Anomaly Detection and Fault Disambiguation in Large Flight Data: A Multi-modal Deep Auto-encoder Approach

Kishore K. Reddy[1], Soumalya Sarkar[2], Vivek Venugopalan[3], Michael Giering[4]

[1,2,3,4] *United Technologies Research Center (UTRC), East Hartford, CT, USA*
reddykk@utrc.utc.com, sarkars@utrc.utc.com, venugov@utrc.utc.com, GierinMJ@utrc.utc.com

Abstract

Flight data recorders provide large volumes of heterogeneous data from arrays of sensors on-board to perform fault diagnosis. Challenges such as large data volumes, lack of labeled data, and increasing numbers of sensors (multiple modalities) exacerbate the challenges of being able to hand-craft the features needed for state-of-the-art PHM algorithms to effectively perform system diagnosis. In this paper, the authors propose leveraging existing unsupervised learning methods based on Deep Auto-encoders (DAE) on raw time series data from multiple sensors to build a robust model for anomaly detection. The anomaly detection algorithm analyzes the reconstruction error of a DAE trained on nominal data scenarios. The reconstruction error of individual sensors is examined to perform fault disambiguation. Training and validation are conducted in a laboratory setting for various operating conditions. The proposed framework does not need any hand-crafted features and uses raw time series data. Our approach is tested on data from the NASA open database and demonstrates high fault detection rates ($\sim 97.8\%$) with zero false alarms. Our paper also demonstrates robust fault disambiguation on two different fault scenarios. Moreover, the paper provides a strong rationale for utilizing deep architecture (multi-hidden-layer neural network) via thorough comparison with a single hidden-layer DAE.

1. Introduction

It is of paramount importance to detect, disambiguate and monitor faults in an extremely complex system like aircraft for maintaining adequate levels of aviation safety and reliability. Effective fault diagnostics and monitoring also help in scheduling condition-based maintenance (CBM) efficiently. However, real-time fault diagnostics in an aircraft is a highly challenging task, mainly due to the presence of multi-modal sensors (e.g., thermocouple, accelerometer, pressure sensor, speed sensor, current/voltage sensors etc), distributed electro-

mechanical couplings, noisy data and various operating conditions. Moreover, the task of fault disambiguation brings more complexities into the picture because it is necessary to analyze signatures from a multitude of sensors in a holistic way to eventually separate faults with overlapping properties. The recent rise of on-board PHM in the presence of ubiquity and redundancy of sensors has also generated the demand for robust diagnostic tools which can handle large amounts of multi-modal and multi-scale data.

Fault diagnostics and disambiguation have been addressed from various points of view for many years. There are multiple techniques in literature and practice, which can be categorized into mainly three types: 1)model-based, 2) data-driven, and 3) hybrid fault diagnostic approaches. Early diagnostics approaches were mainly based on analytical models (Gertler, 1988; Rizzoni & Min, 1991; Simani, Fantuzzi, & Beghelli, 2000) of systems or sensors and they computed the residuals of the runtime signal to detect the presence of fault. There are few types of parity based approaches, namely, Parity Space Approach (PSA) (Gertler, 1997), Parity Equation Approach (PEA) (S. Kim, Kim, & Park, 2004), Generalized Likelihood Ratio Test (GLT), and Least Square Residual Approach (LSRA). Other approaches are based on Bank of observers (state estimators) (Chen & Saif, 2007; Tan & Edwards, 2002; Kiyak, Cetin, & Kahvecioglu, 2008), Kalman filter residual computation (Gustafsson, 2002; Hagenblad, Gustafsson, & Klein, 2003), fault detection filters (Gertler, 1988), and parameter identification (Litt, Kurtkaya, & Duyar, 1994). The main challenges of model-based approaches are centered around the lack of robustness due to insufficient model fidelity and the simplifying assumptions impact on uncertainty.

As system complexities rapidly increase, building adequate models for them is more difficult. This is why, there have been a surge of data-driven fault diagnostic techniques that are built upon the theories of signal processing and machine learning (Duda, Hart, & Stork, 2000). Regular data-driven approaches consist of two steps, (i) feature extraction from signals to capture fault signatures, and (ii) learning and clas-

sification of the features to identify faults. Statistical features or domain knowledge specific features (Seda, Kadir, & Dogru, 2007; S. J. Kim & Lee, 1999), Symbolic Dynamic Filtering (SDF) based features (Das, Sarkar, Ray, Srivastava, & Simon, 2013; Sarkar, Mukherjee, Sarkar, & Ray, 2013; Sarkar, Sarkar, & Ray, 2014), wavelet features, cyclostationary features (Gardner, Napolitano, & Paura, 2006), and few other time domain or frequency domain features have shown significant promise in the area of data-driven diagnostics. In case of the learning step, there are various machine learning techniques such as Artificial Neural Network (ANN) (Balaban, Saxena, Bansal, Goebel, & Curran, 2009), expert system based methods, supervised and semi-supervised classification approaches (Support Vector Machine, k Nearest neighbor classifier) (Duda et al., 2000). Issues with data-driven approaches are related to their high sensitivity to the extracted features and the challenges of generalizing across a range of operating conditions.

