Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 11 (2013) 68 – 73

www.elsevier.com/locate/procedia

2nd International Through-life Engineering Services Conference

Model-based intermittent fault detection

Tabassom Sedighi [a], Paul Phillips* [a], Peter d. Foote [a]

[a]EPSRC Centre for Innovative Manufacturing in Through-life Engineering Services, Cranfield University, MK 43 0AL, UK

* Corresponding author. E-mail address: p.phillips@cranfield.ac.uk

Abstract

Development of reliable fault diagnostics for intermittent failure modes are an important tool to adequately deal with realistic failure behavior within complex systems. A large proportion of previous work utilizing model-based fault diagnostics has focused on persistent faults and often neglects the case where the system intermittently switches between a faulty and non-faulty behavior at discrete random intervals. Such intermittent behavior complicates the diagnostics task, with difficulties in detecting and isolating intermittent faults, which occur with low frequency but yet at high enough frequency to be unacceptable. Accurate assessment of intermittent failure probabilities is critical to diagnosing and repairing equipment and requires the development of models to describe the dynamics of the intermittent failure. This paper presents an overall framework for detecting sensor faults, through the use of nonlinear unknown input observers which are applicable to both persistent and intermittent faults. The work presented demonstrates the detection capabilities of the approach through the use of robust residuals insensitive to system uncertainties and the application of adaptive thresholds.

Keywords: Intermittent Faults; No Fault Found, Model-Based Fault Detection;Robot-Arm Modelling; Through-life Engineering.

1. Introduction

Faults are generally categorized according to whether they have developed slowly during the operation of a system usually characteristic of gradual component wear (incipient fault); arisen suddenly like a step change as a result of a sudden breakage (abrupt faults); or accrued in discrete intervals attributed to component degradation or unknown system interactions (intermittent faults).

Intermittent faults can manifest in any system, mechanical or electronic, in an unpredictable manner, and if left unattended over time they will evolve into serious and persistent faults. The assumed unpredictability of an intermittent fault means that it cannot be easily predicted, detected nor is it necessarily repeatable during maintenance testing, thus faults of this nature raise many concerns in the realm of Through-life Engineering of products. However, an intermittent fault, which is missed during standardized maintenance testing, by its very definition will reoccur at some time in the future.

The intermittent fault case therefore poses an ever increasing challenge in the maintenance of electronic, mechanical and hydraulic equipment. A substantial portion of malfunctions attributed to intermittent faults will test well and will be categorized as No Fault Found (NFF) [1].

When the fault is not intermittent and the fault symptoms are consistent (hard fault), it is not difficult to isolate and repair. However, a fault that persists for a very short duration and manifests itself intermittently and only during a particular set of operational stresses can be extremely difficult to identify and isolate. In general, the intermittent fault typically tends to worsen with time, until it eventually becomes substantial enough that it can be detected with conventional test equipment’s [2]. Hence, developing the capability for early detection and isolation of the intermittent fault will help to avoid major system breakdowns [3].

doi:10.1016/j.procir.2013.07.065

A fault within a system is described as an external input that causes the behavior of a system to deviate from a pre-defined performance threshold. Faults can occur in the actuators, process components or the sensors. Sensor faults are of particular importance as intermittent behavior in sensors can lead to unknown and undesirable behaviors in the system. The impact of sensor faults could be that the system fails to perform its function, or results in a catastrophic mechanical failure.

Model-based fault detection scheme can be powerful tools in determining sensor and actuator faults. The concept is to compare the behavior of an actual process to that of a nominal fault-free model of the process driven by the same input signals. Model-based approaches are more powerful than data-driven signal-processing-based approaches [4, 5], because they rely much more upon physical knowledge of the process and its interactions whereas signal processing techniques rely on large quantities of data to be recorded that may not be practical.

A model-based fault detection scheme consists of two main stages: residual generation and residual evaluation. The objective of designing residuals is to define a signal that can be compared to the appropriate measurements and estimations and then evaluated for possible presence of faults [6]. The purpose of this paper is to demonstrate the development and application of a model-based fault detection method, applied to an intermittent fault case, using a nonlinear observer and adaptive threshold techniques. The method developed in this paper, uses a nonlinear robot-arm model as an example but the approach is extendable for all electronic, mechanical and hydraulic nonlinear systems [7].

The paper is organized as follows: Section 2 presents the mathematical description of the nonlinear system of interest. The design of the observer and residual along with the numerical example and simulation results are addressed in sections 3 and 4 respectively. Conclusions and on-going work are discussed in section 5.

