Computer Controlled Systems II – Diagnosis

Case study on event sequence based diagnosis with clustering of event sequences

Katalin Hangos

University of Pannonia
Faculty of Information Technology
Department of Electrical Engineering and Information Systems

hangos.katalin@virt.uni-pannon.hu

Jan 2020

Lecture overview

1. Basic notions
   - Prediction based diagnosis
   - Traces, trace distances
2. Traces and trace distances revisited
   - ℓ-Neighbourhood of traces
   - Coordinate vector form
3. Clustering of traces
   - Acquiring characteristic traces using clustering
   - Model validation
4. The diagnostic procedure
5. A simple example

Basic notions

Prediction-based diagnosis

General problem statement

Given:

- *The number of faulty modes* N_F (0=normal)
- *Predictive dynamic model* for each faulty mode

$$y^{(Fi)}(k+1) = \mathcal{M}^{(Fi)}(\mathcal{D}[1,k];\ p^{(Fi)}) \quad , \quad k = 1, 2, \ldots$$

- *Measured data record*: $D[0,k] = \{\ (u(\tau), y(\tau) \mid \tau = 0, \cdots, k\}$
- *Loss function* $J^{(Fi)}$, $i = 0, \cdots, N_F$

$$J^{(Fi)}(y - y^{(Fi)}, u) \ = \ \sum_{\tau=1}^{k} [\ r^{(i)T}(\tau) Q r^{(i)}(\tau)\ ] \ , \ \ r^{(i)}(\tau) = y(\tau) - y^{(Fi)}(\tau) \ \ , \ \ \tau = 1$$

Compute: The actual faulty mode of the system, i.e. the fault index i that minimizes the loss function.

Fault isolation

Signal traces – event sequences

The **(signal trace)** of a qualitative signal $[x]$ is the event sequence

$$\mathcal{T}_{(x)}(t_0, t_F) = \{(t_0; [x](t_0) = q_{x0}), (t_1; [x](t_1)] = q_{x1}), ..., (t_F; [x](t_F) = q_{xF})\}$$

defined on the time interval (t_0, t_F) with $q_* \in \mathcal{Q}_x$

A **vector-valued trace of multiple signals** is defined as $\mathcal{T}_{(u,d,y)}(t_0, t_F)$
Simplified notation: by omitting the time, e.g.

$$\mathcal{T}_{(h,T)}(1,3) = \{(N,N),(L,H),(L,e+)\}$$

For diagnostic purposes we define

- **nominal traces** (for describing normal behaviour)
- **characteristic traces** (for describing faulty behaviour)

Norms of traces

Scalar valued trace: discrete time signal with qualitative values

$$\mathcal{T}_{(x)}(t_0, t_F) = \{(t_0; [x](t_0) = q_{x0}), (t_1; [x](t_1)] = q_{x1}), ..., (t_F; [x](t_F) = q_{xF})\}$$

defined on the time interval (t_0, t_F) with $q_* \in \mathcal{Q}_x$ Example:

$$\mathcal{T}_{(h)}(1, 3) = \{(N), (L), (L)\}$$

Norm: based on the norm of discrete time scalar valued real signals using a mapping function $\mathcal{R} : \mathcal{Q}_x \mapsto \mathbb{R}$:

$$\mathcal{R}(q) = \begin{cases} -1 & q = e- \\ 0 & q = 0 \\ 1 & q = L \\ 2 & q = N \\ 3 & q = H \\ 4 & q = e+ \end{cases}$$

Traces and trace distances revisited

Qualitative range spaces

Range spaces with *different granularity*

- Boolean

$$\mathcal{B} = \{0, 1\}$$

- Real valued

$$Q = \{e-, 0, L, N, H, e+\}$$

where "N" is the normal range, "L" and "H" denote the low and high but acceptable interval, "$e-$" and "$e+$" is the unacceptably low and high values

$$Q_{refined} = \{e-, -0, 0, 0L, L, LN, N, NH, H, H+, e+\}$$

with "-0" small negative values, "$0L$" very low, "LN" a bit low, "NH" a bit high, "$H+$" very high

