*Journal of Information Hiding and Multimedia Signal Processing* ©2015 ISSN 2073-4212
**Ubiquitous International** Volume **6**, Number **4**, July **2015**

# Comparison of Two Kinds of Image Scrambling Methods Based on LSB Steganalysis

Chun-Yu Zhang, Wen-Xiang Zhang

School of Information Engineering
Xizang Minzu University
6 Wenhui Road, Xianyang, 712082, China
zcy@xzmy.edu.cn;dino.z@foxmail.com

Shao-Wei Weng

School of Information Engineering
Guangdong University of Technology
100 Waihuanxi Road, Guangzhou HEMC, Guangdong, 510006, China
wswweiwei@126.com

Received January, 2014 ; revised January, 2015

Abstract. *According to the research of LSB steganalysis, LSB steganography could destroy the distribution of 0 and 1 in the lowest bit planes, and this situation will cause the abnormality of histogram and asymmetry statistical, thereby left an opportunity that could be exploited to the steganalysis. From the view of steganalysis, this paper compared the two typical kinds of image scrambling methods which are the Arnold transform from the position-based scrambling and the Logistic chaotic sequence from the gray values of pixels based scrambling, studied the performance of the two typical methods which are used in LSB steganography. The theoretic and experimental results show that the Gray values of pixels based scrambling could effectively improve the histogram abnormality and statistical asymmetry, which is better than the Position-based scrambling.*

**Keywords:** LSB, Steganalysis, Position transforming, Gray values of pixels transform, Image scrambling

1. **Introduction.** As the important part of information hiding technology, Steganography and Steganalysis both are the researching hotspot of field of information security in recent years[1].Steganography aims to covert communication by hiding secret information in the carrier to transfer without causing third-party doubt. As the reverse analysis technology, Steganalysis aims to detect, extract, revert or destruct the secret information hided in the carrier [2].

LSB (Least Significant Bits) embeds secret information into the least important bits of the carrier randomly. As a typical spatial steganography algorithm, the robustness and anti attack ability are worse, but all the advantages like good transparency, simply algorithm, fast speed, large hiding capacity are difficultly to reach for the transform domain algorithms. Therefor currently most of the information hiding software are still using the LSB technology as the main algorithm. Just because of its importance, more and more steganalysis algorithms aimed at LSB ware proposed. Anti steganography field studies have found that even through the LSB replacement algorithm makes small changes to the carrier, it still leaded to some abnormal features in the lowest bit plane that are easy to discover. First, because the lowest bit plane is not always uniform random distribution of 0 and 1, and in some areas it will show some structure which connecting with the content. Unfortunately, LSB replacement will destroy the correlation. Second, the simple LSB replacement will cause statistical asymmetry and provide clues for analysts[3]. Some studies told that the spatial embedding in the JPEG images would change the statistical framework caused by quantization[4]. LSB steganography in DCT quantization coefficient could increase the block effect, so analysts could estimate the DCT coefficient histogram by constructing reference images [5]. In short, the LSB replacement leads to Histogram abnormality and statistical asymmetry and leaves an opportunity for analysts.

These research results on the fields of steganalysis also show that the scrambling algorithms are very important in LSB. Because the scrambling algorithms not only have the effect of encrypting the secret images, but also need to minimize the changes of the stego-images and make it difficult to be detected. Image scrambling is a kind of image encryption methods, the main purpose is to make people can't find out the real meaning which expressed by the image through the human visual system and computer system. In the applications, the image scrambling is often used in the information-hiding algorithm to preprocess the cover image or the secret image before embedding. In the literature[6], the method of scrambling the cover image and then embedding the secret image into is very rare. Differently, most of algorithms scrambled the secret images and then embedded them into the cover images aimed for stronger robustness. Therefore, the good image scrambling method has become one of the most important subjects in the field of information security.

Based on the LSB steganalysis, this paper talked about the typical position-based scrambling and the gray values of pixels based scrambling, for instance, the Arnold transform and the Logistic chaotic sequence. Respectively using the two methods to preprocess the secret image and then used to LSB algorithms, aimed to discuss the performance of the two methods in the histogram abnormality.