2. System Description

Consider the class of nonlinear systems defined by the state-space form

$$\dot{x}(t) = h_x(x,u,g_s,u_d)$$
$$y(t) = h_y(x,f_i)$$
(1)

If the nonlinear function $y(t) = h_y(x,f_i)$ is differentiable with respect to the state x, then this class of the system may be expressed in terms of a linear unforced part, and nonlinear state dependent controlled part [8].

$$\dot{x} = Ax(t) + Bu(t) + Sg_s(x,u,t) + Du_d \qquad (2)$$
$$y(t) = Cx(t) + K_i f_i(t)$$

Where x,u,y present state, input and output vectors respectively. A,B,C,S and K_i are the known distribution matrix and f_i is the corresponding intermittent fault signal. D is the known distribution matrix of the uncertainty and u_d is an unknown bounded vector which describes the uncertainty input and/or any kind of modeling uncertainty such as noise, parameter variation or other time-varying terms. This paper considers general nonlinearities that depend on unmeasured states, but for illustration, a nonlinearity of the form $g_s(x,u,t)$ has been included in the design procedure. To illustrate the application of the results obtained in sections 2-4, consider the robot arm system shown 2-4, consider the robot arm system shown in Figure 1, where two arms are connected together in series [9].

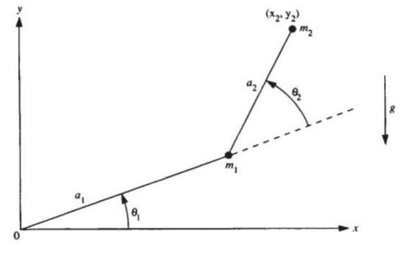

Figure 1: Two-link robot arm

To obtain the equation of motion for a two-link robot arm the Euler-Lagrangian equation is used:

$$L(\theta,\dot{\theta}) = T(\theta,\dot{\theta}) - P(\theta) \qquad (3)$$

where $T(\theta,\dot{\theta})$ and $P(\theta)$ represent the kinetic and the potential energies respectively. Prior to motion equation design, the following assumption is made.

Assumption 1:

The robot arms (links) are rigid, massless and constant.

- The joints are point mass.

The kinetic and potential energies for links 1 and 2 are obtained as follows

$$K_1 = \frac{1}{2} m_1 l_1^2 \dot{\theta}_1^2$$
$$P_1 = m_1 g l_1 \sin\theta_1$$
$$K_2 = \frac{1}{2} m_2 \left[l_1^2 \dot{\theta}_1^2 + l_2^2 (\dot{\theta}_1 + \dot{\theta}_2) + 2 l_1 l_2 \cos(\theta_2) \dot{\theta}_1 (\dot{\theta}_1 + \dot{\theta}_2) \right]$$
$$P_2 = m_2 g \left[l_1 \sin\theta_1 + l_2 \sin(\theta_1 + \theta_2) \right]$$

Substituting the above kinetic and potential energies for both links into equation (3) and taking its derivative

$$\frac{\partial L}{\partial \theta_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}_i} = U_i \qquad (4)$$

After some algebraic manipulation, the equation of motion for the system under investigation will find the following form

$$M(\theta)\ddot{\theta} + C_c(\theta,\dot{\theta})(\dot{\theta}) + K(\theta) + F(\dot{\theta}) + U_d = U \tag{5}$$

where $M(\theta)$, $C_c(\theta,\dot{\theta})$, and $K(\theta)$ represent the manipulator inertia, centrifugal and coriolis force and the gravity matrices respectively with

$$M_{11}(\theta) = (m_1 + m_2)l_1^2 + m_2 l_2^2 + 2m_2 l_1 l_2 \cos(\theta_2)$$
$$M_{12}(\theta) = m_2 l_2^2 + m_2 l_1 l_2 \cos(\theta_2)$$
$$M_{21}(\theta) = m_2 l_2^2 + m_2 l_1 l_2 \cos(\theta_2)$$
$$M_{22}(\theta) = m_2 l_2^2$$

$$C_{c11}(\theta,\dot{\theta}) = -m_2 l_1 l_2 \sin(\theta_2)(2\dot{\theta}_1\dot{\theta}_2 + \dot{\theta}_2^2)$$
$$C_{c21}(\theta,\dot{\theta}) = -m_1 l_1 l_2 \theta_1^2 \sin(\theta_1)$$

$$K_{11}(\theta) = (m_1 + m_2) g l_1 \cos(\theta_1) + m_2 g l_2 \cos(\theta_1 + \theta_2)$$
$$K_{21}(\theta) = m_2 g l_2 \cos(\theta_1 + \theta_2)$$