The hybrid approaches include a combination of PEA and wavelet based signal features (S. Kim et al., 2004), Principal Component Analysis (PCA) based system models (Hagenblad et al., 2003) etc. Faults diagnostics approaches have employed sensor fusion at multiple levels which are data level, feature level and decision level. An example of data level and feature level fusion technique captures the SDF-dependent cross relationship (Sarkar, Sarkar, Mukherjee, Ray, & Srivastav, 2013; Sarkar, Sarkar, Virani, Ray, & Yasar, 2014) among multi-modal sensors for aircraft engine fault diagnostics. For engine fault diagnosis, (Basir & Yuan, 2007) demonstrated a decision level sensor fusion via applying Dempster-Shafer evidence theory. Other than these fusion-based diagnosis approaches, there are other model-based fusion methods built upon Kalman filtering, particle filtering.

The shortcomings with most of the fault diagnosis techniques are generally related to the lack of high-fidelity non-linear models, tedious hand-crafting (domain knowledge) of fault features, lack of scalability to large data, insufficient robustness to noise and the presence of various operating modes, or presence of multi-modal sensors for fault disambiguation. This paper proposes a Deep Auto-encoder (DAE) based fault detection and disambiguation approach, which is built upon the concepts of deep learning (Hinton & Salakhutdinov, 2006; Bengio, Lamblin, Popovici, Larochelle, et al., 2007; Ngiam et al., 2011). Deep learning methods have proven to be an advanced breakthrough in machine learning and are based upon multi-layer neural network and optimized by stochastic back-propagation (Hinton & Salakhutdinov, 2006). The recent (last five years) discoveries in deep learning have greatly reduced the classical problems relevant to neural networks such as over-fitting and lack of generalizability. There are a handful of recent efforts that apply deep learning (e.g., deep belief network, sparse auto-encoder on unimodal sensors) specifi-
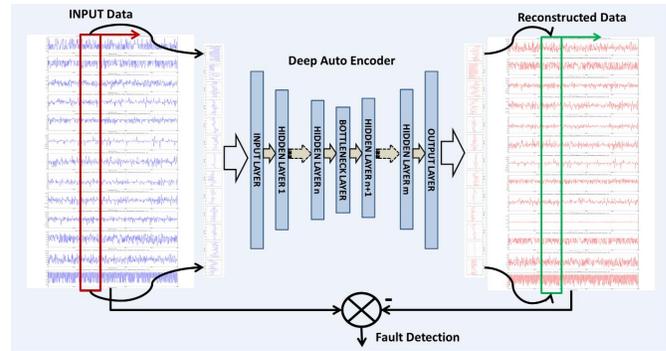
Figure 1. Schematic of the deep auto-encoder framework

cally for PHM applications, namely, fault diagnosis of reciprocating compressor valves (Tran, AlThobiani, & Ball, 2014), CBM of rotating machines (Verma, Gupta, Sharma, & Sevakula, 2013), and early detection of combustion instability faults (Sarkar et al., 2015). The authors propose to capture the nominal signature at various conditions by building and training a large DAE on multi-modal sensor data. The proposed DAE-based framework is trained directly on raw time series from heterogeneous sensors without hand-crafting any specific type of features. The paper demonstrates the functionalities and performance of the proposed technique by testing it on a large set of real data. This paper also successfully disambiguates among different types of faults in an unsupervised fashion. This data was collected from a multi-sensor electro-mechanical actuator for various nominal and fault scenarios (aircraft operation related) by Balaban et. al. (Balaban et al., 2009, 2015).

The paper is organized in five sections, including the present one. Section 2 describes the proposed framework along with its core methodology via explaining the concepts of DAE. Section 3 describes the data collection system in brief, which serves as the data source for experimental validation of the proposed architecture for fault diagnostics and disambiguation. Section 4 presents the performances and advantages of the proposed approach. Finally, the paper is summarized and concluded in Section 5 with selected recommendations for future research.

