

## Fault Diagnosis and Isolation in Aircraft Gas Turbine Engines

Soumik Sarkar    Kushal Mukherjee    Asok Ray    Murat Yasar  
*szs200@psu.edu    kum162@psu.edu    axr2@psu.edu    yasarm@technosci.com*

*Pennsylvania State University  
University Park, PA-16802, USA*

**Keywords:** *Symbolic Dynamic Filter, Fault Detection and Isolation, Gas Turbine Engine*

**Abstract**—This paper formulates and validates a novel methodology for diagnosis and isolation of incipient faults in aircraft gas turbine engines. In addition to abrupt large faults, the proposed method is capable of detecting and isolating slowly evolving anomalies (i.e., deviations from the nominal behavior), based on analysis of time series data observed from the instrumentation in engine components. The fault diagnosis and isolation (*FDI*) algorithm is based upon Symbolic Dynamic Filtering (*SDF*) that has been recently reported in literature and relies on the principles of Symbolic Dynamics, Statistical Pattern Recognition and Information Theory. Validation of the concept is presented and a real life software architecture is proposed based on the simulation model of a generic two-spool turbofan engine for diagnosis and isolation of incipient faults.

### I. INTRODUCTION

PERFORMANCE and reliability of gas turbine engines deteriorate due to degradation of their gas path components such as fan, compressor, combustor, and turbines. Common causes of performance degradation include compressor fouling and increase of blade-tip clearance in the turbine due to wear and erosion, labyrinth seal leakage, and corrosion in the hot sections of the gas path components [1]. These physical faults cause changes in the macroscopic measurable parameters, such as temperature, pressure, rotational speed, and fuel flow rate. In a complex multi-component and interconnected system, such as a gas turbine engine, a single fault in one component produces an anomalous output that might serve as the input excitation to other healthy components, and degrade their performance. Therefore, gradual evolution of small anomalies in individual components may lead to cascaded faults eventually leading to catastrophic failures and forced shutdown of the entire system. As such, advanced warning of emerging faults is essential in an aircraft gas turbine engine for degradation monitoring and mitigation of catastrophic failures.

Current state-of-the-art fault diagnostic and prognostic tools that vary widely in their complexity and applications are primarily built upon both model-based and data-driven analysis. A model incorporating the nonlinearities involved in aircraft engine dynamics combined with conventional

optimization methods was first introduced in early 1990s by Stamatis et al. [2] However, conventional optimization methods have a possibility of convergence to a local minimum; and this shortcoming was addressed in genetic algorithms, first proposed by Zedda and Singh [3]. Concept of neural networks, rule-based fuzzy logic has also been introduced in the literature of engine diagnosis. Sensor-based analysis of time series data provides another approach for fault diagnosis and prognosis in mechanical systems [4], [5]. The objective here is to develop and validate a novel methodology for fault diagnosis and isolation (*FDI*) in aircraft gas turbine engines based on time series of pertinent sensors and analytically derived process variables. The *FDI* algorithm is built upon the concept of Symbolic Dynamic Filtering (*SDF*) [6] that is a recently reported technique of pattern recognition based on time series analysis of process variables and enables compression of information into pattern vectors of low dimension [6]. The *FDI* algorithm is validated on the simulation model of a generic two-spool turbofan engine [7]–[9]. The current paper also outlines a possible implementation procedure of the developed software for real life testing of aircraft engines during maintenance.

The paper is organized in six sections including the present one. Section II provides a brief review of symbolic dynamic filtering (*SDF*). Section III presents the proposed concept of fault diagnosis and isolation in the case of a generic two-spool turbofan engine. Application of the *FDI* software in simulation testbed is discussed and also a real life implementation procedure is proposed in section IV. Description of the pertinent features of the simulation experiments results are presented in Section V. The paper is summarized concluded in Section VI along with recommendations for future research.

### II. OVERVIEW OF SYMBOLIC DYNAMIC FILTERING

The theory of symbolic dynamic filtering (*SDF*) for time series data analysis is built upon the underlying principles of Nonlinear Dynamics [10], Symbolic Dynamics [11], Information Theory [12], and Statistical Pattern Recognition [13]. While details on the theory and construction of *SDF* for anomaly detection are reported in recent publications [6], [14], their salient features are summarized below.

