Robust nuclear signal reconstruction by a novel ensemble model aggregation procedure

P. Baraldi[1], E. Zio[1,*], G. Gola[1], D. Roverso[2], M. Hoffmann[2]

[1]Department of Energy, Polytechnic of Milan, Via Ponzio 34/3, 20133, Milano, Italy

[*]enrico.zio@polimi.it

[2]Institutt for energiteknikk, OECD Halden Reactor Project, 1751, Halden, Norway

Abstract

Monitoring of sensor operation is important for detecting anomalies and reconstructing the correct values of the signals measured. This can be done, for example, with the aid of auto-associative regression models. However, in practical applications difficulties arise because of the need of handling large numbers of signals. To overcome these difficulties, ensembles of reconstruction models can be used. Each model in the ensemble handles a small group of signals and the outcomes of all models are eventually combined to provide the final outcome. In this work, three different methods for aggregating the model outcomes are investigated and a novel procedure is proposed for obtaining robust ensemble-aggregated outputs. Two applications are considered concerning the reconstruction of 920 simulated signals of the Swedish Forsmark-3 Boiling Water Reactor (BWR) and 215 signals measured at the Finnish Pressurized Water Reactor (PWR) situated in Loviisa.

1. Introduction

Sensors are placed at various locations in a production plant to monitor its state and consequently operate its control and protection systems. Early detection of sensor malfunctions and reconstruction of the signals measured is then necessary [1, 2].

In real plants there are thousands of sensors whose signals cannot be effectively handled by a single reconstruction model [2-5]. A possible way to overcome this problem is by subdividing the set of signals into small overlapping groups, building a corresponding number of reconstruction models and aggregating their outcomes to provide the ensemble-aggregated output [6-11]. This latter task is crucial for the good performance of the ensemble.

In this work, three methods of aggregating the outcomes of the models in an ensemble are first investigated: Simple Mean (SM), Median (MD) and Trimmed Mean (TM) [12]. SM amounts to using all the outcomes of the individual models in the ensemble; MD considers only the central value in the distribution of the models outcomes; TM discards the outcomes in the tails of the distribution. A novel procedure based on the combination of the MD and TM aggregation approaches is then proposed.

The groups at the basis of the ensemble are generated by randomly selecting their signals [12-14]. This allows injecting high diversity in the group structure; for each group, a regression model based on Principal Component Analysis (PCA) [15-18] is developed.

The paper is organized as follows. In Section 2, the three methods for aggregating the models outcomes are described. Two applications are illustrated in Section 3: the first one concerns the reconstruction of a data set of 920 signals obtained with the HAMBO code [19] which simulates the Forsmark-3 Boiling Water Reactor (BWR) located in Sweden; the second addresses the reconstruction of 215 signals measured at the Finnish Pressurized Water Reactor (PWR) situated in Loviisa. The novel procedure is applied and tested on both applications. Conclusions on the advantages and limitations of the aggregation methods are drawn in the last Section.

2. Methods of aggregation of the outcomes of the models in the ensemble

Given the set of $n >> 1$ signals f_i, $i = 1,2,...,n$, measured in the plant, a set of K groups of $m << n$ signals are generated by randomly sampling the signals from the n available [12-14]. The procedure is simple, allows a direct and fast group generation suitable for large scale applications, guarantees high signal diversity between the groups (and thus high diversity between the models outcomes, beneficial to ensemble reconstruction) and attains high signal redundancy, upon a reasonable choice of the ensemble parameters m and K [14].

The K diverse signal groups generated are used as bases for developing a corresponding number of PCA reconstruction models. To do this, the data set $\mathbf{X}$ of N signal patterns available is partitioned into a training set $\mathbf{X}_{TRN}$ (made of N_{TRN} patterns) and a test set $\mathbf{X}_{TST}$ (made of N_{TST} patterns). The former is used to train the individual models, whereas the latter is used to verify the ensemble performance in the signal reconstruction task.

Each signal i is present in a number K_i of groups and thus a corresponding number of individual PCA models provide its reconstruction. Three different methods are here considered to aggregate the outcomes of these individual models: Simple Mean (SM), Median (MD) and Trimmed Mean (TM).

SM amounts to considering all the K_i available model outcomes in a simple average [12, 14, 20, 21]; for the generic pattern $t = 1, 2, ..., N_{TST}$:

$$\hat{f}_i^{E,SM}(t) = \frac{1}{K_i}\sum_{k=1}^{K_i} \hat{f}_i^k(t) \qquad i = 1, 2, ..., n \tag{1}$$

where $\hat{f}_i^{E,SM}(t)$ is the ensemble aggregate of the reconstructions $\hat{f}_i^k(t)$ of pattern t of signal i provided by the individual models $k = 1, 2, ..., K_i$ containing signal i.

