A Fast Pseudo Stochastic Sequence Quantification Algorithm Based on Chebyshev Map and Its Application in Data Encryption

Chong Fu¹, Pei-rong Wang¹, Xi-min Ma², Zhe Xu¹, and Wei-yong Zhu¹

¹ School of Information Science and Engineering, Northeastern University,
Shenyang 110004, China
fu_chong@sohu.com

² Computer Department, Liaoning Economic Vocational Technological Institute,
Shenyang 110036, China

Abstract. Chaos theory has been widely used in cryptography fields in recent years and the performance of the pseudo stochastic sequence quantified from chaos map has great influence on the efficiency and security of an encryption system. In this paper, an improved stochastic middle multi-bits quantification algorithm based on Chebyshev map is proposed to enhance the ability of anti reconstruction a chaos system through reverse iteration and improve the performance of the generated sequence under precision restricted condition. The balance and correlation properties of the generated sequence are analyzed. The sequence is proved to be a binary Bernoulli sequence and the distribution of the differences between the amounts of 0 and 1 is analyzed. The side lobes of auto correlation and values of cross correlation are proved to obey normal distribution $N(0, 1/N)$.

1 Introduction

Chaos theory has been widely used in encryption fields nowadays due to its initial value sensitive, nonperiodic, unpredictable, and Gauss like statistical characteristics. The raw generated analog chaotic sequence could not be directly used in digital system and the binary quantification process has a great influence on system security and performance. The classic quantification method mainly has the following two disadvantages: (1) Only one bit can be generated per iteration, the computation load is very heavy. For example, to encrypt a plaintext of 1MB size, we must iterate at least 8×2^{20} times; (2) With the increasing of iteration, the generated raw sequence would be periodic because no processor is precision unrestricted. This will cause the quantified sequence also being periodic and making the auto and cross correlation performance worse [1]. Some scholars proposed a multi-bits quantification algorithm to reduce the computation load and extend the period of the sequence by transforming the analog value of each iteration to its binary form directly, but the security of this algorithm is relatively weak, the analog sequence can be recovered from the quantified binary sequence and the chaotic system can be reconstructed through reverse iteration [2]. In this paper, an improved stochastic middle multi-bits quantification algorithm based on Chebyshev map is proposed to strengthen the ability of anti reverse iteration and improve the performance of the generated sequence under precision restricted condition.

2 The Statistical Properties of Chebyshev Map

Chebyshev map is defined as:

$$x_{n+1} = \cos(k \cos^{-1}(x_n)), \quad x_n \in [-1, 1]. \quad (1)$$

The probability density of chaotic sequence generated by Eq.1 is [3]:

$$\rho(x) = \begin{cases} \frac{1}{\pi\sqrt{1-x^2}} & -1 \leq x \leq 1; \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Let $\{x_i\}$ be chaotic sequence generated by Eq.1, we can get the following three properties based on Eq.2.

Property 1. The mean of $\{x_i\}$ is:

$$\bar{x} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} x_i = \int_{-1}^1 x \rho(x) dx = 0. \quad (3)$$

Property 2. The normalized auto correlation function of $\{x_i\}$ is:

$$AC(m) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \bar{x})(x_{i+m} - \bar{x}) = \int_{-1}^1 x f^m(x) \rho(x) dx - \bar{x}^2 = \begin{cases} 0.5 & m = 0; \\ 0 & m \neq 0. \end{cases} \quad (4)$$

Property 3. For any two different initial values x_{01} and x_{02} , the normalized cross correlation function of the two generated sequences is:

$$\begin{aligned} CC_{12}(m) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} (x_{i1} - \bar{x})(x_{i+m2} - \bar{x}) \\ &= \int_{-1}^1 \int_{-1}^1 x_1 f^m(x_2) \rho(x_1) \rho(x_2) dx_1 dx_2 - \bar{x}^2 = 0. \end{aligned} \quad (5)$$

Property 2 and 3 indicates that Chebyshev map is excellent to be used to generate pseudo stochastic sequence.