ℓ-Neighbourhood of traces

Ordered range spaces: $Q = \{e-, 0, L, N, H, e+\}$ is *ordered* using the ordering of the underlying real interval values such that

$$e - \leq 0 \leq L \leq N \leq H \leq e+$$

Definition: ℓ-neighbourhood of a qualitative value $q \in Q = \{e-, 0, L, N, H, e+\}$ is a set $neighbourhood(q, \ell)$ (with $\ell \in \mathbb{N}_+$ and Q being an ordered range space) with elements from Q not farther that the given ℓ.
ℓ-neighbourhood of a trace $\mathcal{T}_{(x)}(t_0, t_F)$ defined on the time interval (t_0, t_F) with $q_* \in Q = \{e-, 0, L, N, H, e+\}$ is a trace

$$neighbourhood(\mathcal{T}_{(x)}, \ell)(t_0, t_F) = \{(t_0; neighbourhood(q_{x0}, \ell), (t_1; neighbourh$$

Example:

$$\mathcal{T}_{(h)}(1, 3) = \{(N), (L), (L)\}$$

$$neighbourhood(\mathcal{T}_{(h)}, 1)(1, 3) = \{(L, N, H), (0, L, N), (0, L, N)\}$$

Can be used for traces with qualitative measurement error

Different mappings for trace norm computations

Mapping of qualitative values to real ones by *mapping functions*

- linear

$$M_{linear}(q) = \begin{cases} -1 & q = e- \\ 0 & q = 0 \\ 1 & q = L \\ 2 & q = N \\ 3 & q = H \\ 4 & q = e+ \end{cases}$$

- non-linear

$$M_{non-linear}(q) = \begin{cases} -10.0 & q = e- \\ -2.0 & q = 0 \\ -1.0 & q = L \\ 0.0 & q = N \\ 10.0 & q = H \\ 20.0 & q = e+ \end{cases}$$

- refined linear

$$M_{finer}(q) = \begin{cases} -1.0 & q = "e-" \\ -0.5 & q = "0-" \\ 0.0 & q = "0" \\ 0.5 & q = "0L" \\ 1.0 & q = "L" \\ 1.5 & q = "LN" \\ 2.0 & q = "N" \\ 2.5 & q = "NH" \\ 3.0 & q = "H" \\ 3.5 & q = "H+" \\ 4.0 & q = "e+" \end{cases}$$

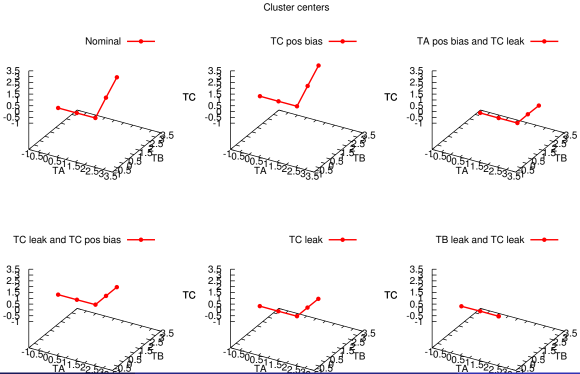
Coordinate vector form of events and traces

Event mapping function $G_{IO} : E \mapsto \mathbb{R}^r$, where $r = no_{imputs} + no_{outpouts}$ using a given mapping function M_X

$\mapsto$ *event coordinate form* , a point in the event state space

Trace coordinate form : a sequence of vectors that represent the events in the trace in their event coordinate form

$\mapsto$ *path of linear segments* in the event state space



Clustering of traces

Clustering – basic definitions

1. Given a distance metric D (for example the Euclidean distance).
2. Given a set X, let us denote the number of elements in X by $|X|$.
3. Given n *diagnostic scenarios* , let i be the scenario index going from 1 to n.
4. Given a *set of traces Y in trace coordinate form* , and *centroids* Z and W.
   Let the relation Y *belongs to* Z denote the set of traces from Y which are closer to Z than W using distance metric D.
   Similarly, Y *belongs to* W denotes the set of traces from Y which are closer to W than Z using distance metric D.
   Consequently, $|Y| \geq |Y \textit{ belongs to } C|$ for every centroid C.

Acquiring characteristic traces

Training phase of diagnosis using characteristic traces

Given: *For every fault scenario i a set of traces in trace coordinate form are provided* for creating and validating the centroids. This set is split into a training set T_i (for creating the centroids) and a validation set V_i (for performing validation of the centroids).

Compute: the centroid trace C_i for each scenario i.

- These centroids are created in single trace coordinate form, and they might not be equal to any specific input trace of the training set.
- A centroid, like a trace is a piece-wise linear trajectory in n dimensional space where n is the number of outputs, and the length of the line is the length of the trace (number of events in the trace).