## 2. Comparison of Two Kinds Image Scrambling Methods.

There are two kinds of image scrambling methods in spatial domain. One is based on the position transform, another is based on the transform of the gray values of pixels. Among them, the Arnold transform and the Logistic chaotic sequence are extremely widespread[7].

### 2.1. The Position Transforming Based Scrambling: Arnold Transform.

Given the two variables function $Z = F(x, y)$ is a two dimensional digital image in the planar domain $D$, where $(x, y)$ is the coordinate of any pixel in $D$, that is $(x, y) \in D$. $F(x, y)$ represent the image information, for example, if $Z$ is a gray image, then $F(x, y)$ is gray level. If $Z$ is color image, then $F(x, y)$ is RGB component value.

For an image with size $n \times n$, let $(x, y)$ is the coordinates of any pixel, its Arnold transform is as follows[8]:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} (mod\ n) \tag{1}$$

Where $(x', y')$ is the new coordinate after transforming for the point $(x, y)$. Repeatedly doing this transform, the image $Z$ can be scrambled.

The Arnold transformation has periodicity, different size of image after a certain number of iterations can be restored to the original image. As shown in formula (1), apparently, it is only change the position of the original pixel in the image instead of changing the values of the original pixel.

Similar to the Arnold transform, all position-based scrambling methods only change the position of pixels, redistribute all pixels, and do not change the values of pixels. Therefore, using them to scramble the secret images cannot change the statistical features of the secret images. And then using them to LSB steganography, the result is the statistical features of the stego-images were so different from the features of the cover images, because the secret images embedded still remain their own features. In other words, using this kind of scrambling method cannot improve the problems of histogram abnormality and statistical asymmetry of the stego-images. That is the handle left to steganalysis.

2.2. **The Gray Values Transforming Based Scrambling: Logistic Chaotic Sequence.** Chaotic sequence is produced by a deterministic equation, as long as the parameter equation and initial values are determined, which can reproduce the chaos phenomenon. Its advantages are sensitive to initial value extremely, sequence length variable, aperiodic, easy to produce and copy[9]. So, chaotic sequence is good at secrecy and easy management. From a statistics standpoint, the statistical properties of chaotic sequence close to the Gauss white noise with great randomness.

Given the Logistic mapping equation as follow:

$$x_{n+1} = \mu x_n(1 - x_n) \tag{2}$$

where $x_n \in (0, 1)$, $\mu$ is the bifurcation parameter. According to different values of $\mu$, the sequence is periodic or chaotic. When $0 < \mu \leq 3.57$ , $x_{n+1}$ is a periodic function with $2^m$ cycle. When $3.57 < \mu \leq 4$ , the sequence is aperiodic and non-convergence, the Logistic mapping works in chaos, its autocorrelation function approximately equal to function $\delta$ and its cross-correlation is 0. Logistic chaotic sequence is very sensitive to initial value $x_0$ , two Logistic chaotic sequences produced by two different initial values are aperiodic, non-convergence and uncorrelated.

When this method is used to scramble the binary image, the steps supposed to be: Firstly, produce a Logistic chaotic sequence, which length equal to the numbers of pixels of the secret image. Second, the sequence is binarized and mapped to a binary matrix which size equal to the size of the secret image also. Finally, do XOR operation between the logistic binary matrix and the secret image.

Obviously, this scrambling method changed the values of pixels in the secret image. For the binary secret image, this method changed the distribution of 0 and 1, and made it more uniform because of the randomness of the logistic chaotic sequence. So using this method into LSB steganography can largely decrease the effect to the distribution of 0 and 1 in the lowest bit plane of the stego-image, and therefore, improve the histogram abnormality and statistical asymmetry of stego-image.

2.3. **Experiments and Analysis.** The secret image was scrambled by Arnold transform and Logistic chaotic sequence, respectively, and then compared the histogram of the secret image before and after scrambling. The result is shown in figure 1 and figure 2. Figure 1 shows the result of the Arnold transform. Clearly, the distribution of 0 and 1 is no changed. But in Figure 2, the distribution is changed more uniform.

3. **The LSB steganography algorithm based on Logistic chaotic sequence.** The randomness of Logistic chaotic sequence determines that it can make the distribution of 0 and 1 of the secret image more uniform and more random. Therefore, when the secret image which was scrambled by Logistic chaotic sequence was embedded into the lowest bit plane of the cover image, the histogram abnormality of the stego-image would be slighter, and this method will improve the problems of statistical asymmetric of the stego-image. In this paper, the LSB steganography algorithm based on Logistic chaotic sequence was presented and the detail step as follow:

Step1: generating the binary Logitstic chaotic matrix.