$U_d^T = [u_{d1}, u_{d2}, \cdots, u_{dn}] \leq d$ Represents the bounded disturbance where d is a positive real number and $U^T = [u_1, u_2, \cdots, u_n]$ is the control input of the system. The matrix $F(\dot{\theta})$ describes the friction in the system and takes the following form:

$$\begin{cases} F_{11}(\dot{\theta}) = \beta_1\dot{\theta}_1 + \alpha_1 \operatorname{sgn}(\dot{\theta}_1) \\ F_{21}(\dot{\theta}) = \beta_2\dot{\theta}_2 + \alpha_2 \operatorname{sgn}(\dot{\theta}_2) \end{cases} \tag{6}$$

$\beta\dot{\theta}$ and $\alpha \operatorname{sgn}(\dot{\theta})$ represent the joints viscous and coulomb frictions respectively where Signum function $\operatorname{sgn}(x)$ is defined as

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{for} \quad x < 1 \\ 0 & \text{for} \quad x = 0 \\ +1 & \text{for} \quad x > 0 \end{cases} \tag{7}$$

The coefficients $\beta_1, \beta_2, \alpha_1, \alpha_2$ in (6) must be measured practically and cannot be found through calculation. Rougher surfaces have higher effective values. Finally the state-space equations of the considered robot arm will find the following form

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}\begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} + \begin{pmatrix} C_{c11}(\theta,\dot{\theta}) \\ C_{c21}(\theta,\dot{\theta}) \end{pmatrix}\begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{pmatrix} + \begin{pmatrix} K_{11}(\theta_1) \\ K_{21}(\theta_2) \end{pmatrix} + \begin{pmatrix} F_{11}(\dot{\theta}_1) \\ F_{21}(\dot{\theta}_2) \end{pmatrix} + \begin{pmatrix} U_{d1} \\ U_{d2} \end{pmatrix} = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \tag{8}$$

Since matrix M (θ) is a symmetric-positive-definite matrix (s.p.d), hence its inverse, M^{-1} (θ) exists. Equivalently equation (8) could be rewritten as follows

$$\ddot{\theta} = -M^{-1}(\theta)(C_c(\theta.\dot{\theta}) + K(\theta) + U_d - U)$$

Where

$$\theta_1 = x_1 \quad \dot{\theta}_1 = x_2 \quad \theta_2 = x_3 \quad \dot{\theta}_2 = x_4 \tag{9}$$

The output of the systems can therefore be considered as

$$y(t) = cx(t) + K_i f_i(t) \tag{10}$$

where $x(t) = [x_1, x_3]^T$, $f_i(t)$ presents the intermittent fault while K_i is the known intermittent distribution matrix.

3. Intermittent Fault Detection

3.1. Residual Generation

Not all states $x(t)$ can be directly measured (as is commonly the case). We can therefore design an observer to estimate them, while measuring only the output $y(t) = Cx(t)$. The observer is basically a model of the plant; it has the same input and follows a similar differential equation. An extra term compares the actual measured output $y(t)$ to the estimated output of the observer $\hat{y}(t)$ minimising this error will cause the estimated states $\hat{x}(t)$ to tend towards the values of the actual real-system states $x(t)$. It is conventional to write the combined equations for the system plus observer using the original state $x(t)$ plus the error state [10]

$$e(t) = x(t) - \hat{x}(t) \tag{11}$$

The fault detection system consists of two parts, the first is the generation of the fault detection residual and the second is the evaluation of the residual against a specified threshold. While a suitable observer is chosen for every case, and the error system stability is satisfied, then the following scalar observer-based residual is generated for each output to detect the intermittent faults

$$r(t) = \xi(y(t) - \hat{y}(t)) = \xi Ce(t) + \xi K_i f_i(t) \tag{12}$$

where $\xi \in \mathbb{R}^{n_\xi \times p}$, is a suitable weighting matrix to be designed. To design ξ, the following assumption is made:

Assumption 2:

- Residual must be insensitive to disturbance or noise.
- Residual must be insensitive to parameter errors or nonlinearities.
- Residual must be sensitive to faults

The object is to show that the residuals are differing from zero when faults have occurred; however, the residual tends to zero in no fault situation [11, 12].

