

Article

Minimum Error Entropy Filter for Fault Detection of Networked Control Systems

Jianhua Zhang *, Lilong Du, Mifeng Ren and Guolian Hou

State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China; E-Mails: dulilong1231@163.com (L.D.); renmifeng@126.com (M.R.); hougulian@sina.com (G.H.)

* Author to whom correspondence should be addressed; E-Mail: zjh@ncepu.edu.cn.

Received: 22 February 2012; in revised form: 27 February 2012 / Accepted: 28 February 2012 / Published: 6 March 2012

Abstract: In this paper, fault detection of networked control systems with random delays, packet dropout and noises is studied. The filter is designed using a minimum error entropy criterion. The residual generated by the filter is then evaluated to detect faults in networked control systems. An illustrative networked control system is used to verify the effectiveness of the proposed approach.

Keywords: networked control systems; fault detection; minimum error entropy

1. Introduction

Major advancements in the area of communication and computer networks have made it possible to insert them into feedback systems to meet real time requirements. Such networked control systems (NCSs) have been used in various areas, such as the automotive industry, tele-autonomy, tele-operation of robots and automated manufacturing systems [1,2]. However, the network leads to several issues: time delays, data dropout and limited bandwidth due to data transmission, and information loss due to encoding and quantization. Consequently, compared with conventional point-to-point control systems, these issues make the analysis and design of NCSs much more complex.

In recent years, some attention has been paid to fault detection of NCSs to improve their safety and reliability [3–6]. In [3], the network-induced delay was assumed to be a Markov chain, and fault detection of NCSs was investigated by combining a parity-equation approach with a fuzzy-observer-based approach. In [4], the induced delay was assumed to be an integer multiple of the “sampling

piece”, a H_∞ fault detection filter was designed for the NCSs modeled by discrete Markovian jump systems. In [5], the influence caused by delay in a NCS was transformed into a structured modeling error, and an existing continuous robust fault detection method was extended to the discrete time domain using a reference model and filter, moreover, the proposed approach was applied to NCSs.

Entropy has been widely used in information, communications, identification, estimation and control theories as a more general measure of uncertainty [7–13]. The Δ -entropy criterion was applied in system identification to search for the optimum parameter set in [11]. The proposed algorithm can be used to deal with coarsely quantized input-output data in digital control systems, networked control systems (NCS), wireless sensor networks, *etc.* An optimal fuzzy control system was investigated using the cross-entropy method and applied to a drilling process over communication networks [12]. Entropy was used to study the achievable sensitivity bounds for MIMO control systems, the amount of information that the communication signals contain was evaluated regarding the plant and the controller [13].

Recently, fault detection filters were proposed for non-linear stochastic systems with non-Gaussian noises using the entropy optimization principle [14–16]. However, it is difficult to extend these approaches to NCSs because non-Gaussian delays induced by networks cannot be modeled as noises.

Since the characteristics of noises, delays and dropouts in NCSs are stochastic in nature, in this paper the fault detection method of NCSs is investigated using the entropy optimization principle. The main contribution resides in designing minimum error entropy filters for NCSs via an information theoretical learning technique. The adaptive filter is then utilized to generate the residual for detecting faults in NCSs.

The rest of this paper is organized as follows: Section 2 describes the schematic diagram of fault detection for an NCS and formulates the problem. Section 3 designs a filter to generate residuals using the minimum error entropy (MEE) criterion. Section 4 introduces the rules for evaluating residuals in order to detect faults in the NCS. Section 5 verifies the efficiency of the proposed approach using an illustrative example, and the last section concludes the paper.

2. System Description and Problem Formulation

Filter-based fault detection approaches have been widely used in [17–19], and usually consist of two stages: residual generation using a filter and residual evaluation. The filter-based fault detection scheme for an NCS is shown in Figure 1, where the one-channel NCS with feedback has been commonly used to investigate the effects of delay and packet dropout in NCSs.

