Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105 (Online)
Volume **6**, Number **1**, February **2021**

Improvement and Implementation of SM4 Algorithm Based on FPGA

Zijing Jiang

Electronic Engineering College
Heilongjiang University
Harbin 150080, China

Longfei Yu

Electronic Engineering College
Heilongjiang University
Harbin 150080, China

Xinyue Tang

Electronic Engineering College
Heilongjiang University
Harbin 150080, China

Renxiu Zhang

Electronic Engineering College
Heilongjiang University
Harbin 150080, China

Wei Ding

Electronic Engineering College
Heilongjiang University
Harbin 150080, China

Qun Ding*

Electronic Engineering College
Heilongjiang University
Harbin 150080, China
* Correspondence: qunding@aliyun.com

Received November 2020; revised December 2020

Abstract. *Sbox is the only nonlinear element in SM4, which largely determines the safety of SM4. As we all know, the Sbox in SM4 is fixed and vulnerable to static Sbox attacks, such as linear cryptanalysis, differential cryptanalysis and CPA attacks. Therefore, we improve SM4 algorithm by introducing Logistic chaotic system to construct dynamic Sbox. The key space, avalanche effect and information entropy of SM4 are compared and analyzed. The results show that the application of dynamic Sbox framework in SM4 improves the quality and security of encryption.*
Keywords: Avalanche effect, Chaotic, Sbox, FPGA, Information entropy, SM4.

1. **Introduction.** Block cipher is to divide plaintext into fixed length blocks, and then convert each plaintext block into ciphertext block of equal length under the effect of key. With the development of cryptographic standards in the world, the design and analysis of cryptographic algorithms are also more and more concerned. Therefore, the State Encryption administration announced SM4 [1]. SM4 algorithm is a block cipher algorithm independently researched and designed in China. It is an encryption algorithm used in WAPI wireless network standard. This algorithm is suitable for the security field of WLAN. Sbox [2–5] is a key module in SM4. By improving the design level of Sbox, we can get high reliability password. In SM4, algebraic method is used to construct Sboxes, which can obtain high nonlinearity, but because of its simple structure and weak differential performance, it can not resist algebraic attacks. Due to the advantages of ergodicity, mixing and sensitivity to initial conditions and parameters, chaotic systems [6–15] can be used to design more ideal Sboxes. Therefore, the research of Sbox based on chaotic system has attracted the attention of researchers in the field of information security. Using the excellent characteristics of chaotic system to construct a new Sbox has rapidly developed into a hot research direction in the field of information security. The idea of using chaos to construct dynamic Sboxes first appeared in the research of Urias et al. [16] in 1998. In recent years, a lot of achievements have been made in the research based on chaotic dynamic Sboxes. In 2016, Fan and Yang [17] proposed a method to construct dynamic Sboxes based on Logistic chaotic map and tent chaotic map. After testing and analyzing, the dynamic Sbox generated by the algorithm can meet the security of cryptographic algorithm. In 2018, Zhu et al. [18] proposed a method to generate new pseudo-random sequence based on multi chaotic system and composite idea, which was applied to the construction process of Sbox. In 2020, Han et al. [19] superposed tent and Henon chaotic maps, and introduced the idea of image scrambling into the design of Sbox, which enhanced the randomness of key sequence of ZUC algorithm.

This paper mainly studies the construction of chaotic dynamic Sbox and its hardware implementation in SM4. After establishing Sbox comparison table through a set of nonlinear mapping, the original Sbox arrangement of SM4 is fixed and cannot be changed. The key expansion and loop functions in SM4 encryption are calculated by a set of Sboxes. If fixed on the hardware, it is easy to cause the problem that the Sbox cannot be updated and the original Sbox is damaged. In addition, ciphertext can be obtained from plaintext. The more choices of permutations, the higher the encryption strength. Because of its simple structure, easy implementation and less resource consumption, this paper improves SM4 algorithm by introducing Logistic chaotic system to construct dynamic Sbox. The key space, avalanche effect and information entropy of SM4 are compared and analyzed. The results show that the application of dynamic Sbox framework in SM4 improves the quality and security of encryption.

The second section of this paper mainly introduces the basic knowledge, including SM4 structure framework and algorithm, Logistic discrete chaotic system. The third section introduces the realization of Logistic sequence generator and the improved SM4 algorithm based on Logistic dynamic Sbox. The fourth section makes a comparative analysis of resource consumption and key space.The fifth section analyzes the security of the algorithm, including avalanche effect and information entropy. Finally, we summarize the paper in the sixth section.