This work has been supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under Grant No. W911NF-07-1-0376 and by NASA under Grant No. NNX07AK49A.

Time series data are obtained from sensors and/or from information fusion of sensor signals and analytical models at the time scale of engine gas path dynamics, which is called the fast scale. Sets of (fast scale) time series data are generated at slow scale (discrete) time epochs, over which statistics of the engine dynamical system may become non-stationary, possibly due to occurrence of faults. In general, a long time span in the fast scale is a very tiny (i.e. several order of magnitude smaller) interval in the slow scale, over which the system dynamics are assumed to be quasi-stationary. For example, evolution of anomalies causing a detectable change in the combustor dynamics occurs on the slow scale of (possibly) hundreds of hours, whereas the statistical behavior of the combustor is essentially invariant on the fast scale of, say, minutes.

Then, in this approach the set of (fast scale) time series data, generated at a (slow scale) time epoch, is converted into a symbol sequence by partitioning a compact region (i.e., closed and bounded) region  $\Omega$  (over which the data evolves) of the phase space into finitely many discrete blocks. Each block is labeled as a symbol  $\gamma_j \in \Gamma$ , where the symbol set  $\Gamma$  is called the alphabet, consisting of  $|\Gamma|$  different symbols. As the system evolves in time, the trajectory of the dynamical system travels through various blocks in its phase space and the corresponding symbol is assigned to it, thus converting the time series data sequence to a symbol sequence  $\gamma_0\gamma_1\dots\gamma_\ell\dots$ . Thus, generation of a symbol sequence represents coarse-graining of the trajectory's time evolution [15].

Partitioning of the phase space is a crucial step in *SDF*. Several partitioning techniques have been reported in literature, primarily based on symbolic false nearest neighbors (*SFNN*) [16] that become cumbersome and extremely computation-intensive if the dimension of the phase space is large. Therefore, symbolic sequences as representations of the system dynamics should be generated by alternative methods because phase-space partitioning might prove to be a difficult task in the case of high dimensionality of the phase space and in presence of noise; both of these difficulties arise in the gas path analysis. A wavelet-based partitioning technique [6], [14] is used for construction of symbol sequences. The wavelet partitioning largely alleviates the shortcomings of *SFNN* partitioning and is particularly effective with noisy data from high-dimensional dynamical systems such as combustion chambers. The coefficients of the wavelet-transformed data are generated at different scales and time shifts. Graphs of wavelet coefficients versus scale, at selected time shifts, are stacked starting with the smallest value of scale and ending with its largest value and then back from the largest value to the smallest value of the scale at the next time shift.

The wavelet space is partitioned with alphabet size  $|\Gamma|$  into segments of coefficients on the ordinate separated by horizontal lines such that the regions with more information are partitioned finer and those with sparse information are partitioned coarser. In this approach, the maximum

entropy is achieved by the partition that induces uniform probability distribution of the symbols in the alphabet. Figure 1 depicts an overview of symbol generation from time series data. A comparison of wavelet partitioning using maximum entropy principle and other partitioning methods such as *SFNN* is reported in recent literature [14], where wavelet partitioning has been shown to yield comparable performance with several orders of magnitude smaller execution time. This feature is well suited for real-time applications for early detection of anomalies.

Once the partitioning is performed with alphabet size  $|\Gamma|$  at the nominal condition of, say, time epoch  $t_0$ , it is kept constant for all subsequent time epochs,  $t_1, t_2, \dots$ . In other words, the partitioning structure generated at the nominal condition serves as the reference frame for data analysis at subsequent slow time epochs that may experience anomaly occurrence and hence have different statistical distribution.