Let us consider the NCS represented by:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + B_{f_a} f_a(k) + B_\omega \omega(k) \\ z(k) &= Cx(k) + D_{f_s} f_s(k) + D_v v(k) \end{aligned} \quad (1)$$

and:

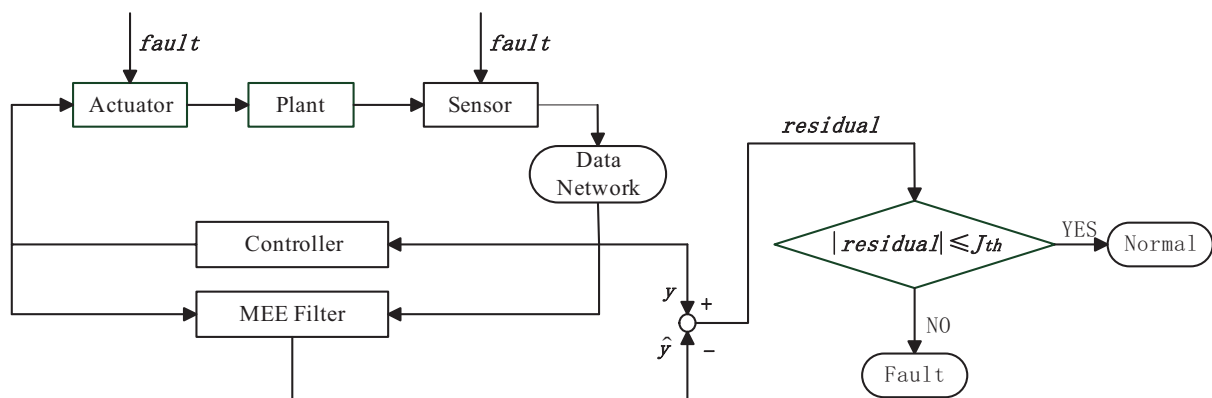
$$y(k) = \begin{cases} z(k - \tau_{sc}(k)), & \zeta(k) = 0 \\ y(k-1), & \zeta(k) = 1 \end{cases} \quad (2)$$

where $x(k) \in R^n$ denotes the state vector, $u(k) \in R^m$ the control input, $z(k) \in R^l$ the measurable output vector and $y(k) \in R^l$ the output vector through network. $\omega(k) \in R^p$ and $v(k) \in R^q$ are system noise and measurement noise, respectively. The unknown actuator fault and sensor fault are denoted by $f_a(k) \in R^l$ and $f_s(k) \in R^r$, respectively. $\tau_{sc}(k)$ is the random delay from sensor to controller. A, B, B_ω, C and D_v are known matrices of appropriate dimensions.

The following assumptions are introduced:

- The actuators and sensors are time-driven, while the controller is event-driven;
- The data packets are transmitted in right order;
- The measurements and control signals are transmitted using single-packet with time-stamp;
- A is nonsingular and (A, C) is observable.

Figure 1. Fault detection scheme for an NCS.



When a continuous-time signal is transmitted over a network, the overall delay between sampling and eventual decoding at the receiver can be highly variable because both the network access delays and the transmission delays depend on highly variable network conditions such as congestion and channel quality [1]. In most cases, the network induced delays are random and not necessarily Gaussian.

Another issue in NCSs is packet dropout caused by both the transmission errors in physical network links and buffer overflows due to congestion. In addition, long transmission delays sometimes also amounts to a packet dropout if the receiver discards “outdated” arrivals [1]. In this paper, packets delayed by more than a given maximum delay $M\tau_{sc}$ will be regarded as dropouts. The transmission of data packets can be described by the following switching logic:

$$\zeta(k) = \begin{cases} 0, & \tau_{sc}(k) \leq M\tau_{sc} \text{ (packet arrived)} \\ 1, & \tau_{sc}(k) > M\tau_{sc} \text{ (packet lost)} \end{cases} \quad (3)$$

Assuming that the probability density function of the random time delay $\tau_{sc}(k)$ is $\gamma_{\tau_{sc}}$, it yields:

$$\begin{aligned} P(\zeta(k) = 0) &= \int_0^{M\tau_{sc}} \gamma_{\tau_{sc}}(\sigma) d\sigma = 1 - \xi \\ P(\zeta(k) = 1) &= \int_{M\tau_{sc}}^{\infty} \gamma_{\tau_{sc}}(\sigma) d\sigma = \xi \end{aligned} \quad (4)$$

where ξ ($0 \leq \xi \leq 1$) is the data packet dropout rate.

Because the random delays induced by the communication network and data dropouts in NCSs are not necessarily Gaussian, and even the system noise and measurement noise are also not Gaussian in

some practical NCSs, it is necessary to design a filter using higher order statistics of the filtering error. Ideally the filter design for this kind of stochastic system should be performed such that the shape of the probability density function (PDF) of the filtering error is made as narrow as possible. A narrow distribution can be represented by a small entropy, which generally indicates that the uncertainty of the filtering error is small and all the moments of PDF of the filtering error are constrained. Therefore, in this paper, the filter to generate residuals is designed using minimum error entropy (MEE) criterion.