The MD aggregation approach amounts to considering for the generic pattern t of signal i only the single outcome $\hat{f}_i^{k_C}(t)$ lying in the centre of the distribution of the K_i model outcomes for that pattern, i.e.:

$$\hat{f}_i^{E,MD}(t) = \hat{f}_i^{k_C}(t) \qquad i = 1, 2, ..., n \tag{2}$$

where k_C denotes the index of the model whose outcome is central with respect to the reconstructed values of the K_i models including signal i.

The TM approach amounts to discarding a fraction ϑ_{TM} of the K_i outcomes of signal i at the tails of the distribution and then simply averaging the $K_i^{TM} = (1-2\vartheta_{TM})K_i$ remaining outcomes:

$$\hat{f}_i^{E,TM}(t) = \frac{1}{K_i^{TM}} \sum_{k=1}^{K_i^{TM}} \hat{f}_i^k(t) \qquad i = 1, 2, ..., n \tag{3}$$

In a sense, the TM approach represents a compromise between the SM and MD methods, for it allows discarding the tails of the distribution while still considering multiple outcomes in the ensemble reconstruction of the signal. In fact, $\vartheta_{TM} \in \left[0,\ \frac{1}{2}\left(1-\frac{1}{K_i}\right)\right]$, the lower limit corresponding to considering all the available outcomes (SM), the upper limit representing the case of using the single central outcome (MD).

Finally, to evaluate the performance of the ensemble aggregates, first the absolute signal reconstruction error is computed[1]:

$$\varepsilon_i^{E,\,(SM,\,MD,\,TM)} = \frac{1}{N_{TST}} \sum_{t=1}^{N_{TST}} \left| f_i(t) - \hat{f}_i^{E,\,(SM,\,MD,\,TM)}(t) \right| \tag{4}$$

Then, the ensemble performance index is retained as the average of the absolute signal reconstruction errors (Eq. 4):

$$\eta^{E,\,(SM,\,MD,\,TM)} = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^{E,\,(SM,\,MD,\,TM)} \tag{5}$$

3. Applications

The applications concern the reconstruction of a set of 920 simulated signals of the Swedish Forsmark-3 BWR and a set of 215 signals measured at the Finnish PWR located in Loviisa. Table 1 reports the main characteristics and parameters adopted in the two case studies.

The PCA models have been constructed with the code http://lib.stat.cmu.edu/multi/pca, adapted to perform the signal reconstruction task of interest here.

The robustness of the ensemble has been tested on the reconstruction of signals when in presence of sensor failures, e.g. random noises or offsets. Within the proposed ensemble approach, a faulty sensor sends a faulty signal in input to the PCA models which contain that signal; each PCA model reconstructs the value of the signal by auto-associating the information

[1] In the application that follows, each signal of the validation set has been previously normalized in the range [0.2, 1], for convenience.

of the non-faulty signals in the model; finally, the ensemble aggregate of the individual outcomes of the models is obtained.

To verify the performance of the ensemble, disturbs are introduced in the test set of both case studies to challenge the corresponding ensemble in the reconstruction of the true, undisturbed signal. More precisely, the signals of a test pattern are randomly affected either by a random noise (with probability $p^{RN} = 0.12$) or by setting their value equal to the offset value of the corresponding sensor (with probability $p^{OF} = 0.08$); with probability 0.8 the signal values remain unchanged.

Case study			Forsmark-3 Boiling Water Reactor (BWR)	Loviisa Pressurized Water Reactor (PWR)
Signals	Number of signals, n		920	215
	Type of signals		Simulated with the HAMBO code [20]	Measured *in situ*
Data set	Number of patterns available, N		5463	12713
	Number of training patterns, N_{TRN}, in X_{TRN}		3600	8000
	Number of test patterns, N_{TST}, in X_{TST}		1863	4713
	Type of measurements [2]		Start up, normal operation and shut down	Normal operation and transients related to two outages
	Sampling rate		1 hour	1 hour
Wrapper-optimized ensemble parameters [14]	Signal redundancy, K_i		7	7
	Number of signals per group, m		70	38
	Number of groups in the ensemble [3], K		92	40
Aggregation parameters	Fraction ϑ_{TM} used in the TM aggregation		0.2	0.2
	Number of models used for the reconstruction of each signal by each aggregation method	SM	7	7
		MD	1	1
		TM	5	5

Table 1. Main characteristics and parameters of the two case studies

Table 2 reports the results for the two case studies in terms of the values of the ensemble reconstruction performance indexes (Eq. 5) on the test patterns obtained by the three aggregation methods on undisturbed ($\eta^{E,U}$) and disturbed ($\eta^{E,D}$) signals. In general, the results show the benefits of excluding some outcomes from the signal reconstruction. Overall, the MD ensemble provides the most accurate and robust signal reconstruction, with a slight improvement with respect to the TM technique.