2. MULTI-MODAL DEEP AUTO-ENCODER (DAE) FRAMEWORK

This section describes the proposed framework for the analysis of multi-sensor time series for fault detection and disambiguation. Figure 1 shows the architecture in a modular fashion. Each of the time series from different sensors are normalized to zero mean and unit variance. Synchronized windows are traversed over multi-modal time series with overlap and windows from each sensor are concatenated. A large vector containing the sliced time series from all sensors are fed

2

into the DAE. The parameters of DAE are learned by back-propagation and stochastic gradient descent (Bengio et al., 2007) based on nominal (no fault) data. More detailed description of DAE is in the following subsection. While testing, the whole tool chain is operated in a feed-forward fashion in real-time and provides the RMS reconstruction error for a time window. The average reconstruction error of a time window over all sensors denotes how far it is from the nominal condition and is used as an anomaly indicator. The distributions of reconstruction error over multi-modal sensors are leveraged to differentiate specific fault signatures. The fault disambiguation in an unsupervised manner (without exposing the model to all fault types in the training phase) is possible due to the advanced representational capability of a DAE. A DAE can capture a large amount of multi-scale and non-linear features of multi-sensor data without significant over-fitting. These statements are supported by adequate visualization and analysis in the section 4.

2.1. Deep auto-encoder (DAE)

DAE are one of the significant variants of deep Learning and put a strong emphasis on modeling multiple levels of data abstraction (from low-level features to higher-order representations, i.e., features of features) from data (Erhan, Courville, & Bengio, 2010). A DAE is constructed by a multi-layer neural network, where there is an input layer, single or multiple hidden layers and an output layer. Each layer can have different number of neural units as shown in the figure 1. A DAE (Bengio et al., 2007) takes an input vector $\mathbf{x} \in \mathcal{R}^d$ and first maps it to the latent representation $\mathbf{h} \in \mathcal{R}^{d'}$ using a deterministic sigmoid function of the type $\mathbf{h} = f_\theta = \sigma(Wx + b)$ with parameters $\theta = \{W, b\}$, where W is the weight and b is the bias. This "code" is then used to reconstruct the input by a reverse mapping of $f : y = f\theta'(h) = \sigma(W'h + b')$ with $\theta' = \{(W'), b'\}$. The two parameter sets are constrained to be of the form $W' = W^T$, using the same weights for encoding the input and decoding the latent representation. Each training pattern x_i is then mapped onto its code h_i and its reconstruction y_i. The parameters are optimized via stochastic gradient descent method (Bengio et al., 2007), minimizing an appropriate cost function over the training set $\mathcal{D}_n = \{(x_0, t_0), ..., (x_n, t_n)\}$. In this paper, the cost function $L(xy)$ is assumed to be the root mean square error between the input vector and reconstructed vector.

$$L(xy) = ||x - y||^2 \qquad (1)$$

For a multi-hidden layer DAE, that is used in this paper, the parameters of each individual layer are first optimized individually according to the method mentioned above. After all the hidden layers are stacked and a mirror version of them is rolled out from bottleneck layer to output layer, the weights of the whole DAE is fine-tuned by deep back-propagation

(Bengio et al., 2007) based on the final reconstruction error for nominal training data.

3. DATA DESCRIPTION FOR VALIDATION

The data for validating the proposed algorithm is downloaded from the NASA DASHlink open database. The original design of experiment and data collection were performed by Balaban et. al. (Balaban et al., 2009, 2015). A brief description of the experimental setup, sensor suits, fault injection methods and fault types are given below. For more details, refer to Balaban et. al. (Balaban et al., 2009, 2015).

A set of electromechanical actuators (EMA), constructed by Moog Corporation, were used by Balaban et. al. (Balaban et al., 2009, 2015) for running different experiments. To increase the horizon of available operating conditions, flyable electromechanical actuator (FLEA) testbed was also constructed. This paper mainly deals with the fault scenarios that are injected in laboratory setting under various operating conditions. Coupling of test actuators to the load actuator is accomplished via an electromagnetic system. Only one test actuator at a time is normally coupled to the load actuator. The data acquisition system consists of two National Instruments 6259 cards and the Galil motor controller. In this paper, the data contains 13 modalities sampled at $100Hz$, which are Time, Actuator Z Position, Measured Load, Motor X Current, Motor Y Current, Motor Z Current, Motor X Voltage, Motor Y Voltage, Motor Y Temperature, Motor Z Temperature, Nut X Temperature, Nut Y Temperature, Ambient Temperature. The accelerometers are connected through custom-fabricated conditioner boards that supply them with excitation voltage and remove the dc portion of the return signal. The current sensors are built in the Galil motor controller. Specific number of nominal scenarios and different fault scenarios are described in section 4.

Two faults under various operating conditions were injected into the test articles in the following manner: 1) A jam fault was injected via a mechanism mounted on the return channel of the ball screw that can stop circulation of the bearing balls through the circuit. 2) A spall fault was injected by introducing cuts of various geometries via a precise electrostatic discharge process. The initial size and subsequent growth of these cuts were confirmed by using an optical inspection and measurement system.

