Cryptanalysis of an Image Scrambling Encryption Algorithm

Chengqing Li, Dongdong Lin
College of Information Engineering, Xiangtan University, Xiangtan 411105, Hunan, China

Abstract
Position scrambling (permutation) is widely used in multimedia encryption schemes and some international encryption standards, like DES and AES. This paper re-evaluated security of a typical image scrambling encryption algorithm (ISEA). Using the internal correlation remaining in the cipher-image, we can disclose some important visual information of the corresponding plain-image under the scenario of ciphertext-only attack. Furthermore, we found the real scrambling domain, position scrambling scope of the scrambled elements, of ISEA, which can be used to support the most efficient known/chosen-plaintext attack on it. Detailed experimental results verified the contribution points and demonstrated some advanced multimedia processing techniques can facilitate the cryptanalysis of multimedia encryption algorithms.

Keywords: Ciphertext-only attack, known-plaintext attack, cryptanalysis, scrambling, template matching.

1. Introduction
Position scrambling (permutation) is one of the most simple and efficient methods for protecting all kinds of multimedia data [1, 2]. More important, it can effectively facilitate an encryption scheme obtain efficient combination of confusion and diffusion properties. In the past two decades, a large number of multimedia scrambling algorithms were proposed by designing different mechanisms to derive the position scrambling relation on scrambling elements from a secret key. As the number of possible scrambling relation is the factorial of that of scrambling elements, there exists a huge number of different scrambling relation in theory when the latter is sufficiently large. As for image, video and audio (speech), various datum can be selected as scrambling elements, such as bit, pixel, and compressing coefficients of image [1, 3], transform coefficients and motion vectors of video [4], frames of audio [5]. If the scrambling elements are spatial blocks of an image, recovery of the original image can be attributed to general problems of solving Jigsaw puzzles discussed in [6].

Opposite to cryptography, design an algorithm to realize secure communication between the sender and the intended recipient, the object of cryptanalysis is to gain as much information on the secret key as possible without any prior knowledge on it. Some specific multimedia scrambling encryption algorithms were successfully cryptanalyzed under different conditions, in terms of the number of available plaintexts and the corresponding ciphertexts, computational complexity of the attack, and the costed storage space [7, 8]. As different scrambling elements may have dramatically different effects on the sensible information of the plaintext, different multimedia scrambling encryption algorithms may own totally different capabilities against ciphertext-only attack. However, their capabilities against plaintext attacks can be analyzed by a uniform model [3]. As cryptanalyzed in [3], any multimedia scrambling encryption algorithm can be efficiently broken with \(O(\lceil \log_2(MN) \rceil)\) plaintexts and \(O(\lceil \log_2(MN) \rceil \cdot MN^2)\) times of computation, where \(MN\) is the size of scrambling domain, position scrambling scope of the scrambled elements, \(L\) is the number of different value levels of the scrambled elements, and \(\lceil \cdot \rceil\) returns the smallest integer greater than or equal to a given number. In [9], the computational load is further reduced to \(O(\lceil \log_2(MN) \rceil \cdot MN)\) by replacing the operations of intersecting sets of position candidates with linear visit of a branch tree, whose node stores information about the position candidates.

In [10], a typical image scrambling encryption algorithm (ISEA) was proposed by scrambling binary presentation of a gray-scale plain-image with a pseudo-random number sequence, generated by iterating a digital chaotic map. In 2011, ISEA was analyzed as a scrambling encryption algorithm exerting on a scrambling domain of size \(M \times (8N)\) [9]. In [10, Sec. 3.1], a weaker version of ISEA was suggested to decrease the computation complexity and save running
time: every row and column of the binary matrix of size \( M \times (8N) \) are scrambled with the same scrambling vector, respectively. In 2012, a set of specific plaintexts were constructed to break equivalent version of the secret key of ISEA \([11]\).

This paper re-evaluated security of the weaker version of ISEA and found the real reason supporting the advantage claimed by the authors of \([11]\): it does not exert on scrambling domain of size \( M \times (8N) \), and just cascade two scrambling algorithms with scrambling domain of sizes \( M \) and \( 8N \). So, we proposed the most efficient known-plaintext attack and a general chosen-plaintext attack on it. Beside these, we found some important visual information of the plain-image can be observed from only one cipher-image encrypted by the analyzed algorithm.

The rest of the paper is organized as follows. The next section briefly introduce ISEA. Security of ISEA against ciphertext-only attack and plaintext attacks are re-evaluated in Sec. 2 with some experimental results. The last section concludes the paper.

