DATA-MINING BASED FAULT DETECTION

Ma Hongguang  Han Chongzhao

(Xi'an Jiaotong University, Xi'an 710049, China)

Wang Guohua  Xu Jianfeng  Zhu Xiaofei

(Research Institute of High Technology, Xi'an 710025, China)

Abstract  This paper presents a fault-detection method based on the phase space reconstruction and data mining approaches for the complex electronic system. The approach for the phase space reconstruction of chaotic time series is a combination algorithm of multiple autocorrelation and T-test, by which the quasi-optimal embedding dimension and time delay can be obtained. The data mining algorithm, which calculates the radius of gyration of unit-mass point around the centre of mass in the phase space, can distinguish the fault parameter from the chaotic time series output by the tested system. The experimental results depict that this fault detection method can correctly detect the fault phenomena of electronic system.

Key words  Chaotic time series; Phase space reconstruction; Data mining; Fault detection

I. Introduction

The process of fault diagnosis is to detect betimes if the abnormal phenomena exist during the operation of the System Under Testing (SUT), and the fault identification is to estimate the type and intensity of fault in case that a fault phenomenon has been found out. The early techniques of fault detection and identification are mainly based on the heuristic models of fault diagnosis and classification[1]. They recur to the practice experiences of engineers and technologists to build up various types of the fault feature vector databases by means of the signal processing technique, the operating states of SUT are detected with the analysis in time or frequency domains, and then compared them with the vectors in the fault feature vector databases, thereout the fault status and type are estimated. The representative methods are the fault-dictionary approach and the fuzzy-logic based fault diagnosis and classification[2]. These kinds of method are simple to implement and efficient for some specific objects, but lack of the theoretic support for the construction of fault feature vector database, and they cannot completely involve all of the fault phenomena, i.e. they usually have a poor generality. It is difficult to find out a functional relation between the changes of system parameters and that of feature vectors. In order to overcome those defects of the heuristic based methods, the model-based fault diagnosis approaches are proposed[3]. These approaches can accurately describe the relationship between the system fault and the internal parameters, and possess a good generality. Unfortunately the effectiveness seriously depends on the precision of math model used to describe the SUT. In fact, the precision

1Manuscript received date: December 29, 2003; revised date: June 29, 2004.
Communication author: Ma Hongguang, born in 1959, male, professor. Xi'an Xingqing Road, Middle Section No.31, Xi'an 710048, China. mhg_xian@163.com
model of a system is hardly obtained in a great deal of application situations. Therefore, the measuring data based nonlinear modeling approaches are presented in recent years, the representatives of which are the Genetic Algorithm (GA) based nonlinear modeling approach[4] and the neural network based nonlinear modeling method[5]. They preferably overcome the defects of aforementioned methods. However they pay the cost of losing the correlativity between the change of system model parameters and that of fault feature vectors.

The method proposed in this paper is based on the chaotic time series analysis and data-mining based fault detection. It effectively overcomes the shortcomings inherited in the all aforementioned approaches. To study the chaos from the chaotic time series begins from the theory of delay coordinates reconstruction of phase space proposed by Packard[6]. As we know, the time evolvement of any variable in the determinate system’s long-term evolvement contains the long-term evolving information of all the system variables. Therefore, we can study the systematic chaos behavior via any mono-variable time series. From the viewpoint of dynamic system theory, a complex electronic system that possesses the chaotic character can be described as

$$
\begin{align*}
    \dot{x} & = f(x, \mu(\Phi), t) \\
    \dot{\Phi} & = \varepsilon g(x, \Phi, t)
\end{align*}
$$

where, $x \in \mathbb{R}^n$ is the state variable of system, and it is the directly observed “fast dynamic” state vector of system. $\Phi \in \mathbb{R}^m$ is a “slow dynamic” state vector that is tightly correlative to the varying of system internal parameters, i.e. the non-observable state vector “hidden” in the “fast dynamic” state vector. $\mu \in \mathbb{R}^s$ is a parameter vector of the function of $\Phi$. $0 < \varepsilon \leq 1$ defines a coefficient of time scale separation between $x$ and $\Phi$. How to find out the “slow dynamic” state vector $\Phi$ from the “fast dynamic” state vector $x$ well and truly is the key point that the system fault diagnosis needs to solve.

II. Phase-space Reconstruction and Data-mining Based Algorithm[7]

1. Phase-space reconstruction method of chaotic time series

Generally speaking, the right-hand item $f(x, \mu(\Phi), t)$ of Eq.(1) does not have the analytic expression due to the complexity of nonlinear complex electronic system. Hence, one should construct the proximal equations that accord with the original circuit system in practice, to realize the accurate real-time measurement of system states.

Suppose $X = \{x_i\}, i = 1, 2, \cdots, N$ is the sampling value of $f(x, \mu(\Phi), t)$ within a certain interval. The theorem of Takens[8] proofs that the regular trajectories (attractor) can be recovered in an embedding space if a suitable embedding dimension $m$ is found, i.e. $m \geq 2d + 1$, $d$ is the order of the original system. Therefore, the trajectories in reconstructed space $\mathbb{R}^m$ keep a differential homeomorphism with the original dynamic system, namely, the chaotic time series $X$ can be constructed to the phase points $\{x_j\}$ with $M = N - (m-1)\tau$ dimensions, where $\tau$ is the time delay and $j = 1, 2, \cdots, M$, $x_j = [x_j \ x_{j+1} \ \cdots \ x_{j+(m-1)\tau}], \ x_j \in \mathbb{R}^m$.

The interconnections of phase points describe the evolving locus of system in $M$-dimensional space.

The reconstruction of phase space is the groundwork for computing the system invariants. It is very important to select a suitable pair of embedding dimension $m$ and time delay $\tau$ when performing the phase space reconstruction. There are two different approaches: one is that $m$ and $\tau$ are not correlated with each other, i.e. they can be selected