Journal of University of Science and Technology Beijing
Volume 11, Number 2, April 2004, Page 183

Automation

An input signal used in process identification — chaos sequence

Geng Zhao[1,2], *Deling Zheng*[2], *and Shaojun Huang*[2,3]

1) The Third Department, Beijing Electronic Science and Technology Institute, Beijing 100070, China
2) Information Engineering School, University of Science and Technology Beijing, Beijing 100083, China
3) Computer Engineering School, Anshan University of Science and Technology, Anshan, 114044, China
(Received 2003-03-16)

Abstract: The sequences which consist of any segment of a chaos sequence are called as C-sequences. These sequences could be used as a kind of input signals to replace M-sequences in the process identification. This substitution is theoretically proved to be feasible. Inverse C-sequences are created in a way similar to inverse M-sequences to solve the problem that C-sequences have non-ideal balance property, that is, the numbers of '0' and '1' are unequal. Besides its good pseudo-random property, the sequences have other advantages such as easy to generate, varieties of the segment and adjustable cycle time.

Key words: process identification; M-sequence; C-sequence

[*This work was financially supported by the National Natural Science Foundation of China (No.70071047, No.10247005) and the Postdoctoral Science Foundation of China (No.2002032129).*]

Theoretical analysis indicates that better identification effect can be obtained when white noise is used as the input signal of process identification. But industrial facility, such as valves, can not act in form of white noise. So the traditional input signal of identification in common use is M-sequence. It has pseudo-random property like white noise. As chaos is researched widely in recent ten years, new theories and concepts of chaos come forth in other fields and are being used [1-5]. Because of the pseudo-random property of chaos sequences, it is used successfully in spread-spectrum communication. Reference [4] researched the distributing characteristic of '0' and '1' and the correlation property of Chebyshevy mapping in spread-spectrum communication. Reference [5] researched the correlation property of chaotic sequences in multi-address communication produced by Lorenz equation. This suggests a use in control theory. So C-sequences and inverse C-sequences are presented in this paper.

1 Random M-sequences and inverse M-sequences

The M-sequence is a pseudo-random binary sequence (PRBS), the relation of its element is

$$x_i = a_1 x_{i-1} \oplus a_2 x_{i-2} \oplus \cdots \oplus a_P x_{i-P} ,$$

Corresponding author: Geng Zhao, **E-mail:** dr_zhaogeng@sohu.com

where $i = P+1, P+2, \cdots$; the values of coefficients $a_1, a_2, \cdots, a_{P-1}$ are either 0 or 1, coefficient a_P is always 1; $\oplus$ is XOR operation (exclusive OR), so as coefficients $a_1, a_2, \cdots, a_{P-1}$ are selected properly, then the sequence can be cycled by the maximal length of $2^P - 1$, generating an M-sequence. The self-correlation function of M sequences is similar to the pulse function. A self-correlation function of M-sequences with bit-length P is presented in reference [6], its length of the cycle is N_P and the clock time is Δt.

$$R_M(\tau) = \begin{cases} a^2\left(1 - \dfrac{N_P+1}{N_P} . \dfrac{|\tau|}{\Delta t}\right) & -\Delta t \le \tau \le \Delta t \\ -\dfrac{a^2}{N_P} & \Delta t < \tau < (N_P - 1)\Delta t \end{cases} \quad (1)$$

where a is the voltage value of logic '0' or logic '1'. It can be seen from equation (1) that the self-correlation function of M-sequences is similar to δ function when $N_P \to \infty$, so it is a good input signal for the identification of processes. But in one cycle period $N_p = (2^P-1)$, the number of logic '0' is $(N_p-1)/2$, the number of logic '1' is $(N_p+1)/2$, one more than logic '0'. In practice, N_p is not easy to be made big, so M-sequences have the component of direct current. This will produce pure bias to the identification object, which is not what we hope. Therefore M-sequences are improved.

Let $M(k)$ be an M-sequence, its cycle is N_p and the value of its element is 0 or 1; $S(k)$ is a square wave sequence, its cycle is 2 and the value of its element is 0 or 1. An M-sequence of $2N_p$ in length and a square wave sequence are made the operation of XOR, then a composite sequence with the cycle period of $2N_p$ is obtained, that is the inverse M-sequence. In one cycle period of $2N_p$ in the inverse M-sequence, probabilities of logic '0' and logic '1' are equal, the sequence in the first cycle period N_p is inverse repetition of the sequence in the second N_p. The correlation function of the inverse M-sequence and the original M-sequence equals 0, and the self-correlation function of the inverse M-sequence resembles that of the M-sequence, its random property is better than that of the M-sequence, so it has wider application in the field of process identification.