2. Description of ISEA

The encryption object of ISEA is a gray-scale image of size \( M \times N \) (height×width), which can be represented as a matrix over integer set \( \{0, 1, \cdots, 255\} \), \( I = [I(i, j)]_{i=0,j=0}^{M-1,N-1} \). The image \( I \) is further decomposed as an \( M \times (8N) \) binary matrix \( B = [B(i, l)]_{i=0,j=0}^{M-1,8N-1} \), where \( \sum_{k=0}^{7} B(i, l) \cdot 2^k = I(i, j) \), \( l = 8 \cdot j + k \). After preforming scrambling on the row and column vectors of \( B \), the cipher-image \( I' = [I'(i, j)]_{i=0,j=0}^{M-1,N-1} \) is obtained, where \( I'(i, j) = \sum_{k=0}^{7} B'(i, 8 \cdot j + k) \cdot 2^k \). The basic concrete parts of ISEA can be described as follows.

- **The secret key:** three positive integers \( m, n, \) and \( T \), and the initial condition \( x_0 \in (0,1) \) and control parameter \( \mu \in (3.569945672,4) \) of Logistic map
  \[
  f(x) = \mu \cdot x \cdot (1 - x).
  \]  

- **Initialization:** 1) run the Logistic map from \( x_0 \) to generate a chaotic sequence, \( \{x_k\}_{k=1}^{L} \), where \( L = \max(m + M, n + 8N) \); 2) produce a vector \( T_M \) of length \( M \), where \( S_M(T_M(i)) \) is the \( (i + 1) \)-th largest element of \( S_M = \{x_{m+k}\}_{k=1}^{M}, i \in \{0,1,\cdots, M-1\} \); 3) produce a matrix \( T_N \) of size \( 1 \times (8N) \), where \( S_N(T_N(j)) \) is the \( (j + 1) \)-th largest element of \( S_N = \{x_{n+k}\}_{k=1}^{N}, j \in \{0,1,\cdots, 8 \cdot N - 1\} \).

- **The encryption procedure:**
  - **Step 1 – vertical permutation:** generate an intermediate matrix \( B' = [B'(i, l)]_{i=0,j=0}^{M-1,8N-1} \), where
    \[
    B'(i,:) = B(T_M(i,:));
    \]  
  - **Step 2 – horizontal permutation:** generate an intermediate matrix \( B'' = [B''(i, l)]_{i=0,j=0}^{M-1,8N-1} \), where
    \[
    B''(:,l) = B''(:, T_N(l));
    \]  
  - **Step 3 – repetition:** reset the value of \( x_0 \) to the current state of map \( (1) \), and repeat the above operations from the \textit{initialization} part for \( (T - 1) \) times.

- **The decryption procedure** is similar to the encryption one except some simple modifications: 1) the multiple rounds of encryption are executed in a reverse order; 2) \textit{Step 2} is exerted first in each round; 3) the left parts and right parts of Eq. (2) and Eq. (3) are swapped, respectively.

Note that the full version of ISEA is given in \([9]\), Sec. 2], where the matrix \( T_N \) is of size \( M \times (8N) \), each row vector \( B'(i,:) \) is scrambled by the \( i \)-th row vector of \( T_N \). As the full version of ISEA was cryptanalyzed comprehensively in \([9]\), only the simpler version is studied in the remainder of this paper.

3. Cryptanalysis of ISEA

3.1. Ciphertext-only Attack on ISEA

Ciphertext-only attack is an attack model for cryptanalysis where the attacker can only access some ciphertexts. Obviously, owning sufficient robustness withstand the attack is a basic requirement for any multimedia encryption algorithm. Unfortunately, ISEA fails to satisfy the requirement seriously.

The essential form of the encryption process of ISEA can be represented as

\[
B' = (T_L)^T \cdot B \cdot (T_R)^T,
\]

where \( T_L \) and \( T_R \) are permutation matrices representing the vertical permutation in Eq. (2) and the horizontal permutation in Eq. (3), respectively \([11]\). As is well known, multiplication product of any number of arbitrary permutation matrices is still a permutation matrix, a square binary matrix owning exactly one entry of 1 in each row and each column. Regardless of the value of \( T \) and whether intermediate matrices \( B' \) in Eq. (2) and \( B'' \) in Eq. (3) are updated in each round, the final essential form of the encryption process of ISEA still is

\[
B' = (\hat{T}_L)^T \cdot B \cdot (\hat{T}_R)^T,
\]

where \( \hat{T}_L \) and \( \hat{T}_R \) are permutation matrices. So, the repetition in \textit{Step 3} has no any influence for cryptanalysis of
ISEA. Without lose of generality, we just set the number of repetition times $T = 1$ in the following cryptanalysis.

