Fault Diagnosis for a Class of Sampled-data Systems via Deterministic Learning

Shaoyan Wang[1], Tianrui Chen[1] and Cong Wang[2]

1. School of Automation, Guangdong University of Technology, Guangzhou 510006, China
E-mail: wsy950409@gmail.com
E-mail: trchen@gdut.edu.cn

2. School of Automation, South China University of Technology, Guangzhou 510640, China
E-mail: wangcong@scut.edu.cn

Abstract: In this paper, an approach for rapid detection of small oscillation faults for a class of sampled-data nonlinear system is proposed based on deterministic learning theory. Firstly, based on the Euler approximate discrete time model of the continuous-time system, a training estimator is constructed to learn the normal mode and the fault modes. By using the deterministic learning theory and stability results of linear discrete time-varying systems, the modeling uncertainty and the fault functions are locally-accurately approximated. The obtained knowledge are stored in constant RBF networks. Secondly, by utilizing learned knowledge, a set of estimators are constructed. One estimator represents the normal mode, whereas the others represent the fault modes. The average L_1 norms of the residuals are taken as the measure of the differences of dynamics between the monitored system and the estimators. The occurrence of a oscillation fault can be rapidly detected according to the smallest residual principle. Simulation study is included to demonstrate the effectiveness of the approach.

Key Words: Fault detection, sampled-data systems, deterministic learning, neural networks, persistent excitation (PE) condition

1 INTRODUCTION

Motivated by the development of digital technology, considerable efforts have been devoted to the problem of fault diagnosis for sampled-data (SD) systems[1]. One common idea to develop the SD fault detection and isolation (FDI) schemes is based on indirect approaches, in which either a continuous-time FDI is designed and then implemented digitally, or a discrete-time FDI is designed and implemented for the discretized process [1]-[4]. The indirect approaches ignore intersample behavior, and approximations exist, so the detection system designed may not work properly [5, 8]. In recent years, by using the operators to capture the intersample behavior of a system, a direct design approach has been introduced for fault detection in SD systems. In [5]-[7], the parity space, H_∞ optimal and H_2 optimal methods have been developed to design residual generators for sampled-data systems. In [8], by defining norms of sampled systems and the so-called norm invariant transformation, a general framework for sampled-data fault detection has been developed, which offers a convenient tool for both design and analysis. Another main approaches for direct design of SD FDI is the hybrid system approach, which is based on the representation of the system in the form of hybrid discrete/continuous model [9, 10]. For the multirate SD systems, different D/A and A/D converters often work at different rates [11]. The design and analysis of fault diagnosis schemes for multirate systems have also attracted considerable attention [12]-[15].

In the literature of fault diagnosis, rapid detection of small faults is important, since it is helpful for condition-based maintenance, fault prediction, and avoiding catastrophic consequences [16]-[18]. To achieve rapid detection of faults, it is essential to capture the fault dynamics and to process the data in a dynamical manner. Another important problem studied in this area is how to deal with the unstructured modeling uncertainties, which cannot be exactly decoupled from the faults. For SD systems, full decoupling of uncertainty from faults becomes more difficult, even if the uncertainty satisfies the matching condition in the original continuous-time system, because uncertainty does lose matching property after discretization [1, 19]. To solve the above problems, one promising approach is using neural networks to compensate the effects of the uncertainty terms and to approximate the fault functions [18]. Nonetheless, convergence of NN weights to their optimal values cannot be guaranteed due to the difficulties in satisfying the persistent excitation (PE) condition [18, 30].

Recently, a deterministic learning theory was proposed to develop a framework for knowledge acquisition, representation, and utilization in uncertain dynamical environments [21]-[22]. It is shown that, by using localized radial basis function (RBF) NNs, almost any periodic or recurrent trajectory can lead to the satisfaction of a partial PE condition and accurate NN approximation of the system dynamics is achieved in a local region along the periodic or re-

This work is supported by National R&D Program for Major Research Instruments (Grant No. 61527811), by the National Natural Science Foundation of China (Grant No. 61403087), and by Natural Science Foundation of Guangdong Province, China (Grant No. 2015A030313484).

current trajectory. With deterministic learning, fundamental knowledge on system dynamics can be accumulated, stored, and represented by constant RBF networks in a deterministic manner. Based on deterministic learning theory, fault detection and isolation schemes have been developed for a class of nonlinear continuous-time systems in the previous work [26, 27]. The unknown system dynamics underlying normal and fault oscillations are locally-accurately approximated and the learned knowledge is used to detect and isolate the faults. Motivated by the digital implementation of monitoring systems, it is of interest to investigate the problems of how to acquire the knowledge of system dynamics and how to reuse the learned knowledge to achieve rapid FDI by using the sampled data.