2. Basic Knowledge.

2.1. SM4 Algorithm Description.

SM4 block cipher algorithm mainly includes encryption algorithm, decryption algorithm and key expansion algorithm. The encryption

key in the algorithm is 128bit, which is expressed as $MK = (MK_0, MK_1, MK_2, MK_3)$ in the algorithm, while the round key in the algorithm is generated by the key in the encryption algorithm, which is expressed as $(rk_0, rk_1, ..., rk_{31})$. $FK = (FK_0, FK_1, FK_2, FK_3)$ is the system parameter, are fixed parameters, $CK = (CK_0, CK_1, ..., CK_{31})$ which are mainly used in key extension algorithm.

The encryption algorithm flow of SM4 block cipher algorithm includes 32 iterative operations and 1 reverse order transformation r to get ciphertext output, as shown in Figure 1. T is a reversible transformation composed of a nonlinear transformation τ and a linear transformation L . The nonlinear transformation τ is composed of four Sboxes in parallel. L is a linear left shift transformation, that is, $T(.) = L(\tau(.))$, as shown in Figure 1. If the input is $(X_i, X_{i+1}, X_{i+2}, X_{i+3}), i = 0, 1, ..., 31$,then the round function is:

$$X_{i+4} = F(X_i, X_{i+1}, X_{i+2}, X_{i+3}, rk_i) = X_i \oplus T(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus rk_i), i = 0, 1, ..., 32 \quad (1)$$

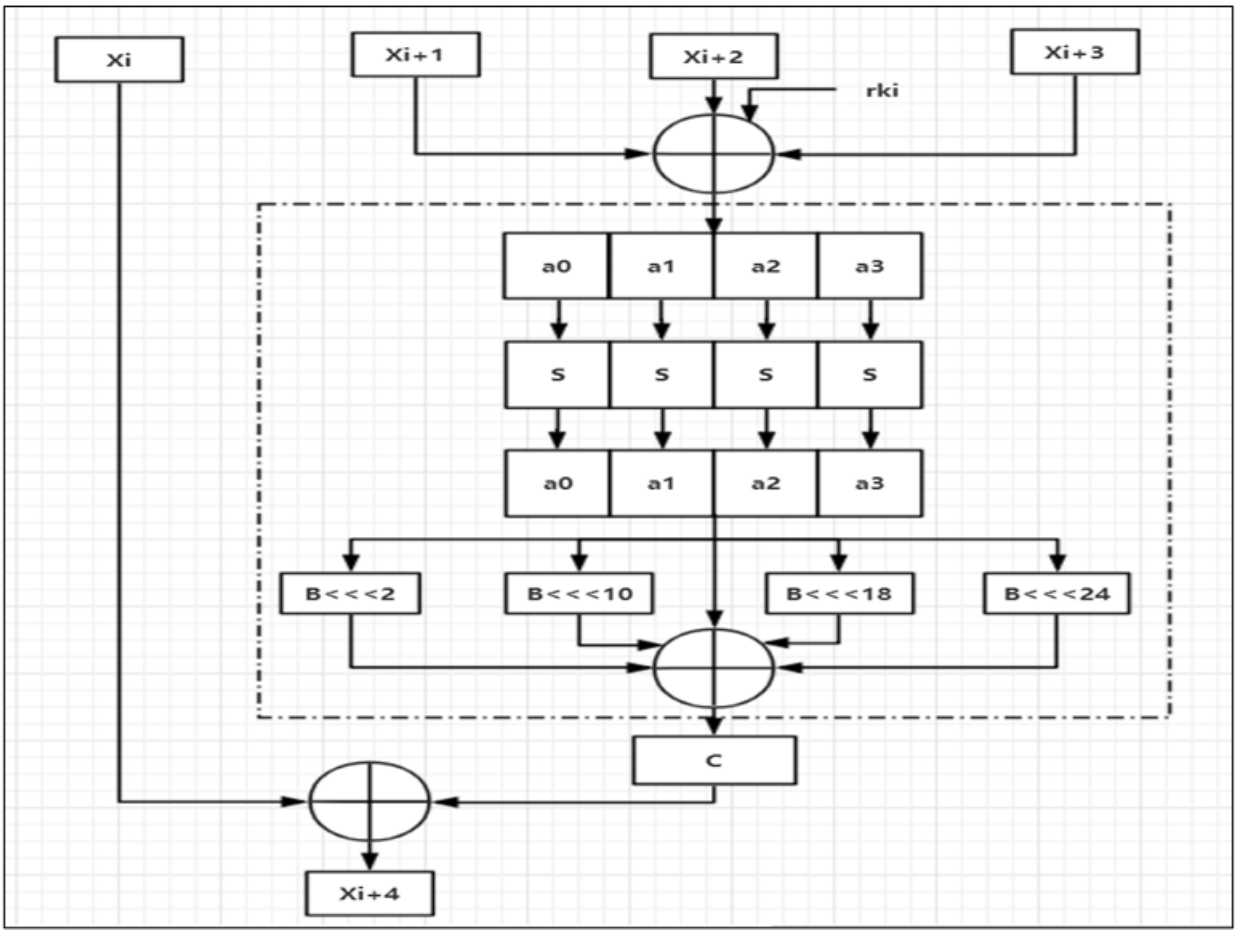

FIGURE 1. T transform

The decryption algorithm of the SM4 block cipher algorithm is the same as that of the encryption algorithm, except that the round key used in the decryption algorithm is $(rk_{31}, rk_{30}, ..., rk_0)$.