A finite state machine is then constructed, where the machine states are defined corresponding to a given alphabet  $\Gamma$  and window length  $D$ . The alphabet size  $|\Gamma|$  is less than or equal to the total number of partitions while  $D$  is the length of symbol strings forming the machine states. That is, the states are chosen as possible words of length  $D$  from the symbol sequence, thereby making the number of states to be less than or equal to the total permutations of the alphabet symbols within a word of length  $D$ . The choice of  $|\Gamma|$  and  $D$  depends on specific experiments, noise level and also on the available computation power. Using the symbol sequence generated from the wavelet coefficients of time series data, the state machine is constructed on the principle of sliding block codes [6], [11]. Similar to the partitioning, the state machine structure is generated at the nominal condition and serves as the reference frame for data analysis at subsequent slow time epochs. Construction of damage patterns from the state machine is described below.

The probability distribution  $p_k$  of damage patterns, represented as histograms in Figure 1, is recursively computed as an approximation of the natural invariant density of the dynamical system at the slow time epoch  $t_k$ , which is a fixed point of the local Perron-Frobenius operator [17]. The computation of the posterior distribution  $p_k$  from time series data at a time epoch  $t_k$  is computationally straightforward and can be executed in real time on an inexpensive platform.

### III. FAULT DIAGNOSIS AND ISOLATION

The proposed architecture for degradation monitoring in an aircraft gas turbine engine, consists of dynamic process models for individual engine components that are augmented with fault diagnosis and isolation (*FDI*) algorithms. As described in section IV, the engine system model is composed of different interconnected components (e.g. fan, compressor, combustor, high-pressure and low-pressure turbines, nozzle, and afterburner) and the engine control system. To illustrate the *FDI* architecture, a complex multi-component and interconnected system is defined

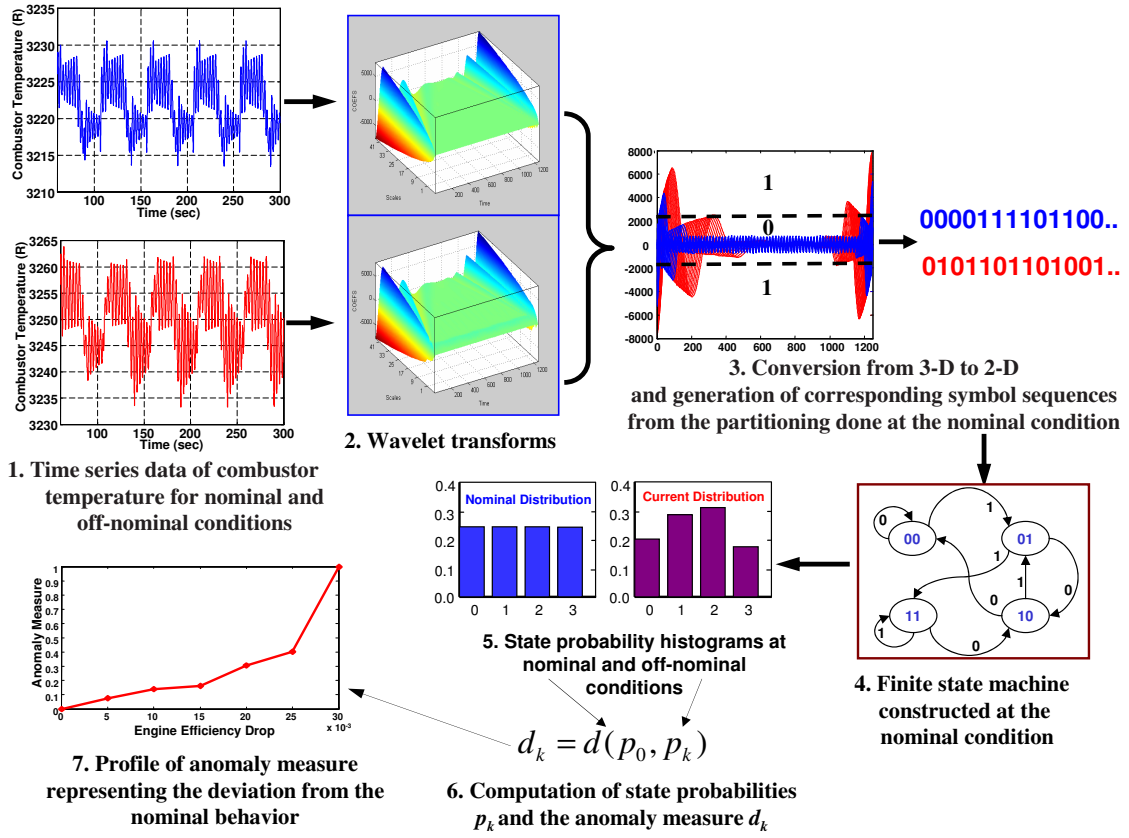


Fig. 1. Overview of Symbolic Dynamic Filtering (SDF)

below, based on the above discussion.

