A Data-Driven Architecture for Sensor Validation Based on Neural Networks

Hossein Darvishi*, Domenico Ciuonzo†, Eivind Rosón Eide‡, Pierluigi Salvo Rossi*
*Department of Electronic Systems, Norwegian University of Science and Technology, 7491 Trondheim, Norway
†University of Naples "Federico II," Naples 80125, Italy
‡Kongsberg Digital AS, Norway
Email: hossein.darvishi@ntnu.no; domenico.ciuonzo@unina.it; eivind.roson.eide@kdi.kongsberg.com; salvorossi@ieee.org

***Abstract*—In this paper, we propose a novel sensor validation architecture, which performs sensor fault detection, isolation and accommodation (SFDIA). More specifically, a machine-learning based architecture is presented to detect faults in sensors measurements within the system, identify the faulty ones and replace them with estimated values. In our proposed architecture, sensor estimators based on neural networks are constructed for each sensor node in order to accommodate faulty measurements along with a classifier to determine the failure detection and isolation. Finally, numerical results are presented to confirm the effectiveness of the proposed architecture on a publicly-available air quality (AQ) chemical multi-sensor data-set.**

***Index Terms*—Fault tolerance, neural networks, sensors.**

I. Introduction

With the new wave of digitalization, digital twins are at the core of the development process within Industry 4.0. Accordingly, sensors constitute the driving force for the accomplishment of this concept [1]. However, sensors are prone to failure and faulty data may negatively affect functionalities of the monitored system. Accordingly, SFDIA is a crucial practice since it can hinder faulty sensors from leading systems to catastrophic consequences. In this context, numerous approaches have been developed in the literature related to the use of analytical redundancy techniques for sensor fault detection and isolation. Such techniques can be mainly categorized into two groups: *model-based* methods and *data-driven* (or more generally model-free) methods.

The most widely used model-based methods comprise (multiple-model) Kalman filter [2], [3] and observer-based [4] approaches. Despite their appeal, model-based methods require an accurate mathematical model of the system, whose constitutive parameters are difficult to apply in the presence of nonlinearities. On the other hand, data-driven methods for SFDIA schemes have attracted significant attention by the scientific community due their ease of implementation and capabilities to capture nonlinear behavior by learning from historical data [5]–[9]. Data-driven methods include neural networks (NNs) and other machine-learning approaches [6], [8]–[10], hidden Markov models [11], fuzzy logic [12] and principal component analysis [13], whose successful application has been demonstrated to manifold systems. These comprise diesel-engines, gas-turbines, wireless sensor networks and air-crafts.

[0]This work was partially supported by the Research Council of Norway under the project SIGNIFY within the IKTPLUSS framework.

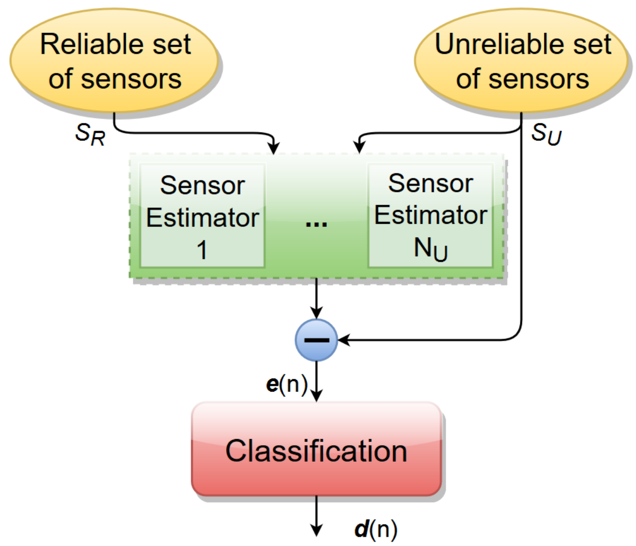


Fig. 1: The proposed system architecture for SFDIA.