Key expansion algorithm: the round key of SM4 block cipher algorithm is generated by key expansion algorithm. The structure of key expansion algorithm is similar to that of round key, and the input of key expansion algorithm is the encryption key used by

encryption algorithm. The round key is expressed as:

$$(K_0, K_1, K_2, K_3) = (MK_0 \oplus FK_0, MK_1 \oplus FK_1, MK_2 \oplus FK_2, MK_3 \oplus FK_3) \tag{2}$$

$$rk_i = K_{i+4} = K_i \oplus T(K_{i+1} \oplus K_{i+2} \oplus K_{i+3} \oplus CK_i) \tag{3}$$

2.2. **Logistic Chaotic System.** In this paper, the chaotic system is used to fuse the original algorithm, which increases the randomness of Sbox. Due to the limitation of logic resources in hardware implementation, the chaotic system uses Logistic map. The chaotic mapping equation is expressed as: $x_{n+1} = 4x_n(1 - x_n)$,in which the initial value $x_n \in (0, 1)$ and the system parameter $\mu \in (0, 1)$ are used. The dynamic behavior of Logistic chaotic map is closely related to the system parameter μ .The bifurcation characteristics of Logistic chaotic map are shown in Figure 2,which shows the two-dimensional relationship between the numerical distribution of iterative chaotic sequence and the system parameter μ . With the increase of μ , the complexity of its dynamic behavior increases, and the Logistic map reaches chaotic state through periodic doubling. Previous studies have shown that Logistic map is in chaotic state when $3.56994568 \leq \mu \leq 4$, and only when the system parameter $\mu = 4$, the iterative value will be mapped in the whole $[0, 1]$ interval, which is called full mapping state.

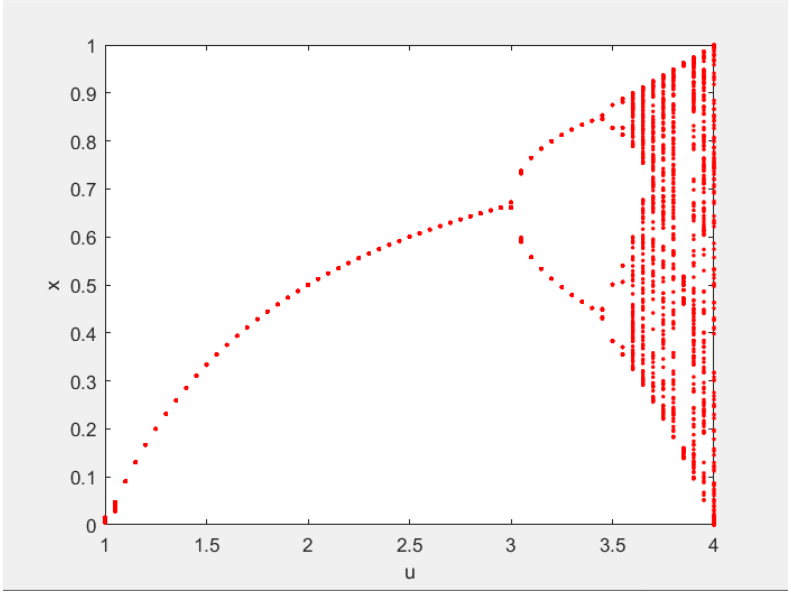

FIGURE 2. Bifurcation characteristics of Logistic chaotic map

3. Improved SM4 Algorithm Based on Logistic Dynamic Sbox.

3.1. **Improvement of Logistic Chaotic Sequence Generator.** The application background of this paper is based on FPGA. The $1-x_n$ module in the Logistic mapping formula can be quickly realized through the inversion module and +1 module. The function of the bitno module is: x_n inversion. Bitno module and adder module are cascaded to complete the calculation of $1 - x_n$. Delay module will delay the input by two clock cycles, because the bitno module and the adder module go through two clock cycles, it is necessary to keep x_n and $1 - x_n$ synchronization; multiplier module mult_gen_0 directly calls the IP core in vivado to complete the calculation of $4x_n(1 - x_n)$. The physical RTL diagram implemented by FPGA hardware is shown in Figure 3.