In this paper, we present a rapid detection scheme for nonlinear SD systems. We consider a class of nonlinear oscillation systems with unstructured modeling uncertainty. The proposed scheme consists of two phases: the training phase and the diagnosis phase. In the training phase, based on the Euler approximate discrete time model of the continuous-time system, a training estimator is constructed to learn the normal mode and the fault modes. By using the deterministic learning theory and stability results of linear discrete time-varying systems, the modeling uncertainty and the fault functions are locally-accurately approximated. The obtained knowledge are stored in constant RBF networks. In the diagnosis phase, by utilizing learned knowledge, a set of estimators are constructed. One estimator represents the normal mode, whereas the others represent the fault modes. The average L_1 norms of the residuals are taken as the measure of the differences of dynamics between the monitored system and the estimators. The occurrence of a oscillation fault can be rapidly detected according to the smallest residual principle.

The rest of the paper is organized as follow: Section 2 presents problem formulation and preliminary results. Section 3 presents the distributed detection schemes. Section 4 presents the simulation results. Section 5 concludes the paper.

2 PROBLEM FORMULATION AND PRELIMINARIES

2.1 Problem Formulation

Consider a class of oscillation faults generated from the following class of nonlinear systems

$$\dot{x} = f(x,u) + \upsilon(x,u) + d(t) + \beta(t-T_0)\phi^s(x,u) \quad (1)$$

where $x = [x_1, \ldots, x_n]^T \in R^n$ is the state vector of the system, $u = [u_1, \ldots, u_m]^T \in R^m$ is the control input vector, $f(x,u) = [f_1(x,u), \ldots, f_n(x,u)]^T$, $\phi^s(x,u) = [\phi_1^s(x,u), \ldots, \phi_n^s(x,u)]^T$, $\upsilon(x,u) = [\upsilon_1(x,u), \ldots, \upsilon_n(x,u)]^T$ are the continuous nonlinear vector fields, with $f(x,u)$ representing the dynamics of the nominal model, $\upsilon(x,u)$ the modeling uncertainty, and $\phi^s(x,u)$ the deviation in system dynamics due to fault s ($s = 1, \cdots, N$); $d(t) = [d_1(t), \ldots, d_n(t)]^T$ is the vector of noises or disturbances; $\beta(t-T_0) = diag(\beta_1(t-T_0), \cdots, \beta_n(t-T_0))$ represents the fault time profile, with T_0 being the unknown fault occurrence time. When $t \le T_0$, $\beta_i(t-T_0) = 0$, when $t > T_0$, $\beta_i(t-T_0) = 1$, $i = 1, \cdots, n$.

Given a sampling period $T_s > 0$, we assume that the control u is constant during sampling intervals $[kT_s, (k+1)T_s)$ and that the state x is measured at sampling instants kT_s, that is $x(k) := x(kT_s)$. The Euler approximation for the sampled-data presentation of (1) can be expressed as

$$\begin{aligned} x(k+1) =& x(k) + T_s(f(x(k),u(k)) + \upsilon(x(k),u(k)) \\ &+ d(k) + \beta(k-K_0)\phi^s(x(k),u(k))) \end{aligned} \quad (2)$$

where T_s denotes the sampling time, K_0 denotes the fault occurrence time. When $k \le K_0$, $\beta_i(k-K_0) = 0$, when $k > K_0$, $\beta_i(k-K_0) = 1$, $i = 1, \cdots, n$.

Assumption 1 *The system states and inputs are in bounded oscillations for normal and fault modes, i.e., $(x(k),u(k)) \in \Omega$, where $\Omega \subset R^{n+m}$ is a compact set.*

Let $\psi_0(k)$ denote the trajectory of the state $x(k)$, input $u(k)$ of system (2) in normal mode, and let $\psi_s(k)$ denote the one after fault mode s occurs.

The oscillation fault considered in this paper is "small" in the sense that (i) the magnitude of the fault function $\phi_i^s(x,u)$ is not larger than the magnitude of the modeling uncertainty $\eta_i(k) = \upsilon_i(x,u) + d_i(k)$; and (ii) the system trajectory does not change much due to fault, i.e.,

$$\begin{aligned} &dist((x^s(k),u^s(k)),\psi_0(k)) < d_\zeta, \\ &\qquad \forall (x^s(k),u^s(k)) \in \psi_s(k) \end{aligned} \quad (3)$$

where $dist((x^s(k),u^s(k)),\psi_0(k))$ denotes the distance from the point $(x^s(k),u^s(k))$ to the trajectory ψ_0 , $d_\zeta > 0$ is a small constant.

The proposed fault diagnosis approach consists of two phases: the training phase and the diagnosis phase. In the training phase, since the modeling uncertainty $\upsilon(x,u)$ and fault $\phi^s(x,u)$ ($s = [0,1,\cdots,N]$) can not be decoupled from each other, we consider the two terms together as an undivided term, and define $\varphi^s(x,u) := \upsilon(x,u) + \phi^s(x,u)$ as the general fault function (for the simplicity of presentation, the normal mode is represented by fault mode $s = 0$, with $\phi^0(x,u) = 0$). Moreover, we use advanced filter technique (e.g. [31]) to reduce noises and disturbances. Therefore, $d(k)$ is decreased to $d_\epsilon(k) = [d_{\epsilon 1}(k), \cdots, d_{\epsilon n}(t)]^T$, where $|d_{\epsilon i}(k)| < \bar{d}_{\epsilon i}$, $\bar{d}_{\epsilon i} > 0$ is a small constant. Consequently, for both normal and fault modes, system (2) is described by

$$\begin{aligned} x(k+1) =& x(k) + T_s(f(x(k),u(k)) + \varphi^s(x(k),u(k)) \\ &+ d_\epsilon(k)) \end{aligned} \quad (4)$$

The objective of the training or learning phase is to identify or approximate the general fault function $\varphi^s(x,u)$ to a desired accuracy via deterministic learning.

