

# A NOVEL ALGORITHM FOR THE DETERMINATION OF DELAY TIME BASED ON MUTUAL INFORMATION

<sup>1</sup>PENG HE, <sup>1</sup>DE-YUN ZHOU, <sup>1</sup>YING ZHOU, <sup>1</sup>JI YU

Department of Electronic Information, Northwestern Polytechnical University, Xi'an, China

E-mail: [hp17@163.com](mailto:hp17@163.com)

## ABSTRACT

We analyze different methods for the determination of delay time in state space reconstruction. Although the method of mutual information is more accurate, its computational process is quite cumbersome and time-consuming. Therefore, based on the pseudo-periodicity of chaotic attractor, we put forward a new algorithm to refine delay time step by step using mutual information. The reliability of this algorithm is proved based on the numerical experiment of Rössler and Lorenz systems.

**Keywords:** *State space reconstruction, Time series, Delay time, Mutual information, Pseudo-periodicity*

## 1. INTRODUCTION

Process monitoring data-based fault diagnosis is an effective means to protect the safety of complex systems process and avoid major accidents. Large amounts of data can be collected in the processing of complex systems, and this signal usually presents a strong non-stationary and non-linear, making its fault feature extraction more difficult. The conventional method uses all of the variables for diagnosis, it takes long calculation time, and diagnostic accuracy is low, which increases the system operation and maintenance costs [1]. With the continuous progress of the nonlinear dynamics research, particularly in the state space reconstruction theory of chaotic time series signal [2] applied to the field of equipment fault diagnosis, so that we can get from the univariate time series with the effectiveness of the system the behavior dynamics indicators of interest rates, such as dimension [3], Lyapunov index [4], reveals the essence of the failure, the more profound understanding of the complex nonlinear dynamical systems from a new perspective [5]. Other people will phase space reconstruction theory is widely used in a variety of engineering fields, such as in the sensor system fault diagnosis [6], power load prediction [7], both unexpected excellence effect.

## 2. STATE SPACE RECONSTRUCTION

State space reconstruction proposed by Takens[2] is always employed as the first step in the analysis of nonlinear time series. The fundamental principle is: because every component

of a system is influenced by the others, we can derive the dynamics from the observation. Take the data at some fixed delay time later as high dimension:

$$X(n) = [x(n), x(n+\tau), x(n+2\tau), \dots, x(n+(m-1)\tau)]$$

, where  $m$  is embedding dimension and  $\tau$  is delay time. It has been proven that the reconstructed attractor is topological equivalent to the original one. When the time series is infinite and noiseless, the choice of delay time is arbitrary. But in practice the sampled series is always noisy and limited, so  $\tau$  is important for the quality of reconstruction.

The principle of choosing  $\tau$  is: the original series and its delayed series are independent to each other but not unrelated at all. There are mainly three algorithms for choosing  $\tau$ : autocorrelation function [8], average displacement [9] and mutual information [10]. Essentially, autocorrelation function reflects the linear correlation of data, so it is not suitable to analyze nonlinear problem. Average displacement is a geometric approach, but it lacks the theoretical basis and implementation criterion. What is more important is that the result is influenced by embedding dimension, so the corresponding joint algorithm [11] needs to be tested repeatedly, which is a cumbersome process. Recently many researchers employ the first local minimum of mutual information as delay time and get satisfactory results. This method is independent of  $m$ , along with high accuracy. While in practice the tedious calculations greatly restrict its application.

### 3. MUTUAL INFORMATION

Suppose we have gotten observed time series  $x(i), i = 1 \cdots N'$ , and its delayed series is  $y(i)$ , wherein  $y(i) = x(i + \tau)$  ( $\tau = n \times t$ ,  $t$  is the sampling interval). According to information science, we measure the interdependence relationship between  $x(i)$  and  $y(i)$  by mutual information [12]. When the function of mutual information reaches the first local minimum, we take this time as suitable delay time  $\tau$ .