In this work, we propose a machine-learning-based framework for sensor validation with different applications. The proposed architecture takes advantage of reliable and unreliable sensors' measurements as well as their temporal correlation. Synthetically-generated weak bias faults were added to a data-set of a chemical multi-sensor device to evaluate the presented SFDIA architecture. The benefits of the proposed approach are the flexibility in terms of the application domain, the capability to promptly deal with weak faults and (not explored here) with simultaneous faults of multiple sensors.

The outline of this manuscript is the following. Sec. II describes the proposed machine-learning based SFDIA architecture. The description of the considered multi-sensor data-set used in this work and numerical results are provided in Sec. III. Some final remarks are given in Sec. IV.

Notation – Lower-case bold letters and bold numbers denote vectors and $(\cdot)^T$, denotes transpose operator.

II. Proposed SFDIA Architecture

In this section, we briefly describe the three-layer system architecture (illustrated in Fig. 1). More specifically, we con-

978-1-7281-6801-2/20/$31.00 ©2020 IEEE

sider a system monitored via $(N_R + N_U)$ different sensors. Sensors measurements constitute the input of the proposed SFDIA system, where measurements are divided into two sets: N_R reliable sensors (set S_R), which represent supportive data, and N_U unreliable sensors (set S_U), which are prone to failure. Still, we underline that the present architecture does not necessarily require the presence of reliable sensors.

A. Estimation Layer

According to Fig. 1, input sensors data enter the first layer with N_U independent sensor estimators, namely *virtual sensors*. Each virtual sensor receives all sensors' data except for the sensor under estimation from time instant n to $n-m$ (i.e. using a sliding window of length $m+1$) as input and estimates the measurement of the sensor under estimation at time n as output. Outputs of the estimators are exerted to replace the isolated faulty data by the SFDIA system at the last layer. A classic multilayer perceptron (MLP) [7] architecture is considered for each virtual-sensor implementation.

B. Error Computation

The estimated measurement from each virtual sensor is then subtracted from the respective unreliable sensor measurement in the second layer to obtain N_U error signals, collected within $\boldsymbol{e}(n)$. Error signals measure the dissimilarity between the normal and faulty status of unreliable sensors, wherein the case of perfect estimation and no faulty sensors $\boldsymbol{e}(n) = \mathbf{0}$.

C. Classification Layer

The last stage of the proposed architecture consists of a classifier which aims at (i) detecting and (ii) identifying faulty measurements from the set of unreliable sensors S_U. In detail, the classification stage accepts the error vectors inputs at time instants n to $n-k$, namely $\boldsymbol{e}(n), \dots, \boldsymbol{e}(n-k)$ (i.e. a sliding window of length $k+1$). The error vectors are used by the classifier as a metric for fault detection and isolation. Accordingly, the decision vector output is in the format $\boldsymbol{d}(n) = [d_0(n), d_1(n), \dots, d_{N_U}(n)]^T$ (with $d(n) \in [0,1]$). Therein, $\{d_0(n) = 1\}$ denotes the event that no sensor failure is present, while other decision elements $\{d_i(n) = 1\}$ with $i = 1, \dots, N_U$ indicate failure on the ith unreliable sensor.

More specifically, the classifier is made of a two-layer MLP with a softmax output activation function and $N_U + 1$ output nodes. The classifier softmax output gives a decision vector representing the probability distributions of the vector of potential outcomes. Thus, decision element with the highest probability represents the occurred event

$$i_m = \underset{i \in 0, \dots, N_U}{\text{argmax}} \ d_i(n) \ ,$$

where i_m points to the largest element of the decision vector (i.e. it represents the event with the highest probability of occurrence). Finally, if an unreliable sensor is declared in failure, its measurements are replaced with the estimated values from the corresponding virtual sensor.

Briefly, $i_m = 0$ vs. $i_m \neq 0$ represents the *detection task*, being equivalent to "no fault detected" $\{d_0(n) = 1\}$ vs. "fault detected" $\{d_0(n) = 0\}$. In the case $i_m \neq 0$, the specific values of i_m performs the *isolation task* and replacing faulty sensor measurements with corresponding virtual sensor estimates employs the *accommodation task*.