$\times 10^{-2}$	Forsmark-3 Boiling Water Reactor (BWR)	Loviisa Pressurized Water Reactor (PWR)

[2] Notice that in the Forsmark-3 case study both $\mathbf{X}_{TRN}$ and $\mathbf{X}_{TST}$ include signal measurements related to start up, normal operation and shut down conditions, whereas in the Loviisa case study transients related to the first outage are included in $\mathbf{X}_{TRN}$, while those of the second outage are in $\mathbf{X}_{TST}$.

[3] Once m and K_i are set, the number of groups to generate by random feature selection is obtained from the relation $mK = nK_i$. For a more detailed explanation of the procedure, the interested reader may refer to [14].

	SM	MD	TM	SM	MD	TM
$\eta^{E,U}$	2.091	1.878	1.894	0.977	0.956	0.957
$\eta^{E,D}$	8.172	7.607	7.638	6.221	5.169	5.261

Table 2. Ensemble reconstruction performance indexes obtained with SM, MD and TM aggregations on undisturbed and disturbed signals for the Forsmark-3 and Loviisa case studies.

To further delve into the MD and TM approaches, their reconstruction capabilities on each signal are analyzed. For the Forsmark-3 and Loviisa applications, respectively, Tables 3a and 3b summarize numerically the comparison on the training and test sets between MD and TM in terms of the number of signals for which one method outperforms the other (i.e. $\varepsilon_i^{E,MD} < \varepsilon_i^{E,TM}$ for MD outperforming TM or viceversa) and the corresponding average error gain in using the best performing one. For the Forsmark-3 case study (Table 3a), approximately half of the signals are better reconstructed by MD and half by TM; nevertheless, on disturbed signals, the average error reduction achieved by MD is larger, i.e. the improvements of using MD are more relevant than those obtained with TM, even though on a slightly smaller number of signals; on the other hand, in the Loviisa case study (Table 3b), results show that on disturbed signals MD performs better on a large number of signals (more than two thirds) with a considerably higher average error reduction than TM.

Aggregation method	Number of signals for which the aggregation method is better / Average error reduction ($\times 10^{-3}$)			
	Training		Test	
	Undisturbed	Disturbed	Undisturbed	Disturbed
MD	356 / 0.14	451 / 2.92	474 / 1.58	420 / 2.97
TM	564 / 0.13	469 / 1.60	446 / 1.34	500 / 1.92

Table 3a. Forsmark-3 case study: comparison of MD and TM aggregation methods on undisturbed and disturbed training and test signals

Aggregation method	Number of signals for which the aggregation method is better / Average error reduction ($\times 10^{-3}$)			
	Training		Test	
	Undisturbed	Disturbed	Undisturbed	Disturbed
MD	78 / 0.79	155 / 5.68	113 / 4.55	163 / 6.12
TM	137 / 0.73	60 / 2.88	102 / 4.89	52 / 2.91

Table 3b. Loviisa case study: comparison of MD and TM aggregation methods on undisturbed and disturbed training and test signals

On the basis of these insights, a novel procedure is proposed for combining MD and TM with the aim of exploiting the advantages of both methods. The idea is to define a reconstruction scheme finalized to discern automatically which method is the most effective for reconstructing a signal. The procedure (sketched in Figure 1) is based on the ensemble reconstruction errors obtained on the training set by the MD and TM methods. The conjecture is that if the training patterns of a signal are better reconstructed by one method (MD or TM), then the same method will better reconstruct also the signal's test patterns. The procedure is tailored with respect to the disturbed training signals for application to the reconstruction of the test signals (be it undisturbed or disturbed).

For each signal i, the ensemble reconstruction errors $\varepsilon_i^{E,MD}$ and $\varepsilon_i^{E,TM}$ by the MD and TM methods, respectively, are computed using the disturbed training set: then, the method by which reconstructing the test patterns of signal i is MD if $\varepsilon_i^{E,TM} > \varepsilon_i^{E,MD}$ or, viceversa, TM if $\varepsilon_i^{E,TM} < \varepsilon_i^{E,MD}$ (Figure 1).

