A data-driven adaptive model-identification based large-scale sensor management system: application to Self Powered Neutron Detectors (SPND)

Nahit Pawar

M. Belur, M.Bhushan, A.P. Tiwari, M.G. Kelkar, M. Pramanik, V. Singh

Electrical Engineering Department, Indian Institute of Technology Bombay, India

IEEE Conf. on Evolving and Adaptive Intelligent Systems Linz, Austria, June 2-4, 2014

June 8, 2014

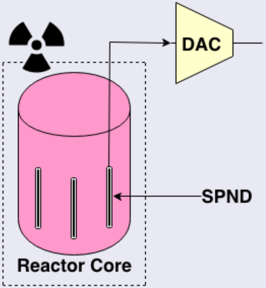
Overview

Introduction

About SPNDs

- SPNDs are widely used to measure neutron flux inside nuclear reactor.
- Flux measurement provides direct measurement of reactor output power.
- Nuclear reactor considered for this work has 144 SPNDs.
- Positioned at different location inside the reactor core.

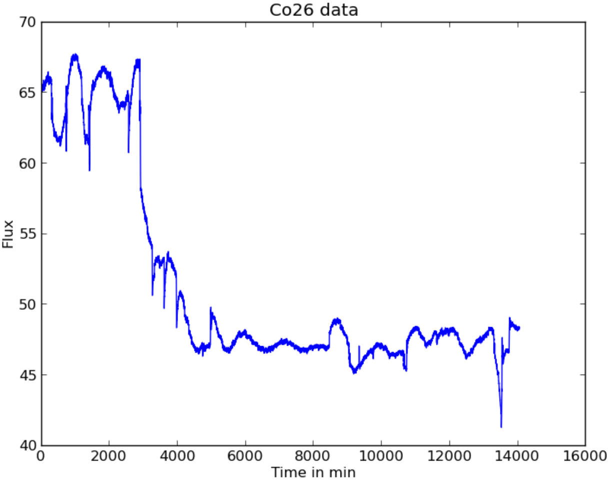
SPNDs Location



Problem:

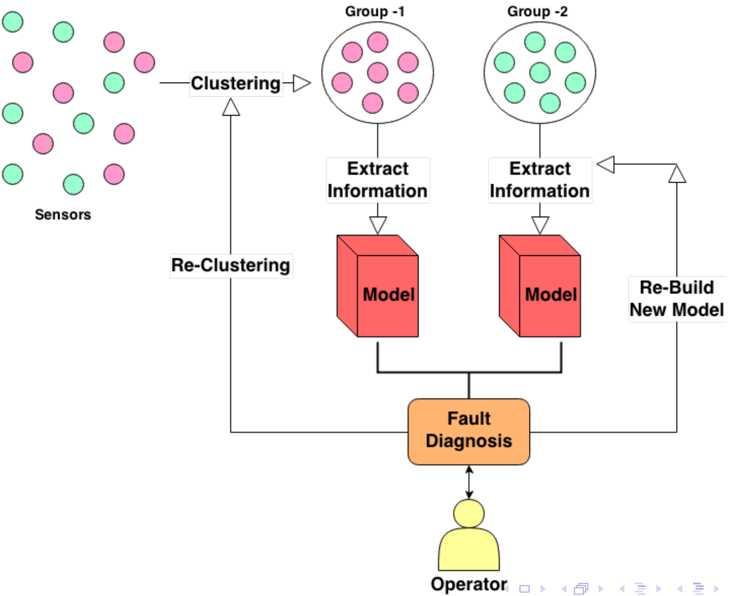
- SPNDs can become faulty due various reasons like - aging, mechanical and electrical failure.
- Results in forced shutdown of reactor to replace faulty sensor.

SPND Output



Co26 data
Flux
Time in min
70
65
60
55
50
45
40
0
2000
4000
6000
8000
10000
12000
14000
16000

Overall Proposed Solution



SQLite and Pytables

SQLite - Sensor database

- SQLite is used to maintain database of all 144 sensors data.
- It is the most widely deployed SQL database engine in the world.
- sqlite3 module provides interface for using SQLite in python.

Pytables - model and cluster database

- Pytables is a python module that is build on top of HDF5 library.
- Hierarchical Data Format (HDF5) is a set of data format and libraries designed to store large amount numerical data.