4. RESULTS AND DISCUSSIONS

This section describes the training parameters and the performance results that are obtained when the proposed framework is applied on the validation data for fault detection and unsupervised fault disambiguation.
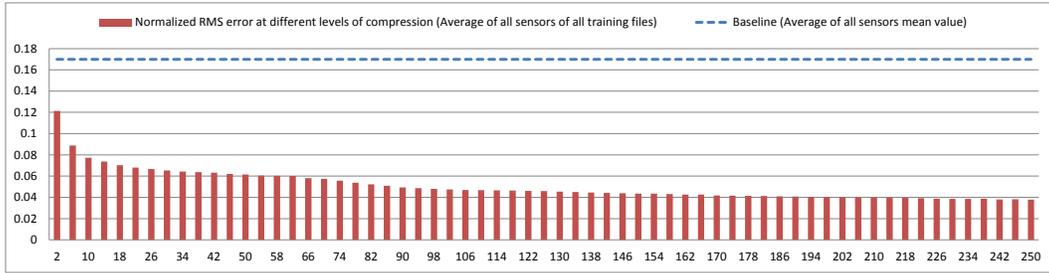
(a)
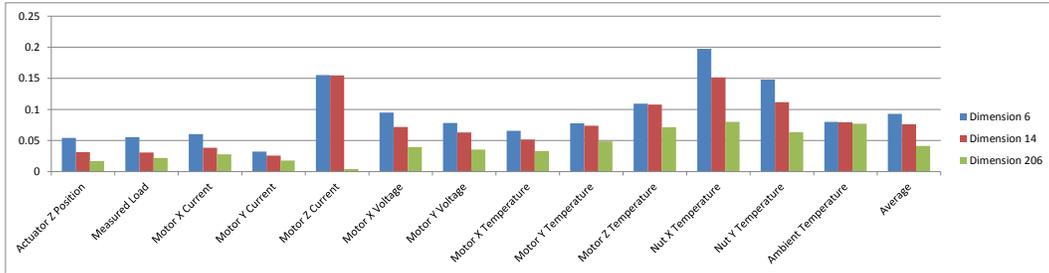
(b)

Figure 3. (a) Variation of normalized root mean square (RMS) error at the reconstructed output layer with increasing dimension of the bottleneck layer, (b) Individual sensor-wise reconstruction errors at the output layer for 3 different bottleneck layer dimensions
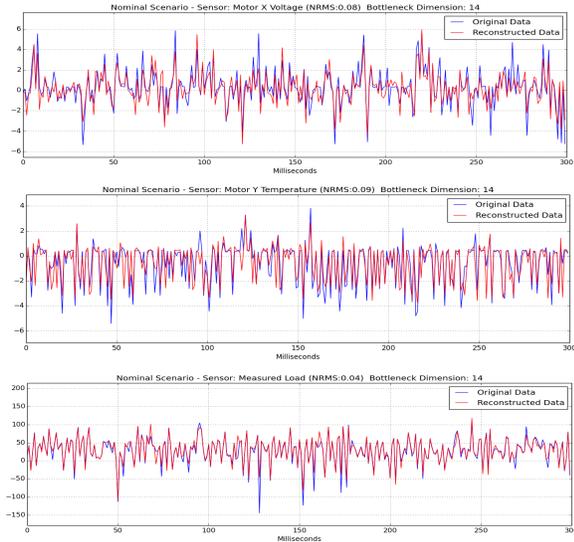
Figure 2. Actual signals and reconstructed signals for Motor X voltage, Motor Y temperature, and load sensors (from top to bottom) with bottleneck layer of 14 dimension

4.1. Training and Parameter Learning

For training, 51 nominal runs of length $\sim 30s$ are considered. The sampling rate of all 13 sensor data is $100Hz$. They are all performed during a certain operating condition in a laboratory setting. To learn the parameters of a nominal DAE we generate batches of nominal samples via a sliding-window. A window length of $0.5s$ (50 data points) with an overlap of $0.49s$ is chosen. After concatenating the windows from all the sensors, the dimension of a single input sample feeding to the DAE is 650 (50×13). The total number of training samples generated from those 51 runs by sliding-window method is $\sim 100,000$. The nominal DAE is trained by mini-batch wise stochastic gradient descent on a GPU with batch size of 128 samples. The dimensions of the first, second, third and fourth hidden layers are 256, 196, 136 and 76 respectively for a bottleneck layer dimension 14. The fifth layer is the bottleneck layer. Starting from the bottleneck layer, the first four hidden layers are repeated in a reverse order to finally arrive at the output layer with the same dimension as the input layer. The DAE has 11 neural layers including one input, 9 hidden layers, and one output layer. Some examples of the actual signal and reconstructed signals for Motor X voltage, Motor Y temperature and load sensors are illustrated in figure 2. The average normalized RMS reconstruction errors for these three sensors are $0.08, 0.09, 0.04$ respectively.