Comparing Fig. 2a) and Fig. 2b), one can see that most visual information in Fig. 2a) are concealed by the scrambling operations. However, the correlations existing among the rows and columns of image are not changed. Given a row vector or a column vector of the cipher-image, its neighbouring vector may be found by searching for the vector owning the highest correlation index. Repeat the process iteratively in the horizontal and vertical directions, the approximate version of $B$ can be obtained. Actually, the vector search problem can be attributed to the binary template matching problem [12]. For simplicity, we select ratio of the same bits existing in two binary vectors, namely Sokal and Michner’s measure discussed in [12], as the similarity measure. For illustration, the image shown in Fig. 2d) is the obtained result from that in Fig. 2b) by the approach. Some visible image blocks can be automatically detected with the image quality metric defined in [13], where the image composed by the eight bit-planes of the image in Fig. 1a) in the weight order is used as the reference. Four cropped images are shown in Fig. 3 which reveals some important visual information, especially the rough sketch, of the original image shown in Fig. 1a). As shown in Fig. 3d), the vector order of the obtained result may be reversible to that of the right version since only relative locations of the vectors can be confirmed.

Figure 1: A pair of plain-image and the corresponding cipher-image: a) image “Lenna”; b) cipher-image of a).

3.2. Known-plaintext attack on ISEA

The known-plaintext attack is an attack model for cryptanalysis where the attacker know every implementation detail of the analyzed encryption algorithm (Shannon’s axiom) and can access to a set of plaintexts and the corresponding ciphertexts encrypted with the same unknown secret key. Because the two basic scrambling parts of ISEA are exerted in two orthogonal directions, its capability known-plaintext attack is worse than other position permutation-only encryption algorithm costing encryption complexity of the same magnitude.

As shown in [9], $O(\log_2(8MN))$ known plain-images can efficiently break ISEA if it is considered as a position scrambling encryption algorithm exerting on the domain of size $M \times 8N$ with permuted elements of two possible values. Essentially, ISEA is composed by cascading two basic permutation-only schemes of following parameters: permutation domain of size $M \times 1$ with permuted elements of $2^{8N} = 256^N$ possible values; permutation domain of size $1 \times (8N)$ with permuted elements of $2^M$ possible values. Due to the cascading, the two basic permutation-only schemes can not be broken separately in a direct way. However, some and even all parts of them can be recovered iteratively. Concrete approaches are described as follows.

- **Step 1**: Compare the number of 1’s elements in each row vector of $B'$ with that in all row vectors of $B$. If $B(j^*, :) = i^*$ is the unique row of $B$ owning the same number of 1’s elements as $B'(i^*, :)$, one can assure $T_M(j^*) = i^*$ and put $i^*$ into set $R$.

- **Step 2**: Similar to **Step 1**, compare the number of 1’s elements in each column vector of $B'$ with that in all column vectors of $B$. If the former is unique in the latter, the corresponding element in $T_N$ can be recovered correctly and put the column number into set $C$ (Initially, sets $R$ and $C$ are both empty).

- **Step 3**: Compare each column vector of $\{B'(i, :)\}_{i=1}^r$
with all column vectors of $\{B'(i,:)\}_{i=1}^{\ell}$, and add its column number into set $C$ if there exist the unique same column vector, where $\{B'(i,:)\}_{i=1}^{\ell}$ is recovered via Eq. (2) and $R = \{i_1, \ldots, i_r\}$.

- **Step 4**: Similar to **Step 3**, compare each row vector of $\{B(:,j)\}_{j=1}^{c}$ with all row vectors of $\{B(:,j)\}_{j=1}^{c}$ and add its row number into set $R$ if there exist the unique same row vector, where $\{B(:,j)\}_{j=1}^{c}$ is obtained via Eq. (3), and $C = \{j_1, \ldots, j_c\}$.

- **Step 5**: Iteratively repeat **Step 3** and **Step 4** till the sizes of sets $R$ and $C$ are both not increased (Only the new-found number is added into $R$ and $C$ in the above steps).