3 Stochastic Middle Multi-bits Quantification Algorithm and Its Performance Analysis

The stochastic middle multi-bits quantification algorithm is as follows:

- ① Let $k = 4$, selecting a proper initial value x_0 , getting the real value x_i from x_{i-1} by iterating Eq.1, $x_i \in [-1, 1]$;
- ② Let $y_i = (x_i + 1)$, so $y_i \in [0, 1]$;
- ③ Transform y_i to binary form: $(0.a_1a_2a_3\dots a_n)_2$, $a_i = 0$ or 1 , getting the multi-bits $a_1a_2a_3\dots a_n$, n is equal or less than the precision that the processor can provide;
- ④ Let $y_i' \in [0, 1]$, whose decimal part is the reverse order of that of y_i , as another initial value, calculating the real value y_i' from y_{i-1}' by iterating Eq.1, $y_i' \in [0, 1]$;

- ⑤ Let the number of bits we want to get by each iteration be l ($l < n$), $m = \lceil (y_i') \times (n - l) \rceil$, m is a stochastic integer between 0 and $(n - l)$ thus;
- ⑥ Skip first m bits of $a_1 a_2 a_3 \dots a_n$, getting the middle multi-bits $a_{m+1} a_{m+2} \dots a_{m+l}$.

The rest may be deduced by analogy, concatenating all the l bits generated by each iterating together, thus getting the pseudo stochastic binary sequence with the length we needed.

Because the position m is determined by chaotic map and also stochastic, the algorithm not only has the advantages such as reducing the computation load by generating l bits per iteration and extending the period of the sequence under precision restricted condition that the original multi-bits algorithm has, but also further strengthen the ability of anti reconstruction of the chaos system. The balance and correlation performance are the two most import properties of stochastic sequence, which greatly affects the security and key space of an encryption system, and will be analyzed in the following.

Lemma 1. The pseudo stochastic sequence generated by stochastic middle multi-bits algorithm is a binary Bernoulli distributed sequence. For sequence with length N , the mean of the differences is 0 and the variance is N .

Lemma 2. For any two different sequence $\{x_{i1}\}$ and $\{x_{i2}\}$ that generated by stochastic middle multi-bits algorithm, they satisfy: When N is large enough and m is relatively small, the side lobes of auto correlation and values of cross correlation obey normal distribution $N(0, 1/N)$.

Fig. 1. The balance and auto/cross correlation performance of generated sequence. (a) balance performance, (b) auto correlation function, (c) cross correlation function, (d) distribution of cross correlation values.

The simulation result of balance property is shown in Fig. 1(a). 1000 sequences with length $N=8192$ are generated and the initial values are selected independently. The mean of differences is -0.9480 and the standard deviation is 92.3074, which are in good accordance with theoretical value 0 and 90.5097. The simulation results of the auto and cross correlation functions and the distribution of cross correlation values are shown in Fig. 1(b) to (d). The sequence length is 8192 and the initial values are selected as 0.60000 and 0.60001 for verify the initial value sensitive property of a chaotic system. The maximum absolute value of auto correlation side lobe is 0.0410 and the value of cross correlation is 0.0905, which are in good accordance with theoretical value 0. Above analysis indicate that the quantified sequences have excellent balance and auto/cross correlation performance.

4 Experimental Results and Conclusions

The comparison of different quantification algorithms is shown in Table 1. The plain-text to be encrypted is 1MB size and parameter l is set to 8. From Table 1 we can see, the efficiency and the sequence performance of the stochastic middle multi-bits quantification algorithm is far better than that of classic binary quantification algorithm. While compared with the multi-bits algorithm, although the encryption time is longer due to it's more complicated algorithm, but the security is high due to its powerful anti reconstruction ability. The algorithm proposed in this paper can be widely used in signal generators for generating high quality PN sequence [4], which can also be used in communication channel simulation, spread spectrum communication, etc.

Table 1. The performance comparison of sequences generated by different quantification algorithms

Algorithms	$\sqrt{\frac{1}{N} \sum_{m=1}^N [AC(m)]^2}$	$\sqrt{\frac{1}{N} \sum_{m=1}^N [CC_{12}(m)]^2}$	Encryption time (S)	Security
Binary	0.0803	0.0247	58.76	High
Multi-bits	0.0781	0.0226	31.72	Low
Stochastic multi	0.0783	0.0223	38.45	High

References

1. Heidari-Bateni, G., McGillem, C. D.: A Chaotic Direct Sequence Spread Spectrum Communication System. IEEE Transaction on Communication, Vol. 42 (1994), 1524-1527
2. Baptista, M. S.: Cryptography with Chaos. Physics Letters A, Vol. 240 (1998), 50-54
3. Geisel, T., Fahren, V.: Statistical Properties of Chaos in Chebyshev Maps. Physics Letters A, Vol. 105 (1984), 263-266
4. Lipton, J. M., Dabke, K. P.: Spread Spectrum Communications Based on Chaotic Systems. International J. Bifurcation and Chaos, Vol. 6 (1996), 2361-2374