Clustering

- SPNDs with strongly correlated measurements are grouped together.
- Used k-means algorithm available in Pycluster module for python.
- High correlation is decided by - absolute Pearson correlation distance.

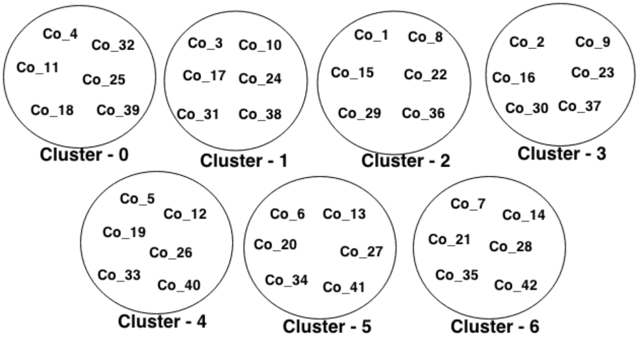
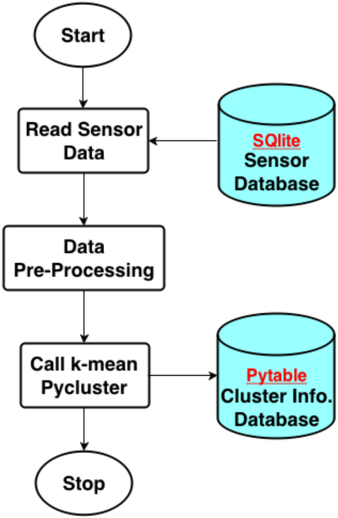
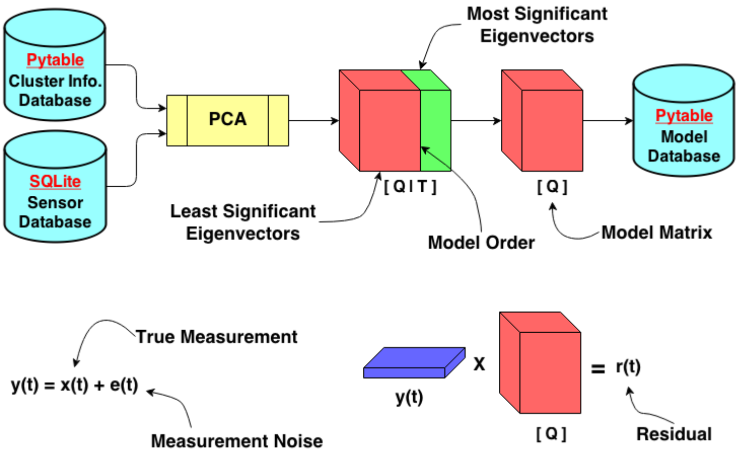


Figure : On of the cobalt SPNDs clustering solution - 7 groups each containing 6 sensors