In Figure 1, the residual $r(k)$ is generated by the MEE filter based on the measured input signals $u(k)$ and output signals $y(k)$, and then the residual is compared with the pre-assigned threshold to decide whether the fault occurs or not.

3. Design of MEE Filter

Considering stochastic delays, data dropouts and noises in NCSs, the residual is generated utilizing the following filter:

$$\begin{aligned}\hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + L(y(k) - \hat{y}(k)) \\ \hat{z}(k) &= C\hat{x}(k)\end{aligned}\quad (5)$$

where:

$$\hat{y}(k) = \begin{cases} \hat{z}(k - \tau_{sc}(k)), & \zeta(k) = 0 \\ \hat{y}(k-1), & \zeta(k) = 1 \end{cases}$$

$\hat{x}(k) \in R^n$ and $\hat{y}(k) \in R'$ are the state and output estimation, respectively.

3.1. Performance Index

The filter gain matrix $L \in R^{n \times r}$ is updated by minimizing the entropy of the filtering error between the measured and estimated output $e(k) = y(k) - \hat{y}(k)$. In this paper, the quadratic Renyi entropy of filtering is utilized. Denote the PDF of the filtering error as $\gamma_e(\xi)$, since minimizing quadratic Renyi entropy $H_2(e) = -\log \int_{-\infty}^{\infty} \gamma_e^2(\xi) d\xi$ is equivalent to maximizing the quadratic information potential $V_2(e) = \int_{-\infty}^{\infty} \gamma_e^2(\xi) d\xi = E[\gamma_e(\xi)]$ (defined in [20,21]), MEE criterion can be replaced by minimizing the reciprocal of the quadratic information potential.

The quadratic Renyi entropy of filtering error $H_2(e)$ is employed as the performance index for designing the adaptive gain of the filter. It can be estimated directly from data samples $\{e_1(k), e_2(k), \dots, e_{\bar{N}}(k)\}$ by utilizing Parzen windowing technique with Gaussian kernel function

$\kappa_{\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x^2}{2\sigma^2}}$, it leads to:

$$\hat{H}_2(e) = -\log \left[\frac{1}{\bar{N}^2} \sum_{i=1}^{\bar{N}} \sum_{j=1}^{\bar{N}} \kappa_{\sigma}(e_i(k) - e_j(k)) \right] \quad (6)$$

and the quadratic information potential can be estimated as follows:

$$\hat{V}_2(e) = \frac{1}{\bar{N}^2} \sum_{i=1}^{\bar{N}} \sum_{j=1}^{\bar{N}} \kappa_{\sigma}(e_i(k) - e_j(k)) \quad (7)$$

Although the quadratic information potential of the filtering error can be estimated using a series of error data, it is preferred to find a practical solution to handle on-line, recursive computation for the estimate $\hat{V}_2(e)$.

The estimate $\hat{V}_2(e)$ starts by estimating the PDF of the filtering error recursively. Initialize the PDF estimate as $\hat{\gamma}_{e_0}(\xi) = \kappa_\sigma(\xi - e_0)$, the PDF estimate can be updated by using the new sample $e(k+1)$ [10]:

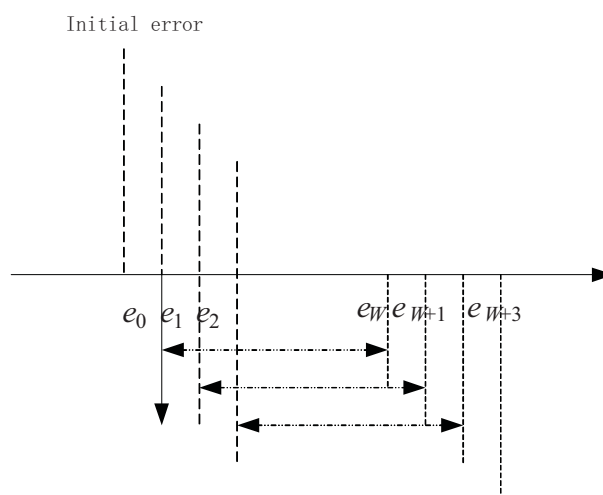
$$\hat{\gamma}_{e_{k+1}}(\xi) = (1-\lambda)\hat{\gamma}_{e_k}(\xi) + \lambda\kappa_\sigma(\xi - e_{k+1})$$