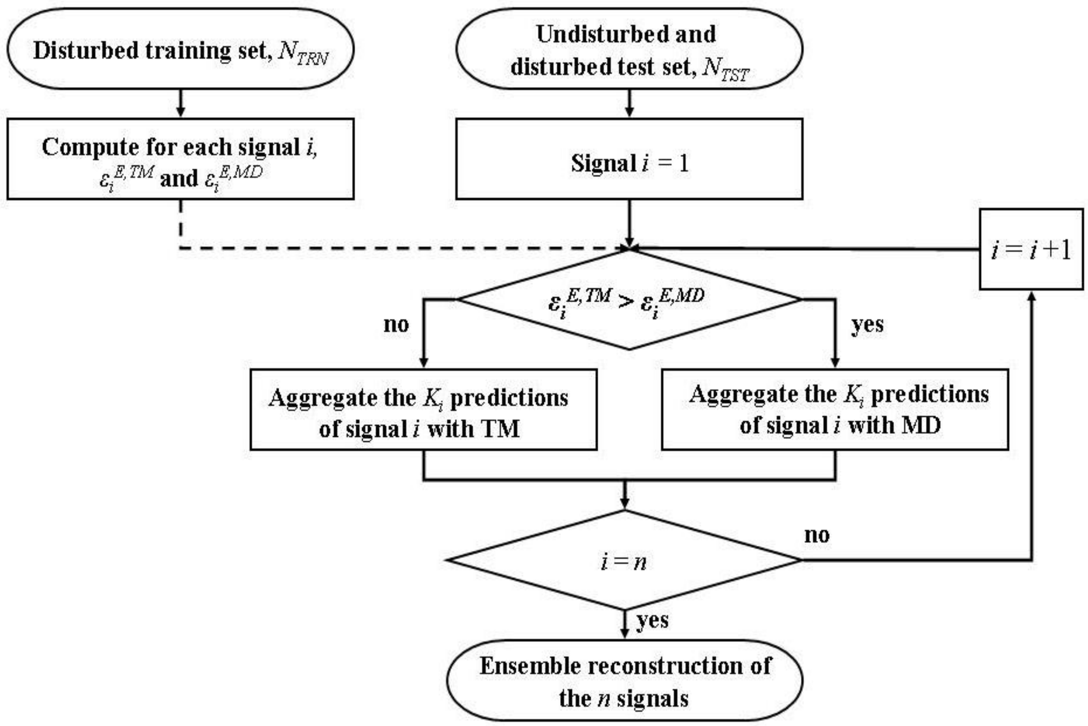

Figure 1. Sketch of the novel procedure for combining the MD and TM aggregation methods

Table 4 reports the ensemble reconstruction performances on undisturbed ($\eta^{E,U}$) and disturbed ($\eta^{E,D}$) test signals obtained by MD, TM and the proposed mixed approach (MX) on the Forsmark-3 and Loviisa case studies. In both, the MX approach provides a more robust reconstruction of the disturbed test set, at the expenses of a negligible loss of accuracy on undisturbed signals.

$\times 10^{-2}$	Forsmark-3 Boiling Water Reactor (BWR)			Loviisa Pressurized Water Reactor (PWR)		
	MD	TM	MX	MD	TM	MX
$\eta^{E,U}$	1.878	1.894	1.890	0.956	0.957	0.958
$\eta^{E,D}$	7.607	7.638	7.298	5.169	5.261	5.023

Table 4. Ensemble reconstruction errors obtained with MD, TM and MX approaches on undisturbed and disturbed signals for the two case studies

5. Conclusions

This work addresses the problem of reconstructing the correct signal values measured by faulty sensors in nuclear power plants. The task is rather complex due to the large number of signals involved. A feasible approach to treat the high dimensionality of the problem is to resort to an ensemble of models for signal reconstruction.

To construct the ensemble, in this work the set of signals is first subdivided into small, overlapping groups by random feature selection. Then, one PCA-based auto-associative reconstruction model is developed based on the signals of each group. The outcomes of the models have finally been aggregated by three methods to obtain the ensemble reconstruction output: Simple Mean (averaging all the available model outcomes), Median (taking only the single outcome in the centre of the distribution of outcomes) and Trimmed Mean (discarding the model outcomes lying in the tails of the distribution).

These methods have been applied to two high-dimensional problems regarding the reconstruction of 920 simulated signals of a nuclear boiling water reactor and 215 signals measured at a pressurized water reactor. The advantages of discarding the outlying outcomes in the ensemble aggregation have been proved in both cases. A novel procedure has then been developed for combining the Median and Trimmed Mean approaches in a way to fully exploit their benefits. The procedure has been applied to both case studies showing an increased robustness in reconstructing signals affected by disturbs.