3.2. Residual Evaluation

A common choice of evaluation signal is the 2-norm:

$$\|r(t)\|_z \triangleq \sqrt{\int_0^{\infty} |r(\tau)|^2 d\tau} \tag{13}$$

The benefits of using the 2-norm for the residual evaluation are that it is straightforward to optimize the residual generator to minimize the influence of disturbance [13]. Since the evaluation function (13) cannot be realized exactly, because the value of $\|r(t)\|_2$ is not known until $t = \infty$, and it is reasonable to assume that the faults could be detected, if occurred over the finite time interval, therefore equation (13) could be modified to

$$\|r(t)\|_z \triangleq \sqrt{\int_0^{t} |r(\tau)|^2 d\tau} \tag{14}$$

where τ is the time window and it is finite . It does not make sense to detect faults over the whole time range. It is reasonable to assume that the fault $f_i(t)$ could be detected, if accrued, over the finite time interval.

3.3. Threshold Generation

The threshold is obtained based on the residual dynamics in a fault-free case. For the evaluation signal (14), the occupancy of faults can be alarmed if

$\|r(t) > T_r\| \Rightarrow$ a fault is detected , (Figure 2)

and

(15)

$\|r(t) \leq T_r\| \Rightarrow$ no fault is detected, (Figure 3).

Where T_r represents the threshold.

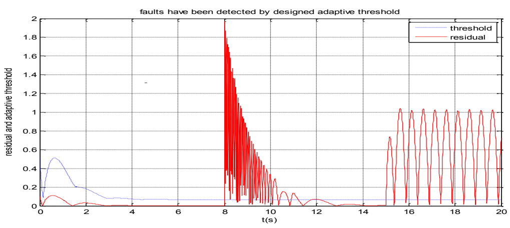

Fig 2: Adaptive threshold design when faults have been detected.

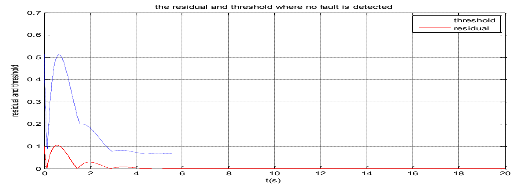

Fig 3: Adaptive threshold design when there is no fault to detect.

There are different methods to define threshold such as adaptive threshold design. The adaptive threshold is related to the main factors including system input, output, disturbance and parameters drifting over time, which are taken into account for non-linear system adaptive threshold modelling [14]. To design an adaptive threshold T_r define

$$r_{fi}(t) = r(t)\,|_{\mu(t)=0}$$

$$r_{\mu}(t) = r(t)\,|_{f_i(t)=0}$$

$$r_e(t) = r(t)\,|_{f_i(t)=0,\mu(t)=0}$$

Then the residual (12) can be rewritten as

$$r(t) = r_e(t) + r_{f_i}(t) \tag{16}$$

and the adaptive threshold design will be as follows

$$T_r = \max\|r_{\mu}(t)\| + \|r_e(t)\| \geq 0 \tag{17}$$

Since the disturbance (unknown input) is bounded to a positive scalar, $\mu(t) \leq d$ then

$$\max\|r_{\mu}(t)\| = \delta_d \geq 0 \tag{18}$$

Substituting (18) into (17),

$$\|r_e(t)\| = T_r - \delta_d$$

Therefore

$$\|r_e(t)\| \leq T_r \tag{19}$$

To show that the T_r is the upper bound of residual $\|r(t)\|$ consider the residual (16) as follows

$$\|r(t) = \|r_e(t)\| + r_{f_i}(t)\| \tag{20}$$

It is clear that in the faulty case

$$\|r_{f_i}(t) > \beta > 0\| \tag{21}$$

Where β a positive scalar. Substituting (19) and (21) into (20) we obtain

$$\|r(t)\| = \beta + T_r > 0$$
$$\|r(t)\| > T_r \quad (22)$$

If there is no fault in the system then the residual (20) becomes

$$\|r(t)\| = \|r_e(t)\| \quad (23)$$

Substituting (19) into (23)

$$\|r(t)\| = \|r_e(t)\| \le T_r \quad (24)$$

This means that no fault will be detected by the designed threshold when there is no fault present in the system. Hence the value of T_r gives an explicit bound on $\|r(t)\|$ in the fault free case and thus provides a valuable guideline for a robust threshold selection.