To select the dimension of the bottleneck layer, it is varied from 2 to 250 and the normalized root mean square (RMS) error at the reconstructed output layer is demonstrated in figure 3(a). This is performed on the training data itself. It is observed that the RMS error reduces as the bottleneck compression is relaxed. Choosing a larger bottleneck layer dimension reduces the RMS reconstruction error, but will eventually lead to over-fitting.The abrupt onset if over-fitting can

4

(a)

(b)
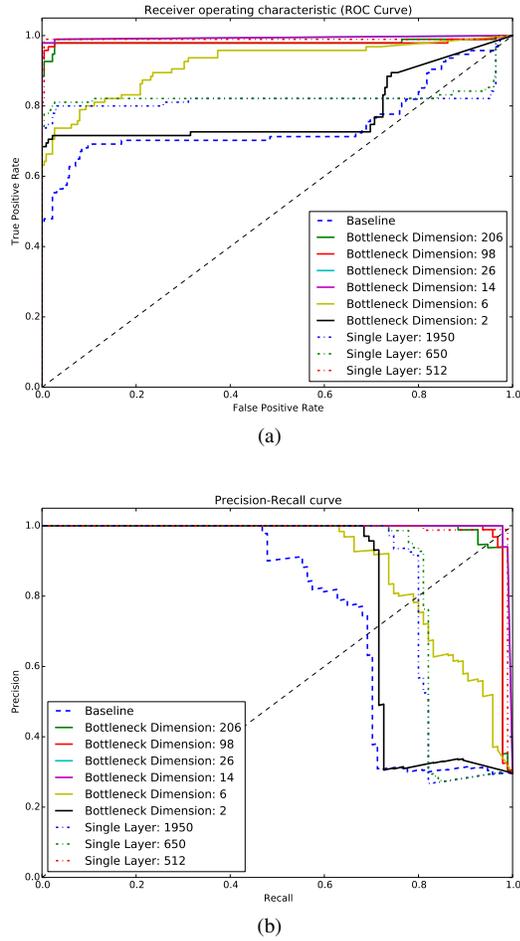
Figure 4. (a) Receiver operating characteristics (ROC) curves via varying detection threshold λ on testing data for different bottleneck layer dimensions of the 11-layer DAE and few single layer DAE models, (b) Precision-Recall curves for the same conditions as (a)

be observed and is discussed in the next subsection in the context of fault diagnostics. We use this information for the selection of an optimal bottleneck layer dimension for fault disambiguation. Figure 3(b) shows sensor-wise reconstruction errors at the output layer for 3 different bottleneck layer dimensions. Other than temperature sensors of nut X and nut Y, and motor Z current sensors, all other sensors have reconstruction errors less than 0.1.

4.2. Fault Diagnostics

During the testing phase, there are a total of 95 nominal runs and 225 faulty runs. Out of 225 faulty runs, there are 15 scenarios of spall faults and 210 scenarios of ballscrew jam fault. The scenarios are typically of length $\sim 30s$. But the test scenarios are performed in few sets of different operating conditions compared to the training scenarios. The sliding-window method is applied on the testing scenarios to generate sam-

ples and those samples are fed through the nominal DAE to obtain the final reconstruction (NRMS) errors averaged over all sensors. A detection threshold λ is applied on the average NRMS error such that a test run with larger NRMS error than λ would be diagnosed as a fault. Varying λ, multiple receiver operating characteristics (ROC) curves are generated and showed in figure 4(a) for different bottleneck layer dimensions and few single hidden-layer DAE models. It is observed that the ROC curve for 14-dimensional bottleneck layer case performs the best in fault detection. For $\lambda = 0.56$, the fault detection rate is 97.8% with 0.0% false alarm. As the number of fault scenarios is much larger than the number of nominal runs in the testing phase, i.e., the testing data set is not balanced, a series of precision-recall curves (Duda et al., 2000) are also shown on figure 4(b). By also considering the proximity to the upper-right corner of precision-recall plots, it can be concluded that the 11-layer DAE model with 14-dimensional bottleneck layer performs best in fault diagnostics. According to the ROC curves, a single hidden-layer DAE model with 512-dimensional bottleneck layer performs similarly to the proposed 11-layer DAE with 14-dimensional bottleneck layer. The reason, for which the proposed DAE is superior compared to a single hidden-layer DAE model, is explained in the following subsections.

4.3. Fault Disambiguation and Comparison with Single Hidden-Layer Auto-encoder

After fault detection, the next step is to disambiguate among multiple fault ensembles or to identify specific fault types. This is a very challenging task because fault signatures often overlap with each other. As shown in figure 5, multiple spider charts visualize the NRMS errors across different sensors for nominal scenarios, spall fault scenarios, and ballscrew jam fault scenarios during testing phase. It is observed that, even without any training on specific type of fault data, the proposed 11-layer DAE with 14-dimensional bottleneck layer can emerge distinguishable patterns of NRMS error distribution (across sensors) for different fault scenarios. The dimension of the NRMS error distribution is same as the number of sensors, i.e., 13.