*Definition 3.1:* Let the  $i^{th}$  component of the system be designated by  $C_i$  and let there be  $m$  components such that  $C = \{C_i : 1 \leq i \leq m\}$  is the set of all components. Let the input excitations to the  $i^{th}$  component be given by a vector  $\mathbf{u}_i$  and let the corresponding output response be given by a vector  $\mathbf{y}_i$ . Let  $\delta_{C_i}(\bullet)$  define the closed loop input-output transition function for component  $C_i$  such that  $\Delta = \{\delta_{C_i}(\bullet) : \mathbf{y}_i = \delta_{C_i}(\mathbf{u}_i), 1 \leq i \leq m\}$  is the set of transition functions of all components of the system. Let  $\tau_{i,j}(\bullet)$  be the interaction between the  $i^{th}$  and the  $j^{th}$  components such that  $\mathcal{T} = \{\tau_{i,j}(\bullet) : \mathbf{u}_j = \tau_{i,j}(\mathbf{y}_i), 1 \leq i \leq m, 1 \leq j \leq m\}$  is the set of all interactions of the system components. A multi-component and interconnected system is then defined as a set  $S = \{C, \Delta, \mathcal{T}\}$ .

The process of fault diagnosis involves detection of anomalies in the system by sensing off-nominal behavior and identification of the causes by eliminating all sources of false alarms including noise and disturbances and operational errors. In a multi-component system, detection and isolation of faulty components require analysis of time series data from local sensors and/or process variables of individual components.

The relevant outputs  $\mathbf{y}_i$  from each component  $C_i$  are analyzed using the *SDF* algorithm to generate a pattern vector  $\mathbf{p}_i^k$  at slow time epoch  $t_k$  as summarized in Section II. The pattern vector  $\mathbf{p}_i^{nom}$  generated from the

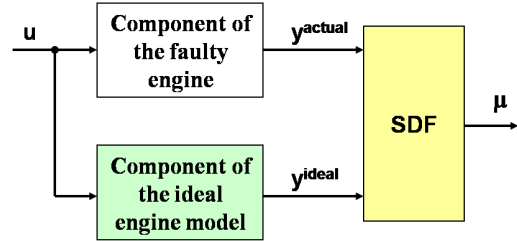


Fig. 2. Fault Detection by Comparison with the Ideal Model

nominal model serves as the reference pattern vector for fault diagnosis using *SDF*. The anomaly measure  $\mu_i^k$  of component  $C_i$  at slow time epoch  $t_k$  can be calculated as  $d(\mathbf{p}_i^k, \mathbf{p}_i^{nom})$ , where  $d(\bullet, \bullet)$  is an appropriately chosen distance function (Euclidean norm in this case [6]). Therefore, the anomaly measure is based on the distance between the reference pattern, generated from *SDF* of  $\mathbf{y}_i^{nom}$ , and the actual pattern at time  $t_k$ , generated from *SDF* of  $\mathbf{y}_i^k$ . This definition of anomaly measure is sufficient for a single component system or a multi-component system with no mutual interactions. Since a multi-component system  $S$  in Definition 3.1 has different mutual interactions  $\mathcal{T}$ , the anomaly measure has to be calculated such that the nominal and the faulty models of each component are excited at the same input excitation. This is due to the fact that the anomaly measure can be non-zero either if there is an actual fault in the component relative to the nominal condition, or if the input to the component is anomalous

