
5. Shamir, A.: How to share a secret. *ACM Comm.* 22(11), 612–613 (1979)
6. Blakley, G.R.: Safeguarding cryptography keys. In: *Proc. of the AFIPS 1979 National Computer Conference*, vol. 48, pp. 313–317 (1979)
7. Menezes, A., Van Oorschot, P., Vanstone, S.: *Handbook of Applied Cryptography*. CRC Press, Boca Raton (1997)
8. Tsai, D., Chen, T., Horng, G.: A cheating prevention scheme for binary visual cryptography with homogeneous secret images. *Pattern Recogn.* 40, 2356–2366 (2007)
9. Chang, C., Hwang, R.: Sharing secret images using shadow codebooks. *Inform. Sciences* 111, 335–345 (1998)
10. Wang, R.Z., Su, C.H.: Secret image sharing with smaller shadow images. *Pattern Recognition Letters* 27(6), 551–555 (2006)
11. Thien, C.C., Lin, J.C.: Secret image sharing. *Computers and Graphics* 26(1), 765–770 (2002)
12. Wu, Y., Thien, L., Lin, J.: Sharing and hiding secret images with size constrain. *Pattern Recogn.* 37, 1377–1385 (2004)
13. Rey, M.D.: A matrix-based secret sharing scheme for images. In: Ruiz-Shulcloper, J., Kropatsch, W.G. (eds.) *CIARP 2008. LNCS*, vol. 5197, pp. 635–642. Springer, Heidelberg (2008)
14. Bard, G.: Algorithms for the solution of linear and polynomial systems of equations over finite fields, with applications to cryptanalysis. PhD thesis, Department of Applied Mathematics and Scientific Computation, University of Maryland at College Park (2007)
15. Studholme, C., Blake, I.F.: Random Matrices and Codes for the Erasure Channel. *Algorithmica* 56(4), 605–620 (2010)
16. William, S.: Sage Mathematics Software (Version 4.4.2). The Sage Group (2010), <http://www.sagemath.org>