Figure 6 shows the comparison between the consolidated spider charts that are obtained by using the single hidden-layer DAE model with 512-dimensional bottleneck layer and the 11-layer DAE with 14-dimensional bottleneck layer. It is clear that the proposed model disambiguates the faults via identifiable patterns in the sensor NRMS distributions, visualized on a spider chart, whereas the single hidden-layer DAE model shows minimal separability of individual fault classes. To further showcase the robustness of the fault disambiguation capability, principal component analysis (PCA) (Duda et al., 2000) is performed on 13-dimensional NRMS error distributions for all the testing scenarios. Figure 7 shows the clusters for all testing scenarios over two largest principal
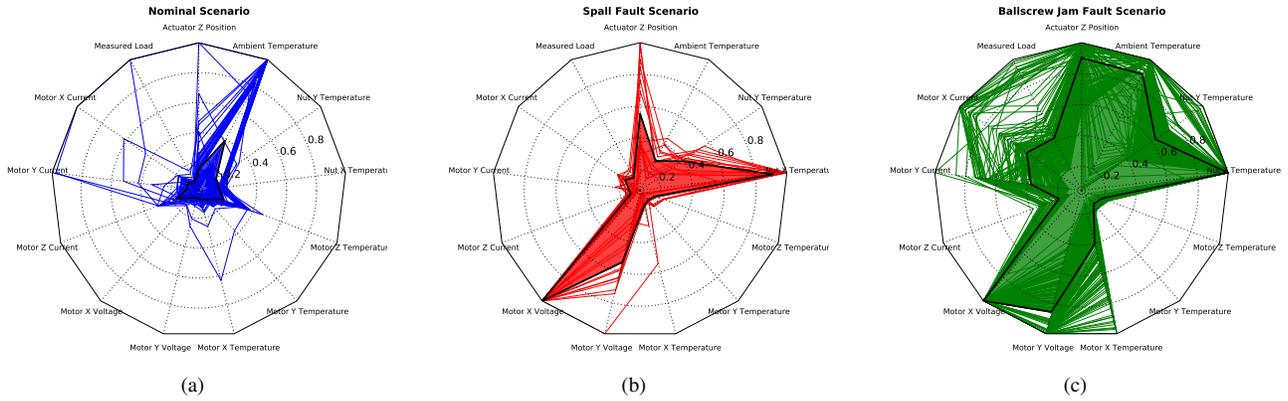
Figure 5. Spider charts showing the NRMS error across different sensors during testing phase for (a) nominal scenarios, (b) spall fault scenarios, and (c) ballscrew jam fault
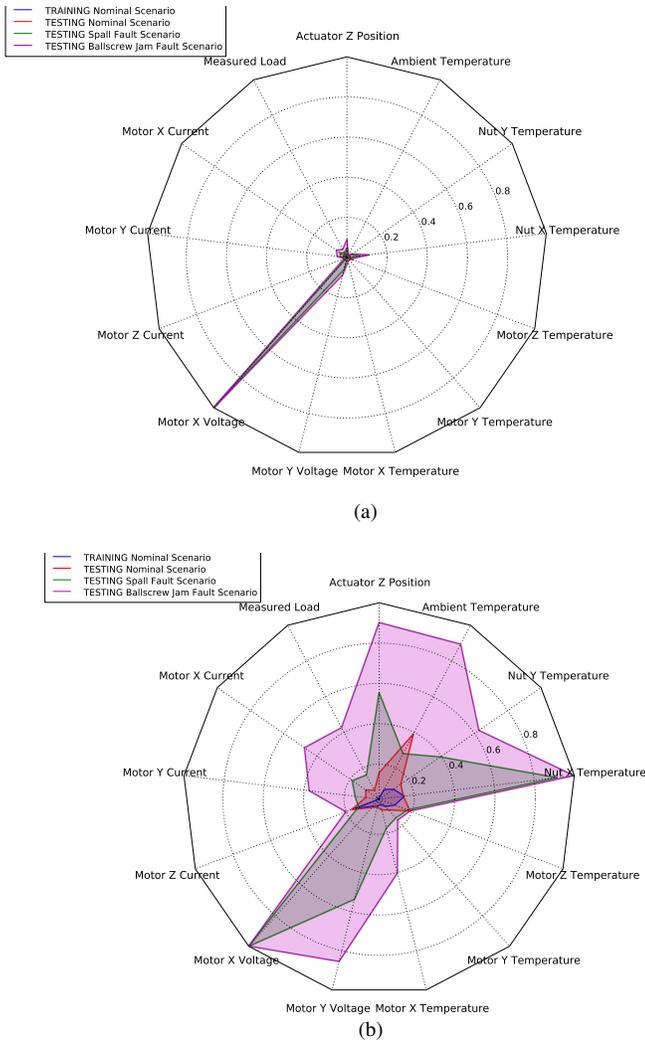
Figure 6. Spider charts of the average (over nominal and fault scenarios) NRMS error across different sensors during testing phase for (a) single hidden-layer DAE model with 512-dimensional bottleneck layer and (b) proposed 11-layer DAE with 14-dimensional bottleneck layer
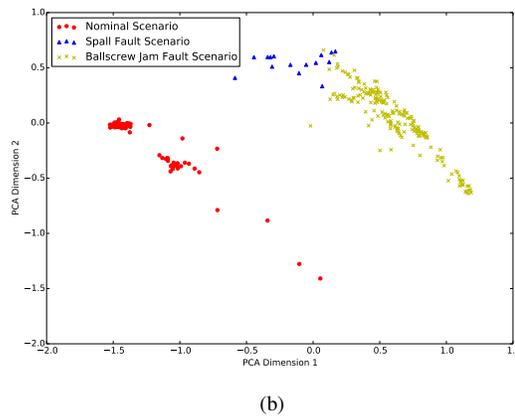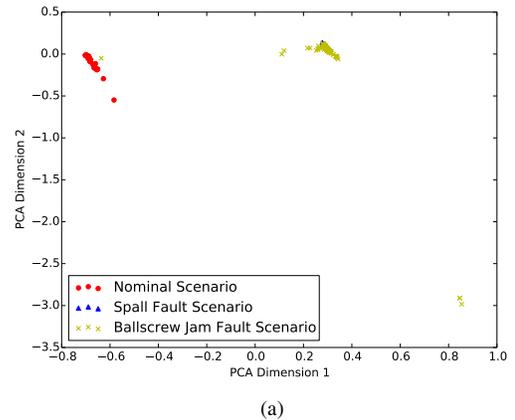
Figure 7. Clusters of two largest principal components obtained from PCA on 13-dimensional NRMS error distributions over all testing scenarios for two models: (a) single hidden-layer DAE model with 512-dimensional bottleneck layer and (b) proposed 11-layer DAE with 14-dimensional bottleneck layer