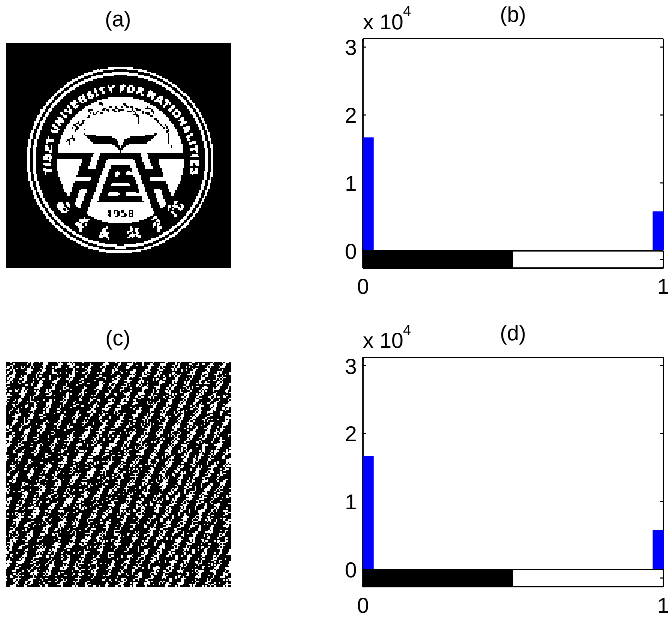

FIGURE 1. Arnold transform ((a) the secret image, (b) histogram of (a), (c) the scrambling secret image, (d) histogram of (c))

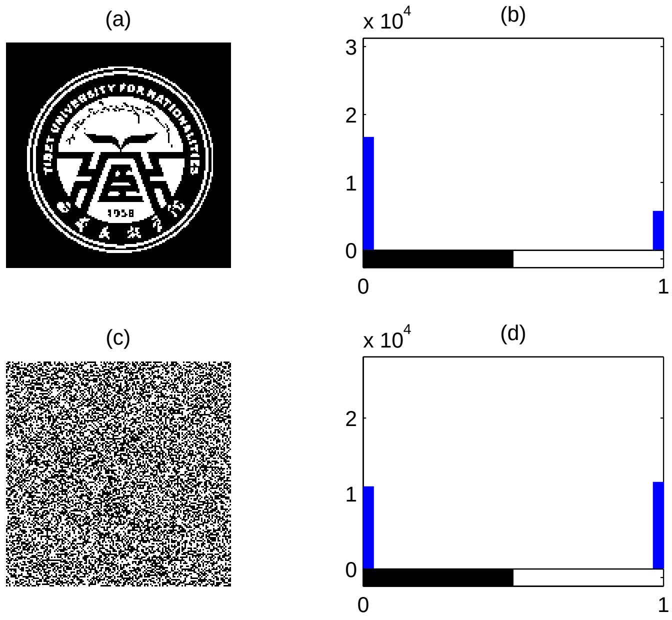

FIGURE 2. Logistic chaotic sequence transform ((a) the secret image, (b) histogram of (a), (c) the scrambling secret image, (d) histogram of (c))

Given $W = \{w(i,j)|w(i,j) \in \{0,1\}, i \in N, j \in N, 1 \leq i \leq m, 1 \leq j \leq n\}$is the binary secret image with size $m \times n$. With the key $key = (x_0, \mu)$, where $x_0$ is the initial, $\mu$ is the bifurcation parameter, to generate a Logistic chaotic sequence $X = \{x(k)|x(k) = \mu x(k-1)(1 - x(k-1)), x(0) = x_0, k \in N, 1 \leq k \leq m \times n\}$.

let $mean = \sum_{k=1}^{m\times n} x(k)/(m\times n)$ , the binary Logistic chaotic sequence is $L = \{l(k)|l(k) \in \{0,1\}, k \in N, 1 \leq k \leq m \times n\}$, where the binarization rule is as follow: if $x(k) \geq mean$ , then $l(k) = 1$; else $l(k) = 0$.