If more pairs of known plain-images and the corresponding cipher-images encrypted with the same secret key are available, the sets $R$ and $C$ can be further expanded with the above steps. The values of $T_M(i^*)$ and $T_N(j^*)$ can be assigned with the the index of the first vector owning the same measure (the number of 1’s elements in a vector or the vector itself) during the above comparisons, where $i^* \in (\mathbb{Z}_M - R)$ and $j^* \in (\mathbb{Z}_N - C)$. As demonstrated in [3, 9], more known plain-images can make the vectors owning the same comparison measure less, and $T_M(i^*)$ and $T_N(j^*)$ can be correctly guessed in higher probability. In addition, they can be further checked with the methods given in the previous subsection.

To verify real performance of the attack given in this sub-section, we illustrate it with three pairs of known plain-images of size $256 \times 256$, shown in Fig. 4 and the corresponding cipher-images encrypted with secret key $m = 20$, $n = 51$, $x_0 = 0.2009$, and $\mu = 3.98$. Using the known plain-image shown in Fig. 4(a), the sizes of $R$ and $C$ can reach 118 and 18 after **Step 2**, respectively. Note that the size of $B$ is $256 \times 2048$. The size of $R$ is further increased to 192 and 256 after the 1-th and 2-rd operation of **Step 4**, respectively. In contrast, the size of $C$ only increases to 1809, 1913, 1921 after the 1-th, 2-rd and 3-th operation of **Step 3**, respectively. As the 4-th operation of **Step 3** has no effect on increasing the size of $C$, we use the second plain-image shown in Fig. 4(b), which increase it to 2002. Its maximal value is obtained with the third plain-image shown in Fig. 4(b). The ratios between sizes of $R, C$ and their respective maximal values during the above attack processes are shown in Fig. 5.

Figure 4: Three pairs of plain-image and the corresponding cipher-image: a) “Fighter Jet”; b) “Candy”; c) “Baboon”.

![Figure 2](image_url)
3.3. Chosen-plaintext attack on ISEA

The chosen-plaintext attack is an enhanced version of the known-plaintext attack, where attacker can arbitrarily choose the plaintexts to optimize the breaking performance, e.g. reduce the number of needed plaintexts, and increase accuracy of the obtained information on the secret key. In [11], a set of specific plain-image and the corresponding cipher-images were selected to accurately obtain the scrambling relation with a small number of plain-images. Here, we further analyze the problem from a general viewpoint and proposed even more efficient chosen-plaintext attack on ISEA than [11].

If $M \leq 8N$, the equivalent secret key of ISEA can be recovered with the following steps:

- **Step 1**: Choose a plain-image satisfying $B = [B_L, B_R]$, where
  \[
  B_L = T_L \cdot \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & \cdots & 1 \\ \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{M \times M} \cdot T_R,
  \]  
  $T_L$ and $T_R$ are any permutation matrices of size $M \times M$, and $B_R$ is an $M \times (8N - M)$ matrix of fixed value zero or one.

- **Step 2**: As the number of $1$’s in every row of $B$ is different and the horizontal permutation has no influence on it, one can exactly recover the vertical permutation vector $T_M$ by comparing the number of $1$’s in each row of $B$ and that of $B'$.

- **Step 3**: Now, ISEA is degenerated as a permutation-only scheme exerted in domain of size $8N$ with permuted elements of $2^M$ possible values. Referring to

Eq. (5) in [8], one can assure that the $8N$ elements of vector $T_N$ can be exactly recovered with $\lceil \log_2(8N) \rceil$ chosen plain-images, where

\[
B_k(i, j) = \lfloor j/2^{M_k+i} \rfloor \mod 2,
\]

$[x]$ gives the largest integer less than or equal to $x$, $k \in \{0, 1, \cdots, \lceil \log_2(8N) \rceil - 1\}$, $i \in \{0, 1, \cdots, M - 1\}$, and $j \in \{0, 1, \cdots, 8N - 1\}$.

If $8N \in \{M, M + 1\}$, the plain-image with $B$ in [6] can be used to accurately confirm $M$ elements of $T_N$ with the same approach in Step 2. The sole remaining element of $T_N$ can be recovered also when $8N = M + 1$. So, the number of required chosen plain-images for breaking ISEA is $\lceil \log_2(8N) \rceil = \lceil \frac{1}{2} (3 + \log_2(N)) \rceil$ if $M < 8N - 1$, and one if $8N \in \{M, M + 1\}$. Similarly, we can deduce that the required number is $1 + \lceil \log_2(M) \rceil = \lceil \frac{1}{2} \log_2(M) \rceil$ if $M > 8N + 1$, and one if $M \in \{8N, 8N + 1\}$. In all, we can conclude that the number of required chosen plain-images for breaking ISEA is