2 Random property of C-sequences and inverse C-sequences

Chaos mapping can be described with the determined difference equation, but in some condition the mapping appears random. This can be illustrated by the simplest chaotic mapping—logistic mapping.

Assuming the logistic mapping is:

$$x_{k+1} = 1-\mu x_k^2, \quad \mu \in (0,2], \quad x \in [-1,1] \tag{2}$$

For the general chaotic mapping $x_{k+1} = f(x_k)$, the probability density $\rho(x)$ can be obtained by Perron-Frobenious equation [7]:

$$\rho(y) = \sum_{\{x_i = f^{-1}(y)\}} \frac{\rho(x_i)}{|f'(x_i)|}.$$

When $\mu = 2$, the probability density of equation (2) is:

$$\rho(x) = \frac{1}{\pi\sqrt{1-x^2}}, \quad x \in [-1,1].$$

Let $f(x)$ be even symmetrical mapping from [−1, 1] to [−1, 1], and its probability density is also even and symmetrical, then the chaotic sequence of $\{C_k\}$ can be obtained from the trace of $\{x_k\}$:

$$\{C_k\} = \{\mathrm{sgn}(x_k)\}, \quad k = 0, 1, \cdots, N-1 \tag{3}$$

where sgn(x) is the sign function, N the cycle of chaotic sequence $\{C_k\}$. Here the sequence $\{C_k\}$ could be selected arbitrarily and is defined as the C-sequence.

The self-correlation function of the C-sequence is:

$$M_C(\tau) = \frac{1}{N}\sum_{k=0}^{N-1} c_k \cdot c_{k+\tau} = \frac{1}{N}\sum_{k=0}^{N-1} \mathrm{sgn}(x_k)\cdot \mathrm{sgn}(x_{k+\tau})$$
$$= \frac{1}{N}\sum_{k=0}^{N-1} \mathrm{sgn}[f^{(k)}(x_0)\cdot f^{(k+\tau)}(x_0)]$$

where $f^{(k)}(x_0), f^{(k+\tau)}(x_0)$ are the kth and $(k+\tau)$th iteration of $f(x)$ beginning with the original value of x_0. From the ergodic theory [4, 5] of chaos, to each integrabel function $\varphi(x)$ and almost all the original value x_0,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} \varphi[f^{(k)}(x_0)] = \int \rho(x)\varphi(x)\mathrm{d}x.$$

Let $\varphi(x) = \mathrm{sgn}[xf^{(\tau)}(x)]$, then

$$\lim_{N\to\infty} M_C(\tau) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} \varphi[f^{(k)}(x_0)] = \int_{-1}^{1} \rho(x)\cdot\varphi(x)\mathrm{d}x \tag{4}$$

when $k \neq 0$, $f^{(\tau)}(x)$ is even and symmetrical, so $\varphi(x)$ is odd and symmetrical; $\rho(x)$ is also even and symmetrical, therefore the integral of formula (4) is 0,

$$\lim_{N\to\infty} M_C(\tau) = \begin{cases} 1 & \tau = 0 \\ 0 & \tau \neq 0 \end{cases} \tag{5}$$

From formula (5) it can be seen that the self-correlation function of the C-sequence is δ function. The function has asymptotic ideal self-correlation property. It must be pointed out that chaos sequences are easy to be produced. If two original values x_0, y_0 are given arbitrarily, from the knowledge that chaos is very sensitive to the original condition [7], two independent sequences will be produced after a certain period of time. Their mutual-correlation function is:

$$\lim_{N\to\infty} M_{CX,CY}(\tau) = \frac{1}{N}\sum_{k=0}^{N-1} x_k \cdot b_{k+\tau}$$
$$= \frac{1}{N}\sum_{k=0}^{N-1} \mathrm{sgn}[f^{(k)}(x_0)]\mathrm{sgn}[f^{(k+\tau)}(y_0)]$$
$$= \int_{-1}^{1}\int_{-1}^{1} \mathrm{sgn}(x)\mathrm{sgn}[f^{(\tau)}(y)]\rho(x,y)\mathrm{d}x\mathrm{d}y$$
$$= \int_{-1}^{1} \mathrm{sgn}(x)\rho(x)\mathrm{d}x \int_{-1}^{1} \mathrm{sgn}[f^{(\tau)}(y)]\rho(y)\mathrm{d}y = 0 \tag{6}$$