According to $IV[127{:}0]$ changes to four 32bit initial_values,*Init_new1*=$IV[31{:}0]$, *Init_new2*=$IV[63{:}32]$,*Init_new3*=$IV[95{:}64]$,*Init_new4*=$IV[127{:}96]$ Under the control of the digital selection signal, the data selector will make four initial values enter the system in turn, and these values will complete the calculation of $4x_n(1-x_n)$. in order, that is, the system outputs the first iteration result of initial value 1 in the fifth clock cycle, and the first iteration result of initial value 2 in the sixth clock cycle, and so on Four Logistic chaotic sequence generators with different initial values start from different time to connect the iterative results.

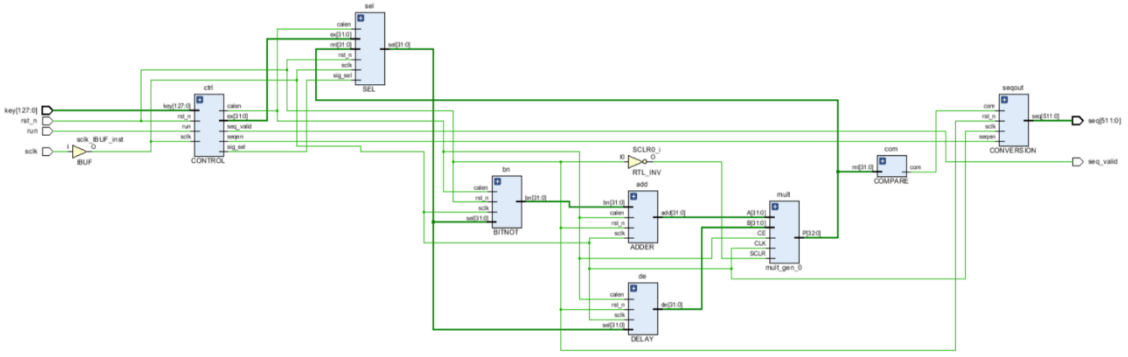

FIGURE 3. FPGA implementation of $4x_n(1-x_n)$ RTL diagram

3.2. **Improved SM4 Algorithm Based on Chaotic Sequence.** From the introduction of SM4 encryption algorithm, we know that the Sbox of SM4 encryption algorithm plays a very important role. On the one hand, it determines the generation of round secret key in the key expansion algorithm, on the other hand, it plays an important role as the only nonlinear component in each round of encryption. Here, the dynamic Sbox based on Logistic is generated by performing operations on the basic Sbox in SM4. Therefore, we do not violate the basic structure of SM4, but improve the security of standard SM4.

The Sbox of SM4 is constructed by multiplication inversion on a finite field and two linear affine transformations.

$$Y = A(Ax + C)^{-1} + C \tag{4}$$

Many SM4 hardware implementations of ASIC and FPGA use look-up tables to perform nonlinear Sbox, such as a 256×1byte ROM. In order to design a small area dynamic Sbox circuit, the complex field decomposition technology is used in this section. The inverse multiplication operation $(Ax+C)^{-1}$ of finite field $GF(2^8)$ is converted to compound field $GF((2^4)^2)$ respectively. Then formula (4) is transformed into the form shown in equation (5). Where T and T^{-1} denote the mapping matrix and inverse mapping matrix of Sbox in SM4, respectively.

$$Y = A(T^{-1}(T(Ax + C))^{-1}) + C \tag{5}$$

According to the expression (5), the Sbox circuit structure of SM4 based on compound field operation is designed as shown in Figure 4. We perform the following steps to execute the SM4 algorithm for dynamic Sboxes

•By introducing 128 bit initial vector IV as the initial value of logistic, a series of 256 bit chaotic pseudo-random sequences are generated.

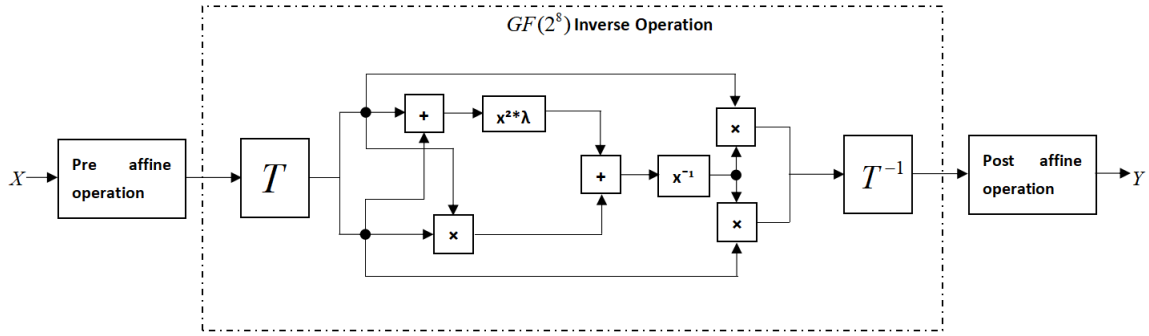

FIGURE 4. Sbox circuit of SM4 based on compound field operation

•The 256 bit pseudo-random sequence is divided into 32 8-bit pseudo-random sequences. chaos_seq=chaos_seq(1),chaos_seq(2),...,chaos_seq(32).