The distance $dist((x^s(k),u^s(k)),\psi_0(k))$ is defined as

$$dist(((x^s(k),u^s(k)),\psi_0) := \min_{(x^0(k),u^0(k))\in\psi_0} \{\|(x^s(k),u^s(k)),(x^0(k),u^0(k))\|\},$$

where $\|\cdot\|$ is the Euclidean norm.

In the diagnosis phase, we assume that $d(t)$ is bounded, i.e., $|d_i(t)| < \bar{d}_i$, where $\bar{d}_i > 0\,(i = 1, \cdots, n)$ is a constant, and the Euler approximation for the sampled-data presentation of the monitored system is described by

$$\begin{aligned} x(k+1) =& x(k) + T_s(f(x(k), u(k)) + \upsilon(x(k), u(k)) \\ &+ d(k) + \beta(k - K_0)\phi'(x(k), u(k))) \end{aligned} \tag{5}$$

where $\phi'(x, u)$ represents the deviation in the system dynamics due to a fault, which is similar to one of the training fault functions $\phi^s(x, u)$, i.e.,

$$\max_{(x,u)\in\Omega} \left|\phi'_i(x, u) - \phi^s_i(x, u)\right| < \varepsilon^*_i, i = 1, \cdots, n \tag{6}$$

where ε^*_i is the similarity measure between the two fault modes.

The objective of the test or diagnosis phase is to rapidly detect the fault occurred by utilizing the learned knowledge obtained in training phase.

2.2 Localized RBF Networks and Partial PE Condition

The RBF networks belong to a class of linearly parameterized networks, and can be described by

$$f_{nn}(Z) = W^T S(Z) = \sum_{i=1}^{Q} w_i s_i(Z) \tag{7}$$

where $Z \in \Omega_Z \subset R^q$ is the input vector, $W = [w_1, \cdots, w_Q]^T$ is the weight vector, Q is the NN node number, $S(Z) = [s_1(Z), \cdots, s_Q(Z)]^T$ is the vector of radial basis functions (RBFs). It has been shown (e.g. [29]) that for any continuous function $f(Z) : \Omega_Z \to R$ where $\Omega_Z \subset R^q$ is a compact set, and the RBF network $W^T S(Z)$ where the node number Q is sufficiently large, there exists an ideal constant weight vector W^* such that for each $\epsilon^* > 0$,

$$f(Z) = W^{*T} S(Z) + \epsilon(Z),\ \forall Z \in \Omega_Z \tag{8}$$

where $\epsilon(Z) < \epsilon^*$ is the approximation error. For the bounded trajectory $Z_\zeta(t)$ within the compact set Ω_Z, $f(Z)$ can be approximated by using the neurons located in a local region along the trajectory:

$$f(Z) = W^{*T}_\zeta S_\zeta(Z) + \epsilon_\zeta \tag{9}$$

where $S_\zeta(Z) \in R^{Q_\zeta}$ is a subvector of $S(Z)$ (see [21] for more details), $Q_\zeta < Q$, $W^*_\zeta \in R^{Q_\zeta}$, and ϵ_ζ is the approximation error, $\epsilon_\zeta = O(\epsilon)$.

For RBF networks, an assumption regarding $S(X(k))$ is given as follows.

Assumption 2 *There exists a constant $S_M > 0$ such that for all $k \geq 0$, the following bound is satisfied:*

$$\|S(X(k))\| \leq S_M \tag{10}$$

2.3 Stability of Linear Discrete Time-varying Systems

Consider the linear discrete time-varying system given by

$$x(k+1) = A(k)x(k) \tag{11}$$

with $A(k)$ are appropriately dimensioned matrices.

The following results are summarized in [23]:

Lemma 1: Define $\psi(k_1, k_0)$ as the state-transition matrix corresponding to $A(k)$ for system (11), i.e., $\psi(k_1, k_0) = \Pi_{k=k_0}^{k_1-1} A(k)$. Then if $||\psi(k_1, k_0)|| \leq 1$, $\forall k_1, k_0 \geq 0$, system (11) is exponentially stable.

Lemma 2: If $A(k) = I - \alpha S(x(k))S^T(x(k))$ in (11), where $0 < \alpha < 2$ and $S(x(k))$ is a vector of basis functions, and there is an $L > 0$ such that $\sum_{k=k_0}^{k_1+L-1} S(x(k))S^T(x(k)) > 0$ for all k. Then Lemma 1 guarantees the exponential stability of the system (11).