components using two models: (i) single hidden-layer DAE model with 512-dimensional bottleneck layer and (ii) proposed 11-layer DAE with 14-dimensional bottleneck layer. It is apparent from the figure 7 that, even in a reduced dimension, the proposed model can identify different faults with high confidence in an unsupervised fashion. However, the single hidden-layer DAE model can not distinguish between the faults in lower dimension because it does not have enough representational capacity for the multi-modal sensor data. Although it has comparable anomaly detection performance, it can not classify different faults as well as the proposed 11-layer DAE model due to over-fitting.

5. CONCLUSION AND FUTURE WORK

The authors propose a Deep Auto-encoder (DAE) based fault detection and disambiguation framework, which is built upon the concepts of deep learning. This work is one of the few recent efforts utilizing the large representational capability of deep architectures for addressing complex PHM problems. The paper attempts to capture the nominal signature under a variety of conditions by building and training an 11-layer DAE on multi-modal sensor data. The proposed framework is trained directly on raw time series from heterogeneous sensors without hand-crafting any features based on domain knowledge. The paper demonstrates the training approach and performance of the proposed technique via testing it on a large set of realistic data. A high fault detection rate ($\sim 97.8\%$) along with zero false alarm are achieved by the the 11-layer DAE model. This paper also successfully disambiguates among different types of faults with high confidence in an unsupervised way, i.e., via clustering. A thorough comparison (via visualization of sensor-specific NRMS errors) of the proposed DAE model with a single hidden-layer DAE model demonstrates the necessity of deep architecture for building a robust algorithm for fault detection and diagnostics. Some future research tasks are:

- further validation on a larger number of fault types,
- introduction of supervised classification of faults,
- testing at simultaneous multi-fault scenarios.

ACKNOWLEDGMENT

REFERENCES

Balaban, E., Saxena, A., Bansal, P., Goebel, K. F., & Curran, S. (2009, Dec). Modeling, detection, and disambiguation of sensor faults for aerospace applica-

tions. *IEEE Sensors Journal*, *9*(12), 1907-1917. doi: 10.1109/JSEN.2009.2030284

Balaban, E., Saxena, A., Narasimhan, S., Roychoudhury, I., Koopmans, M., Ott, C., & Goebel, K. (2015). Prognostic health-management system development for electromechanical actuators. *Journal of Aerospace Information Systems*, *12*, 329-344.

Basir, O., & Yuan, X. (2007). Engine fault diagnosis based on multi-sensor information fusion using dempstershafer evidence theory. *Information Fusion*, *8*(4), 379 - 386.

Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., et al. (2007). Greedy layer-wise training of deep networks. *Advances in neural information processing systems*, *19*, 153.

Chen, W., & Saif, M. (2007). A sliding mode observer-based strategy for fault detection, isolation, and estimation in a class of Lipschitz nonlinear systems. *International Journal of Systems Science*, *38*(12), 943–955.