Lastly, Mapping $L$ into a two-dimensional binary matrix $L_M = \{l_M(i,j)|l_M(i,j) \in \{0,1\}, i \in N, j \in N, 1 \leq i \leq m, 1 \leq j \leq n\}$.

Step2: scrambling the secret image.

Let $W' = \{w'(i,j)|w'(i,j) = w(i,j) \oplus l_M(i,j), i \in N, j \in N, 1 \leq i \leq m, 1 \leq j \leq n\}$ be the scrambled secret image.

Step3: embedding the scrambled secret image.

Let $C = \{c(s,t)|c(s,t) \in N, 0 \leq c(s,t) \leq 255, s \in N, t \in N, 1 \leq s \leq p, 1 \leq t \leq q\}$ be the gray cover image with size $p \times q$. Its $k$th bit plane is $B_k = \{b_k(s,t)|b_k(s,t) \in \{0,1\}, 1 \leq s \leq p, 1 \leq t \leq q\}$, where $k \in \{0,1,2,3,4,5,6,7\}$. Then, to replace the bit plane $B_0$ with $W'$, the stego-image is $C' = \sum\limits_{k=1}^{7} B_k * 2^k + W'$.

4. **The comparison of LSB steganography algorithms based the two scrambling methods.** For the scrambled secret image, when is embedded into the cover image, the more uniform and random its distribution of 0 and 1 is, the less the histogram abnormality of stego-image is.

Using the same secret image and cover image, the LSB steganography algorithm based on the Arnold transform scrambling was compared with the algorithm based on the Logistic chaotic sequence scrambling. The results were shown in figure 3 and figure 4. Figure 3 shows the result of LSB steganography based on the Arnold transform, where (b) is the histogram of the original cover image (a), (d) is the histogram of the stego-image (c). Comparing (b) and (d), it is clear that the difference before and after embedding, because the secret image embedded has sharply changed the distribution of gray level of the cover image. This is the histogram abnormality and statistical asymmetry of stego-image. Figure 4 shows the result of LSB steganography based on the Logistic chaotic sequence transforms, where (b) is the histogram of the original cover image (a), (d) is the histogram of the stego-image (c). Comparing (b) and (d), there is almost no difference, because the secret image embedded has little effect to the distribution of gray level of the cover image.

As shown in figure 5, for more intuitively showing the difference between two methods, this paper counted the values of all pixels for the original cover image and the stego-image which respectively based on the Arnold transform and the Logistic transform, and then plotted the line chart with blue, green and red respectively. Clearly, the blue and the red are almost coincides, but the green is very different with them.

Zhang Jun[10] presented that the local extremum of the histogram of the stego-image will be affected by the LSB steganography and leave chance for steganalysis. This paper respectively counted the local extreme points of the histogram of the original cover image and the stego-image which based on the Arnold method and the Logistic method. As shown in the figure 6, the Arnold method is very different with others. Using the same algorithm in this paper, the extreme points of the Arnold method increased 37, but the Logistic method just increased 5. Based on the NRCS image library, 100 gray images with $256 * 256$ was randomly selected to be as the cover images, and 10 binary image with $150 * 150$ was selected to be as the secret images randomly. The experiments were executed respectively using them based on the two methods. Experimental results show

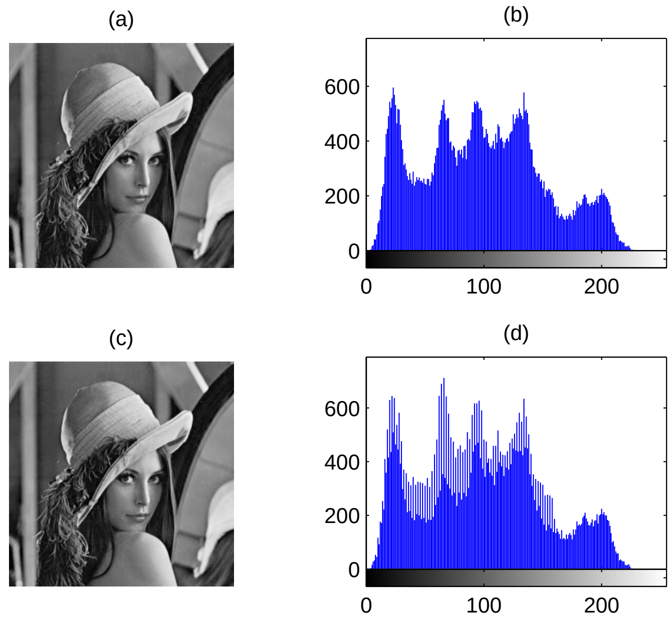

Figure 3. LSB steganography based on the Arnold transform ((a) the cover image, (b) histogram of (a), (c) the stego-image, (d) histogram of (c))