In deduction above, the independence theorem of compound probability in the stochastic process, that is $\rho(x,y) = \rho(x)\cdot\rho(y)$, is applied. It may be seen from the chaotic property [7] and formula (6) that C-sequences are easy to be produced, the number of C-sequences are very big, their cycle period is very long, they are independent mutually and can be selected according to the need of processes identification.

Now suppose voltage '1' be logic '1', voltage '−1' be logic '0' in the C-sequence, the difference between their numbers is L, then

$$\lim_{N\to\infty} \frac{L}{N} = \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} C_k = \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} \mathrm{sgn}[f^{(k)}(x_0)] =$$

$$\int_{-1}^{+1} \mathrm{sgn}(x)\rho(x)dx = 0 .$$

So the number of logic '1' and '0' in the C-sequence will become equal with the increase of N, but N is finite in the process identification. According to formulae (2) and (3), **table 1** shows the simulated distributing characteristic of '0' and '1' with different original values N and μ=2. Table 1 shows that if the C-sequence is used directly, pure bias to the process is possibly greater than that of the M-sequence. The number of difference between logic '1' and '0' in the M-sequence is 1 without relation with the size of N. The number of difference between logic '1' and '0' in the C-sequence is changeable depending on the value of N, that is to say, the bias in the C-sequence is bigger than that of the M-sequence, this limits the number of C-sequences which can be used.

Table 1 Distributing numbers of '0' and '1' simulated by computer

N	logic	Original value						
		0.05	0.25	0.45	0.65	0.75	0.85	0.95
100	1	50	54	50	52	55	51	50
	0	50	46	50	48	45	49	50
1000	1	492	503	513	521	508	497	491
	0	508	497	487	479	492	503	509
10000	1	4918	5034	4974	4997	5001	4995	5082
	0	5082	4966	5026	5003	4999	5005	4918

The following is the inverse C-sequence (IC-sequence) which is constructed according to the IM-sequence.

Let $C(k)$ be a C-sequence which is cut randomly, the cycle is N, the value of its elements are either 0 or 1, and

$$\mathrm{IC}(k) = -\mathrm{IC}(k+N) ,$$

then the inverse C-sequence with a cycle of $2N$ is obtained:

$$\mathrm{IC}_k = \begin{cases} \mathrm{sgn}(x_k) & 0 < k \le N \\ -\mathrm{sgn}(x_k) & N < k \le 2N \end{cases} \quad (7)$$

Obviously the inverse C-sequence in formula (7) shows that the probability of logic '0' and '1' is equal in one cycle and it has the same self-correlation property with the C-sequence, so it is more suitable than the C-sequence to be used in the process identification.

The following is the self-correlation property and mutual-correlation property of inverse-C sequence of Logistic mapping ($\mu = 2, x_0 = 0.56782, y_0 = 0.467234$) which is shown in **figures 1**(a) and 1(b). The distributing characteristic of '0' and '1' is shown in **figure 2**.

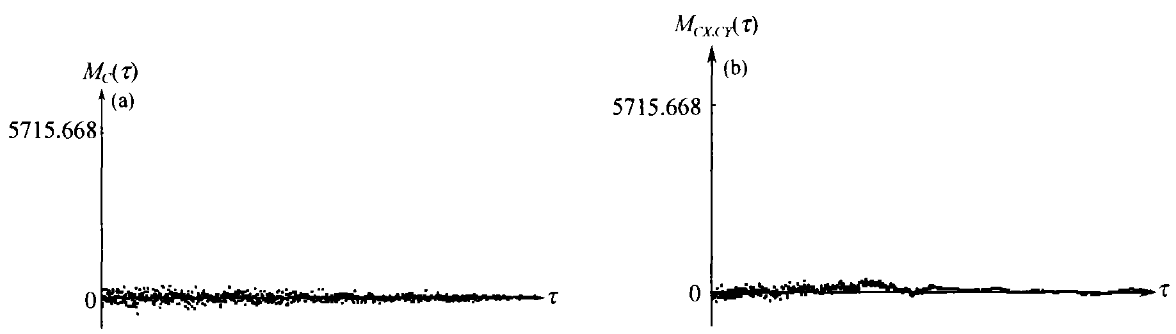

Figure 1 Curves of the correlation property: (a) self-correlation property; (b) mutual-correlation property (The interval τ=0, 1, $\cdots$, 2000).