Table 1. The comparison of the fixed threshold method and the adaptive threshold method by fault diagnosis rate

The selected method of threshold	Fault detection rate	False alarm rate
Fixed threshold	90%	10%
Adaptive threshold	97%	3%

Adaptive thresholds not only reflect the changes of the system signal but can be insensitive to disturbances in its normal state. They are also useful for fast and efficient fault diagnosis. Compared with the fixed threshold method, this method can effectively improve the accuracy of fault detection and reduce false alarms as shown in Table 1.

4. Simulation Results

The two-link manipulator adopted in this study has a maximum of n inputs and n outputs; (where n is the number of joints). It is assumed that the friction parameters are already measured and the arms move in a clockwise direction. The equilibrium position is located at $\theta 1 = -1.57$ Radian and $\theta 2 = 2.96$ radians and the aim is to locate them at $\theta 1 = \theta 2 = 0$ radians.
Intermittent faults are generated as a combination of impulses at different amplitudes are designed to occur at discrete intervals and modelled as:

$$f_i(t) = \begin{cases} 0 \times d_y & \text{for} \quad t < 5s \\ 0.5 \times d_y & \text{for} \quad 5s \le t < 8s \\ 0 \times d_y & \text{for} \quad 8s \le t < 10s \\ 0.8 \times d_y & \text{for} \quad 10s \le t < 15s \\ 0 \times d_y & \text{for} \quad 15s \le t < 16s \\ 1.5 \times d_y & \text{for} \quad 16s \le t < 25s \end{cases} \quad (25)$$

where $dy > 0$ is a constant.

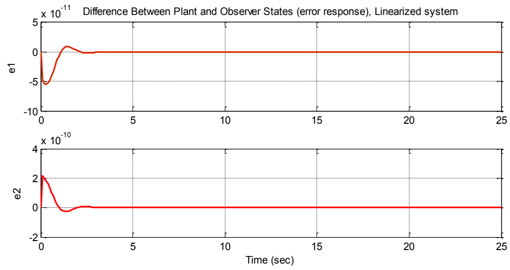

Figure 4: Responses of state errors e(t) when there is no fault in the system.

The numerical values of the different parameters are listed in Table 2.

Table 2. Numerical values of the system parameters.

Parameters	Numerical values
Length of arm 1, l_1	1 m
Length of arm 2, l_2	0.8 m
Mass 1, m_1	0.5 Kg
Mass 2, m_2	0.62 Kg
Gravity, g	9.8 $\frac{N}{m^2}$
Viscous coefficient, β_1	0.049
Viscous coefficient, β_2	0.077
Kinetic coefficient, α_1	0.08
Kinetic coefficient, α_2	0.12

Figure 4 shows the behavior of the state errors and demonstrates that the errors between the estimated and actual states are stable and converge to zero asymptotically even though uncertainties within the system exist. Based on Figure 5, it can be seen that the observer-based residual performs as expected and the intermittent faults have been detected at a very first stage where there is no disturbance, unknown input or parameter errors exist.

Figures 6 and 7 show the residual response when the system is influenced by some unknown input with known bound. In this case faults can be detected more precisely if an adaptive threshold is designed.

Figure 6, shows that with a fixed threshold false alarm can occur due to the dynamics of the system. This can be seen as a breach in the fixed threshold at the beginning of the systems operation where no fault exists. When an adaptive threshold is designed, as shown in Figure 7 the same system dynamics do not breach the threshold. With the adaptive threshold case it is also easy to design the threshold to be insensitive to faults of specific amplitude. Again considering the fixed threshold case the first intermittent fault occurring between 5s and 8s will result in alarm, unlike in the adaptive threshold case. The adaptive threshold approach therefore provides the capability to ignore small intermittent disturbances that manifest as system noise and do not have a serious impact on the system operation.

The simulation results also demonstrate that the proposed design approach minimizing the effects of any uncertainties and will give a straight-forward way to design a robust observer for intermittency fault detection where the bounded uncertainties are existed.