Validating the characteristic traces

"Validating phase" *using independent measured traces* – the same as using the characteristic traces in the *diagnostic phase*

Fault detection rate: for a given centroid (characteristic trace) C_i

$$FDR_i = \frac{|V_i \ belongs \ to \ C_i|}{|V_i|}$$

where V_i is the validation trace set for fault scenario i.

The sequence $\{FDR_i | i = 1...n\}$ also gives an overall fitness of the model composed of the centroids C_i.

The diagnostic procedure

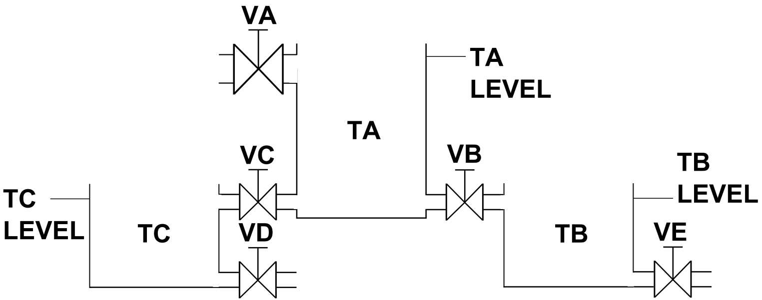
The steps of the diagnostic procedure

1. **Training phase.** Every input trace is converted to trace coordinate form for every training scenario. Only outputs are participating further in clustering (the inputs and the sequence numbers are not present in this form).
2. The training and validating sets T_i, V_i and therefrom the characteristic trace centroids C_i are created for each scenario $i = 1, ..., n$.
3. The diagnostic model is validated using C_i and sets V_i after all centroids are determined. FDR_i values are calculated for every scenario $i = 1, ..., n$.
4. Each cluster centre C_i is labeled with the inputs of scenario i and the particular fault (those are fixed).
5. **Diagnosis phase.** Given an unknown trace which is converted into *trace coordinate form*, the nearest centroid can be determined by computing its distances from centroids C_i. The fault index i which corresponds to the nearest centroid is regarded as the most probable fault mode of the system during the execution of the unknown trace.

A simple example

The 3-tank system

Controlled composite process system with three tanks that is driven by an operational procedure



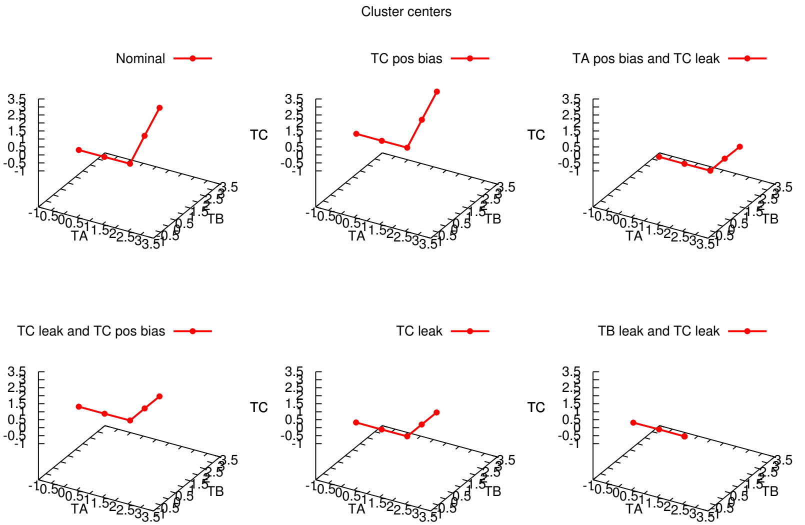
The considered faults

- **The leak/rupture of the tank**. The size of the leak prevents any fluid from staying inside of the tank, therefore fluid level constantly stays at qualitative value 0.
- **The positive bias failure** of the level sensor. The level sensor always detects a qualitative value one degree higher than the actual level of the tank.
- **The negative bias failure** of the level sensor. The level sensor always detects a qualitative value one degree lower than the actual level of the tank.

Training sets: formed by simulation using 1 neighbourhood level measurement error on the traces.

Characteristic trace centroids

Axes represent level output values for the three tanks, with qualitative mapping $e- = -1$, $0 = 0$, $L = 1$, $N = 2$, $H = 3$. Single dot represents a centroid where all values are the same over time.



FDR values for single and dual faults

Results with **linear mapping function**. Scenarios are sorted in ascending order by their *FDR* value. Values for *TA − leak*, *TA and TB leaks* and *TA and TC leaks* are smaller than 0.9 hence not shown