•In this paper, chaotic pseudo-random sequences are introduced to generate random affine functions a and C to generate dynamic Sboxes.Take the round I function as an example,i=1,2,...,32,The 8-bit pseudo-random sequence of chaos_seq(i) is written into the first row of matrix A, and then the upper row of each row is shifted to the right by one bit.Write chaos_seq(i + 1) to C. when I = 32, write chaos_seq (1) to C. According to formula (5), the dynamic Sbox is generated, and the specific process is shown in Figure 5.

•The generated dynamic Sbox replaces the original Sbox in each round and continues to perform the next round.

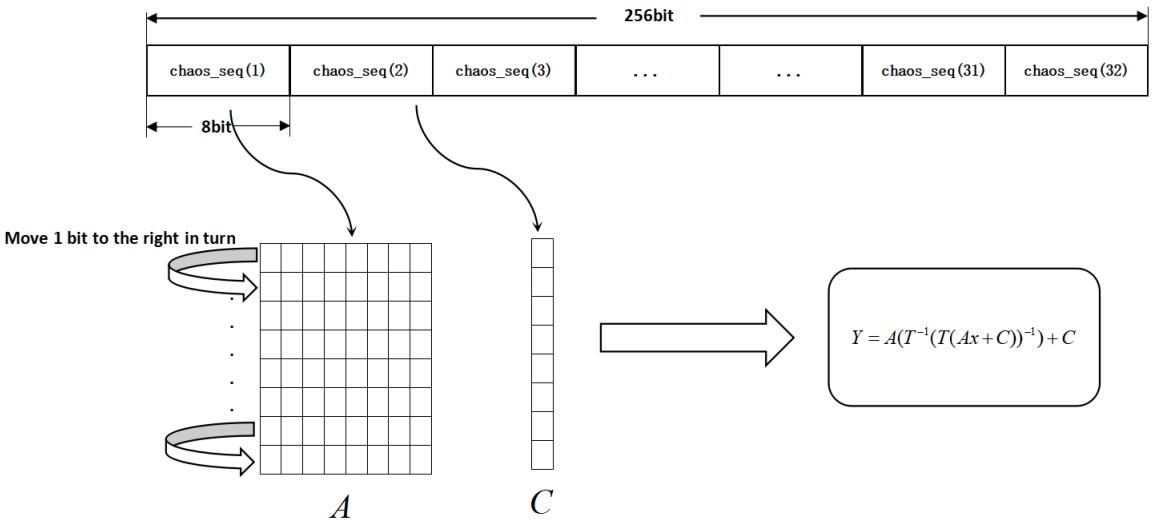

FIGURE 5. Dynamic Sbox based on chaos and order reduction in compound domain

4. **Resource Analysis.** The following table shows the resource consumption of this algorithm at 32 precision, including the SM4 key expansion algorithm based on chaotic system proposed by Wang and Ding's SM4 [20] in 2017. LUT is the basic logic unit in FPGA; multiple is the 32-bit multiplier integrated on FPGA chip; key space is the size of key space. It can be seen from Table 1 that under 32-bit precision, the LUT of chaotic system consumes 63, while the reserved LUT of SM4 algorithm consumes 1324,

TABLE 1. SM4 resource consumption comparison table

Algorithm	Accuracy of Logistic Multiplier	Chaos LUT	SM4 scheme LUT	Key Space
Standard SM4	none	none	1324	128
Wang and Ding's SM4 [20]	32	63	1324	160
Improved SM4	32	63	1324	256

and the key space is 256 bits. It can be seen from table I that although the LUT resource consumption of one-dimensional Logistic chaotic system is 63 more, the SM4 key space is increased compared with the original algorithm, Because of the improvement of the Logistic chaotic sequence generator, compared with the SM4 key expansion algorithm based on chaotic system proposed by Wang and Ding's SM4, the key space is 96 bits larger when the precision of Logistic chaotic multiplier is 32 bits.

5. **Security Analysis.**

5.1. **Avalanche Effect.**

5.1.1. *Strict Avalanche Criterion.* In order to measure the diffusion and confusion of algorithms, SAC attribute has been used. It was first proposed by Webster and Tavares in 1985 . For functions satisfying sac, when the input bit is flipped, it should change the output bit at least 50% probability.

Avalanche effect is an important property of block cipher. It provides the idea of how output can be changed by changing a bit in plaintext. The randomness was calculated by avalanche effect. If the avalanche effect is not greater than that of ciphertext, the analyst can guess the plaintext by monitoring the ciphertext. The avalanche effect is calculated as follows :

$$Avanlancheeffect = \frac{numberoftimescipherbitchanges}{Totalnumberofbitsinplaintext} \tag{6}$$

5.1.2. *Analysis of simulation results.* We simulate SM4 encryption and propose SM4 encryption scheme on vivado 2018.1. In order to study the avalanche effect, we set the same plaintext "0123456789abcdeffedcba9876543210" and the key "0123456789abcdef fedcba9876543210", and the improved SM4 algorithm IV vector is set to "0123456789abcdeffedba9876543210", and generate ciphertext for the two algorithms.