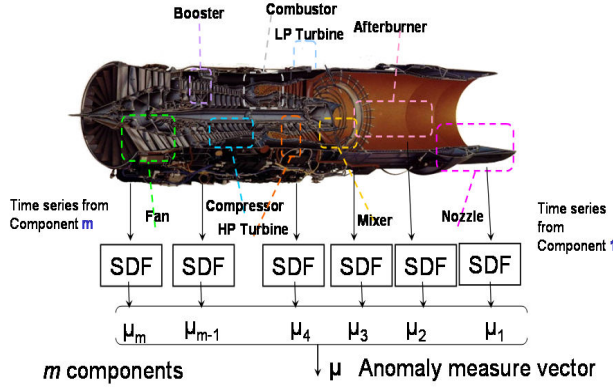


Fig. 3. Fault monitoring in a Gas Turbine Engine

due to a fault in some other component. In both cases, the fault status of the component could be falsely identified due to altered input excitation. Therefore, the input excitation as currently available to the actual component model must excite the nominal model for generation of the reference pattern for *SDF*. Since the input excitation is the same in this case, the distance  $d(\mathbf{p}_i^k, \mathbf{p}_i^{nom})$  between pattern vectors  $\mathbf{p}_i^k$  and  $\mathbf{p}_i^{nom}$  can only be observed if there is a fault in the actual system component, as illustrated in Fig. 2. In this context, the following points are noteworthy:

- The time series data  $\mathbf{y}_i^{nom}$  generated from the nominal model of component  $C_i$ , could change at different time epochs because the input excitation  $\mathbf{u}_i$  to the component may change due to a fault in some other component. However, the reference pattern vector  $\mathbf{p}_i^{nom}$  is always a uniform distribution as a result of maximum entropy partitioning as summarized in Section II.
- Point-by-point comparison of data points in time series data sets  $\mathbf{y}^{nom}$  and  $\mathbf{y}^{actual}$  is not physically useful because of the possible phase difference and noise in the signals, thereby requiring the use of statistical comparison of pattern vectors using the *SDF* method.
- An observation of non-zero value of the anomaly measure  $\mu_i^k$  at time epoch  $t_k$  enables both fault diagnosis and isolation in component  $C_i$ .

In multiple fault scenario, the engine system as shown in Fig. 3 can be divided into critical components such that the relevant output response from each component is analyzed using *SDF* to generate the pattern vectors that are compared with the corresponding reference patterns generated under the same input excitations. Therefore, each component  $C_i$  yields the value of anomaly measure  $\mu_i^k$  at time epoch  $t_k$ . The ensemble of information, generated from *SDF* of individual components is expressed as an  $m$ -dimensional anomaly measure vector  $\boldsymbol{\mu}^k = [\mu_1^k \ \mu_2^k \ \dots \ \mu_m^k]^T$ . A non-zero value of anomaly measure for any component indicates that the component is faulty. The procedure is capable of isolating

both single-component and multiple-component faults. The anomaly measure vector  $\boldsymbol{\mu}^k$  at a time epoch  $t_k$ , calculated as described above, serves the purpose of fault diagnosis and isolation in a multi-component system. In case of cascading failures, a non-zero element of  $\boldsymbol{\mu}^k$  at a time epoch  $t_k$  might lead to other non-zero elements at other slow time epochs. Therefore, the evolution of anomaly measure vector  $\boldsymbol{\mu}^k$  at the slow time scale determines the behavioral changes occurring in the individual components and their affects on the overall health of the engine system.

#### IV. IMPLEMENTATION PROCEDURE

Model of a generic two-spool turbofan engine with the engine controller is used to implement the *SDF* based *FDI* algorithm. A Graphical User Interface(*GUI*)(see Figure 4) is made which couples a data-acquisition software and the *FDI* software module required to detect and isolate faults in its major components.

The engine simulator uses the gas turbine engine model including its continuously varying gain-scheduled feedback control system. This model is described by ordinary differential equations and supporting algebraic equations and look-up tables. The test bed is capable of simulating steady-state and dynamical conditions for the engine under different operating conditions. There are six components in this model Fan, high pressure compressor, combustor, high pressure turbine, low pressure turbine and booster. The simulator also includes a basic Damage model. Damage is simulated by independently scaling the thermal and flow efficiencies of each of the components. This efficiency loss may be the result of several factors related to structural damage, which leads to degraded performance and affects the time series of sensor data in the gas path.