The quadratic information potential can be obtained using samples within the sliding window shown in Figure 2:

$$\begin{aligned}\hat{V}_2(e_{k+1}) &= E[\hat{\gamma}_{e_{k+1}}(\xi)] = (1-\lambda)E[\hat{\gamma}_{e_k}(\xi)] + \lambda E[\kappa_\sigma(\xi - e_{k+1})] \\ &\cong (1-\lambda)\hat{V}_2(e_k) + \frac{\lambda}{W} \sum_{i=k-W+1}^k \kappa_\sigma(e(i) - e(k))\end{aligned}\quad (8)$$

where $0 \leq \lambda < 1$ is the forgetting factor and W the width of sliding window.

Figure 2. Estimation of $J(k)$ using sequences with sliding window.



Since minimizing quadratic Renyi entropy is equivalent to minimizing the reciprocal of the quadratic information potential, in this work, the following performance index is utilized to update the filter gain L :

$$J(k) = \frac{1}{\hat{V}_2(e_k)} \quad (9)$$

3.2. Filter Gain Design

Denote $L = [L_1 \quad L_2 \quad \dots \quad L_t] \in R^{n \times t}$, where $L_i \in R^{n \times 1}$, $i = 1, 2, \dots, t$. The updated rule for the filter gain L can adopt stochastic gradient algorithm [22]. The procedure of calculating the filter gain L is presented as follows.

Utilizing the chain rule and taking the derivatives of performance criteria $J(k)$ with respect to the filter gain L , it can be obtained that:

$$\frac{\partial J(k)}{\partial L_j(k)} = \frac{1}{W} \sum_{i=k-W}^{k-1} \kappa'_\sigma(e(i) - e(k)) \cdot \left(\frac{\partial \hat{y}(k)}{\partial L_j(k)} - \frac{\partial \hat{y}(i)}{\partial L_j(i)} \right) \quad (10)$$

where $\kappa'_\sigma(e(i) - e(k)) = \frac{-(e(i) - e(k))}{\sigma^2} \cdot \kappa_\sigma(e(i) - e(k))$:

$$\frac{\partial \hat{y}(k)}{\partial L_j(k)} = \begin{cases} \frac{\partial \hat{z}(k - \tau_{sc}(k))}{\partial L_j(k)}, & \zeta(k) = 0 \\ \frac{\partial \hat{y}(k-1)}{\partial L_j(k)}, & \zeta(k) = 1 \end{cases} \quad (11)$$

$$\begin{aligned} \frac{\partial \hat{z}(k - \tau_{sc}(k))}{\partial L_j(k)} &= \frac{\partial \hat{z}(k - \tau_{sc}(k))}{\partial \hat{z}(k)} \cdot \frac{\partial \hat{z}(k)}{\partial L_j(k)} \\ &= \frac{\partial \hat{z}(k - \tau_{sc}(k))}{\partial \hat{z}(k - \tau_{sc}(k) + 1)} \cdot \frac{\partial \hat{z}(k - \tau_{sc}(k) + 1)}{\partial \hat{z}(k - \tau_{sc}(k) + 2)} \cdots \frac{\partial \hat{z}(k-1)}{\partial \hat{z}(k)} \cdot \frac{\partial \hat{z}(k)}{\partial L_j(k)} \\ &\approx \frac{\partial \hat{z}(k)}{\partial L_j(k)} = C \cdot \frac{\partial \hat{x}(k)}{\partial L_j(k)} \end{aligned} \quad (12)$$

It can be shown from Equation (5) that:

$$\frac{\partial \hat{x}(k)}{\partial L_j(k)} \approx (A - L \cdot C) \frac{\partial \hat{x}(k-1)}{\partial L_j(k-1)} + I_{n \times n} \cdot (y(k-1) - \hat{y}(k-1)) \quad (13)$$

Using the stochastic gradient algorithm, the updated rule of filter gain L can be then formulated to read:

$$L(k+1) = L(k) + \eta \cdot \frac{\partial J(k)}{\partial L(k)} \quad (14)$$

where η is the step size for adaptation.