Das, S., Sarkar, S., Ray, A., Srivastava, A., & Simon, D. L. (2013). Anomaly detection in flight recorder data: A dynamic data-driven approach. In *American control conference (acc), 2013* (pp. 2668–2673).

Duda, R. O., Hart, P. E., & Stork, D. G. (2000). *Pattern classification, second ed.* (Vol. 654). New York: John Wiley and Sons.

Erhan, D., Courville, A., & Bengio, Y. (2010). Understanding representations learned in deep architectures. *Department d'Informatique et Recherche Operationnelle, University of Montreal, QC, Canada, Tech. Rep*, *1355*.

Gardner, W., Napolitano, A., & Paura, L. (2006). Cyclostationarity: half a century of research. *Signal processing*, *86*(4), 639–697.

Gertler, J. J. (1988). Survey of model-based failue detection and isolation in complex plants. *IEEE Control Systems Magazine*, 11.

Gertler, J. J. (1997). Fault detection and isolation using parity equations. *Control Engineering Practice*, *5*(5), 653–661.

Gustafsson, F. (2002). Stochastic fault diagnosability in parity spaces. In *15th IFAC world congress, barcelona, spain.*

Hagenblad, A., Gustafsson, F., & Klein, I. (2003). A comparison of two methods for stochastic fault detection and principal component analysis. In *13th ifac symposium on system identification (sysid), rotterdam, nl* (pp. 27–29).

Hinton, G., & Salakhutdinov, R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313.5786*, 504-507.

Kim, S., Kim, Y., & Park, C. (2004). Hybrid fault detection and isolation techniques for aircraft inertial measurement sensors. In *Aiaa guidance, navigation and control conference and exhibit, providence, rh.*

Kim, S. J., & Lee, C. W. (1999). Diagnosis of sensor faults in

active magnetic bearing system equipped with built-in force transducer. *IEEE Transactions on Mechatronics*, *4*(3), 180–186.

Kiyak, E., Cetin, O., & Kahvecioglu, A. (2008). Aircraft sensor fault detection based on unknown input observers. *Aircraft Engineering and Aerospace Technology*, *80*(5), 545–548.

Litt, J., Kurtkaya, M., & Duyar, A. (1994). Sensor fault detection and diagnosis simulation of a helicopter. In *Conference of military, government, and aerospace simulation, san diego, ca.*

Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal Deep Learning. In *Proceedings of the 28th international conference on machine learning (icml-11)* (pp. 689–696).

Rizzoni, G., & Min, P. S. (1991). Detection of sensor failures in automotive engines. *IEEE Transactions on Vehicular Technology*, *40*(2), 487–500.

Sarkar, S., Lore, K. G., Sarkar, S., Ramanan, V., Chakravarthy, S. R., Phoha, S., & Ray, A. (2015). Early detection of combustion instability from hi-speed flame images via deep learning and symbolic time series analysis. In *Annual conference of the prognostics and health management society 2015*.

Sarkar, S., Mukherjee, K., Sarkar, S., & Ray, A. (2013). Symbolic dynamic analysis of transient time series for fault detection in gas turbine engines. *Journal of Dynamic Systems, Measurement, and Control*, *135*(1), 014506.

Sarkar, S., Sarkar, S., Mukherjee, K., Ray, A., & Srivastav, A. (2013). Multi-sensor information fusion for fault detection in aircraft gas turbine engines. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, *227*(12), 1988–2001.

Sarkar, S., Sarkar, S., & Ray, A. (2014). *Data-enabled health management of complex industrial systems*. Fault Detection: Classification, Techniques and Role in Industrial Systems, NOVA Science Publishers.

Sarkar, S., Sarkar, S., Virani, N., Ray, A., & Yasar, M. (2014). Sensor fusion for fault detection and classification in distributed physical processes. *Frontiers in Robotics and AI*, *1*, 16.

Seda, P., Kadir, E., & Dogru, B. E. (2007). Intelligent sensor fault detection and identification for temperature control. In *Proceedings of the 11th wseas international conference on computers, agios nikolaos, greece.*

Simani, S., Fantuzzi, C., & Beghelli, S. (2000). Diagnosis techniques for sensor faults of inustrial processes. *IEEE Transactions on Control Systems Technology*, *8*(5), 848–855.

Tan, C. P., & Edwards, C. (2002). Sliding mode observers for detection and reconstruction of sensor faults. *Automatica*, *38*, 1815–1821.

Tran, V., AlThobiani, F., & Ball, A. (2014). An approach to fault diagnosis of reciprocating compressor valves using teager-kaiser energy operator and deeep belief networks. *Expert Systems with Applications*.

Verma, K., Gupta, V., Sharma, M., & Sevakula, R. (2013). Intelligent condition based monitoring of rotating machines using sparse auto-encoders. *IEEE*.