\[
n = 1 + \begin{cases} 0 & \text{if } 8N \in \{M, M + 1\} \land M - 1; \\
\lceil \frac{1}{2} (3 + \log_2(N)) \rceil & \text{if } 8N > M + 1; \\
\lceil \frac{1}{2} \log_2(M) \rceil & \text{if } M > 8N + 1,
\end{cases}
\]

which is much smaller than the estimated number

\[
n' = \begin{cases} \lfloor 8N/M \rfloor + 1 & \text{if } M < N; \\
9 & \text{if } M = N; \\
\leq 9 & \text{if } 8N \geq M > N; \\
\lfloor M/8N \rfloor + 1 & \text{if } M > 8N,
\end{cases}
\]

given in [11] Sec. 5. Concretely, one can calculate $n = 2$ if $8N + 1 < M \leq 2^N$ or $M + 1 < 8N \leq 2^M$.

4. Conclusion

In this paper, security of a typical binary image scrambling encryption algorithm against ciphertext-only attack and known/chosen-plaintext attacks was studied comprehensively. Just as previous cryptanalytic works on the class of permutation-only encryption algorithms illustrated, secret scrambling is incapable of providing a sufficiently high level of security against known/chosen-plaintext attacks alone. Beside this, this cryptanalysis work demonstrate more problems on protecting multimedia data: 1) The correlation existing in multimedia data may be used to support some specific attacks and enhance breaking performance; 2) The size of each independent scrambling domain should be carefully checked to reach expected security requirement; 3) No matter what the permuted (scrambled) elements are, any permutation-only encryption algorithm does not change histogram on them.
Acknowledgement

This research was supported by the Hunan Provincial Natural Science Foundation of China (No. 2015JJ1013), Scientific Research Fund of Hunan Provincial Education Department (No. 15A186), and the National Natural Science Foundation of China (No. 61532020).

References

[1] S. Li, C. Li, K.-T. Lo, and G. Chen, “Cryptanalysis of an image scrambling scheme without bandwidth expansion,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 3, pp. 338–349, 2008.

[2] A. L. A. Dalhoum, B. A. Mahafzah, A. A. Awad, I. Aldamari, A. Ortega, and M. Alfonseca, “Digital image scrambling using 2D cellular automata,” *IEEE Multimedia*, vol. 19, no. 4, pp. 28–36, 2012.

[3] S. Li, C. Li, G. Chen, N. G. Bourbakis, and K.-T. Lo, “A general quantitative cryptanalysis of permutation-only multimedia ciphers against plaintext attacks,” *Signal Processing: Image Communication*, vol. 23, no. 3, pp. 212–223, 2009.

[4] W. Zeng and S. Lei, “Efficient frequency domain selective scrambling of digital video,” *IEEE Transactions on Multimedia*, vol. 5, no. 1, pp. 118–129, 2003.

[5] C. Li, S. Li, D. Zhang, and G. Chen, “Cryptanalysis of a data security protection scheme for VoIP,” *IEEE Proceedings–Vision Image and Signal Processing*, vol. 153, no. 1, pp. 1–10, 2006.

[6] H. Liu, S. Cao, and S. Yan, “Automated assembly of shredded pieces from multiple photos,” *IEEE Transactions on Multimedia*, vol. 13, no. 5, pp. 1154–1162, 2011.

[7] W. Li, Y. Yan, and N. Yu, “Breaking row-column shuffle based image cipher,” in *Proceedings of ACM International Conference on Multimedia*, 2012, p. art. no. 6011939.

[8] C. Li, “Cracking a hierarchical chaotic image encryption algorithm based on permutation,” *Signal Processing*, vol. 118, pp. 203–210, 2016.

[9] C. Li and K.-T. Lo, “Optimal quantitative cryptanalysis of permutation-only multimedia ciphers against plaintext attacks,” *Signal Processing*, vol. 91, no. 4, pp. 949–954, 2011.

[10] G. Ye, “Image scrambling encryption algorithm of pixel bit based on chaos map,” *Pattern Recognition Letters*, vol. 31, no. 5, pp. 347–354, 2010.

[11] L. Zhao, A. Adhikari, D. Xiao, and K. Sakurai, “On the security analysis of an image scrambling encryption of pixel bit and its improved scheme based on self-correlation encryption,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 8, pp. 3303–3327, 2012.

[12] J. D. Tubbs, “A note on binary template matching,” *Pattern Recognition*, vol. 22, no. 4, pp. 359–365, 1989.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.