Theorem 1: The filter gain given by (14) will be convergent if η satisfies the following inequality:

$$0 < \eta \leq \frac{1}{\left(\frac{\partial J(k)}{\partial L(k)} J(k) \right)^2} \quad (15)$$

Proof: Since the information potential $J(k)$ is positive, a Lyapunov function candidate can be selected as:

$$\pi = \frac{2}{J(k)} + (\Delta J(k))^2 \quad (16)$$

then by performing the first order derivatives of this Lyapunov function, it can be formulated that:

$$\frac{\partial \pi}{\partial k} = -2J(k)^{-2} \frac{\partial J(k)}{\partial L(k)} \frac{\partial L(k)}{\partial k} + 2\Delta J(k) \frac{\partial J(k)}{\partial L(k)} \frac{\partial L(k)}{\partial k} \quad (17)$$

Approximately, it can be seen that:

$$\frac{\Delta \pi}{\Delta k} = 2(-J(k)^{-2} + \Delta J(k)) \frac{\partial J(k)}{\partial L(k)} \frac{\Delta L(k)}{\Delta k} \quad (18)$$

Since the filter gain can be obtained by (14), from (17), (18) and $\Delta J(k) = \frac{\partial J(k)}{\partial L(k)} \Delta L(k)$, it can be obtained that:

$$\Delta \pi = 2\eta \left(\frac{\partial J(k)}{\partial L(k)} \right)^2 \left(-J(k)^{-2} + \eta \left(\frac{\partial J(k)}{\partial L(k)} \right)^2 \right) \quad (19)$$

and:

$$E(\Delta \pi) = E \left\{ 2\eta \left(\frac{\partial J(k)}{\partial L(k)} \right)^2 \left(-J(k)^{-2} + \eta \left(\frac{\partial J(k)}{\partial L(k)} \right)^2 \right) \right\} \quad (20)$$

Hence, the condition (15) can be obtained for convergence.

Remark 1: The functional approximation capabilities of neural networks have been exploited in the design of neural observers or filters [23,24]. Likewise, neural networks can be utilized in approximating the filter gain. The weights of the neural networks which can be trained using stochastic gradient algorithm construct the filter gain L , the inputs can be constructed by the error $e(k) = y(k) - \hat{y}(k)$ and delayed signals $e(k-1), \dots, e(k-W)$ of the neural networks. The output is a linear combination of the inputs.

Remark 2: In the extreme case, one might select a single sample in Equation (8), corresponding to $W = 1$. The size of window W usually should be selected to contain the dynamic characteristics of the plant, delays and packet dropouts. Increasing the window width results in smaller estimation variance and the speed of convergence is not affected by variations in this parameter [21]. The accuracy of the estimation increases while more storage space of convergence is required for holding the previous samples in memory. Thus, there is a tradeoff between the accuracy of estimation and the memory requirements.

Remark 3: The error of the filter is taken into account in a probabilistic framework, moreover, the filter is designed by data-driven strategy, the probability distribution functions of delays, packet losses and noises in NCSs are not needed.

In this contribution, the generated residual can be constructed as:

$$r(k) = y(k) - \hat{y}(k) \quad (21)$$

where the error $r(k)$ between the real NCS output and the filter output is called the residual signal used to detect faults in the NCS.

4. Fault Detection

The generated residual will be evaluated to detect faults in NCSs. The residual evaluation scheme is selected as:

$$\begin{cases} |\bar{J}(k)| \leq \bar{J}_{th}, & \text{no fault occurs} \\ |\bar{J}(k)| > \bar{J}_{th}, & \text{fault has occurred} \end{cases} \quad (22)$$

where $|\bar{J}(k)| = \left\{ \sum_{i=k-\bar{L}}^k r^T(i)r(i) \right\}^{\frac{1}{2}}$ is the evaluation function to judge the status of the system, $\bar{L}$ the width of residual estimation window. The threshold $\bar{J}_{th}$ should be selected to not only distinguish the noise and the fault but also decrease the false alarm, generally the threshold should be the maximal value of the evaluated output $|\bar{J}(k)|$ in the nominal operating state (fault-free). In this work, the threshold $\bar{J}_{th}$ is pre-assigned according to prior knowledge.

Remark 3: In order to achieve small false alarm rate with an acceptable sensitivity to faults, robust residual valuation strategies can be investigated and applied, such as statistical data processing, data reconciliation, correlation, pattern recognition, fuzzy logic or adaptive thresholds.