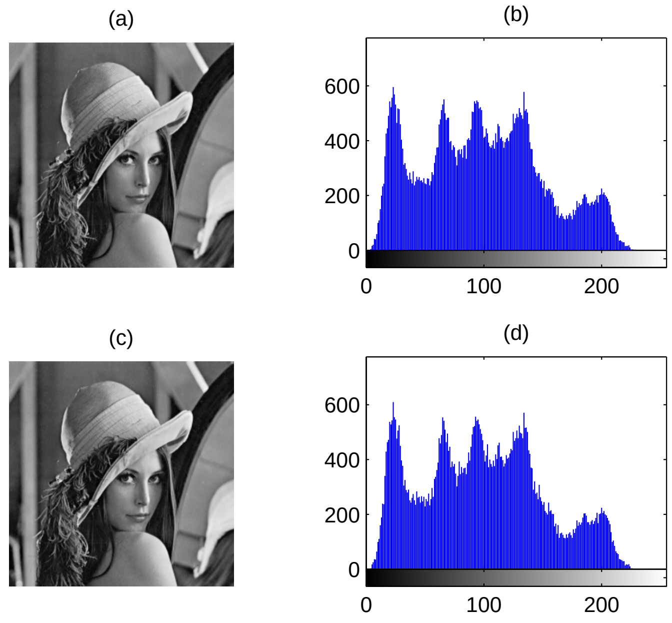

Figure 4. LSB steganography based on the Logistic chaotic sequence transform ((a) the cover image, (b) histogram of (a), (c) the stego-image, (d) histogram of (c))

that the average increment of the Arnold method was 41.7, while the Logistic method was 7.3.