The stability condition needed in Lemma 2 is actually the PE condition [24]. The following definition of the PE condition was presented in [25], which is suitable for RBF networks identification in both continuous and discrete cases:

Definition 1 *Let μ be a positive, Σ-finite Borel measure on $[0, \infty)$. A continuous, uniformly bounded, vector-valued function $S : [0, \infty) \to R^m$ is persistently exciting, if there exist positive constants α_1, α_2, and T_0 such that*

$$\begin{aligned} \alpha_1\|c\|^2 \geq \int_{t_0}^{t_0+T_0} |S(\tau)^T c|^2 d\mu(\tau) \geq \alpha_2\|c\|^2, \\ \forall t_0 \geq 0, \forall c \in R^N \end{aligned} \tag{12}$$

holds for every constant vector $c \in R^N$.

Concerning the relationship between the recurrent trajectories and the PE condition, the following result is given in [22].

Theorem 1 *Consider any recurrent trajectory $Z(t)$. Assume that $Z(t)$ is a continuous map from $[0, \infty)$ into a compact set $\Omega_Z \subset R^q$, and $\dot{Z}(t)$ is bounded within Ω_Z. Then, for the RBF network $W^T S(Z)$ with centers placed on a regular lattice (large enough to cover the compact set Ω_Z), the regressor subvector $S_\zeta(Z(t))$, as defined in (9), is persistently exciting almost always.*

Consequently, we have the following result.

Corollary 1 *Consider any recurrent sequence $Z(k)$. Assume that $Z(k)$ is a discrete map from $[0, \infty)$ into a compact set $\Omega_Z \subset R^q$. For the RBF network $W^T S(Z)$ (8) with centers placed on a regular lattice (large enough to cover the compact set Ω), $S_\zeta(Z(k))$ (9) is persistently exciting in the sense of (12) almost always.*

3 DETECTION OF OSCILLATION FAULTS FOR SAMPLED DATA SYSTEMS

In this section, we present a two-phase detection approach for sampled-data systems based DL theory. A rigorous analysis of the performance of the proposed detection scheme is also provided.

3.1 Training Phase

Consider the oscillation system (4) undergoing a fault mode s ($s \in \{0, 1, \cdots, N\}$, $s = 0$ denotes the normal mode). We use the following training estimator:

$$
\begin{aligned}
\hat{x}_i(k+1) &= x_i(k) + a_i(\hat{x}_i(k) - x_i(k)) \\
&+ T_s \left(f_i(x(k), u(k)) + \hat{W}_i^{sT}(k+1) S_i(x(k), u(k)) \right)
\end{aligned} \tag{13}
$$

where $\hat{x}_i(k)$ is the state of the training estimator, $x_i(k)$ is the state of system (4), $|a_i| < 1$ is the estimator gain, $\hat{W}_i^{sT}(k+1)S_i(x(k), u(k))$ is the localized RBF networks used to approximate the general fault function $\varphi_i^s(x(k), u(k))$, $\hat{W}_i^s(k) = [\hat{w}_{i1}^s(k), \cdots, \hat{w}_{iQ}^s(k)]^T$ is the weight vector, Q is the NN node number, $S_i(x(k), u(k)) = [s_{i1}(x(k), u(k)), \cdots, s_{1Q}(x(k), u(k))]^T$ is the vector of radial basis functions (RBFs).

Define the following error vectors:

$$
\begin{aligned}
e_i(k) &= \hat{x}_i(k) - x_i(k) \\
\tilde{W}_i^s(k) &= \hat{W}_i^s(k) - W_i^{s*}
\end{aligned}
$$

where e_i is the state tracking error, W_i^{s*} is the ideal weight vector defined by (8).

To simplify the discussion, let $e_{ik} = e_i(k)$, $\hat{W}_{ik}^s = \hat{W}_i^s(k)$, $S_{ik} = S_i(x(k), u(k))$, $d_{i\epsilon k} = d_{i\epsilon}(k)$.

From (13), we have

$$
\begin{aligned}
e_{i(k+1)} =& a_i e_{ik} + T_s(\hat{W}_{i(k+1)}^{sT} S_{ik} - \varphi_{ik}^s - d_{i\epsilon k}) \\
=& a_i e_{ik} + T_s(\tilde{W}_{i(k+1)}^{sT} S_{ik} - \epsilon_{ik}^s - d_{i\epsilon k})
\end{aligned} \tag{14}
$$

The NN weight updating law is given by:

$$
\hat{W}_{i(k+1)}^s = \hat{W}_{ik}^s - \frac{\alpha(e_{ik} - a_i e_{i(k-1)}) S_{i(k-1)}}{1 + S_{i(k-1)}^T S_{i(k-1)}} \tag{15}
$$

where $0 < \alpha < 2/T_s$ is the learning gain.

The weight estimate error equation can be written as

$$
\tilde{W}_{i(k+1)}^s = \tilde{W}_{ik}^s - \frac{\alpha(e_{ik} - a_i e_{i(k-1)}) S_{i(k-1)}}{1 + S_{i(k-1)}^T S_{i(k-1)}} \tag{16}
$$

It has been shown in [28] that with update law (15), the tracking error e_{ik} becomes small after a finite step. In following, we will show that along the trajectory φ_s, the neural-weight estimates $\hat{W}_{i\zeta k}^s$ converge to small neighborhoods of their optimal values $W_{i\zeta}^{s*}$, and a locally-accurate approximation for the unknown $\varphi_i^s(x(k), u(k))$ is obtained by $\hat{W}_{ik}^{sT} S_{ik}$.