5. An Illustrative Example

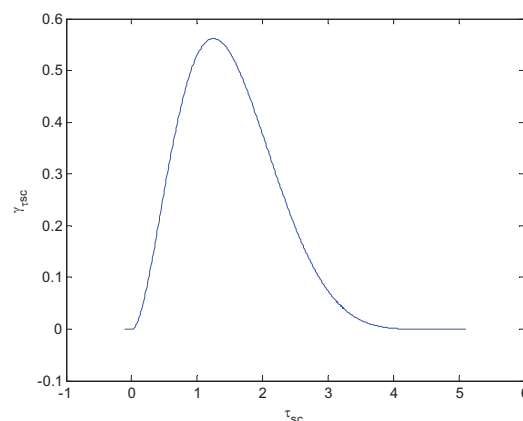
In order to show the applicability of the proposed fault detection method for NCSs, let us consider the following NCS, which consists of the plant represented by:

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 0 & 1 \\ -1 & -0.25 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0.01 \end{bmatrix} u(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} f_a(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \omega(k) \\ z(k) &= [1 \quad 0.01] x(k) + f_s(k) + v(k) \end{aligned} \quad (23)$$

The PDF of the induced delay $\tau_{sc}(k)$ is shown in Figure 3. Both uniformly distributed noises $\omega(k)$ and $v(k)$ lie in $[-0.005, 0.005]$. The kernel size and the forgetting factor used to estimate the information are experimentally set to $\sigma = 0.1$ and $\lambda = 0.02$ respectively. The width of sliding window is $W = 100$, the width of residual estimation window $\bar{L} = 20$, the learning rate $\eta = 0.0001$ and the sampling period $T = 1s$. The threshold is selected to $\bar{J}_{th} = 1.9$ according to prior knowledge. The fault is imposed as follows:

$$f_s(k) = \begin{cases} 0, & t < 150s \\ 1, & 150s < t < 200s \end{cases} \quad (24)$$

Figure 3. PDF of the induced delay τ_{sc} .



Assume that the control input is:

$$u(k) = [-0.05 \quad -0.25] \bar{x}(k) \quad (25)$$

where the state transmitted to controller is:

$$\bar{x}(k) = \begin{cases} x(k - \tau_{sc}(k)), & \tau_{sc}(k) \leq 4 \\ \bar{x}(k-1), & \tau_{sc}(k) > 4 \end{cases} \quad (26)$$

Moreover, the output arrived at controller is:

$$y(k) = \begin{cases} z(k - \tau_{sc}(k)), & \tau_{sc}(k) \leq 4 \\ y(k-1), & \tau_{sc}(k) > 4 \end{cases} \quad (27)$$

and the initial filter gain is set to $L = [-0.1 \quad -0.15]^T$.

The simulation results are shown in Figures 4–7. Figure 4 shows the increasing information potential of the MEE filter, in which the entropy of the filtering error decreases with time. Figures 5 and 6 indicate the evolution of two state errors respectively. The convergent state errors demonstrate better performance of the filter. Figure 7 shows the evolution of the evaluation function $|\bar{J}(r, k)| = \left\{ \sum_{i=k-L}^k r^T(i)r(i) \right\}^{\frac{1}{2}}$ for both the fault case and fault-free case, and it can be seen that fault detection can achieve good performance despite existing delays, packet dropout and noises in the NCS.

Figure 4. Information potential of filter error.

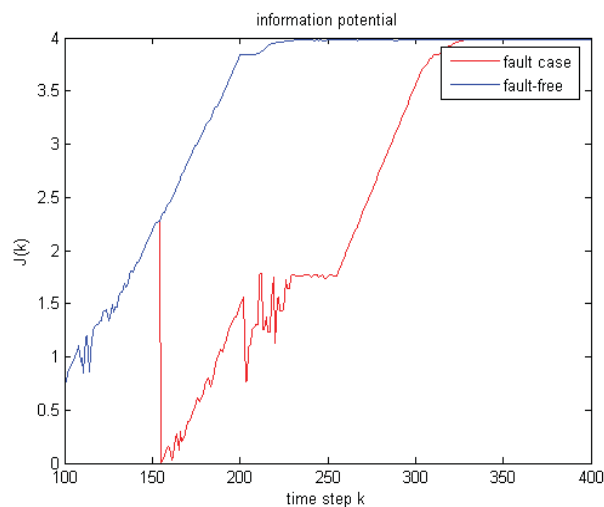


Figure 5. Estimation error of state 1.