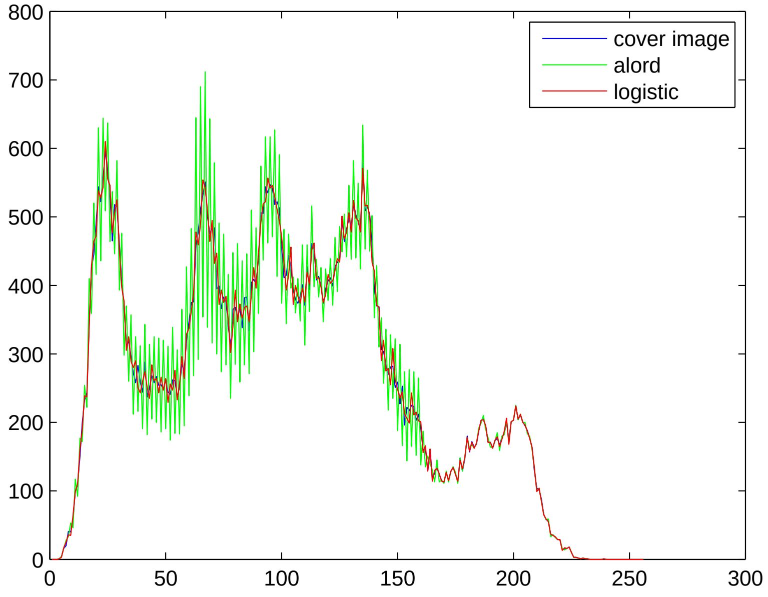

Figure 5. comparison of the statistical results of pixel values

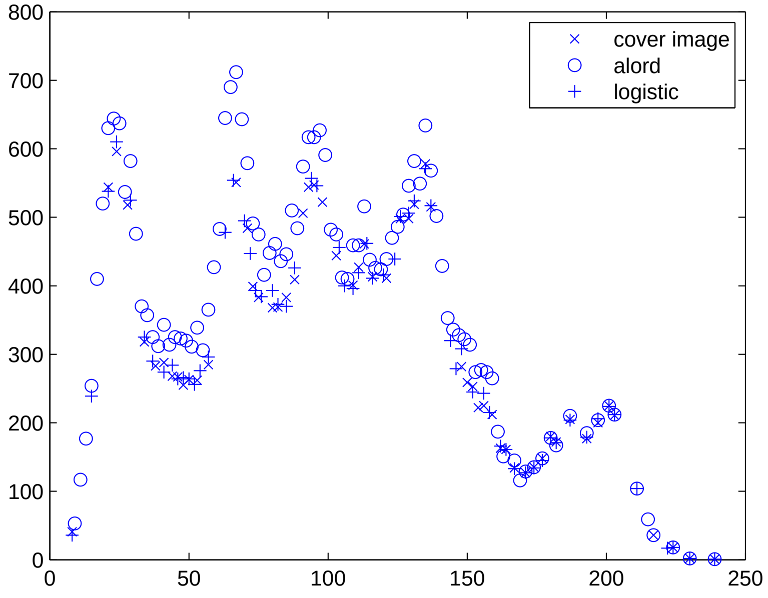

Figure 6. comparison of the local extreme points of histogram

5. **Conclusions.** The LSB replacement leads to histogram abnormality and statistical asymmetry and leaves the chance for steganalysis. Firstly, this paper told about two scrambling methods which are the position-based transform scrambling with an example from the Arnold transform and the gray value of pixel based transform scrambling with an example from the Logistic chaotic sequence, and discussed the results of 0 and 1 distribution by using two methods to scramble the secret image. Secondly, The LSB steganography algorithm based on Logistic chaotic sequence was presented. Thirdly,

from the steganalysis, respectively used two methods into the LSB steganography, compared and analyzed the histogram abnormality of stego-images and the problem of local extremum increase. The experimental results show that the gray value of pixel based transform scrambling is good at improving histogram abnormality, and is more effective at resisting attack of LSB steganalysis.

**Acknowledgment.** This work was partially supported by National NSF of China (No. 61201393) and the Natural Science Foundation of Tibet (No. 2015ZR-14-17).

## REFERENCES

[1] S. Z. Wang, X. P. Zhang, K. W. Zhang, *Digital Steganography and Steganalysis-Information Warfare Technology in the Internet Age*, Tsinghua University Press, Beijing, China, pp. 205–212, 2005. (in Chinese)

[2] J. Zhang, F. Xiong, D. Zhang, Image Steganalysis Technology Overview, *Computer Engineering*, vol. 39, no. 4, pp. 165–172, 2013. (in Chinese)

[3] S. Z. Wang, X. P. Zhang, W. M. Zhang, Recent Advances in Image-based Steganalysis Research, *Chinese Journal of Computer*, vol. 32, no. 7, pp.21–23, 2009. (in Chinese)

[4] J. Fridrich , M. Goljan, and R. Du, Steganalysis based on JPEG compatibility, *SPIE Multimedia Systems and Applications IV*, Denver, CO, USA, pp. 275–280, 2001.

[5] J. Fridrich, M. Goljan and R. Du, Attacking the OutGuess, *Proceedings of the ACM Workshop on Multimedia and Security*, Juan-les-Pins, France, pp.239–242, 2002.

[6] L. F. Zhu, B. Wang, Image Information Hiding Scheme Based on New Anti-Arnold Transformation, *Application Research of Computer*, vol. 26, no. 5, pp. 1926–1928, 2009. (in Chinese)

[7] L. P. Shao, Q. Z. Meng, C. M. Li, Overview of Image Scrambling Transformation, *Netinfo Security*, no. 4, pp. 22–24, 2009. (in Chinese)

[8] W. Ding, W. Q. Yan, Digital Image Scrambling Technology Based on the Arnold Transformation, *Journal of Computer-Aided Design & Computer Graphics*, vol. 13, no. 4, pp. 338–341, 2001. (in Chinese)

[9] B. Q. Shu, *A Study on Chaotic sequence and its application in Spread Spectrum Communication*, M.S.Thesis, Southwest Jiaotong University, Chengdu, China, 2007. (in Chinese)

[10] J. Zhang, steganalsis method for 1 Steganography, *Computer Engineering and Application*, vol. 44, pp.58–60, 2008. (in Chinese)