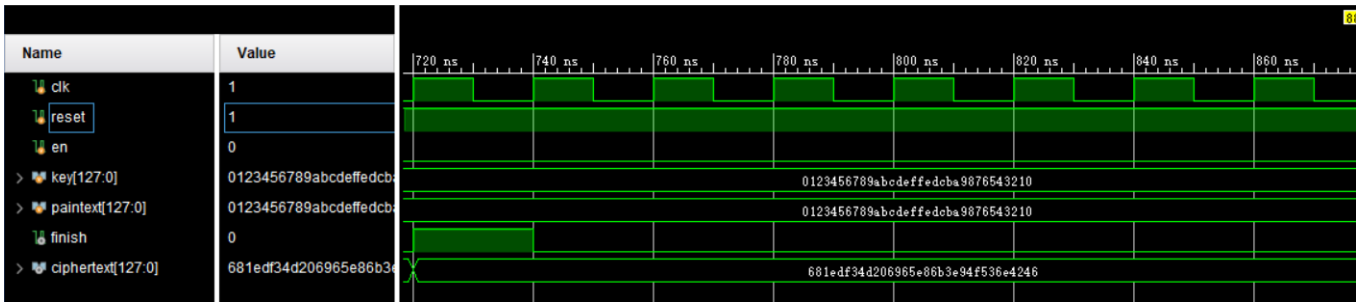

FIGURE 6. Simulation results of standard SM4 encryption

We use the Logistic discrete chaotic system to generate 256 bit pseudo-random sequences, [7:0] to generate the first round of dynamic Sbox, and so on to generate 32 rounds of dynamic Sboxes. Therefore, multiple Sbox can be generated to avoid the static

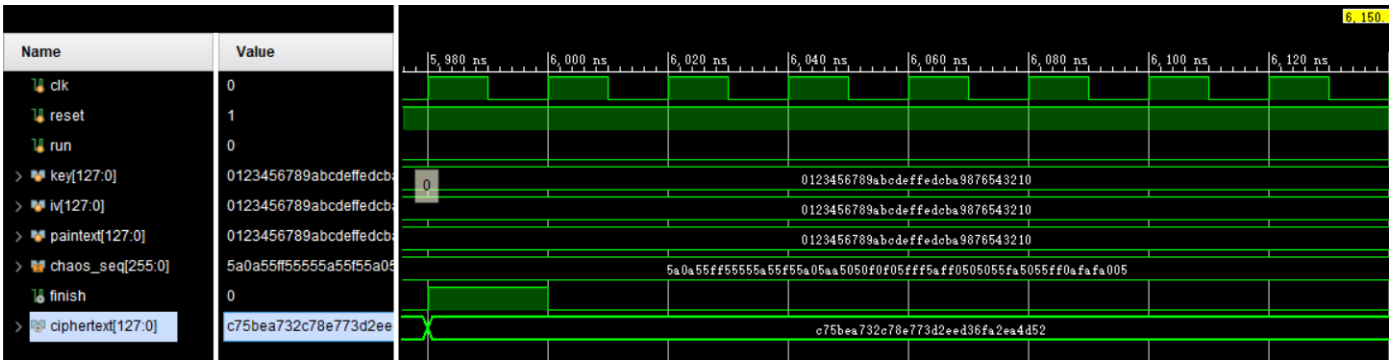

FIGURE 7. Improved SM4 encryption simulation results

behavior of Sbox in SM4, thus improving the security of Sbox. In the Table 2, we show the avalanche effect changes of Standard SM4 -Wang and Ding's SM4 and improved SM4.

TABLE 2.

The avalanche effect changes of standard SM4 and improved SM4

Bit Variation Index	Bit Variations of Standard SM4	Bit Variations of Wang and Ding's SM4	Bit Variations of Improved SM4
0	67	66	69
1	66	62	59
5	57	56	60
9	68	64	65
16	70	69	64
18	55	54	69
23	56	55	55
26	49	48	66
29	58	57	60
36	62	62	63
40	60	63	67
42	65	64	61
45	47	49	51
51	63	60	68
55	61	60	63
59	53	55	61
63	65	63	59
67	59	60	61
69	60	58	59
75	70	65	63
78	65	64	66
84	62	68	72
87	65	62	63
91	60	60	63
94	53	55	69
97	66	65	64
103	50	52	69
107	58	60	65
110	59	62	67
116	58	55	57
119	62	60	68
121	65	60	59
127	60	58	62