In the experimentation procedure, the Engine simulation program runs on one computer and the Data-acquisition and *FDI* software, established on a second computer analyzes the sensor information transported from the first computer over the communication channel. For the purpose of real-life implementation of the *FDI* algorithm, it is proposed to run the physical turbofan engine every time it comes in for preventive maintenance. During this testing process, the input to the engine model would be excited with an input profile of the Power Lever Angle (*PLA*), suitable for analysis. An input profile of the *PLA*, having the mean value of 35 radians and frequency of 10 radians, has been used in the simulation runs. The sensor data collected would be directly sent to the data-acquisition and *FDI* module. In other words, the engine simulator (first computer) (See Figure 4) would be replaced by the actual engine. The *FDI* algorithm would then be used to isolate the faulty components and estimate the magnitude of the fault.

#### V. SIMULATION EXPERIMENTS AND RESULTS

This section presents the results of simulation experiments on diagnosis and isolation of single-component and multi-component faults in the engine test bed. As

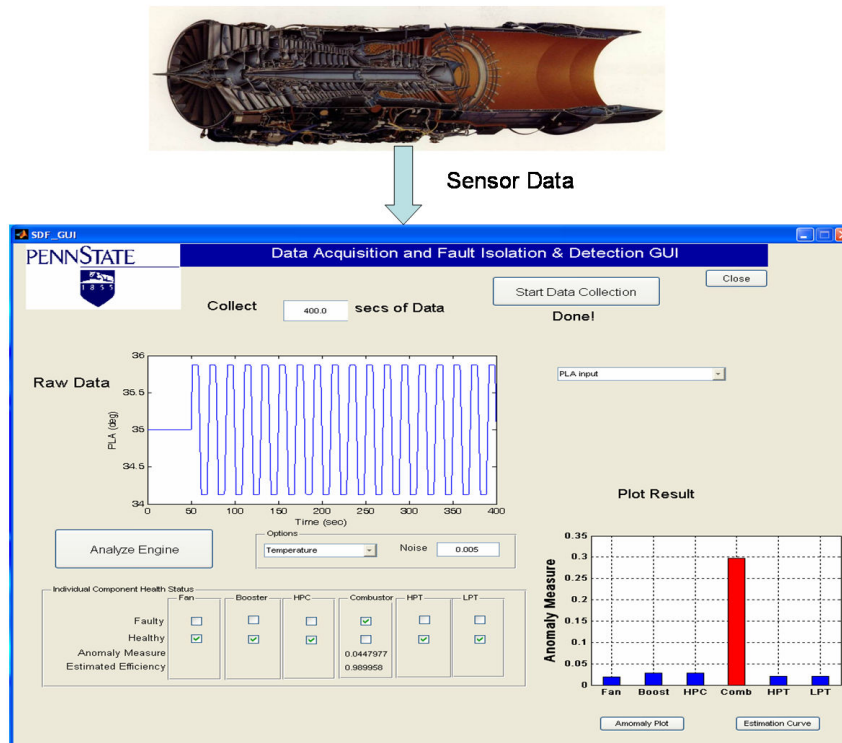


Fig. 4. Software Implementation GUI

explained in the previous section IV, time series of the pertinent sensor data and analytically derived variables are generated under nominal and anomalous conditions for the following two scenarios: (a) *Fault in a single components*, and (b) *Faults in multiple components*. The time series data sets are analyzed to detect and isolate the faulty components. Figure 5 shows the results of diagnosis and isolation of a single faulty component in a multi-component aircraft engine system, where a fault is injected into a single component at a time. The six plots exhibit six cases of individual faulty components as a) fan, b) booster, c) high pressure compressor (HPC), d) combustor, e) high pressure turbine (HPT) and f) low pressure turbine (LPT) in the order. In each of the plots in Fig. 5, the fault is injected by reducing the relative efficiency parameter from 1 to 0.99 that represents a small amount of anomaly in the corresponding component. Time series data sets of temperature, mass flow rate and enthalpy have been analyzed to detect and isolate the faults. The vertical bars in each plot of Fig. 5 represent the values of anomaly measure for each component generated by analysis of the observed time series data. As seen in each plot of Fig. 5, the faulty component has a very high value of anomaly measure as compared to the remaining components that have much smaller (non-zero) values of anomaly measure due to noise. These bar plots indicate that the fault in a single component is successfully detected and isolated from noise corrupted data at an incipient stage. Figure 6 presents the results of detection and isolation of several faulty components in a multi-component aircraft