References

[1] M. Hoffmann, “On-line Monitoring for Calibration Reduction”, HWR-784, OECD Halden Reactor Project, October 2005.

[2] M. Hoffmann, “Signal Grouping Algorithm for an Improved On-line Calibration Monitoring System”, Proceedings of FLINS, Genova, Italy, Aug. 2006.

[3] D. Roverso, M. Hoffmann, E. Zio, P. Baraldi, G. Gola, “Solutions for plant-wide on-line calibration monitoring”, 2007, Proc. ESREL 2007, Stavanger, Norway, Vol. 1, pp. 827-832.

[4] E. Zio, P. Baraldi, G. Gola, D. Roverso, M. Hoffmann, “Genetic Algoritms for Grouping of Signals for System Monitoring and Diagnostics”, 2007, Proc. ESREL 2007, Stavanger, Norway, Vol. 1, pp. 833-840.

[5] P. Baraldi, E. Zio, G. Gola, D. Roverso, M. Hoffmann, “Genetic algorithms for signal grouping in sensor validation: a comparison of the filter and wrapper approaches”, Journal of Risk and Reliability (JRR), 2008, Proc. IMechE, Vol. 222, Part O, pp. 189-206.

[6] M. P. Perrone and L. N. Cooper, “When networks disagree: ensemble methods for hybrid neural networks”, 1992, National Science Fundation, USA.

[7] A. Krogh and J. Vedelsby, “Neural network ensembles, cross-validation and active learning”, in: G. Tesauro, D. S. Touretzky and T. K. Loen, 1995, Advances in newel information processing systems, Vol. 7, pp. 231-238, MIT press, Cambridge, MA, USA.

[8] A. J. C. Sharkey, “On combining artificial neural nets”, 1996, Connection Science, Vol. 8(3), pp. 299-314.

[9] G. Gola, E. Zio, P. Baraldi, D. Roverso, M. Hoffmann, “Signal Grouping for Sensor Validation: a Multi-Objective Genetic Algorithm Approach”, HWR-852, OECD Halden Reactor Project, Feb. 2007.

[10] P. Baraldi, E. Zio, G. Gola, D. Roverso, M. Hoffmann, Reconstruction of faulty signals by an ensemble of principal component analysis models optimized by a multi-objective genetic algorithm, Proceedings of FLINS, Madrid, Spain, Sept. 2008.

[11] G. Gola, E. Zio, P. Baraldi, D. Roverso, M. Hoffmann, Reconstructing signals for sensor validation by a GA-optimized ensemble of PCA models, HWR-894, OECD Halden Reactor Project, Apr. 2008.

[12] R. Polikar, Ensemble based systems in decision making, 2006, IEEE Circuits and Systems Magazine, Vol. 6(3), pp. 21-45.

[13] R. Bryll, R. Gutierrez-Osuna, F. Quek, Attribute bagging: improving accuracy of classifiers ensembles by using random feature subsets, 2003, Pattern Recognition, Vol. 36, pp. 1291-1302.

[14] P. Baraldi, G. Gola, E. Zio, D. Roverso, M. Hoffmann, A randomized model ensemble approach for reconstructing signals from faulty sensors, submitted to…

[15] I.T. Jolliffe, “Principal Component Analysis”, Springer Eds., 2002.

[16] K.I. Diamantaras, S.Y. Kung, “Principal component neural networks: theory and applications”, John Wiley & Sons, Inc. New York, NY, USA, 1996.

[17] B. Scholkopf, A. Smola and K.R. Muller, “Kernel principal component analysis”, Advances in Kernel Methods-Support Vector Learning, 1999.

[18] B. Moore, “Principal component analysis in linear systems: Controllability, observability, and model reduction”, IEEE Transactions on Automatic Control, Vol. 26, Issue 1, Feb. 1981.

[19] T. Karlsson, H. Jokstad, B.D. Meyer and C. Nihlwing, “The HAMBO BWR Simulator of HAMMLAB”, HWR-663, OECD Halden Reactor Project, Feb. 2001.

[20] L. I. Kuncheva, “Classifier Ensembles for Changing Environment”, Multiple Classifier Systems, in: “Lecture Notes on Computer Science, Vol. 3077, 2004, Springer Berlin / Heidelberg Eds.

[21] J. Kittler, M. Hatef, R. P. W. Duin, J. Matas, “On Combining Classifiers”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, Vol. 20 (3), pp. 226-239.