Active Fault Detection for Neural Network Based Control of Non-linear Stochastic Systems

Ivo Punčochář, Ladislav Král and Miroslav Šimandl

Department of Cybernetics
University of West Bohemia
Czech Republic

SAFEPROCESS'09
Spain, Barcelona
1 July, 2009

Contents

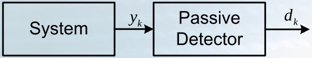
Introduction

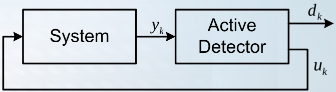
⇒ active fault detection

- area of active fault detection has received great attention during the recent years (*Zhang 1989, Nikoukhah 1993, Niemann 2006*)
- mainly designed for linear systems

⇒ active fault detection for non-linear stoch. systems

- task of active fault detection for stochastic non-linear systems has attracted only minor attention up to now
- a general solution of the problem is extremely difficult (*Šimandl 2006*)
- a special case - active detector design for a given set of controllers for jump Markov non-linear Gaussian models
- nonlinear system is modelled using neural networks





Neural Networks in Fault Detection

⇒ neural networks in fault detection:

- neural networks - challenging problem of fault detection of nonlinear systems
- neural networks as a powerful tool for modelling non-linear functions
- basic approaches:
  - ① neural network working as residue generator (neural network is a model of the system)
  - ② neural network working as residue evaluation tool (neural network is a classifier)
  - ③ combination of ① and ②
- ①, ② and ③ in passive fault detection only

Goal of the Paper

goal

- Main goal of the paper is to design an active detector for a given set of controllers considering jump Markov non-linear Gaussian models.
- In other words to extend the authors previous work on active fault detection (*M. Šimandl and I. Punčochář (2006). Closed loop information processing strategy for optimal fault detection and control. In Preprints of the 14th IFAC Symposium on System Identification, Newcastle, Australia*) to nonlinear systems.

Problem Statement

$$\mathcal{S}: \quad y_k = f_{\mu_k}(\mathbf{x}_{k-1}) + g_{\mu_k}(\mathbf{x}_{k-1})u_{k-1} + e_k, \tag{1}$$

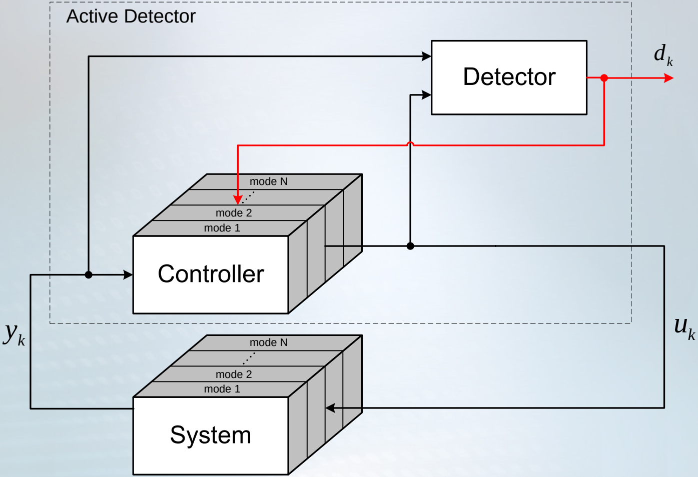
- unknown nonlinear functions $f_{\mu_k}(\cdot)$, $g_{\mu_k}(\cdot)$ can be switched at an arbitrary instant in time taking on any of the functions in the given set $\{(f_1, g_1), (f_2, g_2), \ldots, (f_N, g_N)\}$
- general form of the controller: $u_k = \gamma_k(\mathbf{I}_0^k, d_k)$
- $\mathbf{I}_0^k$ is information state at time k, where $\mathbf{I}_0^k = [y_k, \ldots, y_0, u_{k-1}, \ldots, u_0]$

Active detector design

- d_k is a decision which is provided by a detector,
- general description of the detector: $d_k = \sigma_k\left(\mathbf{I}_0^k\right)$,
- goal of the design is to find a whole sequence of functions $\boldsymbol{\sigma}_0^F$ that minimizes a suitable additive criterion penalizing wrong decisions

$$J(\boldsymbol{\sigma}_0^F) = \mathrm{E}\left\{\sum_{k=0}^{F} L_k^{\mathrm{d}}\left(\mu_k, d_k\right)\right\}, \tag{2}$$