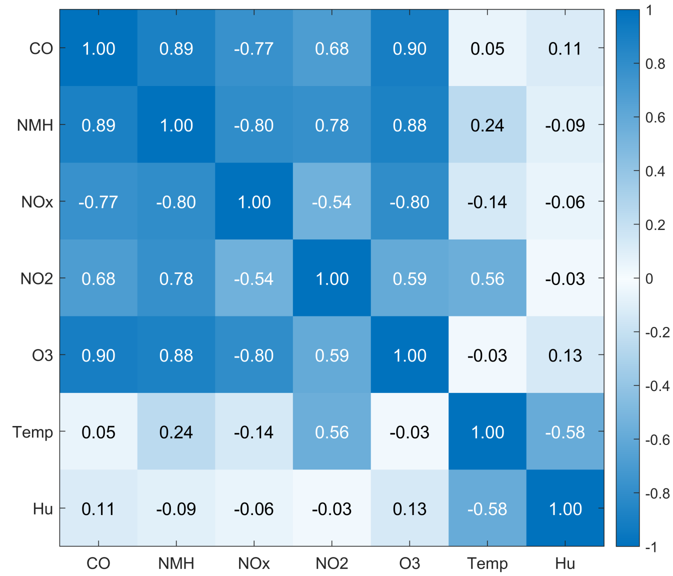


Fig. 2: correlation matrix of sensor pairs for AQ data-set.

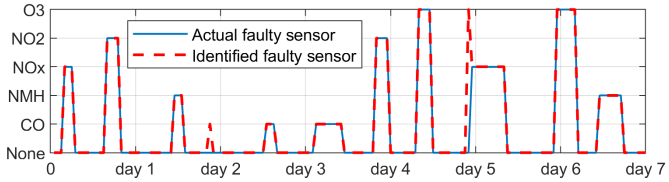
III. Data-Set Description and Numerical Results

The proposed architecture is applied to an air quality (AQ) data-set with 5 metal oxide chemical sensors embedded in an AQ chemical multi-sensor device installed on the field in an Italian city [14]. Hourly averaged measurements of the multi-sensor device consisting of carbon monoxide (CO), Non-Metanic Hydrocarbons (NMH), Nitrogen Oxides (NOx), Nitrogen Dioxide (NO2) and ozone (O3) gas concentrations are considered as unreliable set. Moreover, measurements of temperature (Temp) and humidity (Hu) in the AQ data-set are used within reliable set in this study. Accordingly, we have:

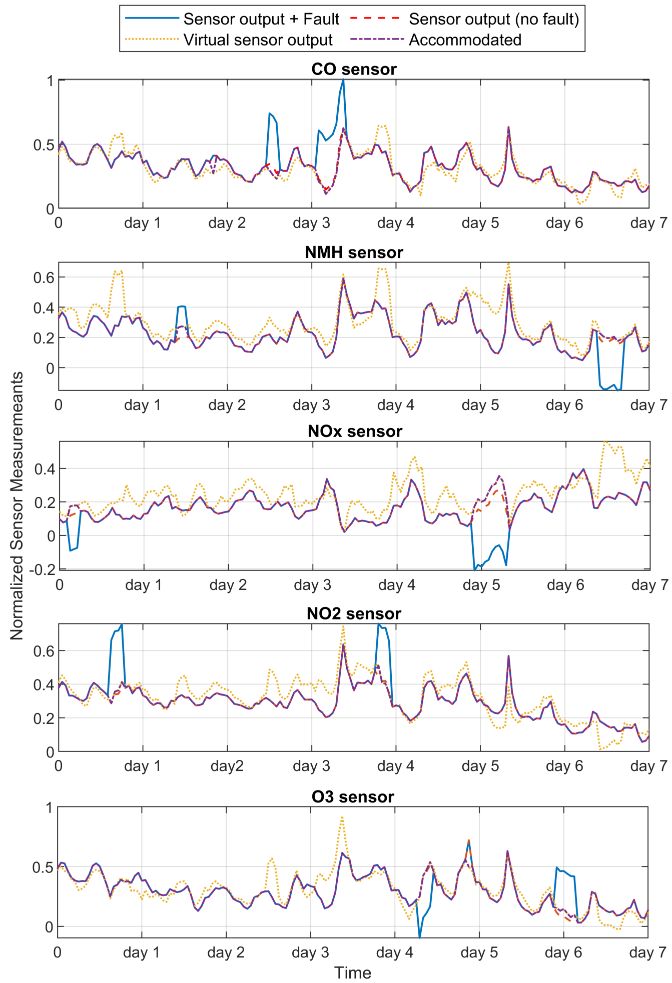
$$\mathcal{S}_U = \{\text{CO}, \text{NMH}, \text{NOx}, \text{NO2}, \text{O3}\} \quad (N_U = 5)$$
$$\mathcal{S}_R = \{\text{Temp}, \text{Hu}\} \quad (N_R = 2)$$