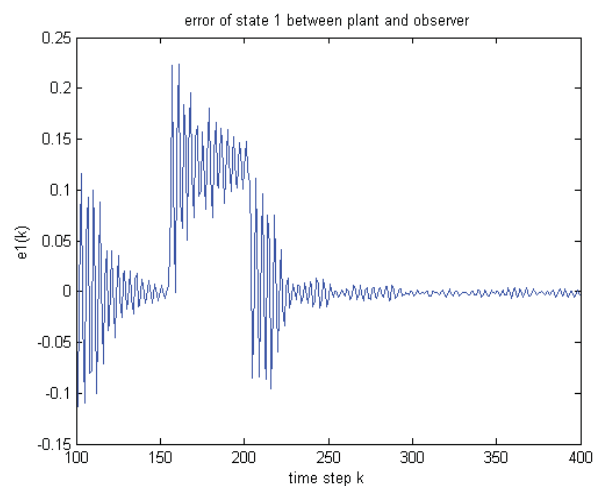
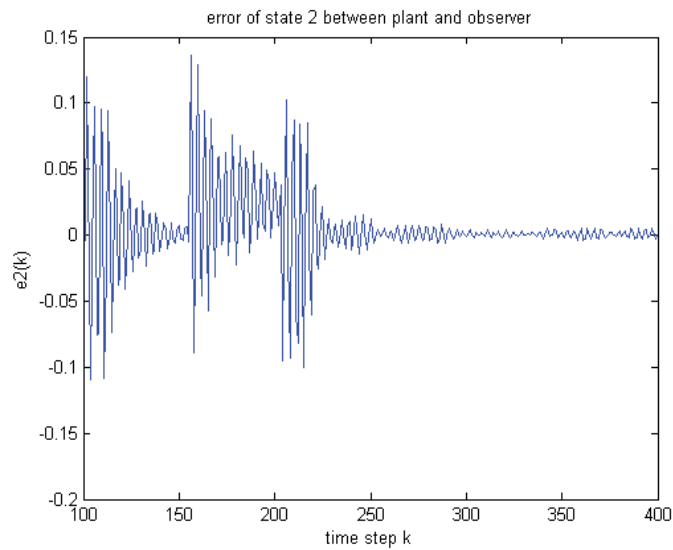
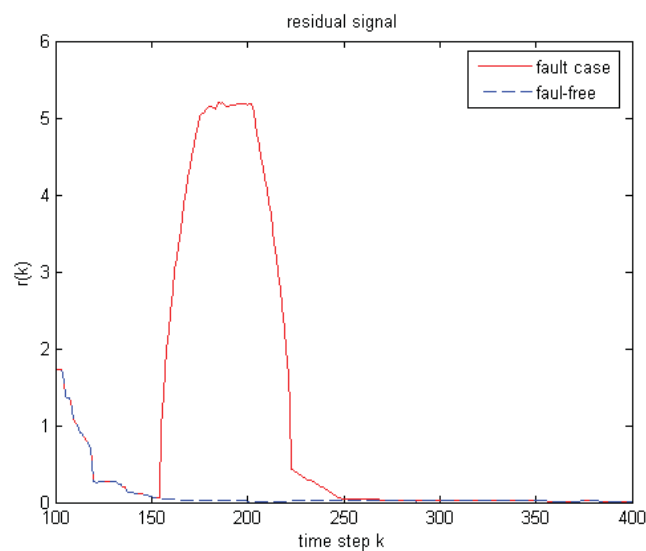


Figure 6. Estimation error of state 2.**Figure 7.** Evolution of $\bar{J}(r, k)$.

6. Conclusions

This paper presents a filter-based fault detection scheme for NCSs with network-induced delays, packet dropout and noises. In most cases, the random delays and noises may not obey Gaussian distribution, and packet dropout occurs stochastically. The filter under entropy framework is designed to generate residuals. The adaptive filter gain is updated using a stochastic gradient algorithm and can be implemented by neural networks. A time-variant threshold is employed to detect faults. Simulations demonstrate the effectiveness of the proposed fault detection scheme. However, the extension of this approach to generally non-linear NCSs calls for further investigation.

Acknowledgments

This work was supported by the National Basic Research Program of China under Grant (973 Program 2011CB710706) and the China National Science Foundation under Grants (60974029 and 61004031). These are gratefully acknowledged.