By using the localization property of RBF networks, the weight estimate error along the state trajectory is given by

$$
\begin{aligned}
\tilde{W}_{i\zeta(k+1)} =& \tilde{W}_{i\zeta k} - \frac{\alpha(e_{ik} - a_i e_{i(k-1)}) S_{i\zeta(k-1)}}{1 + S_{i(k-1)}^T S_{i(k-1)}} \\
=& \left[I - \frac{\alpha S_{i\zeta(k-1)} S_{i\zeta(k-1)}^T}{1 + S_{i(k-1)}^T S_{i(k-1)}} \right] \tilde{W}_{i\zeta k} \\
&- \frac{\alpha S_{i\zeta(k-1)}}{1 + S_{i(k-1)}^T S_{i(k-1)}} \cdot (\epsilon_{i\zeta}^s - d_{i\epsilon(k-1)})
\end{aligned} \tag{17}
$$

$$
\tilde{W}_{i\bar{\zeta}(k+1)} = \tilde{W}_{i\bar{\zeta}k} - \frac{\alpha(e_{ik} - a_i e_{i(k-1)}) S_{i\bar{\zeta}(k-1)}}{1 + S_{i(k-1)}^T S_{i(k-1)}} \tag{18}
$$

where the subscripts $(\cdot)_\zeta$ and $(\cdot)_{\bar{\zeta}}$ are used to stand for the terms related to the regions close to and away from the trajectory ψ_s, $\epsilon_{i\zeta}$ is the approximation error defined by (9).

Let $S_k' = S_{i\zeta(k-1)} / \sqrt{(1 + S_{i(k-1)}^T S_{i(k-1)})}$. The weight estimate error system (17) can be written as

$$
\tilde{W}_{i\zeta(k+1)} = \left[I - \alpha S_k'^T S_k' \right] \tilde{W}_{i\zeta k} - d_k' \tag{19}
$$

where $d_k' = \epsilon_{i\zeta}^s + d_{i\epsilon(k-1)}$.

When the system (4) is undergoing a recurrent trajectory ψ_s, according to Theorem 1, the regressor subvector $S_{\zeta i}(x, u)$ satisfies PE, that is there exist $\lambda > 0$ and $k_1 \geq 1$ integer such that

$$
\lambda_{min}[\sum_{k=k_0}^{k_1} S(x(k)) S^T(x(k))] > \lambda, \forall k_0 \geq 0 \tag{20}
$$

With Assumption 2, we have

$$
\lambda_{min}[\sum_{k=k_0}^{k_1} S_k' S_k'^T] > \frac{\lambda}{1 + S_M^2}, \forall k_0 \geq 0 \tag{21}
$$

which means S_k' satisfies PE. Using Lemma 2, the closed-loop system (19) is globally exponentially stable in the absence of d_k'. Therefore, the NN weight estimate error $\tilde{W}_{\zeta i}^s$ converge to small neighborhoods of zero, with the sizes of the neighborhoods being determined by $\epsilon_{\zeta i}^s$ and $\bar{d}_{\epsilon i}$, both of which are small values.

On the other hand, for the neurons with centers far away from the trajectory φ_ζ, $S_{\bar{\zeta}}(x(k))$ will become very small due to the localization property of Gaussian RBFs. As a result, $W_{\bar{\zeta}(k+1)}$ is updated only slightly and remain bounded. This means that the entire RBF network $\hat{W}_i^{sT} S(x, u)$ can approximate the unknown $\varphi_i^s(x, u)$ along the trajectory ψ_s, i.e.,

$$
\varphi_i^s(x, u) = \hat{W}_i^{sT} S(x, u) + \epsilon_{i1}^s \tag{22}
$$

where $\epsilon_{i1}^s = O(\epsilon_i^s)$. □

Based on the convergence result, we can obtain a constant vector of neural networks $\bar{W}_i^s S_{ik}$, where

$$
\bar{W}_i^s = \frac{1}{k_b - k_a + 1} \sum_{k=k_a}^{k_b} \hat{W}_i^s(k) \tag{23}
$$

where $\{k_a, \ldots, k_b\}$ represent a time segment after the transient process. Therefore, we conclude that accurate approximation of the general fault function $\varphi_i^s(x, u)$ is obtained in a local region Ω_ζ^s along the trajectory ψ_s, i.e.,

$$
\begin{aligned}
\varphi_i^s(x, u) =& \overline{W}_i^{sT} S(x, u) + \xi_i^s(k), \\
& \forall (x, u) \in \Omega_\zeta^s, s \in \{0, 1, \cdots, N\}
\end{aligned} \tag{24}
$$

where $|\xi_i^s(k)| < \xi_i^*$, $\Omega_\zeta^s := \{(x, u) | dist((x, u), \psi^s) < d_\zeta\}$, $d_\zeta > 0$ and $\xi_i^* > 0$ are constants, d_ζ characterizes the size of the NN approximation region, ξ_i^* is the NN approximation error within Ω_ζ^s.