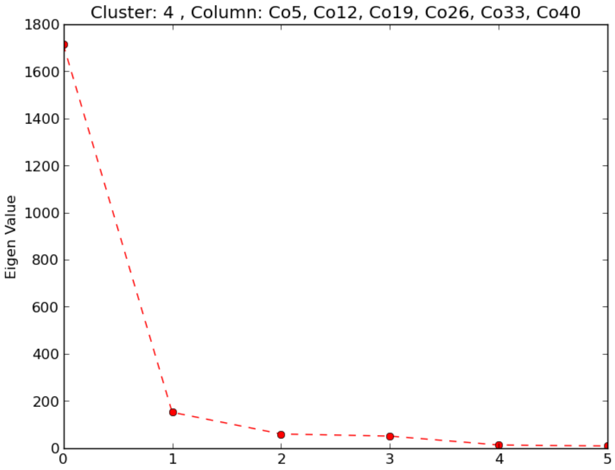
Cluster database



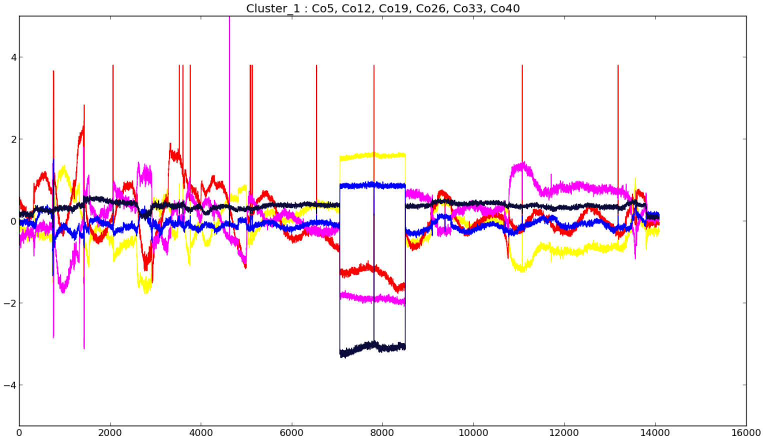
Model Identification using Principal Component Analysis (PCA)



Eigenvalue plot - Co sensors



Residuals plot - Co sensors



Fault Diagnosis - Gross Error Detection

- For each cluster the following test statistic is considered:

$$\gamma(t) = r(t)^T V^{-1} r(t) \tag{1}$$

  V is covariance matrix of residuals

- In the absence of gross error $\gamma(t)$ follows χ^2-distribution with m degrees of freedom and α as confidence level.
- A gross error is detected if $\gamma \geq \chi^2_{1-\alpha,m}$
- Used python's statistical library scipy.stats for χ^2 test.
- Once gross error is detected for a cluster, the variable(s) responsible for gross errors along with estimates of gross error is identified using Generalized Likelihood Ratio (GLR).

Fault Diagnosis - Data Reconciliation

- It is process to obtain more precise estimate of true value which would satisfy the process model.
- The estimate is given by the following weighted least square optimization problem:

$$\min_{x(t)} \; (y(t) - x(t))^T \mathbf{W} (y(t) - x(t)) \tag{2}$$

$$\text{constraint} \quad Ax(t) = 0$$

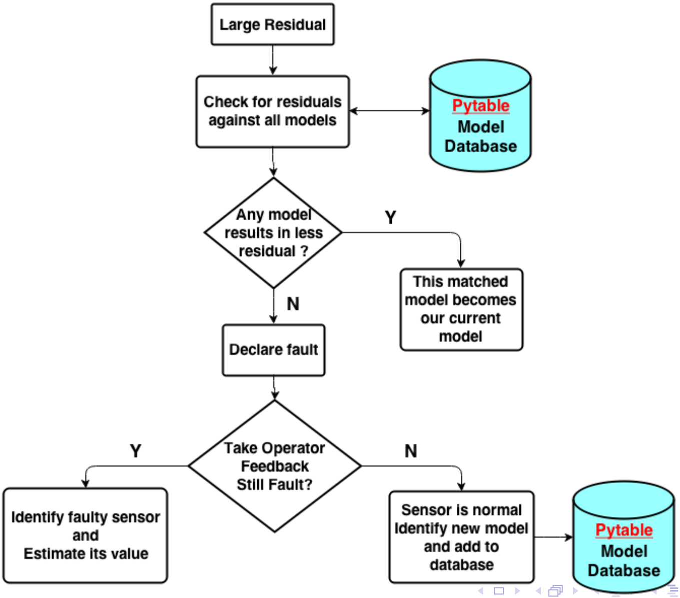
W is weighting matrix taken as inverse of noise covariance matrix.

- The solution to the above optimization problem is given by:

$$\hat{x}(t) = y(t) - W^{-1}A^T(AW^{-1}A^T)^{-1}Ay(t) \tag{3}$$

$\hat{x}(t)$ are reconciled estimate of the true values at time t.