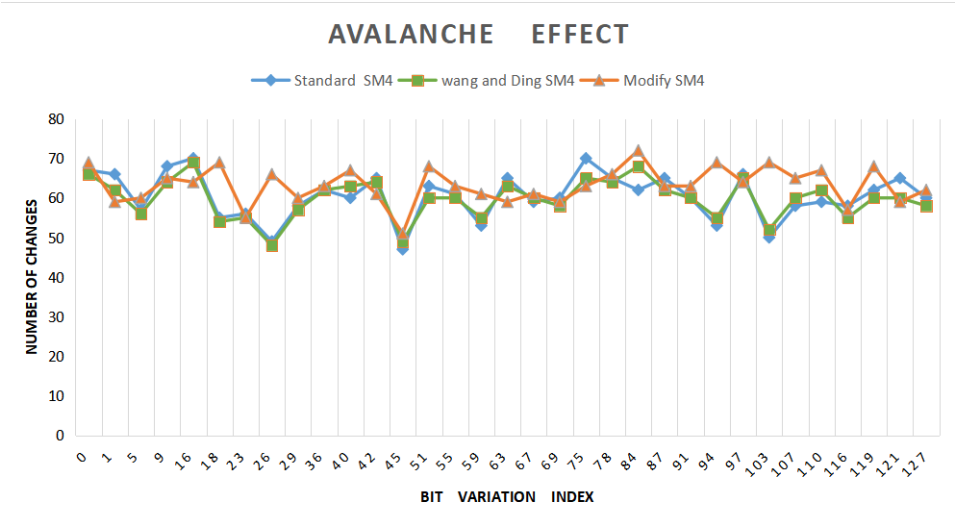

FIGURE 8. Comparison of bit change between Standard SM4 and improved SM4

As can be seen from Figure 9, compared with the Standard SM4 algorithm and Wang and Ding's SM4, the algorithm has better avalanche effect in most cases, and the algorithm using chaotic dynamic Sbox satisfies the strict avalanche criterion more times. This paper does not violate the security of SM4, but takes SM4 as the basic block and does not change its mathematical calculation, so that the Sbox is associated with the pseudo-random sequence generated by logistic. Therefore, this paper does not violate the basic SM4 algorithm.

5.2. **Information Entropy.** Entropy is a physical quantity which represents the disorder degree of physical system in thermodynamics. Shannon, the founder of information theory, formally introduced the concept of entropy into information theory [21] in mathematical theory of communication, which is the founder of information theory. The definition of entropy is as follows:

$$H(x) = -\sum_{i}^{n} p(x_i) \log_2 p(x_i) \tag{7}$$

Where, represents the probability of the symbol Xi appearing in the message in the symbol set, and satisfies the following conditions:

$$\begin{gathered} 0 \leq p(x_i) \leq 1 \\ \sum_{i}^{n} p(x_i) = 1 \end{gathered} \tag{8}$$

The maximum value of information entropy occurs when every symbol appears with equal probability, that is, when $p(x_i) = \frac{1}{n}$, the maximum information entropy is $H_{max} = \log_2 n$.

We set up 50 groups of random plaintext, use the same key and initial value of chaos to generate ciphertext of the three algorithms, and calculate the information entropy of the corresponding algorithms. The average values of 50 groups of information entropy are shown in Table 3. 1 sequence has higher entropy, which is closer to the maximum value 1, and the distribution is more uniform. Therefore, from the perspective of information entropy attack, the encryption algorithm is secure.

TABLE 3. Information entropy comparison table

Algorithm	Information Entropy
Standard SM4	0.985090433
Wang and Ding's SM4	0.985409765
Improved SM4	0.994703367

6. **Conclusions.** The standard SM4 algorithm uses static Sbox, which cannot be updated. In the face of various attacks, it is possible to obtain the key from the Sbox, so as to crack the ciphertext and get the plaintext. In this paper, a dynamic Sbox SM4 encryption algorithm based on Logistic is proposed. The initial vector IV is taken as the initial value of Logistic chaotic system. The random sequence is associated with the original Sbox, and different Sbox is generated in different encryption rounds. Due to the sensitivity of initial value of chaos, small change of initial value will lead to different encryption results, which effectively increases the key space, Compared with the improved algorithm proposed by Wang and Ding's SM4, this paper takes up the same resources under the condition of 32-bit precision, and the key space is 96 larger. Compared with the original algorithm, because Sbox is dynamically generated, linear and differential cryptanalysis can not be carried out on dynamic Sbox, and security analysis is conducted with standard SM4 to find avalanche effect and information The entropy is greater than the standard SM4, so this paper is more secure than the existing improved SM4 algorithm and the original SM4 algorithm.Of course, this paper still has limitations.We should consider more security indexes and randomness of pseudo-random sequences. The next work will focus on two aspects.First, we will analyze more safety indicators, such as correlation, CPA and so on. Next, We will test the randomness and periodicity of the pseudo-random sequence and discuss the properties of the sequence.