References

1. Hespanha, J.P.; Naghshtabrizi, P.; Xu, Y. A survey of recent results in networked control systems. *Proc. IEEE* **2007**, *95*, 138–162.
2. Baillieul, J.; Antsaklis, P.J. Control and communication challenges in networked real-time systems. *Proc. IEEE* **2007**, *95*, 9–28.
3. Zheng, Y.; Fang, H.; Wang, H.O. Takagi-Sugeno fuzzy-model-based fault detection for networked control systems with Markov delays. *IEEE Trans. Syst. Man Cybern. B Cybern.* **2006**, *36*, 924–929.
4. Mao, Z.; Jiang, B.; Shi, P. H_∞ fault detection filter design for networked control systems modelled by discrete Markovian jump systems. *IET Contr. Theor. Appl.* **2007**, *1*, 1336–1443.
5. Wang, Y.; Ye, H.; Ding, X.S.; Wang, G. Fault detection of networked control systems based on optimal robust fault detection filter. *Acta Autom. Sin.* **2008**, *34*, 1534–1539.
6. Fang, H.J.; Ye, H.; Zhong, M.Y. Fault diagnosis of networked control systems. *Annu. Rev. Contr.* **2007**, *31*, 55–68.
7. Wang, H. *Bounded Dynamic Stochastic Systems: Modeling and Control*; Springer-Verlag: London, UK, 2000.
8. Yue, H.; Zhou, J.L.; Wang, H. Minimum entropy control of closed-loop tracking errors for dynamic stochastic systems. *IEEE Trans. Automat. Contr.* **2003**, *48*, 118–122.
9. Yue, H.; Wang, H. Minimum entropy of B-spline PDF systems with mean constraint. *Automatica* **2006**, *46*, 989–994.
10. Erdogmus, D.; Principe, J.C.; Kim, S.P.; Sanchez, J.C. A recursive Renyi's entropy estimator. In Proceedings of the 12nd IEEE Workshop on Neural Networks for Signal Processing, Martigny, Switzerland, 4–6 September 2002; pp. 209–217.
11. Chen, B.; Zhu, Y.; Hu, J.; Principe, J.C. Δ -Entropy: Definition, properties and applications in system identification with quantized data. *Inf. Sci.* **2011**, *181*, 1384–1420.
12. Haber, R.E.; del Toro, R.M.; Gajate, A. Optimal fuzzy control system using the cross-entropy method. A case study of a drilling process. *Inf. Sci.* **2010**, *180*, 2777–2792.
13. Ishii, H.; Okano, K.; Hara, S. Achievable sensitivity bounds for MIMO control systems via an information theoretic approach. *Syst. Contr. Lett.* **2011**, *60*, 111–118.
14. Guo, L.; Wang, H.; Chai, T. Fault detection for non-linear non-Gaussian stochastic systems using entropy optimization principle. *Trans. Instit. Meas. Contr.* **2006**, *28*, 145–161.
15. Yin, L.; Guo, L. Fault isolation for multivariate nonlinear non-Gaussian systems using generalized entropy optimization principle. *Automatica* **2009**, *45*, 2612–2619.
16. Guo, L.; Wang, H. Minimum entropy filtering for multivariate stochastic systems with non-Gaussian noises. *IEEE Trans. Autom. Contr.* **2006**, *51*, 695–700.
17. Gertler, J.J. *Fault Detection and Diagnosis in Engineering Systems*; Marcel Dekker: New York, NY, USA, 1998.

18. Chen, J.; Patton, R.J. *Robust Model-Based Fault Diagnosis for Dynamic Systems*; Kluwer Academic Publishers: Boston, MA, USA, 1999.
19. Blanke, M.; Kinnaert, M.; Lunze, J.; Staroswiecki, M. *Diagnosis and Fault-Tolerant Control*; Springer-Verlag: Berlin, Heidelberg, Germany, 2003.
20. Erdogmus, D.; Pricipe, J.C. An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems. *IEEE Trans. Signal Process.* **2006**, *50*, 1780–1786.
21. Pricipe, J.C.; Xu, D.; Fisher, J.W. Information theoretic learning. In *Unsupervised Adaptive Filtering*; Wiley: New York, NY, USA, 2000, pp. 265–319.
22. Xu, J.W.; Erdogmus, D.; Principe, J.C. Minimum error entropy Luenberger observer. In Proceedings of the 2005 American Control Conference, Portland, OR, USA, 8–10 June 2005; pp. 1923–1928.
23. Resendiz, J.; Yu, W.; Fridman, L. Two-stage neural observer for mechanical systems. *IEEE Trans. Circuits Syst. II* **2008**, *55*, 1076–1080.
24. Huang, S.N.; Tan, K.K.; Lee, T.H. Further result on a dynamic recurrent neural-network-based adaptive observer for a class of nonlinear systems. *Automatica* **2005**, *41*, 2161–2162.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).