engine system where faults were injected into multiple components. This is the representation of a scenario when faulty components are not immediately repaired and such faulty components accumulate. The faults are injected into multiple components at the same time by reducing the relative efficiency parameter to 0.99, which signifies a small amount of efficiency drop for the corresponding components. Three plots in Fig. 6 exhibit three cases of randomly chosen faulty components.

## VI. CONCLUSIONS AND FUTURE WORKS

This paper presents the underlying concept of a fault diagnosis and isolation (FDI) method for degradation monitoring in aircraft gas turbine engines. The FDI methodology has been validated on the simulation model of a generic two-spool turbofan engine. A possible manifestation of the technique in real life scenario has also been outlined. Although the theory of the proposed FDI method is built upon rigorous scientific principles, its software implementation requires application-specific relationships such as selection of the partitioning alphabet size [6] and wavelet basis functions and scale range [14]. Further research is recommended to enhance the SDF, especially in the area of space partitioning and fusion of multiple heterogeneous information from sensors into a single symbol sequence.

A critical assumption (that is trivially valid) in the simulation test runs is availability of the engine component models under different input excitation. Such models may not be available for the necessary input conditions in an



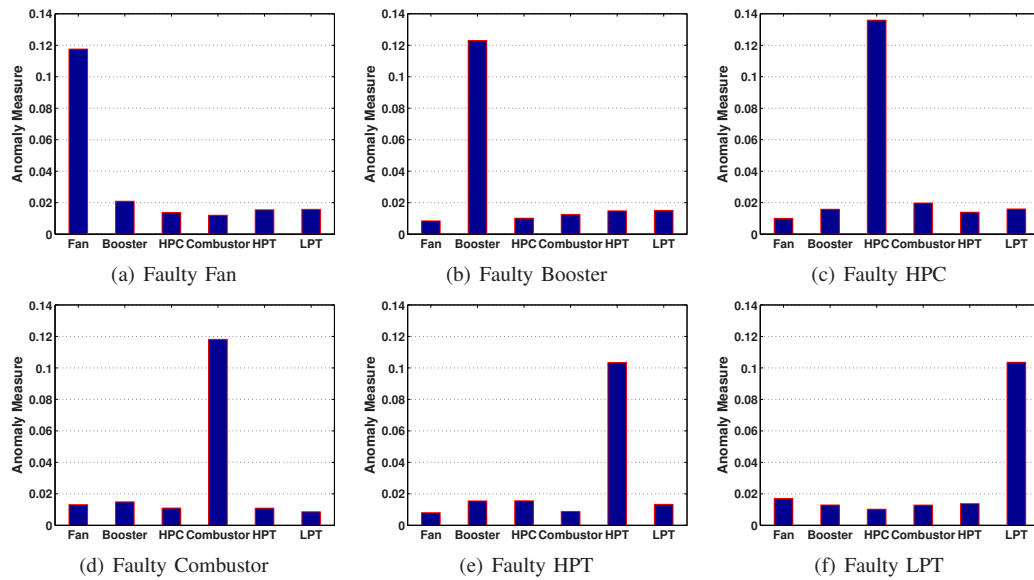


Fig. 5. Single Component Fault Diagnosis and Isolation

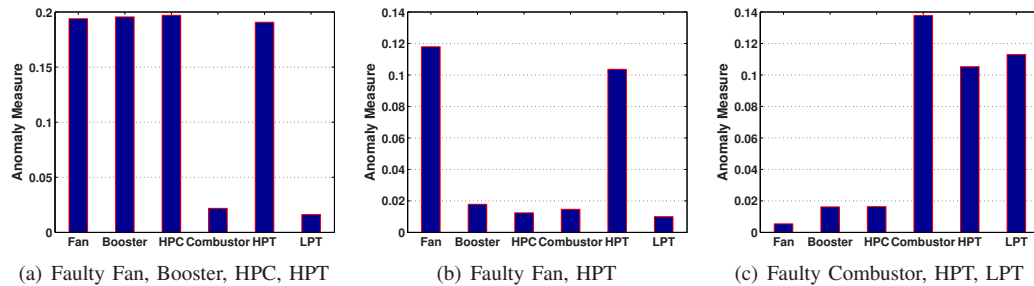


Fig. 6. Multi-Component Fault Diagnosis and Isolation

operating engine. As a further development, the reference behavior of ideal engine components could be identified based on time series data of appropriate sensors. As such data-driven ideal models of engine components could be obtained using black-box system identification such as Artificial Neural Networks.

## REFERENCES

- [1] A. M. Y. Razak and J. S. Carlyle, "An advanced model based health monitoring system to reduce gas turbine ownership cost," *ASME 2000-GT-627*, 2000.
- [2] A. Stamatidis, K. Mathioudakis, M. Smith, and K. Papailiou, "Gas turbine component fault identification by means of adaptive performance modeling," *ASME 90-GT-376*, 1990.