Problem Statement - Graphically



Active Detector - Optimal Solution

optimal solution

$$V_k^*\left(\mathbf{I}_0^k\right) = \min_{\substack{d_k \in \mathcal{M} \\ u_k = \gamma_k(\mathbf{I}_0^k, d_k)}} \mathrm{E}\left\{L_k^{\mathrm{d}}\left(\mu_k, d_k\right) + V_{k+1}^*\left(\mathbf{I}_0^{k+1}\right) | \mathbf{I}_0^k, u_k, d_k\right\},$$

$$d_k^* = \arg \min_{\substack{d_k \in \mathcal{M} \\ u_k = \gamma_k(\mathbf{I}_0^k, d_k)}} \mathrm{E}\left\{L_k^{\mathrm{d}}\left(\mu_k, d_k\right) + V_{k+1}^*\left(\mathbf{I}_0^{k+1}\right) | \mathbf{I}_0^k, u_k, d_k\right\},$$

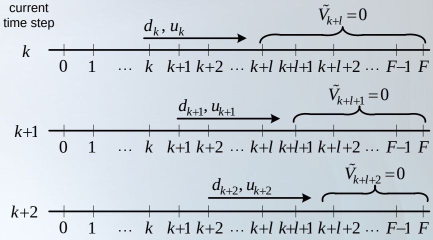
$$u_k^* = \gamma_k\left(\mathbf{I}_0^k, d_k^*\right).$$

- The optimal decision d_k^* is a compromise between minimization the current costs and excitation of the system through the given controller.

- analytical and numerical optimal solution is unmanageable
- approximations have to be taken to obtain a suboptimal feasible solution

Active Detector - Suboptimal Solution

- approximate $V^*_{k+l}\left(\mathbf{I}_0^{k+l}\right)$ by a heuristically computed function $\tilde{V}_{k+l}\left(\mathbf{I}_0^{k+l}\right)$
- rolling horizon approximation $\tilde{V}_{k+l}\left(\mathbf{I}_0^{k+l}\right) = 0$
- choice of a number of steps l - compromise between computational burdens and quality



State (model) estimation; required pdf's

$$P\left(\mu_k|\mathbf{I}_0^k\right) = \frac{p\left(y_k|\mathbf{I}_0^{k-1}, u_{k-1}, \mu_k\right) P\left(\mu_k|\mathbf{I}_0^{k-1}\right)}{p\left(y_k|\mathbf{I}_0^{k-1}, u_{k-1}\right)}$$

$$P\left(\mu_{k+1}|\mathbf{I}_0^k\right) = \sum_{\mu_k \in \mathcal{M}} P\left(\mu_{k+1}|\mu_k\right) P\left(\mu_k|\mathbf{I}_0^k\right),$$

$p\left(y_{k+1}|\mathbf{I}_0^k, u_k, \mu_{k+1}\right) = ?$ $\Rightarrow$ neural networks

Neural Network Model

- unknown non-linear functions $f_{\mu_k}(\cdot)$, $g_{\mu_k}(\cdot)$ are approximated by a couple of one-hidden-layer perceptron networks ➠ **model**
- identification of the model parameters is performed off-line
- dependency of $\hat{y}_k$ on networks parameters is nonlinear
  ➠ utilization of nonlinear estimation method - EKF
- structure optimization - based on pruning algorithm (*Šimandl, 2005*)

j-th local neural network

$$\mathcal{M}: \quad \hat{y}_{k,j} = \hat{f}_j(\mathbf{c}_j^f, \mathbf{x}_{k-1}^a, \mathbf{w}_j^f) + \hat{g}_j(\mathbf{c}_j^g, \mathbf{x}_{k-1}^a, \mathbf{w}_j^g) u_{k-1},$$

$$\hat{f}_j(\mathbf{c}_j^f, \mathbf{x}_{k-1}^a, \mathbf{w}_j^f) = (\mathbf{c}_j^f)^T \phi^f(\mathbf{x}_{k-1}^a, \mathbf{w}_j^f),$$