Acknowledgment: This work is partially supported by National Natural Science Foundation of China (No.61471158).The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] W. Diffie, G. Ledin, SMS4 Encryption Algorithm for Wireless Networks,*IACR Cryptology eprint Archive*, 329, 2008.
[2] P. Agarwal, A. Singh, A. Kilicman, Development of key-dependent dynamic Sboxes with dynamic irreducible polynomial and affine constant, *Advances in Mechanical Engineering*,vol. 10, no. 7, pp. 1-18, 2018.
[3] M. Ahmad, E. A. Solami, Evolving Dynamic Sboxes Using Fractional-Order Hopfield Neural Network Based Scheme,*Entropy*, vol. 22, no. 7, pp. 717-731, 2020.
[4] Q. Lu, C. X. Zhu, G. J. Wang, A Novel S-Box Design Algorithm Based on a New Compound Chaotic System,*Entropy*,vol. 21, no. 10, pp. 1004-1018, 2020.
[5] B. Rashidi, Compact and efficient structure of 8-bit S-box for lightweight cryptography,*Integration*,vol. 76, pp. 172-182, 2021.
[6] K. H. Sun, principle and technology of chaotic secure communication,*Beijing: Tsinghua University Press*, pp. 35-37, 2015.
[7] C. F. Wang, Q. Ding, A Class of Quadratic Polynomial Chaotic Maps and Their Fixed Points Analysis,*Entropy*,vol. 21, no. 7, pp. 658-670, 2019.
[8] C. F. Wang, Q. Ding, A New Two-Dimensional Map with Hidden Attractors, *Entropy*, vol. 20, no. 5, pp. 322-331, 2018.
[9] C. M. Ou, Design of block ciphers by simple chaotic functions, *IEEE Computational Intelligence Magazine*,vol. 3, no. 2, pp. 54-59, 2008.8.

[10] T. Y. Wu, X. N. Fan, K. H. Wang, J. S. Pan, C. Ming Chen, J. M. T. Wu, Security Analysis and Improvement of An Image Encryption Scheme Based on Chaotic Tent Map,*Journal of Information Hiding and Multimedia Signal Processing*,vol. 9, pp. 1050-1057, 2018.

[11] C. M. Chen, L. Xu, K. H. Wang, S. Liu, T. Y. Wu, Cryptanalysis and improvements on three-party-authenticated key agreement protocols based on chaotic maps, *Journal of Internet Technology*,vol. 19, pp. 679-687, 2018.

[12] C. M. Chen, W. Fang, S. Liu, T. Y. Wu, K. H. Wang, Improvement on a chaotic map-based mutual anonymous authentication protocol,*Journal of Information Science and Engineering*,vol. 34, pp. 371-390, 2018.

[13] C. M. Chen, W. Fang, K. H. Wang, T. Y. Wu, Comments on "An improved secure and efficient password and chaos-based two-party key agreement protocol,*Nonlinear Dynamics*,vol. 87, no. 3, pp. 2073-2075, 2017.

[14] K. Feng, X. Huang, Shu-Chuan Chu, John F Roddick, D. Qun, An Implementation of Chaotic Pseudo-Random Sequence Generator Based on Pipelined Architecture,*Journal of Network Intelligence*,vol. 4, no. 2, pp. 71-79, 2019.

[15] J. A. P. Artiles, D. Chaves, C. Pimentel, Image encryption using block cipher and chaotic sequences,*Signal Processing: Image Communication*, vol. 79, pp. 24-31, 2019.

[16] J. Urías, E. Ugalde, G. Salazar, A cryptosystem based on cellular automata,*Chaos*,vol. 8, no. 4, pp. 819-822, 1998.

[17] M. H. Fan, F. F. Yang, Dynamic Sbox construction based on chaotic map in block cipher, *Radio engineering*,vol. 46, no. 3, pp. 33-36, 2016.

[18] H. H. Zhu, X. J. Tong, M. Zhang, A novel method of designing Sbox based on dynamic compound chaotic system,*Journal of Nanjing University:Natural Science*, vol. 54, no. 3, pp. 543-547, 2018.

[19] Y. Y. Han, Y. R. He, P. H. Liu, D. Zhang, Z. Q. Wang, W. C. He, Construction and application of ZUC dynamic S-box based on chaotic system,*Computer research and development*,vol. 10, pp. 2147-2157, 2020.

[20] C. F. Wang, Q. Ding, SM4 key scheme algorithm based on chaotic system,*Acta Physica Sinica*,vol. 66, no. 2, pp. 80-88, 2017.

[21] C. Shannon, A Mathematical theory of communication.*Bell System Technical Journal*,vol. 27, pp. 623-656, 1948