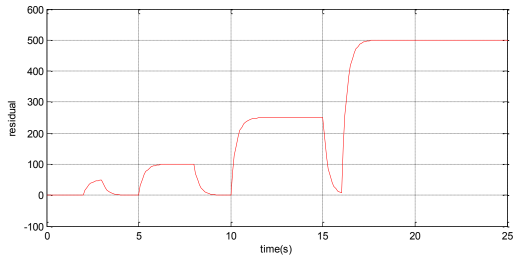

Figure 5: Residual response (no disturbance exists)

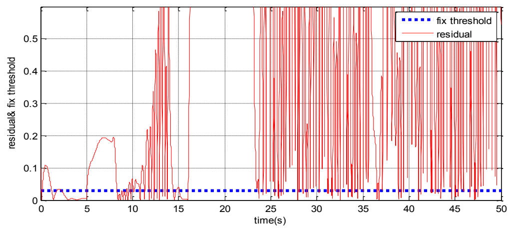

Figure 6: Residual and fix threshold responses.

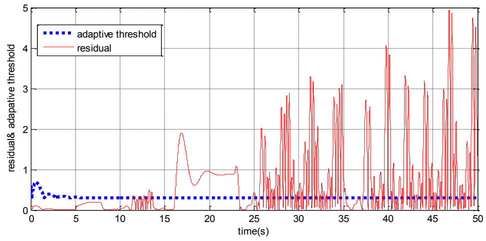

Figure 7: Residual and adaptive threshold responses.

5. Conclusions

A robust nonlinear observer has been designed for a class of nonlinear systems with bounded unknown inputs (uncertainties). We also show that the existing error dynamics between the estimated and actual states are stable. In this method, the non-unique design matrix W has been used to provide extra degrees of freedom to the user to design the residual. The main advantage of the proposed method is the possibility to diagnose the intermittent faults by generating a residual and an adaptive threshold which is highly sensitive to faults and insensitive to any bounded uncertainties. An adaptive threshold as employed in this paper makes the difficult intermittency fault detection an easier task for the considered class of nonlinear systems.

Finally, the effectiveness of the technique is illustrated by the help of a numerical example. The simulation results show that the designed residual and adaptive threshold can indeed detect the intermittent faults regardless of the bounded unknown inputs (uncertainties).

The research presented in this paper is ongoing as part of the EPSRC Centre in Through-life Engineering Services No Fault Found research. Further work is focused on understanding the time-varying dynamics of intermittent faults in order to predict and estimate intermittent faults in existence of unknown inputs (uncertainties) for a general class of nonlinear systems, using available known parameters.

References

[1] Hockley C., Phillips P. The impact of No-Fault-Found (NFF) on through-life engineering services. Int. J. Of Quality in Maintenance Engineering, 2012, 18(2), p.141-153.
[2] Steadman B., Berghout F., Olsen N. Intermittent fault detection and isolation system, IEEE Autotest Conf., 2008.
[3] Qi H., Ganesan S., Pecht M.,No fault faound and intermittent failure in electronic products,2008,(48),p.663-647.
[4] Frank P.M., On-line fault detection in uncertain nonlinear systems using diagnostic observers:a survey. Int J. Systems Sci, 1994, 25(12), p.2129-2154.
[5] Frank P.M., Analytical and qualitative model-based fault diagnosis: a survey and some new results. Eur. J. Control, 1996, (2),p.6-28.
[6] Abid M., Fault detection in nonlinear systems: An observer-based approach, PhD thesis, 2010.
[7] Sedighi T., Shields D. N., The performance of the high gain observer-based residual design for detecting faults in a mass-spring-damper system, UKACC Conf. Glasgow, 2006.
[8] Rajmani R., Observer for Lipschitz nonlinear systems, IEEE Trans. on Automatic control, 1998, (43), p.397-401.
[9] Munro N., Lewis F. L., Robot manipulator control, Theory and practice, Second edition, 2004.
[10] Sedighi T., Koshkouei A.J, Burnham K.J., Nonlinear observer-based residual design for the nonlinear systems. Int. Conf. Of UKACC2008, Manchester, 2008.
[11] Sedighi T., Koshkouei A.J, Burnham K.J., Optimal backstepping fault detection for mass-spring-damper systems. IASTED, USA, 2007.
[12] Sedighi T., Koshkouei A.J, Burnham K.J., Fault detection for mass-spring-damper systems: backstepping approach, Int. Conf. Of ICSE06, Coventry, 2006.
[13] Ding S. X., Model-based fault diagnosis techniques: design schemes, algorithms and tools. 2nd Ed., Springer-Verlag, London, 2013.
[14] Liu H., Lu C., Hou W., Wang S., An adaptive threshold based on support vector machine for fault diagnosis. 8th Int. Conf. On Reliability, Maintainability and Safety, ICRMS2009, P.907-911.