When $\tau = 0$, the value of the self-correlation function is maximal. If $\tau > 0$, the value is small. The former is about 1000 times larger than the latter, which is shown in figure 1(a). When τ changes from 0 to 2000, the mutual-correlation value is very small, which is shown in figure 1(b). With more experiments it can be found that the correlation property gradually approaches 0 as N increases.

The probabilities of '0' and '1' in figure 2 are 0.5, each of them is 16000. The probability of one '0' or one '1' is about 1/2. The probability of two '0' or two '1' is about 1/4. The probability of three '0' or three '1' is about 1/8. The rest can be deduced by analogy. This is very close to the statistics property of white noise.

According to the above analysis and random axiom condition of Golomb, an inverse-C sequence is a random signal which accords with Gauss distributing, so the sequences are favorable to be as input signals for the process identification.

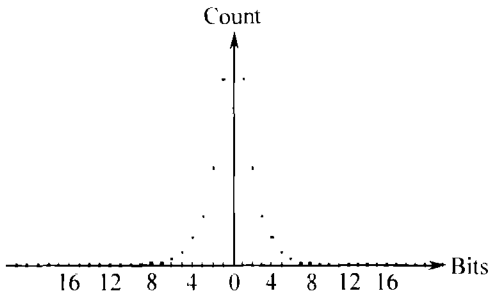

Figure 2 Symmetrical distributing characteristic of '0' and '1'.

3 Comparison between C-sequences and M-sequences

Both the self-correlation functions of a C-sequence and an M-sequence are approximate to the pulse function, have the property of white noise, and can be used as an input signal for process identification. The distributing characteristic of '0' and '1' in a C-sequence is worse than that in an M-sequence when N is finite. But the distributing characteristic and random property of the inverse C-sequence and the inverse M-sequence are same, so the inverse C-sequences are more suitable to be the input signal for process identification.

The selection of coefficients $a_1, a_2, \cdots, a_{P-1}$ in M-sequences is not random. Two characteristic polynomial conditions must be satisfied, otherwise M-sequences will not possibly be produced. From this point, C-sequences are produced easier than M-sequences.

The number of M-sequences is limited by P, so it is finite. C-sequences can be produced with different conditions, and to a given original condition it can be selected arbitrarily if without regard to the length of operation word, so it is infinite.

4 Conclusions

The concept of C-sequences and inverse C-sequences is introduced. The feasibility of the sequences as input signal for process identification is verified. It adds a kind of input signal source to process identification, but its practical effect should be verified by further simulations and studies. In the same way, the feasibility of C-sequences and inverse C-sequences produced by other mappings, such as May mapping, Henon mapping, Tent mapping and Chebyshevy mapping *etc.*, also should be studied more in future.

References

[1] G.H. Beteni and C.D. McGillem, A chaotic direct-sequence spread-spectrum communication system [J], *IEEE Trans. Commun.*,1994, COM-42(2/3/4), p.1524.

[2] C. Ling and S.G. Sun, Chaos spread-spectrum sequence generator [J], *J. Electron.* (in Chinese), 23(1998), No.2, p.235.

[3] H. Wang and JJ.D. Hu, Logistic map chaos spread-spectrum sequence [J], *Acta Electron. Sin.* (in Chinese), 25(1997), No.1, p.19.

[4] T. Kohda and A. Tsuneda, Pseudo-random sequences by chaotic nonlinear maps and their correlation properties [J], *IEICE Trans. Commun.*, E76-B(1993), No.8, p.855.

[5] D.G. Xu, J.X. Yin, and B.Z. Xu, Chaos sequence and multi-address communication [J], *J. Electron.* (in Chinese), 20(1998), No.4, p.515.

[6] C.Z. Fang and D.Y. Xiao, *Process Recognition* [M], Tsinghua University Press, 1989.

[7] X.X. Wu and Z. Chen, *Introduction to Chaos* [M], Shanghai Science and Technology Literature Press, 1997.