- [3] M. Zedda and R. Singh, "Gas turbine engine and sensor fault diagnosis using optimization techniques," *AIAA-99-2530*, 1999.
- [4] D. C. A. J. D. J. P. Grondel, S. and L. Reithler, "Fatigue crack monitoring of riveted aluminium strap joints by lamb wave analysis and acoustic emission measurement techniques," *NDT&E international*, vol. 35, pp. 339–351, 2006.
- [5] R. A. Gupta, S. and E. Keller, "Fatigue damage monitoring by ultrasonic measurements: A symbolic dynamics approach," *International Journal of Fatigue*, vol. 29, no. 6, pp. 1100–1114, 2007.
- [6] A. Ray, "Symbolic dynamic analysis of complex systems for anomaly detection," *Signal Processing*, vol. 84, no. 7, pp. 1115–1130, 2004.
- [7] Y. Diao and K. Passino, "Stable fault-tolerant adaptive fuzzy/neural control of a turbofan engine," *IEEE Transactions on Control Systems Technology*, vol. 9, no. 3, pp. 494–509, 2001.
- [8] K. I. Parker and T. H. Guo, "Development of a turbofan engine simulation in a graphical simulation environment," in *JANNAF Aero-Propulsion Subcommittee Meeting*, Destin, FL, 2002.
- [9] D. K. Tolani, M. Yasar, A. Ray, and V. Yang, "Anomaly detection in aircraft gas turbine engines," *Journal of Aerospace Computing, Information, and Communication*, vol. 3, no. 2, pp. 44–51, 2006.
- [10] J. P. Eckmann and D. Ruelle, "Ergodic theory of chaos and strange attractors," *Reviews of Modern Physics*, vol. 57, no. 3, pp. 617–656, 1985.
- [11] D. Lind and M. Marcus, *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995.
- [12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley, 1991.
- [13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley, 2001.
- [14] V. Rajagopalan and A. Ray, "Symbolic time series analysis via wavelet-based partitioning," *Signal Processing*, vol. 86, no. 11, pp. 3309–3320, 2006.
- [15] C. Beck and F. Schögel, *Thermodynamics of Chaotic Systems: An Introduction*. Cambridge University Press, 1993.
- [16] M. B. Kennel and M. Buhl, "Estimating good discrete partitions from observed data: Symbolic false nearest neighbors," *Physical Review Letters*, vol. 91, no. 8, p. 084102, 2003.
- [17] R. Badii and A. Politi, *Complexity, Hierarchical Structures and Scaling in Physics*. Cambridge, U.K: Cambridge University Press, 1997.