Measurements of both sets are normalized via min-max normalization within the range $[0, 1]$. In addition, we dropped missed data before processing the data-set.

Our experimental analysis is carried out by dividing the data-set into a training set accounting for 85% of the first part of data-set and a test set accounting for the remaining 15%. The holdout validation method is used to prevent over-fitting to some extent. Synthetically-generated bias faults are added to the AQ data set to verify the proposed architecture performance. To represent weak faults, we considered positive and negative additive bias faults. The bias absolute level ranges within $[20, 40]\%$ of each sensor measurements' variation domain on the train set. Five MLP virtual sensors (estimators) with one single hidden layer (made of 10 neurons) are trained to provide estimation of the $N_U = 5$ unreliable



(a) synthetically generated faults



(b) sensors' outputs

Fig. 3: Output of different stages of the proposed SFDIA architecture for bias faults over one week of the test set.

metal oxide chemical sensors. Differently, two hidden layers with 15 neurons per layer are considered for the classifier. Also, the size of the sliding window is assumed to span 10 samples for both the estimators and the classifier (i.e. $m = 10$ and $k = 10$).

Fig. 2 shows the correlation coefficient between different sensor pairs. Indeed, a higher correlation between sensor pairs would lead to more accurate estimators (viz. virtual sensors) in the first layer. As a result, this would imply a higher-precision classifier, since error signals represent difference in actual and virtual sensors' measurements. Results highlight significant

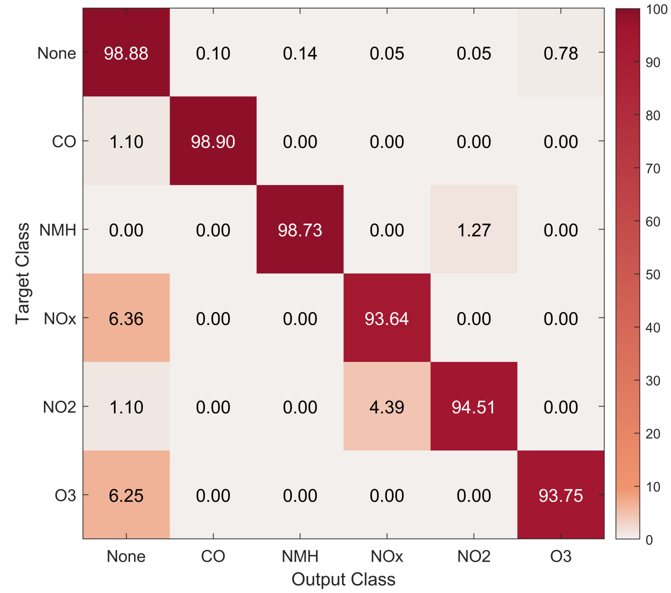


Fig. 4: Normalized confusion matrix for all classes during the test period. Numbers are in percent.

dependencies among different pairs, which indicates the feasibility of our data-driven SFDIA.

The output of several parts of SFDIA architecture for one week of test set is shown in Fig. 3. More specifically, Fig. 3(a) monitors the faults on different sensors where the proposed architecture successfully detects and identifies all faults without delay in the system (dashed line) with only two false declaration samples (false positive) in the first and fifth days. As can be seen in Fig. 3(b), after fault identification, system accommodates isolated faulty data with its estimation to ensure the fault-free performance of the system.