3.2 Diagnosis phase

Consider the monitored system (5). By utilizing the learned knowledge obtained in the training phase, a bank of $N+1$ estimators is firstly constructed for the trained normal mode and oscillation faults as follows:

$$\begin{aligned}\bar{x}(k+1) =& x(k) + B(\bar{x}(k) - x(k)) \\ &+ T_s(f(x(k), u(k)) + \overline{W}^{sT} S(x(k), u(k)))\end{aligned} \tag{25}$$

where $s = [0, 1, \cdots, N]$ is used to stand for the sth estimator, $\bar{x}^s = [\bar{x}^s_1, \ldots, \bar{x}^s_n]^T$ is the state of the estimator, x is the state of the monitored system (1), $B = diag\{b_1, \ldots, b_n\}$ is a diagonal matrix which is kept the same for all normal and fault estimators, $0 < b_i < 1$.

By comparing the monitored system (5) with the set of $N+1$ estimators (25), we obtain the following residual systems:

$$\begin{aligned}\tilde{x}^s_i(k+1) = b_i \tilde{x}^s_i(k) + T_s(&\overline{W}^{sT}_i S_k - \upsilon_{ik} \\ &- \beta_i(k - K_0)\phi'_{ik} - d_{ik})\end{aligned} \tag{26}$$

where $\tilde{x}^s_i = \bar{x}^s_i - x_i$ is the state estimation error (residual), $i = 1, \ldots, n, s = 0, 1, \cdots, N$, $S_k = S(x(k), u(k))$, $\upsilon_{ik} = \upsilon_i(x(k), u(k))$, $\phi'_{ik} = \phi'_i(x(k), u(k))$, $d_{ik} = d_i(k)$.

The following average L_1 norm is used to make decision

$$\|\tilde{x}^s_i(k)\|_1 = \frac{1}{K+1} \sum_{k_\tau = k-K}^{k} |\tilde{x}^s_i(k_\tau)|, k \geq K \tag{27}$$

where K is the period of the monitored oscillation system. Based on the smallest residual principle, we have the following decision schemes of fault detection.

Fault detection decision scheme: Compare $\|\tilde{x}^s_i(k)\|_1$ with $\|\tilde{x}^0_i(k)\|_1$, $i = 1, \cdots, n$. If, for some $s \in \{1, \ldots, N\}$ and some $i \in \{1, \ldots, n\}$, there exists some finite time k^s_i such that $\|\tilde{x}^s_i(k^s_i)\|_1 < \|\tilde{x}^0_i(k^s_i)\|_1$, then the occurrence of a fault is deduced.

Note that the absolute fault detection time is defined as $K_d = \min\{k^s_i, i \in \{1, \ldots, n\}\}$, and the fault detection time k_d is defined as the difference between the occurrence time K_0 and the absolute fault detection time K_d, i.e., $k_d = K_d - K_0$.

4 SIMULATIONS

4.1 Training Phase

To verify the effectiveness of the proposed schemes, the van der Pol oscillator is considered as the monitored system [18]:

$$\begin{aligned}\dot{x}_1 =& f_1(x_1, x_2) \\ \dot{x}_2 =& f_2(x_1, x_2) + d(t) + \Gamma(t; t_f)\phi^f(x_1, x_2)\end{aligned} \tag{28}$$

where x_1, x_2 are the states, $f_1(x_1, x_2) = x_2$, $f_2(x_1, x_2) = 2\omega\zeta x_2 - \omega^2 x_1 - 2\omega\zeta c x_1^2 x_2$, $d(t) = (sin(5t) + 3cos(7t))/50$ is the noise, and ω, ζ, c are positive constant parameters. The fault function set is described as follow

$$\mathcal{F} \triangleq \left\{\phi^1((x_1, x_2)), \phi^2((x_1, x_2))\right\} = \{\theta_1 sin(x_2), \theta_2 cos(x_2)\}$$

Parameters used in the simulation studies are as follows: $\omega = 0.9, \zeta = 0.6, c = 0.95$. It is assumed there exists a 10% inaccuracy in the parameter ζ. The sampling period h is set as 0.01. The following training identifier is employed to learn the unknown sampled dynamics $\Psi^s_2(x_{1k}, x_{2k})$ in the 2th state equation of the system (28)

$$\begin{aligned}\hat{x}_{2(k+1)} =& x_{2k} + a(\hat{x}_{2k} - x_{2k}) + h(f_2(x_{1k}, x_{2k}) \\ &+ \hat{W}^{sT}_{2(k+1)} S_i(x_{1k}, x_{2k}))\end{aligned} \tag{29}$$

where $s = 0, 1, 2$, $\hat{x}_{2k}$ is the state of the identifier, x_{1k} and x_{2k} are the sampled data generated from the system (28), $h = 0.01$ is the sampling period, $\hat{W}^{sT}_{2\zeta}(k+1) S_{i\zeta}(x_{1k}, x_{2k})$ is the RBF NN with $N = 21 \times 21 = 441$ fixed centers evenly spaced on $[-4.5, 4.5] \times [-4.5, 4.5]$. The widths of neurons are chosen as $\eta_i = 0.45$.