Adaptive model updating



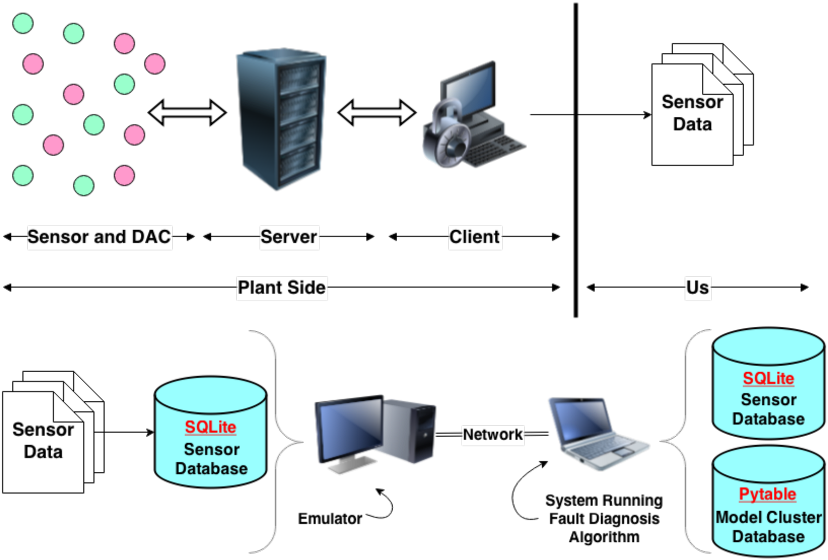
Adaptive re-clustering

- Re-clustering decision is at higher level than updating model.
- Cluster configuration may change depending on normal operating condition of plant.

Rules used for re-clustering sensors:

1. If faults are detected in multiple clusters simultaneously.
2. Once again take operator feedback before re-clustering.

Emulating server-client interaction



Results - Profiling - Cluster building

- **cProfile** module of python is used for profiling.
- **Cluster building**:
  - Data points = 7000

SPND	**No. of clusters**	**npass**	**nfound**
Cobalt	7	100	13
Vanadium	12	100	1

Table : Clustering Parameters

npass = No. of times clustering algorithm is run.
nfound = No. of times optimal solution is found.

SPND	Time(sec) - **build_cluster()**	Time(sec) - **kcluster()**
Cobalt	**4.192**	**3.292**
Vanadium	**25.278**	**23.277**

Table : Timing analysis - Cluster building

Results - Profiling - Model building and Gross error detection

- **Model Building**:
  - Data points = 7000

SPND	Time(sec) - Serial Code	Time(sec) - Parallel Code
Cobalt	**0.428**	**0.015**
Vanadium	**1.314**	**0.058**

Table : Timing analysis - Model building

- **Gross Error Detection**:
  - Data Points = 14061

SPND	Time(sec) - **error_detect()**	Time(sec) - per cluster
Cobalt	**20.055**	**2.867**

Table : Timing analysis - Gross error detection

Conclusion and Future work

Real-Time

The key challenge would be to implement the proposed techniques in real-time.

Real-Time Communication

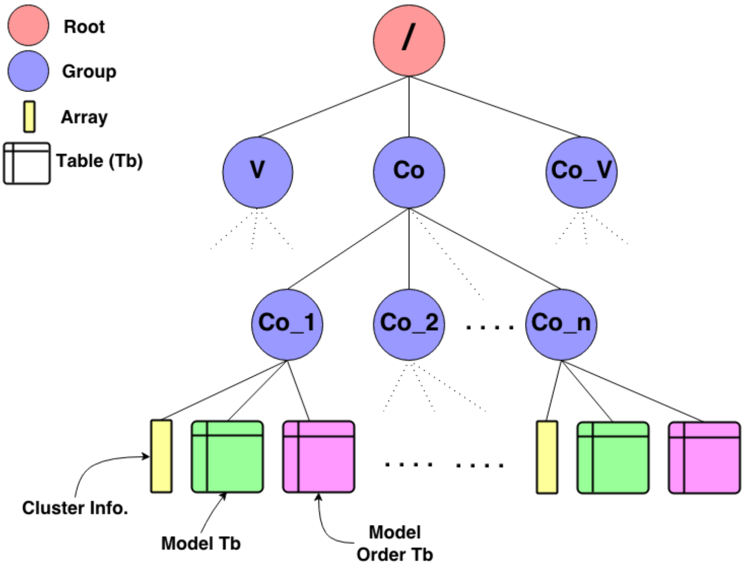
Actual interaction with the client system and remotely situated server requires knowledge of network protocol and thus establishment of real time communication.

Fully-supervised

The decisions for re-clustering and re-building of models are required to be ratified by the operating personal.

Thank you!

Pytable - cluster and model database



GLR

- S. Narasimhan and R.S.H Mah. *Generalized Likelihood Ratios for gross error identification in dynamic process.* AIChE Journal, 34:1321-1331, 1998.
- S. Narasimhan and C. Jordache. Data reconciliation and gross error detection: intelligent use of process data. Gulf Pub. Co., 2000.