The calculation of mutual information is quite complicated. And the algorithms are classified by different partition methods on X and Y coordinates axes. In this paper we employed the method proposed by Fraser [10], which divides the coordinate axes into grids with equal marginal probability:

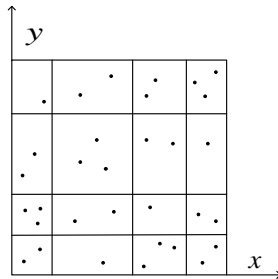


Fig.1 Schematic Of A Uniform Probability Distribution

First we divide X and Y axes into two parts with equal marginal probability, and judge the sparseness of the four grids according to  $\chi^2$  distributed criterion [10]. If it still has substructure, we continue the partition; if the grid is sparse, we stop our partition in this region and calculate the probability. After dividing for  $m$  times, each axes is divided into  $2^m$  parts and the number of all the grids is  $4^m$ . The formula for calculation mutual information is:

$$I_m = \sum_n P_{xy}(R_m(n)) \log \left[ \frac{P_{xy}(R_m(n))}{P_x(R_m(n))P_y(R_m(n))} \right] \quad (1)$$

Where  $P_{xy}(x_i, y_j)$  means the probability  $x_i, y_j$  both happen,  $P_x(x_i)$  is the probability  $x_i$  happens, and  $P_y(y_j)$  is the probability  $y_j$

happens.  $R_m(n)$  Means the partition has been implemented for  $m$  times.

The recursion formula is:

First install:

$$I(X, Y) = (1/N_0)F(R_0) - \log(N_0) \quad (2)$$

$N_0$  Is the total number of points;

If  $R_m(n)$  is sparse, then

$$F(R_m(n)) = N(R_m(n)) \log[N(R_m(n))] \quad (3)$$

If  $R_m(n)$  still has substructure, then

$$F(R_m(n)) = N(R_m(n)) \log(4) + \sum_{j=0}^3 F(R_{m+1}(n, j)) \quad (4)$$

### 4. IMPROVED ALGORITHM TO CHOOSE DELAY TIME $\tau$ BY MUTUAL INFORMATION

In practice all the chaotic systems are dissipative [13], so their trajectories will finally shrink into a capture area and form an attractor. While inspecting the local region, the motion inside the attractor is quite unstable. Adjacent tracks reject from each other and separate at an exponential speed. In short attractor in chaotic system is formed by much separation and folds of the trajectory. And we can find much self-similarity shapes in the structure if we amplify any local region of the attractor, which is called fractal. Meanwhile it has been proven that the chaotic motion is ergodic. All these properties can explain why there are many troughs and crest one after another in the function of mutual information:

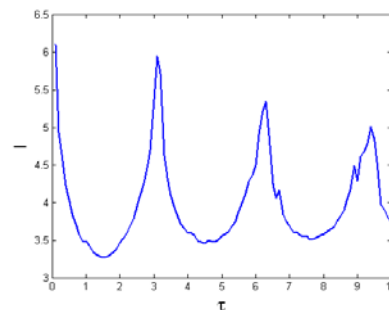


Fig.2 The Mutual Information Of Rössler System

Because of the pseudo-periodicity of chaotic systems we put forward a fast algorithm to determine  $\tau$  by calculating mutual information at some particular delay time. First we extend the calculation range and set the step at a relatively large value. Take the Rössler system:

$$\begin{cases} \dot{x} = -y - z \\ \dot{y} = x + ay \\ \dot{z} = z(x - c) + b \end{cases} \quad (a = 0.15, b = 0.2, c = 10) \text{ as}$$

an example. Suppose we have gotten 4096 observed points of x component, and the sampling interval is 0.1s. In the first step we choose the feasible range at  $\tau = 1 \sim 10s$ , the Sep length is 1s, and calculate the mutual information. The result is as follows:

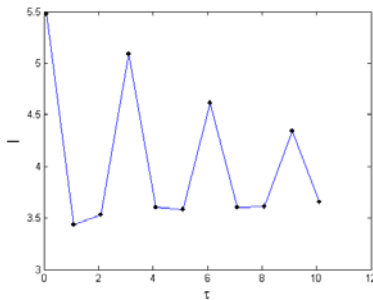


Fig.3 Mutual Information Of Rössler System ( $\tau = 0 \sim 1s$ )

From figure (3) we determine that first extreme interval occurs at  $\tau = 0 \sim 0.2s$ , so we set our calculation region at this region and reduces the step length to 0.1s. The result is:

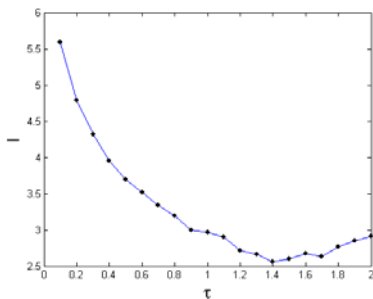


Fig.4 Mutual Information Of Rössler System ( $\tau = 0 \sim 0.2s$ )

From the above figure we obviously get the first local minimum of mutual information at  $\tau = 1.4$ , and take this value as optimal delay time in phase space reconstruction. In order to verify the feasibility of our algorithm we applied it to Lorenz system:

$$\begin{cases} \dot{x} = -\sigma(x - y) \\ \dot{y} = -xz + rx - y \\ \dot{z} = xy - bz \end{cases} \quad (\sigma = 16, r = 45.9, b = 4)$$

The sampling interval is  $t = 0.01s$ . As shown in figure (5) the result ( $\tau = 0.1$ ) is agreed with empirical choice [14].