In the diagnosis process, according to (25), a bank of estimators are established as follow,

$$\begin{aligned}\bar{x}^s_{2(k+1)} =& x_{2k} + b_2(\bar{x}^s_{2k} - x_{2k}) + h(f_2(x_{1k}, x_{2k}) \\ &+ \bar{W}^{sT}_2 S(x_{1k}, x_{2k}))\end{aligned} \tag{30}$$

where $b_2 = 0.1$. The estimation error norms are generated by using (27), in which $K = 200$. Fig. 1 shows the simulation results in the case that fault of type 1 occurs with $\theta_1 = 0.28$ and $K_0 = 6000$. From Fig. 1(a), it can be seen that the estimation error $\tilde{x}^0_{2k}$ corresponding to normal mode is small before the fault occurs, and become large after the fault occurs, while the estimation error $\tilde{x}^1_{2k}$ corresponding to fault of type 1 is large before the fault occurs, and become small after the fault occurs. In Fig. 1(b), it is shown that the estimation error norm $\|\tilde{x}^0_{2k}\|_1$ corresponding to normal mode exceeds the estimation error norm $\|\tilde{x}^0_{2k}\|_1$ corresponding to fault of type 1 at the instant $K_d = 6220$ ($T_d = 62.20s$). Thus, the fault of type 1 is detected at the instant $K_d = 6220$ ($T_d = 62.20s$). The FD instant is $k_d = K_d - K_0 = 220$ ($t_d = T_d - T_0 = 2.20s$).

REFERENCES

[1] P. Zhang and S. X. Ding, "On fault detection in linear discrete-time, periodic, and sampled-data systems," *Journal of Control Science and Engineering,* 2008.

[2] J. Chen and R. Patton, Robust model-based fault diagnosis for dynamic systems, Kluwer Academic Publishers, 1999.

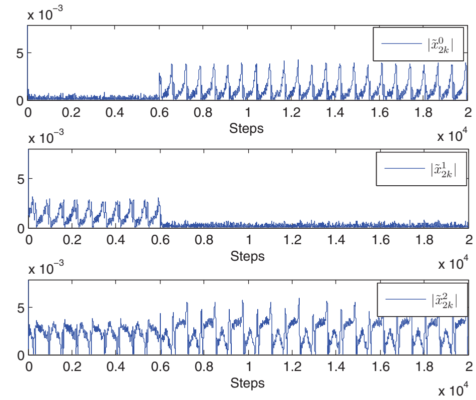
[3] P. M. Frank and X. Ding, "Survey of robust residual generation and evaluation methods in observerbased fault detection systems," J of Process Control, vol. 7, no. 6, pp.403-424, 1997.

[4] J. J. Gertler, *Fault Detection and Diagnosis in Engineering Systems,* Marcel Dekker, 1998.

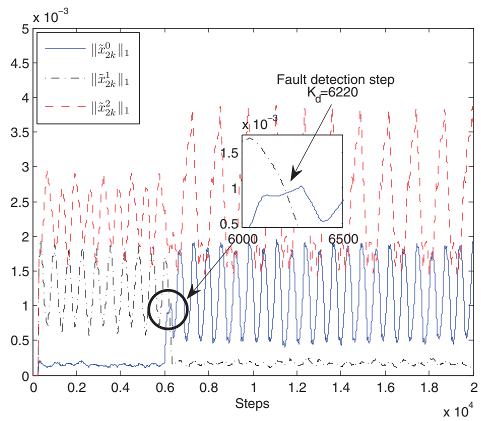
[5] P. Zhang, S. X. Ding, G. Z. Wang, D. H. Zhou, An FDI approach for sampled-data systems, in: Proceedings of the 2001 American Control Conference, Arlington, VA, 2001, pp. 2702-2707.

[6] P. Zhang, S. X. Ding, G. Z. Wang, and D. H. Zhou, "A frequency domain approach to fault detection in sampled-data systems," Automatica, vol. 39, no. 7, pp. 1303-1307, 2003.

[7] P. Zhang, S. X. Ding, G. Z. Wang, and D. H. Zhou, "An approach to fault detection of sampled-data systems," Tech. Rep., Institute of Automatic Control and Complex System-s(AKS), University of Duisburg, Duisburg, Germany, 2003.



(a) Estimation errors $|\tilde{x}^0_{2k}|$, $|\tilde{x}^1_{2k}|$ and $|\tilde{x}^2_{2k}|$.



(b) Estimation error norms $\|\tilde{x}^0_{2k}\|_1$, $\|\tilde{x}^1_{2k}\|_1$ and $\|\tilde{x}^2_{2k}\|_1$.

Figure 1: Detection of the fault of type 1 in the SD system.

[8] I. Izadi, T. W. Chen, and Q. Zhao, "Norm invariant discretization for sampled-data fault detection," Automatica, vol. 41, no. 9, pp. 1633-1637, 2005.