Finally, the (normalized) confusion matrix on the test set is presented in Fig. 4. The confusion matrix shows excellent accuracy of the proposed architecture, i.e. classification rate about 96.5%. All classes show high precision over 90%, with the lowest precision exhibited on O3 and NOx sensors with values 93.75% and 93.64%, respectively.

IV. Conclusions and Future Directions

This manuscript presented a machine-learning based architecture for SFDIA scheme in real-time operation. MLP-based virtual sensors provided appropriate estimates of unreliable sensors to replace corresponding corupted measurements in presence of faults, while an MLP-based classifier was responsible for detection and isolation of faults. The proposed architecture is validated by real-world data from AQ monitoring sensors, and results illustrate the prompt detection, isolation and accommodation of sensors' failures with less than 2.6% of faults on average remained undetected on the test set. Future directions will include the use of deep networks for the modules of the proposed SFDIA and type-of-fault classification.

REFERENCES

[1] S. Yin, S. X. Ding, and D. Zhou, "Diagnosis and prognosis for complicated industrial systems—Part I," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2501–2505, 2016.

[2] I. Samy, I. Postlethwaite, and D.-W. Gu, "Detection and accommodation of sensor faults in UAVs-a comparison of NN and EKF based approaches," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 4365–4372.

[3] B. Pourbabaee, N. Meskin, and K. Khorasani, "Sensor fault detection, isolation, and identification using multiple-model-based hybrid kalman filter for gas turbine engines," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 4, pp. 1184–1200, 2016.

[4] S. Amin, X. Litrico, S. Sastry, and A. Bayen, "Cyber security of water SCADA Systems—Part I: analysis and experimentation of stealthy deception attacks," *IEEE Transactions on Control Systems Technology*, vol. 21, pp. 1963–1970, 09 2013.

[5] M. R. Napolitano, Y. An, and B. A. Seanor, "A fault tolerant flight control system for sensor and actuator failures using neural networks," *Elsevier Aircraft Design*, vol. 3, no. 2, pp. 103–128, 2000.

[6] S. Hussain, M. Mokhtar, and J. M. Howe, "Sensor failure detection, identification, and accommodation using fully connected cascade neural network," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1683–1692, 2015.

[7] S. Gururajan, M. L. Fravolini, H. Chao, M. Rhudy, and M. R. Napolitano, "Performance evaluation of neural network based approaches for airspeed sensor failure accommodation on a small UAV," in *21st IEEE Mediterranean Conference on Control and Automation*, 2013, pp. 603–608.

[8] G. Campa, M. Thiagarajan, M. Krishnamurty, M. R. Napolitano, and M. Gautam, "A neural network based sensor validation scheme for heavy-duty diesel engines," *ASME Journal of dynamic systems, measurement, and control*, vol. 130, no. 2, 2008.

[9] S. Gutiérrez and H. Ponce, "An intelligent failure detection on a wireless sensor network for indoor climate conditions," *MDPI Sensors*, vol. 19, no. 4, p. 854, 2019.

[10] S. Mandal, B. Santhi, S. Sridhar, K. Vinolia, and P. Swaminathan, "Nuclear power plant thermocouple sensor-fault detection and classification using deep learning and generalized likelihood ratio test," *IEEE Transactions on Nuclear Science*, vol. 64, no. 6, pp. 1526–1534, 2017.

[11] C. Alippi, S. Ntalampiras, and M. Roveri, "Model-free fault detection and isolation in large-scale cyber-physical systems," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 1, no. 1, pp. 61–71, 2017.

[12] S. Ogaji, L. Marinai, S. Sampath, R. Singh, and S. Prober, "Gas-turbine fault diagnostics: a fuzzy-logic approach," *Applied Energy*, vol. 82, no. 1, pp. 81 – 89, 2005.

[13] M. Z. Sheriff, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Fault detection using multiscale PCA-based moving window GLRT," *Journal of Process Control*, vol. 54, pp. 47 – 64, 2017.

[14] S. De Vito, E. Massera, M. Piga, L. Martinotto, and G. Di Francia, "On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario," *Sensors and Actuators B: Chemical*, vol. 129, no. 2, pp. 750–757, 2008.