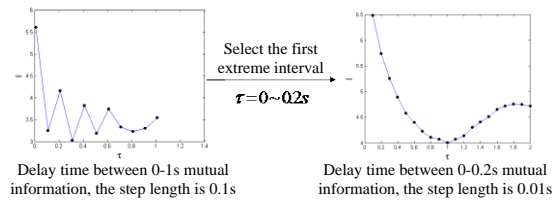


Fig.5 The Flowchart Of Determining  $\tau$  For Lorenz System

## 5. VERIFICATION OF FAST ALGORITHM

In order to illustrate the validity and reliability of our method we realized the phase state reconstruction of Rössler and Lorenz systems. According to [2] we choose the sufficiently large embedding dimension  $m = 5$  and set delay time  $\tau$  as our selected value in section 3. We calculated the correlation dimension ( $D_v$ ) [15-16] of both the systems and compared the results with theoretical value. The relative errors are both less than 5%, which demonstrates the chosen parameters are reasonable.

Table 1 the calculation results of correlation dimension

System	Sampling interval	Embedding dimension	Delay time $\tau$
Rössler	0.1s	5	1.4
Lorenz	0.01s	5	0.1
System	Calculated value of $D_v$	Theoretical value of $D_v$	Relative error
Rössler	1.9258	2.01	4.2%
Lorenz	2.0330	2.06	1.3%



## 6. CONCLUSION

According to the pseudo-periodicity and ergodicity of chaotic systems, we put forward a fast algorithm to determine delay time  $\tau$  by mutual information. The calculation of mutual information is time-consuming and difficult to realize, which limits its application. Our algorithm makes the use of the property of chaotic system and overcomes the shortcoming of mutual information successfully. We progressively reduce the feasible ranges and step length, which increase the accuracy, until get satisfactory  $\tau$ . This algorithm needs much less computation and can get high reliability, which can be popularized in most chemical industry.

## REFERENCES:

- [1] Liu SL, Zhang JZ, Xu MQ. Wavelet packet and neural network based approach for fault diagnosis of reciprocating compressors. *Oil Field Equipment* 2002; 31:1-3.
- [2] Takens F. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics* 1981;366-381..
- [3] Rolando C, Tim S. Correlation dimension of attractors through interspike intervals *Phys.Rev.E* 1997;287-290.
- [4] Alan W, Jack B.S, Determining Lyapunov exponents from a time series. *Phys.D.* 1985;16:285-317.
- [5] Gans R F. When is cutting chaotic ?[J]. *Journal of sound and vibration. Journal of sound and vibration.* 1995; 188: 75-83.
- [6] Qian SX, Jiao WD, Yang SX. Method of Independent Component Analysis and Its Application to Fault Diagnosis. *Proceedings of the CSEE*; Vol.26 No.5 Mar.2006.
- [7] Li CS, Wang YN, Chen GH. Combination Forecast Model Based on Mutual Information and Its Application to Power Load Prediction. *Journal of Hunan University(Natural Sciences)*; Vol.35 No.9 Sep.2008.
- [8] Albano A M, Muench J, Schwartz C. Singular-value decomposition and the Grassberger-Procaccia algorithm *Physical Review A*, 1988;38: 3017-3026.
- [9] Rosenstein M T, Collins J J, De Luca C J. Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D* 1994; 73: 82-98.
- [10] Fraser A M, Swinney H L. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 1986;33: 1134-1140.
- [11] Wang HY, Sheng ZH. Choice of the Parameters for the Phase Space Reconstruction of Chaotic Time Series. *Journal of Central South University*, 2000;30: 113-117 .
- [12] Zhou MQ. The foundation of information theory. Beijing: the publishing house of Beijing University of Aeronautics and Astronautics.1993.
- [13] Liu BZ, Peng JH. *Nonlinear dynamics*. Beijing: Higher Education Press. 2004.
- [14] J.M.Martinerie, A.M.Albano. Mutual information, strange attractors, and the optimal estimation of dimension. *Phys.Rev.A*1992;45: 7058~7064.
- [15] Caostro R, Sauer T. Correlation dimension of attractors through inters pike intervals. *Phys Rev E*1997; 55: 287~291.
- [16] Hentschel H G E, ocaccia I. The infinite number of generalized dimensions of fractals and strange attractors. *Physica D*1983;8: 433~435.