[9] P. Shi, C. E. de Souza, L. H. Xie. "Bounded real lemma for linear systems with finite discrete jumps," International Journal of Control, vol. 66, no. 1, pp. 145-160, 1997.

[10] A. Qiu, C. Wen, B. Jiang, A hybrid system approach to robust fault detection for a class of sampled-data systems, Acta Automatica Sinica, vol. 36, no. 8, pp. 1182-1188, 2010.

[11] T. Chen, B. Francis, Optimal Sampled-Data Control Systems, Springer, New York, 1995.

[12] P. Zhang, S. X. Ding, G. Z. Wang, D. H. Zhou, "Fault detection for multirate sampled-data systems with time delays," International Journal of Control, vol. 75, no. 18, pp. 1457-1471, 2002.

[13] M. Y. Zhong, H. Ye, S. X. Ding, and G. Z. Wang, "Observer-based fast rate fault detection for a class of multirate sampled-data systems," IEEE Transactions on Automatic Control, vol. 52, no. 3, pp. 520-525, 2007.

[14] I. Izadi, Q. Zhao, T. Chen, "An optimal scheme for fast rate fault detection based on multirate sampled data," Journal of Process Control, vol. 15, no. 3, pp. 307-319, 2005.

[15] I. Izadi, Q. Zhao, and T. W. Chen, "An H_∞ approach to fast rate fault detection for multirate sampled-data systems," Journal of Process Control, vol. 16, no. 6, pp. 651-658, 2006.

[16] Z. Xu, Y. Ji, and D. Zhou, "A new real-time reliability prediction method for dynamic systems based on on-line fault prediction," *IEEE Transactions on Reliability*, vol. 58, no. 3, pp. 523-538, Sep. 2009.

[17] M. A. Demetriou, M. M Polycarpou, "Incipient fault diagnosis of dynamical systems using online approximators," *IEEE Transactions on Automatic Control*, vol. 43, no. 11, pp. 1612-1617, Nov. 1998.

[18] X. Zhang, M. M. Polycarpou, and T. Parisini,"A robust detection and isolation scheme for abrupt and incipient faults in nonlinearsystems," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 576-593, Apr. 2002.

[19] Silvia Pettinari, "Fault diagnosis techniques for linear sampled data systems and a class of nonlinear systems",

[20] X. Zhang, M. M. Polycarpou, T. Parisini. "Fault diagnosis of a class of nonlinear uncertain systems with Lipschitz nonlinearities using adaptive estimation," Automatica, 46, 290-299, 2010.

[21] C. Wang and D. J. Hill, "Deterministic learning and rapid dynamical pattern recognition," *IEEE Transactions on Neural Networks*, vol. 18, no. 3, pp. 617-630, May 2007.

[22] C. Wang and D. J. Hill, *Deterministic learning theory for identification, recognition and control*, CRC Press, 2009.

[23] S. Jagannathan and F. L. Lewis, "Multilayer Discrete-Time Neural-Net Controller with Guaranteed Performance," *IEEE Transactions on Neural Networks,* vol. 7, no. 1, pp. 107-130, Jan. 1996.

[24] N. Sadegh, "A perceptron network for functional identification and control of nonlinear systems," *IEEE Trans. Neural Networks,* vol. 4, no. 6, pp. 982-988, Nov. 1993.

[25] A. J., Kurdila, F. J. Narcowich and J. D. Ward, "Persistancy of excitation in identification using radial basis function approximants," *SIAM J. Control and Optimization*, vol. 33, no. 2, pp. 625-642, 1995.

[26] C. Wang and T. Chen, "Rapid detection of small oscillation faults via deterministic learning," *IEEE Transactions on Neural Networks* vol. 22, no. 8, pp. 1284-1296, Aug. 2011.

[27] T. Chen and C. Wang "Rapid isolation of small oscillation faults via deterministic learning," *International Journal Of Adaptive Control And Signal Processing*, in press.

[28] C. Yuan and C. Wang, "Design and performance analysis of deterministic learning of sampled-data nonlinear systems", Science China.

[29] M. J. D. Powell, "The theory of radial basis function approximation in 1990," in *Advances in Numerical Analysis II: Wavelets, Subdivision, Algorithms, and Radial Basis Functions*, W. A. Light, Ed. London, U.K.: Oxford Univ. Press, 1992, pp. 105-210.

[30] D. Gorinevsky, "On the persistancy of excitation in radial basis function network identification of nonlinear systems," *IEEE Transactions on Neural Networks,* vol. 6, no. 5, pp. 1237-1244, Sep. 1995.

[31] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[32] S. R. Samantaray, P. K. Dash, S. K. Upadhyay, "Adaptive Kalman filter and neural network based high impedance fault detection in power distribution networks," *Electrical Power and Energy Systems,* vol. 31, no. 3, pp. 167-172, May 2009.

[33] F. Da, "Decentralized sliding mode adaptive controller design based on fuzzy neural networks for interconnected uncertain nonlinear system," *IEEE Transactions on Neural Networks,* vol. 11, no. 6, pp. 1471-1480, Nov. 2000.