$$\hat{g}_j(\mathbf{c}_j^f, \mathbf{x}_{k-1}^a, \mathbf{w}_j^f) = (\mathbf{c}_j^g)^T \phi^g(\mathbf{x}_{k-1}^a, \mathbf{w}_j^g),$$

$$\mathbf{\Theta}_j = \left[(\mathbf{c}_j^f)^T, (\mathbf{w}_j^{f1})^T, \ldots, (\mathbf{c}_j^g)^T, (\mathbf{w}_j^{g1})^T, \ldots\right]^T ➠ p(\mathbf{\Theta}_j | \mathbf{I}_0^k)$$

$$➠ p\left(y_{k+1} | \mathbf{I}_0^k, u_k, \mu_{k+1}\right)$$

Controller Design

feedback linearization controller

- control action based on criterion

$$J(u_k) = \mathrm{E}\{[r_{k+1} - y_{k+1}]^2 | \mathbf{I}_0^k\}$$

- r_{k+1} chosen reference signal
- controller is obtained by minimization of the criterion using the certainty equivalence principle
- final control law

$$u_k = \gamma_k(\mathbf{I}_0^k, d_k) = \frac{r_{k+1} - \hat{f}_{d_k}(\mathbf{c}_{d_k}^f, \mathbf{x}_k, \mathbf{w}_{d_k}^f)}{\hat{g}_{d_k}(\mathbf{c}_{d_k}^f, \mathbf{x}_k, \mathbf{w}_{d_k}^f)}$$

Numerical Example

nonlinear system with two modes

$\boldsymbol{\mu_k}$**=1**: $$y_k=\frac{1.5y_{k-1}y_{k-2}}{1+y_{k-1}^2+y_{k-2}^2}+0.35\sin(y_{k-1}+y_{k-2})+1.2u_{k-1}+e_k,$$

$\boldsymbol{\mu_k}$**=2**: $$y_k=\frac{1.5y_{k-1}y_{k-2}}{1+y_{k-1}^2+y_{k-2}^2}+0.35\cos(y_{k-1}+y_{k-2})+0.3u_{k-1}+e_k,$$

- e_k is white noise with zero mean and variance $\sigma^2 = 0.09$.
- the initial probabilities of modes
  $$P(\mu_0 = 1) = 0.49 \text{ a } P(\mu_0 = 2) = 0.51,$$
- transition probabilities
  $$P_{1,1} = P_{2,2} = 0.9 \text{ a } P_{1,2} = P_{2,1} = 0.1.$$
- reference signal r_k is zero.

Numerical Example - Identification

⇒

- appropriate models of the system were obtained by off-line identification process
- parameter estimation + structure optimization of neural networks
- input signal u_k for identification was chosen to be a zero mean white noise with variance 2
- 10 000 time instants data for identification, 1 000 unused time instants data for model validation

	$\mu_k = 1$	$\mu_k = 2$
nf^*	13	10
ng^*	3	2
MSE_{train}	0.096	0.102
MSE_{test}	0.110	0.118

Numerical Example - Results

- cost function of the active detector is chosen as zero-one cost function

$$L_k^{\mathrm{d}}(\mu_k, d_k) = \begin{cases} 0 & \text{if } d_k = \mu_k, \\ 1 & \text{if } d_k \neq \mu_k. \end{cases}$$

- suboptimal active detector with given controller uses just one-step rolling horizon technique
- control for 50 steps, 3 000 Monte Carlo simulations
- comparison with passive detector based on open-loop feedback strategy

$$d_k^{\mathrm{passive}} = \min_{d_k \in \mathcal{M}} \mathrm{E}\left\{ L_k^{\mathrm{d}}(\mu_k, d_k) \,|\mathbf{I}_0^k, d_k \right\}$$

	$\hat{J}$	$\mathrm{var}\,\hat{J}$
PFD	6.5990	0.0040
AFD	6.1010	0.0031

Conclusion

- Novel active fault detector for non-linear stochastic system was proposed.
- The jump Markov non-linear Gaussian model was considered for system description.
- Neural networks were utilized for modelling the individual modes of the system.
- The numerical example showed that the sub-optimal active detector can improve the quality of detection in comparison to the passive detector, which was designed using open-loop feedback strategy.